April 2008 Volume 87 Number 1

ELSEVIER

Journal of Environmental *Management*

Journal of

Environmental Management

Aims and Scope

The Journal of Environmental Management is a journal for the publication of peer reviewed, original research for all aspects of management and the managed use of the environment, both natural and man-made. Critical review articles are also welcome; submission of these is strongly encouraged.

As governments and the general public become more keenly aware of the critical issues arising from man's use of his environment, this journal provides a forum for the discussion of environmental problems around the world and for the presentation of management results. It is aimed not only at the environmental manager, but at anyone concerned with the sustainable use of environmental resources.

Research Areas Include, but are not exclusive to:

• resource quality, quantity and sustainability • economics of environmental management • transport and fate of pollutants in the environment • spill prevention and management • remediation of contaminated sites • process modification for pollution prevention • improved energy efficiency • waste treatment and disposal

Papers submitted should address environmental management issues using a range of techniques e.g. case studies, observational and theoretical analyses, the application of science, engineering and technology to questions of environmental concern or mathematical and computer modeling techniques with the aim of informing both the researcher and practitioner.

Note to contributors

Online submission is available for this journal, to submit your article online, go to http://ees.elsevier.com/jema. Instructions on how to submit a manuscript to the journal can be found in the Guide for Authors on the journal homepage at http://www.elsevier.com/locate/jenvman.

Journal of Environmental Management

Editor-in-Chief

A. GILL

ELSEVIER

Orders, claims, and journal enquiries: please contact the Customer Service Department at the Regional Sales Office nearest you:

Orlando: Elsevier, Customer Service Department, 6277 Sea Harbor Drive, Orlando, FL 32887-4800, USA; phone: (877) 8397126 or (800) 6542452 [toll free numbers for US customers]; (+1) (407) 3454020 or (+1) (407) 3454000 [customers outside US]; fax: (+1) (407) 3631354 or (+1) (407) 3639661; e-mail: usjcs@elsevier.com or elspcs@elsevier.com

Amsterdam: Elsevier, Customer Service Department, PO Box 211, 1000 AE Amsterdam, The Netherlands; phone: (+31) (20) 4853757; fax: (+31) (20) 4853432; e-mail: nlinfo-f@elsevier.com

Tokyo: Elsevier, Customer Service Department, 4F Higashi-Azabu, 1-Chome Bldg, 1-9-15 Higashi-Azabu, Minato-ku, Tokyo 106-0044, Japan; phone: (+81) (3) 5561 5037; fax: (+81) (3) 5561 5047; e-mail: jp.info@elsevier.com

Singapore: Elsevier, Customer Service Department, 3 Killiney Road, #08-01 Winsland House I, Singapore 239519; phone: (+65) 63490222; fax: (+65) 67331510; e-mail: asianfo@elsevier.com

USA mailing notice: Journal of Environmental Management (ISSN 0301-4797) is published monthly, with additional copies in January, April, July, and October, by Elsevier Ltd. (The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK). Annual subscription price in the USA US$ 1223 (valid in North, Central and South America), including air speed delivery. Periodical postage paid at Rahway NJ and additional mailing offices.

USA POSTMASTER: Send address changes to Journal of Environmental Management, Elsevier, 6277 Sea Harbor Drive, Orlando, FL 32887-4800

AIRFREIGHT AND MAILING in the USA by Mercury International Limited, 365, Blair Road, Avenel, NJ 07001.

Printed by the Alden Group, Oxford, UK

Author enquiries
For enquiries relating to the submission of articles (including electronic submission where available) please visit this journal's homepage at http://www.elsevier.com/locate/jenvman. You can track accepted articles at http://www.elsevier.com/trackarticle and set up e-mail alerts to inform you of when an article's status has changed. Also accessible from here is information on copyright, frequently asked questions and more.

Contact details for questions arising after acceptance of an article, especially those relating to proofs, will be provided by the publisher.

Funding body agreements and policies
Elsevier has established agreements and developed policies to allow authors who publish in Elsevier journals to comply with potential manuscript archiving requirements as specified as conditions of their grant awards. To learn more about existing agreements and policies please visit http://www.elsevier.com/fundingbodies

:LSEVIER

Journal of Environmental Management 87 (2008) 1–13

Journal of
Environmental
Management

www.elsevier.com/locate/jenvman

Strategies for sustainable development of industrial park in Ulsan, South Korea—From spontaneous evolution to systematic expansion of industrial symbiosis

Hung-Suck Park[a,*], Eldon R. Rene[a], Soo-Mi Choi[a], Anthony S.F. Chiu[b]

[a]*Department of Civil and Environmental Engineering, University of Ulsan, P.O. Box 18, Ulsan 680-749, Republic of South Korea*
[b]*Department of Industrial Engineering, De La Salle University-Manila, 1004, Philippines*

Received 15 June 2006; received in revised form 22 November 2006; accepted 31 December 2006
Available online 6 March 2007

bstract

The Korea National Cleaner Production Center (KNCPC) affiliated to the Korea Institute of Industrial Technology (KITECH) has arted a 15 year, 3-phase EIP master plan with the support of Ministry of Commerce, Industry, and Energy (MOCIE). A total of 6 dustrial parks, including industrial parks in Ulsan city, known as the industrial capital of South Korea, are planning projects to find the asibility of shifting existing industrial parks to eco-industrial parks. The basic survey shows that Ulsan industrial complex has been ntinuously evolving from conventional industrial complexes to eco-industrial parks by spontaneous industrial symbiosis. This paper scribes the Korean national policies and the developmental activities of this vision to drive the global trend of innovation for nverting the existing industrial parks to eco-industrial parks through inter-industry waste, energy, and material exchange in Ulsan dustrial complexes. In addition, the primary and supportive components of the Ulsan EIP pilot project, which will be implemented for years is elaborated with its schedules and economic benefits.

ywords: Eco-industrial parks; Industrial symbiosis; Sustainable development; Policies; Industrial networks; South Korea

Introduction

In 1962, Ulsan was appointed as a special industrial park South Korea and since then, several industrial complexes ve been established without considering the environental impacts. Ulsan has grown up to be the industrial pital of Korea, with a strong hold of petrochemical, nferrous metal, ship building and automobile industries ig. 1). There are 504 industrial parks in South Korea, out which 34 are large scaled industrial facilities accounting r 2/3rd of the total land for industrial parks (164,440 res). One among them is a cluster of complexes located in san City, which was been often described as the ollutant's department store" and "the country's most lluted city". Recently, awareness campaigns followed by ingent environmental standards and legislations have motivated these industries to adopt suitable technologies that lead to cleaner production and abatement of all forms of pollution (Park et al., 2004). After the Rio Earth Summit in 1992, a comprehensive approach was taken to improve the environmental, social and business performance in Korean industry by applying the concepts of cleaner production and industrial ecology. Ulsan industrial complexes has been continuously evolving from conventional industrial complexes to eco-industrial parks, based on sustainable development policies adopted by existing industries (Lowe, 2001; Chiu, 2003; Park et al., 2004; Lowe and Chiu, 2005).

An eco industrial park (EIP) is "an industrial system which conserves natural and economic resources; reduces production, material, energy, insurance and treatment costs and liabilities; improves operating efficiency, quality, worker health and public image; and provides opportunities for income generation from use and sale of wasted materials" (Cote and Hall, 1995). Lowe (1996) pointed out

*Corresponding author. Tel.: +82 52 259 1050; fax: +82 52 221 0152.
E-mail address: parkhs@ulsan.ac.kr (H.-S. Park).

)1-4797/$ - see front matter
:10.1016/j.jenvman.2006.12.045

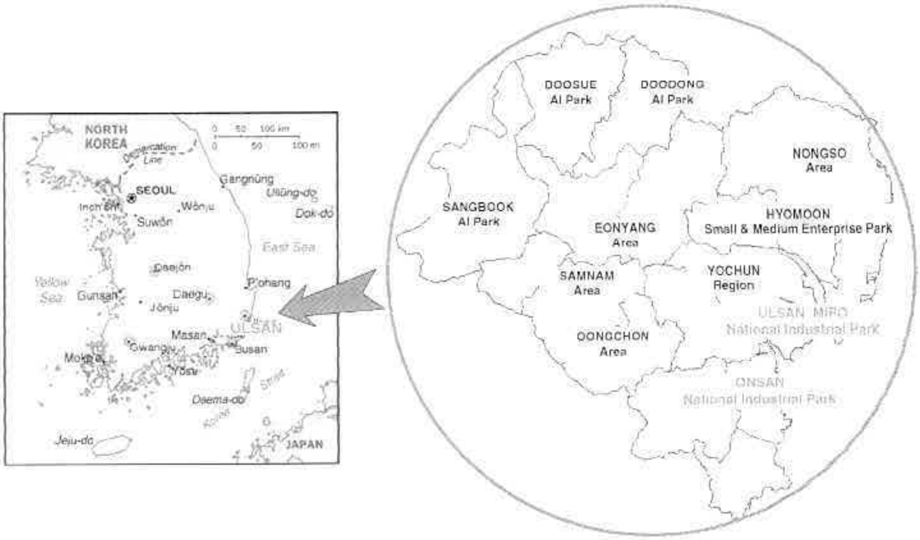


Fig. 1. Map of Ulsan industrial complexes.

that "EIP conjoins the principles of mixed use development, recycling business and by-product exchange in coordination with green technology companies that makes eco friendly products". There have been substantial efforts to formulate strategies and enforce regulations that could change this scenario. Additionally, there are many combinations of sustainable development programs, which basically take the following forms (Lowe, 2001):

By-product exchange (BPX)—a set of companies seeking to utilize each other's by-products (energy, water, and materials) rather than disposing them as waste.
Eco-industrial network (EIN)—a set of companies collaborating to improve their environmental, social, and economic performance in a region.
Eco-industrial park or estate (EIP/EIE)—an industrial park developed and imaged as a real estate development enterprise but seeking higher and balanced environmental, economic, and social benefits as well as business excellence.

The Kalundborg industrial symbiosis has been envisioned as a standard model for the sustainable development of EIPs in the world (Ehrenfeld and Gertler, 1997). This had a strong effect on the rapid planning and implementation of EIPs in Europe and America in the early 80 s. The prime objectives behind this have been "to reduce, recover, reuse and recycle wastes keeping in mind the economic and social benefits". Furthermore, it also improves relationships with external parties, facilitates development of new products, opens new ventures, seeks opportunities and maintains a safe and clean work environment (Ayres and Ayres, 1996; Esty and Porter, 1998; Chertow, 2000; Mirata, 2004). Industrial Symbiosis (IS) in UK was inspired by the Mexican by-product synergy programme (Business Council for Sustainable Development (BCSD), 1999). The case of the Landskrona industrial symbiosis programme in Sweden was well appreciated by managers who concluded that "being a part of the network has helped them with valuable ideas on alternative ways for addressing their current environmental concerns and that they now anticipate the same for their potential future problems" (Mirata and Emtairah, 2005). In developing countries, the concept of IS has shown to improve international cooperation, introduce new concepts, tools and technologies, teach new lessons from experiences and involves government participation in international environmental treaties and conventions (Chiu and Yong, 2004). In Australia, a Synergy Industrial Park master plan was laid using a consultative process targeting on the layout of a food processing industry. This layout featured a common denominator in order to share the infrastructure, which took nearly 5 years to develop the concept. However, some lessons were learned that could be common to the development of any EIP. They include "the importance of an industrial catalyst, a synergy trust, coordination of key industries and utilizing players, the need for strategic planning that has foresight and flexibility and developing community oriented development" (Roberts, 2004). Three projects, viz., ecology programme, mini parks and alternative fuel through long range planning constitute the integrated component of EIP development in Singapore's Jurong Island (Yang and Lay, 2004).

This article describes the strategies adopted by the Ulsan Industrial Complex in response to the Korean national EIP program. The strategy covers two levels, the first level referred to the national environmental policies stimulating sustainable development and the industrial policies promoting eco-friendly industrial structures in South Korea. The second level of strategy focuses on the regional level of IS in Ulsan industrial complex together with its economic effects. The regional level adopted the policy-structure-technology (PST) approach (Chiu, 2003, 2005) wherein success factors for EID were classified into primary and supportive components. The different development strategies for its successful implementation are also addressed comprehensively.

. National level: Korean policies supporting EIP and their salient features

.1. Environmental policies stimulating sustainable development

The policies pertaining to sustainable environmental evelopment focuses on fulfilling the goals of environmental conservation and to develop an investment plan to nprove the existing environmental infrastructure facilities. he environmental policy entitled "Environmental Vision I" strives to provide a pleasant living space and pursue to alance conservation and development within the scope of nvironmental capacity and to prepare for assistance ructure (MOE, 2004). The five principles that were roposed to enable their implementation are mentioned s follows: (i) principle of prevention, (ii) principle of armony and integrity, (iii) principle of charging polluter nd recipient, (iv) principle of utilizing economic incenves, and (v) principle of opening information and getting sidents involved (Lee et al., 2005). Further, this policy renghened its initiatives in preserving the natural sources by implementing steps that considers the water, r and soil environment. An aggregate volume based gulation for sufficient supply of clean water, introducing gional based management for air pollution and building ean production system for reducing and recycling solid astes were proposed. For the nationwide effective nplementation of environmental policies, various reform easures are taken. It is proposed to establish a green ross National Product (GNP) and benchmark for stainability to achieve eco-friendly development through: troduction of an environmental agreement, strengthening e regional environmental management systems (EMS) rough standards and promotion of regional Agenda 21. he updated version of this policy sets the goal as "to aximize the conservation of the natural environment by inimizing consumption of resources and to establish a ciety of justice that equally guarantees quality of life for l people". The primary objective was henceforth to hieve economic stability for better living standards, cial maturity that provides equal opportunity and a ealthy environment that conserves the ecological capacity. echnically, the strategies target in establishing an vironmentally sustainable national innovation system, ster a market foundation for environmental industry and vironmental science and technology through the rational form of environmental regulation. The first Presidential ommission on Sustainable Development (PCSD) was rmed as an inter ministry regulatory authority to focalize rmony between environmental policies and other national policies, constitute rapid action plans for Agenda21 d assay strategies to cope up with global environmental regulations. A long-term vision named The Third Comprehensive National Environmental Plan was also prepared by the Ministry of Environment (MOE) for the year 2006–2015 with four major goals having different strategies. These strategies focus on the following themes: providing safe and quality living conditions, conserving the natural eco-system, effective utilization of natural resources, seeking eco-friendly economy, establishing environmental justice, strengthening cooperation with Asian countries and initiating global sustainable development. The environment-friendly business management policy promotes preventive solutions to environmental pollution that arises during production processes. Additionally, a corporate environmental information disclosure system and an eco-labeling system has been introduced and implemented to stimulate the environment-friendly production and consumption.

2.2. Policies promoting eco-friendly industrial structures

2.2.1. Industrial policies

The manufacturing industry would play a significant role in the growth of the Korean economy in the future. Hence, the MOCIE has been concentrating its policy capabilities to expand the development base such as industry supply and demand and the market base (technology, manpower and deregulation), so as to achieve world class competitiveness. A rapid shift in economic growth strategy from a capital and external growth driven strategy to innovation and qualitative growth driven strategy has been planned. The latter includes: (i) developing source and core technologies; (ii) strengthening the combination of basic sciences with industrial technology; (iii) knowledge and information intensive techniques; (iv) high-quality batch production with a "Korea Premium" brand; and (v) interlinking technical groups for joint research. The 2010 Industrial Vision of Korea also focuses on strategies to address the salient features that would enable Korea to emerge as one of the four industrial superpowers by 2011. The most striking feature of this policy pertaining to our focus is chapter 30 of Agenda 21, in which the role of business for the support of Agenda 21 is defined. It states that the involvement and cooperation of business organizations are vital factors in achieving the objectives of Agenda 21, sustainable development.

Industrial environmental policy has drastically changed after the Ministry of Commerce, Industry, and Energy (MOCIE) enacted 'APEFIS', an Act to Promote Environmental Friendly Industrial Structure, in December 1995. Based on the APEFIS, the MOCIE established an institutional system for cleaner production (CP) and EMS based on ISO 14001 as an implementing tool. The first comprehensive master plan for environment friendly industrial development was made and operated based on APEFIS. This plan includes: streamlining the supporting system, CP transfer and dissemination, promoting environmental industry, and stimulating environmental

Table 1
Details of the six industrial complexes chosen for the present EIP study

Location of industrial complex	Land area (acres)	No. of companies	Typical industries
Banwol and Siwha, Seoul	7860	5400	Textile, dying, chemical processing, incinerators, pulp and paper mill etc.
Mipo and Onsan, Ulsan	13700	700	Non ferrous metals, steel, metal manufacture, automobile, ship building, petrochemicals, refinery, incinerators, etc.
Yeousu	7736	149	Petrochemical industry, refinery, etc.
Cheongju	1010	200	Pulp and paper, electronics, non ferrous metals, metal processing, food processing, petrochemicals, etc.
Machun, Chilso and Sangpyeong in Jinhae, Haman and Jinju	1450	550	Steel, metal finishing, non ferrous metals, food processing, chemical industries, etc.
Pohang	4970	220	Cement, steel industry, metal processing, fine chemical industry, waste disposal, etc.

management. The CP transfer and dissemination deals with technology transfer, international collaborative projects, supply chain environmental management (SCEM), EMS and eco-industrial parks (EIP).

2.2.2. *Energy policies and newer strategies*

The Korean government has made significant efforts to regulate energy supply and demand and energy prices. Since Korea has few natural resources, dependence on overseas sources has risen from 87.9% in 1990 to 98.8% in 2005. The share of energy imports among total imports also increased from 15.6% in 1990 to 25.8% in 2005. The five major industries that contribute to this policy are: petroleum, natural gas, electricity, nuclear power and coal industry. The total energy demand is expected to increase by 3.2% annually by the end of 2010. This figure is substantially lower than the expected annual economic growth rate of 6%. There has been a new shift in paradigm in maintaining the existing energy policy guidelines of Korea. The strategies adopted include: (i) shifting from government-led to market-driven policy, (ii) management of energy demand, and (iii) industry energy characteristics. The visions are to follow international environmental regulations, open new ventures with private sectors, reform existing technologies and to block economic globalization.

2.3. *The Korean EIP development strategies—transition from IP to EIP*

The Korea National Cleaner Production Center (KNCPC) and the Korea Institute of Industrial Technology (KITECH) started a 15-year, 3-phase project titled, 'Eco-industrial Park (EIP): construction for establishing infrastructure of cleaner production in Korea' with the support of MOCIE. The six industrial complexes that are included under this closed domain of EIP study are given in Table 1 (Lee, 2005).

The first phase (2006–2010) of the developmental plan strives to perform trial projects for two industrial parks in order to shift them to EIPs, with prior understanding of the material and energy flow analysis, input and output of raw materials, products, by products and wastes. An energy efficient BPX network would be created using the basic concepts of industrial ecology. Pollution monitoring systems are to be installed to envisage the existing wastewater and waste treatment systems. Additionally, an integrated EMS would emerge together with detailed analysis of the infrastructure. Sustainable education and awareness campaign would also be conducted. The development so far has envisioned for further phases that would upgrade the existing manpower resources in conjunction with an organizing group that manages the operation on a timely basis.

The second phase (2011–2015) would provide conceptual ideas and disseminate understanding of the designed concept to 20 other industrial parks. It would also help in spreading the EMS and sustain a balance between the different key factors that are likely to influence economic growth. A system of common sharing and practice, common purchase and common transportation system would be organized to establish an enlarged infrastructure that is capable of handling joint ventures. The third phase (2016–2020) would overview the flaws and constraints envisioned in the earlier phases and strive to rework and reinvent the existing system of practice. The performance indicators would be analyzed and evaluated by an expert committee to redesign any missing components and infrastructure. The ultimate aim would idealize and pave the way to provide zero discharge in all process industries within the EIP.

The Ulsan EIP would expand on these themes and provide ample opportunities to establish a more convenient and user friendly symbiotic system.

Table 2
Korean EIP construction plan

Categories	1st year	2nd year	3rd year	4th year	5th year	Total
Technology development						
By-product exchange	174	348	348	348	435	1652
Process analysis and optimization	87	174	130			391
Alternative raw material	174	348	435	435	435	1826
Material cycle modeling	43	43	43			130
Water pinch expansion	43	43	43			130
Technology transfer						
Process diagnosis and analysis	234	348	348			930
Integrated EMS				87	261	348
Inter-industry CMS				43	157	200
Infrastructure						
By-product and waste recycle DB		87	87	87	87	348
EIP professional education		87	87	43	43	260
Integrated Recycling Pilot Plant		174	867	867	1739	3652
Integrated resource recycling system			130	870	1739	2739
Comprehensive Water reuse Network				870	1304	2174
Total	757	1652	2521	3652	6200	14,783

Note: Unit: 10,000\$, 1 USD,\$ = ~1000 Wons.

3. Regional level: the Ulsan EIP pilot project development strategy

The 'Korean Eco-Industrial Park Model' is characterized by a cluster of inter-networking businesses, which perform individual and collective cleaner production program prior to by-products exchange network, within an EMS framework (Chiu, 2005).

Under such framework, the KNCPC and Korea Industrial Complex Corporation will be the main national actors in implementing the different phases and strategically supervising the development and implementation programmes. Meanwhile, the regional eco-industrial development (EID) team has ample resources for planning, decision making and implementation of an EIP project comprising of personnel with appropriate management and research skills, which include experts having industrial, academic and administrative experiences. At this phase an open approach is taken to explore the precondition required for allocation of budget and seeking potential investors. Moreover, the active participation from different investing groups such as existing industries, civic community groups, governmental and non-governmental organizations and institutional groups are sought to enhance the process of planning and implementing. The local government however would play a significant role in the developmental activities since it eventually manages the extensive housing, municipal services and major infrastructure in Ulsan. It also acts as a developer and a regulatory authority as evident from the TEDA Urban Industrial Complex in China (Koenig, 2005). The involvement of stakeholders represents a broad domain and a diversified group to collaborate and improve issues pertaining to create effective and appropriate solutions. "People preconceived notions of industrial zones as towering monoliths belching clouds of black smoke from their peaks are difficult to overcome" (Koenig, 2005).

The compendious construction plan of the Korean EIP programme is shown in Table 2, and the EIP plan charted out for Ulsan City focuses on the following key strategies:

Primary component of EID deals with resources interaction wherein resources (material, energy, water, and information flow) are optimally directed for utilization in the system (Chiu, 2003, 2005).

- Implementing ecological constraints to protect indigenous flora and fauna and to maintain natural ecological stability.
- Providing appropriate sewer system, proper maintenance and clean constructional practices.
- Pollution monitoring and abatement programs—systematic effort.
- Appropriate rainwater harvesting methods to be practiced together with runoff management.
- Utilizing renewable energy sources and conversion of waste into energy.

Supportive component of EID deals with the system element inter-relationship. This component generates sufficient conditions for the EID primary actions to take place (Chiu, 2003, 2005).

- Clearly understanding the policies outlined by the MOE in Korea and seeking relevant information about international cooperation.
- Specifying the project site and evaluating the potential of collaborators.

- Establishment of a work group that strategically evaluates the economics.
- Creation of a multi faceted community that lives in social harmony and understanding with each other.
- Public litigation measures and awareness concerning legislations, resource recovery and natural environment together with environmental education.
- Organizing different groups that would supervise, manage and provide adequate analysis and support for the above-mentioned strategies.

The most critical factors evolved from the consultation work done by Indigo Development (Lowe and Chiu, 2005) states that the success of the Korean EIP programme can be distinguished based on policy, structural (management and organizational), and technological based (PST) factors . They are highlighted as follows:

- Good cooperation among the national agencies with responsibility for EIPs.
- Each pilot industrial park requires an adequate management structure for coordination and cooperation supporting the transition to an EIP.
- Both public management authorities and business associations require capacity development and education so they can participate effectively in the EIP initiative.
- Businesses in the park need to be involved from the beginning of the planning process. They are the ultimate actors in the system.
- The high level planning process for the transition to EIPs must be supported by a strong bottom up planning process, i.e., a dialogue between top down and bottom up.
- An evolving long-term vision of the whole system is required to make effective decisions about the specific strategies used in each phase of the transition.
- An EIP is much more than an exchange of by-products among companies.
- Strong support to the growth of the environmental technology and services cluster will provide Korean industrial parks with many of the solutions they require.
- Green chemistry is an important field for petrochemical EIPs as well as customers using chemicals.
- Resource based policies.
- Policy in support of the EIP initiative should take an integrated view of all aspects of cleaner production as complementary to eco industrial strategies.
- National policy should support excellent management of the eco-industrial park initiative and individual industrial parks.

3.1. *Status of industrial complexes*

There are two national industrial complexes, Ulsan Mipo and Onsan industrial complexes, four agricultural and industrial estates in Ulsan. Ulsan industrial complex, the core site for automobile, shipbuilding and petrochemical industries and Onsan complex come together to form a horizontally integrated industrial system (Fig. 2, Tables 3 and 4). These complexes have been performing together since 1975, and have played a significant role in improving the Korean economy.

3.2. *Pollutants generated from Ulsan industrial complexes*

Table 5 summarizes the different pollutants generated from Ulsan industrial complexes. The rate of generation of these pollutants is found to be alarming as evident from their relatively high emission loads. It is clearly evident that, apart from the conventional air pollutants like SO_x and NO_x, the volatile organic compounds (VOCs) also contribute significantly in depleting the atmospheric quality in Ulsan region. VOCs are defined as "any compound of carbon, excluding CO, CO_2, carbonic acid, metallic carbides or carbonates, which participates in atmospheric photochemical reactions" (Nunez, 1998). These VOCs, even in low concentration range have been found to be carcinogenic, cause damage to the liver and kidney and paralyze the central nervous system (Martin et al., 1998; Murata et al., 1999). This accounts for a very high percentage of the total air pollutants emitted from Korean industries. Though significant efforts have been made in many fields to improve the environmental quality, Ulsan is still suffering daily from pertinent environmental issues. The strict environmental regulations enforced on process industries have also led to the shutdown of some companies who have violated the law. At this articulate moment, a urgent changeover of industrial complexes to EIP would keep motivating these companies and provide sustainable development of industrial parks.

3.3. *Current industrial symbiosis at Ulsan industrial parks*

Ulsan industrial complex were originally developed as conventional industrial estate, with collective energy providers for electricity, steam and water. However, this spontaneous transition of industrial park to EIP was incited by economic benefits and regulatory constraints. The introduction of the concept of EIP and demonstrating the BPX system evoked nation-wide interest in Korea, especially among municipalities, business managers, and citizens. The different materials that were transferred to and from different companies and their economic benefits are highlighted in Table 6. The present collaboration among industries in Ulsan includes seven partners, whose industrial profile is described below:

(i) *Koentec Ltd.*: Koentec is an environmental management company that specializes in incineration and sanitary landfill operations of wastes in Ulsan. It has an incineration capacity of 300 tons/day together with a landfill volume of 2,245,128 m^3 which covers a land area of 138,460 m^2. A total of 39 different types of wastes are handled using suitable pre-processing

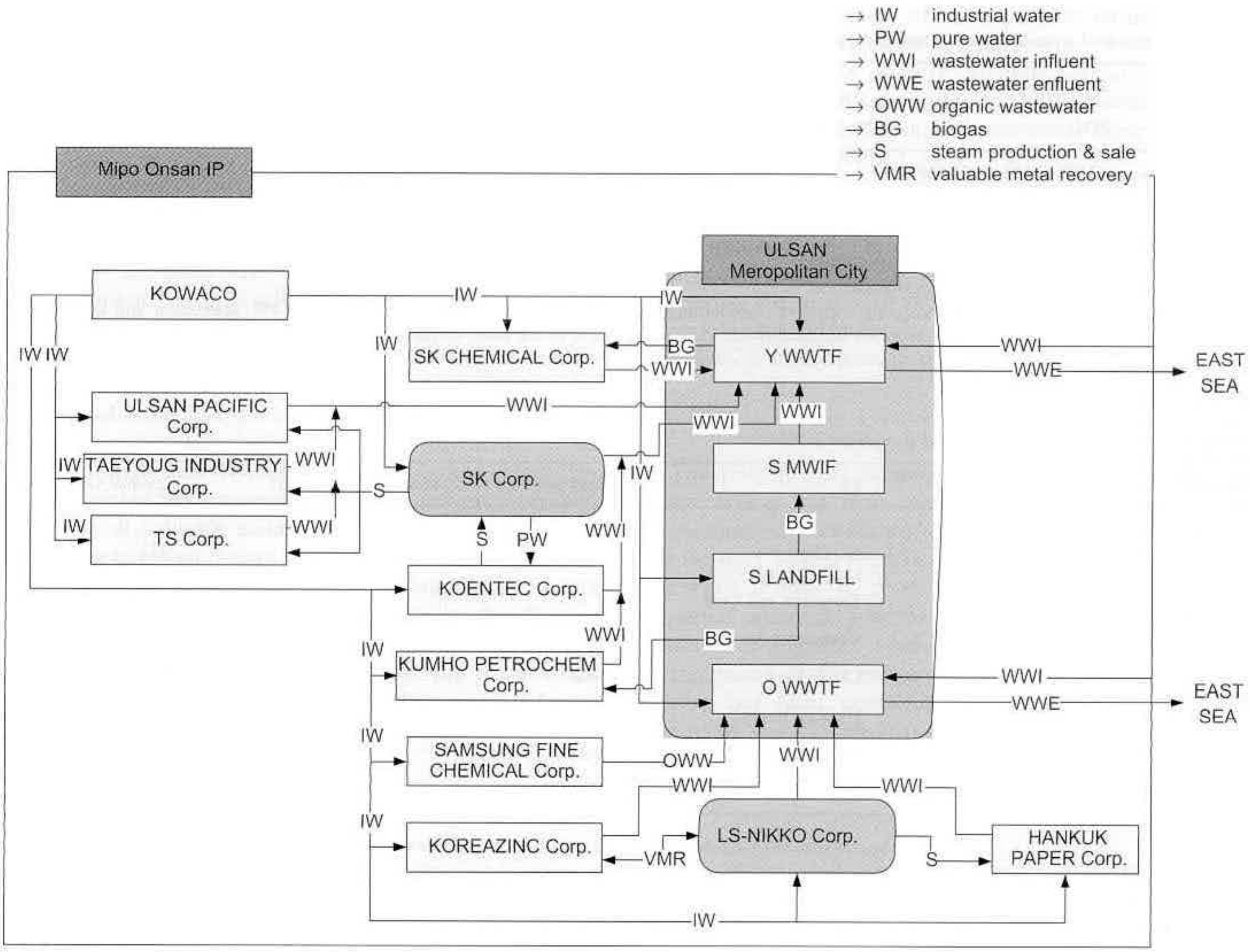


Fig. 2. The present status of industrial symbiosis at Mipo, Onsan industrial complexes in Ulsan.

Table 3
Status of Ulsan/Mipo and Onsan national industrial estates

Items	Ulsan Mipo	Onsan
Area (1000 m^2)		
Total (km^2)	40,340	15,237
Factory area (km^2)	34,567	13,422
Construction	1975. 6. 23	1974. 4. 1
No. of companies		
Moved in	558	212
In operation	485	158
Water supply capacity (m^3/day)	641,000	340,000
Wastewater (m^3/day)	250,000	150,000
Production (100 million $)	443	99
Employee (person)	87,494	9716

Table 4
Classification of industries in Mipo and Onsan national industrial park of Ulsan

Category	Ulsan Mipo	Onsan	Total
Food products	9	–	9
Textile products	5	1	6
Wood papers	13	3	16
Petrochemical	105	52	157
Nonmetallic	24	9	33
Steel	8	13	21
Machinery	149	43	192
Electric, electronic	45	4	49
Transport equipment	97	28	125
Others	13	2	15
Services	54	23	77
Total	522	178	700

Data as on April 2004.

techniques. The incineration facility includes two rotary kilns facilitated with appropriate air pollution control equipments like electrostatic precipitators, wet scrubbers and selective catalytic reduction systems. On the other hand, it has a wastewater treatment plant that has a field capacity to treat 441.5 m^3/day of wastewater that includes leachates from landfills. Its contribution to environmental awareness and concern

Table 5
Pollutants generated from Ulsan industrial complexes

	Air[a,b] (unit: ton)				Wastewater (unit: m^3/day)		Solid waste[a] (unit: ton/day)	
	SO_2	NO_2	TSP	VOC	Generation	Discharge	General	Hazardous
Ulsan (ratio)	82,971 (15.8)	63,569 (6.1)	22,849 (24.6)	82,666 (11.8)	367,216 (8.7)[c]	345,234 (23.4)[c]	6672 (3.3)	296,611 (10.2)
Korea	526,599	1,045,332	92,720	699,214	4,226,321[c]	1,477,166[c]	204,428	2,914,546

Source: MOE (2003).
[a]Data for air and general solid waste are from 2001.
[b]Air emissions are based on fuel consumption.
[c]Neglecting generation ratio of Kwangyang complexes, accounting for over 50% of total and discharge ratio 3.4%, generation and discharge of wastewater from Ulsan industrial complexes are accounting for 27.0% and 24.2% of the total, respectively.

Table 6
Industrial symbiosis in Ulsan industrial complexes in 2004

Material	From	To	Sold/free	Investment (10,000$)	Revenue (10,000$/yr)
Pure water	SK Corp.	Koentec			
Steam	Koentec	SK Corp.	Sold	209	411
Steam	SK Corp.	Ulsan Pacific Taeyoung Ind Corp			
Zn recovery	LS-Nikko	Koreazinc	Sold		461
Cu recovery	Koreazinc	LS-Nikko	Sold		1739
Steam	LS-Nikko	Hankuk Paper	Sold	696	300
Biogas	Y-WWT	SK Chemical	Sold		26
Waste MeOH	Samsung	O-WWT	Free		130

has paved the way to get accredited with ISO-14001 and their research facility has obtained several patents to its credit, the most acknowledged one is the invention of the leachate leakage detection system.

(ii) *SK Corp.*: SK Corp. is a world class company with more than 30 years of experience in the global oil market. It has a 34% share in the domestic market with total annual sales of 212 million barrels, including 60 million exported barrels. In addition, their lubricant brand named *ZIC*, was named No. 1 for the sixth consecutive year in the National Industrial Power Brand survey conducted by the Korea Management Association Consulting. SK Corp supplies nearly 620,000 tons of asphalt products to the domestic market. It has developed and commercialized a modified polymer asphalt product called *Superphalt* representing a new dimension in the growing construction technology in Korea. The operation of Korea's first naphtha cracker (capable of producing 100,000 tons of ethylene a year), located at the Ulsan Complex was initiated in March 1973 by SK Corp. It produces MTBE (200,000 tons), ethylene (730,000 tons), propylene (480,000 tons) and butadiene (100,000 tons) per year, contributing significantly to the international petrochemical industry.

(iii) *SK Chemicals Corp.*: SK Chemicals focuses on specialty chemicals and life science products in addition to its strong petrochemical base. It manufactures the high performance *PETG* resin (*SKYGREEN*®) and is making rapid and steady progress in the polyurethane business. It produces specialty chemicals for environmental, industrial and information technology applications as well as in the area of biotechnology and life science for developing synthetic and natural drugs. Some of them include: polyester acoustic and thermal insulations, biocides, flame-retardant polyester core materials, coolants/anti-freeze, chemicals for water treatment, carbon fiber composite materials, liquid photopolymers, anti-cancer agents, anti-ulcerates, etc. It was chosen as the best company in 2004 for Energy Saving and Collaboration Energy Saving and was awarded the Minister Prize for Environment Management by MOCIE.

(iv) *LS-Nikko Corp.*: LS Nikko produces 99.5% pure copper anode and 99.99% pure electrolytic copper cathode through the process of electrolytic refining. In addition, with other elements extracted from copper concentrate such as iron, sulfur, gold and silver, it also produces *LS-Ferrosand* through smelting and high purity precious metals through a precious metal process. This company currently has the capacity to produce a total 510,000 tons of electrolytic copper cathode a year. The Onsan Smelter and Refinery, is the only plant in the world that produces pure sulfuric acid by using sulfur dioxide gas generated during the copper smelter process. It produces sulfuric acid, liquid sulfur trioxide, liquid sulfur dioxide, gypsum and liquid argon.

(v) *Koreazinc Corp.*: The major products of Koreazinc Corp are gold, silver, palladium, platinum, copper, cadmium, sulfuric acid and bismuth ingot, etc. The Onsan Refinery operates the refining facilities for both zinc and lead. This refinery adopts the zinc residue-processing technology that maximizes the recovery of valuable metals from zinc residue and changes the slag environmentally sustainable. The construction of the zinc and lead refinery together in one plant has generated the synergy effects such as higher recovery ratio of valuable metals, use of various types of raw materials and lower waste discharge. Improvised technologies such as the Top Submerged Lance (TSL) systems are being used to solve the disposal problem of residues generated during the refining process. All these efforts for preserving the environment collectively contributed to the award of Environmental New Technology by the Ministry of Environment in 2002.

(vi) *The Ulsan metropolitan city*: This government organization has a major hold in implementing the visions of the Ulsan EIP. It promotes to review successive ways in executing standards/emission limits and to control the total amount of emission by evaluating and analyzing the capacity of the local atmospheric environment. They provide superior companies with mileage through evaluating their investment on environment management and autonomous conventions. It promotes acquisition of ISO-14001 authentication over the environmental administration and holds nationwide nature preservation seminars and awareness campaigns. Additionally, it also helps to expand the existing drainage treatment facilities, strengthen the inspection of the quality of water from different sources, inspect water purification plants, ameliorate the quality of water around the sea areas and on the shores of rivers, facilitate the reduction of household waste by imposing graded charges, promote recycling and improve independent financial capability of cleaning budgets.

(vii) *KOWACO*: The Korea Water Resources Corporation is an integrated water management enterprise focusing on water-utilization and flood control having both ISO-9001/14001. KOWACO is also targeting its operations for the implementation of remote control operations in electricity generation facilities, development of multi-purpose dams and the operation of water gates throughout the nation by using remote control facilities. Apart from these, it is also involved in the development of technology to ensure sustainable water resource management, water supply systems, water quality analysis, water treatment equipments and facilities. In 2005, it was awarded the prestigious Korea Environmental Prize for water quality management.

Over a very short period of time, these partners spontaneously developed a series of bilateral exchanges, which also include a number of other small companies (presently 39). It simply evolved as a collection of one-to-one deal that made environmental and economic benefits for pairs of participant in each.

3.4. Action plan for Ulsan EIP

The pilot project focuses on the symbiotic transfer of waste arising out from industries such as; spent caustic wastewater, waste plastics, steam and sludge. The forecast for the next 5 fiscal years in terms of the implementation of the pilot project is given in Table 7(A, B), while the expected economic benefits are highlighted in Table 8. A total of 39.9 billion won would be saved through the participation of five major industries in Ulsan. Additionally it has been predicted that 56 projects would be completed by the 5th year (2010) through the transfer of water, energy, solid and other industrial wastes.

3.4.1. Primary component of the pilot study—the metabolism on resources

The pilot project is to expand the feasible IS by intra-plant innovations and approaches, and inter-plant collaboration for improving environmental and economic performance, by both individual companies and collective industrial system. Fig. 3 is the proposed EIP to be implemented with the existing IS network.

3.4.2. Supportive component of the pilot study—the stakeholder interaction, cooperation, and coordination

An EIP project not only comprises of the hardware factors, e.g. technology; but also the software component, i.e. stakeholder involvement. At an estate level, cooperation among the tenants is an essential success factor; however, coordination at a higher level is also of great contribution to the success. This project was involved by several key stakeholders who are very helpful in the understanding and coordination of this industrial estate. The core team was led by the Ulsan University, and consisted of the KICOX (Estate Management Authority of Korea) and the city government's environmental division. Additionally, local and regional NGOs showed special interest to participate in the developmental activities of this project.

Anchor tenants provide a very strong supply chain network for vertical integration of industrial parks. Moreover, many basic internal management systems such as Responsible Care and compliance with local laws already

Table 7
Yearly forecast of industrial symbiosis network through the EIP project

(A) *Status of individual projects*

Resource	1st year	2nd year	3rd year	4th year	5th year	Total projects (Networks)
Water						4(6)
Waste						5(7)
Energy						2(2)
Other						2(2)

(B) *Cumulative yearly symbiosis for the next 5 years*

Resource	Current	1st year	2nd year	3rd year	4th year	5th year
Water	31	31	32	35	36	37
Waste	2	2	2	6	7	9
Energy	6	6	6	6	7	8
Other	0	0	0	0	0	2
Total	39	39	40	47	50	56

Table 8
Economic benefits through the additional industrial symbiosis network planned (Unit : billion won)

Participation	Target material	Content	Expected profit (year)
SK Co.	Spent caustic wastewater	Liquid incineration cost saving Replacement carbon source in sewage treatment plant	14.2 1.45
SK Chemicals Co.	Industrial wastewater Industrial water	Reuse of boiler water Line up 100,000 ton/day	15 —
Energy Co.	Combustible waste plastic	Incineration cost Product sale	3.7 5.17
Hanwha Corp. Ltd	VCM (material) Steam	Material reuse Reduction	0.11 0.09
SGR Tech. Co.	Petrochemical wastewater sludge	Processing cost saving via recycling Sale of biological resource	0.18 —
Total			39.9

exist in the supply chain. The location of Ulsan to the neighboring city, Kyeongju, an UNESCO world heritage site, is also considered to be a positive asset for this industry-nature symbiosis model. The call for a harmonious built-environment and natural ecosystem is very strong in the civil society. It is further strengthened with the local government's initiative of creating "Eco-Polis Ulsan". Aside from the principal component, the following areas could be explored in the process of developing the Ulsan EIP:

- Considering the large-scale business units in the region, there is sizeable opportunity to introduce environmental service providers into the region; this ignites another rise of business life cycle in the region.
- Considering the strong supply chain network of the anchor tenants, i.e. the two conglomerates; the network can evolve to include stakeholder interactions on the resource flow of the general industries in the complex, stand-alone industries around the complex and the industries within proximity of the region/eco-polis Ulsan.
- Information network should be in place to support the metabolic resource flow in the primary component.
- For the retrofitting component, green design of the existing complexes (or landscape ecology and green

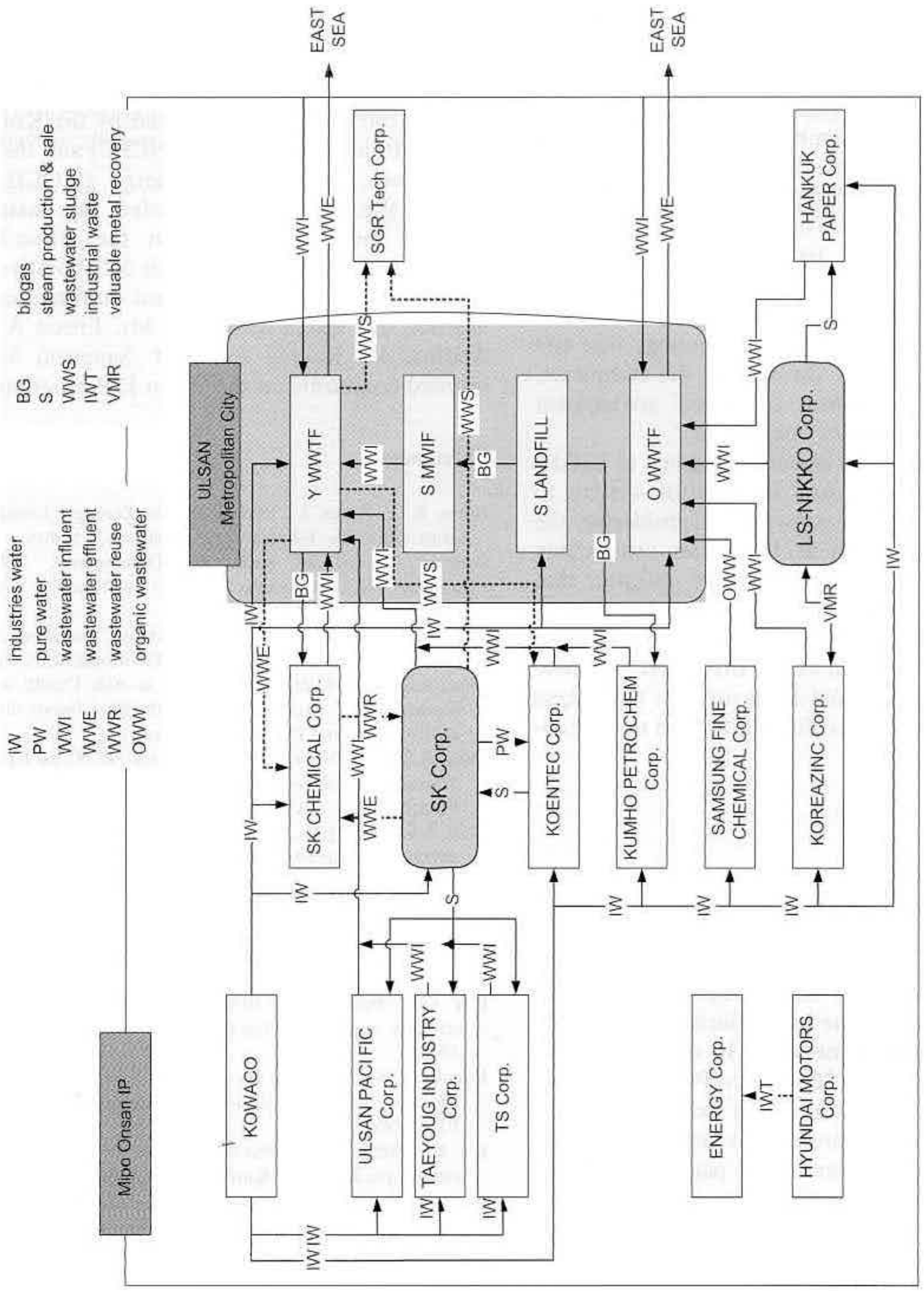


Fig. 3. The proposed EIP based on industrial symbiosis at Mipo, Onsan industrial complexes in Ulsan.

architecture) has to be planned in conjunction with eco-tourism development projects of the neighboring city, Kyeongju.

- Awareness and Preparedness for Emergency at Local Level (APELL) can be implemented in parallel to responsible care in petrochemical complexes. APELL is not a solution and not the beginning to the end, instead a stepping stone for integrating local community partnership programmes.

Propped by the existing symbiosis network and the gradually evolving sustainable development concept in the region, Ulsan industrial complexes are expected to be able to better materialize future opportunities by incorporating three success factors: (i) *economic incentives to participating industries*: reduction of cost or creation of new revenue through BPX practices, (ii) *innovation and technology diffusion*: new or existing technology that can substantiate resource synergy throughout the complexes, and (iii) *institutional supports*: continued government supports for efficient transition toward EIP.

Further research directions would be targeted to build a collective user-friendly data and information system to expand material and energy networks. Considering the diverse clusters of industries in Ulsan, an overarching performance metrics such as eco-efficiency indicator that includes productivity, emission reduction and societal prospects should be developed and implemented, both at the individual industry as well as the EIP levels. All these aforementioned strategies would endeavor after the success of integrating the presently envisioned EIP plan to the Eco-Polis Ulsan master plan.

4. Conclusions

Ulsan industrial complex was originally developed as conventional industrial estates, with a collective energy provider such as electricity, steam and water, and IS evolved due to stringent environmental regulations and economic benefits. Since mid 1990s, this network developed spontaneously on a one-to-one basis which resulted in the present industrial symbiotic network. In order to be the industrial capital of Korea, in 2004, "Eco-Polis Ulsan" was declared based on "The Master plan of Eco-Polis Ulsan", in which the Ulsan EIP Pilot project was included as one of the action plans. Ulsan eco-industrial park development can be achieved by expanding the existing BPX and eco-industrial network (EIN). Thus by considering the metabolism on resources and the stakeholder's interaction, cooperation and coordination, system analysis including industrial metabolism, input–output analysis, environmental evaluation and flexibility analysis must be conducted in detail for potential networking. In addition, Ulsan EIP project must be associated with the regional strategic environmental technologies and businesses to upgrade environmental technologies. To get the public support, education, publicity and leadership for Ulsan EIP are also highly required. Additionally, this developmental phase and lesson learnt hitherto, requires priority based planning system, rapid action with right intellectual support to understand the economics and managerial aspects of the EIP.

Acknowledgements

This project work was initiated through the research Grants (2005-B052-01) supported by the Korea National Cleaner Production Center (KNCPC) and the Ministry of Commerce, Industry, and Energy (MOCIE) of South Korea. We greatly acknowledge the management of different industries located in the Ulsan/Mipo-Onsan industrial park complex for their collaborative efforts and cooperation in sharing relevant information. We also mention our special thanks to Mr. Ernest A. Lowe, Mr. Andreas W. Koenig and Prof. Sangwon Suh for their inspired comments on the Ulsan EIP programme.

References

Ayres, R.U., Ayres, L., 1996. Industrial Ecology: Towards Closing the Materials Cycle. Edward Elgar Publishers, London.

Business Council for Sustainable Development, 1999. By-product Synergy: A Demonstration Project in Tampico, Gulf of Mexico, Altamira.

Chertow, M.R., 2000. Industrial symbiosis: literature and taxonomy. Annual Review of Energy and the Environment 25, 313–337.

Chiu, A.S.F., 2003. Strategy for EID in Asia Pacific: an overview for Korean NCPC. In: Proceedings of the First International Conference on Eco-Industrial Park, Seoul, Korea.

Chiu, A.S.F., 2005. Training Package for NCPC on CP/Environmental Management of Industrial Estates, Ver 3.3. UNEP/InWEnt, Paris, France. ⟨http://www.uneptie.org/pc/ind-estates/event⟩.

Chiu, A.S.F., Yong, G., 2004. On the industrial ecology potential in Asian developing countries. Journal of Cleaner Production 12, 1037–1045.

Cote, R.P., Hall, J., 1995. Industrial parks as ecosystems. Journal of Cleaner Production 3, 41–46.

Ehrenfeld, J., Gertler, N., 1997. Industrial ecology in practice—the evolution of interdependence at Kalundborg. Journal of Industrial Ecology 1, 67–79.

Esty, C.D., Porter, M.E., 1998. Industrial ecology and competitiveness: strategic implications for the firm. Journal of Industrial Ecology 2, 35–43.

Koenig, A.W., 2005. The eco-industrial park development—a guide for Chinese government officials and industrial park managers. EMCP, EU-China.

Lee, K., 2005. Establishment of eco-industrial parks for infrastructure of cleaner production in Korea. In: Presentation Manuals at the Korea National Cleaner Production Center, South Korea.

Lee, C., Koh, J., Jeong, H., 2005. Vision and changes of sustainable development: experiences of Korea. In: Proceedings of the International Workshop on Climate Change and Sustainable Development, Seoul, South Korea, pp. 23–37.

Lowe, E.A., 2001. Eco-industrial Park Handbook for Asian Developing Countries. A Report to Asian Development Bank, Environment Department, Indigo Development, Oakland, CA.

Lowe, E.A., Chiu, A.S.F., 2005. Consultant's Report on Korean EIP Project. Korea National Cleaner Production Center.

Lowe, W.J., 1996. The Source of Value: An Executive Briefing and Sourcebook on Industrial Ecology (PNNL-10943). Pacific Northwest National Laboratory, Richland, Washington.

Martin, M.A., Keuning, S., Janssen, D.B., 1998. Handbook on Biodegradation and Biological Treatment of Hazardous Organic Compounds, second ed. Academic Press, Dordrecht.

Ministry of Environment, MOE, 2003. Strategies for Sustainable Development of Eco-Industrial Park. Republic of South Korea.

Ministry of Environment, MOE, 2004. Green Korea 2004—Building an Eco-community for the 21st Century. Republic of South Korea.

Mirata, M., 2004. Experiences from early stages of a national industrial symbiosis programme in the UK: determinants and coordination challenges. Journal of Cleaner Production 12, 967–983.

Mirata, M., Emtairah, T., 2005. Industrial symbiosis networks and the contribution to environmental innovation: the case of the Landskrona industrial symbiosis programme. Journal of Cleaner Production 13, 993–1002.

Murata, M., Tsujikawa, M., Kawanishi, S., 1999. Oxidative DNA damage by minor metabolites of toluene may lead to carcinogenesis and reproductive dysfunction. Biochemical and Biophysical Research Communications 261, 478–483.

Nunez, C., 1998. VOCs: sources, definitions and considerations for recovery. In: EPA, Volatile Organic Compounds (VOC) Recovery Seminar, September 16–17, Cincinnati, OH.

Park, H.-S., Choi, S.-M., Lee, S.-Y., 2004. Pilot project for the development of eco-industrial park in Ulsan. In: Proceedings of the First International Environmental Symposium for Eco-Polis Ulsan, Ulsan, Korea.

Roberts, B.H., 2004. The application of industrial ecology principles and planning guidelines for the development of eco-industrial parks: an Australian case study. Journal of Cleaner Production 12, 997–1010.

Yang, P.P.-J., Lay, O.B., 2004. Applying ecosystem concepts to the planning of industrial areas: a case study of Singapore's Jurong Island. Journal of Cleaner Production 12, 1011–1023.

Journal of Environmental Management 87 (2008) 14–25

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Water stress, water transfer and social equity in Northern China—Implications for policy reforms

Ximing Cai

Ven Te Chow Hydrosystems Laboratory, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Received 5 May 2006; received in revised form 27 November 2006; accepted 31 December 2006
Available online 6 March 2007

Abstract

Water stress in Northern China is characterized with major, inefficient irrigation water use and rapidly growing non-agricultural water demands, as well as limited water quantity and declining water quality. Water use in the region is undergoing transfer from agricultural to municipal and industrial sectors. Currently, part of the economic loss and environmental damage due to water stress can be considered as a consequence of water transfer failures, including the current transfers, which hurt farmers' livelihood and income, and the needed transfers, which industry and cities have been waiting for but have not received. This paper starts with a discussion of the causes of water stress in Northern China, which is fundamental to understand the necessity and complexity of agricultural water transfers. Following that, it reviews water transfers in Northern China as a cause for concern over the social stability, economy and environment of the region. Based on an integrated analysis of economic, environmental, fiscal and social implications, this paper begins by identifying critical barriers to smooth water redistribution; and ends with implications for policy reforms, ensuring that farmers can and will save water. It is concluded that the decisions of water reallocation under water stress should be shared by communities at all levels, from the local to the national, to ensure equal access of water, especially the availability of the basic water need for all groups.

Keywords: China; Water stress; Water transfer; Social equity; Policy

1. Introduction

While China's rapid economic development is impressive, water stress in Northern China is a widely recognized crisis. Water stress occurs when the demand for water exceeds the available amount during a certain period or when poor quality restricts its use. One of the common indicators of water stress is per capita renewable water; $1000\,m^3$ per capita is recognized as a critical level for severe water scarcity (Engelman and LeRoy, 1993). The three major basins in Northern China, Huai, Hai, and Huang/Yellow (the 3-H basins) as shown in Fig. 1, have 520, 470, and $530\,m^3$ per capita, respectively, which are less than half of the water scarcity threshold (Cai and Rosegrant, 2005). Water stress in the region is also reflected in other indicators. In 1997 the 3-H basins contained about 40% of the country's agricultural land and produced 32% of the country's GDP, but had less than 8% of the country's water resources (CAE, 2001).

Water stress in Northern China has intensified water use conflicts between upstream and downstream areas and also between agriculture, which is still the largest water consumer, and the municipal and industrial sectors (M&I), which have been growing fast. Water stress has also caused deterioration of fresh water resources in terms of quantity (aquifer over-exploitation and dry rivers, etc.) and quality (eutrophication, organic matter pollution, saline intrusion, etc.). If left unresolved, such problems may worsen until they threaten social stability. The Chinese Government has initiated strategic changes to reform water resources allocation, especially to transfer agricultural water to other sectors. During the coming adjustment period, agriculture will collide head-on with industry, cities, and ecosystems. Agriculture faces the most difficult adjustment, in order to ensure that sufficient water is available for essential uses in cities and industries. The critical question is how to

E-mail address: xmcai@uiuc.edu.

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2006.12.046

Fig. 1. The 3-H Basins (Huai, Hai and Huang) in Northern China.

implement new allocation policies, which cause the least disruption to agricultural output and the livelihoods of farmers (especially poor farmers).

This paper reviews water transfers in Northern China as a cause for concern over the social stability, economy and environment of the region. Based on an integrated analysis of economic, environmental, fiscal and social implications, this paper attempts to begin by identifying critical barriers to smooth water redistribution; and ends with implications for policy reforms from a perspective of social equity. Social equity implies that the decisions of water reallocation under water stress should be shared by communities at all levels, from the local to the national, to ensue equal access of water, especially the availability of the basic water need (Gleick, 1996), for all groups. However, to understand the necessity and complexity of the agricultural water transfers, it is first necessary to discuss the causes of water stress in Northern China.

2. Causes of water stress in Northern China

The limited water availability due to low rainfall and runoff and the uneven temporal distribution has been recognized as a natural cause of water stress in Northern China (CAE, 2001). Research has also found that runoff generation in the region shows a declining trend during recent decades. Compared to the annual average during 1956–1979, average precipitation in the 3-H basins decreased by 9.6%, runoff decreased by 23.8%, and flow to the ocean decreased by 58.6% (CAE, 2001). At the regional scale, surface runoff generation within Hebei Province has declined by 65% and inflow declined by 72% compared to the average in the 1950s (Li and Wei, 2003). The impact of global climate change on water availability in this region needs further monitoring. As in many other regions in the world, the impacts of human activities on runoff generation include deforestation, agricultural and urban development, and groundwater overdraft. According to the assessment of the Chinese Academy of Science (CAS, 2005), in Northern China during 1988–1999, vegetation coverage (percentage of the total land) decreased from 38% to 32%, and urbanized area increased from 5.0~8.2% to 6.5~9.5% (varying over eastern, western and central parts). Around urban areas agricultural land has been converted to urban; on the other hand, in mountain areas, which were the primary areas of runoff generation, forest and grassland is being replaced by cropland due to growing population in rural area. This land use conversion reduces soil water storage and increases flooding.

Starting in the early 1950s, the practice of developing engineering facilities to catch and redistribute water has dominated water resources management in Northern China. Huge engineering efforts have been made to catch the precipitation and runoff in the region. In 1995, the reservoir storage of the 3-H basins was 1.6 times of their total annual runoff (Rosegrant et al., 2002). Water storage and water diversion and extraction engineering facilities can easily control all renewable water in Northern China. One can understand the scope of the stream flow regulation by imagining that gates crossing many rivers in this region are used to divert floodwater for groundwater recharge in the dry floodplains. Since no more water can be caught for use within the region, engineering now turns to trans-boundary water transfer projects. For example, the south–north water transfer (under construction), will transfer 38–48 billion km^3 of water (comparable to the total annual runoff in the Hai River and more than that of the Colorado River) from the Yangtze River in southern China to Northern China (mainly for the urban and industrial region in the Hai River Basin) over a distance of more than 1000 km.

Water resources engineering development in Northern China is impressive and has indeed supported growing water demands. However, it is widely believed that there is less potential today than before for engineering left within the region. Even the Ministry of Water Resources (MWR) has attempted to change the engineering-dominated water management. A new strategy has been proposed, called resource-based water management, which is an economic-theory approach centered on water demand management (Wang, 2003).

Indeed, the water stress problem in Northern China should be re-examined from a water demand management perspective given excessive and still growing demand due to population growth, food demand increase, and industrial and municipal development. According to the assessment of CAE (2001), between 1980 and 1999, water use in Northern China (withdrawal) has increased by 41.7 km^3 (CAE, 2001), which is about the average total annual runoff in the Hai River Basin, and 1.3 times of that in the Colorado River Basin. In 2004, the total water withdrawal in the 3-H basins was 137 billion m^3, which was 70% of the total annual runoff in the basins. Water demand, particularly M&I, will continue to increase (Hai River Conservancy Commission, 2005). According to a medium projection made by CAE (2001), in the 3-H basins, between

1997 and 2010, water demand will increase by 13.0 km^3 (8%); and between 1997 and 2030, 31.4 km^3 (20%). For the same region, Rosegrant et al. (2002) made a business-as-usual projection of the water demand increase between 1995 and 2025 as 49 km^3, close to the high projection of CAE. The continuous increase will make the water supply system more vulnerable.

Agriculture is the largest water user in Northern China. In the 3-H basins, agriculture was responsible for 84% of total water consumption in the region; the fraction declined in recent years but it is still over 75% (Rosegrant et al., 2002). Like in many other countries and regions, low irrigation water-use efficiency is accused to be responsible for water stress at present. According to the statistics in the Annual Book of Water Resources in China over 85% of the irrigation is done through overflow methods and open water channels (MWR, 2004). Starting from the middle 1990s, the Chinese government has greatly increased investment for irrigation system updating. However, the implementation of advanced irrigation technologies faces many technical difficulties (Henry, 2004) and social-economic blocks, which will be discussed later in this paper.

Moreover, many Chinese researchers believe that water pollution is the biggest cause of water stress, particularly Northern China (e.g., Jiang et al., 2004). Due to rapid industrial and municipal development and a large increase in agricultural fertilizer and pesticide use in the past two decades, cases of downstream users receiving polluted water from upstream lands have dramatically increased. During 1980–2004, sewage water discharge doubled in the 3-H basins, and it increased by 160% and 140% in the Huai and Huang River, respectively, according to the annual water resources bulletins of the 3-H basins (Huai River Conservancy Commission, 2005; Yellow River Conservancy Commission, 2005).

3. Water transfer and social equity

Under water stress, conflicts in water allocation and the damages to the public environment have intensified, and a pressing concern for "social equity" has appeared in Northern China. Social equity implies fair access to resources and livelihood; the concept of what is "fair" reflects the ethical values shared by the society, as well as economic values associated with resource uses. Social equity also reflects a principle that each citizen regardless of economic statuses or personal traits deserves and has a right to be given fair treatment by the political system, giving special attention to the needs of weak and vulnerable populations. In the context of resource allocation, social equity refers to a bundle of rights and duties of government, collective, and/or individuals, which are applied to protect weak and vulnerable populations in society. With social equity in water resources in mind, it is interesting to see that Chinese Government is promoting a national strategy called "building a harmonious society". Challenges of social equity in water allocation under aggravated water stress are worth great attention, since water is a basic resource for life and production in society, and is an essential part of any harmonious society.

3.1. Setoral water use conflicts

With water supplies in Northern China declining in absolute and quality-adjusted terms and water resources already over-allocated, one of the key questions facing water managers and residents will be how to meet growing demand in the M&I sectors. According to CAE's assessment based on various statistics, during 1980–1999, in Northern China, the fraction of agricultural use decreased from 84% to 73%, while the fraction of M&I water use increased from 22% to 28%. The assessment and projection of M&I fraction in the 3-H basins are shown in Figs. 2(a, b)

One clear opportunity to meet the growing M&I demand is through change in inter-sectoral allocations. In the competition for scarce water, local authorities generally give higher priority to industry than agriculture, while municipal and domestic uses receive the highest priority. This is because water used in industry has a much higher economic value, and China's local officials want to facilitate growth in industrial production more than agriculture. Hence, when there is a decision to be made on whether water should be sent to an industrial facility or

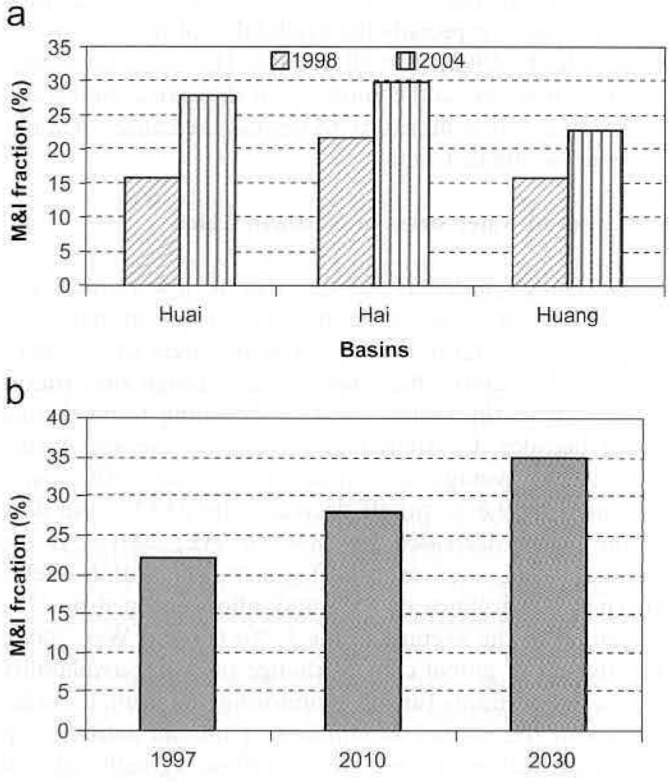


Fig. 2. M&I fraction in the 3-H basins: (a) individual basins in 1998 and 2004 (sources:bulletins of 3-H basins); (b) 3-H basins together in 1997 (assessment) 2010, and 2030 (projections) (*Source*: CAE, 2001).

kept for agriculture, industry often wins out (Lohmar et al., 2002). Sectoral transfers have already occurred within large basins, provinces, cities, or counties. Table 1 compares water uses by sector in two periods (1988–1992 and 2002–2004) in the Yellow River Basin. In the middle and lower parts of the basin, agricultural water use decreases significantly while industrial and domestic increases dramatically. The only increase of agricultural water occurs the upper basin, but the increase is relatively small. Table 1 also illustrates a common upstream–downstream water allocation problem. In the Yellow River, more than half of the runoff generated by rainfall comes from the upper reach and fewer people live there, therefore, upstream water users suffer less water stress. Furthermore, upstream provinces such as Ningxia and Inner Mongolia (autonomous region) are less urbanized and irrigation has a more important economic role than in the middle and downstream provinces. This explains why there is an increase of irrigation water use in the upper reach (Table 1). The double threat of continuous increase of water use in the upstream provinces and increasing local M&I water use has escalated the agricultural water shortage in downstream provinces. In 1987, the State Council of China issued the "Water Allocation Programme" along the Yellow River, which specifies water withdrawal quotas for each province along the basin. For example, the downstream Shandong Province can withdraw 7.0 km^3, about 19% of the total allowable water withdrawal in the basin. However, the actual water withdrawal by Shandong depends on how much those upstream provinces withdraw, as well as the runoff generated in a specific year. The Yellow River Conservation Commission (YRCC) has been responsible for the realization of this policy, but actual water allocation has not followed the rules until the policy was strengthened in most recent years to prevent flow cutoffs in downstream main channel. YRCC has also improved the monitoring system including monitoring facilities and institutions along the River.

Table 1
Yellow river water uses by sector (billion cubic meters), 1988–1992 and 2002–2004

Years	Reach	Total	Agricultural	Industrial	Domestic
1988–1992[a]	Upper	13.11	12.38	0.51	0.22
	Middle	5.44	4.77	0.38	0.28
	Lower	12.18	11.24	0.55	0.38
	Basin	30.72	28.39	1.45	0.89
2002–2004[b]	Upper	17.54	15.71	1.42	0.41
	Middle	5.71	4.16	0.97	0.58
	Lower	8.44	7.04	0.82	0.58
	Basin	31.69	26.91	3.21	1.57
Difference	Upper	34%	27%	179%	84%
	Middle	5%	−13%	155%	108%
	Lower	−31%	−37%	49%	54%
	Basin	3%	−5%	121%	77%

[a]Data from Chen (2002).
[b]YRCC Water Resources Bulletins of 2002–2004.

3.2. Agricultural water transfers

Agricultural water transfers have occurred in large irrigation districts of China. Hong et al. (2001) presented such an example with the Zhanghe Irrigation District located in northern Hubei Province. Although this area belongs to Central China, it is on the boundary of Northern China and represents typical agricultural water transfers in large irrigation districts in China. As shown in Fig. 3, there has been substantial reallocation of water from agriculture to hydropower generation and industrial and domestic uses over the past several decades, especially during the 1980s and 1990s. In late 1970s, agriculture covered 80% of total water use in the region; while around the middle 1990s, agriculture only used less than 25% of the water and non-agriculture used over 75%. This sharp decline in water supply for agriculture led to a 40% decline in irrigated rice area (Fig. 4). Although, the irrigated area decline is considerable (40%), it is much less than the decline in deliveries of irrigation water (about 67%). Moreover, the crop yield per hectare has doubled from the 1960s to 1990s (mainly due to the use of hybrid rice), and the yield per cubic meter of water supplied has tripled, which shows that there have been considerable water savings. A number of factors have contributed to the water saving, including new irrigation techniques (e.g., alternate wetting and drying at the farm level, canal lining, etc.), development of alternate sources such as rainfall harvest by small reservoirs, groundwater, and reuse of return flow, as well as crop pattern change from two to one crops of rice. It is believed that the sharp reduction of irrigation water has forced farmers to adopt cost-effective water-saving techniques. Meanwhile, the government investment in the improvement of irrigation systems and adoption of the new technology (e.g., hybrid rice) has played a critical role in maintaining the livelihood of farmers in the area. A major challenge is to identify those practices that could be successfully extended to other regions, both inside and outside China.

Large and medium irrigation districts in China often use reservoirs for flow regulation (i.e., storing water in wet periods and releasing water in dry periods) for irrigation water supply, and some of the reservoirs also have other functions such as hydropower generation and flooding control. Water transfers are often implemented through modification of reservoir operations. For the Zhanghe Irrigation District, reservoirs and diversion canals were originally designed for irrigation, but now larger priorities are given to hydropower generation and urban water supply. Consequently, reservoir release is driven by energy demand and domestic water use timing instead of crop water requirements. Such reservoir function and operation changes might have far-reached impacts if they are extended to other parts of China. In the summer of 2000,

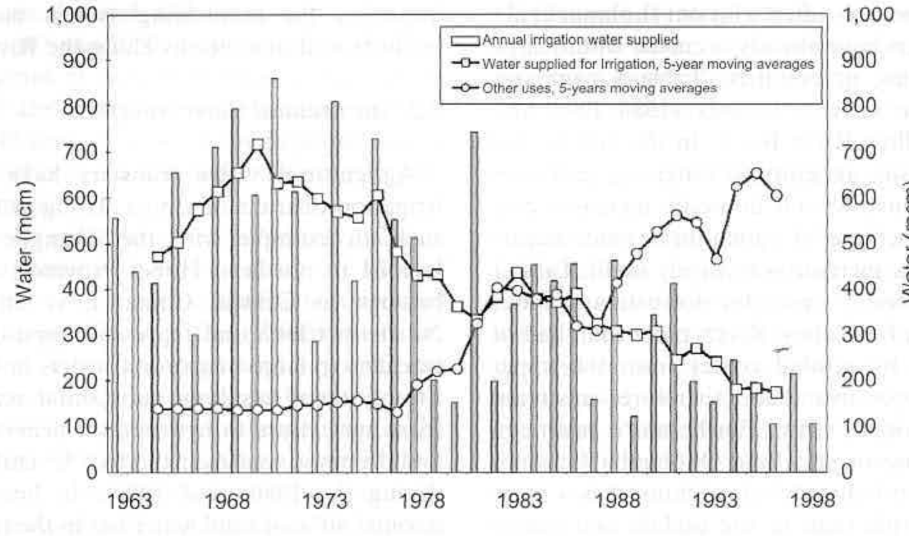


Fig. 3. Annual water allocations for irrigation and other uses (in the unit of million cubic meter [mcm]) Zhanghe Irrigation District, 1965–1998 (*Source*: Hong et al., 2001).

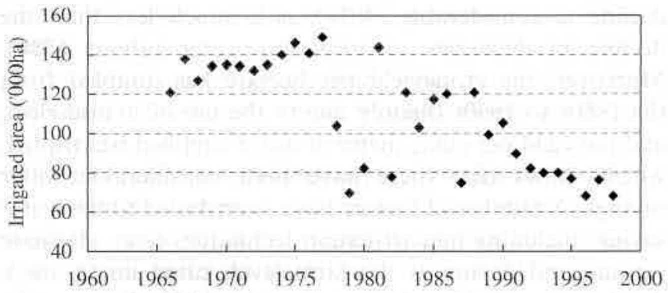


Fig. 4. Irrigated area (mainly rice) in Zhanghe Irrigation District, 1966–1998 (*Source*: Hong et al., 2001).

the author of this paper visited a county in northeastern Hubei Province, which was hit by a serious drought in that year. The county has a medium-size irrigation district fed by a reservoir. Farmers used to get water from the reservoir through canals when drought occurred. However, in that year, farmers did not get water for quite many days. Ponds for drinking water dried up in most villages. Farmers complained that the reservoir authority tried to hold water as long as possible for the newly developed tour business at the reservoir site, unfairly withholding the farmers' water.

Agricultural water transfers are even more needed and have been undertaken much more extensively around urban areas. A typical case about Beijing municipality and neighboring agricultural counties was described by Peisert and Sternfeld (2004). The administration of Beijing includes several agricultural counties such as Miyun, Huairou and Yanqing. To protect water quantity and quality supplied to the Beijing municipality, restrictions have been established for land and water uses and agricultural development in those counties. The cause of the urban–rural water conflicts sparked by Beijing's growing municipal water demands is the burden of restructuring the counties' agricultural sector. In 2001, the government of Miyun county announced it would completely abandon growing cereals, and instead develop perennial cultures, mainly fruit trees. Another major shift in the county's agricultural sector coming in 2001 was a complete ban on chemical fertilizers within the next 5 years, in favor of organic fertilizers. Miyun's land conversion and organic programs are in line with a general restructuring of Beijing's agricultural policy, as announced in March 2001: Grain-growing areas will be reduced to save ground water, and more trees will be planted. Animal breeding and 'highly efficient' agriculture with modern water-saving irrigation methods will be developed. It is predicted that the water used in agriculture will drop to 35% of the city's water consumption in 2010 from 43% in 1998, and the figure will continue to drop to 28–30% in 2020. Regulations lay out some compensation and alternative economic development strategies for the counties, including funds for cereal to fruit field conversion and organic programs. However, these have not yet been fully carried out. Peisert and Sternfeld (2004) argued that there is still a long way to go for the transition of land use and other agricultural practices to be completed, due to the current lack of an institutionalized compensation system to replace the current ad hoc downstream–upstream compensation payment occurring annually. Peisert and Sternfeld (2004) think that any format of payment in the system "must be transparent and include mechanisms to directly reward upstream farmers and cities for protecting the watershed".

Agricultural water use reduction also has other causes. For example, On the North China Plain, prolonged extraction of ground water for industry has greatly lowered the water table under many urban districts, which enforced farmers in and around these urban regions to draw their water from deeper and deeper wells (Lohmar et al., 2002). Another cause is the quality degradation of agricultural water caused by industrial wastewater discharge. Although industrial wastewater treatment capacity has grown tremendously in the past several years, in most cities a large portion of industrial effluents are still discharged directly into rivers. Surface-water is often too polluted to be used for irrigation. Following the Environmental Protection Law of China (approved in 1989), all provinces have established quotas of pollutant discharge in wastewater for urban communities and various industrial sectors. However, "local officials often sidestep legislation and regulations designed to curb such pollution in order to keep local industries profitable" (Lohmar et al., 2002). Effective inspection and monitoring is needed to ensure the full execution of those regulations.

There is evidence that agricultural water transfers requested by local industry were denied despite the fact that industry is generally a higher value use (Lohmar et al., 2002). Upstream irrigation districts that have built their own reservoirs and canal systems have little incentive to provide water to industrial centers, since their own agricultural activities would be adversely affected. For example, several major cities in Henan Province failed to access water from reservoirs controlled by local agricultural authorities; as a result they had to shut down some factories even as rice production was expanded in the region. These conflicts usually represent the divergent goals and interests of the urban construction authorities versus the local water resource bureaus and agriculture bureaus.

Finally, it is noticed that several large-scale agricultural water transfers have been undertaken to restore ecological systems in Northern China. The most high-profile example occurs with the Yellow (Huang) River. Due to excessive river diversion and a prolonged drought cycle, flow cutoffs had been experienced in the main channel downstream from 1972 to 1998. Flow in the main channel was cutoff in 21 of 30 year with both the duration of time and the distance from the river mouth increasing each year. The flow interruption left users in Shandong and Henan Provinces without their traditional sources of surface water, and more seriously, it precluded sediment flushing to the ocean and threatened the downstream ecosystems (Li, 2002). Starting from the year 2000, the Yellow River Conservation Commission (the basin management authority under MWR) has undertaken firm administrative measures to prevent the river cutoffs, including water withdrawal monitoring and more strict execution of the "Water Allocation Programme" with the provinces in the upstream of the basin. The status of the river has since been improved, and there has been no absolute flow cutoff in recent years. Meanwhile, irrigation water withdrawal by the upstream provinces has declined. According to a news report from MWR (MWR, 2005), irrigation districts in Ningxia reduced water withdrawals by $2.2\,km^3$ during 2001–2005, which is about one-quarter of the total withdrawal by Ningxia in 2000.

Another example of agricultural water transfer to ecosystem has occurred in the Hei River Basin, an inland basin in northwestern China (Liu et al., 2005). The Hei River flows from the upstream Gansu Province to downstream Inner Mongolia region, where the river water feeds the Erjina Oasis and traditional pasturelands around the Oasis. Starting from the early 1980s, large expansion in irrigation water diversion in the middle stream of the basin has resulted in downstream ecosystem degradation including a shrinking oasis area, declining vegetation coverage, and a complete dry-up of the outlet lake. The environmental change threatens the homeland of more than one million people who depend on the pastureland for livestock breeding. In the late 1990s, the Chinese government initiated activities for the ecological restoration of the region, including annual flow releases to downstream by administrative order. These activities followed institutional development including activating the "Gansu-Inner Mongolia Water Allocation Plan (GIWAP)" approved by State Council and establishing the Bureau of Hei River Management, an agency that is authorized to coordinate and supervise the implementation of the new water allocation plan. These releases have prevented further degradation of the downstream ecosystem. Meanwhile middle stream irrigation water withdrawal is required to decline by one-quarter, which has led to significant crop pattern change and also stimulated irrigation system updating with financial support from the central government (Liu et al., 2005). In Zhangye City, the major agricultural district in the middle stream of the Hei River Basin, the shares of agricultural land were 65% cropland, 16% forestland, and 19% pastureland in 1999; and recently the shares were adjusted to 52%, 26%, and 22%; during 2000–2003, respectively; canal lining was improved for 585 km of main canals and 1160 km of branch canals, and advanced irrigation systems such as low-pressure pipe, drip, and sprinkler systems were established for 29,000 ha (10% of the total irrigated area in the district). Moreover, in 2000 the MWR approved an experimental program for water conservation in Zhangye, which is the first program for testing comprehensive water demand management measures in China. The program is intended for determining both primary water use rights and water use permits and facilitating both water trade and public participation in water allocation.

3.3. Dilemmas in water transfers

Given the growing supply/demand imbalance in Northern China, it will be increasingly difficult if not impossible to meet new water demands from one sector without decreasing supplies to another. Since agriculture is now by far the largest consumer of water resources in the region and appears to have relatively low economic output levels, meeting growing industrial and domestic demands is likely going to mean a reduction in supplies to the agricultural sector. However, agriculture in China, as in many other countries, holds a special place in rural livelihoods and national food security, so it remains a high priority. Furthermore, the rural agricultural sector that is most impoverished and has probably benefited least from recent economic growth. Shifting water away from those already relatively disadvantaged has clear implications for equity and, perhaps, social stability. At the same time, it is industrial growth, which is dependent on increasing water supplies, which is seen as the driving force in powering China's transformation to a modern, world-class economy.

The premise for successful water transfers out of agriculture is that farmers can use less water while food production is not affected. There is no doubt that advanced irrigation technology is needed to replace traditional gravity systems. Starting from the mid-1990s, the Chinese government has significantly increased investment for irrigation system updating, and low-flow irrigation such as drip and sprinkler is now gaining ground in China (Henry, 2004). However, the implementation of advanced irrigation technologies means a higher cost for water, and it is up to farmers to pay at least part of the cost. The question is whether farmers can afford the cost of advanced irrigation technology, and if farmers will and can pay a water price that is closer to the true value of water. The answer might be no, given the current low income of farmers due to low food prices and high costs of agricultural inputs such as fertilizer. If water prices cause lower or negative net profits from irrigated agriculture, it may force farmers simply to use less water and give up high crop yields.

There is a "dead lock" with agricultural water management in China, which blocks effective agricultural water savings and smooth water transfers. The dead lock exists with the following conflicts, which are inter-related to each other: (1) farmers are wasting water even as they suffer water shortage themselves and leave increasing industrial and municipal sectors thirsty (Henry, 2004); (2) the society needs agricultural water savings, but farmers may not afford the costs of water savings and because of low profits from crops their willingness-to-pay for water is much lower than the true water value; (3) the national policy of food self-sufficiency requires the maintenance of high crop yields and production that depend on irrigation, even when the government can not maintain the required water for agriculture given the growing requirements of non-agricultural sectors. China must break this dead lock, which will continue to damage the country's growing economy.

Difficulties also exist with the governance of water transfers in China. Current water transfers follow the priorities of local government that are often driven by short-sighted economic profits. There is no formal institution for mediation, so large water disputes between agriculture and other sectors (and also between regions) often are resolved by the government in an ad hoc, case-by-case fashion. Ideally, water transfers would go through markets based on a deal made by both water buyers and sellers and without negative social impacts (Rosegrant and Binswanger, 1994). The establishment of water rights is the pre-request of water markets. In recent years, researchers and government officers begin to embrace the modern concept of water right and water market (e.g., Wang, 2003; Liu, 2004). In China, water right is a kind of usufruct right and only the nation has the property right of natural resources. There is no clear delineation of the usufruct right, formal and consistent institutional support for water rights does not exist, and thus no mechanism for economically efficient water rights trading exists. Although, the China Water Law approved in 1988 and amended in 2002 provides codes for water allocation and use, current water use rights, both their identification and supervision, are heavily manipulated by the powerful water bureaucracy. Conflict resolution related to water rights is usually handled by administrative measures as individual cases, following some government guidelines such as the Implementation Measures for Water Use Permit Systems (issued by the State Council in 1993). This system, which lacks a consistent legislative framework, may not guarantee the security of water rights and is subject to high transaction costs, as well as social–economic loss with the two sides involved in water transfer. Furthermore, the current water use permit for agriculture is defined as a collective right for farmers in a town or even in a county, and individual farmers need to share the water right within their group. This may add to the difficulty of water right management and affect the efficiency of water right in water allocation. This situation has given rise to grass-root institutions in China, called water users associations (World Bank, 1997). Moreover, in recent years, in rural regions of the country, individual farmers are involved in small pilot trials of water right transfers, which are different from sporadic trades at the village level. The purpose of those experiments is to develop formal and institutionalized in nature from the bottom level to higher levels (Turner and Hildebrandt, 2005). Several specific examples of the water right trade developments are discussed by Turner and Hildebrandt (2005) and Liu (2004), which will be summarized later in this paper. Local government is involved in the initial market development, but the challenge remains to build a bottom-up mechanism that allows farmers at the root level to be actively involved.

3.4. Consequences and risks in water transfers

On one hand, farmers lose water that they once used without appropriate compensation; on the other hand, cities and industry are thirsty and expect to access more water. Both situations have resulted in significant negative consequences to the regional economy and environment. Huge losses of food production and economic losses in industry due to water shortage have been reported every year by national and provincial statistics. According to the estimate by MWR, food production loss in the country is 25–30 million tons, and economic loss in industry is about 200 billion RMB. For the future, risk of even more serious socio-economic and environmental loss is apparent if current practices continue (Jiang et al., 2004).

The government has made various efforts in adequate and/or equitable water supply across the country, particularly in compensating farmers who have to reduce irrigation water use significantly and in assisting the agricultural areas to undertake new development strategies. Examples include the compensation to farmers in the middle stream of the Hei River Basin (Liu et al., 2005) and farmers in Miyun County, Beijing (Peisert and Sternfeld, 2004). However, whether the compensation is sufficient to maintain farmers' income and livelihood and whether it can be fully carried out, remains a concern. It has been noticed that farmers in Northern China have been struggling to reduce the income loss due to water shortage. One effort is using wastewater for irrigation. Irrigated area using waste water increased from 42 thousand hectares in 1963 to 3.6 million hectares in 1998, of which the major is in Northern China, surrounding large cities such as Beijing and Tianjin (Li et al., 2005). Using waste water for irrigation has brought up a high risk of soil contamination (CER, 1999). Moreover farmers sell most of the agricultural products such as vegetables and fruits irrigated by wastewater to urban consumers, which closes the "pollution cycle" between cities and nearby agricultural areas.

Another effort for farmers to reduce food production loss is to seek help from private entrepreneurs to use deep wells in order to restore irrigation (Lohmar et al., 2002). Obviously such approaches are neither sustainable nor adequate regarding public health. When there is no way to sustain irrigation, the problem is not only with farmers' income and the way they depend for life, but also, more seriously, with the national food production target. China never wants to see a large reduction of food production in Northern China as predicted by Brown (1995).

Moreover, drinking water supply in rural area of Northern China is connected with the irrigation systems. According to Li et al. (2005), water quality for drinking water in China is worsening due to reasons including excessive use of chemicals. In those regions where irrigation water is polluted or is largely reduced in quantity, drinking water will be affected in both quantity and quality. The most notorious case is the extremely high cancer rate in a number of villages located along some tributaries of the Huai River. Farmers have used the rivers running around their villages for drinking and irrigation hundreds of years, but now they find that not only the water directly taken from the rivers, but also the water pumped from wells nearby the rivers, is poisonous (news report, visiting a cancer village in the Huai River Basin, http://www.people.com.cn/GB/14576/33320/33321/33754/2919537.html, in Chinese). One could not question the importance of industrial water users, which have contributed to local economy; however, industrial water use that ignores the existence of downstream farmers could deplete the basic needs of clean, healthy water for the farmers, giving the fact that currently the farmers do not have any treatment facilities for drinking water supply.

When effective water transfer does not occur under water stress, water stress usually causes deterioration of both water quantity and quality. It is not difficult to understand why instream flow and deep groundwater has been depleted and polluted in Northern China given the following facts: although agricultural water savings have not yet been implemented, part of the irrigation water has already been transferred, and although M&I has received some water from agriculture, it is still not fully supplied. Therefore both sides, to sustain irrigated agriculture under the current irrigation technology level and to maintain fast growth of economy, drive excessive water withdrawal from rivers and aquifers. For local governments, the strong wish for economic development always overrides the responsibility of environment protection.

So far in Northern China the environment, not the farmers, has borne the largest sacrifices from water stress. From the limited text of this paper, one sees the drying Yellow River, the second largest river in China, the nearly "closed" Hai River (little flow discharge to the ocean), groundwater depletion in the North China Plain, and ecosystem degradation in the downstream of the Hei River. In addition, for many rivers, there is not enough water available for pollutant dilution, which has partially caused the water quality problem, and groundwater overdraft in coastal areas has caused seawater intrusion. Hebei Province, located in north of the Yellow River and covering a major part of the Hai River Basin, is a typical case of the changed water environment in Northern China. According to Li and Wei (2003), the province changed itself from a water-rich region in the 1950s to a water-poor region at present. Fifty years ago, Hebei had perennial rivers with over 3000-km long navigation channels, large lakes including the well-known Baiyangdian Lake, and widely distributed wetlands, and the region suffered frequent disasters of flooding and waterlogging (associated with land salinization). Today, rivers are dry most of the time, lakes and wetlands have shrunk and even disappeared (including the Baiyangdian Lake, "pearl" on the land of North China), and the region suffers damages caused by water shortage every year. Such a change is caused by excessive water consumption, as well as regional climate variability and environmental change.

4. Implications for policy reforms

Some choices must be made. Even if initially painful, properly conceived policies executed correctly can bring tremendous long-term benefits. Since farmers are the least able to cope with the policy changes needed to rationalize water supply, the key solution is then to help farmers to get ready for the changes, which essentially means that farmers *can* and *will* save water. First of all, it is important to replace traditional low-efficiency irrigation systems by advanced irrigation systems, which has been occurring since the 1990s (Henry, 2004). This will strengthen the physical feasibility for farmers to save water while maintaining production. The concern is whether the traditional focus on water supply engineering during the past 50 year can be switched to water management centered on efficient water uses (Boxer, 2001). Second, farmers may not volunteer to save water unless they must pay higher water prices that at least partially reflect the costs of irrigation system updates and the economic value of water. The economic value of water is much higher when water is allowed to transfer between agricultural and other sectors than when water use is constrained within agriculture (Briscoe, 1997). Currently what farmers can afford for water is far below the true economic value of water in Northern China. Under their current income level, farmers are the least able to cope with the price hikes to rationalize water supply. Therefore, agricultural water saving in China is not only a technological problem, but also a socio-economic problem, which is related to national and regional agricultural policies and markets and even the international food markets. For policy makers in China, efficient agricultural water use needs to be considered within the framework of the nation's gross economy. A research question is that at what level of income, farmers' willingness-to-pay to water will match the economic value of water assessed from the gross economy of Northern China. Starting in 2006, the Chinese government stops levying agricultural taxes. The impact of this change will be positive and broad although it will take time to emerge. It is expected that such measures, if they are stable, will enable farmers to pay a reasonable price for water so that water prices can become an effective economic incentive for water saving.

Third, an institutional establishment is needed to allow farmers and other groups to fairly exchange water. The role of government in water transfers has been demonstrated in this paper. At least for now and in the foreseeable future, administrative means of water allocation will take the lead with other means such as water markets as a supplement (Turner and Hildebrandt, 2005). Appropriate governance is still needed to guide and manage the ongoing water transfers. In particular, the mechanism is expected to provide a negotiation and mediation mechanism for officials and representatives from competing sectors to resolve the conflicts, and to protect the benefits of all, especially the weak groups. Many provinces and municipalities are promoting reforms to merge the functions of different water management units into a single authority, called Water Affairs Bureau (WAB). At the root level of irrigation management, the government has been sponsoring the development of water users associations (WUA), who not only take charge of water allocation among individual farmers within a WUA, but also represent a group of farmers (World Bank, 1997).

Fourth, delineation of secure and consistent water rights for various water users will be the basis for equitable water transfers among farmers and other groups. According to Wang (2003), implementing water rights and water markets seems to be a long-term target for the so-called "resource-based water resources management" being promoted by the MWR in recent years. Although formal water markets are still in the experimental stage in China, they are increasingly playing an important role to solve water conflicts (Turner and Hildebrandt, 2005; Liu, 2004). Water researchers and managers in the western world might be curious about the rapid spread of the western economics based water management approaches in China, which have even not been very successful in their countries. Given the extensive studies and practices sponsored by the government of China, the results and impacts might be worth of attention as time goes on. The involvement of government and administrative institutions in the markets under development is particularly interesting. There are some demonstration examples for market-based water allocations, which have been widely posted and discussed in China (Liu, 2004). A typical example of agricultural water right transfer to M&I in China has occurred in areas along the Yellow River, the Inner Mongolia Autonomous Region and Ningxia Hui Autonomous Region. In these regions, agriculture uses more than 95% of the water, with the aged irrigation infrastructure resulting in lower water use efficiency. During the past decade, as many other regions in China, there has been growing M&I demand under rapid economy development, and the total water usage in these regions has reached their allocated quota determined by Water Allocation Programme applied to the main channel of the Yellow River. Therefore, water transfer becomes a critical issue for the economic development in the region. To guide, formalize, and promote water rights transfer, the MWR issued the Guidelines for Tentative Water Rights Transfer in Mongolia Autonomous Region and Ningxia Hui Autonomous Region along the Yellow River Main Stream. Following the Guidelines, the Yellow River Water Resources Commission formulated the "Yellow River Water Rights Transfer Management Procedures (Tentative)," particularly applied to the upstream regions. The procedures include guaranteeing the domestic needs, matching the basic ecological water requirements, adapting advanced irrigation systems, and realizing compensations. By May 2004, eight large industrial projects in the two regions have signed agreements of intent for water rights transfer, with a total compensation of 360 million RMB.

Other transfers have occurred between administrative regions. In Zhejiang province, water use right trading was introduced between two cities, Dongyang and Yiwu in 2002. Dongyang is relatively rich in water resource while Yiwu is short in water supply. The two cities, through consultations, signed an agreement on transferring the water use right of Hengjin Reservoir of Dongyang to Yiwu at an agreed-upon price. Another example is with the Zhanghe flowing through Shanxi, Henan and Hebei. The three provinces, have settled their dispute over water allocation through an agreement. Changzhi City in Shanxi province, located in the upper stream of the river, agreed to supply water from its five reservoirs to Hebei and Henan at a mutually acceptable price. Through these demonstration transfers, the involvement of government as a regulator is a sign that local governments start to build up more formal markets based on the informal markets, which have existed for long time. On the other hand, it also shows the important role of government in the water transfers between sectors and between regions. Local governments usually act as buyers or sellers in the trades between two administrative regions. Although water markets are certain to develop further in China, there is no doubt for people both inside and outside China that water markets will simply supplement the administrative allocation methods in the foreseeable future before consistent water rights are mature in China.

Fifth, given the current serious situation of environmental flow depletion, it is critical to restore, at least partially, environmental flow to prevent irreversible environmental disasters. Facing the risk of environmental damage, China is beginning to recognize the spirituality of "human harmony with nature," which is rooted with Chinese historical ethics of Confucius. This is a philosophy emphasizing education, mastery of natural phenomena, discipline, and social harmony, as well as Taoism, a philosophy of non-interference (Li et al., 2004). As illustrated by the environmental flow restoration in the Yellow River and the Hei River, projects have been initialized in Northern China. Although current measures under administrative orders impose large transaction costs, they show some positive signs for the restoration of environment. The challenge will be to convert the short-term emergency measures into long-term, sustainable water management and regional development plans (Liu et al., 2005). To balance environmental flow requirements and offstream water consumption will be a long-term strategy in Northern China from generation to generation.

Finally, there will be some challenges for national agricultural policies facing unavoidable agricultural water transfer. A major tension behind the water stress is the premise of food self-sufficiency in China, which is important for China with a population of 1.4 billion, but also for the entire world given the concern of Brown (1995)—"who will feed China?" However, there is no doubt that with the growth of China's economy, the solutions to the problems related water and food must be flexible. When there is no way to hit all targets of food production, environment restoration and non-agricultural water demand, a compromise solution must be found. In a preliminary study, Cai and Rosegrant (2004) explored some strategies for sustainable groundwater use in Northern China given the fact that groundwater depletion is already a serious problem in the region. It was found that to reduce the current groundwater use to a sustainable level by 2025, food production in the Hai and Huang (Yellow) River Basins would decline by 15% and 9%, respectively. Although improvements of water use efficiency directed at overdrafting basins could, in theory, compensate the food production declines, they would be unlikely in reality due to technology constraints. In light of this fact, will it be possible for the Chinese government to reconsider the food self sufficiency policy, taking into account food production conditions in China, food security, environmental sustainability, as well as international food markets? This paper concludes with a potential alternative that could eliminate the growing water stress in Northern China and provide smooth water transfers between agriculture, thirsty cities, industry, and the dried-up environment. It should be taken as a research suggestion, not a policy to challenge the national food self-sufficiency policy. The international community is now exploring the policy implications of "virtual water flow" (Allan, 1998), which basically means water is transferred from one country or region to another with food export/import. The concept of virtual water links a large range of sectors and issues that revolve around relieving pressures on water resources, ensuring food security, developing global and regional water markets. In the future, higher international food demand may increase food commodity prices, which will hurt poorer countries that must import food. The ideal result of a global strategy of virtual water flow is that more food will be produced in those countries with enough water for crop growth, so that the impact of higher demand (increasing prices) will be balanced by the impact of higher production (decreasing prices) in the international market. Hoekstra and Hung (2002) identified China as one of the top ten countries that import virtual water. Regarding such an international virtual water market, China faces both challenges and opportunities. China's recent ascension to the World Trade Organization may provide larger opportunities to begin discussing such options. A better understanding of the role of water in economic growth in Northern China, the trade-offs between equity and efficiency in sectoral water allocation, and the range of possibilities for institutionalizing water allocation decisions would serve to better inform upcoming critical decisions on inter-sectoral transfers.

5. Conclusions

Water stress in Northern China is characterized with major, inefficient irrigation water use and rapidly growing non-agricultural water demands, as well as limited water quantity and declining water quality. Water use in the region is undergoing transfer from agricultural to M&I sectors. Currently part of the economic loss and environmental damage due to water stress can be considered as a consequence of water transfer failures, including the current transfers which hurt farmers' livelihood and income, and the needed transfers, which industry and cities have been waiting for but have not received. Successful water transfers in Northern China imply that farmers use less water to produce even more food. However, currently farmers cannot afford the costs of water savings and their willingness-to-pay to water is far below the economic value of water assessed from the gross economy of Northern China.

Policy implications for successful water transfers include the improvement of irrigation systems, the use of feasible water prices, establishment of effective water management institution, delineation of secure and consistent water rights, restoration of environmental flow, and reconsideration of national agricultural policies. These long-term planning strategies allow people in China to take time to perform trial-and-errors. Due to the strong water bureaucracy in China and the need of capacity building in water management, a "smooth" water transfer will take a long time—although some gradual progress can be expected. Whereas some actions are needed right now to avoid serious social instability. First of all, the basic water needs, including those for drinking and ordinary living (Gleick, 1996), should be protected firmly for farmers and also for residents in cities. For this purpose, some transfers that deplete farmers' water below their basic needs should be prohibited, and for the same purpose, some transfers must be undertaken to ensure the basic needs of residents in cities during drought periods. It will be necessary for government agencies to conduct careful monitoring and quick responses to water transfer events that impact the basic needs of humans.

Acknowledgments

This work with this paper is funded by the United Nations Development Programme (UNDP). The author is grateful for the helpful discussions and suggestions from Dr. Daniel Coppard at the Human Development Report Office, UNDP.

References

Allan, T., 1998. Moving water to satisfy uneven global needs: trading water as an alternative to engineering it. International Commission for Irrigation and Drainage (ICID) Journal 47 (2), 1–8.

Boxer, B., 2001. Contradictions and challenges in China's water policy development. Water International 26 (2), 335–341.

Briscoe, J., 1997. Managing water as an economic good: rules for reformers. In: The Paper Presented to the International Commission on Irrigation and Drainage (ICID), Oxford, UK, 25pp.

Brown, L., 1995. Who Will Feed China? Wake-Up Call for a Small Planet. W.W. Norton/Worldwatch Institute, New York, 160pp.

Cai, X., Rosegrant, M.W., 2005. Water management and food production in China and India: a comparative assessment. Water Policy 7, 643–663.

Cai, X., Rosegrant, M.W., 2004. In: Tao, C.T., et al. (Eds.), China's Water and Food: Future Perspectives in China's Agricultural in 2050. Chinese Agricultural Press, Beijing, PRC.

Chen, Z.K., 2002. Impact of the successive drought on the North China Plain ecological environment. Working Paper of the China Institute of Water and Hydropower Research (in Chinese).

China Environmental Review (CER), 1999. The Development of Soil/Groundwater, Quarterly Newsletter for China Environment, Health & Safety Managers, Winter Issue 〈http://www.environmental-expert.com/magazine/aer/china/winter/article2.htm〉, assessed in December 2005.

Chinese Academy of Engineering (CAE), 2001. Water Resources Assessments and Supply and Demand Projections, vol. 2: Reports on Water Resources Development Strategies/Sustainable Development in China, Beijing, China.

Chinese Academy of Science (CAS), 2005. The Impacts of Human Activities on Droughts in Arid Regions 〈http://pd973.tea.ac.cn/download/middle/kt4.pdf〉, assessed in December 2005 (in Chinese).

Engelman, R., LeRoy, P., 1993. Sustaining Water: Population and the Future of Renewable Water Supplies. Population and Environment Program, Washington, DC.

Gleick, P.H., 1996. Basic water requirements for human activities: meeting basic needs. Water International 21 (2), 83–92.

Hai River Conservancy Commission, 2005. 1998–2004, Hai River Water Resources Bulletins 〈http://www.hwcc.com.cn/haiwei/static/szygb.asp〉, assessed in December 2005 (in Chinese).

Henry, E., 2004. Water scarcity in the North China Plain: water saving irrigation and its implications for water conservation. Working Paper. School of Forestry, Michigan State University.

Hoekstra, A.Y., Hung, P.Q., 2002. Virtual water trade: a quantification of virtual water flows between nations in relation to international crop trade. Research Report. IHE, Delft.

Hong, L., Li, Y., Deng, L., Chen, C., Dawe, D., Barker, R., 2001. Analysis of changes in water allocations and crop production in the Zhanghe Irrigation System and District, 1966–98. In: Proceedings of International Workshop on Water-saving Irrigation for Paddy Rice, Wuhan, China, 23–25 March 2001.

Huai River Conservancy Commission, 2005. 1998–2004, Huai River Water Resources Bulletins 〈http://www.hrc.gov.cn/index_zhyw.aspx?LM_ID=LANMU_SZYGB〉, assessed in December 2005 (in Chinese).

Jiang, R., et al., 2004. A report water scarcity in China—the emerging dead lock of economic development. May 8, 2004, China Business Weekly, Beijing, China.

Li, B., Wang, D., Zhu, Y., 2005. Causes and Countermeasures of Water Environment Deterioration in Rural Regions of China. Chinese Hydraulic Engineering Society (in Chinese).

Li, G.Y., 2002. Regulation of water flow and sediment flushing in the downstream of the Yellow River. Chinese Water Magazine 474 (12) (in Chinese).

Li, R., Beek, E.V., Gijsbers, P., 2004. Integrated water resources management for the Yellow River in China: discussion on scientific and ethical approaches. In: The Proceedings of the International Association of Hydrological Sciences Symposium: The Basis of Civilization—Water Science? IAHS Press.

Li, Z.Q., Wei, Z.M., 2003. Water resources changes in Heibei Province and implications for the development of Western China. Science and Technology in South–North Water Transfer Project, No. 1.

Liu, B., 2004. Institutional design considerations for water rights development in China, in water rights reform: lessons for institutional design. In: Randolph Bruns, B., Ringler, C., Meinzen-Dick, R. (Eds.), International Food Policy Research Institute (IFPRI), Washington, DC.

Liu, H., Cai, X., Geng, L., Zhong, H., 2005. Restoration of pastureland ecosystems: a case study of western Inner Mongolia. Journal of Water Resources Planning and Management 131 (6), 420–430.

Lohmar, B., Wang, J.X., Rozelle, S., Huang, J.K., Dawe, D., 2002. China's Agricultural Water Policy Reforms: Increasing Investment, Resolving Conflicts, and Revising Incentives. Market and Trade Economics Division, Economic Research Service, US Department of Agriculture. Agriculture Information Bulletin, No. 782.

Ministry of Water Resources of China (MWR), 2004. The Annual Book of Water Resources in China. China Water Press, Beijing, China (in Chinese).

Ministry of Water Resources of China (MWR), 2005. News release 〈http://www.mwr.gov.cn/index/20060103/63451.asp〉, assessed in December 2005 (in Chinese).

Peisert, C., Sternfeld, E., 2004. Quenching Beijing's thirst: the need for integrated management for the endangered Miyun reservoir. China Environment Series 7, 33–45.

Rosegrant, M., Binswanger, H., 1994. Markets in tradable water rights: potential for efficiency gains in developing country water resource allocation. World Development 22 (11), 1613–1625.

Rosegrant, M.W., Cai, X., Cline, S., 2002. World Water and Food to 2025: Dealing with Scarcity. The International Food Policy Research Institute, Washington, DC.

Turner, J., Hildebrandt, T., 2005. Navigating peace: forging new water partnerships US–China water conflict resolution water working group. China Environment Series 7, 89–98.

Wang, S.C., 2003. Resource-Based Water Management—Human Harmony with Nature. China Water Press, Beijing, China.

World Bank, 1997. At China's Table. World Bank, Washington, DC.

Yellow River Conservancy Commission, 2005. 1998–2004 Yellow River Water Resources Bulletins, YRCC Website 〈http://www.yrcc.gov.cn/〉. December 2005 (in Chinese).

Journal of Environmental Management 87 (2008) 26–36

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Risk and opportunity in upgrading the US drinking water infrastructure system

Jeffrey W. Rogers*,[1], Garrick E. Louis[2]

Department of Systems and Information Engineering, SEAS, University of Virginia, Olsson Hall, 151 Engineers Way, Charlottesville, VA 22903, USA

Received 17 February 2006; received in revised form 12 December 2006; accepted 1 January 2007
Available online 27 March 2007

Abstract

This paper presents a practical risk assessment methodology to provide drinking water infrastructure (DWI) decision-makers with an objective risk assessment tool. The purpose of this risk assessment tool is to maintain the desired level-of-service or systems reliability [$r(f)$], while managing the financial uncertainty of the expected budgetary impact within the capital improvement program (CIP). The goal of this paper is to demonstrate the value of an objective risk assessment tool for estimating the DWI decision-maker's sensitivity to the risk of systems failure (R). The objectives are to: (1) incorporate probability of systems failure [$p(f)$] into the CIP budgetary analysis process and (2) evaluate the affects of $p(f)$ on the expected CIP budgetary outcome. The magnitude of the expected budgetary impact is managed through the DWI decision-maker's sensitivity to R, which is represented by the level of the rate of reinvestment (RR). The expected result of the proposed risk assessment tool demonstrates that by proactively managing R to maintain a desired $r(f)$ will effectively manage the impact of uncertainty on the expected budgetary outcome within the CIP. The expected contribution of the practical risk assessment methodology is to provide DWI decision-makers with the ability to reduce budgetary uncertainty when allocating limited financial resources among competing operational, repair, maintenance, and expansion activities within the CIP. The conclusions of the paper reveal that if DWI decision-makers assume risk-avoidance positions through proactive asset management (AM) strategies, they will achieve positive affects on expected budgetary outcomes.

Keywords: Risk assessment; Asset management

1. Introduction

Under their current philosophy drinking water infrastructure (DWI) decision-makers attempt to manage the risk of systems failure (R) through deterministic trial and error approaches that provide inefficient solutions (Wu et al., 2001). These deterministic trial and error approaches typically exclude uncertainty and access individual DWI component failures as a function of age, length, and service for a given distribution network configuration at a specific point in time (Silinis and Stewart, 2003). The algorithms for these methods generally use simple performance deficiency prioritization rules to setup skeletal preventative maintenance (PM) programs (Quimpo and Wu, 1997; McKay et al., 1999). These skeletal PM programs are then further enhanced by incorporating linear or power laws that enable PM activities to be calibrated against historic DWI component failures (Franks, 1999).

Historically, Hastak and Baim (2001) found that the chosen methods for managing a systems performance were driven by politically sensitive solutions that chose DWI expansion alternatives for economic development activities rather than optimizing tradeoffs among competing DWI maintenance, repair, and rehabilitation activities to maintain a desired level-of-service or systems reliability [$r(f)$]. This lack of interest for managing $r(f)$ was acerbated by the practical issues of actually estimating R that include: (1) evaluating the benefits for optimal financial tradeoffs among competing DWI reactive and proactive maintenance activities (Silinis and Stewart, 2003); (2) selecting the optimal rate of reinvestment (RR) over the life cycle of

*Corresponding author. Tel.: +434 384 7845; fax: +434 386 5696.
E-mail addresses: jwr@centralva.net, jwr9f@virginia.edu (J.W. Rogers), gel7f@virginia.edu (G.E. Louis).
[1]Present address: 102 Saddle Rock Rd., Lynchburg, VA, 24503, USA.

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.002

critical DWI components (Booth and Rogers, 2001); (3) identifying a set of optimal rehabilitation improvements for a given distribution network configuration (Wu et al., 2001); (4) prioritizing the efficient allocation of limited financial resources to maintain a desired $r(f)$ and extend the service life of critical DWI components; and (5) analyzing the interrelationships among complex risk factors and associated DWI component costs (Hastak and Baim, 2001).

The lack of ability to actually estimate R within the existing DWI methods requires developing new methods that: (1) facilitate integrated solutions through systems and multi-objective decision analysis techniques (Stern and Kendall, 2001); (2) incorporate complex environmental influences such as soil type, trench construction, water chemistry, piping material, etc. within existing methods (Franks, 1999); and (3) optimize tradeoffs between $r(f)$ and associated costs among competing DWI scheduled maintenance (SM) and emergency repair (ER) activities (Wu et al., 2001). Accurately estimating R will provide DWI decision-makers with a time-dependent snapshot of the physical deterioration of the critical DWI components. This time-dependent snapshot ultimately provides decision-makers with an efficient maintenance and rehabilitation roadmap for reliable systems operations at the lowest possible cost (McKay et al., 1999).

Although history suggests that few DWI decision-makers were interested in a probability formulation for estimating R, new financial rules, such as the General Accounting Standards Board (GASB) Rule 34 issued in June 1999, are changing their mindset. GASB-34 requires public services, such as DWI utilities, to develop viable asset management (AM) programs (Booth and Rogers, 2001). The purpose of these AM programs are to preserve the life and function of critical assets by optimizing the allocations of limited DWI operational, maintenance, and capital resources. Stern and Kendall (2001) and Booth and Rogers (2001) concluded that the implementation of a viable AM program for DWI will result in fiscal savings of 20–40%.

These potential fiscal savings, alone, provide sufficient justification for DWI decision-makers to develop new AM programs that incorporate attributes for uncertainty. These AM programs utilize methods, such as integrated quality approaches, to facilitate coordinated and comprehensive processes that estimate R for managing DWI assets (Stern and Kendall, 2001). The integrated quality approach of AM programs seeks to maintain a desired $r(f)$, while minimizing the total cost for owning and operating the DWI assets. However, before an integrated quality approach can be implemented by DWI decision-makers, they must understand the role that uncertainty plays in the estimation of R.

This paper presents a practical risk assessment methodology for estimating R to provide the DWI decision-maker with an objective risk assessment tool. The purpose of this objective risk assessment tool is to maintain the desired $r(f)$, while managing the financial uncertainty of the expected budgetary outcome within the CIP. The magnitude of the expected budgetary outcome is managed through the DWI decision-maker's sensitivity to R, which is represented as the level of the systems' RR. The expected result of the proposed risk assessment tool demonstrates that by proactively managing R to maintain the systems' $r(f)$ will effectively manage the uncertainty of the DWI systems' expected budgetary outcome within its' CIP.

The limitations of the proposed objective risk assessment tool are that the: (1) underlying principles are applicable to only DWI. Additional theories would need to be evaluated for other industries; (2) decision response alternatives only evaluate three discrete risk sensitivities rather than evaluating a stochastic distribution of risk sensitivities; and (3) value is demonstrated through a generic example calculation rather than actual case study. The objective risk assessment tool needs to be verified and validated through the application of extensive case study analyzes.

The paper's discussions of the practical risk assessment methodology begin with an overview of the concepts of risk analysis. Next, the existing methods for estimating R are evaluated and combined within the proposed risk assessment methodology. Next, the applicable theories for estimating R and $r(f)$ are formulated into the objective risk assessment tool. Finally, the value of the objective risk assessment tool is demonstrated through a set of example calculations for a conceptual DWI system to show the ability of the DWI decision-maker to manage the expected budgetary outcome within the CIP.

2. Risk analysis concepts

Traditional risk analysis is undertaken to improve the quality of information that the DWI decision-maker utilizes by examining the uncertainty of decision response alternatives. The risk analysis is performed by calculating the magnitude of risk for an individual risk factor and/or cumulative risk. The cumulative risk is represented by the summation of the individual risk factors times the relative weight of the respective risk factors (Hastak and Baim, 2001). A comprehensive risk analysis should quantitatively understand the risk factors, their intermediate/final impacts on cost, and the actions to mitigate their impact (Ezell et al., 2000).

Haimes (1998) suggests that risk assessment looks at what can go wrong as well as its likelihood and consequences. Hastak and Baim (2001) define infrastructure risk as the product of the probability of system failure $p(f)$ and associated costs of returning the system to service. This mathematical relationship is represented as $R = p(f)xC$, where R is defined as the risk of infrastructure system failure; $p(f)$ is defined as the likelihood or probability of the systems' failure; and C is defined as the economic-value of returning the failed system to service. The economic-value of an DWI's $p(f)$ is calculated as the total service expenditures in terms of equipment, material, and labor costs that are necessary to return the system to its normal operating condition.

DWI systems' uncertainty or R is defined as the likelihood or probability that the DWI utility fails to provide water on-demand to its customers. The DWI decision-maker either: (1) waits until critical DWI components fail and then expeditiously completes an ER activity; or (2) performs SM activities on critical DWI components prior to its failure. These are receptacle decision response alternatives that provide the DWI decision-maker the ability to balance the expected CIP budgetary outcome on an annual basis.

The decision response alternatives for the deterministic SM activities are the known or controllable costs that can be delayed by the DWI decision-maker in response to their affect on the expected CIP budgetary outcome. However, delaying SM activities on critical DWI components adversely affects the $p(f)$ over the life of the DWI system. Correspondingly, as the $p(f)$ rises, the level of the stochastic ER activities increases. Lauer (2001) found that balancing stochastic ER activities with deterministic SM activities is the key to minimizing the overall DWI maintenance and repair (MR) costs.

The decision response alternatives for the stochastic ER activities are uncertain or uncontrollable costs that are determined by the $p(f)$. It is the randomness in the level of the ER activity costs that create the R for the DWI decision-maker regarding its cumulative affect on the expected CIP budgetary outcome. This situation is acerbated further by the fact that ER activity costs are paid at overtime or premium rates. These premium rates are charged for the ER activities because they require immediate attention causing increased costs for shifting, expediting, and/or providing extra personnel, materials, and equipment on a moments notice. The C of the premium cost of the stochastic ER activity typically range from 1.5 to 2.0 times the cost of the deterministic SM activity.

This proposed objective risk assessment tool defines the DWI decision-maker's sensitivity to R by specifying corresponding levels of RR. A risk assessment is performed utilizing three decision response alternatives that are based on desired risk-sensitivity in terms of low, medium, or high RR. The goal of the objective risk assessment tool is to demonstrate the value of estimating the DWI decision-maker's sensitivity to R. The objectives are to: (1) incorporate $p(f)$ into the CIP budgetary analysis process; and (2) evaluate the affects of $p(f)$ on the expected CIP budgetary outcome.

3. Existing methods

Many professional and governmental agencies have published assessments and projections for the nation's DWI needs over the next 20 years. These need assessments have shown that the current DWI is in a state of general disrepair and substantial funding above the current levels will be required to maintain the DWI at acceptable $r(f)$ through cost-effective CIP. To further compound the current needs dilemma the DWI competes for limited resources with the other critical public infrastructures for transportation, schools, wastewater, solid-waste, and energy.

The American Society of Civil Engineers (ASCE) published their *Report Card on America's Infrastructure* in March 2001. This ASCE (2001) report gave America's infrastructures an overall grade of D+ and identified a 5-year total infrastructure investment need of $1.3 trillion. Specifically, the DWI consisting of approximately 54,000 community water systems (CWS) will face an annual investment shortfall of $11 billion for replacement and compliance requirements. The ASCE report recommended immediately increasing the federal funding levels by an additional $5 billion per year under the Safe Drinking Water Act (SDWA). The ASCE (2001) report also recommended the creation of a water trust fund to ensure sustainable funding for the projected infrastructure needs as well as a source of associated innovative financing mechanisms to develop public/private partnerships to encourage a steady influx of new capital for technological advancement.

The Environmental Protection Agency (EPA) completed its *1999 Drinking Water Infrastructure Needs Survey* (DWINS) in February 2001. The EPA (2001) report surveyed approximately 54,000 CWS and 21,400 not-for-profit non-community water systems (NCWS) for their respective infrastructure needs. The EPA (2001) report also documented a 20-year capital investment need of $150.9 billion for public water systems that are eligible to receive funding under the State Revolving Fund (SRF) program. The EPA (2001) report found that $31.2 billion is directly attributable to specific SDWA regulations. The EPA (2001) report also documented that the average age of the DWI network is estimated to be between 50 and 100 years.

The continued aging or deterioration of the DWI network is the primary reason for the projected financial needs as documented in the Water Infrastructure Network (WIN) Coalition report entitled, "*Safe and Clean Water for the 21st Century: A Renewed National Commitment to the Water and Wastewater Infrastructure*" in April 2000 and in its follow-up report entitled, "*Water Infrastructure NOW: Recommendations for Clean and Safe Water in the 21st Century*" in February 2001. These WIN (2000, 2001) reports argue for a combination of federal, non-federal, and private solutions for the resolution of the infrastructure-funding shortfall. The WIN (2000) report suggests that the economic and political history of general infrastructure investment has precedent for providing clean water under the Clean Water Act (CWA) and safe water under the SDWA.

The WIN (2001) report recommends a number of ways the federal government might take a leadership role, including research grants, water trust funds, low-interest loans, and incentives for private investment. Without a significantly enhanced federal role in financing the drinking water and wastewater infrastructure, critical investments will not occur. Failure to meet the clean and safe water investment needs of the next 20-years risks reversing the environmental, public health, and economic gains of the last three decades. Since federal government subsidy programs are down 75% from 1986 levels, local governments and ratepayers must fund approximately 90% of the clean and safe water infrastructure costs while grappling with competing needs for new schools, law enforcement, and social services (WIN, 2001).

The current financial health of the nation's CWS sector is documented by the EPA's *1995 Nation-wide Community Water System Survey* which was completed in January 1997. The results of the EPA (1997) survey revealed that the CWS sector spent \$32 billion from 1986 to 1995 on CIPs. On average the 1995 CWS sector's AM activities spent approximately 20% on water quality (WQ) compliance, 30% on MR activities, and 50% on scheduled replacement (SR) improvements. On average the 1995 CWS sector completed MR activities on approximately 2% of their respective piping networks. In addition, the 1995 CWS sector completed SR activities on approximately 2% of their respective piping networks. The WQ, MR, and SR activities within the CIP compete for the same AM program dollars.

The aging of the DWI system, along with increasing competitive pressures and customer demands, are forcing public water services to develop effective AM programs such as the GASB-34 requirements for maintaining service efficiency and product reliability while meeting quality standards and containing service costs (Booth and Rogers, 2001). The public water services are looking to AM programs to provide integrated solutions to improve the service efficiency and product reliability by balancing the MR activities as a function of the SM and ER needs (Stern and Kendall, 2001). It is generally accepted that as the DWI system ages the occurrence of stochastic ER activities increases due to the adverse impact of physical, chemical, and biological deterioration forces. Additionally, as the frequency of stochastic ER activities increases, the SR improvements of the aging DWI system become the more viable option (Silinis and Stewart, 2003).

DWI needs may range from stochastic ER activities for maintaining the $r(f)$ to SR improvements for creating new capacity for present and projected demand. The ability to pay may range from public drinking water utilities with strong bases of commercial and affluent users to those with large bases of government services and low-to-moderate-income residential users. Managing the DWI requires decision-makers to manage the old as well as to build the new. They must know about governance, public health, engineering, operations, and politics as well as economics and finance. The DWI problems discussed in the public forum are only 5% technical; 95% of the focus is on finance, public acceptance, and environmental impacts. There are abundant opportunities in the DWI arena to take lead roles in defining issues, creating solutions, explaining to public, and leading the holistic implementation process. Grigg (1999) and Means (2001) stressed that communicating the critical nature of DWIs WQ, MR, and SR activities under the CIP is paramount for the fundamental survival, public-health protection, economic development, and quality of life of America's people.

Craun and Calderon (2001) emphasized the importance of adequately maintaining DWI systems. Craun and Calderon (2001) found that since 1995 DWI deficiencies have been responsible for 45% of all waterborne disease outbreaks reported by the nation's 54,000 CWS. The DWI system deficiencies involved chemical and microbial contamination through cross-connections, back siphonage, inadequate separation, contaminated storage facilities, water main repairs, water main construction, and meter installation (Craun and Calderon, 2001). To reduce the risk of waterborne disease outbreaks more attention needs to be placed on the preventing of contamination of the distribution system through an optimal combination of corrective, preventative, and proactive maintenance policies.

The DWI system is aging, which contributes to waterborne disease outbreak and reduced $r(f)$. Sufficient CIP funds consisting of WQ, MR, and SR budgetary components should be allocated for the routine inspection, repair, and expansion of water mains and storage facilities as well as the SR of the older infrastructure. To reduce the potential for distribution system outbreaks and failures, drinking water utilities must maintain adequate water pressures, repair leaking mains, maintain chlorine residual, adopt cross-connection programs, include inspection programs, adequately disinfect after construction work, and increase corrosion control efforts (Craun and Calderon, 2001).

Scharfenaker (2001) complements the growing body of related research by estimating the nation's projected investment need to replace buried DWI pipes and associated appurtenances. Scharfenaker (2001) forecasts DWI repair and replacement expenditures based on how such assets wear out over their expected life spans. The historic investment pattern for US drinking water utilities shows: (1) local revenue financing is resulting in a steady accumulation of "invisible replacement liability" through deferred MR decisions; (2) changes in pipe manufacturing techniques has resulted in a significant drop in life expectancy; (3) specific short-term facility upgrades concentrated on compliance with the stringent SDWA water quality standards over the last 25-years; and (4) demographic change places a severe financial strain on public DWI systems and has resulted in the projected "replacement gap" in many areas of the country. Scharfenaker (2001) estimates DWI expenditures to be approximately \$250 billion over the next 30 years. Drinking water utilities must develop innovative ways to close this "replacement gap" to ensure the continued delivery of safe water by transitioning from a path of repairing their pipe networks to cost-effectively replacing huge amounts of buried infrastructure that is now at the end of its life (Scharfenaker, 2001).

Scharfenaker (2001) recommends three steps to close the "replacement gap" and associated "affordability gap": (1) developing a comprehensive local strategy to assess the condition of the infrastructure; (2) reforming state programs so that they are required to use federal funds in a timely, efficient, and effective manner; and (3) increasing federal assistance by expanding the existing SRF programs and creating new funds to stimulate research for more efficient management and technologies. Failure to accurately predict the appropriate RR for the DWI system risks severe $r(f)$ deterioration adversely affecting the financial stability of the public drinking water utilities.

Optimizing the financial tradeoffs between the competing corrective, preventative, and proactive maintenance activities of the AM program subject to minimizing the adverse impacts on the expected CIP budget outcome emphasizes the importance of DWI system maintenance in the effectiveness of the overall system safety and reliability (Stern and Kendall, 2001; Silinis and Stewart, 2003). Many quantitative analysis methodologies utilize mathematical representations to model the complex physical, chemical, and biological deterioration processes that adversely affect the DWI's $r(f)$ (Quimpo and Wu 1997; McKay et al., 1999; Franks, 1999; Wu et al., 2001; Stern and Kendall, 2001; Silinis and Stewart, 2003). These quantitative analysis methodologies provide objective and repeatable procedures to justify CIP budgets that effectively allocate dollars, extend asset life, and plan programs to maintain the systems' $r(f)$. However, they do not incorporate risk assessment into their respective quantitative analysis methodologies.

More recent quantitative analysis methodologies incorporate risk assessment and management techniques to facilitate decision analysis by highlighting tradeoffs among individual risks factors and associated mitigation cost impacts to effectively manage infrastructure assets (Hastak and Baim, 2001; Ezell et al., 2000). However, the stochastic nature of the DWI systems' deterioration processes adds considerable complexity to these quantitative analysis methodologies requiring substantial technical and modeling expertise. While the required DWI system investment is daunting in size, it creates an opportunity to sensitize the consumers and decision-makers of the opportunities of strategic risk management as communities differ in their DWI needs, ability to pay, and ability to manage $r(f)$. This paper demonstrates an objective risk assessment tool for maintaining a desired $r(f)$ while managing the expected CIP budgetary outcome. The ability to manage $r(f)$ by minimizing the $p(f)$ may range from instituting high reinvestment strategies to no reinvestment strategies due to limited budgets.

4. Theory

The discontinuation of the current state of disrepair of the DWI system requires the prioritization of corrective, preventative, and proactive maintenance activities within the CIP budgeting process to ensure the effective allocation of limited resources (Li and Haimes, 1992). The DWI system is now nearly 100 years old with at least 26% of the drinking water mains constructed of unlined cast iron and steel piping. These drinking water mains have greatly reduced carrying capacities and are rated in poor condition adversely impacting the current DWI systems' $r(f)$ (Kirmeyer et al., 1994).

The prioritization of corrective, preventative, and proactive maintenance activities within the CIP budgeting process is the basis for the development of the DWI risk assessment tool. The DWI risk assessment tool is characterized by three risk-sensitive decision response alternatives, which are described in Table 1.

The *Null* or corrective decision response assumes that the $p(f)$ is high and decision-maker's risk sensitivity is low because low levels of SM or PM activities are planned to mitigate the $p(f)$ due to infrastructure deterioration. The associated $r(f)$ is adversely affected because the ER or corrective maintenance activities are performed only after the DWI system fails.

The *Traditional* or preventative decision response assumes that the $p(f)$ is medium and decision-maker's risk sensitivity is medium because medium levels of SM or PM activities are planned to mitigate the $p(f)$ due to infrastructure deterioration. The associated $r(f)$ remains stable since the level of ER or corrective maintenance activities is marginally reduced by the SM activities.

Finally, the *Strategic* or proactive decision response assumes that the $p(f)$ is low and decision-maker's risk sensitivity is high because high levels of SM or PM activities are planned to prohibit the $p(f)$ due to infrastructure deterioration. The $r(f)$ increases since the level of ER or corrective maintenance activities is substantially reduced by the SM activities.

The basic relationships between an infrastructures' maintenance frequency (MF) and systems' performance are discussed in detail by Nelson et al. (1999) in an EPA report entitled: "*Optimization of Collection System Maintenance Frequency and System Performance.*" These relationships are diagramed in Fig. 1.

The Nelson et al. (1999) methodology involves the following assumptions:

(1) The average infrastructures' life expectancy is assumed to be 100 years with a salvage value of 25%. The

Table 1
Description of the drinking water infrastructure (DWI) decision response alternatives as a function of the level of the probability of system failure $p(f)$, risk sensitivity of the decision-maker, and expected level of the emergency repair activities

Decision responses	$p(f)$	Risk sensitivity	Emergency repairs
Null (corrective)	High	Low	High
Traditional (preventative)	Medium	Medium	Medium
Strategic (proactive)	Low	High	Low

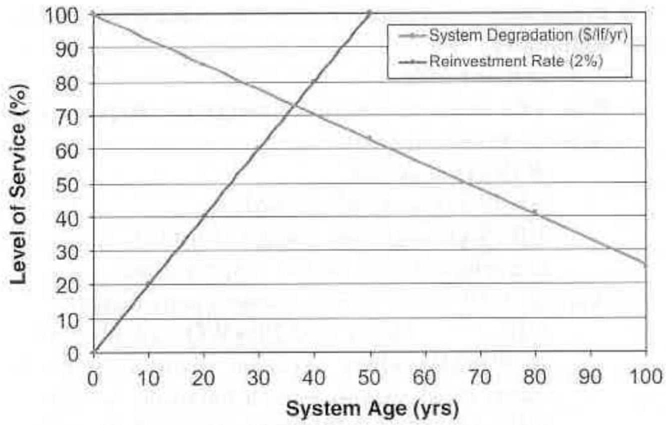


Fig. 1. Graphical representation of the linear relationships between the DWI systems' age (years) and level-of-service (% like-new performance) as a function of the system degradation and reinvestment rates. *Source*: Nelson, R.E., Hsiung, P.H., and Witt, A.A., *Optimization of collection system maintenance frequency and system performance*. EPA Cooperative Agreement No. CX-824902-01-0, Washington, DC, February 1999.

average infrastructures' unit value is assumed to be $100 per linear foot of pipe. The rate of the infrastructures' deterioration is assumed to be $0.75 per linear foot per year. Based on these assumptions the rate of system degradation (SD) is represented by the following linear relationship: SD = $100−($0.75*X), where the (*x*-axis) value is equal to the average age of the infrastructure network in years.

(2) The *r*(*f*) or associated (*y*-axis) value is assumed to be the percentage of the "like new" systems' performance. The "remaining" systems' value is a function of the average MF and systems' age. The percentage of the "like-new" systems' performance is assumed to represent the infrastructures' current *r*(*f*). The infrastructures' MF is assumed to be the number of years to return the infrastructure to 100% or a "like-new" systems' performance level given a specific level-of-maintenance activities as represented by the chosen level of RR.

(3) Infrastructures' MR program is defined as ongoing ER and/or SM activities that include actions that retard or correct the deterioration of the infrastructure system. The infrastructure systems' level of maintenance or RR is defined as the percentage of systems' rehabilitated within a given time period or MF.

(4) The infrastructures' RR is assumed to be the inverse of its' maintenance frequency (MF). The RR or 1/MF determines the level-of-maintenance activities or percentage of the total infrastructure rehabilitated to return the system to "like-new" condition in a given period of time. For example, a RR = 2% will return the system to a "like-new" systems' performance condition in MF = 1/0.02 = 50 years.

Fig. 1 is utilized to estimate the expected infrastructure systems' *r*(*f*) given the average rate of system degradation (SD) and reinvestment (RR) assumptions that are specified in the EPA (1997) survey. The methodology steps include: (1) drop a vertical line from the end of the 2% RR line to the corresponding point on the SD line; (2) draw a horizontal line to the desired level-of-service [% "like-new" systems' performance or *r*(*f*)] axis; and (3) note that the *r*(*f*) [*y*-axis] value is intersected at 65%, which represents the current infrastructure systems' *r*(*f*). Therefore, given the assumed rate of SD with a 2% RR the overall infrastructure systems' *r*(*f*) is expected to perform at 65% of the performance level of a similar "like-new" infrastructure system.

The basic infrastructure relationships defined by Nelson et al. (1999) is applicable to critical infrastructures, such as DWI systems. The *r*(*f*) for DWI systems may be defined as the probability that the DWI system operates correctly throughout a specified period of time given the system is operating at the start of the time period (Haimes, 1998). The *p*(*f*) for DWI systems may be defined as the probability that the DWI system fails during a specified period of time given the system is operating at the start of the time period. The relationship between DWI systems' *r*(*f*) and *p*(*f*) values may therefore be expressed as the $p(f) = 1-r(f)$. Given this simple relationship, the systems' *r*(*f*) or *y*-axis value calculated in Fig. 1 can be expressed as the systems' *p*(*f*) value over a given 1/MF or RR. For example, given the DWI's average rate of SD in Fig. 1 with a 2% RR from the previous example, a DWI systems' *r*(*f*) value is estimated to be 65% of the "like-new" systems' performance level. The systems' *p*(*f*) value within the 50-year MF period is estimated to be 35%. This means that a DWI system with a 2% RR is operating at a *r*(*f*) that is 35% below the 100% systems' performance level. In other words, the DWI system has a 35% chance of failing to operate at least once throughout the associated 50-year MF period. Therefore, the level of financial reinvestment plays a crucial role in the maintenance of the systems' *r*(*f*).

Guignier and Madanat (1999) found that, historically, the level of financial reinvestment through the infrastructures' CIP budget is broken down into competing maintenance and improvement activities. Additionally, the EPA (1997) survey found the CIP budget for DWI systems is broken down into: (1) WQ activities for meeting the systems' regulatory requirements; (2) MR activities for meeting the systems' customer requirements; and (3) SR improvement activities for meeting the systems' economic development requirements within the AM program. The MR activities consists of CIP budgetary line-items that are broken down as SM and/or ER activities. The ER activities retard and/or correct infrastructure deterioration after the system has failed. The ER activities are stochastic in nature and vary widely from one year to the next because they are based on risk factors such as infrastructure age, material, leakage, water quality, etc. The WQ, SM, and SR activities consist of CIP budgetary line-items that are compliance, maintenance, and replacement activities that improve and/or functionally alter the infrastructure system returning it to a "like-new" condition state. The budgetary line-item amounts for these WQ, SM, and SR activities are determined by the DWI decision-maker for each budgetary cycle.

The DWI decision-making problem involves balancing the financial tradeoffs between the deterministic WQ, SM, and SR activities and expected adverse impact from the stochastic level of the ER activities, while maintaining the current $r(f)$. As explained earlier, the C of the stochastic ER activities are quite expensive relative to the more cost-effective WQ, SM, and SR activities. The DWI decision-maker must answer the following question: What is the optimum level of the SM activities that will maintain the current $r(f)$, while minimizing the adverse affects of the C from the expected level of ER activities on the CIP budgetary outcome? Therefore, the DWI decision-maker must accurately estimate the expected level of ER activities relative to a chosen level of SM activities within the MR component of the CIP budget.

To answer the above question a practical risk assessment methodology is used to develop an objective DWI risk assessment tool to estimate the impact of $p(f)$ on the expected CIP budgetary outcome. The following CIP budgetary assumptions were made for a conceptual DWI system:

(1) Annual Water Quality (WQ) Compliance Activity Costs
 WQ (\$/yr) = Labor + Material + Equipment aspects.
 WQ (\$/yr) = (\$/h, \$/lf, \$/event) × (# h, # lf, #/events).
(2) Annual Scheduled Replacement (SR) Improvement Activity Costs
 SR (\$/yr) = Labor + Material + Equipment aspects,
 SR (\$/yr) = (\$/h, \$/lf, \$/event) × (# h, # lf, #/events).
(3) Annual SM Activity Costs
 SM (\$/yr) = Labor + Material + Equipment aspects,
 SM (\$/yr) = (\$/h, \$/lf, \$/event) × (# h, # lf, #/events).
(4) Annual Emergency Repair (ER) Activity Costs with $C = 1.5$ premium factor
 ER (\$/yr) = 1.5 × SM.
(5) Average CIP Costs with *no* risk assessment factors
 CIP (\$/yr) = (0.20)*WQ + (0.3)*(ER + SM) + (0.50)*SR, where assumed component coefficients must sum to 1 for a balanced budget.
(6) Annual Maintenance and Repair (MR) Activity Costs
 MR (\$/yr) = 0.3 (ER + SM), where coefficient constrained in assumption 5,
 MR (\$/yr) = (0.3−b)*ER + (b)*SM, where b is the selected level of SM activity costs,
 MR (\$/yr) = (0.3−b)*(1.5*SM) + (b)*(SM),
 MR (\$/yr) = [(0.3−b)*(1.5) + (b)]*(SM).
(7) Probability of Systems' Failure as a function Systems' Reliability
 $p(f) = 1 - r(f)$.
(8) Risk (R) from Stochastic Emergency Repair (ER) Activity Premium Costs
 R (\$/yr) = $p(f)$*C,
 ER (\$/yr) = $p(f)$*[1.5*SM],
 ER (\$/yr) = [$p(f)$*1.5*(0.3−b)]*SM, where b is the selected level of SM activity costs.
(9) Average CIP Costs *with* risk assessment factors
 CIP (\$/yr) = (0.20)*WQ + (0.30)*MR + (0.50)*SR, where assumed component coefficients must sum to 1 for a balanced budget,
 CIP (\$/yr) = (0.20)*WQ + (0.3)*(ER + SM) + (0.50)*SR,
 CIP (\$/yr) = (0.20)*WQ + (0.3−b)*(1.5*SM) + (b)]*SM + (0.50)*SR,
 CIP (\$/yr) = (0.20)*WQ + (0.3−b)*{$p(f)$*1.5}*SM + (b)*SM + (0.50)*SR,
 CIP (\$/yr) = 0.2*T + 0.3*T + 0.5*T, where T represents the \$/yr values of the WQ, SM, and SR components.
(10) Objective DWI Risk Assessment Tool

$$\mathrm{CIP}(T) = (0.2)*T + [p(f)*1.5*(0.3 - b)]*T + (b)*T + (0.5)*T, \text{ where } T \text{ in \$/yr.} \quad (1)$$

Eq. (1) provides the mathematical foundation for the objective risk assessment tool. The WQ, ER, SM, and SR component coefficients for the conceptual DWI systems' CIP budget are based on national averages from the EPA (1997) survey. Therefore, the assumed values of these coefficients are specific to the drinking water sector. Actual values for these coefficients for an individual drinking water utility will need to be verified and validated on a case-by-case basis. The WQ and SR component coefficients are assumed to be constant AM activities within the CIP budget of the conceptual DWI system. The ER and SM component coefficients of the MR component are assumed to be competing activities within the CIP budget of the conceptual DWI system.

The objective risk assessment tool assumes that under normal conditions the DWI decision-maker will re-allocate limited financial resources between the competing ER and SM activities of the MR component within the CIP budget of the conceptual DWI system. The efficient allocation of the financial resources for the MR component is based on the decision-maker's understanding of the competing needs to maintain the current $r(f)$ and balance the expected total CIP budget. The objective risk assessment tool allows the DWI decision-maker to set the RR in accordance with his/her level of risk-sensitivity (low, medium, and high) and determine its impact on the expected CIP budgetary outcome.

5. Example calculation

To demonstrate the utilization of the objective risk assessment tool for the conceptual DWI system, as illustrated by Eq. (1), three typical DWI risk-sensitive (low, medium, and high) decision response alternatives are outlined in Table 1 and defined as follows:

(1) NULL (Corrective): The decision-maker's risk-sensitivity level is low and selects a RR of 1% that equates to a systems' MF of 100 years. Then using Fig. 2, the y-axis value for the $r(f)$ equals 25%. The corresponding $p(f)$ is calculated as $1-0.25 = 0.75$ or 75%.
(2) TRAD (Preventative): The decision-maker's risk-sensitivity level is medium and selects a RR of 2% that equates to a systems' MF of 50 years. Then using Fig. 2, the y-axis value for the $r(f)$ equals 65%. The corresponding $p(f)$ is calculated as 35%.
(3) STRAT (Proactive): The decision-maker's risk-sensitivity level is high and selects a RR of 5% that equates to a systems' MF of 20 years. Then using Fig. 2, the y-axis value for the $r(f)$ is 85%. The corresponding $p(f)$ is calculated as 15%.

Next, the DWI decision-maker sets the level (b) of SM activities and calculates the associated level ($0.3-b$) of ER activities relative to his/her risk-sensitivity. It is assumed that the more risk-sensitive the DWI decision-maker the higher the level of SM activities within the MR component of the CIP budget. The three DWI risk-sensitive decision response alternatives with associated levels of SM and ER activity are shown in Table 2.

The selected level (b) of the SM activities represents the ability of the DWI decision-maker to manage the magnitude of adverse affects on the expected total budgetary outcome. The magnitude of adverse affects on the expected total budgetary outcome is mitigated by decreasing the expected level ($0.3-b$) of the ER activities within the MR component of the CIP budget. In other words,

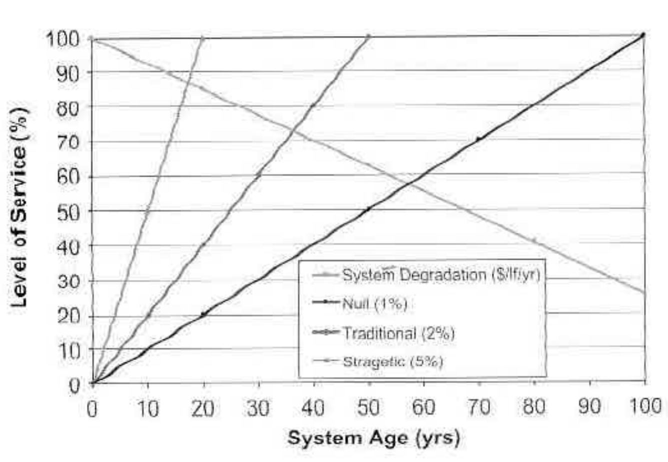


Fig. 2. Graphical representation of the three DWI risk-sensitive decision response alternative reinvestment rate (%) assumptions as a function of the systems' age (years), level-of-service (% like-new performance), and rate of system degradation. *Source*: Nelson, R.E., Hsiung, P.H., and Witt, A.A., *Optimization of collection system maintenance frequency and system performance*. EPA Cooperative Agreement No. CX-824902-01-0, Washington, DC, February 1999.

Table 2
Three DWI risk-sensitive decision response alternatives as a function of the level (b) of scheduled maintenance, and associated level ($0.3-b$) of emergency repair activities within the maintenance and repair component of the capital improvement program budget

Decision response alternative	$SM = b$	$ER = 0.30-b$
Null (low risk-sensitivity)	0.10	0.20
Traditional (medium risk-sensitivity)	0.20	0.10
Strategic (high risk-sensitivity)	0.25	0.05

there exists a correlation such that as the selected level (b) of SM activities increases the level ($0.3-b$) of expected ER activities decreases within the MR component of the CIP Budget. By increasing the level of SM activities through a higher RR, the DWI decision-maker is effectively managing the systems' $r(f)$ by decreasing the expected level of the costly ER activities.

Finally, using Eq. (1), as developed under the conceptual DWI system's CIP budgetary assumptions, the expected total CIP budgetary outcome is calculated for each of the DWI risk-sensitive decision response alternatives:

$$\text{Expected Total CIP Budgetary Outcome} \\ => T(\$/\text{yr}) = 0.2{*}T + [p(f){*}1.5{*}(0.3-b)]{*}T \\ + (b){*}T + 0.5{*}T,\ \text{given } 0.0<b<0.3. \quad (2)$$

The objective risk assessment tool calculates the adverse affects of three risk-sensitive (low, medium, and high) decision response alternatives on the expected total CIP budgetary outcome. The results of this demonstration of the DWI risk assessment tool are shown in Table 3.

6. Discussion

The results of the above demonstration reveal that the implementation of SM activities as part of the MR component can have a favorable impact on the expected total CIP budgetary outcome. An important result to note from Table 3 is that under the DWI risk assessment tool assumptions increasing the level of SM activities from $b = 0.20$–0.25 imposes very little marginal impact 0.95–0.96 on the expected total CIP budgetary outcome. This result suggests that a threshold level exists for which changes to the level of SM activities have insignificant impact on the expected total CIP budgetary outcome. This result coincides with the Lauer (2001) rule-of-thumb, which suggests that the overall costs of the MR component can be minimized when the following breakdown is utilized: The MR component breakdown should be 2/3 of the costs are for SM activities and 1/3 of the costs are for ER activities within the DWI CIP budget.

Intuitively, the DWI decision-maker assumes that as the level of RR increases the $r(f)$ increases and the $p(f)$ decreases proportionally. However, for the $p(f)$ to decrease the expected level, of ER activities must be managed to an

Table 3
Demonstration results of the DWI risk assessment tool for evaluating the impact of the three DWI risk-sensitive decision response alternatives on the expected total capital improvement budgetary outcome

Decision response alternative	Rate of reinvestment	$r(f)$	$p(f)$	SM	ER	SR	WQ	CIP budget impact
Null (low risk-sensitivity)	1%	25%	0.75	0.10	0.20	0.50	0.20	$1.03 \times T$
Traditional (medium risk-sensitivity)	2%	65%	0.35	0.20	0.10	0.50	0.20	$0.95 \times T$
Strategic (high risk-sensitivity)	5%	85%	0.15	0.25	0.05	0.50	0.20	$0.96 \times T$

acceptable level to have a favorable impact on the expected total CIP budgetary outcome. The DWI decision-maker manages R through optimal tradeoffs between the levels of the ER and SM activities within the MR component of the CIP budget. The DWI decision-maker utilizes the drinking water sectors' infrastructure and budgetary relationships that are illustrated in Fig. 2 and Eq. (1), respectively, to assume different levels for SM activities for each risk-sensitive decision response alternative under constant $r(f)$ assumptions. The DWI risk assessment tool clearly offers the decision-maker a practical risk assessment methodology to proactively manage the expected total CIP budgetary outcome.

Now suppose that the DWI risk-sensitive decision response alternative and associated level (b) of SM activities are held constant. How will the systems' $r(f)$ and associated $p(f)$ affect the expected total CIP budgetary outcome? This sensitivity analysis is graphically displayed in Fig. 3. The sensitivity analysis results show that under the DWI risk assessment tool assumptions, the $p(f)$ does not negatively impact the expected total CIP budget outcome until the $p(f)$ exceeds 0.67 indicating that a minimal $r(f)$ of 0.33 should be maintained at all times. This phenomenon is explained because as the $p(f)$ approaches unity the impact of the premium rates (1.5–2.0) from the expected ER activity costs become more significant thereby offsetting the benefits of the risk-reduction through the SM activities.

Now suppose the systems' $r(f)$ and associated $p(f)$ are held constant across the DWI risk-sensitive decision response alternatives. How will the DWI risk-sensitive decision response alternative and associated level (b%) of SM activities affect the expected total CIP budgetary outcome? This sensitivity analysis is graphically displayed in Fig. 4. The sensitivity analysis shows under the DWI risk assessment tool assumptions, the impacts from the level of SM activity costs on the expected total CIP budgetary outcome over the complete range the $p(f)$ values are partitioned by the risk-sensitive decision response alternatives. As the level of the SM activities increase within the MR component, the overall magnitude of the expected CIP budgetary impact stabilizes over the entire $p(f)$ range. This phenomenon occurs because the variability in the stochastic ER activity costs is mitigated by risk-reduction through the SM activities.

Under the DWI risk assessment tool assumptions, the magnitude of the expected total CIP budgetary impacts can

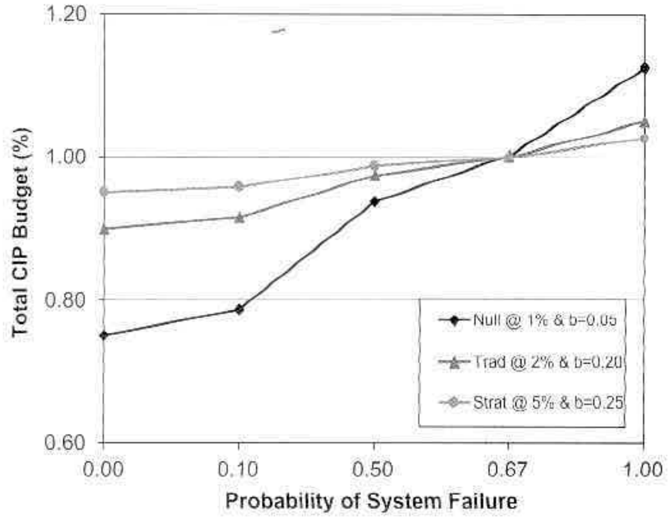


Fig. 3. Sensitivity analysis for the DWI risk assessment tool showing the impacts of the probability of system failure $p(f)$ on the expected total capital improvement program budgetary outcome for each of the DWI risk-sensitive decision response alternatives.

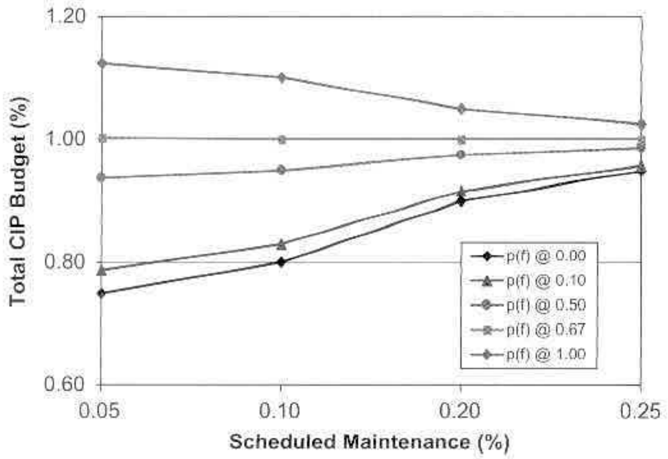


Fig. 4. The sensitivity analysis for the conceptual water infrastructure risk assessment tool showing the impacts of the level of scheduled maintenance activities on the total capital improvement program budget over the complete range the probability of system failure values for each of the risk-sensitive decision response alternatives.

be migrated by increasing the level of SM activities within the MR component regardless of the $p(f)$. In the same vain, the magnitude of the expected total CIP budgetary impacts can be migrated by increasing the overall DWI's RR regardless of the $p(f)$. Therefore, the DWI decision-maker may elect to take either a pay now or pay later approach to ensure that the appropriate $r(f)$ is maintained through an adequate RR throughout the life of the DWI.

7. Conclusions

The goal of the paper is to demonstrate the value of an objective risk assessment tool for estimating the DWI decision-maker's sensitivity to R. The purpose of the DWI risk assessment tool is to provide the ability for the decision-maker to maintain a desired $r(f)$, while managing adverse impacts on the expected total CIP budgetary outcome. The objectives of this paper are to: (1) incorporate $p(f)$ into the CIP budgetary analysis process and (2) evaluate the affects of $p(f)$ on the expected CIP budgetary outcome.

The usefulness of the objective risk assessment tool was demonstrated by defining three risk-sensitive (low, medium, and high) decision response alternatives that are encountered by the typical DWI decision-maker. To use the DWI risk assessment tool the decision-maker manages R by selecting the RR and determining the associated $r(f)$ from Fig. 2. Next, the associated $p(f)$ is calculated as $1-r(f)$. Then, the decision-maker sets the level (b) of SM activities such that $(0.00 \leqslant b \leqslant 0.30)$ and calculates the associated level of the ER activities as $\{p(f)*1.5*(0.3-b)\}$. Finally, using Eq. (1), the expected total CIP budgetary outcome is determined as a percentage of the total CIP (T) budget.

The CIP budgetary assumptions of the conceptual DWI system extends the MF and systems' performance relationships developed by Nelson et al. (1999) to the public drinking water sector. The CIP budgetary assumptions of the conceptual DWI system also extends the utilization of the EPA (1997) survey's average budgetary coefficient breakdowns for the WQ, MR, and SR components of the CIP to include $p(f)$ as a risk assessment aspect within the DWI decision-making process. The DWI risk assessment tool validates Lauer (2001)'s rule of thumb that the optimal breakdown of the MR component of the total CIP budget should be 2/3 SM activity costs and 1/3 ER activity costs.

This paper concludes that the DWI systems' $r(f)$ directly affects the $p(f)$ which in turn, affects the expected level of ER activities. The DWI decision-maker can minimize his expected total CIP budgetary impact by maintaining a minimum $r(f)$ or maximum $p(f)$ threshold through the selection of an adequate RR. Optimizing the associated tradeoffs between the corresponding levels of ER and SM activities within the MR component will mitigate the adverse impacts of the expected total CIP budgetary outcome. When applied to the DWI risk-sensitive three decision response alternatives, the DWI risk assessment tool reveals that by selecting an adequate RR to ensure an acceptable $r(f)$, the decision-maker can anticipate the expected impact on the total CIP budgetary outcome.

Every DWI decision-maker must meet annual budgetary constraints under their respective CIP budgets. The DWI decision-maker effectively manages $p(f)$ by selecting an appropriate level of the RR for the DWI system. The selected level of the RR maintains a corresponding level of SM activities, which has a significant impact on the expected total CIP budgetary outcome. DWI decision-makers must use practical risk assessment methods, such as the objective DWI risk assessment tool, as a means to effectively reduce R by efficiently allocating limited financial resources. This means developing risk-avoidance positions through optimal tradeoffs among competing corrective, preventative, and proactive maintenance activities within the context of a cost-effective AM program over the entire lifecycle of the critical components of the DWI system. Finally, the authors acknowledge that the proposed objective risk assessment tool needs additional verification and validation to assess its applicability at the systems level.

Acknowledgments

Work completed with support from EPA Star Fellowship Award U91594101 and National Science Foundation Career Award 9984318.

References

American Society of Civil Engineers (ASCE), 2001. The 2001 Report Card for America's Infrastructure. ASCE Headquarters, Reston, VA.

Booth, R., Rogers, J., 2001. Using GIS technology to manage infrastructure capital assets. Journal American Water Works Association (AWWA) 93 (11), 62–68.

Craun, G., Calderon, R., 2001. Waterborne disease outbreaks caused by distribution system deficiencies. Journal AWWA 93 (9), 64–75.

Environmental Protection Agency (EPA), 2001. 1999 Drinking Water Infrastructure Needs Survey: Second Report to Congress. EPA-816-R-01-004, Washington, DC, February 2001.

EPA, 1997. 1995 Community Water System Survey, vols. I and II: Overview and Detailed Survey Result Tables and Methodology Reports. EPA-815-R-97-001a/b, Washington, DC, January 1997.

Ezell, B., Farr, J., Wiese, I., 2000. Infrastructure risk analysis of municipal water distribution system. Journal of Infrastructure Systems, ASCE 6 (3), 118–122.

Franks, S., 1999. Disaggregation of environmental factors affecting sewer pipe failures. Journal of Infrastructure Systems, ASCE 3 (1), 150–158.

Grigg, N., 1999. Infrastructure: integrated issue or tower of babel, viewpoint. Journal of Infrastructure Systems, ASCE 5 (4), 115–117.

Guignier, F., Madanat, S., 1999. Optimization of infrastructure systems maintenance and improvement policies. Journal of Infrastructure Systems, ASCE 5 (4), 124–134.

Haimes, Y., 1998. Risk Modeling, Assessment and Management. Wiley, New York, NY.

Hastak, M., Baim, E., 2001. Risk factors affecting management and maintenance cost of urban infrastructure. Journal of Infrastructure Systems, ASCE 7 (2), 67–75.

Kirmeyer, G., Richards, W., Smith, C., 1994. An Assessment of Condition of North America Water Distribution Systems and Associated Research Needs. AWWA Research Foundation.

Lauer, W., 2001. Excellence in Action: Water Utility Management in the 21st Century. AWWA, Denver, CO.

Li, D., Haimes, Y., 1992. Optimal maintenance-related decision making for deteriorating water distribution systems, 1 and 2. Water Resource Research 28, 1053–1061 and 1063–1070.

McKay, D., Rens, K., Greimann, L., Stecker, J., 1999. Condition index assessment for US army corps of engineers civil works. Journal of Infrastructure Systems, ASCE 5 (2), 52–60.

Means, E., 2001. Meeting future head on. Journal AWWA 93 (7), 36–48.

Nelson, R., Hsiung, P., Witt, A., 1999. Optimization of collection system maintenance frequency and system performance. EPA Cooperative Agreement No. CX-824902-01-0, Washington, DC, February 1999.

Quimpo, R., Wu, S., 1997. Condition assessment of water supply infrastructure. Journal of Infrastructure Systems, ASCE 3 (1), 15–21.

Scharfenaker, M., 2001. Drawn of the replacement era: reinvesting in the drinking water infrastructure. Journal AWWA 93 (7), 20–32.

Silinis, P., Stewart, F., 2003. Development of a diagnostic index to guide maintenance. Journal AWWA 95 (5), 105–111.

Stern, C., Kendall, D., 2001. Improved utility infrastructure and operations management. Journal AWWA 93 (11), 69–73.

Water Infrastructure Network (WIN) Coalition, 2000. Clean & Safe Water for the 21st Century: A Renewed National Commitment to the Water and Wastewater Infrastructure Report, April 2000.

WIN Coalition, 2001. Water Infrastructure NOW: Recommendations for Clean and Safe Water in the 21st Century Report, February 2001.

Wu, Z., Boulos, P., Orr, C., Ro, J., 2001. Using genetic algorithms to rehabilitate distribution systems. Journal AWWA 93 (11), 74–85.

Journal of Environmental Management 87 (2008) 37–45

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Removal of copper (II) from aqueous solution by adsorption onto low-cost adsorbents

Haluk Aydın[a,*], Yasemin Bulut[a], Çiğdem Yerlikaya[b]

[a]Department of Chemistry, Faculty of Science and Arts, University of Dicle, 21280 Diyarbakır, Turkey
[b]Graduate School of Sciences, University of Dicle, 21280 Diyarbakır, Turkey

Received 13 October 2005; received in revised form 5 October 2006; accepted 2 January 2007
Available online 8 March 2007

Abstract

The use of low-cost adsorbents was investigated as a replacement for current costly methods of removing metals from aqueous solution. Removal of copper (II) from aqueous solution by different adsorbents such as shells of lentil (LS), wheat (WS), and rice (RS) was investigated. The equilibrium adsorption level was determined as a function of the solution pH, temperature, contact time, initial adsorbate concentration and adsorbent doses. Adsorption isotherms of Cu (II) on adsorbents were determined and correlated with common isotherm equations such as Langmuir and Freundlich models. The maximum adsorption capacities for Cu (II) on LS, WS and RS adsorbents at 293, 313 and 333 K temperature were found to be 8.977, 9.510, and 9.588; 7.391, 16.077, and 17.422; 1.854, 2.314, and 2.954 mg g^{-1}, respectively. The thermodynamic parameters such as free energy (ΔG^0), enthalpy (ΔH^0) and entropy changes (ΔS^0) for the adsorption of Cu (II) were computed to predict the nature of adsorption process. The kinetics and the factors controlling the adsorption process were also studied. Locally available adsorbents were found to be low-cost and promising for the removal of Cu (II) from aqueous solution.

Keywords: Adsorption; Adsorption modeling; Low-cost adsorbents; Heavy metals; Adsorption of copper (II)

1. Introduction

The contamination of water by toxic heavy metals through the discharge of industrial wastewater is a worldwide environmental problem (Ajmal et al., 2003). Their presence in streams and lakes has been responsible for several health problems with animals, plants and human beings. Numerous metals such as Sb, Cr, Cd, Cu, Pb, Hg, etc. have toxic effects on human and environment (Taty-Costodes et al., 2003). Since copper is a widely used material, there are many actual or potential sources of copper pollution. Copper may be found as a contaminant in food, expecially shellfish, liver, mushrooms, nuts, and chocolate. Briefly, any processing or container involving copper material may contaminate the products such as food, water or drink. Copper is essential to human life and health but, like all heavy metals, is potentially toxic as well. For example, continued inhalation of copper-containing spray is linked with an increase in lung cancer among exposed workers (Yu et al., 2000). The permissible limit of Cu is 2.5 mg L^{-1} in water (Prasad and Freitas, 2000).

Various technologies exist for the removal of such metals. They include filtration, chemical precipitation, ion exchange, adsorption using activated carbon, electrodeposition and membrane process. All these methods are generally expensive (Low et al., 2000). Recently, Bailey et al. (Bailey et al., 1999) reviewed a wide variety of low-cost adsorbents for the removal of heavy metals. A low-cost adsorbent is defined as one which is abundant in nature, or is a by-product or waste material from another industry. The review concluded that these low-cost processes involved the use of activated carbon and ion-changers. The cost of these biomaterials is negligible compared with the cost of activated carbon or ion-exchange resins that are in the range of approximately \$2.0–\$4.0/kg (Ferro-Garcia et al., 1988). The removal of heavy metals from industrial

*Corresponding author. Tel.: +90 412 2488550x3164; fax: +90 412 2488039.
E-mail address: halukaydin@dicle.edu.tr (H. Aydın).

0301-4797/\$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.005

wastewater is considered an important application of adsorption processes using suitable adsorbent.

At present, there is growing interest in using low-cost, commercially available materials for the adsorption of heavy metals. A wide variety of materials such as rice husk (Ajmal et al., 2003), modified cellulosic materials (Okieimen et al., 1985), fly ash (Panday et al., 1985), wheat brean (Bulut and Baysal, 2006), modified bark (Gloaguen and Morvan, 1997), sawdust (Ajmal et al., 1998), pine bark (Al-Asheh and Duvnjak, 1998), animal bones (Al-Asheh et al., 1999), holly oak (Prasad and Freitas, 2000), modified sawdust of walnut (Bulut and Tez, 2003) and tea-industry waste (Çay et al., 2004) are being used as low-cost alternatives to expensive adsorbents.

The aim of this paper is to assess the ability of low-cost adsorbents such as shells of lentil (*Lens culinaris Medik*), wheat (*Triticum durum Desf*) and rice (*oryza sativa L.*) to adsorb Cu (II) from aqueous solution. The effect of the solution pH, temperature, contact time, initial adsorbate concentration and adsorbent doses on the removal of Cu (II) was studied. The thermodynamic parameters for the adsorption Cu (II) have also been computed and discussed. The kinetics and the factors controlling the adsorption process were also studied.

2. Materials and methods

2.1. Adsorbents

Shells of lentil (LS), wheat (WS) and rice (RS) used as adsorbent were supplied from the South-eastern Anatolia of Turkey. The chemical and physical characteristics of the adsorbents are presented in Table 1. For the experimental studies, only the adsorbents were washed several times with double distilled water to remove surface impurities. The process was followed by air drying at 383 K for 24 h. The adsorbents were sieved through 0.6 mm sieve and used as such without any pretreatment.

Table 1
Physical and chemical properties of adsorbents used in the experiments

Chemical characteristic	Percent		
	LS	WS	RS
Moisture content	6.58	6.40	5.46
Water soluble components (inorganic matter)	18.12	22.33	6.46
Insoluble components (organic matter)	78.85	75.14	89.41
Ash content	9.18	2.58	22.80
Total loss of ignition	89.28	88.45	72.55
C content	33.53	44.59	51.37
H content	4.67	6.56	6.93
pH	5.03	6.05	5.98
Physical characteristics			
Surface area (BET) ($m^2 g^{-1}$)	0.19	0.67	0.83
Bulk density ($g cm^{-3}$)	0.49	0.36	0.38
Particle size (mm)	0.60	0.60	0.60

2.2. Chemicals

All chemicals [$Cu(NO_3)_2 \cdot 3H_2O$, HCl, NaOH, $NaHCO_3$ and Na_2CO_3)] were purchased from Merck. All the compounds used to prepare reagent solutions were of analytic reagent grade. The 1000 mg L^{-1} of stock solution of Cu (II) was prepared by dissolving $Cu(NO_3)_2 \cdot 3H_2O$ in double-distilled water.

2.3. Apparatus

Absorbance values were measured with a Unicam Model 929 atomic absorption spectrometer (AAS). A pH-meter (Jenway 3010) and shaker (Nuve ST 400) were used for pH adjustment and shaking, respectively. Elemental analysis was carried out with an EA 1108 Fisons instruments. The surface area was determined by a single-point N_2 gas adsorption method by using a model micromeritics Flow Sorb II, 2300.

2.4. Procedure

The effect of initial concentration on the adsorption by adsorbents from aqueous solution was investigated for a kinetic study, whose initial concentration was known. In addition, 2.0 g samples of adsorbent were mixed with 100 mL samples of Cu (II), and then mixtures were shaken (agitation speed = 150 rpm). Suspension was centrifuged in centrifuge test tubes for 5 min at 2000 rpm. Absorbance was measured with AAS. Moreover, the effects of temperature, adsorbent doses, contact time and pH on the adsorption of Cu (II) by adsorbents from solution aqueous were similarly investigated. The effect of pH on the adsorption phenomenon was studied by adding either 0.1 N HCl or NaOH in Cu (II) solution. Then, an isotherm study for Cu (II) was investigated. First, 1.0 g samples of adsorbent were mixed with 100 mL samples of solutions of different initial concentration (25–500 mg L^{-1}) prepared from stock solutions of Cu (II) and shaken for their equilibrium contact times at different temperatures (293, 303, and 333 K) and 150 rpm. The equilibrium concentration of Cu (II) was determined by AAS. The experiments were repeated twice for each point.

3. Results and discussion

3.1. C, H analysis

Elemental analysis was carried out with an EA 1108 Fisons instruments. A sample of adsorbents was put in an oven at 1273 K under oxygen in order to obtain a quick and complete combustion. H_2O and CO_2 were released and conducted in a copper oven at 923 K, then passed through a 2 m column with helium vector gas, and analyzed by a catharameter detector. The results of analysis are shown in Table 1.

3.2. Other chemical–physical characteristics

Characteristics of the adsorbents such as bulk density, moisture content, ash content, solubility in water (inorganic and organic matter), total loss of ignition were determined. The ash content was determined by burning it in a furnace at 1173 K (Ricordel et al., 2001). The results are summarized in Table 1.

3.3. pH of adsorbent surface

The sample of 1.0 g of dry adsorbent powder was added to 30 mL of water and the suspension was stirred overnight to reach equilibrium. Then, the pH of slurry was measured. The results are shown in Table 1.

3.4. Functional groups dosage

Knowing the chemical structure of the used adsorbent is important to understand the adsorption processes. The adsorption capacity of adsorbent is strongly influenced by the chemical structure of their surface (Ricordel et al., 2001). The oxygenated surface groups were determined according to the method by Bohem. First, 1.0 g of adsorbent was placed in 50 mL of each of the following 0.05 N solutions: sodium hydroxide, sodium carbonate, sodium bicarbonate, and hydrochloric acid. Vials were sealed and shaken for 24 h and then 5 mL of each filtrate was pipetted; the excess of the base or acid was titrated with HCl or NaOH, as required. The numbers of acidic sites of various types were calculated on the assumption that NaOH neutralizes carboxylic, phenolic, and lactonic; Na_2CO_3 neutralizes carboxylic and lactonic groups; and $NaHCO_3$ neutralizes only carboxylic groups (Adib et al., 1999). The number of surface basic sites was calculated from the amount of hydrochloric acid which reacted with the adsorbents; the results of Bohem titration (numbers of acidic and basic groups) are summarized in Table 2.

3.5. Contact time

The contact time experiments were carried out at 293 K temperature. The dependence of adsorption of Cu (II) with time is presented in Fig. 1. The adsorption increased with increasing contact time, and the equilibrium was attained after shaking for 3 h. Therefore, in each experiment, the shaking time was set to 3 h. The reason for LS that not constant after 3 h is due to the experimental mistake.

Table 2
Results of Boehm titration ($mmol\,g^{-1}$)

Sample	Carboxylic	Lactonic	Phenolic	Basic	Acidic
LS	0.434	0.051	0.586	0.00	1.071
WS	0.146	0.020	0.842	0.00	1.008
RS	0.166	0.076	1.212	0.00	1.454

3.6. Effect of pH

The pH of the aqueous solution is an important controlling parameter in the adsorption process (Taty-Costodes et al., 2003). Therefore, the effect of hydrogen ion concentration was examined from solutions at pH ranging from 2 to 9. Fig. 2 summarizes the uptake of Cu (II) shells of LS, WS and RS used as adsorbent at various pH values.

At lower pH values, the H_3O^+ ions compete with the metal ions for the exchange sites in the adsorbent. There is an increase in metal ions uptake as the pH value increases from 2 to 6. The peak percentage adsorption of Cu (II) was attained at pH 6.0. After that, the sharp decrease adsorption of Cu (II) is due to the Cu(II) begins to precipitate as $Cu(OH)_2$.

One of the reasons for the metal ions adsorption behavior of the bio-adsorbent (LS, WS, and RS) is that the adsorbents surface contains a large number of active

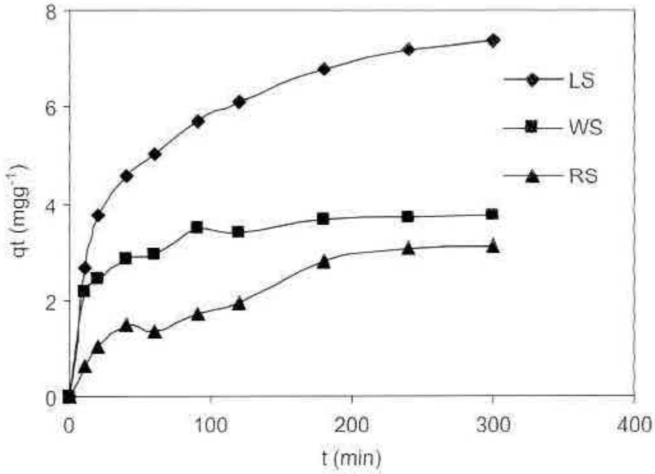


Fig. 1. Effect of contact time on the adsorption of Cu (II).

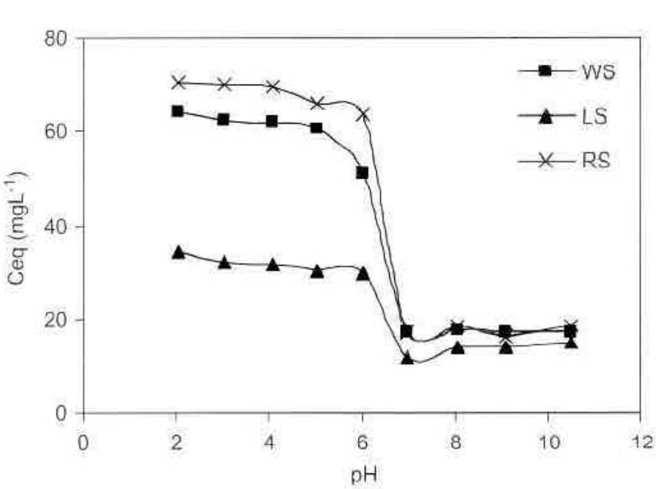


Fig. 2. Effect of pH on the adsorption of Cu (II).

sites. The metal uptake depends on these active sites as well as on the nature of the metal ions in solution. The affinities of the surface sites of LS, WS and RS depend on the nature of ions. This may be further explained in relation to a competition effect between the H_3O^+ and Cu (II) ions. At low pH values, the concentration of metal ions, hence the former ions occupy the binding sites on the LS, WH and RS, leaving metal ions free in solution. When the pH increases, the concentration H_3O^+ ions decreases, and the sites on the adsorbent surface mainly turn into dissociated forms (Fig. 3a) and can exchange H_3O^+ ions with metallic ions in solution (Fig. 3b) (Taty-Costodes et al., 2003).

In other respects, the property of the adsorbents to uptake metal ions can also be explained by the carbonyl groups (C=O) or the hydroxyl groups (OH) of polyphenols. The oxygen of each carbonyl and hydroxyl group is considered a strong Lewis base because of the presence of its vacant double electrons. Thanks to this doublets, the oxygen base makes a complex of coordination with the chemical entities low in electrons (metal ions, for example) (Fig. 4) (Prasad and Freitas, 2000)

3.7. Effect of adsorbent dosage

Adsorbent dosage is an important parameter because it determines the capacity of an adsorbent for a given initial concentration of the adsorbate. The effect of adsorbent dosage (adsorbent prepared in different batches) was studied on Cu(II) removal, keeping all other experimental conditions constant (Table 3). The results show that, as the adsorbent concentration increases, percentage adsorption generally increases, but the amount adsorbed per unit mass of the adsorbent decreases considerably. The decrease in unit adsorption with increase in the dose of adsorbent is basically due to adsorption sites remaining unsaturated during the adsorption reaction (Raji and Anirudhan, 1997).

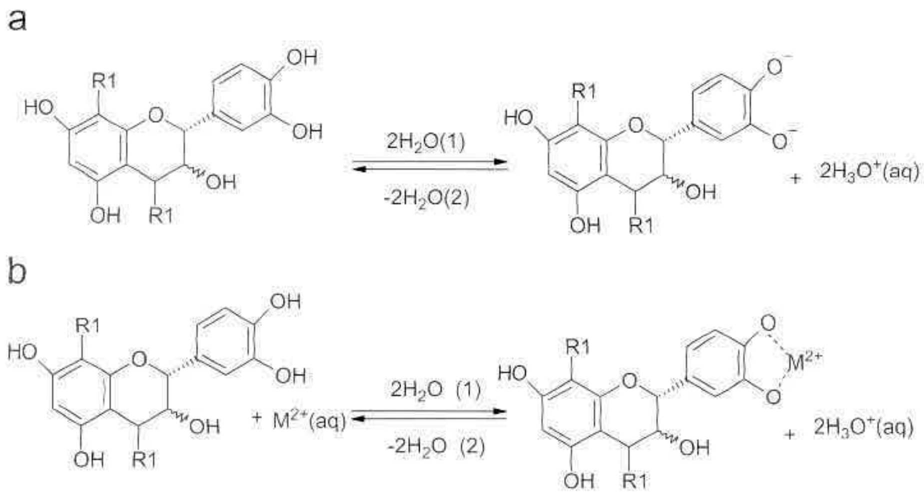


Fig. 3. Mechanism of biosorption: (a) represents the first stage of ion-exchange (deprotonation), while (b) shows attachment (adsorption) of the metal cations onto the deprotonated active sites on the sawdust surface (Taty-Costodes et al., 2003). Symbol M is a metal ion of charge 2^+.

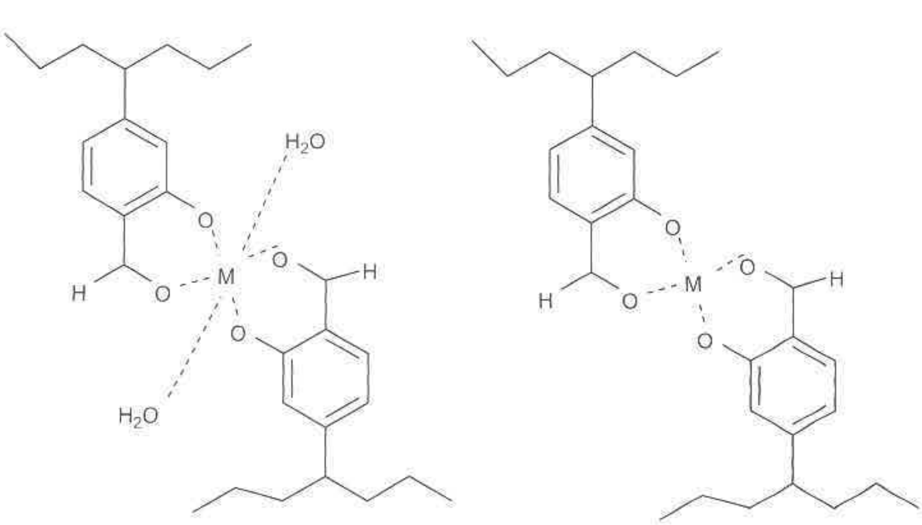


Fig. 4. Formation of metal complexes with the functional groups on the adsorbent surface.

Table 3
Effect of adsorbent dosage on the adsorption of Cu (II) with the different initial concentrations ($T = 293$ K, contact time = 3 h)

Adsorbent	Adsorbent dosage (g)	$C_o = 100\,\mathrm{mg\,L^{-1}}$		$C_o = 200\,\mathrm{mg\,L^{-1}}$		$C_o = 500\,\mathrm{mg\,L^{-1}}$	
		q_{eq} (mg g^{-1})	% A	q_{eq} (mg g^{-1})	% A	q_{eq} (mg g^{-1})	% A
	0.5	9.56	47.79	9.40	23.50	15.62	15.62
	1.0	6.72	67.20	8.44	42.20	14.54	29.08
LS	2.0	3.92	78.48	5.18	51.78	14.56	58.22
	4.0	2.14	85.76	2.68	53.60	10.38	83.07
	8.0	1.09	87.36	1.37	54.62	5.57	89.13
	0.5	6.45	21.23	6.50	16.25	11.00	11.00
	1.0	2.71	22.05	5.64	28.20	9.14	18.28
WS	2.0	1.19	23.74	4.97	49.65	4.74	18.94
	4.0	0.61	24.35	2.29	45.8	4.29	34.28
	8.0	0.39	31.10	1.17	46.70	3.21	51.28
	0.5	0.46	2.32	0.72	1.80	1.78	1.78
	1.0	0.74	7.39	0.28	1.40	3.60	7.20
RS	2.0	1.09	21.81	0.93	9.30	2.33	9.30
	4.0	0.89	35.40	0.91	18.25	1.45	11.62
	8.0	0.44	35.23	0.72	28.65	0.93	14.92

3.8. *Effect of initial Cu(II) concentration*

The equilibrium sorption capacity of the biomass for Cu(II) ion increased with a rise in initial ion concentration, as also shown in Table 3. Ion removal is higly concentration dependent. This increase in loading capacity of the bioadsorbent with relation to metal ions concentration is probably due to a high driving force for mass transfer. In fact, the more concentrated the solution is, the better the adsorption. At 293 K, when the initial metal ions concentration was increased from 100 to 500 mg L^{-1}, the loading capacity of dried adsorbent increased from 6.72 to 14.54 mg of Cu(II) per gram of LS, 2.71 to 9.14 mg of Cu(II) per gram of WS and 0.74 to 3.60 mg of Cu(II) per gram of RS (Table 3).

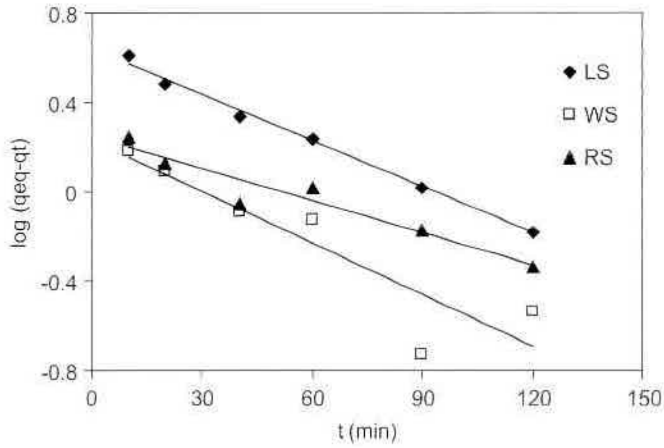


Fig. 5. Pseudo first-order sorption kinetics of Cu (II) on LS, WS and RS.

3.9. *Adsorption kinetics*

It is well known that the rate of the sorption process is modified by several parameters such as the structural properties of the adsorbent (i.e. porosity, specific area, particle size, etc.), the properties of the metallic ions (ionic radius, number of coordination, and speciation), metallic ions concentration, chelates formation between metallic ions, and the adsorbent, etc. This is why, intraparticular diffusion was also studied in the present work. Temperature and stirring speed were constant for all experiments. In order to determine the adsorption kinetics of Cu(II) ions, the first-order and second-order kinetics models were checked. The first-order rate expression of Lagergren based on solid capacity is generally expressed as follows (Taty-Costodes et al., 2003).

$$\log(q_{eq} - q_t) = \log q_{eq} - k_{pf}t/2.303, \tag{1}$$

where q_{eq} is the amount adsorbed (mg g^{-1}) at equilibrium, q_t the amount adsorbed (mg g^{-1}) at time t and k_{pf} the adsorption rate constant (min^{-1}). The straight line of the plot of $\log(q_{eq}-q_t)$ versus time (Fig. 5) suggest the applicability of the Lagergren equation for the present system (Bulut and Tez, 2003). The values of k_{pf} were determined from the slope of the plots and are given in Table 4.

Recently, Ho and Mc Kay (Mc Kay and Ho, 1999) have reported that most of the sorption systems followed a second-order kinetic model which can be expressed as

$$t/q_t = 1/k_{ps}q_{eq}^2 + t/q_{eq}, \tag{2}$$

where k_{ps} is the adsorption rate constant (g mg^{-1} min^{-1}). The plot of t/q_t versus time (Fig. 6) is similar as shown in (Weber and Morris).

The values of k_{ps} were determined by the slope of the plot and are given in Table 4.

The rate constant for intraparticle diffusion (k_{id}) is given as (Doğan et al., 2004):

$$q_t = k_{id}t^{1/2}, \tag{3}$$

Table 4
Adsorption kinetic parameters of Cu (II) on LS, WS and RS ($T = 293$ K, adsorbent $= 20\,\mathrm{g\,L^{-1}}$, $C_o = 100\,\mathrm{mg\,L^{-1}}$)

Adsorbent	First-order kinetic model			Second-order kinetic model			Intraparticle diffusion	
	k_{pf} (min^{-1})	q_{eq} (mg g^{-1})	R^2	k_{ps} (g mg^{-1} min^{-1})	q_{eq} (mg g^{-1})	R^2	k_{id} (mg g^{-1} min$^{-1/2}$)	R^2
LS	0.016	4.449	0.993	4.33×10^{-3}	7.962	0.997	0.343	0.995
WS	0.018	1.707	0.834	18.44×10^{-3}	3.925	0.999	0.149	0.937
RS	0.011	1.747	0.917	3.189×10^{-3}	3.88	0.933	0.187	0.938

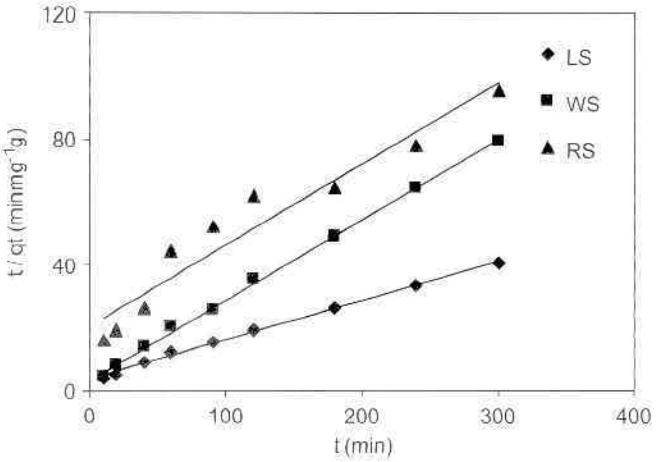


Fig. 6. Pseudo second-order sorption kinetics of Cu (II) on LS, WS and RS.

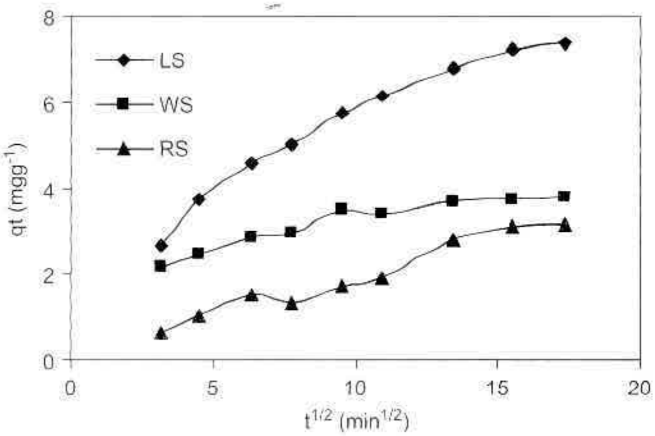


Fig. 7. Intraparticle diffusion plots of Cu (II) on LS, WS and RS.

where q_t is the amount adsorbed (mg g^{-1}) at time t (min). The plots of q_t versus $t^{1/2}$ for different adsorbents are shown in Fig. 7. All plots have a similar general trend and initial curved part, followed by a linear one and then a plateau. The initial curved part may be attributed to the bulk diffusion; the linear one to the intraparticle diffusion and the plateau to the equilibrium. This indicates that the transport of Cu(II) ions from solution through the particle solution interface into the pores of the particles as well as the adsorption on the available surface of adsorbents are both responsible for the uptake of Cu(II) ions (Yu et al., 2000).

k_{id} (mg g^{-1} min$^{-1/2}$) is characteristic for the rate of adsorption in the region, where intraparticle diffusion is rate-controlling. The extrapolation of the linear curve parts of the plots back to the axis provides the intercepts, which are proportional to the extent of the boundary layer thickness, i.e. the larger the intercept, the greater the boundary layer effect (Khraisheh et al., 2004). The deviation of the curves from the origin also indicates that intraparticle transport is not the only rate limiting step (Yu et al., 2000). The values of rate constants (k_{id}) of LS, WS and RS were 0.343, 0.149, and 0.187 mg g^{-1} min$^{-1/2}$, respectively (Table 4).

3.10. Adsorption model

The equilibrium adsorption of Cu (II) was studied as a function of concerntration (the initial Cu(II) concentration is in the range of 25–500 mg L^{-1}). The amount of Cu (II) adsorbed (q_{eq}) has been plotted against the equilibrium conceration (C_{eq}) for the adsorbents and is shown in Fig. 8 for Cu (II).

Several isotherm models are available. In this study the Langmuir and Freundlich adsorption isotherms were employed (Magdy and Daifullah, 1998).

The Langmuir isotherm can be represented by following equation:

$$q_{eq} = Q_m b C_{eq}/1 + bC_{eq}, \tag{4}$$

where q_{eq} (mg g^{-1}) is the amount of Cu (II) adsorbed per unit mass of adsorbent particles at equilibrium and C_{eq} (mg L^{-1}) is the equilibrium liquid phase concentration of Cu (II), b is the equilibrium constant (L mg^{-1}) and Q_m is the amount of adsorbate required to form a monolayer (mg g^{-1}).

The linearized equation of Langmuir is represented as follows:

$$C_{eq}/q_{eq} = 1/bQ_m + C_{eq}/Q_m. \tag{5}$$

Hence, a plot C_{eq}/q_{eq} versus C_{eq} should be a straight line with a slope $1/Q_m$ and intercept as $1/Q_m b$. Their values are summarized in Table 5.

The Freundlich model was also used to model the observed phenomena as given by the following equation:

$$q_{eq} = K_f C_{eq}^{1/n}, \tag{6}$$

where q_{eq} is the amount of Cu (II) adsorbed per unit mass of the adsorbent and n, K_f are constants. Eq. (6) may be

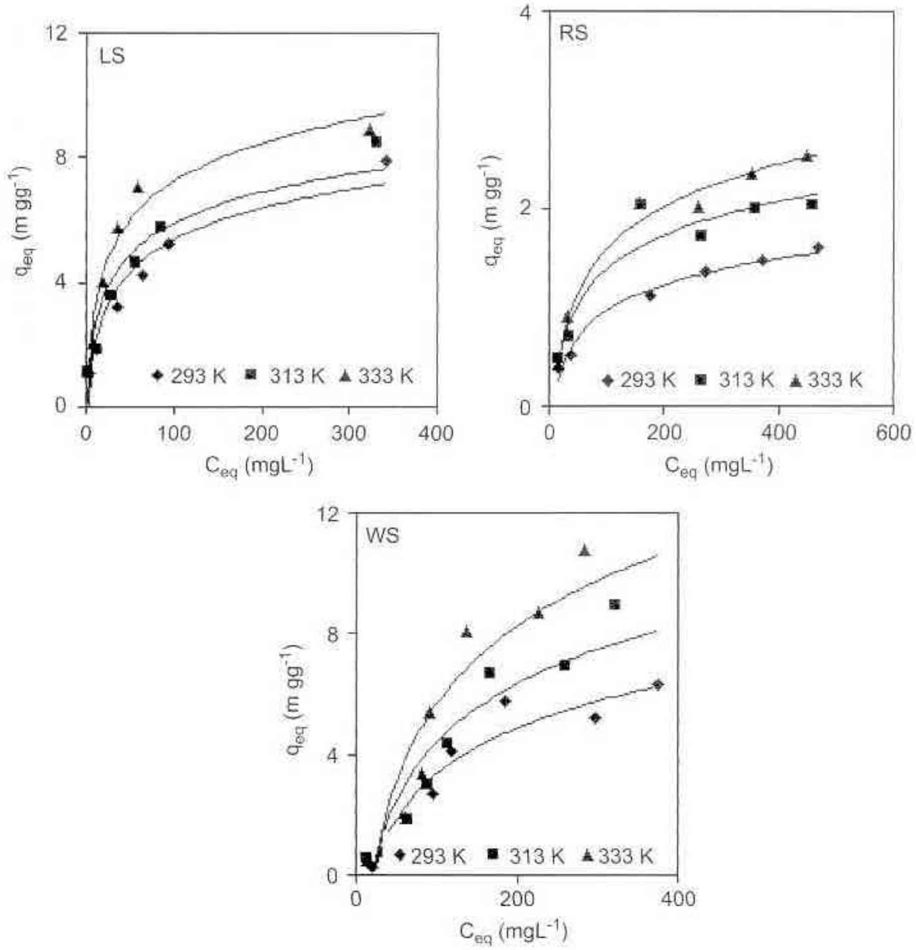


Fig. 8. Isotherms of Cu (II) adsorption from solution at different temperatures.

Table 5
Freundlich and Langmuir isotherm constants for adsorption of Cu (II) onto LS, WS and RS

Adsorbent	T (K)	Freundlich costants			Langmuir constants			
		K	n	R^2	Q_m (mg g^{-1})	b (L mg^{-1})	b (L mol^{-1})	R^2
	293	0.648	2.769	0.993	8.977	0.019	1207	0.980
LS	313	1.202	2.578	0.964	9.510	0.022	1398	0.981
	333	1.138	2.867	0.901	9.588	0.041	2605	0.999
	293	0.019	10.909	0.912	7.391	0.002	127	0.925
WS	313	0.034	9.594	0.955	16.077	0.004	254	0.791
	333	0.022	13.428	0.945	17.422	0.005	318	0.883
	293	0.108	2.787	0.996	1.854	0.011	699	0.991
RS	313	0.148	2.816	0.929	2.314	0.016	1017	0.982
	333	0.130	3.188	0.938	2.954	0.018	1143	0.987

linearized via a logarithmic plot as follows:

$$\log q_{eq} = \log K_F + (1/n) \log C_{eq}. \quad (7)$$

The parameters K_F and n were calculated and the results are presented in Table 5.

Adsorption isotherms of Cu (II) on LS, WS and RS are shown in Fig. 8. The Langmuir isotherms for RS and LS are obeyed better than the Freundlich isotherms at 313 and 333 K, which is evident from the value of regression coefficients (R^2). The adsorption isotherm of WS is obeyed by the Freundlich isotherm at 313 and 333. The adsorption of Cu (II) increased with increase in temperature (Table 5). Langmuir model is an indication of surface homogeneity of the adsorbent.

Table 6
Previously reported adsorption capacities of various low-cost adsorbents for Cu (II)

Low-cost adsorbents	Adsorption capacity ($mg\,g^{-1}$)	Reference
Diatomite	27.55	Gaballah et al., 1997
Modified diatomite	55.56	Gaballah et al., 1997
Fly ash	1.54	Panday et al., 1985
Sawdust	1.79	Yu et al., 2000
Peat	12.07	Babel and Kurniawan, 2003
Low-rank Turkish coals	1.62	Karabulut et al., 2000
Tea-industry waste	8.64	Çay et al., 2004
LS (Shells of lentil)	9.59	In this study
WS (Shells of wheat)	17.42	In this study
RS (Shells of rice)	2.95	In this study

Table 7
Values of thermodynamic parameters for the adsorption of Cu (II) on LS, WS and RS

Adsorbent	T (K)	$-\Delta G^0$ ($kJ\,mol^{-1}$)	ΔH^0 ($kJ\,mol^{-1}$)	ΔS^0 ($J\,mol^{-1}\,K^{-1}$)
	293	17.286		
LS	313	18.848	15.373	111
	333	21.775		
	293	11.800		
WS	313	14.401	18.791	105
	333	15.953		
	293	15.956		
RS	313	18.021	10.054	89
	333	19.493		

As seen in Table 5 and Fig. 8, the maximum adsorption capacities for Cu (II) on LS, WS and RS adsorbents at 293, 313 and 333 K temperature were found to be 8.977, 9.510 and 9.588; 7.391, 16.077 and 17.422; 1.854, 2.314 and 2.954 $mg\,g^{-1}$, respectively.

By comparison of the results obtained from this study to the previously reported work (Table 6) on adsorption capacities of various low-cost adsorbent in aqueous solution for Cu (II) ions, it can be stated that our findings are well.

The free energy of adsorption (ΔG^0) can be related with the equilibrium constant K ($L\,mol^{-1}$), corresponding to the reciprocal of the Langmuir constant, b, by the following equation (Silva et al., 2004):

$$\Delta G^0 = -RT \ln b, \tag{8}$$

where R is the gas universal constant ($8.314\,J\,mol^{-1}\,K^{-1}$) and T is the absolute temperature. Also, enthalpy (ΔH^0) and entropy (ΔS^0) changes can be estimated by the following equations, respectively.

$$\ln b = \Delta S^0/R - H^0/RT \tag{9}$$

Thus, a plot of $\ln b$ versus $1/T$ should be a straight line. ΔH^0 and ΔS^0 values were obtained from the slope and intercept of this plot, respectively. ΔG^0, ΔH^0 and ΔS^0 were obtained from the Eqs. (8) and (9) and are given in Table 7. Negative values of ΔG^0 indicate the feasibility of the process and spontaneous nature of the adsorption with a high preference of Cu (II) for all adsorbents. Positive value of ΔH^0 indicates the endothermic nature of the process, while positive value of ΔS^0 reflects the affinity of the adsorbents for the Cu (II) ions and suggests some structual changes in adsorbate and adsorbent (Panday et al., 1985).

4. Conclusion

The following conclusions can be drawn based on investigation of Cu (II) removal by LS, WS and RS.

First, LS, WS and RS appears to be a promising adsorbent for removal of Cu (II) from aqueous solution. At these adsorption levels, a process using LS, WS, and RS for the removal and recovery of a heavy metal is potentially more economical than current process technology.

Secondly, adsorption of Cu (II) is dependent on its initial concentrations, the amount of adsorbent, time of contact, temperature and pH of the metal solution, maximum removal of Cu (II) on LS, WS, and RS are at pH about 6.0. The removal of Cu (II) increased with increase in temperature.

Thirdly, isothermal data of Cu (II) adsorption on adsorbents can be modeled by both Freundlich and Langmuir isotherm. The capacity of adsorbents for adsorption of Cu (II) can be calculated by using these models. The maximum adsorption capacities for Cu (II) on LS, WS, and RS adsorbents at 293, 313, and 333 K temperature were found to be 8.977, 9.510, and 9.588; 7.391, 16.077, and 17.422; 1.854, 2.314, and 2.954 $mg\,g^{-1}$, respectively. The maximum adsorption capacities for other metals may be very different.

As it can be seen from these values, removal efficiency of the adsorbents increases and follows the order WS > LS > RS.

The results show that adsorbents which have a very low economical value may be used effectively for removal of Cu (II) ions from aqueous systems for environmental cleaning purpose.

Acknowledgments

We extend our thanks to the Dicle University, Research and Project Coordination for their support to our Project (No: DÜPAK-FF-02-105).

References

Adib, F., Bagreev, A., Bandosz, TJ., 1999. Effect of pH and surface chemistry on the mechanism of H_2S removal by activated carbons. Journal of Colloid and Interface Science 216, 360–369.

Ajmal, M., Khan, A.H., Ahmad, S., Ahmad, A., 1998. Role of sawdust in the removal of copper(II) from industrial wastes. Water Research 32 (10), 3085–3091.

Ajmal, M., Rao, R.A.K., Anwar, J.A., Ahmad, R., 2003. Adsorption studies on rice husk: removal and recovery of Cd (II) from wastewater. Bioresource Technology 86, 147–149.

Al-Asheh, S., Duvnjak, Z., 1998. Binary metal sorption by pine bark: study of equilibria and mechanisms. Seperation Science and Technology 33 (9), 1303–1329.

Al-Asheh, S., Banat, F., Mohai, F., 1999. Sorption of copper and nickel by spent animal bones. Chemosphere 39 (12), 2087–2096.

Babel, S., Kurniawan, T.A., 2003. Low-cost adsorbents for heavy metals uptake from contaminated water: a review. Journal Hazardous Materials B 97, 219–243.

Bailey, S.E., Olin, T.J., Bricka, R.M., Adrian, D.D., 1999. A review of potentially low-cost sorbents for heavy metals. Water. Research 33 (11), 2469–2479.

Bulut, Y., Baysal, Z., 2006. Removal of Pb (II) from wastewater using wheat brean. Journal of Environmental Management 78, 107–113.

Bulut, Y., Tez, Z., 2003. Removal of heavy metal ions by modified sawdust of walnut. Fresenius Environmental Bulletin 12 (12), 1499–1504.

Çay, S., Uyanık, A., Özışık, A., 2004. Single and binary component adsorption of copper (II) and cadmium (II) from aqueous solutions using tea-industry waste. Seperation and Purification Technology 38 (3), 273–280.

Doğan, M., Alkan, M., Türkyılmaz, A., Özdemir, Y., 2004. Kinetics and mechanism of removal of methylene blue by adsorption onto perlite. Journal of Hazardous Materials B 109, 141–148.

Ferro-Garcia, M.A., Rivera-Utrilla, J., Rodriquez-Gordillo, J., Autista-Toledo, I., 1988. Adsorption of zinc, cadmium and copper on activated carbon obtained from agricultural by-products. Carbon 26, 363–373.

Gaballah, I., Goy, D., Allain, E., Kilbertus, G., Thauront, J., 1997. Recovery of copper through decontamination of synthetic solutions using modified barks. Metallurgical and Materials Transactions B 28B, 13–23.

Gloaguen, V., Morvan, H., 1997. Removal of heavy metal ions from aqueous solution by modified barks. Journal of Environental Science and Health Part A 32 (4), 901–912.

Karabulut, S., Karabakan, A., Denizli, A., Yürüm, Y., 2000. Batch removal of copper(II) and zinc(II) from aqueous solutions with low-rank Turkish coals. Seperation and Purification Technology 18, 177–184.

Khraisheh, M.A.M., Al-degs, Y.S., Mchinn, W.A.M., 2004. Ramediation of wastewater containing heavy metals using raw and modified diatomite. Chemical Engineering Journal 99, 177–184.

Low, K.S., Lee, C.K., Liew, S.C., 2000. Sorption of cadmium and lead from aqueous solutions by spent grain. Process Biochemistry 36, 59–64.

Magdy, Y.H., Daifullah, A.A.M., 1998. Adsorption of a basic dye from aqueous solutions onto sugar-industry-mud in two modes of operations. Waste Management 18, 219–226.

Mc Kay, G., Ho, Y.S., 1999. Pseudo-second order model for sorption processes. Process Biochemistry 34, 451–465.

Okieimen, F.E., Ogbeifun, D.E., Nwala, G.N., Kumsah, G.A., 1985. Binding of cadmium, copper, and lead ions by modified cellulosic materials. Bulletin of Environmental Contamination and Toxicology 34, 866–870.

Panday, K.K., Prasad, G., Singh, V.N., 1985. Cu (II) removal from aqueous solutions by fly ash. Water Research 19 (7), 869–873.

Prasad, M.N.V., Freitas, H., 2000. Removal of toxic metals from solution by leaf, stem and root phytomass of *Quercus ilex* L. (holly oak). Environmental Pollution 110, 277–283.

Raji, C., Anirudhan, T.S., 1997. Kinetics of Pb(II) adsorption by polyacrylamide grafted sawdust. Indian Journal of Chemical Technology 4, 157–162.

Ricordel, S., Taha, S., Cisse, I., Dorange, G., 2001. Heavy metals removal by adsorption onto peanut husks carbon: characterization, kinetic study and modeling. Seperation and Purification Technology 24, 389–401.

Silva, J.P., Sousa, S., Rodrigues, J., Antunes, H., Porter, J.J., Gonçalves, I., Dias, S.F., 2004. Adsorption of acid orange 7 dye in aqueous solutions by spent brewery grains. Seperation and Purification Technology 40 (3), 309–315.

Taty-Costodes, V.C., Fauduet, H., Porte, C., Delacroixs, A., 2003. Removal of Cd (II) and Pb(II) ions from aqueous solutions by adsorption onto sawdust of Pinus Sylvestris. Journal of Hazardous Materials 105 (1-3), 121–142.

Weber Jr., W.J., Morris, C.J., Advances in water pollution research. In: Proceedings of the First International Conference On Water Pollution Research, vol. 2. Pergamon Press, Oxford, p. 261.

Yu, B., Zhang, Y., Shukla, A., Shukla, S.S., Dorris, K.L., 2000. The removal of heavy metal from aqueous solutions by sawdust adsorption-removal of copper. Journal of Hazardous Materials Part B 80, 33–42.

Journal of Environmental Management 87 (2008) 46–58

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Immobilization of Pb(II), Cd(II) and Ni(II) ions on kaolinite and montmorillonite surfaces from aqueous medium

Susmita Sen Gupta[a], Krishna G. Bhattacharÿya[b,*]

[a]Department of Chemistry, B N College, Dhubri 783324, Assam, India
[b]Department of Chemistry, Gauhati University, Guwahati 781014, Assam, India

Received 23 November 2005; received in revised form 4 December 2006; accepted 4 January 2007
Available online 11 May 2007

Abstract

The present study investigates the immobilization of Pb(II), Cd(II) and Ni(II) on clays (kaolinite and montmorillonite) in aqueous medium through the process of adsorption under a set of variables (concentration of metal ion, amount of clay, pH, time and temperature of interaction). Increasing pH favours the removal of metal ions till they are precipitated as the insoluble hydroxides. The uptake is rapid with maximum adsorption being observed within 180 min for Pb(II) and Ni(II) and 240 min for Cd(II). A number of available models like the Lagergren pseudo first-order kinetics, second-order kinetics, Elovich equation, liquid film diffusion and intra-particle diffusion are utilized to evaluate the kinetics and the mechanism of the immobilization interactions. Two isotherm equations due to Langmuir and Freundlich showed good fits with the experimental data. Kaolinite and montmorillonite have considerable Langmuir monolayer capacity with respect to Pb(II), Cd(II) and Ni(II), the values being in the range of 6.8–11.5 mg/g (kaolinite) and 21.1–31.1 mg/g (montmorillonite). The Freundlich adsorption capacity follows a similar order. The thermodynamics of the immobilization process indicates the same to be exothermic with Pb(II) and Ni(II), but endothermic with Cd(II). The interactions with Pb(II) and Ni(II) are accompanied by decrease in entropy and Gibbs energy while the endothermic immobilization of Cd(II) is supported by an increase in entropy and an appreciable decrease in Gibbs energy. The results have established good potentiality for kaolinite and montmorillonite to remove heavy metals like Pb(II), Cd(II) and Ni(II) from aqueous medium through adsorption-mediated immobilization.

Keywords: Kaolinite; Montmorillonite; Adsorption kinetics; Metal ions; Enthalpy

1. Introduction

Heavy metals are detrimental to the environment because of their non-biodegradable and persistent nature. The toxicity of these metals is enhanced through accumulation in living tissues and consequent biomagnification in the food chain (An et al., 2001). Lead is an important metal from the viewpoint of environmental toxicology as the metal finds its way into air and water from a multitude of sources (ATSDR, 1999a) like lead smelter, battery manufacturer, paper and pulp industry, boat and ship fuel and ammunition industry, etc. Lead toxicity effects on the nervous system, the blood circulation system, cardiovascular system, vital organs like the brain and the kidneys, and on restricted development of IQ, etc. have been well documented (Marino et al. 1989).

The toxic effects of cadmium have received as much attention as those of lead. Its sources in the environment are metal plating, smelting and mining industries, cadmium-nickel battery manufacture, phosphate fertilizer, paints and pigments manufacture, and alloy industries (Kadirvelu and Namasivayam, 2003). Cadmium poisoning leads to lung, liver and kidney damage, bone lesions, cancer and hypertension, and the dreaded itai-itai disease, which are accentuated in case of calcium deficiency (ATSDR, 1999b). The other heavy metal, Ni(II) is released into the atmosphere during nickel mining and by industries that make or use nickel, nickel alloys, or nickel compounds, oil-burning power plants, coal-burning power

*Corresponding author. Tel.: +91 361 2571529; fax: +91 361 2570599.
E-mail address: krishna2604@sify.com (K.G. Bhattacharyya).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.048

plants, and trash incinerators. Chronic bronchitis, reduced lung function, and cancer of the lung and nasal sinus, etc., are consequences of Ni(II) contamination. Carcinogenicity of nickel has also received international attention (ATSDR, 2003).

Adsorption is one of the most effective and economical techniques used to remove heavy metals from water (Tran et al., 1999). Adsorption immobilizes the contaminants separating them from the aqueous phase and the contaminant-loaded adsorbent can be safely disposed off or the contaminant could be recovered. The clay minerals, being important constituents of soil, have been playing this role always by taking up various contaminants as water flows over soil or penetrates underground. The process of immobilization takes place either through ion exchange or adsorption, or a combination of both. Clays are hydrous aluminosilicates which make up the colloid fraction ($<2\,\mu m$) of soil, sediment, rocks and water (Pinnavaia, 1983). The high specific surface area, chemical and mechanical stability, layered structure, high cation exchange capacity (CEC), Brönsted and Lewis acidity, etc., have made the clays excellent materials for adsorption (Tanabe, 1981).

A large number of studies have been reported on use of clays for metal ion removal from aqueous solution: Zn(II) with natural bentonite (Mellah and Chegrouche, 1997), and illite (Echeverria et al., 2002), Mn(II), Co(II), Ni(II) and Cu(II) with raw kaolin (Yavuz et al., 2003), Co(II) with sepiolite (Kara et al., 2003), etc. Various forms of modified clays have also been used: Na-exchanged bentonites for Cr(III), Ni(II), Zn(II), Cu(II) and Cd(II) (Alvarez-Ayuso and Garcia-Sanchez, 2003), surfactant modified montmorillonite for Cu(II) and Zn(II) (Lin and Juang, 2002), 1:10 phenanthroline-grafted Brazilian bentonite for Cu(II) (De Leon et al., 2003).

The present work was undertaken to explore the feasibility of using kaolinite and montmorillonite for removal of Pb(II), Cd(II) and Ni(II) in single batch system from aqueous solution by adsorption under various environmental conditions.

2. Materials and methods

2.1. Adsorbate solutions

The adsorption experiments were conducted by using synthetic effluents containing Pb(II), Cd(II) and Ni(II) separately. The stock solution containing 1000 mg of Pb(II), Cd(II) and Ni(II) per liter was prepared by dissolving Pb $(NO_3)_2$ (Glaxo, Mumbai, India), Cd $(NO_3)_2 \cdot 4H_2O$ (Qualigens, Mumbai, India) and $Ni(NO_3)_2 \cdot 6H_2O$ (Qualigens, Mumbai, India), respectively, in 1 l of double distilled water and were used to prepare the adsorbate solutions by appropriate dilution. The pH of the aqueous solution of Cd(II) as prepared was 5.5 while that of the aqueous solution of Ni(II) and Pb(II) as prepared was 5.7.

2.2. Clay adsorbents

Kaolinite, KGa-1b (K) and Montmorillonite, SWy-2 (M) were obtained from the University of Missouri-Columbia, Source Clay Minerals Repository, USA. Both the clays were calcined at 773 K for 10 h before using them as adsorbents.

2.3. Characterization of clay adsorbent

2.3.1. XRD measurement

XRD measurements were taken with Phillips Analytical X-ray spectrometer (PW 1710) using CuKα radiations.

2.3.2. Surface area

The surface areas for the clay adsorbents were estimated utilizing the Sears' method (Sears, 1956). A sample containing 0.5 g of clay was acidified with 0.1 N HCl to pH 3.0–3.5. The volume was made up to 50 ml with distilled water after addition of 10.0 g of NaCl. The titration was carried out with standard 0.1 M NaOH in a thermostatic bath at 298 ± 0.5 K to pH 4.0, and then to pH 9.0. The volume, V, required to raise the pH from 4.0 to 9.0 was noted and the surface area was computed from the following equation:

$$S(m^2/g) = 32V - 25. \tag{1}$$

2.3.3. Cation exchange capacity

The CEC of the clay adsorbents was estimated by using the copper bisethylenediamine complex method (Bergaya and Vayer, 1997). Fifty milliliter of 1 M $CuCl_2$ solution was mixed with 102 ml of 1 M ethylenediamine solution to allow for the formation of the $[Cu(en)_2]^{2+}$ complex. The slight excess of the amine ensures complete formation of the complex. The solution is diluted with water to one liter to give a 0.05 M solution of the complex. 0.5 g of a dry clay sample was mixed with 5 ml of the complex solution in a 100 ml flask, diluted with distilled water to 25 ml and the mixture was shaken for 30 min in a thermostatic water bath and centrifuged. The concentration of the complex remaining in the supernatant is determined by iodometric method. For this, 5 ml of the supernatant was mixed with 5 ml of 0.1 M HCl to destroy the $[Cu(en)_2]^{2+}$ complex and KI salt was added at 0.5 g per ml of solution. The mixture was titrated with 0.02 M $Na_2S_2O_3$ solution with starch as indicator. The CEC was calculated from the following formula:

$$CEC(meq/100\,g) = MSV(x - y)/1000m \tag{2}$$

such that M is the molar mass of the complex, S the concentration of the thio solution, V the volume (ml) of the complex taken for iodometric titration, m the mass of adsorbent taken (g), x the volume (ml) of thio required for blank titration (without the adsorbent), and y the volume (ml) of thio required for the titration (with the adsorbent).

2.4. Adsorption experiments

The batch adsorption experiments were carried out in 100 ml Erlenmeyer flasks by mixing clay and 50 ml aqueous solution of metal ions and agitating the mixture in a constant temperature water bath thermostat for a desired time interval. The mixture was centrifuged (Remi R 24, 20 000 rpm) and the metal ions remaining unadsorbed in the supernatant liquid were determined with Atomic Absorption Spectroscopy (Varian SpectrAA 220, air-acetylene oxidizing flame, Pb(II): lamp current 5 mA, wavelength 217.0 nm, slit width 1.0 nm, optimum working range 0.1–30.0 μg/ml; Cd(II): Lamp current 4 mA, wavelength 228.8 nm, slit width 0.5 nm, optimum working range 0.02–3.0 μg/ml, Ni(II): lamp current 4 mA, wavelength 232 nm, slit width 0.5 nm, optimum working range 0.02–3.0 μg/ml). The experiments were repeated with different clay amounts, adsorbate concentration, pH of the medium, interaction time and temperature. For maintaining pH of the medium, either 0.01 N NaOH or 0.01 N HNO_3 was added dropwise before carrying out adsorption and the pH was monitored both before and after adsorption. The following conditions were maintained for the different sets of experiments:

(i) Effects of pH:
Pb(II): Clay 2 g/l, Pb(II) 50 mg/l, temperature 303 K, interaction time 180 min, pH 1.0–6.0 at unit intervals
Cd(II): Clay 2 g/l, Cd(II) 50 mg/l, temperature 303 K, interaction time 240 min, pH 1.0–10.0 at unit intervals
Ni(II): Clay 2 g/l, Ni(II) 50 mg/l, temperature 303 K, interaction time 180 min, pH 1.0–10.0 at unit intervals
(ii) Kinetics:
Pb(II): Clay 2 g/l, Pb(II) 50 mg/l, temperature 303 K, pH 5.7, time 20, 40, 60, 90, 120, 150, 180, 240, 300, 360 min
Cd(II): Clay 2 g/l, Cu(II) 50 mg/l, temperature 303 K, pH 5.5, time 20, 40, 60, 90, 120, 180, 240, 300, 360 min
Ni(II): Clay 2 g/l, Ni(II) 50 mg/l, temperature 303 K, pH 5.7, time 20, 40, 60, 90, 120, 150, 180, 240, 300, 360 min
(iii) Effects of adsorbent amount:
Pb(II): Pb(II) 50 mg/l, temperature 303 K, pH 5.7, time 180 min, Clay 2, 3, 4, 5, 6 g/l
Cd(II): Cd(II) 50 mg/l, temperature 303 K, pH 5.5, time 240 min, Clay 2, 3, 4, 5, 6 g/l
Ni(II): Ni(II) 50 mg/l, temperature 303 K, pH 5.7, time 180 min, Clay 2, 3, 4, 5, 6 g/l
(iv) Effects of adsorbate concentration and adsorption isotherm:
Pb(II): Clay 2 g/l, temperature 303 K, pH 5.7, time 180 min, Pb(II) 10, 20, 30, 40, 50 mg/l
Cd(II): Clay 2 g/l, temperature 303 K, pH 5.5, time 240 min, Cd(II) 10, 20, 30, 40, 50 mg/l
Ni(II): Clay 2 g/l, temperature 303 K, pH 5.7, time 180 min, Ni(II) 10, 20, 30, 40, 50 mg/l
(v) Thermodynamics:
Pb(II): Clay 2 g/l, time 180 min, pH 5.7, temperature 303, 308, 313 K, Pb(II) 10, 20, 30, 40, 50 mg/l
Cd(II): Clay 2 g/l, time 240 min, pH 5.5, temperature 303, 308, 313 K, Cd(II) 10, 20, 30, 40, 50 mg/l
Ni(II): Clay 2 g/l, time 180 min, pH 5.7, temperature 303, 308, 313 K, Ni(II) 10, 20, 30, 40, 50 mg/l

The amount, q, of metal ions adsorbed per unit mass of the adsorbent and the extent of adsorption (%) are computed from the expression

$$q = (C_0 - C_t)/m,$$
$$\%\text{adsorption} = (C_0 - C_t) \times 100/C_0$$

where C_0 (mg/l) and C_t (mg/l) are metal ion concentrations before and after adsorption for time t, and m(g) is the amount of clay adsorbent taken for 1 l of solution.

2.5. Theoretical foundation

The adsorption equilibrium is usually described by an isotherm equation whose parameters express the surface properties and affinity of the adsorbent, at a pre-set temperature and pH. The following two widely used isotherms (Freundlich, 1906; Langmuir, 1918) are usually applied:

$$\text{(a)Freundlich isotherm}: q_e = K_f C_e^n, \tag{3}$$

$$\text{(b)Langmuir isotherm}: C_e/q_e = 1/(bq_m) + (1/q_m)C_e, \tag{4}$$

where C_e is the concentration of the adsorbate at equilibrium in the liquid phase and q_e is the corresponding concentration of the adsorbate in the solid phase. The Freundlich coefficients, K_f and n, are, respectively, related to adsorption capacity and adsorption intensity of the solid adsorbent. Similarly, b and q_m, are Langmuir coefficients representing the equilibrium constant for the adsorbate–adsorbent equilibrium and the monolayer capacity of the solid. The linear Freundlich and Langmuir plots are obtained by plotting (i) $\log q_e$ vs. $\log C_e$ and (ii) C_e/q_e vs. C_e, respectively, from which the adsorption coefficients could be evaluated.

The Langmuir equation is also used to obtain R_L, the dimensionless equilibrium parameter or the separation factor (Hall et al., 1966) from the expression

$$R_L = 1/(1 + bC_0), \tag{5}$$

where C_0 is the initial concentration of the adsorbate. The shapes of the isotherms are indicated by R_L values (Ho, 2003) as $0 < R_L < 1$ indicates favorable adsorption, and $R_L > 1$, $R_L = 1$ and 0 represent unfavorable, linear and irreversible isotherm, respectively.

Adsorption normally is two-stage process consisting of (a) rapid removal of the adsorbate from the liquid phase through adsorption, and (b) much slower uptake continuing till the equilibrium is established (Ho and McKay, 1999a). Kinetics for clay metal interaction was studied thoroughly in this work by applying the following equations and models. Lagergren equation (Lagergren, 1898; Ho, 2004) was applied first assuming pseudo first-order kinetics, where the number of metal ions outnumbers the number of adsorption sites on clay surface. The expression for the first-order rate constant, k_1, is given by the differential rate law

$$dq_t/dt = k_1(q_e - q_t) \tag{6}$$

which on integration under the boundary conditions of $t = 0$ to $t = t$ and $q_t = 0$ to $q_t = q_t$, gives a linear expression

$$\ln(q_e - q_t) = \ln q_e - k_1 t. \tag{7}$$

The values of k_1 can be obtained from the slope of the plot of $\log(q_e - q_t)$ vs. t. The validity of the first-order kinetics and hence the Lagergren equation could be tested by comparing q_e values obtained from the intercepts of the plots with those obtained. If the validity is weak, the kinetics can be tested for following second-order mechanism (Ho and McKay, 1999b; Ho et al., 2001) as per the rate law:

$$dq_t/dt = k_2(q_e - q_t)^2, \tag{8}$$

where k_2 is the second-order rate constant. For the boundary conditions, $t = 0$ to $t = t$ and $q_t = 0$ to $q_t = q_t$, the integrated linear form of the equation is written as

$$t/q_t = 1/(k_2 q_e^2) + (1/q_e)t. \tag{9}$$

The plot of t/q_t vs. t gives a linear relationship, allowing for computation of q_e and k_2. The validity can again be tested by comparing values of q_e obtained from the plots and from experiment.

When the adsorbate ions and the surface sites interact chemically through a second-order mechanism, the application of the Elovich equation, given by

$$dq_t/dt = \alpha \exp(-\beta q_t) \tag{10}$$

may be more appropriate (Ho and McKay, 1998). The Elovich coefficients, α and β, are known to represent the initial adsorption rate (g/mg min^2) and the desorption constant (mg/g min), respectively. These coefficients are computed from the plots of q_t vs. $\ln t$, the integrated form of Eq. (10).

For a porous substrate, there is also the possibility that the diffusion of adsorbate species into the pores is the rate determining process. The intra-particle diffusion rate constant (k_i) is given by the equation (Weber and Morris, 1963)

$$q_t = k_i t^{0.5}. \tag{11}$$

The plots of q_t vs. $t^{0.5}$ yield straight lines passing through the origin and the rate constant, k_i can be obtained from the slope. The mechanism of the interactions may also be governed by the slow transport of the adsorbate ions from the liquid phase up to the adsorbent surface. In such cases, the liquid film diffusion model (Boyd et al., 1947) represented as

$$\ln(1 - F) = -k_{fd} t \tag{12}$$

should have validity. F is the fractional attainment of equilibrium ($F = q_t/q_e$), k_{fd} is the adsorption rate constant. The plot of $-\ln(1-F)$ vs. t should be a straight line with zero intercept.

The thermodynamic parameters for the adsorption process, ΔH (kJ/mol), ΔS (J/K mol) and ΔG (kJ/mol) were evaluated using the equations (Abou-Mesalam, 2003)

$$\Delta G = -RT \ln K_d, \tag{13}$$

$$\Delta G = \Delta H - T\Delta S, \tag{14}$$

$$\ln K_d = \Delta S/R - \Delta H/RT, \tag{15}$$

where K_d is the distribution coefficient of the adsorbate ($= q_e/C_e$), T the absolute temperature (K), R (gas constant) $= 8.314 \times 10^{-3}$ kJ/K mol. The plots of $\ln K_d$ vs. $1/T$ could be utilized to find ΔH and ΔS which help to obtain ΔG.

3. Results and discussion

3.1. XRD study

The XRD spectra of the clays are given in Fig. 1. The calcined kaolinite (K) and montmorillonite (M) samples yielded, respectively, 13 and 7 peaks in the range of 1–30° (2θ). The basal spacing is 7.15 Å (2θ = 12.35°) and 10.09 Å (2θ = 8.75°) with an intensity of 69.7% and 12.05%, respectively, for kaolinite and montmorillonite and the

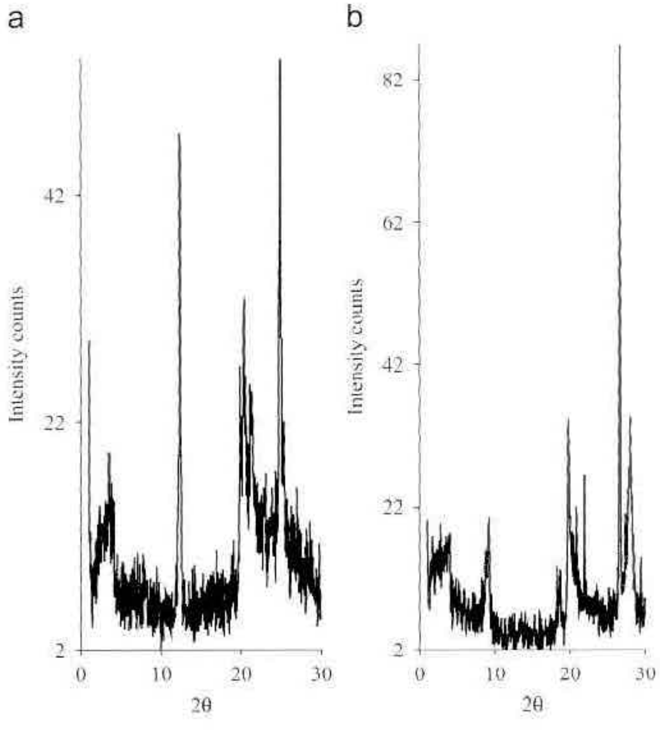


Fig. 1. XRD patterns for (a) kaolinite (K) and (b) montmorillonite (M).

corresponding tip widths were 0.10 ($2\theta = 12.35°$) and 0.24 ($2\theta = 8.75°$). These values are in agreement with the reported values (Grim, 1968, pp. 126–159) showing that the structural features of both kaolinite and montmorillonite have not undergone any significant change after calcination at 773 K (Grim, 1968, pp. 298–329; Ghosh and Bhattacharyya, 2002).

3.1.1. Surface area and cation exchange capacity of the clays

The specific surface areas of kaolinite (K) and montmorillonite (M) (both calcined at 773 K) were found to be 3.8 and 19.8 m^2/g, respectively, compared to the values of 3.1 and 18.7 m^2/g for the uncalcined kaolinite and montmorillonite. The surface area of montmorillonite is almost five times that of kaolinite and calcination has resulted in marginal increase in the specific surface area. The values are in agreement with those reported in the literature, i.e. 5–25 m^2/g for kaolinite (Volzone et al., 1999) and 15.5–82.0 m^2/g for montmorillonite (Grim, 1968, pp. 463–465; Ravichandran and Sivasankar, 1997). It is to be noted that specific surface area depends on particle size distribution, particle shape, and number distribution of cracks and pores in the material, and therefore, cannot be represented as a general characteristic of a particular type of material (Grim, 1968, pp. 463–465).

The CEC values for kaolinite were measured as 13.4 meq/100 g (uncalcined) and 11.3 meq/100 g (calcined), while those for montmorillonite were found to be 225.0 meq/100 g (uncalcined) and 153.0 meq/100 g (calcined). These values are within the range specified by Grim (1968, pp. 188–189). It is clear that heat treatment has resulted in a decrease in the CEC indicating that some of the cation exchange sites are lost.

3.2. Adsorption of Pb(II), Cd(II) and Ni(II)

All the adsorption experiments were done in triplicate under the same conditions and the concentrations of the adsorbate in the aqueous phase and in the solid phase were computed from the averages of experimentally determined metal ion concentration in the aqueous phase after adsorption. These values were then used in all subsequent treatment and analysis of data.

3.2.1. Effects of pH

Adsorption of Pb(II), Cd(II) and Ni(II) on kaolinite and montmorillonite showed a continuous increase right from pH 1.0 (Fig. 2). The experiments with Pb(II) could not be continued beyond pH 6.0 due to the low solubility of Pb(II) hydroxide, which has also been observed by other workers (e.g. Jain and Ram, 1997). With Ni(II), the amount adsorbed on unit mass increased slowly up to pH 8.0 after which there was a much rapid rise indicating precipitation of Ni(II)-hydroxide and related species (Bayat, 2002).

Cd(II) adsorption showed a gradual increase in the pH range of 1.0–10.0 with no indication of precipitation of

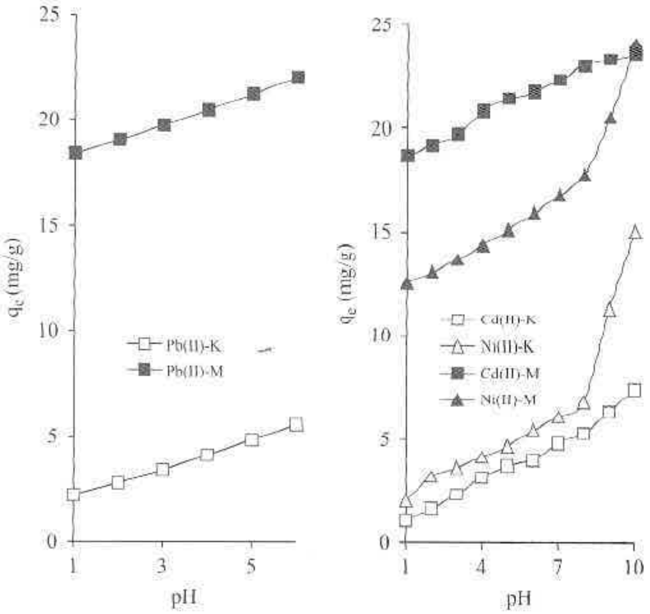


Fig. 2. Amount of metal ion adsorbed per unit mass (q_e) on clays at 303 K at different pHs (K kaolinite, M montmorillonite).

Cd(II)-hydroxide. This was in conformity with the results of a blank experiment without the presence of the clays. Mathialagan and Viraraghavan (2002) carried out adsorption of Cd(II) on perlite up to pH 9.0 without observing any significant precipitation of the hydroxide. Similar observation has been recorded by Evans et al. (2002) when carrying out Cd(II) adsorption on chitosan-based crab shells. It has been reported that Cd(II) uptake increases above pH 6.0 when the ions are likely to be bound strongly to hydroxyl groups on the edges of the clay minerals (Angove et al., 1998; Coles and Yong, 2002). However, in the present work, such high rate of adsorption of Cd(II) at pH > 6.0 was not observed.

The adsorption experiments were carried out down to pH 1.0 when the possibility of dissolution of the clay substrate had to be considered. Such possibility can be ruled out as the amount adsorbed showed the same pattern of variation even at the low pHs (as otherwise, with the clay gradually disappearing due to dissolution, adsorption would have come down more rapidly). This is in agreement with the observations of (i) Khan et al. (1995) on adsorption of Sr(II) on bentonite in the pH range 1.5–8.5, (ii) Lin and Juang (2002) for adsorption of Cu(II) and Zn(II) on sodium dodecylsulfate modified montmorillonite at pH 1.5 and above, and (iii) De Leon et al. (2003) who carried out adsorption of Cu(II) on pillared and delaminated bentonite in the pH range 1.5–12.5.

Fig. 2 shows that (i) montmorillonite takes up more of the metal ions than kaolinite at a particular pH (for example, the amount of Cd(II) adsorbed on montmorillonite was almost 17 times of that adsorbed on kaolinite at pH 1.0, but with increase in pH, this came down to ~3 times only at pH 10.0), (ii) uptake of Pb(II) on both the clays is more in comparison to Cd(II) and Ni(II),

(iii) kaolinite takes up more Ni(II) than Cd(II), while montmorillonite has the reverse order, i.e. it takes up more Cd(II) than Ni(II).

The variations in amount adsorbed with pH could be explained on the basis of competition between the metal ions and H_3O^+ ions for adsorption sites on clay surface. At very low pH, the number of H_3O^+ ions exceeds that of the metal ions several times and the surface is most likely covered with H_3O^+ ions, which account for less adsorption. It is also possible that oxygen atoms of the clay surface interact with water in an acidic medium forming some aqua complexes (Mathialagan and Viraraghavan, 2002) as follows:

$$-\mathrm{MO} + \mathrm{H} - \mathrm{OH} \xrightarrow{\mathrm{H}^+} \mathrm{M} - \mathrm{OH}_2^+ + \mathrm{OH}^-.$$

This surface charge interacts repulsively with approaching metal ions and prevents them from reaching the surface and thus, the adsorption was not much at low pH. In an alkaline medium, the clay surface becomes negatively charged favouring M^{+2} uptake

$$-\mathrm{MOH} + \mathrm{OH}^- \rightleftarrows -\mathrm{MO}^- + \mathrm{H_2O},$$

$$-\mathrm{MO}^- + \mathrm{M}^{2+} \rightleftarrows -\mathrm{M} - \mathrm{O}^- \cdots\cdots \mathrm{M}^{2+}.$$

When the pH increases, more and more H_3O^+ ions leave the clay surface making the sites available to the metal ions, which now increasingly bind to clay surface (Taty-Costodes et al., 2003) through a mechanism similar to that of exchange interactions (H^+/Pb^{+2}, H^+/Cd^{+2}, H^+/Ni^{+2}) (Singh et al., 1993). As the pH becomes alkaline, precipitation of the insoluble metal-hydroxides may appear as apparently higher metal ion removal (Yu et al., 2001; Bayramoglu et al., 2003).

3.2.2. *Effects of adsorbent amount and adsorbate concentration*

Amount of metal ion adsorbed per unit mass of clay decreased with high clay loading (Fig. 3). Similar results have been reported by other authors (e.g. Cu(II) by sawdust, Yu et al., 2000; Cr (III) by ion exchange resins, Rengaraj et al., 2003). This may be attributed to two reasons: (i) a large adsorbent amount effectively reduces the unsaturation of the adsorption sites and correspondingly, the number of such sites per unit mass comes down resulting in comparatively less adsorption at higher adsorbent amount, and (ii) higher adsorbent amount creates particle aggregation, resulting in a decrease in the total surface area and an increase in diffusional path length both of which contribute to decrease in amount adsorbed per unit mass (Shukla et al., 2002).

Amount of metal ions adsorbed per unit mass of clay adsorbents increased gradually with more and more metal ions in the adsorbate solution (Fig. 4). At low metal ion loading, the ratio of the number of metal ions to the number of available adsorption sites is small and consequently, adsorption is independent of initial concentration, but as the concentration of metal ions increases, the situation changes and the competition for adsorption sites becomes fierce. As a result, the extent of adsorption comes down considerably, but the amount adsorbed per unit mass of the adsorbent rises. Similar results have also been reported by other workers (e.g., Ni(II) by bagasse fly ash, Gupta et al., 2003).

All the results show that montmorillonite stands apart due to its large adsorption capacity with respect to all the three metals. Being 2:1 clay type, montmorillonite has high

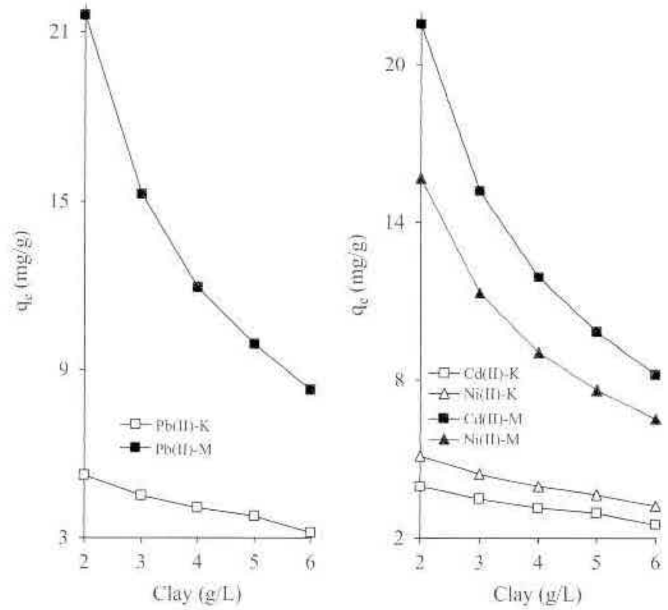


Fig. 3. Amount of metal ions adsorbed per unit mass (q_e) on clays for five different clay amounts (2, 3, 4, 5, 6 g/l) at 303 K (K kaolinite, M montmorillonite).

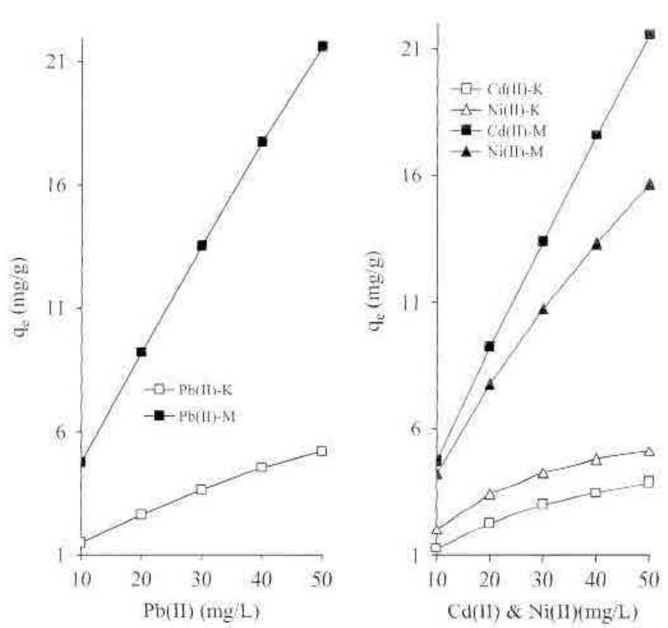


Fig. 4. Amount of metal ions adsorbed per unit mass (q_e) on clays for five different initial metal concentrations (10, 20, 30, 40, 50 mg/l) at 303 K (K kaolinite, M montmorillonite).

surface charges resulting from the spread of isomorphous substitution in tetrahedral and octahedral sheets, whereas 1:1 layered kaolinite has little isomorphous substitution. This explains the higher capacity of montmorillonite to adsorb cations (Chantawong et al., 2003).

3.2.3. Effects of interaction time and kinetics of adsorption

The metal ions interacted with the clays rapidly and within 40 min, the maximum uptake was observed (Fig. 5). Afterwards, the interactions slowed down and approached equilibrium in nearly 180 min for Pb(II) and Ni(II), but took almost 240 min for Cd(II) under the given set of experimental conditions.

Attainment of equilibrium is influenced by several factors including the nature of the adsorbent and the adsorbate, and the interactions between them. Jain and Ram (1997) have reported that a quasi-stationary state is obtained within 45 min of shaking river bed sediments with Pb(II). Tran et al. (1999) have reported 60–70% Cd(II) removal by silica gel within 50 min of interaction although equilibration required another 135 min. Similar observations have been reported by other authors (Bayat, 2002; Kadirvelu and Namasivayam, 2003). For Ni(II) adsorption on almond husk activated carbon (Hasar, 2003) and on IRN77 cation-exchange resin (Rengaraj et al., 2002), the equilibrium was obtained in almost 180 min.

Initially, the rate of adsorption on the bare surface was very high, but as the sites got covered with the metal ions, the rate decreased. The rate now becomes predominantly dependent on the rate at which metal ions are transported from the bulk liquid phase to the adsorbent-adsorbate interface. The kinetics of the interactions is thus likely to be dependent on different rate processes as the interaction time increases (Yu et al., 2000).

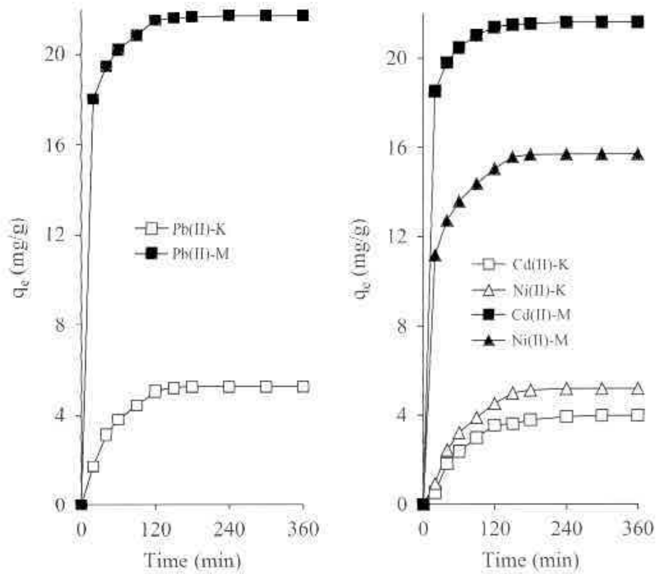


Fig. 5. Amount of metal ion adsorbed per unit mass (q_e) on clays for different time intervals at 303 K (*K* kaolinite, *M* montmorillonite).

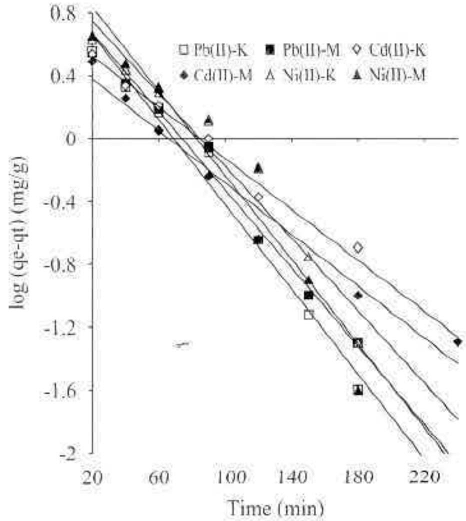


Fig. 6. Lagergren pseudo first-order plots for metal ions adsorbed on clays at 303 K (K kaolinite, M montmorillonite, experimental pH for Pb(II): 5.7, Cd(II): 5.5, Ni(II): 5.7).

The pseudo first-order plots of log (q_e–q_t) vs. t as per Lagergren equation are given in (Fig. 6). The plots had good linearity and yielded the first-order rate constant in the ranges of 1.9×10^{-2}–$2.6 \times 10^{-2}\,\mathrm{min}^{-1}$ for kaolinite and 1.9×10^{-2}–$3.1 \times 10^{-2}\,\mathrm{min}^{-1}$ for montmorillonite (Table 1) with respect to the three metals. The rate constants are in the order of Cd(II) < Ni(II) < Pb(II) for kaolinite and Cd(II) < Pb(II) < Ni(II) for montmorillonite.

On testing the validity of the first-order kinetics, it turned out that q_e values obtained from the Lagergren plots compare poorly with the experimental q_e values (Table 2) with both positive and negative deviations (range: −83.8 to +60.4%). Therefore, despite having high linearity of the Lagergren plots, the interactions do not appear to have followed first-order kinetics. The second-order plots of t/q_e vs. t (Fig. 7) are linear (R: 0.99) and the rate constant, k_2 varied from 2.2×10^{-2}–4.0×10^{-2} g/mg min for kaolinite and from 3.0×10^{-2}–8.4×10^{-2} g/mg min for montmorillonite, respectively (Table 1). From the consideration of the clay–metal interactions, the second-order rate constants are in the order of Ni < Pb(II) < Cd(II) for kaolinite and Cd(II) < Ni(II) < Pb(II) for montmorillonite.

The comparison of q_e values (experimental and those obtained from the slopes of the second-order plots) is given in Table 2 and the two sets now show much better agreement (deviations from −15.7 to +25.0%). Montmorillonite gave the least deviation for Pb(II) and Ni(II), while kaolinite showed 100% agreement for interactions with Cd(II). The deviations still existing might be due to the uncertainty inherent in obtaining the experimental q_e values and also due to the actual process being not in exact conformity with either first-order or second-order kinetics.

Interactions of Pb(II), Cd(II) and Ni(II) with kaolinite and montmorillonite also yielded good, linear Elovich plots (q_t vs. ln t, Fig. 8) with correlation coefficients of 0.98–0.99

Table 1
First-order rate constant, second-order rate constant, Elovich coefficients, intra-particle diffusion rate constant and liquid film diffusion rate constant for adsorption of metal ions at 303 K (clay 2 g/l, initial metal concentration 50 mg/l, pH 5.7 for Pb(II), 5.5 for Cd(II), 5.7 for Ni(II), K kaolinite, M montmorillonite)

Parameter		Pb(II)		Cd(II)		Ni(II)	
		K	M	K	M	K	M
First-order	$k_1 \times 10^2$ (min^{-1})	3.1	2.8	1.9	1.9	2.6	3.1
	R	0.99	0.99	0.98	0.74	0.97	0.97
Second-order	$k_2 \times 10^2$ (g mg^{-1} min^{-1})	3.5	8.4	4.0	3.0	2.2	5.3
	R	0.99	0.99	0.99	0.99	0.99	0.99
Elovich	$\alpha \times 10^3$ (g mg^{-1} min^2)	14.8	29.4	36.8	93.9	23.4	382.8
	β (mg g^{-1} min^{-1})	1.5	1.6	1.3	1.1	1.8	1.9
	R	0.98	0.98	0.98	0.98	0.99	0.99
Intra-particle diffusion	$k_i \times 10^{-1}$ (mg g^{-1} min$^{-0.5}$)	3.2	3.3	2.5	2.1	3.9	4.2
	Intercepts	1.0	17.3	0.2	18.5	−1.0	10.1
	R	0.93	0.93	0.92	0.88	0.95	0.95
Liquid film diffusion	$k_{fd} \times 10^{-2}$ (min^{-1})	3.2	2.8	1.8	1. 9	2.6	3.0
	Intercepts	−0.5	1.1	−0.2	1.8	−0.6	0.3
	R	0.99	0.99	0.99	0.98	0.98	0.97

Table 2
Experimental and computed q_e values from Lagergren and second-order plots for adsorption of metal ions at 303 K (clay 2 g/l, initial metal concentration 50 mg/l, pH 5.7 for Pb(II), 5.5 for Cd(II), 5.7 for Ni(II), K kaolinite, M montmorillonite)

Metal ion	Clay	Expt	q_e (mg/g)		Deviation (%)	
			Lagergren	Second-order	Lagergren	Second*order
Pb(II)	K	5.3	8.5	6.4	+60.4	+20.8
	M	21.7	7.5	22.3	−65.4	+2.7
Cd(II)	K	4.0	4.9	4.0	+22.5	0
	M	21.6	3.5	18.2	−83.8	−15.7
Ni(II)	K	5.2	6.9	6.5	+32.7	+25.0
	M	15.7	12.7	16.5	−19.1	+5.1

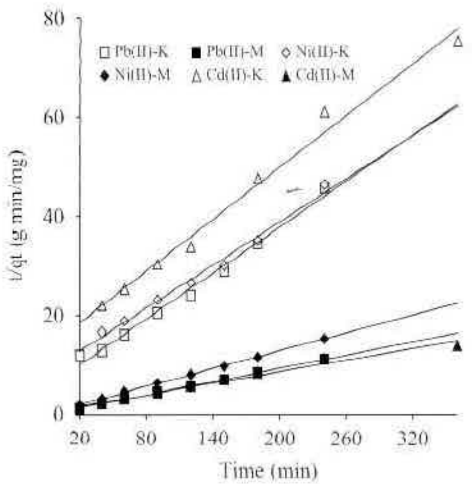


Fig. 7. Second-order plots for metal ions adsorbed on clays at 303 K (K kaolinite, M montmorillonite, experimental pH for Pb(II): 5.7, Cd(II): 5.5, Ni(II): 5.7).

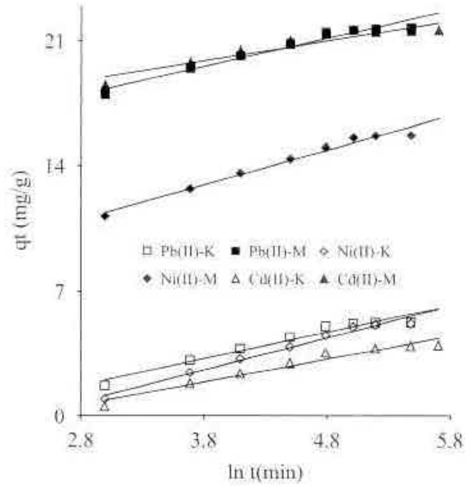


Fig. 8. Elovich plots for metal ions adsorbed on clays at 303 K (K kaolinite, M montmorillonite, experimental pH for Pb(II): 5.7, Cd(II): 5.5, Ni(II): 5.7).

(Table 1). The Elovich equation describes predominantly chemical adsorption on highly heterogeneous adsorbents, but the equation does not propose any definite mechanism for adsorbate-adsorbent interaction (Ho and McKay, 1998). The coefficients depended significantly on the amount of adsorbent with α being much more sensitive. As α represents the initial rate of adsorption and it is seen that with all the three metal ions, montmorillonite had a much higher initial rate of uptake than kaolinite. This might be due to very high surface area and CEC of montmorillonite compared to those of kaolinite. Similar results are also obtained by Ho and McKay (2002) for adsorption of Cu(II) on peat.

In adsorption from solution, diffusion from the solid–liquid interface to the interior of the solid particles plays a

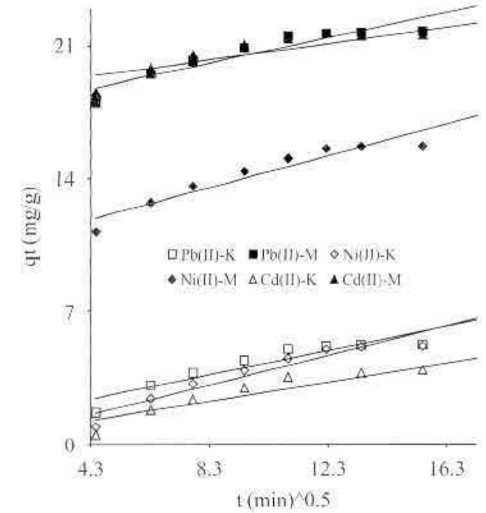


Fig. 9. Intra-particle diffusion plots for metal ion adsorbed on clays at 303 K (K kaolinite, M montmorillonite, experimental pH for Pb(II): 5.7, Cd(II): 5.5, Ni(II): 5.7).

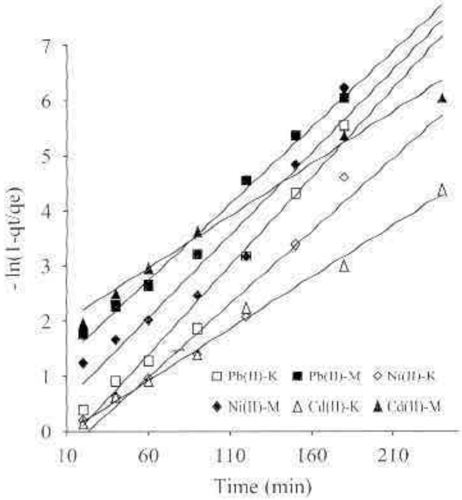


Fig. 10. Liquid film diffusion plots for metal ions adsorbed on clays at 303 K (K kaolinite, M montmorillonite, experimental pH for Pb(II): 5.7, Cd(II): 5.5, Ni(II): 5.7).

very important role. Whether the process of adsorption is controlled by this type of intraparticle diffusion, is tested by plotting q_t vs. $t^{0.5}$ (Fig. 9) as in Eq. (11). The plots were linear with regression coefficient of 0.93 for Pb(II), 0.88 and 0.92 for Cd(II) and 0.95 for Ni(II) adsorption (Table 1). Significantly, however, the plots do not have zero intercept (varies from -1.0 to $+18.5$) and thus, despite being linear, they do not fit into Eq. (11). Thus, intraparticle diffusion is not likely to be the controlling factor in determining the kinetics of the processes and the large intercepts suggest that the process was largely of surface adsorption.

Applying the liquid film diffusion model, it is observed that the plots of $-\ln(1-F)$ vs. t (Fig. 10) are linear (R: 0.97–0.99) (Table 1) with very small intercepts (-0.6 to $+1.1$). Thus, although the plots do not exactly pass through the origin, the small intercepts indicate that liquid film diffusion might have some role to play in the kinetics of adsorption of Pb(II), Ni(II) and Cd(II) on kaolinite and montmorillonite.

Testing of different kinetic models is important because the valid model indicates the possible mechanism of adsorption of the heavy metals on kaolinite and montmorillonite. In the present case, the second-order model gives the best possible fit to the experimental data and therefore, it is likely that adsorption of the metal ions on kaolinite and montmorillonite follows a second-order kinetic mechanism.

3.2.4. Adsorption isotherm

The adsorption data follow the empirical Freundlich isotherm (R: 0.99) (Fig. 11). The Freundlich coefficient, n, remains between 0.49 and 0.72 with respect to interaction of Pb(II), Cd(II) and Ni(II) with both the clays in conformity with the requirement of $n < 1.0$. The Freundlich adsorption capacity, K_f (Table 3), is large for montmorillonite (Pb(II): 7.5, Cd(II): 6.8, Ni(II): 3.4 $\mathrm{mg}^{(1-1/n)}\mathrm{l}^{1/n}/\mathrm{g}$)

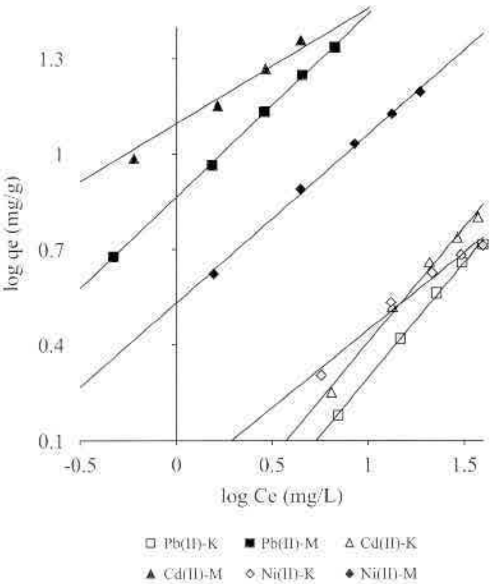


Fig. 11. Freundlich plots for metal ions adsorbed on clays at 303 K (initial metal ions 10, 20, 30, 40, 50 mg/l) (K kaolinite, M montmorillonite, experimental pH for Pb(II): 5.7, Cd(II): 5.5, Ni(II): 5.7).

compared to that of kaolinite (Pb(II): 0.4, Cd(II): 0.4, Ni(II): 0.9 $\mathrm{mg}^{(1-1/n)}\mathrm{l}^{1/n}/\mathrm{g}$). This is in agreement with the experimental results where montmorillonite has been found to adsorb much more than kaolinite.

The Langmuir isotherm, applicable strictly to chemisorptive monolayer formation, has also been found to be equally applicable (Fig. 12) (R: 0.98–0.99) for the present work. The equilibrium coefficient, b, is large with values of 20.7–70.1 (Table 3) for metal ion—kaolinite systems, and even larger for metal ion–montmorillonite systems (29.6–137.7) l/g. The large values show that the equilibrium

$$\text{Clay(solid phase)} + \text{M(II)(aqueous phase)}$$
$$= \text{Clay} \cdots\cdots \text{M(II)}$$

Table 3
Freundlich and Langmuir coefficients for adsorption of metal ions at 303 K (clay 2 g/l, initial metal concentration 10, 20, 30, 40, 50 mg/l, pH 5.7 for Pb(II), 5.5 for Cd(II), 5.7 for Ni(II), time 180 min for Pb(II), 240 min for Cd(II), 180 min for Ni(II), K kaolinite, M montmorillonite)

Metal ion	Clay	Freundlich coefficients			b ($l\,g^{-1}$)	Langmuir coefficients		
		K_f ($mg^{1-1/n}\,L^{1/n}\,g^{-1}$)	n	R		q_m ($mg\,g^{-1}$)	R_L	R
Pb(II)	K	0.4	0.7	0.99	20.7	11.5	0.0021	0.99
	M	7.5	0.7	0.99	31.0	31.1	0.0014	0.98
Cd(II)	K	0.4	0.6	0.99	32.3	6.8	0.0013	0.99
	M	6.8	0.6	0.99	29.6	30.7	0.0014	0.99
Ni(II)	K	0.9	0.5	0.99	70.1	7.1	0.0007	0.99
	M	3.4	0.5	0.99	137.7	21.1	0.0003	0.99

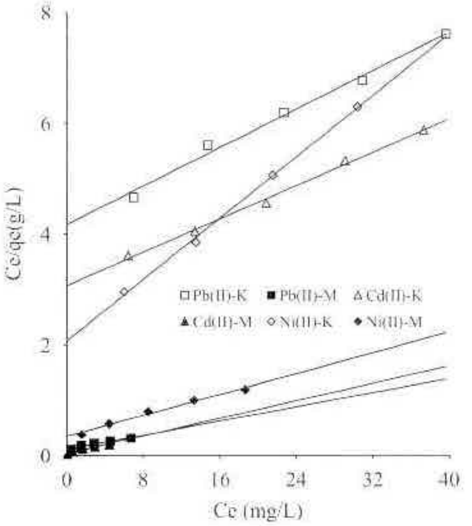


Fig. 12. Langmuir plots for metal ions adsorbed on clays at 303 K (initial metal ions 10, 20, 30, 40, 50 mg/l) (K kaolinite, M montmorillonite, experimental pH for Pb(II): 5.7, Cd(II): 5.5, Ni(II): 5.7).

is shifted predominantly to the right-hand side, i.e. towards the formation of the adsorbate–adsorbent complex. The largest value of 'b' was obtained for Ni(II)–montmorillonite indicate that the interactions were the strongest between Ni(II) and montmorillonite. The monolayer capacity, q_m, is appreciably large with values of 6.8 (kaolinite ... Cd(II)) to 31.1 (montmorillonite ……Pb(II)) mg/g. The separation factor, R_L (values 0.0003–0.0021), also indicates that the adsorption of the metal ions is favoured on the clay adsorbents.

Values of the adsorption coefficients in similar ranges have been reported by other workers. Taty-Costodes et al. (2003) from studies of adsorption of Pb(II) and Cd(II) on sawdust, found Langmuir adsorption capacity, q_m, between 8.45 and 22.22 mg/g in the pH range 7.0–4.0 ($R^2 = 0.91$–0.99) for Pb(II) adsorption. In case of Cd(II), q_m was from 6.72 to 15.27 mg/g for the same pH range ($R^2 = 0.93$–0.99) Bayramoglu et al. (2003) have reported Freundlich adsorption capacity, K_f as 0.32–1.29 l/g for adsorption of Pb(II) on white-rot fungus. In another work, Al-Subu (2002) studied the adsorption of Pb(II) on decaying leaves of cypress, cinchona, and pine and reported the Freundlich adsorption capacity as 0.416, 1.347 and 0.019 l/g and that of n as 1.027, 1.652 and 0.634, respectively. Chen et al. (2001) have reported q_m of 28.26 mg/g ($R = 0.999$) and K_f of 5.321 l/g ($R = 0.962$) for adsorption of Ni(II) on peat. These values are quite close to the values found in the present work.

3.2.5. Effect of temperature and thermodynamic parameters

Adsorption of Pb(II) and Ni(II) decreased as the temperature was increased from 303 to 313 K (Fig. 13). The interactions are thus exothermic and the metal ions leave the solid phase at higher temperature (Echeverria et al., 2003). With increase in temperature, the solubility of the metal ions in the aqueous phase is likely to increase with the resultant decrease in metal ion concentration in the solid phase. The trends demonstrated a tendency to escape from the solid phase (clay adsorbent) to the bulk solution phase with the rise in temperature (Singh et al., 1998) the excess energy promoting desorption rather than adsorption.

Cd(II) adsorption on the clays follows a different pattern. Increase in temperature from 303 to 308 K has the result of enhancing Cd(II) adsorption, but the trend reverses between 308 and 313 K. The process is controlled by the adsorbate–adsorbent and adsorbate–adsorbate forces and it is clear from the results that the first one becomes weak in comparison to the latter as the temperature increases.

Exothermic adsorption of Pb(II) ions has been observed on china clay and wollastonite (Yadava et al., 1991). However, the adsorption of Pb(II) on peat (Ho and McKay, 1999b) have been reported as endothermic. The endothermic interaction of Cd(II) on bagasse fly ash (Gupta et al., 2003) and kaolinite (Angove et al., 1998) have been reported earlier. But Singh et al. (1998) reported exothermic nature of Cd(II) adsorption on hematite. The exothermic nature of Ni(II) adsorption has been observed by Abou-Mesalam (2003) using synthesized silico-antimonate ion exchanger as an adsorbent. Endothermic adsorption of Ni(II) on canola meal has also been reported (Al-Asheh and Duvnjak, 1999).

The thermodynamic parameters, ΔH, ΔS and ΔG, for the adsorption process, are computed from the plots of $\ln K_d$ vs. $1/T$ (Table 4). For Pb(II)–clay and Ni(II)–clay systems, ΔH values are in accordance with the exothermic nature of

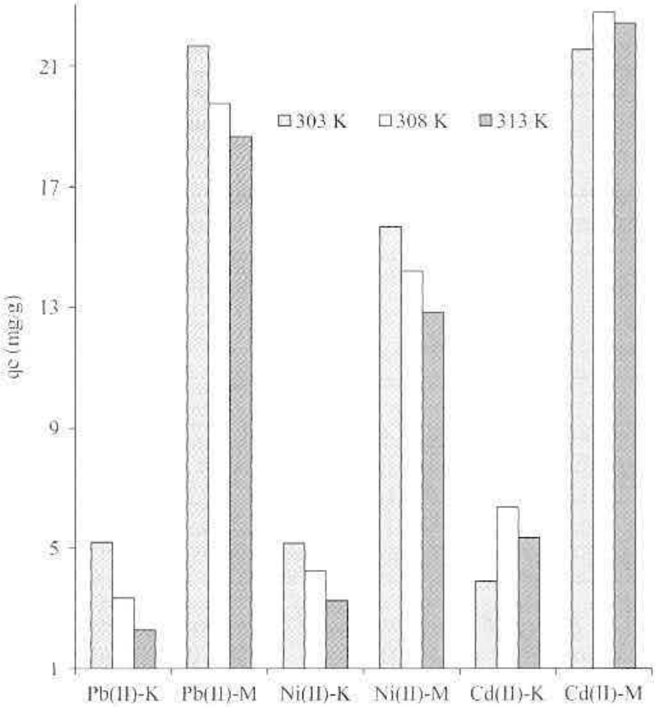


Fig. 13. Effect of temperature on amount of metal ions adsorbed per unit mass (q_e) on clays (K kaolinite, M montmorillonite).

Table 4
Thermodynamic data for adsorption of metal ions (clay 2 g/l, initial metal concentration 10, 20, 30, 40, 50 mg/l, pH 5.7 for Pb(II), 5.5 for Cd(II), 180 min for Ni(II), time 180 min for Pb(II), 240 min for Cd(II), 180 min for Ni(II), K kaolinite, M montmorillonite)

Metal ion	Clay	ΔH (kJ/mol)	ΔS (J/K mol)	$-\Delta G$ (kJ/mol)			
				303 K	308 K	313 K	Mean
Pb(II)	K	−58.9	−209.7	63.5	64.6	65.6	64.6
	M	−31.5	−116.3	35.2	35.8	36.4	35.8
Cd(II)	K	25.0	66.3	20.0	20.4	20.7	20.4
	M	46.2	147.3	44.6	45.4	46.1	45.4
Ni(II)	K	−37.9	−118.2	41.9	42.6	43.2	42.6
	M	−45.1	−146.4	44.4	45.1	45.9	45.1

interactions (Pb(II)–kaolinite: −58.9 kJ/mol; Pb(II)–montmorillonite: −31.5 kJ/mol; Ni(II)–kaolinite: −37.9 kJ/mol; Ni(II)–montmorillonite: −45.1 kJ/mol). In accordance with the endothermic nature of Cd(II)–clay interactions, ΔH values were found to be positive (Cd(II)–kaolinite: 25.0 kJ/mol; Cd(II)–montmorillonite: 46.2 kJ/mol). From the magnitude of the adsorption enthalpy, ΔH, moderately strong metal ion–clay bonding could be expected as was observed by other workers (Gupta et al., 2001).

For Pb(II) and Ni(II) adsorption, entropy decreased during the process (ΔS: −209.7 to −116.3 J/K mol) leading to a stable configuration. Since stability is associated with an ordered arrangement, it is obvious that Pb(II) and Ni(II) ions in aqueous solution are in much more chaotic distribution than they are in the adsorbed state. Thus, Pb(II) and Ni(II) have strong affinity towards the clays (Echeverria et al., 2003). For Cd(II), however, adsorption is accompanied by an increase in entropy (Cd(II)–kaolinite: 66.3; Cd(II)–montmorillonite: 147.3 J/K mol) indicating considerable change in surface configuration due to the interactions and the strong affinity of Cd(II) towards the clay surface is supported by entropy increase(Abou-Mesalam, 2003).

Gibbs energy decreases in all the cases of adsorption of Pb(II), Cd(II) and Ni(II) on kaolinite and montmorillonite and therefore, the interactions are spontaneous. The mean decrease in Gibbs energy in the temperature range 303–313 K was from −20.4 to −64.6 kJ/mol considering all the systems (Table 4).

Thermodynamic data on metal adsorption on clays are scarce. Yavuz et al. (2003) have found that ΔH, ΔS and ΔG for adsorption of Cu(II) on Turkish kaolinite are 39.52 kJ/mol, 11.7 J/K mol and 4.61 kJ/mol, respectively. Echeverria et al. (2003) have reported that ΔH, ΔS and ΔG for adsorption of Ni(II) on illite have values of +16.8 kJ/mol, +58 J/mol K and −1.04 kJ/mol, respectively. ΔH, ΔS and ΔG for Cu(II) adsorption on surfactant-modified montmorillonite were reported as 7.05 kJ/mol, 9.09 J/K mol and −9.66 kJ/mol, respectively (Lin and Juang, 2002). These values are comparatively smaller than the values obtained in the present study indicating the Pb(II)–clay, Cd(II)–clay and Ni(II)–clay adsorption complexes to be much more stable with the clays holding these metal ions strongly to the surface. It is to be noted that the present experiments were carried out only over a small temperature interval of just 10° (303–313 K) due to the comparatively high ambient temperature, which however should not introduce much uncertainty in interpretation of the thermodynamic results.

4. Conclusion

Both kaolinite and montmorillonite are capable of removing metal ions [Pb(II), Cd(II) and Ni(II)] from aqueous solution. The adsorption capacity of montmorillonite is much more than kaolinite for all the three metal ions. Adsorption increases with pH till the metal ions are precipitated out as metal hydroxides. Although different kinetic models were tested for clay-metal ion interactions, it was found that the second-order kinetic model gave the best fit with the experimental data and therefore, the adsorption followed a mechanism based on second-order kinetics.

Isotherm-fitting yields good results with both Langmuir and Freundlich equations indicating that while the interactions are predominantly chemical in nature, the adsorption sites are non-uniform and non-specific in nature. This is in conformity with the existence of different types of possible adsorption sites on clay surface with considerable difference in energy if the site is on an edge or is located in a defect position. The adsorption coefficients agree well with the conditions supporting favorable adsorption. The

Langmuir monolayer capacity is large particularly for montmorillonite in conformity with its capacity to take up more of the metal ions. The interactions are thermodynamically favorable and are accompanied by decrease in Gibbs energy.

Acknowledgments

The authors are very grateful to all the three reviewers for pointing out various errors and omissions. The comments of the reviewers have resulted in a substantial improvement of the manuscript. One of the authors (SSG) is grateful to the University Grants Commission, New Delhi for providing assistance under the FIP scheme for this work.

References

Abou-Mesalam, M.M., 2003. Sorption kinetics of copper, zinc, cadmium and nickel ions on synthesized silico-antimonate ion exchanger. Colloids and Surfaces A: Physicochemical and Engineering Aspects 225, 85–94.

Al-Asheh, S., Duvnjak, Z., 1999. Sorption of heavy metals by canola meal. Water, Air and Soil Pollution 114, 231–276.

Al-subu, M.M., 2002. The interaction effects of cypress (*Cupressus sempervirens*), cinchona (*Eucalyptus longifolia*) and pine (*Pinus halepensis*) leaves on their efficiencies for lead removal from aqueous solutions. Advances in Environmental Research 6, 569–576.

Alvarez-Ayuso, E., Garcia-Sanchez, A., 2003. Removal of heavy metals from wastewaters by natural and Na-exchanged bentonites. Clays and Clay Minerals 51, 475–480.

An, H.K., Park, B.Y., Kim, D.S., 2001. Crab shell for the removal of heavy metals from aqueous solution. Water Research 35, 3551–3556.

Angove, M.J., Johnson, B.B., Wells, J.D., 1998. The influence of temperature on the adsorption of cadmium(II) and cobalt(II) on kaolinite. Journal of Colloid and Interface Science 204, 93–103.

ATSDR, 1999a. Toxicological profile for lead, US Department of Health and Human Services, Public Health Services (Agency for Toxic Substances and Disease Registry), Atlanta, GA.

ATSDR, 1999b, Toxicological profile for cadmium, US Department of Health and Human Services, Public Health Services (Agency for Toxic Substances and Disease Registry), Atlanta, GA.

ATSDR, 2003. Draft Toxicological profile for nickel, US Department of Health and Human Services, Public Health Services Agency for Toxic Substances and Disease Registry, Atlanta, GA.

Bayat, B., 2002. Comparative study of adsorption properties of Turkish fly ashes I. The case of nickel(II), copper(II) and zinc(II). Journal of Hazardous Materials B 95, 251–273.

Bayramoglu, G., Bekta, S., Arica, M.Y., 2003. Biosorption of heavy metal ions on immobilized white-rot fungus *Trametes versicolor*. Journal of Hazardous Materials B 101, 285–300.

Bergaya, F., Vayer, M., 1997. CEC of clays: measurement by adsorption of a copper ethylendiamine complex. Applied Clay Science 12, 275–280.

Boyd, G.E., Adamson, A.W., Myers Jr., L.S., 1947. The exchange adsorption of ions from aqueous solutions on organic zeolites. Kinetics II. Journal of American Chemical Society 69, 2836–2842.

Chantawong, V., Harvey, N.W., Bashkin, V.N., 2003. Comparison of heavy metal adsorptions by thai kaolin and ballclay. Water, Air and Soil Pollution 148, 111–125.

Chen, B., Hui, C.W., McKay, G., 2001. Film-pore diffusion modeling for the sorption of metal ions from aqueous effluents onto peat. Water Research 35, 3345–3356.

Coles, C.A., Yong, R.N., 2002. Aspects of kaolinite characterization and retention of Pb and Cd. Applied Clay Science 22, 39–45.

De Leon, A.T., Nunes, D.G., Rubio, J., 2003. Adsorption of Cu ions onto a 1:10 phenanthroline-grafted Brazilian bentonite. Clays and Clay Minerals 51, 58–64.

Echeverria, J.C., Churio, E., Garrido, J., 2002. Retention mechanisms of Cd on illite. Clays and Clay Minerals 50, 614–623.

Echeverria, J., Indurain, J., Churio, E., Garrido, J., 2003. Simultaneous effect of pH, temperature, ionic strength, and initial concentration on the retention of Ni on illite. Colloids and Surfaces A: Physicochemical and Engineering Aspects 218, 175–187.

Evans, J.R., Davids, W.G., MacRae, J.D., Amirbahman, A., 2002. Kinetics of cadmium uptake by chitosan-based crab shells. Water Research 36, 3219–3226.

Freundlich, H.M.F., 1906. Über die adsorption in lösungen. Zeitschrift fur Physikalische Chemie (Leipzig) 57A, 385–470.

Ghosh, D., Bhattacharyya, K.G., 2002. Adsorption of methylene blue on kaolinite. Applied Clay Science 20, 295–300.

Grim, R.E., 1968. Clay Mineralogy. McGraw-Hill, New York.

Gupta, V.K., Gupta, M., Sharma, S., 2001. Process development for the removal of lead and chromium from aqueous solutions using red mud—an aluminium industry waste. Water Research 35, 1125–1134.

Gupta, V.K., Jain, C.K., Ali, I., Sharma, M., Saini, V.K., 2003. Removal of cadmium and nickel from wastewater using bagasse fly ash—a sugar industry waste. Water Research 37, 4038–4044.

Hall, K.R., Eagleton, L.C., Acrivos, A., Vermeulen, T., 1966. Pore- and solid-diffusion kinetics in fixed-bed adsorption under constant-pattern conditions. Industrial and Engineering Chemistry Fundamentals 5, 212–223.

Hasar, H., 2003. Adsorption of nickel(II) from aqueous solution onto activated carbon prepared from almond husk. Journal of Hazardous Materials B 97, 49–57.

Ho, Y.S., McKay, G., 1998. A comparison of chemisorption kinetic models applied to pollutant removal on various sorbents. Transactions of the Institution of Chemical Engineers 76B, 332–339.

Ho, Y.S., McKay, G., 1999a. The sorption of lead(II) ions on peat. Water Research 33, 578–584.

Ho, Y.S., McKay, G., 1999b. Batch lead(II) removal from aqueous solution by peat: equilibrium and kinetics. Transactions of the Institution of Chemical Engineers 77B, 165–173.

Ho, Y.S., McKay, G., 2002. Application of kinetic models to the sorption of copper(II) on to peat. Adsorption Science and Technology 20, 797–815.

Ho, Y.S., Ng, J.C.Y., McKay, G., 2001. Removal of lead(II) from effluents by sorption on peat using second-order kinetics. Separation Science and Technology 36, 241–261.

Ho, Y.S., 2003. Removal of copper ions from aqueous solution by tree fern. Water Research 34, 2323–2330.

Ho, Y.S., 2004. Citation review of Lagergren kinetic rate equation on adsorption reactions. Scientometrics 59, 171–177.

Jain, C.K., Ram, D., 1997. Adsorptions of lead and zinc on bed sediments of the river kali. Water Research 34, 154–162.

Kadirvelu, A., Namasivayam, C., 2003. Activated carbon from coconut coirpith as metal adsorbent: adsorption of Cd(II) from aqueous solution. Advances in Environmental Research 7, 471–478.

Kara, M., Yuzer, H., Sabah, E., Celik, M.S., 2003. Adsorption of cobalt from aqueous solutions onto sepiolite. Water Research 37, 224–232.

Khan, S.A., Rehman, R., Khan, M.A., 1995. Sorption of strontium on bentonite. Waste Management 15, 641–650.

Lagergren, S., 1898. Zur theorie der sogenannten adsorption gelöster stoffe. Kungliga Svenska Vetenskapsakademiens Handlingar, Band 24 (4), 1–39.

Langmuir, I., 1918. The adsorption of gases on plane surfaces of glass, mica and platinum. Journal of the American Chemical Society 40, 1361–1403.

Lin, S-H., Juang, R-S., 2002. Heavy metal removal from water by sorption using surfactant-modified montmorillonite. Journal of Hazardous Materials B 92, 315–326.

Marino, P.E., Franzblau, A., Lilis, R., 1989. Acute lead poisoning in construction workers: the failure of current protective standards. Archives of Environmental Health 44, 140–145.

Mathialagan, T., Viraraghavan, T., 2002. Adsorption of cadmium from aqueous solutions by perlite. Journal of Hazardous Materials 94, 291–303.

Mellah, A., Chegrouche, S., 1997. The removal of zinc from aqueous solutions by natural bentonites. Water Research 31, 621–629.

Pinnavaia, T.J., 1983. Intercalated clay catalysts. Science 220, 365–371.

Ravichandran, J., Sivasankar, B., 1997. Properties and catalytic activity of acid-modified montmorilonite and vermiculite. Clays and Clay Minerals 45, 854–858.

Rengaraj, S., Yeon, K-H., Kang, S-Y., Lee, J-U., Kim, K-W., Moon, S-H., 2002. Studies on adsorptive removal of Co(II), Cr(III) and Ni(II) by IRN77 cation-exchange resin. Journal of Hazardous Materials B 92, 185–198.

Rengaraj, S., Joo, C.K., Kim, Y., Yi, J., 2003. Kinetics of removal of chromium from water and electronic process wastewater by ion exchange resins: 1200H, 1500H and IRN97H. Journal of Hazardous Materials B 102, 257–275.

Sears, G., 1956. Determination of specific surface area of colloidal silica by titration with sodium hydroxide. Analytical Chemistry 28, 1981–1983.

Shukla, A., Zhang, Y.-H., Dubey, P., Margrave, J.L., Shukla, S.S., 2002. The role of sawdust in the removal of unwanted materials from water. Journal of Hazardous Materials 95, 137–152.

Singh, D.K., Tiwari, D.P., Saksena, D.N., 1993. Removal of lead from aqueous solutions by chemically treated tealeaves. Indian Journal of Environmental Health 35, 169–177.

Singh, D.B., Rupainwar, D.C., Prasad, G., Jayaprakas, K.C., 1998. Studies on the Cd(II) removal from water by adsorption. Journal of Hazardous Materials 60, 29–40.

Tanabe, K., 1981. Solid acid and base catalysis. In: Anderson, J.R., Boudart, M. (Eds.), Catalysis Science and Technology. Springer, New York, p. 231.

Taty-Costodes, V.C., Fauduc, H., Porte, C., Delacroix, A., 2003. Removal of Cd(II) and Pb(II) ions, from aqueous solutions, by adsorption onto sawdust of *Pinus sylvestris*. Journal of Hazardous Materials B 105, 121–142.

Tran, H.H., Roddick, F.A., O'Donell, J.A., 1999. Comparison of chromatography and desiccant silica gels for the adsorption of metal ions—I. Adsorption and kinetics. Water Research 33, 2992–3000.

Volzone, C., Thompson, J.G., Melnitchenko, A., Ortiga, J., Palethorpe, S.R., 1999. Selective gas adsorption by amorphous clay-mineral derivatives. Clays and Clay Minerals 5, 647–657.

Weber, W.J., Morris, J.C., 1963. Kinetics of adsorption of carbon from solutions. Journal of the Sanitary Engineering Division American Society of Civil Engineering 89, 31–63.

Yadava, K.P., Tyagi, B.S., Singh, V.N., 1991. Effect of temperature on the removal of lead(II) by adsorption on china clay and wallastonite. Journal of Chemical Technology and Biotechnology 51, 47–60.

Yavuz, O., Altunkaynak, Y., Guzel, F., 2003. Removal of copper, cobalt and manganese from aqueous solution by kaolinite. Water Research 37, 948–952.

Yu, B., Zhang, Y., Shukla, A., Shukla, S.S., Dorris, K.L., 2000. The removal of heavy metal from aqueous solutions by sawdust adsorption–removal of copper. Journal of Hazardous Materials B 80, 33–42.

Yu, B., Zhang, Y., Shukla, A., Shukla, S.S., Dorris, K.L., 2001. The removal of heavy metals from aqueous solutions by sawdust adsorption–removal of lead and comparison of its adsorption with copper. Journal of Hazardous Materials B 84, 83–94.

ELSEVIER

Journal of Environmental Management 87 (2008) 59–72

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Critical analysis of the waste management performance of two uranium production units in Brazil—part I: Poços de Caldas production centre

Horst Monken Fernandes[a,b,*], Mariza Ramalho Franklin[a,b], Luiz Alberto Gomiero[c]

[a]*Institute of Radiation Protection and Dosimetry (IRD), Av. Salvador Allende s/n, Recreio, Rio de Janeiro, RJ, CEP: 22780-160, Brazil*
[b]*University of Central Florida (UCF), P.O. Box 162450, Orlando, FL, USA*
[c]*Nuclear Industries of Brasil (INB), Fazenda Cachoeira s/n, Caixa Postal 7, 46400-000–Caetité, BA, Brazil*

Received 3 October 2005; received in revised form 23 October 2006; accepted 4 January 2007
Available online 12 June 2007

Abstract

Waste management strategies in mining projects will depend to a large extent on the characteristics of the operational process, the type of ore and prevailing socio-environmental conditions, amongst other issues. The expenditures required by the management scheme and the implementation of remediation programs will be determined by the extent that the above issues were considered in the planning phase of the project. Several works have been published in the literature concerning the analysis of waste management programs and environmental impacts associated with uranium projects around the world. However, the vast majority do not report a comprehensive assessment integrating the various relationships among operational process, environmental impact, remediation strategy and costs. This study, divided into two papers, presents a detailed critical analysis of the waste management strategies adopted in two uranium production centres in Brazil, i.e., the Poços de Caldas Project (Part I) and the Caetité Project (Part II). The operational processes are described and the environmental impacts of the generated wastes as well as the adopted management strategies and costs are examined. Also, in Part II, a comparison between both production centres is made emphasizing the impacts of environmental and social-economical issues on the overall assessment.

Keywords: Uranium mining; Environmental management; Remediation; Acid drainage

1. Introduction

Uranium production plants may give rise to huge amounts of wastes that encompass both mining and milling processes. Mining waste can be defined as the materials that result from the exploration, mining and processing of substances from mines and quarries. It may consist of natural materials without any modification other than crushing (ordinary mining waste, unusable mineralized materials) or of materials generated during the ore-processing and enrichment phases, possibly containing chemical, inorganic and organic additives; overburden and topsoil may also classified as waste (NEA, 2002).

Generally speaking, the volume of waste rock will usually exceed the volume of the ore extracted. For underground mines, a much smaller amount of waste rock will typically be generated. While the uranium content in these materials may not be significant in economic or mining terms, the radionuclide (including its long-lived progenies) content may be sufficient to pollute surface and groundwater, or present a direct exposure hazard (dust, radon) to the adjacent community.

After the mining operations take place uranium has to be extracted from the ore. Both alkaline and acid leaching may be used in the extraction of uranium from the ore using one of a variety of leaching systems. The most frequently used ones are:

- Leaching in tanks,
- *In situ* leaching, or
- Heap leaching.

*Corresponding author. Institute of Radiation Protection and Dosimetry (IRD), Av. Salvador Allende s/n, Recreio, Rio de Janeiro, RJ, CEP: 22780-160, Brazil. Tel.: +55 21 34118087; fax: +55 21 24422699.
E-mail address: hmf@ird.gov.br (H.M. Fernandes).

0301-4797/$ - see front matter © 2007 Published by Elsevier Ltd.
doi:10.1016/j.jenvman.2007.01.045

The heap leach and *in situ* leach systems are, generally speaking, more appropriate when low grade ore deposits are involved (Diehl, 2004). Under these conditions the profits arising from the operations will not match the capital investments to be made in a leaching tank. Conversely, this situation is changed when one deals with high grade deposits, in which a decrease in the leaching efficiency would result in significant economic losses.

Leaching in tanks involves crushing and grinding of the ore to a very fine grain size (70 μm) in order to form a slurry when the ore is mixed with an acid or alkaline solution. The leaching process may involve a series of up to 12 tanks. One of the most significant disadvantages of this technique (besides the high energy consumption) is the type of wastes that result from the leaching operations. The adequate management of these wastes (tailings) is still a challenge to be faced, especially when long-term aspects are taken into account.

The *in situ* leach involves the direct injection of a leaching solution into the ore body and subsequent recovery of the solution enriched in uranium. The advantage of this method is that the excavation of an open pit or underground mine is not necessary. As a result, the amount of mining wastes produced is minimal. However the potential contamination of nearby aquifers is a problem to be considered.

Heap leaching is a method for recovering uranium from mined or crushed ores of typically low grade, without going through the milling process. Pads with an impervious bottom liner to collect the leaching solution are constructed. These heaps may either be left in place after leaching is complete or the material may be removed to a disposal site to make room for fresh ore. Although the relatively coarse-sized materials, comprised of rocks ranging in size from boulders or cobbles to gravel, may be less reactive than the milled materials, many of the radioactive and chemically active constituents, such as sulfides will remain in the residues.

When tailings are generated in the uranium production process, terrestrial deposition is the predominant method of disposal utilizing geomorphological depressions or valleys. The principle of tailing dams (or ponds) is to dispose of the tailings in an accessible condition that provides for their future reprocessing (once improved technology or a significant increase in price makes it profitable). Other disposal methods such as underground backfilling or deep water disposal (lakes and sea) may also take place. Actually, the vast majority of tailing facilities are designed as permanent disposal facilities. In the past, there was little or no care taken to isolate the tailing materials from their environment.

Identification of the environmental risks associated with the exploitation of mines and ore processing plants not only requires the characterization and quantification of the different types of waste, and a knowledge of the processes used, but also an assessment of the vulnerability of the specific environments, dependent on the geological and hydrogeological conditions and peripheral targets. For instance, typical environmental problems arising from mill tailings are radon emissions, windblown dust dispersal, and the leaching of contaminants, including radionuclides, heavy metals and arsenic, into surface and ground waters (IAEA, 2004).

Mining and milling wastes can affect the environment through one or more of the intrinsic characteristics, e.g., (i) its chemical and mineralogical composition, (ii) its physical properties, (iii) its volume and the surface occupied, or (iv) the waste disposal method.

Besides these parameters, one must also take into account extrinsic ones such as (i) climatic conditions liable to modify the disposal conditions, (ii) geographic and geological location, or (iii) existing targets liable to be affected (human health and his environment).

Management of mining waste disposal facilities must take into consideration long-term environmental issues, because these structures will more than likely survive both the mine and the mining company. This situation raises a legal problem with regard to the responsibility for maintenance and repair of these facilities since liability, under most laws, cannot be endless. Even where the facility becomes a long-term structure, it is still necessary to provide permanent monitoring and inspection system. Closure and after-care operations are therefore of paramount importance to reduce, as far as possible, the long-term environmental risks.

Most of the works in the literature treat these aspects reported above in an isolated way. This study presents a critical analysis of the waste management strategies employed in two uranium production centres in Brazil. The objective of this work is to perform an integrated and comparative assessment of the waste disposal methods used by two uranium mining and milling facilities in Brazil. It was aimed to show how the operational process, environmental conditions and social-economical settings may affect dramatically the management options and also to stress the need for an integrated and comprehensive evaluation of these aspects. Part I extends the study of the Poços de Caldas Site. Part II covers the Caetité Production Centre.

2. Methodology

The methodology used in this paper is depicted in Fig. 1.

Initially, the data from the operational process were collected to allow us to understand, the distribution and fate of the different potential contaminants amongst the generated wastes as well as their chemical form in these materials. The second step was to perform an inventory of the amount of wastes generated by past and present operations. The wastes were characterized regarding the radioactive and non-radioactive contaminant content and the probable mobilization and transporting mechanisms were defined. Subsequently, the potential impacts of the deposited wastes were assessed. Environmental media to be

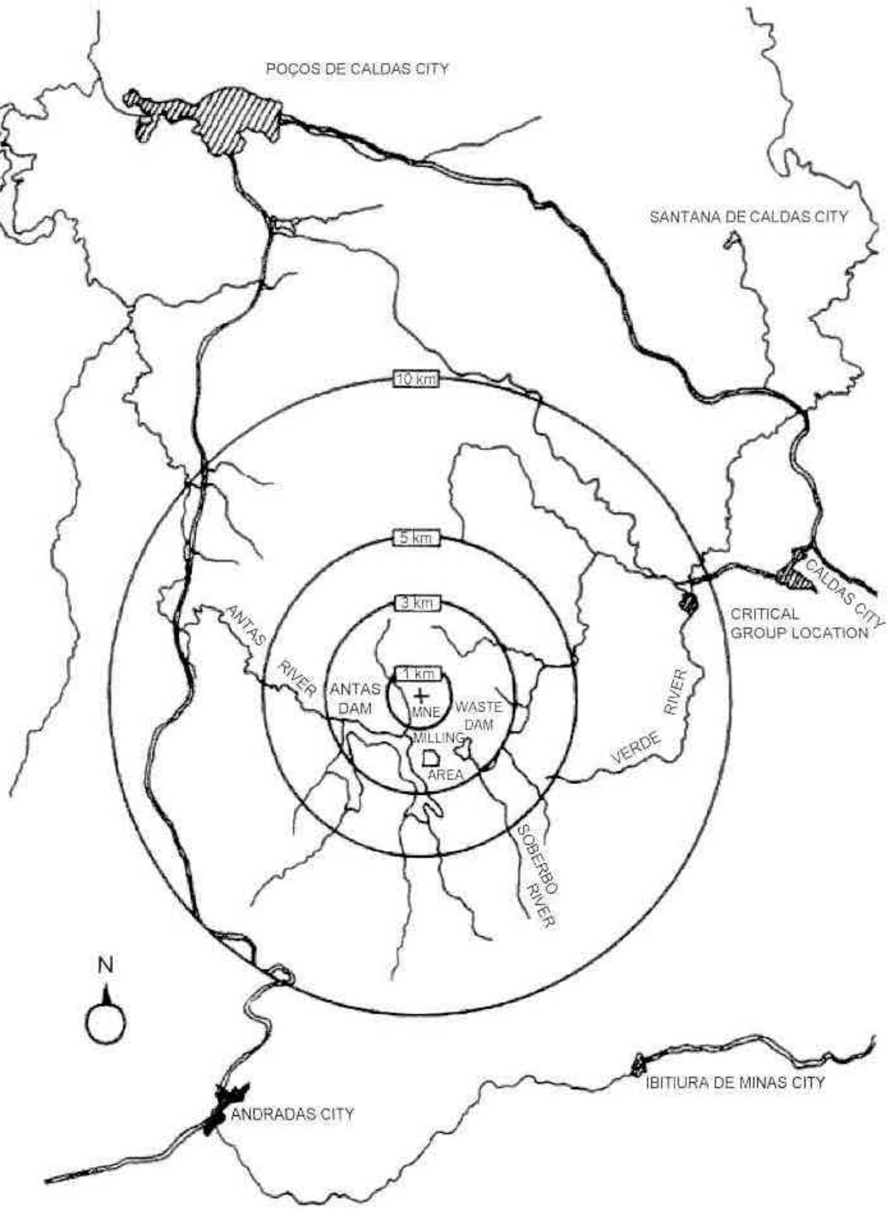


Fig. 1. Location map.

affected were taken into consideration and the potential exposures of members of the public were assessed, both in the operational and post-operational scenarios. Finally, remediation strategies were proposed for each particular situation under the perspective of a cost-effectiveness analysis.

3. Site description and assesment of quantities of generated mining wastes

The Poços de Caldas mining site is located in the Minas Gerais state, in the southern region of Brazil (latitude 21° 45′ S and longitude 46° 35′ W), 180 km northwest from São Paulo city and 360 km southwest from Rio de Janeiro the two major cities in the country. It occupies an area of about 15 km^2. The location map is shown in Fig. 2 where the city of Poços de Caldas (200,000 inhabitants) can be identified being located 20 km north from the mining site. The two major water courses which receive the releases of the mining and milling operation are the Antas River that flows in the direction of Poços de Caldas city and the Soberbo river which flows in the direction of the city of Caldas Average annual precipitation is 1800 mm/year. The mine covers an area of 2.0 km^2. The mineralized zone was located at about 200 m below surface and the mine area was divided into three different ore bodies (A, B and E) for the purpose of mining operations. The ore grade in respect to uranium varied between 675 to 1705 ppm. The uranium deposit is defined as being of low grade associated with a primary mineralization of Zr-rare earth elements (REE)-U-Th-Mo (bodies A and B) and a secondary mineralization caused by hydrothermal processes (body E). Uranium

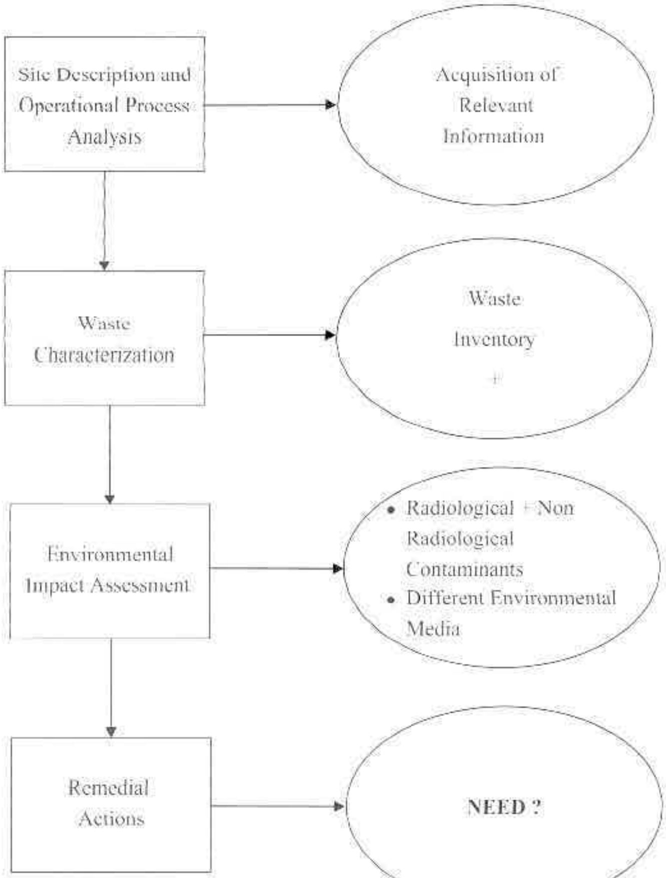


Fig. 2. Methodological approach.

occurs in the form of pitchblende, although brannerite ($UTiO_6$) was also identified in marginal proportions (Waber et al., 1991). The rocks have varying proportions of mafic and nefeline syenite minerals. Chemical composition of the rocks is shown in Table 1. Attention must be called to the high contents of sulfur, occurring as pyrite that varies from 5637 to 18,961 ppm. The occurrence of pyrite in the rock has an important bearing in the generation of acid drainage as it will be discussed below in this text.

The Poços de Caldas Project was intended to produce 500 t U_3O_8/year and 275 t/year of calcium molybdate as a by-product. The operations gave rise to two main sources of contaminants to the environment: the waste rock piles (WRP) and the tailing dam.

After 15 years (1982–1997) the uranium mining and milling operations have ceased. However, the chemical plant in charge of the liquid effluent treatment is still active. The overall management system of these effluents is shown in Fig. 3. It can be seen that the drainage from the WRP is pumped to the mine pit. From there, the drainage is pumped to the chemical treatment plant. In the past the slurry from the drainage treatment was deposited in the tailing dam. Recently, due to the exhaustion of the capacity of the tailing dam to receive additional wastes, the

Table 1
Average composition and standard deviation of rocks from the three ore bodies of the Poços de Caldas mine

Element	Body A	Body B	Body E
SiO_2 (%)	55±0.53 ($n=7$)	53±2.85 ($n=20$)	55±2.64 ($n=14$)
Al_2O_3 (%)	21.7±0.64 ($n=7$)	20±2.96 ($n=20$)	23±1.14 ($n=14$)
Fe-tot (%)	2.6±0.60 ($n=7$)	4.88±3.68 ($n=20$)	2.61±1.06 ($n=14$)
F (mg/kg)	1488±172 ($n=7$)	4178±2957 ($n=13$)	2013±803 ($n=4$)
Th (mg/kg)	60±46 ($n=7$)	96±89 ($n=20$)	318±962 ($n=13$)
U (mg/kg)	89±57 ($n=7$)	538±958 ($n=20$)	279±619 ($n=12$)
Zn (mg/kg)	253±47 ($n=7$)	570±646 ($n=20$)	592±1360 ($n=9$)
S (mg/kg)	8616±2544 ($n=7$)	18,961±18,025 ($n=19$)	5637±−5321 ($n=9$)
Zr (mg/kg)	1708±873 ($n=7$)	4334±7115 ($n=20$)	1009±828 ($n=14$)

Source: Waber et al. (1991).

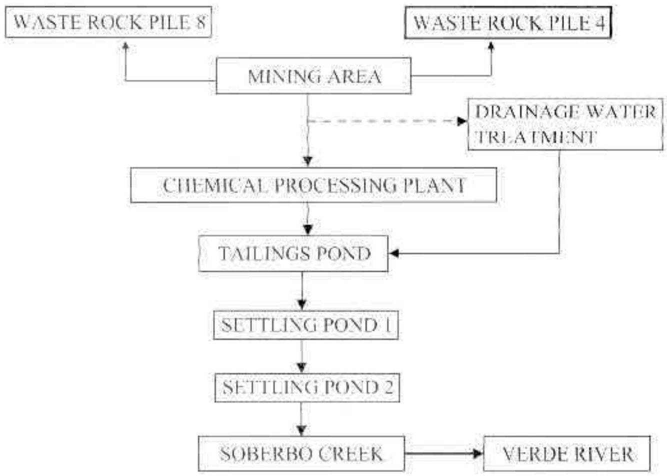


Fig. 3. Overall effluent management streams of Poços de Caldas mining site.

precipitate from the chemical treatment has been deposited in the mine open pit. The effluent from the tailing dam is treated with $BaCl_2$ to remove radium isotopes from the solution. The solids (precipitate) settle in two holding tanks and the overflow is discharged into the environment.

Table 2 shows the amount of waste rocks deposited in the different piles, piles number 4 and 8 being the largest. These two piles together contain about 60% of the total amount of rocks removed during mining operations. The relevant aspect of the WRP 4 is that all the drainage percolating from the system is collected in a holding pond at the toe of the dump. Table 3 shows the chemical composition of the drainage coming from WRP 4. The oxidation of pyrite in the rock material explains the low pH values in the drainage. It can also be seen that uranium

Table 2
Volume, mass and areas of the waste rock piles at the Poços de Caldas Mining site

Waste-rock pile	Volume ($10^6\,m^3$)	Mass ($10^6\,t$)	Area ($10^6\,m^2$)
WRP-1	4.4	10.6	2.5
WRP-3	9.8	23.5	2.0
WRP-4	12.4	29.8	5.7
WRP-7	2.4	5.8	5.3
WRP-8	14.8	35.5	6.4
Total	43.8	105.2	21.9

Table 3
Chemical composition of WRP 4 drainage ($n = 21$)

Parameter	Average	Minimum value	Maximum value
^{226}Ra (Bq/L)	0.29	0.14	0.58
^{238}U (Bq/L)	175	71	315
Al (mg/L)	96	61	161
F (mg/L)	99	5.1	167
Mn (mg/L)	75	6.6	105
pH	3.30	2.9	3.7

activity concentration much higher than that of ^{226}Ra. It means that the first is being preferentially mobilized relatively to the second assuming that both of them have identical activity concentration in the rock (i.e. they are in secular equilibrium).

Zenaro (1989) studied the speciation of U (VI) in the mining waters of Poços de Caldas mining site. The dominant species of U^{+6} were the uranyl–fluoride complexes and uranyl ion. However, the author suggested that in lower pH values the sulphate complexes would be of more relevance. At higher pH values the formation of carbonate and phosphate complexes would play a significant role. The simulations conducted by the author was in pH 5.2 and the concentrations of F^-, PO_4^{3-}, SO_4^{2-} and UO_2^{2-} were, respectively, 0.5 mg L^{-1}, 5.7 μg L^{-1}, 24 mg L^{-1} and 5.1 μg L^{-1} respectively.

In order to investigate the speciation of the U (VI) compounds in the acid drainage of WRP 04 the computational code MINEQL (Westall et al., 1976) was used. The input data were the same as those reported in Table 3. The results reveal that a precipitate of $BaSO_4$ would be formed. This would explain the low concentrations of ^{226}Ra in the acidic waters as it would be co-precipitating as $Ba(Ra)SO_4$. As for uranium the modeling results indicated the following distribution of uranium species: UO_2 (53.9%); UO_2F (8.5%), $UO_2(SO_4)_2$ (9.7%) and UO_2SO_4 (27%). Fernandes et al. (1998) reported ^{238}U concentrations in sampling location out of the influence of the mining and milling plant discharges in the range of 0.02–0.25 Bq L^{-1}. As the average ^{238}U concentrations in the acid drainage is orders of magnitude higher than the background values the release of these waters into the environment without prior treatment cannot be warranted. The high concentrations of Al and F and the low pH values reinforce the need of chemical treatment of these waters.

The occurrence of high radionuclide concentrations in the drainage from WRP 4 was first reported by Amaral et al. (1988). Fernandes et al. (1998). Fernandes and Franklin (2000) investigated the main processes responsible for contaminant mobilization and transport from the WRP 4 and discussed sound management options for these drainages. The dissolution of K-feldspar and kaolinite, along with the oxidation of pyrite by O_2 and Fe^{3+} were simulated by means of geochemical modeling. The results revealed intrinsic oxygen rates (IOR) on the order of $10^{-9}\,kg(O_2)\,m^{-3}\,s^{-1}$ (being the IOR the rate of consumption of oxygen by the material in the waste under the conditions which apply to that material). It was estimated that more than 500 years would be necessary for the contaminant concentrations in the drainage to decrease to acceptable values. The significance of these findings to the management strategies is that permanent solutions will have to be applied instead of the collection and treatment scheme presently in place.

Regarding the tailings, it was estimated that in the period between 1982 and 1992 approximately 2.05×10^6 t of wastes were deposited in the tailing dam. The amounts of the different materials deposited in the tailing dam are depicted in Table 4. Sulphate is precipitated as $CaSO_4$ and plays an important role in the immobilization of $^{226,228}Ra$. It results from the neutralization of the acid effluents with lime. Pyrolusite used to be added to the process to aid in the oxidation of U^{4+}–U^{6+}. The addition of phosphatic rock was introduced to recover Mo as a by-product of the uranium production. The average elemental composition of the deposited wastes is shown in Table 5. Fernandes et al. (1996) submitted the tailings to a simple leaching test with HCl 0.5 M to assess the potential mobility of the contaminants present in the tailings. They found out that virtually all the Mn present in the tailings was solubilized. As it will be discussed further in this text Mn is a critical contaminant amongst the chemical elements present in the tailings. Amongst the radioactive contaminants the solubilization of ^{226}Ra, ^{228}Ra, ^{210}Pb, ^{238}U and ^{232}Th were, respectively, 59%, 16%, 65%, 7%, 1% and 21%. The higher solubilization of ^{232}Th in relation to ^{238}U results from the fact that uranium was already removed from the

Table 4
Inventory of the wastes deposited in the tailings dam

Material	Mass (t)
Milled ore	1,764,976
Sulfate	135,168
Pyrolusite	35,049
Phosphatic rock	6770
Lime	109,950
Total	2,052,913

Table 5
Average composition of the wastes in the tailings dam

Species	Concentration (%)
ZrO_2	0.15
MoO_3	0.02
Al_2O_3	23.4
K	11.2
SiO_2	54.0
CaO	0.25
SO_4^{2-}	2.3
S^{2-}	0.5
Fe_2O_3	3.9
Mn	0.02
P_2O_5	0.09
U	0.018
Th	0.004
^{226}Ra (Bq/g)	2.5
^{210}Pb (Bq/g)	3.4
^{228}Ra (Bq/g)	1.4

Source: Fernandes et al. (1996).

ore in the leaching process and selectively extracted by an organic solvent.

Fernandes et al. (1996) described the main geochemical mechanisms taking place in the tailings environment. It was suggested that oxygen diffusion would be taking place through the first 1-m layer of the tailings. In addition to this it was observed that metals and radionuclides are not uniformly distributed in the solid material from the tailing dam. The possible mechanisms accounting for this observation are: (a) the accumulation of fines, including precipitates of gypsum and metal hydroxides, in the lower zone because of settling processes during initial deposition of the tailings; (b) the precipitation, co-precipitation and adsorption of major metals and trace radionuclides from neutralized process water; and (c) pyrite oxidation, leaching and desorption of metals and radionuclides in the top upper zone, producing highly acidic, high total dissolved solids (TDS) pore water which predominantly migrates downwards (solute transport) with precipitation, co-precipitation and absorption, occurring in the neutralized buffer zone (re-deposition). In summary, as infiltrating water moves down through the tailings, it attains a basic pH because of excess lime not yet leached from the system.

The depletion of metals and radionuclides like Fe, Al and U in the upper layers was probably related to acid leaching resulting from the residual pyrite oxidation of the tailings. In fact, the authors reported X-ray diffractograms showed the existence of pyrite in the tailings even after the H_2SO_4 attack. Oxidation of sulphide minerals in the tailings requires the availability of both oxygen (air) and water. Thus it is suggested that most of the sulphide oxidation will be concentrated in the aerated portion of the tailing dam.

Differently from Fe, Al and U, Ra isotopes and ^{210}Pb presented a different profile caused by the increase of sulphate concentration in solution due to the dissolution of sulphate from tailings, which induces the re-precipitation of radium. The direct precipitation of Ra, in the form of sulphate, however, does not seem to be the most probable mechanism to explain the radionuclide activity increase in solid phase, being the co-precipitation with Ca a much more probable mechanism. This might happen in a gypsum-containing soil and particularly uranium tailings, which is the case in the Poços de Caldas. Thus, the different shapes of ^{210}Pb and ^{226}Ra behavior in the tailings may be attributed to the immobilization caused by sulphate.

In the upper zone, pH would be as low as 3.0. However, pH values increase with depth (varying in the range of 5.0–6.0), a consequence of the introduction of the milling effluent with high pH values (in the range between 10 and 12) during the operation of the industrial plant. It was postulated that this region of the tailing dam would function as a buffer zone that would neutralize the acidic percolating waters moving downward. As a result, the migrating elements would precipitate in this higher pH zone. It was also estimated that the release of untreated waters from the tailing dam into the environment, would result in doses as high as 8.0 mS/y. This value is unacceptably higher than the primary limit established by the Brazilian regulatory authority, i.e., 1.0 mS/y.

4. Identification and analysis of the potential environmental impacts associated with mining and milling wastes

Deposited wastes may cause environmental impacts through different pathways as shown in Fig. 4. The extent of these impacts will determine the need for remedial actions to reduce the exposures to acceptable levels. Both radiological and non-radiological impacts must be taken into account. As for example, the removal of heavy metals from the effluents will also imply in the decrease of radionuclides from these materials.

Regarding the Poços de Caldas site, the release of acid drainage from the WRP directly into the environment will affect the quality of the receiving waters. Regarding the tailing dam, it has already been mentioned that the direct release of untreated effluents into the receiving waterbodies will result in unacceptable doses to members of the public. In addition, one can expect contamination of groundwater due to the infiltration of contaminated discharges. Exposure to radon and to gamma radiation may also result if houses are constructed directly over the tailings in a future scenario of unrestricted use of the site. As a result, all of these exposure pathways shall be considered in the definition of the best strategy of remediation.

The impacts of the release of acid drainage from the WRP into rivers may be assessed by a simple model of dilution (mixing zone model). One can assume that if the concentrations are low enough at this point, further populations would not be unduly exposed to the released contaminants. As a result, it was considered that a critical

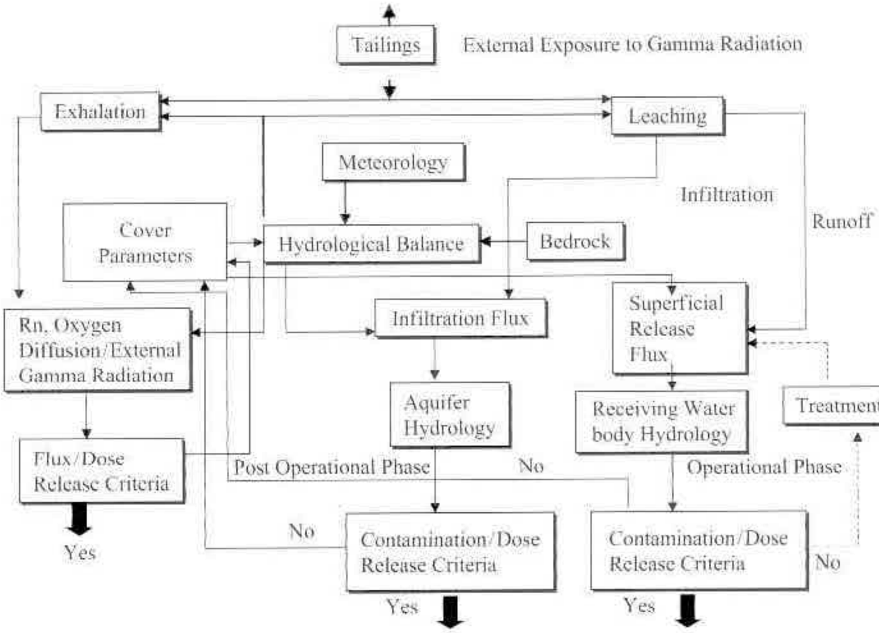


Fig. 4. Pathway analysis scheme for environmental impact assessment of disposed wastes.

group (i.e. group of the most exposed individuals) would be located within the zone of mixture of the acid drainage and river waters. The mixing zone model is represented by

$$C_{MZ} = \frac{(C_{\text{river}} * Q_{\text{river}}) + (C_{\text{eff}} * Q_{\text{eff}})}{Q_{\text{river}} + Q_{\text{eff}}} \quad (1)$$

where C is the concentration (mg/L or Bq/L), Q the flow ($m^3\,s^{-1}$).

Table 6 shows the results of the simulation and makes a comparison with Brazilian Standards. Average contaminants concentrations in the effluents were the same as those presented in Table 3. It can be seen that sulfate concentrations would not exceed the value established in the legislation but F^- and Al would.

Due to the location of the discharge point, the only exposure pathway predicted is the direct ingestion of waters. Assuming an ingestion rate of 2 L/d by an individual and the dose conversion factor (DCF) of 4.5×10^{-8} Sv/Bq for an adult (IAEA, 2001) the radiological dose due to ^{238}U would be equal to 0.21 mSv/y. This dose cannot be regarded as negligible, as it corresponds to 21% of the 1 mSv/y dose limit established in the Brazilian legislation. As per ^{226}Ra a dose as low as 0.0022 mSv/y can be estimated (using a DCF equal to 2.8×10^{-7} Sv/Bq). As a result, the remediation strategy would be governed by the concentrations of uranium and the non-radioactive contaminants present in the drainage and will have to produce reductions in the contaminant loads in such a way that the standards would not be exceeded.

Regarding the tailing dam, the contamination of groundwater in its area of influence was assessed by means of the code GWSCREEN (Rood, 1994). The code was developed for the assessment of the groundwater pathway from leaching of radioactive and non-radioactive substances from surface or buried sources. It simulates the transit of contaminants through different media, i.e., the source itself, the unsaturated zone and the saturated zone.

Table 6
Comparison between estimated contaminant concentrations in river waters and legal Brazilian standards

Contaminant	Estimated value ($mg\,L^{-1}$)	Brazilian standard ($mg\,L^{-1}$)
SO_4	38	250
F^-	4.67	1.4
Al	4.53	0.10

For modeling purposes, it was assumed that the tailings environment (i.e. the contaminated zone) was homogenous. For the sake of conservatism, the values of the coefficient of distribution, K_d for the whole contaminated zone were then represented by the ratios of radionuclide concentrations in the tailings (in $Bq\,kg^{-1}$) to those in the seepage water (in $Bq\,L^{-1}$) in the upper layers of the tailings, i.e., giving the lowest values K_d. The adopted K_d values for the contaminated zone were those reported by Fernandes et al. (1996): $^{238}U = 50\,L\,kg^{-1}$; $^{226}Ra = 800\,L\,kg^{-1}$ and $^{210}Pb = 200\,L\,kg^{-1}$.

For the saturated zone the K_d values used for ^{238}U and ^{226}Ra were those reported in IAEA (2001): 50 and 500 $L\,kg^{-1}$, respectively. As this publication does not present K_d value for ^{210}Pb in freshwater it was assumed as being the same of ^{226}Ra.

In fact, our aim in this assessment was not to achieve accurate estimates of the concentrations in the aquifer underneath the tailing dams. Instead, we were mainly interested in knowing the transit time from the contaminated zone to the aquifer. If this time were short (i.e. decades) more realistic simulations would have to be done. However, if the transit time was long enough (above 1000 years), even when conservative assumptions are made to preclude the need of the implementation of remedial actions, then the simulations should not need to be repeated and the contamination pathway would be excluded from further consideration. The assumption of the same value of K_{ds} for all the radionuclides leads to a conservative approach.

The transit time through the unsaturated zone was calibrated based on the knowledge of the time it took for sulfate anions to reach the aquifer. Since the operator runs a monitoring program that includes the sampling of underground water beneath the tailing dam, it was possible to establish the time since the initial deposition of the tailings for the anion to reach the aquifer. Sulfate concentrations in the aquifer have varied from 21 to 296 mg L^{-1}. Table 7 shows the input parameter values to the model and Table 8 shows the results obtained for the studied radionuclides.

It can be seen that a very long period of time will be necessary for the contaminants to reach the aquifer. It was found that the ^{210}Pb already exiting in the tailings, due to its relatively short half-life (22 y), will never reach the aquifer. However, the GWSCREEN software does not correctly take into account the decay of radionuclides as they move through the environment. If the time for the observation of the initial concentration of ^{226}Ra in the aquifer is taken into account, i.e., 4500 years, it can be estimated that roughly the 85% of the initial amount of

Table 7
Input values for the simulation of groundwater contamination due to the seepage water of the Poços de Caldas tailings dam

Parameter	Description	Value
L	Source length	432 (m)
W	Source width	432 (m)
T	Source thickness	12 (m)
P	Percolation rate	0.58 (m y^{-1})
Θc	Volumetric moisture content in contaminated zone	0.5 $m^3 m^{-3}$
Θu	Volumetric moisture content in unsaturated zone	0.1 $m^3 m^{-3}$
P	Bulk density of the contaminated zone	1.44 (g cm^{-3})
P	Bulk density of unsaturated zone	1.5 g (cm^{-3})
A	Longitudinal dispersivity	45 (m)
A	Transversal dispersivity	37 (m)
U	Saturated pore velocity	520 m y^{-1}
Q_0	Initial activity in source volume	$^{238}U = 0.01$ Ci
		$^{226}Ra = 138$ Ci
		$^{210}Pb = 186$ Ci
K_{ds}	Distribution coefficient in saturated zone	11 L kg^{-1} (assumed as being equal for all the radionuclides)[a]

[a]The assumption of the same value of K_{ds} for all the radionuclides leads to a conservative approach.

Table 8
Predicted peak concentration values in groundwater under the influence of the Poços de Caldas Tailings dam and time for observation of the peak concentration and transit time of contaminants to reach groundwater

Radionuclide	Concentration peak (Bq/L)	Time for peak concentration (y)	Contaminant transit time (y)
^{238}U	0.28	7700	4400
^{226}Ra	7.43	7800	4500
^{210}Pb	–	–	–
Mn			800

^{226}Ra will be converted leading to the generation of ^{210}Pb. This amount will be distributed through the water flow path and part of it may eventually reach the aquifer and will be added to the amount of ^{210}Pb to be observed in groundwater resulting from the decay of ^{226}Ra. For the non-radioactive contaminants, it was estimated that Mn is expected to reach the aquifer in 800 years.

The contamination of groundwater by contaminants in tailings deposited in impoundments will depend on a series of local variables turning the comparison between different sites not straightforward. However, some works available in the literature came to fairly similar results as those obtained in the present study. Davey and Grenn (1994) report that the movement of radioactive contaminants in the tailings of the Olympic Dam site ceases within 40 cm from the contact of the tailings with the soil being the ^{226}Ra and ^{210}Pb concentrations in the tailings (500–1500 Bq/kg) equivalent to those observed in the present work. Similarly, Silhanek and Kurokawa (1984) report ^{226}Ra contamination is not expected to take place in aquifers beneath uranium tailing dams, however, contamination from As, Se, Mo, cyanide and Cr (which are not an issue at Poços de Caldas) may take place. Rogers et al. (1982), adopted a leaching model that took into account meteorological data, infiltration rates, solubility limits and equilibrium coefficients. Contrarily to the results obtained in this study it was estimated that radium peak concentration would be observed in 105 years. The leaching rates these authors determined for Ra and U were in the range of 10^{-4}–10^{-5} y^{-1}, while those predicted by the model in this work varied between 10^{-3} and 10^{-4} y^{-1}.

The other pathways to be considered are exposure to radon and gamma radiation. The scenario considered in these cases was a house built directly over the tailings. In such a situation, an individual might inhale radon gas released from the tailings as well one might be exposed to external radiation from the gamma emitters present in the tailings. The external gamma and radon exposure assessment of the code RESRAD (Yu et al. 1993) was used. The results from the exposure to radon are shown in Fig. 5.

It can be seen that initial doses do not allow the unconditional use of the area for dwelling purposes. The decrease of the dose with time is a consequence of ^{226}Ra decay. Regarding the exposure to external gamma radiation, the shape of the dose vs. time curve is very similar to

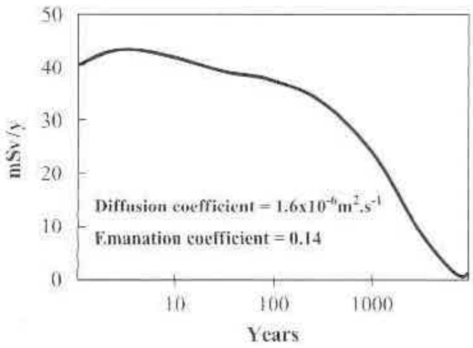


Fig. 5. Time dose evolution due to radon exposure.

Table 9
Main approaches to remediation of waste contaminated areas

Removal of source	• Bulk removal • Surface scraping • Turf cutting
Containment	• Capping • Subsurface barriers
Immobilization	• Cement-based solidification (*in situ and ex situ*) • Chemical immobilization (*in situ and ex situ*)
Separation	• Soil washing • Flotation • Chemical/solvent extraction

that of the radon. The effective dose estimated for the first year is 8.0 mSv (i.e. 20% of that estimated for radon exposure). Veska and Eaton (1991) reported effective dose to individuals exposed to tailings in a similar scenario to that used in this work as varying between 42 and 73 mSv y^{-1}. However ^{226}Ra concentrations in the tailings were 4–20 times higher than those observed in the present work. These assessments suggest as opposed to groundwater contamination observation that, remediation should be employed at the site to reduce potential exposures to radon and external gamma radiation.

5. Application of potential remediation strategies

5.1. General concepts

A number of strategies are available in the remediation of contaminated land. Some of them can be considered to be well-established and have been successfully applied to treat radioactive waste sites, others show considerable promise in laboratory and/or field trials. Table 9 summarizes some of these strategies.

In determining the feasibility of applying technologies in Table 9 for a particular contaminated site, a number of factors must be taken into account including:

- The characteristics of the site.
- The risk to the public.
- The performance and cost of the technique to be applied, and
- The exposure of the workforce during remediation work.

The impact of the different categories of remediation technologies may be described as follows:

- Removal of sources and separation technologies reduce the input of radionuclides into groundwater pathways and reduce the level of radon emissions by removing the parent nuclides with a proportionate decrease in the magnitude of all exposure pathways;
- Immobilization and containment (except capping) technologies reduce input to groundwater, and therefore have a significant impact on the off-site terrestrial and aquatic exposure pathways (ingestion and external irradiation), but have little impact on radon emissions;
- Capping reduces the input of radionuclides into groundwater and also reduces radon emissions in a manner proportional to the thickness of the surface barrier.

Removal of source material and capping of the waste area reduce external irradiation both on-site and off-site, in the case of capping this presumes that the covering material remains intact. For extreme intrusion scenarios, where penetration of the cap may be assumed, the on-site external radiation pathway will not be reduced. Immobilization and containment approaches, other than capping, tend to reduce off-site external irradiation, but may have little effect on-site as they may not provide adequate shielding against gamma radiation.

5.2. Costs evaluation

Table 10 depicts the costs of remediation activities that were carried out from 1997 to 2001 regarding only one of the waste rock piles (WRP 4). It can be seen that the high expenditures were practiced in relation to the pile reshaping, clay layer application (at the top of the pile) and revegetation (grass + trees). Because of the lack of a more accurate evaluation of the aspects involved on the acid drainage generation, it turned out that the adopted measures did not prove to be effective.

In addition to remediation costs of rock pile those related to acid water neutralization have to be taken into account. As reported by the mine operator these expenditures involving acid drainage originated in the mining site (all waste rock piles included) amounted to about US$ 3.0 million from 1984 to 2004.

5.3. Remediation options

Removal of the source, i.e., returning the tailings and the waste-rock material back to the open pit is one of the possible solutions to be assessed. The costs related to this option based on MEND (1995) are: hauling of the materials—US$ 1.5/tonrock; landfill operations—US$ 3.0/tonrock, and revegetation—US$ 5000/ha. A rough estimation of the overall cost for this option, taking only into account the material deposited in the WRP 4, would be about US$ 70 million. This seems to be a very high expenditure, therefore this option would not be recommended as an effective strategy to remediate the site.

Immobilization techniques, as stated above, may be effective in reducing migration of radionuclides from tailings, but not from the WRPs. In addition, radon exhalation and gamma exposure would not be addressed by this strategy. As a result of the above considerations, capping would be the preferred strategy. In order to evaluate the effectiveness of capping, the impact of the cap on acid drainage must be determined. First, one has to take into consideration that acid drainage is a result of pyrite oxidation. The supply of oxygen is the limiting force in pyrite oxidation. Bacterial activity may also play an important role in the process (Eqs. (2) and (3))

$$FeS_2 + 8H_2O + 14Fe^{3+} \rightarrow 2SO_4^{2-} + 15Fe^{2+} + 16H^+ \quad (2)$$

$$Fe^{2+} + \tfrac{1}{4}O_2 + H^+ \rightarrow Fe^{3+} + \tfrac{1}{2}H_2O \quad (3)$$

The oxygen removed from the system in the process of pyrite oxidation will create a concentration gradient in the rock pore spaces. A high oxygen demand will promote a pressure gradient inside the pile that will lead to the transport of air mass into the pile as a result of oxygen removal. At lower oxygen demands the transport of oxygen into the pile will be governed by diffusion. Ritchie (1994) introduced the concept of IOR, which is simply the rate of oxygen consumption in the material forming the pile in the prevailing conditions. The most practical unit to express the IOR is kg $(O_2)\,m^{-3}\,s^{-1}$. Table 11 shows some values of IOR and relevant characteristics associated with these values.

As has been demonstrated by Fernandes et al. (2000), IOR values applicable to WRP 4 of the Poços de Caldas mining site are of the order of 10^{-9} kg$(O_2)\,m^{-3}\,s^{-1}$. In such a situation diffusion dominates the process. Providing more oxygen to the system, in an attempt to speed up the oxidation process will not be an effective strategy since not all the oxygen provided to the system would be consumed under these conditions, i.e., low IOR values. In such cases capping the WRP with a material with a lower oxygen diffusion coefficient than the rocks forming the pile would be an effective manner to halt the input of oxygen into the system and consequently reduce the concentration of contaminants in the drainage. According to Ritchie (1995) the attenuation factor will depend more on the properties of the capping material than on the properties of the waste material. The practical problem to be addressed is to design an appropriate capping system in such a way

Table 10
Summary of the expenditures of the remedial actions of the WRP - 4

Activity	Cost (US$)
Reshaping	287,191
Clay Layer	171,442
Revegetation—grass	175,812
Revegetation—trees	66,419
Drainage work	53,092
Garden maintenance	75,643
Wells drilling	55,612
Technical consultancy	35,944
Dikes construction	14,618
Support activities	60,330
Total	996,103

Table 11
Significance of the IOR value magnitude

IOR kg $m^{-3}\,s^{-1}$	Time for pyrite consumption (year)	Time for the consumption of rock bearing carbonates	Time for the consumption of O_2 initially present in rock pore spaces	Heat Generation Rate (W m^{-3})	Sulfate concentration in the drainage (mg/L)	Relevance of the IOR
10^{-5}	0.167	2.4 d	–	129	–	Rate observed in laboratory experiments of biological oxidation
10^{-6}	1.67	3.4 w	0.9 d	12.9	–	Rate necessary for running bio-oxidation heaps.
10^{-7}	16.7	0.66 y	1.3 w	1.29	–	Effective extremely high rates for some waste-rock piles.
10^{-8}	167	6.57 y	13.1 w	0.129	21,600	Rate typically found in waste-rock piles
10^{-9}	1,670	65.7 y	2.52 y	0.013	2160	Typically low rates found in some waste-rock piles
10^{-10}	16,700	657 y	25.2 y	–	216	Rate associated with marginal environmental problems

Source: Ritchie (1995).

that its physical stability can be maintained in the long term.

For a cap of inactive material with thickness equal to X_c and diffusion coefficient D_c the oxidation rate can be determined using the following expression (see Ritchie, 1995):

$$\mathrm{GOR} = \sqrt{2C_0DS^*}(\sqrt{\alpha + n} - \sqrt{\alpha + n - 1}), \tag{4}$$

where

$$\alpha = \left(\frac{X_c}{D_c}\right)^2\left(\frac{S^*D}{2C_0}\right). \tag{5}$$

GOR is the global oxidation rate ($\mathrm{kg_{O_2}\,m^{-2}\,s^{-1}}$), X_c the thickness of the inactive material cap (m), S^* the Intrinsic Oxidation Rate ($\mathrm{kg_{O_2}\,m^{-3}\,s^{-1}}$), C_0 the oxygen concentration in air ($\mathrm{kg\,m^{-3}}$), D the Total oxygen diffusion coefficient in the pile ($\mathrm{m^2\,s^{-1}}$), D_c the Oxygen diffusion coefficient in the cap material ($\mathrm{m^2\,s^{-1}}$).

With the aid of Eq. (4), values of GOR were simulated as well as the sulfate load for different capping thicknesses while varying values of oxygen diffusion coefficients. It was assumed that the infiltration rate would be of the order of 50% of the total precipitation (equivalent to $0.85\,\mathrm{m\,a^{-1}}$). It has been demonstrated, however, that the infiltration rates may be reduced to values as low as 1–5% of the total precipitation rates with an effective cap (Harries, 1990). The resulting values are presented in Table 12 and pictured in Fig. 6.

It is well known that costs of remediation increase linearly with the attenuation of the environmental impacts. Fig. 6 suggests that capping the WRP 04 with a material with an oxygen diffusion coefficient of $10^{-9}\,\mathrm{m^2\,s^{-1}}$ and a thickness of 0.5 m (option 4) would be the most effective solution to reduce the concentrations of contaminants in the drainage leaving the system because increasing the thickness of the capping material or using materials with lower oxygen diffusion coefficients will not be more effective than what is obtained from option 4. Regarding ^{238}U activity concentrations, it was assumed that the radionuclide varies linearly with sulfate as suggested by Fernandes (2000). If the same reduction level is applied to the other contaminants (e.g. Al and F-) the concentrations in the mixing zone will be lower than the Brazilian legislation standards.

One of the issues regarding the capping of waste rock piles has to do with the erosion of the capping material. If clay is used as the capping material some sort of protective barrier should be applied. A general scheme to be put in place in those situations is a three layer barrier consisting

Table 12
Model results of ^{238}U and SO_4^{2-} in the drainage of the WRP 04 for varying oxygen diffusion coefficient and thickness of the capping material

Remedial Option	1	2	3	4	5	6	7	8	9
Dif. Coef. ($m^2 s^{-1}$)	1×10^{-8}	1×10^{-8}	1×10^{-8}	1×10^{-9}	1×10^{-9}	1×10^{-9}	1×10^{-10}	1×10^{-10}	1×10^{-10}
Thick (m)	0.5	1.0	2.0	0.5	1.0	2.0	0.5	1.0	2.0
GOR $kg/m^2 s^{-1}$ ($\times10^{-10}$)	180	2.2	1.1	0.20	0.1	0.02	2.0	1.0	0.05
[SO_4] $mg\,L^{-1}$	287	154	78	32	16	8.0	3.7	1.58	0.79
[^{238}U] Bq/L	49	26	13	5.4	2.7	1.3	0.63	0.27	0.13

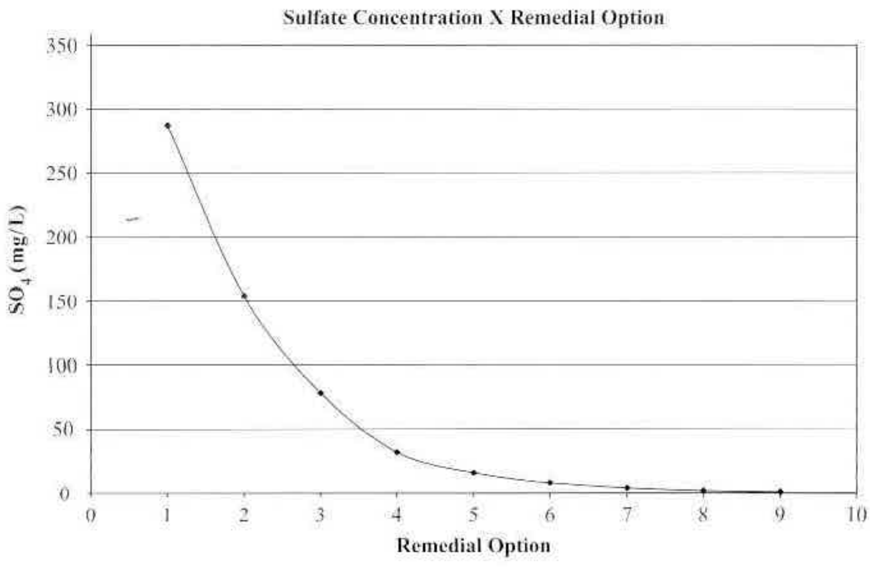


Fig. 6. Sulphate concentration variation in the drainage of the WRP-4 with different remedial options.

of a bottom granular layer, an intermediate one—the capping material itself, and a third layer consisting of gravel to protect the capping layer against erosion. Table 13 depicts the costs related to the application of each one of these layers to WRP 4. The costs are based on MEND (1995). The total cost would be about US$ 10 million for the remediation of only one of the existing waste rock piles in the Poços de Caldas mining site. These costs may be reduced if a less costly capping material (clay) can be obtained from neighbor borrowing areas.

Clearly, the figures from MEND (1995) are based on Canadian unit costs that may not be applicable to the Brazilian context. The merit of this approach is to give a cost perspective based on the experience of a developed country that holds a closer relationship to other similar countries in the world. Basing the calculations on Brazilian unit costs could lead to biased conclusion due to the eventual associated lower values of the used material and labor costs. Despite this a rough estimate can be made and a figure of US$ 3 million would be achieved by using a cost rate of US$ 5 per m^3 of manually compacted clay. This value represents 30% of the cost estimate based on the Canadian experience.

Harries (1990) reports that rehabilitation costs at the Woodlawn site in Australia, with similar problems to those of Poços de Caldas were of the order of US$22,000/ha. IAEA (1994) reports that decommissioning strategies adopted in different countries show average capital costs in the range of US$ 3–15 million depending on the complexity of the installation. Reported costs for remediation of waste rock piles were US$ 0.30/t of deposited material. Applying this figure to waste rock pile 04 we would have a value of about US$ 7.0 million, which is parallel with the estimated costs presented in Table 13.

Another strategy that could be applied to mitigate the impacts of acid drainage generated by pyrite oxidation in WRP 04 would be the economical recovery of uranium present in the drainage. Presently these drainages, along with the drainage from the other waste rock piles are being pumped to three main ponds and then neutralized with lime. The resulting slurry is deposited in the open pit. Table 14 presents the volumes being treated; the average uranium concentration in those waters and the resulting loads of ^{238}U.

Table 13
Formation costs for a three layer capping system to be applied to the WRP-4 as a remediation solution for the abatement of acid drainage generation

	US$/$m^2$	m^2	Total (US$)
Lower layer (sand type material)	3.5	56.5×10^4	1.98×10^6
Intermediate layer (compacted clay)	11	56.5×10^4	6.22×10^6
Upper layer (gravel)	3.0	56.5×10^4	1.70×10^6
Total	–	–	9.9×10^6

Table 14
U-238 activity concentrations, volume of drainage treated and U-238 load in the acid drainage of Poços de Caldas mining site

Pond	^{238}U (mg L^{-1})	Volume of water treated per year (L)	^{238}U Load (t)
1	16.6	8.9×10^8	16
2	24.8	2.7×10^8	8.4
3	80.2	3.9×10^7	4.0
Total	–		28.4

It can be seen from Table 14 that approximately 30 t of ^{238}U are disposed as waste per year. This figure represents 30% of the average annual production of the Poços de Caldas facility. The challenge here is to establish an efficient way to extract uranium from these waters. The ionic exchange technique has been adopted in similar situations elsewhere and is a mature technology. The most important exchangers use synthetic polymers to which the active functional groups are attached. Ionic inorganic exchangers, or zeolites, are also available (MEND, 1995). Critical to the success of ionic exchange resins in treating acid drainages is to identify the exchangers that will exhibit a specific selectivity for the metal of interest over those that are not to be separated like Fe, Al and Ca. MEND (1995) makes an assessment of the removal of metals from acid mine drainage. It is concluded that most of the investigated resins do not show a well-defined selectivity for the metals of interest—Sb, Cd, Cu, Ni and Zn—and that only the co-extraction of Fe was an obstacle for the application of resins to acid drainage.

The use of ionic exchange resins in the extraction of uranium from water has been reported by several authors (Rosenbaum and George, 1971; Garrido et al., 1981; Gow, 1985; Rhumer, 1985). Naden et al. (1985) estimated that capital costs involving the construction of an ion extraction plant would be about US$ 1.6 million.

The proposed scheme to be applied in the case of the Poç os de Caldas site would have to include the removal of iron from solution. That should be achieved by means of the aeration of the solution and the elevation of the pH to 4.5. Of concern here is the fact that iron removal by oxidative precipitation is likely to remove also at least some of the uranium (by e.g. sorption on the fresh precipitates). This issue will have to be examined in detail when the tests of the recovery scheme. The overall treatment process is shown in Fig. 7.

If it is assumed that the process has an efficiency of 90% and if the costs of uranium in the international market are taken into account, i.e., US$ 40/kg, it can be roughly estimated that about US$ 1.2 million would be gained with the recovery of uranium from the acid drainage. It was reported by the operator that about US$ 2.6 million were spent on water treatment or approximately $ 145,000 per year. As a result it can be proposed that the revenue obtained with uranium recovery could be deposited in

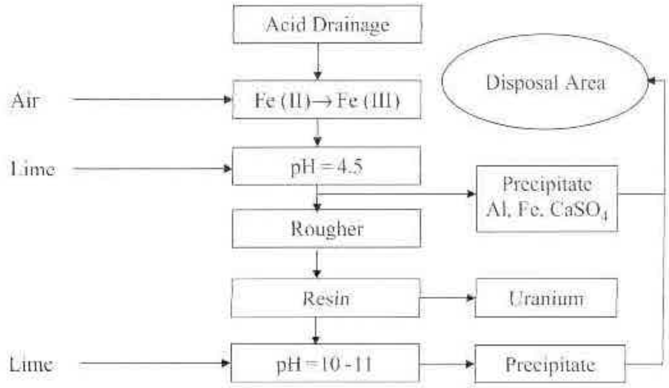


Fig. 7. Uranium recovery scheme from acid drainage.

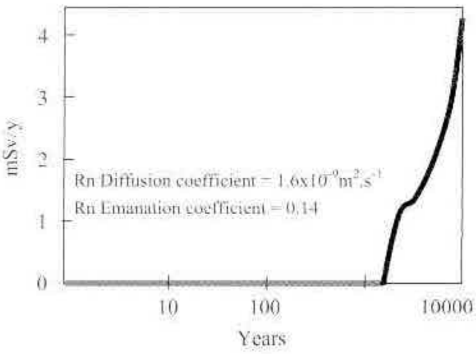


Fig. 8. Effect of capping the tailing dam with a compacted clay layer.

a fund to be used in future remediation of the site. It must be appreciated that with the termination of the secondary supply of uranium in the international market (represented by the highly enriched uranium—HEU—detained by the former Soviet Union Countries) prices of the commodity will increase, making the recovery of the metal even more attractive.

Regarding the remediation of the tailing dam, capping seems to be the most appropriate course of action. Capping will reduce radon emissions, will shield against the exposure to gamma radiation and additionally will prevent the diffusion of oxygen into the tailings. As a result, oxidation of pyrite will be reduced.

The computational code RESRAD (Yu et al., 1993) was used to examine the effect of application of a cap material over the tailings. As input data, a material with the same properties of that used in the waste rock remediation was tested, i.e., a material with an oxygen diffusion coefficient of $10^{-9}\,m^2\,s^{-1}$. For the purpose of our estimates radon diffusion coefficient was assumed as being equal to $10^{-6}\,m^2\,s^{-1}$ (Chauhan and Chakarvarti, 2002). Fig. 8 displays the results.

It can be seen that the cap would reduce the doses (both exposure to radon and gamma radiation) to values virtually equal to zero. The increase after 1000 years is due to the erosion of the cap material. In the absence of site-specific information the erosion rate of $6.0 \times 10^{-5}\,m\,a^{-1}$ was adopted. According to Yu et al. (1993) this value would be typical of a soil not subjected to any agricultural practice. Values in the range of 10^{-7}–$10^{-4}\,m\,a^{-1}$ are reported in the literature and depend, besides the vegetation, on the slope of the land, its use and type of cap material.

The costs associated with this strategy are estimated as being on the order of approximately US$ 3.7 million if the same scheme used in the application of a cover in the WRP 04 is adopted in the tailing dam. Additional geotechnical work would be required for the tailing dam as opposed to the WRP, and this additional work may result in extra costs not projected in this work. Details of the application of dry cover to a tailing dam can be seen in Leoni et al. (2004).

6. Conclusions

This paper presented an analysis of the waste management strategies adopted as it proposes course of action to be put in place in the future in the uranium production center of Poços de Caldas. It was observed that due to a lack of a careful initial planning, significant expenditures are necessary to mitigate the environmental impacts associated with two of the most relevant contaminant sources, i.e., the tailing dam and the waste-rock piles. Additional expenditures will have to be put in place to remediate the site from an effectiveness and long-term perspective. It was demonstrated that the characteristics of the ore (occurrence of pyritic material), climatological conditions (high precipitation rates) and social-economical issues (existence of population groups and intensive agriculture activities in the neighborhood of the mining site) together contribute to the existence of potential scenarios that may lead to unacceptably high exposures to radiation not only in the present but also in the future. The long-term component of the contaminant emissions into the environment, for instance preclude the adoption of active controls, e.g., collect-and-treat strategies due to economic consideration. Instead, passive controls have to be put in place. In the case of the waste-rock piles it was demonstrated that covering these entities with a material layer with an oxygen diffusion coefficient of $10^{-9}\,m^2\,s^{-1}$ and 0.5 m thick will bring the contaminant releases down to marginal levels. The same principle can be applied to the tailing dam (in which pyrite is still present). Covering the tailings with a 1.0 m thick clay layer will also prevent radon emissions as well as providing shield against gamma radiation. Modeling of the transport of contaminants to groundwaters showed that contamination of this medium by radionuclides and heavy metals will not be relevant. As an alternative route to deal with the remediation of the contaminant sources, it was demonstrated that the recovery of uranium from the acid drainage waters would be economically feasible and the revenues coming from this process should be invested in the overall remediation plan of the mining site.

Acknowledgement

The authors gratefully acknowledge Dr. Debra Reinhart for reviewing the manuscript of this paper.

References

Amaral, E.C.S., Godoy, J.M., Rochedo, E.R.R., Vasconcelos, L.M.H., Pires do Rio, M.A., 1988. The Environmental Impact of the Uranium Industry: is the Waste Rock a Significant Contributor? Radiation Protection Dosimetry 22 (3), 165–171.

Chauhan, R.P., Chakarvarti, S.K., 2002. Radon diffusion through soil and fly ash: effect of compaction. Radiation Measurements 35, 143–146.

Davey, J.F., Grenn, S.F., 1994. Natural clay based soil as an effective barrier against radionuclides migration in a uranium tailings structure. Radiation Protection in Australia 12 (4), 135–138.

Diehl, P., 2004. Uranium mining and milling wastes: an introduction. WISE Uranium Project 〈http://antenna.nl/wise/uranium/uwai.html〉.

Fernandes, H.M., Franklin, M.R., 2000. Assessment of acid rock drainage pollutants release in the uranium mining site of Poços de Caldas—Brazil. Journal of Environmental Radioactivity 54 (1), 5–25.

Fernandes, H.M., Franklin, M.R., Veiga, L.H., Freitas, P., Gomiero, L.A., 1996. Management of uranium mill tailings: geochemical processes and radiological risk assessment. Journal of Environmental Radioactivity 30 (1), 69–95.

Fernandes, H.M., Franklin, M.R., Veiga, L.H., 1998. Acid rock drainage and radiological environmental impacts. A study case of the uranium mining and milling facilities at Poços de Caldas. Waste Management 18 (3), 169–181.

Garrido, P., Andalafat, E., Soto, R., 1981. Planta Piloto de Recuperacion de Uranio a Partir de Soluciones de Lixiviacion de Cobre en Chuquicamata Sur, Chile. In: Proocedings Uranium Deposits in Latin America: Geology and Exploration. International Atomic Energy Agency, Vienna, pp. 581–594.

Gow, A., 1985. Recent advances in uranium ore processing. In: Proceedings of a Technical Committee Meeting Advances in Uranium Ore Processing and Recovery from Non-Conventional Resources. International Atomic Energy Agency, Vienna, pp. 3–14.

Harries, J.R., 1990. Acid drainage from waste rock dumps at mine sites. Technical Report ANSTO/E692. Australian Nuclear Science and Technology Organisation.

IAEA, 1994. Decommissioning of facilities for mining and milling of radioactive ores and closeout residues. International Atomic Energy Agency. Technical Report Series No. 362. Vienna, 186p.

IAEA, 2001. Generic Models for Use in Assessing the Impact of Discharges of Radioactive Substances to the Environment. Safety Reports Series No. 19. International Atomic Energy Agency, Vienna, 229p.

IAEA, 2004. The long term stabilization of uranium Mill tailings. IAEA-TECDOC-1403. Vienna, 311p.

Leoni, G.L.M., Almeida, M.S.S., Fernandes, H.M., 2004. Computational modeling of final covers for uranium mill tailings impoundments. Journal of Hazardous Materials 110 (1–3), 139–149.

MEND, 1995. Metals Removal from Acid Mine Drainage by Ion Exchange. Mine Environment Neutral Drainage Program, MEND Report 3.2.1.1 (b). Ontario, 55p.

Naden, D., Rowden, G.A., Lunt, D.J., Wilson, D., 1985. Development and application of solvent extraction, resin ion exchange (RIP) and liquid membrane processes for uranium recovery. In: Proceedings of a Technical Committee Meeting Advances in Uranium Ore Processing and Recovery from Non-Conventional Resources. International Atomic Energy Agency, Vienna, pp. 53–72.

NEA, 2002. Environmental remediation of uranium production facilities. A Joint Report by the OECD Nuclear Energy Agency and the International Atomic Energy Agency, Paris. 323p.

Ritchie, A.I.M., 1994. Sulfide oxidation mechanisms: controls and rates of oxygen transport. In: Jambor, J.S., Blowes, D.W. (Eds.), Short Course Handbook on Environmental Geochemistry of Sulfide Mine-Wastes. Mineralogical Association of Canada, Ontario, pp. 210–244.

Ritchie, A.I.M., 1995. Application of oxidation rates in rehabilitation design. In: Grundon, N.J., Bell, L.C. (Eds.), Second Australian Acid Mine Drainage Workshop. Charles Town, pp. 101–116.

Rogers, V.C., Nielson, K.K., Rich, D.C., Grant, M.W., Mauch,M.L., Sandquist, G.M., 1982. Uranium mill tailings containment systems performance and cost. In: Management of Wastes from Uranium Mining and Milling. IAEA SM-262/43. International Atomic Energy Agency, Vienna, pp. 197–212.

Rood, A.S., 1994. GWSCREEN: A Semi-Analytical Model for Assessment of the Groundwater Pathway from Surface or Buried Contamination. Theory and User's Manual Version 2.0. EGG-GEO-10797. Idaho National Engineering Laboratory.

Rosenbaum, J.B., George, D.R., 1971. Cost reduction in ion exchange processing of uranium ores. In: Proceedings Symposium—The Recovery of Uranium. International Atomic Energy Agency, Vienna, pp. 297–307.

Silhanek, J.S., Kurokawa, Y., 1984. Pollution abatement. In: The Behaviour of Radium in Waterways and Aquifers. IAEA-TECDOC-301. International Atomic Energy Agency. Vienna, pp. 237–243.

Veska, E., Eaton, R.S., 1991. Abandoned rayrock uranium mill tailings in the northwest territories: environmental conditions and radiological impact. Health Physics 60 (3), 399–409.

Waber, N., Schorscher, H.D., Peters, E., 1991. Mineralogy, petrology and geochemistry of the Pocos de Calda analogue study sites, Minas Gerais. Technical Report TR 90-11. Swedish Nuclear Fuel and Waste Management Co., 513p.

Westall, J.C., Zachary, J.L., Morel, F.N.M., 1976. MINEQL—A Computer Program for the Calculation of Chemical Equilibrium Composition of Aqueous Systems. Technical Note, No. 18. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge.

Yu, C., Loureiro, C., Cheng, J.J., Jones, L.G., Wang, Y.Y., Chia, Y.P., Failace, E., 1993. Data Collection Handbook to Suport Modeling the Impacts of Radioactive Material in Soil. ANL/EAIS-8. Environmental Assessment and Information Science Division. Argone National Laboratory, Illinois, 149p.

Zenaro, R., 1989. Contribuição ao Estudo da Migração de Urânio e de Alguns Metais Traço em Solução na Mina de Urânio de Poços de Caldas. Rio de Janeiro. 152p. Msc. Dissertation. Departamento de Química. Pontifícia Universidade Católica.

ELSEVIER

Journal of Environmental Management 87 (2008) 73–79

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Options for management of municipal solid waste in New York City: A preliminary comparison of health risks and policy implications

Pearl Moy[a], Nikhil Krishnan[b], Priscilla Ulloa[b], Steven Cohen[c], Paul W. Brandt-Rauf[a,b,*]

[a]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, 60 Haven Avenue, NY 10032, USA
[b]Department of Earth and Environmental Engineering, Fu Foundation School of Engineering and Applied Science, Columbia University, NY 10027, USA
[c]Program in Environmental Science and Policy, School of International and Public Affairs, Columbia University, NY 10027, USA

Received 24 December 2005; received in revised form 14 November 2006; accepted 4 January 2007
Available online 26 March 2007

Abstract

Landfill disposal and waste-to-energy (WTE) incineration remain the two principal options for managing municipal solid waste (MSW). One critical determinant of the acceptability of these options is the different health risks associated with each. In this analysis relying on published data and exposure modeling, we have performed health risk assessments for landfill disposal versus WTE treatment options for the management of New York City's MSW. These are based on the realistic scenario of using a waste transfer station (WTS) in Brooklyn and then transporting the untreated MSW by truck to a landfill in Pennsylvania or using a WTE facility in Brooklyn and then transporting the resultant ash by truck to a landfill in Pennsylvania. The overall results indicate that the individual cancer risks for both options would be considered generally acceptable, although the risk from landfilling is approximately 5 times greater than from WTE treatment; the individual non-cancer health risks for both options would be considered generally unacceptable, although once again the risk from landfilling is approximately 5 times greater than from WTE treatment. If one considers only the population in Brooklyn that would be directly affected by the siting of either a WTS or a WTE facility in their immediate neighborhood, individual cancer and non-cancer health risks for both options would be considered generally acceptable, but risks for the former remain considerably higher than for the latter. These results should be considered preliminary due to several limitations of this study such as: consideration of risks only from inhalation exposures; assumption that only volume and not composition of the waste stream is altered by WTE treatment; reliance on data from the literature rather than actual measurements of the sites considered, assuming comparability of the sites. However, the results of studies such as this, in conjunction with ecological, socioeconomic and equity considerations, should prove useful to environmental managers, regulators, policy makers, community representatives and other stakeholders in making sound and acceptable decisions regarding the optimal handling of MSW.

Keywords: Municipal solid waste; Landfill disposal; Waste-to-energy incineration; Health risk assessment

1. Introduction

Despite increased efforts to prevent, reduce, reuse and recycle waste, the appropriate management of municipal solid waste (MSW) remains a major environmental issue (Landreth and Rebers, 1997; Williams, 2005). Currently, there are two principal options for managing such MSW—landfill disposal or incineration in waste-to-energy (WTE) facilities (Landreth and Rebers, 1997; Williams, 2005). However, concerns have been raised in the past that emissions from both landfills and incinerators may pose environmental health risks that make both options less than optimal (Rushton, 2003). Both of these technologies have been improved in the last 20 years. Modern landfills are required by Subtitle D rules (Lee et al., 2000) to include a non-permeable liner at the bottom, be capped at the top, and contain and treat emissions as much as possible (Landreth and Rebers, 1997; Williams, 2005). WTE

*Corresponding author. Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, 60 Haven Avenue, NY 10032, USA. Tel.: +1 212 305 3959; fax: +1 212 305 4012.
E-mail address: pwb1@columbia.edu (P.W. Brandt-Rauf).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.032

facilities, through the implementation of EPA Maximum Achievable Control Technology (MACT) standards, have reduced emissions of certain hazardous materials including heavy metals and dioxins by a factor of almost 100 (Williams, 2005). Nevertheless, there is a continuing debate over which option, landfill disposal or WTE treatment, poses less risk to the environment and human health, the latter concern usually being the most important for affected populations (Rushton, 2003).

The present study is a preliminary attempt to quantify and compare the health risks from landfill disposal and WTE treatment using the principles of risk assessment. The study focuses on one hypothetical scenario of MSW management in New York City (NYC), which generates large amounts of MSW and is searching for more effective methods for its handling. NYC currently exports most of its MSW to out-of-state landfills that are constrained by decreasing capacity and thus charge increasing tipping fees (Tammemagi, 1999). On the other hand, there is considerable community resistance to siting a WTE facility in NYC due in large part to concerns over associated potential health risks (Tammemagi, 1999). Therefore, a comparison of the health risks for these two options could be useful for environmental managers, regulators and policy makers, as well as other concerned stakeholders including the affected communities, in terms of reaching consensus regarding the most acceptable option for the future handling of MSW.

2. Methods

The objective of this study was to use risk assessment methodology (NRC, 1994) to estimate and compare the human health impacts of inhalation exposure from emissions from landfill disposal and WTE treatment of managing one million tons of MSW in 1 year in NYC. Based on the literature (Landreth and Rebers, 1997; Rushton, 2003; Williams, 2005), it was assumed that inhalation represents the most significant route of exposure; although indirect pathways of exposure are know to exist, they could not be adequately considered in this study due to a lack of relevant data on the sites.

The study thus considered two specific options: (1) siting of a waste transfer station (WTS) in Greenpoint, Brooklyn, for collection of MSW with transportation via diesel trucks to final disposal in a landfill in Dunmore, PA; or (2) siting a WTE facility in Greenpoint, Brooklyn, for combustion and electricity generation, with transportation via diesel trucks for the resultant ash to final disposal in a landfill in Dunmore, PA (see Fig. 1). The first option represents the predominant method of MSW disposal in NYC. Currently, 71% of all of NYC's MSW and all of Brooklyn's MSW is transported via trucks to landfill disposal in PA, OH or VA (Tammemagi, 1999). The second option has the advantage of reducing the volume of the waste stream that requires landfill disposal by up to 90% (Tammemagi, 1999); in this study, we assumed a relatively conservative reduction of the volume of MSW by 80% using the WTE option.

2.1. WTS scenario

Greenpoint, Brooklyn, was chosen as the particular site for facility placement because it represents a typical mixed-use municipal neighborhood containing multiple industrial facilities and residential dwellings. More importantly, it has been considered by the city as a potential location for a large marine WTS that could handle this scale of MSW volume; in anticipation of such a siting, the NYC Department of Sanitation has recently prepared an environmental impact statement for this facility which includes an assessment of the hazardous exposures and estimated resultant health risks associated with the facility which we could use for direct comparison to the risks associated with a WTE facility at the same site (NYC DOS, 2005). This is one of the few, if not only, health-risk assessments for a WTS available in the literature. This study estimated typical exposures to pertinent hazardous agents from a WTS (benzene, formaldehyde, 1,3-butadiene, acetaldehyde, benzo(a)pyrene, propylene, acrolein, toluene, xylenes, anthracene, benzo(a)anthracene, chrysene, naphthylene, pyrene, phenanthrene, dibenz(a,h)anthracene) and used accepted unit cancer and non-cancer risk factors for each to calculate individual cancer and acute and chronic non-cancer health risks (NYC DOS, 2005). Thus, the individual cancer risk and the sum of the individual acute and chronic non-cancer risk values from this report were used directly in this study in combination with similar risk values for landfilling (from the literature as described below) and truck transportation (calculated as described below) in the evaluation of the first option.

2.2. Landfilling scenario

Although no health-risk assessments specific to landfill facilities for NYC MSW such as the one in Dunmore, PA, have been performed, health-risk assessments have been performed previously for typical modern MSW landfill facilities that include impermeable liners, leachate collection and gas emission collection for power generation. A study by Manca et al. (1997) was chosen for use in this case because of similarities to the Dunmore site in terms of the technology employed, environmental conditions and receptor distribution. This study estimated typical exposures from pertinent hazardous agents from a landfill facility (1,1-dichloroethane, vinyl chloride, bromodichloromethane, 1,1-dichloroethylene, 1,1,2,2-tetrachloroethane, methylene chloride, trichloroethylene, 1,2-dichloroethane, 1,1,2-trichloroethane, chloroethane, benzene, methyl mercaptan, ethyl mercaptan, hydrogen sulfide, iron, zinc, lead) and used accepted cancer and non-cancer unit risk factors for each to calculate individual cancer and non-cancer health risks (Manca et al., 1997). Thus, the individual cancer and non-cancer risk values from this study were used in the present study, adjusted for the difference in waste volume between the two studies (6.6 versus 1.0 million tons for the Manca study and our study,

Fig. 1. Map illustrating the siting of a WTS or WTE facility at Greenpoint, Brooklyn (center triangle) in New York City and a landfilling facility in Dunmore, PA (inset) for disposal of MSW or WTE ash (adapted from DeAngelo, 2004).

respectively), in combination with similar risk values for a WTS (as described above) and truck transportation (as described below) in the evaluation of the first option. In addition, similar cancer and non-cancer risk values were used in the present study, adjusted for the difference in waste volume reduction for WTE treatment (as noted above, assumed to be 80%), in combination with similar risk values for WTE (calculated as described below) and truck transportation (as described below) in the evaluation of the second option. In this case, an additional assumption was made that only the volume, not the nature of the hazardous components, of the MSW was altered by the WTE treatment; this is clearly an over-simplification because in some cases, e.g., embedding of the ash in lime would raise the pH and inhibit metal contaminant mobility. Thus, one might expect the total concentration of hazardous materials to vary somewhat after WTE treatment, but these potential changes were ignored in this first approximation due to a lack of relevant data. It should be noted that these changes would likely decrease the risks associated with landfilling ash so that not taking this into account represents a worst case scenario, providing an upper bound on the associated risks.

2.3. Truck transport scenario

One of the current sites for landfill disposal for NYC MSW is in Dunmore, PA, approximately 190 miles from Greenpoint. A box model incorporating vertical and horizontal dispersion and wind speed was used to estimate the exposures and resultant health impacts from diesel truck transport of the MSW or WTE ash (estimating 20 ton transported per truck) over this route (Derwent et al., 1995). The model assumes a uniform distribution of emission of pollutants along the route with vertical and horizontal dispersions of 1000 and 3000 m, respectively, chosen based on estimates from studies of mixing depths over the northeastern USA (Berman et al., 1999). Average wind speed for NY, NJ and PA was obtained from the 1996 Wind Energy Resource Atlas of the USA (NREL, 1996). The diesel pollutants of major concern were considered to be nitrogen oxides (NO_x) and particulate matter (PM) with emission rates based on 1998 EPA emissions standards for new trucks (US EPA, 2002). Ground level exposure concentrations for NO_x and PM were calculated from these emission rates applied to the box model, and average annual concentrations were converted to individual cancer and non-cancer health-risk values through multiplication by unit cancer (for PM) and non-cancer (for NO_x and PM) risk factors taken from the published literature (CA OEHHA, 2005; US EPA, 2005).

2.4. *WTE scenario*

For the WTE facility, major pollutants of concern were considered to be: dioxins and related compounds (combined as TEQ), mercury, cadmium, lead, PM, HCl, SO_2 and NO_x (NRC, 2000; Rushton, 2003). Emission factors for each were taken from the EPA Performance Data for Large Municipal Waste Combustors at MACT (US EPA, 2000) to generate emission rates for 1 million tons MSW processed. These emission rates, along with typical WTE facility characteristics (e.g., a stack height of 85 m)(EBI USA, 2005), were used to calculate ground level concentrations using a standard Gaussian plume air dispersion model (Boudet et al., 1999) with reflection, using the McElroy–Pooler values and incorporating wind rose data (speed, direction and temperature) for nearby LaGuardia airport from the National Climate Data Center and NOAA, in 100 m^2 grids around the WTE facility (Wark et al., 1998; NOAA, 2005). Average annual concentrations for each pollutant were converted to individual cancer and non-cancer risk values through multiplication by unit cancer and non-cancer risk factors taken from the published literature (US EPA, 1997, 2005; CA OEHHA, 2005).

For each specific step above (WTS, landfill of untreated MSW or ash, transport of untreated MSW or ash, WTE), individual cancer and non-cancer risk values for each pollutant of concern were summed to yield total individual cancer and non-cancer risk values. Then, for each treatment option (WTS/untreated MSW transport/landfill or WTE/ash transport/landfill), total individual cancer and non-cancer health risks were calculated and compared.

3. Results

The individual cancer and non-cancer health risks for each specific step (WTS, landfill of untreated MSW or ash, transport of untreated MSW or ash, WTE) are presented in Table 1. Acceptable excess cancer risks for the general population are usually considered to be less than one excess cancer case per 10,000–1,000,000 people (1.0E−04 to 1.0E−06)(NRC, 1994). For the steps of WTS, untreated MSW transport, ash transport, and WTE, the excess cancer risks are well below this range, and for untreated MSW landfill and ash landfill, they are still well within this acceptable range. Acceptable non-cancer risks are usually considered to have a Hazard Quotient less than 1.0 (i.e., Chronic Daily Intake is less than the Reference Dose, which is defined as that chronic dose that is unlikely to result in deleterious health effects with an adequate margin of safety) (NRC, 1994). Once again, for the steps of WTS, untreated MSW transport, ash transport, and WTE, the non-cancer risks are well below this value. However, in this case, the untreated MSW landfill and ash landfill are above this value, and the former is well above.

The total individual cancer and non-cancer health risks for each treatment option (WTS/untreated MSW transport/untreated MSW landfill or WTE/ash transport/ash landfill) derived by summing the appropriate items from Table 1 are presented and compared in Table 2. For the both options, the individual cancer risk is within the acceptable range (4.0E−05 and 7.9E−06, respectively), but the individual non-cancer risk is well above the acceptable

Table 1
Individual health risks associated with various handling methods for NYC MSW[a]

Method		Individual cancer risk	Individual non-cancer risk
WTS in Brooklyn		1.3E−07	4.6E−01
Transport to PA Landfill	Untreated MSW	3.9E−07	5.2E−04
	WTE Ash	7.8E−08	1.0E−04
Landfilling in PA	Untreated MSW	3.9E−05	1.2E+01
	WTE Ash	7.8E−06	2.3E+00
WTE in Brooklyn		6.6E−08	6.3E−04

[a]Based on handling one million tons of MSW in 1 year.

Table 2
Comparison of individual health risks associated with landfill disposal versus WTE treatment for NYC MSW[a]

Option	Individual cancer risk	Individual non-cancer risk
Landfill disposal	4.0E−05	1.2E+01
WTE treatment	7.9E−06	2.3E+00
Ratio of landfill disposal risk/WTE treatment risk	5.0	5.2

[a]Based on handling one million tons of MSW in 1 year.

value of 1.0 (1.2E+01 and 2.3E+00, respectively). However, for both individual cancer and non-cancer risks, the first option yields risks that are higher than the second option, approximately by a factor of 5 (5.0 for cancer risk and 5.2 for non-cancer risk).

4. Discussion and conclusions

Based strictly on the outcome of the health-risk assessments, one would conclude that WTE treatment is a better option than landfilling for NYC MSW due to the differences in non-cancer and cancer health risks noted above. Furthermore, it should be noted that more expensive technology currently exists (and is mandated for use to meet the more stringent European Union emissions standards) that would make the WTE emissions even lower, thus further favoring this option from a health risk perspective.

However, the limitations of this study should be kept in mind. First, the landfill risk estimates were not based on a NYC landfill but were based on literature estimates of typical municipal landfill chemicals of concern, although the particular study chosen is similar to others in the literature (as described below) and there is no reason to assume that a NYC landfill would differ significantly from these. Second, the box model used for truck transport exposures is relatively simple and does not incorporate the impact of meteorological conditions (other than wind) and terrain type which could alter the exposures to pollutants. In addition, many NYC landfill sites are farther away than Dunmore, PA, so exposed populations at risk could be larger. Conversely, as diesel truck fleets for transport of untreated MSW or WTE ash age and are replaced by trucks that meet more stringent 2004 emissions standards, the corresponding exposures and health risks should decrease. As noted, we have used a relatively conservative assumption of only 80% reduction in waste volume from WTE treatment. Larger reductions in volume from WTE treatment, which are clearly achievable with state-of-the-art facilities (Landreth and Rebers, 1997; Williams, 2005), would result in even bigger differences between the two MSW management options, further favoring the second option. Also, as noted, it is likely that WTE treatment would alter not only the volume of waste but also the composition of the waste to be landfilled, and this may conceivably result in even lower health risks from landfilling WTE ash compared to untreated MSW (Landreth and Rebers, 1997; Williams, 2005), again further favoring the second option. The health-risk assessment for the WTS was based on a study for a proposed marine WTS, which would involve less truck traffic, at least in the immediate vicinity, and alteration in the transportation-associated risks because subsequent truck-based transport routes would be different. Finally, other plume dispersion models are available (Derwent et al., 1995) that might provide different, and perhaps better, estimates of exposures and risks from a WTE facility.

Nevertheless, despite the uncertainties and assumptions involved in this study, the results are in reasonable agreement with those from other studies in the literature. For example, a health-risk assessment of a WTE facility in Montgomery County, MD, with a similar pollution emission profile, estimated cancer risks ranging from 1.07E−07 to 2.41E−08 based on actual emissions data, compared to a cancer risk of 6.55E−08 in this study, which is right in the middle of this range (Rao et al., 2003). As noted above, health-risk values from other landfill studies also agree reasonably with those used for this study (e.g., a cancer risk of 2.0E−05 compared to 4.0E−05 here) (Redfearn and Roberts, 2002). Eschenroeder and von Stackelberg (1999) compared the health risks from landfills and WTE facilities and found cancer risks for the former of 1.1E−05 and for the latter of 4.0E−06, similar to those from this study. Finally, a study by the Ontario Ministry of the Environment (1999) estimated cancer risks ranging from 4.0E−06 to 1.0E−05 for landfills and 4.7E−08 to 2.3E−07 for WTE facilities, similar to those from this study. Therefore, although this study cannot be considered definitive in terms of the health risks for landfill or WTE facilities, its results are consistent with others and thus plausible.

However, it must also be recognized that the cancer and non-cancer disease risks in this study are distributed across different populations (those around the landfill, those along the truck route, and those around the WTS or WTE facility). A more critical issue for environmental managers, regulators, policy makers and community representatives in NYC could be the difference in health risks posed only by siting a WTS as opposed to a WTE facility to the population in their immediate neighborhood. Even by this comparison, a WTS is a less healthy option than a WTE facility for the neighboring population, although for both the risks are within acceptable limits. For example, the individual non-cancer risk from a WTS (4.6E−01) is 727 times higher than that from a WTE facility (6.3E−04), and the individual cancer risk from a WTS (1.3E−07) is 2 times higher than that from a WTE facility (6.6E−08), in this population.

Given the limitations of the current study noted above, it is still possible that these results alone are insufficient for making decisions on MSW handling options that are acceptable to all stakeholders. Additional studies of this nature, however, can be used to help further refine the decision-making process. Future studies could address the limitations of the current study, for example, by using different dispersion models, multi-compartment and/or scenario analyses, inclusion of secondary pathways of exposure to pollutants, emissions data from actual facilities, sensitivity analyses that vary the parameters of the models, and statistical uncertainty analyses such as Monte Carlo methods for better quantification of the range of uncertainties in the risk estimates, among others. Finally, innovative approaches to health risk assessment based on molecular epidemiologic methods could incorporate biomarker measurements as estimates of the actual absorbed doses of pollutants of concern from these facilities in exposed populations. For example, in previous studies of workers at NYC's former MSW incinerators, we were able to identify elevated levels of biomarkers for metal (lead) and organic (dioxins) contaminants of incinerator ash in some individuals (Schecter et al., 1991; Malkin et al., 1992); more recent studies have followed up on this approach and confirmed elevated levels of dioxins in incinerator workers in other countries (Kumagai and Koda, 2005; Shih et al., 2006). Although it should be recognized that emissions for lead and dioxins have likely decreased in recent years (Chang and Lin, 2001; Chang et al., 2001; Domingo et al., 2001; Meneses et al., 2004), applying similar approaches for measuring biomarkers in exposed populations around current state-of-the-art facilities may still be useful in providing more persuasive evidence for (or against) the choice of particular options for MSW management based on health considerations (Gonzalez et al., 2000). For example, it would be useful to recruit a population in Greenpoint now before a facility is sited there and measure the relevant biomarkers and then again after the facility is in operation; this could help define more precisely the contaminant burden due to the facility as opposed to other confounding sources.

In the final analysis, though, such health considerations will be only one, albeit an important one, of the considerations of environmental managers, regulators, policy makers, community representatives and other stakeholders must take into account in choosing options for MSW handling. Ecological impact on non-human populations may be a significant concern in certain situations. Broader economic and social impacts and considerations of environmental justice and equity will also remain central issues in any future decision-making on MSW management. Thus, as has been discussed in detail elsewhere, the issue of NYC's MSW has multiple dimensions, including ethical, political, technological, policy and managerial elements (Cohen, 2006). However, progress to date has been stymied primarily because the politics of siting of MSW facilities, particularly "not-in-my-back-yard" self-interest politics, tends to dominate the other aspects. Therefore, it is unlikely that a resolution of NYC's MSW problems will be forthcoming in the absence of considerable community-based environmental education efforts. As out-of-state landfill options become more restricted and their tipping fees increase further, such educational efforts will be useful as communities are forced to face the trade-offs of the various MSW management options, such as hosting a WTS versus a WTE facility (Cohen, 2006). Scientific considerations, such as the comparative health risk assessment presented here, will need to be part of these educational efforts so that they can play a meaningful role in influencing the political dialogue in arriving at solutions for the optimal handling of MSW in NYC and elsewhere.

Acknowledgments

This work was supported in part by funding from the Waste-to-Energy Research and Technology Council to Prof. N. Themelis and the Earth Engineering Center and by funding from NIEHS (P30-ES09089).

References

Berman, S., Ku, J.Y., Rao, S.T., 1999. Spatial and temporal variation in the mixing depth over the Northeastern United States during the summer of 1995. Journal of Applied Meteorology 38, 1661–1673.

Boudet, C., Zmirou, D., Laffond, M., Balducci, F., Benoit-Guyod, J.L., 1999. Health risk assessment of a modern municipal waste incinerator. Risk Analysis 19, 1215–1232.

CA OEHHA, 2005. Toxicity Criteria Database. California Office of Environmental Health Hazard Assessment, Sacramento.

Chang, M.B., Lin, J.J., 2001. Memory effect on the dioxin emissions from municipal waste incinerator in Taiwan. Chemosphere 45, 1151–1157.

Chang, B.C., Wey, M.Y., Huang, S.M., 2001. Heat treatment of incinerator retired sorbents containing heavy metals. Science of the Total Environment 273, 83–92.

Cohen, S., 2006. Understanding Environmental Policy. Columbia University Press, New York.

DeAngelo, M., 2004. Siting of waste-to-energy facilities in New York City using GIS technology. Master of Science Thesis, Department of Earth and Environmental Engineering, Columbia University, New York. Available online at http://www.seas.columbia.edu/earth/wtert/sofos/DeAngelo_Thesis_final.pdf.

Derwent, R.G., Middleton, D.R., Field, R.A., Goldstone, M.E., Lester, J.N., Perry, R., 1995. Analysis and interpretation of air quality data from an urban roadside location in Central London over the period from July 1991 to July 1992. Atmospheric Environment 29, 923–946.

Domingo, J.L., Schuhmacher, M., Llobet, J.M., Muller, L., Rivera, J., 2001. PCDD/F concentrations in soil and vegetation in the vicinity of a municipal waste incinerator after a pronounced decrease in the emissions of PCDD/Fs from the facility. Chemosphere 43, 217–226.

EBI USA, 2005. The Directory of Waste-to-Energy Facilities. Environmental Business International, San Diego.

Eschenroeder, A., von Stackelberg, K., 1999. Health risks of landfilling versus combustion of municipal solid waste: an Illinois comparison. Proceedings of Air and Waste Management Association 99–57, 1–16.

Gonzalez, C.A., Kogevinas, M., Gadea, E., Huici, A., Bosch, A., Bleda, M.J., Papke, O., 2000. Biomonitoring study of people living near and working at a municipal solid-waste incinerator before and after two years of operation. Archives of Environmental Health 55, 259–267.

Kumagai, S., Koda, S., 2005. Polychlorinated dibenzo-p-dioxin and dibenzofuran concentrations in serum samples of workers at an infectious waste incinerator plant in Japan. Journal of Occupational and Environmental Hygiene 2, 120–125.

Landreth, R.E., Rebers, P.A., 1997. Municipal Solid Waste: Problems and Solutions. CRC Press, Boca Raton.

Lee, C.C., Huffmann, G.L., Mao, Y.L., 2000. Regulatory framework for the thermal treatment of various waste streams. Journal of Hazardous Materials 76, 13–22.

Malkin, R., Brandt-Rauf, P., Graziano, J., Parides, M., 1992. Blood lead levels in incinerator workers. Environmental Research 59, 265–270.

Manca, D., Birmingham, B., Raha, D., 1997. Toxicological screening of chemical emissions from municipal solid waste landfills: application of a predictive framework to a state-of-the-art facility. Human and Ecological Risk Assessment 3, 257–286.

Meneses, M., Schuhmacher, M., Domingo, J.L., 2004. Health risk assessment of emissions of dioxins and furans from a municipal waste incinerator: comparison with other emission sources. Environment International 30, 481–489.

NOAA—National Oceanic and Atmospheric Administration, 2005. Climate Report for NY LaGuardia Airport. US Department of Commerce, Washington, DC.
NRC—National Research Council, 1994. Science and Judgment in Risk Assessment. National Academy Press, Washington, DC.
NRC—National Research Council, 2000. Waste Incineration and Public Health. National Academy Press, Washington, DC.
NREL—National Renewable Energy Laboratory, 1996. Wind Energy Resource Atlas of the United States. US Department of Energy, Washington, DC.
NYC DOS—New York City Department of Sanitation, 2005. Final Environmental Impact Statement (FEIS) for the New York City Comprehensive Solid Waste Management Plan. Bureau of Public Information and Community Affairs, New York.
Ontario Ministry of the Environment, Standards Development Branch, Environmental Sciences and Standards Division, 1999. Environmental Risks of Municipal Non-Hazardous Waste Landfilling and Incineration. Queen's Printer for Ontario, Toronto.
Rao, R.K., Chaudhuri, I., Garcia, M., Stormwind, B., Ruffle, B., 2003. Multiple pathway health risk assessment of a municipal waste resource recovery facility in Maryland. EM. August 21–26.
Redfearn, A., Roberts, D., 2002. Health effects and landfill sites. Issues of Environmental Science and Technology 18, 103–140.
Rushton, L., 2003. Health hazards and waste management. British Medical Bulletin 68, 183–197.
Schecter, A.J., Malkin, R., Papke, O., Ball, M., Brandt-Rauf, P.W., 1991. Dioxin levels in blood of municipal incinerator workers. Medical Science Research 19, 331–332.
Shih, T.S., Chen, H.L., Wu, Y.L., Lin, Y.C., Lee, C.C., 2006. Exposure assessment of polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs) in temporary municipal-waste-incinerator maintenance workers before and after annual maintenance. Chemosphere 64, 1444–1449.
Tammemagi, H.Y., 1999. The Waste Crisis: Landfills, Incinerators and the Search for a Sustainable Future. Oxford University Press, New York.
US EPA, 1997. Health Effects Assessment Summary Table. US Environmental Protection Agency, Washington DC.
US EPA, 2000. Performance Data for Large MWC at MACT. US Environmental Protection Agency, Washington DC.
US EPA, 2002. Diesel Exhaust in the United States. US Environmental Protection Agency, Washington DC.
US EPA, 2005. Integrated Risk Information System (IRIS). US Environmental Protection Agency, Washington, DC.
Wark, K., Warner, C.F., Davis, W., 1998. Air Pollution: Its Origin and Control, third ed. Addison Wesley, San Francisco.
Williams, P.T., 2005. Waste Treatment and Disposal, second ed. Wiley, Chichester.

ELSEVIER

Journal of Environmental Management 87 (2008) 80–94

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Governance of water supply systems in the Palestinian Territories: A data envelopment analysis approach to the management of water resources

Kamal Alsharif[a,*], Ehsan H. Feroz[b], Andrew Klemer[c], Raymond Raab[d]

[a]*Department of Biology, 211 Life Sciences Building, University of Minnesota Duluth, Duluth, MN 55812-2496, USA*
[b]*Professor of Accounting, Milgard School of Business, University of Washington, Tacoma, 1900 Commerce Street, Campus Box 358420, Tacoma, WA 98402-3100, USA*
[c]*Professor of Biology, Department of Biology, 211 Life Sciences Building, University of Minnesota Duluth, 10 University Dr., Duluth, MN 55812, USA*
[d]*Professor of Economics, Department of Economics, 175 School of Business and Economics, 10 University Dr., Duluth, MN 55812, USA*

Received 7 November 2005; received in revised form 19 November 2006; accepted 4 January 2007
Available online 26 March 2007

Abstract

This study demonstrates that data envelopment analysis (DEA) can be a useful tool to assess the relative efficiencies of water supply systems and to establish benchmarks with which to measure progress in the management of water resources. Frontier efficiency models measure the efficiency of water use in the Palestinian Territories (West Bank and the Gaza Strip). At the municipality level, sufficient data for the years 1999–2002 were available to estimate efficiency and stability scores. The Gaza Strip efficiency scores were considerably lower than those of the West Bank. Water losses were the major source of the inefficiency as indicated by the large slacks of this input. The relative sizes of the municipalities affect efficiency scores little. Palestinian policy makers should focus on rebuilding the infrastructure of the water networks, beginning with the most DEA inefficient municipalities in order to minimize water losses.

Keywords: Governance of Palestinian Territories; Data envelopment analysis; Water management benchmarks; Water use efficiency

1. Introduction to Palestinian water management

The purpose of this study is to evaluate the efficiencies of the water supply systems by applying data envelopment analysis (DEA) to 33 Palestinian municipalities. It is shown that water shortage crises can be alleviated by improvements in the management of the domestic water sector. The needs for improvements in water network infrastructure are assessed to aid water management decisions. The evaluation of water use efficiency factors allow policy makers to monitor improvement programs and to make better management decisions. Critically, the alleviation of the water supply crisis will improve the prospects for a peaceful settlement of other issues involving the Palestinian Territories. Gestures by the Israeli and Palestinian authorities toward political stability can also stimulate more progressive and enlightened water management policies. Rather than being perceived as an obstacle to progress and permanent peace, solutions to these problems can help motivate broader commonalities of interest.

The rest of this paper is organized as follows. Section 2 describes the physical geography and the water management by the Palestinian Authority. Section 3 describes the DEA methodology. Section 4 describes the data and Section 5 reports the empirical results. Finally, Section 6 provides conclusions and policy recommendations.

2. Physical geography and water management

The Palestinians took control of their major cities and small towns in the beginning of the 1990s; they did not

*Corresponding author. Tel.: +1 218 726 6262; fax: +1 218 726 8142.
E-mail addresses: alsh0018@umn.edu (K. Alsharif), ehf2@u.washington.edu (E.H. Feroz), aklemer@d.umn.edu (A. Klemer), rraab1@d.umn.edu (R. Raab).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.008

control the aquifers under the West Bank, unlike Gaza, where they did control the aquifer. Salt water from the Mediterranean Sea was polluting the wells in Gaza (Sabbah and Isaac, 1995). In addition, the aquifer in Gaza was polluted from pesticides, fertilizers, and raw sewage (Green, 1993). It is evident that more water is being used than can be replenished. Many studies have found water shortages in the area and have focused on the water conflicts and ways to split the water among the Middle Eastern countries (Dolatyar and Gray 2000, Rowland, 2000). With the lack of waste treatment plants and the lack of storage of annual rainfall (besides the aquifers), a huge water crisis looms.

The Palestinians need to think in a "sustainable way," and they need to protect their water resources (Sabbah and Isaac, 1995). The population is increasing at an alarming rate, while the aquifers are being depleted and polluted. The Palestinian sector faces limited access to water and a seriously deteriorating water distribution problem (e.g., losses of 40–60% in most municipalities estimated by the World Bank, 1993). The Palestinians and the Israelis simply do not have enough water for their current populations and agricultural production (Rouyer, 2000). Current water management practices need to be revised. Too little is known or pursued with regard to the potential for improvements in the efficiency of domestic water use and management in this region. In other countries in the region, such as Saudi Arabia, water management conservation changes need to be made immediately, since groundwater resources are being depleted (Al-Ibrahim, 1990). In general, Wood (2003) also suggests the focus should be on efficiency of water use.

The Palestinian Territories consist of the West Bank and Gaza Strip. The West Bank has an area of 5545 km^2, and Gaza Strip has an area of 365 km^2, for a total area equal to 5910 km^2, thus slightly smaller than the State of Delaware (Elmusa, 1997). West of the Gaza Strip is the Mediterranean Sea, east is Israel, and to the south is Egypt. On the other hand, the West Bank is surrounded by Israel to the north, south, and west, and by Jordan to the east. The Palestinian areas account for about 22% of historical Palestine. Many Israeli settlements, which were built after the 1967 war, exist among the Palestinian cities, villages, and refugee camps.

The West Bank is mostly mountainous in the middle section with dry valley country in the Jericho area. More than three million Palestinians live in cities, villages, and refugee camps in the Palestinian Territories, and over 200,000 Israeli settlers live in settlements. About 2 million people live in the West Bank and 1.2 million people live in the Gaza Strip. The Israeli settlers had considered the West Bank as part of the greater Israel. Presently these settlers have full Israeli citizens' rights, and they enjoy the protection and support of the Israeli government.

The area has a short rainy winter season and is considered semiarid. The annual range of rainfall in the West Bank is 400–600 mm, and the annual rainfall in Gaza ranges from 350 to 400 mm in the north and 150–200 in the south (Elmusa, 1997). The major water sources are the coastal aquifer under Gaza and three mountain aquifers under the West Bank (the Northern Aquifer, Eastern Aquifer, and Western Aquifer). They have a safe reliable yield of 140, 125, and 350 mcm, respectively. The potential of ground water from Gaza is 65 mcm/year, and currently more than 100 mcm of fresh water is being pumped, which has resulted in pollution of the aquifer from seawater intrusion. Rain infiltrates directly through the fractured rocks and the porous soil (Haddad, 1998). The Palestinians have been denied access to the surface water from the Jordan River.

According to the Palestinian Water Authority (2002), there are multiple agencies that manage the water supplies in the West Bank on micro- and macrolevels. The main agency in charge of water management is the West Bank Water Department. This department was formerly under the control of the Israeli Civil Administration; that is, under the control of the military governor. The origin of the West Bank Water Department was the Jordanian Natural Resources Authority, which controlled the West Bank from 1948 to 1967. After signing the Oslo II agreement in 1995, the Palestinian Water Authority, which is part of the Palestinian National Authority, was given the charge of administrating, operating, and maintaining the water department. In addition, other non-profit municipal and local councils manage water. In the main cities, there are water departments that are in charge of water supplies. The main municipalities are Nablus, Qalqilia, Salfit, Jericho, Jenin, Tulkarm, and Hebron. There also exist independent institutions that are owned by the municipalities, and they include Jerusalem Water Undertaking for Ramallah and Al-Bireh Districts; Water and Sanitation Authority for Bethlehem, Beit Jala, and Beit Sahur; and finally, Joint Services Water Council for West Jenin Villages. In addition, there are some other small municipalities and councils that have their own wells and springs. In the Gaza Strip, the Palestinian Water Authority is the main agency in charge of water. Many municipalities administer the distribution of water to the public. The municipalities and village councils have their own wells, and they distribute water to the public. United Nations Relief and Works Agency (UNRWA) manages water from wells for the refugees in the camps (World Bank, 1993).

There are variations in water consumption in each area especially between the West Bank and the Gaza Strip. The agencies responsible for the management of water resources are also different. The Palestinian Water Authority lacks the central control of water distribution in the municipalities. Because of the current political instability the PWA only provides general guidelines and works with the different local municipalities.

3. DEA methodology

As water resources decline and the population increases, managers operate under tighter budget constraints and must use these resources more efficiently. Performance or efficiency evaluation can be done using DEA. As a frontier estimation technique developed by Charnes et al. (1981), performance measures are simple output-to-input ratios that are used as efficiency measures (Stolp, 1990). DEA can use multiple inputs and outputs to determine summary efficiency scores that can be used to evaluate the relative efficiency of the Palestinian municipalities. These municipalities have responsibility for transferring the inputs into outputs in order to become more efficient (Cooper et al., 2002). Unlike regression analysis, DEA does not assign a uniform weight for each input and output (Cooper et al., 2002). DEA does assign separate weights for inputs and outputs for each individual municipality such that the efficiency score of that municipality is maximized. The Charnes Cooper Rhodes (CCR) model is the basic DEA model which assumes constant returns to scale (CRS). It is also input oriented (i.e., inefficiency measured as surplus inputs) or output oriented (i.e., inefficiency measured as deficient outputs) (Charnes et al., 1994).

The BCC model (Banker et al., 1984) is also input oriented or output oriented, but it utilizes a variable returns to scale frontier (VRS) and additionally can indicate returns to scale (RTS). We selected the input orientation so that inefficiencies are measured as excess inputs. The third model, the additive model (Charnes et al., 1985) also utilizes the VRS frontier, but it is non-oriented since it has both input and output slacks (i.e., minimizes the inputs and maximizes the outputs simultaneously). The slacks are defined as the excess in the inputs and the shortage in the outputs. These twoVRS models measure only technical efficiency, not scale efficiency. DEA measures relative efficiency of the different municipalities under consideration, and tests the decision cases with slightly different criteria (Cook et al., 1996). The additive model allows for the improvements associated with the input reduction and the output augmentation. A decrease in an input and/or an increase in an output are considered an improvement, and they can be calculated by identifying the radial distance toward the calculated frontier (Charnes et al., 1994).

The BCC and the additive model have the same frontiers, but provide different pieces of information. The BCC provides input-orientated efficiency scores, while the additive model provides efficiency rankings of the municipalities. The BCC input-oriented model can be expressed in the following equations; where x_{ij} and y_{rj} are inputs and outputs, and S_i^- and S_r^+ are the corresponding slacks:

$$
\begin{aligned}
\min \quad & z_o = \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right) \\
\text{s.t.} \quad & \\
& \theta x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_i + s_i^- \quad i = 1,\ldots,m, \\
& y_{ro} = \sum_{j=1}^{n} y_{ij}\lambda_i - s_r^+ \quad r = 1,\ldots,s, \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& \theta \leqslant \lambda_j, s_i^-, s_r^+ \quad \forall i,j,r.
\end{aligned} \tag{1}
$$

Here $\varepsilon > 0$ is a non-Archimedean element smaller than any real number. Therefore, it cannot be represented as a real number. See Ali and Seiford (1993). The software utilized did not require the Archimedean infinitesimal to be specified. The DMU_o is considered efficient if and only the two conditions are met (Cooper et al., 2002)

(i) $\theta^* = 1$ and
(ii) all slacks are zero.

RTS can be identified in the following dual form of model (1), above.

$$
\begin{aligned}
\max \quad & \omega = \sum_{r=1}^{s} w_r y_{ro} - \mu_o \\
\text{s.t.} \quad & \\
& \sum_{r=1}^{s} w_r y_{rj} - \sum_{i=1}^{m} \mu_i x_{ij} - \mu_o \leqslant 0, \\
& \sum_{i=1}^{m} \mu_i x_{io} = 1, \\
& w_r, \mu_i \geqslant \varepsilon, \quad r = 1,\ldots,s, \; i = 1,\ldots,m,
\end{aligned} \tag{2}
$$

where RTS is increasing, constant or decreasing according to whether the coefficient $\mu^* \gtreqless 0$. (A good discussion of RTS can be found in Cooper et al., 2006.)

The additive model primal below dichotomizes the municipalities as efficient (i.e., on the frontier) or inefficient (i.e., below the frontier):

$$
\begin{aligned}
\min \quad & (-e^T s^+ - e^T s^-) \\
\text{s.t.} \quad & \\
& Y\lambda - s^+ = Y_j, \\
& X\lambda + s^- = X_j, \\
& e^T\lambda = 1, \\
& \lambda, s^+, s^- \geqslant 0,
\end{aligned} \tag{3}
$$

where Y and X are the matrices of the outputs and inputs, and s^+ and s^- are the shortfall in production and excess consumption slacks. For this model efficiency is achieved if and only if all slacks are zero. Efficiency scores also supplied in Table 7 measure the amount of inefficiency in each influenced output in the form of non-zero scale. The e^T is the summation vector.

In order to test for stability or sensitivity of the solutions with respect to data changes we modify (3) and then use the results for ranking the DMUs with respect to this robustness. Two additional linear programs must be computed in order to rank the dichotomized municipalities, one for efficient and one for inefficient municipalities. For these purposes two additional linear programs are computed, following Charnes et al. (1992, 1996).

For an efficient municipality the stability index defines the largest "radii" in which all simultaneous perturbations (increase in inputs and decrease in output components) that will not cause a change to the efficiency status from technically efficient to technically inefficient. The larger the stability index, the more robustly efficient the municipality is said to be, or, the more the efficiency frontier has been "pushed out" by this efficient municipality. The following linear program can be used to measure the stability index in Eqs. (4) and (5), or θ, for the technically "efficient" municipalities:

$$\begin{aligned} \min \quad & \theta \\ \text{s.t.} \quad & \\ & Y^{(E)}\lambda - s^{+} - \theta d_0 = Y_j, \\ & X^{(E)}\lambda + s^{-}\theta d_I = X_j, \\ & e^{T}\lambda = 1, \\ & \lambda, s^{+}, s^{-}, \theta \geqslant 0, \end{aligned} \tag{4}$$

where $Y^{(E)}$ and $X^{(E)}$ are the matrices of inputs and outputs, respectively, with the test municipality j omitted; d_0 and d_I are vectors of weights set equal to unity such $(d_0^{\mathrm{T}} = (1,1,\ldots,1)$ and $d_I^{\mathrm{T}} = (1,1,\ldots,1))$. For the purpose of illustration see Cooper et al. (2004, pp. 78–80) for discussion. The stability index value, θ^{*} is referenced as a solution value for the individual municipality being measured as in Eqs. (4) and (5).

For the inefficient municipalities the stability index can be determined by solving the following linear program:

$$\begin{aligned} \max \quad & \theta \\ \text{s.t.} \quad & \\ & Y\lambda - s^{+} - \theta d_0 = Y_j, \\ & X\lambda + s^{-}\theta d_I - X_j, \\ & e^{T}\lambda = 1, \\ & \lambda, s^{+}, s^{-}, \theta \geqslant 0, \end{aligned} \tag{5}$$

where all notations are defined above. For inefficient municipality$_j$, the stability index (θ), defines the necessary minimum favorable perturbations (decreases in inputs and increases in outputs) that must be undertaken to cause the municipality to become virtually efficient. The larger the stability index is, the further the municipality lies from the frontier and the more robustly inefficient it is. By negating the stability index for inefficient municipalities, they can be rank ordered from highest to lowest stability index values. See Charnes et al. (1989) for specifying the principle on which DEA rankings can be based and see Cooper et al. (2001) for supporting this method of ranking. Raab and Lichty (2002) provide an explanation and a recent application of this approach. See Table 7 for the efficiency scores and stability index results.

Extant DEA applications have addressed water resource management problems in other countries such as Ghana, Japan, Mexico, Texas, and the UK (Akosa et al., 1995; Aida et al., 1998; Anwandter, 2000; Sanders, 1999; Thanassoulis, 1999, 2000, 2002). No DEA water applications have been applied to the Palestinian Territories. As Akosa et al. (1995) indicated, the DEA method can combine the input and output factors to define a single efficiency score. This score was used to evaluate the water supplies and wastewater projects in Ghana. Resource poor countries need to manage their domestic water supply efficiently. Akosa et al. (1995) recommended DEA as a method to be used in poor countries as a tool to compare the efficiency of different types of water technologies.

Aida et al. (1998) used the DEA additive model in the form of RAM (range adjusted measure) model to evaluate the water systems in 108 cities in the Kanagawa Prefecture and Kanto Region in Japan. Their inputs included the number of employees, operating expenses before depreciation, net plant and equipment, population, and the length of pipes. Their outputs included water billed and operating revenues. They showed that DEA could help in conducting performance and efficiency audits, and demonstrated that it can be useful in evaluating the efficiency of the water sectors in industrialized countries.

Some management scientists favor the DEA approach over regression analysis (Cubbin and Tzanidakis, 1998; Feroz et al., 2001). DEA concentrates on the individual observation as represented by the optimization for each observation instead of the central tendency estimation of parameters that are associated with "single optimization statistical approaches" (Charnes et al., 1994). DEA estimates relative efficiency by employing observed data to establish benchmarks on the frontier using a criterion of Pareto optimality. This means that efficiency is attained if and only if any increase in inputs or decrease in outputs is not possible without worsening the levels of other inputs or outputs (Cooper et al., 2002; Charnes et al., 1994, 1985; Akosa et al., 1995; Chang et al., 2004).

An additional strength of this approach is its use of multiple inputs and outputs (Stolp, 1990; Cubbin and Tzanidakis, 1998). Traditional regression approaches do specify a unique functional relationship for all of the municipalities examined while DEA allows each municipality to formulate its own optimum production relationship (Stolp, 1990; Cubbin and Tzanidakis, 1998; Bowlin, 2004). Inefficiency occurs only in terms of excess inputs or a shortfall of outputs (Cooper et al., 2002; Bowlin, 2004). It also identifies the benchmark or "best practice" entities of the efficiency set used to evaluate each decision making unit (DMU) or entity and the benchmarks are useful as means of setting standards (Cooper et al., 2002). One disadvantage of DEA is that it is subject to significant error when outliers are present (Charnes et al., 1994). Because of these differences regression analysis and DEA may give somewhat different results (Cubbin and Tzanidakis, 1998).

4. Data

The data of this simple model is provided by the Palestinian Water Authority as follows:

Input (I):	Water losses (I)	Water and energy (I)	Maintenance and others (I)	Salary of Workers (I)
Output (O):	Total revenue (O)			

There are four inputs to our DEA model. The (I) stands for Input and the (O) stands for Output. Total revenue represents the single output. The cost of water bought and energy costs were combined because they both measure the municipality's expenses by buying its water from Mekoroth (the Israeli water company), private wells, and shared wells. Energy cost is also added since it reflects the energy expended to produce the water. Maintenance and other expenses include the cost of maintaining the operations, materials, and other miscellaneous expenses. Salary represents the wages for both the managerial and non-managerial staff. The water losses are defined as the differences between the water supplied to the customers and the actual amount consumed. This is usually measured in cubic meters. The water loss values in cubic meters were converted to New Israeli Shekels (NIS). These inputs were chosen since they were available for most of the municipalities and they act as indicators of the operational performance of the municipalities. For example, minimizing water losses and related expenses help insure less scarce resources are being wasted relative to the maximum value of output. The model's sole output of revenue equals the value of water according to water billing, not the revenue collected combined with other revenues.[1]

The raw data used to generate the efficiency scores are given in Tables 1–4. These cover the years from 1999 to 2002, and they were collected in formats used by the Palestinian Water Authority and other agencies. The water losses were converted to NIS instead of using the physical measurement of cubic meters. In this cross-sectional analysis, each year is considered as a separate cross section. The data were for 31 municipalities that existed in 1999, 30 existed in 2000, 27 existed in 2001, and 33 existed in 2002. Identical data sets were not available for all the years. However, the input and output data are consistent for each year. The Gaza Strip municipalities are the same throughout the period of analysis, 1999–2002, but data for some municipalities in the West Bank were missing and unobtainable. The size of the municipalities varied considerably since both very small and very large municipalities were included. For example, the water revenues range from hundreds of thousands to millions of NIS. There was an increase in the water revenue for each of the different municipalities from 1999 to 2002.

4.1. Data analysis using the BCC model

Table 5 shows the BCC Eq. (1) scores, population, water losses, and scale efficiency coefficients for the different municipalities from 1999 to 2002. The BCC model measures technical efficiency and scale efficiency. The RTS analysis is based on the sign of μ^* in Eq. (2), and represents the intercept on the frontier. The negative sign for μ^* indicates increasing returns to scale (IRS) or the hyperplane intercept the output axes below the origin. The positive sign for μ^* indicates decreasing returns to scale (DRS) or the hyperplane intercepts the output axes above the origin. Although both small and large communities were just as likely to be on the frontier, the larger municipalities overwhelmingly exhibit IRS and the smaller ones exhibit DRS.

Table 6 shows the efficiency scores and the slacks input for the different municipalities from 1999 to 2002. In this cross-sectional analysis, each year is independent. The most frequently appearing slacks were for water losses in almost every single year. Gaza Strip municipalities' slacks for water losses were more dominant than in the West Bank.

4.2. Data analysis using the additive model

The same data in Tables 1–4 were analyzed using the additive model. Table 7 shows the efficiency scores and the radius of stability values (θ^*'s in Eqs. (4) and (5)) for the West Bank and Gaza Strip for 1999 to 2002. For efficient municipalities on the frontier the stability index defines the largest symmetric increases in inputs and decreases in outputs that would just maintain the status of technical efficiency (i.e., by what distance the frontier was pushed out by that particular municipality) as mentioned previously in the DEA methodology, Section 3. For inefficient municipalities below the frontier, the stability index defines decrease in inputs or increase in outputs that would just maintain the status of technical inefficiency (i.e., by what distance the data can be altered in order for it to reach the frontier). Once the stability index is known for each municipality, they can be ranked from most robustly technically efficient to most robustly inefficient. To do so, the stability indexes for inefficient municipalities are first determined. These firms can then be rank ordered from highest to lowest based on their stability index values. A general mathematical and graphical representation of the stability index for both efficient and inefficient municipalities can be found in Raab and Lichty (2002). Cooper et al. (2001) shows that this sensitivity model legitimately can be used for efficiency rankings.

4.3. Discussion of the empirical results

The differences in results may reflect changes in the political atmosphere. In 2000, the Intifada (the Palestinian uprising) started in the West Bank and Gaza Strip, and this led to civil unrest in the Occupied Territories. People who used to work in Israel lost their jobs, and the whole Palestinian economy was placed in a vulnerable position. The unemployment rate in the West Bank has increased during the study period from 9.5% to 28.2%, from 1999 to 2002, respectively. On other hand, the unemployment rate in the Gaza Strip rose from 16.9% to 38% from 1999 to

[1]Other revenues include the income from replacement of the water meters, income from new services, income from private contributions, revenue from stakeholder contributions, and income from water sold in tanks and bulk. We preferred a fairly narrowly defined output measure to determine the performance of the DMUs.

Table 1
Water Data for the West Bank and Gaza Strip in 1999

Municipality	(I) Water losses[a] (NIS)	(I) Water and energy[b] (NIS)	(I) Maint and others[c] (NIS)	(I) Salary of workers[d] (NIS)	(O) Total revenue[e] (NIS)
Jenin	1,669,909	1,324,283	439,985	961,744	4,662,473
Qabatiya	168,250	673,000	43,000	168,000	893,300
Tubas	191,763	268,195	39,856	158,148	577,214
Tulkarm	1,839,831	1,650,000	742,400	1,070,708	4,368,246
Anabta	461,503	140,000	103,440	203,347	663,097
Qalqilya	787,866	650,000	187,000	490,000	2,301,688
Azzun	43,187	254,259	206,060	51,831	581,694
Nablus	1,899,031	12,188,436	1,778,201	4,587,387	25,044,490
Salfit	373,106	720,818	60,718	84,000	931,896
Ramallah	6,434,216	25,987,067	14,287,353	780,339	43,708,100
Beitunya	470,832	1,530,169	190,200	267,964	2,153,381
W.Bani Zaid	55,464	318,862	35,674	36,132	519,216
Jericho	4,632,518	283,878	1,749,404	449,304	1,318,398
Bethlehem	3,336,909	6,811,077	561,550	1,011,432	9,987,543
Dura	162,159	535,641	79,223	263,939	549,392
Gaza	8,421,715	4,124,034	969,952	2,277,949	14,003,170
Rafah	1,132,013	829,672	241,946	814,836	4,205,446
Beit Hanun	536,788	238,507	66,964	180,804	1,040,253
Beit Lahia	616,681	457,485	171,225	360,623	1,779,823
Jabalia	2,029,891	1,344,963	362,423	811,279	3,715,985
El Bureij	342,935	975,792	96,938	39,925	1,072,884
Nuseirat	685,048	2,282,463	44,763	261,102	2,014,662
Zuwadia	38,161	788,079	9061	53,790	948,952
Maghazi	123,309	475,005	15,645	63,625	687,013
Dier Al Balah	1,340,802	1,314,485	131,910	371,315	2,339,463
Khan Yunis	4,593,862	3,046,442	403,809	919,812	4,348,744
Qarara	403,144	144,528	36,736	113,774	484,008
Abassan Jadida	53,046	266,519	5260	24,278	281,348
Abassan Kubra	190,351	776,427	15,531	69,883	815,887
Khazaa	56,723	402,014	12,349	57,904	497,619
Bani Suhalia	338,243	1,044,711	16,828	117,821	1,131,496
Mean	1,400,944	2,317,639	745,336	552,355	4,439,577
Total	43,429,257	71,846,811	23,105,404	17,122,995	137,626,881

NIS is defined as New Israeli Shekel (the Israeli currency).
[a]Water loss values are defined as the differences between the water supplied to the customers and the actual amount consumed.
[b]Cost of water bought and energy used were combined because they measure how much the municipality spends on buying its water.
[c]Maintenance and other expenses include the cost of maintaining the operations, materials, and other miscellaneous expenses.
[d]Salary represents the wages for both the managerial and non-managerial staff.
[e]The total revenue equals the value of water according to the water billing (see footnote 1).

2002 (Palestinian Central Bureau of Statistics, 2004). People could not travel between the cities, and tourism was devastated. This decline was accelerated by reductions in investment in the infrastructure of the Palestinian cities, including the water sector. In 18 of 33 municipalities the maintenance and other expenses declined in 2002 compared to 1999, and these declines were mostly in the larger municipalities. The percentage reduction ranges from 7% to 98%. The percent of people paying their billings also declined. This sector, already in need of repair and rebuilding, worsened. The impact of civil unrest is more obvious in the Gaza Strip where the economy is weaker, the unemployment higher, and the population more crowded than in the West Bank.

The Palestinian sector faces a deteriorating water distribution problem in the West Bank and the Gaza Strip with losses of 40%–60% in most municipalities (World Bank, 1993). Beginning the 1990s, the World Bank (1993) estimates Palestinians need at about \$500–600 million dollars to rehabilitate and expand the water and sewerage systems. Urgent needs include improvement in the quality of water supply especially in Gaza, reduction in water pollution and rehabilitation and expansion in water supply in the West Bank.

Excessive water losses or "Unaccounted for Water" (UFW) is the greatest contributor to inefficiency. Many municipalities exhibited excesses in this input that appear to be the result of old and deteriorating water networks according to the water losses in Table 5 and the slacks in Table 6. Actual water losses values in cubic meter per year are given in Table 5. However, these values are not parentages and some municipalities with high water losses are still efficient. For example, Gaza water losses in 1999 were 8,149,956 cubic meters per year, but percentage wise

Table 2
Water Data for the West Bank and Gaza Strip in 2000

Municipality	(I) Water losses[a] (NIS)	(I) Water and energy[b] (NIS)	(I) Maint and others[c] (NIS)	(I) Salary of workers[d] (NIS)	(O) Total revenue[e] (NIS)
Jenin	1,011,838	771,697	862,572	1,313,224	4,337,092
Qabatiya	292,679	830,446	5840	121,620	635,052
Tubas	169,752	354,298	69,193	158,381	521,232
Tulkarm	1,808,510	1,877,609	381,492	1,240,099	5,877,378
Anabta	502,178	133,440	99,763	164,876	696,197
Qalqilya	595,043	759,850	221,148	385,740	2,239,875
Azzun	359,779	321,123	29,092	43,371	617,716
Nablus	8,678,546	10,420,874	608,795	5,509,969	20,557,691
Salfit	332,326	646,645	25,686	154,152	970,356
Ramallah	6,797,131	27,300,655	9,743,916	12,595,321	48,030,735
Beitunya	551,132	1,871,108	229,521	229,700	2,526,678
Jericho	1,052,106	140,940	198,973	366,814	1,911,743
Bethlehem	4,432,141	10,341,555	855,595	1,399,746	12,988,763
Halhul	160,846	696,767	40,774	87,922	815,969
Gaza	8,283,180	3,796,788	1,082,097	2,292,151	13,592,779
Rafah	1,274,190	789,828	209,448	754,351	4,491,907
Beit Hanun	715,567	312,959	102,186	178,382	1,167,853
Beit Lahia	1,235,467	440,772	171,098	395,382	1,932,572
Jabalia	1,354,623	1,380,725	1,427,048	646,528	4,112,853
El Bureij	338,066	1,518,545	13,239	51,585	1,165,661
Nuseirat	758,056	2,451,672	38,456	233,470	2,126,176
Zuwadia	80,882	847,446	21,702	66,277	921,508
Maghazi	474,525	649,858	31,383	74,118	773,102
Dier Al Balah	1,157,129	1,398,308	181,630	392,402	2,474,602
Khan Yunis	6,056,215	1,308,721	346,166	1,337,353	4,505,283
Qarara	352,839	133,542	41,967	123,301	530,981
Abassan Jadida	83,205	326,584	6970	52,849	297,112
Abassan Kubra	45,145	654,427	13,997	72,477	939,631
Khazaa	40,971	504,029	18,211	71,980	560,324
Bani Suhalia	434,119	1,241,459	43,020	172,708	1,255,352
Mean	1,647,606	2,474,089	570,699	1,022,875	4,785,806
Total	49,428,185	74,222,670	17,120,978	30,686,249	143,574,173

NIS is defined as New Israeli Shekel (the Israeli currency).
[a]Water loss values are defined as the differences between the water supplied to the customers and the actual amount consumed.
[b]Cost of water bought and energy used were combined because they measure how much the municipality spends on buying its water.
[c]Maintenance and other expenses include the cost of maintaining the operations, materials, and other miscellaneous expenses.
[d]Salary represents the wages for both the managerial and non-managerial staff.
[e]The total revenue equals the value of water according to the water billing (see footnote 1).

they were 38% of the total water supplied. In Khan Yunis the actual water losses were 2,753,448 cubic meters per year, but that was 51% of the total water supplies. Water losses in this region are very high compared to those typical of developed countries and less or about the same as those in less developed countries. For example, a survey of 47 municipalities in California revealed that the average water losses were about 10%, and the range was 5% to less than 30% of the total water supplied (Department of Water Resources, 2004). Southern California has a Mediterranean-like climate similar to that of the Middle East. According to the World Bank Group (2004), the types of water losses are physical and administrative losses due to water theft, for example. In Ramallah and the surrounding communities the causes of water losses or unaccounted for water from 1991 to 1994 included damage to main pipelines (0.38%), damage to distribution networks (50.74%), inaccuracy of meters (42.90%), water losses due to theft (5.80%), and washing of new lines and reservoirs (0.18%). The main losses occur from losses between the main supply lines and the customers, called network losses (Jerusalem Water Undertaking, 2003). In the West Bank, deteriorating water networks mainly caused the water losses. Recent reductions in these excess inputs in this region are considered an improvement. In the Gaza Strip, water losses decreased in 2002 compared to 1995 when they were about 45% (World Bank Group, 2004). In 2002, the average was around 33% and this improvement coincided with the establishment of the Palestinian Water Authority in 1994 when improvements were made to the water networks.

The size of the municipality is not correlated with efficiency in our study. In some cases, the smaller municipalities were efficient though they were a fraction of the size of the larger municipalities. These findings are similar to those reported by Aida et al. (1998) when they

Table 3
Water Data for the West Bank and Gaza Strip in 2001

Municipality	(I) Water losses[a] (NIS)	(I) Water and energy[b] (NIS)	(I) Maint and others[c] (NIS)	(I) Salary of workers[d] (NIS)	(O) Total revenue[e] (NIS)
Tubas	106,307	665,502	55,705	109,903	722,432
Tulkarm	1,599,921	1,539,357	290,516	1,299,642	4,098,452
Anabta	298,993	142,990	217,702	214,426	708,679
Qalqilya	460,931	742,273	146,871	378,278	2,142,888
Azzun	181,153	304,660	166,253	43,711	606,589
Nablus	11,677,637	12,224,203	2,799,098	5,887,818	25,924,680
Salfit	370,837	759,982	42,769	152,328	985,746
Ramallah	7,497,737	28,513,139	5,486,154	14,229,940	47,386,843
Beir Zeit	69,559	529,113	16,832	57,963	828,969
Beitunya	705,127	2,017,858	94,969	284,900	2,440,481
Bethlehem	7,131,992	11,147,421	1,331,511	2,593,155	5,652,764
Gaza	6,660,262	1,819,923	635,070	2,364,173	10,047,227
Rafah	3,339,482	1,274,412	293,459	1,463,092	2,149,210
Beit Hanun	866,169	307,406	66,474	201,179	574,368
Beit Lahia	1,390,252	564,711	160,177	393,045	870,965
Jabalia	1,577,700	1,143,096	1,420,853	762,143	1,746,136
El Bureij	601,324	1,813,828	13,702	41,274	638,482
Nuseirat	1,355,565	2,306,606	34,620	181,734	1,093,708
Zuwadia	239,589	527,388	24,223	67,799	350,408
Maghazi	491,605	891,782	23,283	79,278	484,238
Dier Al Balah	668,459	1,086,574	126,706	412,127	844,910
Khan Yunis	5,893,708	1,050,987	747,988	1,758,275	2,071,405
Qarara	395,195	146,784	43,500	178,569	312,894
Abassan Jadida	69,945	315,763	6033	50,149	84,520
Abassan Kubra	197,593	915,877	11,378	92,507	353,174
Khazaa	71,402	656,359	12,284	70,075	254,854
Bani Suhalia	254,604	1,043,120	28,342	155,801	398,697
Mean	2,006,409	2,757,449	529,499	1,241,603	4,213,841
Total	54,173,047	74,451,114	14,296,472	33,523,284	113,773,719

NIS is defined as New Israeli Shekel (the Israeli currency).
[a]Water loss values are defined as the differences between the water supplied to the customers and the actual amount consumed.
[b]Cost of water bought and energy used were combined because they measure how much the municipality spends on buying its water.
[c]Maintenance and other expenses include the cost of maintaining the operations, materials, and other miscellaneous expenses.
[d]Salary represents the wages for both the managerial and non-managerial staff.
[e]The total revenue equals the value of water according to the water billing (see footnote 1).

examined the large water suppliers versus the smaller ones in Japan. However, they used RAM in their DEA estimates, which is a modified type of the additive model to compensate for the differences in sizes, and they also used different sets of inputs and outputs. However, we used the basic BCC model, and additive model for our case study. The BCC model produced the RTS values (μ^*'s) as shown in Table 5. Municipalities with larger populations showed IRS and municipalities with smaller populations showed DRS. The decreasing RTS for the smaller cities means that they had proportionately less revenue, and increasing RTS means increasing the inputs by 1% result in revenue increase of greater than 1%. There was low negative correlation between the BBC efficiency scores in Table 5 and the percent water losses for all the different years.

In summary, water losses were the main cause of inefficiency, and these losses must be remedied to allow for the Palestinian water sector to meet demands at present and in the future. Water use inefficiency was more a result of physical factors and less a result of inadequate management. According to the slacks generated by the BCC model, most of the inefficiencies were caused by excess losses of water. The lack of political stability may have contributed to the destabilization of the water sector through lack of revenue collections and resulting deficits in the water budget (especially prevalent in the Gaza Strip) and through the physical destruction of the water network. The latter will likely contribute to continuing water losses and damage the future infrastructure of the Palestinian cities. In addition, since personal incomes are jeopardized, investment in the water sector will be minimal. All of these factors will tend to either maintain, or, more likely, lead to an increase in water losses.

5. Conclusions and policy recommendations

This study found that most of the inefficiencies in the water supply systems of the Palestinian Territories were caused by excess losses of water. Therefore, the main focus

Table 4
Water Data for the West Bank and Gaza Strip in 2002

Municipality	(I) Water losses[a] (NIS)	(I) Water and energy[b] (NIS)	(I) Maint and others[c] (NIS)	(I) Salary of workers[d] (NIS)	(O) Total revenue[e] (NIS)
Jenin	515,862	1,757,182	299,506	1,381,258	4,908,642
Qabatiya	365,540	543,665	30,000	172,000	953,004
Tubas	270,948	592,952	45,704	119,505	756,277
Tulkarm	1,785,989	1,161,039	524,596	826,362	5,532,066
Anabta	730,023	198,720	132,134	176,803	650,914
Qalqilya	738,325	816,637	421,473	403,694	2,460,918
Azzun	222,221	407,467	129,992	35,883	727,658
Nablus	14,025,111	12,257,088	560,285	5,233,742	29,298,834
Salfit	422,492	792,105	42,769	152,328	1,015,133
Ramallah	7,761,274	29,231,136	4,421,844	16,359,757	41,691,400
Beir Zeit	99,180	714,000	19,297	72,248	1,146,133
Beitunya	653,838	1,949,020	73,254	290,892	2,349,424
W.Bani Zaid	51,019	336,744	72,432	90,691	776,559
Jericho	994,244	230,692	21,095	581,937	2,525,161
Bethlehem	4,843,865	9,527,241	453,855	1,075,787	10,435,730
Halhul	140,410	651,215	48,969	278,584	921,155
Dura	147,041	629,399	48,295	182,097	807,976
Gaza	6,668,998	3,380,537	899,203	2,227,775	11,670,305
Rafah	2,916,694	1,219,631	303,307	1,635,311	1,788,287
Beit Hanun	822,510	419,581	74,388	240,289	703,348
Beit Lahia	1,594,892	608,749	190,737	403,990	898,912
Jabalia	1,904,526	1,227,539	276,864	830,096	1,897,367
El Bureij	464,934	1,187,631	21,641	44,404	504,629
Nuseirat	1,107,971	2,757,670	31,757	197,392	1,100,329
Zuwadia	220,323	510,896	27,117	70,949	256,484
Maghazi	617,394	996,144	23,070	88,184	403,079
Dier Al Balah	708,110	1,124,334	144,407	400,572	558,430
Khan Yunis	6,562,139	1,022,023	357,710	1,863,030	1,842,084
Qarara	400,106	95,694	42,811	204,910	247,803
Abassan Jadida	117,273	470,537	8986	55,827	93,275
Abassan Kubra	209,143	1,165,976	12,816	108,819	251,604
Khazaa	54,965	740,903	8461	38,477	163,744
Bani Suhalia	401,088	1,484,591	31,741	197,853	247,449
Mean	1,773,892	2,430,568	296,985	1,092,165	3,926,791
Total	58,538,450	80,208,738	9,800,516	36,041,446	129,584,113

NIS is defined as New Israeli Shekel (the Israeli currency).
[a]Water loss values are defined as the differences between the water supplied to the customers and the actual amount consumed.
[b]Cost of water bought and energy used were combined because they measure how much the municipality spends on buying its water.
[c]Maintenance and other expenses include the cost of maintaining the operations, materials, and other miscellaneous expenses.
[d]Salary represents the wages for both the managerial and non-managerial staff.
[e]The total revenue equals the value of water according to the water billing (see footnote 1).

of Palestinian water management should be on repairing the infrastructure so that the water supply systems function properly. Water losses need to be reduced to about 10%, which is common in the developed countries, within the next five years. The water consumption of the Palestinian people should be raised to meet or exceed the minimum water requirement of 100 L per person per day as recommended by the World Health Organization. Palestinian water management authorities should attempt to establish a better mechanism to monitor water losses and to define specific water losses according to the cause. In some cases, the cause could be the deteriorating networks, and in others, it might be inaccuracy in water meter readings. A water data bank located in a centralized location for all of the districts in the West Bank and Gaza Strip would also lead to appropriate water management strategies.

Deficiencies in the data available for this study may have erroneously influenced results, especially water losses. For example, Jericho was only 94.3% efficient, and had 80% water losses in 1999 when you convert 1,587,617 cubic meters to percentages. These water losses disappear by 2000. The low water loss values for 2000 and 2002, however, put the Jericho municipality on the frontier with 100% efficiency and zero slacks (see Table 5). Note that no water data for Jericho existed for 2001. Water loss data may be too high or too low due to metering problems. On one hand the meter readers may be charging people for air going through the pipes when water is not continuously flowing. On the other hand, they may be accurately

Table 5
The population, water losses, BCC efficiency scores, and scale (u^*) values from 1999 to 2002

Municipality	Population	Water losses (cubic meters)	Efficiency score	RTS[a] (u^*)
1999				
Gaza	394,815	9,748,195	1	C
Ramallah	210,000	2,753,448	1	C
Khan Yunis	152,893	2,931,818	0.637	I
Rafah	120,296	1,806,173	1	I
Jabalia	112,339	3,222,008	0.763	I
Nablus	108,819	2,295,151	1	C
Bethlehem	78,000	1,645,914	1	I
Nuseirat	46,234	416,719	1	I
Dier Al Balah	44,682	1,160,130	0.830	I
Tulkarm	36,860	1,742,998	0.681	C
Qalqilya	34,997	628,175	0.840	D
El Bureij	34,997	227,232	1	D
Beit Lahia	33,936	628,672	0.895	D
Jenin	27,000	810,772	0.851	I
Bani Suhalia	26,040	261,886	0.963	I
Maghazi	25,444	98,859	1	D
Abassan Kubra	19,907	132,690	0.816	D
Beit Hanun	18,837	467,288	1	D
Qabatiya	15,946	73,785	0.770	D
Zuwadia	14,081	18,838	1	I
Jericho	12,855	1,587,617	0.943	D
Tubas	11,600	81,386	1	D
Qarara	11,224	305,393	1	D
Beitunya	10,354	146,594	0.827	I
Khazaa	9386	43,048	0.994	D
W.Bani Zaid	8790	24,224	1	I
Anabta	8105	146,986	1	D
Salfit	7767	159,804	0.853	I
Azzun	6467	26,762	1	D
Abassan Jadida	5739	37,901	1	D
Dura	1200	43,753	0.570	D
2000				
Gaza	408,633	9,888,052	1	I
Ramallah	215,000	2,812,096	1	I
Khan Yunis	158,244	3,252,065	0.608	I
Nablus	148,622	2,748,628	1	I
Jabalia	116,271	2,425,689	0.954	I
Bethlehem	84,710	1,545,687	1	I
Tulkarm	65,077	1,459,716	0.946	I
Nuseirat	47,853	417,569	1	I
Dier Al Balah	46,246	1,141,501	0.802	I
Jenin	42,690	425,060	1	D
Beit Lahia	35,123	1,235,650	0.802	I
Qalqilya	34,997	534,945	0.880	D
El Bureij	34,997	179,956	1	D
Bani Suhalia	26,952	312,044	0.685	I
Maghazi	26,335	272,520	0.762	I
Abassan Kubra	20,604	34,403	1	D
Beit Hanun	19,496	608,776	0.942	I
Halhul	18,154	68,739	0.778	D
Jericho	17,068	673,396	1	D
Qabatiya	17,010	126,361	1	D
Zuwadia	14,574	38,818	1	D
Tubas	13,626	104,439	0.927	D
Qarara	11,617	292,832	1	D
Beitunya	10,871	154,414	1	I
Khazaa	9715	27,414	1	D
Salfit	8223	139,630	0.894	D
Anabta	7074	134,510	1	D
Azzun	6796	117,555	1	D
Abassan Jadida	5939	53,393	1	D
Rafah	125	1,820,272	1	D

Table 5 (*continued*)

Municipality	Population	Water losses (cubic meters)	Efficiency score	RTS[a] (u^{-})
2001				
Gaza	422,936	8,149,956	1	I
Ramallah	235,600	3,100,027	1	I
Nablus	165,000	2,192,449	1	I
Khan Yunis	163,783	3,670,151	0.370	D
Rafah	128,864	2,346,523	0.452	D
Jabalia	120,341	2,571,227	0.449	D
Bethlehem	80,000	3,140,071	0.405	I
Tulkarm	62,000	1,551,735	0.999	I
Nuseirat	49,528	1,471,575	0.857	I
Dier Al Balah	47,864	863,668	0.384	D
Qalqilya	40,000	460,631	1	I
Beit Lahia	36,353	1,369,280	0.560	D
El Bureij	34,997	302,563	1	D
Bani Suhalia	27,895	194,971	0.387	D
Maghazi	27,257	281,484	0.660	D
Abassan Kubra	21,325	132,532	0.867	D
Beit Hanun	20,179	743,620	0.851	D
Beitunya	17,000	201,460	1	I
Zuwadia	15,084	112,483	0.767	D
Tubas	14,000	62,010	0.731	D
Qarara	12,024	329,328	1	D
Khazaa	10,055	41,490	0.977	D
Salfit	9000	155,814	0.707	D
Anabta	8000	185,032	1	D
Azzun	6800	107,228	1	D
Beir Zeit	6458	26,754	1	D
Abassan Jadida	6147	45,509	1	D
2002				
Gaza	437,739	8,158,497	1	I
Ramallah	247,983	3,197,225	1	I
Nablus	170,000	2,353,389	1	I
Khan Yunis	169,514	3,787,043	0.227	D
Rafah	133,374	2,056,569	0.228	D
Jabalia	124,552	3,352,671	0.378	D
Bethlehem	80,000	2,022,758	1	I
Tulkarm	68,000	1,528,405	1	I
Nuseirat	51,261	700,757	0.556	D
Dier Al Balah	49,540	1,277,879	0.302	D
Jenin	37,915	272,165	1	I
Beit Lahia	37,625	1,567,820	0.441	D
Qalqilya	34,997	578,503	0.858	I
El Bureij	34,997	287,638	1	D
Bani Suhalia	28,872	229,206	0.332	D
Maghazi	28,211	315,683	0.617	D
Abassan Kubra	22,071	114,556	0.697	D
Beit Hanun	20,885	740,722	0.674	D
Jericho	18,239	671,464	1	I
Qabatiya	18,000	175,595	0.884	D
Halhul	18,000	58,996	0.799	D
Beitunya	17,000	185,915	0.960	I
Zuwadia	15,612	109,907	0.878	D
Dura	15,000	45,227	0.784	D
Tubas	14,400	119,919	0.803	D
Qarara	12,445	340,554	1	D
BirZeit	12,000	28,337	1	I
W.Bani Zaid	12,000	21,422	1	D
Salfit	11,000	172,233	0.709	D
Khazaa	10,407	29,218	1	D
Anabta	10,000	215,907	1	D
Azzun	7,140	103,210	1	D
Abassan Jadida	6362	61,910	1	D

[a]Constant Returns to Scale (C), Increasing Returns to scale (I), and Decreasing returns to scale (D).

Table 6
The BCC efficiency scores and the slacks in the inputs for the different DMUs from 1999 to 2002

DMU	1999					2000					2001					2002				
	Excess water losses	Excess water and energy	Excess Maint. and others	Excess salaries of workers	Efficiency score	Excess water losses	Excess water and energy	Excess Maint. and others	Excess salaries of workers	Efficiency score	Excess water losses	Excess water and energy	Excess Maint. and others	Excess salaries of workers	Efficiency score	Excess water losses	Excess water and energy	Excess Maint. and others	Excess salaries of workers	Efficiency score
Jenin	0	0	0	0	0.851	0	0	0	0	1						0	0	0	0	1
Qabatiya	0	0	0	31112	0.770	0	0	0	0	1						61862	0	0	0	0.884
Tubas	0	0	0	0	1	0	0	11215	24282	0.927	0	0	18297	17914	0.731	82701	0	0	0	0.803
Tulkarm	122358	0	78692	0	0.681	0	0	0	84880	0.946	0	0	0	425326	0.999	0	0	0	0	1
Anabta	0	0	0	0	1	0		0	0	1	0.880	0	0	0	1	0	0	0	0	1
Qalqilya	0	0	0	0	0.840	0	0	96213	0	0.880	0	0	0	0	1	0	0	148528	0	0.858
Azzun	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
Nablus	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0.707	140625	0	0	0	0.709
Salfit	164707	0	0	0	0.853	174402	0	0	32771	0.894	50956	0	0	0	1	0	0	0	0	1
Ramallah	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0				
Beir Zeit											0	0	0	0	1	0	0	0	0	1
Beitunya	0	0	0	0	0.827	0	0	0	0	1	0	0	0	0	1	0	359805	0	0	0.960
W. Bani Zaid	0	0	0	0												0	0	0	0	1
Jericho	3781060	0	1519922	107046	0.943	0	0	0	0	1						0	0	0	0	1
Bethlehem	0	0	0	0	0			0	0	1	685390	1E+06	74957	0	0.405	0	0	0	0	1
Halhul		0				0	0	13702	0	0.778						0	0	0	127924	0.799
Dura	0	0	0	80645	0.570											0	0	0	49454	0.784
Gaza	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
Rafah	0	0	0	0	1	0	0	0	0	1	186887	0	0	161941	0.452	60274	0	24064	0	0.228
Beit Hanun	0	0	0	0	1	77748	0	11321	0	0.942	343828	0	0	0	0.851	301567	0	0	0	0.674
Beit Lahia	0	0	0	0	0.895	174872	0	5615	0	0.802	224104	0	0	0	0.560	314581	0	0	0	0.441
Jabalia	0	0	0	0	0.763	156867	0	1131653	0	0.954	0	0	437063	0	0.449	188876	0	0	0	0.378
El Bureij	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
Nuseirat	0	0	0	0	1	0	0	0	0	1	987709	1E+06	0	60497	0.857	448523	855757	0	0	0.556
Zuwadia	0	0	0	0	1	0	0	0	0	1	98884	0	0	0	0.767	86752	0	0	0	0.878
Maghazi	0	0	0	0	1	127531	0	0	0	0.762	187744	0	0	0	0.660	284808	0	0	0	0.617
Dier Al Balah	320405	0	0	0	0.830	355598	0	0	0	0.802	0	0	0	2461	0.384	26129	0	0	0	0.302
Khan Yunis	903514	0	0	0	0.637	2398021	0	0	55685	0.608	947547	0	0	121374	0.370	660599	0	24779	0	0.227
Qarara	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
Abassan																				
Jadida	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
Abassan		0																		
Kubra	80329	0	0	0	0.816	0	0	0	0	1	0	145437	0	30000	0.867	55833	90612	0	17135	0.697
Khazaa	0	0	0	19841	0.994	0	0	0	0	0	1	207425	0	13972	0.977	0	0	0	0	1
Beni Suhalia	184207	0	0	23124	0.963	103119	0	0	0	0.685	25155	0	0	5434	0.387	4448	0	0	0	0.332

Table 7
The efficiency scores and the radius of stability values according to the additive model from 1999 to 2002

Efficiency ranking	1999			2000			2001			2002		
	Municipality	Efficiency score	Stability value	Municipality	Efficiency score	Stability value	Municipality	Efficiency score	Stability value	Municipality	Efficiency score	Stability value
1	Ramallah	1	5.1281	Ramallah	1	5.7405	Ramallah	1	5.0933	Ramallah	1	3.3245
2	Nablus	1	1.8677	Nablus	1	1.2141	Nablus	1	0.5897	Nablus	1	2.8765
3	Gaza	1	0.4319	Bethlehem	1	0.4867	Gaza	1	0.5459	Bethlehem	1	0.3625
4	Bethlehem	1	0.3426	Gaza	1	0.4607	Beitunya	1	0.0717	Jericho	1	0.1943
5	Rafah	1	0.1127	Rafah	1	0.1206	Qalqilya	1	0.0518	Jenin	1	0.1938
6	Zuwadia	1	0.0428	Jericho	1	0.0614	Beir Zeit	1	0.0436	Gaza	1	0.1516
7	Abassan Jadida	1	0.0323	Jenin	1	0.0494	Qarara	1	0.0434	Tulkarm	1	0.1436
8	Nuseirat	1	0.0259	Qarara	1	0.0328	Abassan Jadida	1	0.0411	Qarara	1	0.0731
9	Qarara	1	0.0248	Abassan Jadida	1	0.031	Anabta	1	0.0345	Beir Zeit	1	0.066
10	Azzun	1	0.021	El Bureij	1	0.0271	Azzun	1	0.0274	W.Bani Zaid	1	0.0607
11	El Bureij	1	0.0149	Azzun	1	0.0247	El Bureij	1	0.0111	Azzun	1	0.033
12	Beit Hanun	1	0.0132	Abassan Kubra	1	0.0224	Tulkarm	0.912	−0.0002	Abassan Jadida	1	0.0323
13	Anabta	1	0.0113	Nuseirat	1	0.0199	Khazaa	0.714	−0.0008	Khazaa	1	0.0276
14	W.Bani Zaid	1	0.0102	Beitunya	1	0.0173	Abassan Kubra	0.547	−0.0026	Anabta	1	0.0043
15	Maghazi	1	0.0012	Khazaa	1	0.0112	Nuseirat	0.376	−0.0067	El Bureij	1	0.0031
16	Tubas	1	0.0005	Qabatiya	1	0.0059	Zuwadia	0.695	−0.0128	Beitunya	0.851	−0.0069
17	Khazaa	0.911	−0.0003	Zuwadia	1	0.0019	Beit Hanun	0.594	−0.0133	Zuwadia	0.789	−0.0115
18	Bani Suhalia	0.809	−0.0015	Anabta	1	0.0004	Tubas	0.707	−0.0183	Abassan Kubra	0.559	−0.0122
19	Jericho	0.209	−0.0051	Beit Hanun	0.844	−0.0056	Maghazi	0.526	−0.0196	Qabatiya	0.811	−0.0165
20	Abassan Kubra	0.753	−0.0094	Tubas	0.713	−0.0067	Salfit	0.703	−0.0229	Tubas	0.754	−0.0287
21	Salfit	0.680	−0.0157	Salfit	0.709	−0.0085	Bani Suhalia	0.510	−0.0302	Halhul	0.700	−0.0296
22	Beit Lahia	0.865	−0.0186	Maghazi	0.610	−0.0144	Beit Lahia	0.311	−0.076	Dura	0.749	−0.0328
23	Qabatiya	0.689	−0.0224	Jabalia	0.554	−0.0162	Dier Al Balah	0.344	−0.1334	Maghazi	0.553	−0.034
24	Dier Al Balah	0.687	−0.0304	Halhul	0.750	−0.0196	Jabalia	0.257	−0.1831	Qalqilya	0.632	−0.0342
25	Beitunya	0.816	−0.0415	Tulkarm	0.936	−0.0263	Khan Yunis	0.231	−0.1887	Nuseirat	0.329	−0.0443
26	Qalqilya	0.831	−0.0445	Qalqilya	0.798	−0.0269	Rafah	0.365	−0.2028	Salfit	0.659	−0.0491
27	Dura	0.405	−0.0478	Bani Suhalia	0.637	−0.0285	Bethlehem	0.376	−0.6862	Beit Hanun	0.520	−0.0553
28	Jenin	0.835	−0.0744	Beit Lahia	0.713	−0.0375				Bani Suhalia	0.470	−0.0716
29	Jabalia	0.696	−0.1192	Dier Al Balah	0.678	−0.0476				Beit Lahia	0.304	−0.1272
30	Khan Yunis	0.492	−0.1734	Khan Yunis	0.446	−0.1269				Dier Al Balah	0.358	−0.2194
31	Tulkarm	0.666	−0.1806							Khan Yunis	0.179	−0.275
32										Jabalia	0.370	−0.2789
33										Rafah	0.254	−0.3338

recording the volume of water being consumed. Some of the municipalities were being supplied from their own wells and others from outside sources. For this reason both the energy used to pump water and purchased water costs were combined as one input. However, these inaccuracies might differ between municipalities, introducing inaccuracies in the final analysis. There is a need for a unified water cost system to correct for these differences, and such a system currently does not exist.

The amount of piped water consumed is not a very good indicator of efficiency because the customers in this region often rely on tanks of water for drinking because some of the piped water is highly saline. In the Gaza Strip, the model should contain an input to measure the drinkable quality of the water. This could help to identify which of the municipalities have excess pollutants and lead to better water management. At present, such data were not obtainable from the municipalities. The districts may have collected data but applying these data to smaller units is troublesome. Further research needs to be done to look at other variables in addition to the inputs and outputs used in this study. One major potential input in the empirical model might be a water pollution indicator to be minimized. Another input used in other studies is the length of pipes as a means of preventing water losses from evaporation. Finally, the total revenue used as an output has some bias because it was based on the value of the water sold (billed) and not the actual collections (albeit the two variables could be good proxies for each other as in the case of corporate accounting). Thus, the differences in the revenue collected and the money owed by the customers should be considered when checking for efficiency.

This research has highlighted the advantage of using DEA as a tool to evaluate water use efficiency in the Palestinian Territories. While improving the efficiency of water use in the Palestinian Territories is recommended as an essential part of the water management, it should be considered along with other solutions such as desalination projects and recycling. This study's contribution is its focus first on the West Bank and Gaza as a whole, and then on comparing them to each other as separate units. It contributes some DEA-based efficiency evaluations in water resources management that will help Palestinian policy makers address such difficult resource allocation issues as proposals to merge small municipalities.

References

Aida, K., Cooper, W.W., Pastor, J.T., Sueyoshi, T., 1998. Evaluating water supply services in Japan with RAM: a range-adjusted measure of inefficiency. OMEGA, International Journal of Management Science 26, 207–232.

Akosa, G., Franceys, R., Baker, P., Weyman-Jones, T., 1995. Efficiency of water-supply and sanitation projects in Ghana. Journal of Infrastructure System 1, 56–65.

Ali, I.A., Seiford, L.M., 1993. Computational accuracy and infinitesimals in data envelopment analysis. INFOR 31, 290–297.

Al-Ibrahim, A.A., 1990. Water use in Saudi Arabia: problems and policy implications. Journal of Water Resources Planning and Management 116, 375–388.

Anwandter, L., 2000. Can public sector reforms improve the efficiency of public water utilities? An empirical analysis of the water sector in Mexico using Data Envelopment Analysis. Doctoral dissertation, University of Maryland, College Park.

Banker, R.D., Charnes, A., Cooper, W.W., 1984. Models for the estimation of the technical and scale inefficiencies in data envelopment analysis. Management Science September, 1078–1092.

Bowlin, W.F., 2004. Measuring performance: an introduction to data envelopment analysis (DEA). Retrieved March 1, 2004 from 〈http://llanes.panam.edu/edul8305/papers/introtodea.pdf〉.

Chang, H., Cheng, W., Das, S., Li, S., 2004. Health care regulation and operating efficiency of hospitals: evidence from Taiwan. Journal of Accounting and Public Policy 23, 483–510.

Charnes, A., Cooper, W.W., Rhodes, E., 1981. Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. Management Science 27, 668–697.

Charnes, A., Cooper, W.W., Golany, B., Seiford, L., Stutz, J., 1985. Foundations of data envelopment analysis for Pareto–Koopmans efficiency empirical production functions. Journal of Econometrics 30, 91–107.

Charnes, A., Cooper, W.W., Li, S., 1989. Using DEA to evaluate efficiency in the economic performance of Chinese cities. Socio-Economic Planning Sciences 23, 325–344.

Charnes, A., Haag, S., Jaska, P., Semple, J., 1992. Sensitivity of efficiency classification in the additive model of data envelopment analysis. International Journal of Systems Science 23, 789–798.

Charnes, A., Cooper, W.W., Lewin, A.Y., 1994. Data Envelopment Analysis: Theory, Methodology, and Application. Kluwer Academic Publishers, Norwell.

Charnes, A., Rousseau, J., Semple, J., 1996. Sensitivity and stability of efficiency classifications in data envelopment analysis. Journal of Productivity Analysis 7, 5–18.

Cook, W.D., Kress, M., Seiford, L.M., 1996. Data envelopment analysis in presence of both quantitative and qualitative factors. Journal of Operational Research Society 47, 945–953.

Cooper, W.W., Li, S., Seiford, L.M., Tone, K., Thrall, R.M., Zhu, J., 2001. Sensitivity and stability analysis in DEA: some recent developments. Journal of Productivity Analysis 15, 217–246.

Cooper, W.W., Seiford, L.M., Tone, K., 2002. Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References, and DEA-Solver Software. Kluwer Academic Publishers, Norwell.

Cooper, W.W., Seiford, L.M., Zhu, J., 2004. Handbook on Data Envelopment Analysis. Kluwer Academic Publishers, Norwell.

Cooper, W.W., Seiford, L.M., Tone, K., 2006. Introduction to Data Envelopment Analysis and its Uses: With DEA-solver Software and References. Springer Science and Business Media, Inc, New York.

Cubbin, J., Tzanidakis, G., 1998. Regression versus data envelopment analysis for efficiency measurement: an application to the England and Wales regulated water industry. Utilities Policy 17, 75–85.

Department of Water Resources, State of California, 2004. Water use efficiency. Leak detection frequently asked questions. Retrieved January 14, 2004 from 〈http://www.owue.water.ca.gov/faq/faq.cfm〉.

Dolatyar, M., Gray, T., 2000. Water Politics in the Middle East. St. Martin's Press, Inc, New York.

Elmusa, S., 1997. Water Conflict: Economics, Politics, Law, and the Palestinian–Israeli Water Resources. Institute for Palestinian Studies, Washington, DC.

Feroz, E.H., Raab, R., Haag, S., 2001. An income efficiency model approach to the economic consequences of the OSHA cotton dust regulation. Australian Journal of Management 26, 69–89.

Green, E.A., 1993. Hydropolitics in the Middle East and US policy (Report). US Naval War College, Newport.

Haddad, M., 1998. Planning water supply under complex and changing political conditions: Palestine as a case study. Water Policy 1, 177–192.

Jerusalem Water Undertaking. 2003. Unaccounted for Water. Retrieved March 1, 2003 from ⟨http://www.jwu.org/jwubk_09.htm⟩.

Palestinian Water Authority. 2002. Water supply in the West Bank for year 2001 (Water Resources and Planning Directorate Report). Ramallah, Palestinian Water Authority.

Palestinian Central Bureau of Statistics, 2004. Labour Force Survey: (April–June, 2004) Round (Q2/2004) (Press Conference on the Labour Force Survey Results). Ramallah, Palestine.

Raab, R., Lichty, R., 2002. Identifying subareas that comprise a greater metropolitan Area: The criterion of county relative efficiency. Journal of Regional Science 42, 579–594.

Rouyer, A.R., 2000. Turning Water into Politics: The Water Issues in the Palestinian–Israeli Conflict. Palgrave, New York.

Rowland, M.A., 2000. The evolution of two water resource management systems: case studies of Tampa Bay and the Middle East. Colorado Journal of International Environmental Law & Policy 11, 411–496.

Sabbah, W., Isaac, J., 1995. June. Toward a Palestinian water policy. Retrieved March 20, 2003, from, ⟨http://www.arij.org/pub/index.htm⟩. Presented at the Seminar on Options and Strategies for Freshwater Development and Utilization in Selected Arab Countries, organized by the Center for Environment and Development for the Arab Region and Europe (CEDARE), Amman.

Sanders, S.D., 1999. Data envelopment analysis for benchmarking public sector utilities operations. Doctoral dissertation, Southern Methodist University.

Stolp, C., 1990. Strength and weaknesses of data envelopment analysis: an urban and regional perspective. Computers, Environment, and Urban Systems 14, 103–116.

Thanassoulis, E., 1999. DEA and its use in the regulation of water companies. European Journal of Operational Research 127, 1–13.

Thanassoulis, E., 2000. The use of data envelopment analysis in the regulation of UK water utilities: water distribution. European Journal of Operational Research 126, 436–453.

Thanassoulis, E., 2002. Comparative performance measurement in regulation: the case of English and Welsh sewerage services. Journal of European Research Society 53, 292–302.

Wood, W.W., 2003. A fresh water odyssey: some observations on the global resource. Ground Water 41, 300–305.

World Bank, 1993. Developing the Occupied Territories: An Investment in Peace. World Bank, Washington, DC.

World Bank Group, 2004. Urban and water sanitation. Retrieved January 13, 2004, from ⟨http://lnweb18.worldbank.org/mna/mena.nsf/⟩. All/6C586003928975DE85256949006FC441?OpenDocument.

ELSEVIER

Journal of Environmental Management 87 (2008) 95–105

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Temperature and relative humidity distributions in a medium-size administrative town in southwest Nigeria

O.M. Akinbode[a], A.O. Eludoyin[b,*], O.A. Fashae[c]

[a]Department of Geography and Planning Sciences, Adekunle Ajasin University, Akungba–Akoko, Nigeria
[b]Department of Geography, Obafemi Awolowo University, Ile-Ife, Nigeria
[c]Department of Geography, University of Ibadan, Nigeria

Received 29 October 2004; received in revised form 21 September 2006; accepted 4 January 2007
Available online 7 May 2007

Abstract

This study was carried out in one of the medium-sized public administrative towns in the southwestern part of Nigeria. Its aim is to highlight the effect of spatial distribution of settlements, population, and socio-economic activities on urban air temperature and humidity in the town.

Temperature and relative humidity data from 1992 to 2001 were obtained from three meteorological stations in Akure, the Administrative Capital of Ondo State, Nigeria. The stations are located within the Federal Ministry of Aviation, Akure Airport (FMA), Federal University of Technology, Akure (FUTA) and Federal School of Agriculture (SOA). Air temperature and relative humidity measurements were also obtained from 27 points, which were cited to include road junctions, markets, built up areas, etc., using sling psychrometer. The data were subsequently analysed for spatial and temporal variations using statistical packages (SPSS and Microsoft Excel) and isolines. Actual vapour pressure and dew point temperature were computed using Magnus conversion formulae.

The results obtained showed that spatial variation was insignificant, in terms of the temperature and humidity variables. The annual mean temperature (T_{mean}) ranged between 21.9 and 30.4 °C while minimum (T_{min}) and maximum (T_{max}) temperatures varied from 13 to 26 and 21.5–39.6 °C, respectively. Relative humidity (RH), actual vapour pressure (E_s) and dew point temperature (T_d) values also varied from 39.1% to 98.2%, 19.7–20.8 g m^{-3}, and 17.3–17.8 °C, respectively. A significant relationship ($p>0.6$; $r<0.05$) between T_{min}, E_s and T_d was observed while the daytime 'urban heat island' intensity (UHI) ranged between 0.5 and 2.5 °C within the study period.

The study concluded that there is influence of urban canopy on the microclimate of Akure, and hypothesizes that the urban dwellers may be subjected to some levels of weather related physiological disorderliness.

Keywords: Administrative town; Temperature; Relative humidity; Spatial and temporal variation; Spatial distribution and 'urban heat island'

1. Introduction

1.1. Situation

Nigeria is one of the few countries in tropical Africa where urbanization is not a recent phenomenon. Some Nigerian towns are as old as some of the ancient cities in Europe. These include cities such as Kano, Katsina and Sokoto in the northern part, which participated vigorously in the worldwide commercial activities in the medieval period. Others include the large settlements in the Southwestern region such as Ado-Ekiti, Benin City, Ibadan, Oyo, Iwo, Oshogbo, Ogbomosho, which were pre-colonial towns (Onakerhoraye and Omuta, 1994). Some recent towns have also emerged as medium-sized public administrative towns. Such public administrative towns include Akure (Ondo State), Uyo (Akwa Ibom State), Asaba (Delta State), etc., whose declaration as administrative capital cities have precipitated their growth. These towns have served as functional zones to the areas outside them, performing commercial, administrative, residential,

*Corresponding author. Tel.: +234 805398631.
E-mail addresses: baynick2003@yahoo.com (O.M. Akinbode), oaeludoyin@yahoo.com, oaeludoyin@oauife.edu.ng (A.O. Eludoyin), toyinafashae@yahoo.com (O.A. Fashae).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.018

religious and educational roles. Such roles have endeared immigrants to them, and sometimes leading to over-population, over-utilization of the resources at the region of highest population concentration, and stratification of such towns into economic and occupational classes and sub-classes. The central region, termed 'core' is always the most developed and often acquires the lion's share of the community's resource. The periphery, on the other hand, suffers from inadequate resource allocation, relative under-development and under-population (Onakerhoraye and Omuta, 1994). The city core, as a result of population concentration, intense transportation activity and surface concretization within it could therefore exhibit a micro-climate different from other part of the same town, giving rise to a condition of greater core temperature difference termed the 'urban heat island' intensity (UHI).

The UHI is the characteristic warmth of urban areas compared to their outskirts. It is also often referred to as the increase of air temperature in the near-surface layer of the atmosphere within cities relative to their surrounding countryside (Voogt, 2002). It is so named because the isotherms forms an island shaped pattern. The built-up environment has been found to exacerbate heat stress, particularly at night, during heat waves and provides a preferential site for spread of vector borne diseases (Samuels, 2004; Svensson and Tarvainen, 2004).

The rapid pace of urbanization has been shown to be a global problem present in most of the developing countries. For instance, the urban populations in these countries have grown by 40% between 1900 and 1975. Furthermore, there is every indication that the trend will continue for the next 30 years, adding approximately two billion people to the urban population of the presently less-developed nations (United Nations Environment Programme, 2002). Nigeria's urbanization rate was conservatively put at 5% as at 1973 (Olowu, 1985) and by 1980 nearly 30% of the total population has already been residing in urban centres. By 2005, approximately 60% of the populations in the Africa in general are expected to be urban. Nigeria's urban centres have been growing at double the national population (estimated at 2.5%) and at double the rate of urban information in several countries of Europe and America (Olowu, 1985).

This paper presents a preliminary assessment of the variations in the temperature and humidity in a Capital city in southwest Nigeria. The area has been selected because of its rapid urbanization from 1976 until present. The results presented are 10 years (1992–2002) of air temperature and humidity data from three established meteorological stations, and daytime, periodic, observations taken from 18 stations in the city in December 2002.

1.2. Problem

Previous studies (e.g. Adebayo, 1990 and Omogbai, 1985) have shown a local climate with spatial structure such as heat island is formed within some medieval Nigerian settlements compared to outside open spaces. Few of these studies are published. On the other hand, studies of urban climate are important because they investigate the alterations that the human actions promote in the low atmosphere (Maitelli et al., 2002). For instance, the heat emissions from many anthropogenic sources add to the warming of the built environment while the vegetation in the suburban area has been revealed to cool the environment (e.g. Givoni, 1991; Samuels, 2004). The high-rise buildings in the urban areas magnify the impact of the built environment (Santamouris, 2001; Arnfield, 2003), trapping heat and distorting air movement. The warming effect that results from this phenomenon could affect, significantly, the comfort and the liveability of the urban people. For instance, over 10,000 people died of heat wave in France in 2003 (Samuels, 2004) on which urban heat islands add an additional stress. Heat islands in the atmosphere are best expressed at night under calm and clear conditions when differential rates of radiative cooling are maximized between urban areas and their surroundings, with cities cooling more slowly than their surroundings. Heat island magnitudes decrease with increasing wind speed and cloud cover. Daytime heat islands may be positive or negative depending on the particular characteristics of the urban area and their surroundings (e.g. Runnals and Oke, 2000).

While much information is available on this phenomenon and its effects on the people in the European and American countries (e.g. Kilbourne, 1989; Kalkstein and Greene, 1997; Hajat et al., 2002), the microclimate of Nigerian cities has not been well documented, and little is known about their effects, especially in the growing state capitals, which form the present focus of socio-economic growth in the country.

1.3. Objective and hypothesis

The objective of this investigation is to assess the impact of urbanization on the spatial and temporal variations of temperature and humidity in Akure, a medium size, growing area that is the administrative and economic capital of Ondo State in Nigeria. The conclusions are based on the hypotheses that air in the urban canopy is usually warmer than that in the countryside (Oke, 1987), and that there has been no net warming of the countryside (Nordli, 2001).

1.4. Study area

Akure is located on latitude 7°17′N and longitude 5°18′E. The urbanized land area is approximately 41.2 km^2. It became an administrative and economic seat to Akure South Local Government Authority, and Ondo State with the latter's creation in 1976 from the old Western Region. Since then, the city has witnessed immense growth in the size of built-up areas, number of immigrants, transportation, and commercial activities. It experiences warm humid tropical climate, with average rainfall of about 1500 mm per annum. Annual average temperatures range between 21.4 and 31.1 °C, and its mean annual relative humidity is about 77.1%. Its vegetation is the tropical rainforest type (Ayoade, 1993; Oguntoyinbo, 1982). Akure lies on a relatively flat plain of about 250 m above sea level within the Western Nigerian plains. The pattern of land use distribution shows that about 65.7% is used for residential purposes while the remaining portions are shared for industrial (2.1%), commercial (1.6%), public offices (14.4%), cultural or recreational (0.2%) and educational purposes (14.8%). Unused or vacant land was 1.1%.

The centrally located Oba or Erekesan market remains the most important centre of commercial activities in the town. The area also forms a node for a number of roads linking other parts of the town and nearby settlements. It thus remains the primary central business district (CBD) of the city. Secondary CBDs include Isikan and NEPA markets located, respectively, at the southwest and southeastern parts of Oba market. The residential zone extends into the commercial zones and is densely populated around these markets, with modern multi-storey buildings from the post 1976 development. Another category of residential areas comprises the planned estates, including Ijapo, Ala and Alagbaka, which are government owned estates, and Oluwatuyi Quarters, a private estate property.

The three meteorological stations are located surrounding the city: the Federal Ministry of Aviation, Akure Airport, Federal University of Technology, and the Federal School of Agriculture, Akure. These have been designated as FMA, FUTA and SOA stations, respectively, in this study (Fig. 1). The FMA and SOA are located to the northeast while FUTA is at the northwestern part of the town. They are all at the outskirts of the town, characterized by light forest vegetation and scattered trees. A *Gmelina* sp. plantation and experimental farms (for student) exist around the SOA. The natural vegetation is regrowth following the clearing of the previous dense forest cover.

Tall buildings and tarred surfaces are found within the environment of the stations, and there are substantial populations surrounding FUTA and SOA stations, that average between 10,000 and 5000, respectively, but which vary with the programme schedules of the institutions. FMA, on the other hand, is surrounded by fewer buildings mainly offices.

2. Materials and methods

2.1. Sample design, collection, and field observations

Field data were collected in December 2002. Twenty-seven sampling stations were established based on the land use map of Akure: three from each of Oke-Aro, and Araromi districts (residential), Oba, Isikan and NEPA markets (commercial), COOP (industrial), Oba Ile and Owo roads (vegetated and relatively undisturbed sites), and Ondo-Ore Motor Park. Temperature and relative humidity readings were obtained at road junctions, under overhead bridges and open spaces to reflect the various characteristics of the roads in nine stations, along the two primary

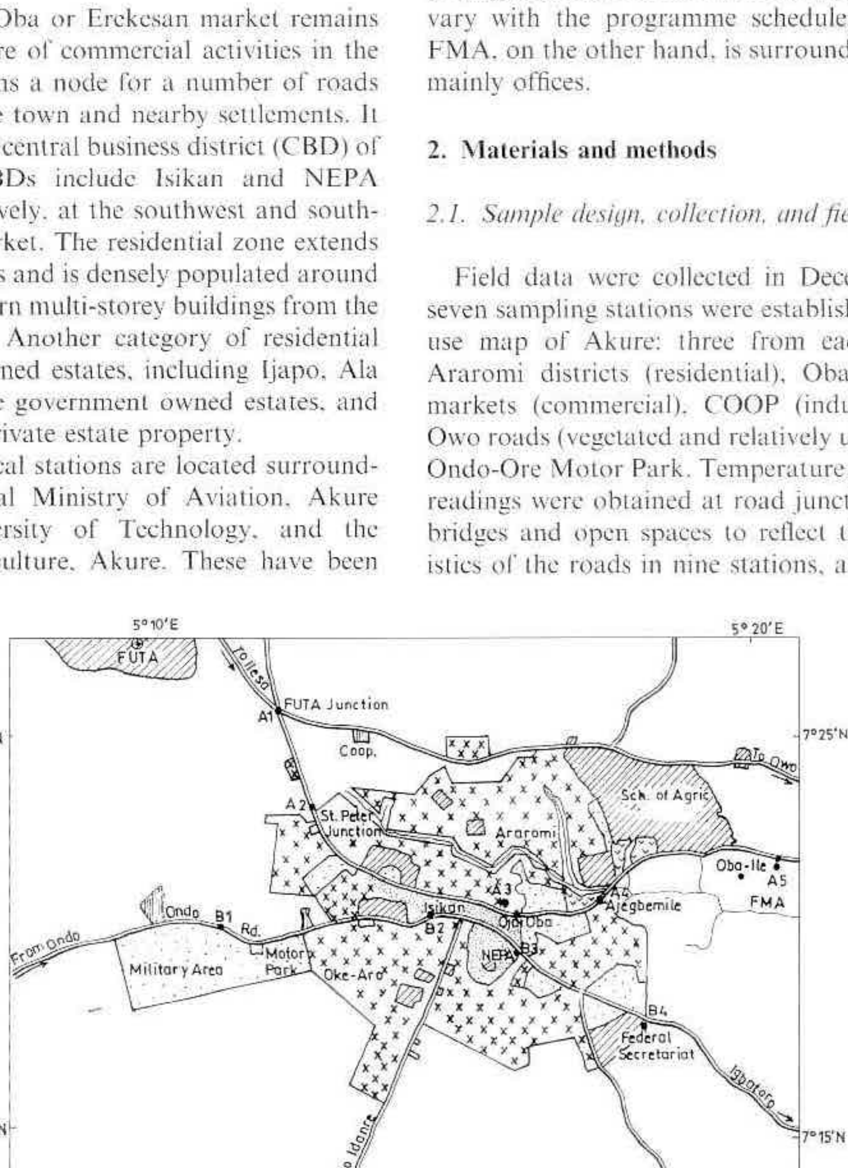


Fig. 1. Land use map of Akure showing sampling stations.

(Ilesa to Oba-Ile and Ondo to Igbatoro) roads that traverse the city (Fig. 1).

Air temperature and relative humidity at these stations were made for 12 days between 09:00 and 18:00 local standard time (LST) from 10 to 21 December 2002. Although studies have revealed that UHI is more of a nocturnal phenomenon, data for night periods could not be collected for security reason and lack of standard unmanned recording instrument for such measurements, at the time of study. The instrument used for the daytime temperature and relative humidity was the sling psychrometer, consisting of both a dry and wet bulb thermometer. To take readings the instrument was whirled for a minimum of 5 min and watched at intervals until at least two consecutive readings were equal. The difference between the dry bulb (air temperature) and the wet bulb readings was checked under appropriate columns in the hygrometric tables to determine the equivalent relative humidity. Readings were made bearing in mind all the precautionary measures in the instrument manual to minimize possible errors.

In addition, a set of temperature and relative humidity data from 1992 to 2001 were collected at the FMA, SOA, and FUTA meteorological stations in Akure. This was the only period when data were commonly available in the three stations. Actual vapour pressure and dew point temperature were computed using the Magnus formula documented elsewhere (e.g. www.ist-ag.com/docs/basics_humidity_e_pdf).

2.2. Statistical analysis

Data were generally subjected to statistical analysis using SPSS (Version 11) and Microsoft Excel software while descriptions of spatial variations are drawn as isolines. Data used for the computation of regression equations were first tested for normality as expected for a parametric statistic (Zar, 1992), and those that did not conform to the assumptions were transformed to logarithms. Missing values at some stations, particularly FUTA and the SOA were replaced by linear interpolation of the stations' time series.

The comparison of sites, land uses, and time were achieved using the one-way analysis of variance (ANOVA). Linear trends for each of the variables for the years were determined with least square linear regression method. Means, standard deviations and ranges, which are measures of central tendencies and dispersions, were used to derive the mean annual temperature and relative humidity from the mean monthly records and to test for dispersion of the distribution of the data for analysis. In addition, spatial variability/heterogeneity of the variables were determined using Principal Component Analysis (PCA) (Zar, 1992). The results derived from these analyses were subjected to F-test statistics and the significance level was accepted at $p = 0.05$.

3. Results

3.1. Variations in the temperature and humidity around the meteorological stations

The mean and standard deviation values of the temperature and humidity obtained from the three meteorological stations in Akure are presented in Table 1. Table 1 shows that T_{mean} ranged from 21.9 to 30.4 °C, with the highest values at FMA and FUTA, respectively. T_d ranged between 17.3 and 18.2 °C, while the E_s varied from 19.7 to 20.8 g m^{-3} (Table 1). The results of the one-way ANOVA reveals that all the station pairs are different in terms of relative humidity and vapour pressure, except FMA and SOA. Significant differences occurred only in the T_{mean} between the FMA and the SOA ($p = 0.01$; $r = 0.05$), and between SOA and FUTA ($p = 0.001$; $r = 0.05$) (Table 2). RH also varied significantly ($p < 0.1$; $r = 0.05$) between FMA and FUTA on the one hand, and SOA and FUTA on the other hand (Table 2).

Fig. 2 shows the mean monthly temperatures and humidity between 1992 and 2001. The monthly value of mean temperature was highest (28.9 °C) in February and lowest (24.4 °C) in August. Conversely, the highest relative humidity was recorded in August (84.2%) and the lowest in February (55.6%). T_d varied from 18.2 °C in April and May to 13.2 °C in January. Comparison of the variables using the Pearson correlation coefficients shows that the temperature variables (except T_{min}) were inversely correlated with the humidity variables. T_{min} on the other hand, correlated significantly with RH, E_s and T_d (Table 3).

In terms of variability, the first two factors accounted for 96.7% of the total variance among the variables. E_s, T_d, RH and T_{min}, which are related to physiologic stress, clustered together (Fig. 3), suggesting possible discomfort among the people in the study area during some time of the year. The dispersion from the T_{max} and T_{min} suggest that E_s and T_d may not be temperature dependent (Table 4).

Table 1
Means and standard deviations of temperature and humidity values from three meteorological stations in Akure between 1992 and 2001

Location		T_{max} (°C)	T_{min} (°C)	T_{mean} (°C)	RH (%)	E(g m^{-3})	T_d (°C)
FMA	$\bar{\chi}$	31.3	21.4	26.3	75	19.7	17.3
	δ	2.4	1.5	1.4	10.9	0.15	−42.0
FUTA	$\bar{\chi}$	30.7	21.4	26.1	76.6	20.0	17.5
	δ	2.1	1.8	1.4	7.4	0.10	−45.7
SOA	$\bar{\chi}$	31.1	21.5	27.0	76.8	20.8	18.2
	δ	2.6	1.4	2.8	1.1	0.03	−54.0
Total	$\bar{\chi}$	31.1	21.4	26.4	76.7	20.3	17.8
	δ	2.4	1.6	1.7	9.2	0.16	−41.4

T_{max} = maximum temperature; T_{min} = minimum temperature; T_{mean} = mean temperature; RH = relative humidity; E = actual vapour pressure; T_d = dew point temperature.

Table 2
ANOVA and Scheffe multiple comparisons of the temperature and relative humidity values from three meteorological stations in Akure

Variables	Overall ANOVA		Scheffee Multiple Comparisons		
	F ratio	F probability[a]	FMA	FMA	SOA
			FUTA	SOA	FUTA
Temperature					
Maximum	2.784	ns[b]	ns	ns	ns
Minimum	0.485	ns	ns	ns	ns
Mean	7.591	0.001	ns	0.011	0.001
Relative humidity	17.116	0.000	0.000	ns	0.007
Actual vapour pressure	14.894	0.037	0.002	ns	ns
Dew point temperature	0.982	ns	ns	ns	ns

[a]The F-probability is significant for the mean differences at the $p \leqslant 0.05$ level.
[b]ns indicates that mean difference is not significant at $p \leqslant 0.05$.

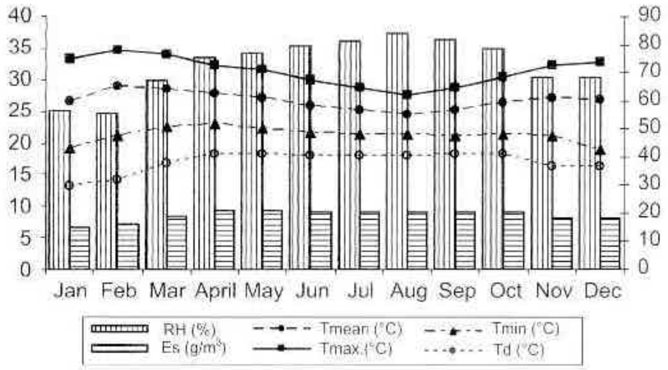


Fig. 2. Monthly distribution of temperature and humidity variables in Akure between 1992 and 2001.

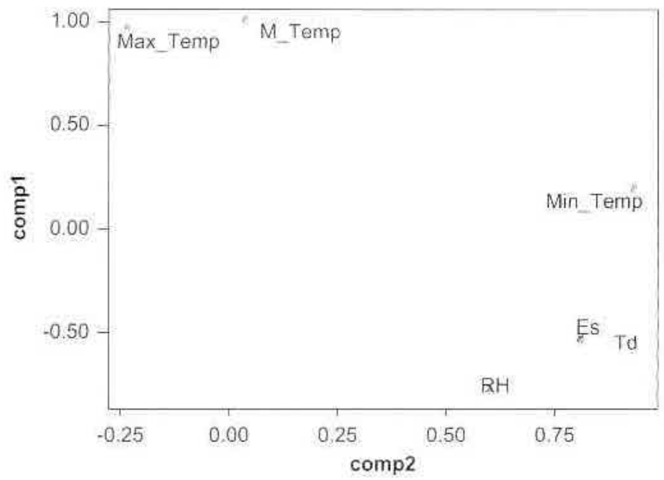


Fig. 3. Result of PCA for temperature and humidity data between 1992 and 2001.

Table 3
Results of the correlation of temperature and humidity variables in Akure, Nigeria between 1992 and 2001

	T_{mean}	T_{max}	T_{min}	RH	E_s	T_d
T_{mean}		0.950	ns	−0.756	−0.498	−0.487
T_{max}			ns	−0.889	−0.690	−0.679
T_{min}				ns	0.624	0.620
RH					0.943	0.938
E_s						0.999

Table 4
Linear trends in annual temperature and relative humidity values between 1992 and 2001

Dependant	R^2	b_0 (intercept)	b_1 (slope)	F probability	F-value
Maximum temperature	0.005	30.8634	0.0264	ns[a]	1.95
Minimum temperature	0.003	21.5394	−0.0124	ns	0.94
Mean temperature	0.002	26.1067	0.0108	ns	0.65
Relative Humidity	0.044	80.0186	−0.3310	0.000	16.28
Dew point temperature	0.208	−9.889	0.968	ns	2.67
Actual vapour pressure	0.190	−9.067	0.804	ns	2.33

[a]ns indicates that mean difference is not significant at $p \leqslant 0.05$.

3.2. *Variations in the temperature and relative humidity within Akure*

Figs. 4a and b show the diurnal variations in the mean temperature and relative humidity values in different land use in Akure in December 2002. Highest mean temperature values were obtained at 15:00 LST in the residential, industrial and commercial zones and 16:00 LST at the Motor Park and the outskirts. Highest T_{mean} (32.5 °C) was obtained at the around the CBD while the residential landuse recorded the lowest (28 °C). Fig. 3a shows that the CBD, Motor Park and saw mill heated more than the outskirts. Conversely, RH was highest at the outskirt (Fig. 4b), which appeared to contain lesser concretized surfaces, automobile patronage and number of tall buildings. Similarly, temperature fluctuated periodically along

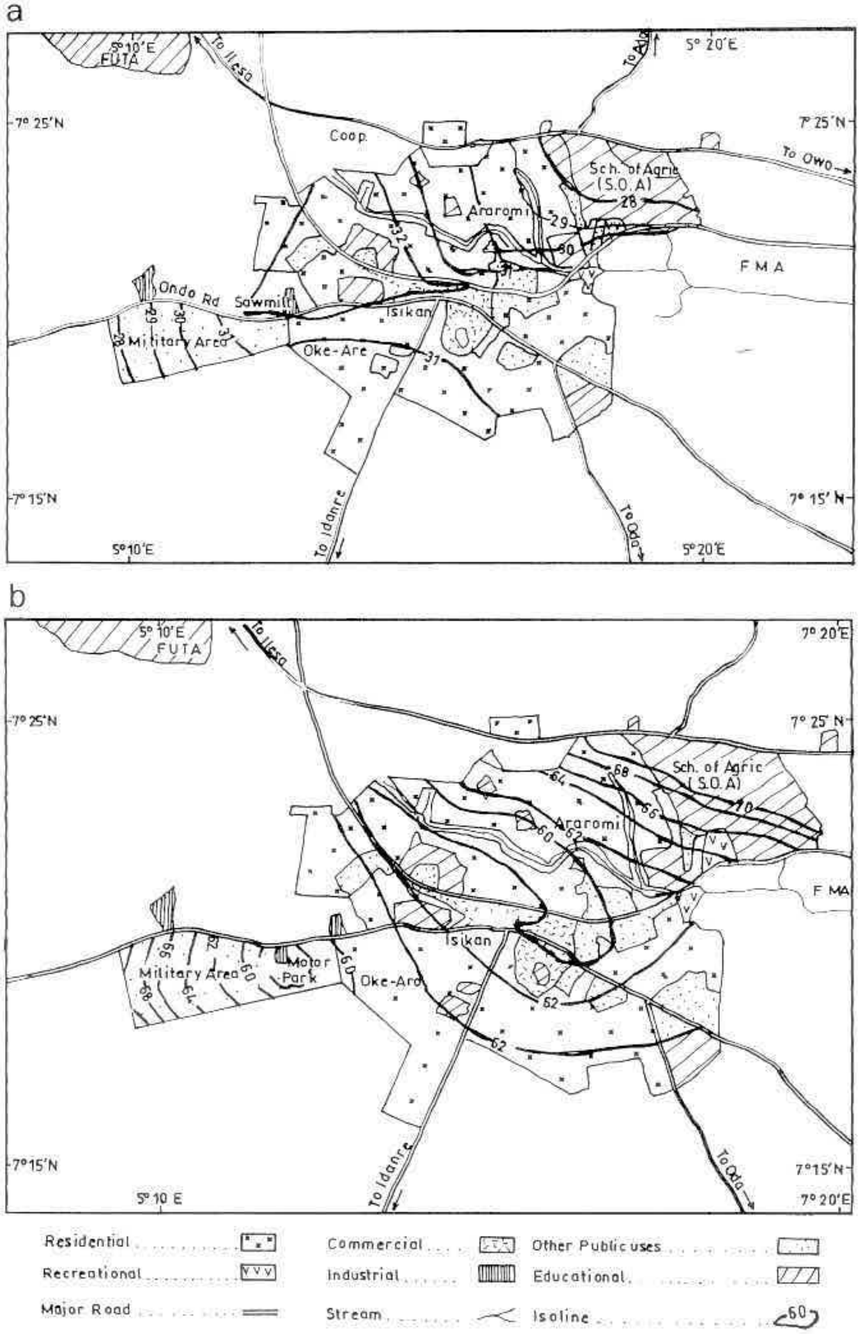


Fig. 4. Variation patterns of temperature (a) and relative humidity (b) in Akure in December 2002.

the roads within the CBD and at junctions (Fig. 5a–e) and its distribution was inversely related to RH (Fig. 6a–e, compare Fig. 1).

3.3. 'Urban heat island' phenomenon

Fig. 7a shows the summary of the daytime UHI pattern observed. The close isotherms around the CBD suggested warmer air temperature. The range of heat intensity was 0.5–2.5 °C. The markets, which incidentally are the CBD, formed the peak as expected. They receive over 70% population of the population daily. The RH field formed its peak towards the northern extremes (northwest and northeast) (Fig. 7b).

4. Discussion

This investigation aimed at assessing the impact of urbanization on the relative humidity and temperature in Akure, Nigeria. The higher temperature and lower relative humidity observed in the commercial area compared to that obtained at the countryside, as seen in this study is

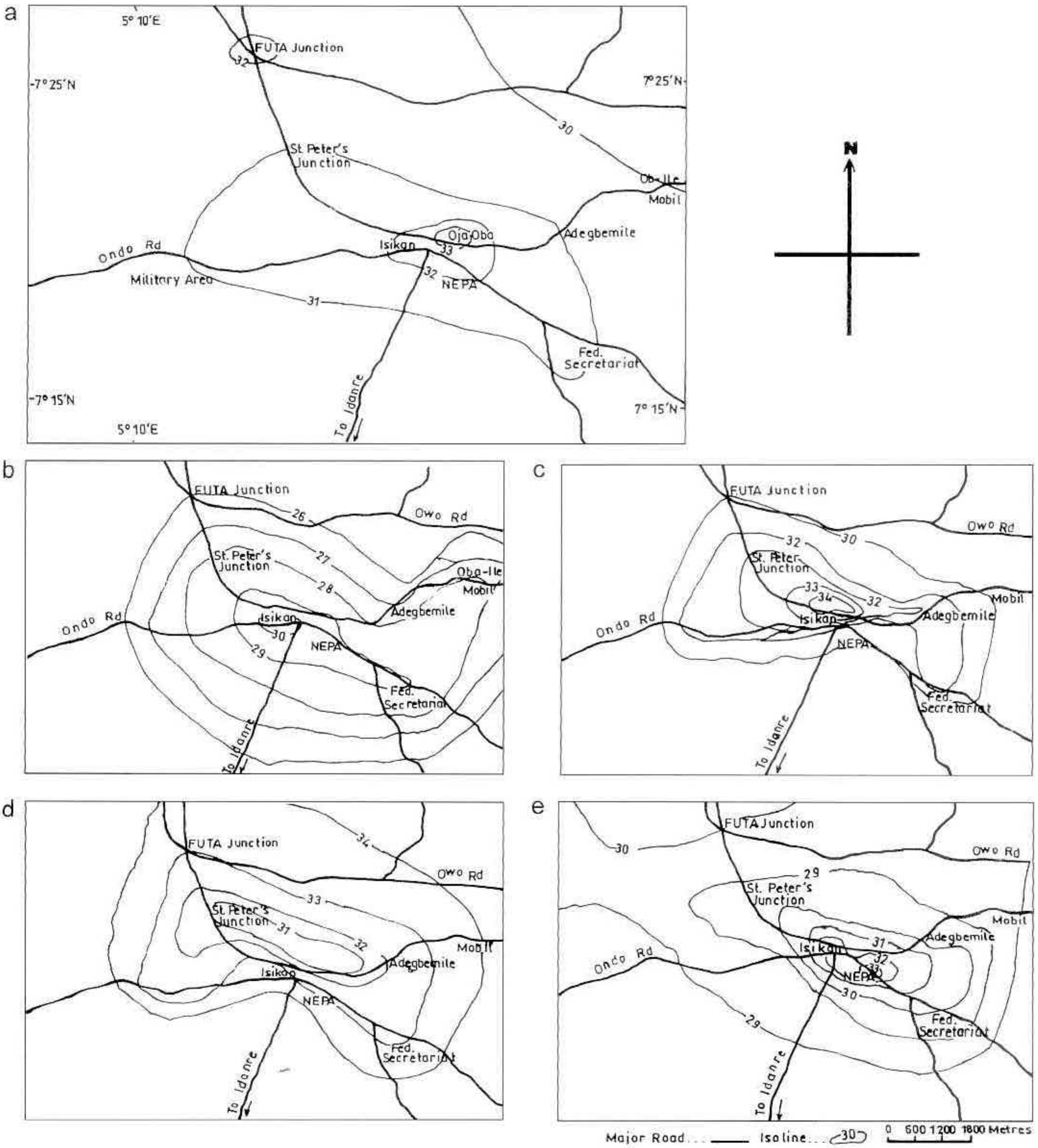


Fig. 5. Diurnal temperature along the road traverses in December 2000: (a) shows mean diurnal variation and (b–e) are variations at 09:00, 12:00 15:00 and 18:00 h, respectively).

similar to the pattern obtained in cities in other countries such as Dhaka, Bangladesh (Hossain and Nooruddin, 1994), Brazil (Maitelli et al, 2002) and Beijing, China (Svensson and Tarvainen, 2004). From this study, urbanization played a significant role in the distribution of temperature and relative humidity. The Motor Park and commercial areas have higher mean average temperature than the outskirts. These two areas also exhibited lower relative humidity. Similar patterns of temperature and relative humidity were demonstrated in the temperature

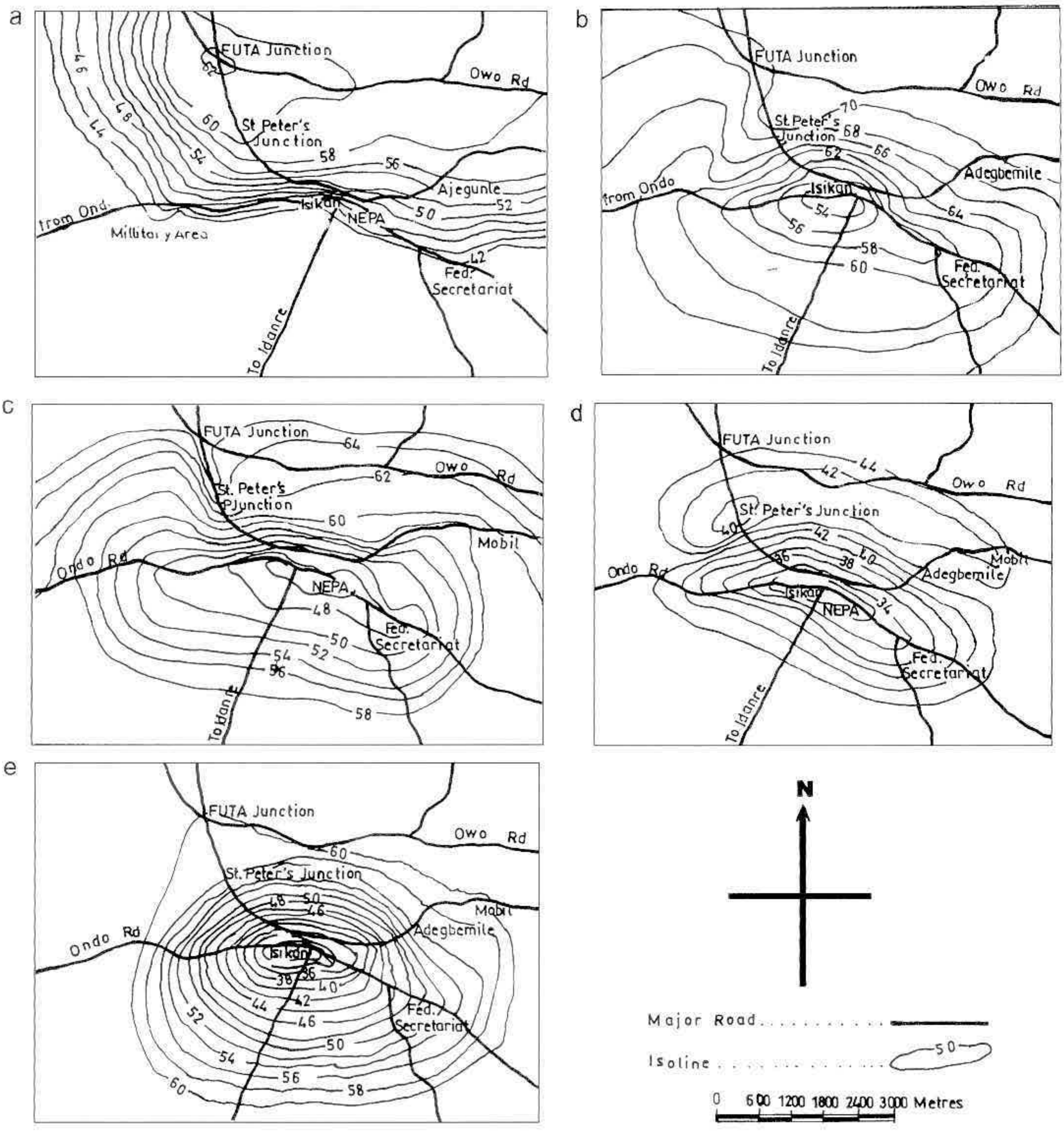


Fig. 6. Relative humidity along the road traverses in December 2000: (a) shows mean diurnal variation and (b–e) are variations at 09:00, 12:00 15:00 and 18:00 h, respectively).

changes along the two major roads, which traversed the city. The areas at the interior of the city were generally warmer than the exterior. Oke (1987) suggested that the structure and composition of the urban canopy and the thermal properties of urban construction materials are the main factors that determine the urban–rural thermal contrasts. The fact that urban areas tend to be warmer has been well studied and documented (e.g. Oke, 1982; Arnfield, 2003; Samuels, 2004; Svensson and Tarvainen, 2004).

It is also known that the higher temperatures are often associated with higher density urban dwellings, where it forms an urban heat island. The daytime 'heat island' intensity at the time of this study is within the ranges

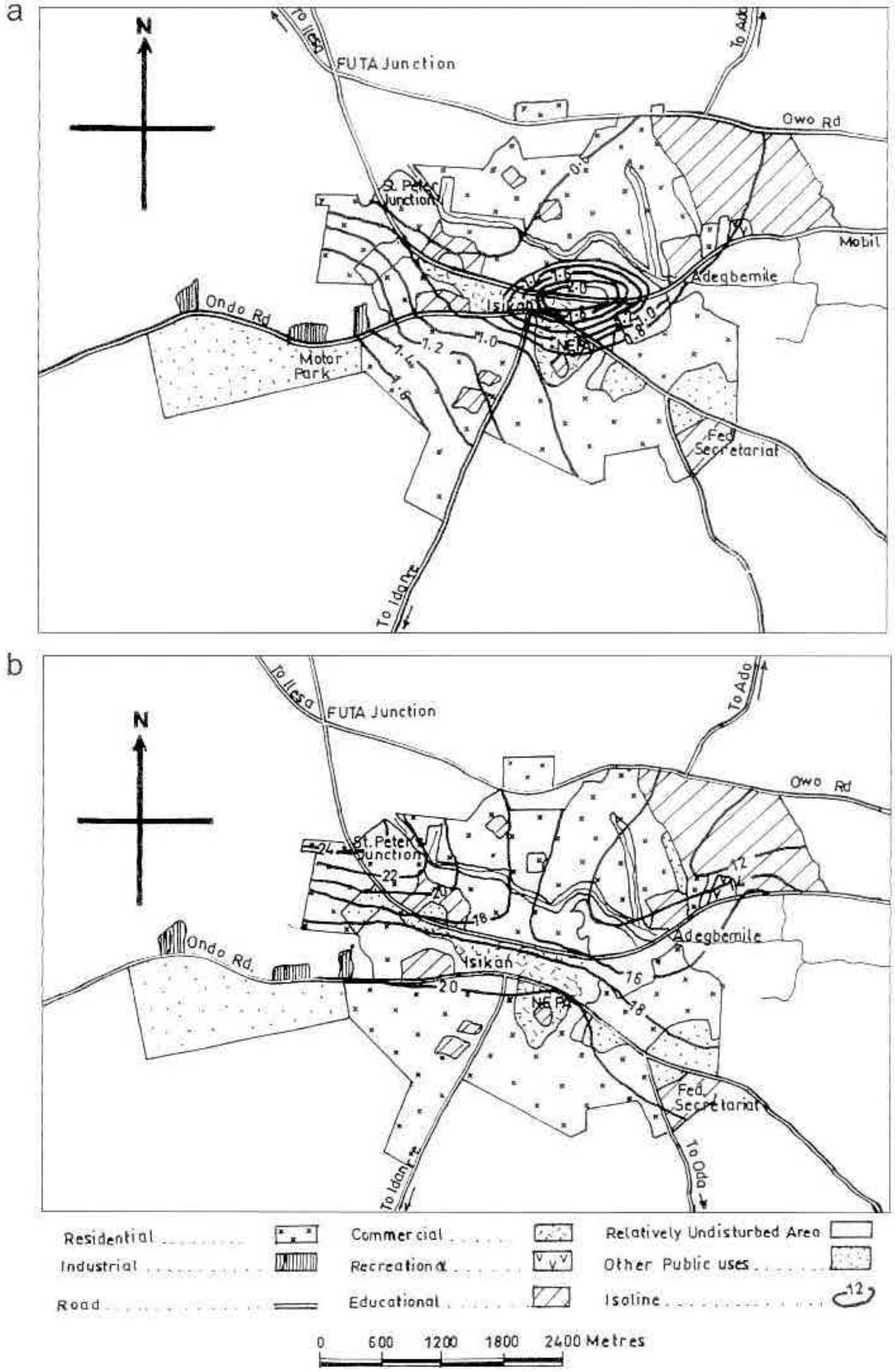


Fig. 7. Generalized isotherms using the average temperature to show the pattern of UHI and RH field in Akure.

reported of bigger and older cities in Nigeria (e.g. Ibadan; 3–5 °C (Adebayo, 1990), Benin City: 4 °C (Omogbai, 1985)). The small magnitude of heat intensity in this study may be explained by the phenomenon at the daytime hours in which this study was conducted (e.g. Oke, 1982; McGregor and Nieuwolt, 1998; Runnals and Oke, 2000).

The main 'heat island' characteristics described in some studies outside this environment (Voogt, 2002) were also observed in Akure at the time of this study. The CBD contains closely spaced buildings, which ranged from 2 to 4 storeys and open market spaces where over 20,000 people trade daily. These form the thermal sinks and urban hot spots in the study area. Such properties have been observed to be capable of absorbing and emitting heat (Bornstein and Lin, 1999; Samuels, 2004).

The results of the one-way ANOVA performed separately on each land use type suggested that the mean temperature did not vary significantly at different locations within the town. Relative humidity and vapour pressure did show, however. Generally, there was a relative uniformity in the mean temperature in Akure while the distribution of the relative humidity suggested local influence. For example, less humid surfaces were obtained in the commercial areas and Motor Park. This result is similar to that obtained in the two cities of Cuiaba and Verzea Grande in Brazil (Maitelli et al., 2002) in which the temperature trend was found to be inversely proportional to the relative air humidity. On the other hand, the results of the PCA showed that the measures of absolute humidity (vapour pressure and dew point temperature) are not patterned by temperature influence. The appearance of the minimum temperature, vapour pressure and dew point temperature in the PCA cluster confirms Adebayo's (1991) hypothesis that minimum temperature may have been a better indicator of urban heat island in southwestern Nigeria. It may also suggest that the people may be discomfited physiologically during some period of the year. Although the scope of this study did not include this, it may form an hypothesis for subsequent study in the region.

In addition, diurnal variation existed in both the temperature and relative humidity fields within the city. The trend is similar to those obtained in Dhaka, Bangladesh (Samendra and Ayesha, 1994) where 15:00 h was the warmest. While average temperature has increased from 09:00 to 15:00 and declined towards 18:00 LST, its relative humidity has decreased. The trends observed suggest that the temperatures in Akure have not significantly changed in the last two decades. The explanation for this is not obvious but theoretical speculation may suggest that effect of urban growth on temperature in this area is probably insignificant. It may also be due to the location of the monitoring network and the source area of their measurements. The true explanation of this will, however, not be known until the temperature data before Akure became an administrative capital city in 1976 is analysed. Nevertheless, the observed relative stability of temperature variables does not imply that urbanization has no impact on temperature. Relative humidity has, on the other hand, reduced significantly in the past two decades, a phenomenon that may be explained by the increasing trend of built up areas in Akure.

5. Conclusions

The study has shown that Akure experiences a daytime urban heat island. Its relative humidity has also significantly declined from 1980 to 2001. The mean temperature and relative humidity are influenced by commercial activities in the CBD and urban growth over the years, suggesting an influence of urbanization. It has equally upheld the hypothesis that relates minimum temperature to UHI and perhaps physiological stress in the tropics. Equipment for collecting some data, especially nocturnal weather data has been shown as a bane of some meteorological researches in Nigeria. It is one of the factors that have limited the scope of this study.

Akure is typical of most medium size state capitals, in developing nations, to which the applicability of these findings is anticipated. However, most urban settlements in Nigeria lack reasonable number of meteorological stations, a condition that has hindered investigations into the urban climate change. Inadequate and frivolous data from poorly monitored few existing meteorological stations may have also discouraged long-term studies in urban climatology in most Nigerian cities (Ayoade, 1993). Nonetheless, it is hoped that this study will adequately fill a gap.

Acknowledgements

The authors wish to give thanks to Dr. (Mrs.) I.O. Adelekan, Department of Geography, University of Ibadan, who supervised the thesis that forms the foundation of this study; Adeniyi. B. Eludoyin, who assisted in sampling; J.O. Faleye and Omopariola (Mrs.) of the Department of English Language and Literature, EAC, Ibadan, and Abiodun Emmanuel of the Department of Geography, Obafemi Awolowo University, Ile-Ife, for editing the manuscript at the early stage, and the anonymous reviewers for raising issues and making suggestions that have significantly improved the paper.

References

Adebayo, Y.R., 1990. Aspects of the variation in some characteristics of radiation budget within the urban canopy of Ibadan. Atmospheric Environment 24B (1), 9–17.

Adebayo, Y.R., 1991. Heat island" in a humid tropical city and its relationship with potential evaporation. Theoretical and Applied Climatology 43 (3).

Arnfield, A.J., 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. International Journal of Climatology 23, 1–26.

Ayoade, J.O., 1993. Introduction to Climatology for the Tropics. Spectrum Books Ltd., Ibadan, Nigeria.

Bornstein, R., Lin, Q., 1999. Urban convergence zone influences on convective storms and air quality. In: Proceedings of Congress of Biometeorology and International Conference on Urban Climate. WMO, Sydney.

Givoni, B., 1991. Impact of planted areas on urban environmental quality: a review. Atmospheric Environment 25B (3), 289–299.

Hajat, S., Kovats, R., Atkinson, R.W., Harner, A., 2002. Impact of hot temperatures on death in London: a time series approach. Journal of Epidemiology and Community 56, 367–372.

Hossain, M.E., Nooruddin, M., 1994. Some aspects of urban climates of Dhaka city. In: Reports of the Technical Conference on Tropical Urban Climates, March 28–April 2, 1993, Dhaka, Bangladesh.

Kalkstein, L.S., Greene, J.S., 1997. An evaluation of climate/mortality relationships in large US cities and the possible impacts of a climate change. Environmental Health Perspective 105, 84–93.

Kilbourne, E.M., 1989. Heat waves In: Gregg, M.B (Ed), The Public Health Consequences of Disasters, CDC. US Department of Health and Human services, Atlanta, pp. 51–61.

Maitelli, G.T., Souza, S.C., de Pinho, J.G., 2002. The magnitude of urban heat island in the tropical continental areas in Brazil, Department of Geography, University of Mato Grosso, Brazil, pp. 1–4.

McGregor, G.R., Nieuwolt, S., 1998. Tropical Climatology. 2nd ed., John Wiley and Sons Ltd.

Nordli, P.O., 2001. Reconstruction of nineteenth century summer temperatures in Norway by proxy data from farmers' diaries. Climate Change 48, 201–218.

Oguntoyinbo, J.S., 1982. Climate characteristics. In: Filani, M.O. (Ed.), Ibadan Regions. Department of Geography, University of Ibadan, Nigeria.

Oke, T.R., 1982. The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society 108, 1–24.

Oke, T.R., 1987. Boundary Layer Climates. Methuen, London.

Olowu, D., 1985. Local government and urban development in Nigeria. In: Abiodun, J.O. (Ed.), Urban and Regional Planning Problems in Nigeria. University of Ife Press Ltd., Nigeria, pp. 272–298.

Omogbai, B.E., 1985. Some aspects of urban climates of Benin City, Nigeria. Unpublished M.Sc. Thesis, University of Ibadan, Nigeria.

Onakerhoraye, A.G., Omuta, G.E., 1994. Urban Planning Systems and Planning for Africa, Benin Social Science Series for Africa, University of Benin, Nigeria.

Runnals, K.E., Oke, T.R., 2000. Dynamics and control of the near-surface heat Island of Vancouver, British Columbia. Physical Geography 21 (4), 283–304.

Samendra, K., Ayesha, K., 1994. The variability and probability extremes of some climatic elements over Dhaka. Reports of the Technical Conference on Tropical Urban Climates, March 28–April 2, 1993, Dhaka, Bangladesh.

Samuels, R., 2004. Urban Heat Islands. Australian House of Representative Standing Committee on Environment and Heritage Sustainable Cities 2025 Enquiry.

Santamouris, M., 2001. The canyon effect. In: Santamouris, M. (Ed.), Energy and Climate in the Urban Built Environment. James and James, London.

Svensson, D., Tarvainen, L., 2004. The Past and Present Urban Heat Island of Beijing. B416 Projektarbete Goteborg, Earth Sciences, Goteborg University, Sweden.

United Nations Environment Programme, 2002. Global Environment Outlook 3 〈www.unep.org/GEO/geo3/index.htm〉, 2004-03-03.

Voogt, J.A., 2002. Urban heat island. In: Munn, T. (Ed.), Encyclopedia of Global Change. Wiley, New York, pp. 660–666.

Zar, J.H., 1992. Biostatical Analysis, second ed. Prentice-Hall, Eaglewood Cliffs, NJ.

ELSEVIER

Journal of Environmental Management 87 (2008) 106–116

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Geological and technical characterisation of Iscehisar (Afyon-Turkey) marble deposits and the impact of marble waste on environmental pollution

Mustafa Yavuz Çelik[a,*], Eyüp Sabah[b]

[a] Afyon Kocatepe University, Afyon Vocational High School, 03100 Afyon, Turkey
[b] Afyon Kocatepe University, Afyon Engineering Faculty, 03100, Afyon, Turkey

Received 9 March 2006; received in revised form 24 November 2006; accepted 4 January 2007
Available online 8 March 2007

Abstract

Turkey, due to its location in the Alpine-Himalayan belt, has numerous marble deposits. More than 250 marble types with different colours and patterns have been produced from these deposits and one hundred of these are well known around the world. One such well-known marble type is Afyon-Iscehisar. Afyon-Iscehisar is Palaeozoic in age and has been quarried since the era of the Roman empire. The Afyon region is known as one of the most important marble production and processing centres in Turkey. The Afyon province, which possesses 3.5% of exploitable marble reserves (3,872,000,000 tonne) in Turkey, yields 9% of the total marble block production. The 409 marble processing plants in Afyon produce 19% of the total slab in Turkey. As a result of marble production activities, approximately 340,000 tonne of marble waste has accumulated in the area. While some of these unshaped marble blocks are re-used and returned to the economy, the majority are discarded. There are two waste marble storage fields located in the Afyon-Iscehisar region. All of the solid and fine-grained marble waste is stored in waste marble storage fields in Susuz Boğazı and in the nearby Iscehisar marble quarries. The ecological effects of the marble waste, which were once discharged everywhere and exhibited visual pollution, has now been reduced to a minimal level.

Keywords: Iscehisar-Afyon; Marble; Marble waste; Environmental pollution; Solid waste

1. Introduction

Many rich marble deposits in countries like Portugal, Spain, Italy, Greece, Turkey, Iran and Pakistan are located in Alpine-Himalayan belt. As a result of its geological location, Turkey possesses very rich, natural stone reserves in various colours and patterns. There are approximately 3,872,000,000 m^3 of exploitable marble reserves of Turkey (DPT, 2001) as summarised in Table 1.

The Turkish Natural Stone Industry has proved its consistency and continuity in the commercial arena by supplying over 250 natural stone types in rich colours and patterns. Approximately one hundred of these are well known in the international market and are regularly demanded. Some of these well-known commercial marbles include Afyon White, Afyon Sugar, Afyon Tiger skin, Supreme, Salome, Bordo Grizo (Eskisehir), Elazıg Cheery, Hazar Pink (Elazığ), Aegeen Bordeux, Ayhan Black, Milas Lilac, Mugla White (Muğla), Uşak Green, Uşak White (Uşak), Dove (Balıkesir), Vize Pink (Kırklareli).

There are approximately 700 Turkish marble quarries in operation, 90% of which are located in West Anatolia, Afyon, Balıkesir, Eskişehir, Uşak, Kütahya, Muğla, Bursa, Aydın and Izmir. The Turkish Natural Stone Industry has shown steady development and has played an important role in stone exporting. While the value of Turkish marble and natural stone exports was only 40 million USD in 1990, it reached 188.7 million USD in 2000, 300 million USD in 2002 and 626 million USD in 2004 (İMMİB, 2002, 2005).

The Iscehisar (Afyon) marble quarry is one of the largest, ancient, Anatolian marble quarries. During the

*Corresponding author. Tel.: +90 2722141949; fax: +90 2722135405.
E-mail address: mycelik@aku.edu.tr (M. Yavuz Çelik).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.004

Table 1
Exploitable marble reserves in Turkey (DPT, 2001)

Location	Exploitable reserve (m^3)
Balıkesir	1,300,000,000
Eskişehir	960,000,000
Uşak	500,000,000
Niğde	250,000,000
Kütahya	200,000,000
Muğla	181,000,000
Kırşehir	165,000,000
Afyon	135,000,000
Bursa	135,000,000
Kırklareli	33,500,000
Aydın	9,000,000
Ankara	2,000,000
İzmir	1,500,000
Total	3,872,000,000

Roman era, quarried marble was sent either as processed marble or as blocks to coastal Aegean and Mediterranean cities by river and sea. Therefore, numerous marble monuments and artwork made from Iscehisar-Afyon marble have been displayed in ancient cities in Turkey and Europe and in museums today (İleri et al., 2001). Iscehisar marbles were first quarried on the foundation of the ancient Dokimeion (now Iscehisar) settlement and have been well known for 2300 years. Iscehisar marbles were known in the world as "Synnada marbles" (now Şuhut) by the Romans and "Dokimeion marbles" in Anatolia (Gönç er, 1971). During the later stages of the operation of the Dokimeion quarries (235–236 BC), numerous white marble blocks, and half- or fully-processed marble artwork were left in waste rubble. The most interesting of these are human figures (Sodini, 2002).

In 2000, 110,350 m^3 of block marble and 650,000 m^3 of fragmented marble were quarried at Iscehisar-Afyon. During the same time period, the total block marble production of Turkey was 1,100,000 m^3, which means that approximately 9% of block marble production took place at Iscehisar-Afyon. Over 400 marble processing plants in the region use at least one gang-saw or one S/T (disc-cutter). The slab production in these plants was 6,611,000 m^2 in 2000, which accounts about 25% of the total slab production in Turkey (Kuşcu et al., 2001). These figures underscore the importance of Iscehisar-Afyon marble products in the Turkish marble industry.

2. The characteristic of Iscehisar-Afyon marbles

2.1. The geology of Iscehisar-Afyon region

There are two Iscehisar-Afyon marbles deposits located 25 km northeast of Afyon on the Ankara highway in the Iscehisar district (Fig. 1). The first deposit is located on Dangıç Tepe and occupies a $500 \times 1300\,m^2$ area. The second deposit near Bacakale occupies a $1000 \times 4500\,m^2$ area. The thickness of the marble deposits at Dangıç Tepe is 100 m, and that at Bacakale reaches 260 m (Güleç, 1973; Anil et al., 1996). Quarrying operations are concentrated in the northern part of the region (Fig. 5).

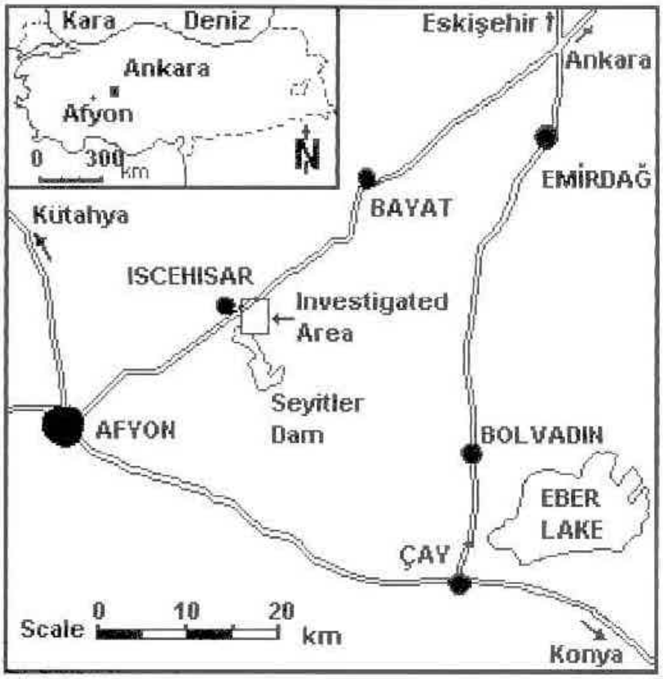


Fig. 1. Map of Iscehisar marble deposits.

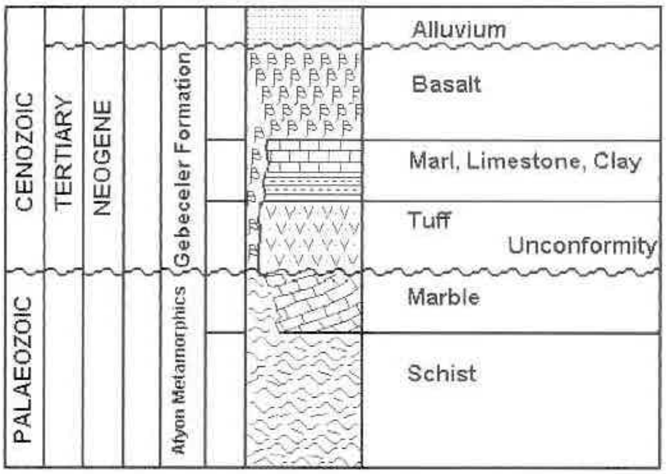


Fig. 2. Generalised columnar section of an Iscehisar marble deposit.

Palaeozoic aged rocks, referred to as Afyon metamorphites, are the oldest in the region. These rocks, usually sedimentary in origin, completed their metamorphic evolution in the pre-Mesozoic age (Tolluoğlu et al., 1997). Afyon metamorphites are usually brown and green coloured, with a highly folded albite–chlorite–muscovite–biotite quartz schist at the base. Occasionally, metaconglomerate marble and calc-schist appear in these schists in the form of large lenses and belts (Fig. 2). The thickness of schist is estimated to be 2000 m, and the upper contact is discordant to the Tertiary units (Metin et al., 1987).

Iscehisar marbles form in the upper layers of metamorphic rocks (Fig. 2). The marbles, which originate from limestone, are geologically indicated by different colours, particle sizes and mineral compositions. In this respect, marbles are divided into three units (Figs. 3 and 4). Afyon Grey and Tigerskin marbles form the bottom unit, Afyon Sugar and Afyon Kaymak (Cream) form the central unit, and the upper unit is formed by Afyon Violet marbles (Sümer et al., 1997). The bottom unit exhibits grey coloured marbles and white coloured, large calcite lenses that form unique patterns. The central unit, in contrast, is composed of fine calcite minerals with the dominant colour being white. The Afyon Sugar marbles are primarily white in colour with yellow veins. The upper unit marbles have purple and lilac colours. The cracks formed by tectonic movements were filled with yellow iron-oxide and white coloured veins of calcite minerals, and have breccias structure.

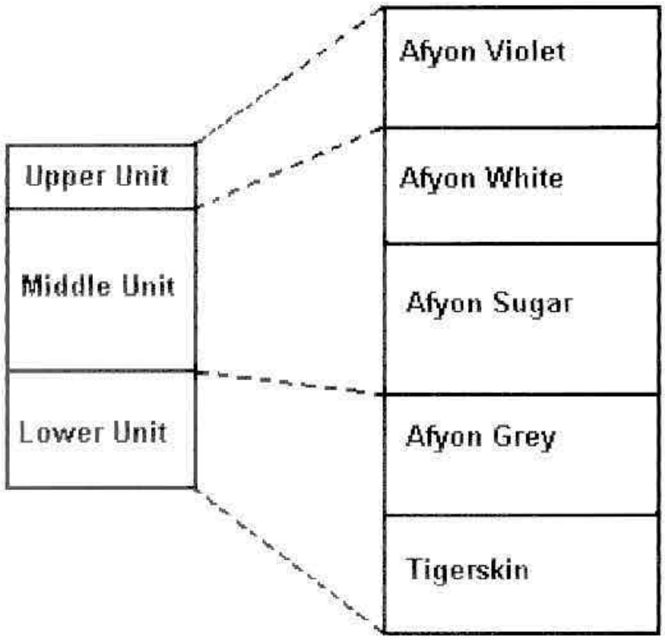


Fig. 3. Vertical stratigraphical column-section of Iscehisar-Afyon marbles and commercial names (Sümer et al., 1997).

The Gebeciler Formation is named after the Gebeciler district, located just outside the study area. After the Mesozoic age, this area was subjected to erosion processes. The Upper–Middle Miocene-aged crystalline metamorphic base of the Gebeceler Formation sedimented over during this period. Deposits in this area generally appear as a greyish to white colour, and are composed of fine to thick-layered conglomerate, sandstone, agglomerate, tuff, tuffite, marl, clay-limestone and silica-limestone units (Fig. 5) (Metin et al., 1987; Çelik and Sabah, 2002).

In the study area, tuffs occupy a vast area to the northwest of the marble deposits. Tuffs are dasitic with white and cream colour and indicate thick layers in the range of 50–150 m. Tuffs are composed of a mineral assemblage of various crystals including quartz, plagioclases, biotite lamella and opak particles with glass cement.

The Karakaya basalt formations are composed of basalt lava, which represents the youngest volcanism in the field, and they are positioned on the surface of the youngest fault zones. From the late Middle through the late Upper Miocene, widespread terrestrial volcanism of different stages dominated this region. Besang et al. (1997) used the K/Ar method to radiometrically date these volcanic rocks and concluded that their ages range from 14.75 million years to 8 million years (Çevikbaş et al., 1988).

Alluviums represent the youngest units in the study area.

2.2. Mineralogy and petrography

Iscehisar marbles originate from limestone rocks that metamorphosed under heat and pressure. Examination of the marbles under the microscope indicates that the main mineral component is calcite (more than 90%). The sizes of calcite crystals vary in the range of 0.2–0.5 mm, and between 1000 and 1700 crystals per cm^2 area. Medium-sized calcite minerals are tightly clamped with mosaic texture, and occasionally, granoblastic texture and

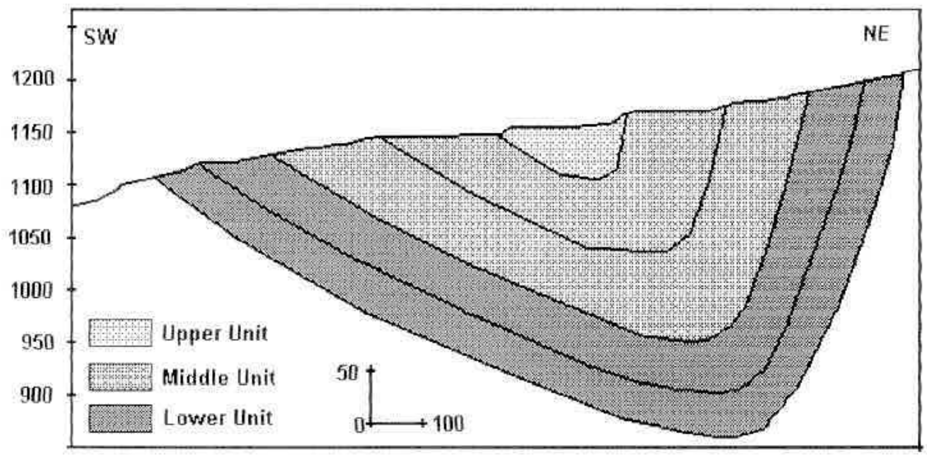


Fig. 4. Geological section of an Iscehisar-Afyon marble deposit (modified from Sümer et al., 1997).

polysynthetic twinning are observed. Chlorite, sericite, quartz and hematite are accessory minerals.

Iscehisar marbles have been divided into three units based upon their colour and particle size. The light and dark grey coloured units at the base consist of calcite and muscovite minerals. Besides these compounds, chloride, quartz and opak minerals are also present. Finely crystallised white and yellowish-white marbles in the central unit are composed of pure calcite minerals. Quartz constitutes approximately 0.2–0.8% of the central unit marbles. Purple and violet upper unit marbles are made mainly of calcite, and dolomite. muscovite; chlorite and quartz are also observed as secondary minerals.

2.3. Chemical characteristics

The results of the X-ray fluorescence spectrometry (XRF) chemical analysis of Iscehisar-Afyon marbles are summarised in Table 2. The main chemical composition of these marbles was found to be 55.00% CaO, 43.56% CO_3 and 1.43% other constituents. Therefore, calcium (or lime) is the main component of these marbles. Elemental analysis revealed that the most prevalent metallic element of Iscehisar-Afyon marbles is Mg (0.18%), with Fe, Mn, As and Pb being present in lesser proportions (Table 3). The physical, mechanical and technological specifications of Iscehisar-Afyon marbles as described by İMMİB (1990) are provided in Table 4.

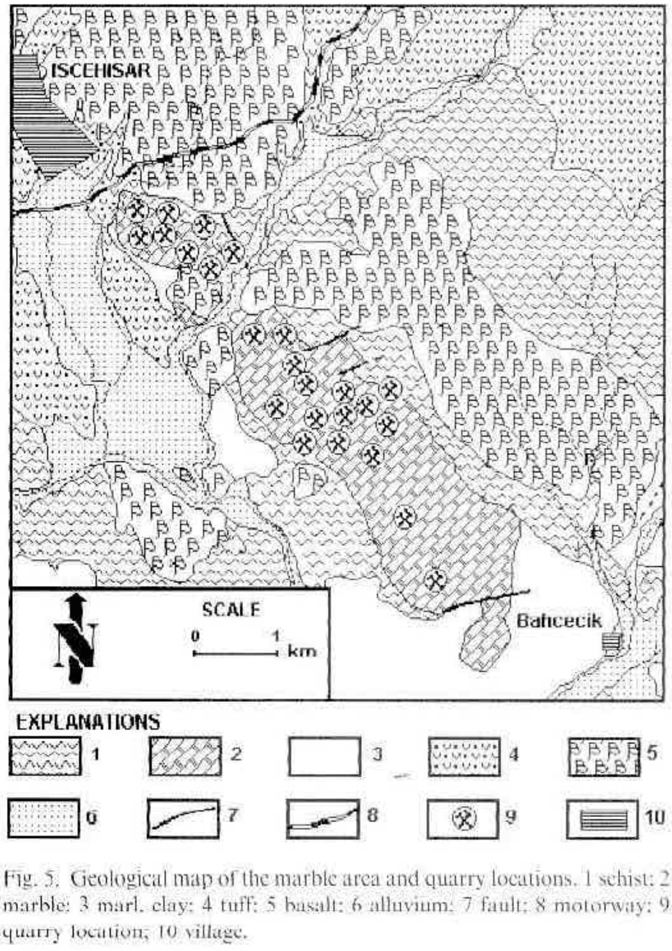


Fig. 5. Geological map of the marble area and quarry locations. 1 schist; 2 marble; 3 marl, clay; 4 tuff; 5 basalt; 6 alluvium; 7 fault; 8 motorway; 9 quarry location; 10 village.

3. Marble industry in the Iscehisar-Afyon region and generation of marble waste

Marble waste, generated by quarries and processing plants, can be divided by size into two main groups, particles and pieces. Types of marble waste pieces include Palladian, flat stone and slope rubble.

The proportion of marble discharged as waste during block production at the quarries is equal to 40–60% of the overall production volume (Çelik, 1996). According to 2001 statistical data, there were 24 quarries in operation in the Iscehisar-Afyon region, but the number of operating quarries fluctuates depending on market conditions. Reserve problems have arisen for marbles produced in regions where many companies operate for a long period in a very small area. The block productivity rate at Iscehisar marble quarries ranges between 10% and 20%. Despite this very low block productivity rate, quarrying operations still continue because even small, irregular-shaped, low-standard marble blocks can be processed and are in demand all around the world.

As shown in Fig. 6, the marble processing plants in the Afyon region are concentrated in three areas, e.g., Iscehisar and surroundings, Susuz Boğazı, and the Afyon industrial region. Around 409 marble processing plants of different sizes operate in Afyon and its surroundings. These companies process a variety of marbles from other parts of Turkey, as well as those quarried from the Iscehisar marble deposits. Many of these plants are equipped for

Table 3
Elemental analysis of Iscehisar-Afyon marbles

Elements	%
Mg	0.18
Fe	0.03
Pb	<0.005
As	<0.008
Mn	0.01

Table 2
The chemical analysis results of Iscehisar-Afyon marbles

	CaO	MgO	SiO_2	Al_2O_3	Fe_2O_3	Na_2O	K_2O	SO_3	CO_2
%	55.00	0.62	0.36	0.28	0.04	0.00	0.07	0.06	43.56

Table 4
Physical, mechanical and technological specifications of Iscehisar-Afyon marbles (İMMİB, 1990)

	Afyon grey and tigerskin	Afyon sugar and white
Hardness (Ʊ)	4	3
Unit volume weight (g/cm^3)	2.71	2.73
Density (g/cm^3)	2.73	2.75
Porosity (%)	0.2	0.2
Water absorp. at atmospheric pressure (%)	0.1	0.1
Compressive strength (kgf/cm^2)	648	701
Comp. strength after freezing (kgf/cm^2)	447	590
Strength to blow (kgf/cm^2)	11	23
Strength to bending (kgf/cm^2)	65	151
Ratio of fullness (%)	99.3	99.3
Degree of pores (%)	0.7	0.7
Average abrasion (cm^3/50 cm^2)	33.3	25.4
Average of tensile strength (kgf/cm^2)	46	39

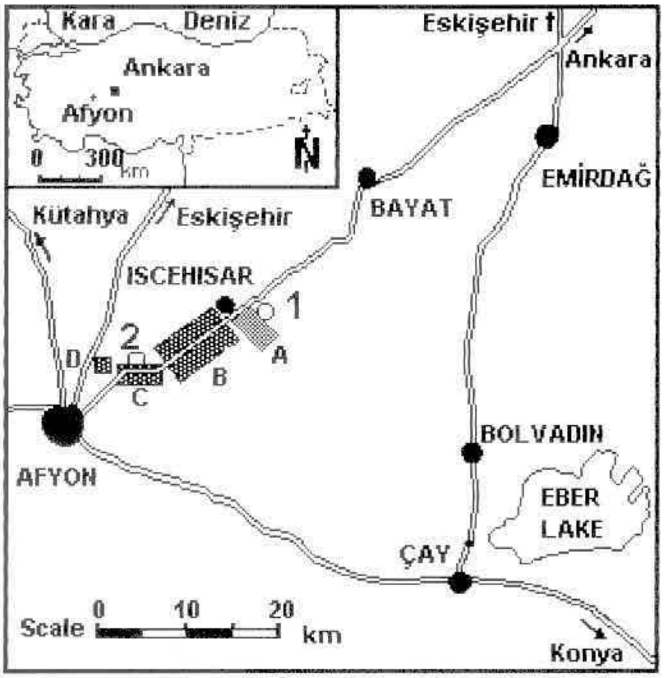


Fig. 6. Map of marble waste storage field locations in the Iscehisar-Afyon and Susuz Boğazı regions. (A) Iscehisar marble deposit and quarry; (B) Iscehisar area; (C) Susuz Boğazı area; (D) Afyon organised industrial area; (1) Iscehisar marble waste storage field; (2) Susuz Boğazı marble waste storage field.

slab production with simple machinery, like splitting machines. Over 200 companies have at least one gang-saw or one S/T (disc cutter).

The waste generation rate at marble processing plants is around 30–35% and this varies according to shape and kind of blocks being cut (Köse and Onargan, 1992). For example while the proportion of waste in the form of dust and palladien in 305 × 305 × 10 mm tile production per 1 m^3 block is 53%, it is reduced to 39% in 300 × 20 mm × free length floor tile production (Önenç, 2001). Therefore, when the thickness of the product is increased, the proportion of dust and palladien is reduced. Some of this waste has been used for the production of palladien, mosaic and square cement floor tile. However, as most of it cannot be sold, the unit cost for marble quarrying and processing plants is high. It is estimated that 340,000 tonne of marble waste are produced annually by the quarries and processing plants in the Iscehisar-Afyon region (Sabah and Çelik, 2001). In the long run, continued production at this rate could result in the accumulation of millions of tons marble waste in the area.

3.1. Environmental impact of marble waste

There are five major areas in which the environment is impacted by marble waste that is generated from extraction and processing work. These are (1) topography alteration, (2) land occupation, (3) surface and subterranean water degradation, (4) air pollution and (5) visual pollution. Visual pollution is the product of the overburden and production fragments generated at the quarrying sites. However, marble waste in the form of dust, palladien and flat stone can be utilised as by-products (Kaliampakos and Panagopoulos, 1995; Kaliampakos et al., 1996). In general, marble waste includes non-radioactive by-products. While the waste does not induce climate changes, it does destroy plant life.

Marbles usually contain the chemical compounds CaO, MgO, SiO_2, Al_2O_3, Fe_2O_3, Na_2O, TiO_2 and P_2O_5. During the cutting process, chemical compounds release no gases that contribute to global warming and climate changes. Water can be used in the cutting process to capture dust. This water does not produce wastewater because it is recycled after storage in sedimentation pools or tanks. Topographical changes and disruptions take place in small areas during marble quarrying operations; however, these areas will be rearranged by field improvement works at a later stage. Therefore, the environmental effect is not permanent. Furthermore no explosives or chemicals are used during quarrying, and there is no radioactive substances released during the quarrying and cutting processes (Güngör and Önenç, 1999).

3.1.1. Dust waste of marble

One of the major problems for the marble industry is the production of fine particles (<2 mm) while cutting marble. When a 1 m^3 marble block is cut into 2 cm thick slabs, the proportion of fine particle production is approximately 25%. This proportion increases with 1 cm thick products and decreases with 3 cm products (Kun, 2000). The majority of the marble processing plants in the Afyon region use sedimentation ponds to separate fine particle from water. Few marble processing plants have wastewater treatment facilities.

Unless it is stored properly, fine particles can cause more pollution than other forms of marble waste. The fine particles can be easily dispersed after losing humidity, under some atmospheric conditions, such as wind and rain. The white dust particles usually contain $CaCO_3$ and thus can cause visual pollution. In order to ensure that marble producers discharge of and store marble waste in appropriate locations in the region, the Union of Marble Producers, Governorships of Afyon Province and the Municipality of Afyon jointly created two marble waste storage facilities around the Susuz Boğazı and Iscehisar Marble quarries (Figs. 6 and 7).

3.1.2. Utilisation of marble waste chunks

Slope rubbles are very small, amorphous, marble pieces generated during quarrying. They are stored in certain areas around quarries. Some of this waste is sent to crushing and sieving plants to produce marble mosaics that are classified as 3 different aggregate sizes. There are 14 mosaic production plants in the Iscehisar district, all of which use a crusher followed by a trommel sieve.

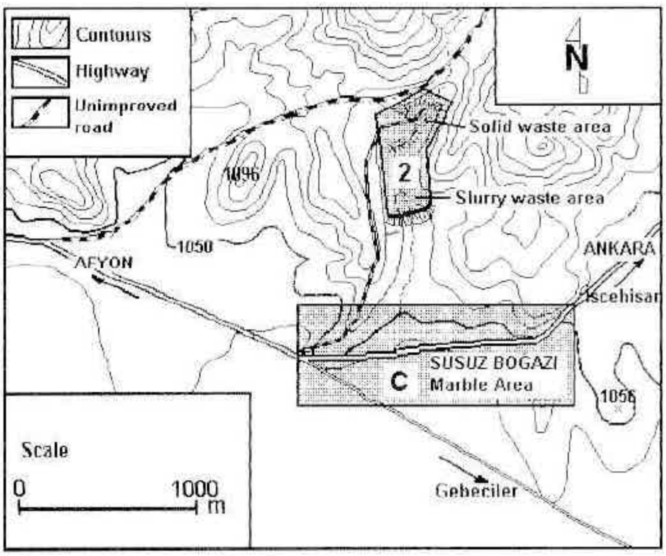


Fig. 7. Topographical map of marble waste storage fields in Iscehisar-Afyon and Susuz Boğazı regions.

Flats stones are cut with diamond discs from the bottoms and sides of the marble blocks from which no slabs can be produced. This type of waste has a considerable negative impact on the environment, as nicely coloured and patterned flat stones are purchased by small marble workshops to produce kitchen sinks, vases, candlesticks and marble souvenirs. The remaining parts are used in aggregate production.

Palllediens, on the other hand, are the amorphous remains left after cutting smooth marble slabs. The majority of pallediens are used in buildings as flooring material. Very small and unusable pallediens, however, are collected from processing plants and sent to marble waste storage fields (Figs. 8 and 9).

4. Marble waste storage field

The current marble storage field was brought into existence in 1995 by the joint efforts of the Union of Susuz Boğazı Marble Producers, Governorship of Afyon Province and the Municipality of Afyon in order to prevent storage of marble waste in inappropriate areas and visual pollution. The marble waste storage area is 2 km away from the Afyon-Ankara highway (Fig. 7). Twenty-eight marble processing plants discharge their liquid and solid marble waste in this storage field, which was originally a riverbed. A few marble processing plants in the region have water treatment facilities. The majority of processing plants separate liquid from solid wastes in sedimentation ponds, which creates a larger amount of liquid marble wastes than water treatment facilities.

The geology of the marble waste storage fields and its surroundings are shown in Fig. 9. The Susuz Boğazı field is located on the Palaeozoic-age Afyon metamorphites and the Neogene-age Gebeceler Formation, which is composed of marl and limestone. Nearly the entire surface of the storage field is covered by clayey soil with a thickness ranging between 0.5 and 1.0 m. Although faults exist in the area, there are none known to be in the marble storage field. Generally, solid and relatively dry water treatment wastes are dumped in the area to the north of the storage field (Fig. 10). In the southern section of the storage field, slurry from the sedimentation ponds is dumped. There is a 2 m-high clayey soil barrier to hold the liquid marble wastes.

While clay and soils have a high cation exchange capacity and can absorb a high proportion of heavy metals and cations, such as Ca, Mg, K and Na, the soils are not as effective as in holding pollutants like Cl (Gönüllü et al., 1988). Although, no surface improvement work has been carried out, since the particle size of the slurry is less than 80 μm (Fig. 11), it is later consolidated as a result of accumulation. The waste in the water does not completely sink to the ground, and much of it remains on the surface. As the water on the surface evaporates, then the liquid wastes solidify. Meanwhile, relatively wet marble waste, which is subjected to rain and snow will sink down into the ground over time. The waste marble storage fields fulfill an

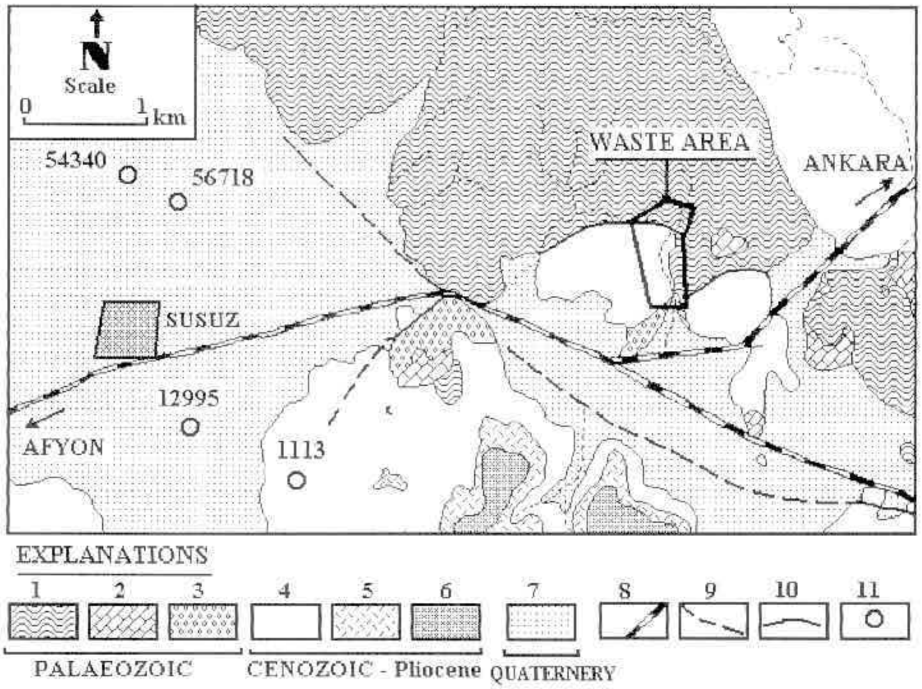


Fig. 8. Geological map of marble waste storage fields in the Iscehisar-Afyon and Susuz Boğazı regions and the location of underground water wells around the region. 1. Afyon metamorphites (schist); 2. Marble; 3. Meta-conglomerate; 4. Marl; 5. Tuff; 6. Limestone; 7. Alluvium; 8. Road; 9. Probable fault; 10. Fault; 11. Underground water well.

Fig. 9. Overall view of the marble slurry waste storage field at Susuz Boğazı.

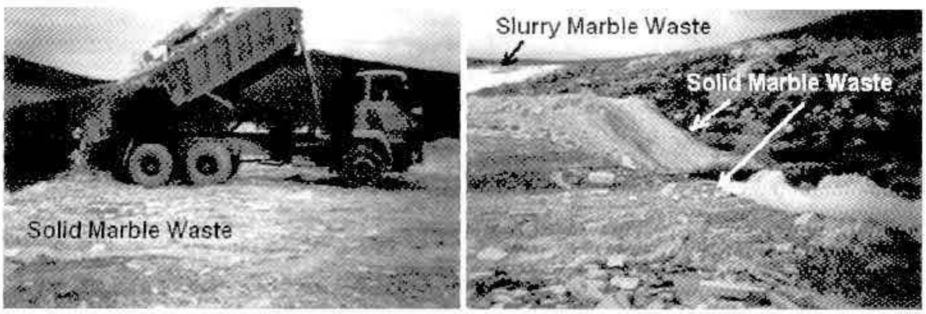


Fig. 10. Overall view of the solid marble waste storage field at Susuz Boğazı and storage of solid waste.

important function to prevent storage of waste in inappropriate areas and prevent visual pollution in the region.

4.1. Features and characterisation of marble waste

Four different samples were taken from the Susuz Boğazı marble waste storage field. One sample was a liquid sample taken from accumulation in the waste marble barrier and contained suspended solid particles, and the other three samples were slurry samples, which contained a high proportion of solid waste. The analysis of the wastewater sample, which has a high probability to leak into groundwater, indicated that the suspended solids amount to 34,800 mg/l. The analytical results of marble

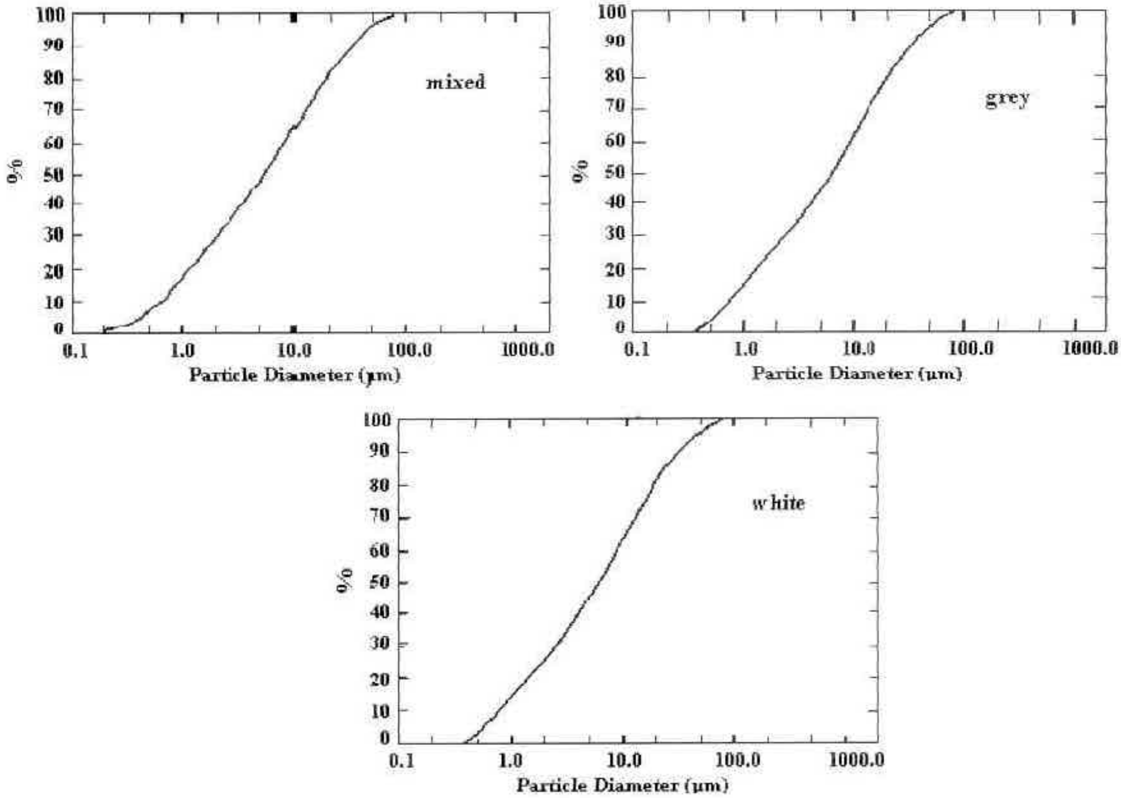


Fig. 11. Particle size distribution of dust waste of marble stored at wastewater storage fields.

Table 5
Chemical analyses results of marble wastewater by using ICP

	Ca	Mg	Na	K	Fe	Pb	Cd	Zn	H_2CO_3	Cr
ppm	5317.64	199.34	288.03	0.0	13.83	0.16	0.25	0.57	39,000	0.0

Table 6
Chemical analyses results of marble slurry waste

	Sample 1 (mixed)	Sample 2 (grey)	Sample 3 (white)
SiO_2	0.27	0.84	0.00
Al_2O_3	0.32	0.48	0.25
Fe_2O_3	0.05	0.09	0.00
CaO	55.14	54.02	55.07
MgO	0.29	0.40	0.25
SO_3	0.10	0.23	0.27
K_2O	0.07	0.07	0.04
Na_2O	0.00	0.00	0.00
CO_2	43.88	43.42	43.81

Table 7
Pollutant characteristics of leaked water (Gönüllü et al., 1988)

Parameter	Unit	General interval	Typical data
Ca	mg/l	200–3000	1000
Mg	mg/l	50–1500	250
Cl	mg/l	100–3000	500
pH	mg/l	5.3–8.5	6.0
Total hardness $CaCO_3$	mg/l	300–10,000	3500
Total solid materials in suspension		200–1000	500

wastewater, carried out by means of inductively coupled plasma (ICP) spectrometry, are shown in Table 5. The proportion of solid waste from the mixed grey and white coloured slurry samples were 79%, 77% and 77%, respectively, and chemical analysis results of these samples are given in Table 6.

The pollutant values of leakage water are given in Table 7. The dilution of leaked wastewater depends on the distance it has travelled to reach groundwater resources. This is affected by factors such as dispersion, adsorption and chemical conversions, which depend on the flow rate of underground water. In this respect, while the amount of Ca in liquid marble waste at the waste storage area was 5317 mg/l, it did not exceed 17.48 mg/l in groundwater (Table 8).

The results of particle size analysis of marble wastewater, which contains high proportion of solid materials (mixed, grey and white), using the Malvern Master Sizer are given

Table 8
The analysis of the well waters drilled by DSI for irrigation purposes (DSI 1970, 1999, 2001)

	Well number	DSİ 1113	DSİ 12995	DSİ 54340	DSİ 56718
	Year	1970	1970	1999	2001
	pH	7.20	7.40	6.30	7.00
	EC (µm/cm)	—	—	2474	2600
Cation (mg/l)	Na	1.03	2.65	11.50	14.55
	K	0.62	0.34	0.80	0.53
	Ca+Mg	5.50	6.90	14.28	17.48
Anions (mg/l)	CO_3	0.00	0.00	0.00	0.00
	HCO_3	3.82	5.25	13.81	17.20
	Cl	1.12	0.80	6.31	9.21
	SO_4	2.21	3.57	6.46	6.22

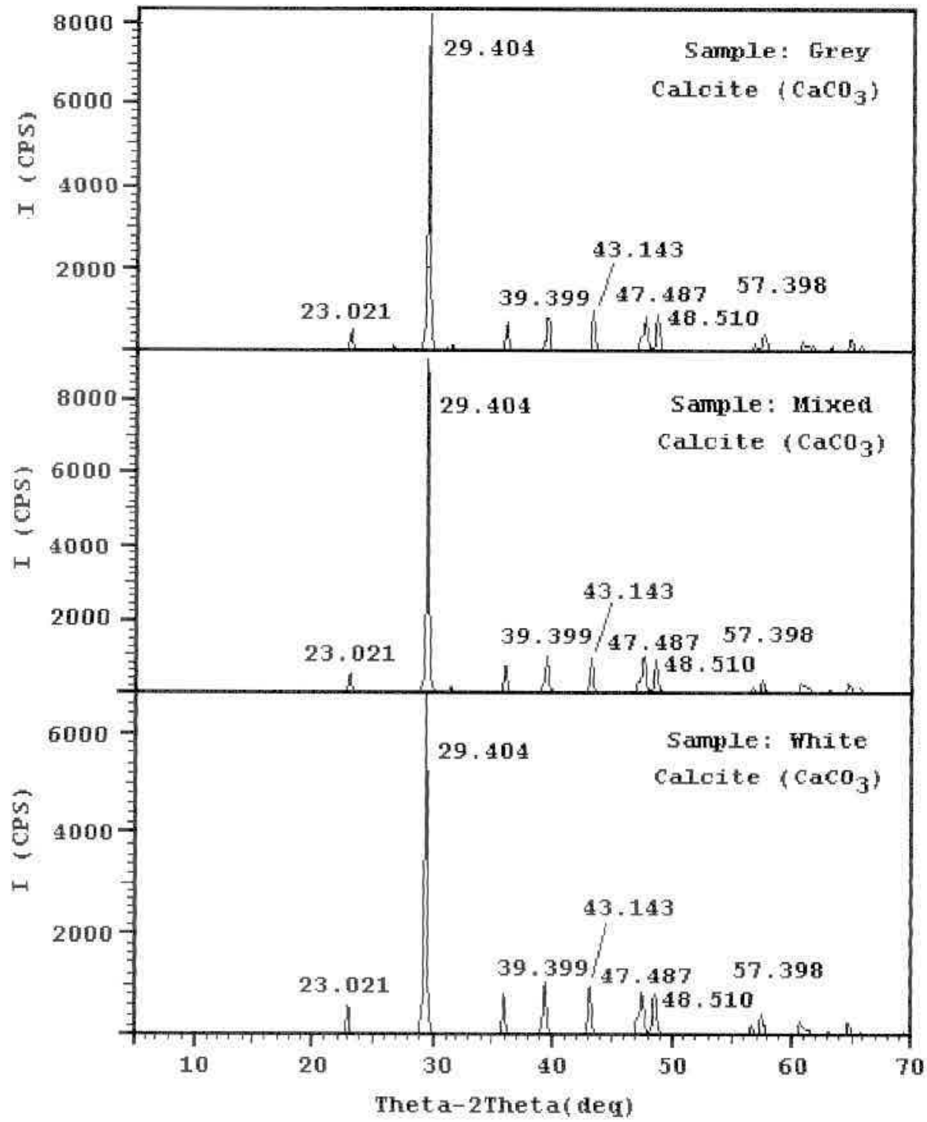


Fig. 12. XRD analyses of mixed, grey and white coloured marble waste samples.

in Fig. 11. Eighty per cent of the particles were found to be less than 20 μm in diameter.

X-ray diffraction (XRD) analyses on three samples taken from the Susuz Boğazı waste marble storage field where the solid material proportion was considerably high revealed that calcite mineral was the main component for mixed, grey and white coloured samples; no other minerals were present at detectable levels (Fig. 12).

4.2. The effects of marble waste on groundwater

The Susuz Boğazı waste marble storage field is located on a dry riverbed. The field has a wavy surface topographical profile, and the south and west areas of the storage field are plains covered with alluvium. The Susuz Boğazı marble processing plants are situated on the flat area to the south. The Susuz village is situated 4 km west of the storage field on a flat area where several underground water wells were opened by the DSI (State Waterworks Department) for irrigation purposes. All of these wells were drilled inside alluvium with depths ranging between 50 and 150 m. The nearest wells to the marble waste storage field are those numbered 1113, 12.995, 54.340 and 56,718, located to the north and west of the Susuz Village. These wells were drilled in 1970, 1972, 1999 and 2001, respectively. Lithologies of the wells numbered 54.340 and 56.718 are represented graphically in Fig. 13, and the chemical analysis results of samples taken from these wells are summarised in Table 8. According to the analytical results, there have been considerable increases in the anion and cation amounts in the past 30 years. While the amounts of Ca and Mg were 5.50 and 6.90 mg/l in 1970, they have increased to 14.28 and 17.48 mg/l, respectively. The amount of Cl, an important indicator of water quality, increased from 1.12 and 0.80 mg/l to 6.31 and 9.21 mg/l. These data indicate that underground pollution has increased due to the marble industry in the region.

5. Conclusions

The Iscehisar-Afyon region is one of the most important centres for the marble industry. These marbles are well known in Turkey and throughout the world, and have been used since ancient times. The Iscehisar-Afyon region has 24 marble quarries and 400 marble processing plants. During marble quarrying and processing, dust and pieces of marble wastes are generated. Some of this waste has been used locally, and two marble waste storage fields are located in Susuz Boğazı and Iscehisar.

The marble waste storage fields examined in this study are categorised as "open solid waste storage fields," and they have not been subjected to any ground treatments. Although the water analysis results taken from the underground wells 4 km away from the Susuz Boğazı storage field indicate a significant increase in the amount of anions and cations in the underground water, these levels are not high enough to restrict water use. Most of the waste leakage into the groundwater is prevented. In the future, increasing marble production and marble waste will create new storage fields and contribute to increased underground water pollution. By organising and controlling storage, waste leakage into groundwater can be prevented or reduced to a minimum level. Therefore, any new storage fields constructed in the coming years should undergo a careful ground treatment.

As a result of major development trends in the marble industry, there has been a significant increase in the number of marble quarries and processing plants.

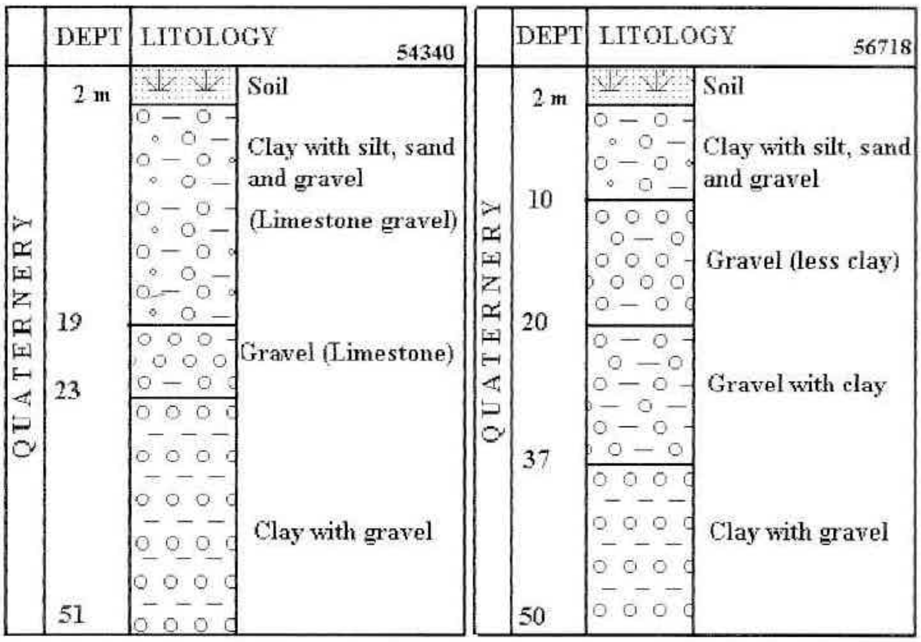


Fig. 13. Logs of the wells numbered 56,718 and 54,340 drilled closest to the Iscehisar-Afyon Susuz Boğazı waste marble storage field.

Consequently, parallel to this development trend, the amount of marble waste has increased. Increases in marble waste are inevitable as long as marble quarrying and processing continues. When one considers that the establishment and management of marble quarries and processing plants require a high investment cost, marble waste products are economical losses. Research on possible uses of marble waste products in various sectors as raw material and additives should contribute to both the economy of the marble industry and the environment.

References

Anil, M., Kiliç, A., Kahraman, S., Sandikçi, M., Akyildiz, M., 1996. İscehisar (Afyon) Mermerlerinin Petrografik ve Mekanik Özellikleri. K.T.Ü. Jeoloji Müh. Bölümü 30.Yıl Sempozyumu Bildirileri, Trabzon, pp. 436–447.

Besang, C., Eckhardt, F.J., Harre, W., Kreuzer, H., Müller, P., 1997. Radiometrische Altersbestimmungen An Neogenen Eruptivgesteinen Der Türkei. Geol. Jb. B 25, 3–36.

Çelik, M.Y., 1996. Mermer Artıklarının (Parça-Tozların) Değerlendirilmesi. A.K.Ü. Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, Afyon, 119pp.

Çelik, M.Y., Sabah, E., 2002. The geological and technical characterisation of Ömer-Gecek geothermal area and environmental impact assessment of geothermal heating system in Afyon, Turkey. Environmental Geology 41 (48), 942–953.

Çevikbaş, A., Ercan, T., Metin, S., 1988. Geology and regional distribution of neogene volcanics between Afyon-Şuhut. METU (Middle East Technical University). Journal of Pure and Applied Sciences 21 (1–3), 479–499.

DPT, 2001. Devlet Planlama Teşkilatı. Madencilik Özel İhtisas Komisyonu Raporu. Endüstriyel Hammaddeler Alt Komisyonu. Yapı Malzemeleri II Çalışma Grubu Raporu. Sekizinci Beş Yıllık Kalkınma Planı. Yayın No: DPT: 2616-ÖİK:627, Ankara, 191pp.

DSI, 1970. Akarçay Havzası Hidrojeolojik Etüt Raporu. Devlet Su İşleri Jeoteknik Hizmetler ve Yer Altı Suları Bölümü. DSİ Basımevi, Ankara, 64pp.

Gönçer, S., 1971. Afyon İli Tarihi. Cilt 1. Karınca Matbaacılık. İzmir, 400pp.

Gönüllü, M.T., Samsunlu, A., Baştürk, A., 1988. Açık Katı Atık Depolama Yerlerinde Oluşan Sızıntı Suların Yer Altı Sularını Kirletmesi. Ulusal I. Hidrojeoloji Simpozyumu, Ankara Üniversitesi fen Fakültesi Jeoloji Mühendisliği Bölümü, Ankara, pp. 229–235.

Güleç, K., 1973. Afyon Mermerlerinin Mühendislik Jeolojisi ve Fiziko Mekanik Özelliklerinin Ayrışma ile İlişkileri. Doktora tezi. İ.T.Ü. Maden Fak. İstanbul, 60pp.

Güngör, N., Önenç, D.İ., 1999. Mermer İşletmeciliğinde "Atık" mı Yoksa "Artık" mı Var!. I.Batı Anadolu Hammadde Kaynakları Sempozyumu, 8–14 Mart, İzmir, pp. 63–69.

İleri, S., Kun, N., Türkmen, F., 2001. Antik Dönemde Anadolu Mermerciliğinde Jeolojinin Önemi. Mermer Doğal Taş Sektörünün Dergisi Yıl 6, Sayı 28, Mart 2001, pp. 16–20.

İMMİB, 2002. Türkiye Mermer Endüstrisi (2001 Yılı). Başbakanlık Dış Ticaret Müsteşarlığı İstanbul Maden İhracatçıları Birliği Bülteni, İstanbul.

İMMİB, 1990. İstanbul Maden İhracatçıları Birliği Mermer Katalogu, İstanbul.

İMMİB, 2005. 〈http://www.immib.org.tr〉.

Kaliampakos, D., Panagopoulos, K., 1995. Recovery of economic value from the marble wastes of Thassos Island, Greece. In: Recovery, Recycling, Re-integration. EMPA—Swiss Federal Institute of Technology—Chalmer University of Technology—Swedish National and Research Institute, Geneva, Switzerland, International Congress Proceedings, vol. II, pp. 187–192.

Kaliampakos, D., Moutsatsou, A., Skotaras, J., 1996. Comparative evaluation of the wastes of the marble quarries, as paint fillers. Transactions of the Institution of Mining and Metallurgy, Section C 105, 37–42.

Köse, H., Onargan, T., 1992. Mermer Üretimi, Kesme Teknolojisi, Ekonomisi. Dokuz Eylül Üniversitesi Mühendislik-Mimarlık Fakültesi Yayınları No. 220, İzmir, 204pp.

Kun, N., 2000. Mermer Jeolojisi Ve Teknolojisi. Tezer Matbaası, İzmir, 149pp.

Kuşcu, M., Bağci, M., Yildiz, A., 2001. Afyon Mermer Endüstrisi ve Türkiye Mermer Sektöründeki Konumu. Türkiye III. In: Mermer Sempozyumu (Mersem'2001) Bildiriler Kitabı, 3–5 Mayıs 2001, Afyon, pp. 361–368.

Metin, S., Genç, Ş., Bulut, V., 1987. Afyon ve Dolayının Jeolojisi. MTA Rapor No. 8103, Ankara.

Önenç, D.İ., 2001. Tozlaşan Bloklar ve Umutlar, Mermer Doğal Taş Sektörünün Dergisi. Yıl 7, Sayı 30, Aralık 2001, pp. 66–68.

Sabah, E., Çelik, M.Y., 2001. İscehisar (Afyon) Mermer Artıklarının Hayvan Yemi Katkı Maddesi Olarak Kullanılabilirliğinin Araştırılması. Türkiye III. In: Mermer Sempozyumu (Mersem' 2001) Bildiriler Kitabı, 3–5 Mayıs 2001 Afyon, pp. 309–316.

Sodini, J.-P., 2002. Marble and stoneworking in Byzantium seventh–fifteenth centuries. In: Laiou, A.E. (Editor-in-Chief), The Economic History of Byzantium from the Seventh Through the Fifteenth Century. Dumbarton Oaks Research Library and Collection, Washington, DC, pp. 129–146.

Sümer, E.Ö., Tolluoğlu, A.Ü., Erkan, Y., 1997. Mermer Üretiminde Jeolojik Verilerin Önemi ve Afyon-Iscehisar Mermerlerinde Üç Boyutlu Modelleme Yöntemi. In: Türkiye II. Mermer Sempozyumu Bildiriler Kitabı 1–2 Mayıs 1997, Afyon, pp. 35–41.

Tolluoğlu, A.Ü., Erkan, Y., Sümer, E.Ö., Boyaci, M.N., Yavaş (Bektaş), F., 1997. Afyon Metasedimanter Grubunun Mesozoik Öncesi Metamorfik Evrimi. Türkiye Jeoloji Bülteni c.40, sayı 1, Şubat 1997, pp. 1–17.

ELSEVIER

Journal of Environmental Management 87 (2008) 117–124

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Land use change and population growth in the Morobe Province of Papua New Guinea between 1975 and 2000

Tine Ningal[a,b], A.E. Hartemink[a,*], A.K. Bregt[c]

[a]ISRIC—World Soil Information, P.O. Box 353, 6700 AJ Wageningen, The Netherlands

[b]Department of Surveying and Land Studies, PNG University of Technology, PMB, Lae, Morobe Province, Papua New Guinea

[c]Center for Geo-Information, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands

Received 9 February 2006; received in revised form 15 November 2006; accepted 4 January 2007

Available online 13 March 2007

Abstract

The relation between human population growth and land use change is much debated. Here we present a case study from Papua New Guinea where the population has increased from 2.3 million in 1975 to 5.2 million in 2000. Since 85% of the population relies on subsistence agriculture, population growth affects agricultural land use. We assessed land use change in the Morobe province (33,933 km^2) using topographic maps of 1975 and Landsat TM images of 1990 and 2000. Between 1975 and 2000, agricultural land use increased by 58% and population grew by 99%. Most new agricultural land was taken from primary forest and the forest area decreased from 9.8 ha person^{-1} in 1975 to 4.4 ha person^{-1} in 2000. Total population change and total land use change were strongly correlated. Most of the agricultural land use change occurred on Inceptisols in areas with high rainfall (>2500 mm year^{-1}) on moderate to very steep slopes (10–56%). Agricultural land use changes in logged-over areas were in the vicinity of populated places (villages), and in close proximity to road access. There was considerable variation between the districts but districts with higher population growth also had larger increases in agricultural areas. It is concluded that in the absence of improved farming systems the current trend of increased agriculture with rapid population growth is likely to continue.

Keywords: Land use change; Population growth; Agricultural expansion; Agro-ecological conditions; Papua New Guinea

1. Introduction

Global land use has significantly changed in the past decades. Historically, the driving force for most land use changes is population growth (Ramankutty et al., 2002) although there are several interacting factors involved (Lambin et al., 2001, 2003). At the global and supra-national scales, population growth is often used as a proxy for land use change (Kok, 2004) but at lower scales a set of complex drivers are important (Lambin et al., 2001). Land use change is mainly caused by human activities.

Objectives for land use change differ between the developed and developing countries. In developed countries, land use change is based on economic reasons such as large-scale farming or urban development and an increasing need to conserve biodiversity and environmental quality for current and future generations (Bouma et al., 1998), whereas in the developing countries, rapid population growth, poverty and the economic situation are the main driving forces (Lambin et al., 2003; Meertens et al., 1996; Ramankutty and Foley, 1999).

The need for increased food production results, amongst others, in the conversion of forest and grassland to cropland. A region or country's ability to supply food is determined by productive cropland, the ability to maintain crop yields, or the ability to purchase imported food. Globally, cropland decreased from 0.75 ha person^{-1} in 1900 to 0.35 ha person^{-1} in 1990 (Ramankutty et al., 2002) and most of the production increases in the past decades have resulted from higher crop yields (Greenland, 1997). There are, however, large differences between different countries and regions—both in the extent that land use has changed and human population and crop yields have increased.

*Corresponding author. Tel.: +31 317 471 713; fax: +31 317 471 700.
E-mail address: Alfred.Hartemink@wur.nl (A.E. Hartemink).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.006

Some areas in the world appear to be more affected by rapid land-cover change because they are studied more intensively (Lambin et al., 2003).

Papua New Guinea is a developing country where population has grown exponentially in the past decades. About 85% of the population lives in rural areas and depends on subsistence agriculture for their livelihoods. Forests cover more than 70% of the land area and 97% of the land is privately owned (Pat, 2003; Saunders, 1993). The average population density in the highlands region is 22 persons km^{-2}; whereas the national population density is 11 persons km^{-2}. A national study showed little relationship between population growth and agricultural land use change (McAlpine et al., 2001). Environmental and socio-economic conditions differ markedly between provinces and to obtain insight into land use changes at a lower scale, we selected the Morobe province to investigate changes in population in relation to land use change.

2. Study area, data and methods

2.1. Study area

The Morobe Province has nine districts and covers an area of 33,933 km^2, accounting for 7% of the total land area in Papua New Guinea (Fig. 1). The population was approximately 539,000 in 2000. A major part of the province is covered by primary forest. The topography ranges from sea level to over 4000 m a.s.l. and plate tectonics are active. The vast Markham Valley is dominated by grassland that spans from Lae city westward through Huon to Kaiapit district dividing the Saruwaged, and Finisterre mountains to the north and the Highlands mountain systems to the south. These mountains and hills comprise 77% of the total land area. In the lowlands, the climate is hot and humid with an average temperature of 30° and a mean annual rainfall of about 2900 mm $year^{-1}$. Wau, Lae, Siassi, parts of Menyamya, Huon and Finschhafen districts are some of the wettest areas with up to 5000 mm rain $year^{-1}$. Dominant soils are Humitropepts, Dystropepts, Troporthents and Rendolls (Bleeker, 1983). The provincial capital, Lae city, has 15% of the total population in the province.

2.2. Topographic maps and Landsat images

Topographic maps from 1975 and Landsat thematic mapper images of 1990 and 2000 were used to investigate changes in land use. One topographic map sheet covers an area of 55 × 55 km and 25 sheets covered the whole province. The map scale is 1:100,000 and spatially referenced to transverse mercator projection on Australian geodetic datum of 1966 (AGD66). The geometric accuracy is up to 40 m horizontally and 17 m vertically with contour intervals at 40 m. The topographic map sheets were scanned at 300 dpi scan resolution and georeferenced to AGD66, which results in a spatial resolution of 8.5 m.

The Landsat images fully covered the study area for both 1990 and 2000. Their overall spatial resolution was 28.5 m, however; the resolution for band 8 of ETM+ for 2000 was 14.25 m, which sharpens the other bands. Both images were

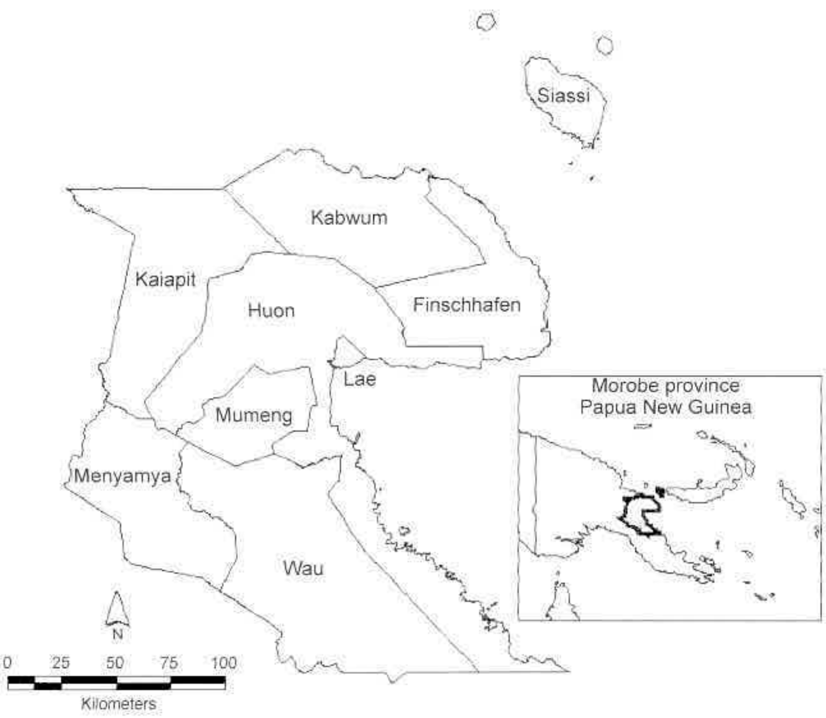


Fig. 1. Districts in the Morobe province of Papua New Guinea.

geometrically corrected and orthorectified by NASA to form seamless mosaics or tiles of 6° longitude and 5° latitude. The tiles were spatially referenced to universal transverse mercator projection of 1984 (UTM84). The spectral bands in both images were reduced to 2 (green), 4 (near infrared) and 7 (short wave infrared). Band 2 is useful for human-induced and vegetation differentiation; band 4 is useful for vegetation types determination and band 7 is sensitive to rock types. The images were reprojected to AGD66 and clipped to fit the study area. Both images contained less than 10% cloud cover along the higher altitude and mountainous areas which are mostly under forest. The areas under cloud cover were updated using a combination of Papua New Guinea resource information system (PNGRIS), forest information system (FIMS) and agriculture systems (AGSYS) datasets (see Section 2.3) to derive land use.

2.3. Auxiliary data

In addition to the topographic maps and Landsat images, we used various data sets to improve the assessment of land use change. Auxiliary data sets for this study were the PNGRIS, AGSYS, FIMS and population data. PNGRIS is a polygon-based natural resource data set with spatial units called resource mapping units (RMUs) in MapInfo and ESRI shape formats. RMUs are based on common set of geographic attributes which include landform, geology, climate and administrative boundaries (Bellamy and McAlpine, 1995). The RMU boundaries are derived from 1:500,000 base maps called tactical pilotage charts (TPC) which contain appropriate topographical information for mapping the distribution of natural resources. The department of agriculture and livestock (DAL) is the custodian of the PNGRIS database and other government departments and agencies acquire license to use. Census data points are overlaid with RMU polygons and through GIS spatial operations the RMU attribute table is updated with census data. The population in the RMU is assumed to be equally distributed.

AGSYS is a polygon-based agricultural systems data set of Papua New Guinea which is also under the custody of DAL. The mapping units are based on agricultural system on a common set of attributes such as crop types, cultivation intensity, cropping period, fallow type, fallow period, cash crop activity, soil fertility maintenance technique and other aspects related to agriculture (Bourke et al., 1998, 2002). AGSYS data were used for checking digitized land use and identifying land use that was under cloud cover on the satellite images.

FIMS is a polygon-based forest data set compiled from 1:100,000 topographic maps and contains forest data such as major forest species and forest zones of Papua New Guinea in Mapinfo format. FIMS contains land use as a land cover type and this was used to check land use and also for areas under cloud cover.

2.4. Assessing land use change

The land use classification for this study was based on the topographic maps of 1975 as follows: (i) agriculture, (ii) forest, (iii) grassland, (iv) plantation, (v) urban and (vi) water. The 1975 topographic maps were derived from aerial photographs at scale 1:100,000 through photogrammetry. These topographic maps show the six land use types for 1975. Land use for 1975 has been assessed by screen digitizing the scanned topographic images as background maps, similar to Li et al. (2004). This was preferred above adopting the existing 1975 land use map as a base map (McAlpine and Freyne, 2001). A copy of the digitized 1975 land use was used to update 1990 land use by using the 1990 Landsat image as background. Land use changes were assessed visually and edited to 1990 land use in ArcGIS. The procedure was repeated to assess land use changes for 2000 by making a copy of 1990 land use and the 2000 Landsat image as background.

The 3 band Landsat images when displayed in default true colour showed shades of green for vegetation, and lavender, magenta or pale pink for urban areas and bare soils. Water is shown black to dark blue and cloud cover shows white while cloud shadows as black. From the image, water, urban, grassland and forest could easily be distinguished and also plantation and agriculture covering larger areas (>0.5 ha) are easily identified, but areas smaller than 0.5 ha were difficult. To update land use boundaries for 1990 and 2000, image enhancements were used to recognize roads, built-up areas and plantations. Band combinations and spectral enhancements of image properties such as colour, tone, brightness, structure, location and association (Dekker, 2004; Lillesand and Kiefer, 2000) were used to aid land use boundary adjustments. Other topographic layers such as contours, rivers, roads, and villages were used to refine boundary determinations. The adjusted boundaries were overlaid with AGSYS, PNGRIS and FIMS layers to detect and resolve conflicts such as geometric overlaps, attribute mismatch, or boundary inconsistencies and edited in ArcMap.

The land use layers and RMU were overlaid through union GIS spatial operation. The composite of all layers contained data for querying and analyses. The attribute table of the composite layer was imported into MS Access and SQL queries were designed to extract information on population and agricultural land use changes by province, districts and for different environmental factors such as soil quality, slope, altitude and rainfall.

Calculating the spatial correlation surfaces between population change and transitions from one land use category to another (hereafter referred to as "from → to") was done using custom avenue script in ArcView. This script uses a moving window to compute surface of correlation coefficients between two grid themes. The size of the moving window was set to 16 km (156 cells or measurements) and the overlapping area was 5 km between windows. The percentage of maximum no-data for each moving window was 10%. The grid cell size was 1 × 1 km.

3. Results

3.1. General agricultural change

Agricultural land use increased considerably between 1975 and 2000 (Fig. 2). Most of the expansion occurred at the expense of primary forest, which decreased over the same period (Fig. 3). Large areas remain under forest in the Morobe Province. There is expansion in most land use types but agriculture increased more than others (Table 1). The annual gain in agricultural land use between 1975 and 1990 was 3%, compared to 0.9% between 1990 and 2000. Annual loss in forest was 0.8% between 1975 and 1990, but 0.4% between 1990 and 2000. Between 1975 and 2000, agriculture gained 15% from forest, which is the highest compared to gains in other land use types.

The population of Papua New Guinea more than doubled from 2.3 million in 1975 to 5.2 million in 2000. Average annual growth was 5.0%. The population of the Morobe Province grew from 270,700 to 539,400: an increase of 99% or 4% growth per year. The average agriculture area was 1.8 ha person^{-1} in 1975, but decreased to 1.5 ha person^{-1} in 2000. The forest area was more than halved, from 9.8 ha person^{-1} in 1975 to 4.4 ha person^{-1} in 2000. There was little change in grassland and plantations area when expressed in hectare per person.

3.2. Changes on correlation surfaces between population and land use

The relationship between population change and from→to land use changes were compared by running custom Avenue script in ArcView (Fig. 4). The correlation coefficient is scaled from −1 to 1. The results were classified in three intervals: −1 to 0.5 (strong negative correlation), −0.5 to 0.5 (no significant correlation), and 0.5 to 1 (strong positive correlation). Correlation between total population change and total land use change is strongly positive (64%) for Morobe Province. The correlation between total population and forest to agriculture change is 31%. The change between total population and forest to grassland was strong in certain districts. Total population change and grassland to agriculture change were positively correlated in highly populated areas.

3.3. Land use change and environmental conditions

Inceptisols are the dominant soils and cover about 54% of the Morobe province. We classified the soils using expert knowledge in which soil physical and chemical

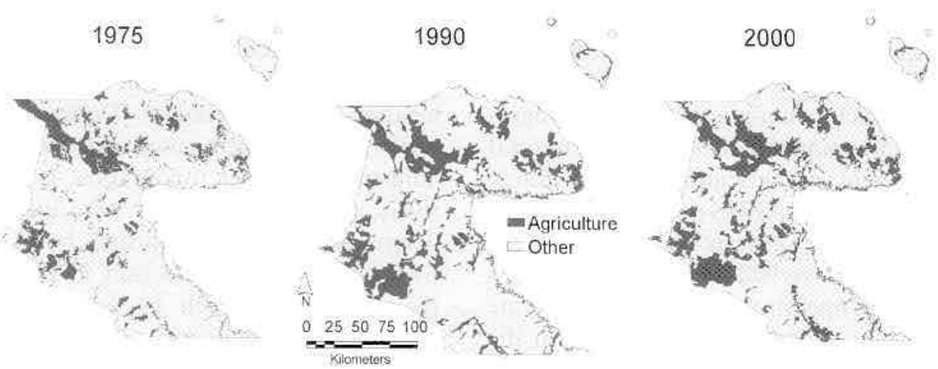


Fig. 2. Agricultural land use in 1975, 1990 and 2000 in the Morobe province of Papua New Guinea.

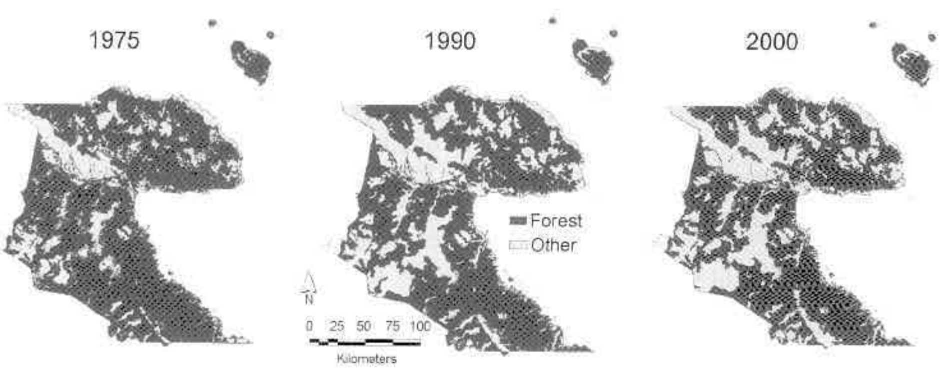


Fig. 3. Forest cover in 1975, 1990 and 2000 in the Morobe province of Papua New Guinea.

properties were classified in three categories (poor, moderate, good), and found that about 41% of the Morobe province has soils of good quality and 53% soils of moderate quality. More than 50% of the Morobe province has steep to very steep slopes; 55% of the province is below 1200 m a.s.l. Precipitation in about 55% of the province is moderately wet (1000–2500 mm year^{-1}), 33% wet (2500–4000 mm year^{-1}) and 12% very wet (>400 mm year^{-1}) (Table 2).

Table 1
Land use (in 1000 ha) in the Morobe Province of Papua New Guinea in 1975, 1990 and 2000, based on topographic maps (1975) and Landsat images (1990 and 2000)

Land use	1975		1990		2000	
	× 1000 ha	%	× 1000 ha	%	× 1000 ha	%
Agriculture	500.1	15	726.6	21	791.8	23
Forest	2646.3	78	2317.0	68	2228.7	66
Grassland	216.9	6	276.4	8	294.5	9
Plantation	23.7	1	63.2	2	67.8	2
Urban	2.5	<0.1	6.4	<0.1	6.7	<0.1
Water	3.7	<0.1	3.8	<0.1	3.9	<0.1
Total	3393.3	100	3393.3	100	3393.3	100

Relating agricultural change to the environmental conditions shows that most of the change occurred in moderately wet to very wet areas (>2500 mm rain year^{-1}), with fertile soil and on moderate to very steep slopes (10–56%). Agricultural areas that are converted to other land use are mostly Mollisols, in areas with about 1000–2500 mm rain per year.

3.4. Agricultural changes by districts

Expansion of agriculture was observed in all districts of the Morobe province. Rates of annual agricultural change ranged from 0.1% in Kaiapit to 10.3% in Siassi (Table 3). Other districts that showed a significant change in agriculture are Wau, Mumeng, Lae, Huon and Menyamya. While agricultural areas per person decreased with increasing population for most districts, Siassi and Wau districts showed an increase in agricultural area per person. Significant increases in agricultural changes on steep slopes are observed in Huon, Menyamya and Mumeng districts, while Siassi experienced a decline in land under agriculture in steeper areas.

The relationship between population growth and agricultural change by district showed that the largest change occurred in the Huon District. Menyamya, Wau and

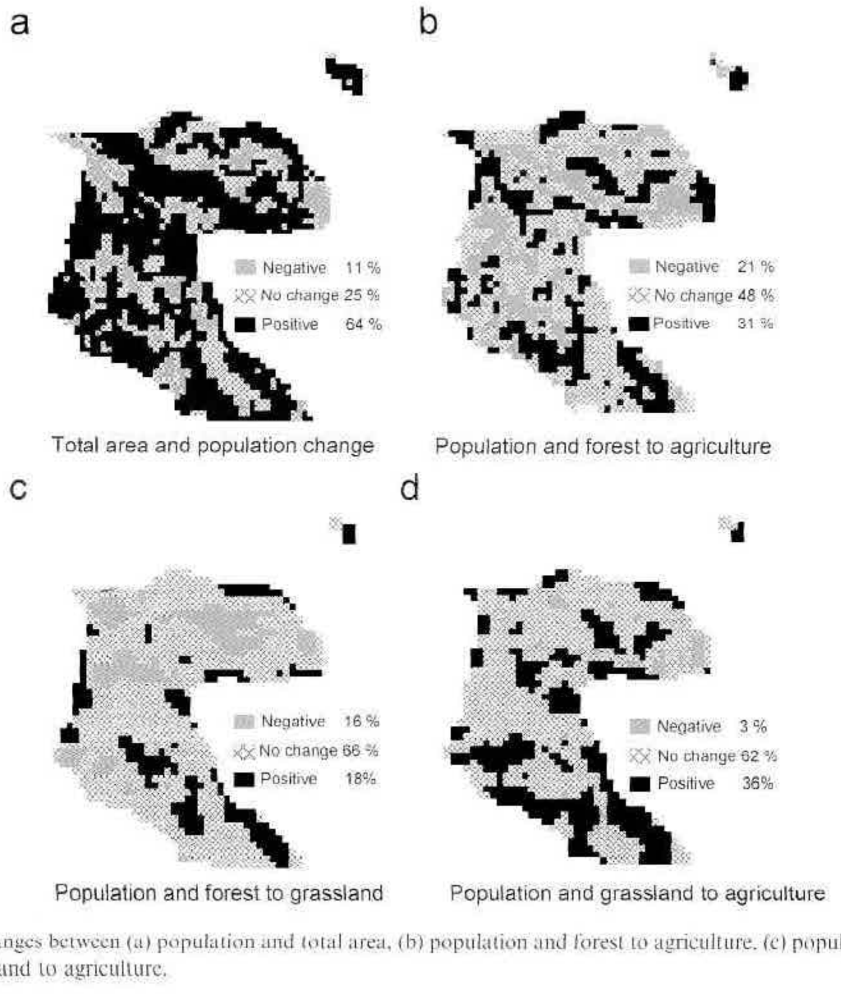


Fig. 4. Surface correlation changes between (a) population and total area, (b) population and forest to agriculture, (c) population and forest to grassland, and (d) population and grassland to agriculture.

Table 2
Environmental conditions and land use change in the Morobe Province of Papua New Guinea between 1975 and 2000

	Classes		Morobe province (× 1000 ha)		Agriculture in 1975 (× 1000 ha)		Agriculture in 2000 (× 1000 ha)		Change in agriculture (%)	
			Area	%	Area	%	Area	%	Total	Annual
Soil order	Alfisols		40.9	1	8.2	2	10.0	1	23	0.9
	Entisols		1170.9	35	181.0	36	261.8	33	45	1.8
	Inceptisols		1833.0	54	196.5	39	403.8	52	106	4.2
	Mollisols		340.0	10	113.4	23	110.8	14	−2	−0.1
	Other soils		8.4	<1						
Soil quality	Poor		199.9	6	26.3	5	40.9	5	56	2.2
	Moderate		1788.1	53	247.4	50	447.18	57	81	3.2
	Good		1395.2	41	225.4	45	298.5	38	32	1.3
Slope	Gentle	Up to 10%	408.9	12	141.9	28	160.7	20	13	0.5
	Moderate	10–34%	519.2	15	45.9	9	82.7	11	80	3.2
	Steep	35–56%	1622.0	48	218.7	44	371.4	47	70	2.8
	Very steep	>56%	843.2	25	92.6	19	171.8	22	85	3.4
Altitude	Low	0–1200 m	1875.0	55	327.8	66	520.4	66	59	2.4
	Moderate	1200–2400 m	1295.5	38	169.6	34	262.9	33	55	2.2
	High	>2400 m	222.8	7	1.7	<1	3.2	<1	83	3.3
Rainfall (mm)	Dry	1000–2500	1859.3	55	350.7	70	460.1	58	31	16.8
	Moderate	2500–4000	1138.2	33	88.3	18	193.00	25	119	38.7
	Wet	>4000	395.8	12	60.1	12	133.4	17	122	40.9

Data source: PNGRIS.

Table 3
Agricultural area in 1975 and 2000 with rate of changes by districts in the Morobe Province of Papua New Guinea

District	1975		2000		Rate of change (% year^{-1})	
	ha	ha person^{-1}	ha	ha person^{-1}	ha	ha person^{-1}
Finschhafen	50,497	1.3	84,793	1.3	2.7	0.0
Huon	120,344	1.5	212,185	1.2	3.1	−0.7
Kabwum	42,427	1.3	61,349	1.2	1.8	−0.4
Kaiapit	120,386	5.4	123,061	3.5	0.1	−1.4
Lae	3258	0.2	6119	0.2	3.5	−1.1
Menyamya	95,298	2.6	157,679	2.4	2.6	−0.4
Mumeng	15,105	1.2	28,824	1.2	3.6	−0.2
Siassi	3763	0.5	13,472	1.0	10.3	4.1
Wau	47,980	1.8	99,297	1.9	4.3	0.1

Finschhafen districts show moderate changes while the other districts show less moderate changes. Population growth with no corresponding significant changes in agriculture is observed in Lae and Kaiapit districts (Fig. 5). The total agricultural change for the Morobe province between 1975 and 2000 was 58%, of which 6% occurred within logged-over areas where forest has been cleared for commercial logging. Agricultural changes within the logged areas tend to appear at gentler slopes, in the vicinity of populated places (villages) and in close proximity to transport access (road/sea). Fig. 6 shows a typical example on Siassi Island where agricultural changes occurred on the fringes of logged-over areas and on gentle slopes.

4. Discussion

Significant agricultural expansion was found between 1975 and 2000 and this coincides with rapid population growth. Forest is cleared for agriculture as a result of increasing population density, migrations, general economic situation or access to land resources. About 6% of the agricultural changes occur within logged-over areas. Marked agricultural changes are in close proximity to villages where land use intensity tends to be high (Bourke et al., 1998) or within easy access to transport (Kok, 2004). Clearing of forest takes place in populated areas, or that are closer to transportation access and are suitable for agriculture.

There was a significant difference in the annual agricultural land use change between 1975 and 1990 (3%) compared to 1990 and 2000 (0.9%). Land use changes are caused by a series of factors and often show a non-linear pattern over time (Lambin et al., 2003; Lepers et al., 2005). Although we have no data to investigate the differences between 1975–1990 and 1990–2000, the lower agricultural expansion between 1990 and 2000 was possibly caused by

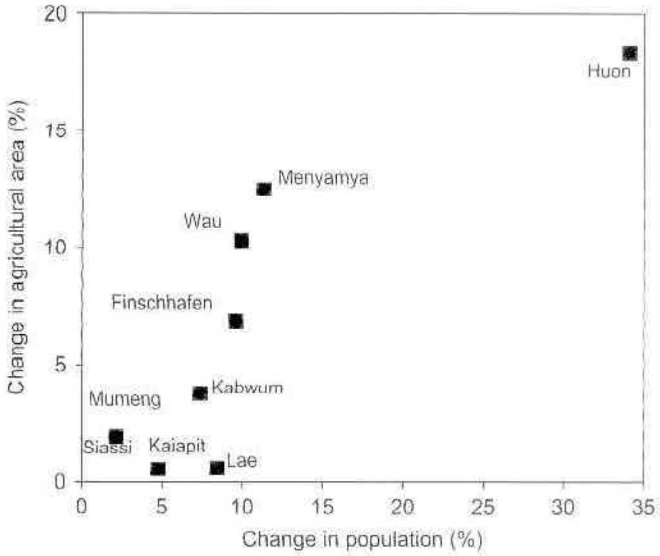


Fig. 5. Relation between changes in agricultural land use and population between 1975 and 2000 in 9 districts of the Morobe Province of Papua New Guinea.

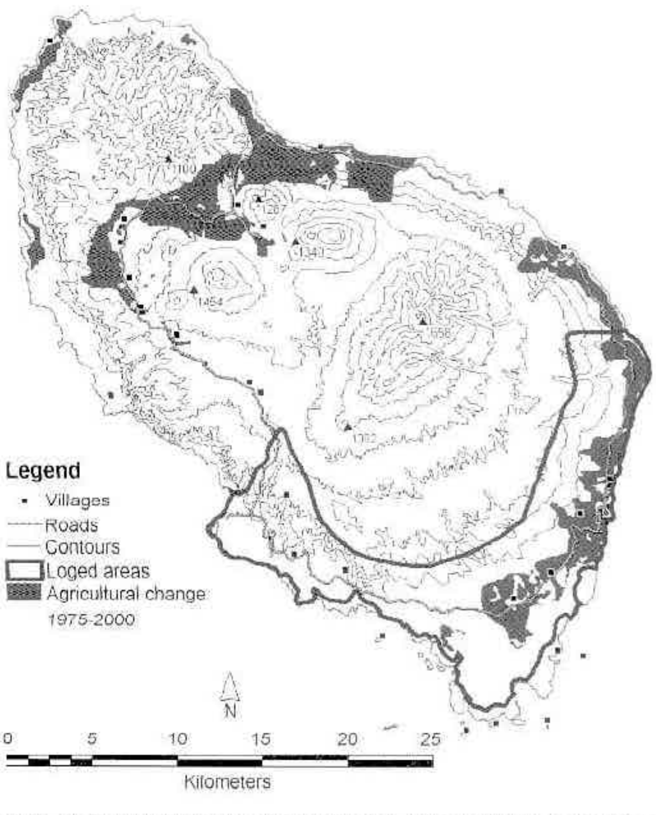


Fig. 6. Agricultural land use change between 1975 and 2000 on the fringes of logged over areas on Siassi Island in the Morobe Province of Papua New Guinea.

increased access to alternate food sources such as store and market, and increased off-farm income leading to less crop cultivation. Much of Papua New Guinea was affected by frost and drought during the 1997–1998 El Niño phenomena (Bourke, 2000) which concomitantly increased import of grains (Blakeney and Cough, 2000; Bourke, 2000; Gwaiseuk, 2000) and perhaps reduced the area under cultivation.

The correlation of total population and forest to grassland conversion is positive in certain areas. When superimposed with forestry data, most of the changed areas appear to have been logged, which means that logged-over areas are succeeded by grassland. Also correlation between total population change and grassland to agriculture conversion shows a positive relationship in highly populated areas. In the logged-over areas, grassland succeeds logging but it is subsequently converted to agriculture land. Such sequence of conversions (forest–grassland–cropland) are common in many tropical regions (Lepers et al., 2005).

The strong positive correlation between total population change and total land use change could imply lack of technological development as new areas are cleared to increase crop production rather than improving current farming techniques. People adapt to environmental conditions in the absence of technology by improving their farming techniques such as introducing new crops or mixed cropping (Bourke, 2001; Ohtsuka, 1995). Without improved farming or soil fertility management techniques, the observed land use change may affect the sustainability of agricultural production as a result of soil fertility depletion (Hartemink and Bourke, 2000). There was also increased cultivation on steeper slopes with increased risk of soil erosion; it confirms other studies in tropical regions (e.g. Kok, 2004; Pahari et al., 2001; Pfeffer et al., 2005).

There was variation in land use change between the districts of the Morobe Province and three patterns were found. Firstly, Huon district had the largest change in both population growth and agricultural land use. This could be due to better access to transport and agro-ecological conditions which allowed for a rapid population growth and an increase in the area under agriculture. Most of the district is suitable for agriculture especially along the coast and Markham valley. It is commonly found that areas with good transport access (sea, land, air) experience more change (Verburg et al., 2004).

The second group of districts (Menyamya, Wau and Finschhafen) show moderate changes. Generally, these districts have difficult terrains for accessibility but also have high agricultural potential. Menyamya is known for coffee production with other food crops due to its favourable climatic conditions, high rainfall and some fertile soils. The same applies for parts of Finschhafen and Wau districts. Apparently, population growth keeps pace with agricultural expansion in such areas.

The third group (Kabwum, Mumeng, Siassi and Kaiapit) show varying changes between 5% and 10% for agriculture and population, respectively. Mumeng, Kabwum and Kaiapit experienced more population growth compared to agriculture change which can be caused by location specific agro-ecological situations, access to transport and off-farm activities. Kaiapit is the district with the lowest agriculture change (0.1%) and the mountainous terrain constraints agricultural expansion.

5. Conclusion

The area under agriculture in the Morobe province increased by 58% while population almost doubled between 1975 and 2000. The area under agriculture was 1.5 ha person^{-1} in 2000, compared to the global average of 0.35 ha person^{-1} (Ramankutty et al., 2002). Most of the agricultural growth was observed in wet fertile areas but also on steep slopes. Increase in population density slightly reduced the per capita arable land but imposes pressure on the environment. In the absence of improved soil fertility management, improved farming system and increased crop production, the current trend of increasing the agriculture area as a result of human population growth is likely to continue.

Acknowledgements

We thank Mr. John Stuiver from the Centre of Geo-Information in Wageningen University in the Netherlands for the help in data processing and Dr. Michael Govorov from Malaspina University College, Canada, for the surface correlation avenue script. Comments on the draft of this paper by Mr. Niels Batjes of ISRIC and Dr. Bryant Allen of the Australian National University are highly appreciated.

References

Bellamy, J.A., McAlpine, J.R., 1995. Papua New Guinea Inventory of Natural Resources, Population Distribution and Land Use Handbook. Australian Agency for International Development, Canberra, 155pp.

Blakeney, M., Cough, R., 2000. Assessment of Grain Production and Imports in PNG, 1975–2000. Huon Seminar—Food Security in Papua New Guinea. PNG University of Technology, Lae, Papua New Guinea, pp. 23–29.

Bleeker, P., 1983. Soils of Papua New Guinea. Macarthur Press, Canberra, 352pp.

Bouma, J., Varallyay, G., Batjes, N.H., 1998. Principal land use changes anticipated in Europe. Agriculture Ecosystems and Environment 67, 103–119.

Bourke, R.M., 2000. An Overview of Food Security in Papua New Guinea. Huon Seminar—Food Security in Papua New Guinea. PNG University of Technology, Lae, Papua New Guinea, pp. 5–14.

Bourke, R.M., 2001. Intensification of agricultural systems in Papua New Guinea. Asia Pacific View Point 42, 219–235.

Bourke, R.M., Allen, B.J., Hobsbawn, P., Conway, J., 1998. Papua New Guinea: text summaries. Agricultural systems of Papua New Guinea working Paper No. 1., Department of Human Geography, Research School of Pacific and Asian Studies, The Australian National University, Canberra, pp. 221.

Bourke, R.M., et al., 2002. Morobe Province, second edition, Agricultural Systems of Papua New Guinea, Working Paper No. 19, ANU, Canberra.

Dekker, R.J., 2004. Object-based updating of land-use maps of urban areas using satellite remote sensing. In: 12th International Conference on Geoinformatics—Geospatial Information Research: Bridging the Pacific and Atlantic. University of Gävle, Sweden, pp. 623–630.

Greenland, D.J., 1997. Inaugural Russell-memorial-lecture—soil conditions and plant growth. Soil Use and Management 13, 169–177.

Gwaiseuk, W.R., 2000. The Role of Agriculture in the PNG Economy. Huon Seminar—Food Security in Papua New Guinea. ACIAR Publications. PNG University of Technology, Lae, Papua New Guinea, pp. 30–36.

Hartemink, A.E., Bourke, R.M., 2000. Nutrient deficiencies of agricultural crops in Papua New Guinea. Outlook on Agriculture 29, 97–108.

Kok, K., 2004. The role of population in understanding Honduran land use patterns. Journal of Environmental Management 72, 73–89.

Lambin, E.F., et al., 2001. The causes of land-use and land-cover change: moving beyond myths. Global Environmental Change 11, 261–269.

Lambin, E.F., Geist, H.J., Lepers, E., 2003. Dynamics of land-use and land-cover change in tropical regions. Annual Review of Environment and Resources 28, 205–241.

Lepers, E., et al., 2005. A synthesis of information on rapid land-cover change for the period 1981–200. BioScience 55, 115–124.

Li, Z., Li, X., Wang, Y., Ma, A., Wang, J., 2004. Land-use change analysis in Yulin prefecture, northern China using remote sensing and GIS. International Journal of Remote Sensing 25, 5691–5703.

Lillesand, T.M., Kiefer, R.W., 2000. Remote Sensing and Image Interpretation. Wiley, New York, 715pp.

McAlpine, J., Freyne, D.F., Keig, G., 2001. Land use and population change in PNG: 1975–1995. In: Bourke, R.M., Allen, M., Salisbury, J. (Eds.), Food Security in Papua New Guinea. ACIAR, Canberra.

McAlpine, J.R., Freyne, D.F., 2001. Land use and intensification in Papua New Guinea 1975–1996. Asia Pacific View Point 42, 209–218.

Meertens, H.C.C., Fresco, L.O., Stoop, W.A., 1996. Farming systems dynamics: impact of increasing population density and the availability of land resources on changes in agricultural systems. The case of Sukumuland, Tanzania. Agriculture, Ecosystems and Environment 56, 203–215.

Ohtsuka, R., 1995. Agricultural sustainability and food in Papua New Guinea. In: Uitto, J.I., Ono, A. (Eds.), Population, Land Management and Environmental Change. UNU Global Environmental Forum IV. The United Nations University (UNU), Tokyo, Japan.

Pahari, K., Umezaki, M. and Ohtsuka, R., 2001. Land use dynamics in Tari basin of Papua New Guinea using multi sensor satellite data. In: 22nd Asian Conference on Remote Sensing, Asian Association on Remote Sensing (AARS), Singapore, 4pp.

Pat, R.L.K., 2003. Customary Land Tenure in a Changing Context. Department of Lands and Physical Planning, Port Moresby.

Pfeffer, M.J., Schlelhas, J.W., DeGloria, S.D., Gomez, J., 2005. Population, conservation, land use change in Honduras. Agriculture, Ecosystems and Environment 110, 14–28.

Ramankutty, N., Foley, J.A., 1999. Estimating historical changes in global land cover: croplands from 1700 to 1992. Global Biogeochemical Cycles 13, 997–1027.

Ramankutty, N., Foley, J.A., Olejniczak, N.J., 2002. People on the land: changes in global population and croplands during the 20th century. Ambio 31 (3), 251–257.

Saunders, J.C., 1993. Forest Resources of Papua New Guinea. Australian International Development Assistance Bureau (AIDAB), Canberra, 17pp.

Verburg, P.H., Overmars, K.P., Witte, N., 2004. Accessibility and land-use patterns at the forest fringe in the northeastern part of the Philippines. Geographical Journal 170, 238–255.

ELSEVIER

Journal of Environmental Management 87 (2008) 125–131

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Alkaline hydrothermal conversion of fly ash precipitates into zeolites 3: The removal of mercury and lead ions from wastewater

Vernon Somerset[a,*], Leslie Petrik[b], Emmanuel Iwuoha[a]

[a]*Sensor Research Laboratory, Department of Chemistry, University of the Western Cape, Private Bag X17, Bellville 7535, South Africa*
[b]*Environmental and Nano Sciences Group, Department of Chemistry, University of the Western Cape, Private Bag X17, Bellville 7535, South Africa*

Received 31 January 2006; received in revised form 16 September 2006; accepted 4 January 2007
Available online 19 March 2007

Abstract

In this paper, the utilisation of zeolites synthesised from fly ash (FA) and related co-disposal filtrates as low-cost adsorbent material were investigated. When raw FA and co-disposal filtrates were subjected to alkaline hydrothermal zeolite synthesis, the zeolites faujasite, sodalite and zeolite A were formed. The synthesised zeolites were explored to establish its ability to remove lead and mercury ions from aqueous solution in batch experiments, to which various dosages of the synthesised zeolites were added. The test results indicated that when increasing synthesised zeolite dosages of 5–20 g/L were added to the acid mine drainage (AMD) wastewater, the concentrations of lead and mercury in the wastewater were reduced accordingly. The lead concentrations were reduced from 3.23 to 0.38 and 0.17 μg/kg, respectively, at an average pH of 4.5, after the addition of raw FA zeolite and co-disposal filtrate zeolite to the AMD wastewater. On the other hand, the mercury concentration was reduced from 0.47 to 0.17 μg/kg at pH = 4.5 when increasing amounts of co-disposal filtrate zeolite were added to the wastewater. The experimental results had shown that the zeolites synthesised from the co-disposal filtrates were effective in reducing the lead and mercury concentrations in the AMD wastewater by 95% and 30%, respectively.

Keywords: Fly ash; Co-disposal reaction; Hydrothermal zeolites; Lead; Mercury; Absorption

1. Introduction

In South Africa large quantities of fly ash (FA) are produced annually in the combustion of coal for electricity generation. This is due to the fact that coal with high-ash content is used in power generation producing approximately 25 Mt of FA in 1997. Since only 5% of this FA was re-used in 1997, there are increasing concerns about the fate of the FA and the environmental consequences if appropriate utilisation avenues are not investigated. The main constituents of FA are primarily aluminosilicate glass, mullite ($Al_6Si_2O_{13}$) and quartz (SiO_2) (Woolard et al., 2000; Somerset et al., 2004). With these quantities of FA produced annually the possibility of zeolite production from waste material that is readily available at the source can be explored for FA utilisation in South Africa.

Specific applications for the utilisation of FA are being explored annually. Some of the strategies include: (i) utilisation of FA in zeolite synthesis for wastewater processing (Woolard et al., 2000; Somerset et al., 2004); (ii) as an additive in the manufacturing of cement, concrete, construction materials and road pavements (Kao et al., 2000; Horiuchi et al., 2000); (iii) use of FA in the removal of organic compounds from aqueous solution (Kao et al., 2000).

Zeolites containing the faujasite structure play an important role in petrochemical industries. Sites containing natural deposits of the zeolite are scarce, with only six small deposits in the world and a relatively bigger site in Jordan. In view of this, commercially useful faujasite zeolites are typically manufactured (Singh and Dutta, 1998). When faujasites are prepared in laboratory experiments, it sometimes tends to be metastable, quickly disappearing to form more condensed structures. This is also a contributing factor to the scarcity of faujasite. Faujasite may also contain cations of Ca^{2+} and Mg^{2+},

*Corresponding author. Tel.: +27 21 9592792; fax: +27 21 9593055.
E-mail address: vsomerset@uwc.ac.za (V. Somerset).

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.01.033

accounting for 70–80% of the neutralizing cations in its structure (Singh and Dutta, 1998).

Several studies (Woolard et al., 2000; Seames, 2003; Querol et al., 2001; Hollman et al., 1999; Cheng-Fang and Hsing-Cheng, 1995) reported the conversion of FA into zeolites by treatment of the FA with concentrated NaOH solutions at elevated temperatures and pressures. Other modern techniques employ microwave radiation and fusion with NaOH solutions, followed by hydrothermal treatment of the FA and NaOH solution (Woolard et al., 2000; Somerset et al., 2004).

Lead is classified as a hazardous waste and is highly toxic to humans, plants and animals. When animals and plants are exposed to it, it causes death of these species, whereas in humans it causes anaemia, brain damage, mental deficiency, anorexia, vomiting and malaise. Calcium present in the bony tissues of mammals can be substituted by lead and it will accumulate there. Lead is also commonly found in drinking water, causing various types of illnesses and serious health problems. Researchers have shown that various low-cost adsorbents such as onion skin, tea leaves and peat moss can be used to absorb lead (II) ions from solutions in their native state, followed by suitable further chemical treatment for enhanced adsorption (Bulut and Baysal, 2005; Rahmana et al., 2005).

An increase in mercury pollution of the environment has been observed due to the increasing industrial use of mercury, another toxic metal that is hazardous to mammals, plant and animals. Mercury is included in the US EPA priority list of pollutants and has received great attention for many years. When humans are exposed to excessive mercury concentrations, it causes carcinogenic, mutagenic, teratogenic side effects and also promotes tyrosinemia. Furthermore, high concentrations of mercury cause impairment of pulmonary and kidney function, chest pain and dyspnousea in humans. To effectively deal with this increase in mercury pollution, clean-up technologies are required that are capable of treating large volumes of soil, water or sediment contaminated with relatively low levels of mercury. Not only are effective technologies required, but these technologies also need to reduce mercury levels in a cost-effective manner (Von Canstein et al., 1999; Zhang et al., 2005).

Mercury and lead were chosen for the absorption studies presented here, since they are known for their toxicity to the environment and to human health. Several techniques such as solvent extraction, ion exchange, precipitation, membrane separation, reverse osmosis, coagulation and photo-reduction have been applied for effective reduction of lead and mercury concentrations from various aqueous solutions. These listed techniques are not always cost-effective since they require either high amounts of energy or large quantities of chemicals. Adsorption is known to be effective in removing lead or mercury from wastewater by using activated carbon; however, it is expensive for large-scale application. Recently, more attention is being paid to the use of municipal sewage sludge, FA and zeolites (natural or synthetic) for the adsorption of pollutants from wastewater (Zhang et al., 2005; Turan et al., 2005; Woolard et al., 2000).

In a previous paper (Somerset et al., 2004) the results obtained for the conversion of raw FA and co-disposal FA filtrates into zeolites by alkaline hydrothermal treatment with sodium hydroxide (NaOH) were established and reported. Following this it was reported in another paper (Somerset et al., 2005b) how efficient the synthesised zeolites were as compared to commercial zeolites, in absorbing certain metal ion species from wastewater as the zeolite dosages were increased in batch experiments. Using these results, further investigations were carried out in this follow-up study to assess how the zeolitic materials synthesised from different FA starting material are different in removing metal ion species from acid mine drainage (AMD) wastewater. This paper thus focuses on the absorption efficiency of the different zeolites obtained during synthesis, to remove lead and mercury metal ion species from mine wastewater. The metal ion concentrations of lead and mercury were investigated as increased dosages of the different zeolites were added to the wastewater in a batched process. The ultimate aim was to investigate if the different zeolitic material can lower these heavy metal concentrations as increased dosages of the zeolites were added to the wastewater.

2. Materials and methods

2.1. Zeolite preparation

A novel approach was used in this study, which involved the collection of co-disposal filtrates by using a co-disposal reaction wherein FA was reacted with AMD in a specific FA:AMD ratio (e.g. 1:3.5, 1:4, 1:5) (Somerset et al., 2004, 2005a).

Fresh FA and co-disposal filtrates were then prepared for zeolite synthesis. Thoroughly dried FA and co-disposal filtrates were subjected to alkaline hydrothermal zeolite synthesis. This involved fusing the sample with sodium hydroxide (NaOH) in a 1:1.2 ratio at 600 °C for approximately 1–2 h. The fused product was then mixed thoroughly with distilled water and the slurry was subjected to aging for 8 h. After aging the slurry was subjected to crystallisation at 100 °C for 24 h. In the next step, the solid product was recovered by filtration and washed thoroughly with deionised water until the filtrate had a pH of 10–11. The recovered product was then dried at a temperature of 70 °C and prepared for characterisation (Rayalu et al., 2000; Somerset et al., 2004).

2.2. Elemental characterisation using X-ray fluorescence (XRF) spectrometry

The chemical composition of the raw FA and co-disposal filtrates that were used for zeolite synthesis was evaluated using XRF spectrometry. A Phillips 1404 XRF

Wavelength Dispersive Spectrometer equipped with an array of six analysing crystals and fitted with a rhodium X-ray tube target was used. A vacuum was used as the medium of analyses to avoid interaction of X-rays with air particles (Somerset et al., 2004).

2.3. X-ray diffraction (XRD) analysis

The mineralogy of each prepared zeolite sample was evaluated with a Phillips Analytical XRD spectrometer. This instrument was equipped with a graphite monochromator and Cu-Kα radiation samples were scanned for 2θ ranging from 7 to 70. The data files presented by X'Pert Graphics & Identify data collection software were used to identify the minerals present in the samples (Somerset et al., 2004).

2.4. Mercury and lead adsorption experiments

Mercury and lead absorption experiments were carried out on navigation AMD wastewater samples. Increasing dosages of the synthesised zeolites were added to the AMD aliquots and batch experiments were carried out in duplicate at room temperature, with continuous stirring in PVC plastic containers. To aliquots of 100 ml of AMD, synthesised zeolite material was added in dosages ranging from 0 to 20 g/L. After 1 h of continuous stirring, the mixtures were filtered and the pH measured, while the lead concentration was determined, using inductively coupled plasma mass (ICP-MS) spectrophotometry. The cold-vapour atomic absorption spectroscopy (AAS) technique was used for mercury detection using the 253.7 nm analytical resonance line on the instrument.

3. Results and discussion

3.1. XRF spectrometry results of starting materials

The raw FA and co-disposal filtrates used for zeolite synthesis were analysed by XRF spectrometry for quantitative determination of its SiO_2 and Al_2O_3 content. From the results in Table 1 it can be seen that the $[SiO_2]/[Al_2O_3]$ ratio is 1.95 for both the raw FA and co-disposal filtrate samples. Rayalu et al. (2000) have indicated and found in their zeolite synthesis experiments that a $[SiO_2]/[Al_2O_3]$ ratio higher than 1.5 is successful in delivering faujasite as zeolitic material from FA starting material.

The results in Table 1 further show that the SiO_2 content is higher in the co-disposal filtrate than in the raw FA, while the Al_2O_3 content of the co-disposal filtrate material was lower than in the raw FA. The CaO content of the co-disposal filtrate was found to be lower than in the raw FA, which was attributed to its consumption in the AMD neutralisation reaction. The MgO content was found to be relatively similar for the raw FA and co-disposal filtrate samples.

Table 1
Chemical composition of the fresh fly ash and co-disposal filtrate samples as analysed by XRF spectrometry

Oxide (wt%)	Samples	
	FA	CD filtrate
SiO_2	51.49	57.03
TiO_2	1.45	1.30
Al_2O_3	26.37	23.52
Cr_2O_3	0.03	0.02
Fe_2O_3T	5.03	5.70
MnO	0.06	0.16
NiO	0.01	0.02
MgO	2.52	2.61
CaO	7.83	6.22
Na_2O	0.53	0.00
K_2O	0.65	0.59
P_2O_5	0.38	0.24
H_2O^-	0.11	0.49
Total	99.0	99.9
Si/Al	1.95	1.95

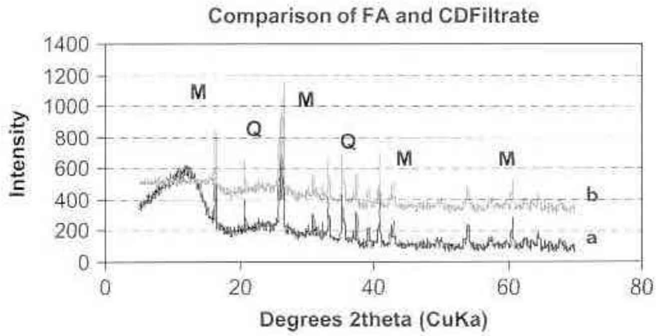


Fig. 1. XRD patterns for (a) raw FA and (b) co-disposal filtrate samples. The phases M (mullite) and Q (quartz) were present in the sample material.

3.2. XRD spectrometry results

XRD analysis was conducted on the raw FA and co-disposal filtrate samples in order to determine the crystalline phases present in the initial material before and after zeolite synthesis was performed. The XRD patterns of the studied raw FA and co-disposal filtrate are shown in Fig. 1. Fig. 1(a) shows that the predominant crystalline phases present in the FA were quartz (SiO_2) and mullite ($Al_6Si_2O_{13}$). For the co-disposal filtrate sample, quartz and mullite were also present as shown in Fig. 1(b). These results also showed that the mineralogies of the raw FA and co-disposal filtrate samples are relatively similar and that the co-disposal filtrates are a rich source of SiO_2 and Al_2O_3 for hydrothermal zeolite synthesis.

After hydrothermal zeolite synthesis, the initial material was again subjected to XRD analysis in order to determine whether hydrothermal zeolite synthesis of the initial material was successful. The XRD pattern of the raw FA sample subjected to zeolite synthesis is shown in Fig. 2, indicating that the crystalline zeolite phase called faujasite (F) has been formed. For this sample the results have shown that a single zeolite phase can be obtained through hydrothermal zeolite synthesis.

For the co-disposal filtrate sample used in the hydrothermal zeolite synthesis, the following XRD spectrometry results were obtained. In Fig. 3 it can be seen that more than one crystalline zeolite phase has been formed. These crystalline phases consist of faujasite (F), sodalite (S) and zeolite A (A). Again, the co-disposal sample was also successfully transformed into faujasite as zeolite crystalline phase. The $[SiO_2]/[Al_2O_3]$ ratio was thus a successful precursor in determining whether the chosen samples could be transformed into faujasite as zeolite material.

In both Figs. 2 and 3 it can be seen that the mineralogical phase mullite (M) is still present in both the zeolites synthesised hydrothermally with the use of a FA sample: NaOH ratio of 1:1.2. Other researchers (Singer and Berkgaut, 1995; Lin and Hsi, 1995; Querol et al., 1997) have shown that when the NaOH concentration is less than 3.5 M, considerable quantities of mullite are still present in the zeolites, due to the fact that mullite is relatively stable under alkaline treatment. However, only when the NaOH concentration is more than 4 M the mullite is indeed digested (Singer and Berkgaut, 1995; Querol et al., 1997).

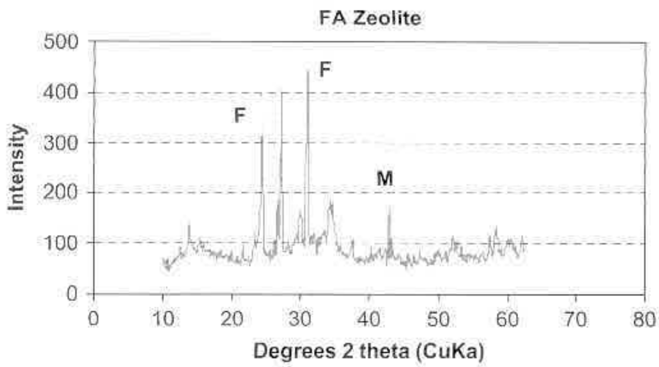


Fig. 2. XRD pattern for zeolite sample obtained from raw FA. The zeolitic phase F (faujasite) was present in the sample material. M = mullite.

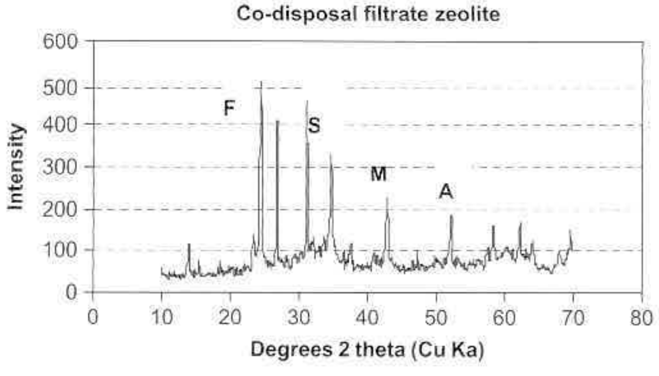


Fig. 3. XRD pattern obtained for zeolite sample obtained from co-disposal filtrate sample. The zeolitic phases F (faujasite), S (sodalite) and A (zeolite A) were present in the sample material. M = mullite.

3.3. Lead absorption results

Two representative AMD sources were used in the whole study, but only navigation AMD was explored in evaluating the absorption capacity of the synthesised zeolites to reduce metal ion species. One of the major aims of the total project studied was to use AMD and FA sources that are close to each other so as to reduce transport costs when industrial-scale studies of laboratory bench experiments were to be launched. This was in line with the objective of reducing cost while exploring the use of FA waste for the synthesis of a valuable product. For navigation AMD it was found that the pH and electrical conductivity (EC) are typical of an AMD source (Burgess and Stuetz, 2002). The results of the ICP-MS, Hg-vapour and IC analysis of this AMD source, as shown in Table 2, indicate that high concentrations of sulphate, iron, calcium and aluminium species are present.

Since the pH of the solution plays an important role in the absorption process, the pH of the solution after contact with the synthesised zeolites was monitored with the experiments done at room temperature (21 °C). With the AMD at an initial pH of 2.64 and the synthesised zeolites obtained at alkaline pH, the resulting pH after absorption experiments was found to be in the region of 4.25–4.88.

Before the absorption experiments were carried out, the cation exchange capacity of the synthesised zeolites was determined in a similar method as reported in a previous paper, and the results were reported as the amount of exchangeable Na^+ cations present in the zeolitic material (Somerset et al., 2005a). For the current synthesised zeolites, it was found that zeolites synthesised from raw FA had an amount of exchangeable Na^+ cations of 71.55 meq/100 g, while the value for zeolites synthesised from the co-disposable filtrate was 75.81 meq/100 g.

Table 2
ICP-MS, Hg-vapour and IC analysis of the chemical composition of the navigation AMD source

Constituent		Conc. (μg/kg)
pH	2.64	
EC (mS/cm)	9.45	
B		124.51
Cd		2.19
Cr		17.83
Co		640.30
Cu		320.31
Pb		3.36
Hg		0.49
Ni		921.91
Zn		4910.80
Al		13,840.00
Ca		497,220.00
Fe		3,522,920.00
SO_4		18,888,620.00

Although the amount of exchangeable Na^+ cations is high for both types of zeolites, it was found for this study that zeolites from raw FA had a lower amount of exchangeable Na^+ cations compared to zeolites from the co-disposable filtrate material. These results are important as they are indicative of the extent to which the synthesised zeolites will be able to exchange some of the lead and mercury ions in the absorption experiments. A possible mechanism whereby the ion-exchange process occurs in the zeolite structure is one where the Pb^{2+} (or Hg^{2+}) cations replace the Na^+ ions as shown below:

$$\text{Z-Na} + Pb^{2+} \rightarrow \text{Z-}Pb^{2+} + Na^+$$

This process is controlled by Le Chatelier's principle and is pH and Na^+ concentration dependent (Woolard et al., 2000).

The results obtained for the removal of lead ions from the navigation AMD wastewater are displayed in Fig. 4. The results in Fig. 4 indicate that when the zeolite material synthesised from raw FA, in Pb[A], was added to the AMD, a sudden drop in the lead concentration was observed followed by a relatively gradual decrease as more zeolite material was added to the wastewater. The zeolite material was effective in reducing the lead concentration from 3.23 to 0.38 μg/kg.

Similarly, a sudden initial drop in the lead concentration was observed when zeolite material synthesised from co-disposal filtrate, in Pb [B], was added to the AMD. This was followed by a gradual decrease as the zeolite dosage was increased. The zeolite material synthesised from the co-disposal filtrate was effective in reducing the lead concentration from 3.23 to 0.17 μg/kg. The fact that a lower lead concentration was obtained with this zeolite material can be attributed to the fact that more than one zeolite phase was present in the material after synthesis.

The fact that the resulting pH of the solution after interaction of the navigation AMD with the synthesised zeolites was found to be in the region of 4.25–4.88 can be related to the amount of alkaline material present in the synthesised zeolites.

Since the synthesised zeolites contain sodium ions that have been added during synthesis and some calcium from the FA starting material, these alkaline metals consume acid by neutralisation during the absorption experiments. With the final pH above 3.5 a favourable environment is created for the precipitation of Pb^{2+} ions in the solution as $Pb(OH)_2$ at the end of the contact period. Therefore, the increase in Pb^{2+} absorption as the zeolite dosage is increased can be attributed to the precipitation of Pb^{2+} as $Pb(OH)_2$ at a pH above 3.5. Similarly, it was shown by Erdem and Ozverdi (2005) that a pH of 3.5 and higher is favourable for the absorption of Pb^{2+} using siderite in their absorption studies.

3.4. Mercury absorption results

Fig. 5 displays the results obtained for the removal of mercury ions from the navigation AMD wastewater.

In the mercury absorption experiments, the effect of pH was also taken into account. With the pH after absorption experiments in the region of 4.25–4.88, it definitely affected the absorption results for mercury with the use of synthesised zeolites.

It was observed that as the zeolite material synthesised from raw FA, in Hg[A], was added to the AMD, an initial increase in the mercury concentration was observed, followed by a gradual decrease as zeolite dosage was

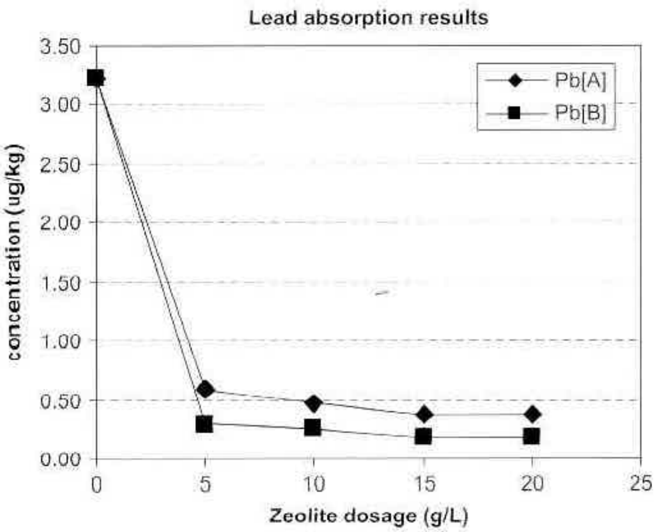


Fig. 4. Results shown for lead absorption from AMD wastewater vs zeolite dosage (g/L) added to wastewater. In Pb[A] the results are shown for the addition of the zeolites synthesised from raw FA, and in Pb[B] the results obtained for the zeolites synthesised from the co-disposal filtrate material are shown.

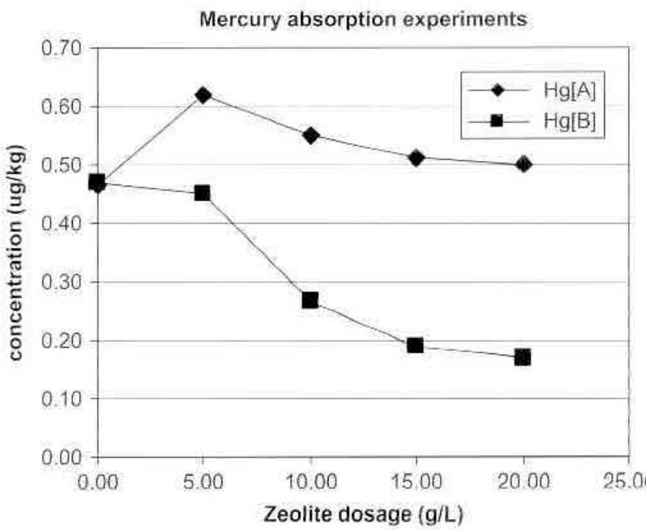


Fig. 5. Results shown for mercury absorption from AMD wastewater vs. zeolite dosage (g/L) added to wastewater. In Hg[A] the results are shown for the addition of the zeolites synthesised from raw FA, and in Hg[B] the results obtained for the zeolites synthesised from the co-disposal filtrate material are shown.

increased. The initial increase can be attributed to mercury ions already present in the wastewater or other possible ion exchange effects not focused on in this study. The wastewater had a final mercury concentration slightly higher at 0.50 μg/kg as when no zeolite material was added to the wastewater. These results indicate that the zeolites synthesised from raw FA were not able to lower the mercury concentration in the absorption experiments, while the use of higher zeolite dosages to observe a lowering effect needs to be further explored in a follow-up study.

Better results were obtained when the zeolite material synthesised from the co-disposal filtrate, in Hg[B], was added to the AMD as shown in Fig. 5. A gradual decrease in the mercury concentrations was observed as the zeolite dosage was increased from 0.5 to 20 g/L. This synthesised zeolite material was effective in reducing the mercury concentrations from 0.47 to 0.17 μg/kg. The results obtained can also be attributed to the fact that more than one zeolite phase was present in this material after synthesis.

Furthermore, Zhang et al. (2005) found that the predominant species of mercury in solution is $Hg(OH)_2$ at a pH of 4.5, and at a pH between 2 and 6 small amounts of $Hg(OH)^+$ are present. These data thus indicate that the pH obtained during the absorption experiments was favourable for mercury precipitation in the batch experiments conducted.

4. Conclusions

This work has demonstrated that faujasite, sodalite and zeolite A can be prepared from coal combustion FA and co-disposal filtrates by hydrothermal treatment of the starting material with NaOH. A novel approach was followed by using the FA to first neutralise the AMD in a co-disposal reaction, thus treating the AMD, and then using the obtained co-disposal filtrates in synthesising value-added zeolite material from the by-products. Furthermore, the zeolite material synthesised from the co-disposal filtrate material has proved to be effective in reducing the lead and mercury concentrations in navigation AMD wastewater, when increasing dosages of the zeolite material were added to the wastewater.

It was further illustrated that at an average pH of approximately 4.5 at room temperature, excellent conditions exist for the absorption of lead and mercury ions from the navigation AMD using the synthesised zeolites obtained from the co-disposal filtrate sample material. The zeolite material synthesised from the co-disposal filtrate was effective in reducing the lead concentration by 95% from 3.23 to 0.17 μg/kg, while the mercury concentration was reduced by 30% from 0.47 to 0.17 μg/kg.

It is envisaged that this conversion of FA and co-disposal filtrates into a beneficial product for further wastewater treatment could prove to be effective in removing heavy metals such as lead and mercury from industrial contaminated effluent waste streams.

Acknowledgements

The authors wish to express their gratitude to the Water Research Commission (WRC), Coaltech 2020, and the National Research Foundation (NRF) for funding and financial support to perform this study. The assistance provided by the CSIR, Pretoria and Eskom, Witbank in the collection of FA samples, iThemba labs for XRD analysis and the Chemistry department of the University of the Western Cape are also appreciated.

References

Bulut, Y., Baysal, Z., 2006. Removal of Pb(II) from wastewater using wheat bran. Journal of Environmental Management 78 (2), 107–113.

Burgess, J.E., Stuetz, R.M., 2002. Activated sludge for the treatment of sulphur-rich wastewaters. Minerals Engineering 15, 839–846.

Cheng-Fang, L., Hsing-Cheng, H., 1995. Resource recovery of waste fly ash: synthesis of zeolite-like materials. Environmental Science and Technology 29 (4), 1109–1117.

Erdem, M., Ozverdi, A., 2005. Lead adsorption from aqueous solution onto siderite. Separation and Purification Technology 42, 259–264.

Hollman, G.G., Steenbruggen, G., Janssen-Jurkovičová, M., 1999. A two-step process for the synthesis of zeolites from coal fly ash. Fuel 78, 1225–1230.

Horiuchi, S., Kawaguchi, M., Yasuhara, K., 2000. Effective use of fly ash slurry as fill material. Journal of Hazardous Materials 76, 301–337.

Kao, P.-C., Tzeng, J.-C., Huang, T.-L., 2000. Removal of chlorophenols from aqueous solution by fly ash. Journal of Hazardous Materials 76, 237–249.

Lin, C., Hsi, H., 1995. Resource recovery of waste fly ash: Synthesis of zeolite-like materials. Environmental Science and Technology 29, 1109–1117.

Querol, X., Plana, F., Alastuey, A., Lopez-Soler, A., Andres, J.M., Juan, R., Ferrer, P., Ruiz, C.R., 1997. A fast method of recycling fly ash: microwave assisted zeolite synthesis. Environmental Science and Technology 31, 2527–2532.

Querol, X., Umaña, J.C., Plana, F., Alastuey, A., López-Soler, A., Medinaceli, A., Valero, A., Domingo, M.J., Garcia-Rojo, E., 2001. Fuel 80, 857–865.

Rahmana, M.A., Ahsan, S., Kaneco, S., Katsumata, H., Suzuki, T., Ohta, K., 2005. Wastewater treatment with multilayer media of waste and natural indigenous materials. Journal of Environmental Management 74, 107–110.

Rayalu, S., Meshram, S.U., Hasan, M.Z., 2000. Highly crystalline faujasitic zeolites from fly ash. Journal of Hazardous Materials B 77, 123–131.

Seames, W.S., 2003. An initial study of the fine fragmentation fly ash particle mode generated during pulverized coal combustion. Fuel Processing Technology 81, 109–125.

Singer, A., Berkgaut, V., 1995. Cation exchange properties of hydrothermally treated coal fly ash. Environmental Science and Technology 29, 1748–1753.

Singh, R., Dutta, P., 1998. Stabilization of natural Faujasite zeolite: possible role of alkaline earth metal ions. Microporous and Mesoporous Materials 21, 103–109.

Somerset, V.S., Petrik, L.F., White, R.A., Klink, M.J., Key, D., Iwuoha, E.I., 2004. The use of X-ray fluorescence (XRF) analysis in predicting the alkaline hydrothermal conversion of fly ash precipitates into zeolites. Talanta 64, 109–114.

Somerset, V.S., Petrik, L.F., White, R.A., Klink, M.J., Key, D., Iwuoha, E.I., 2005a. Alkaline hydrothermal zeolites synthesized from high SiO_2 and Al_2O_3 co-disposal fly ash filtrates. Fuel 84, 2324–2329.

Somerset, V., Petrik, L., Iwuoha, E., 2005b. Alkaline hydrothermal conversion of fly ash filtrates into zeolites 2: utilization in wastewater treatment. Journal of Environmental Science and Health Part A 40 (8), 1627–1636.

Turan, M., Mart, U., Yüksel, B., Çelik, M.S., 2005. Lead removal in fixed-bed columns by zeolite and sepiolite. Chemosphere 60, 1487–1492.

Von Canstein, H., Li, Y., Timmis, K.N., Deckwer, W.D., Wagner-Döbler, I., 1999. Removal of mercury from chloralkali electrolysis wastewater by a mercury-resistant *Pseudomonas putida* strain. Applied Environmental Microbiology 65 (12), 5279–5284.

Woolard, C.D., Petrus, K., van der Horst, M., 2000. The use of a modified fly ash as an adsorbent for lead. Water SA 26 (4), 531–536.

Zhang, F.-S., Nriagu, J.O., Itoh, H., 2005. Mercury removal from water using activated carbons derived from organic sewage sludge. Water Research 39, 389–395.

Journal of Environmental Management 87 (2008) 132–138

Journal of Environmental Management

www.elsevier.com/locate/jenvman

LCA: A decision support tool for environmental assessment of MSW management systems

Chalita Liamsanguan, Shabbir H. Gheewala*

The Joint Graduate School of Energy and Environment, King Mongkut's University of Technology Thonburi, 126 Pracha-Uthit Road, Bangmod, Tungkru, Bangkok 10140, Thailand

Received 19 January 2006; received in revised form 1 December 2006; accepted 4 January 2007
Available online 12 March 2007

Abstract

Life cycle assessment (LCA) can be successfully applied to municipal solid waste (MSW) management systems to identify the overall environmental burdens and to assess the potential environmental impacts. In this study, two methods used for current MSW management in Phuket, a province of Thailand, landfilling (without energy recovery) and incineration (with energy recovery), are compared from both energy consumption and greenhouse gas emission points of view. The comparisons are based on a direct activity consideration and also a life cycle perspective. In both cases as well as for both parameters considered, incineration was found to be superior to landfilling. However, the performance of incineration was much better when a life cycle perspective was used. Also, landfilling reversed to be superior to incineration when methane recovery and electricity production were introduced. This study reveals that a complete picture of the environmental performance of MSW management systems is provided by using a life cycle perspective.

Keywords: Energy consumption; Greenhouse gas emission; Life cycle assessment; Municipal solid waste management

1. Introduction

Currently, Thailand is confronted with a high amount of municipal solid waste (MSW) and its inappropriate management, especially open dumping and non-sanitary landfill. These problems pose harm to the environment as well as human health. At the moment, major concerns associated with waste management are not only public health and safety but also sustainable development. For sustainable development, MSW management has to be balanced between environmental effectiveness, economic affordability and social acceptability to ensure the quality of life now and for coming generations. Concerning the environmental sustainability of MSW management systems, energy and resource conservation and reduced environmental impacts are desirable. To evaluate the performance of MSW management systems, life cycle assessment (LCA) is a useful tool.

LCA has been defined as a technique for assessing the environmental aspects and potential impacts associated with a product, by compiling an inventory of relevant inputs and outputs of a product system; evaluating the potential environmental impacts associated with those inputs and outputs; and interpreting the results of the inventory analysis and impact assessment phases in relation to the objectives of the study (ISO, 1997).

LCA is a methodology considering the entire life cycle of products and services—from cradle to grave (from raw material acquisition through production, use, and disposal). It is thus a holistic assessment methodology of products and services. LCA has been proven to be a valuable tool to document the environmental considerations that need to be part of decision making towards sustainability (UNEP, 2003).

LCA has been successfully utilized in the field of solid waste management to assess environmental impacts of solid waste management systems (Harrison et al., 2000), to compare the environmental performance of different scenarios for management of mixed solid waste (Denison, 1996; Mendes et al., 2004; Finnveden et al., 2000; Arena

*Corresponding author. Tel.: +66 2 4708309 10x4139; fax: +66 2 8729805.
E-mail address: shabbir_g@jgsee.kmutt.ac.th (S.H. Gheewala).

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.01.003

et al., 2003; Chaya and Gheewala, 2006; Wanichpongpan and Gheewala, 2006) as well as of specific waste fractions (Finnveden and Ekvall, 1998; Ross and Evans, 2003).

A systems approach does not always need to use impact assessment. In many cases, inventory data alone are sufficient for an evaluation (McDougall and White, 1998). The term life cycle inventory (LCI) is used to indicate that a study has excluded the impact assessment phase (Friðriksson et al., 2002).

Using LCA, an MSW management system is evaluated based on a system wide or life cycle perspective. A system that generates energy, such as incineration with energy recovery, is credited with reducing the amount of energy (and the associated resource use and emissions) that would otherwise need to be generated, typically at a power plant. If MSW management systems are compared in isolation without accounting for the system-wide environmental impacts, referred in the study as a direct activity consideration, such a limited perspective may not provide a complete picture of environmental impacts.

This study demonstrates a life cycle perspective evaluation of MSW management systems. Phuket, a province in the southern part of Thailand, was selected as the study site. Two methods currently used for MSW management in Phuket, landfilling (without energy recovery) and incineration (with energy recovery) are compared from both the energy consumption and the greenhouse gas emission points of view. The comparisons are based on a direct activity consideration as well as a life cycle perspective. The results of this study reveal the advantage of using a life cycle perspective in MSW system evaluation.

2. Current Phuket MSW management

Phuket is an island province in the south of Thailand stretching 49 km from north to south and 19 km from east to west with a total area of 570 km^2. With beautiful beaches along the western and southern parts of the island, Phuket is a major tourist attraction.

MSW in Phuket is collected and transported to the treatment and disposal center, where it is weighed and separated based on source and characteristics of the waste, to be managed by three methods—incineration, recycling, and landfilling. Flow of current Phuket MSW in a 1-year period (July 2003–June 2004) obtained from Phuket Municipality is illustrated in Fig. 1. Of the 133 374 tons of MSW collected in the 1-year period, an estimated 71% was sent for burning in incinerator, 26% landfilled, and 3% sorted and recovered for recycling.

3. Methodology

In this study, a comparison between the two methods used for current Phuket MSW management, landfilling (without energy recovery), and incineration (with energy recovery) is performed. The environmental burdens considered in the evaluation are energy consumption and greenhouse gas emission. To compare the two methods of MSW management, a fixed reference point for the environmental evaluation, called functional unit, is defined as 1 ton of MSW treated. For fairness of comparison, the same characteristics of waste are assumed to be treated by both landfilling and incineration. Waste characterization information obtained from monthly reports of Phuket incineration plant is illustrated in Table 1.

The evaluation includes activities that are of direct concern in MSW management and also activities that supply services to or interact with MSW management methods as illustrated in Fig. 2. Direct and indirect activities associated with MSW management methods contributing to energy consumption and greenhouse gas emission are listed in Table 2 and Fig. 2. Energy consumption for ash management is included in the calculations, however, greenhouse emissions are not since the ash is inorganic in nature. Based on the existing practice that there is no gas collection and flaring system in Phuket landfill and with the assumption of 10% methane oxidation in landfill cover (IPCC, 2001), 90% of the methane produced is released to the atmosphere. Although carbon dioxide is also emitted from the landfill, it is not considered because, being of biomass origin, it does not contribute to global warming. The landfill leachate is treated by pond system, which is the common method in Thailand. The energy and resource requirements are thus negligible. The main impact from this system would be on land use, which is not within the scope of this study. Transportation is not included in the system boundary as the collection and transportation of waste is common to both the waste management systems and hence will not influence the comparative result.

Findings from the study are presented based on two sets of boundary conditions (Table 2):

(1) a direct activity consideration, limited to only those processes that lie within MSW management method itself and
(2) a life cycle perspective, considering direct activities as well as other processes interacting with MSW management system. Since the function of landfilling is solid waste management, whereas the function of incineration is solid waste management with electricity production as a supplementary function, to make the systems comparable, the incineration is credited with the avoided emissions from the alternative process of producing an equivalent amount of electricity. The average electricity mix of Thailand is used for calculating the credits. Environmental burdens in the modified system are the environmental loads from the incineration minus those from the conventional power plants. In this way, both MSW management methods can be compared based on the same function which is solid waste management as illustrated in Fig. 3 (Finnveden and Ekvall, 1998).

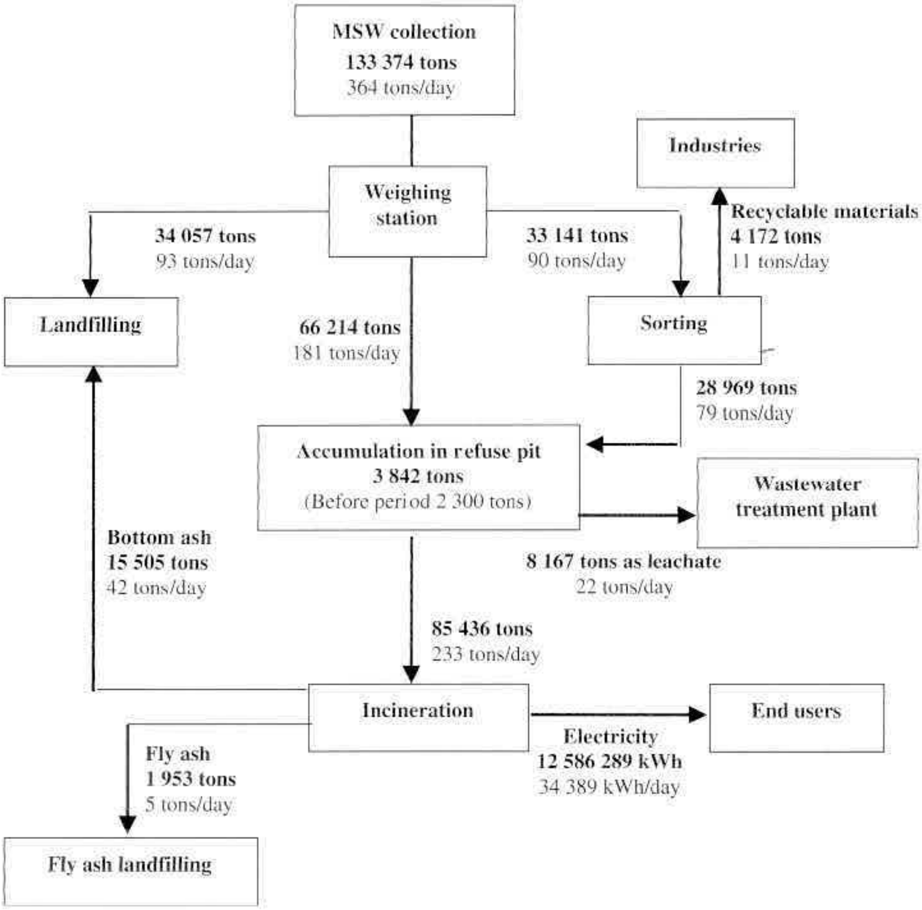


Fig. 1. Current Phuket MSW management system.

Table 1
Phuket waste characteristic

Waste composition (%)	
Plastic	27.71
Food	18.12
Wood/ grass	13.65
Paper	11.45
Cloth	3.06
Rubber/ leather	1.85
Incombustible	15.44
Others	8.71
Waste property	
Density (kg/m^3)	379
Moisture content (%)	41
Low heating value (kcal/kg)	1750

Information about energy consumption of the MSW management systems was collected from the actual processes at the study site. The environmental parameters contributed by direct activities, which were not available in the study area, were derived from literature. Fossil CO_2 emitted from incineration system was calculated by using emission factors for various types of plastics in the waste (Harrison et al., 2000). CH_4 emitted from landfill was calculated by using CH_4 emission factors of different waste fractions in MSW (Sandgren et al., 1996). Energy consumption and emissions to air from the Thai electricity production were obtained from the study of the Thailand Environment Institute (TEI, 2003). In the TEI study, the system boundary includes only the power generation process in power plants, i.e. fuel combustion and air pollution control. Extraction, processing and transportation fuels, and construction of power plants are not included. Data of energy and greenhouse gas emissions relating to diesel production and lime production were derived from BUWAL 300 and ETH-ESU (1996), respectively.

4. Results

The comparative information in this section includes the energy consumption and the greenhouse gas emissions of the two MSW management methods, landfilling (without energy recovery), and incineration (with energy recovery). The comparisons are based on a direct activity consideration and a life cycle perspective. The environmental parameters are expressed per functional unit, which is 1 ton of MSW treated.

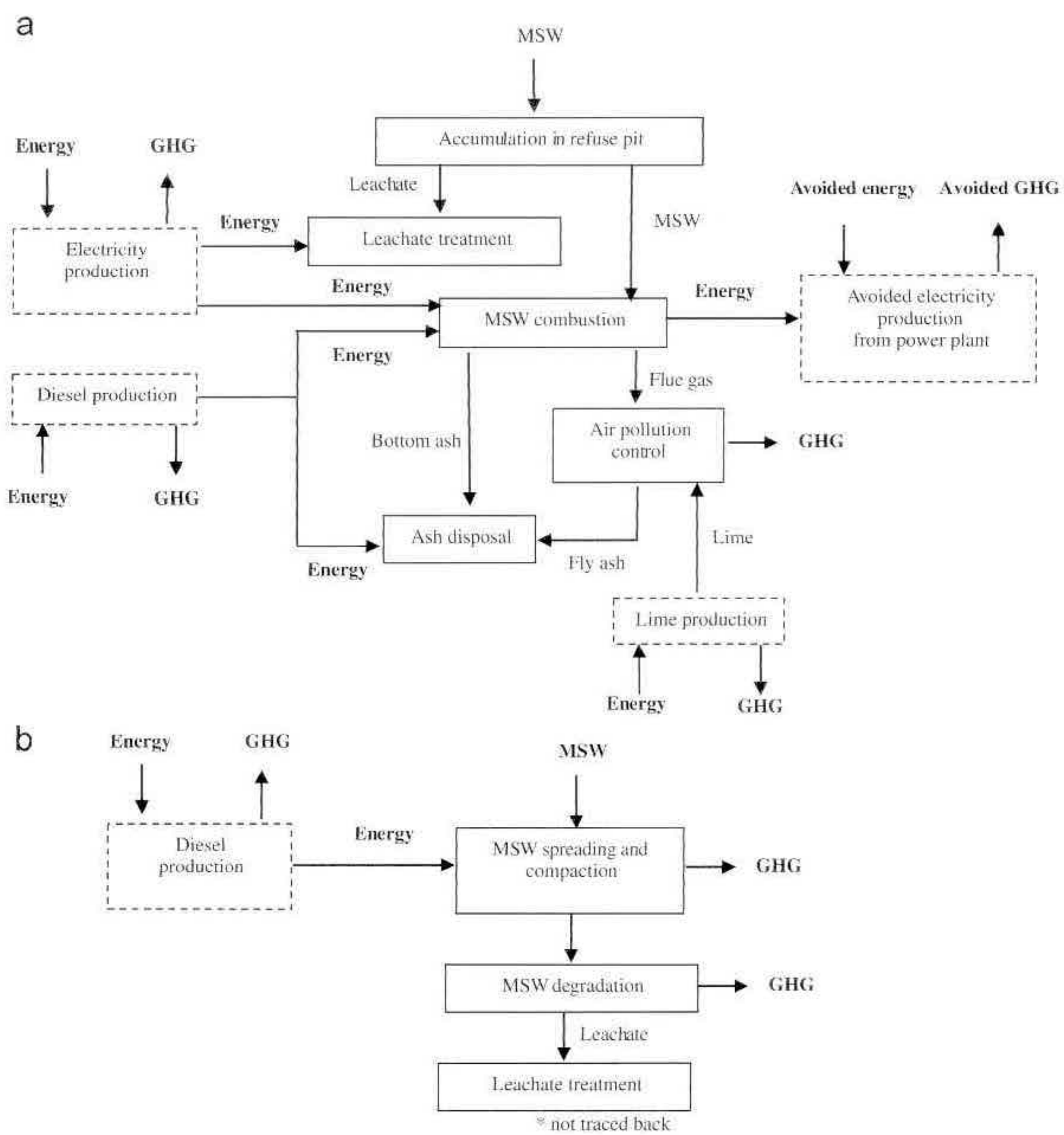


Fig. 2. Flow diagram for energy and greenhouse gas emission. (a) flow diagram of incineration (indirect activities are indicated by dashed lines), (b) flow diagram of landfilling (indirect activities are indicated by dashed lines).

4.1. Energy consumption

The comparison between the energy consumption for landfilling and incineration using a direct activity and a life cycle perspective is illustrated in Table 3.

From Table 3, based on the direct activity consideration, energy in the form of electricity and diesel required by all processes related with the incineration is 354 MJ/ton of MSW treated, whereas the electricity recovered from waste combustion is 530 MJ/ton of MSW treated. The net energy consumption equals to −176 MJ/ton of MSW treated.

Although the energy consumed at the landfill by engines for MSW spreading and compaction is less than that for incineration, the effect of energy recovered from incineration compensates for the energy used thus making it superior to landfilling.

With a life cycle perspective, the energy consumption is quantified by adding the net amount of energy consumed as in the direct activity consideration with the energy consumed by all related activities listed in Table 2. Since the methane gas generated in landfill is not recovered and utilized for energy, no credits from energy recovery are accounted for evaluation. Thus, the energy consumption in landfilling increases to be 30 MJ/ton of MSW treated.

In the case of incineration, 1458 MJ of primary energy at power plants is avoided due to electricity recovered from the incineration. This is much more than that added by other related activities, resulting in a net energy

Table 2
Direct and indirect activities associated with MSW management methods contributing to energy consumption and greenhouse gas emission

Environmental parameter	Description of direct and indirect activities
Energy consumption	*Incineration*
	Direct activity
	Electricity and diesel oil used for MSW combustion
	Electricity used for leachate treatment
	Indirect activity
	Energy used for electricity production in power plant
	Energy used for diesel production
	Energy used for lime production
	Landfilling
	Direct activity
	Diesel oil used for MSW spreading and compaction
	Indirect activity
	Energy used for diesel production
Greenhouse gas emission	*Incineration*
	Direct activity
	Emissions from the incinerator arising from MSW combustion
	Emissions from combustion of diesel used to operate incinerator
	Indirect activity
	Emissions from conventional power plant
	Emissions from diesel production
	Emissions from lime production
	Landfilling
	Direct activity
	Landfill gas from MSW degradation
	Emissions from diesel engines during MSW spreading and compaction
	Indirect activity
	Emissions from diesel production

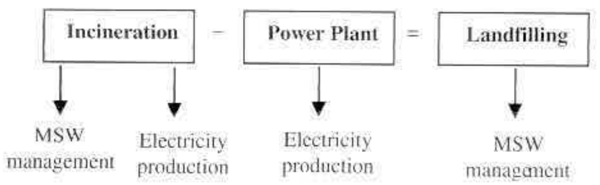


Fig. 3. System comparison from a life cycle perspective.

consumption of −1048 MJ/ton of MSW treated. This is much less than that in the case of direct activity consideration.

In both cases, the incineration is the preferred method over the landfilling based on energy consumption. However, incineration works out to be even more advantageous when a life cycle perspective is used.

4.2. Greenhouse gas emission

The greenhouse gases from the processes of MSW management considered in the evaluation are fossil CO_2, CO, CH_4, and N_2O. For global warming, the reference substance is CO_2. All greenhouse gases are expressed in the units of kg CO_2 equivalents/ton MSW treated. The equivalency factor for potential contributions from greenhouse gases to global warming over a time horizon of 100 years are 1 for carbon dioxide, 2 for carbon monoxide, 23 for methane, and 296 for nitrous oxide (Ramaswamy et al., 2001). The comparison between the greenhouse gas emission from landfilling and incineration using a direct activity and a life cycle perspective is illustrated in Table 4.

For direct activity consideration, amount of greenhouse gases emitted from the MSW management system itself is accounted. With this consideration, landfilling produces more greenhouse gas emission (in CO_2 equivalents) than incineration. For landfilling, the maximum contribution, 1311 kg CO_2 eq./ton MSW treated, is from methane produced during MSW degradation. For incineration, the most important greenhouse gas is fossil carbon dioxide from MSW (plastic portion) combustion contributing 736 kg CO_2 eq./ton MSW treated, the remaining being from CO. Although biodegradable waste fraction is also burnt in the incinerator, the carbon dioxide emitted is not accounted for in the evaluation because, being of biomass origin and hence part of the global carbon cycle, it does not contribute to global warming.

With a life cycle perspective, amount of greenhouse gases emitted from the MSW management itself along with those from all related activities listed in Table 2 are accounted.

For incineration, the amount of greenhouse gas emissions in the life cycle perspective case is less than that in the direct activity case because of the avoided electricity production from conventional power production resulting in avoided greenhouse gas emissions. After adding the greenhouse gas emissions from the indirect activities and the credits from avoided conventional electricity production, incineration produced the net greenhouse gas emission of 637 kg CO_2 eq./ton of MSW treated.

For landfilling, the greenhouse gas emissions in the life cycle perspective case is equal to that in the direct activity, 1313 kg CO_2 eq./ton of MSW treated, due to a very slight effect on greenhouse gas emission from the indirect activities.

Incineration performs better than landfilling for both direct activity as well as life cycle perspective. However, as before, the life cycle perspective offers more advantage to incineration due the conventional electricity production displaced. Landfilling would similarly benefit from methane collection and energy conversion as discussed later in this study.

5. Interpretation

A closer analysis of the results from the previous section revealed that incineration required a large amount of energy for waste combustion. This may be due to high moisture content of the waste to be burnt (41%). The performance of the incineration in the case of energy consumption might be better if more biodegradable waste

Table 3
Energy consumption of incineration and landfilling.

Activities	MJ/ton of MSW treated			
	Direct activity consideration		Life cycle perspective	
	Incineration	Landfilling	Incineration	Landfilling
MSW combustion (electricity)	311		311	
MSW combustion (diesel)	20		20	
Leachate treatment (electricity)	12		12	
Ash management (diesel)	11		11	
Energy recovery (electricity)	−530			
MSW spreading and compaction (diesel)		22		22
Electricity production (primary energy)			−1458	
Lime production			46	
Diesel production			10	7
Net energy consumption	−176	22	−1048	30

Table 4
Greenhouse gas emission of incineration and landfilling.

Activities	kg CO_2 eq./ton of MSW treated			
	Direct activity consideration		Life cycle perspective	
	Incineration	Landfilling	Incineration	Landfilling
MSW combustion	737		737	
MSW degradation		1311		1311
MSW spreading and compaction		2		2
Lime production			9	
Diesel production			0.2	0.3
GHG avoidance due to energy recovery			−109	
Net greenhouse gas emission	737	1313	637	1313

is separated before sending to the incinerator. In terms of greenhouse gas emissions, high amount of fossil carbon dioxide is emitted due to plastic burning. This performance might be improved by increasing the percent efficiency of plastic separation. The energy recovered from waste incineration, however, would be decreased because a large fraction with high energy content is separated out. Thus, the strategy of plastic separation to be treated by either landfilling or recycling should be assessed against the strategy of incineration including plastic fraction.

The emission of greenhouse gas from landfilling, which is dominated by methane, would be substantially reduced by introduction of gas collection and flaring system to convert the methane gas to carbon dioxide which, being of biomass origin, will not contribute to global warming. Additional credits can be obtained if the collected methane is utilized as an energy source.

For the purpose of estimation, the heating value of methane is set at 50 MJ/kg. It is assumed that 50% of CH_4 produced is collected to produce electricity by an engine with 35% efficiency and 10% of uncollected CH_4 is oxidized in landfill cover. Therefore, 45% of CH_4 is emitted to the atmosphere. Results from the analysis of landfilling with energy recovery indicate that 50% of greenhouse gases are reduced resulting in a net greenhouse gas emission of only 657 kg CO_2 eq./ton of MSW treated, based on direct activity consideration, as a consequence of reduction of CH_4 emitted from landfill. Credits obtained from electricity production from the recovered methane leads to further reduction of greenhouse gas when considered in life cycle perspective. About 59% of greenhouse gas reduction is obtained resulting in net greenhouse gas emissions of 544 kg CO_2 eq./ton of MSW treated. Reduction of greenhouse gas emissions by recovering landfill gas and producing electricity makes landfilling preferable to incineration. The same trend reversal is observed for the case of energy. When the methane is collected and electricity produced, landfill turns out to perform better than incineration in a life cycle perspective. In this case, the net energy consumption of landfill is −1494 MJ/ton of MSW treated, as compared to −1048 MJ/ton of MSW treated for incineration.

6. Conclusions

This study compares energy consumption and greenhouse gas emissions from landfilling (without energy recovery) against incineration (with energy recovery) based on a direct activity (base case) consideration and a life cycle perspective. For both cases and both parameters, incineration was found to be superior to the landfilling. However, landfilling reversed to be superior when landfill gas is recovered for electricity production.

This study demonstrates that a complete picture of the environmental performances of MSW management systems is provided by using a life cycle perspective. Life cycle assessment could serve as an invaluable tool for assessing environmental sustainability of waste management systems, single as well as integrated ones.

The results of this study are dependent on the actual MSW characteristics and management in Phuket province. The results of the environmental evaluation in other areas may be different due to MSW characteristics, technology, spatial and temporal factors, and related information. Also, an integrated waste management system based on separation of waste, with different waste fractions going to different waste management technologies, would be more efficient than a single option such as incineration or landfilling. However, in the case of small communities where only a single option may be economically viable, the results of the study are of direct relevance.

Acknowledgments

The authors acknowledge the Phuket Municipality and Kumjornkit Construction Company Limited for providing data. The Joint Graduate School of Energy and Environment is also acknowledged for all support. Comments of the anonymous reviewers are gratefully acknowledged.

References

Arena, U., Mastellone, M.L., Perugini, F., 2003. The environmental performance of alternative solid waste management options: a life cycle assessment study. Chemical Engineering Journal 96, 207–222.

Chaya, W., Gheewala, S.H., 2006. Life cycle assessment of MSW to energy schemes in Thailand. Journal of Cleaner Production, in press.

Denison, R., 1996. Environmental life-cycle comparisons of recycling, landfilling, and incineration: a review of recent studies. Annual Review of Energy and Environment 21, 191–237.

ETH-ESU, 1996. "Öko-inventare von Energiesystemen" FRISCH-KNECHT et al. third ed. 1996.

Finnveden, G., Ekvall, T., 1998. Life cycle assessment as a decision-support tool—the case of recycling vs. incineration of paper. Resource, Conservation and Recycling 24, 235–256.

Finnveden, G., Johansson, J., Lind, P., Moberg, Å., 2000. Life Cycle Assessments of Energy from Solid Waste. Stockholms Universitet, Sweden.

Friðriksson, G. B., Johnsen, T., Bjarnasóttir, H. J., Slentnes, H., 2002. Guidelines for the use of LCA in the waste management sector. Nordtest Project no. 1537–01.

Harrison, K.W., Dumas, R.D., Barlaz, M.A., 2000. Life-cycle inventory model of municipal solid waste combustion. Journal of the Air and Waste Management Association 50, 993–1003.

IPCC, 2001. Third Assessment Report: Climate Change. Intergovernmental Panel on Climate Change.

ISO, 1997. Environmental standard ISO 14040, Environmental management—life cycle assessment—principal and framework, Reference Number: ISO 14040: 1997(E).

McDougall, F., White, P. R., 1998. The use of life cycle inventory to optimize integrated solid waste management system: a review of case studies. Paper presented at "systems engineering models for waste management" International Workshop in Göteborg, Sweden, 25–26 February 1998, ⟨http://www.entek.chalmers.se/~josu/art-fmc.htm⟩.

Mendes, R.M., Aramaki, T., Hanaki, K., 2004. Comparison of the environmental impact of incineration and landfilling in São Paulo City as determined by LCA. Resource, Conservation and Recycling 41, 47–63.

Ramaswamy, V., Boucher, O., Haigh, J., Hauglustaine, D., Haywood, J., Myhre, G., Nakajima, T., Shi, G.Y., Solomon, S., 2001. Radiative forcing of climate change. In: Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K., Johnson, C.A. (Eds.), Climate Change 2001: The Scientific Basis. Contribution to Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Ross, S., Evans, D., 2003. The environmental effect of reusing and recycling a plastic-based packaging system. Journal of Cleaner Production 11, 561–571.

Sandgren, J., Heie, A., Sverud, T., 1996. Utslipp ved håndtering av kommuanalt avfall. Statens forurensningstilsyn (SFT). TA-number 1336/1996. In: Friðriksson, G.B., Johnsen, T., Bjarnasóttir, H.J., Slentnes, H. (Eds.) Guidelines for The Use of LCA in the Waste Management Sector. Nordtest Project no. 1537–01. 2002.

Wanichpongpan, W., Gheewala, S.H., 2006. LCA as a decision support tool for landfill gas-to-energy projects. Journal of Cleaner Production, in press.

TEI, 2003. Final Report for the Project on Life Cycle Assessment for Asian Countries—Phase III. Thailand Environmental Institute, Thailand.

UNEP, 2003. Environmental management tools: life cycle assessment, available from: ⟨http://www.uneptie.org/pc/pc/tools/lca.htm⟩.

ELSEVIER

Journal of Environmental Management 87 (2008) 139–153

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Combining GIS with fuzzy multicriteria decision-making for landfill siting in a fast-growing urban region

Ni-Bin Chang[a,*], G. Parvathinathan[b], Jeff B. Breeden[c]

[a]Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL 32816, USA
[b]Department of Environmental Engineering, Texas A&M University - Kingsville, Kingsville, TX 78363, USA
[c]Caesar Kleberg Wildlife Research Institute Texas A&M University - Kingsville, Kingsville, TX 78363, USA

Received 6 June 2006; received in revised form 28 October 2006; accepted 4 January 2007
Available online 23 March 2007

Abstract

Landfill siting is a difficult, complex, tedious, and protracted process requiring evaluation of many different criteria. This paper presents a fuzzy multicriteria decision analysis alongside with a geospatial analysis for the selection of landfill sites. It employs a two-stage analysis synergistically to form a spatial decision support system (SDSS) for waste management in a fast-growing urban region, south Texas. The first-stage analysis makes use of the thematic maps in Geographical information system (GIS) in conjunction with environmental, biophysical, ecological, and socioeconomic variables leading to support the second-stage analysis using the fuzzy multicriteria decision-making (FMCDM) as a tool. It differs from the conventional methods of integrating GIS with MCDM for landfill selection because the approach follows two sequential steps rather than a full-integrated scheme. The case study was made for the city of Harlingen in south Texas, which is rapidly evolving into a large urban area due to its vantage position near the US–Mexico borderlands. The purpose of GIS was to perform an initial screening process to eliminate unsuitable land followed by utilization of FMCDM method to identify the most suitable site using the information provided by the regional experts with reference to five chosen criteria. Research findings show that the proposed SDSS may aid in recognizing the pros and cons of potential areas for the localization of landfill sites in any study region. Based on initial GIS screening and final FMCDM assessment, "site 1" was selected as the most suitable site for the new landfill in the suburban area of the City of Harlingen. Sensitivity analysis was performed using Monte Carlo simulation where the decision weights associated with all criteria were varied to investigate their relative impacts on the rank ordering of the potential sites in the second stage. Despite variations of the decision weights within a range of 20%, it shows that "site 1" remains its comparative advantage in the final site selection process.

Keywords: Fuzzy MCDM; Geographic information systems; Decision support system; Landfill siting; Solid waste management

1. Introduction

New approaches to the sustainable planning, design, and management of urban regions will depend upon improvements in our knowledge of the causes, chronology, and impacts of the urbanization process and their driving forces (Klostermann, 1999; Longley et al., 2001). Worsening conditions of crowding, housing shortages, insufficient or obsolete infrastructure, increasing urban climatological and ecological problems, and the issues of urban security underline a greater than ever need for effective management and planning of urban regions (O'Meara, 1999). Innovative approaches to urban land use planning and management, such as sustainable development and smart growth, have been proposed and widely discussed (American Planning Association, 2002; Kaiser et al., 1995). Landfill selection in an urban area is a critical issue in the urban planning process because of its enormous impact on the economy, ecology, and the environmental health of the region. Landfill site selection can generally be divided into two main steps: the identification of potential sites through preliminary screening, and the evaluation of their suitability based on environmental impact assessment,

*Corresponding author. Tel.: +1 407 7547521.
E-mail address: nchang@mail.ucf.edu (N.-B. Chang).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.011

Nomenclature

S_{itj}	average fuzzy appropriateness index rating of alternative
W_t	average importance weight of criterion
$q_{itj}, o_{itj}, p_{itj}, c_{tj}, a_{tj}, b_{tj}$	triangular fuzzy numbers
$\oplus$ and $\otimes$	fuzzy addition and fuzzy multiplication operator
F_i	fuzzy appropriate indices of m alternatives
$U_M(F_i)$	optimistic utility for each appropriate index F_i
$U_G(F_i)$	pessimistic utility for each appropriate index F_i
R	total index of rating attitude
Y	the index of rating attitude of an individual decision maker
m	total number of alternatives
k	total number of criteria
n	total number of decision makers
α	index of rating attitude
L	candidate sites for landfill
E	experts
C	criteria

economic feasibility, and engineering design, and cost comparison (Charnpratheep et al., 1997). As a consequence, landfill siting is a difficult, complex, tedious, and protracted process (Allanach, 1992). Many siting factors and criteria should be carefully organized and analyzed. An initially chosen candidate site may be later abandoned because opposition arises due to previously neglected but important factors. Such a delay increases costs and postpones the final decision of a landfill site. The "not in my backyard" (NIMBY) and "not in anyone's backyard" (NIABY) phenomena is becoming popular nowadays creating a tremendous pressure on the decision makers involved in the selection of a landfill site. Other issues related to the availability of land, public acceptance, increasing amounts of waste generation complicate the process of selection of a suitable site for landfill. An inappropriate waste facility may adversely affect the surrounding environment and other economic and socio-cultural aspects.

Criteria and methodologies used for the initial screening are so pragmatic that areas are excluded as matters of social and environmental significance without removing large numbers of technically advantageous sites from consideration. The criteria used for preliminary screening are primarily to examine the proximity of potential sites with respect to geographic objects that may be affected by the landfill siting (e.g., groundwater wells) or that may affect landfill operations (e.g., areas with steep slopes). Methodologies used are normally based on a composite suitability analysis using thematic map overlays (O'Leary et al., 1986) and their extension to include statistical analysis (Anderson and Greenberg, 1982). With the development of geographical information systems (GIS), the landfill siting process is increasingly based on more sophisticated spatial analysis and modeling. Jensen and Christensen (1986) demonstrated the use of a raster-based GIS with its associated Boolean logic map algebra to identify potential waste sites based on suitability of topography and proximity with respect to key geographic features, while Keir et al. (1993) discussed the use of both raster-based and vector-based GIS for the full-scale site selection process. Sener et al. (2006) integrated GIS and multicriteria decision analysis (MCDA) to solve the landfill site selection problem and developed a ranking of the potential landfill areas based on a variety of criteria. The utilization of GIS for a preliminary screening is normally carried out by classifying an individual map, based on selected criteria, into exactly defined classes or by creating buffer zones around geographic features to be protected. All map layers are then intersected so that the resulting composite map contains two distinct areas. For example, if screening criteria involve the provision of a protective buffer around certain types of spatial objects, the area outside the intersected boundary is considered suitable and that inside is unsuitable. The two distinct classes separated by a sharp boundary reflect the representation of, and GIS operations on, geo-referenced data based on a binary true or false Boolean logic. With the aid of this functionality, GIS have been used in order to facilitate and lower the cost of the process of selection of sites for building sanitary landfills in the last few years (Siddiqui et al., 1996; Kao et al., 1997).

Advanced algorithms, however, may further help justify the uncertainty in siting new landfills. Several approaches were proposed for multicriteria decision-making (MCDM) and the relevant methods were developed and applied with more or less success depending on the specific problem. In the past, analytic hierarchy process (AHP) introduced by Saaty (1980), was one of the useful methodologies, which plays an important role in selecting alternatives (Fanti et al., 1998; Labib et al., 1998; Chan et al., 2000). AHP is an analytical tool enables people to explicitly rank tangible and intangible criteria against each other for the purpose of selecting priorities. The process involves structuring a problem from a primary objective to secondary levels of criteria and alternatives. Once the hierarchy has been established, a pair-wise comparison matrix of each element within each level is constructed. The AHP allows group decision-making, where group members can use their experience, values and knowledge to break down a problem into a hierarchy and solve it by the AHP steps. Participants can weigh each element against each other within each level, each level is related to the levels above and below it, and the entire scheme is tied together mathematically. For evaluating the numerous criteria, AHP has become one of the most widely used methods for the practical solution of MCDM problems (Cheng, 1997; Akash et al., 1999; Chan et al., 2000). The main difficulty arises in the estimation of the required input data that express qualitative observations and preferences. The AHP is mainly used in nearly crisp decision applications. It does not take into account the uncertainty associated with the mapping of people's judgment to an evaluation scale (Chen, 1996; Hauser and Tadikamalla, 1996; Cheng, 1997). In order to overcome the shortcomings of the AHP, fuzzy set principle is used to integrate AHP to determine the best alternative (Chen, 1996; Hauser and Tadikamalla, 1996; Levary and Ke, 1998).

Fuzzy set theory was developed and extensively applied in previous decade (Zadeh, 1965). It was designed to supplement the interpretation of linguistic or measured uncertainties for real-world uncertain phenomena. These uncertainties could originate with non-statistical characteristics in nature that refer to the absence of sharp boundaries in information. However, the main source of uncertainties involving in a large-scale complex decision-making process may be properly described via fuzzy membership functions. The practical applications of fuzzy multicriteria decision-making (FMCDM) reported in the literature have shown advantages in handling unquantifiable/qualitative criteria and obtained quite reliable results (Altrock and Krause, 1994; Teng and Tzeng, 1996; Baas and Kwakernaak, 1997; Mcintyre and Parfitt, 1998; Tang et al., 1999). Fuzzy linguistic models permit the translation of verbal expressions into numerical ones, thereby dealing quantitatively with imprecision in the expression of the importance of each criterion. FMCDM utilizes linguistic variables and fuzzy numbers to aggregate the decision makers' subjective assessment about criteria weightings and appropriateness of alternative candidate sites versus selection criteria to obtain the final scores—fuzzy appropriateness indices.

Extended application can be found in developing a decision support system (DSS) (Kuo et al., 2002). The proposed DSS consists of four components: (1) hierarchical structure development for fuzzy AHP, (2) weights determination, (3) data collection, and (4) decision-making (Kuo et al., 2002). Therefore, the integration of fuzzy set and AHP gives a much better and more exact representation of relationship between criteria and alternatives (Lee et al., 1998; Chiadamrong, 1999; Yu and Skibniewski, 1999; Choi and Oh, 2000; Karsak and Tolga, 2001). FMCDM methods have been used in environmental planning and decision-making processes in order to clarify the planning process, to avoid various distortions, and to manage all the information, criteria, uncertainties, and importance of the criteria. This paper presents an integrated approach to construct a spatial decision support system (SDSS) for the selection of landfill sites via a two-stage analysis synergistically. The first-stage analysis makes use of the thematic maps in GIS in conjunction with environmental, biophysical, ecological, and socioeconomic variables leading to support the second-stage analysis using FMCDM as a tool. In essence, in the first stage, the geographical data were analyzed using GIS and a data matrix was created that combines the environmental, transportation, public health, social, and economic criteria for the selection of seven-candidate sites. It eventually generates the ranking of all seven-candidate sites in a preferential order based on different criteria involved collectively in a FMCDM analysis. The case study was made for the city of Harlingen in south Texas, which is rapidly evolving into a large urban area due to its vantage position near the US–Mexico borderlands.

2. Background and study site

The Lower Rio Grande Valley (LRGV or Valley), comprised of Cameron, Willacy, Hidalgo, and Starr counties, is located at the southernmost tip of Texas along the US Mexico border. The Office of Management and Budget (OMB) ranks Metropolitan statistical area (MSA) according to their population and economic growth. Cameron County, at the tip of Texas, comprises 3266 km^2 (1276 square miles) and includes the 28th MSA, Brownsville-Harlingen-San Benito. Hidalgo County, the largest of the three LRGV counties, covers the western half of the region with an area of 3963 km^2 (1548 square miles). This county is mostly urbanized, containing the McAllen-Edinburg-Mission MSA, the 4th fastest growing areas in Texas. Both of the LRGV's MSAs are experiencing a developmental change due to their strategic location and economic ties with the US–Mexico borderland. The North American Free Trade Agreement (NAFTA) that was enacted in 1994 has increased trade throughout America. The Valley, with a total area of 9216 km^2 (3600 square miles), has emerged as a warehouse and transportation center between Central America and the US (TSHA, 2003). The increasing number of maquiladoras, or twin plants, having manufacturing industries both in the MSAs of the Valley and in nearby Reynosa and Matamoros, Mexico, are positively influencing the economic development in the region. This has been a catalyst for further growth in other Valley cities located in between these two MSAs. As a result, the population of the LRGV is growing at a tremendous pace and yard waste, food waste, and biosolids waste production is increasing over time. Fig. 1 indicates the study area along with the waste disposal sites. The area's population has increased by 39.8% in the last 10 year due to the NAFTA's economic impact. It is expected to continue growing at an estimated rate of 4% per year in the coming years. The population is projected to be over 1.7 million people in 2022 (LRGVDC, 2002).

Solid waste management (SWM) is at the forefront of environmental concerns in the LRGV, South Texas. The complexity in SWM drives area decision makers to look for innovative and forward-looking solutions to address various waste management options. The LRGV is facing the difficult reality of siting new landfills due to their large capital costs and local protests, like those seen as a result of

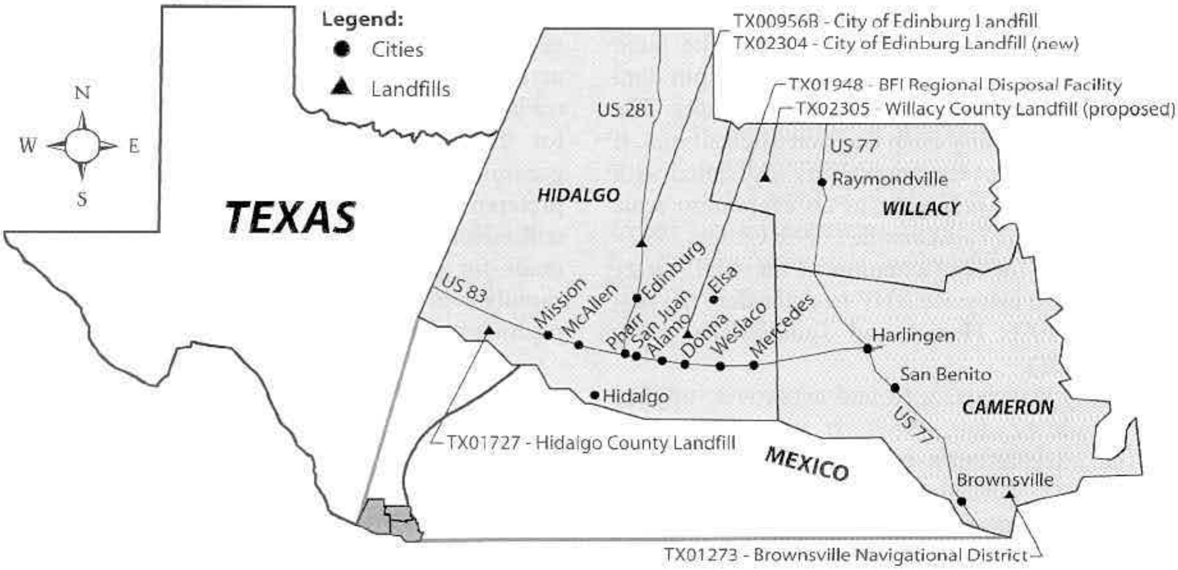


Fig. 1. Map of South Texas indicating the facility location of the study area.

Willacy County's intentions to site a new landfill. The hotly contested landfill permit process culminated in a hearing August 1, 2005 with the decision pending whether to allow the process to continue amid community resistance (Del Valle, 2005). Adding to the complexity of the issue, the realization by local residents of the economic value of their ecosystem from tourism dollars generated from bird watching enthusiasts means siting future landfills could become more contested.

The ability to give planners more options will allow for compromise surrounding potential disposal alternatives. Privately owned landfills have been viewed in a negative light as of late due to litigation against (Browning-Ferris Industries) BFI that has essentially jeopardized its long-term ability to operate and that closed the C&T landfill (Pierson, 2004). These underpinnings place an emphasis on giving regional planning partners like the council of government "Lower Rio Grande Development Council" as many possible alternatives to consider in their planning. The political and environmental climate surrounding SWM in the LRGV lends itself to analysis that embraces uncertainty in as many possible decision making levels as possible. In order to better inform area decision makers about their options to cope with the mounting municipal solid waste (MSW) generation and a general lack of landfill space, a SDSS needs to be developed to address regional planning around issues of waste routing and the hotly contested SWM facility site selection. The study location (Harlingen) is one of the fastest growing cities in the LRGV. Presently the city generates huge amount of solid waste that is disposed off at B.F.I. Landfill in Donna at high cost per ton, which is expected to increase in the next 7 year of period of contract. The transportation cost of ton of solid waste is also high. Due to these huge costs of waste disposal, the city has plans on starting its own landfill.

Development of a landfill in Harlingen can possibly cause environmental impacts on the soil, groundwater, surface water, regional air quality, atmosphere, biodiversity, and landscape. Besides these environmental impacts, there are those related to the economy, employment, attainability and valuation of different areas, services, safety, and health. A landfill in this region can also affect many of the endangered and threatened species that occur at their northernmost limit in the LRGV. In light of such circumstances, there is acute necessity for a careful selection of a landfill site in order to preserve the ecological and environmental quality that's unique to the LRGV.

3. Case study

Landfill siting is a complicated process requiring a detailed assessment over a vast area to identify suitable location for constructing a landfill subject to many different criteria. GIS offers the spatial analysis capabilities to quickly eliminate parcels of land unsuitable for landfill site. This study employed GIS to perform a screening process that led to identification of a couple suitable candidate sites based on given criteria. The suitability criteria are defined with the focus to minimize any potential health risks from direct or indirect contamination due to the proximity of a landfill site with respect to key geographic features. Thus, the first-stage analysis using GIS is essential for the initial identification of a couple suitable landfill sites prior to undertaking further analyses or field investigations. Although, the initial screening is based on criteria related to environmental and ecological factors involved in the site selection process, there are certain criteria, such as impact on historical markers, public comfort, and economic factors for which data are not always readily available, which cannot be included in the first stage. A second-stage analysis based on a handful of suitable sites from the initial GIS screening was performed with the objective of including the opinions of domain experts in the region through a FMCDM approach. FMCDM was useful in addressing the issue of lack of availability of data for certain important criteria as well as to incorporate human judgment into the selection process that can prove useful in solving political debates in the future. The second-stage analysis using FMCDM was applied to rank the proposed candidate sites and summarize the final selection. Such method followed in the process of identifying the most suitable landfill site is described in the next two sub-sections. To ease the illustration, the following sub-sections would delineate or review the methods briefly and then come up with the results and discussions directly. The list of variables and parameters that were used in the FMCDM analysis is summarized in Nomenclature.

3.1. Data collection and analysis

GIS data sets of land-use, rivers, wetlands, roads, demography, wildlife parks, airports, soil types, groundwater wells, and digital elevation models (DEMs) were collected for the Cameron, Hidalgo and Willacy counties from different sources, such as Texas Natural Resources Information Systems (TNRIS), Texas Department of Transportation, US Geological Survey (USGS), and US Environmental Protection Agency (USEPA). They were summarized as shown in Table 1. Geographical features required fir the first-stage analysis could be extracted by using ArcGIS® software. For example, to obtain GIS data sets of buffer zone, the land in the LRGV was classified by creating buffer zones around geographic features to be protected using literature values widely used in landfill selection process. The buffer maps were then converted into raster maps of uniform grid sizes and the raster calculator available in spatial analyst tool in ArcGIS® was utilized to eliminate unsuitable land parcels based on the different criteria leading to identification of seven potential landfill sites in the first stage.

Table 1
GIS map layers used in the study

Data	Scale	Data source
Rivers	1:500 000	EPA
Lakes	1:250 000	EPA
Wetland	1:250 000	EPA
Land use/land cover	1:250 000	EPA
Roads	1:100 000	EPA
Ground water wells		TWDB (Texas Water Development Board)
Urban areas	1:24 000	EPA
Soil map STATSGO	1:250 000	USGS
Digital elevation model	1:250 000	EPA basins
County census data	1:2 000 000	Tiger data

3.2. Application of GIS in landfill candidate site selection

Fig. 2 illustrates the typical procedure applying the GIS for initial landfill siting. The landfill site selection process was completed in two stages with the first stage utilizing GIS to identify a few candidate sites that were later ranked using FMCDM method in the second stage. There are several different criteria involved in the selection of a landfill site in the first stage. Literature review was conducted to identify the most important criteria. According to Dikshit et al. (2000), a landfill site must be situated at a fair distance away from biophysical elements such as water, wetlands, critical habitats, and wells to reduce the risk of contamination from landfill. Different studies used different buffer distances from stream and rivers based on the size of the watershed, such as buffer of 0.8 km (Siddiqui et al., 1996), 180 m (Zeiss and Lefsrud, 1995) and 2–3 km (Lin and Kao, 1999). Considering the size of Harlingen city, a buffer distance of 1 k was used for river system in this study.

Proximity of a landfill to a groundwater well is an important environmental criterion in the landfill site selection so that wells may be protected from the runoff and leaching of the landfill. For this study, groundwater wells data were obtained from Texas Water Development Board (TWDB), and a buffer distance of 50 m from the wells was used to prevent contamination from landfill due to leaching of pollutants. Slope is also an important factor when siting a landfill since higher slopes would increase runoff of pollutants from the landfill, and thereby increasing the contamination zone area (Lin and Kao, 1999). Lin and Kao's (1999) study suggested that a slope less than 12% would be suitable for the prevention of contaminant runoff. Based on this study, regions with slope greater than 12% were defined as unsuitable for a landfill site. DEM data sets with 30 M resolution obtained from USEPA basins data source were used to calculate the slope percentage area wide. In addition, the landfill should be situated at a significant distance away from urban residential areas due to public concerns, such as aesthetics, odor (Tagaris et al., 2003), noise, decrease in property value (Zeiss and Lefsrud, 1995), and health concerns, which may avoid contamination of freshwater aquifers through leaching (Nagar and Mizra, 2002). Urban buffers may range from 150 m (Lin and Kao, 1999) to 5 km (Zeiss and Lefsrud, 1995). A buffer distance of 3 km was chosen for the study area.

Economic considerations include finding the most cost effective route for transporting wastes and locating the most suitable land for the candidate sites based on land value (Siddiqui et al., 1996). Developments on or too close to existing road and rail networks would hinder transportation and may have an impact on tourism in the region (Zeiss and Lefsrud, 1995). Baban and Flannagan (1998) used a 50-m buffer for roads, while Dikshit et al. (2000) used a 1-km buffer in his study. However, a study done by Lin and Kao (1999) stated that a 1 km buffer was too far

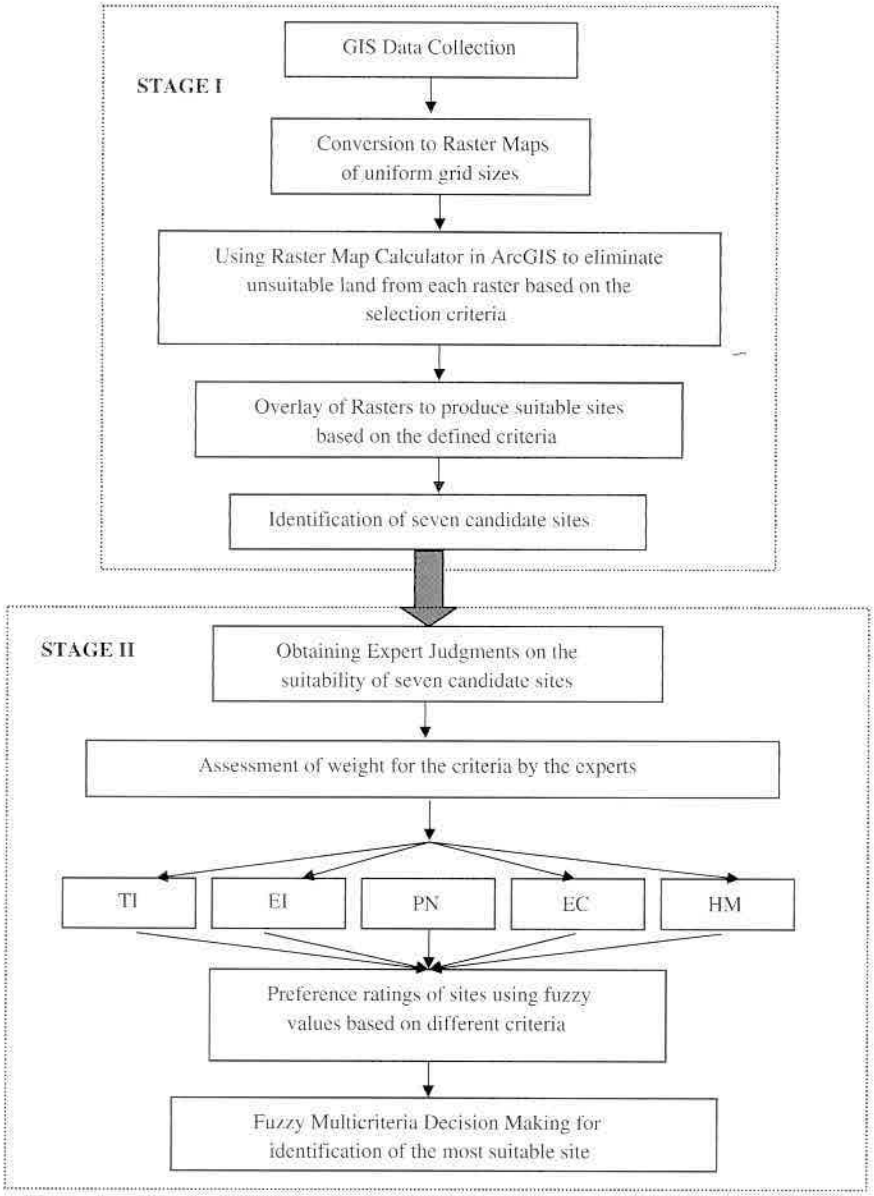


Fig. 2. Flowchart of the methodology. *Note*: EI = environmental and ecological impact, TI = transportation issues, PN = public nuisance, EC = economical Impact, HM = impact on historical markers.

from roadways, and would result in incurring more economic costs to the project over the long term by constructing new roads. Considering the huge cost of transportation, a 75-m buffer for roads was finally selected for this study.

The different constraint maps developed in this study include an environmental constraint map, a stream constraint map, a wells constraint map, a slope constraint map, an urban constraint map, a water body constraint map, and a transportation constraint map. The obtained constrained map layers are overlaid as shown in Fig. 3, and final constraint maps were developed with the candidate sites, as shown in Fig. 4. Fig. 5 shows the seven-candidate sites in GIS, which are subject to advanced assessment in the second-stage analysis.

Besides, ecological assessment study states that the region is divided into several ecoregions based on topographic, climatic and edaphic factors, and plant community similarities. These ecoregions are characterized by high summer temperatures, high evaporation rates, and periodic droughts. The seven-candidate sites are currently in use as agricultural cropland and have been cleared of native vegetation. Soils have a direct effect on the types of vegetation and ultimately the animal species that will occur in an area. The US Department of Agriculture (1977 and 1982) rated the potential for soil types throughout Cameron and Willacy counties to provide elements of habitat for various species of wildlife. The soils are also rated on their potential for wildlife species to occur. Based on the criteria of the US Department of Agriculture (1977,

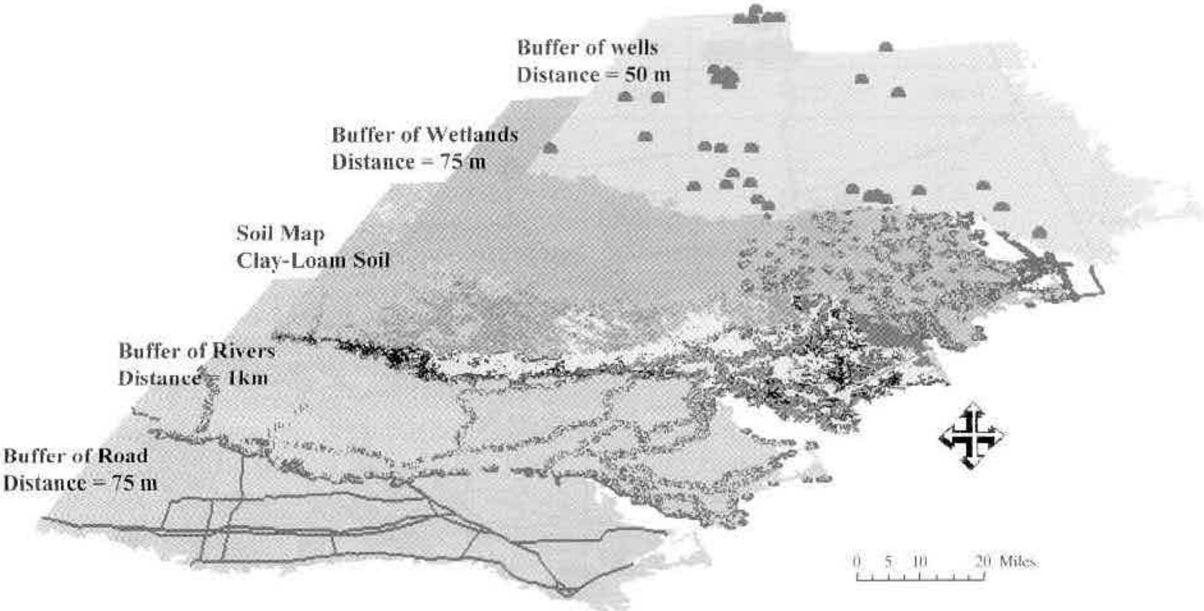


Fig. 3. Overlay of different constrained maps.

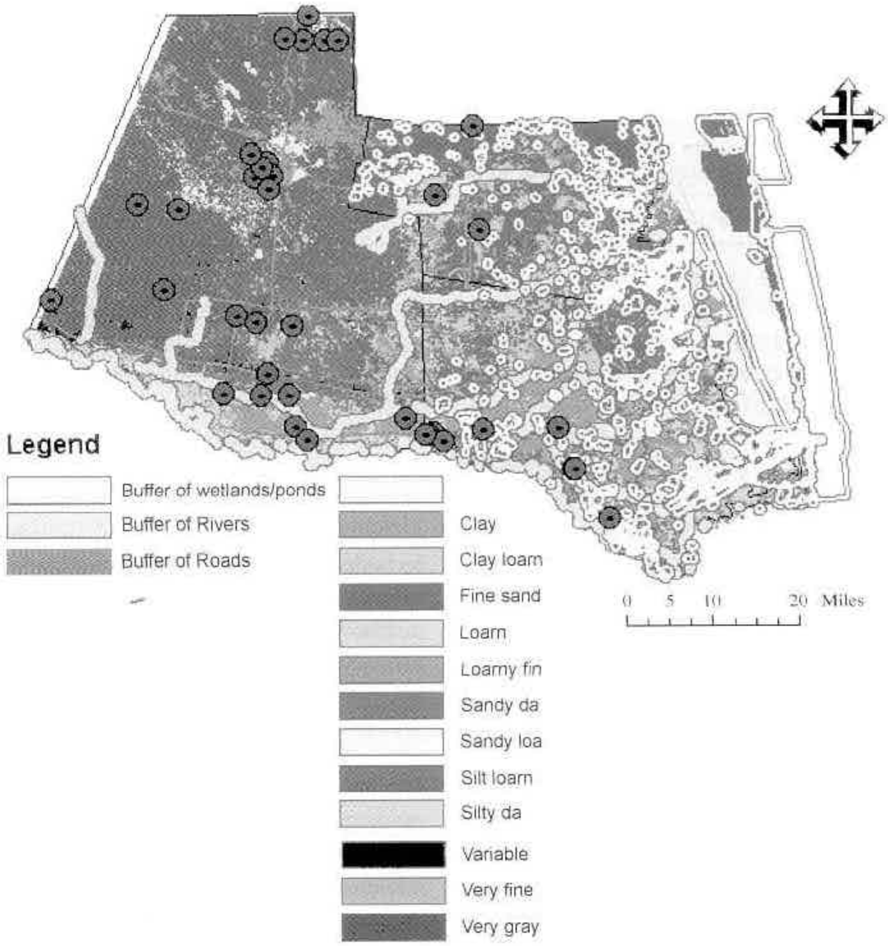


Fig. 4. Final map showing different constrained maps.

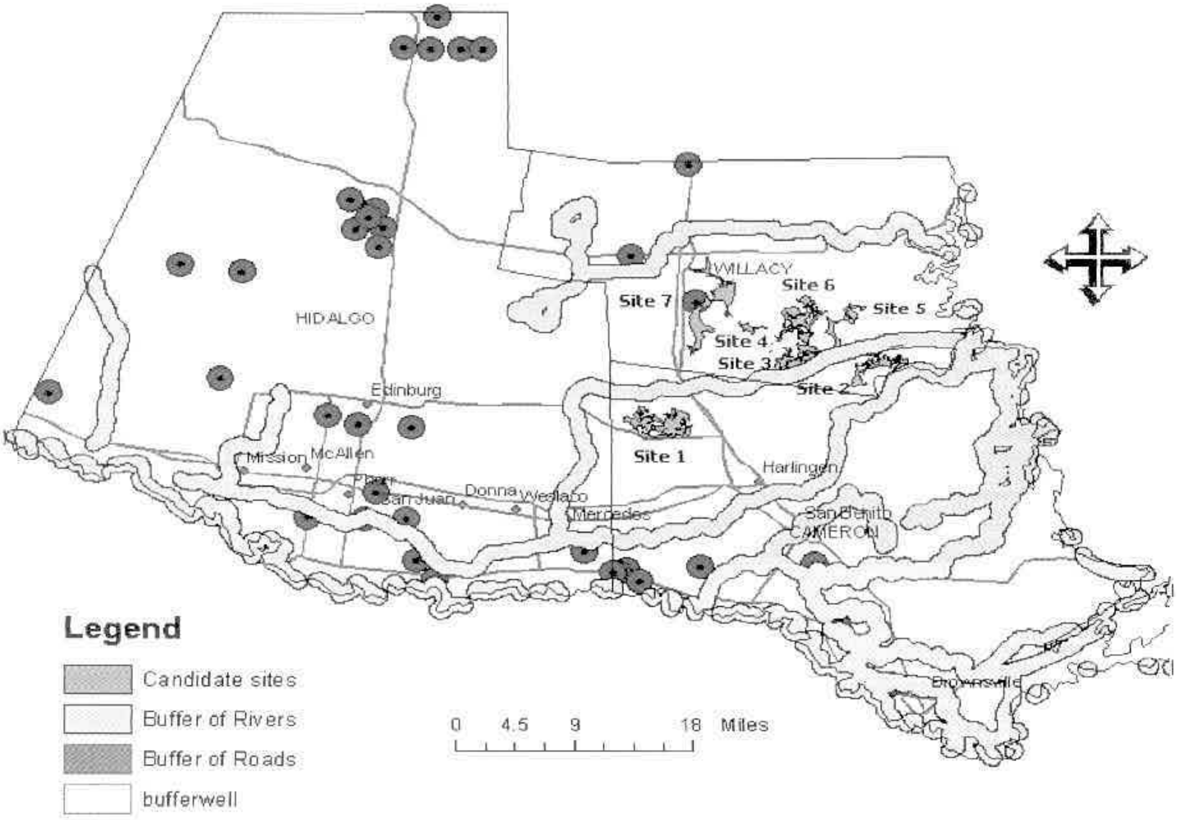


Fig. 5. Map showing the candidate sites for landfill with different constraints.

Table 2
Comparison of seven-candidate sites using potential for elements of wildlife habitat

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	
Soil type[a]	1	1	1	1	1	1	1	2
Coverage %	95	99	99	100	100	100	79	17
Grains	Good	Good	Good	Good	Good	Good	Good	Fair
Grasses	Good	Good	Good	Good	Good	Good	Good	Fair
Herbaceous plants	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Shrubs	Good	Good	Good	Good	Good	Good	Good	Fair
Wetland plants	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Good
Shallow wetlands	Very poor	Very poor	Very poor	Very poor	Very poor	Very poor	Very poor	Good

[a] 1—Raymondville clay loam, 2—Mercedes clay.

1982), a rating of good indicates that the kind of habitat is easily established and maintained. A rating of fair indicates that the kind of habitat can be established with moderately intensive management. A poor rating indicates that the habitat type can be established, but with intensive and difficult management. A very poor rating indicates that creating or maintaining the habitat type is impractical or impossible.

The terrain in south Texas is quite flat and all candidate sites are managed as agricultural land at present except site 7. Future landfill to be built in this area should be designed as a plain-type rather than a gully-type landfill so that soil thickness was not an obvious issue on site. Thus, the proposed criteria did not include soil thickness and depth to bedrock, which may hamper the excavatibility of the site in some cases.

The potential for elements of wildlife habitat to occur and their ratings are compared across the seven-candidate sites in Table 2. Table 3 lists and compares the potential for types of wildlife species to occur in the seven-candidate sites. Based on the ecological assessment study, all candidate sites are similar in soil type and similar in the potential for wildlife habitat and wildlife species. Because of the similarity between all sites, the potential effects on endangered and threatened species are the same for all candidate sites. Candidate sites 1–6 would result in the same ecological effect of any actions. Candidate site seven is slightly different than the other six sites because an

Table 3
Comparison of potential types of wildlife species occur in the seven-candidate sites

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	
Soil type[a]	1	1	1	1	1	1	1	2
Coverage %	95	99	99	100	100	100	79	17
Rangeland Wildlife	Fair	Fair	Fair	Fair	Fair	Fair	Fair	Fair
Open land Wildlife	Good	Good	Good	Good	Good	Good	Good	Fair
Wetland Wildlife	Very poor	Very poor	Very poor	Very poor	Very poor	Very poor	Very poor	Good

[a] 1—Raymondville clay loam. 2—Mercedes clay.

additional soil type occurs there. This soil type is better suited for the potential occurrence of wetland habitat and wetland wildlife species than the predominant soil type found in the other candidate sites. It is commonly known that wetlands are an important component of the ecosystem and are diminishing across the country. Therefore, candidate site seven would be the most ecologically sensitive site by any action because of the potential impact on wetland habitat.

3.3. Fuzzy multicriteria decision-making

The second-stage analysis for landfill site selection requires having a careful evaluation of the advantages and disadvantages of different candidate sites with respect to different predetermined criteria because landfill siting is a complicated process that leads to different impacts in the area. Due to lack of crisp data, the evaluation of different alternatives against different criteria requires assessment using fuzzy numbers. FMCDM method is therefore chosen for ranking different landfill sites for Harlingen city based on decisions given by a group of experts. Experts or planners were called on for participating in a questionnaire survey using linguistic variables or fuzzy numbers to give the preference ratings for each individual candidate sites.

Chang and Chen (1994) proposed a new MCDM method to solve the distribution center location selection problem under fuzzy environment. The ratings of each alternative and the weight of criterion are described by linguistics variables that can be expressed in triangular fuzzy numbers. The evaluation value of each facility site is also expressed in a triangular fuzzy number. By calculating the difference of evaluation value between each pair of candidate sites, a fuzzy preference relation matrix is constructed to represent the intensity of the preferences of one plant location over another. Then, a stepwise ranking procedure is proposed to determine the ranking order of all candidate locations. When conducting the inference, triangular fuzzy numbers (TFN) are commonly used by the experts to describe vagueness and ambiguity in the real-world system. Many methods, such as max, min, median, addition, multiplication, and mixed operators, are available to aggregate TFNs. Related literature can be found in (Kaufmann and Gupta, 1988; Paek et al., 1992).

The experts can employ an assumed weighting set W = {Very poor, Poor, Fair, Good, and Very good} to evaluate the appropriateness of the alternatives versus various criteria. The membership functions of the linguistic values in the weighting set W represented by the approximate reasoning of triangular fuzzy numbers are shown in Fig. 6. If one does not agree with the assumed preference rating system, one can give his own rating by using the triangular fuzzy number, showing perception of the linguistic variables, 'importance' and 'appropriateness'.

The different criteria that were selected for evaluating the merits of the different landfill sites are: (1) environmental and ecological impact, (2) transportation issues, (3) impacts on historical markers, (4) economic impacts of the landfill and (5) public nuisance. These criteria are described below. Transportation of waste loads from the hauling station to the landfill causes disruption of traffic within the city limits that cannot be clearly quantified in the decision-making process, thereby requiring fuzzy description of the criteria. Similarly, the possible impacts that can be caused by landfill on historical markers in terms of aesthetical impairment; bad odors etc. are critical and vague and hence, require fuzzy concepts to represent the importance of historical makers on the landfill selection process. The criterion of economical impact reflects the possibility of decrease in land value in the neighborhood and also in the farming productivity of the region, thereby affecting the economy of the city directly, which is also vague in many other ways. Public nuisance is another vague but important factor that refers to the feeling of discomfort caused to the public due to the construction and operation of a landfill in the middle of a populous place.

The decision objective is to select the most appropriate landfill from seven different candidate sites. The different alternatives are defined as $L = \{L1, L2, L3, L4, L5, L6, L7\}$ and the decision criteria are defined as C = {TI, EI, PN, EC, HM}, where {TI = transportation issues, EI = environmental and ecological impact, PN = public nuisance, EC = economical impact, HM = historical markers}. Linkage between different alternatives with different criteria is shown in Fig. 7. There is a committee of two experts (E1 and E2) who are called on for assessing the appropriateness of 'm' alternatives ($\{L1, L2, L3, L4, L5, L6, L7\}$) under each of '$k$' criteria ({TI, EI, PN, EC, HM}) as well as the importance weight of the criteria.

Let S_{itj} ($i = 1, 2, \ldots, m$; $t = 1, 2, \ldots, k$; $j = 1, 2, \ldots, n$) be the rating assigned to alternative A_i by expert E_j under criterion C_t. Let W_{tj} be the weight given to C_t by decision

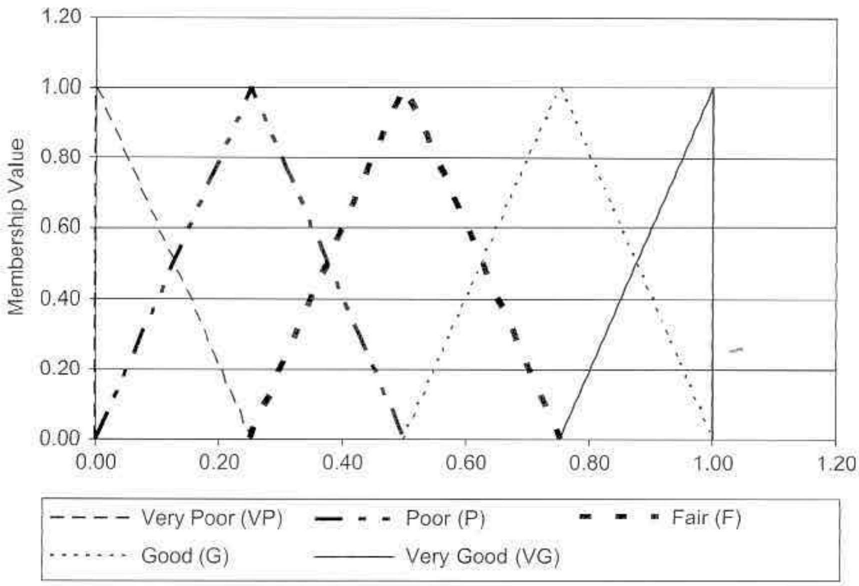


Fig. 6. Fuzzy membership functions.

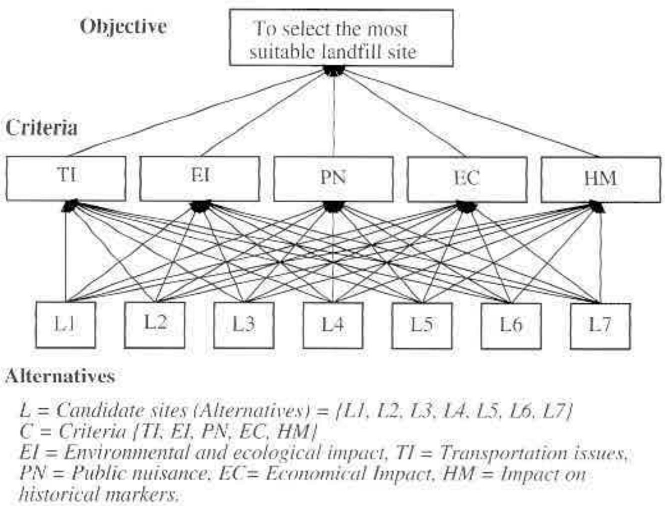


Fig. 7. Description of decision-making process.

maker E_j. The rating S_{itj} of n experts for each alternative vs. each criterion is aggregated. Each pooled rating is further weighted by weight W_t according to the relative importance of the k criteria. Then the final score F_i, fuzzy appropriate index, of alternative A_i is obtained by aggregating S_{itj} and W_t, which is finally ranked to obtain the most suitable alternative (Chang and Chen, 1994). The experts give their own preference rating for the different alternatives and weights for different criteria by using the triangular fuzzy numbers. Tables 4 and 5 present the rating done by the two experts comparing the seven alternatives (i.e., candidate sites) against the five criteria. The weights assigned to the different criteria for decision-making are presented in Table 6.

Following the method developed by Chang and Chen (1994), this paper utilizes mean fuzzy operator to aggregate the expert's assessment. Let $\oplus$ and $\otimes$ be the fuzzy addition and fuzzy multiplication operator, respectively. The aggregation of the different ratings is given by

$$S_{itj} = (S_{it1} \oplus S_{it2} \oplus \cdots S_{itn}) \otimes (1/n), \quad (1)$$

$$W_t = (W_{t1} \oplus W_{t2} \oplus \cdots W_{tn}) \otimes (1/n), \quad (2)$$

where S_{itj} is the average fuzzy appropriateness index rating of alternative A_i under criterion C_j, and W_t is the average importance weight of criterion C_j. Thus, the fuzzy appropriateness index F_i of the ith alternative can be obtained by aggregating S_{itj} and W_t, expressed as

$$F_i = [(S_{i1} \oplus W_1) \oplus (S_{i2} \oplus W_2) \oplus \cdots \oplus (S_{ik} \otimes W_k)] \otimes (1/k). \quad (3)$$

Let $S_{itj} = (q_{itj}, o_{itj}, p_{itj})$ and $W_{itj} = (c_{tj}, a_{tj}, b_{tj})$ be triangular fuzzy numbers. Then F_i can be expressed as

$$F_i = (Y_i, Q_i, Z_i), \quad (4)$$

where

$$Y_i = \sum_{t=1-k} (q_{it}c_t/k), \quad Q_i = \sum_{t=1-k} (o_{it}a_t/k),$$
$$Z_i = \sum_{t=1-k} (p_{it}b_t/k),$$

$$o_{it} = \sum_{j=1-n} (o_{itj}/n), \quad p_{it} = \sum_{j=1-n} (p_{itj}/n),$$
$$q_{it} = \sum_{j=1-n} (o_{itj}/n),$$

$$c_t = \sum_{j=1-n} (c_{tj}/n), \quad p_t = \sum_{j=1-n} (p_{tj}/n),$$
$$a_t = \sum_{j=1-n} (a_{tj}/n)$$

for $i = 1,2,\ldots,m$; $t = 1,2,\ldots,k$; $j = 1,2,\ldots,n$.

Table 4
Evaluation of different alternative against all criteria by expert E1

Criteria	Alternatives						
	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
TI	(0.6, 0.7, 0.8)	(0.6, 0.7, 0.8)	(0.6, 0.7, 0.8)	(0.6, 0.7, 0.8)	(0.6, 0.7, 0.8)	(0.6, 0.7, 0.8)	(0.6, 0.7, 0.8)
PN	(0.9, 0.95, 1.0)	(0.4, 0.5, 0.6)	(0.4, 0.5, 0.6)	(0.4, 0.5, 0.6)	(0.4, 0.5, 0.6)	(0.4, 0.5, 0.6)	(0.4, 0.5, 0.6)
EI	(0.7, 0.85, 0.9)	(0.6, 0.65, 0.75)	(0.4, 0.45, 0.55)	(0.50, 0.55, 0.60)	(0.4, 0.45, 0.55)	(0.5, 0.55, 0.65)	(0.5, 0.55, 0.65)
EC	(0.65, 0.75, 0.8)	(0.40, 0.45, 0.5)	(0.5, 0.55, 0.65)	(0.55, 0.6, 0.65)	(0.55, 0.6, 0.65)	(0.50, 0.60, 0.75)	(0.50, 0.60, 0.75)
HM	(0.50, 0.60, 0.75)	(0.55, 0.6, 0.65)	(0.55, 0.6, 0.65)	(0.55, 0.6, 0.65)	(0.55, 0.6, 0.65)	(0.55, 0.6, 0.65)	(0.55, 0.6, 0.65)

Table 5
Evaluation of different alternative against all criteria by expert E2

Criteria	Alternatives						
	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
TI	(0.55, 0.6, 0.70)	(0.35, 0.4, 0.45)	(0.4, 0.45, 0.5)	(0.45, 0.5, 0.55)	(0.4, 0.45, 0.5)	(0.5, 0.55, 0.6)	(0.3, 0.35, 0.4)
PN	(0.4, 0.45, 0.5)	(0.5, 0.55, 0.6)	(0.4, 0.45, 0.5)	(0.35, 0.4, 0.45)	(0.5, 0.55, 0.60)	(0.5, 0.55, 0.6)	(0.35, 0.4, 0.45)
EI	(0.75, 0.8, 0.85)	(0.5, 0.55, 0.60)	(0.55, 0.60, 0.65)	(0.30, 0.35, 0.40)	(0.50, 0.55, 0.60)	(0.30, 0.35, 0.40)	(0.55, 0.60, 0.65)
EC	(0.7, 0.75, 0.8)	(0.4, 0.45, 0.5)	(0.6, 0.65, 0.7)	(0.5, 0.55, 0.60)	(0.6, 0.65, 0.7)	(0.5, 0.55, 0.60)	(0.5, 0.55, 0.6)
HM	(0.45, 0.5, 0.55)	(0.45, 0.5, 0.55)	(0.45, 0.5, 0.55)	(0.45, 0.5, 0.55)	(0.45, 0.5, 0.55)	(0.45, 0.5, 0.55)	(0.45, 0.5, 0.55)

Table 6
Weights of different criteria by two experts

Criteria	Experts	
	E1	E2
TI	(0.8, 0.9, 0.95)	(0.8, 0.9, 0.95)
PN	(0.9, 0.95, 1)	(0.75, 0.8, 0.9)
EI	(0.7, 0.75, 0.8)	(0.85, 0.9, 0.99)
EC	(0.8, 0.9, 0.95)	(0.7, 0.75, 0.8)
HM	(0.45, 0.55, 0.6)	(0.45, 0.55, 0.6)

Based on the aggregation functions, the fuzzy appropriate indices are obtained and presented in Table 7. This information may help justify the final ranking among these seven-candidate sites. Therefore, the ranking values of fuzzy appropriate indices for the alternatives were computed based on the method developed in Chang and Chen (1994).

Let F_i $(i = 1,2,\dots,m)$ be the fuzzy appropriate indices of m alternatives. The maximizing set $M = \{(x, f_m(x)) | x \in R\}$ with

$$f_m(x) = \begin{cases} (x - x_1)/(x_2 - x_1), & x_1 < x \leqslant x_2, \\ 0 & \text{otherwise} \end{cases}$$

and minimizing set $G = = \{(x, f_g(x)) | x \in R\}$ with

$$f_g(x) = \begin{cases} (x - x_2)/(x_1 - x_2), & x_1 \leqslant x < x_2, \\ 0 & \text{otherwise,} \end{cases}$$

where $x_1 = \inf S$, $x_2 = \sup S$, $S = \cup_{i=1,m} F_i$, $F_i = \{x | f_{Fi}(x) > 0\}$, for $i = 1,2,\dots,m$.

Defining the optimistic utility $U_M(F_i)$ and pessimistic utility $U_G(F_i)$ for each appropriate index F_i as

$$U_M(F_i) = \sup(f_{Fi}(x) \wedge f_M(x)) \text{ and} \tag{5}$$

$$U_G(F_i) = 1 - \sup(f_{Fi}(x) \wedge f_G(x)). \tag{6}$$

For $i = 1,2,\dots$, where $\wedge$ means min.

Ranking value $U_T(F_i)$ of fuzzy appropriate indices is defined as:

$$U_T(F_i) = \alpha U_M(F_M) + (1 - \alpha) U_G(F_i), \quad 0 \leqslant \alpha \leqslant 1. \tag{7}$$

The value α is an index of rating attitude. It reflects the expert's risk-bearing attitude. Let $B = (c, a, b)$ be a normal triangular fuzzy number. The index of rating attitude of an individual expert is defined as $Y = (a-c)/(b-c)$ (Chang and Chen, 1994). If $Y > 0.5$, it implies that the expert is a risk lover. If $Y < 0.5$, the expert is a risk averter. If $Y = 0.5$, the attitude of expert is neutral to the risk. Thus, the total index of rating attitude, R, with the evaluation data of individuals can is shown as

$$R = \left\{ \sum_{i=1}^{k} \sum_{j=1}^{n} (a_{ij} - c_{ij})/(b_{ij} - c_{ij}) + \sum_{i=1}^{m} \sum_{t=1}^{k} \sum_{j=1}^{n} (o_{itj} - q_{itj})/(p_{itj} - q_{itj}) \right\} \Big/ (kn + mkn). \tag{8}$$

From Eqs. (4), (6) and (8), the ranking values $U_t(F_i)$ can be approximately expressed as

$$U_T(F_i) \cong R[(Z_i - x_1)/x_2 - x_1 - Q_i + Z_i) + (1 - R)[1 - (x_2 - Y_i)/(x_2 - x_1 + Q_i + Y_i)]. \tag{9}$$

Table 7
Fuzzy appropriateness indices for the seven alternatives

Alternatives	Fuzzy appropriateness index
Site 1	(0.45563, 0.55988, 0.65963)
Site 2	(0.3405, 0.42463, 0.51308)
Site 3	(0.34713, 0.43275, 0.52553)
Site 4	(0.53163, 0.41638, 0.49888)
Site 5	(0.35525, 0.4415, 0.53055)
Site 6	(0.34425, 0.434, 0.5311)
Site 7	(0.33525, 0.4235, 0.52023)

Table 8
Ranking values of the different alternatives

Alternatives	Ranking values
Site 1	0.786689
Site 4	0.580556
Site 5	0.371734
Site 3	0.342253
Site 6	0.340457
Site 7	0.310792
Site 2	0.266668

And the ranking values of the fuzzy appropriateness indices for alternatives are presented in Table 8. Site 1 exhibits the highest potential in this site selection process.

4. Discussion

4.1. Advantages and disadvantages of the approach

The methodology followed in this paper differs from the conventional methods of integrating GIS with MCDM for landfill selection because the approach follows two sequential steps rather than a full-integrated scheme. In the first stage, GIS-based analysis of spatial data has been a new specialized process, capable of analyzing complex problem of evaluating various geospatial features for targeting potential areas for siting landfills. While GIS offers unique capacities for automating geospatial analysis for screening all possible sites, data availability can prove to be a limiting factor in its application for selection of a landfill. Landfill selection process can lead to situations in which certain criteria, such as public nuisance, economic factors and impacts on historical markers, may cause increased ambiguities in the decision making process due to lacking sufficient information. The candidate sites obtained in the first stage can be narrowed down using a prescribed MCDM process. Multicriteria evaluation is primarily concerned with how to combine the information from several criteria to form a single index of evaluation. In case of Boolean criteria, the solution usually lies in the union (logical OR) or intersection (logical AND) of conditions. However, for continuous factors in crisp MCDM process, a weighted linear combination is a usual technique (Voogd, 1983). As the criteria are measured at different scales, they are standardized and transformed such that all factor maps are positively correlated with suitability. Establishing factor weights is the most complicated aspect, for which the most commonly used technique is the pair-wise comparison matrix. In response to the vague (fuzzy) conditions, domain experts in the second stage got involved. By including the expert opinion and combining them with the power of fuzzy and MCDA yielded a crystal structure very much dependent of the screening values of data sets. This can be enormously advantageous in solving controversial political debates in the future. The advantage of this method is therefore placed upon the capability to incorporate the knowledge of the domain experts in the uncertain decision making process when there is a lack of crisp information related to certain criteria, such as public nuisance and impact of landfill on historical markers. However, the disadvantage of this method is that the selection of the best candidate site is dependent on the judgments of the domain experts and can be sensitive to changes in the decision weights associated with criteria. In certain situations, two experts may have contradicting judgments about suitability of a candidate site. Hence, it is required to assess the extent of difference or similarity between the two experts in association with decision weights. Where the experts are forced to give ranks to the pre-defined candidate 7 sites, the selections are only made among these (i.e., site 1 is better than all other 6 sites). But some might argue that it may not be the very best ideal case. If the screening process is loosened a little bit may be a candidate site 8 will appear and may be at some criterion it will score more. To respond to this challenge, a field check was done in the early stage and in the middle of this study to ensure that site 1 would also be the approved one at the field eventually.

In this study the decision weights are provided by the decision makers as a triangular fuzzy number and therefore overlap of the triangular fuzzy sets (shown in Fig. 8) is used as a similarity measure between the two experts. The overlap measure indicates the extent to which the two experts agree upon each other for the importance of a particular criterion in the selection of a landfill. The overlap measure for the each of the criteria was estimated

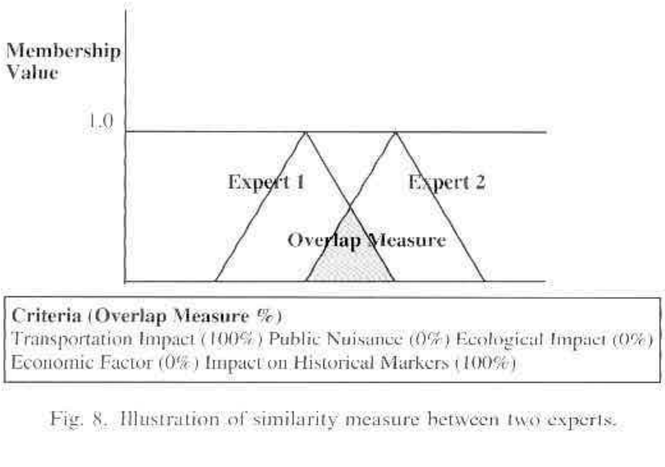


Fig. 8. Illustration of similarity measure between two experts.

mathematically as illustrated in the figure. The values of such overlap measure for the criteria TI and HM are 100% whereas there was no overlap between the weights for the criteria PN, EI and EC indicating that there is a marked difference between the experts in their judgments of the importance of the three criteria (PN, EI and EC) in the selection of a landfill. The discrepancy in the judgment between two experts can have a significant impact on the selection process, which can be minimized by having more experts to provide assessment of the decision criteria weights. The overlap measure can thus provide vital information related to the similarity between the experts involved in the decision making process.

4.2. Sensitivity analysis

In a landfill selection process, it becomes necessary to assess the reliability of the method involved in identification of the best candidate site. A small perturbation in the decision weights may have a significant impact on the rank ordering of the sites and subsequently change the best choice. Therefore, sensitivity analysis using Monte Carlo simulation was performed to determine the probability of changes in rank ordering. Hence, the decision weights were systematically varied to investigate the relative impacts of the weights on the rank ordering of the landfill sites. The weights for the five different criteria provided by the experts as triangular fuzzy set were varied within a range of 20% provided that a latin hypercube sampling of the inputs was used to conduct such a simulation. The results of 100 simulations were shown in Fig. 9. It can be observed that site1 still completely dominate all the other sites despite a certain degree of variations in the decision weights. With the aid of 100 simulation runs, it indicates that the ranks of the seven-candidate sites remained the same as shown in Table 8 except for candidate sites 2 and 7. Candidate site 7 occupied the last rank 49 times out of the 100 simulations replacing site 2, thus demonstrating the fact that site 7 and site 2 perform identically with respect to the decision criteria selected according to the two experts. The fact that the perturbation of the decision weights has a small impact on the ranking of the candidate sites reveals that the degree of domination of the candidate sites is almost independent of changes in the decision weights associated with selected criteria.

5. Conclusions

The increasing generation of MSW in the LRGV is one of the greatest challenges faced by governmental authorities. In order to mitigate the impacts on the environment and public health, a claim, which requires a fast decision-making process regarding the final disposal of the MSW, motivates this study. Research findings show that a SDSS, featuring a well-structured architecture and the computational power, improves the application potential in urban and regional planning, and gives essential support to the decision-maker in the assessment of the waste management problem so that a higher level of understanding can be reached in regard to environmental decisions. In order to gain an all-inclusive perspective, the process of decision-making consisted of a two-stage analysis, beginning with an initial site screening followed by a detailed assessment of the suitability of the candidate sites using a FMCDM approach guided by a panel of experts in the site selection process. The first-stage analysis was successful in preliminary landfill site screening leading to exclude the sensitive areas while retaining sufficient areas for further evaluation at the same time. Within the recovered fuzzy region in the second-stage analysis, MCDM method smoothly incorporated the information provided by two experts leading to fulfill the ranking of the seven

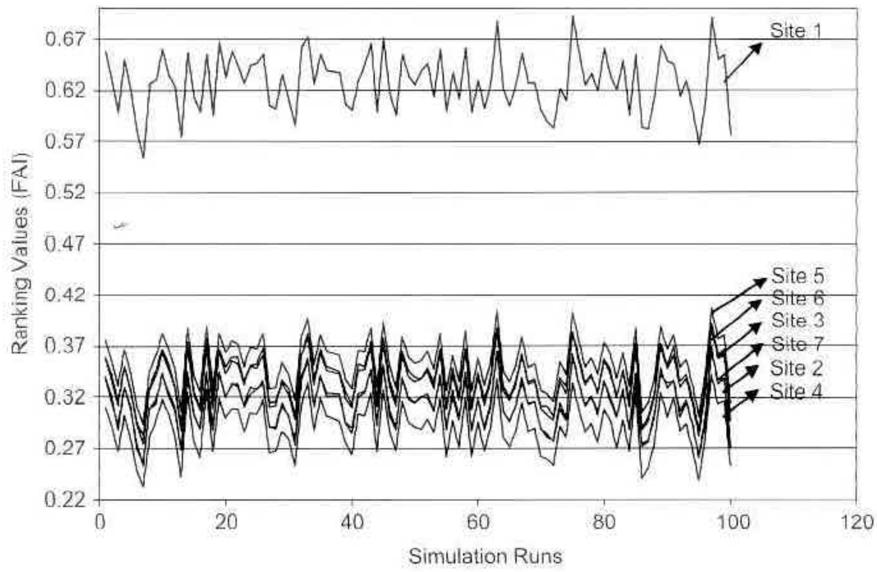


Fig. 9. Monte Carlo simulation showing the changes in ranking values of the candidate sites.

alternatives with respect to five different criteria. All the criteria were eventually aggregated to select the most suitable site in terms of ratings given the fact that fuzzy set theory may aid in justification of the uncertainty in decision-making. In consequence, a SDSS may strengthen the generation and evaluation of alternatives by providing an insight of the problem among the varied objectives and granting essential support to the process of decision-making under uncertainty (Malcezwki, 1999; Sharifi and Van Herwijnen, 2003). With such an effort, it is concluded that "site 1" located near highway 77 closer to Cameron–Willacy boundary is the most suitable site for landfill based on an integrated GIS and FMCDM analysis. A sensitivity analysis was conducted to assess the reliability of the ranking of the candidate sites using a Monte Carlo simulation by changing the decision weights associated with selected criteria. The results indicated that the candidate site 1 still completely dominate the other sites despite variations of the decision weights within a range of 20%. Overall, GIS thus offered the means to identify seven potential landfill sites based on well-defined criteria, which were later ranked according to the preferences provided by two domain experts that were based on their experiences and knowledge of the dynamics of the Lower Rio Grande Valley using FMCDM. FMCDM offered the capacity to incorporate the opinions of the domain experts that can be useful in the future to settle political debate regarding the site selection. Such procedure was eventually proved useful in the case study identifying favorable areas for waste disposal in a fast-growing urban region in south Texas.

Acknowledgements

The authors acknowledge the financial support from the City Government of Harlingen, Texas and the data reports cited and used in this analysis.

References

Akash, B.A., Mamlook, R., Mohsen, M.S., 1999. Multi-criteria selection of electric power plants using analytical hierarchy process. Electric Power Systems Research 52, 29–35.

Allanach, W.C., 1992. Regional landfill planning and siting. Public Works, December, 48–50.

Altrock, C.V., Krause, B., 1994. Multi-criteria decision-making in German automotive industry using fuzzy logic. Fuzzy Sets and System 63, 375–380.

American Planning Association, 2002. Growing Smart—Legislative Guidebook: Model Statutes for Planning and the Management of Change. APA's Planners Book Service, 〈http://www.planning.org/guidebook/Login.asp〉 (access:September 2003).

Anderson, R.F., Greenberg, M.R., 1982. Hazardous waste facility siting: a role for planners. Journal of the American Planning Association 48, 204–218.

Baas, S.M., Kwakernaak, H., 1997. Rating and ranking of multiple aspect alternative using fuzzy sets. Automatica 13, 47–58.

Baban, S.M.J., Flannagan, J., 1998. Developing and implementing GIS-assisted constraints criteria for planning landfill sites in the UK. Planning Practice and Research 13, 139–151.

Chan, F.T.S., Chan, M.H., Tang, N.K.H., 2000. Evaluation methodologies for technology selection. Journal of Materials Processing Technology 107, 330–337.

Charnpratheep, K., Zhou, Q., Garner, B., 1997. Preliminary landfill site screening using fuzzy geographic information systems. Waste Management and Research 15, 197–215.

Chen, S.M., 1996. Evaluating weapon systems using fuzzy arithmetic operations. Fuzzy Sets and Systems 77, 265–276.

Cheng, C.H., 1997. Evaluating naval tactical systems by fuzzy AHP based on the grade value of membership function. European Journal of Operational Research 96, 343–350.

Chiadamrong, N., 1999. An integrated fuzzy multi-criteria decision making method for manufacturing strategies selection. Computers and Industrial Engineering 37, 433–436.

Choi, D.Y., Oh, K.W., 2000. ASA and its application to multicriteria decision making. Fuzzy Sets and Systems 114, 89–102.

Del Valle, F., 2005. Both sides confident at landfill hearing. Valley Morning Star. Retrieved from 〈http://www.valleystar.com/archives〉.

Dikshit, A.K., Padmavathi, T., Das, R.K., 2000. Locating potential landfill sites using geographic information systems. Journal of Environmental Systems 28, 43–54.

Fanti, M.P., Maione, B., Naso, D., Turchiano, B., 1998. Genetic multi-criteria approach to flexible line scheduling. International Journal of Approximate Reasoning 19, 5–21.

Hauser, D., Tadikamalla, P., 1996. The analytic hierarchy process in an uncertain environment: a simulation approach. European Journal of Operational Research 91, 27–37.

Jensen, J.R., Christensen, E.J., 1986. Solid and hazardous waste disposal site selection using digital geographic information system techniques. The Science of the Total Environment 56, 265–276.

Kaiser, E., Godschalk, D., Chapin Jr., S.F., 1995. Urban Land Use Planning. University of Illinois, Urbana, IL, 493pp.

Kao, J., Lin, H., Chen, W., 1997. Network geographic information system for landfill siting. Waste Management and Research 15, 239–253.

Karsak, E.E., Tolga, E., 2001. Fuzzy multi-criteria decision-making procedure for evaluating advanced manufacturing system investments. International Journal of Production Economics 69, 49–64.

Kaufmann, A., Gupta, M.M., 1988. Fuzzy Mathematical Models in Engineering and Management Science. Elsevier, Amsterdam, Netherlands.

Keir, A.W., Doucett, J.A., Oliveri, T., 1993. Landfill siting using GIS technology: the case of the Peel Landfill site search. In: Proceedings of the Canadian Conference on GIS-1993, Ottawa, Canada, pp. 13–24.

Klostermann, R.E., 1999. The "what if?" collaborative planning support system. Environment and Planning. B, Planning and Design 26, 393–408.

Kuo, R.J., Chi, S.C., Kao, S.S., 2002. A decision support system for selecting convenience store location through integration of fuzzy AHP and artificial neural network. Computers in Industry, in press.

Labib, A.W., O'Connor, R.F., Williams, G.B., 1998. Effective maintenance system using the analytic hierarchy process. Integrated Manufacturing Systems 9, 87–98.

Lee, S.J., Lim, S.I., Ahn, B.S., 1998. Service restoration of primary distribution systems based on fuzzy evaluation of multi-criteria. IEEE Transactions on Power Systems 13, 1156–1163.

Levary, R.R., Ke, W., 1998. A simulation approach for handling uncertainty in the analytic hierarchy process. European Journal of Operational Research 106 (1), 116–122.

Lin, H.Y., Kao, J.J., 1999. Enhanced spatial model for landfill siting analysis. Journal of Environmental Engineering, ASCE 125 (9), 845–851.

Longley, P.A., Barnsley, M.J., Donnay, J.P., 2001. Remote sensing and urban analysis: a research agenda. In: Donnay, J.P., Barnsley, M.J., Longley, P.A. (Eds.), Remote Sensing and Urban Analysis. Taylor & Francis, London, pp. 245–258.

Lower Rio Grande Valley Development Council (LRGVDC), 2002. Regional Solid Waste Management Plan Amendment 2002–2020. June 30, 2003.

McIntyre, C., Parfitt, M.K., 1998. Decision support system for residential land development site selection process. Journal of Architect Engineering ASCE 4 (4), 125–131.

Malczewki, J., 1999. Spatial multicriteria decision analysis. In: Thill, J.C. (Ed.), Spatial Multicriteria Decision Making and Analysis. Ashgate Publishing, Hamshire.

Nagar, B.B., Mizra, U.K., 2002. Hydrogeological environmental assessment of sanitary landfill project at Jammu City, India. Green Journal 17 (8), 223–245.

O'Leary, P.R., Canter, L., Robinson, W.D., 1986. Land disposal. In: Robinson, W.D. (Ed.), The Solid Waste Handbook: A Practical Guide. Wiley, NY, USA p. 271.

O'Meara, M., 1999. Reinventing cities for people and the planet. Washington, DC, Worldwatch Institute, 68pp. 〈http://www.worldwatch.org/pubs/paper/147/〉 (Access: September 2003).

Pierson, E., 2004. Valley could run out of landfill space in coming decades, reports say. The Brownsville Herald.

Paek, J.H., Lee, Y.W., Napier, T.R., 1992. Selection of design/build proposal using fuzzylogic system. Journal of Construction Engineering and Management 118, 303–317.

Saaty, T.L., 1980. The Analytic Hierarchy Process. McGraw-Hill, New York.

Sener, B., Suzen, L., Doyuran, V., 2006. Landfill site selection by using geographic information systems. Environmental Geology 49, 376–388.

Siddiqui, M.Z., Everett, J.W., Vieux, B.E., 1996. Landfill siting using geographical information systems: a demonstration. Journal of Environmental Engineering, ASCE 122 (6), 515–523.

Sharifi, A., Van Herwijnen, M., 2003. Spatial Decision Support System. ITC, Esnchede, p. 64.

Tagaris, E., Sotiropolou, R.E., Pilinis, C., Halvadakis, C.P., 2003. A methodology to estimate odors around landfill sites: the use of methane as an odor index and its utility in landfill siting. Journal of the Air and Waste Management Association 53 (5), 629–634.

Tang, M.T., Tzeng, G.H., Wang, S.W., 1999. A hierarchy fuzzy MCDM method for studying electronic marketing strategies in the information service industry. Journal of International Information Management 8 (1), 1–22.

Teng, J.Y., Tzeng, G.H., 1996. Fuzzy multicriteria ranking of urban transportation investment alternatives. Transportation Planning Technology 20 (1), 15–31.

TSHA Online. The Handbook of Texas Online—Rio Grande Valley. Available at 〈http://www.tsha.utexas.edu/handbook/online/articles/view/RR/ryr1.html〉, accessed on September 12, 2003.

US Department of Agriculture, 1977. Soil Survey of Cameron County Texas. Soil Conservation Service.

US Department of Agriculture, 1982. Soil Survey of Willacy County Texas. Soil Conservation Service.

Voogd, H., 1983. Multi-criteria Evaluations for Urban and Regional Planning. Princeton University, NJ, USA.

Yu, W.D., Skibniewski, M.J., 1999. Quantitative constructability analysis with a neuro-fuzzy knowledge-based multi-criterion decision support system. Automation in Construction 8 (5), 553–565.

Zeiss, C., Lefsrud, L., 1995. Analytical framework for facility waste siting. Journal of Urban Planning and Development, ASCE 121 (4), 115–145.

Zadeh, L.A., 1965. Fuzzy sets. Information and Control 8, 338–353.

Journal of Environmental Management 87 (2008) 154–164

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Who should manage protected areas in the Swedish mountain region? A survey approach to co-management

Anna Zachrisson*

Department of Political Science, Umeå University, 901 87 Umeå, Sweden

Received 3 December 2005; received in revised form 8 November 2006; accepted 9 January 2007
Available online 23 March 2007

Abstract

This article investigates attitudes towards co-management of protected areas in Sweden, at the national, county and local level. In Sweden, protected areas are still primarily designated and managed hierarchically—a practice increasingly contested by people living close to them, including indigenous Sámi reindeer herders whose economic activities are located within protected areas. The general view could, on the contrary, be anticipated to be pro-state since protected areas are considered to be of national interest. For democratic reasons, however, the opinions of the whole population should be considered. In order to measure both local and general views, this study is based on a two-sample survey of 8868 respondents. The objectives are to map and explain attitudes regarding who should manage protected areas in Sweden, and to test the usefulness of a multi-level quantitative method. Such an approach is unusual in co-management literature that is empirically mainly based on local case studies. The explanatory ambition sets out to test three hypotheses drawn from common-pool resource theory: resource dependency, common understanding, and trust. Perhaps surprisingly, the results show that a considerable majority of the respondents (at all levels) wish to see self- or co-management. All three hypotheses are important to understand attitudes toward the management of protected areas, but not always in the way that the theory anticipates.

Keywords: Co-management; Protected areas; Swedish mountain region

1. Introduction

The overall aim of this study is to map and explain the attitudes towards who should be involved in the management of Swedish protected areas—28 national parks and about 2700 nature reserves (Naturvårdsverket, 2006a), covering in total about 10% of the land surface (SCB, 2006). About 90% of the national park surface and 84% of the nature reserve surface is found in the mountain region (Naturvårdsverket, 2006b). The first national parks were designated in 1909, and since the mid-twentieth century the amount of protected areas has increased tenfold. Even though recent government recognition of the importance of local involvement (Skr. 2001/2002, p. 173), nature conservation is still a traditional, hierarchic business run by central and county authorities. This practice is increasingly contested as small, relatively marginalized communities and the indigenous people, the Sámi, are demanding more influence on how large nature reserves and national parks are managed (Rådelius, 2002).

This situation must be viewed in the light of international developments during the last decades; something that can be described as a paradigm shift has actually taken place in nature conservation ideology. It is slowly being recognized that for a protected area to be legitimate in the eyes of local populations, they need to be involved in the decision-making process (Nepal and Weber, 1995; Ghimere and Pimbert, 1997). This is emphasized in, for instance, the Agenda 21 (Section 15.3), the UN Convention on Biological Diversity (article 8 (j)) and in the ILO Convention 169 (concerning indigenous peoples' rights). Co-management procedures are believed to have the potential to reconcile clashing interests by establishing conflict resolution arenas (Pinkerton, 1989). In this way, both conservation and development goals could potentially be achieved. One important question is, however, whether these local demands are representative of the general

*Tel.: +46 90 786 6185; fax: +46 90 786 6681.
E-mail address: anna.zachrisson@pol.umu.se.

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.01.010

opinion in the country. Nature conservation is not done first and foremost to protect the environment for local people, but to protect important natural assets, both for biodiversity reasons and for future generations. Hence, the opinions of all citizens, not only the local population, should be heard and considered.

One common way to elicit opinions is through random sample surveys. Such a quantitative approach is rather unusual in the context of studying and analysing pre-requisites for co-management, but could possibly result in new, important insights. Co-management literature relies mainly on numerous local case studies, either designed as single-case studies or as comparisons of multiple cases. General attitudes, however, have been much less analysed, even though these ought to influence the success of co-management initiatives. Understanding the role of opinions is crucial to the realization of fully democratic processes in cases dealing with complex, multiple-use natural resources of national and global interest.

The co-management field of research actually has a general lack of theory, not least due to the lack of a clear and concise definition of the notion of co-management itself (Pinkerton, 1989; Campbell, 1996; Berkes, 2002). Nevertheless, researchers such as Jentoft (2003) state that a broad definition is necessary since co-management will not look the same in different settings. In Sweden, co-management of protected areas is a rather new phenomenon and its forms yet to be established, and therefore it makes sense to ask the citizens who they wish to include in the decisions about management rules for protected areas.[1] In this survey this has been done, as a way to narrow down the broad concept of co-management in order to apply it on Swedish protected areas.

Quantitative methods, such as regression and factor analyses, are used to bring together multiple variables in order to control for spurious relationships. Here, logistic regression analysis is applied to investigate what factors might explain the preferred management solution in Sweden. What characterizes opponents and advocates of different management forms? In answering this question, three hypotheses will be tested, derived from the more general field of common-pool resource research (which co-management is part of), and hereby the potential of a quantitative survey approach for analyses of common-pool resource management can be evaluated.

Thus, the objectives in this article are twofold: to map and explain the attitudes regarding who should manage protected areas, and to test the usefulness of quantitative methods in a research field where such an approach is unusual. Section 1 outlines the theoretical point of departure, while the methodological considerations are found in Section 2. In Section 3, the empirical results are exposed: the support for the principal management solutions is reported and logistic regression analyses are used to explain these choices. Section 4 contains the general discussion of the results and the conclusions.

[1]It should be noted though that co-management is not that new as a phenomenon in Sweden, there are examples of co-managed lakes, forests and other natural resources.

2. Protected areas, local resistance and co-management

The establishment and management of protected areas are statutory and international obligations for states, as shown for example by the UN Convention on Biological Diversity and the Habitats Directive of the European Union. It is still one of the major strategies to achieve wildlife conservation and protection of biodiversity, since protected areas are seen as a powerful means for ensuring the continuity of fragile ecosystems and the future availability of limited natural resources (Mulder and Coppolillo, 2005; Gorenflo and Brandon, 2006). Often, however, proposals for establishing protected areas are met with protests from local residents living in, or close by, the area in question (Stevens, 1997). Since the majority of land areas proposed for protection are so-called pristine ecosystems and areas with high nature values (for recreation and/or biodiversity) they are mostly situated in rural areas (Brandon et al., 1998). Some people in these areas may fear that the protected area will severely restrict their economic activities, which are often based on natural resources, and also their freedom of access (Brechin et al., 2003). The first empirical question to be answered is, thus, whether there is a difference in local populations' concerns in urban compared to rural areas.

2.1. *Protected areas as common-pool resources*

Ostrom (1990, p. 30) defines common-pool resources as "a natural or man-made resource system that is sufficiently large as to make it costly (but not impossible) to exclude potential beneficiaries from obtaining benefits from its use". These resources all share two important characteristics: (1) subtractability or rivalry, which means that consumption of resource units removes those units from those available to others and (2) difficulty in excluding potential beneficiaries from access to the resource system, which creates a risk of free riders who may use the resource without contributing to its provision (Berkes, 1989; Ostrom, 1990; Ostrom et al., 1994). Preventing access by users who do not follow the rules is costly, which makes exclusion costs a core problem for the management of commons. There is also a limit to the number of resource units that can be produced by a common-pool resource. When this limit is approached, crowding effects are produced, and in the long run the reproduction capability of the resource may be destroyed (Ostrom, 1990).

Protected areas may be considered as common-pool resources. Anttila (1999) has suggested that the vast, sparsely populated mountain landscape in northern Sweden fits Ostrom's definition of a commons. These protected areas are non-excludable since they are nearly impossible to parcel and fence in, not only physically and aesthetically, but also politically, since the mountains are thought of as belonging to everyone, according to the traditional right of public access.[2] These areas also have the characteristic of subtractability, since they are defined areas of land that could potentially become overused—in particular in the sense of being too crowded.

According to Hardin (1968), degradation is inevitable whenever many individuals use a scarce resource in common. He termed this situation a "tragedy of the commons". The tragedy was illustrated with the example of the rational herder who adds more and more animals to the common grazing lands. The herder will immediately receive the direct benefit of his own animals while he will only bear a small share of the costs resulting from overgrazing. Almost a decade earlier, Gordon (1954, p. 135) concluded that "everybody's property is nobody's property". The solutions to avoid this tragedy that both Hardin and Gordon came up with were privatization or centralization of management decision-making. Until the 1980s other scholars generally agreed with their analyses, but then a shift in research priorities began to occur. Many book-length studies and edited volumes concentrating on both examples of self-management and co-management have been produced since then, leading to a serious rethinking of the common-pool resource problem (McCay and Acheson, 1987; Berkes, 1989; Pinkerton, 1989; Ostrom, 1990; Bromley, 1992; Ostrom et al., 2002). A rich case-study literature has also evolved to document cooperation on common-pool resources (Lam, 1998; Wade, 1988; Gibson et al., 2000). Much of this research has been aimed at showing under what circumstances more *locally based forms of management* appear and succeed, and it is clear that alternatives exist to privatization and centralization.

2.2. Self-management and co-management

In the groundbreaking "Governing the Commons" Ostrom (1990, p. 25) discusses self-organization, which is defined as "a group of principals can organize themselves voluntarily to retain the residuals of their own efforts". Jentoft and McCay (2003, pp. 298–299) describe the same type of management as "community-based management [...] a highly decentralized management system whereby some or all authority rests at the local community level". In this article, such a management solution will be referred to as *self-management*. In the Swedish context, it seems reasonable to include management at the municipal and sub-municipal level as self-management (specifically, municipalities, Sámi communities, if any, and within local populations). The municipalities are included because they have in general a rather independent position in relation to the central government, and local groups have quite good possibilities to influence this level.

Co-management research can be considered as a sub-field of common-pool resource research (Pinkerton, 1989; Baland and Platteau, 1996; Jentoft, 1998). Co-management has been defined as "the sharing of power and responsibility between government and local resource users" (Berkes et al., 1991), which means that it covers a wide range of arrangements at the interface between the state, the civil society and the private sector. It is often pointed to as a solution for the future by common-pool resource theorists (see for instance Berkes, 1989; Berkes and Folke, 1998; Dietz et al., 2002). In Sweden, co-management of protected areas would, in accordance with the discussion above about the role of the municipalities, mean that the Swedish Environmental Protection Agency and/or county administrative boards would share power with municipalities and/or local users (and potentially also other actors such as research institutions).

2.3. The IAD-framework

The Institutional and Development Analysis (IAD) framework has been used extensively in the efforts to better understand common-pool resources (see for instance Oakerson, 1992; Imperial, 1999; Carlsson, 2000; Rudd, 2004). It outlines favourable conditions for robust common-pool resource management in terms of five contextual components: biophysical and technological attributes, exogenous attributes (the macro-economy for instance), organizational/institutional attributes and, finally, community attributes (Edwards and Steins, 1999; Dolšak and Ostrom, 2003). Since, the focus of this study is the opinions of the Swedish people on management issues, it only concerns the community attributes, and therefore the others will not be further discussed here.

The community attributes are: (U1) users are dependant on the resource system for a major portion of their livelihood, (U2) users have a common understanding of the resource and of how their actions affect each other and the resource, (U3) users' relations are built on trust and reciprocity (direct communication), (U4) users are able to determine access and harvesting rules without external authorities countermanding them, and (U5) users have prior organizational experience and local leadership (Ostrom, 2000). Two more attributes are often discussed as well, but the results on their impact are ambiguous. These are the extent of heterogeneity in the community (ethnicity, gender, and interests) (Baland and Platteau, 1996; Bardhan and Dayton-Johnson, 2002) and the size of the community (Ostrom, 2005). In this study, only the first three attributes; (U1), (U2), and (U3), will be tested since the data material does not include any measures for the other conditions.

2.3.1. The dependency hypothesis

Resource dependency, U1, is considered a crucial factor for understanding local common-pool resource

[2] The Swedish term "the right of public access" means that everyone has the right to use the countryside, but this freedom must not infringe upon the freedom of others. The landscape or animal life is not to be damaged, and consideration for both landowners and for others who are out in the countryside must be shown (Naturvårdsverket, 2006c).

management. Baland and Platteau (1996, p. 287) express this criteria as follows: "the more vital the resource for survival the greater the chances of success". Carrus et al. (2005, p. 240) propose that the designation of a protected area may be conceived as "a type of commons dilemma, particularly for those local stakeholders who derive their daily sustenance from the exploitation of natural resources within the protected area". Residents must choose whether to comply with the regulations and limit their individual exploitation of natural resources within the protected areas. Research on rural resistance against carnivores confirms this idea; cattle owners, such as sheep farmers and reindeer herders who lose animals to carnivores, and people involved in primary sectors (agriculture, forestry, fishing, reindeer herding, and mining), are the strongest opponents of carnivores (Wilson, 1997; Svarstad, 2003). Stoll-Kleeman (2001) advances the idea that it may just be the *impression* of facing restrictions on traditional day-to-day activities due to nature conservation regulations that triggers emotional and cultural drivers. To summarize this argument, it seems plausible that individuals with a high degree of dependence (real or imagined) on the resources in question would be favouring self- or co-management in which they could expect a higher degree of influence.

Hypothesis 1: Those involved in, or dependent on, the exploitation of natural resources are more positive towards self-management/co-management than those who are not.

2.3.2. The common understanding hypothesis

That resource users have a shared image of how the resource system operates and how it is affected by their actions, U2, is another important asset of successful common-pool resource management (Gibson et al., 2000). The 'use' of protected areas is traditionally non-consumptive; it is about activities such as hiking and skiing for the purpose of recreation. At the same time, these areas have often been used for hunting, fishing, agriculture, etc. before becoming protected, and with the new paradigm of how protected areas are to be managed the acceptance for continued consumptive use has increased (Folke, 2006). Such use should, nevertheless, not threaten the aims of protection. In conservation debates, urban people are often believed to represent the 'purist' view on protected areas as protected from all kinds of activities, while rural people represent the consumptive users. This view is highly simplistic; neither rural communities nor urban centres can be considered as unitary actors or groups (Skogen and Krange, 2003). Heberlein and Ericsson (2005) have suggested that the categories urban and rural are too broad as there are various views within both that are related to previous experiences, such as where you grew up, affecting your opinion. However, Stoll-Kleeman (2001) has shown how local resistance against the designation and management of protected areas is formed by group processes encouraging social identity together with communication and perception barriers between the involved groups. She gives evidence that "there are powerful emotional and cultural drivers that divide nature conservationists and local land users and residents into two camps, maintained by stereotyping and group bonding" (Stoll-Kleeman, 2001).

In the present study, it is difficult to measure the degree of common understanding correctly since different user groups cannot be discerned. Respondents living in urban and rural areas who could be expected to represent different uses to some degree (or at least opinions as discussed above) can, however, be distinguished. Local and national—that will be primarily urban—views will thus be analysed here. This simplification does not imply that these two groups are considered to be homogenous, but that a majority within each group probably shares a similar view on the particular issue of protected areas.

Hypothesis 2: If people living close to protected areas and people living in cities have a common understanding of whether protected areas should be designated, why they are important, and what activities should be allowed in them, self- or co-management will be preferred.

2.3.3. The trust hypothesis

The criteria U3 states that users who trust each other are more likely to succeed to create institutions for managing common-pool resources (Dolšak and Ostrom, 2003). Ostrom (1998, p. 14) has stated that "the individual attributes that are particularly important in explaining behaviour in social dilemmas include the expectations individuals have about others' behaviour (trust)". This further underlines the essential role that trust may play. Trust affects human relations so that an individual who expects that her initiative to cooperate will be reciprocated is more willing to do just that, which in turn would foster self-management regimes. Trust is not only necessary for self-management to succeed, but also for democratic practices in general. For instance, Rothstein (2003) has found a correlation between trust in other people and trust in public institutions such as the parliament and the police. From his results it is, however, not possible to decide in which direction this correlation works, i.e. if it is trust in other people that leads to more trust in institutions or the other way round. He also shows that individuals with a high level of trust in other people have a more positive view on democracy, while those who have a low level of trust only have confidence for a very small number of persons. This could be interpreted as though the 'low-trusters' would prefer self-management since it presumably could involve people they trust. Co-management, on the other hand, requires cooperation between different levels and therefore it would imply that its supporters trust not only other people in general, but also politicians and the democracy. The hypothesis that can be formed with the help of Ostrom and Rothstein is thus:

Hypothesis 3: Persons with a low level of trust in other people in general, politicians and democracy are more positive towards self-management, while those with a high level of trust in fellow human beings, politicians and democracy are more positive towards co- or state management.

3. Methodology

The analysis in the present study is based on a mail survey conducted by the Mountain Mistra Programme, carried out in collaboration with the five northernmost counties and the Swedish Environmental Protection Agency during March and April 2004. Respondents were contacted a maximum of four times (pre-warning, questionnaire, 1st and 2nd reminder). The survey covered opinions about protected areas, carnivores, and fishing/hunting. It consisted of two random samples taken from the national population register; the first consisting of 150 permanent inhabitants 16–65 years old from each municipality in the four northernmost counties Dalarna, Jämtland, Västerbotten, Norrbotten (in total 7800), and the second included 1067 persons from the rest of Sweden. The response frequency was 65% in the mountain counties and 57% in the second sample of 1067 individuals. As it is used in this text, the term "national sample" actually refers to *both* samples since together they represent the general view of Swedes, when the answers from the four mountain counties are weighed in relation to their proportion of the population. Answers from the municipality sample are also weighed according to population when they are used for regional comparisons.

The dependent variable is constructed from the following question: Which of the following institutions should decide management rules for protected areas—the state, the county administrative board, the municipality, the local population, and/or the Sámi communities?[3] Three outcomes will be examined and, in accordance with the management solutions outlined in the theory section, they are: self-, co- and state management. These are constructed as dichotomous or dummy variables (0 for absence of a management alternative and 1 for its presence). As *logistic regression analysis* is the preferred tool to use when dealing with categorical, non-continuous variables such as dummy variables (Pampel, 2001), this is the method employed here. It can, for instance, be used to analyse how different factors influence individuals' choices between different alternatives.

[3]The Sámi communities are the legislative units for the organization of reindeer herding, which is the only way to access immemorial Sámi rights of not only reindeer herding but also fishing and hunting on 'traditionally used land'. Sámi communities are not allowed to engage in any other economic activities than reindeer herding. All this is specified in the Reindeer Herding Act (SFS 1971, p. 437).

The response frequency for this particular question was somewhat lower than for the survey as a whole; 2.3% of those who replied did not answer this question and 14.4% did not have any opinion. Drop out analyses show, first of all, that women and unemployed are somewhat over-represented among those answering that did not have any opinion. Secondly, highly educated and people with high incomes are under-represented. Among those who did not answer the question at all, many had not replied to the demographic questions either, but women and people with high incomes are over-represented.

4. Results

In order to investigate whether Swedes *want* co-management at all, dummy variables were constructed for all 31 theoretically possible combinations of the actors mentioned in the survey as potential protected area managers. These were labelled under three main categories according to the definitions above: self-management (all combinations including the municipality and/or the local population and/or the Sámi communities), co-management (state and/or county administrative board AND municipality and/or population and/or Sámi communities), and state management (state and/or county administrative board). These three 'aggregated' management solutions were then used as dependent variables for the regression analyses. The results are summarized in the continuum in Fig. 1, which shows that self- and co-management solutions have a strong support among Swedes—about 65%. The county administrative board is a slightly more popular co-management partner than the state; 17% of the respondents include the county administrative board, 13.5% the state and 9.5% both. State management, which is the general rule today, is preferred only by 18%, thus constituting a clear minority. This must be interpreted as though Swedes want more diverse actors to be involved in the management of protected areas; 25% even wish to have self-management and thus an entirely locally controlled management.

4.1. Urban/rural cleavages?

The basic assumption behind the discussion about protected areas is that popular resistance develops in rural areas where protected areas are most often established. In Sweden, protected areas are highly concentrated to the mountain region, which means that land use conflicts due to the designation of protected areas ought to be more acute there than in other parts of the country. The mountain region is located in the western municipalities of the four northern-most counties of Sweden; namely in Dalarna, Jämtland, Västerbotten, and Norrbotten. In order to test the assumption about urban/rural cleavages, the frequencies of resistance are compared between the mountain region, the mountain counties and the national sample. The analysis thus aims to reveal possible urban/rural

Self-management	Co-management	State management
n = 1190 (25%)	n = 1874 (40%)	n = 840 (18%)

Fig. 1.

cleavages between the mountain region and the national level, but also within the mountain counties where urban centres are located at the coast.

The results show that the attitudes in the mountain region are indeed different from the national sample (Table 1). First, it should be noted that the dropout is about 9% lower in the mountains than nationally. Thus, one explanation to why the mountain region distinguishes itself could be that its inhabitants have more experiences of protected areas and therefore more firm opinions about its management. While self-management is preferred by 38% of the respondents in the mountain region, only about 25% in the national sample and 29% in the samples from the counties support this alternative. By contrast, the state management alternative is preferred by slightly lower numbers in the mountain region sample—about 3% less than in the other samples. However, the support for co-management is more or less the same in all samples—just above 40%.

In order to identify outliers in the mountain region, municipality frequencies are also analysed (see Table 1). For comparison, two non-mountain municipalities from each county are included (marked with italics in the table). Two outliers can be discerned among the mountain municipalities; Sorsele, situated in Västerbotten, and Malung, in Dalarna, with 55% and 54% of respondents, respectively, supporting self-management. Åre, a ski resort municipality in Jämtland, shows the lowest support (27%) for self-management. Moreover, the support for self-management is about 10–20% lower in the non-mountain municipalities.

Table 1
Support for the three management alternatives in Sweden, the mountain counties and the mountain municipalities (%)

Sample	Self-management	Co-management	State management
Sweden	25	40	18
Mountain region	38	41	14
Average counties	29	42	17
Dalarna	32	41	14
Älvdalen	39	42	6
Malung	54	27	11
Mora	*36*	*31*	*17*
Falun	*26*	*42*	*18*
Jämtland	26	46	15
Härjedalen	40	37	16
Åre	27	45	15
Berg	34	37	17
Krokom	30	40	22
Strömsund	32	48	10
Östersund	*20*	*51*	*14*
Bräcke	*22*	*51*	*11*
Västerbotten	29	40	19
Storuman	42	37	11
Sorsele	55	35	4
Dorotea	43	40	9
Vilhelmina	40	40	7
Umeå	*28*	*38*	*22*
Skellefteå	*30*	*43*	*16*
Norrbotten	29	42	21
Arjeplog	34	52	10
Jokkmokk	37	44	13
Gällivare	38	42	20
Kiruna	40	41	12
Luleå	*27*	*39*	*23*
Piteå	*26*	*40*	*25*

4.2. *Explaining the differences and testing the hypotheses*

4.2.1. *Urban/rural impact*

To further examine the results of the above descriptive analysis, the regression analyses were begun by testing the explanatory power of the assumed urban/rural divide between primarily the mountain region and the rest of the country. First, a general variable measuring the degree of urbanization where respondents live was created. The results of these logistic regressions are shown in Table 2 (round a), showing that the degree of urbanization is only significant for state management. Thereafter a dummy variable for respondents living in and outside of the mountain region was constructed (round b), which proved to be significant for the self-management solution. When comparing both variables at the same time (round c), the urbanization variable becomes significant for state management and the mountain variable for self-management. State management supporters thus tend to live in more urbanized areas, while self-management supporters tend to live in the mountain region. It is, however, obvious that there is a need for complementary explanations due to the very low R^2-values (ranging from 0.000 to at most 0.004) that measures the overall fit of the tested model.

4.2.2. *Overall regression model*

Departing from the theoretical considerations presented above, three subsequent logistic regression analyses were conducted to test the hypotheses. In a very first test of the overall regression model, both the urbanization and mountain variables were included in order to control their real effects when compared to alternative hypotheses. Only the urbanization variable turned out to be significant and therefore the mountain variable was excluded in the model

Table 2
Logistic regression analyses—urbanization and mountain variables

Independent variables	Self-management			Co-management			State management		
	a	b	c	a	b	c	a	b	c
Urbanization	−0.018 (0.028)		−0.011 (0.028)	−0.044 (0.026)		−0.047 (0.026)	0.090 (0.032)**		0.085 (0.033)**
Mountain		0.480 (0.232)*	0.475 (0.235)*		−0.168 (0.229)	−0.229 (0.232)		−0.458 (0.319)	−0.362 (0.323)
Constant	−0.766 (0.104)***	−0.836 (0.035)***	−0.803 (0.106)***	0.078 (0.096)	−0.076 (0.032)*	−0.094 (0.098)	−1.612 (0.122)***	−1.286 (0.125)***	−1.588 (0.124)***
R^2 (Nagelkerke)	0.000	0.001	0.002	0.001	0.000	0.001	0.003	0.001	0.004
N	5027	5027	5027	5027	5027	5027	5027	5027	5027

Three subsequent rounds were conducted for each dependent variable: a—urbanization, b—mountain, and c—both urbanization and mountain. (*B*-values, SE within brackets.)
Levels of significance: *0.05-level; **0.01-level; ***0.001-level.

presented here (see Table 3). Background demographic variables were included in order to validate the results.

The effect of the urbanization model is actually strengthened in the overall model, but for *other* management solutions. When tested alone, urbanization was only significant for the state management outcome, but when tested together with the other hypotheses it is significant at the 0.001-level for co-management and at the 0.05-level for self-management. Hence, co-management seems to have more support in less urbanized areas, while self-management quite surprisingly tends to be more frequently preferred in more urbanized areas.

Dependency (hypothesis 1): Three variables were used to measure the effect of dependency: current or former employment in primary sectors (agriculture, forestry, fisheries, reindeer herding or mining), how often households eat game meat that someone in the household has hunted, and how important respondents find fishing and hunting to be for household economies. Employment in primary sectors tends to make people more supportive of co-management, while those employed in other sectors seem to support self-management. The latter result is somewhat surprising, in particular since the effect is high (*B*-value −0.490) as is also the significance level (0.001). However, the frequency of consumption of game meat is only significantly related to state management; the tendency is that the less often game meat is consumed, the greater the odds for approval of state management. Co-management advocates tend to state that fishing and hunting are economically important, while advocates of state management seem to think the contrary. The hypothesis is thus verified for co-management, which is shown to be preferred by resource dependent people, but not for self-management. It is also supported by the results for state management supporters; they do not consume much game meat and do not think that fishing and hunting are important for household economies.

Common understanding (hypothesis 2): Another three variables were used to measure the common understanding hypothesis. The first one considers the opinions on the quantity of protected areas in Sweden—whether there is too little, enough or too much protected nature. The findings indicate that self-management supporters think that there is too much protected nature in Sweden. Those respondents who prefer state management oppose this view and tend to answer that there is too little protected nature. These results are significant at the 0.001-level and *B*-values are rather high (0.389 and −0.308 for self- and state management respectively), which proves an important effect of this particular variable. The second variable measures the degree of importance respondents assign to protected areas *in the mountain region* for satisfying their general need of nature experiences. Co-management supporters assign a high and significant importance to those areas (*B*-value 0.272), while state management supporters tend to consider them less important (−0.319). For self-management advocates this variable is not

Table 3
Logistic regression analyses—the three management solutions (*B*-values, SE within brackets)

Independent variables	Self-management	Co-management	State management
Urban/rural			
Urbanization	0.090 (0.039)*	−0.122 (0.033)***	0.076 (0.041)
Dependency			
Primary sectors	−0.490 (0.151)***	0.291 (0.127)*	0.151 (0.154)
Use of game meat	0.060 (0.058)	0.027 (0.054)	−0.165 (0.076)*
Importance hunting and fishing	−0.054 (0.037)	0.097 (0.032)**	−0.096 (0.038)**
Common understanding			
PA quantity	0.389 (0.060)***	−0.094 (0.052)	−0.308 (0.064)***
Importance PA mountains	−0.079 (0.043)	0.272 (0.037)***	−0.319 (0.046)***
PA snowmobiling	0.257 (0.048)***	−0.119 (0.044)**	−0.162 (0.058)**
Trust			
Trust in other people	0.311 (0.115)**	−0.552 (0.109)***	0.461 (0.130)***
Trust in S politicians	−0.481*** (0.076)	0.323 (0.066)***	0.070 (0.079)
Satisfaction S democracy	−0.093 (0.078)	−0.169 (0.071)*	0.382 (0.090)***
Demographic variables			
Sex	0.420 (0.096)***	0.074 (0.083)	−0.570 (0.101)***
Age	0.019 (0.003)***	0.001 (0.003)	−0.023 (0.004)***
Education	−0.419 (0.049)***	0.075 (0.043)	0.384 (0.055)***
Constant	−1.550 (0.483)***	−0.525 (0.427)	−0.025 (0.522)
R^2 (Nagelkerke)	0.171	0.083	0.132
N	5291	5291	5291

Levels of significance: *0.05-level; **0.01-level; ***0.001-level.

significant. The third variable measures opinions about whether snowmobiling should be allowed in protected areas. It is significant for all three management solutions. Co- and state management advocates are against this, while self-management supporters would like to allow snowmobiling.

These results imply that the hypothesis cannot be verified, since the differences are seemingly considerable. However, these results must be interpreted with caution: the hypothesis is actually supported when simple frequencies are analysed separately for the mountain region and the national samples. The majority views are more or less the same in both samples, which imply that the degree of common understanding is high, and co- and self-management are also supported by both samples. Also when frequencies for supporters of the different management solutions are compared separately, the differences appear to be rather negligible. Even so, there are important minorities in the mountain region, in particular with regard to whether snowmobiling and hunting should be allowed in protected areas (35% for snowmobiling and 40% for hunting). For other forms of use of protected areas (such as fishing, commercial activities, fire and tenting) the support is about 10% higher in the mountain region than nationally. It is worth noting also that a majority of Swedes want riding and fishing to be allowed in protected areas, and that these areas were designed to facilitate visits.

Trust (hypothesis 3): Also the effect of trust was measured by three variables: general trust for other people, confidence for Swedish politicians, and satisfaction with democracy. Co-management is supported by respondents who tend to trust other people, while self-management *and* state management supporters have less trust in others. The significance levels are high (mostly 0.001), as well as the effects (*B*-values ranging from 0.311 to 0.552), which are among the highest in the model. Co-management advocates also appear to have significant confidence in Swedish politicians, although they are less satisfied with the Swedish democratic system. In contrast, self-management supporters tend not to trust politicians, and state management supporters appear satisfied with the democratic system. Most of these results are also highly significant and with rather high *B*-values. The hypothesis is partly verified for all three management solutions, but for co- and state management the results contradict theory to some extent.

Demographic variables: Sex, age and education level are all highly significant to explain preferences for self- and state management, but not for co-management. Women, older people and less educated tend to prefer self-management, while men, younger people and highly educated choose state management.

Finally, it is also needed to take into consideration how well the model describes, or fits, the observed data. Several measures of 'goodness of fit' exist, and each of them may give different results. There is considerable disagreement about which is best, and none of them has the seemingly straightforward interpretation that the multiple regression coefficient, R^2, has (Buttolph-Johnson, 2005). According to Pampel (2001, p. 50), "Researchers should use these measures as only rough guides without attributing great importance to a precise figure". I have chosen to rely on the Nagelkerke R^2, which certainly does not represent explained variation. For my model it suggests that the proportional improvement when adding the independent variables is 0.169 (or about 17%) for the self-management solution, 0.132 (about 13%) for the state management solution, and 0.085 (about 8.5%) for the co-management solution. These values show that this model can partly explain why people want self-, co- or state management, but since they are quite low further studies are needed to fully understand what characterizes supporters of different management solutions.

5. Discussion

Urban/rural divisions are often discussed in relation to issues of nature conservation such as protected areas and carnivores, and the results of this study first pointed in this direction too. Simple frequencies showed that people in the mountain region, where the great majority of protected areas are found, actually had a slightly different view. Here, the support for self-management is almost ten percent higher than in the rest of the country, including the counties where the mountains are situated. Then, however, the logistic regressions showed that the degree of urbanization is actually more important: co-management tend to be preferred by people living in more rural areas while self-management actually seems to be supported by urban dwellers. These results are contradictory, but strengthen the view of researchers such as Heberlein and Ericsson who suggest that the categorization urban/rural is too broad. Cities, in particular, are home to such diverse populations that generalizations are difficult. The effects (*B*-values) of the urbanization variable are also rather low, which implies that the other three hypotheses might better explain which characteristics are important to understand support for different management solutions.

The first hypothesis concerned the impact of resource dependency and it was only supported to a certain extent, for co- and state management. It has been proposed that employment in primary sectors has an impact on attitudes toward conservation issues, and it can be seen as a dependency-related factor. For example, if a person is working in forestry, a ban on cutting down trees will restrict his ability to find work. This factor was shown to have some impact; co-management supporters are more often employed in these sectors compared to self-management supporters. Co-management advocates also think that hunting and fishing is important for the households' economies, which strengthens the result that they are to some extent dependent on natural resources. Nevertheless, the only direct measure of dependency—namely, the consumption of meat that someone in the family has hunted—proved not significant for co-management.[4] It was reversely significant for state management; the less game meat is consumed; the higher is the support for the state. Self-management supporters, who were assumed to be resource dependent, appear less dependent on natural resources. This result corresponds with the previous one that they tend to live in more urban areas.

Rural location and resource dependency are common explanations in the literature to why people oppose protected areas, but in this study, this kind of people tend to choose co-management. One reason to why the theory is not being verified could be that the salience of natural resources after all is quite low in a welfare state such as Sweden. Another reason could be that *the nature* of the resource—whether it is complex and multiple-use or more simple and single-use—may affect which solutions are preferred. If the resource-dependent users in this survey have realized that nature conservation is a global interest, they might accept the need for government to be involved *together* with them in a co-management arrangement. With a less complex common-pool resource, self-management might have had more support among the resource-dependent users.

The second hypothesis of common understanding, as formulated in Section 2.3.2, is being undermined by the result that urban/rural is a too broad categorization. It needs to be reformulated to capture whether common understanding exists among all the relevant user groups (such as hunters, hikers, reindeer herders, etc). Nevertheless, the regressions did underline the major differences between advocates of different management solutions: their opinions on the amount of protected areas and their usage. To recall, self-management supporters tend to think that there are too many protected areas and that more activities should be allowed, while co- and state management supporters tend to think the opposite. In general, however, the Swedish population wants to change the current protected area policy only to some extent; they would like riding and fishing to be allowed in these areas and that they were designed to facilitate visits. More substantial change is demanded for in management structure—the majority wants protected areas to be co- or self-managed.

While this view supports that protected areas are important and that their primary objective is to protect nature, it also challenges the traditional 'fortress' policy of protected areas stating that they should be kept as untouched by humans as possible so that 'pristine wilderness' can be conserved. Perhaps it is time to stop just saying that the new conservation policy has moved beyond

[4]It is somewhat surprising that this variable is not more significant, since the use of game meat is important. For instance, close to half of the households in the mountain region regularly consume game meat provided by someone in the family, between 21% and 41% in the mountain counties, and 9% in the country as a whole (see also Ericsson et al., 2005).

the wilderness idea, and actually start implementing it. This would require the establishment of regulations for protected areas that allow for certain use and of monitoring mechanisms that give feedback on how the nature protection objectives develop.

The third hypothesis considered the effect of trust and was partly verified by the results in this study; respondents who prefer self-management seem to have a low level of trust in fellow human beings as well as in politicians, co-management supporters tend to trust others and politicians even though they are not satisfied with how Swedish democracy works, and state management advocates trust others less but are more satisfied with democracy. The assumptions that are not verified are that co-management should imply trust in, or satisfaction with, the democratic system, and that state management supporters should trust other people. Why co-management advocates tend not to be satisfied with democracy is probably related to the fact that this policy domain has been (and still is) run by largely centralized authorities in Sweden. When the democratic system does not correspond to one's ideals, it might be natural to become dissatisfied.

It is, however, more difficult to understand why Rothstein is proved wrong in this case, namely that general trust in democracy is not accompanied by trust in fellow human beings. A possible explanation could be that the central administration has implemented a policy, which satisfies state management supporters, but not a majority of Swedes. Therefore, state management supporters might want the state to keep the responsibility in order to guarantee a continuation of 'their' policy. However, as already discussed, the differences between these supporters and others are not that large. In Swedish media though, the reporting about nature conservation issues often emphasizes a polarization between the supposed two camps of local, rural people and nature conservationists. In accordance, general perceptions on who thinks what about nature conservation might be distorted.

6. Conclusions

This study modifies the current emphasis on rural versus urban views in nature conservation issues suggesting that other variables are more important to explain why different management solutions are preferred. Three such explanations were tested. First, resource dependency was shown to be related to support for co-management, but not self-management. Second, common understanding was confirmed as important, and co- and self-management are correspondingly preferred by an overwhelming majority of 65%. Demands for increased local influence on how protected areas are managed thus not only come from marginalized communities, but are supported by the general opinion in Sweden. Third, trust was proven significant as well, although not entirely as prescribed by theory. Overall, common understanding and trust seem to be more important than resource dependency.

Even if these explanations were shown to be important, the theoretical assumptions were not entirely fulfilled. Self- and co-management supporters were assumed to have similar characteristics, but they actually differed substantially. This suggests that it should be clarified that self- and co-management are analytically separate entities, although both are common property regimes. Self-management, whose supporters only seem to trust their closest, would then be suitable for very small-scale situations. Co-management, on the other hand, would be suitable for more complex situations where user groups need to interact with governmental organizations whose goals are decided by politicians. Self- or co-management is thus a question of scales.

Finally, the second objective in this study was to test the usefulness of quantitative survey methods in co-management research. It did indeed prove to be an interesting approach for undertaking investigation of a resource of national and local interest. However, it was difficult to construct hypotheses to test on management of protected areas, which need to address multiple user groups at multiple scales, when the theoretical prescriptions seem to apply primarily to self-organization on the local level. In order to further increase our understanding of the vertical linkages that connect self-organized layers to the polycentric system in which it is embedded, more research is needed. Surveys matching the scales of the problem may be one useful way to move forward!

Acknowledgments

I gratefully acknowledge Mistra, the Swedish Environmental Protection Agency, and the county administrative boards of Norrbotten, Västerbotten, Dalarna, and Västernorrland for funding of the survey, and the Mountain Mistra Programme for processing the data. Thanks also to the participants at the Methodology Seminar in Lund, 2005, and to three anonymous reviewers for valuable comments.

References

Anttila, S., 1999. The Snowmobile Issue as a Commons Dilemma. Rapport 1999:2. Fjällforskningsinstitutet, Östersund.

Baland, J.-M., Platteau, J.-P., 1996. Halting Degradation of Natural Resources: Is there a Role for Rural Communities? Clarendon Press, Oxford, NY.

Bardhan, P., Dayton-Johnson, J., 2002. Unequal irrigators: heterogeneity and commons management in large-scale multivariate research. In: Ostrom, E., et al. (Eds.), The Drama of the Commons. National Academies Press, Washington, DC, pp. 87–112.

Berkes, F. (Ed.), 1989. Common Property Resources. Ecology and Community-Based Sustainable Development. Belhaven Press, London.

Berkes, F., 2002. Cross-scale institutional linkages: perspectives from the bottom up. In: Ostrom, E., et al. (Eds.), The Drama of the Commons. National Academy Press, Washington, DC.

Berkes, F., Folke, C. (Eds.), 1998. Linking Social and Ecological Systems. Management Practices and Social Mechanisms for Building Resilience. Cambridge University Press, Cambridge.

Berkes, F., George, P., Preston, R.J., 1991. Co-management: the evolution in theory and practice of the joint administration of living resources. Alternatives 18 (2), 12–18.

Brandon, K., Redford, K.H., Sanderson, S.E. (Eds.), 1998. Parks in Peril. The Nature Conservancy Island Press, Washington, DC.

Brechin, S.R., Wilshusen, P.R., Fortwangler, C.L., West, P.C. (Eds.), 2003. Contested Nature. State University of New York Press, Albany, NY.

Bromley, D.W. (Ed.), 1992. Making the Commons Work. Theory, Practice, and Policy. Institute for Contemporary Studies Press, San Fransisco.

Buttolph-Johnson, J., 2005. Political Science Research Methods. CQ Press, Washington, DC.

Campbell, T., 1996. Co-management of aboriginal resources. Information North 22 (1).

Carlsson, L., 2000. Policy networks as collective action. Policy Studies Journal 28 (3), 502–520.

Carrus, G., Bonaiuto, M., Bonnes, M., 2005. Environmental concern, regional identity, and support for protected areas in Italy. Environmental Behaviour 37 (2), 237–257.

Dietz, T., Dolšak, N., Ostrom, E., Stern, P.C., 2002. Introduction: the drama of the commons. In: Ostrom, E., et al. (Eds.), The Drama of the Commons. National Academy Press, Washington, DC, pp. 3–35.

Dolšak, N., Ostrom, E. (Eds.), 2003. The Commons in a New Millennium. MIT Press, London.

Edwards, V., Steins, N.A., 1999. Special issue introduction: the importance of context in common pool resource research. Journal of Environmental Policy and Planning 1, 195–204.

Ericsson, G. et al., 2005. Delrapport om jakt och fiske—omfattning, betydelse och förvaltning. FjällMistrarapport 14. The Mountain Mistra Programme, Umeå.

Folke, C., 2006. The economic perspective: conservation against development versus conservation for development. Conservation Biology 20 (3), 686–688.

Ghimere, K., Pimbert, M.P., 1997. Social Change and Conservation. Earthscan Publications Limited and UNRISD, London.

Gibson, C.C., McKean, M.A., Ostrom, E. (Eds.), 2000. People and Forests. Communities, Institutions, and Governance. MIT Press, London.

Gordon, S.H., 1954. The economic theory of a common-property resource. Journal of Political Economy 62, 124–142.

Gorenflo, L.J., Brandon, K., 2006. Key human dimensions of gaps in global biodiversity conservation. BioScience 56 (9), 723–731.

Hardin, G., 1968. The tragedy of the commons. Science 162, 1243–1247.

Heberlein, T.A., Ericsson, G., 2005. Ties to the countryside: accounting for urbanites attitudes toward hunting, wolves, and wildlife. Human Dimensions of Wildlife 10 (3), 1–23.

Imperial, M.T., 1999. Institutional analysis and eco-system-based management: the institutional analysis and development framework. Environmental Management 24 (4), 449–465.

Jentoft, S. (Ed.), 1998. Commons in a Cold Climate. UNESCO and Parthenon Publishing, Paris, NY.

Jentoft, S., 2003. Co-management—the way forward. In: Wilson, D.C., Nielsen, J.R., Degnbol, P. (Eds.), The Fisheries Co-management Experience: Accomplishments, Challenges and Prospects. Fish and Fisheries Series. Kluwer Academic Publishers, Dordrecht, pp. 1–14.

Jentoft, S., McCay, B.J., 2003. The place of civil society in fisheries management. In: Wilson, D.C., Nielsen, J.R., Degnbol, P. (Eds.), The Fisheries Co-management Experience. Fish and Fisheries Series. Kluwer Academic Publishers, Dordrecht, pp. 293–307.

Lam, W.F., 1998. Governing Irrigation Systems in Nepal: Institutions, Infrastructure, and Collective Action. ICS Press, San Fransisco, CA.

McCay, B.J., Acheson, J.M. (Eds.), 1987. The Question of the Commons. The Culture and Ecology of Communal Resources. University of Arizona Press, Tucson, USA.

Mulder, M.B., Coppolillo, P., 2005. Conservation. Princeton University Press, Princeton.

Naturvårdsverket, 2006a. Skyddad natur. ⟨http://www.naturvardsverket.se/index.php3?main=/dokument/natur/naturres/reservat.html⟩, 17/10/2006.

Naturvårdsverket, 2006b. Natur och naturvård. ⟨http://www.naturvardsverket.se/index.php3?main=/dokument/natur/naturres/reservat.html⟩, 17/10/2006.

Naturvårdsverket, 2006c. Right of Public Access. ⟨http://www.allemansratten.se/templates/Page.asp?id=2060⟩, 17/10/2006.

Nepal, S.K., Weber, K.E., 1995. Managing resources and resolving conflicts: national parks and local people. International Journal of Sustainable Development and World Ecology 2, 11–25.

Oakerson, R., 1992. Analyzing the commons: a framework. In: Bromley, D.W. (Ed.), Making the Commons Work. Theory, Practice, and Policy. The Institute for Contemporary Studies Press, San Fransisco.

Ostrom, E., 1990. Governing the Commons. Cambridge University Press, New York.

Ostrom, E., 1998. A behavioral approach to the rational choice theory of collective action. American Political Science Review 92 (1), 1–22.

Ostrom, E., 2000. Reformulating the commons. Swiss Political Science 6 (1), 29–52.

Ostrom, E., 2005. Understanding Institutional Diversity. Princeton University Press, Princeton, Oxford.

Ostrom, E., et al. (Eds.), 2002. The Drama of the Commons. National Academies Press, Washington, DC.

Ostrom, E., Gardner, R., Walker, J., 1994. Rules, Games, and Common-Pool Resources. The University of Michigan Press, Ann Arbor, MI.

Pampel, F.C., 2001. Logistic Regression: A Primer. Sage Quantitative Applications in the Social Sciences Series, 132. Sage Publications, Thousand Oaks.

Pinkerton, E. (Ed.), 1989. Co-operative Management of Local Fisheries: New Directions for Improved Management and Community Development. University of British Columbia Press, Vancouver.

Rådelius, C., 2002. Självstyre eller samförvaltning? Problem och möjligheter utifrån en studie av världsarvet Laponia. Licentiate Thesis, Luleå University of Technology, Luleå.

Rothstein, B., 2003. Sociala fällor och tillitens problem. SNS Förlag, Kristianstad.

Rudd, M.A., 2004. An institutional framework for designing and monitoring ecosystem-based fisheries management policy experiments. Ecological Economics 48, 109–124.

SCB, 2006. Skyddad natur 31 dec 2004. Sveriges officiella statistik. Statistiska meddelanden MI 41SM 0501. ⟨http://www.scb.se/statistik/MI/MI0603/2004A02/MI0603_2004A02_SM_MI41SM0501.pdf⟩, 07/11/2006.

Skogen, K., Krange, O., 2003. A wolf at the gate: the anti-carnivore alliance and the symbolic construction of community. Sociologia Ruralis 43 (3), 309–325.

Skr 2001/02:173. En samlad naturvårdspolitik. Miljö- och samhällsbyggnadsdepartementet.

Stevens, S. (Ed.), 1997. Conservation through Cultural Survival. Indigenous Peoples and Protected Areas. Island Press, Washington, DC.

Stoll-Kleeman, S., 2001. Barriers to nature conservation in Germany: a model explaining opposition to protected areas. Journal of Environmental Psychology 21, 369–385.

Svarstad, H., 2003. Rovviltkonflikter i noen europeiske land. NINA Fagrapport 68, NINA, Trondheim.

Wade, R., 1988. Village Republics: Economic Conditions for Collective Action in South India. ICS Press, San Fransisco.

Wilson, M.A., 1997. The wolf in yellowstone: science, symbol or politics? Deconstructing the conflict between environmentalism and wise use. Society & Natural Resources 10 (5), 453–468.

Journal of Environmental Management 87 (2008) 165–176

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Environmental performance indicators: An empirical study of Canadian manufacturing firms

Jean-François Henri*, Marc Journeault

École de Comptabilité, Université Laval, Que., Canada G1K7P4

Received 31 January 2006; received in revised form 2 October 2006; accepted 9 January 2007
Available online 19 March 2007

Abstract

The aim of this exploratory study is to examine the importance of measurement and use of environmental performance indicators (EPIs) within manufacturing firms. Two research questions are investigated: (i) To what extent are firm characteristics associated with the importance of measurement of various categories of EPIs? (ii) To what extent are firm characteristics associated with global and specific uses of EPIs? More specifically, this paper examines four uses of EPIs (i.e. to monitor compliance, to motivate continuous improvement, to support decision making, and to provide data for external reporting) as well as four characteristics of firms, namely environmental strategy, International Organization for Standardization (ISO) 14001 compliance, size, and ownership. This study contributes to the environmental management accounting literature by collecting and analyzing empirical evidence that provides a better understanding of the associations among firm characteristics and EPIs.

Keywords: Environmental performance indicators; Environmental strategy; ISO 14001

1. Introduction

The aim of this exploratory study is to examine the importance of measurement and the use of environmental performance indicators (EPIs) within manufacturing firms in Canada. EPIs represent numerical measures providing key information related to environmental issues. Several reasons justify the importance of EPIs as a significant component of the environmental management system (EMS) (Eckel et al., 1992; Figge et al., 2002; Schaltegger and Burritt, 2000; Epstein, 1994). First, organizations are increasingly being held responsible for environmental actions, as reflected by the growing number of laws, regulations, and penalties in this area. Consequently, organizations are now obliged to measure, control, and disclose their environmental performance. Second, reliable EPIs are necessary to supply information for decision making while ensuring the attainment of environmental objectives. Third, the allocation of the organization's limited resources to environmental problem solving requires persuasive evidence supporting the benefits of such actions. The environmental system must therefore be able to supply information concerning the cost of reducing risks and concerning the measurement of this reduction. Lastly, as several studies have demonstrated, performance indicators are effective tools for improving business practices and organizational performance (e.g., Hoque and James, 2000; Baines and Langfield-Smith, 2003; Ittner et al., 2003; Said et al., 2003). Although no clear empirical evidence has been provided, it is believed that EPIs may also have the capacity to improve environmental performance.

In this study, we examine specifically the association among firm characteristics and two dimensions of EPIs, namely the importance of measurement and use. This choice has been motivated by numerous studies in the management accounting literature that have examined those two aspects of performance indicators (e.g., Scott and Tiesen, 1999; Ittner et al., 2003; Hoque and James, 2000; Henri, 2006a, b; Bisbe and Otley, 2004; Chenhall, 2005). Indeed, the importance of measurement and use of indicators constitute fundamental dimensions of any

*Corresponding author. Tel.: +1418 656 7737; fax: +1418 656 7746.
E-mail address: Jean-Francois.Henri@ctb.ulaval.ca (J.-F. Henri).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.009

information system. While the former refers to the content of the information system (i.e., what is measured), the latter refers to the manner in which the information is used by managers (i.e., how it is used).

Two research questions are investigated: (i) To what extent are firm characteristics associated with the importance of measurement of various categories of EPIs? (ii) To what extent are firm characteristics associated with global and specific uses of EPIs? More specifically, this paper examines the importance of measurement of EPIs based on two classifications (i.e. financial/non-financial indicators, and International Organization for Standardization (ISO) 14031 classification) and four uses of EPIs (i.e. to monitor compliance, to motivate continuous improvement, to support decision making, and the provision of data for external reporting). Moreover, four firm characteristics are analyzed, namely environmental strategy, ISO 14001 compliance, size, and ownership. Two purposes motivate the choice of these factors.

First, while numerous studies have examined the relationships among those factors and various organizational practices, systems or routines (e.g., Melnyk et al., 2003; Sharma and Vredenburg, 1998; Marshall and Brown, 2003; Aragon-Correa, 1998), scant attention has been devoted to their association with EPIs. Globally, as mentioned by Burritt (2004), the empirical literature is lacking in the area of environmental management accounting (EMA).

Secondly, from a theoretical standpoint, opposite viewpoints can be provided to argue the presence or absence of relationships among the four firm characteristics and EPIs. For instance, organizations following a more active environmental strategy may place greater importance on the measurement and use of EPIs to favor the alignment of actions toward the attainment of environmental objectives. However, strategy and performance measurement are two different concepts that are not automatically linked. The literature offers various examples of lack of coherence between strategy and performance indicators within different organizations (e.g., Kerr, 1995; Ittner and Larcker, 2003). Moreover, firms implementing ISO 14001 standards may have the organizational resources and structure to more effectively collect and report environmental practices data. However, a firm may have developed various EPIs and used them intensively while not implementing ISO 14001. Similarly, ISO 14001 standards do not warrant the diversity of EPIs as their extensive use for various purposes either.

In terms of size, while larger organizations may have more resources to invest in the development of EPIs, those resources are not necessarily allocated to environmental issues. Instead, top management may assign those resources to other organizational priorities or critical uncertainties. Lastly, while public firms may face more pressure from various stakeholders to develop and report EPIs, both public and private firms face similar regulatory regimes. Hence, both types of firms may use EPIs to manage their environmental impacts. In summary, there is a need to collect and analyze empirical evidence to provide a better understanding of the associations among the four firm characteristics and EPIs.

2. Theory

2.1. Definition of EPIs

EPIs represent numerical measures, financial or non-financial, that provide key information about environmental impact, regulatory compliance, stakeholder relations, and organizational systems (Veleva and Ellenbecker, 2000; Ilinitch et al., 1998; Chinander, 2001). EPIs refer to the measurement of the interaction between the business and the environment (Olsthoorn et al., 2001). They represent the quantification of the effectiveness and efficiency of environmental action with a set of metrics (Neely et al., 1995). The indicators act as surrogates or proxies for organizational phenomena (Ijiri, 1975).

EPIs are one component of EMA. EMA can be defined as the management of environmental and economic performance through the development and implementation of appropriate environmental-related accounting systems and practices. While this may include reporting and auditing in some companies, EMA typically involves life-cycle costing, full-cost accounting, benefits assessment, and strategic planning for environmental management. The EMA is considered in turn as one component of EMS. The latter refers to the formal systems and databases that integrate procedures and processes for the training of personnel, monitoring, summarizing, and reporting of specialized environmental performance information to internal and external stakeholders of the firm (Melnyk et al., 2003). The two dimensions of EPIs that will be examined in this study, namely the importance of measurement and use, are described next.

2.2. EPIs—importance of measurement

The importance of measurement refers to the attention devoted by firms to the quantification of various environmental issues. Prior research has identified four dimensions of environmental performance that EPIs should measure: internal, external, process, and result (Lober, 1996; Ilinitch et al., 1998). To capture those four dimensions, two classifications are examined in this study: (i) financial and non-financial indicators and (ii) ISO 14031 guidelines. The distinction between financial and non-financial indicators is a widely used classification in the current literature that covers the four dimensions described above (e.g., Baines and Langfield-Smith, 2003; Said et al., 2003; Kaplan and Norton, 2001; Gosselin, 2005). The ISO 14031 guidelines also refer to the four dimensions and represent international standards that are well recognized, accepted, and implemented around the world. These two classifications are described in further detail below.

2.2.1. Financial and non-financial indicators

In spite of their capacity to present results of decisions in a comparable measurement unit, to capture the cost of trade-offs between resources and the cost of spare capacity, and to support contractual relationships and capital markets, financial measures have been criticized for several reasons (Atkinson et al., 1997; Epstein and Manzoni, 1997; Fisher, 1992; Kaplan and Norton, 1992, 1996). Those criticisms have led to the emergence of non-financial measures. In comparison with non-financial measures, financial measures are perceived as too historical and backward looking, lacking predictive ability to explain future performance, rewarding short-term or incorrect behavior, lacking actionability, lacking timely signals, being too aggregated and summarized to guide managerial action, reflecting functions instead of cross-functional processes, and providing inadequate guidance to evaluate intangible assets (Ittner and Larcker, 1998). However, the link between improvement in non-financial measures and profits is unclear and sometimes impossible to assess directly.

In summary, both types of information capture different aspects of the various facets of organizational effectiveness. The ends and outputs are revealed by financial measures while the means and processes are reflected by non-financial measures. Hence, both types of information are useful for managers. In terms of environmental performance, given the nature and diversity of the factors measured (e.g. monetary resources invested in the environment, atmospheric emissions, waste water, the number of environmental audits, the number of environmental non-compliance incidents), EPIs generally integrate both financial and non-financial measures.

2.2.2. ISO 14031 guidelines

A second possible grouping of EPIs is the classification according to ISO 14031 guidelines. This standard, a sub-category of ISO 14001, concerns the evaluation of environmental performance. It proposes guidelines for the development of monitoring and measurement tools that evaluate the efficiency of an environmental system. This standard proposes three categories of EPIs (Bennett and James, 1998; Marshall and Brown, 2003):

(1) Environmental condition indicators (ECIs): defined as specific expressions that provide information about the local, regional, national, or global condition of the environment. Those measures include (i) receptor indicators (e.g. ecotoxicity, biological oxygen demand, (ii) sustainability indicators (e.g. emissions of a substance per volume of production or per unit of value added), and (iii) proxy ECIs (i.e. indicators that express emissions and waste data in terms of their capacity to cause environmental damage).
(2) Operational performance indicators (OPIs) provide information about the environmental performance of an organization's operations. They include (i) input of materials, energy, and services, (ii) operation of facilities and equipment and logistics, and (iii) output of products, services, waste, and emissions.
(3) Management performance indicators (MPIs) provide information about management's efforts to influence an organization's environmental performance. Four sub-categories are identified: (i) implementation of policies and programs, (ii) conformity of actions with requirements or expectations, (iii) community relations, and (iv) environment-related financial performance.

2.3. EPIs—use

Despite considerable interest in the different uses of performance indicators in the management accounting literature (e.g. Atkinson et al., 1997; Henri, 2006a, b; Ittner et al., 2003; Simons, 2000), little attention has been devoted to the various types of uses of EPIs in the EMA literature (notable exceptions include Bennett and James (1998) and Briassoulis (2001). Indeed, despite a considerable body of literature examining the generic use of various EPIs (i.e., global and undifferentiated use of indicators), the specific manner in which those indicators are used by managers as control mechanisms, motivation tools, or communication devices has been overlooked empirically. Four types of uses are reflected in the accounting literature, namely monitoring, attention-focusing and signaling, decision making, and external reporting. From the overlap between the accounting and environment literature, four main uses are reflected and examined in this study: (i) to monitor compliance with environmental policies and regulation, (ii) to motivate continuous improvement, (iii) to provide data for internal decision making, and (iv) to provide data for external reporting.

2.4. Firm characteristics

In this study, the relationships among four contextual factors and EPIs are examined, namely environmental strategy, ISO 14001 compliance, size and ownership. First, the environmental strategy is operationalized in the literature using several typologies (e.g., Ullmann, 1985; Hunt and Auster, 1990; Roome, 1994; Hart, 1995; Clarkson, 1995). These classifications group organizations according to the level of deployment of their organizational strategy at the environmental level. Although some classifications are more detailed, the categorization used by Ullmann (1985) advantageously presents a dichotomy of antithetical levels of corporate environmental strategy: active and passive. Accordingly, passive organizations are described as having little or no managerial involvement, little or no environmental management and integration, little or no employee involvement and training, and few or no resources allocated to environmental performance. Conversely, active organizations have medium or high managerial involvement, partial or complete integration of the environment function, moderate or substantial employee involvement and training, and moderate or considerable resources allocated to the attainment of environmental objectives.

More specifically, various elements can be used to determine whether an organization follows an active or passive environmental strategy based on components such as filters and controls on emissions and discharges, residue recycling, use of environmental arguments in marketing, environmental aspects in administrative work, periodical environmental audits, purchasing manuals with ecological guidelines, environmental seminars for executives, environmental training for firms' employees, total quality program with environmental aspects, pollution damage insurance, environmental management manuals for internal use, environmental analysis of a product's life cycle, participation in government-subsidized environmental programs, and sponsorship of environmental events (Aragon-Correa, 1998).

Secondly, developed by the ISO, the ISO 14000 series of standards addresses various aspects of environmental management (Veleva and Ellenbecker, 2000). The ISO 14001 specifies the need for an EMS that represents a structured approach to setting and attaining environmental objectives and targets, and demonstrating that they have been achieved by management (Veleva and Ellenbecker, 2000). This voluntary set of standards is intended to encourage organizations to systematically address the environmental impacts of their activities (Pringle et al., 1998). In this study, we examine whether the firms are ISO-compliant or not. Lastly, in this study, size is defined as the total number of employees within the firm. Ownership refers to the public or private nature of the organization.

3. Study design

3.1. Data collection

Data were collected using a survey design. A random sample comprised of 1500 Canadian manufacturing organizations was formed based on the Scott's Manufacturing 2004 database. The sample contains organizations that have 100 employees or more, and report sales of over $20 million annually. These criteria are intended to ensure that organizations are large enough for organizational and strategic variables to apply (Miller, 1987) and that management control systems are sufficiently developed (Bouwens and Abernethy, 2000).

The questionnaire was first validated using a pre-test administered to various academics and managers. This pre-test validated an understanding of each of the measurement instruments. Then, the questionnaire was sent to the CEO or another member of the top management team (COO or senior vice-president). A letter presenting the purpose of the study and a self-addressed stamped envelope was included with the questionnaire. Three weeks after the initial mailing, 500 organizations randomly selected from among the non-respondents received a reminder by telephone. All the other organizations that did not respond to the questionnaire following the initial mailing and that were not selected for the telephone follow-up received a replacement questionnaire.

The final sample comprised 1447 organizations (considering wrong addresses, organizations that moved, etc). In total, 303 usable questionnaires were received, for a response rate of 20.9%. On average, organization size was 710 employees and the respondents had on average 13.7 years of experience working for their organization. Appendix A presents a complete profile of the respondents. An analysis of the non-response bias was performed to confirm the validity of the data. Initially, the comparison between respondents and non-respondents with respect to size, industry and geographical region did not reveal any significant differences. Moreover, the comparison between the first and last 10% of respondents (the latter being used as a proxy for the non-respondents) did not reveal any significant differences in the responses obtained for the main constructs of the study (e.g. EPIs and firm characteristics).

3.2. Measurement of constructs

Table 1 presents the instruments used to measure the various constructs as well as the descriptive statistics and the correlations matrix. Respondents were asked about the importance of measurement of EPIs using an instrument developed based on the ISO 14031 standard (panel A). The instrument included 13 items from the three categories described above ranging on a seven-point Likert-type scale. Respondents were asked about the use of EPIs with an instrument developed by Bennett and James (1998) containing four items (panel B) and measured on a seven-point Likert-type scale. Prior work suggests that the various uses of management control systems are likely to be correlated or exhibit overlap (Shields and Shields, 1998; Hansen and Van der Stede, 2004). Considering that our objective is to examine the association among firm characteristics and specific uses of EPIs, we remove the common factor among the various uses of EPIs and focus on its "unique" element. In order to focus on the uniqueness associated with each use, we follow the work of Hansen and Van der Stede (2004). Specifically, we use the residuals from regressing each use of EPIs on the other three as key variables in the analyses. Appendix B presents the initial work to determine whether the four uses were correlated yet sufficiently unique to justify analysis by itself.

Respondents were asked about the type of environmental strategy implemented by their organizations using an instrument developed by Aragon-Correa (1998), which includes 14 items (panel C). Answers were measured using a seven-point Likert-type scale. Organizations that have an above average mean score of environmental practices compared with the total respondent population are

Table 1
Measurement of variables and descriptive statistics

(A) EPIs—importance of measurement

(i) Specific measures (*Note*: 1 = not important at all, 7 = very important)

EPIs	Mean	Std. Dev.	Median	Min.	Max.	ISO 14031 classification			Fin/non-fin classification	
						ECI	OPI	MPI	Fin	Non-Fin
Conformity with requirements or expectations	5.70	1.92	6.0	1	7			X		X
Inputs of energy	5.50	1.67	6.0	1	7		X			X
Community relations	5.20	2.06	6.0	1	7			X		X
Outputs of solid waste	5.13	1.83	6.0	1	7		X			X
Outputs of air emissions	5.10	2.02	6.0	1	7		X			X
Financial impact	4.92	1.94	5.0	1	7			X	X	
Installation, operation, and maintenance of the physical facilities and equipment	4.91	1.68	5.0	1	7		X			X
Outputs of waste water	4.79	2.37	6.0	1	7		X			X
Inputs of raw materials	4.76	2.22	5.0	1	7		X			X
Inputs of water	4.76	2.12	5.0	1	7		X			X
Implementation of environmental policies and programs	4.72	2.04	5.0	1	7			X		X
Inputs of auxiliary materials	4.60	2.08	5.0	1	7		X			X
Indicators providing information on the local, regional, or national condition of the environment	3.75	2.31	4.0	1	7	X				X
Mean	4.91	1.39	5.20	1	7					

(ii) Categories of measures (significant at the .05 level, **significant at the .01 level)

ISO 14031 indicators	Mean	Std. Dev.	Median	Min.	Max.	Correlation matrix (Pearson)		
						MPI	OPI	ECI
Management performance indicators (MPI)	5.13	1.68	5.50	1	7	1		
Operational performance indicators (OPI)	4.97	1.40	5.29	1	7	.70**	1	
Environmental condition indicators (ECI)	3.75	2.31	4.0	1	7	.61**	.54**	1

Financial/non-financial indicators	Mean	Std. Dev.	Median	Min.	Max.	Correlation matrix (Pearson)	
						Fin	Non fin
Financial	4.92	1.94	5.0	1	7	1	
Non-financial	4.91	1.39	5.08	1	7	.65**	1

(B) EPIs—use (*Note*: 1 = not used at all, 7 = used to a very great extent)

Use of environmental performance indicators	Mean	Std. Dev.	Median	Min.	Max.
Monitor internal compliance with environmental policies and regulations	5.31	1.91	6.0	1	7
Motivate continuous improvement	5.16	1.78	6.0	1	7
Provide data for internal decision-making	5.01	1.79	5.0	1	7
Provide data for external reporting	4.48	2.05	5.0	1	7
Mean	4.99	1.68	5.50	1	7

(C) Environmental strategy (*Note*: 1 = not important at all, 7 = very important)

Environmental strategy	Mean	Std. Dev.	Median	Min.	Max.
Filters and controls on emissions and discharges	5.76	1.49	6.0	1	7
Residue recycling	5.40	1.60	6.0	1	7
Use of natural environmental arguments in marketing	3.63	1.77	4.0	1	7
Natural environmental aspects in administrative work	3.80	1.62	4.0	1	7
Periodic natural environmental audits	4.44	1.96	5.0	1	7

Table 1 (*continued*)

(C) Environmental strategy (*Note*: 1 = not important at all, 7 = very important) Environmental strategy	Mean	Std. Dev.	Median	Min.	Max.
Purchasing manual with ecological guidelines	3.31	1.75	3.0	1	7
Natural environmental seminars for executives	3.30	1.60	3.0	1	7
Natural environmental training for firm's employees	3.81	1.75	4.0	1	7
Total quality program with natural environmental aspects	4.16	1.81	4.0	1	7
Pollution damage insurance	3.84	2.01	4.0	1	7
Natural environmental management manual for internal use	4.17	2.06	4.0	1	7
Natural environmental analysis of product life cycle	3.32	1.78	3.0	1	7
Participation in government-subsidized natural environmental programs	2.94	1.72	3.0	1	7
Sponsorship of natural environmental events	2.87	1.67	3.0	1	7
Mean	3.91	1.25	4.0	1	6.50

(D) ISO 14001	N	%
Step 1: Not being considered	93	30.7
Step 2: Future consideration	50	16.5
Step 3: Assessing suitability	17	5.6
Step 4: Planning to implement	15	4.9
Step 5: Currently implementing	22	7.3
Step 6: Successfully implemented	106	35.0
Total	303	100

(E) Size	Mean	Std. Dev.	Median	Min.	Max.
Number of employees (log)	2.57	.42	2.48	1.08	4.34

(F) Ownership	N	%
Public	119	39.3
Private	184	60.7
Total	303	100

considered to be active, whereas organizations with a mean score below the average are considered to be passive. The ISO 14001 compliance was measured using an instrument developed by Melnyk et al. (2003). Respondents were asked to identify the stages of development of the ISO 14001 standard into six steps. Accordingly, organizations situated in steps five (*currently implementing*) and six (*successfully implemented*) are considered ISO 14001 compliant (panel D). Size was measured by computing the natural log of the number of employees (panel E). This measure is preferable to financial measures that complicate the comparison of organizations with differing accounting methods (Chenhall, 2003). Organizations situated above the average size are considered larger, whereas those situated below the average size are considered smaller. Lastly, respondents were asked to indicate whether their organization was privately owned or publicly traded (panel F).

3.3. *Data analysis*

Two types of analysis are used to examine the associations among the four contextual factors and the importance of measurement and use of EPIs. First, a correlation matrix is used to provide preliminary evidence of the relationship between constructs. Then, analyses of variance (ANOVAs) are performed to compare the mean score of EPIs using two groups for each contextual factor: active or passive environmental strategy; ISO compliance or not; large or small business; private or public ownership.

4. Results

4.1. *Descriptive results*

Table 1 presents the importance of measurement of various EPIs and their classifications. The results presented in panel A suggest that overall, organizations devote moderate importance to the various EPIs (4.91/7; s.d. 1.39). Specifically, the indicators on which the managers devote the most importance to are: those that measure conformity with requirements or expectations (5.70), inputs of energy (5.50), community relations (5.20), outputs of solid waste (5.13), and outputs of air emissions (5.10). Indicators that are considered least important are those providing information on the local, regional or national condition of the environment (3.75), measuring the inputs of auxiliary materials (4.60), implementation of environmental policies and programs (4.72), inputs of water (4.76), inputs of raw materials (4.76), and outputs of waste water (4.79). The large standard deviation for the majority of the indicators indicates the dispersion of results in the sample.

According to the ISO 14031 classification, the results suggest that organizations devote most importance to measurement of MPIs (5.13; s.d. 1.68) followed closely by OPIs (4.95; s.d. 1.39). Organizations place little importance on measurement of ECIs (3.75; s.d. 2.31). The correlation matrix (Table 1, panel A) shows that these three categories of indicators are significantly correlated. These results suggest that these categories are unique but also complementary. The results for the financial and non-financial classifications show that organizations place similar importance on the measurement of these two groups of indicators (4.92 s.d. 1.94 vs. 4.91 s.d. 1.39). The correlation matrix reveals a significant correlation between those categories, which also reflects their uniqueness and complementarities.

The results of panel B suggest that overall, organizations tend to make moderate use of EPIs (4.99/7; s.d. 1.68). More specifically, monitoring internal compliance with environmental policies and regulations (5.31) is the most frequent use of environmental indicators in organizations, followed by continuous improvement (5.16) and providing data for internal decision making (5.01). The use of environmental performance measurement systems for external reporting purposes (4.48) is less common in organizations.

4.2. *Comparative results*

Table 2 presents the correlation matrix between EPIs dimensions and organizational characteristics while Table 3 presents the results of ANOVAs.

4.2.1. *Environmental strategy*

First, the correlation matrix indicated that environmental strategy has a strong correlation with the importance of measurement (.675, $p<.01$) and use of EPIs (.668, $p<.01$). Moreover, the results of ANOVAs also suggest that the type of environmental strategy is associated with the importance of measurement and the use of EPIs. Accordingly, panel A suggests that organizations that adopt an active environmental strategy place more importance on the measurement of EPIs than organizations that adopt a passive environmental strategy (5.59 vs. 4.08; $p<.01$). The same finding can also be observed in relation to different components of ISO 14031 classifications and financial/non-financial indicators ($p<.01$). Regarding the use of EPIs (panel B), the results suggest that organizations which adopt an active environmental strategy *globally* use EPIs significantly more than organizations with a passive strategy (5.86 vs. 3.97; $p<.01$). Specifically, using the residuals, the results suggest that two *individual* uses differ significantly. Indeed, organizations reflecting an active environmental strategy appear to use EPIs more intensively to motivate continuous improvement ($p<.05$) and to provide data for decision making ($p<.05$) than organizations reflecting a passive strategy. No difference is observed for the two other uses (i.e. to monitor compliance and to provide data for external reporting).

Those results are in line with two streams of research. First, the management accounting literature provides considerable evidence supporting the relationship between organizational strategy and the measurement and use of performance indicators (e.g. Abernethy and Guthrie, 1994; Gosselin, 2005; Hoque, 2004; Said et al., 2003). Second, the environmental management literature provides insights into the link between environmental strategy and organizational routines, such as the measurement and use of EPIs (e.g. Hunt and Auster, 1990; Doonan et al., 2002;

Table 2
Correlation matrix

	EPIs—importance of measurement	EPIs—use	Environmental strategy	ISO 14001	Size	Ownership
EPIs—importance of measurement	1.0					
EPIs—use	.758**	1.0				
Environmental strategy	.675**	.668**	1.0			
ISO 14001	.331**	.446**	.410**	1.0		
Size	.187**	.190**	.162**	.203**	1.0	
Ownership	.278**	.237**	.195**	.300**	.131*	1.0

ISO 14001 and ownership are dichotomous variables.
*Significant at the .05 level; **Significant at the .01 level.

Table 3
Associations among firms' characteristics and EPIs

(A) Importance of measurement of EPIs (*$p<.05$, **$p<.01$)

	Environmental strategy			ISO 14001 compliant			Size			Ownership		
	Active	Passive	Sig.	Yes	No	Sig.	Larger	Smaller	Sig.	Public	Private	Sig.
Number of firms	161	142		128	175		133	170		119	184	
Environmental condition indicators	4.46	2.87	**	4.24	3.39	**	3.90	3.63	n.s.	4.15	3.49	*
Operational performance indicators	5.56	4.19	**	5.42	4.65	**	5.13	4.80	*	5.37	4.67	**
Management performance indicators	5.94	4.15	**	5.87	4.60	**	5.47	4.87	**	5.76	4.73	**
Financial indicators	5.50	4.21	**	5.43	4.53	**	5.20	4.71	*	5.34	4.65	**
Non-financial indicators	5.60	4.06	**	5.45	4.51	**	5.14	4.73	*	5.40	4.59	**
Global	5.59	4.08	**	5.44	4.52	**	5.14	4.73	*	5.40	4.60	**

(B) Use of EPIs (*$p<.05$, **$p<.01$)

	Environmental strategy			ISO 14001 compliant			Size			Ownership		
	Active	Passive	Sig.	Yes	No	Sig.	Larger	Smaller	Sig.	Public	Private	Sig.
Number of firms	161	142		128	175		133	170		119	184	
Monitor internal compliance with environmental policies and regulations[a]	.07	−.08	n.s.	.15	−.11	*	.06	−.05	n.s.	−.02	.01	n.s.
Motivate continuous improvement[a]	.10	−.11	*	.08	−.06	n.s.	.11	−.09	*	−.02	.02	n.s.
Provide data for internal decision-making[a]	.09	−.10	*	.07	−.05	n.s.	−.02	.02	n.s.	.09	−.06	n.s.
Provide data for external reporting[a]	.01	−.02	n.s.	.001	−.0008	n.s.	−.08	.07	n.s.	.25	−.17	*
Global[b]	5.86	3.97	**	5.80	4.37	**	5.40	4.70	**	5.55	4.66	**

[a]Considering the objective to analyze the specific uses of EPIs, those comparisons are conducted using the 'unique' portion of each use. As previously mentioned, the uniqueness is obtained by using residuals from regressing each use of EPIs on the other three uses.
[b]Considering the objective to analyze the global use of EPIs, those comparisons are conducted using both 'unique' and 'common' portion of each use, i.e. the data provided by the respondents.

Henriques and Sadorsky, 1999). In sum, those streams of research suggest that organizations that adopt an active environmental strategy are more likely to place greater importance on the measurement of EPIs and to use them extensively (i) to favor the alignment of actions toward the attainment of environmental objectives, (ii) to motivate their senior executives to become involved in environmental management, (iii) to promote employee involvement, and (iv) to support the decision-making process.

4.2.2. ISO 14001 compliance

The correlation matrix indicated a correlation between ISO 14001 compliance and both the importance of measurement (.331 $p<.01$) and the use of EPIs (.446 $p<.01$). The results of ANOVAs also suggest that compliance to the ISO 14001 standard is associated with the importance of measurement and use of EPIs. Accordingly, the ISO-compliant firms place greater importance on the measurement of EPIs than do the non-ISO-compliant firms (5.44 vs. 4.52; $p<.01$). The results obtained are similar for the two classifications ($p<.01$). Regarding the use of EPIs, the results suggest that the ISO-compliant firms *globally* use EPIs significantly more than the non-ISO-compliant firms (5.80 vs. 4.37; $p<.01$). Specifically, the results suggest that only one *individual* use differs significantly (to monitor compliance $p<.05$) while no difference is observed for the three other uses.

Thus, firms implementing ISO 14001 standards will likely have the organizational resources and structure to more effectively collect environmental practices data (Marshall and Brown, 2003). The existence of an EMS thus suggests a higher level of top management commitment to environmental issues, and a greater likelihood that the firm will establish goals for environmental impact reduction as well as measuring programmatic investments focused on achieving those goals (Marshall and Brown, 2003). In addition, the ISO-compliant firms are more likely to be aware of EPIs than those that are not. Indeed, they are more likely to be aware of the ISO 14031 guidelines that propose a wide range of EPIs. Furthermore, since the ISO 14001 standard requires that each of the organizational processes has at least one objective and an indicator that measures the attainment of this objective it is probable that ISO-compliant firms place more importance on the measurement of EPIs, which they use more than non-ISO-compliant firms to monitor internal compliance with policies and regulations.

4.2.3. Size

The correlation matrix indicated a correlation between size and both importance of measurement (.187, $p<.01$) and use of EPIs (.190, $p<.01$). The results of ANOVAs also suggest that size is associated with the importance of measurement and the use of EPIs. Thus, larger firms place more importance on the measurement of EPIs than do smaller firms (5.14 vs. 4.73; $p<.05$). Regarding classifications, only the difference in ECI scores is not significant. In the case of use, the results suggest that *globally* larger firms use EPIs more than smaller firms do (5.40 vs. 4.70; $p<.01$). However, this finding does not apply to the four types of use. Specifically, only the use of EPIs to motivate continuous improvement is significantly different between the two groups ($p<.05$).

Those results are in line with common findings in the management accounting literature suggesting that larger organizations are more likely than smaller organizations to develop and use management control systems (Chenhall, 2003). This phenomenon also seems to be observed with respect to environmental management control systems, particularly EPIs. Indeed, large organizations generally have more resources to invest in pollution control, prevention technologies, and the development of environmental performance measurement systems (Marshall and Brown, 2003). Furthermore, larger size is associated with a rise in problems related to social control, communication, and coordination (Merchant, 1981). Organizations thus face an exponential increase in the number of channels that require a flow of information for coordination purposes, which hinders communication (Merchant, 1981). Consequently, as organizations grow, they tend to implement a more administratively oriented control strategy that involves increased structuring of activities, more formalized communication and greater use of standardized information for the evaluation of managerial performance (Bruns and Waterhouse, 1975). These factors encourage large organizations to devote more attention to the measurement of performance indicators and to use them as part of organizational routines. Externally, large organizations are more visible to external stakeholders, such as environmental interest groups, the community and governments, groups which may all exert pressure on these organizations (Ullmann, 1985). Hence, large firms may devote more attention to the measurement of EPIs to help manage environmental issues.

4.2.4. Ownership

The correlation matrix indicated that ownership is correlated with importance of measurement (.278, $p<.01$) and use of EPIs (.237, $p<.01$). Moreover, the results of ANOVAs also suggest that ownership is associated with the importance of measurement and use of EPIs. Consequently, publicly owned organizations place more importance on the measurement of EPIs than privately owned organizations (5.40 vs. 4.60; $p<.01$). This observation also applies to all the classifications. Regarding the use, publicly owned organizations *globally* use EPIs more than private organizations do (5.55 vs. 4.66; $p<.01$). Specifically, the results suggest that one *individual* use differs significantly, namely external reporting ($p<.05$). No difference is observed specifically for the three other uses.

Those results concur with arguments and conclusions provided in past research. For instance, Klassen and McLaughlin (1996) note that strong environmental management results in significant positive stock market performance. This result is corroborated by Deutsch (1998), who observes that eco-efficient organizations reward shareholders with good financial performance. Thus, stronger environmental performance can improve the value of the firm and attract new stockholders (Melnyk et al., 2003). This phenomenon is confirmed by the advent of several mutual funds that select stocks for portfolios based on their environmental performance. Consequently, public organizations that want to implement a green strategy that can help create value for shareholders are likely to control and measure the attainment of their environmental objectives by developing EPIs.

In addition, Doonan et al. (2002) show that two of the three main sources of pressure reported by environmental managers are the government and the general public. Public organizations are increasingly responding to these stakeholders' concerns and transmitting more information through obligatory disclosure of financial statements to shareholders, than are, for example, private organizations. Therefore, in response to pressures from these external stakeholders, public organizations are more likely than private organizations to want to disclose environmental information. Environmental performance measurement systems are thus an excellent means for public organizations to gather environmental information and disclose it to external stakeholders.

5. Conclusion

The objective of this exploratory study was to identify associations among firm characteristics and the importance of measurement of EPIs and their use. The results of this study suggest three main conclusions. First, the importance of measurement of EPIs is associated with (i) firms having a more active environmental strategy, (ii) ISO 14001 compliant firms, (iii) larger firms, and (iv) public firms. Second, the *global* use of EPIs is also associated with a more active environmental strategy, ISO 14001 compliance, larger firms and public firms. Third, the *specific* uses of EPIs are associated with different firm characteristics: (i) to monitor compliance is associated with ISO-compliant firms, (ii) to motivate continuous improvement is associated with an active environmental strategy and larger firms, (iii) decision making is associated with an active environmental strategy, and (iv) external reporting is associated with public firms.

This study contributes to the development of knowledge in the field of EMA. First, it enriches the empirical literature, which is currently underdeveloped in this area (Burritt, 2004), by providing evidence of associations among various firm characteristics and EPIs. In addition, this study provides a better understanding of EPIs by differentiating and specifically examining two fundamental dimensions: the importance of measurement and use. Also, this study differentiates four uses of EPIs and suggests that global use and specific use of EPIs are not associated with the same firm characteristics.

This study has also important business implications. Indeed, managers should be aware that the measurement and use of EPIs: (i) support and communicate the environmental strategy throughout the organization, (ii) support and ensure conformity of environmental processes helping organizations to obtain and maintain the ISO 14001 certification, (iii) formalize complex environmental processes and procedures, (iv) decentralize and support environmental information systems, and (v) contribute to meeting stakeholders expectations.

This study is subject to potential limitations. First, this study encompasses four organizational factors influencing the measurement and use of EPIs in manufacturing firms. Measurement of different firm characteristics could lead to other interesting results. Moreover, the interaction among firm characteristic has not been examined and could lead to different results. Second, the number of items used to measure the diversity of measurement of EPIs may appear small and imprecise. However, considering the scope of the sample, a generic definition of indicators (i.e. output of solid waste, inputs of energy) has been used to allow comparison between different industries. The use of more precise indicators in specific industries could lead to different results. Third, using the survey method to collect data can create a potential bias due to common response and social desirability. Finally, considering differences in the regulations and market settings among different countries, results may not be generalized outside the scope of the current sample (i.e., small-to-medium sized manufacturing firms in Canada).

This study opens many avenues for future research. As previously mentioned, several other firm characteristics such as organizational structure, organization life cycle, presence of a green leader in the organization, pressure from stakeholders, global strategy of the organization, and the structure of the board of directors could be investigated to determine their influence on EPIs. Moreover, the interaction among the different firm characteristics could also be explored. In addition, other dimensions of EPIs such as information quality and EPIs updating process could also be examined to determine how they are influenced by firm characteristic. Furthermore, the link between EPIs and existing database compatibility, information providers, software tools, and information characteristics (e.g. aggregation, timeliness) could also be investigated. Considering the particular economic structure in Canada, specific attention could also be devoted to small businesses to extend the discussion related to the measurement and use of EPIs in this type of organization. Lastly, in addition to considering firm characteristics, future research could assess the influence of EPIs on environmental actions put forth by managers and employees, along with the environmental and economic performance of organizations.

Appendix A

Description of the sample (Table A1).

Table A1
Description of the sample

Industry SIC code	#
20 Food and kindred products	26
21 Tobacco manufacturers	1
22 Textile mill products	5
23 Apparel and other textile products	5
24 Lumber and wood products	42
25 Furniture and fixture	15
26 Paper and allied products	27
27 Printing and publishing	4
28 Chemicals and allied products	17
29 Petroleum and coal products	6
30 Rubber and misc. plastics products	19
31 Leather and leather products	2
32 Stone, clay, glass, and concrete products	9
33 Primary metal industries	15
34 Fabricated metal products	25
35 Industrial machinery and equipment	31
36 Electrical and electronic equipment	20
37 Transportation equipment	26
38 Instrument and related products	4
39 Misc. manufacturing industries	4
Total	303
Size Number of employees	#
<100	7
Between 100 and 499	192
Between 500 and 999	65
Between 1000 and 4999	35
>5000	4
Total	303
Average	710

Appendix B

Analysis of the uses of EPIs (Table B1).

Table B1
Analysis of the uses of EPIs

(A) Correlation matrix (Pearson) (significant at the .05 level, **significant at the .01 level.)

	(i)	(ii)	(iii)	(iv)
(i) Monitor compliance	1.0			
(ii) Continuous improvement	.811**	1.0		
(iii) Decision-making	.761**	.830**	1.0	
(iv) External reporting	.636**	.671**	.658**	1.0

Preliminary conclusion: All four uses are significantly correlated ($p<.01$).

(B) Factor analysis

(i) Principal components

Components	Eigenvalue	Variance (%)	Cumulative (%)
1	3.189	79.73	79.73
2	.413	10.32	90.05
3	.240	6.0	96.05
4	.158	3.95	100.0

Preliminary conclusion: Data suggest a one-component factor structure.

(ii) One-factor structure

EPIs use	Factor loading	Uniqueness
(i) Monitor compliance	.901	.188
(ii) Continuous improvement	.932	.132
(iii) Decision-making	.913	.166
(iv) External reporting	.822	.325

Conclusion: Although the four use load on one factor, each use has some "uniqueness", i.e. a portion of the variance unexplained by the common factor.

References

Abernethy, M.A., Guthrie, C.H., 1994. An empirical assessment of the "fit" between strategy and management information system design. Accounting and Finance November, 49–66.

Aragon-Correa, J.A., 1998. Strategic proactivity and firm approach to the natural environment. Academy of Management Journal 41 (5), 556.

Atkinson, A.A., Waterhouse, J.H., Wells, R.P., 1997. A stakeholder approach to strategic performance measurement. Sloan Management Review Spring, 25–37.

Baines, A., Langfield-Smith, K., 2003. Antecedents to management accounting change: a structural équation approach. Accounting, Organizations and Society 28, 675–698.

Bennett, M., James, P., 1998. Environment Under the Spotlight—Current Practice and Future Trends in Environment—Related Performance Measurement for Business. Association of Chartered Certified Accountants, London.

Bisbe, J., Otley, D., 2004. The effects of the interactive use of management control systems on product innovation. Accounting, Organizations and Society 29 (8), 709–737.

Bouwens, J., Abernethy, M.A., 2000. The consequences of customization on management accounting system design. Accounting, Organizations and Society 25, 221–241.

Briassoulis, H., 2001. Sustainable development and its indicators: through a (planner's) glass darkly. Journal of Environmental Planning and Management 44 (3), 409.

Bruns, W.J., Waterhouse, J.H., 1975. Budget control and organization structure. Journal of Accounting Research Autumn 1985, 177–203.

Burritt, R., 2004. Environmental management accounting: roadblocks on the way to green and pleasant land. Business Strategy and the Environment 13, 13–32.

Chenhall, R.H., 2003. Management control systems design within its organizational context: Findings from contingency-based research and directions for the future. Accounting, Organizations and Society 28 (2,3), 127.

Chenhall, R.H., 2005. Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study. Accounting, Organizations and Society 30 (5), 395–422.

Chinander, K.R., 2001. Aligning accountability and awareness for environmental performance in operations. Production and Operations Management 10 (3), 276.

Clarkson, M.B.E., 1995. A stakeholder framework for analyzing and evaluating corporate social performance. Academy of Management Review 20 (1), 92.

Deutsch, C.H., 1998. For Wall Street, Increasing Evidence that Green Begets Green. New York Times, New York, NY, p. 7.

Doonan, J., Lanoie, P., Laplante, B., 2002. Environment Performance of Canadian Pulp and Paper Plants: Why Some Do Well and Others Do Not? Montréal, Scientifics Series.

Eckel, L., Fisher, K., Russell, G., 1992. Environmental performance measurement. CMA Management 66 (2), 16.

Epstein, M., 1994. Environmentally responsible corporations. Management Accounting April, 74.

Epstein, M., Manzoni, J.-F., 1997. The balanced scorecard and tableau de bord: translating strategy into action. Management Accounting August, 28–36.

Figge, F., Hahn, T., Schaltegger, S., Wagner, M., 2002. The sustainability balanced scorecard – linking sustainability management to business strategy. Business Strategy and the Environment 11 (5), 269.

Fisher, J., 1992. Use of non-financial performance measures. Journal of cost management Spring, 31–38.

Gosselin, M., 2005. An empirical study of performance measurement in manufacturing firms. The International Journal of Productivity and Performance Management 54 (5/6), 419.

Hansen, S.C., Van der Stede, W.A., 2004. Multiple facets of budgeting: an exploratory analysis. Management Accounting Research 15 (4), 415–439.

Hart, S.L., 1995. Natural-resource-based view of the firm. Academy of Management Review 20 (4), 986–1014.

Henri, J.-F., 2006a. Management control systems and strategy: a resource-based perspective. Accounting, Organizations and Society 31 (6), 529–558.

Henri, J.-F., 2006b. Organizational culture and performance measurement systems. Accounting, Organizations and Society 31 (1), 77–103.

Henriques, I., Sadorsky, P., 1999. The relationship between environmental commitment and managerial perceptions of stakeholder importance. Academy of Management Journal 42 (1), 87.

Hoque, Z., 2004. A contingency model of the association between strategy, environmental uncertainty and performance measurement: impact on organizational performance. International Business Review 13, 485–502.

Hoque, Z., James, W., 2000. Linking balanced scorecard measures to size and market factors: impact on organizational performance. Journal of Management Accounting Research 12, 1.

Hunt, C.B., Auster, E.R., 1990. Proactive environmental management: avoiding the toxic trap. Sloan Management Review 31 (2), 7–18.

Ijiri, Y., 1975. Theory of accounting measurement. Studies in Accounting Research 10.

Ilinitch, A.Y., Soderstrom, N.S., Thomas, T.E., 1998. Measuring corporate environmental performance. Journal of Accounting and Public Policy 17 (4,5), 383.

Ittner, C.D., Larcker, D.F., 1998. Innovations in performance measurement: trends and research implications. Journal of Management Accounting Research 10, 205–238.

Ittner, C.D., Larcker, D.F., 2003. Coming up short on nonfinancial performance measurement. Harvard Business Review November, 88–95.

Ittner, C.D., Larcker, D.F., Randall, T., 2003. Performance implications of strategic performance measurement in financial services firms. Accounting, Organizations and Society 28 (7–8), 715–741.

Kaplan, R., Norton, D., 1992. The balanced scorecard—measures that drive performance. Harvard Business Review Jan.–Feb., 71–79.

Kaplan, R., Norton, D., 1996. The Balanced Scorecard: Translating Strategy into Action. Harvard Business School Press, Boston, MA.

Kaplan, R.S., Norton, D.P., 2001. Transforming the balanced scorecard from performance measurement to strategic management: part I. Accounting Horizons 15 (1), 87–104.

Kerr, S., 1995. On the folly of rewarding A, while hoping for B. The Academy of Management Executive 9 (1), 7.

Klassen, R.D., McLaughlin, C.P., 1996. The impact of environmental management on firm performance. Management Science 42 (8), 1199.

Lober, D., 1996. Evaluating the environmental performance of corporations. The Journal of Managerial Issues 8 (2), 184–205.

Marshall, R.S., Brown, D., 2003. Corporate environmental reporting: what's in a metric? Business Strategy and the Environment 12 (2), 87.

Melnyk, S.A., Sroufe, R.P., Calantone, R., 2003. Assessing the impact of environmental management systems on corporate and environmental performance. Journal of Operations Management 21 (3), 329.

Merchant, K.A., 1981. The design of the corporate budgeting system: influences on managerial behavior and performance. The Accounting Review 56 (4), 813.

Miller, D., 1987. The structural environmental correlates of business strategy. Strategic Management Journal 8 (1), 55–76.

Neely, A., Gregory, M., Platts, K., 1995. Performance measurement system design: a literature review and research agenda. International Journal of Operations and Production Management 15 (4), 80–116.

Olsthoorn, X., Tyteca, D., et al., 2001. Environmental indicators for business: a review of the literature and standardisation methods. Journal of Cleaner Production 9, 453–463.

Pringle, J., Lueteritz, K., Fitgerald, M., 1998. ISO 14001: a discussion of implications for pollution prevention. National Pollution Prevention Roundtable, ISO 14000 Workgroup.

Roome, N., 1994. Business strategy, R&D management and environmental imperatives. R&D Management 24 (1), 65–82.

Said, A.A., Elnaby, H., Wier, B., 2003. An empirical investigation of the performance consequences of nonfinancial measures. Journal of Management Accounting Research 15, 193–223.

Schaltegger, S., Burritt, R., 2000. Contempory Environmental Accounting: Issues, Concepts and Practice. Greenleaf Publishing Limited, Sheffield, UK.

Scott, T.W., Tiesen, P., 1999. Performance measurement and managerial teams. Accounting, Organizations and Society 24, 263–285.

Sharma, S., Vredenburg, H., 1998. Proactive corporate environmental strategy and the development of competitively valuable organizational capabilities. Strategic Management Journal 19 (8), 729–753.

Shields, J.F., Shields, M.D., 1998. Antecedents of participative budgeting. Accounting, Organizations and Society 23 (1), 49–76.

Simons, R., 2000. Performance Measurement and Control Systems for Implementing Strategy. Prentice Hall, Upper Saddle River, NJ.

Ullmann, A.A., 1985. Data in search of a theory: a critical examination of the relationships among social performance, social disclosure, and economic performance of US firms. Academy of Management Review 10 (3), 540–557.

Veleva, V., Ellenbecker, M., 2000. A proposal for measuring business sustainability: addressing shortcomings in existing frameworks. Greener Management International 31, 101–119.

ELSEVIER

Journal of Environmental Management 87 (2008) 177–192

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Relationships between anaerobic consortia and removal efficiencies in an UASB reactor degrading 2,4 dichlorophenol (DCP)

Delia Teresa Sponza*, Canan Cigal

Faculty of Engineering, Department of Environmental Engineering, Dokuz Eylül University, Buca Kaynaklar Campus, İzmir, Turkey

Received 10 November 2005; received in revised form 19 October 2006; accepted 9 January 2007
Available online 26 March 2007

Abstract

To gain more insight into the interactions between anaerobic bacteria and reactor performances (chemical oxygen demand—COD, 2,4 dichlorophenol—2,4 DCP removals, volatile fatty acid—VFA, and methane gas productions) and how they depended on operational conditions the microbial variations in the anaerobic granular sludge from an upflow anaerobic sludge blanket (UASB) reactor treating 2,4 DCP was studied. The study was composed of two parts. In the first part, the numbers of methanogens and acedogens in the anaerobic granular sludge were counted at different COD removal efficiencies. The relationships between the numbers of methanogens, the methane gas production and VFA production were investigated. The COD removal efficiencies increased to 74% from 30% while the number of total acedogens decreased to 10 from 30 cfu ml^{-1}. The number of total methanogens and acedogens varied between 11×10^3 and 10×10^9 MPN g^{-1} and 10 and 30 cfu ml^{-1} as the 2,4 DCP removal efficiencies were obtained between 60% and 99%, respectively. It was seen that, as the number of total acedogens decreased, the COD removal efficiencies increased. However, the number of total methanogens increased as the COD removal efficiencies increased. Correlations between the bacterial number and with the removal efficiencies obtained in different operational conditions were investigated. From the results presented in this paper a high correlation between the number of bacteria, COD removals, methane gas percentage, 2,4 DCP removals and VFA was observed. In the second part, methanogen bacteria in the anaerobic granular sludge were identified. Microbial observations and biochemical tests were applied to identify the anaerobic microorganisms from the anaerobic granular sludge. In the reactor treating 2,4 DCP, *Methanobacterium bryantii*, *Methanobacterium formicicum*, *Methanobrevibacter smithii*, *Methanococcus voltae*, *Methanosarcina mazei*, *Methanosarcina acetivorans*, *Methanogenium bourgense* and *Methanospirillum hungatei* were identified.

Keywords: Upflow anaerobic sludge blanket reactor (UASB); Methanogens; Acedogens; COD; VFA; Methane gas; 2, 4 DCP

1. Introduction

The upflow anaerobic sludge blanket reactor (UASB) reactor is an advanced anaerobic treatment in biological treatment of wastewaters that is characterized by an anaerobic granular sludge with a notably high metabolic activity and good biosolids settleability (Dolfing, 1987; Grotenhuis, 1992).

The anaerobic treatment converts the organic pollutants (COD, BOD) in wastewater into a small amount of sludge and a large amount of biogas. In anaerobic conditions, many different groups of anaerobic bacteria work together to degrade the complex organic pollutants to methane and carbondioxide (Speece, 1996). The biological degradation of complex organic compounds takes place in several consecutive biochemical steps, each performed by different groups of specialized bacteria. The acedogenic and methanogenic phases are most often the rate-limiting steps (Grotenhuis, 1992).

The principal genera of microorganisms that have been identified in anaerobic sludges include: the rods (*Methanobacterium*, *Methanobacillus*) and spheres (*Methanococcus*, *Methanosarcina*). They have very slow growth rates;

*Corresponding author. Tel.: +90 232 453 1153; fax: +90 232 453 1143.
E-mail address: delya.sponza@deu.edu.tr (D.T. Sponza).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.01.034

therefore, their metabolism is usually considered as rate limiting in the anaerobic treatment. The non-methanogenic and methanogenic bacteria must be in a state of dynamic equilibrium to maintain an anaerobic treatment system which stabilizes the organic wastes efficiently (Doddema and Vogels, 1978).

Hulshoff and Pol (1989) showed that the majority of reactors treating wastes with high volatile fatty acid (VFA) levels contained two genera of acetoclastic methanogens, namely *Methanosarcina* and *Methanothrix*. The bacterial population in the biomass is important during the reactor start up since it significantly affects the acidification and the methanogenesis (Mink and Patrick, 1977). Dolfing (1987) studied the microbiological aspects of granular methanogenic sludge treating sugar refinery wastewater. It was shown that the biomass consisted of a significant part of acetoclastic methanogens. The microbiological composition consisted of the acetate utilizing methanogen *Methanothrix soehngenii*, which is the predominant organism in UASB reactor depending on the composition of the wastewater. The VFA concentration is also of importance with respect to the ultimate bacterial composition of the granular sludge. When using a concentrated VFA solution as feed a granular sludge was developed consisting mainly of *Methanosarcina* rather than *Methanothrix* (Mink and Patrick, 1977). With lower substrate concentrations, *Methanosarcina* may develop under conditions of continuous overloading. The granules developed on VFA substrates with low VFA concentrations consisting mainly of filamenteous *Methanothrix*. Anderson et al. (1994) studied the changes in the microbial population of a two-stage anaerobic digestion system. The numbers of *Methanosarcina* and filamentous bacteria species decreased, whereas in the pre-acidification stage, only a few *Methanosarcina* species were identified. Zhu et al. (1997) studied the bacterial numeration in a laboratory scale single- and two-phase UASB reactors. They isolated the fermentative bacteria, H_2-producing acedogenic bacteria and methanogenic bacteria. Plumb et al. (2001) observed that there was a methanogenic population dominated by *Methanosaeta*, *Methanobacterium* and *Methanosprillum* in textile industry wastewater. Morgan et al. (1991) studied the methanogenic population dynamics during the start up of a full-scale anaerobic sequencing batch reactor treating swine waste. The reactor performance improved at a volumetric sludge-loading rate (VSLR) of $1.7\,g\,VSS\,l^{-1}\,day^{-1}$. The number of *Methanosarcina* genus decreased, while the levels of *Methanosaeta* from the acetate utilizing methanogens remained low. Sponza (2002) studied the enhancement of granule formation and sludge retainment for tetrachloroethylene (TCE) removal in an UASB. Microbiological examinations indicated that the granules were mainly comprised of *Methanothrix*- and *Methanosarcina*-like bacteria. They were the main methanogens on the inner surface of the granules, plus a small amount of *Methanobrevibacterium* sp. and *Methanococcus* sp. Smith and Mah (1978) studied the characterization of microbial consortia in a terephthalate degrading anaerobic granular sludge system. The microbial composition and spatial distribution in a terephthalate degrading anaerobic granular sludge contained acetoclastic *Methanosaeta* and the hydrogenotrophic *Methanospirillum* and some members of Methanobacteriacea.

Chlorophenols are widespread toxic compounds that cause serious environmental problems all over the world (Zinder and Mah, 1979). They are used in several industries such as production of pesticides, in preservative chemicals for wood, glue, paint, leather, and pulp materials (Montenegro et al., 2001; Schauer et al., 1982; Schmidt and Ahring, 1999). Additionally, the chlorophenolic compounds are often found in discharges of many industries including petrochemical, oil refinery. In 1988, EPA was notified of the death of a worker acutely exposed to 2,4 dichlorophenol (2,4 DCP) and three more cases concerning toxicity of 2,4 DCP (EPA US Environmental Protection Agency (EPA), 2003).

Little is known about the relations between reactor performances and anaerobic bacteria in UASB reactors treating different toxic substances. No study was found investigating the methanogenic consortia in UASB reactors treating 2,4 DCP at different operational conditions. It is important to note that the IC_{50} value (The 2,4 DCP concentration reduce the 50% of gas production compared to control) of 2,4 DCP was found as $25\,mg\,l^{-1}$ (Speece, 1996). Therefore, in this study $10\,mg\,l^{-1}$ 2,4 DCP was added to synthetic wastewater. As the OLR increased the 2,4 DCP concentration will be increased and toxicity will be reduced the removal efficiencies with COD loading rate. Therefore $10\,mg\,l^{-1}$ 2,4 DCP cause toxicity and reduce the removal efficiencies together with high COD loading rates. Since the COD, 2,4 DCP removal efficiencies and methane percentage decreased as the COD and 2,4 DCP loading rates increased, it was aimed to observe the effects of removal efficiencies on the number of acedogen and methanogen bacteria. Furthermore, the relationships between the reactor performance and microorganism numbers were evaluated. Ever run was operated under steady-state conditions. However, the reactor performance (COD, 2,4 DCP removals, VFA production and methane percentage) varied at different organic loading rate (OLR) and HRTs (hydraulic retention time). Therefore, the specific objectives of this research were:

1. to obtain the bacterial number of methanogens and acedogens in different COD and 2,4 DCP removal efficiencies depending to operational conditions and their relationships with COD, 2,4 DCP removal efficiencies, VFA concentration and methane gas productions,
2. to identify the methanogenic species in the anaerobic reactor treating 2,4 DCP and to determine the variations of methanogenic consortia.

2. Materials and methods

2.1. Used microorganisms

The sludge samples taken from the UASB reactors treating 2,4 DCP under different operational conditions were analyzed for the number of methanogenic and acedogenic bacteria. Furthermore, the dominant methanogenic bacteria were identified.

2.2. Properties of UASB reactor treating 2,4 DCP

Stainless-steel UASB reactors which had an internal diameter of 90 mm and a height of 1000 mm with a volume of approximately 2.2 l were used in this study. Five evenly distributed sampling ports were installed over the top of the reactors (see Fig. 1). The studies were conducted at 35±2 °C by means of a temperature-controlled heater in the reactor, since the anaerobic removal efficiencies (COD and VFA levels and decreased methane gas productions) were reduced at low temperatures in UASB reactor.

2.3. Operating conditions and reactor performances

The 2,4 DCP treating UASB reactor was operated through 125 days. Glucose was used as a carbon source in this reactor and it was kept constant in influent between 3000 and 3200 mg l^{-1} and Vanderbilt mineral medium was used as a source of trace metals (Speece, 1996). The 2,4 DCP concentrations were kept constant at 10 mg l^{-1} while the flow rates increased to 29 from 6 l day^{-1}. Therefore, the organic loading rates increased to 34.99 from 7.24 g COD l^{-1} day^{-1}. Similarly, the 2,4 DCP loading rates increased to 0.116 from 0.024 g l^{-1} day^{-1}. The operating conditions and the performances of the UASB reactor are summarized in Table 1.

2.4. Seed

Partially granulated anaerobic sludge was used as seed in UASB reactor and was taken from the methanogenic and acedogenic reactors of Pakmaya Yeast Baker Factory in Izmir. These sludges were mixed in a ratio of 1:1 in UASB reactor. The number of methanogens and acedogens varied according to organic loadings applied to the UASB reactor. Anaerobic sludge was acclimated during the start-up period through the continuous operation of the UASB reactor.

2.5. Operation conditions and composition of synthetic wastewater

One hundred and ten days following the start-up period, when approximately 75% COD removal and 60% methane production efficiencies were obtained at a flow rate of 6 l

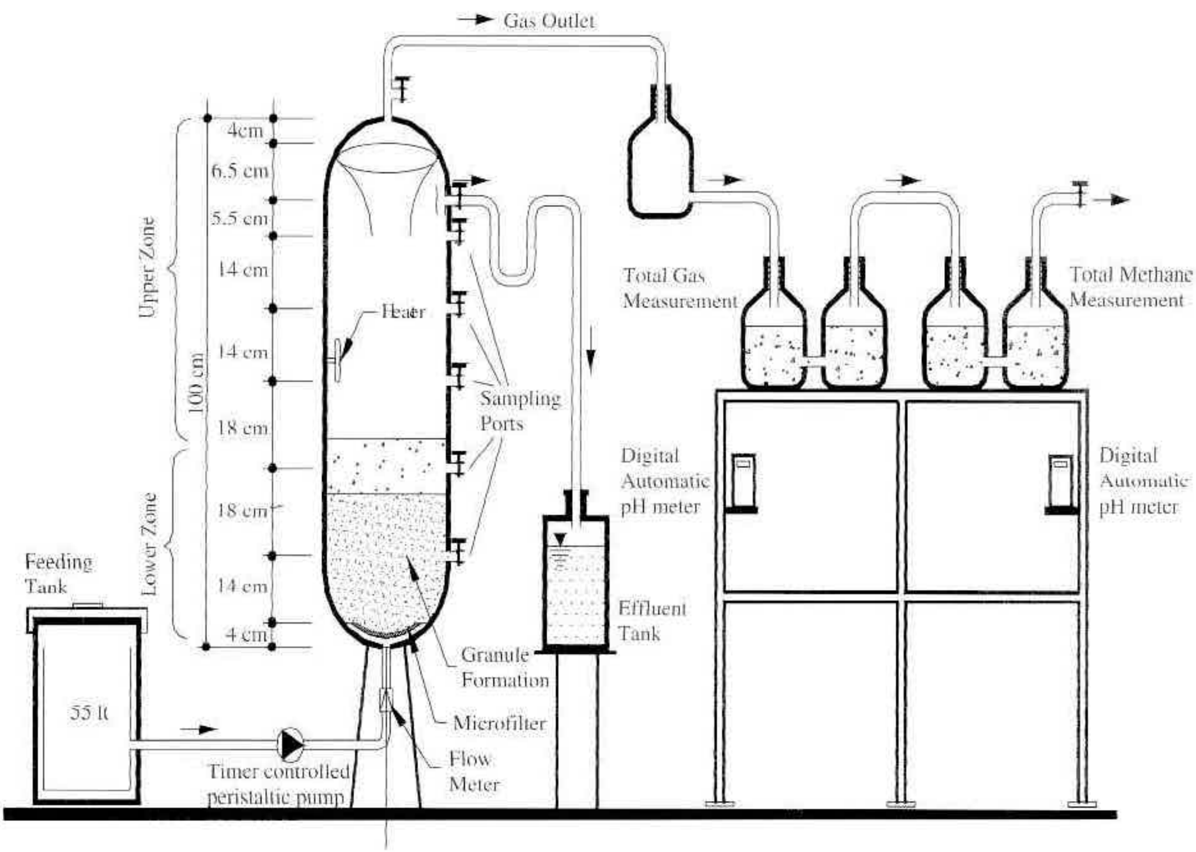


Fig. 1. Schematic configuration of UASB reactor.

Table 1
Operational conditions and reactor performances for 2,4 DCP treating UASB reactor 1 (mean values, $n = 3$)

Parameter	Sampling dates				
	9 September 2002	30 September 2002	25 October 2002	19 November 2002	13 December 2002
COD removal efficiency (%)	74	65	54	47	30
VFA concentrations (mg $CH_3COOH\,l^{-1}$)	331.43	508.15	713.01	898.04	1284.74
2-4 DCP concentrations (mg 2,4 $DCP\,l^{-1}$)	10	10	10	10	10
2,4 DCP loading rate (g 2,4 $DCP\,l^{-1}.day^{-1}$)	0.024	0.06	0.062	0.076	0.116
Q (l day^{-1})	6	15	15	19	29
HRT (h)	10	4	3.87	3.16	2.07
Organic loading rate (g $COD\,l^{-1}$ day^{-1})	7.24	18.1	18.7	22.92	34.99
F/M ratio (mg COD mg^{-1} day^{-1})	0.515	0.715	2.293	3.375	6.734
Methane gas (l day^{-1})	4.488	5.874	8.118	8.710	9.042
Methane gas percentage (%)	98.88	92.56	91.89	86.27	84.24
2,4 DCP removal efficiency (%)	99.72	76.44	73.44	70.00	60.60
Mean acedogen number (cfu ml^{-1})	30	25	20	15	10
Mean methanogen number (MPN g^{-1})	10×10^{9}	8×10^{8}	8×10^{7}	2.5×10^{6}	11×10^{3}
pH	7.7	7.4	7.0	6.5	6.0

day^{-1}; the UASB reactor was started to operate with $10\,mg\,l^{-1}$ of 2,4 DCP concentration through 21 days. The upflow rates were then increased from 15 to $29\,l\,day^{-1}$. The UASB reactor was operated through 20–25 days in every upflow rate under steady-state conditions. The steady-state conditions were defined as constant COD removal and methane percentage in 5 consecutive days through the operation. In every upflow rate run the reactor is under steady-state conditions. However, the COD, 2,4 DCP removals, VFA production and methane percentage varied at different upflow rates. The number of methanogen and acedogen also varied depending to operational conditions. The operation conditions and reactor performances are illustrated in Table 1.

2.6. Sampling period

Biomass samples were taken from the sludge of the UASB reactor at different COD removal efficiencies; VFA concentrations, methane gas production and methane gas percentages were obtained under different operational conditions. Organic loading rates representing the different removal efficiencies in UASB reactors are given in Table 1. The chemical and the microbiological analysis results are the mean of triplicate samplings. Three samples were taken from the sludge and effluent samples during each run. One milliliter of biomass was taken from the UASB reactor and diluted, if necessary, during microbiological analysis. The samples were analyzed immediately in a half hour.

2.7. Analytical procedures

COD in influent and effluent samples were measured using closed Reflux colorimetric method (APHA–AWWA, 1992). Cumulative and methane gases were measured by liquid displacement method (Beydilli et al., 1998; Razo-Flores et al., 1997). VFA were measured by a two-stage titration method developed by Anderson and Yang (1992). The temperature was measured automatically with a digital heater held in the UASB reactor. 2,4 DCP was analysed with a spectrophotometer at 500 nm wavelength in the samples centrifuged at 14 000 rpm for 25 min (APHA–AWWA, 1992).

2.8. Sample preparation

After sampling, the mixed liquor samples were immediately transferred to vials closed with a rubbed stopper, transported to the Microbiology laboratory and stored at 4 °C in a refrigerator.

2.9. Bacterial enumeration studies

2.9.1. Counting of viable methanogens

The viable methanogenic bacteria were counted by the Most Probable Number (MPN) method as described by Sekiguchi et al. (1999), Zehnder and Brock (1979) and Balch et al. (1979) using the media proposed for methanogens. A series of 10-fold dilutions were made using Hungate tubes. The prepared dilutions from 10^{-1} up to 10^{-10} were inoculated into triplicate Hungate tubes containing butyl rubber septa. The inoculated tubes were incubated at 37 °C for 4–6 weeks. The positive tubes were determined by measuring methane gas and the presence of growth (DSMZ, 2003).

2.9.2. Counting of viable acedogens

The number of acid-forming bacteria in the anaerobic granular sludge was enumerated using anaerobic deep agar technique. DSMZ culture media were prepared for the growth of these bacteria (Bryant, 1972). Standard colony count method was used to count the number of microorganisms in the solid culture media. When bacteria are placed on a solid medium, ideally each colony is founded by only a single bacterial cell. A single cell or a clump of cells growing into an isolated colony is termed a colony-forming unit (cfu) in milliliter sample.

2.10. Enrichment and isolation procedure for identification of methanogenic bacteria

Enrichment cultures of methanogens were carried out in glass vessels under methanogenic growth conditions through 3 months of incubation period at 37 °C using 11 different methanogenic-specific culture media. Cultures with active methane formation were identified using Zeiss light microscope. All enrichment and isolation procedures were carried out under strictly anaerobic conditions (Williams et al., 1989).

The composition of the media are given in Section 2.7 and 2.10 for the identification of the genus given below (Morgan et al., 1990):

Methanosarcina sp., *Methanosphera* sp., *Methanococcus* sp., *Methanobacterium* sp., *Methanospirillum* sp., *Methanogenium* sp., *Methanothrix* sp., *Methanobrevibacter* sp., *Methanococcoides* sp. and *Methanolobus* sp.

2.10.1. Culture media for Methanosarcina sp.

The weight of the ingredients of the media is given as gram per liter in brackets: K_2HPO_4 (0.348), KH_2PO_4 (0.227), NH_4Cl (0.500), $MgSO_4 \cdot 7H_2O$ (0.500), $CaCl_2 \cdot 2H_2O$ (0.250), NaCl (2.250), $FeSO_4 \cdot 7H_2O$ (0.002), yeast extract (2), casitone (2), resazurine (0.001), $NaHCO_3$ (0.850), cysteine $HCl \cdot H_2O$ (10.30), $Na_2S \cdot 9H_20$ (0.300), vitamin solution 10 ml, trace element solution 1 ml, methanol 10 ml and distilled water 1000 ml.

2.10.2. Culture media for Methanosphera sp.

The weight of the ingredients of the media is given as gram per liter in brackets: $Fe(NH_4)_2(SO_4)_2 \cdot 7H_2O$ (0.002), $NiCl_2 \cdot 6H_2O$ (0.0002), Na-acetate $\cdot 3H_2O$ (3), yeast extract (1), trypticase (1), resazurin (0.001), $NaHCO_3$ (3), cysteine–$HCl \cdot H_2O$ (0.30) $Na_2S \cdot 9H_2O$ (0.30), distilled water 900 ml, methanol 5 ml, mineral solution (containing K_2HPO_4, $MgSO_4 \cdot 7H_2O$, $CaCl_2 \cdot 2H_2O$) 80 ml, $Na_2SeO_3 \cdot 5H_2O$ 0.50 ml, trace elements 10 ml, vitamin solution 10 ml.

2.10.3. Culture media for Methanogenium sp.

The weight of the ingredients of the media is given as gram per liter in brackets: KCl (0.335), $MgCl_2 \cdot 6H_2O$ (4), $MgSO_4 \cdot 7H_20$ (3.450), NH_4Cl (0.250), $CaCl_2 \cdot 2H_2O$ (0.140) g K_2HPO_4 (0.140), NaCl (18), $Fe(NH_4)_2(SO_4)_2 \cdot 7H_2O$ (0.002), $NaHCO_3$ (5), Na-acetate (1) yeast extract (2), trypticase (2), resazurin (0.001), cysteine–$HCl \cdot H_2O$ (0.5), $Na_2S \cdot 9H_2O$ (500), distilled water (1000 ml), trace element solution containing $FeSO_4 \cdot 7H_2O$, $CoSO_4 \cdot 7H_2O$, $CaCl_2 \cdot 2H_2O$, $ZnSO_4 \cdot 7H_20$ and $CuSO_4 \cdot 5H_2O$ (10 ml) and vitamin solution containing folic acid and biotin (10 ml).

2.10.4. Culture media for Methanococcus sp.

The media used for *Methanogenium* sp. was used without yeast extract, trypticase or sodium bicarbonate.

2.10.5. Culture media for Methanobacterium sp.

The weight of the ingredients of the media is given as gram per liter in brackets: KH_2PO_4 (0.5), $MgSO_4 \cdot 7H_2O$ (0.4), NaCl (0.4), NH_4Cl (0.4), $CaCl_2 \cdot 2H_2O$ (0.05), $FeSO_4 \cdot 7H_2O$ (0.002), yeast extract. (1), Na-acetate (1.0), Na-formate (2.0), $NaHCO_3$ (4), resazurin (0.001), cysteine–$HCl \cdot H_2O$ (0.5), $Na_2S \cdot 9H_2O$ (0.5), distilled water (940 ml), trace element solution and 1 ml fatty acid mixture containing a mixture of valeric, isovaleric acid and isobutyric alpha-methylbutyric acid solution (20 ml).

2.10.6. Culture media for Methanospirillum sp.

The medium used for *Methanobacterium* was used and 7.5 g l^{-1} isoproponal was added to this culture media. The medium was reduced by adding 0.4 g l^{-1} of sodium sulfide (0.4 g l^{-1})

2.10.7. Culture media for Methanothrix sp.

The weight of the ingredients of the media is given as gram per liter in brackets: KH_2PO_4 (0.3), NaCl (0.6), $MgCl_2 \cdot 6H_2O$ (0.1), $CaCl_2 \cdot 2H_2O$ (0.08), sodium acetate (6.8), resazurine (0.001), $KHCO_3$ (6.8), cysteine $HCl \cdot H_2O$ (0.3), $Na_2S \cdot 9H_2O$ (0.3), NH_4Cl (1), trace element solution containing $FeCl_3 \cdot 6H_2O$ (1.35), $MnCl_2 \cdot 4H_2O$ (.0.1), $CoCl_2 \cdot 6H_2O$ (0.024), $CaCl_2 \cdot 2H_2O$ (0.1), $ZnCl_2$ (0.025) and $CuCl_2 \cdot 2H_2O$ (0.01) (10 ml) vitamin solution containing vitamin B12 (0.1), *p*-amibobenzoic acid (0.08) and, D(+)-biotin 0.02 and nicotinic acid (0.2) (10 ml), distilled water (1000 ml).

2.10.8. Culture media for Methanobrevibacter sp.

The weight of the ingredients of the media is given as gram per liter in brackets: NaCl (1), KCl (0.50), $MgCl_2 \cdot 6H_2O$ (0.4), $CaCl_2 \cdot 2H_2O$ (0.10), NH_4Cl (0.30), KH_2PO_4 (0.20), Na_2SO_4 (0.15), casoamino acids (0.50), yeast extract (0.50), resazurine (0.005), $NaHCO_3$ (5.80), trace element solution SL-10 (1 ml), selenite–tungstate solution (1 ml), vitamin solution (1 ml) and distilled water (1000 ml).

2.10.9. Culture media for Methanococcoides sp.

Methanogenium culture media was used. Three grams per liter of trimethylamine hydrochloride was added to the aforementioned media.

2.10.10. Culture media for Methanolobus sp.

Methanogenium culture media was used.

2.11. Identification of anaerobic bacteria species

Bacterial morphology was studied in a Zeiss light microscope. The dimensions of methanogenic bacteria were determined as bacterial cell width and length with micrometric slide and ocular. Gram and capsule stainings were carried out for identification of methanogenic species.

2.12. Biochemical tests

In order to determine the biochemical characteristics different biochemical tests was applied to the growth colonies of methanogens (Morgan et al., 1990).

2.13. Statistical analysis

The regression and multiple regression analysis between y (dependent) and x (independent) variables and ANOVA test were carried out using Windows Excel (1998).

3. Results and discussion

3.1. COD, 2,4 DCP removal efficiencies and methane percentages in UASB reactors

It was observed that the COD and 2,4 DCP removal efficiencies decreased when the organic loading rates (OLR) were increased in UASB reactor. The COD and DCP removal efficiencies varied between 30% and 74% while the OLRs increased from 7.24 to 34.99 g COD l^{-1} day^{-1} (see Fig. 2). The 2,4 DCP removal efficiency decreased from 99% to 60% while the OLRs were increased to 34.99 from 7.24 g COD l^{-1} day^{-1}. The methane gas percentage increased from 84.24% to 98.88% when the organic loading rate rates decreased from 34.99 to 7.24 g COD l^{-1} day^{-1} (see Table 1). When the organic loading rates were increased from 7.24 to 34.99 gCOD l^{-1} day^{-1}, the methane gas productions increased from about 4.48 to 9.04 l day^{-1}, respectively (see Fig. 2). The high organic loading rates increased the methane gas productions while decreasing the percentage of methane gas.

3.2. Total number of methanogens at varying methane gas percentages and methane productions

As the 2,4 DCP loading rate increased to 0.116 from 0.024 g l^{-1} day^{-1} the methane gas percentages decreased to about 83% from 98%. The corresponding COD loading rates were 34.99 and 7.24 g l^{-1} day^{-1}. The total number of methanogens at varying methane gas percentages is depicted in Fig. 3. Methane gas percentages were measured at between 84.24% and 98.88% while the number of total methanogen bacteria was found to be between 11×10^3 and 10×10^9 MPN g^{-1}. Increases in the number of total methanogens were observed as the methane gas productions increased. The highest number of total methanogens was determined when the methane gas percentage was measured as 98.88%. The organic loading rate and the environmental conditions of the anaerobic reactor affected the number of methanogen bacteria. A strong exponential reaction kinetic between the number of methanogens and methane gas percentages was obtained with a high regression coefficient ($y = 3E+08x-5E+08$, $R^2 = 0.87$). The number of methanogens increased from 11×10^3 to 10×10^9 as the methane gas production increased from 4.48 to 9.04 l day^{-1}. A strong exponential reaction kinetic between the number of methanogens and methane gas

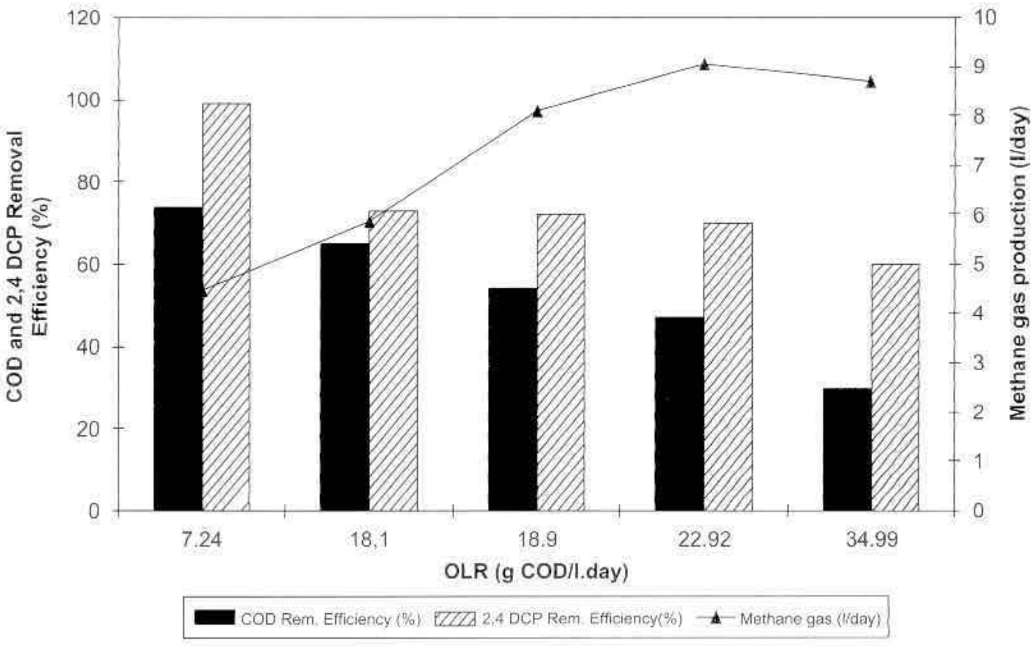


Fig. 2. COD, 2,4 DCP and methane gas productions versus organic loading rate in UASB reactor.

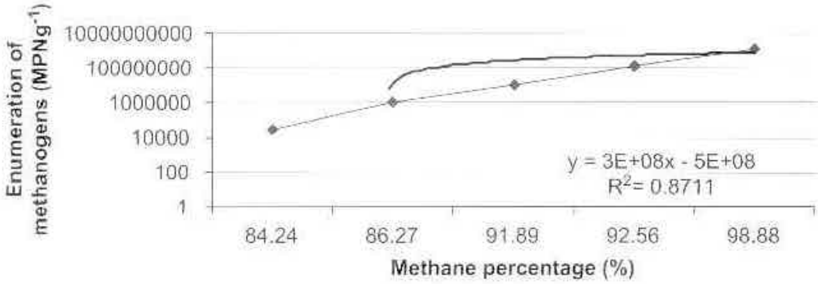


Fig. 3. Number of methanogens versus methane percentage.

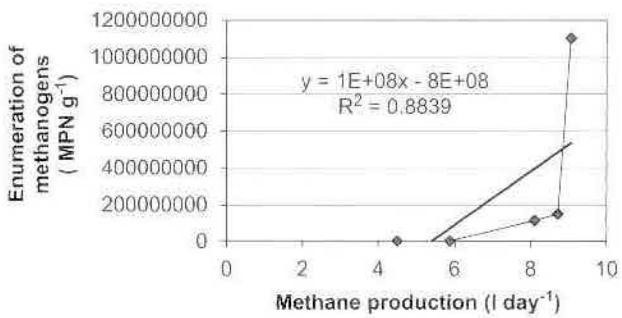


Fig. 4. Number of methanogens versus methane production.

productions was obtained with a high regression coefficient ($y = 1E+08x-8E+08$, $R^2 = 0.88$) (see Fig. 4).

3.3. Total number of methanogens at different VFA levels

The total number of methanogens at different VFA levels is shown in Fig. 5. VFAs concentration varied between 331.43 and 1284.74 mg l^{-1}. VFA production is an indicator of the presence of acedogenic bacteria. As shown in this figure the VFA production decreased to 331.43 from 1284.74 mg l^{-1} when the number of methanogens increased to 10×10^9 from 11×10^3 MPN g^{-1}. Therefore, the highest number of methanogens was observed at lowest VFA productions. At highest VFA concentration, the pH of the anaerobic UASB reactor decreased to 6.0–6.5 (see Table 1). As the VFA increased the number of methanogens also decreased. The growth of methanogens was inhibited under acidic conditions. A strong exponential correlation between the number of total methanogens and VFA production was observed with a regression coefficient of $R^2 = 0.86$ ($y = -1E+06X+1E+09$).

3.4. Enumeration of total acedogens at varying VFA levels

Enumeration of total acedogens at varying VFA levels is illustrated in Fig. 6. It was found that the VFA concentration in the UASB system was related to the number of acedogenic bacteria. A high correlation coefficient ($R^2 = 0.87$) was obtained between the number of acedogens and VFA production. As can be seen from Fig. 6, as the number of acedogenic bacteria increased to 30 from 10 cfu ml^{-1}, the VFA production increased to 1284.74 from 331.43 mg l^{-1}. The results of this study showed that the high VFA concentrations decreased the pH, and increased the number of methanogens at high COD and 2,4 DCP loadings (see Table 1). A strong linear correlation between the number of total acedogens and VFA production was observed ($y = 0.021x + 4.97$). Increases in VFA concentration indicated that the activity of acid producing bacteria and VFA accumulation. High VFA concentration indicates acidogenic activity besides methanogenic activity, which shows that methane gas is produced slowly in the UASB reactor.

3.5. Total number of methanogen and acedogen bacteria at varying COD removal efficiencies

The number of methanogens are an indicator of COD removal efficiency and methane gas production.The number of methanogens increased from 11×10^9 to 10×10^9 MPN g^{-1} as the COD removal efficiencies increased from 30% to 74% (see Fig. 7). A strong exponential correlation between the number of methanogens and COD removal efficiency was observed with a high regression coefficient ($y = 2E+07x-7E+08$, $R^2 = 0.96$). However, it can be concluded that an increase in acedogens in anaerobic UASB reactor affected the growth of methanogenic bacteria negatively. Therefore, the treatment efficiency of the anaerobic system decreased. In other words, the steady state conditions failed since high number of acedogens affected the pH of anaerobic UASB reactor. Generally, the pH of the system declined below 7.0 (see Table 1). Acidic conditions in the reactor affected the performance of the system for methanogenic bacteria. However, the existence of acedogens are necessary for anaerobic system. As shown in Fig. 7, the number of acedogenic bacteria was 10 cfu ml^{-1} in UASB reactor while the highest COD removal efficiency ($E = 74\%$) was obtained. A low number of acedogens was useful for anaerobic system performance.

The number of total acedogens in reactor treating 2,4 DCP at varying COD removal efficiencies is shown in Fig. 8. Acedogens are probably important in anaerobic ecosystems since they may produce acetate from lactate, H_2, CO_2, and methanol from the methoxyl or phenolic

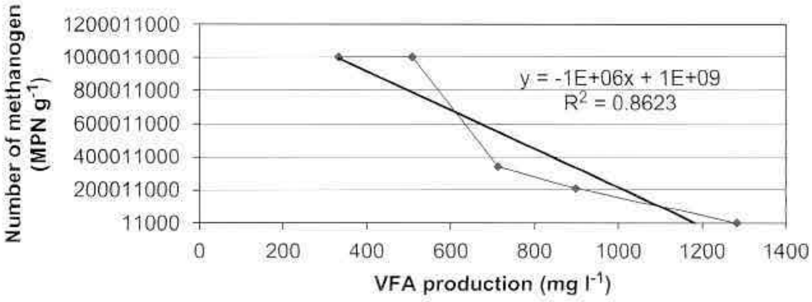


Fig. 5. Number of methanogens versus VFA production.

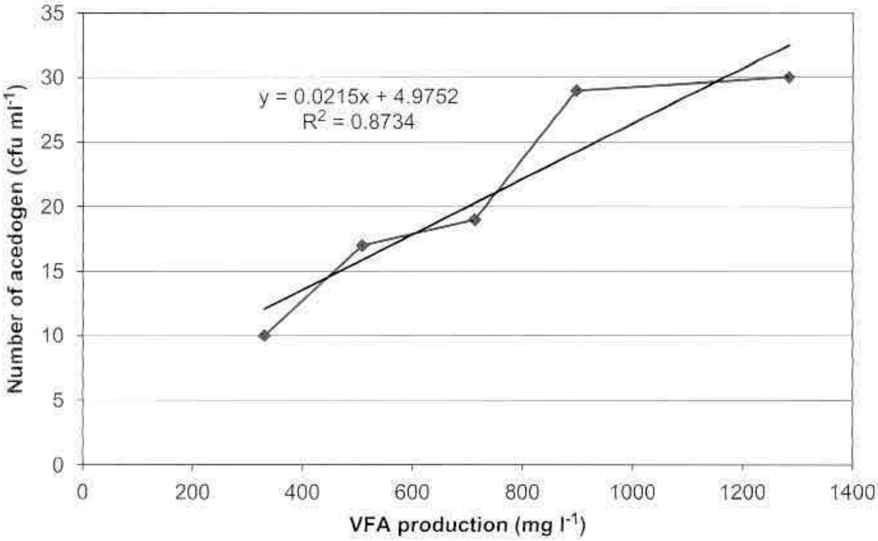


Fig. 6. Number of acedogens versus VFA production.

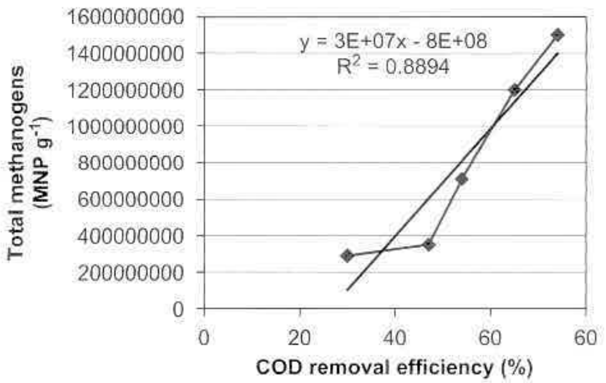


Fig. 7. Number of methanogens at different COD removal efficiencies.

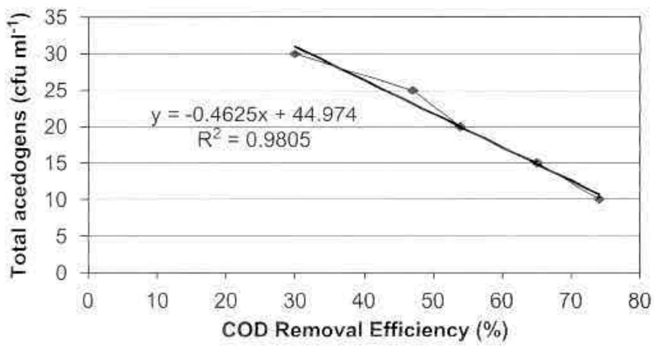


Fig. 8. Number of acedogens at different COD removal efficiencies.

compounds. Methanogens use these products as substrate through methanogenesis.

The number of acedogens decreased to 10 from 30 cfu ml^{-1} as the COD removal efficiencies decreased from 30% to 74% (see Fig. 7). In other words, when the number of total acedogens increased to 30 from 10 cfu ml^{-1}, the COD removal efficiency decreased to 30% from 74%. A strong linear correlation between the number of acedogens and COD removal efficiency was observed with a regression coefficient R^2 of 0.98 ($y = 0.4625x + 44.971$).

3.6. Enumeration of total methanogens at varying 2,4 DCP removal efficiencies

As the 2,4 DCP removal efficiencies increased from 60% to 99% the number of total methanogens increased from 11×10^3 to 10×10^9 MPN g^{-1}. Increases in the number of total methanogens was observed as the 2,4 DCP removal efficiencies increased from 60% to 99% (see Fig. 9).

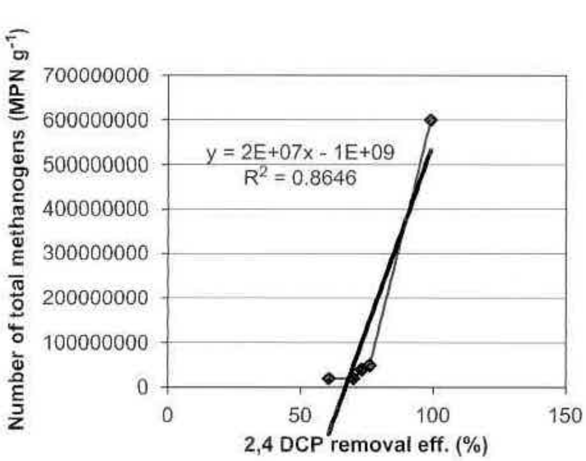


Fig. 9. Number of methanogens at varying 2,4 DCP removal efficiencies.

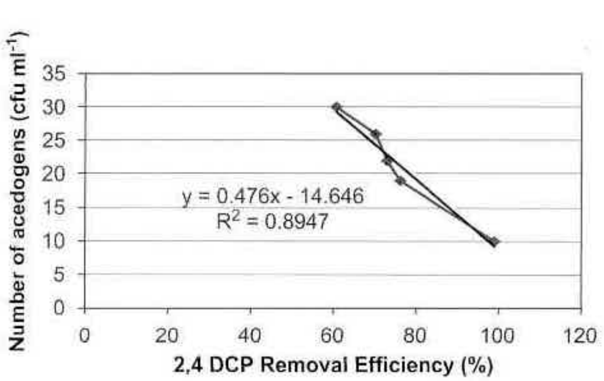


Fig. 10. Number of acedogens at varying 2,4 DCP removal efficiencies.

The 2,4 DCP loading rate and the environmental conditions of anaerobic reactor affects the number of methanogens. A strong exponential reaction kinetic between the number of methanogens and 2,4 DCP removal efficiencies was observed with a high regression coefficient ($y = 3E+06x-2E+08$, $R^2 = 0.97$).

3.7. Enumeration of total acedogens at varying 2,4 DCP removal efficiencies

The number of acedogens decreased to 10 from 30 cfu ml^{-1} as the 2,4 DCP removal efficiencies decreased from 99.72% to 60.60% (see Fig. 10). In other words, when the number of total acedogens increased the 2,4 DCP removal efficiency decreased. The number of methanogens are an indicator of COD and 2,4 DCP removal efficiencies and methane gas production. However, it can be concluded that an increase in acedogens in anaerobic UASB reactor negatively affected the growth of methanogenic bacteria. Therefore, the steady state conditions failed since high number of acedogens affected the pH of anaerobic UASB reactor. Acidic conditions in the reactor affected the performance of the system for methanogenic bacteria. However, the existence of acedogens are necessary for anaerobic system. A low number of acedogens was useful for anaerobic system performance in order to hydrolyze the VFA to methane and carbondioxide as end products. A strong linear correlation between the number of acedogens and 2,4 DCP removal efficiency was observed with a regression coefficient R^2 of 0.89.

As the 2,4 DCP removal efficiencies decreased from 99.72% to 60.60% the total number of acedogens decreased to 10 from 30 cfu ml^{-1}. A strong linear correlation between acedogen numbers and 2,4 DCP removal efficiencies was observed ($y = 0.476x-14.64$).

3.8. Biochemical test results

3.8.1. Substrate utilization tests

The growth on methanol, methylamine, formate, 2 proponol, 2 butanol, ethanol, 2 methyl butyrate, butanol was taken into consideration to determine the methanogenic species. For identification of *Methanobacterium bryantii*, *Methanobacterium formicicum*, *Methanobrevibacter smithii* and *Methanobrevibacterter arboriphilicus* species formate, acetate and methanol tests were carried out in order to determine whether these organic substrates could be used as carbon source.

3.8.2. Aminoacid biochemical tests

Amino acid biochemical tests were used for the identification of the members of Methanococcaceae family. Leucine and isoleucine biochemical test was applied on the methanogen cultures. Leucine and isoleucine are necessary for growth *of Methanococcus voltae*. Therefore, the leucine biochemical test proved positive (see Table 2). However, leucine is not required for the growth of *Methanococcus vannielii*. Therefore, the leucine biochemical test proved negative.

3.8.3. Utilization of sulfur sources

Very different sulfur resources such as sulfide, elemental sulfide were essayed in culture media to identify the methanogen species, especially the *Methanococcus* through methanogenesis (see Table 3). Growth of methanogen bacteria in sulfur containing media was assessed as positive through the utilization of sulfur sources.

Table 2
Biochemical characteristics of *Methanococcus* species

Characteristics	*M. vannielii*	*M. voltae*	*M. maripuladis*	*M. jannaschii*
Irregular coccus	+	+	+	+
Cell diameter (μm)	1	1.5	1	1
H_2 and CO_2	+	+	+	+
Formate	+	+	+	–
Acetate	–	+	–	–
Isoleucine	–	+	–	–
Leucine	–	+	–	–
Ca	–	+	–	–
NH_3	+	+	+	+
N_2	–	–	+	None
Alanine	–	–	+	None

Table 3
Biochemical characteristics of *Methanococcus* species

Characteristics	*M. vannielii*	*M. voltae*	*M. maripaludis*	*M. jannaschii*
Autotrophic growth	+	−	+	+
Sulfide	+	+	+	+
Elemental sulfur	+	+	+	+
Thiosulfide	−	−	−	
Sulfite	−	−	−	
Sulfate	−	−	−	
Temperature range (°C)	20–40	20–45	20–45	50–86
pH range	7–9	6.5–8	6.5–8	5–7

Table 4
Biochemical characteristics of *Methanosarcina* species

Characteristics	*M. barkeri*	*M. mazei*	*M. acetivotans*	*M. vacoulata*
H_2 and CO_2	+	+	+	+
Acetate	+	+ −	+	+
Methylamines	+	+	+	+
Methanol	+	+	+	+
Single cells or small aggregates	+	+	+	+
Large aggregates	−	+	+	−
Cysts	−	+	+	−
Life cycle	−	+	+	−
Gas vacuoles	−	+	+	+
Optimum temperature (°C)	30–50	30–40	35–40	40
Optimum NaCl	−	−	+	−
Heteropolysaccharide in outer layer	+	+	−	+

3.8.4. Differences in physiological properties of methanogens in different environmental conditions

Some differences in the physiological properties of methanotrophs (optimal values of pH, moisture, salinity, optimum temperature) were take into consideration for determination of the methanogenic species. For identification of *Methanosarcina* species the growth of this bacteria in the medium containing salinity was taken into consideration. *Methanosarcina acetivorans* lives in the medium containing NaCl concentration higher than 0.2 molar (see Table 4). *Methanosarcina acetivorans* does not have heteropolysaccharide in the outer layer of the cell. However, other species of *Methanosarcina* have heteropolysaccharide in the outer layer of the cell. For identification of *Methanosarcina mazei* and *Methanosarcina barkeri* formation of cyst was taken into consideration. Small aggregates of *Methanosarcina mazei* developed into cysts containing individual coccoid elements with a common wall. Cyst formation was not observed in the growth of *Methanosarcina barkeri* (see Table 4).

Identification of *Methanogenium* species was done by morphological observation. Capsule and flagella formation were taken into consideration for identification of *Methanogenium aggregans* and *Methanogenium bourgense*

Table 5
Biochemical characteristics of *Methanogenium* species

	M. cariaci	*M. marisnigri*	*M. olentangyi*	*M. tationis*	*M. aggregans*	*M. bourgense*
Flagella	+	+	−	+	−	−
Capsule	−	−	−	−	+	−
Pili	+	−	−	−	−	−
Requires acetate	+	−	+	+	+	+
Growth on formate	+	+	−	+	+	+
Growth at 6 °C	−	−	−	−	−	−
Growth at 15 °C	+	+	−	−	−	−

Table 6
Biochemical characteristics of *Methanobacterium* species

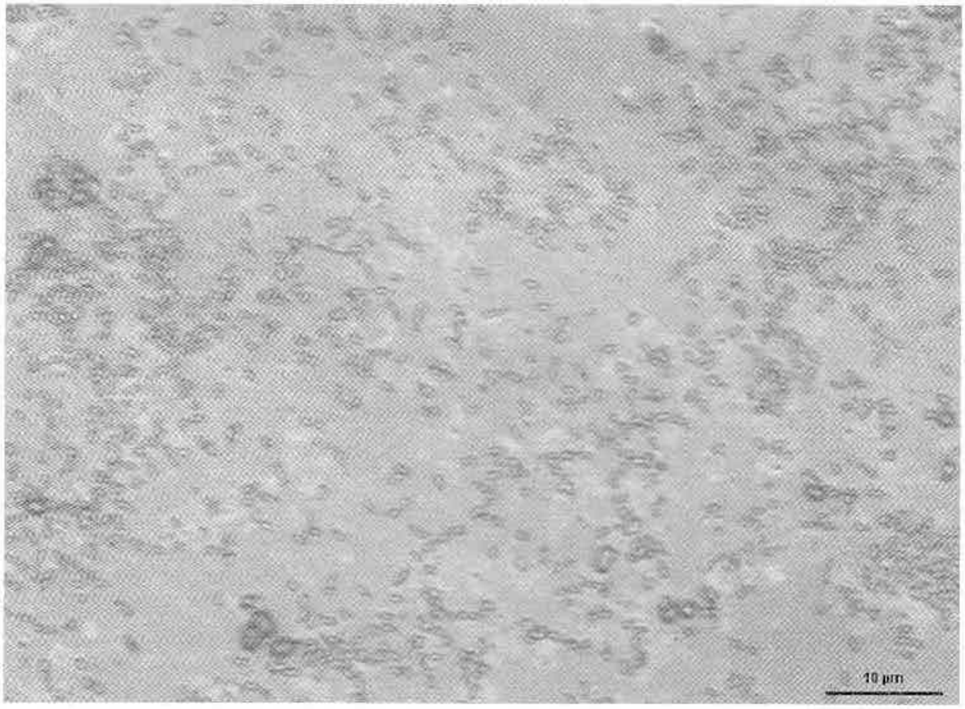
	M. formicicum	*M. wolfei*	*M. uliginosum*	*M. alcaliphilum*	*M. bryantii*
H_2 and CO_2	+	+	+	+	+
Formate	+	−	−	−	−
Long rods	+	+	+	+	+
Filaments	+	−	−	+	+
Gram staining	Variable	+	+	−	Variable
Motility	−	−	−	−	−
Optimum temperature (°C)	37–45	55–65	37–40	37	37–39
Optimum pH	6.6–7.8	7.0–7.5	6.0–8.5	8.0–9.1	6.9–7.2

Table 7
Biochemical characteristics of *Methanobrevibacter* species

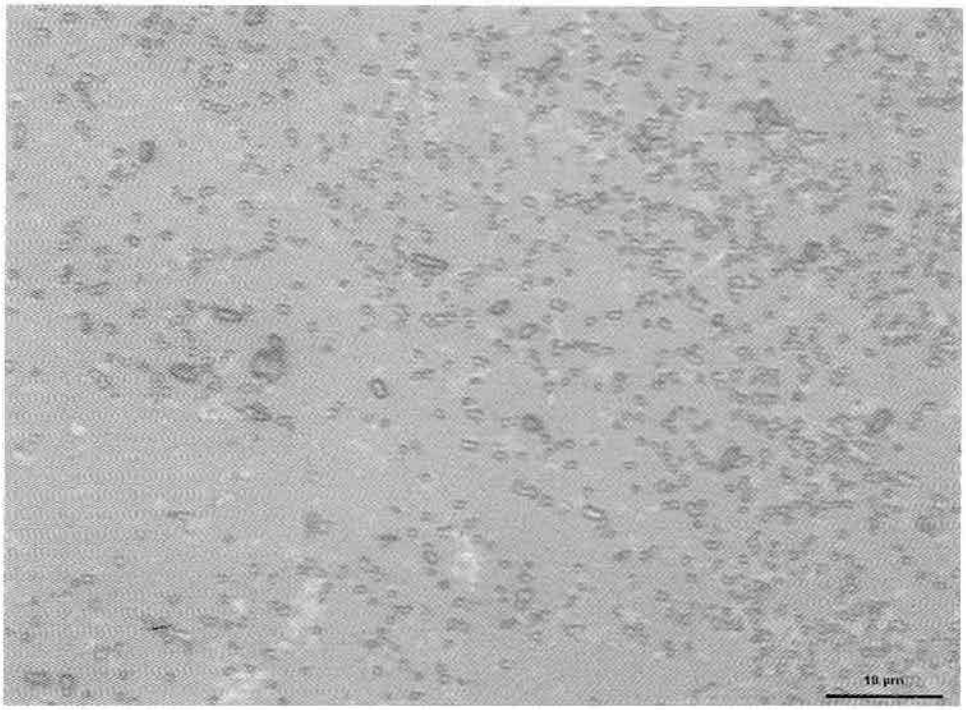
Characteristics	*M. ruminantium*	*M. smithii*	*M. arboriphilicus*
Morphology	Oval rod or coccus	Oval rod or coccus	Short rod
Acetate	+	−	+
B vitamins	+	+	+
Coenzyme M	+	−	−
2 Methylbutyrate	+	−	−
Aminoacids	+	−	+ −
NH_4	+	+	+
H_2 and CO_2	+	+	+
Formate	+	+	−
Bile inhibition	+	−	+

(Table 5). *Methanogenium aggregans* has a capsule while *Methanogenium bourgense* does not have a capsule. The biochemical test results for *Methanobacterium and Methanobrevibacter* species are shown in Tables 6 and 7.

According to biochemical test results the identification of *Methanobacterium bryantii* (Picture 1), *Methanobacterium formicicum* (Picture 2), *Methanococcus vannielii*, *Methanosarcina mazei*, *Methanobrevibacter smithii* (Picture 3), *Methanococcus voltae*, *Methanosarcina acetivorans* (Pictures 4–6), *Methanosarcina acetivorans* (Fig. 5), *Methanosarcina barkeri* (Fig. 6) were performed.



Picture 1. *Methanobacterium bryantii*, long rods or filaments cells. Magnification 500 ×.

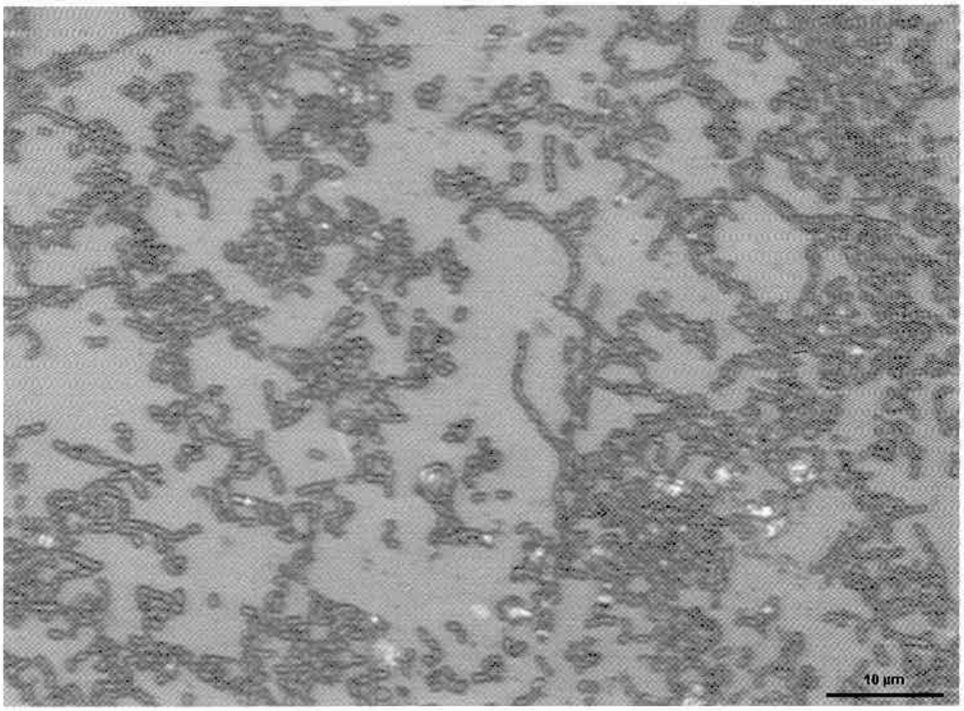


Picture 2. *Methanobacterium formicicum*, long rods or filaments cells. Magnification 500 ×.

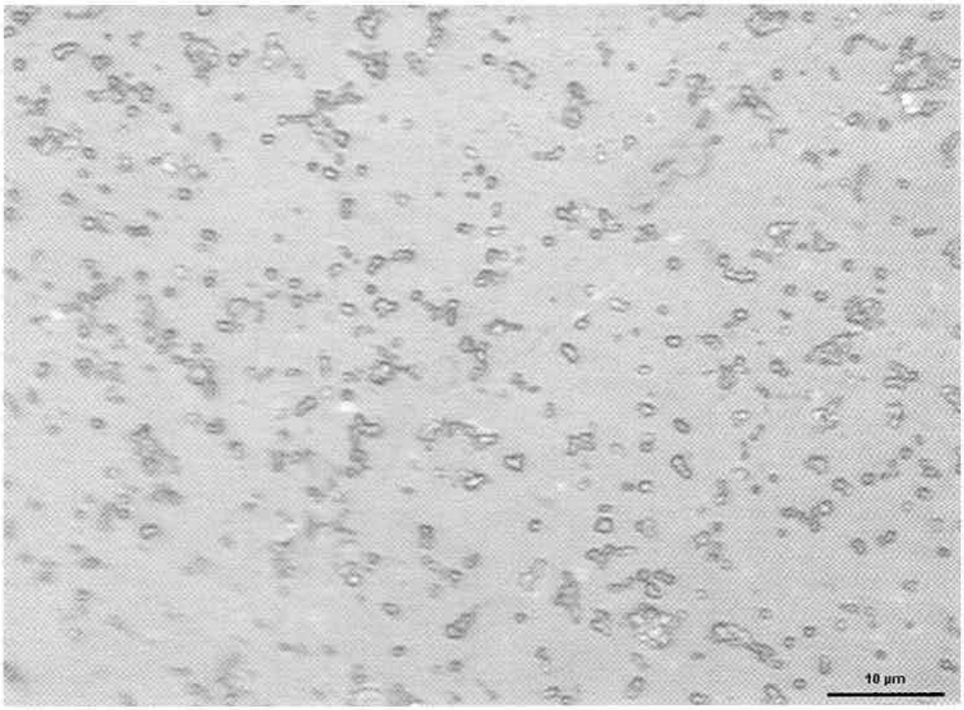
3.9. Multiple regression analysis

The multiple regression analysis was performed to find out the correlation between y (dependent variables; methanogen ($MPN\,g^{-1}$), acedogen numbers ($cfu\,ml^{-1}$)) and x (independent variables; COD removal efficiency (%), methane percentage (%), VFA concentration ($mg\,l^{-1}$) and 2,4 DCP removal efficiency (%)).

The multiple regression analysis results showed that high linear correlations were observed between x and y variables. A high linear relationship between methanogen number, COD removal efficiency, methane percentage,



Picture 3. *Methanobrevibacter smithii*, oval rods or cocci to short rods cells. Magnification 500 ×.



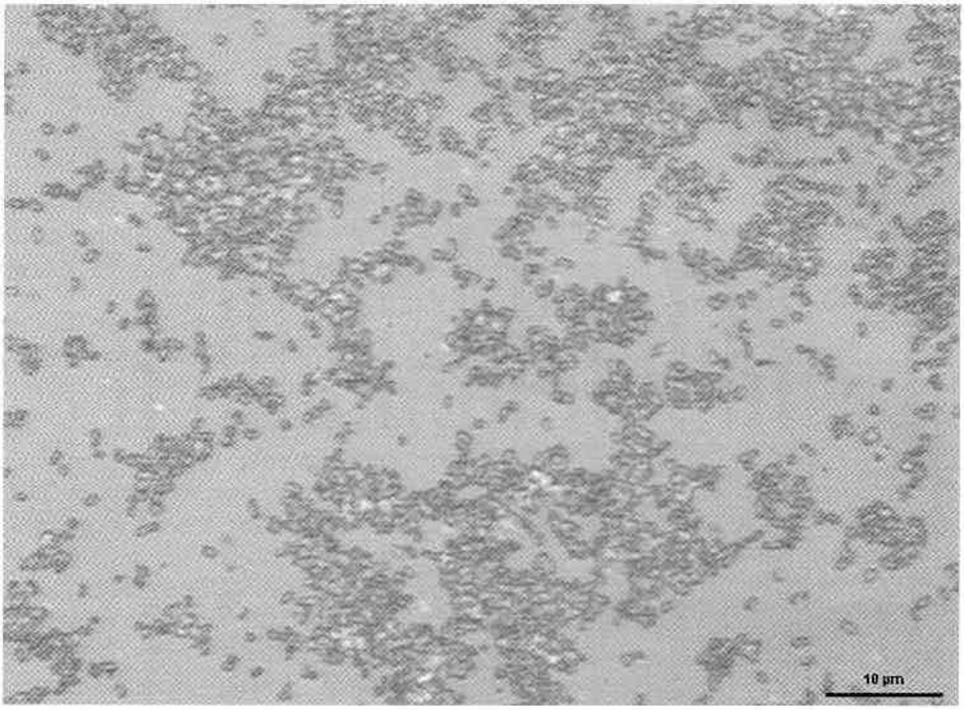
Picture 4. *Methanococcus voltae*, coccoidal cells. Magnification 500 ×.

VFA and 2,4 DCP concentrations was observed in reactor treating 2,4 DCP ($R^2 = 0.97$, $R = 0.98$, df $= 4$, $F = 30.08$). The relationships between x and y variables in this Reactor can be expressed as

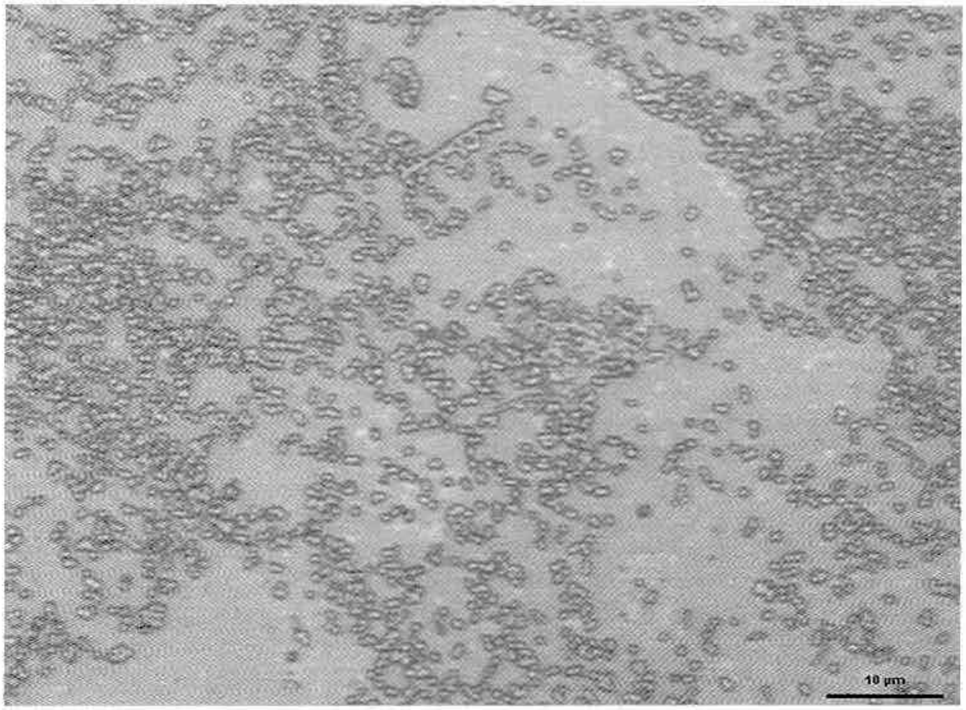
$$y = -38\,139\,530\,623 + 527\,447\,064.5\,X_1 + 22\,450\,579\,X_2 - 75\,163\,479\,X_3 + 6\,754\,321\,X_4. \quad (1)$$

Methanogen number (MPN g^{-1}) $= -38\,139\,530\,623 +$ 527 447 064.5 COD removal efficiency (%) + 22 450 579 methane percentage (%) −75 163 479 VFA production (mg l^{-1}) + 6754321 2,4 DCP removal efficiency (%) (Eq. (1)).

A high linear relationship between acedogen number, COD removal efficiency, VFA concentration, 2,4 DCP removal efficiency and methane percentage was observed in



Picture 5. *Methanosarcina acetivorans*, individiual, irregular coccoid cells forming small and large aggregates. Magnification 500 ×.



Picture 6. *Methanosarcina barkeri*, individiual, irregular coccoid cells forming small aggregates. Magnification 500 ×.

reactor treating 2,4 DCP ($R^2 = 0.79$, $R = 0.80$, $\mathrm{df} = 4$, $F = 18.09$).

The relationships between x and y variables in this Reactor can be expressed as

$$y = 196.642 - 2.178X_1 - 8.4E - 02X_2 + 7.9E - 02X_3 + 6.2E - 01X_4. \quad (2)$$

Acedogen number ($\mathrm{cfu\,ml^{-1}}$) = 196.642–2.178 COD removal efficiency (%)–8.4E02 VFA production ($\mathrm{mg\,l^{-1}}$)+7.9E–02 2,4 DCP removal efficiency (%)+ 6.2E–0.1 methane percentage (%) (Eq. (2)).

Table 8 summarizes the number of methanogen bacteria and the dominated bacteria identified in different industrial wastewaters.

Table 8
Number of methanogens and dominated bacteria in different industrial wastewater through anaerobic degradation

Industry	Type of reactor	Organic loading rate	VFA ($mg l^{-1}$)	COD removal efficiency (%)	Methane yield	Number of methanogens	Dominated bacteria	References
Synthetic ice cream	UASB	$2 kg COD m^{-3} day^{-1}$	342	60	$0.4 m^3 kg COD^{-1}$ removed.	3.1×10^5 MPN mg VSS^{-1}	*Methanosarcina* species, *Methanobacterium formicicum*	Morgan et al. (1990)
Synthetic ice cream	Anaerobic contact process	$3 kg COD m^{-3} day^{-1}$	241	80	$0.4 m^3 kg COD^{-1}$ removed	4×10^3 MPN mg VSS^{-1}	*Methanobacterium formicicum*, *Methanosarcina* sp.	Morgan et al. (1990)
Synthetic ice cream	Anaerobic flagella filter	$5 kg COD m^{-3} day^{-1}$	225	75	$0.4 m^3 kg COD^{-1}$ removed	2.3×10^5 MPN mg VSS^{-1}	*Methanobacterium formicicum*, *Methanosarcina* sp.	Morgan et al. (1990)
Synthetic ice cream	Fluidized bed reactor	$1.8 kg COD m^{-3} day^{-1}$	231	60	$0.1 m^3 kg COD^{-1}$ removed	5.2×10^7 MPN mg VSS^{-1}	*Methanobacterium formicicum*, *Methanobrevibacter* sp.	Morgan et al. (1990)
Synthetic wastewater	Anaerobic upflow filter	$4 kg COD m^{-3} day^{-1}$	500	80	$0.35 m^3 CH_4 kg COD^{-1}$ removed	5.1×10^5 MPN g^{-1}	*Methanosarcina*, *Methanococcus*, *Methanobacterium*, *Methanobrevibacter*	Suihko et al. (2005)
Brewery wastewater	Membrane anaerobic reactor system	$2.5 kg COD m^{-3} day^{-1}$	1600	80	$0.4 m^3 CH_4 kg COD^{-1}$ removed	4.6×10^6 MPN g^{-1}	*Methanosarcina*, *Methanococcus*, *Methanobacterium*, *Methanobrevibacter*	Pearce et al. (2003)
Treatment of pentachloro phenol	Hybrid reactor	$4.96 g COD l^{-1} day^{-1}$	5000	97	$0.861 day^{-1}$	10^4–10^6 cells ml^{-1}	*Methanosarcina*, *Methanosaete* sp.	Montenegro et al. (2001)
Treatment of 2,4 DCP	UASB	$7.24 g COD l^{-1} day^{-1}$	331	74	$4.481 day^{-1}$	15×10^8 MPN g^{-1}	*Methanobacterium formicicum*	In this study, Sponza and Uluköy (2005)

At a COD removal efficiency of 74% the number of methanogen bacteria was higher in 2,4 DCP treating reactor (10×10^9 MPN g^{-1}) compared to the ice cream, brewery and penta chlorophenol wastewater (4×10^3–5×10^5 MPN g^{-1}) (Zhu et al., 1997; Montenegro et al., 2001; Suihko et al., 2005; Pearce et al., 2003). In the study performed by Sponza and Uluköy (2005), the highest methanogenic bacteria number (8×10^7 MPN g^{-1}) was obtained at an organic loading rate of 7.24 g COD l^{-1} day^{-1}.

4. Conclusions

The aim of this study was to gain more insight into the interactions between methanogen, acedogen bacteria and the reactor performances depending on different operational conditions in a UASB reactor treating 2,4 DCP. It was observed that the reactor performances were dependent on operational conditions and there were high correlations between reactor performances and anaerobic microorganisms. Operational conditions in anaerobic wastewater treatment plants treating the complex organic compounds strongly influence the number of methanogens and acedogens. There was evidence that in the UASB reactor treating 2,4 DCP was hydrolyzed by acedogens to VFA together with COD. Methanogens convert these intermediate organic substances to methane and carbondioxide and glucose was preferred as primary substrate. The characterization of the bacteria in granular sludge helped to get a better understanding of the role of bacterial group in the sludge of UASB for 2,4 DCP degradation in different reactor performances. The maximum methanogen number was as 10×10^9 MPN g^{-1} and obtained at maximum COD, 2,4 DCP removals (74% and 99.72%, respectively) at low VFA (331.43 mg $CH_3COOH\ l^{-1}$) and at high methane percentage (98.88%). The maximum acedogen number was obtained as 30 cfu ml^{-1} at low COD, 2,4 DCP removals (30% and 60.60%, respectively), at high VFA (1284.74 43 mg $CH_3COOH\ l^{-1}$) and at low methane percentage (82.24%). The reactor performances was adjusted with decreasing the HRT or increasing the upflow rates. This results with increasing of COD and 2,4 DCP loading rates which affect the number of anaerobic microbial consortia in a UASB reactor treating 2,4 DCP with glucose.

A high linear relationship between methanogen number and COD removal efficiency, methane percentage, VFA concentration and 2,4 DCP removal efficiency was observed in UASB reactor. According to biochemical test the identification of six methanogen bacteria was performed. The operational conditions affect the microbial population in UASB reactor. Therefore, the dominant bacteria numbers will be varied in the reactors containing high toxic substances and operated at high organic loading rates depending to reactor performance.

Acknowledgments

This study was executed as a part of the research activities of the Environmental Microbiology Laboratory of Environmental Engineering Department of projects numbered

Fen 021, Fen 051, and 03.KB.Fen.017, which were partially funded by the Dokuz Eylül University Research Foundation. The authors would like to thank this body for the financial support given to the projects.

References

Anderson, G., Kasapgil, B., Ince, O., 1994. Microbiological study of two-stage anaerobic digestion during start-up. Water Research 28 (11), 2383–2392.

Anderson, G.K., Yang, G., 1992. Determination of bicarbonate and total volatile acid concentration in anaerobic digesters using a simple titration. Water Environment Research 64, 53–59.

APHA–AWWA, 1992. Standard Methods for the Examination of Assoc/American Water and Wastewater, 17th edition. American Publ Health Assoc/American Water Works Association, Washington, DC, USA, pp. 440–670.

Balch, W.E., Fox, G.E., Magram, L., Woese, C.R., Wolfe, R.S., 1979. Methanogens: re-evaluation of a unique biological group. Microbial Review 43, 260–296.

Beydilli, M.I., Pavlosathis, S.G., Tincher, W.C., 1998. Decolorization and toxicity screening of selected reactive azo dyes under methanogenic conditions. Water Science and Technology 38 (4–5), 225–232.

Bryant, M.P., 1972. Commentary on Hungate technique for culture of anaerobic bacteria. American Journal of Clinical Nutrition 251, 324–1327.

Doddema, H.J., Vogels, G.D., 1978. Improved identification of methanogenic bacteria by fluorescence microscopy. Applied and Environmental Microbiology 36 (5), 752–754.

Dolfing, J., 1987. Microbiological aspects of granular methanogenic sludge. W.A.U. Dissertation No. 1138.

DSMZ. Deutsche Sammlung von Microorganismen und Zell Kulturen GMBH, 2003. Braunschweig, Germany, web site ⟨http://DSMZ⟩, 10 February 2003).

EPA US Environmental Protection Agency (EPA), 2003. US Environmental Protection Agency (EPA) This World Wide Web (WWW) site, the home page of the Environmental Protection Agency (EPA).

Grotenhuis, J.T.C., 1992. Structure and stability of methanogenic granular sludge. W.A.U. Dissertation No. 1520.

Hulshoff, N., Pol, L.W., 1989. The phenomenon of granulation of anaerobic sludge. W.A.U. Dissertation No. 1323.

Mink, R.W., Patrick, R.D., 1977. Tentative identification of methanogenic bacteria by fluorescence microscopy. Applied and Environmental Microbiology 33 (3), 713–717.

Montenegro, M.A.D., Mattos, M., Soares, H.M., Vazoller, V.F., 2001. Hybrid reactor performance in pentachlorophenol removal by anaerobic granules. Water Science and Technology 44 (4), 137–144.

Morgan, J.W., Goodwin, J.A.S., Wase, D.A., Forster, J.C.F., 1990. The effects of using various types of carbonaceous substrate on UASB granules and on reactor performance. Biological Wastes 4 (1), 55–71.

Morgan, J.W., Evison, L.M., Forster, C.F., 1991. Changes to the microbial ecology in anaerobic digesters treating ice cream wastewater during start-up. Water Research 25 (6), 639–653.

Pearce, C.I., Lloyd, J.R., Guthrie, J.T., 2003. The removal of colour from textile wastewater using whole bacterial cells. Dyes and Pigments 58 (3), 179–196.

Plumb, J.J., Bell, J., Stuckey, D.C., 2001. Microbial population associated with treatment of an industrial dye effluent in an anaerobic baffled reactor. Applied and Environmental Microbiology 67 (7), 3226–3235.

Razo-Flores, E., Luijten, M., Donlon, B.A., Lettinga, G., Field, J.A., 1997. Biodegradation of selected azo dye under methanogenic conditions. Water Science and Technology 36 (6–7), 65–72.

Schauer, D.P., Brown, J., Ferry, G., 1982. Kinetics of formate metobolism in *Methanobacterium formimicum* and *Methanospirillum hungatei*. Applied and Environmental Microbiology 44 (3), 549–554.

Schmidt, B., Ahring, K., 1999. Immobilization patterns and dynamics of acetate utilizing methanogens immobilized in sterile granular sludge in U.A.S.B. Applied and Environmental Microbiology 65 (3), 1050–1054.

Sekiguchi, Y., Kamagata, K., Nakamura, A., Harada, H., 1999. Fluorescence in situ hybridization using 16S rRNA targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. Applied and Environmental Microbiology 65 (3), 1280–1289.

Smith, R., Mah, A., 1978. Growth and methanogenesis by *Methanosarcina* strain 227 on acetate and methanol. Applied and Environmental Microbiology 6, 870–879.

Speece, R., 1996. Anaerobic Biotechnology for Industrial Wastewaters. Archae Press, Tennesse, USA.

Sponza, D., 2002. Enhancement of granule formation and sludge retainment for tetrachloroethylene (TCE) removal in an upflow anaerobic sludge blanket (UASB) reactor. Advances in Environmental Research 7, 453–462.

Sponza, D.T., Ulukőy, A., 2005. Treatment of 2,4-dichlorophenol (DCP) in a sequential anaerobic (upflow anaerobic sludge blanket) aerobic (completely stirred tank) reactor system. Process Biochemistry 40 (11), 419–3428.

Suihko, M.L., Partanen, L., Sandholm, T.M.O., Raaska, L., 2005. Ocurrrence and molecular characterization of cultivable mesophilic and thermophilic obligate anaerobic bacteria isolated from paper mills. Systematic and Applied Microbiology 28 (1), 555–561.

Williams, S., Sharpe, E., Holt, J. (Eds.), 1989. Bergey's Manual of Systematic Bacteriology. Williams and Wilkins, Baltimore, USA.

Zehnder, A.J.B., Brock, T.D., 1979. Methane Formation and Methane Oxidation by Methanogenic Bacteria. University Park Press.

Zhu, J., Hu, J., Gu, X., 1997. The bacterial numeration and an observation of a new syntrophic association for granular sludge. Water Science and Technology 36 (6–7), 133–140.

Zinder, R., Mah, A., 1979. Isolation and characterization of a thermophilic strain of *Methanosarcina* unable to use H–CO for methanogenesis. Water Research 9, 340–349.

Journal of Environmental Management 87 (2008) 193–200

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Adsorption of basic dye from aqueous solution onto fly ash

J.X. Lin[a,b], S.L. Zhan[a,*], M.H. Fang[a], X.Q. Qian[a], H. Yang[b]

[a]College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310027, PR China
[b]College of Materials and Chemical Engineering, Zhejiang University, Hangzhou 310027, PR China

Received 6 May 2006; received in revised form 17 December 2006; accepted 10 January 2007
Available online 20 February 2007

Abstract

The fly ash treated by H_2SO_4 was used as a low-cost adsorbent for the removal of a typical dye, methylene blue, from aqueous solution. An increase in the specific surface area and dye-adsorption capacity was observed after the acid treatment. The adsorption isotherm and kinetics of the treated fly ash were studied. The experimental results were fitted using Langmuir and Freundlich isotherms. It shows that the Freundlich isotherm is better in describing the adsorption process. Two kinetic models, pseudo-first order and pseudo-second order, were employed to analyze the kinetic data. It was found that the pseudo-second-order model is the better choice to describe the adsorption behavior. The thermodynamic study reveals that the enthalpy (ΔH^0) value is positive (5.63 kJ/mol), suggesting an endothermic nature of the adsorption.

Keywords: Fly ash; Methylene blue; Adsorption; Wastewater; Kinetics model

1. Introduction

Dyes are widely used in various industries, such as textiles, paper, plastics, cosmetics and leather, for coloring their final product. The release of colored wastewater from these industries may present an eco-toxic hazard and introduce the potential danger of bioaccumulation, which may eventually affect man through the food chain. Many techniques have been used to remove harmful dyes from colored wastewater, and one of the powerful and convenient treatment processes is adsorption. Activated carbon is the most popular adsorbent, which is capable of adsorbing many dyes with a high adsorption capacity (Lin, 1993). But it is expensive and the costs of regeneration are high because desorption of the dye molecules is not easily achieved (McKay et al., 1987). Various low-cost adsorbents were therefore investigated as an alternative to activated carbon (Ramakrishna and Viraraghavan, 1997).

Fly ash is a waste material originating in large quantities from modern power stations. Although it has been successfully used as a mineral admixture in concrete and brick production, there are still superfluous fly ashes in some countries, causing environmental and disposal problem. The utilization of fly ash as adsorbent for dye removal from industrial wastewater could be rewarding to both environment and economy. Thus many efforts were attempted to make fly ash a cheap adsorbent for dye removal in recent years (Li et al., 2006; Wang et al., 2006; Talman and Atun, 2006; Eren and Acar, 2006; Mall et al., 2005, 2006; Weng and Pan, 2006; Fungaro et al., 2005; Kumar et al., 2005; Janos et al., 2003). Many works were focused on the absorption property of fly ash, and some treatments were carried out to improve its adsorption capacity, such as sonochemical treatment (Wang and Zhu, 2005), microwave heating (Wang et al., 2005a) and H_2O_2 treatment (Gupta et al., 2000). However, because of variations in coals from different sources, as well as differences in the design of coal-fired boilers, not all the fly ash is the same. For example, it may exhibit acid property (Wang et al., 2005b) or show basic property (Mohan et al., 2002). Different fly ashes may be different in their adsorption behaviors, and more efforts should be made to improve their adsorption capabilities.

*Corresponding author. Tel.: +86 571 87951397; fax: +86 571 87952483.
E-mail address: jzcl@zju.edu.cn (S.L. Zhan).

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.01.001

In this work, one type of coal fly ash with basic property was treated using H_2SO_4 for the adsorption of basic dye from aqueous solution. Methylene blue (MB), a typical basic dye usually used for dyeing and medical purpose, was used as adsorbate. After investigating the effect of pH on the adsorption of MB, we studied the adsorption isotherm and kinetics of MB onto the treated fly ash, and then simplified the pseudo-first-order and pseudo-second-order kinetic models to fit the experimental findings.

2. Materials and methods

2.1. Fly ash

Raw fly ash (R-FA) was provided from Ningbo Thermal Power Plant in Zhejiang province, China. The R-FA was in the form of spherical grayish particles with bulk density of 980 g/L and average particle size of 25 μm. This material was treated with sufficient H_2SO_4 solution (1 mol/L) at 50 °C for 24 h; the sample was then washed several times with distilled water, filtrated, dried at 105 °C for 20 h, and sieved to the particle sizes through 450–700 mesh. Finally, the resulting product (H-FA) was stored in a closed container for further tests.

2.2. Characterization techniques

The $pH_{solution}$ of fly ash was measured by using the method recommended by Al-Ghouti et al. (2003) as follows: 3 g of fly ash was mixed with 30 mL of distilled water and agitated for 24 h. Then the pH value of the mixture was recorded with a pH meter. The zero point of charge (pH_{zpc}) of fly ash was estimated by using the alkalimetric titration method (Noh and Schwarz, 1989; Al-Ghouti et al., 2003). Major oxide of fly ash was analyzed by X-ray florescence (XRF) technique. SEM micrographs were obtained using model SIRION microscope (Fei Company, Holland). For the specific surface area and pore size evaluation, nitrogen gas adsorption method was used. N_2 gas adsorption–desorption isotherms of fly ashes were measured at −196 °C using a model Autosorb-1 nitrogen-adsorption apparatus (Quantachrome Corp., USA).

2.3. Evaluation of MB

MB (CI No. 52015) was AR-grade chemical. Its pH value was 6.13 at the concentration of 2.675×10^{-5} M. The concentration of MB in each aqueous solution was measured on an ET99731 UV–visible spectrophotometer (Tintometer GmbH, Germany) by measuring absorbance at λ_{max} of 665 nm. A calibration plot was made in the concentration range of 3×10^{-6}–3×10^{-5} M for determination of the dye concentration.

To study the effect of pH value on the visible spectra of MB solution, 2.675×10^{-5} M of MB concentration was prepared at various pH values using 1 M HCl and NaOH solution. The visible spectra of MB at different pH values were obtained on the spectrophotometer once for several hours after the pH adjusted.

2.4. Adsorption studies

Adsorption isotherm studies were carried out in a set of 250 mL PE flask, where 1 g of fly ash and 100 mL of MB solution with different initial concentrations (2.675×10^{-5}–26.75×10^{-5} M) were added. The solution with fly ash was shaken at 130 rpm in an isothermal shaker for 72 h to reach the equilibrium at constant temperatures (15, 35 and 55 °C). The flasks were then taken from the shaker and the equilibrium concentration of dye was measured. The amount of dye uptake by fly ash, q_e (mol/g), was obtained as follows:

$$q_e = \frac{C_0 - C_e}{m_s}, \tag{1}$$

where C_0 and C_e (mol/L) are the initial and equilibrium concentrations of dye in solution, respectively, and m_s is the concentration of fly ash (g/L).

Kinetic studies were performed in a 2 dm^3 glass beaker, where 17 g of fly ash was added into 1.7 dm^3 of MB solution with different initial concentrations at varying temperature. The glass beaker was equipped with baffles, and an impeller driven by motor was adjusted on top of the beaker to stir the dye solution and fly ash. The agitation speed was fixed at 130 rpm during the experiments. At preset time intervals, the samples of 5 cm^3 were taken from the solution and the concentration was analyzed.

The dye solution was centrifuged at 2000 rpm for the separation of fly ash from the aqueous solution.

3. Results and discussion

3.1. Characterization of fly ash

The $pH_{solution}$ values of R-FA and H-FA were found to be 10.96 and 6.82, respectively. The cause of the decrease of $pH_{solution}$ value after the treatment is the removal of basic impurities from the R-FA. Major oxide analysis shows that the R-FA consists mainly of SiO_2 (48.5%), Al_2O_3 (24.7%) and Fe_2O_3 (9.0%). The pH_{zpc} is reported to be 8.2 and 2.3 for alumina and silica, respectively (Mohan et al., 2002). The composite pH_{zpc} of R-FA and H-FA was found to be 5.9 and 6.4, respectively. The SEM photographs of R-FA, H-FA and adsorbed H-FA with MB are shown in Fig. 1. It can be seen that the surface morphology of R-FA is slightly changed by the H_2SO_4 treatment, which is quite similar to that of H-FA. Fig. 1(a) and (b) clearly reveals the porous texture of fly ash, while Fig. 1(c) shows that the porous surface is greatly changed upon adsorbing MB.

The textural characteristics of fly ash were studied by using the N_2 adsorption method. The well-known Brunauer–Emmett–Teller (BET) equation was used for the calculation of specific surface area, S_{BET}, pore volume, V, and the average pore diameter, R. The results for R-FA

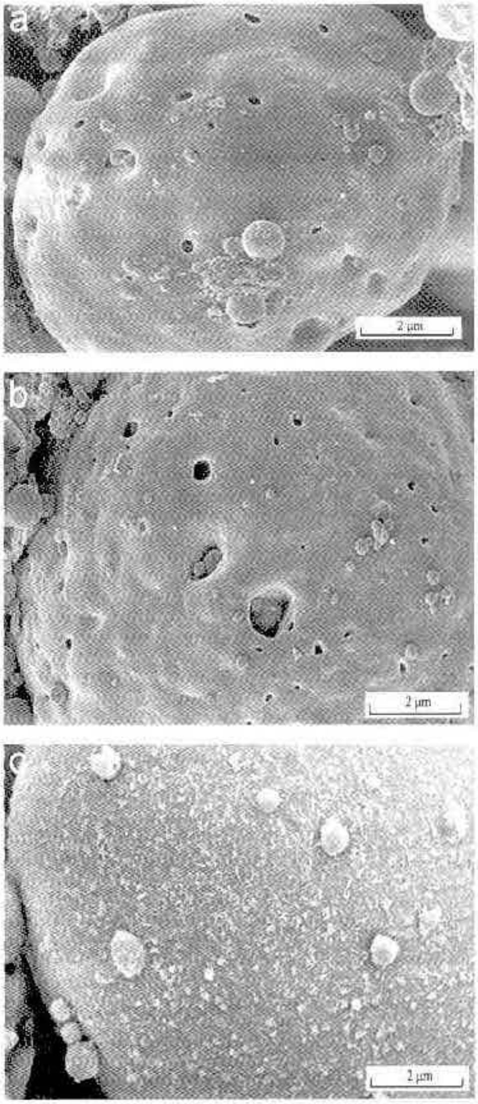


Fig. 1. SEM paragraphs of (a) R-FA, (b) H-FA and (c) H-FA adsorbed MB.

Table 1
Surface area and pore size of R-FA and H-FA

Adsorbent	V (10^{-3} cm^3/g)	S_{BET} (m^2/g)	R (nm)
R-FA	5.413	1.786	12.12
H-FA	12.59	6.236	8.075

and H-FA are given in Table 1. It reveals that the acid treatment enhances the surface area and pore volume of the R-FA, while decreases its average pore diameter. This may be explained by the increasing micropore content of the H-FA after the impurities and organics were removed by H_2SO_4 and water.

3.2. *Effect of pH on the MB adsorption*

Since the pH of the adsorbate–adsorbent system plays an important role in the whole adsorption process, many researchers studied the changes of dye adsorption on adsorbent over a broad range of pH values. Janos et al. (2003) and Wang et al. (2005a,b) investigated the effect of pH on the removal of MB from aqueous solution using fly ash. Their reports showed that the adsorption capacity enhanced with increasing pH. Especially, when the pH value was higher than 10, the adsorption of MB would increase steeply. Khraisheh et al. (2005) reported that the maximum percentage removal of MB using diatomite occurred at basic pH (10–12). The phenomenon above was usually explained by considering the pH_{zpc} of the fly ash or diatomite, i.e. the solution pH above the pH_{zpc} favors the adsorption of MB. However, the possible change of MB in basic solution without any adsorbents may also lead to the steep increase of adsorption of MB. Thus the possibility was tested by investigating the effect of pH on the visible spectra of MB solution without adsorbent.

As the spectra variation at given pH may be time dependence, we recorded the visible spectra of MB solution at interval time after the pH adjusted. It was found that the absorbance of the spectra was stable with pH varying in acid solution while decreased with increasing pH value in basic solution. Fig. 2 shows the spectra of MB solution at basic pH value (10.4–12) after the pH adjusted for 25 h. The value of absorbance decreases from 1.83 at pH 6.13 to 1.03 at pH 12 for the monomer peak (665 nm).

The visible spectra of MB solution at pH 12 recorded at different times are shown in Fig. 3. It can be seen that the absorbance of the MB solution decreases with time. The λ_{max} of the MB solution at pH 12 shifts from the monomer peak (665 nm) to the dimer peak (615 nm) after 40 h, and the absorbance decreases to 0.78 after 70 h. However, the same investigations at pH 2 and 4 showed that the MB in acid solution was stable. Further, the study of MB in a higher concentration of NaOH solution was carried out by Shawabkeh and Tutunji (2003). They mixed 2.7 mmol/L of MB solution with equal volume of 2 M NaOH solution

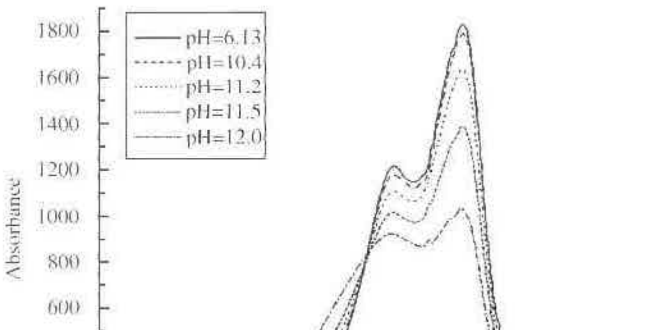


Fig. 2. Effect of pH on visible spectra of MB solution. MB concentration: 2.675e−5 M, recorded time: 25 h.

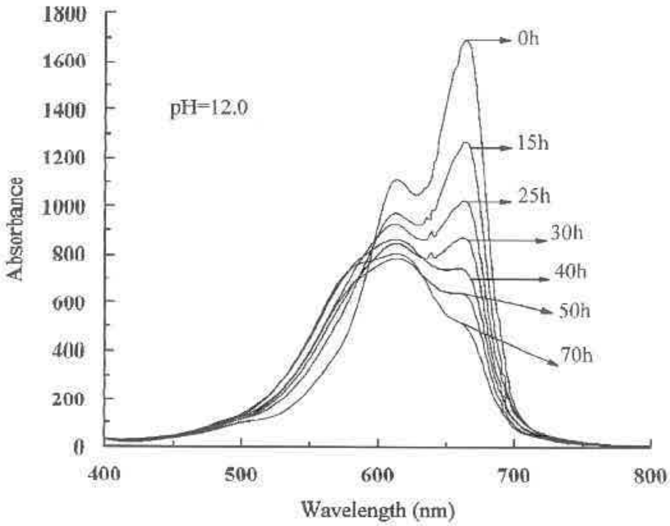


Fig. 3. Effect of time on visible spectra of MB solution at pH = 12.0. MB concentration: 2.675e–5 M.

without adsorbent. It showed that the color of the solution changed with time until it became clear after 5 days.

In short, the absorbance of MB solution at basic pH decreases with time; the higher the pH value, the greater it decreases. The reason for spectra variation of MB in NaOH solution may be the reaction of MB with NaOH and the appearance of dimer $(MB^+)_2$.

From the above results, it may be concluded that the above-mentioned steep increasing adsorption of MB at basic pH (10–12) attributes not only to the pH_{zpc} of the adsorbent but also to the self-change of MB in the basic solution.

It was found that the pH value of 10 mg/L MB solution mixed with 10 g/L R-FA was 8.65, and it was 6.71 for the H-FA. To compare the adsorption capability of H-FA with the capability of R-FA, the pH value of the dye solution with H-FA was adjusted to be 8.65 by NaOH solution. Fig. 4 shows the dye removal of H-FA and R-FA at the same pH value at various MB concentrations. The dye removal of H-FA is higher than that of R-FA in the same pH solution, which indicates that the H_2SO_4 treatment of R-FA enhances its adsorption capacity. In addition, the dye removal of H-FA without alkali added is lower than the dye removal of R-FA due to the basic property of R-FA.

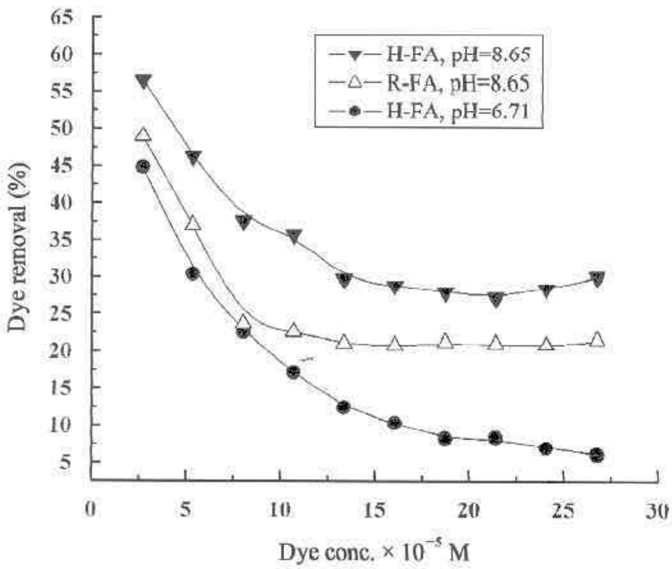


Fig. 4. Effect of solution pH on dye removal of R-FA and H-FA. Adsorbent concentration: 10 g/L, equilibrium time: 72 h, shaking speed: 130 rpm. T: 288 K.

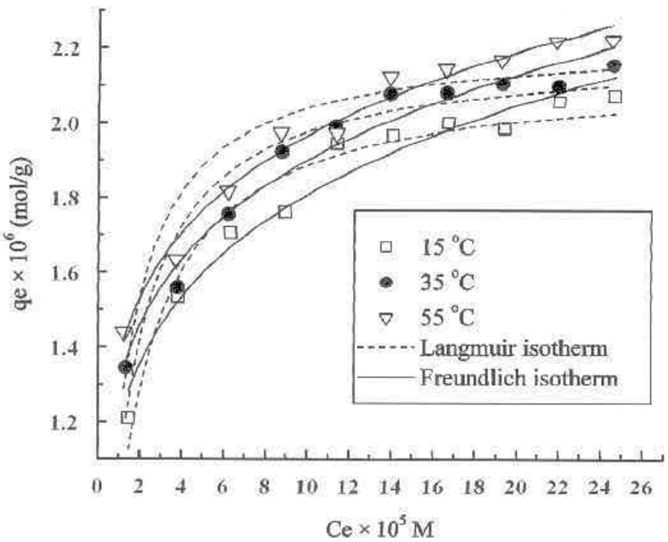


Fig. 5. Fitting the experimental data by Langmuir and Freundlich isotherms at various temperatures. Adsorbent concentration: 10 g/L, equilibrium time: 72 h, shaking speed: 130 rpm.

3.3. Adsorption isotherms

Fig. 5 depicts the amounts of MB uptake by H-FA against the equilibrium concentrations at various temperatures. It reveals that the dye uptake increases on increasing the dye concentration from 0 to 0.2675 mmol/L. Two widely used isotherm models, Langmuir and Freundlich isotherms, were employed to investigate the adsorption behavior. The Langmuir isotherm can be written as

$$q_e = \frac{QbC_e}{1 + bC_e}, \tag{2}$$

where Q is the adsorption capacity corresponding to form a complete monolayer and b is the Langmuir constant.

The Freundlich isotherm, which is empirical for heterogeneous surface energy, is written as

$$q_e = KC_e^{1/n}, \tag{3}$$

where K is the extent of the adsorption and n is the degree of nonlinearity between dye concentration and adsorption.

The above two equations were then used for nonlinear fitting of the experimental data in Fig. 5 by employing a commercial software Origin. The model parameters obtained from the nonlinear regression are given in Table 2.

It can be seen from the correlation coefficient (R^2) that the Freundlich model is better than the Langmuir model in fitting the experimental isotherms, suggesting the heterogeneous surface or pores for the adsorption of MB onto fly ash.

It is known that the increasing temperature may decrease the adsorption capacity of fly ash if it is exothermic, and increase the capacity if it is endothermic. As shown in Fig. 5 and Table 2, both the adsorbed amount of dye (q_e) and the adsorption capacity (Q) enhance with temperature, indicating that the adsorption process was endothermic.

A comparison of adsorption capacities of MB on various fly ashes is presented in Table 3. It can be seen that the fly ash in this study shows lower adsorption capacity than others. Therefore, further work should be carried out to improve the adsorption capacity.

3.4. Adsorption kinetics

A number of experimental parameters are usually considered in kinetics studies for one adsorption process. The effect of dye concentration and temperature on the kinetics of MB adsorption onto H-FA was focused on in this study.

Two well-known kinetic models, pseudo-first-order and pseudo-second-order model, were employed to describe the adsorption process. The pseudo-first order, proposed by Lagergren (1898) for adsorption analysis, is expressed in the form

$$\frac{dq_t}{dt} = k_1(q_e - q_t), \tag{4}$$

where k_1 is the rate constant of pseudo-first-order model and q_t is the amount of solute adsorbed on the adsorbent at time t. When rearranging q_t as the concentration form as Eq. (1), and integrating over the concentration with time, the following equation can be obtained:

$$\frac{C_t}{C_0} = \left(1 - \frac{m_s q_e}{C_0}\right) + \frac{m_s q_e}{C_0} e^{-k_1 t}, \tag{5}$$

where C_t is dye concentration at time t.

The pseudo-second-order equation developed by Ho (1995) and Ho and Mckay (1999) can be expressed as

$$\frac{dq_t}{dt} = k_2(q_e - q_t)^2, \tag{6}$$

where k_2 is the rate constant of pseudo-second-order model. When performing as the previous model, the following equation is obtained:

$$\frac{C_t}{C_0} = 1 - \frac{m_s q_e}{C_0} \frac{q_e k_2 t}{1 + q_e k_2 t}. \tag{7}$$

The above Eqs. (5) and (7) were used for the curve fitting of the experimental results employing software Origin.

Fig. 6 shows the effect of initial MB concentration on the adsorption rate. It can be seen that an increase in the initial concentration leads to a faster decrease of the MB concentration in the solution. The rate of all the adsorption decreases with time until it gradually reaches a plateau. The higher the concentration, the faster the adsorption capacity reaches and the plateau forms. This may attribute to the gradual decrease in the concentration driving force with time.

Table 2
Langmuir and Freundlich constants for the adsorption of MB

Temperature (°C)	Langmuir model			Freundlich model		
	$Q \times 10^6$ (mol/g)	$b \times 10^{-6}$ (L/mol)	R^2	$K \times 10^6$ (mol/g)	$1/n$	R^2
15	2.1344	0.7602	0.9575	1.1979	0.1788	0.9655
35	2.1885	0.9310	0.9094	1.3026	0.1640	0.9663
55	2.2217	1.1193	0.8573	1.3757	0.1550	0.9818

Table 3
Comparison of adsorption capacities of methylene blue on various fly ashes

Adsorbent	Temp. (°C)	Adsorption capacity (m mol/g)	Reference
Sludge ash	24	0.005	Weng and Pan (2006)
Bagasse fly ash	30	0.0202	Gupta et al. (2000)
CFA	25	0.0144	Viraraghavan and Ramakrishna (1999)
CFA	22	0.0189	Janos et al. (2003)
CFA	22	0.0046	Janos et al. (2003)
CFA treated with HNO_3	30	0.025	Wang et al. (2005b)
CFA heated at 800 °C	30	0.014	Wang et al. (2005b)
CFA after sonochemical treatment with NaOH	–	0.016–0.040	Wang and Zhu (2005)
CFA treated with H_2SO_4	15	0.0021	This study

CFA: coal fly ash.

Fig. 7 displays the effect of temperature on the dye adsorption. As shown, the rate of adsorption increases with increasing temperature. Increase of temperature is known to increase the rate of diffusion of the adsorbate molecules across the external boundary layer and in the internal pores of the adsorbent particle, owing to the decrease in the viscosity of the solution (Al-Qodah, 2000).

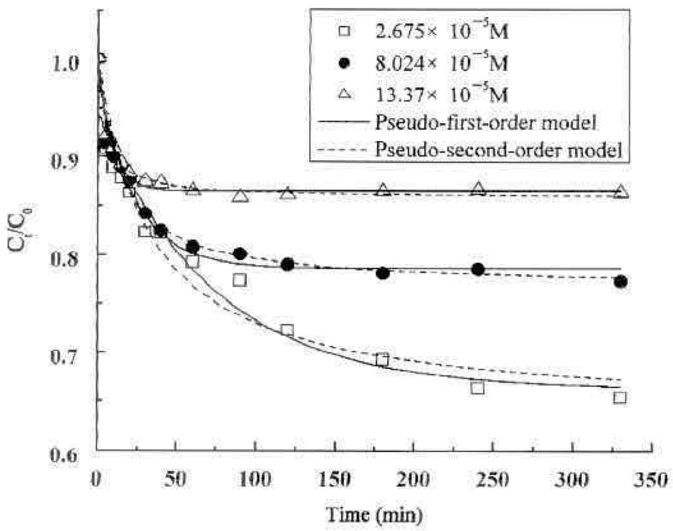


Fig. 6. Effect of initial concentration on the adsorption. Temperature: 15 °C.

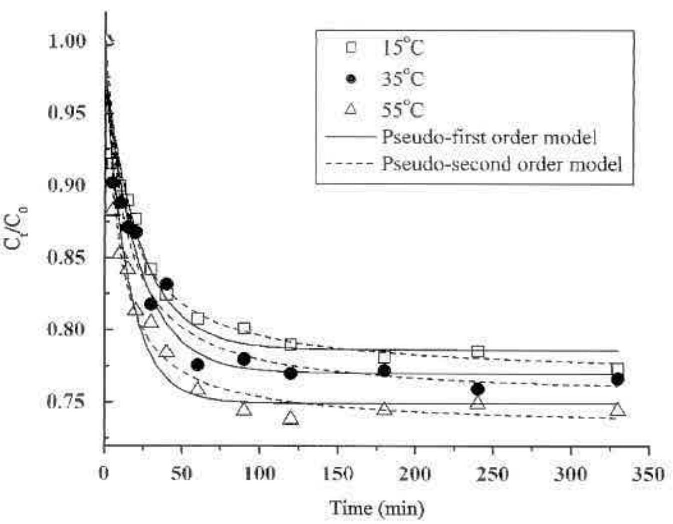


Fig. 7. Effect of temperature on the adsorption. Initial MB concentration: 8.024e−5 M.

The kinetic parameters obtained from the nonlinear regression are listed in Tables 4 and 5. It can be seen that the experimental adsorption capacity ($q_{e,exp}$) is in good agreement with the calculated capacity ($q_{e,cal}$). The adsorption rate parameters (k_1 and k_2) and the adsorption capacity ($q_{e,exp}$ and $q_{e,cal}$) increase with the increasing dye concentration and temperature. Moreover, it shows that the correlation coefficients (R^2) of the pseudo-second-order model are higher than that of the pseudo-first-order model, which indicates a pseudo-second-order mechanism for the adsorption of MB onto the fly ash.

3.5. *Thermodynamic parameters of adsorption*

As the pseudo-second-order kinetic model is better in describing the adsorption, the activation energy of adsorption, E, was obtained from the pseudo-second-order constant, k_2, by using the following Arrhenius type relationship (Ho and Mckay, 1999; Al-Ghouti et al., 2005):

$$k_2 = k_0 \exp\left(\frac{-E}{RT}\right), \tag{8}$$

where k_0 is the rate constant of adsorption, R is the gas constant (8.314 J/mol K) and T is the solution temperature.

The corresponding plot of the values of $\ln(k_2)$ against $1/T$ was drawn and then the data were regressed to obtain the activation energy. The physisorption processes usually have low activation energies (5–40 kJ/mol), while higher activation energies (40–800 kJ/mol) suggest chemisorption (Nollet et al., 2003). As given in Table 6, the value of E was found to be 5.42 kJ/mol. We can conclude that the adsorption of MB may be physical in nature considering that the value is in the typical activation energy range for physisorption.

The thermodynamic parameters, such as enthalpy (ΔH^0), entropy (ΔS^0) and free energy (ΔG^0) of activation, were determined by using the following equations (Khan et al., 1995; Shawabkeh and Tutunji, 2003):

$$\log \frac{q_t}{C_t} = \frac{\Delta S^0}{2.303R} - \frac{\Delta H^0}{2.303RT}. \tag{9}$$

Table 4
Parameters in the adsorption kinetics at various initial MB concentrations

C_0 (10^{-5} M)	$q_{e,exp}$ (10^{-6} mol/g)	Pseudo-first-order model			Pseudo-second-order model		
		$q_{e,cal}$ (10^{-6} mol/g)	k_1 (10^6 g/mol min)	R^2	$q_{e,cal}$ (10^{-6} mol/g)	k_2 (10^6 g/mol min)	R^2
2.675	1.209	0.903	0.0139	0.9532	0.963	0.0313	0.9540
8.024	1.704	1.685	0.0425	0.9674	1.862	0.0391	0.9823
13.37	1.944	1.738	0.1181	0.9714	1.893	0.1183	0.9888

Temperature: 15 °C.

Table 5
Parameters in the adsorption kinetics at various temperatures

T (°C)	$q_{e,exp}$ (10^{-6} mol/g)	Pseudo-first-order model			Pseudo-second-order model		
		$q_{e,cal}$ (10^{-6} mol/g)	k_1 (10^6 g/mol min)	R^2	$q_{e,cal}$ (10^{-6} mol/g)	k_2 (10^6 g/mol min)	R^2
15	1.704	1.685	0.0425	0.9674	1.862	0.0391	0.9823
35	1.755	1.846	0.0446	0.9507	1.987	0.0409	0.9654
55	1.817	2.006	0.0645	0.9580	2.144	0.0516	0.9836

Initial MB concentration: 8.024e−5 M.

Table 6
The thermodynamic parameters of MB adsorption onto fly ash treated by various chemicals

Treated by	E (kJ/mol)	ΔH^0 (kJ/mol)	ΔS^0 (J/mol K)	ΔG^0 (kJ/mol)			Reference
H_2O_2	–	−47.082	82.5	−22.245 (30 °C)	−22.286 (40 °C)	−20.643 (50 °C)	Gupta et al. (2000)
HNO_3	–	76.1	370.2	−38.4 (30 °C)	−37.4 (40 °C)		Wang et al. (2005b)
H_2SO_4	5.42	5.63	46.1	−7.65 (15 °C)	−8.57 (35 °C)	−9.50 (55 °C)	This study

$$\Delta G^0 = \Delta H^0 - T\Delta S^0. \tag{10}$$

The linear plot of the values of $\log(q_t/C_t)$ against $1/T$ revealed a straight line with regression coefficient (R) of 0.98466, intercept of 2.4065 and slope of −293.81. The values of ΔH^0, ΔS^0, and ΔG^0 were then obtained and the results are given in Table 6. The negative values of ΔG^0 reflect that the adsorption of MB onto fly ash is feasible and spontaneous. The ΔG^0 value decreases from −7.65 to −9.50 kJ/mol when the temperature increases from 15 to 55 °C, suggesting the more adsorbable of dye with increasing temperature.

Table 6 compares the thermodynamic parameters of adsorption of MB onto treated fly ash with various chemicals investigated by us and other researchers. It can be seen that the ΔH^0 value reported by Gupta et al. (2000) is negative (−47.082 kJ/mol), indicating the exothermic nature of the adsorption. However, the ΔH^0 values reported by Wang et al. (2005b) and in this study are 76.1 and 5.63 kJ/mol, respectively; the positive values suggest the endothermic nature of the adsorption. In addition, the low value of ΔH^0 in the present paper implies that there was loose bonding between the adsorbate molecules and the adsorbent surface (Singh, 2000).

4. Conclusion

The H_2SO_4 treatment of fly ash increases the surface area and pore volume, while decreases its average pore diameter. The adsorption capacity of the fly ash enhances with increasing pH value in basic solution, which may be explained by considering the pH_{zpc} of the fly ash and by the changes of MB in the basic solution. The isotherm study shows that the Freundlich model in fitting the adsorption isotherm is better than the Langmuir model. The kinetics study reveals that the adsorption behavior can be described using the pseudo-second-order model. Moreover, the positive value of enthalpy (ΔH^0) suggests an endothermic nature of the adsorption, and the low value of E (5.42 kJ/mol) indicates that the adsorption process might be mainly physical.

References

Al-Ghouti, M., Khraisheh, M.A.M., Allen, S., Ahmad, M.N.M., 2003. The removal of dyes from textile wastewater: a study of the physical characteristics and adsorption mechanisms of diatomaceous earth. Journal of Environmental Management 69, 229–238.

Al-Ghouti, M., Khraisheh, M.A.M., Ahmad, M.N.M., Allen, S., 2005. Thermodynamic behaviour and the effect of temperature on the removal of dyes from aqueous solution using modified diatomite: a kinetic study. Journal of Colloid and Interface Science 287, 6–13.

Al-Qodah, Z., 2000. Adsorption of dyes using shale oil ash. Water Research 34 (17), 4295–4303.

Eren, Z., Acar, F.N., 2006. Adsorption of reactive black 5 from an aqueous solution: equilibrium and kinetic studies. Desalination 194 (1–3), 1–10.

Fungaro, D.A., Izidoro, J.C., Almeida, R.S., 2005. Removal of toxics compounds from aqueous solution by adsorption with zeolite synthesized from coal ashes. Eclectica Quimica 30 (2), 31–35.

Gupta, V.K., Mohan, D., Sharma, S., Sharma, M., 2000. Removal of basic dyes (rhodamine B and methylene blue) from aqueous solutions using bagasse fly ash. Separation Science and Technology 35, 2097–2113.

Ho, Y.S., 1995. Adsorption of heavy metals from waste streams by peat. Ph.D. Thesis, University of Birmingham, Birmingham, UK.

Ho, Y.S., McKay, G., 1999. Comparative sorption kinetic studies of dye and aromatic compounds onto fly ash. Journal of Environmental Science Health A 34, 1179–1204.

Janos, P., Buchtova, H., Ryznarova, M., 2003. Sorption of dyes from aqueous solutions onto fly ash. Water Research 37, 4938–4944.

Khan, S.A., Rehman, R., Khan, M.A., 1995. Adsorption of Cr (III), Cr (VI) and Ag (I) on bentonite. Waste Management 15, 271–282.

Khraisheh, M.A.M., Al-Ghouti, M.A., Allen, S.J., Ahmad, M.N., 2005. Effect of OH and silanol groups in the removal of dyes from aqueous solution using diatomite. Water Research 39, 922–932.

Kumar, K.V., Ramamurthi, V., Sivanesan, S., 2005. Modeling the mechanism involved during the sorption of methylene blue onto fly ash. Journal of Colloid and Interface Science 284, 14–21.

Lagergren, S., 1898. About the theory of so-called adsorption of soluble substances. Kungliga Svenska Vetenskapsakademiens Handlingar 24 (4), 1–39.

Li, L., Wang, S., Zhu, Z., 2006. Geopolymeric adsorbents from fly ash for dye removal from aqueous solution. Journal of Colloid and Interface Science 300 (1), 52–59.

Lin, S.H., 1993. Adsorption of disperse dye by powdered activated carbon. Journal of Chemical Technology and Biotechnology 57, 387–391.

Mall, I.D., Srivastava, V.C., Agarwal, N.K., Mishra, I.M., 2005. Removal of Congo red from aqueous solution by bagasse fly ash and activated carbon: kinetic study and equilibrium isotherm analyses. Chemosphere 61 (4), 492–501.

Mall, I.D., Srivastava, V.C., Agarwal, N.K., 2006. Removal of orange-G and methyl violet dyes by adsorption onto bagasse fly ash—kinetic study and equilibrium isotherm analyses. Dyes and Pigments 69 (3), 210–223.

McKay, G., Ramprasad, G., Mowli, P., 1987. Desorption and regeneration of dye colours from low-cost materials. Water Research 21, 375–377.

Mohan, D., Singh, K.P., Singh, G., Kumar, K., 2002. Removal of dyes from wastewater using fly ash, a low-cost adsorbent. Industrial and Engineering Chemistry Research 41, 3688–3695.

Noh, J.S., Schwarz, J.A., 1989. Estimation of the point of zero charge of simple oxides by mass titration. Journal of Colloid and Interface Science 130 (1), 157–164.

Nollet, H., Roels, M., Lutgen, P., Van der Meeren, P., Verstraete, W., 2003. Removal of PCBs from wastewater using fly ash. Chemosphere 53, 655–665.

Ramakrishna, K.R., Viraraghavan, T., 1997. Dye removal using low cost adsorbents. Water Science and Technology 36, 189–196.

Shawabkeh, R.A., Tutunji, M.F., 2003. Experimental study and modeling of basic dye sorption by diatomaceous clay. Applied Clay Science 24, 111–120.

Singh, D., 2000. Studies of the adsorption thermodynamics of oxamyl on fly ash. Adsorption Science Technology 18, 741–748.

Talman, R.Y., Atun, G., 2006. Effects of cationic and anionic surfactants on the adsorption of toluidine blue onto fly ash. Colloids and Surfaces A: Physicochemical and Engineering Aspects 281 (1–3), 15–22.

Viraraghavan, T., Ramakrishna, K.R., 1999. Fly ash for colour removal from synthetic dye solutions. Water Quality Research Journal of Canada 34, 505–517.

Wang, S., Zhu, Z.H., 2005. Sonochemical treatment of fly ash for dye removal from wastewater. Journal of Hazardous Materials B 126, 91–95.

Wang, S., Boyjoo, Y., Choueib, A., 2005a. A comparative study of dye removal using fly ash treated by different methods. Chemosphere 60, 1401–1407.

Wang, S., Boyjoo, Y., Choueib, A., Zhu, Z.H., 2005b. Removal of dyes from aqueous solution using fly ash and red mud. Water Research 39, 129–138.

Wang, S., Soudi, M., Li, L., Zhu, Z.H., 2006. Coal ash conversion into effective adsorbents for removal of heavy metals and dyes from wastewater. Journal of Hazardous Materials 133 (1–3), 243–251.

Weng, C.H., Pan, Y.F., 2006. Adsorption characteristics of methylene blue from aqueous solution by sludge ash. Colloids and Surfaces A: Physicochemical and Engineering Aspects 274, 154–162.

Continued from outside back cover

Journal of
Environmental Management

April 2008 Volume 87 Number 1

Contents

Continued on inside back cover

Abstracted/indexed in Gale Group Business A.R.T.S., Periodical Abs Plustext, Zoological Record Online®, Science Citation Index, Science Citation Index Expanded, CC/Agriculture,-Biology & Environmental Sciences, BIOSIS Previews, Biological Abstracts, GEOBASE®, Pascal, Embase, Environmental Sciences, CAB Abstracts, Social Sciences Abstracts, Elsevier BIOBASE, CSA Life Sciences Abstracts, Uncover, Medline®, Pollution Abstracts, Agricola, Biosis Toxline®, CSA Technology Research Database, ToxFile. Also abstracted/indexed in the abstract and citation database SCOPUS®. Full text available on ScienceDirect®.

Note to contributors. Online submission is available for this journal; to submit your article online, go to http://ees.elsevier.com/jema.Instructions on how to prepare a manuscript to the journal can be found in the Guide for Authors on the journal homepage at http://www.elsevier.com/locate/jenvman.

This journal is printed on acid free paper

0301-4797(200804)87:1;1-Q