July 2008 Volume 88 Number 2

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 [toll free within the USA]; (+1) (407) 5636022 [outside the USA]; fax: (+1) (407) 3631354; e-mail: JournalCustomerService-usa@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

ELSEVIER

Journal of Environmental Management 88 (2008) 195–202

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Quantification of ammonia and hydrogen sulfide emitted from pig buildings in Korea

Ki Youn Kim[a,*], Han Jong Ko[b], Hyeon Tae Kim[b], Yoon Shin Kim[a], Young Man Roh[a], Cheol Min Lee[a], Chi Nyon Kim[c]

[a]Institute of Environmental and Industrial Medicine, Hanyang University, Seoul, Republic of Korea
[b]Graduate School of Agriculture, Kyoto University, Kyoto, Japan
[c]Institute for Occupational Health, College of Medicine, Yonsei University, Seoul, Republic of Korea

Received 17 February 2006; received in revised form 1 December 2006; accepted 8 February 2007
Available online 27 March 2007

Abstract

The aim of this field study was to determine the concentrations and emissions of ammonia and hydrogen sulfide in different types of pig buildings in Korea to allow objective comparison between pig housing types in Korea and other countries. Concentrations of ammonia and hydrogen sulfide in the pig buildings averaged 7.5 ppm and 286.5 ppb and ranged from 0.8 to 21.4 ppm and from 45.8 to 1235 ppb, respectively. The mean emissions of ammonia and hydrogen sulfide per pig (normalized to 75 kg liveweight) and area (m^2) from pig buildings were 250.2 and 37.8 mg/h/pig and 336.3 and 50.9 mg/h/m^2, respectively. Ammonia and hydrogen sulfide concentrations and emissions were higher in the pig buildings managed with deep-pit manure systems with slats and mechanical ventilation than in other housing types.

Keywords: Pig building; Ammonia; Hydrogen sulfide; Concentration; Emission

1. Introduction

Extensively modern pig production practices degrade the air quality in pig buildings and thereby pose a health risk to the pigs and farm workers (Whyte, 1993; Pearson and Sharples, 1995). To address these environmental problems, the U.S. and Europe are imposing more stringent regulations on aerial contaminants both within and from pig houses (Wathes et al., 1998; Gay et al., 2003).

Aerial contaminants generated in the pig buildings are classified into gases, particulates and airborne microorganisms (Wathes, 1994). Representative gaseous compounds in the pig building are ammonia and hydrogen sulfide which are generated by the pigs themselves or by decomposing manure and easily adsorbed onto airbourne dust particles derived from feedstuffs (Bottcher, 2001). If ammonia and hydrogen sulfide are adsorbed onto fine dust particles such as $PM_{2.5}$, they pose seriously adverse health effects on both pigs and workers because they penetrate the respiratory system (Donham et al., 1986). Ammonia in the atmosphere can combine with some other compounds to form $PM_{2.5}$, which has been shown to pose a health threat to humans (Erisman and Schaap, 2004). These gases promote a deterioration of equipment in the pig buildings (Ni et al., 2000b), elicit serious complaints from neighbors if the odors are emitted outdoors (Ni et al., 2000a), and can potentially damage ecosystems by soil acidification, water eutrophication and global warming (Harssema et al., 1981; van Breemen et al., 1982; Buijsman and Erisman, 1988).

For pigs, ammonia concentrations above 100 ppm have been shown to cause a decline in feed consumption and daily weight gain (Stombaugh et al., 1969) and concentrations in the range of 50–75 ppm lowered the ability of the pigs to clear bacteria from their lungs (Curtis et al., 1977). Humans experience eye, nose and throat irritation if exposed to 20–25 ppm for 8 h (NIOSH, 1974). Concentrations of hydrogen sulfide below 10 ppm have no adverse health effect on pig growth (Curtis et al., 1977), but concentrations of 50–100 ppm, 100–1000 ppm, and

*Corresponding author. Tel.: +82 2 2220 1510; fax: +82 2 2292 2510.
E-mail address: kkysnu5@daum.net (K.Y. Kim).

0301-4797/$ - see front matter © 2007 Published by Elsevier Ltd.
doi:10.1016/j.jenvman.2007.02.003

1000 ppm cause chronic, subacute, and acute intoxication, respectively (Smith et al., 1979). To protect workers from exposure to ammonia and hydrogen sulfide, the Korea Occupational Safety and Health Agency established 8-h time-weighted average exposure limit values of 25 and 10 ppm limits, respectively (KOSHA, 1999). The Korea Ministry of Environment restricts livestock facilities, including pig buildings, to maximum levels of 1 ppm for ammonia and 0.02 ppm for hydrogen sulfide, based on a mandate to lessen odor emissions (KME, 2004).

Prior to devising abatement techniques to prevent these environmental and hygiene problems caused by ammonia and hydrogen sulfide emitted from pig buildings, it is essential to accurately determine their concentrations. Most field research to quantify the concentrations and emissions of ammonia and hydrogen sulfide in pig buildings to date has been performed in other developed countries, mainly in the U.S. and E.U., as shown in Tables 1 and 2. According to these reports, the mean concentrations of ammonia and hydrogen sulfide were approximately 15 and 0.2 ppm, respectively. The ranges of emissions per pig (75 kg) and unit area (m^2) were 0.2–3000 and 0.3−4000 mg/h for ammonia and 0.5–70 and 0.6−100 mg/h, respectively, for hydrogen sulfide, indicating large variations for each gas between countries and pig housing types. These variations are in part attributed to differences in climatic conditions and the environmental management practices in the pig buildings.

To date, there is no such field data from pig buildings in Korea. Therefore, the aim of this study was to provide both pig producers and governmental regulators in Korea empirical information on the concentrations and emissions of ammonia and hydrogen sulfide in the different types of pig buildings, and to compare values from Korean pig buildings to those in other countries.

2. Materials and methods

2.1. Selection of pig buildings

The pig buildings investigated in this study were selected based on three criteria: manure removal system, ventilation mode, and growth stage of the pigs (Table 3).The three types of manure removal systems found in Korean pig buildings were (a) manure removal system by scraper, (b) a deep-litter bed system and (c) a deep-pit manure system. The manure removal system by scraper, called the Haglando system (Groenestein, 1993), consists of a shallow manure pit with scrapers under a fully slatted floor. The floor of the pit has a smooth finish and is covered with an epoxy coating, allowing the manure to be completely removed from the pig building several times a day. The deep-litter bed system is a housing system where pigs are kept on a 40 cm-thick layer of a mixture of manure and litter composed of sawdust, straw or woodshavings. The manure mixes with the litter, is fermented in the bed, dries up during the pigs' growing period, and is cleaned out once a month. This system has an advantage in reducing the labor in manure handling and rapid manure drying, while its disadvantage is that it generates a lot of dust, bioaerosols, and parasites. Both nitrification and denitrification occur in this system, which can prevent the emission of ammonia by producing N_2 from NH_4^+ instead of NH_3. When conditions for nitrification and denitrification are suboptimal, NO and N_2O, both volatile intermediates, can be emitted (Betlach and Tiedje, 1981; Lipshcultz et al., 1981). The deep-pit manure system, which has become popular in Korea in recent years, is composed of a deep manure pit under a fully or partially slatted floor. Manure stored in the pit for long periods is removed by pulling the pit plug, and letting the manure drain into a storage compartment located outside the pig building.

The ventilation in Korean pig buildings is either the mechanical type, using wall exhaust fans, or natural ventilation by operation of a winch-curtain. Confinement-style pig buildings are usually mechanically ventilated, and open pig buildings are usually naturally ventilated. Most pig buildings with the deep-litter bed system used natural ventilation, while both mechanical and natural ventilation were common in deep-pit and mechanical scrape pig buildings.

Only pig production sites housing growing/finishing pigs weighing approximately 50−100 kg were included in this study in order to compare objectively the emissions and concentrations among pig housing types.

2.2. Measurements

This field study was carried out from May to June and from September to October of 2002–2003. During this period, the mean ranges of outdoor temperature and relative humidity were 13−17 °C and 63−72%, respectively. Table 1 shows that research into the emissions of ammonia and hydrogen sulfide in pig buildings was done among five pig housing types with 30 sites visited for each housing type. The pig buildings investigated were randomly selected and situated in the central Korean provinces of Kyung-gi, Chung-buk and Chung-nam.

Concentrations of ammonia and hydrogen sulfide were measured at three locations in the central alleys of the pig building under moderate weather conditions. The 15−30 min samples were taken once every hour between 9 am and 5 pm to span the normal work day. Emission rates of ammonia and hydrogen sulfide were estimated by multiplying the average concentration (mg/m^3) measured near the air outlet by the mean ventilation rate (m^3/h) and expressed either on a per pig (75 kg liveweight) (mg/h/pig) or per area ($mg/h/m^2$) basis. The total weight of pigs in the pig building was calculated by multiplying the number of pigs by the averaged weight of one pig (75 kg) as estimated by the stockman. The total area of the pig building was calculated from direct measurements, or in cases where

Table 1
Literature review of ammonia concentrations and emissions in pig buildings in various countries

Country	Housing type	Ammonia concentration (ppm)		Reference
		Mean	Range	
U.S.	Slats	26.0		Zhang et al. (1998)
	Slats	–	0.2–19.0	Stowell and Foster (2000)
	Slats	7.4	0.3–32.1	Attwood et al. (1987)
	Slats	–	1.5–13.2	Crook et al. (1991)
	Slats	7.8	0.1–50.0	Pedersen (1992)
	Slats	18.0	–	Nicks et al. (1993)
E.U.	Slats	–	11.0–14.7	Hendriks et al. (1998)
	Slats	–	10.0–35.0	Hinz and Linke (1998)
	Slats	–	12.1–18.2	Koerkamp et al. (1998)
	Litter		4.3–9.1	
	Slats	18.2	11.7–26.0	Louhelainen et al. (2001)
Canada	Slats	22.9	6.5–64.9	Morrison et al. (1993)
	Slats	19.6	1.9–25.9	Duchaine et al. (2000)
Taiwan	Slats	3.2	1.9–4.2	Chang et al. (2001)
Mean		15.4	0.1–64.9	

Country	Housing type	Ammonia emission mg/h/[a]pig		mg/h/[b]m^2		Reference
		Mean	Range	Mean	Range	
	Slats	–	12.6–792.8	–	16.9–1,066	Heber et al. (1997)
	Slats	–	198.7–1,384	–	267.1–1,861	Ni et al. (1998)
	Slats	–	165.0–210.0		221.8–282.3	Ni et al. (2000a)
U.S.	Slats	–	216.8–3,037	–	291.4–4,081	Stowell and Foster (2000)
	Slats		0.2–0.8	–	0.3–1.1	Zhu et al. (2000)
	Slats	114.0	–	153.2	–	Heber et al. (2001)
	Slats	142.8	–	191.9	–	Gay et al. (2003)
	Litter	29.1	–	39.1	–	Kowalewsky (1981)
	Slats	288.5	–	387.8	–	Eerden et al. (1981)
	Slats	278.5	–	374.3	–	Oosthoek et al. (1991)
	Slats	29.6	–	39.8	–	Aarnink et al. (1995)
	Slats	102.7	–	138.0	–	Gastel et al. (1995)
	Slats	208.3	–	280.0	–	Aarnink et al. (1996)
E.U.	Slats	227.1	–	305.2		Aarnink et al. (1996)
	Slats	–	247.5–741.0		332.7–996.0	Hendriks et al. (1998)
	Slats	375.0	–	504.0	–	Hinz and Linke (1998)
	Slats	299.3	255.0–315.0	402.2	342.7–423.4	Koerkamp et al. (1998)
	Litter	251.0	180.0–465.0	337.4	241.9–625.0	
	Slats	660.0	–	887.1	–	Demmers et al. (1999)
Mean		231.2	0.2–3037	310.8	0.3–4081	

[a]Based on growing/finishing pig (75 kg).
[b]Assuming 0.75 m^2 of floor area per pig.

measurement was not possible, from estimations by the stockman.

The ventilation rate of the pig building was estimated using different methods depending on the type of ventilation used. According to Gay et al. (2003), the use of static pressure readings and fan curve data for mechanically ventilated pig buildings and tracer gas, heat balance, or carbon dioxide measurements for naturally ventilated pig buildings provide the best estimates of ventilation rate. Additionally, numerous factors that affect airflow, including diurnal pig activity, dust accumulation on fan shutters and blades, loose fan belts, and changes of static pressure in the pig building, are not accounted for by existing methods of ventilation rate estimation (Bicudo et al., 2002). This study estimated ventilation rates of the mechanically ventilated pig buildings by multiplying the area of the sidewall exhaust fans by the averaged value of air velocity measured five times across five points near exhaust fans

Table 2
Literature review of hydrogen sulfide concentrations and emissions in pig buildings in various countries

Country	Housing type	Hydrogen sulfide concentration (ppm)		Reference
		Mean	Range	
	Slats	90.0	–	Muehling (1970)
	Slats	624.7	120.0–2,174	Avery et al. (1975)
	Slats	166.0	–	Heber et al. (1997)
U.S.	Slats	380.0	–	Zhang et al. (1998)
	Slats	206.0	154.0–378.0	Ni et al. (2002)
E.U.	Slats	19.5	6.5–52.0	Louhelainen et al. (2001)
Taiwan	Slats	101.7	30.0–180.0	Chang et al. (2001)
Mean		226.8	6.5–2,174	

Country	Housing type	Hydrogen sulfide emission mg/h[a]/pig		mg/h/[b]m^2		Reference
		Mean	Range	Mean	Range	
	Slats	–	0.5–21.9	–	0.6–29.5	Heber et al. (1997)
	Slats	39.4	–	52.9	–	Ni et al. (1999)
U.S.	Slats	–	5.2–72.1	–	7.0–96.8	Ni et al. (2000b)
	Slats	–	12.3–47.9	–	16.5–64.4	Zhu et al. (2000)
Mean		39.4	0.5–72.1	52.9	0.6–96.8	

[a]Based on growing/finishing pig (75 kg).
[b]Assuming 0.75 m^2 of floor area per pig.

Table 3
Characteristics of the pig buildings investigated in this study

Housing type		Pig type	n
Manure collection system	Ventilation mode		
Deep-pit manure system with slats	Natural	Growing/Finishing	30
	Mechanical	Growing/Finishing	30
Manure removal system by scraper	Natural	Growing/Finishing	30
	Mechanical	Growing/Finishing	30
Deep-litter bed system	Natural	Growing/Finishing	30

using a hot-wire anemometer (Model 444, Kurz, Inc., Monterey, USA), once every 2 h during an 8 h sampling period. For naturally ventilated pig buildings, the CO_2 balance method described by Albright (1990) was applied to calculate grossly the ventilation rate based on literature estimates of carbon dioxide produced by pigs (van Ouwerkerk and Pedersen, 1994). This method estimates ventilation rate by the difference between indoor and outdoor CO_2 concentrations, as measured with a detector tube (No. 126B, Gastech Corp., Japan), yielding measurements in the range of 100–7000 ppm. Although it provides only rough values, several researchers applied this method to estimate the ventilation rate of naturally ventilated pig buildings (Koerkamp et al., 1998; Schauberger et al., 2000; Blanes and Pedersen, 2005).

Analyses of ammonia and hydrogen sulfide were performed according to the methods recommended by NIOSH (1998). Using an air sampling pump (Model 71G9, Gilian Instrument Corp., NJ, USA) adjusted to 2.0 l/min of flow rate, air was collected for 10–20 min into an all-glass impinger 30 (Ace Glass Inc., Vineland, USA) including the absorption fluid (10 ml, 0.1 N-H_2SO_4 solution) for ammonia and into charcoal tubes (Gilian Instrument Corp., NJ, USA) for hydrogen sulfide, respectively. After sampling, the absorption fluids and charcoal tubes were analyzed by Ion Chromatography (761 Compact IC, Metrohm, Switzerland).

2.3. Statistical analysis

The SAS package program (2002) was used to analyze experimental data to determine any significant differences in ammonia and hydrogen sulfide concentration between the variables in the selection criteria. Duncan's multiple range tests were performed to indicate significant differences among group means in ANOVA.

3. Results

3.1. Concentrations of ammonia emissions

Mean concentrations and emissions of ammonia from different pig housing types in Korea are given in

Table 4
Ammonia concentrations and emissions in various types of Korean pig houses

Housing		Ammonia concentration (ppm)			
Manure collection system	Ventilation mode	Mean*	Range		
Deep-pit manure system with slats	[a]N.V.	6.9[a]	2.1–10.2		
	[b]M.V.	12.1 [b]	7.3–21.4		
Manure removal system by scraper	N.V.	5.1[a]	3.1–9.5		
	M.V.	11.4[b]	8.1–15.2		
Deep-litter bed system	N.V.	2.2[c]	0.8–5.1		
Mean		7.5	0.8–21.4		
Housing		Ammonia emission mg/h[c]/pig		mg/h/m²	
Manure collection system	Ventilation mode	Mean*	Range	Mean*	Range
Deep-pit manure system with slats	N.V.	284.1[a]	52.5–482.1	381.9[a]	81.1–514.3
	M.V.	320.1[a]	24.2–826.5	430.3[a]	113.8–1,068
Manure removal system by scraper	N.V.	263.5[a]	38.2–524.2	354.1[a]	123.6–678.2
	M.V.	298.3[a]	83.3–725.6	400.9[a]	213.4–820.4
Deep-litter bed system	N.V.	84.9[b]	8.2–210.1	114.1[b]	23.3–352.3
Mean		250.2	8.2–826.5	336.3	23.3–1,068

[a]N.V. : Natural ventilation.
[b]M.V. : Mechanical ventilation.
[c]Based on growing/finishing pig (75 kg).
*a,b,c,d means that averaged values within the row by the same letter do not differ significantly.

Table 4. Regardless of the type of manure collection system, the ammonia concentrations in mechanically ventilated pig building were significantly higher than those in naturally ventilated pig buildings ($p<0.05$). Of the pig housing types examined, the pig buildings with a deep-litter bed system showed the lowest ammonia concentration ($p<0.05$). However, the ammonia emissions did not differ significantly among the other types of pig buildings.

3.2. Concentration of hydrogen sulfide emissions

Table 5 presents the mean concentrations and emissions of hydrogen sulfide from different pig housing types in Korea. The highest levels of hydrogen sulfide were found in the mechanically ventilated pig buildings with slats ($p<0.05$), while the naturally ventilated pig buildings equipped with manure removal system by scraper and the pig buildings equipped with deep-litter bed system had the lowest values ($p<0.05$). There were no significant differences in hydrogen sulfide concentrations between the naturally ventilated pig buildings with slats and the mechanically ventilated pig buildings with scrapers ($p>0.05$). The lowest levels of hydrogen sulfide emissions were seen in pig buildings with the deep-litter bed system ($p<0.05$), and did not differ significantly among the other types of pig buildings ($p>0.05$).

4. Discussion

The concentrations of ammonia and hydrogen sulfide were generally higher in the deep-pit manure system with slats and mechanical ventilation than in other pig housing types. This is consistent with previous reports (Muck and Steenhuis, 1982; Swierstra et al., 1995; Aarnink et al., 1996) that concentrations of gaseous contaminants, especially ammonia, were highest in pig buildings with a deep-pit manure system with slats that did not completely cover the surface of the slurry pit. The high level of ammonia can also be explained by nitrification and denitrification in the bed, which can prevent ammonia emission by producing N_2 from NH_4^+ instead of NH_3 (Groenestein, 1993). Based on the experimental results, the concentrations of ammonia and hydrogen sulfide are dominant in the ventilation mode, rather than in the manure collection system. There were no significant differences in ammonia and hydrogen sulfide emissions among the pig housing types except in the naturally ventilated pig building with a deep-litter bed system, which were lower likely due to increased airflow compared to mechanical ventilated systems. Considering the moderate climatic conditions experienced through the study period, this assumption was validated by previous findings (Albright, 1990).

Large variations in ammonia and hydrogen sulfide emissions among pig housing types were observed in this field survey. Considerable variations in the data were also found in previous studies (Koerkamp et al., 1998; Gay et al., 2003). There are two possible reasons for this. First, there were variations in ambient air temperature and relative humidity which would strongly affect gas volatilization levels in pig buildings (Heber and Stroik, 1988; Korthals et al., 1988; Olesen and Sommer, 1993; Arogo et al., 2002). Secondly, several very different and imprecise

Table 5
Hydrogen sulfide concentrations and emissions in various types of Korean pig houses

Housing		Hydrogen sulfide concentration (ppm)			
Manure collection system	Ventilation mode	Mean*	Range		
Deep-pit manure system with slats	[a]N.V.	296.3[a]	74.2–672.4		
	[b]M.V.	612.8[b]	121.6–1235		
Manure removal system by scraper	N.V.	115.2[c]	46.8–313.1		
	M.V.	270.3[a]	86.9–912.5		
Deep-litter bed system	N.V.	137.8[c]	45.8–289.2		
Mean		286.5	45.8–1235		
Housing		Hydrogen sulfide emission mg/h/[c]pig		mg/h/m^2	
Manure collection system	Ventilation mode	Mean*	Range	Mean*	Range
Deep-pit manure system with slats	N.V.	42.1[a]	8.0–120.2	56.5[a]	19.2–93.2
	M.V.	53.4[a]	18.6–192.5	71.7[a]	24.2–224.3
Manure removal system by scraper	N.V.	36.0[a]	6.2–92.2	48.4[a]	11.2–108.2
	M.V.	39.2[a]	10.2–112.3	52.7[a]	16.3–186.3
Deep-litter bed system	N.V.	18.5[b]	8.3–73.3	24.9[b]	6.3–64.3
Mean		37.8	6.2–192.5	50.9	6.3–224.3

[a]N.V.: Natural ventilation.
[b]M.V.: Mechanical ventilation.
[c]Based on growing/finishing pig (75 kg).
*a,b,c,d means that averaged values within the row by the same letter do not differ significantly.

Table 6
Comparison of ammonia and hydrogen sulfide concentrations and emissions between this study and earlier reports

Aerial contaminants	Unit		Reported data		Data in Korea	
			Mean	Range	Mean	Range
Ammonia	Concentration	ppm	15.4	0.1–64.9	7.5	0.8–21.4
	Emission	mg/h/[a]pig	231.2	0.2–3038	250.2	8.2–826.5
		mg/h/m^2	310.8	0.3–4082	336.3	23.3–1068
Hydrogen sulfide	Concentration	ppb	226.8	6.5–2174	286.5	45.8–1235
	Emission	mg/h/[a]pig	39.4	0.5–72.1	37.8	6.2–192.5
		mg/h/m^2	52.9	0.6–96.8	50.9	6.3–224.3

[a]Based on growing/finishing pig (75 kg).

methods were used to estimate the ventilation rate. These methods provide only rough estimates of house airflow, which in turn contributes to the appreciable variation in the emissions data (Bicudo et al., 2002). Emissions of ammonia and hydrogen sulfide are estimated by multiplying the concentrations of aerial pollutants near the air outlet by the ventilation rate. While concentrations of ammonia and hydrogen sulfide in the pig buildings can be measured fairly precisely, it is difficult to estimate accurately the ventilation rate of the pig buildings. In addition, differences in housing management practices and animal diets adopted by the farms likely contributed to the variations in emissions data.

Compared to pig buildings in the U.S. and E.U. (Table 6), concentrations of emissions in Korean pig buildings were generally lower for ammonia but higher for hydrogen sulfide. The most likely explanations for these differences are erroneous sampling techniques and ventilation rate estimates. Differences between sampling sites and periods, weather conditions during sampling, analytical techniques for assessing ammonia and hydrogen sulfide concentrations, estimation methods for ventilation rates, and differences in swine management practices in Korea compared to U.S. and Europe could all contribute to variations in the final values.

The field survey data reported in this article will help pig producers in Korea by providing starting values for emissions from which reduction strategies for ammonia and hydrogen sulfide can be devised and executed. For regulatory agencies, this data could be utilized to establish pertinent regulations objectively and develop improved setback evaluation methods, models, and tools to promote the health of workers, improve the welfare of confined pigs, and protect the environment through the reduction of ammonia and hydrogen sulfide emitted from the pig buildings.

5. Conclusions

Concentrations and emissions of ammonia and hydrogen sulfide in pig buildings were higher in the pig buildings managed with deep-pit manure systems with slats and mechanical ventilation than in other pig housing types. Compared to other countries, concentrations and emissions in pig buildings in Korea were generally lower for ammonia and higher for hydrogen sulfide, respectively. To present objective and accurate data for concentrations and emissions of ammonia and hydrogen sulfide in Korean pig buildings, additional farms should be investigated and an accurate statistical database for the pig housing types in Korea should be established.

References

Aarnink, A.J.A., Keen, A., Metz, J.H.M., Speelman, L., Verstegen, M.W.A., 1995. Ammonia emission patterns during the growing periods of pigs housed on partially slatted floors. Journal of Agricultural Engineering Research 62, 105–116.

Aarnink, A.J.A., van der Berg, A.J., Keen, A., Hoeksman, P., Verstegen, M.W.A., 1996. Effect of slatted floor area on ammonia emission and on the excretory and lying behavior of growing pigs. Journal of Agricultural Engineering Research 64, 299–310.

Aarnink, A.J.A., van der Berg, A.J., Keen, A., Hoeksman, P., Verstegen, M.W.A., 1996. Effect of type of slatted floor and degree of fouling of solid floor on ammonia emission rates from fattening piggeries. Journal of Agricultural Engineering Research 64, 299–310.

Albright, L.D. 1990. Environmental Control for Animals and Plants. St. Joseph, Mich. ASAE.

Arogo, J., Westerman, P.W., Heber, A.J., Robarge, W.P., Classen, J.J., 2002. Ammonia emissions from animal feeding operations. National Center for Manure and Animal Waste Management White Papers. North Carolina State University, Raleigh, NC.

Attwood, P., Ruigewaard, R., Versloot, P., DeWit, R., Heederik, D., Boleij, J.S.M., 1987. A study of the relationship between airborne contaminants and environment factors in Dutch swine confinement buildings. American Industrial Hygiene Association Journal 48, 745–751.

Avery, G.L., Merva, G.E., Gerrish, J.B., 1975. Hydrogen sulfide production in swine confinement units. Transactions of the ASAE 18, 149–151.

Betlach, M.R., Tiedje, J.M., 1981. Kinetic explanation for accumulation of nitrite, nitric oxide and nitrous oxide during bacterial denitrification. Applied and Environmental Microbiology 42, 1074–1084.

Bicudo, J.R., Schmidt, D.R., Gay, S.W., Gates, R.S., Jacobson, L.D., Hoff, S.J., 2002. Air quality and emissions from livestock and poultry production waste management systems. National Center for Manure and Animal Waste Management White Papers. North Carolina State University, Raleigh, NC.

Blanes, V., Pedersen, S., 2005. Ventilation flow in pig houses measured and calculated by carbon dioxide, moisture and heat balance equations. Biosystems Engineering 92, 483–493.

Bottcher, R.W., 2001. An environmental nuisance: odor concentrated and transported by dust. Chemical Senses 26, 327–331.

Buijsman, E., Erisman, J.W., 1988. Wet deposition of ammonium in Europe. Journal of Atmospheric Chemistry 6, 265–280.

Chang, C.W., Chung, H., Huang, C.F., Su, H.J.J., 2001. Exposure assessment to airborne endotoxin, dust, ammonia, hydrogen sulfide and carbon dioxide in open style swine houses. Annals of Occupational Hygiene 45, 457–465.

Crook, B., Robertson, J.F., Glass, S.A., Botheroyd, E.M., Lacey, J., Topping, M.D., 1991. Airborne dust, ammonia, microorganisms, and antigens in pig confinement houses and the respiratory health of exposed farm workers. American Industrial Hygiene Association Journal 52, 271–279.

Curtis, S.E., Drummond, J.G., Simon, J., 1977. Atmospheric ammoniaaffects swine health. Illinois Research 19, 8–9.

Demmers, T.G.M., Burgess, L.R., Short, J.L., Phillips, V.R., Clark, J.A., Wathes, C.M., 1999. Ammonia emissions from two mechanically ventilated UK livestock buildings. Atmospheric Environment 33, 217–227.

Donham, K.J., Scallon, L.J., Popendorf, W., 1986. Characterization of dusts collected from swine confinement buildings. American Industrial Hygiene Association Journal 47, 404–410.

Duchaine, C., Grimard, Y., Cormier, Y., 2000. Influence of building maintenance, environmental factors, and seasons on airborne contaminants of swine confinement buildings. American Industrial Hygiene Association Journal 61, 56–63.

Eerden, L.J., Harssema, H., Klarenbeek, J.V. 1981. Air from livestock buildings and plants. The relationship between the herd size and the risk of damage to plants growing around intensive cattle units. IPO-Rapport 254, Wageningen, Netherlands.

Erisman, J.W., Schaap, M., 2004. The need for ammonia abatement with respect to secondary PM reductions in Europe. Environmental Pollution 129, 159–163.

Gastel, J.P.B.F., Verdoes, N., Voermans, J.A.M., 1995. Processing pig slurry on farm scale to lower the ammonia emission and to reduce the volume of the slurry. In: Proceedings of the 7th International Symposium on Agricultural and Food Processing Wastes, Chicago, IL. June 18-20, pp. 599–608.

Gay, S.W., Schmidt, D.R., Clanton, C.J., Janni, K.A., Jacobson, L.D., Weisberg, S., 2003. Odor, total reduced sulfur, and ammonia emissions from animal housing facilities and manure storage units in Minnesota. Applied Engineering in Agriculture 19, 347–360.

Groenestein, C.M., 1993. Animal waste management and emission of ammonia from livestock housing systems: Field studies. Proceedings of International Livestock Environment Symposium IV, Warwick, England, 6–9 July. ASAE, pp. 1169–1175.

Harssema, H., van Eerden, W.J., Klarenbeek, J.V., 1981. The relation between intensity of husbandry farming and damage of vegetation. IMAG Report 32, Wageningen, The Netherlands.

Heber, A.J. Stroik, M., 1988. Influence of environmental factors on dust characteristics in swine finishing houses. Proceedings of International Livestock Environment Symposium III, Ontario, Canada, 25–27 April. ASAE, pp. 291–298.

Heber, A.J., Duggirala, R.K., Ni, J.Q., Spence, M.L., Haymore, B.L., Adamchuck, V.I., Bundy, D.S., Sutton, A.L., Kelly, D.T., Keener, K.M., 1997. Manure treatment to reduce gas emissions from large swine houses. In: Voermans, J.A.M. Monteny, G.R. (Eds.), Proceedings of the International Symposium on Ammonia and Odour Control from Animal Production Facilities, vol. 2, Dutch Society of Agricultural Engineering (NVTL), The Netherlands, pp. 449–457.

Heber, A.J., Lim, T.T., Ni, J.Q., Kendall, D., Richert, B., Sutton, A.L. 2001. Odor, ammonia and hydrogen sulfide emission factors for grow-finishing buildings. Final Report, National Pork Producers Council, Clive, IA.

Hendriks, J., Berckmans, D., Vinckler, C., 1998. Field tests of bio-additives to reduce ammonia emission from and ammonia concentration in pig houses. ASHRAE Transactions 104, 1699–1705.

Hinz, T., Linke, S., 1998. A comprehensive experimental study of aerial pollutants in and emissions from livestock buildings. Part 2: results. Journal of Agricultural Engineering Research 70, 119–129.

KME, 2004. The Act of Odor Prevention. Korea Ministry of Environment.

Koerkamp, P.W.G., Metz, J.H.M., Uenk, G.H., Phillips, V.R., Holden, M.R., Sneath, R.W., Short, J.L., White, R.P., Hartung, J., Seedorf, J., Schroder, M., Linkert, K.H., Pederson, S., Takai, H., Johnsen, J.O., Wathes, C.M., 1998. Concentrations and emissions of ammonia in livestock buildings in Northern Europe. Journal of Agricultural Engineering Research 70, 79–95.

Korthals, R.L., Christianson, L.L., Muehling, A.J., 1988. Assessing and improving environmental conditions in commercial swine facilities. Proceedings of International Livestock Environment Symposium III, Ontario, Canada, 25–27 April. ASAE.

KOSHA, 2001. Health effects discussion and determination of final PEL. Korea Occupational Safety and Health Agency, Korea Ministry of Labor.

Kowalewsky, H.H., 1981. Measuring and evaluating odour emissions. KTBL Schrift 216, Landwirtschaftsverlag GmbH, Munster, Germany.

Lipshcultz, F., Zafiriou, O.C., Wofsy, S.C., McElroy, M.B., Valois, F.W., Watson, S.W., 1981. Production of NO and N_2O by soil nitrifying bacteria. Nature 294, 641–643.

Louhelainen, K., Kangas, J., Veijanen, A., Viilos, P., 2001. Effect of in situ composting on reducing offensive odors and volatile organic compounds in swineries. American Industrial Hygiene Association Journal 62, 159–167.

Morrison, W.D., Pirie, P.D., Perkins, S., Braithwaite, L.A., Smith, J.H.,Waterfall, D., Doucett, C.M., 1993. Gases and respirable dust in confinement buildings and the response of animals to such airborne contaminants. Proceedings of International Livestock Environment Symposium IV, Warwick, England, 6–9 July. ASAE, pp. 735–746.

Muck, R.E., Steenhuis, T.S., 1982. Nitrogen losses from manure storages. Agricultural Wastes 4, 41–54.

Muehling, A.J., 1970. Gases and odors from stored swine wastes. Journal of Animal Science 30, 526–530.

Ni, JQ., Heber, A.J., Lim, T.T., Duggirala, R.K., Haymore, B.L., Diehl, C.A., Sutton, A.L., 1998. Hydrogen sulfide emissions from a large mechanically-ventilated swine building during warm weather. ASAE Paper No. 984051. St. Joseph, Mich. ASAE.

Ni, JQ., Heber, A.J., Lim, T.T., Duggirala, R.K., Haymore, B.L., Diehl, C.A., 1999. Continuous measurement of hydrogen sulfide emission from two large swine finishing buildings. ASAE Paper No. 994132, St. Joseph, MI 49085. ASAE.

Ni, J.Q., Heber, A.J., Diehl, C.A., Lim, T.T., 2000a. Ammonia, hydrogen sulfide and carbon dioxide release from pig manure in under-floor deep pits. Journal of Agricultural Engineering Research 77, 53–66.

Ni, J.Q., Heber, A.J., Lim, T.T., Diehl, C.A., Duggirala, R.K., Haymore, B.L., Sutton, A.L., 2000b. Ammonia emission from a large mechanically ventilated swine building during warm weather. Journal of Environmental Quality 29, 751–758.

Ni, J.Q., Heber, A.J., Diehl, C.A., Lim, T.T., Duggirala, R.K., Haymore, B.L., 2002. Characteristics of hydrogen sulfide concentrations in mechanically ventilated swine buildings. Canadian Biosystems Engineering 44, 611–619.

Nicks, B., Marlier, D., Canart, B., 1993. Air pollution levels in pig houses. Proceedings of International Livestock Environment Symposium IV, Warwick, England, 6–9 July. ASAE, pp. 635–641.

NIOSH (National Institute for Occupational Safety and Health), 1974. Criteria for a recommended standard — Occupational Exposure to Ammonia.DHEW(NIOSH) Pub. No. 74-136; NTIS No. PB-246-699/A06. National Technical Information Service, Springfield, VA.

NIOSH (National Institute for Occupational Safety and Health), 1998. Sampling and characterization of bioaerosols. In: Paul, A.J., Schafer, M.P. (Eds.), Manual of Analytical Method. Cincinnati, Ohio, USA.

Olesen, J.E., Sommer, S.G., 1993. Modelling effects of wind speed and surface cover on ammonia volatilization from stored pig slurry. Atmospheric Environment 27, 2567–2574.

Oosthoek, J., Kroodsma, W., Hoeksma, P., 1991. Ammonia emission from dairy and pig housing systems. In: Nielsen, V.C., Voorburg, J.H., Hermite (Eds.), Odour and Ammonia Emissions from Livestock Farming. Elsevier Applied Science, New York, pp. 31–41.

Pearson, C.C., Sharples, T.J., 1995. Airborne dust concentrations in livestock buildings and the effect of feed. Journal of Agricultural Engineering Research 60, 145–154.

Pedersen, S., 1992. Dust and gases. In: Christiaens, J.P.A., Sallvik, K. (Eds.), Climatization of Animal Houses: 2nd Report of CIGR Section II Working Group, second revised ed. Gent, Belgium: State University of Gent, Gent, Belgium, pp. 111–147.

SAS 2002. SAS/STAT User's Guide, Rel. 6.03. Cary, N.C.:SAS Institute, Inc.

Schauberger, G., Piringer, M., Petz, E., 2000. Steady-state balance model to calculate the indoor climate of livestock buildings, demonstrated for finishing pigs. International Journal of Biometeorology 43, 154–162.

Smith, R.P., Cooper, R.C., Engen, T., Hendrickson, E.R., Katz, M., Milby, T.H., Mudd, J.B., Rossano, A.T., Remond, J., 1979. Hydrogen sulfide. University Park Press, Baltimore.

Stombaugh, D.P., Teague, H.S., Roller, W.L., 1969. Effects of atmospheric ammonia on the pig. Journal of Animal Science 28, 844–847.

Stowell, R.R., Foster, S., 2000. Ammonia emissions from a high-rise swine finishing facility. ASAE Paper No. 00-4080. St. Joseph, Mich. ASAE.

Swierstra, D., Smits, M.C.J., Kroodsma, W., 1995. Ammonia emission from cubicle houses for cattle with solid floors. Journal of Agricultural Engineering Research 62, 127–132.

van Breemen, N., Burrough, P.A., Velthorst, E.J., van Dobben, H.F., de Wit, T., Ridder, T.B., Reijnders, H.F.R., 1982. Soil acidification from atmospheric ammonium sulphate in forest canopy throughfall. Nature 299, 548–550.

van Ouwerkerk, E.N.J., Pedersen, S., 1994. Application of the carbon dioxide mass balance method to evaluate ventilation rates in livestock buildings. Proceedings XII CIGR World Congress on Agricultural Engineering, Milan, vol. 1, pp. 516–529.

Wathes, C.M., 1994. Air and surface hygiene. In: Wathes, C.M., Charles, D.R. (Eds.), Livestock Housing CAB International, Wallingford, pp. 123–148.

Wathes, C.M., Jones, C.D.R., Webster, A.J.F., 1998. Ventilation, air hygiene and animal health. Veterinary Record 10, 554–559.

Whyte, R.T., 1993. Aerial pollutants and the health of poultry farmers. World's Poultry Science Journal 49, 139–156.

Zhang, Y., Tanaka, A., Dosman, J.A., Senthilselvan, A., Barber, E.M., Kirychuk, S.P., Holfeld, L.E., Hurst, T.S., 1998. Acute respiratory responses of human subjects to air quality in a swine building. Journal of Agricultural Engineering Research 70, 367–373.

Zhu, T., Pattey, E., Deesjardins, R.L., 2000. Relaxed eddy-accumulation technique for measuring ammonia volatilization. Environmental Science and Technology 34, 199–203.

ELSEVIER

Journal of Environmental Management 88 (2008) 203–210

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Risk aversion and compliance in markets for pollution control

John K. Stranlund*

Department of Resource Economics, University of Massachusetts-Amherst, 214 Stockbridge Hall, 80 Campus Center Way, Amherst, MA 01003, USA

Received 12 September 2006; received in revised form 26 January 2007; accepted 17 February 2007
Available online 30 April 2007

Abstract

This paper examines the effects of risk aversion on compliance choices in markets for pollution control. A firm's decision to be compliant or not is independent of its manager's risk preference. However, non-compliant firms with risk-averse managers will have lower violations than otherwise identical firms with risk-neutral managers. The violations of non-compliant firms with risk-averse managers are independent of differences in their profit functions and their initial allocations of permits if and only if their managers' utility functions exhibit constant absolute risk aversion. However, firm-level characteristics do impact violation choices when managers have coefficients of absolute risk aversion that are increasing or decreasing in profit levels. Finally, in the equilibrium of a market for emissions rights with widespread non-compliance, risk aversion is associated with higher permit prices, better environmental quality, and lower aggregate violations.

Keywords: Emissions trading; Compliance; Enforcement; Pollution markets; Risk aversion

1. Introduction

One of the most important design elements of any regulatory policy is how compliance with the policy will be enforced. Within the context of designing market-based pollution control policies, several authors have provided theoretical analyses of compliance incentives, the consequences of non-compliance, and the design of enforcement strategies (e.g., Keeler, 1991; Malik, 1990, 1992, 2002; van Egteren and Weber, 1996; Stranlund and Dhanda, 1999; Stranlund and Chavez, 2000; Chavez and Stranlund, 2003; Stranlund et al., 2005). Taken as a whole, this literature suggests that firms' incentives toward non-compliance under market-based regulations, as well as the design of enforcement strategies to counteract these incentives, are quite different from compliance and enforcement of other policy instruments, particularly command-and-control regulations.

An important question for enforcers of environmental policies is whether differences in the characteristics of firms generate different compliance choices. One may suspect that firms with different production processes, abatement technologies, or initial allocations of emissions rights may have different compliance incentives. If this is true, then regulators will be motivated to choose a targeted enforcement strategy, in particular targeted monitoring effort, which is conditioned on firm-level characteristics.

However, Stranlund and Dhanda (1999) have shown that the individual compliance choices of risk-neutral competitive firms in emissions trading programs are independent of differences in any firm-level characteristic. Consequently, regulators have no reason to condition their enforcement effort on firm-level characteristics. Their reasoning is straightforward. Since compliance in emissions trading programs means that a firm holds enough permits to cover its emissions, a risk-neutral competitive firm's marginal benefit of non-compliance is what it has to spend for permits to make sure it is compliant; that is, the prevailing permit price. A firm's compliance decision is made by comparing this permit price with the expected marginal penalty for emissions in excess of permits. Since this marginal benefit–cost comparison does not depend on anything unique to a particular firm, the compliance decision is independent of any firm characteristic.

*Tel.: +1 413 545 6328; fax: +1 413 545 5853.
E-mail address: stranlund@resecon.umass.edu.

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.02.012

This independence result contrasts sharply with the effects of firm-level characteristics on compliance with command-and-control standards, under which firms with higher marginal emissions control costs or that face stricter emissions standards will have a greater incentive to be non-compliant. In this way, firm-level characteristics are important determinants of compliance with fixed standards (Garvie and Keeler, 1994). A recent paper by Gray and Shadbegian (2005) finds strong support for this conclusion in their analysis of compliance behavior by pulp and paper manufacturers.

The independence of firms' violation choices on their individual characteristics under market-based regulations clearly depends, at least in part, on the assumption of risk-neutrality. However, no one has addressed the question of whether this independence result holds with risk-averse decision makers. In the theoretical literature on compliance and enforcement of emissions trading only Malik (1990) allows for non-neutral risk preferences. However, he does not provide the qualitative impacts of parametric differences among firms on their violation choices. Doing so is the primary objective of this paper. Moreover, since market-based regulations are unique in that the compliance decisions of firms are linked together through the market for property rights, it is important to understand how risk aversion affects equilibrium decisions and market outcomes.[1]

Several new results about compliance behavior under tradable property rights policies are derived in this paper. After laying out a model of a firm's compliance decisions under a competitive emissions trading program in the next section, Section 3 contains an analysis of the effects of risk aversion on compliance behavior. A firm's compliance decision can be thought of as two distinct decisions. First, a firm must decide whether or not it will comply with its permits. Second, if it chooses to be non-compliant it must choose the level of its violation. With regard to the first decision, I demonstrate that a firm's decision about whether to comply or not is independent of its manager's risk preference. Thus, if the enforcement objective is to achieve full compliance with an emissions trading program, the distribution of risk preferences among the managers of firms has no bearing on the strategy required to achieve this objective. Moreover, a targeted enforcement strategy with which firms with certain characteristics are monitored more closely than others is not justified.

Risk preferences do play a role in determining the levels of violation of non-compliant firms. Not surprisingly, non-compliant firms with risk-averse managers will have lower violations than otherwise identical firms with risk-neutral managers. However, the violations of non-compliant firms with risk-averse managers are independent of any difference that affects their profit functions and differences in their initial allocations of permits if and only if their managers' utility functions exhibit constant absolute risk aversion. On the other hand, firm-level characteristics do impact violation choices when managers have coefficients of absolute risk aversion that are increasing or decreasing in profit levels.

In principle, then, a regulator could target its enforcement effort when there is widespread non-compliance in an emissions trading program and firms' managers are risk averse. This targeting could be based directly on differences in the risk preferences of individual managers, or indirectly on differences in the firms' characteristics. However, a regulator must have detailed information about managers' risk preferences, including their coefficients of absolute risk aversion. Unfortunately, it seems unlikely that a regulator could screen individual managers on the basis of their risk preferences. It also seems unlikely that a regulator could infer this information from observed behavior. Thus, while a targeted monitoring strategy may be justified when firm managers are risk averse, the information requirements for forming such a strategy are rather severe.

Finally, it is important for regulators to understand how risk preferences affect the performance of markets for pollution control. I examine the market effects of risk aversion in Section 4. Risk aversion will have no impact on pollution rights markets when all firms are compliant, but may have significant impacts when there is widespread non-compliance. In these cases, risk aversion is associated with higher permit prices, better environmental quality, and lower aggregate violations.

2. A model of compliance under emissions trading

The analysis of this paper is largely based on a standard model of the decisions of a firm that operates under a competitive emissions trading program. It is important to note that the model of this paper can be applied to other tradable property rights programs with minor modifications. In fact, recent papers by Hatcher (2005) and Chavez and Salgado (2005) are direct applications of the literature on compliance and enforcement of emissions trading to individual transferable fishing quotas (ITQs). Thus the results of this paper apply to ITQ policies, as well as to other policies that seek to limit some activity through a market for the rights to engage in the activity.

The firm's gross profit in terms of emissions is $b(q,\alpha)$, which is strictly concave in its emissions q.[2] Absent an inducement to control its emissions, the amount of pollution the firm releases is the solution to $b_q(q,\alpha) = 0$. (Subscripts on a function denote derivatives of the function throughout the paper; for example, $b_q(q,\alpha) = \partial b(q,\alpha)/\partial q$

[1]Risk-seeking behavior is not considered in this paper, because such behavior is unlikely to be empirically important in most pollution control settings.

[2]This is a restricted profit function; that is, it is the profit a firm obtains when it makes all of its input and output choices optimally while holding its emissions to $q \leq q^0$. See Montgomery (1972) for a demonstration of the concavity of profit in emissions for firms that are price-takers in input and output markets. Since the formulation of $b(q,\alpha)$ is quite general, strict concavity can be guaranteed in many non-competitive settings as well.

and $b_{qq}(q,\alpha) = \partial^2 b(q,\alpha)/\partial q^2$). Denote this value of q as q^0, and let us limit the analysis to emissions less than q^0. For $q<q^0, b_q(q,\alpha)>0$. One can think of $b_q(q,\alpha)$ as the firm's marginal abatement costs, because it reveals the firm's reduction in profit from reducing its emissions by one unit. Assume that $b(q,\alpha)$ is increasing in the parameter α so that $b_\alpha(q,\alpha)>0$. Parametric differences in firms' gross profit functions are captured by differences in α.

The firm receives l_0 permits initially and holds l permits after trading in a compliance period is complete. Competitive behavior in the permit market establishes a constant price per permit p. Net expenditure or revenue from trading in the permit market is $p(l-l_0)$.

If the firm is non-compliant, its emissions exceed the number of permits it holds and the magnitude of its violation is $v=q-l>0$. If the firm is compliant, $q-l\leqslant 0$ and $v=0$. To check for compliance, the firm is audited with a known probability π and is assessed a penalty $f(v)$ if it is found to be non-compliant. There is no penalty for a zero violation, but the penalty is positive, strictly increasing, and strictly convex for positive violations. That is, $f(0)=0, f_v(0)>0, f_v(v)>0$, and $f_{vv}(v)>0$ for $v>0$.

Suppose that the manager of the firm is risk neutral or risk averse, and therefore has a concave (perhaps weakly) utility function u. Given the monitoring uncertainty and the fact that the firm's demand for permits is $l=q-v$, the manager's expected utility is

$$U(w) = (1-\pi)u(w^0) + \pi u(w^1), \tag{1}$$

where

$$w^0 = b(q,\alpha) - p(q-v-l_0) \tag{2}$$

and

$$w^1 = b(q,\alpha) - p(q-v-l_0) - f(v) \tag{3}$$

are the firm's net profit when it is not monitored, w^0, and when it is, w^1. The manager chooses the firm's emissions, q, and its violation, v, to maximize (1).

This description of a manager's decision problem is easily modified to analyze compliance with an emissions tax (Malik, 1990). To do so, let l be the firm's report of its emissions, let p be the fixed unit tax on emissions, and set $l_0=0$ to reflect the fact that a firm is required to pay for each unit of emissions. The results of this section and the next are obtained under the assumption of a fixed permit price. (In Section 4, the permit price is determined endogenously so that we can examine the market effects of risk aversion.) Therefore, all of the results of these two sections, except those involving the effects of l_0, can be directly applied to an analysis of compliance behavior under an emissions tax.[3]

[3]Recent papers that investigate compliance behavior under an emissions tax include Macho-Stadler and Perez-Castrillo (2006) and Sandmo (2002). Only Sandmo allows for risk aversion, but he does not analyze the effects of risk aversion and firm characteristics on compliance behavior, which is the main topic of this paper.

Assuming that the firm has positive emissions, its level of emissions is determined by the first-order condition

$$\partial U(w)/\partial q = \left[(1-\pi)u'(w^0) + \pi u'(w^1)\right]\left(b_q(q,\alpha) - p\right) = 0,$$

which holds if and only if $b_q(q,\alpha)=p$. That is, the manager chooses the firm's emissions so that the marginal gross profit from increased emissions is equal to the prevailing permit price. Note that this choice depends only on the profit parameter α and the going permit price p. Therefore, let us write the firm's optimal emissions as $\bar{q}=\bar{q}(\alpha,p)$. Our results about firms' emissions choices are summarized in our first proposition:

Proposition 1. *A firm's choice of emissions is independent of its manager's risk preference, its endowment of permits, and the enforcement strategy it faces.*[4]

These independence results are not new. Malik (1990) appears to have been the first to derive them in the case of emissions trading. See Harford (1978) and Sandmo (2002) for similar results in the case of an emissions tax.

That a firm's choice of emissions is independent of its manager's risk preference, its initial allocation of permits, and the enforcement strategy imply that a competitive permit market will maximize an industry's gross profit given the aggregate emissions that result. In other words, a competitive emissions trading policy will minimize an industry's aggregate costs of emissions control.[5] However, there are situations in which Proposition 1 will fail to hold, yielding an inefficient distribution of pollution rights. Obviously, the assumption of competitive permit trading is crucial.[6] Furthermore, Malik (1990) has shown that if a non-compliant firm's subjective probability of detection is $\pi(e_i,l_i)$, with $\partial\pi/\partial e_i + \partial\pi/\partial l_i \neq 0$, it chooses its emissions so that its marginal profit from increased emissions differs from the permit price. If this is the case, then it is unlikely that the aggregate costs of emissions control will be minimized. In this paper the detection probability a firm faces is common knowledge between the firm and the regulator, and is independent of the firm's permit demand and choice of emissions.

[4]This last independence result does not imply that equilibrium levels of emissions are independent of the enforcement strategy. Enforcement will have an indirect effect on equilibrium emissions through the permit price. However, given a permit price, each firm's choice of emissions are independent of the enforcement strategy it faces. Murphy and Stranlund (2006) use laboratory experiments of emissions trading when subjects can be noncompliant to confirm the zero direct effect of enforcement on emissions, and a negative indirect effect through the impact of enforcement on permit prices.

[5]This has always been an important objective for analysts and policy makers alike. Montgomery's (1972) seminal work on the efficiency of competitive emissions trading takes this approach, as have many papers that have followed in the literature on emissions trading. Moreover, the ability of competitive markets to distribute emissions control responsibilities cost-effectively is the main justification for the widespread implementation of these markets.

[6]See van Egteren and Weber (1996), Malik (2002), and Chavez and Stranlund (2003) for analyses of compliance behavior under emissions trading programs in the presence of market power.

We can now turn to a firm's compliance decision, given its optimal choice of emissions. Since the condition under which the firm is compliant is an important aspect of this study, its violation choice needs to be constrained to be non-negative. The first-order condition for the violation level that maximizes (1) subject to this constraint is

$$\begin{aligned}\partial U(w)/\partial v &= (1-\pi)u'(w^0)p + \pi u'(w^1)(p - f_v(v)) \\ &\leqslant 0 \quad \text{if} < 0 \quad \text{then } v = 0.\end{aligned}$$

This condition can be rewritten as:

$$p - \pi f_v(v)R(v, \alpha, l_0, \pi, p) \leqslant 0 \quad \text{if} < 0 \quad \text{then } v = 0, \tag{4}$$

where

$$\begin{aligned}R(v, \alpha, l_0, \pi, p) &= u'(w^1)/\left[(1-\pi)u'(w^0) + \pi u'(w^1)\right] \\ &= u'(w^1)/U'(w).\end{aligned} \tag{5}$$

The implicit solution to (4) is the firm's optimal violation, which is denoted

$$\bar{v} = \bar{v}(\alpha, l_0, \pi, p). \tag{6}$$

The second order condition that guarantees that (4) identifies a unique optimal violation requires that $\pi f_v(v)R(v, \alpha, l_0, \pi, p)$ is strictly increasing in v. That is, $f_{vv}R + f_vR_v > 0$. It is straightforward to demonstrate that this condition holds as long as the firm's manager is not a risk seeker. The consequences of risk-seeking behavior are not examined in this paper. Note that I have written $R(v, \alpha, l_0, \pi, p)$ as being independent of the firm's level of emissions, $\bar{q}(\alpha, p)$. This follows from the fact that $w_q^1 = w_q^0 = b_q(\bar{q}, \alpha) - p = 0$, which implies $R_{\bar{q}} = 0$.

$R(v, \alpha, l_0, \pi, p)$ is an adjustment of the marginal expected penalty, $\pi f_v(v)$, that accounts for the manager's attitude toward risk. If the manager is risk neutral, his or her utility function is linear, implying that $u'(w^1) = u'(w^0)$ and $R = 1$. If the firm is compliant, $R = 1$ as well, because $w^0 = w^1$. However, if the manager is risk averse, his or her utility function is strictly concave. Therefore, if the firm is non-compliant, then $w^0 > w^1$ and $u'(w^1) > u'(w^0)$. Since $(1-\pi)u'(w^0) + \pi u'(w^1)$ is a linear combination of $u'(w^1)$ and $u'(w^0)$, $u'(w^1) > (1-\pi)u'(w^0) + \pi u'(w^1)$. This implies that $R > 1$ for a non-compliant firm with a risk-averse manager.

In the case of risk neutrality, it is straightforward to demonstrate Stranlund and Dhanda's (1999) result that the decision to comply and the choice of violation level are independent of any firm-specific characteristics. From (4) it is straightforward to establish that a firm with a risk-neutral manager is compliant if and only if $p \leqslant \pi f_v(0)$; that is, a firm is compliant if and only if the permit price is not greater than the expected marginal penalty of a slight violation. Note that the firm's gross profit, as reflected in the parameter α, and its initial allocation of permits, l_0, do not affect this decision rule. Therefore, a manager's decision about whether the firm should be in compliance is independent of these parameters. This independence extends to the violation level of a non-compliant firm as well. To see this, let (4) hold with equality, set $R = 1$ and substitute the firm's optimal violation (6) to obtain $p - \pi f_v(\bar{v}(\alpha, l_0, \pi, p)) \equiv 0$. Differentiate this identity with respect to α to obtain $-\pi f_{vv}\bar{v}_\alpha = 0$, which implies $\bar{v}_\alpha = 0$ This result indicates that firms' violations are independent of parametric differences in their profit functions when their managers are risk neutral. The same is true for differences in their initial permit allocations.

Therefore, as long as managers are risk neutral, there is no reason for regulators to believe that some firms will be more likely to be non-compliant, or tend toward higher violations, even though they may be very different in ways that affect their profit functions or in their initial permit allocations. Moreover, there is no reason to believe that the marginal productivity of increased enforcement in reducing violations will differ among firms.[7] Hence, a regulator that is motivated to target its enforcement resources to reduce incidences of non-compliance cannot do so on the basis of firm-level characteristics when managers are risk neutral. In Section 3, we examine whether this result continues to hold when managers are risk averse.

3. Risk aversion and compliance behavior

In this section, let us analyze the compliance behavior of firms when their managers are risk averse. We begin with a firm's decision to be compliant or not. Perhaps surprisingly, a manager's risk preference has no bearing on whether his or her firm will be compliant. That is, regardless of the manager's attitude toward risk, the firm is compliant if and only if $p \leqslant \pi f_v(0)$. To prove the "only if" part of this assertion, first note that if $v = 0$, then $w^0 = w^1$ and $R = 1$. From (4), then, it is clear that $v = 0$ requires $p \leqslant \pi f_v(0)$. To prove that $v = 0$ if $p \leqslant \pi f_v(0)$, recall that the second order condition for the determination of v requires that $\pi f_v(v)R$ is strictly increasing in v. This and $R = 1$ when $v = 0$ imply $\pi f_v(v)R > \pi f_v(0)$ for $v > 0$. In turn, if $p \leqslant \pi f_v(0)$, then $p - \pi f_v(v)R < 0$ for $v > 0$. From (4), however, $p - \pi f_v(v)R < 0$ for $v > 0$ implies that the optimal choice of violation is $v = 0$. Therefore, we conclude that $v = 0$ if $p \leqslant \pi f_v(0)$.

Since a manager's decision about whether his or her firm should comply depends only on the prevailing permit price and the enforcement strategy, we have the following proposition:

Proposition 2. *A firm is compliant if and only if $p \leqslant \pi f_v(0)$. Therefore, whether a firm is compliant or not is independent of its gross profit function, its endowment of permits, and its manager's risk preference.*

Proposition 2 implies that firms with risk-averse managers are not more (or less) likely to be compliant than firms with risk-neutral managers. This result has important implications. Suppose that a regulator's objective is to use its enforcement strategy to induce full

[7] Murphy and Stranlund (2007) find strong support for this hypothesis in laboratory experiments of emissions trading.

compliance to an emissions trading policy. One may suspect that it would be easier to induce full compliance by firms with risk-averse managers than to induce compliance by firms with risk-neutral managers, but this is clearly not the case. That a firm's decision to comply or not is also independent of its gross profit function and its initial allocation of permits implies that a regulator with the objective of inducing complete compliance should not pursue a targeted monitoring strategy. Minimizing the enforcement costs of inducing full compliance implies uniform monitoring of firms so that $p = \pi f_v(0)$.

While a firm with a risk-averse manager is not more likely to be non-compliant than a firm with a risk-neutral manager, a non-compliant firm with a risk-averse manager will choose a lower violation than an otherwise identical non-compliant firm with a risk-neutral manager. This follows from the fact that $R = 1$ for a risk-neutral manager and $R>1$ for a risk-averse manager; that is, the expected marginal disutility of being penalized for a particular violation level is higher for a risk-averse manager than for a risk-neutral manager. In principle, with information on the risk preferences of individual managers, a regulator could target its enforcement effort based on this information. However, it seems unlikely that a regulator could categorize or screen managers on the basis of their risk preferences.

There remains the possibility, however, that a regulator can use observable information about firms to target its monitoring effort when firms have risk-averse managers. This information may include observable characteristics of production technologies, levels of inputs and outputs, abatement equipment, or their initial allocations of permits. To determine whether it is possible for a regulator to do this, we need to determine whether a firm's violation depends on the parameter α, or on its initial allocation of permits, l_0. Using (4) and (6), write the identity

$$p - \pi f_v(\bar{v})R(\bar{v}, \alpha, l_0, \pi, p) \equiv 0. \tag{7}$$

From (7) obtain $\bar{v}_\theta = -R_\theta f_v/S$, where $\theta \in (\alpha, l_0)$ and $S = f_{vv}R + f_vR_v > 0$, which is required by the second-order condition for determining an optimal violation. Since $f_v>0$, the sign of $\bar{v}_\theta$ is equal to the sign of $-R_\theta$. Using (5), one can calculate

$$-R_\alpha = Ab_\alpha\left[-u''(w^1)/u'(w^1) + u''(w^0)/u'(w^0)\right], \tag{8}$$

and

$$-R_{l_0} = Ap\left[-u''(w^1)/u'(w^1) + u''(w^0)/u'(w^0)\right], \tag{9}$$

where $A = (1-\pi)u'(w^1)u'(w^0)[U'(w)]^2>0$. Since $b_\alpha>0$, $-R_\alpha$ and $-R_{l_0}$ have the same sign as $-u''(w^1)/u'(w^1)+u''(w^0)/u'(w^0)$. Therefore, $\bar{v}_\alpha$ and $\bar{v}_{l_0}$ have the same sign as this term as well.

However, $-u''(w)/u'(w)$ is the Arrow-Pratt coefficient of absolute risk aversion and the sign of $-u''(w^1)/u'(w^1) + u''(w^0)/u'(w^0)$ is simply a statement of whether the manager's utility function exhibits decreasing, constant, or increasing absolute risk aversion (Mas-Colell et al., 1995, p. 193). Since from (2) and (3), $w^0>w^1$, if the manager's utility function exhibits decreasing absolute risk aversion, then $-u''(w^1)/u'(w^1) + u''(w^0)/u'(w^0)>0$. With constant absolute risk aversion this term is equal to zero, and with increasing absolute risk aversion this term is negative. Therefore, for a non-compliant firm with a risk-averse manager we have the following proposition:

Proposition 3. *The violation of a non-compliant firm with a risk-averse manager is independent of its gross profit function and its endowment of permits if and only if the manager's utility function exhibits constant absolute risk aversion. If the manager's utility function exhibits decreasing (increasing) absolute risk aversion, the firm's violation increases (decreases) with a parametric increase in its gross profit or an increase in its endowment of permits.*

Violation levels of non-compliant firms are independent of their parametric differences and their initial allocations of permits if the firms' managers have utility functions that exhibit constant absolute risk aversion. This is a generalization of Stranlund and Dhanda's (1999) independence result that they obtained under the assumption of risk-neutral managers. However, these results do not hold with non-constant absolute risk aversion. A firm's violation is increasing (decreasing) in its gross profit and its initial allocation of permits if and only if its manager's utility exhibits decreasing (increasing) absolute risk aversion. Managers with decreasing absolute risk aversion take on more risk as their firms become more profitable. Given a choice of emissions and permit demand, a higher value of α or a larger initial allocation of permits both imply that a firm is more profitable. If the firm's manager has decreasing absolute risk aversion, then higher values of these parameters (higher profitability) lead the manager to take on more risk, and hence, to choose a higher violation. The opposite is true if the manager has increasing absolute risk aversion, because then he or she is motivated to take on less risk when the firm is more profitable.

The policy implication of Proposition 3 is that firms' violations are not independent of parametric differences in their gross profit functions or initial allocations if their managers are risk averse with utility functions that exhibit non-constant absolute risk aversion. In principle, then, a regulator would be able to form a targeted monitoring strategy based on observable differences in firm's initial allocations of permits and on characteristics that determine their gross profit functions. To do so, however, requires rather detailed information about individual managers' risk preferences. It is important to note that it is simply not enough know, or assume, that managers are risk averse. If risk-averse managers have constant absolute risk aversion, there is no justification for a targeted monitoring strategy that is based on firm-level characteristics. A targeted strategy is justified only when a regulator knows whether managers have increasing or decreasing absolute risk aversion. Unfortunately, it does not seem likely that a regulator could ever obtain this information, or infer it from observations of firm behavior. Thus, although a targeted monitoring strategy may be justified when firm managers are risk averse, the information requirements for forming such a strategy are rather severe.

Recall from Proposition 2 that no such asymmetric information problem exists if the regulatory objective is to induce full compliance. Allowing imperfect compliance without the information required to form a targeted enforcement strategy implies that the regulator cannot be certain of individual violations when risk aversion is prevalent among the managers of firms in an emissions market. Thus, designing an emissions trading policy that induces full compliance allows regulators to avoid this uncertainty.

4. Market effects of risk aversion

Up to this point the analysis has been conducted under the assumption of a fixed price for emissions rights. However, any analysis of compliance behavior under tradable property rights must ultimately make the price of these rights endogenous. Perhaps the most important reason for doing so is that firms' compliance choices are linked together by the market for property rights. This is one of the features of market-based environmental and natural resource policies that distinguish them from command-and-control regulations and from taxes.

Let us consider the following thought experiment. Start from a situation in which all the firms in an emissions trading program are non-compliant and all their managers are risk neutral. Then, let us replace a significant number, but not all, of the firms' managers with risk-averse managers. Doing so allows us to trace out the effects of risk aversion on the equilibrium permit price, aggregate emissions, aggregate violations, as well as the violations of the firms with their new risk-averse managers and the violations of those that keep their risk-neutral managers.

Suppose that a fixed number of emissions permits are in circulation, and that when all the managers are risk neutral, all firms violate their permits so that aggregate emissions exceed the number of permits in circulation. Now replace a significant number of the firms' managers with risk-averse individuals. Recall that risk-averse managers will choose lower violations for their firms than their risk-neutral counterparts. Thus, holding the permit price constant, the firms that acquire a risk-averse manager will reduce their violations. In principle, these firms could reduce their violations by purchasing more permits or by reducing their emissions. However, recall that a firm's level of emissions is independent of its manager's preference for risk (Proposition 1). Therefore, holding the permit price constant, firms with risk-averse managers will reduce their violations by demanding more permits, not by reducing their emissions.

Clearly the increased permit demand by the firms who now have risk-averse managers will increase the equilibrium permit price. In response to this increase in the permit price, all firms will reduce their emissions. To see why, recall that a firm's choice of emissions, $\bar{q} = \bar{q}(\alpha, p)$, is determined from the first-order condition $b_q(q, \alpha) = p$. From this condition obtain $\bar{q}_p = 1/b_{qq} < 0$, which implies that each firms' emissions are decreasing in the price of emissions permits. (Recall that a firm's gross profit function is strictly concave so that $b_{qq} < 0$). Thus, relative to a situation involving non-compliant firms with risk-neutral managers, replacing a significant number of these managers with risk-averse individuals will induce a higher equilibrium permit price and lower aggregate emissions. Of course, given a fixed supply of emissions permits, lower aggregate emissions imply lower aggregate violations. We conclude, therefore, that under market-based environmental policies with widespread non-compliance, risk aversion is associated with higher permit prices, better environmental quality, and higher aggregate compliance.

However, higher aggregate compliance does not imply lower violations by all firms. In fact, a significant number of risk-averse managers will lead to higher violations by the firms with risk-neutral managers. To demonstrate this, recall that Eq. (7) is the first-order condition for a positive violation by a firm evaluated at its optimal violation $\bar{v}(\alpha, l_0, \pi, p)$. Also recall that if the firm's manager is risk neutral, then $R(\bar{v}, \alpha, l_0, \pi, p) = 1$. From $p - \pi f_v(\bar{v}) \equiv 0$, then, obtain $\bar{v}_p = 1/\pi f_{vv} > 0$, which indicates that the violation of a firm with a risk-neutral manager is increasing in the price of permits. Since a significant number of risk-averse managers will tend to push the equilibrium permit price up because they demand more permits, their behavior will also lead the firms with risk-neutral managers to increase their violations.

Our conclusions about the effects of risk aversion on transferable permit market outcomes are summarized in our final proposition.

Proposition 4. *Relative to a market involving non-compliant firms with risk-neutral managers, an otherwise identical market but with a significant number of risk-averse managers will have* (1) *a higher equilibrium permit price,* (2) *lower emissions by all firms,* (3) *lower aggregate violations, but* (4) *higher violations by the firms with risk-neutral managers.*

In contrast, risk aversion will have no impact on permit markets when all firms are compliant. That is, suppose we started from a situation involving compliant firms with risk-neutral managers. Replacing some or all of the managers with risk-averse managers will have no affect on equilibrium outcomes, because these new managers will also choose compliance for their firms. This follows from Proposition 2, which states that a firm's decision to be compliant is independent of the risk preference of its manager.

Therefore, the market effects of risk aversion are limited to situations involving widespread non-compliance. In these cases, risk aversion produces better environmental quality and lower aggregate violations. However, regulators need to always be aware of indirect price effects under tradable rights regulations. The lower violations of firms with risk-averse managers put upward pressure on the market price of rights, which in turn motivates firms with risk-neutral managers toward higher violations.

5. Conclusion

We have examined the consequences of risk aversion on compliance behavior in markets for pollution control, and have obtained several new results with significant policy implications. First, a firm's decision about whether to comply or not is independent of its manager's risk preference; that is, firms with risk-averse managers are not more (or less) likely to be compliant than firms with risk-neutral (or even risk-loving) managers. Thus, if the objective of a regulator is to enforce an emissions trading policy so that all firms that operate under the policy are fully compliant, the distribution of risk preferences among the managers of the firms has no bearing on what is required to achieve this objective. Moreover, one cannot justify a targeted enforcement strategy that involves monitoring firms with certain characteristics more closely than others.

However, risk preferences do play a role in determining the violation levels of non-compliant firms. While non-compliant firms with risk-averse managers will have lower violations than otherwise identical firms with risk-neutral managers, the effects of parametric differences in firms' gross profit functions and their initial permit allocations on their violation choices depend critically on their managers' coefficients of absolute risk aversion. A firm's violation is independent of these characteristics if its manager's coefficient of absolute risk aversion is constant over profit levels. However, higher (lower) violations are associated with parametrically higher gross profits and higher initial permit allocations when a manager's coefficient of absolute risk aversion is decreasing (increasing) in profit levels.

In principle, a regulator could target its enforcement effort based directly on the risk preferences of individual managers, or indirectly on differences in the firms' characteristics. However, doing so first requires that there is significant non-compliance. Moreover, the regulator must have detailed information about managers' risk preferences, including whether their utility functions exhibit increasing, decreasing, or constant absolute risk aversion. Although, a targeted monitoring strategy may be justified when firm managers are risk averse, the information requirements for forming such a strategy appear prohibitive.

It is important, however, for regulators to understand how risk preferences affect the performance of tradable rights regulations. We have seen that risk aversion will have no impact on markets for pollution control when all firms are compliant, but may have significant impacts when there is widespread non-compliance. In these cases, risk aversion is associated with higher permit prices, lower aggregate violations, and better environmental quality.

Acknowledgments

Primary funding for this research was provided by the US EPA—Science to Achieve Results (STAR) Program grant #R829608. Additional support was provided by the Cooperative State Research Extension, Education Service, US Department of Agriculture, Massachusetts Agricultural Experiment Station, and the Department of Resource Economics under Project No. MAS00871.

References

Chavez, C.A., Salgado, H., 2005. Individual transferable quota markets under illegal fishing. Environmental and Resource Economics 31 (3), 303–324.

Chavez, C.A., Stranlund, J.K., 2003. Enforcing transferable permit systems in the presence of market power. Environmental and Resource Economics 25 (1), 65–78.

Garvie, D., Keeler, A., 1994. Incomplete enforcement with endogenous regulatory choice. Journal of Public Economics 55, 141–162.

Gray, W., Shadbegian, R., 2005. When and why do firms comply? Paper mills in the 1980s. Law and Policy 27 (2), 238–261.

Harford, J., 1978. Firm behavior under imperfectly enforceable pollution standards and taxes. Journal of Environmental Economics and Management 5 (1), 26–43.

Hatcher, A., 2005. Non-compliance and the Quota Price in an ITQ Fishery. Journal of Environmental Economics and Management 49 (3), 427–436.

Keeler, A., 1991. Noncompliant firms in transferable discharge permit markets: some extensions. Journal of Environmental Economics and Management 21, 180–189.

Macho-Stadler, I., Perez-Castrillo, D., 2006. Optimal enforcement policy and firm's emissions and compliance with environmental taxes. Journal of Environmental Economics and Management 51 (1), 110–131.

Malik, A.S., 1990. Markets for pollution control when firms are noncompliant. Journal of Environmental Economics and Management 18, 97–106.

Malik, A.S., 1992. Enforcement cost and the choice of policy instruments for controlling pollution. Economic Inquiry 30, 714–721.

Malik, A.S., 2002. Further results on permit markets with market power and cheating. Journal of Environmental Economics and Management 44 (3), 371–390.

Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford University Press.

Montgomery, W.D., 1972. Markets in licenses and efficient pollution control programs. Journal of Economic Theory 5 (3), 395–418.

Murphy, J.J., Stranlund, J.K., 2006. Direct and market effects of enforcing emissions trading programs: an experimental analysis. Journal of Economic Behavior and Organization 61 (2), 217–233.

Murphy, J.J., Stranlund, J.K., 2007. A laboratory investigation of compliance behavior under tradable emissions rights: implications for targeted enforcement. Journal of Environmental Economics and Management 53 (2), 196–212.

Sandmo, A., 2002. Efficient environmental policy with imperfect compliance. Environmental and Resource Economics 23 (1), 85–103.

Stranlund, J.K., Chavez, C.A., 2000. Effective enforcement of a transferable emissions permit system with a self-reporting requirement. Journal of Regulatory Economics 18 (2), 113–131.

Stranlund, J.K., Dhanda, K.K., 1999. Endogenous monitoring and enforcement of a transferable emissions permit system. Journal of Environmental Economics and Management 38 (3), 267–282.

Stranlund, J.K., Costello, C., Chavez, C.A., 2005. Enforcing emissions trading when emissions permits are bankable. Journal of Regulatory Economics 28 (2), 181–204.

van Egteren, H., Weber, M., 1996. Marketable permits, market power, and cheating. Journal of Environmental Economics and Management 30, 161–173.

ELSEVIER

Journal of Environmental Management 88 (2008) 211–218

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Microwave-assisted stripping of oil contaminated drill cuttings

J.P. Robinson*, S.W. Kingman, O. Onobrakpeya

School of Chemical, Environmental and Mining Engineering, University of Nottingham, University Park, Nottingham NG7 2RD, UK

Received 23 August 2006; received in revised form 26 January 2007; accepted 17 February 2007
Available online 12 April 2007

Abstract

The application of microwave heating technology to conventional gas stripping processes has been investigated in the remediation of contaminated drill cuttings. The technical feasibility and limitations of nitrogen and steam stripping processes are demonstrated, and it is shown that the combination of microwave heating with the stripping process offers a step change in performance. Order of magnitude improvements in processing time are shown for the microwave-assisted processes, as well as greatly improved levels of remediation. The mechanisms of contaminant removal are discussed, along with the phenomena which occur with microwave heating processes. The energy requirements of each of pure gas and microwave-assisted processes are also discussed, and the potential applications of each technology are highlighted relative to the overall remediation requirements.

Keywords: Microwave heating; Drill cuttings; Gas stripping; Steam stripping; Remediation

1. Introduction

Oil contaminated drill cuttings arise from exploratory drilling operations within the oil and gas industry worldwide, and can contain significant quantities of oil from the drilling fluids. The oil levels within the cuttings pose a significant risk (Melchor et al., 2002) to both marine and onshore environments and ecosystems, so the industry has made significant advances in the development of water-based drilling fluids to alleviate environmental concerns. Oil-based muds are still utilised in more demanding drilling applications such as extended-reach and horizontal wells, and when boring through poorly consolidated rock formations. Such drilling is inevitable in mature oil fields such as the UK and Norwegian North Sea sectors, where the past exploitation of the 'easy' reserves means that oil companies are now chasing the more difficult reservoirs accessible only with the latest drilling technology. A recent UK and EU legislation has prevented the discharge of cuttings directly into the sea unless the oil levels are below 1% w/w, although the operating companies effectively view this as 'zero' discharge due to the current technological limitations in achieving this threshold. Operators in the Gulf of Mexico are subject to a discharge limit of 6.9% w/w; however, in future this is likely to become more strict.

1.1. Existing treatment technologies

The only commercially viable treatment strategy used at present is the transport of cuttings to shore and treatment in a thermal desorption plant. Alternative shore-based processes such as bioremediation (Pankhania, 2004), solvent extraction (Guigard and Odusanya, 2004) and surfactant treatment (Harrison and Zwinderman, 2004) have proven to be technically viable; however, the difficulties lie in transporting the drill cuttings to land. Transporting large quantities of drilling wastes from offshore platforms to land poses significant safety and operational hazards to oil field personnel and also requires a large buffer storage capacity due to frequent bad weather. This practice is commonplace due to the lack of proven techniques for remediation of the cuttings on the offshore platforms, and it is for this reason that in recent years significant attention has been paid to the development of offshore treatment methods. Re-injection of drill cuttings into dedicated offshore disposal wells has been

*Corresponding author. Tel.: +44 115 951 4092; fax: +44 115 951 4115.
E-mail address: john.p.robinson@nottingham.ac.uk (J.P. Robinson).

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.02.009

explored (Jones, 2004), but does not offer a sustainable solution in the long term.

The major factor which prevents successful shore-based treatment processes being used on offshore platforms is space, with most platforms never designed to house a cuttings-handling system. The drill cuttings are granular in nature, and these materials exhibit extremely poor heat-transfer properties. Supplying energy within the depth of a granular material using conventional heating processes is very inefficient, and thermal desorption processes which use this technique require long residence times which translates to large plant footprints. Microwave heating has the potential to supplement these technologies by dissipating energy within the depth of the granular material, and selectively heating only the microwave-receptive components without bulk-heating the surrounding inert materials.

1.2. Theory of microwave heating

In conventional thermal processing, energy is transferred to a material through conduction, convection and radiation. In contrast, microwave energy is delivered directly to materials through molecular interactions with an electromagnetic field. The internal temperature distribution of a material subject to conventional heating is limited by its thermal conductivity, whereas microwave heating results in all individual elements of the material being heated individually. Consequently, heating times using microwaves can often be reduced to less than 1% of those required using conventional heating methods (Meredith, 1998).

There are three generic classifications for the behaviour of materials upon interaction with a microwave field:

1. Transparent (low-dielectric-loss materials)—microwaves pass through the material with little absorption.
2. Opaque (conductors)—microwaves are reflected by the material and do not penetrate.
3. Absorbing (high-dielectric-loss materials)—microwave energy is absorbed based on the electric field strength and the dielectric loss factor.

Microwave processing has distinct advantages in the treatment of materials which contain a mixture of absorbers and transparent components. Microwave energy is absorbed by the substances with a high dielectric loss whilst passing through the low-loss transparent material, resulting in selective heating. In this case, significant energy savings are possible since the dielectric material can be heated without heating the entire matrix. The power absorbed per unit volume of material, known as power density (Pd), is dependent on the dielectric properties of the material and can be represented by

$$\mathrm{Pd} = 2\pi f \varepsilon_0 \varepsilon'' |E|^2, \qquad (1)$$

where f is the microwave frequency, ε_0 the permittivity of free space (8.85×10^{-12} F/m), ε'' the relative dielectric loss factor and $|E|$ the magnitude of the electric field. A number of heating mechanisms may occur depending on the dipolar nature of the material and whether any free electrons or ions are present. From Eq. (1) it is evident that the microwave energy absorbed by a dielectric material is proportional to the square of the electric field strength. The design of the microwave cavity is critical in that it can allow very well-defined electric fields in a relatively small volume (single-mode cavity), or can permit the electric fields to encompass a much larger volume, albeit with a compromise in the field definition (multimode cavity). With single-mode cavities the high electric field strength results from superposition of reflected waves, hence the geometry of the cavity is dependent on the microwave frequency.

Few workers have specifically studied the remediation of drill cuttings; however, there has been significantly more research carried out into the treatment of contaminated soils, a system which also contains a mineral phase, water and an organic phase. By far the most common method of soil remediation is gas stripping with a heated gas stream, with Lord (1998) and Braass et al. (2003) demonstrating steam stripping for the removal of organic contaminants. Enhanced soil stripping processes have also been reported, with Buettner and Daily (1995) demonstrating an air stripping system which utilised electrical heating of the bulk contaminated soil. Higher temperatures have been shown to enhance contaminant removal in soil stripping processes (Pina et al., 2002), and the limitations of conventional heating have been identified by Di et al. (2000), who demonstrated a microwave-assisted stripping process for soil remediation. In the latter case, microwaves were used to generate steam from wet soil, with the steam forming a medium to desorb or entrain the organic contaminants. Our previous work has studied a range of parameters such as electric field strength and treatment time on the microwave processing of contaminated drill cuttings (Shang et al., 2005, 2006); however, the potential of microwaves to enhance a stripping process was not studied. The significance of previous studies was the identification that the contaminant molecules are generally microwave transparent. In both the treatment of drill cuttings and contaminated soils it is the water within the media which absorbs microwaves, with the energy then transferred to the contaminants to promote their removal. This is intuitively a somewhat inefficient desorption mechanism; however, the advantages result from volumetric heating rather than selective heating. Microwave energy can be absorbed very rapidly within the depth of the contaminated material, without relying on conventional heat transfer.

Many stripping processes utilise a hot gas to supply energy to a bed of solids, due to the low thermal conductivity of the bed. Such processes therefore have the potential to benefit from the introduction of microwave energy, as the stripping gas could be used only to enhance the mass transfer rather than provide thermal energy. This study focuses on the use of microwave treatment to enhance steam and nitrogen stripping of contaminated drill cuttings in particular, and the applicability to wider desorption and stripping processes is also discussed as well as scale-up issues.

2. Experimental

2.1. Apparatus

The stripping experiments were carried out using fixed bed of contaminated drill cuttings within a single-mode microwave apparatus. A schematic of the experimental set-up is shown in Fig. 1.

Microwaves were generated at 2.45 GHz using a magnetron connected to a 0–1 kW power supply. From the magnetron the microwaves travel down a waveguide (WR 340) to the single-mode applicator, which consists of a tower of circular cross-section positioned towards the end of the waveguide. The sample bed comprised 30 g of contaminated drill cuttings contained within a glass reactor, which was located within the tower and clamped in position. The stripping gas was introduced from under the bed of sample, which rested on a sintered support material. Nitrogen was supplied from compressed cylinders, and steam supplied from a portable steam generator. The stripping gas to be used was passed through a heater and a flow meter before entering the microwave cavity. In both cases the temperature and flow rate were maintained at 120 °C and 10 l/min, respectively. Whilst not the optimal design in terms of stripping, the equipment had to be compatible with the 2.45 GHz single-mode microwave system. This effectively placed limitations on the width of the glass reactors that could be employed, with 40 mm the maximum due to the wavelength of the microwaves at this frequency. The moveable piston and stub-tuners were positioned such that the region of high electric field intensity was located within the tower where the sample was located, and this was achieved by monitoring the amount of power reflected. Any microwave energy not absorbed by the sample travels back down the waveguide and is absorbed by the circulator, which is connected to a cooling water supply.

Experimental sequences were performed in triplicate, with one representative set of results shown in each of the following figures. The variability in oil contents at equivalent conditions in different experiments was less than 0.2% w/w.

2.2. Materials

Drill cuttings samples were supplied by BP from their commercial North Sea drilling activities, and contained 12% water and 10% oil (weight basis). The water content is below its saturation level. Oil and water contents were measured using solvent extraction techniques and verified by thermo-gravimetric analysis (TGA). Solvent extraction was performed using dichloromethane (DCM) in a Soxhlet-extractor, where up to 10 g of drill cuttings were contacted with 100 ml of boiling DCM for 6 h. The composition of the oil is consistent with its function as a base fluid for the drilling industry, and is composed primarily of *n*-alkanes in the C_{12}–C_{20} range, with less than 1% aromatics and an end-point of 250 °C. The contaminated drill cuttings used for this study were of the order of 3–10 mm in diameter.

3. Results and discussion

To fully assess the effect of the microwave energy, stripping experiments were performed with both nitrogen and steam without the application of microwaves. In both cases the temperature of the gas entering the reactor was 120 °C, and gas flow rates set to 10 l/min. A modified Reynolds Number, $Re' = 15$, was used based on the average particle diameter within the bed and the superficial gas velocity. This corresponds to the transition region between laminar and turbulent flow (Perry and Green, 1997), and was the maximum velocity which could be attained without fluidising the bed. Although fluidised bed

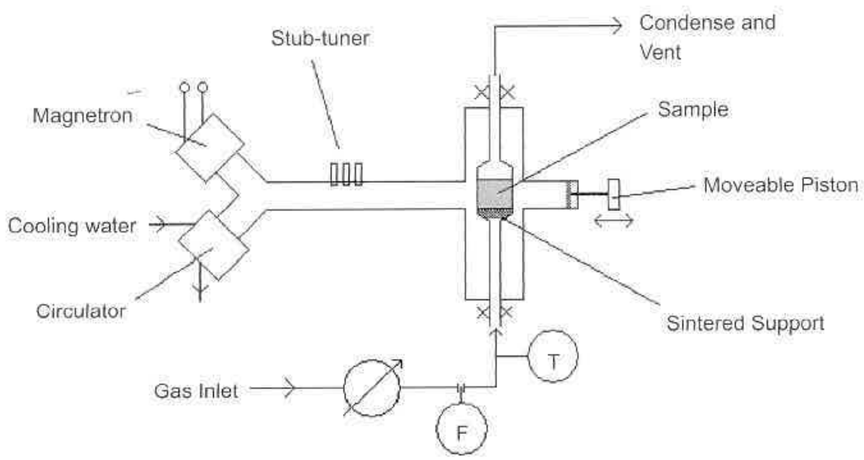


Fig. 1. Schematic of single-mode microwave-assisted stripping apparatus.

systems are clearly of interest, this study was restricted to fixed-bed gas stripping systems in order that the interaction between the material and the electric field could be more easily controlled.

3.1. Hot gas stripping results

Using gas stripping alone yields the results shown in Fig. 2, where it is apparent that a significant fraction of oil can be removed from the contaminated drill cuttings using a simple fixed-bed arrangement. In order that the current environmental discharge limit for drill cuttings is achieved, the oil removal efficiency needs to be of the order of 90–95%, and Fig. 2 shows that neither nitrogen nor steam is able to reach this value. Of note is that a very large number of bed volumes of gas are required to remove the oil, and the economic implications of this are discussed later in this paper.

Steam appears to be more beneficial than nitrogen in removing oil from the contaminated drill cuttings, and this is due to unique phenomena associated with steam stripping. When two immiscible phases are present together (e.g. oil and water in contaminated drill cuttings), the boiling point of the mixture can be considerably lower than the boiling point of either of the pure components. The vapour pressure of immiscible liquids is equal to the sum of the vapour pressures of the individual components, which means that the boiling point of such a mixture depends only on temperature and not on the concentration of either component. In this case the boiling range of oil and water in drill cuttings can be estimated by assuming that each liquid is well mixed within the contaminated cuttings. However, this analysis pertains only to 'free' water and oil and not that adsorbed to the surface of the rock or trapped within capillaries. The oil used in drill cuttings contains a wide range of hydrocarbons, so the properties of both a heavy and light fuel oil have been used as a reference (Spiers, 1962). Fig. 3 shows the resultant vapour pressure for mixtures of water and two fuel oils.

It can be seen in Fig. 3 that the boiling point of the oil/water phases is much less than the temperature of the stripping gas, hence thermal desorption can be induced from the hot gas. This will occur with both nitrogen and steam as water is present within the contaminated cuttings; however, the nitrogen will remove water as well as oil whereas the steam will partially replenish the lost moisture. Once the free water is lost the remaining contaminant oil must be stripped via a conventional solubility process or desorbed at temperatures in excess of its end-point, i.e. 250 °C. Referring to Fig. 2, when the nitrogen stripping process approaches 60% oil removal, it appears that this point has been reached as the further removal of oil becomes more difficult. When steam is used, however, the water is replenished meaning that more of the oil can be removed below its normal boiling point. In this particular case around 80% of the oil can be removed with steam compared to just over 60% with an equivalent flow rate of nitrogen. This finding is not surprising given the widespread use of steam stripping in the industrial extraction of high-molecular-weight organic molecules such as perfumes and 'essential oils'. The advantage of steam stripping is that extractions can be performed without inducing thermal decomposition of the organic molecules of interest, particularly when the organics are of high value. In the case of contaminated drill cuttings the organic phase is more robust, not undergoing thermal decomposition until temperatures in excess of 600 °C are reached.

It has been shown that the technical limitations of nitrogen and steam stripping are their inability to remove sufficient quantities of oil for disposal of the drill cuttings, and this is most likely to arise from poor heat and mass transfer. Water in the contaminated drill cuttings occurs primarily from that trapped within the rock formations which are being drilled, and hence the water is present both

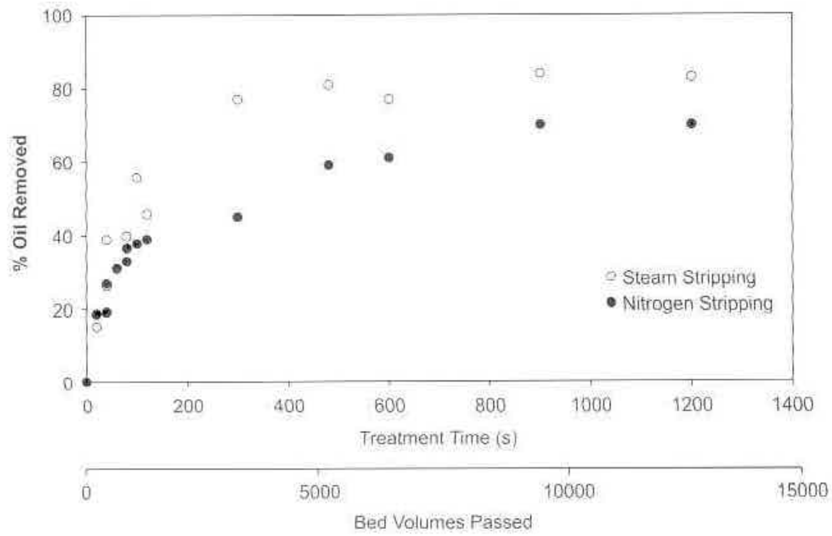


Fig. 2. Removal of oil in fixed-bed nitrogen and steam stripping as a function of treatment time and number of bed volumes of gas passed.

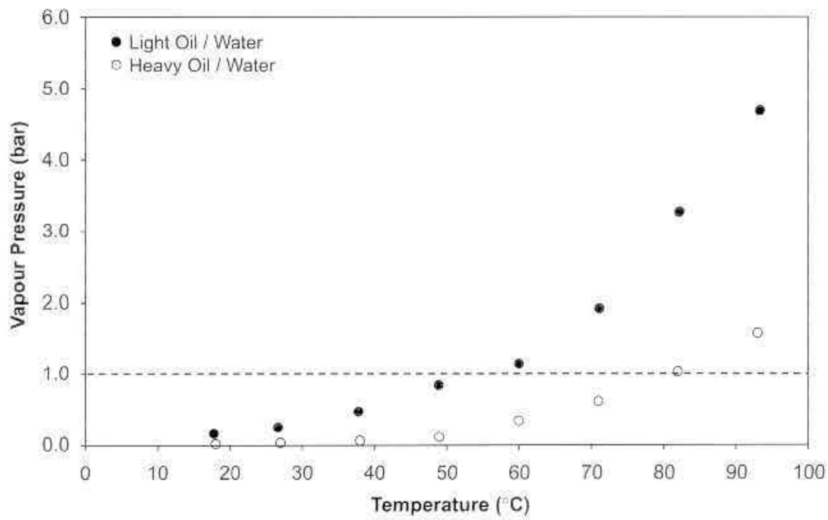


Fig. 3. Vapour pressure of oil and water mixtures with temperature. The boiling point at atmospheric pressure is indicated by the dashed line.

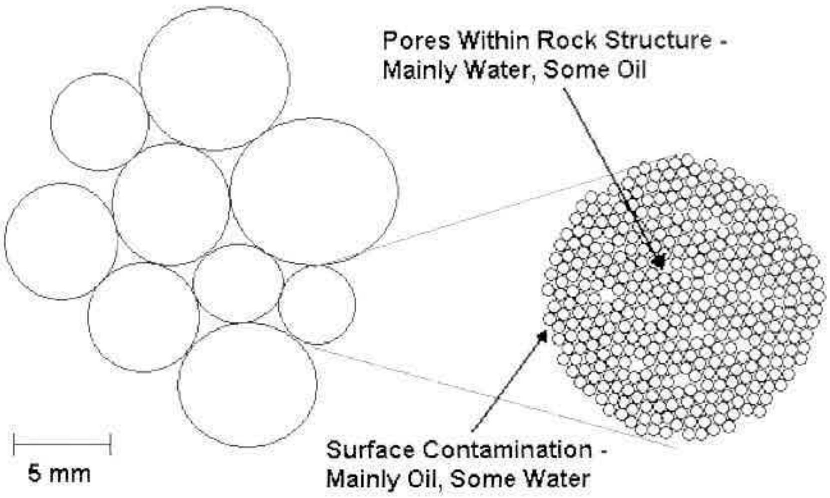


Fig. 4. Diagram showing the make-up of oil contaminated drill cuttings.

externally and within the pores of the cuttings. The oil is not present in the rock formations, and occurs only as a contaminant from the drilling fluid. For this reason most of the oil occurs on the surface of the contaminated cuttings, with small amounts likely to have diffused into the pore structure during storage or been forced into the pores in high-pressure well sections. Fig. 4 shows the structure of the drill cuttings and the areas where oil and water occur. Fixed-bed gas stripping processes appear to be adequate for the removal of the surface contaminants, but are inefficient at removing the oil which is trapped within the pores of the material. The bulk of the gas stream passes between particles in the bed and not within the pore structure of the cuttings, therefore any oil contained in the pores is not heated appreciably by the gas passing through the bed. Without the addition of heat the oil cannot thermally desorb, relying only on diffusion toward the surface of the material where it can be removed in the stripping gas. If energy was supplied within the pore structure rather than at the surface, thermal desorption could be induced and rates of diffusion could be enhanced. It is here that the volumetric heating phenomenon associated with microwave heating offers considerable advantages over conventional gas stripping processes.

3.2. *Microwave-assisted stripping*

Equivalent experiments were performed in which contaminated drill cuttings were treated with a stripping gas but also subjected to a single-mode microwave field. A microwave power of 0.8 kW was applied throughout the duration of the test, with treatment time as the variable. The results are demonstrated in Figs. 5 and 6 for nitrogen and steam, respectively.

In both cases the application of microwaves results in improved levels of oil removal and also much faster desorption kinetics. After 2 min of treatment the microwave-assisted nitrogen stripping process was able to

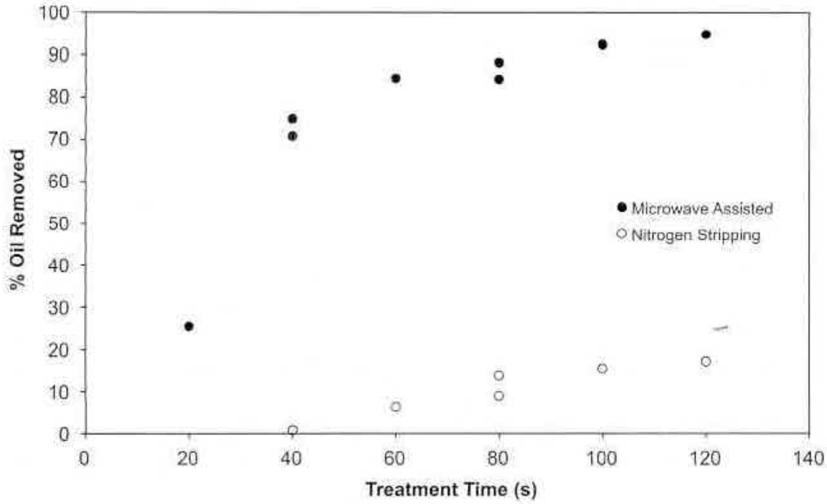


Fig. 5. Comparison of fixed-bed nitrogen stripping process with and without the application of microwaves.

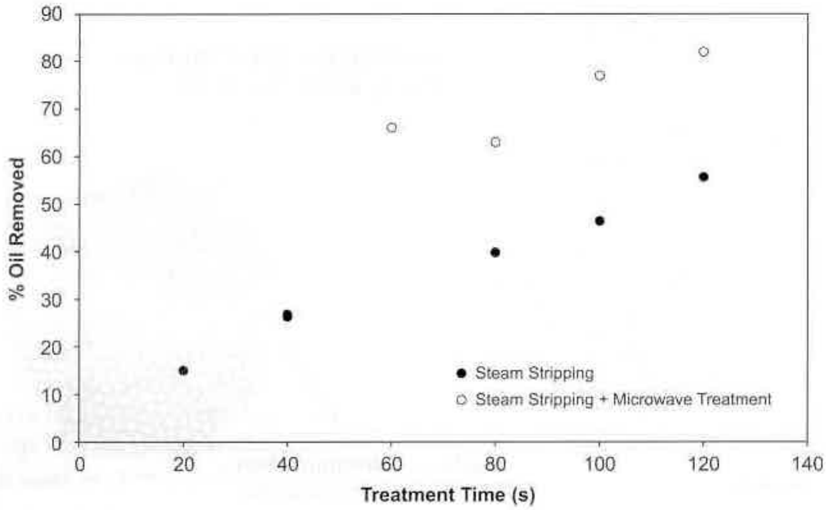


Fig. 6. Comparison of fixed-bed steam stripping process with and without the application of microwaves.

remove 95% of the oil compared to just 18% for the gas alone. This is due to the rapid dissipation of microwave energy within the depth of the bed of contaminated material and results in an order of magnitude increase in the rate of desorption. The water trapped within the pores absorbs the microwave energy and converts to steam, which then entrains the contaminant oil as it passes from the pores into the bulk gas. Measurement of temperatures within the bed in a microwave environment is not a trivial exercise, and could not be performed in this particular experimental set-up. Given that water is the only microwave absorbing species within the contaminated matrix (see Table 1), it is unlikely that the temperature will exceed 100 °C as the formation of steam will leave behind a material which is depleted in its microwave absorbing capability. In the case of steam stripping the microwave treatment again improved the desorption kinetics but not to the same extent as with the nitrogen process, with 80% oil removal achieved after 2 min compared to 95% with nitrogen. Intuitively one would expect an even greater improvement when microwaves are applied given that steam alone is able to remove 55% of the oil; however, the findings can potentially be explained by considering the extra deposition of water caused by the passing steam.

Table 1
Dielectric properties and penetration depths of various substances at 25 °C and 2.45 GHz

Substance	ε'	ε''	D_P (cm)
Fuel oil	2.0	0.002	1378
Feldspar	2.6	0.02	157
Quartz	3.8	0.001	3799
Mica	1.6	0.005	493
Water	77	13	1.3

The extra water deposited within the sample increases the overall dielectric loss factor (ε'') of the material, so from Eq. (1) the indication is that the Pd will be increased. This is likely to be the case, however, the deposited water occurs on the surface of the sample and within the macropores between agglomerates; however, the majority of the oil which occurs in these regions will have been removed so the microwave energy dissipated in this area is being expended on vapourising more water rather than desorbing the contaminant oil. The extra water present on the material surface and within the macropores causes some of the microwave energy to be absorbed, meaning that less penetrates into the depth of the sample where it is required. Referring to Eq. (1), the presence of surface water means that $|E|$ is diminished within the depth of the material and hence the Pd and desorption rate are also decreased. Any dielectric material will cause the electric field intensity to diminish throughout the material depth as energy is dissipated. A nominal quantity termed the penetration depth (D_P) is defined as the depth at which the electric field intensity falls to $1/e$ (0.368) of its surface value, and may be approximated by

$$D_P \approx \frac{\lambda\sqrt{\varepsilon'}}{2\pi\varepsilon''}. \quad (2)$$

The electric field intensity within the bed of contaminated drill cuttings therefore depends on the material dielectric properties and also the wavelength of the microwaves. Substances with relatively high values of ε' and ε'' yield lower penetration depths than materials which are more microwave transparent. Table 1 details the penetration depths for the components of the contaminated drill cuttings used in this work and at the operating frequency of 2.45 GHz.

In fuel oils and those minerals present in shales, clays and sandstones, the penetration depth is orders of magnitude greater than in water. Although the value of 1.3 cm for water is appreciably lower than that for the surrounding rock material, it does not totally prevent the microwaves from reaching into the depth of the bed within the reactor. What the penetration depth calculations do suggest, however, is that the electric field intensity in the centre of a wet bed of material will be less than that in a bed with no surface water present. The extra surface water deposited by the steam therefore causes microwave energy to be dissipated at the surface of the bed, and reduces the Pd within the pores. For this reason, a process which uses a dry gas (nitrogen in this case) will allow the microwaves to target the oil trapped deep within the material pores and result in both greater removal efficiency and improved desorption kinetics.

3.3. Energy requirements

The energy consumptions of the pure gas and microwave-assisted stripping processes were calculated based on the sensible and latent heats of the gas and the applied microwave power. Since microwaves are used in combination with the stripping gases it is clear that more power will be required for the microwave process albeit for less time. The energy consumption values are shown plotted against the amount of contaminant oil removed so that a meaningful comparison can be made, and these data are shown in Fig. 7.

It is shown that steam stripping is the most inefficient process, with the microwave-assisted and nitrogen stripping processes requiring approximately the same energy input to recover 60% of the oil. To attain oil recoveries above 60% the microwave-assisted process appears to be more efficient. The data shown in Fig. 7 can be combined with the desired process specifications to enable the appropriate process route to be chosen. For example, to recover thermally sensitive compounds of a high value then steam stripping may be the only technically viable option, despite the higher energy costs. To remove over 60% of the contaminant, or if rapid kinetics are required, then

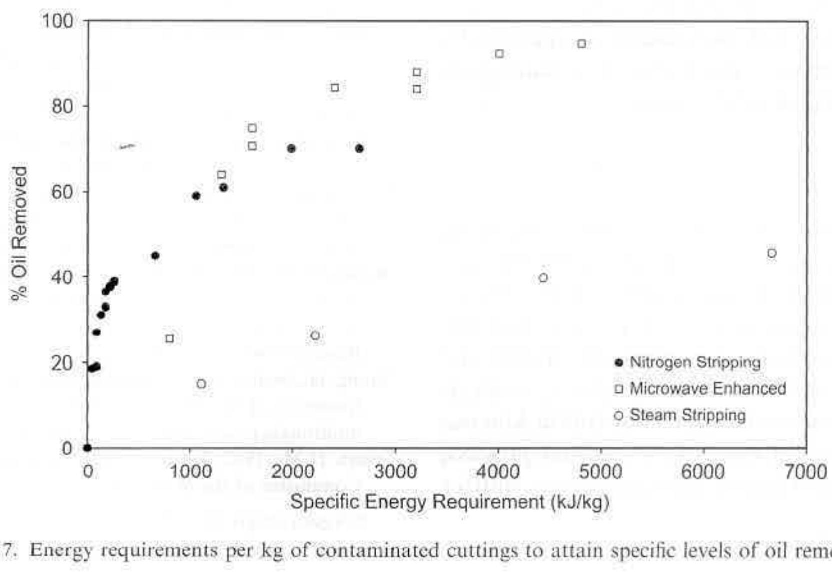


Fig. 7. Energy requirements per kg of contaminated cuttings to attain specific levels of oil removal.

microwave-assisted stripping processes offer significant advantages over conventional gas stripping. Finally, if only a small fraction of the contaminants are to be removed and if slow desorption kinetics are acceptable then nitrogen stripping will offer similar efficiencies as the microwave treatment process but without the need for the installation of microwave hardware. Of note is that the temperature of the stripping gas will also influence the kinetics and efficiency of oil remediation, and this will be addressed in future work.

4. Future development and applications

This article has shown how microwave treatment can be applied to gas stripping processes, and the example of contaminated drill cuttings can be applied to a range of other systems where energy needs to be supplied within the depth of solid or granular materials which are heterogeneous in nature. Potential examples include the remediation of contaminated soils, the extraction of oil from tar sands and the rapid drying of powders. Of particular interest is the molecular weight of the contaminant molecules that can be removed using microwave heating as this study has focused on a relatively light organic contaminant. The issue of penetration depth highlighted in this article presents a significant challenge in the scale-up of microwave processes for industrial applications, where the quantities of material for processing will be such that the penetration depth could be exceeded. In this case the process mechanisms and knowledge obtained in batch experiments at 2.45 GHz can be used to design specific microwave application systems at lower frequencies (typically 896 or 433 MHz), where the penetration depth is increased. The design and implementation of such systems requires extensive knowledge of the dielectric properties of the process materials, along with detailed modelling of electric field distributions and interactions with the microwave cavity and the process material. The importance of materials handling and process engineering issues is also a major factor in scale-up, and the authors' current work involves the construction and testing of continuous microwave treatment systems at pilot scale. The findings of that study will be reported at a later date.

5. Conclusions

It has been shown that microwave-assisted stripping offers clear process advantages over conventional gas stripping operations. For removal of more than 75% of the contaminant the microwave-assisted process was the only one which was technically feasible, and offered the most efficient process route for contaminant removal levels over 60%. It is also demonstrated that desorption kinetics are greatly improved using a microwave-assisted process, with the potential benefit of higher throughputs or shorter process times. Where only a small amount of remediation is required a nitrogen stripping process is shown to be adequate provided the slow desorption kinetics are acceptable. Steam stripping was shown to be the least energy efficient technology, and was also less technically capable than nitrogen stripping when used in conjunction with a microwave process due to issues of penetration depth. The ability of steam stripping processes to desorb thermally sensitive organic contaminants at low temperatures is a clear advantage; however, the application of microwaves to this process is shown to improve the desorption rate.

References

Braass, O., Tiffert, C., Hohne, J., Luo, X., Niemeyer, B., 2003. Decontamination of polyaromatic hydrocarbons from soil by steam stripping: mathematical modelling of the mass transfer and energy requirement. Environmental Science and Technology 37, 5001–5007.

Buettner, H.M., Daily, W.D., 1995. Cleaning contaminated soil using electrical heating and air stripping. Journal of Environmental Engineering 121 (8), 580–589.

Di, P., Chang, D.P.Y., Dwyer, H.A., 2000. Heat and mass transfer during microwave steam treatment of contaminated soils. Journal of Environmental Engineering 126 (12), 1108–1115.

Guigard, S.E., Odusanya, O.O., 2004. Treatment of oil contaminated drill cuttings using supercritical carbon dioxide. Petroleum Technology Alliance Canada.

Harrison, J., Zwinderman, M., 2004. The practical application of microemulsion technology for waste minimisation offshore. In: Proceedings of Drilling Operations—the Environmental Challenge, SPE International, Aberdeen.

Jones, P., 2004. Drilling waste management—the potential for zero landfill. In: Proceedings of Drilling Operations—the Environmental Challenge, SPE International, Aberdeen.

Lord Jr., A.E., 1998. Steam stripping organics from "fluidized" soils (versus confined soils). Journal of Environmental Engineering 124 (4), 390–394.

Melchor, A.E., da Costa, A., Rodriguez, C., Pena, J.R., 2002. E&P waste management in the Orinoco delta. SPE Drilling and Completion 17 (3), 164–173.

Meredith, R., 1998. Engineers Handbook of Industrial Microwave Heating. The Institution of Electrical Engineers, London.

Pankhania, M., 2004. The eco-bio process—a breakthrough technology in the treatment of oily wastewaters using biocatalysis. In: Proceedings of Drilling Operations—the Environmental Challenge, SPE International, Aberdeen.

Perry, R.H., Green, D.W., 1997. Perry's Chemical Engineers Handbook, seventh ed. McGraw-Hill, New York.

Pina, J., Merino, J., Errazo, A.F., Bucala, V., 2002. Thermal treatment of soils contaminated with gas oil: influence of soil composition and treatment temperature. Journal of Hazardous Materials B 94, 273–290.

Shang, H., Snape, C.E., Kingman, S.W., Robinson, J.P., 2005. Treatment of oil contaminated drill cuttings by microwave heating in a high power single mode cavity. Industrial & Engineering Chemistry Research 44, 6837–6844.

Shang, H., Snape, C.E., Kingman, S.W., Robinson, J.P., 2006. Microwave treatment of oil-contaminated North Sea drill cuttings in a high power multimode cavity. Separation and Purification Technology 49, 84–90.

Spiers, H.M., 1962. Technical Data on Fuel, sixth ed. In: British National Committee of the World Power Conference, London.

ELSEVIER

Journal of Environmental Management 88 (2008) 219–252

Journal of Environmental Management

www.elsevier.com/locate/jenvman

A holistic framework for design of cost-effective minimum water utilization network

S.R. Wan Alwi[a], Z.A. Manan[a,*], M.H. Samingin[b], N. Misran[b]

[a]*Chemical Engineering Department, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia*
[b]*MIMOS Semiconductor (M) Sdn. Bhd., Technology Park Malaysia, 57000 Kuala Lumpur, Malaysia*

Received 22 August 2006; received in revised form 25 January 2007; accepted 17 February 2007
Available online 20 April 2007

Abstract

Water pinch analysis (WPA) is a well-established tool for the design of a maximum water recovery (MWR) network. MWR, which is primarily concerned with water recovery and regeneration, only partly addresses water minimization problem. Strictly speaking, WPA can only lead to *maximum water recovery targets* as opposed to the *minimum water targets* as widely claimed by researchers over the years. The *minimum water targets* can be achieved when all water minimization options including elimination, reduction, reuse/recycling, outsourcing and regeneration have been holistically applied. Even though WPA has been well established for synthesis of MWR network, research towards holistic water minimization has lagged behind. This paper describes a new holistic framework for designing a cost-effective minimum water network (CEMWN) for industry and urban systems. The framework consists of five key steps, i.e. (1) Specify the limiting water data, (2) Determine MWR targets, (3) Screen process changes using water management hierarchy (WMH), (4) Apply *Systematic Hierarchical Approach for Resilient Process Screening* (SHARPS) strategy, and (5) Design water network. Three key contributions have emerged from this work. First is a hierarchical approach for systematic screening of process changes guided by the WMH. Second is a set of four new heuristics for implementing process changes that considers the interactions among process changes options as well as among equipment and the implications of applying each process change on utility targets. Third is the SHARPS cost-screening technique to customize process changes and ultimately generate a minimum water utilization network that is cost-effective and affordable. The CEMWN holistic framework has been successfully implemented on semiconductor and mosque case studies and yielded results within the designer payback period criterion.

Keywords: Minimum water network; Water management hierarchy; Water pinch analysis; Maximum water recovery; SHARPS

1. Introduction

The world is on the brink of a water crisis. United Nations in World Water Day 2002 has warned more than 2.7 billion people will face severe water shortages by 2025 (BBC News, 22 March 2002). Shortage of quality raw water, rising costs of water management and drive towards environmental sustainability have encouraged widespread water conservation efforts and stimulated the development of systematic techniques for water minimization.

Over the past decade, the advent of water pinch analysis (WPA) as a tool for the design of a maximum water recovery (MWR) network has been one of the most significant advances in the area of water minimization. Since its introduction by Wang and Smith (1994), various noteworthy WPA developments on targeting, design and improvement of an MWR network have emerged. These include works on processes with fixed flow rate and fixed concentration (Dhole et al., 1996; Sorin and Bédard, 1999; Hallale, 2002; Manan et al., 2004a), regeneration targeting (Kuo and Smith, 1998; Castro et al., 1999), numerical water targeting (Manan et al., 2004), network design to

*Corresponding author. Tel.: +60 7 5535512 or +60 7 5535609; fax: 60 7 5581463.
E-mail addresses: shasha@fkkksa.utm.my (S.R. Wan Alwi), zain@fkkksa.utm.my (Z.A. Manan).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.02.011

Nomenclature

Acronym

AHU air handling units
BOD biological oxygen demand
CEMWN cost effective minimum water network
COD chemical oxygen demand
CT cooling tower
D demand
DI deionized water
EDI electrodeionization
Fab fabrication plant
FW freshwater
HF hydrogen fluoride
IAS net capital investment vs. net annual savings plot
IPA isopropyl-butanol
IWT industrial wastewater treatment
MAU make-up air units
MMF multimedia filter
MWN minimum water network
MWR maximum water recovery
MySem semiconductor company
NAS net annual savings
NCI net capital investment
ppm parts per million
RO reverse osmosis
RW rainwater
S source
SHARPS systematically hierarchical approach for resilient process screening
TDS total dissolved solids
TSS total suspended solids
UF ultra filtration
UPW ultra pure water
UV ultraviolet
WB wet bench
WCA water cascade analysis
WPA water pinch analysis
WMH water management hierarchy
WPA water pinch analysis
WW wastewater

Symbols

C concentration
C_C costs per unit time for chemicals used by water system
$CC_{\text{base case}}$ capital cost of base case water system
$CC_{\text{new system}}$ capital cost of new water system
C_{EOC} costs per unit time for energy for water processing
C_{FW} costs per unit time for freshwater
C_{IC} instrumentation cost
C_{PE} purchased equipment cost
C_{PEI} purchased equipment installation cost
C_{piping} piping cost
C_{WW} costs per unit tile for wastewater disposal
F flow rate
$F_{\text{Demand initial}}$ initial demand flow rate
F_{DI} deionized water flow rate
$F_{\text{EDI initial}}$ initial electrodeionization flow rate before analysis
$F_{\text{EDI new}}$ new electrodeionization flow rate after analysis
F_{FW} freshwater flow rate
$F_{\text{FW initial}}$ initial freshwater flow rate before analysis
$F_{\text{FW new}}$ new freshwater flow rate after analysis
$F_{\text{HeaterWB101 initial}}$ initial heater WB101 flow rate before analysis
$F_{\text{HeaterWB101 new}}$ new heater WB101 flow rate after analysis
$F_{\text{Internal new}}$ new internal pumping flow rate after analysis
$F_{\text{Internal initial}}$ initial internal pumping flow rate before analysis
F_{IWT} industrial wastewater flow rate
$F_{\text{IWT initial}}$ initial industrial wastewater flow rate before analysis
$F_{\text{IWT new}}$ new industrial wastewater flow rate after analysis
$F_{\min}$ minimum point
$F_{\text{MMF initial}}$ initial multimedia filter inlet flow rate before analysis
$F_{\text{MMF new}}$ new multimedia filter inlet flow rate after analysis
F_{MU} minimum utility flow rate
F_{opt} optimum point
$F_{\text{outsource}}$ outsource flow rate
F_{reg} regeneration flow rate
F_{reuse} reuse flow rate
m gradient
m_i gradient of line i
$OC_{\text{base case}}$ operating cost of base case water system
OC_{new} operating cost of new water system
P purity
TPP total payback period
TPP_{AS} total payback period after SHARPS
TPP_{BS} total payback period before SHARPS
TPP_{set} desired payback period specified by designer
$\sum$ summation

achieve water targets (Wang and Smith, 1994; Takama et al., 1980; Olesen and Polley, 1997; Castro et al., 1999; Polley and Polley, 2000; Hallale, 2002; Prakash and Shenoy, 2005), problems with multiple contaminants (Huang et al., 1999; Benkó et al., 2000; Xu et al., 2003; Koppol et al., 2003; Ullmer et al., 2005), water network retrofit (Tan, 2005) and water targeting for batch systems (Wang and Smith, 1995; Foo et al., 2004, 2005).

Wan Alwi et al. (2004) had recently made the first attempt to implement WPA on urban system by using their *Water Cascade Analysis* (WCA) technique to establish water targets and design an MWR network for a mosque. Geldermann et al. (2005) introduce the multi objectives pinch analysis (MOPA) which combines multi criteria targets for minimum energy, wastewater and volatile organic compounds. Mariano-Romero et al. (2005) incorporates the minimum freshwater consumption and infrastructure target for multiple contaminant systems. Most researchers claim that their methods ultimately lead to the minimum freshwater and wastewater targets.

It is important to note that the concept of MWR which relates to maximum reuse, recycling and regeneration (partial treatment before reuse) of spent water has two limitations. Firstly, MWR only partly addresses the issue of water minimization which should holistically consider all conceivable methods to reduce water usage through elimination, reduction, reuse/recycling, outsourcing and regeneration (Manan and Wan Alwi, 2006). Process modifications such as elimination and reduction should be among the prime strategies to consider for water minimization. Regenerating wastewater without considering the possibility of elimination and reduction may lead to unnecessary treatment units. Secondly, since MWR focuses on water reuse and regeneration, strictly speaking, it does not lead to the *minimum water targets* as widely claimed by researchers over the years. We hereby term the water targets associated with MWR network as the *maximum water recovery targets* (MWR targets). On the other hand, the *minimum water targets* can only be achieved when all options for water minimization have been holistically applied.

Even though WPA has been well established for synthesis of MWR network, research towards water conservation from the holistic water minimization viewpoint as mentioned previously has significantly lagged behind. The use of water minimization strategies beyond recycling was first introduced by El-Halwagi (1997) who proposes a targeting technique involving water elimination, segregation, recycle, interception and sink/source manipulation. Hallale (2002) introduces guidelines for reduction and regeneration based on WPA. However, the piece-meal water minimization strategies proposed do not consider interactions among the process changes options as well as the "knock-on effects" of process modifications on the overall process balances, stream data and the economics. There is a clear need to develop a framework to address water minimization holistically, systematically and cost effectively.

This work describes a new holistic framework for water minimization applicable to industry and urban sectors. Two key features of the new framework are the water management hierarchy (WMH) as a guide to prioritize process changes and the *Systematic Hierarchical Approach for Resilient Process Screening* (SHARPS) strategies as a new cost-screening technique. We began by explaining WMH as a foundation for the holistic framework. This is followed by descriptions of a five-step methodology for designing a cost-effective minimum water utilization network (CEMWN). We then demonstrate the step-wise application of CEMWN methodology on a semi-conductor plant and a mosque. The paper concluded by comparing the outcomes of applying various approaches for water minimization.

2. The water management hierarchy

Fig. 1 shows the WM hierarchy consisting of five levels, namely (1) source elimination, (2) source reduction, (3) direct reuse/outsourcing of external water, (4) regeneration, and (5) use of freshwater. Each level represents various water management options. The levels are arranged in order of preference, from the most preferred option at the top of the hierarchy (level 1) to the least preferred at the bottom (level 5). Water minimization is concerned with the first to the fourth level of the hierarchy.

Source elimination at the top of the hierarchy is concerned with the complete avoidance of freshwater usage. Sometimes it is possible to eliminate water rather than to reduce, reuse or recycle water. Examples include using alternative cooling media such as air instead of water. Even though source elimination is the ultimate goal, often, it is not possible to eliminate water completely. One must then try to reduce the amount of water being used at the source of water usage, i.e., certain equipment or process. Such measure is referred to as *source reduction*, which is the next best option in the WM hierarchy (level 2). Examples of source reduction equipment include water saving toilet flushing system and automatic tap.

When it is not possible to eliminate or reduce freshwater at source, wastewater recycling should be considered. Levels 3 and 4 in the WM hierarchy represent two different

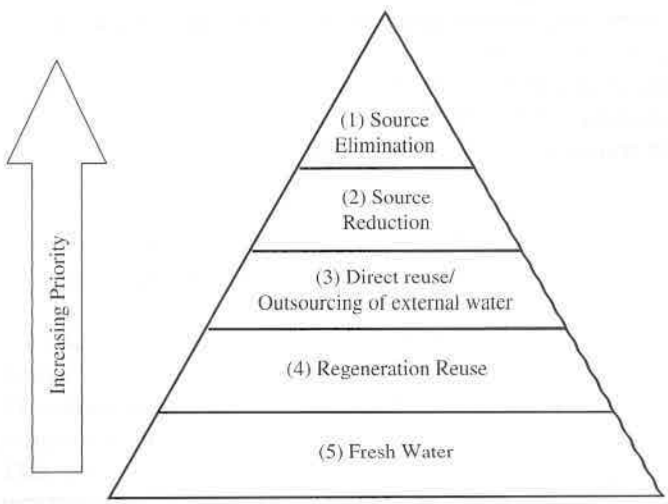


Fig. 1. The water management hierarchy.

modes of water recycling—*direct reuse/outsourcing* (level 3) and *regeneration reuse* (level 4). Direct reuse or outsourcing may involve using spent water from within a building or using an available external water source (e.g. rainwater or river water). Through direct reuse (level 3), spent water or external water source is utilized to perform tasks which can accept lower quality water. For example, wastewater from a toilet wash basin may be directly channeled to a toilet bowl for toilet flushing. Rainwater, on the other hand may be used for tasks which need higher quality water such as for ablution. However, in most domestic applications, regeneration (level 4) may be necessary prior to recycling.

Regeneration refers to treatment of wastewater or even external water source to match the quality of water required for further use. There are two possible cases of regeneration. Regeneration–recycling involves reuse of treated water in the same equipment or process after treatment. Regeneration-reuse involves reuse of treated water in other equipment after treatment. To increase water availability, the water composite curves and the pinch concentration can be used to guide regeneration of water sources as follows (Hallale, 2002):

1. *Regeneration above the pinch*: Water source(s) in the region above the pinch are partially treated to upgrade its purity.
2. *Regeneration across the pinch*: Water source(s) in the region below the pinch are partially treated to achieve purity higher than the pinch purity.
3. *Regeneration below the pinch*: Water source(s) in the region below the pinch are partially treated to upgrade its purity. However, the resulting water source is still maintained below the pinch.

Note that, regeneration above and across the pinch will reduce the freshwater consumption and wastewater generation while regeneration below the pinch will only reduce wastewater generation. Fig. 2 defines the regions above, below and across a pinch point in a water cascade table.

Freshwater usage (level 5) should only be considered when wastewater cannot be recycled or when wastewater needs to be diluted to obtain a desired purity. Note that wastewater has to undergo *end-of-pipe* treatment before discharge to meet the environmental guidelines. Use of freshwater is the least desirable options from the water minimization point of view and is to be avoided whenever possible. Through the WM hierarchy, the use of freshwater may not be eliminated, but it will become economically legitimate.

3. A holistic framework for design of cost effective minimum water network (CEMWN)

The cost effective minimum water network (CEMWN) design procedure is a holistic framework for water management applicable to industry and urban sectors. Fig. 3 illustrates five key steps involved in generating the CEMWN, i.e. (1) Specify limiting water data, (2) Determine the maximum water recovery (MWR) targets, (3) Screen process changes using water management hierarchy (WMH) and (4) Apply *Systematic Hierarchical Approach for Resilient Process Screening* (SHARPS) strategies (5) Design CEMWN. The first step is to identify the appropriate water sources and water demands having potential for integration. The next step is to establish the MWR targets using water cascade analysis (WCA) technique by Manan et al. (2004). The WMH along with a set of new process screening heuristics is then used to guide process changes to achieve the minimum water targets. The fourth step is to use SHARPS strategies to economically screen inferior process changes. The CEMWN is finally designed using established techniques for design of water network. The step-wise approach is described in detail next.

3.1. Step 1: specify the limiting water data

The first step was to specify the limiting water data. This involved process line-tracing, establishing process material

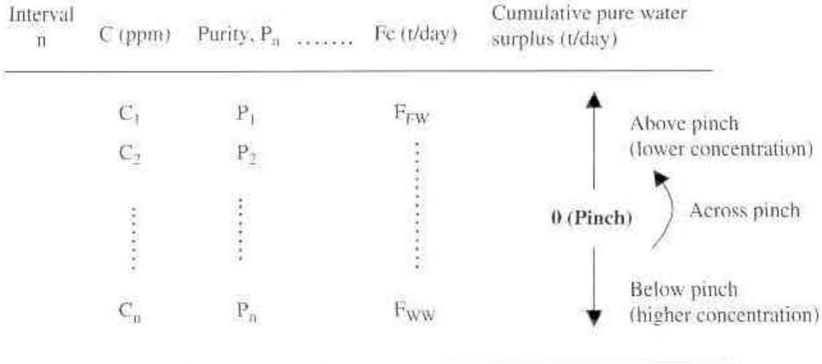


Fig. 2. Pinch location and concentration regions in a water cascade table.

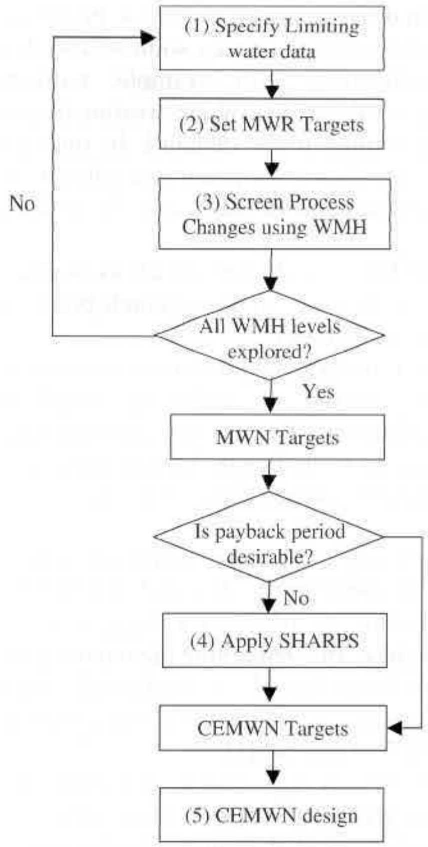


Fig. 3. A holistic framework to achieve CEMWN.

balances and isolating the appropriate water sources (outlet streams with potential to be recycled) and water "demands" (inlet streams representing process water requirements) having potential for integration. The water sources and demands were listed in terms of quantity (flow rate) and quality (contaminant concentration). In a water-intensive process plant, specifying the limiting data is a very tricky and time-consuming exercise and is typically the bottleneck, and more importantly, the critical success factor for a water minimization project. To isolate the relevant limiting data, readers are referred to Liu et al. (2004). Practical steps and rule-of-thumbs for selecting candidate process units for water-saving projects, extracting the right data, preparing a water balance diagram and isolating the candidate water sources and demands are discussed in detail.

The problem can be modelled as a single contaminant or a multiple contaminant system based on the water quality requirement of a process plant. Most of the mathematical modeling-based multi objective water pinch analysis (MOPA) solution by previous workers implements the multiple contaminant approach. However, the problem involving multiple contaminants involves complex modelling procedure and may be difficult to set up especially by industrial users.

An alternative is to use aggregated contaminants such as total suspended solids (TSS) or total dissolved (TDS) solids that allows multiple quality factors to be modeled as a single contaminants system (The Institution of Chemical Engineers, 2000). The aggregated contaminants modelled as a single contaminant are known as pseudo-single contaminant system. The proposed network then needs to be reassessed by checking that all other contaminant concentrations not considered are still within allowable limits before implementation. The single contaminant approach is used in this work. Note that the MOPA approach can be incorporated in the CEMWN framework to consider multiple contaminant systems and to address problem dimensionality.

3.2. *Step 2: determine the maximum water recovery (MWR) targets*

The second step was to establish the *base-case* maximum water recovery (MWR) targets, i.e. the overall freshwater requirement and wastewater generation. Note that the *base-case* MWR targets exclude other levels of WMH except re-use and recycling of available water sources and mixing of water sources with freshwater to satisfy water demands.

Established graphical and numerical techniques for setting the MWR targets are widely available. Some popular ones like the concentration composite curves (graphical approach—Wang and Smith (1994) and Liu et al. (2004)), concentration interval table for mass exchange network (numerical approach—El-Halwagi and Manousiouthakis, 1989) and mass problem table (numerical approach—Castro et al., 1999) however are only ideal for fixed flow rate cases where water-using processes are modelled as mass-transfer based operations involving water as a lean stream or a mass separating agent (MSA). For an industrial project where flow rate gains and losses are quite common, it may be necessary to analyze these streams separately and modify the stream data as done by Liu et al. (2004) if the fixed-flow rate approach is used. A resilient tool should be able to handle not just mass-transfer based but also non-mass transfer-based water using-operations involving flow rate gain or losses which include water used as a solvent or withdrawn as a product or a byproduct in a chemical reaction, or utilized as heating or cooling media. The water cascade analysis (WCA) technique by Manan et al. (2004) which fit the latter category was used in this work.

3.3. *Step 3: screen process changes using water management hierarchy (WMH)*

Changes can be made to the flow rates and concentrations of water sources and demands to reduce the Maximum Water Recovery (MWR) targets and ultimately achieve the Minimum Water Network (MWN) benchmark. This was done by observing the basic pinch rules for process changes and by prioritizing as well as assessing all possible process changes options according to the WM hierarchy. The fundamental rules to change a process depend on the location of water sources and demands relative to the pinch point of a system:

(i) Above the pinch—beneficial changes can be achieved by either increasing the flow rate or purity of a source or by decreasing the flow rate or purity requirements of a demand. These changes will increase water surplus above the pinch thereby reducing the amount of freshwater required.
(ii) Below the pinch—there is already a surplus of water below the pinch, hence any flow rate change made there will not affect the target. An exception to this rule of thumb is for a case where a source purity is increased so that it moves to the region above the pinch as in the case of regeneration.
(iii) At the pinch point—increasing the flow rate of a source at the pinch concentration will not reduce the targets.

It is vital to note that implementation of each process change option will yield new pinch points and MWR targets. In addition, interactions and "knock-on effects" between the process change options should also be carefully considered. It is therefore important that each process change be systematically prioritized and assessed with reference to the revised pinch points instead of the original pinch point so as to obey the fundamental rules for process changes listed previously and to guarantee that the MWN benchmark is attained. Bearing in mind these constraints, the core of step 3 was the level-wise hierarchical screening and prioritization of process changes options using the water management hierarchy (WMH) and the following four new option-screening heuristics which was sequentially applied to prioritize process changes at each level of WMH. As described below, not all four heuristics are applicable at each level of WMH.

Heuristic 1: Begin process changes at the core of a process.

Heuristic 1 was formulated from the *onion model* for process creation (Smith, 1995). Due to interactions among reaction, separation and recycle, heat and mass exchange network and utility layers, any changes, such as demand elimination should be implemented beginning from the core of a process (reaction system) to the most outer layer (utilities). Excessive water usage at the core of the system causes wastage at the outer layers. Hence improving the core of the system first will eliminate or reduce wastage downstream.

Heuristic 1 strictly applied to the process change options at levels 1 and 2 of WMH. Applying heuristic 1 to various source elimination options at level 1 of the WMH will lead to new targets and pinch points. For mutually exclusive options, the one giving the lowest revised MWR targets was selected. Heuristic 1 was repeated to reduce water at WMH level 2 once all elimination options were explored.

Note also that it is quite common for processes to have independent and non-interacting sources and demands at various concentrations. For example, reducing water demand for a scrubber in a waste treatment system does not affect the cooling tower demand. In such a case, the demand flow rate above the pinch (see rule (i) for process changes mentioned previously) can be reduced using heuristic 2.

Heuristic 2: Successively reduce all available demands with concentration lower than the pinch point, beginning from the cleanest demand.

Note that if a dirtier demand were reduced first followed by a cleaner demand, it might be found later that subsequent reduction of a cleaner demand might cause the dirtier demand to lie below the new pinch point. Such situation makes the earlier changes to the dirtier demand meaningless.

If a few demands exist at the same concentration, it is best to begin by reducing the demand that yields the most flow rate reduction to achieve the biggest savings. Then, proceed to reduce the remaining demands that exist at concentrations lower than the revised pinch concentration, as stated in heuristic 3. Heuristics 2 and 3 were applicable to levels 1 and 2 of the WMH.

Heuristic 3: Successively reduce the demands starting from the one giving the biggest flow rate reduction if several demands exist at the same concentration.

There exists a maximum limit for adding new water sources (utilities) either obtained externally such as rain water, river water, snow and borehole water or by regenerating wastewater in order to minimize freshwater in a water distribution system. It is therefore necessary to.

Heuristic 4: Harvest outsourced water or regenerate wastewater only as needed.

Note that, the limit for adding utilities through outsourcing and regeneration corresponds to the minimum utility flow rate (F_{MU}) which leads to minimum freshwater flow rate. To calculate F_{MU}, the WCA method by Manan et al. (2004) was used to produce a plot of freshwater flow rate (F_{FW}) vs. the flow rate of external utilities (F_U) as shown in Fig. 4. External utilities added were increased until the minimum freshwater flow rate became constant. This point corresponded to the maximum possible

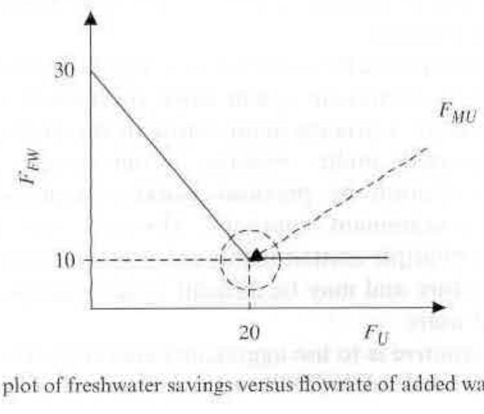


Fig. 4. A plot of freshwater savings versus flowrate of added water source.

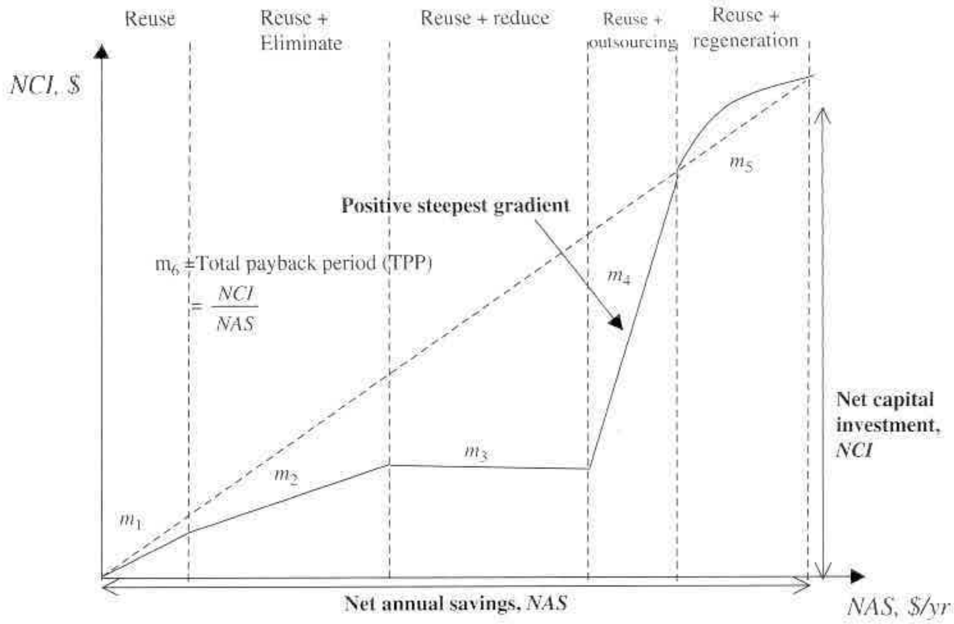


Fig. 5. IAS plot covering all levels of WM hierarchy. m_4 is the positive steepest gradient and TPP is the total payback period for a water network.

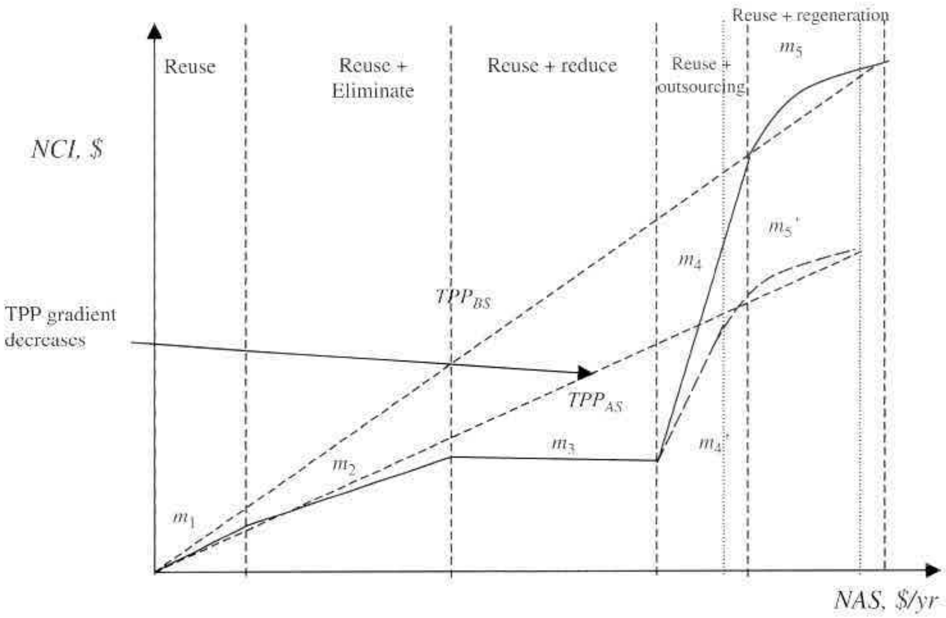


Fig. 6. IAS plot showing the revised total payback period when the magnitude of the steepest gradient is reduced using SHARPS substitution strategy.

freshwater reduction through addition of external utilities F_{MU}. A detail explanation on obtaining multiple F_{MU} target using WCA are explained in Wan Alwi (2006). Heuristic 4 only applies to levels 3 and 4 of the WMH.

The revised MWR targets as well as the four new option-screening heuristics were used as process selection criteria. The screening and selection procedure was hierarchically repeated down the WMH levels to establish the MWN targets which yielded the maximum scope for water savings. SHARPS strategy was used next to ensure that the savings achieved was cost-effective and affordable.

3.4. Step 4: apply SHARPS strategy

SHARPS screening technique involves cost estimation associated with water management (WM) options prior to detailed design. It includes a profitability measure in terms of payback period; i.e. the duration for a capital investment to be fully recovered. Note that the payback period

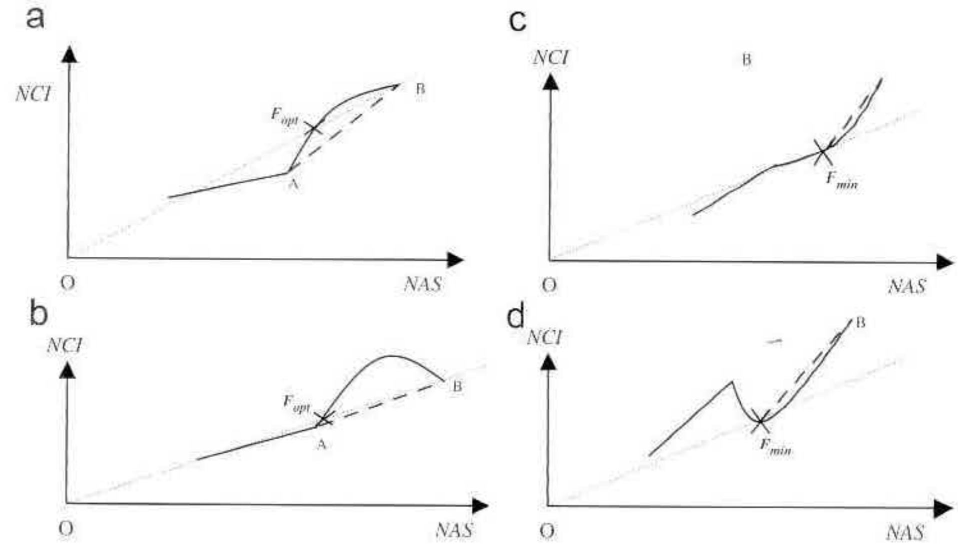


Fig. 7. Linearization of concave curves moving upwards (a) without peak (b) with peak. Convex curves moving upwards linearization (c) without valley (d) with valley.

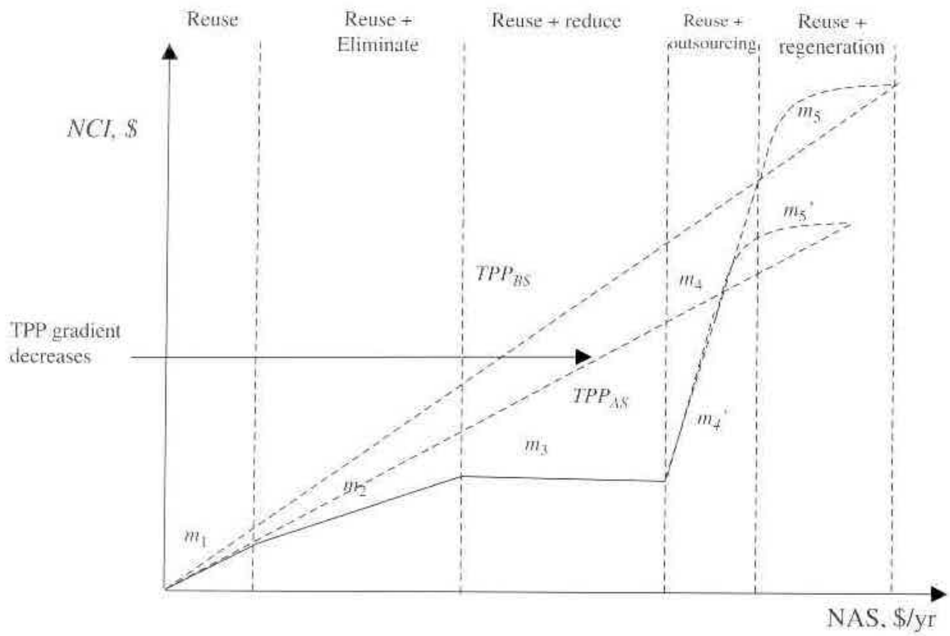


Fig. 8. IAS plot showing the revised total payback period with a shorter steepest gradient curve.

calculations for SHARPS as given by Eq. (1) only concerns the economics associated with design of a minimum water network as opposed to the design of an entire plant

$$\text{Payback period(yr)} = \frac{\text{Net Capital Investment(\$)}}{\text{Net Annual Savings(\$/yr)}}. \quad (1)$$

Since SHARPS is a cost-screening tool, standard plant design *preliminary cost estimation* techniques were used to assess the capital and operating costs of a proposed water system. The equipment, piping and pumping costs built in Eq. (2) are the three main cost components considered for a building or a plant water recovery system (Takama et al., 1980; Olesen and Polley, 1997; Hallale and Fraser, 1998; Alva-Argáez et al., 1998; Jödicke et al., 2001; Bagajewicz and Savelski, 2001; Koppol et al., 2003; Feng and Chu, 2004; Tan, 2005; Gunaratman et al., 2005). The explanation on how to obtain this cost values are described in Appendix A

$$\sum \mathrm{CC} = C_{\mathrm{PE}} + C_{\mathrm{PEI}} + C_{\mathrm{piping}} + C_{\mathrm{IC}}, \quad (2)$$

where C_{PE} is the total capital cost for the equipment; C_{PEI} the equipment installation cost; C_{piping} the water reuse piping cost investment; C_{IC} the instrumentation and controls cost investment.

The economics of employing the WM options for grassroots design as well as retrofit cases were evaluated

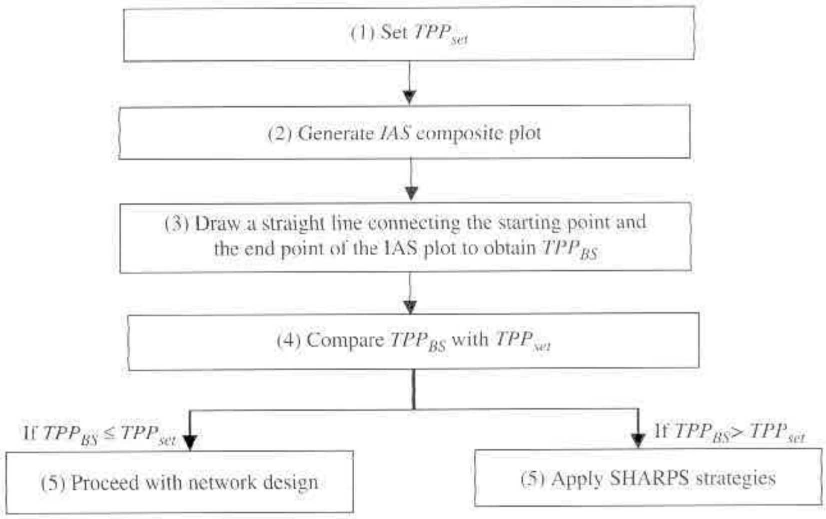


Fig. 9. The overall SHARPS procedure.

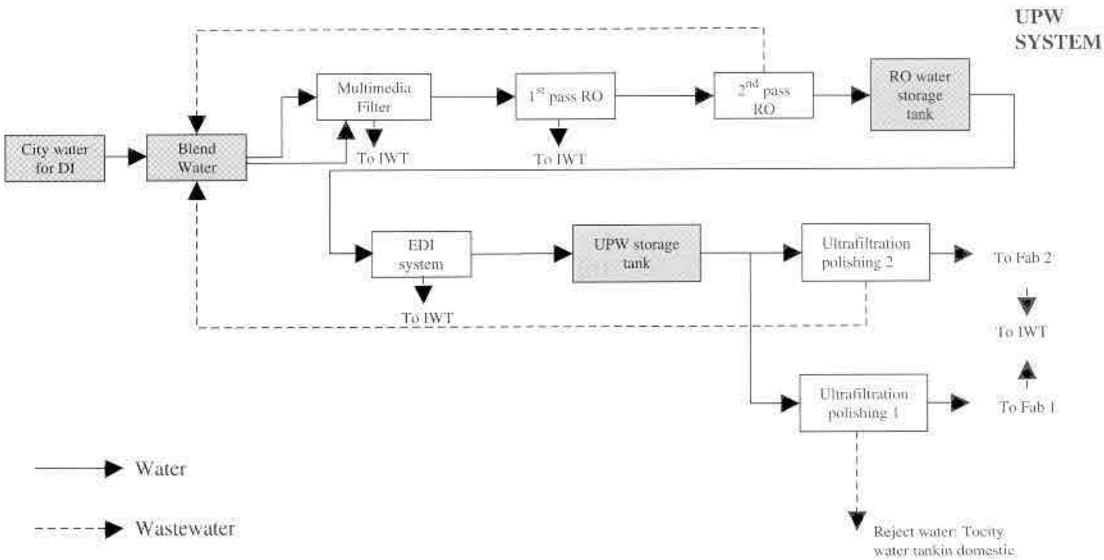


Fig. 10. MySem de-ionized (DI) plant process flow diagram.

by calculating the net capital investment (NCI) for the minimum water network (MWN) using Eqs. (3) and (4) as well as the net annual savings (NAS) using Eqs. (5) and (6). In the context of SHARPS, the NCI for grassroots refers to the cost difference between the new (substitute) equipment and the base-case equipment. The base-case equipment is the initial equipment used before CEMWN analysis. For retrofit case, the NCI covers only the newly installed (substitute) system

$$\text{Net Capital Investment, \$(grassroots)} = \sum CC_{\text{new system}} - \sum CC_{\text{base case}}, \quad (3)$$

$$\text{Net Capital Investment, \$(retrofit)} = \sum CC_{\text{new system}}, \quad (4)$$

where $CC_{\text{new system}}$ is the capital cost associated with new equipment; $CC_{\text{base case}}$ the capital cost for base-case equipment.

For example, a new \$300, 6 l toilet flush gives water savings of 6 l per flush as compared to a \$200, 12 l toilet flush (base case system). For grassroots design, the payback period is therefore based on the NCI given by Eq. (3), i.e. \$100. For retrofit, the payback period is based on the NCI given by Eq. (4), i.e. \$300.

The net annual savings (NAS) is the difference between the base-case water operating costs from the water operating costs after employing WM options as in

$$NAS = OC_{\text{base-case}} - OC_{\text{new}}, \quad (5)$$

where NAS is the Net annual savings (\$/yr); $OC_{\text{base-case}}$ the base-case expenses on water (\$/yr); and OC_{new} the new expenses on water after modifications (\$/yr).

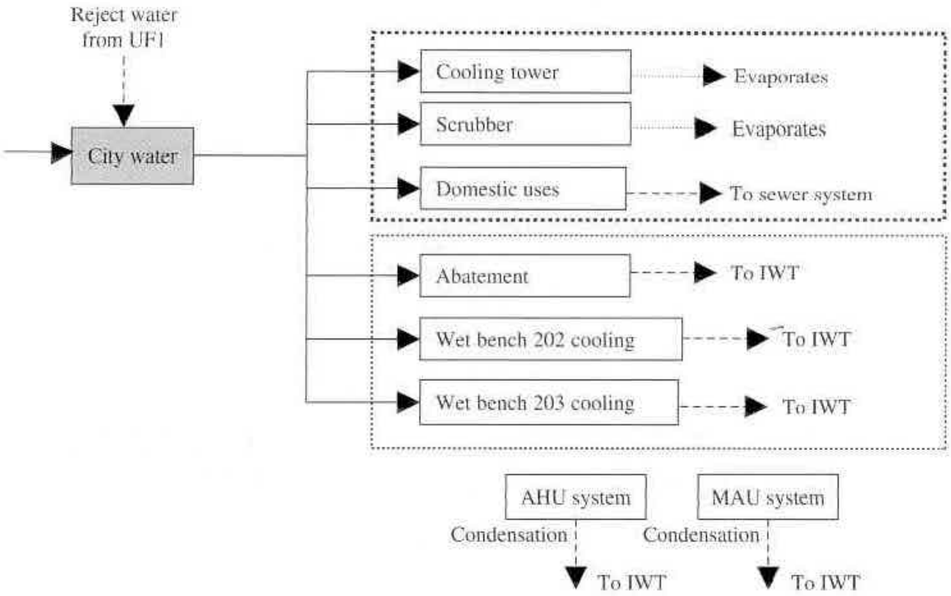


Fig. 11. Domestic and non-process water uses for MySem.

The total operating cost of a water system includes freshwater cost, effluent disposal chares, energy cost for water processing and the chemical cost as given by

$$\mathrm{OC} = C_{\mathrm{FW}} + C_{\mathrm{WW}} + C_{\mathrm{EOC}} + C_{\mathrm{C}}, \tag{6}$$

where OC is the total water operating cost; C_{FW} the costs per unit time for freshwater; C_{FD} the costs per unit time for wastewater disposal; C_{EOC} the costs per unit time for energy for water processing; C_{C} the costs per unit time for chemicals used by water system.

In order to obtain a cost-effective and affordable water network that achieves the minimum water targets (hereby termed the *cost-effective minimum water network* (CEMWN)) within a desired payback period, the new SHARPS technique was implemented as follows.

Step 1: Set the desired total payback period ($\mathrm{TPP_{set}}$). The desired payback period can be an investment payback limit set by a plant owner, e.g. 2 yr.

Step 2: Generate an investment vs. annual savings (IAS) composite plot covering all levels of the WM-hierarchy. Fig. 4 shows a sample of the IAS plot. The gradient of the plot gives the payback period for each process change. The steepest positive gradient (m_4) giving the highest investment per unit of savings represents the most costly scheme. On the other hand, a negative slope (m_3) indicates that the new process modification scheme requires lower investment as compared to the grassroots equipment.

Note that, since most equipment cost is related to equipment capacity through a power law, one is more likely to generate a curved line such as m_5. Hence, in such cases, several data points should be taken to plot a curve for each process change. As in the case of a linear line, a curve moving upward shows that more investment is

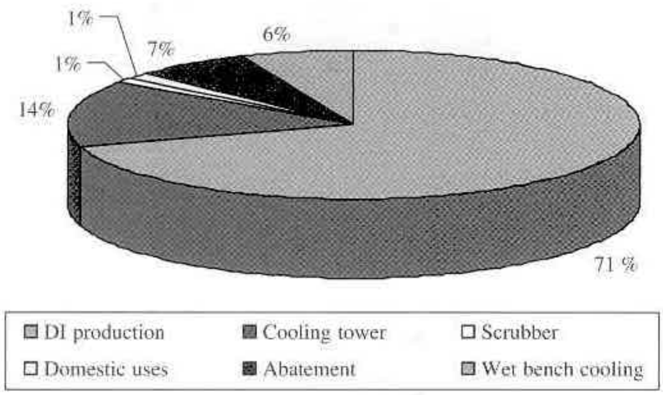


Fig. 12. Breakdown of MySem freshwater usage (October 2005).

needed and a curve moving downward shows less investment is needed with increase in annual savings.

Step 3: Draw a straight line connecting the starting point and the end point of the IAS plot (Fig. 5). The gradient of this line is a preliminary cost estimate of the total payback period (TPP) for implementing all options in line with the WM hierarchy. The $\mathrm{TPP_{BS}}$ is the total payback period before implementing SHARPS.

Step 4: Compare the $\mathrm{TPP_{BS}}$ with the $\mathrm{TPP_{set}}$ (the desired total payback period set by a designer).

The total payback period ($\mathrm{TPP_{BS}}$) should match the maximum desired payback period set ($\mathrm{TPP_{set}}$) by a designer. Thus, it is possible to tailor the minimum water network as per the requirement of a plant/building owner.

If $\mathrm{TPP_{BS}} \leqslant \mathrm{TPP_{set}}$, proceed with network design.

If $\mathrm{TPP_{BS}} > \mathrm{TPP_{set}}$, two strategies may be implemented.

Strategy 1—*Substitution*: This strategy involved replacing the equipment/process that resulted in the steepest

Table 1
Limiting water data for MySem

	Demand	F (t/h)	C (ppm)		Source	F (t/h)	C (ppm)
D1	MMF inlet	32.0	52	S1	MMF rinse	1.33	48.0
D2	Cooling tower	6.00	100	S2	RO reject 1st pass	9.80	70.4
D3	Abatement	2.73	100	S3	EDI reject	3.36	48.6
D4	Scrubber	0.54	100	S4	WB101 rinse water, idle	0.38	0
D5	Toilet flushing	0.08	100	S5	WB101 rinse water, operation	0.07	4608
D6	Wash basin	0.01	52	S6	WB102 rinse water, idle	0.22	0
D7	Ablution	0.15	52	S7	WB102 rinse water, operation	0.07	4480
D8	Toilet pipes	0.12	52	S8	WB201 rinse water, idle	0.76	0
D9	Office cleaning	0.05	52	S9	WB201 rinse water, operation	0.03	23360
D10	MMF backwash	2.08	52	S10	WB202 rinse water, idle	3.48	0
D11	MMF rinse	1.33	52	S11	WB202 rinse water, operation	0.07	163.2
D12	WB203 cooling	1.47	52	S12	WB203 rinse water, idle	3.63	0
D13	WB202 cooling	1.22	52	S13	WB203 rinse water, operation	0.28	928
Total water demands	47.78	t/h	S14	MAU	1.11	6.4	
				S15	AHU	0.36	11.5
				S16	Cassette cleaner	0.08	0
				S17	Abatement	2.73	105.6
				S18	Wafer scrubber	0.54	12.8
				S19	RO reject 2nd pass	4.50	19.2
				S20	UF1 reject	1.54	19.2
				S21	UF2 reject	1.80	0
				S22	Heater WB101	0.46	0
				S23	Wash basin	0.01	60
				S24	Ablution	0.15	40
				Total water sources	36.76	t/h	

positive gradient with an equipment/process that gave a less steep gradient. Note that this strategy did not apply to reuse line since was no equipment to replace. To initialize the composite plot, the option that gave the highest total annual water savings should be used regardless of the total investment needed. Hence, to reduce the steepest gradient according to *Strategy* 1, the process change option giving the next highest total annual savings but with lesser total investment was selected to substitute the initial process option and trim the steepest gradient. Fig. 6 shows that substituting the option causing the steepest positive gradient (m_4) with an option that gives a less steep gradient (m_4') yields a smaller TPP value. For example, a separation toilet may be changed to a much cheaper dual-flush toilet that uses a bit more water. TPP_{AS} is the TPP after implementing SHARPS strategies.

In the case of a curvature, linearization is necessary to determine the line of steepest gradient. For a projecting concave or convex curve, the linearization of a curve moving upwards is as follows:

Concave curves: Connect a straight line to the start (point A) and end (point B) points of the concave curve to obtain a positive gradient (line AB in Figs. 7a and b). Connect a line from the graph origin (point O) going through point F_{opt} to the end point (point B) of the concave curve (line O–F_{opt}–B). To have beneficial TPP reduction, the concave curve must be reduced below point F_{opt} (for Strategy 2).

Convex curve: Connect a line from the graph origin (point O) to the minimum point of the convex curve (F_{min}). Connect a line from F_{min} to the end point (point B) of the convex curve to obtain positive gradient (line F_{min}–B in Figs. 7c and d). Do not reduce further the line on the left hand side of F_{min} since this will increase TPP (for Strategy 2).

When the linearized line was the steepest positive gradient, *Strategy* 1 was implemented to yield a linearized line with a smaller gradient. Note that, the proposed linearization is only a preliminary guide to screen the most cost effective option that satisfies a preset payback period.

Strategy 2—Intensification: The second strategy involved reducing the length of the steepest positive gradient until TPP_{AS} was equal to TPP_{set}. This second strategy was also not applicable for reuse line since there was no equipment to replace. Fig. 8 shows that when the length of the steepest positive gradient (m_4) is reduced, the new gradient line (m_4') gives a less steep gradient, and hence, a smaller TPP. This means that instead of completely applying each process change, one can consider eliminating or partially applying the process change that gives the steepest positive gradient, and hence, a small annual savings compared to the amount of investment). For example, instead of changing all normal water taps to infrared-type, only 50% of the water taps were changed. If TPP_{AS} was still more than the TPP_{set} even after adjusting the steepest gradient, the length for the next steepest gradient was reduced until TPP was equal to TPP_{set}.

Table 2
Base-case maximum water recovery targets for MySem (without process changes)

Conc. C (ppm)	Purity, P	dP	Sum F demand (t/h)	Sum F source (t/h)	Total F (t/h)	Cum water flow rate (t/h)	Water surplus (t/h)	Cum water surplus (t/h)
						$F_{DI}=0$		
0	1			10.808	10.808			
		0.000006				10.808	6.92E−05	
6.4	0.999994			1.11	1.11			6.92E−05
		0.000005				11.918	6.1E−05	
11.52	0.999988			0.36	0.36			0.00013
		0.000001				12.278	1.57E−05	
12.8	0.999987			0.54	0.54			0.000146
		0.000006				12.818	8.2E−05	
19.2	0.999981			6.04	6.04			0.000228
		0.000011				18.858	0.000204	
30	0.99997			$F_{FW}=11.04$	11.04			0.000432
		0.00001				29.898	0.000299	
40	0.99996			0.15	0.15			0.000731
		0.000008				30.048	0.00024	
48	0.999952			1.33	1.33			0.000971
		0.000001				31.378	2.01E−05	
48.64	0.999951			3.36	3.36			0.000991
		0.000003				34.738	0.000117	
52	0.999948		−38.43	0	−38.43			0.001108
		0.000008				−3.692	−3E−05	
60	0.99994			0.01	0.01			0.001078
		0.00001				−3.682	−3.8E−05	
70.4	0.99993			9.8	9.8			0.00104
		0.00003				6.118	0.000181	
100	0.9999		−9.35		−9.35			0.001221
		0.000006				−3.232	−1.8E−05	
105.6	0.999894			2.73	2.73			0.001203
		0.000058				−0.502	−2.9E−05	
164	0.999836			0.069	0.069			0.001174
		0.000764				−0.433	−0.00033	
928	0.999072			0.278	0.278			0.000843
		0.003552				−0.155	−0.00055	
4480	0.99552			0.069	0.069			0.000292
		0.000128				−0.086	−1.1E−05	
4608	0.995392			0.071	0.071			0.000281
		0.018752				−0.015	−0.00028	
23,360	0.97664			0.034	0.034			0 (pinch)
		0.97664				$F_{IWT}=0.019$	0.018558	
								0.018558

Table 3
Amount of IWT and domestic wastewater not used and used for integration initially

Utility	Before MWR (t/h)	After MWR (t/h)	Reduction (%)
Total freshwater	39.94	11.04	72.4
Total IWT wastewater	34.45	5.71	83.4
Total domestic WW	0.41	0.25	39.0

Similarly, for the case of a projecting concave and convex curves moving upward, it was desirable to reduce the length of the curve until TPP_{set} was achieved, if linearization of the curve gave the steepest gradient. Both strategies 1 and 2 should be tested or applied together to yield the best savings. The overall procedure for SHARPS is summarized in Fig. 9.

3.5. Step 5: network design

Once the CEMWN targets have been established, the next step was to design a cost effective minimum water network (CEMWN) to achieve the CEMWN targets. The water network could be designed using one of the established techniques such as the one from Polley and Polley (2000); Hallale (2002) and Prakash and Shenoy (2005). The CEMWN in this work was designed using the technique from Polley and Polley (2000).

Systematic applications of the CEMWN framework on an industrial complex (a semiconductor plant) and an urban building (a mosque) are demonstrated next.

Table 4
Various process changes options applicable for MySem

WMH	Strategy	Option selected based on NAS	Option selected based on MWN procedure
Elimination	Abatement		
	Option 2 (decommisioning)	X	X
	WB 202 and 203 cooling	✓	✓
Reduction	WB reduction in Fab 1 and 2	✓	✓
	Heater reduction	✓	✓
	Fab 1 return reduction	✓	✓
	Abatement		
	Option 1 (0.5 gpm during idle)	X	X
	Option 3 (recirculation)	✓	X
	Option 4 (on demand)	X	✓
	Option 5 (pH analysis)	X	X
	Increase RO system recovery/ install 3rd stage	✓	✓
	EDI return reduction		
	Option 1 (decommissioning)	X	✓
	Option 2 (run intermittent)	✓	X
	Domestic reduction	✓	X
	Cooling tower reduction using N2	✓	✓
	MMF reduction by NTU analysis	✓	✓
Reuse	Total reuse	✓	✓
Outsourcing	RW harvesting	✓	✓
Regeneration	Treat all WB water	X	✓

(✓) for selected option. (X) for eliminated option by MySem.

4. Semiconductor plant case study

4.1. Process description

The CEMWN framework was successfully applied on a semiconductor company in Malaysia (MySem). MySem which involved a combination of domestic and process water usages represented an ideal application of the CEMWN framework for both urban and industrial sectors. Figs. 10 and 11 show the water distribution network for MySem whose primary function was research and development (R&D) as opposed to production. MySem mainly fabricated 6 and 8 in wafers as its main products. Wafers were fabricated using different recipes to meet customer demand. Water demands in MySem included DI water production, domestic uses (toilet flushing, office cleaning, wash basin, toilet pipes and ablution) and non-process uses such as abatement, scrubber, cooling tower and wet bench cooling. A breakdown of the various water demands for MySem is shown in Fig. 12.

The estimated total freshwater consumption for MySem was 42.6 t/h for the month of October. Of this value, 31.78 t/h was used for deionized (DI) water production and the rest for domestic and non-process uses. The total water consumption varied throughout the year according to wafer production and equipment conditions. MySem DI water was used for wet cleans, solvent processes, acid processes and tools cleaning.

4.2. Step 1: specify the limiting water data

This step involved detailed process survey and line-tracing, establishing process stream material balances and conducting water quality tests. Stream flow rates were either extracted from plant distributed control system (DCS) data or from online data-logging using ultrasonic flow meter. Depending on the stream audited, tests for total suspended solids (TSS), biological oxygen demand (BOD), chemical oxygen demand (COD), total dissolved solids (TDS) were made on-site. For MySem process uses which comprised entirely of ultra-pure water, TSS was found to be very negligible. BOD was eliminated since there were no biological contaminations. COD was a component of TDS. TDS was ultimately chosen as the dominant water quality parameter for MySem. TDS was monitored using a conductivity meter. Some of the key constraints considered included:

- Water streams with hydrogen fluoride (HF), isopropyl butanol (IPA) and dangerous solvents were not considered as water source.
- Multimedia filter (MMF) backwash was not considered as water source since it contained high TSS.
- WB202 and 203 cooling were not reused since it involved acid spillage.
- Black water i.e. toilet pipes, toilet flushing and office cleaning wastewater were not reused.
- Greywater could only be reused for processes which did not involve body contact.

WMH levels	Specific process changes considered	New FW target, t/hr	New IWT+WW [based on limiting data] target, t/hr	New pinch point concentration, ppm	New total IWT (all IWT considered) target, t/hr
Initial	None	39.94	34.85	22360	34.4500
Reuse	Base case (MWR)	11.0400	0.0190	22 360	5.7090
+ Elimination	Eliminate WB cooling 202 (D13) and 203 (D12)	8.3525	0.0215	22 360	3.0215
+ Reduction	WB reduction in Fab 1 and Fab 2	6.7518	0.0258	4608	1.4216
+	Heater WB201 reduction	6.7314	0.0264	4608	1.4454
	Fab 1 return reduction	6.6094	0.0354	4608	1.3109
	Option 1: EDI decommissioning	6.3038	0.0378	4608	1.0112
	Increase RO rate of recovery	6.2110	0.0380	4608	0.9132
	Reduce multimedia filter backwash and rinsing time	6.0857	0.0387	4608	0.7879
	Option 4: Reduce abatement pollution system (D3 = 0.57 t/hr and S17 = 0.57 t/hr).	6.0831	0.0361	4608	0.7853
	Cooling tower reduction (D2 = 5.86 t/hr)	5.9452	0.0382	4608	0.7874
Outsourcing	Add a source' S25=0.11 t/hr of C = 16 ppmby harvesting rainwater	5.8349	0.0379	4608	0.7871
+ Regeneration	Regenerate remaining IWT to the maximum flowrate for a source from to C=52 ppm.	5.7970	0	4608	0.7492
= **Minimum water network (MWN) targets**		**5.7970**	**0**	**4608**	**0.7492**

Fig. 13. The effects of WMH-guided process changes on the maximum water reuse/recovery targets and pinch location.

Table 5
Various effects of EDI options on water targets

EDI system	FW target (t/h)	IWT target (t/h)
Initial EDI flow rate	6.6094	0.0354
Option 1(decommission 3 EDI unit)	6.3038	0.0378
Option 2(run intermittently)	6.9281	0.0351

The relevant water streams having potential for recycling were extracted into a table of *limiting water data* comprising of process sources and demands. Table 1 shows the water sources and demands extracted for MySem listed in terms of flow rate and contaminant concentration. The contaminant (C, in ppm) in Table 1 represents TDS.

4.3. Step 2: determine the maximum water recovery (MWR) targets

This step involved establishing the base-case MWR targets using WCA technique that was incorporated in Water-MATRIX© software developed in Universiti Teknologi Malaysia (UTM). Table 2 is the water cascade table (WCT) generated by Water-MATRIX© for MySem showing the freshwater and wastewater flow rate targets at $F_{FW} = 11.04\,t/h$ and $F_{IWT} = 0.02\,t/h$, respectively. Note from Table 2 that the cleanest water targeted water at 0 ppm concentration actually referred to DI water (F_{DI}) needed to be supplied to the blend water tank instead of freshwater. This was because freshwater for MySem had a concentration of 30 ppm. The source water flow rate at 30 ppm shown in Table 2 was actually the amount of freshwater supply needed. Water-MATRIX© computes $F_{DI} = 0\,t/h$ and $F_{FW} = 11.04\,t/h$.

The F_{IWT} of 0.02 t/h shown in Table 2 only comprised of the IWT wastewater that was considered for reuse. The total IWT should also include 5.69 t/h IWT wastewater not available for reuse due to chemical contamination, giving a total IWT of 5.71 t/h to be discharged for MySem as shown in Table 3. Considering the present integration schemes implemented by MySem, the initial FW and total IWT flow rates were at 39.94 and 34.45 t/h, respectively. Eqs. (7) and (8) gave FW and IWT reduction of 72.4% and 83.4%,

Table 6
Effects of abatement options on water targets

Abatement system	F_D (t/h)	F_S (t/h)	FW target (t/h)	IWT target (t/h)
Initial abatement flow rate	2.73	2.73	6.0857	0.0387
Option 1 (0.5 gpm during idle)	1.11	1.11	6.0837	0.0367
Option 2 (decommissioning)	1.36	1.36	6.0840	0.0370
Option 3 (recirculation)	0.14	0.00	6.2183	0.0333
Option 4 (on demand)	0.57	0.57	6.0831	0.0361
Option 5 (pH analysis)	0.79	0.79	6.0833	0.0363

Table 7
MySem water targets after implementation of MWN technique

Conc. C (ppm)	Purity, P	dP	Sum F demand (t/h)	Sum F source (t/h)	Total F (t/h)	Cum water flowrate (t/h)	Water surplus (t/h)	Cum water surplus (t/h)
						$F_{DI} = 0$		
0	1.000000			1.7270	1.7270			
		0.000006				1.727000	0.000011	
6.40	0.999994			1.1100	1.1100			0.000011
		0.000005				2.837000	0.000015	
11.52	0.999988			0.3600	0.3600			0.000026
		0.000001				3.197000	0.000004	
12.80	0.999987			0.5400	0.5400			0.000030
		0.000003				3.737000	0.000012	
16.00	0.999984			0.1100	0.1100			0.000042
		0.000003				3.847000	0.000012	
19.20	0.999981			0.7760	0.7760			0.000054
		0.000011				4.623000	0.000050	
30.00	0.999970			$F_{FW} = 5.797$	5.7970			0.000104
		0.000010				10.420000	0.000104	
40.00	0.999960			0.1500	0.1500			0.000208
		0.000008				10.570000	0.000085	
48.00	0.999952			0.1690	0.1690			0.000293
		0.000001				10.739000	0.000007	
48.64	0.999951			0.8400	0.8400			0.000300
		0.000003				11.579000	0.000039	
52.00	0.999948		−6.5280	0.0381	−6.4899			0.000338
		0.000008				5.089100	0.000041	
60.00	0.999940			0.0100	0.0100			0.000379
		0.000010				5.099100	0.000053	
70.40	0.999930		–	1.1530	1.1530			0.000432
		0.000030				6.252100	0.000185	
100.00	0.999900		−7.0500		−7.0500			0.000617
		0.000006				−0.797900	−0.000004	
105.60	0.999894			0.5700	0.5700			0.000613
		0.000058				−0.227900	−0.000013	
164.00	0.999836			0.0240	0.0240			0.000599
		0.000764				−0.203900	−0.000156	
928.00	0.999072			0.0810	0.0810			0.000444
		0.003552				−0.122900	−0.000437	
4480.00	0.995520			0.0690	0.0690			0.000007
		0.000128				−0.053900	−0.000007	
4608.00	0.995392			0.0539	0.0539			0.000000
		0.995392				$F_{IWT} = 0$	0.000000	
								0.000000

respectively,

FW savings (%)

$$= \frac{\text{FW flowrate before WPA} - \text{FW flowrate targets after WPA}}{\text{FW flowrate before WPA}} \times 100. \tag{7}$$

IWT savings (%)

$$= \frac{\text{IWT flowrate before WPA} - \text{Total IWT to be discharged}}{\text{IWT flowrate before WPA}} \times 100. \tag{8}$$

4.4. Step 3: WMH-guided screening and selection of process options

After calculating the base-case MWR targets, all potential process changes to improve MySem water system were listed according to the various WMH levels as shown in Table 4. Central to MWN approach is the level-wise hierarchical screening and prioritization of process changes options using the water management hierarchy (WMH) and four new option-screening heuristics which was sequentially applied to prioritize process changes. MySem had initially selected process changes based on the net annual savings (NAS) associated with each process change as shown with check marks in column 3 (MySem, 2005) of Table 4. The ultimate minimum water targets obtained after MWN analysis were given check marks in column 4 of Table 4. The steps for screening the options according to the WMH are described next.

4.4.1. Source elimination

The pinch point obtained from the base-case MWR targeting stage was at 23,360 ppm (see the water cascade table, i.e., Table 2). In order to maximize freshwater savings, the top priority was to consider eliminating water demands above the pinch point in the cascade table, i.e. any demand with concentration less than 23,360 ppm. Note that all the water demands in Table 1 met this criterion. All possible means to change process or the existing equipment to new equipment to eliminate water demands were considered. From Table 4, it was possible to eliminate D12 and D13 by changing wet bench 202 and 203 quartz tanks that initially needed continuous water for cooling, to teflon tanks. Not only did this option eliminated water requirement, it will also avoids tank cracking as a result of sudden temperature drop. Elimination of D12 and D13 resulted in new water targets at 8.3525 t/h freshwater and 0.0215 t/h IWT (see third row of Fig. 13). The pinch point was maintained at 23,360 ppm.

4.4.2. Source reduction

After eliminating D12 and D13, the next process change considered according to the WM hierarchy was to reduce demands above the pinch point in the cascade table; i.e. any demand with concentration lower than 23,360 ppm. Table 4 lists a few process changes options related to source reduction. Following heuristic 2, the demand at the lowest contaminant concentration (52 ppm) was reduced first. Follows are possible source reduction process changes (listed in Table 4) affecting demands D1, D10 and D11 (all located at 52 ppm) and sources S1–S13 and S19–S22 simultaneously due to the interactions between equipment in the water system of the DI plant (Fig. 9):

- wet bench flow rate reduction to the minimum during idle mode,
- recirculating hot water and switching heater on demand for heater WB201,

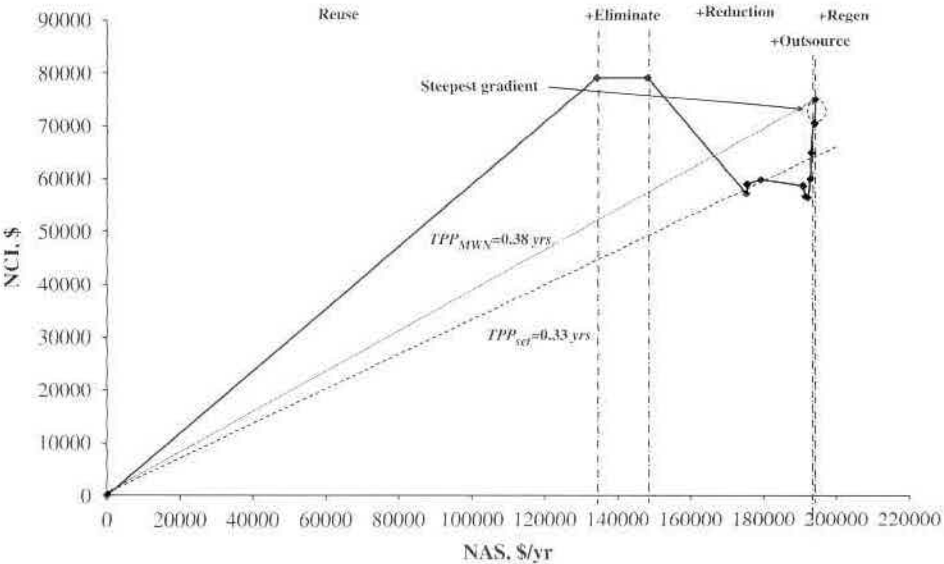


Fig. 14. IAS plot for MWN retrofit.

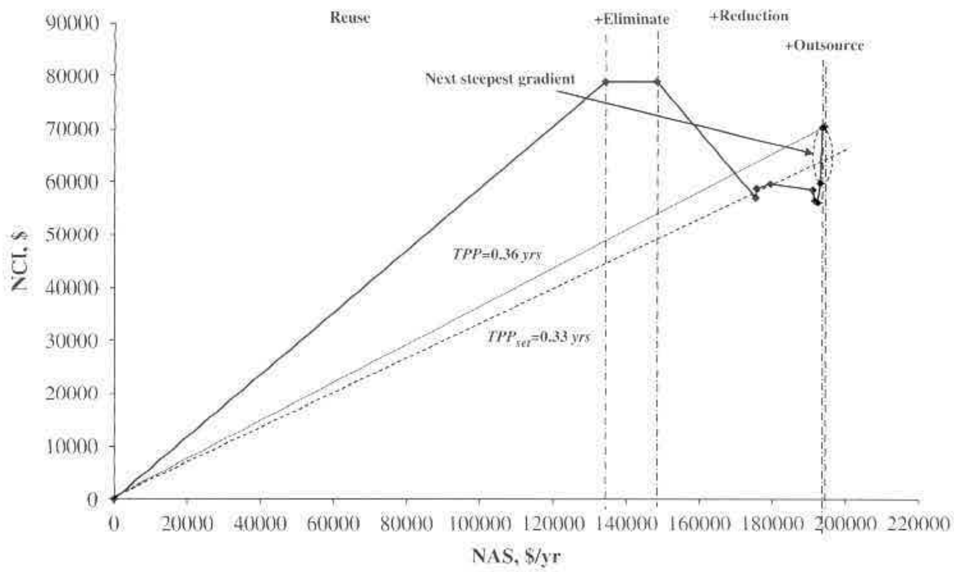


Fig. 15. IAS plot after eliminating regeneration curve.

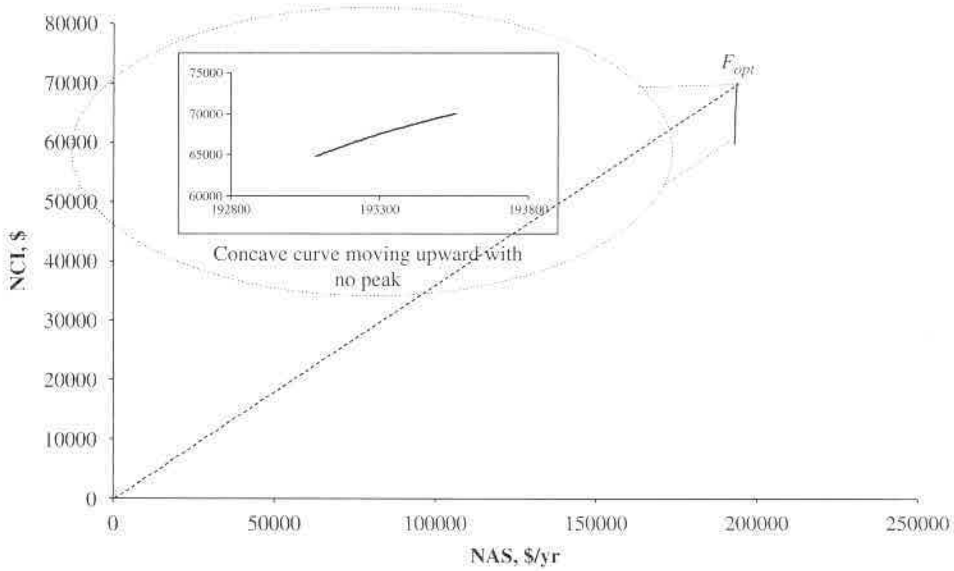


Fig. 16. F_{opt} for cooling tower concave curve moving upwards (without peak).

- reduction of Fab 1 return flow rate by changing to variable speed pump,
- decommissioning three EDI units instead running four units. Note that a sharp decrease in flow rate due to upstream process changes made it possible to reduce three EDI units (Option 2 from Table 4 was rejected due to the increase in FW and IWT targets. Option 1 i.e decommissioning three EDI unit reduced the FW and IWT targets to 6.3038 and 0.0378 t/h, respectively (Table 5) and hence implemented),
- increase rate of recovery for reverse osmosis system, and
- decrease multimedia filter backwash and rinsing time.

The freshwater (FW) and industrial wastewater treatment (IWT) flow rates after application of each process change are summarized in Fig. 13.

Using heuristic 1, the source reduction process changes for the DI water system shown in Fig. 9 were implemented from the core of the process (wet bench systems) to the most outer layer (multimedia filter). Excessive water usage at the core of the system was the main reason for water wastage at the outer layers, hence, improving the core of the system first will reduce wastage downstream. Implementation of the entire range of process changes, from wet bench to MMF (multimedia filter) listed in Table 3 gave

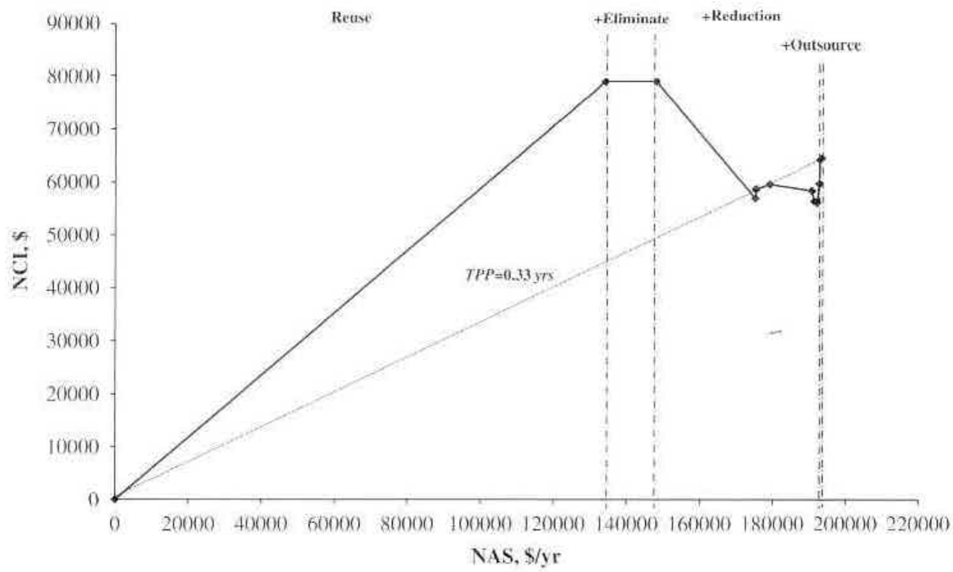


Fig. 17. Final IAS plot after SHARPS analysis.

revised freshwater and wastewater targets at 6.0857 and 0.0387 t/h and a new pinch concentration at 4608 ppm (refer to the ninth row of Fig. 13). Since there were no other demands at 52 ppm, following heuristic 2, the demands with the next lowest contaminant concentration (100 ppm) were considered next. For MySem, D2, D3 and D6 existed at the same concentration of 100 ppm. D3 which yielded the biggest flow rate reduction was chosen first followed by D2 and D6 according to heuristic 3, as described next.

The pollution abatement system demand (D3) existed at 100 ppm. Initially, pollution abatement system demanded 2.73 t/h of water (D3) and produced 2.73 t/h of IWT (S17). Table 6 shows five possible options to reduce the abatement system demand. Option 3 which was predicted to yield the highest savings was initially chosen by MySem prior to MWN approach (Table 4 column 3). However, as shown in Table 6, option 3 actually increased the freshwater target by 2.2% to 6.2183 t/h. This was because introduction of a recirculation system that produces no wastewater but relied on makeup water demand (option 3) had reduced the amount of wastewater that could potentially be reused for MySem as a whole, thereby leading to increased freshwater target. Option 4 in Table 6 gave the highest freshwater and IWT savings. Choosing option 4 led to new freshwater and IWT targets at 6.0831 and 0.0361 t/h, respectively, with pinch point maintained at 4608 ppm (refer to the tenth row of Fig. 13). It was also possible to reduce demand D2 which also existed at 100 ppm. D2 which was the cooling tower makeup, had the second highest flow rate reduction. Heat exchange between cooling tower circuit and liquid nitrogen circuit had potential to reduce D2 to 5.86 t/h, and ultimately the water targets to 5.9452 t/h freshwater and 0.0382 t/h wastewater (see eleventh row of Fig. 13). The pinch point was maintained at 4608 ppm.

Demands D6 (wash basin) and D7 (ablution) were reduced to 0.002 and 0.035 t/h, respectively, by changing the normal water taps to laminar taps. This also reduced sources S23 and S24. However, when targeted using Water-MATRIX©, the freshwater and wastewater targets increased slightly to 5.9455t/h and 0.0385 t/h. Hence this process change was rejected.

4.4.3. External water sources

The next process change according to the WM hierarchy was to add external water source at concentration lower than the new pinch point concentration of 4608 ppm. Based on MySem available roof area and rain distribution, it was possible to harvest 0.11 t/h (maximum design limit, $F_{max\ design}$) of rainwater at concentration of 16 ppm as a new water source. This option had potential to reduce the freshwater and IWT targets to 5.8349 and 0.0379 t/h, respectively (refer to the 12th row of Fig. 13). The pinch point was maintained at 4608 ppm.

4.4.4. Regeneration

Regeneration was the final process change considered according to the WM hierarchy. Freshwater savings could only be realized through regeneration above or across the pinch. Regenerating all 'WB201 in-operation' (S9) at 23,360 ppm and 0.0201 t/h (maximum utility flow rate, F_{MU} obtained using trial and error method in WCA) of 'WB101 in-operation' (S5) at 4608–52 ppm by carbon bed, EDI and ultraviolet (UV) treatment systems reduced the freshwater and IWT targets to 5.797 and 0 t/h, respectively (Table 7). Considering the IWT excluded from integration, the new IWT flow rate after regeneration was 0.7492 t/h. This corresponded to 85.5% freshwater and 97.8%

WMH levels	Specific process changes considered	New FW target, t/hr	New IWT+WW [based on limiting data] target, t/hr	New pinch point concentration, ppm	New total IWT (all IWT considered) target, t/hr
Initial	None	39.94	34.85	22360	34.4500
Reuse	Base case (MWR)	11.0400	0.0190	22360	5.7090
Elimination	Eliminate WB cooling 202 (D13) and 203 (D12)	8.3525	0.0215	22360	3.0215
Reduction	WB reduction in Fab 1 and Fab 2	6.7518	0.0258	4608	1.4216
	Heater WB201 reduction	6.7314	0.0264	4608	1.4454
	Fab 1 return reduction	6.6094	0.0354	4608	1.3109
	Option 1: EDI decommissioning	6.3038	0.0378	4608	1.0112
	Increase RO rate of recovery	6.2110	0.0380	4608	0.9132
	Reduce multimedia filter backwash and rinsing time	6.0857	0.0387	4608	0.7879
	Option 4: Reduce abatement pollution system (D3 = 0.57 t/hr and S17 = 0.57 t/hr).	6.0831	0.0361	4608	0.7853
	Cooling tower reduction (D2 = 5.966 t/hr)	6.0496	0.0366	4608	0.7858
Outsourcing	Add a source' S25=0.11 t/hr of C = 16 ppm by harvesting rainwater	5.9392	0.0362	4608	0.7854
CEMWN targets		**5.9392**	**0.0362**	**4608**	**0.7854**

Fig. 18. Final CEMWN targets after SHARPS analysis.

industrial wastewater reductions. The pinch point was maintained at 4608 ppm.

The *minimum water targets* were ultimately obtained after considering all options for process changes according to the WM hierarchy. Note that, targeting the maximum water recovery through reuse and regeneration only resulted in savings of up to 72.4% freshwater and 83.4% wastewater for MySem. Instead, following the holistic framework guided by the WM hierarchy enabled potential freshwater and wastewater reductions of up to 85.5% and 97.8%, respectively, towards achieving the minimum water network (MWN) design. Once the benchmark target was achieved the minimum water network was designed using various network design approaches such as source-sink mapping diagram (Polley and Polley, 2000), nearest neighbor algorithm (Prakash and Shenoy, 2005) or mathematical modeling [Takama et al., 1980; Alva-Argáez et al., 1998; Huang et al., 1999; Benkó et al., 2000; Bagajewicz and Savelski, 2001; Xu et al., 2003; Koppol et al., 2003; Ullmer et al., 2005; Gunaratman et al., 2005).

4.5. Step 4: apply SHARPS strategy

The desired payback period (TPP_{set}) was set at 4 month (0.33 years) by MySem management. Fig. 14 shows the IAS plot after MWN analysis for MySem retrofit. (Refer to Appendices A and B for cost formula used to generate the IAS plot. Since this is a retrofit case, Eqs. (1), (4) and (5) mentioned previously were used.) The total payback period to attain the MWN targets before SHARPS screening was

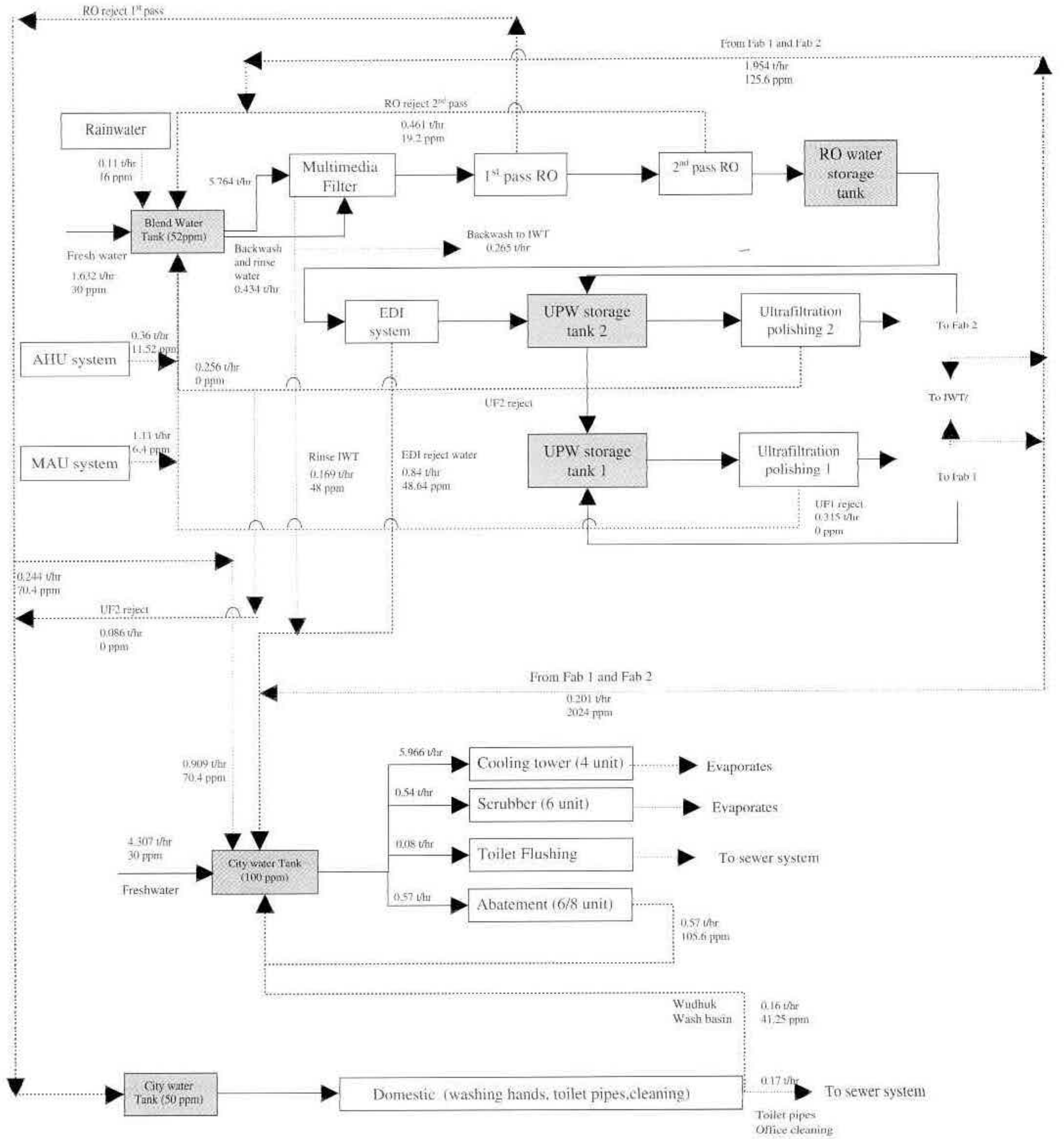


Fig. 19. MySem retrofit DI water balance and non-process water balance after CEMWN analysis, achieving 85.5% freshwater and 97.7% IWT reductions within 4 months payback.

0.38 years. Since, the initial total payback period was more than the TPP_{set}, i.e. 0.33 years, SHARPS strategies were applied to fulfill the TPP_{set} specified.

Fig. 14 shows that *regeneration* process change gives the steepest gradient. Focusing on *Strategy* 1, there was no other option for *regeneration* process change. Hence, *Strategy* 1 which called for equipment substitution could not be implemented. Focusing on *Strategy* 2, when no *regeneration* was applied, the total payback period reduces to 0.36 yr as illustrated in Fig. 15, which still do not achieve TPP_{set}. Thus, the next steepest gradient was observed.

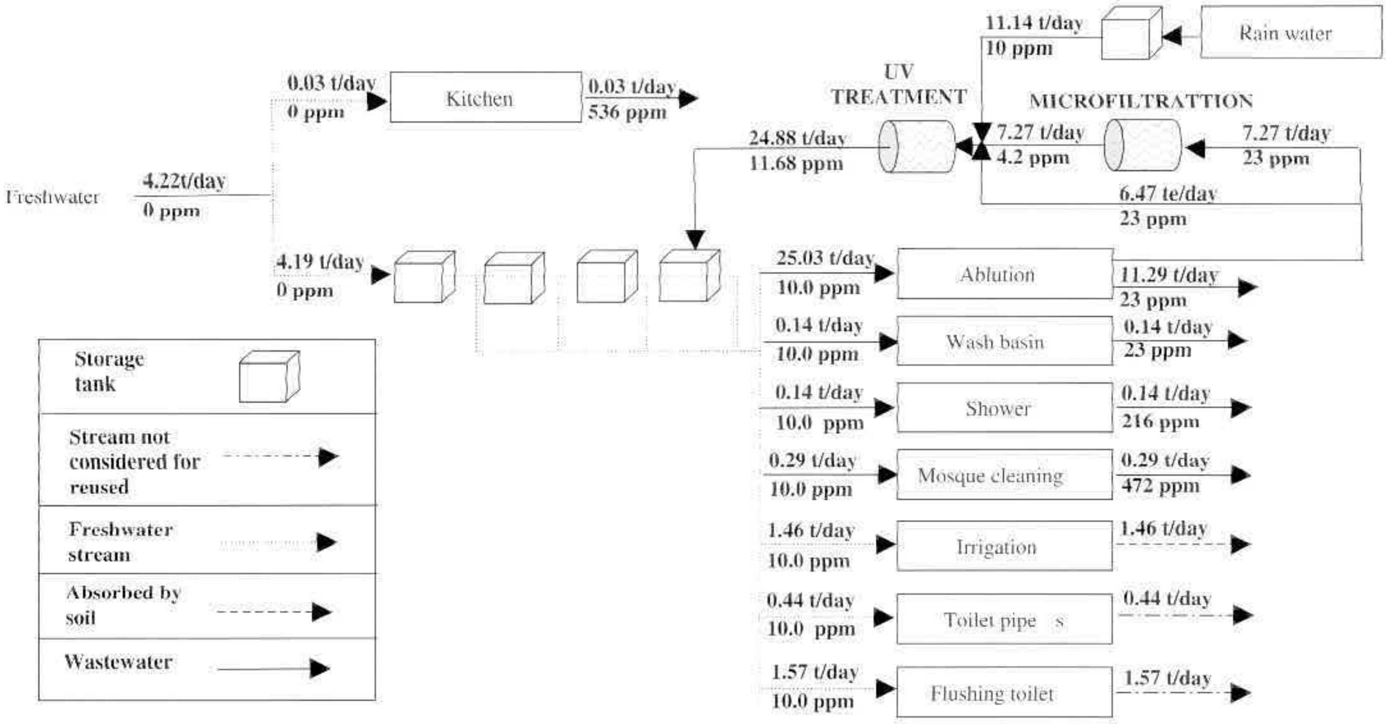


Fig. 20. Final water distribution network for Sultan Ismail Mosque with regeneration and rainwater harvesting by Manan et al. (2006).

Table 8
Limiting water data for SIM

	Demand	F (t/day)	C (ppm)		Source	F (t/day)	C (ppm)
D1	Kitchen	0.03	0	S1	Ablution	25.03	23
D2	Ablution	25.03	10	S2	Wash basin	0.14	23
D3	Wash basin	0.14	10	S3	Showering	0.14	216
D4	Showering	0.14	10	S4	Mosque cleaning	0.29	472
D5	Mosque cleaning	0.29	10	S5	Kitchen	0.03	536
D6	Irrigation	1.46	10	Total water sources	25.63	t/h	
D7	Toilet pipes	0.44	10				
D8	Flushing toilet	1.57	10				
Total water demands	29.10	t/h					

The next steepest gradient was noted to be *cooling tower* process change. Again, *Strategy* 1 cannot be applied since only one option exists for that process change. It can be noted that cooling tower line is a concave curve without peak. F_{opt} is noted to be at the end of the curve as shown in Fig. 16. Hence, based on the linearization rule explained earlier, any reduction of the cooling tower curve will be beneficial. Using *Strategy* 2, reducing the *cooling tower* process change curve by only applying partially nitrogen cooling effect yielded the final IAS plot shown in Fig. 17 that achieved the specified payback of 0.33 yr. D2 was reduced only to 5.966 t/h instead of 5.86 t/h initially. This scheme had reduced 5.94 t/h of freshwater and 0.79 t/h IWT total flow rate.

Hence, the application of SHARPS screening has successively achieve the TPP_{set} of 4 month with 85.1% and 97.7% of freshwater and wastewater reduction respectively prior to design. This is the final cost effective minimum water network (CEMWN) target for MySem. The effect of applying each scheme on freshwater and IWT are illustrated in Fig. 18. Using the pre-design cost estimate method, the system needs approximately a net total investment of $64,500 and will give a net annual savings of $193,550 yearly.

4.6. Step 5: network design

Fig. 19 shows the cost effective minimum water network for MySem after retrofit.

5. Sultan Ismail Mosque case study

This case study compares the results of applying CEMWN framework and water pinch analysis (WPA) on

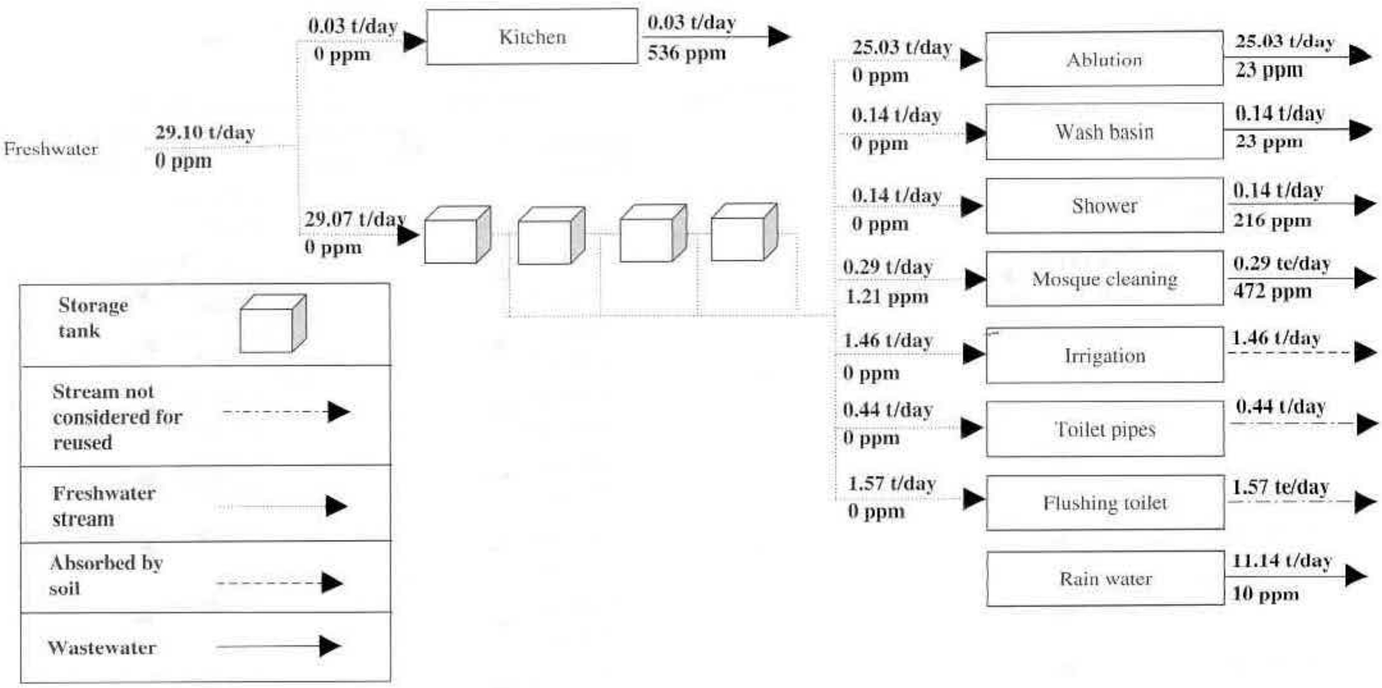


Fig. 21. Water distribution network for Sultan Ismail Mosque.

Table 9
Various process changes options for SIM

WMH	Strategy	Option selected based on MWN procedure
Elimination	Toilet option 1: Change 12 *l* flushing toilet to composting toilet	✔
Reduction	Change normal ablution tap to laminar flow tap	✔
	Toilet option 2: Change 12 *l* flushing toilet to dual flush toilet	X
Reuse	Total reuse	✔
Outsourcing	RW harvesting	✔
Regeneration	Treat ablution as required	✔

(✔) For selected option, (X) for eliminated option by SIM.

Sultan Ismail Mosque (SIM) in Malaysia which is an urban building. Manan et al. (2006) uses WPA on the SIM case study which include maximum water recovery, regeneration and rainwater harvesting to achieve 85.5% freshwater and 67.7% wastewater reductions. The final maximum water recover (MWR) network is shown in Fig. 20. The limiting data for the study is shown in Table 8. Biological oxygen demand (BOD) is used as the key water quality factor. Fig. 21 shows the base-case water distribution network for SIM before integration.

CEMWN framework was next applied to SIM to cost-effectively maximize water savings. Possible process changes options are listed in Table 9. The minimum water network targeted 99.9% freshwater and 63.8% wastewater savings after implementing WMH-guided process changes (see Fig. 22). Note that, in some cases, though the freshwater target decreased, the wastewater target increased. For example, when toilet demand was eliminated, some of the wastewater initially allocated for reuse in D8 had to be discharged.

Fig. 23 shows the initial IAS plot generated after MWN analysis for both grassroots and retrofit cases. The total payback period for grassroots design was 8.0 yr and retrofit case 10.2 yr. TPP_{set} were set at 3 and 5 years for grassroots and retrofit cases respectively for SIM. It is important to note that in the case of urban sector, a payback period of up to 10 years for retrofit cases are typically considered to be on the lower side due to the much cheaper urban freshwater tariff as compared to industrial tariff and the lack of economy of scale. Burkhard et al. (2000); Naisby (1997); Sayers (1998) and Mustow et al. (1997) estimate payback periods for domestic graywater and rainwater reuse systems in the range between 34 to 890 yr in the UK. Thus, CEMWN implementation is encouraged for grassroots design more than for retrofit cases for urban sector.

Elimination of demand D8 by changing from a 12-l-flush toilet to a composting toilet (option 1) led to the steepest gradient on the IAS composite plot. SHARPS *Strategy* 1 was then applied to remove the steepest gradient. Changing to a dual-flush toilet instead (option 2) yielded lower TPPs of 4.43 years for grassroots and 6.69 years for retrofit cases but the dual-flush toilet option then became the steepest gradient and the TPP_{set} was still exceeded (Fig. 24).

The base-case toilet option which gave a TPP of 4.01 years for grassroots and 5.19 years for retrofit (Fig. 25) was finally selected. Since TPP_{set} was not achieved by trimming the steepest gradient, hence, intensifying the regeneration

WMH levels	Specific process changes considered	New FW target, t/day	New WW target, t/day	New pinch point concentration, ppm
Initial	None	29.1	25.6	
Reuse	Base case	16.5	13.0	23
+ Eliminate	Eliminate a demand at C = 10ppm by changing 12 *l* flushing toilet to composting toilet	15.6	13.7	23
+ Reduction	Reduce by half the flowrate of demand at C =10 ppm by changing the normal ablution water tap to low flowrate water tap.	8.5	6.6	23
+ Reuse/ outsourcing	Add a source* of C = 10 ppm by harvesting rainwater.	2.2	11.5	23
+ Regeneration	Regenerate to the maximum flowrate* for a source from C=23 ppm to C=4.2 ppm using a microfiltration, activated carbon and UVsystem.	0.03	9.27	4.2
= **Minimum water network (MWN) target**		**0.03**	**9.27**	**4.2**

Fig. 22. The effects of WMH-guided process changes on MWR targets and pinch location.

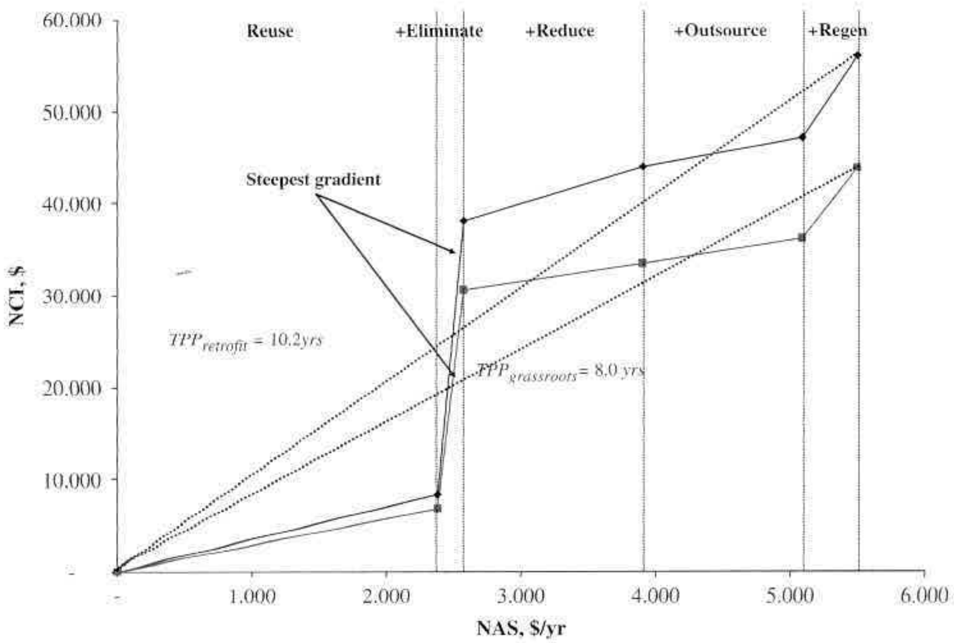


Fig. 23. IAS plot for SIM.

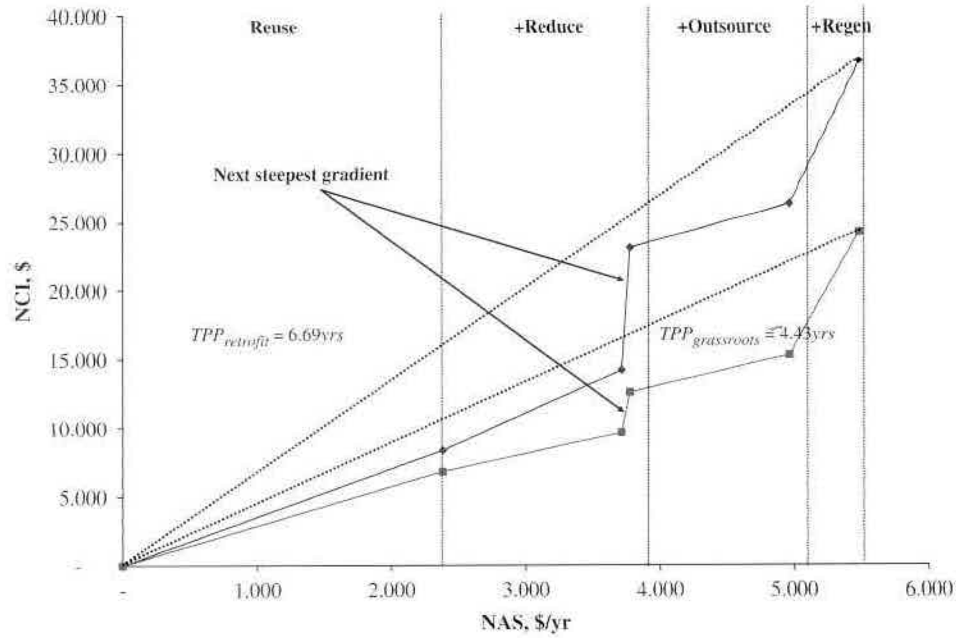


Fig. 24. IAS plot after changing from composting toilet to dual-flush toilet for SIM.

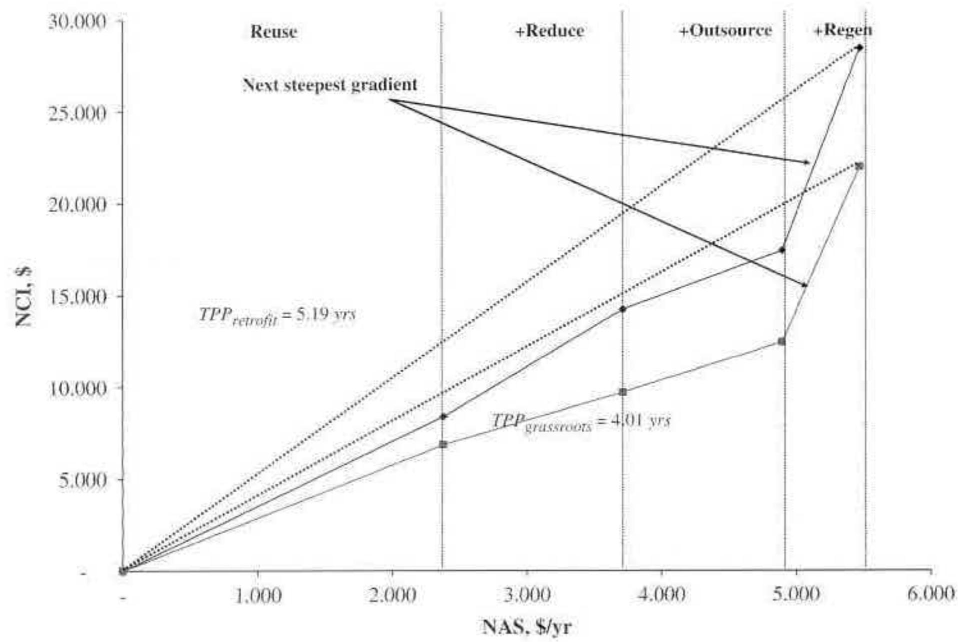


Fig. 25. IAS plot after eliminating toilet flush process change for SIM.

option which formed the next steepest gradient was considered (SHARPS *Strategy* 2). Regenerating only 0.39 t/day of ablution for grassroots and 2.89 t/day for retrofit achieved the TPP_{set}. This gave reductions of 90.5% freshwater and 59.3% wastewater for grassroots and 97.5% freshwater and 67.2% wastewater for retrofit. The final IAS plots that achieved the TPP_{set} are shown in Fig. 26. The cost correlations used and the results are given in Appendix C. Figs. 27 and 28 show the WMH-guided process changes after CEMWN implementation for grassroots and retrofit cases. The final network that achieved the CEMWN targets are shown in Figs. 29(a) and (b).

6. Results comparison

Table 10 compares the results of using techniques such as MWR, MWN and MWN after SHARPS screening

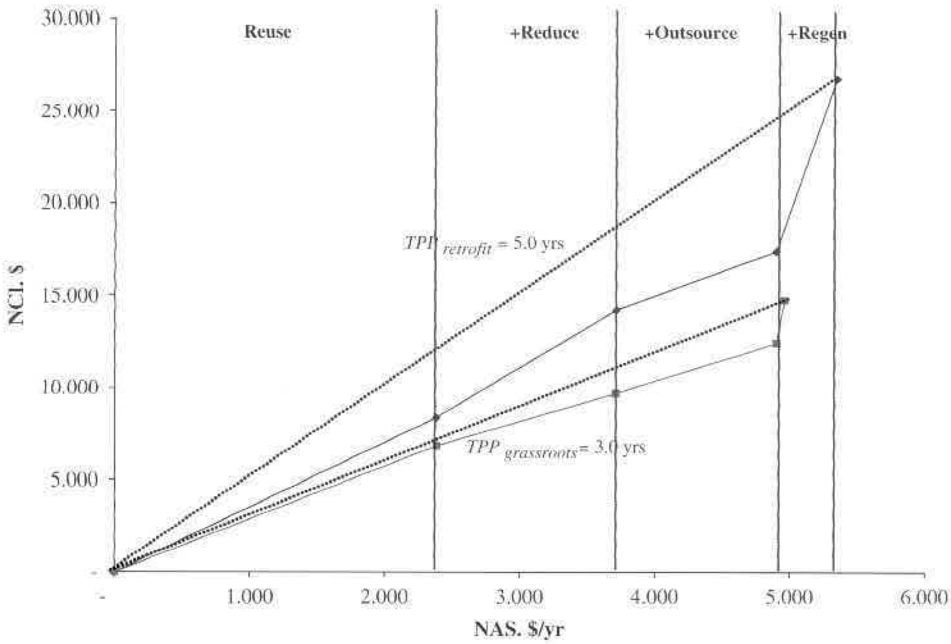


Fig. 26. IAS plot after eliminating toilet flush process change and reducing regeneration for SIM.

WMH levels	Specific process changes considered	New FW target, t/day	New WW target, t/day	New pinch point concentration, ppm
Initial	None	29.1	25.6	
Reuse	Base case	16.5	13.0	23
+ Reduction	Reduce by half the flowrate of demand at C = 10 ppm by changing the normal ablution water tap to low flowrate water tap.	9.39	5.92	23
+ Reuse/ outsourcing	Add a source* of C = 10 ppm by harvesting rainwater.	3.09	10.8	23
+ Regeneration	Regenerate 0.393 t/day of ablution wastewater from C=23 ppm to C=4.2 ppm using a microfiltration, activated carbon and UV system.	2.77	10.4	23
= **CEMWN target**		**2.77**	**10.4**	**23**

Fig. 27. WMH-guided process changes after CEMWN analysis (grassroots).

WMH levels	Specific process changes considered	New FW target, t/day	New WW target, t/day	New pinch point concentration, ppm
Initial	None	29.1	25.6	
Reuse	Base case	16.5	13.0	23
+ Reduction	Reduce by half the flowrate of demand at C = 10 ppm by changing the normal ablution water tap to low flowrate water tap.	9.39	5.92	23
+ Reuse/ outsourcing	Add a source* of C = 10 ppm by harvesting rainwater.	3.09	10.76	23
+ Regeneration	Regenerate 2.891 t/day of ablution wastewater from C=23 ppm to C=4.2 ppm using a microfiltration, activated carbon and UV system.	0.73	8.40	23
= **CEMWN target**		**0.73**	**8.40**	**23**

Fig. 28. WMH-guided process changes after CEMWN analysis (retrofit).

(termed as CEMWN target) for MySem retrofit. It can be seen that the CEMWN target gives a higher net annual savings (NAS) compared to MWR but lower than MWN in order to accommodate the payback period specified by the plant owner. Table 11 compares the results of applying all the approaches to SIM case study. As can be seen, without the CEMWN holistic framework, only 85.5% and 67.7% of freshwater and wastewater reduction was identified using MWR, rainwater harvesting and regeneration process changes options by Manan et al. (2006). The proposed process changes did not satisfy the TPP_{set}. On the other hand, CEMWN approach enabled reductions of 94.2% freshwater and 63.5% wastewater for grassroots and 97.5% freshwater and 67.2% wastewater reductions for retrofit cases within a targeted payback period. TPP_{set}.

The SHARPS technique provides clear quantitative insights to screen various water management options. By applying the SHARPS technique in accordance with the water management hierarchy, it is possible to decide the schemes to partially apply or completely eliminate in order to satisfy a desired payback period, thereby allowing a designer to estimate the maximum potential annual savings ahead of design. SHARPS is a novel cost-screening technique that enables a designer to customize a cost-effective water network design that attains the minimum water targets as per the requirement of a plant or a building owner.

7. Using CEMWN target as reference benchmark

CEMWN targets can be used for own performance and international water reduction benchmarking guide. For example, the CEMWN target for MySem, which also corresponds to the best achievable benchmark targets, are a target of freshwater flow rate of 5.94 t/h and total IWT flow rate of 0.79 t/h. This represented 85.1% freshwater and 97.7% IWT reduction. Hence, these were the best performance benchmark targets (Fig. 30) that MySem needed to achieve. The application of total reuse only using water pinch analysis (WPA) method yielded a lower water savings potentials of 72.4% freshwater and 83.4% wastewater reduction with a 0.59 year payback period. November 2005 water bills had shown that all the conventional water reductions strategies applied by MySem had only managed to reduce freshwater usage from 42.6 to 40.24 t/h representing a savings of $880 per month. An estimated total savings of $193,550 per year was predicted with the implementation of CEMWN method. A preliminary cost

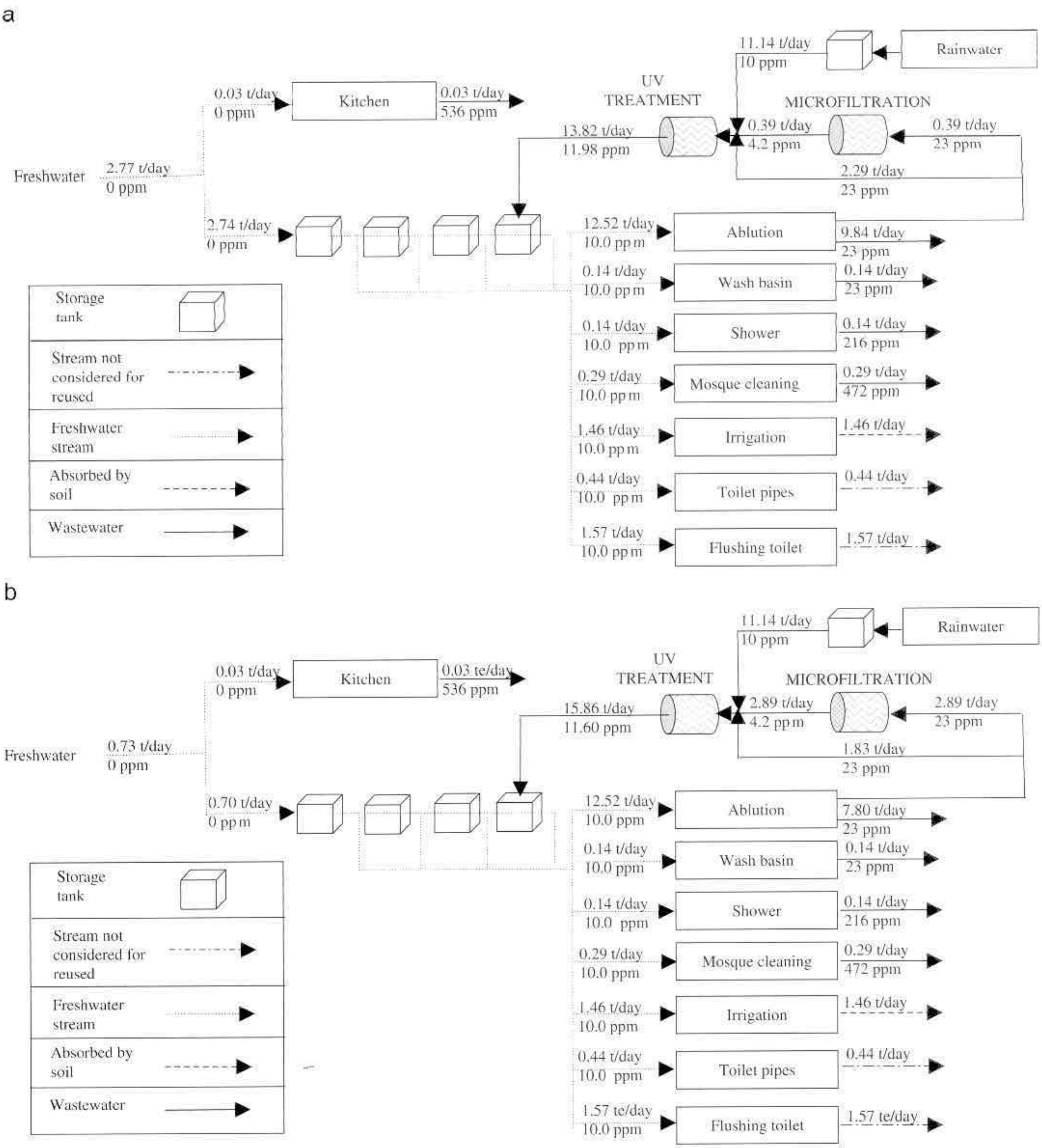


Fig. 29. (a) CEMWN design for SIM (grassroots). (b) CEMWN design for SIM (retrofit).

estimate indicated that this best performance required an investment of approximately $64,507 with a pay-back period of 0.33 years. Note that the schemes proposed and listed in Fig. 18 could also be gradually implemented as part of the company's longer-term utility savings program in line with its quality management practices.

Once the best performance benchmark was established through CEMWN method, the predicted maximum savings of MySem was then compared with the

Table 10
Comparison of MWR, MWN and CEMWN results for MYSEM case study

Method	FW_{target} (t/h)	WW_{target} (t/h)	IWT total (kg/h)	FW savings (%)	IWT savings (%)	NAS (\$/yr)	NCI (\$)	TPP (yr)
Initial	39.94	34.85	34.45					
MWR	11.04	0.019	5.71	72.4%	83.4%	134,250	78,952	0.59
MWN	5.80	0	0.79	85.2%	97.9%	194,242	74,708	0.38
CEMWN	5.94	0.036	0.79	85.1%	97.7%	193,550	64,507	0.33

Table 11
Comparison of MWR, MWR with rainwater harvesting and regeneration by Manan et al. (2006) and CEMWN for SIM case study

System	Water minimization method used	FW_{target} (t/day)	WW_{target} (t/day)	FW savings (%)	WW savings (%)	NAS (\$/yr)	NCI (\$)	TPP (yr)
Initial	None	29.1	25.63					
Grassroots	MWR	16.46	12.99	43.4%	49.3%	2380	6837	2.87
	MWR + RW + Regeneration	4.22	11.89	85.5%	67.7%	4686	24,465	5.22
	MWN	0.03	9.27	99.9%	63.8%	5500	43,915	7.98
	CEMWN	2.77	10.44	90.5%	59.3%	4959	14,739	2.97
Retrofit	MWR	16.46	12.99	43.4%	49.3%	2380	8371	3.52
	MWR + RW + Regeneration	4.22	11.89	85.5%	67.7%	4686	28,864	6.16
	MWN	0.03	9.27	99.9%	63.8%	5500	35,915	10.20
	CEMWN	0.73	8.40	97.5%	67.2%	5343	26,757	5.01

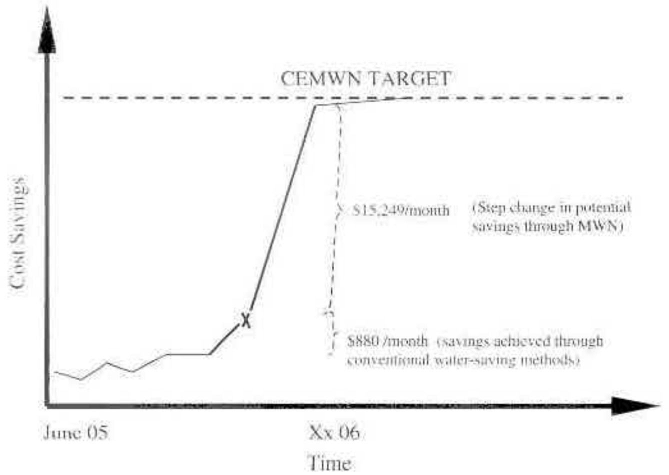


Fig. 30. Savings achieved by MySem in comparison to savings predicted through CEMWN technique.

international benchmark. The International Technology Roadmap for Semiconductor (ITRS, 2001) had aimed to reduce high purity water (HPW) consumption from the current rate of 6–8 m^3 in 2005 to 4–6 m^3 per wafer by 2007 (Wu et al., 2004). After CEWN analysis, MySem had potential to use 4.06 m^3 of DI water per wafer for Fab 1 and 13.73 m^3 of DI water per wafer from Fab 2, down from its previous consumption of 6.3 and 72.4 m^3 of DI water per wafer, respectively. Fab 1 had potential to meet the ITRS 2001 target. Fab 2 however was far from this ITRS target due to its wafer production rate of well below the design capacity.

Table 12
$OC_{base\ case}$ formula for MYSEM case study

Process	Type of OC	Cost formula	Unit
Freshwater cost, C_{FW}	Freshwater	$0.518F_{FW\ initial}$	\$/h
Industrial wastewater cost, C_{IWT}	Wastewater	$0.042F_{IWT\ initial}$	\$/h
EDI cost	Other	$0.017F_{EDI\ initial}$	\$/h
Heater WB101 electrical cost	Other	$0.016F_{Heater\ WB101\ initial}$	\$/h
Overall DI water treatment	Chemical	$0.016F_{MMF\ initial}$	\$/h
Internal pumping cost	Electrical	$0.018F_{internal\ initial}$	\$/h

8. Conclusion

The cost effective minimum water network (CEMWN) technique can help a company realize its' best achievable water savings target and assess its true potential for continuous improvement to fulfill its quality management requirement using the WMH hierarchy with four new heuristics and the new SHARPS technique. Application of CEMWN technique on a semiconductor plant showed that savings of up to 85.1% freshwater and 97.7% industrial wastewater were achievable with an estimated payback period of 4 months. Application of the same method on a mosque building gives a savings of up to 90.5% freshwater and 59.3% wastewater achievable within a payback period of 3 years for grassroots case and 97.5% freshwater and 67.2% wastewater achievable within a payback period of 5 years for retrofit case.

Table 13
OC_{new} formula for MYSEM case study

Process	Type of OC	Cost formula	Unit
JBA cost, C_{FW}	Freshwater	$0.518F_{FW\ new}$	$/h
Industrial wastewater cost, C_{IWT}	Wastewater	$0.042F_{IWT\ new}$	$/h
EDI chemical and pumping cost	Chemical and electrical	$0.017F_{EDI\ new}$	$/h
Heater WB101 electrical cost	Electrical	$0.018F_{HeaterWB101\ new}$	$/h
Overall DI water treatment	Chemical	$0.016F_{MMF\ new}$	$/h
Reuse, outsource and treatment pumping operating cost, C_{EOC}	Electrical	$0.018F_{reuse/outsource/reg}$	$/h
Internal pumping cost	Electrical	$0.018F_{internal\ new}$	$/h
Treatment OC	Chemical	$0.018F_{reg}$	$/h

Table 14
CC_{new} formula for individual equipment for MYSEM case study

Process	New equipment	No of unit	Cost formula ($)	Unit	Cost/system ($)
Wudhuk	Tap 1.9 lpm with installation	7	10.8	$/unit	75.6
Wash basin	Tap 1.9 lpm with installation	15	10.8	$/unit	162
Abatement	Option 1: Run each abatement at 0.5 gpm during idle. Need control system.	6	7020	$/system	7020
	Option 2: Decommissioning 3 abatement unit and new duct with installation	3	8640	$/system	8640
	Option 3: Recirculation of water. Need treatment, piping and control system with installation	6	34,290	$/system	34,290
	Option 4: On demand abatement system. Need control system with installation.	6	7020	$/system	7020
	Option 5: Abatement flow reduction based on pH analysis.	6	Nil	$/system	Nil
WB 202 cooling	Teflon tank with installation	1	1350	$/unit	1350
WB 203 cooling	Teflon tank with installation	1	1350	$/unit	1350
WB in Fab 1 and Fab 2	Adjust WB flow valve during idle mode to minimum. No investment needed.	—	Nil	$/system	Nil
Heater WB101	Heater unused water recirculation and on demand heating. Need piping and control system.	1	1836	$/unit	1836
Fab 1 return	Variable speed pump with installations	1	4050	$/unit	4050
RO system	Increase RO system rate of recovery (RR) to maximum (RR = 80% 1st pass, RR = 90% 2nd pass). No investment needed.	1	Nil	$/system	Nil
EDI	Option 1: Decommissioning extra EDI unit and 3 variable speed pumps	3	6750	$/system	6750
	Option 2: Run EDI intermittently. Need control system.	—	1890	$/system	1890
Cooling tower	Using N2 to cool cooling water return. Heat exchanger and piping with installations	1	$-1\times10^6F_{CT\ makeup}{}^2+2\times10^7F_{CT\ makeup}-5\times10^7$	$/system	
MMF backwash and rinse	Reduce time of backwash and rinse to minimum. No need investment.	2	Nil	$/system	Nil
Total reuse	Reuse diversion system and pumps with installations	—	$(52634*(F_{reuse}/F_{demand\ initial})\ 0.6)*150\%$	$/system	
Rainwater harvesting (10 ppm)	RW diversion system and pumps	—	$(52634*(F_{RW}/\ F_{demand\ initial})\ 0.6)*170\%$	$/system	
Treatment (Treat WB WW to 52 ppm)	Treat all WB WW by using carbon bed, EDI and UV. Need treatment system, installations, control and piping	—	$[261{,}900*(F_{reg}/45.5)^{0.6}*120\%]+[52{,}634*(F_{reg}/F_{demand\ initial})\ 0.6)*150\%]$	$/system	

Appendix A. Obtaining pre-design capital cost estimation

An equipment capital cost is typically a function of the equipment capacity, and, in the context of SHARPS, is related to a flow rate increment or reduction associated with a WM option. The method to estimate the four types of cost used in Eq. (2) for capital cost calculations will be elaborated next.

A.1. Estimation of equipment purchased cost and installation cost (C_{PE} and C_{PEI})

The capital cost of an equipment of a given size can be predicted using the *six-tenth factor rule* (Peters et al., 2003). According to this rule, if the cost of an equipment b at a given capacity is known, the cost of a similar equipment a at X times the capacity of b is $X^{0.6}$ times the cost of equipment b as given by Eq. (9) (Peters et al., 2003). The 0.6-rule of thumb is only used when the actual cost exponent is unknown. The typical exponents for equipment cost as a function of capacity can be obtained from most literatures on plant economics. For example, the exponential value of a flat-head, carbon steel tank is 0.57 (Peters et al., 2003)

$$\text{Cost of equipment } a = (\text{cost of equipment } b)X^{0.6}. \tag{9}$$

Eq. (9) is a capital cost correlation for a biological treatment unit (Gunaratman et al., 2005). The capital cost of a 30 t/h wastewater treatment unit (F_{TU}), is \$136,256

$$CC_{T1}(\$) = 12,600F_{TU}(\mathrm{t/h})^{0.7}. \tag{10}$$

Note that the capacity factor rule for equipment costing is applicable only for similar equipment type of up to 10 times the base-equipment capacity. The cost must also be updated as necessary using the Marshall and Swift equipment cost index or the *Chemical Engineering* cost index. The sum of individual equipment cost gives the total capital cost for the equipment, C_{PE}. The equipment installation cost (C_{PEI}) includes labor cost, foundations, supports, cost of construction, and other factors directly related to the erection of the purchased equipment. The purchased equipment costs may vary between 20% and 90% of the total installed cost depending on the equipment complexity and the type of plant the equipment is installed in (Peters et al., 2003).

A.2. Estimation of piping and plumbing cost (C_{piping})

For piping cost, if the *base-case* plumbing and sanitation piping cost is available, Eq. (9) can also be used to estimate the reuse or outsource or regeneration piping cost using Eq. (11) as follows:

Table 15
MWN targets for MySem

WMH level	Strategy	FW target (t/day)	IWT target (t/day)	NAS (\$/yr)	NCI (retrofit) (\$)	TPP (retrofit) (years)
Initial	None	39.94	34.85			
Reuse	Reuse	11.04	0.0190	134,250	78,952	0.59
+ Elimination	WB cooling	8.353	0.0215	148,273	78,954	0.53
+ Reduction	WB reduction	6.752	0.0258	175,285	56,976	0.33
	Heater reduction	6.731	0.0264	175,466	58,682	0.33
	Fab 1 return reduction	6.609	0.0354	179,293	59,558	0.33
	EDI return option 2	6.304	0.0378	190,843	58,383	0.31
	RO recovery	6.211	0.038	191,492	56,419	0.29
	MMF	6.086	0.0387	192,198	56,102	0.29
	Abatement option 4	6.083	0.0361	192,886	59,694	0.31
	Reduce cooling tower to 5.86 t/day	5.945	0.0382	193,555	69,955	0.36
+ Outsourcing	Add 0.11 t/h RW with 16 ppm TDS	5.835	0.0379	194,056	70,231	0.36
+ Regeneration	Regenerate to 52 ppm	5.797	0	194,242	74,708	0.38

Table 16
CEMWN targets for MySem

WMH level	Strategy	FW target (t/day)	IWT target (t/day)	NAS (\$/yr)	NCI (retrofit) (\$)	TPP (retrofit) (yr)
Initial	None	39.94	34.85			
Reuse	Reuse	11.04	0.0190	134,250	78,952	0.59
+ Elimination	WB cooling	8.353	0.0215	148,273	78,954	0.53
+ Reduction	WB reduction	6.752	0.0258	175,285	56,976	0.33
	Heater reduction	6.731	0.0264	175,466	58,682	0.33
	Fab 1 return reduction	6.609	0.0354	179,293	59,558	0.33
	EDI return option 2	6.304	0.0378	190,843	58,383	0.31
	RO recovery	6.211	0.038	191,492	56,419	0.29
	MMF	6.086	0.0387	192,198	56,102	0.29
	Abatement option 4	6.083	0.0361	192,886	59,694	0.31
	Reduce cooling tower to 5.966 t/day	6.050	0.0366	193,049	64,232	0.33
+ Outsourcing	Add 0.11 t/h RW with 16 ppm TDS	5.939	0.0362	193,550	64,507	0.33

Cost of piping and plumbing

$$= (\text{Cost of base case plumbing and sanitation}) \times \left(\frac{F_{\text{reuse/outsource/regen}}}{F_{\text{demand initial}}}\right)^{0.6}. \quad (11)$$

Table 17
$OC_{\text{base case}}$ formula for SIM case study

Process	Type of OC	Cost formula	Unit
Freshwater cost, C_{FW}	Freshwater	$0.56F_{\text{FW initial}}$	$/t

Table 18
OC_{new} formula for SIM case study

Process	Type of OC	Cost formula	Unit
Freshwater cost, C_{FW}	Freshwater	$0.56F_{\text{FW new}}$	$/t
UV lamp	Treatment	$0.03F_{\text{reuse/outsource/reg}}$	$/t
Pumping	Electrical	$0.014F_{\text{reuse/outsource/reg}}$	$/t

Table 19
$CC_{\text{base case}}$ formula for individual equipment for SIM case study

Process	Base case equipment	No of unit	Cost formula, $	Unit	Cost/system ($)
Toilet	12 l toilet flush with installations	30	200	$/unit	6000
Ablution tap	Tap 13.5 lpm with installation	126	20	$/unit	2520
Plumbing and sanitation	Piping		8000	$/system	8000

Table 20
CC_{new} formula for individual equipment for SIM case study

Process	New equipment	No of unit	Cost formula ($)	Unit	Cost/system ($)
Toilet	Option 1: composting toilet with installations	30	1000	$/unit	30,000
	Option 2: dual flush toilet with installations	30	300	$/unit	9000
Ablution tap	Laminar tap with installations	126	25	$/unit	3150
Total reuse	Reuse diversion system and pumps with installations (Grassroots)	—	$[(499(F_{\text{reuse}}/22.71)^{0.6})+(30F_{\text{reuse}})+8000(F_{\text{reuse}}/F_{\text{demand initial}})^{0.6}]130\%$	$/system	
	Reuse diversion system and pumps with installations (Retrofit)	—	$[(499(F_{\text{reuse}}/22.71)^{0.6})+(30F_{\text{reuse}})+8000(F_{\text{reuse}}/F_{\text{demand initial}})^{0.6}]150\%$	$/system	
Rainwater harvesting (10 ppm)	RW diversion system and pumps (Grassroots)	—	$[(499(F_{\text{RW}}/22.71)^{0.6})+(30F_{\text{RW}})+8000(F_{\text{RW}}/F_{\text{demand initial}})^{0.6}]150\%$	$/system	
	RW diversion system and pumps (Retrofit)	—	$[(499(F_{\text{RW}}/22.71)^{0.6})+(30F_{\text{RW}})+8000(F_{\text{RW}}/F_{\text{demand initial}})^{0.6}]170\%$	$/system	
Treatment (Treat ablution WW to 4.2 ppm)	Treat all ablution WW by using microfiltration, activated carbon and UV. Need treatment system, installations, control and piping (Grassroots)	—	$[(10{,}000(F_{\text{reg}}/7.27)^{0.6})130\%]+[(499(F_{\text{reg}}/22.71)^{0.6})+(30F_{\text{reg}})+8000(F_{\text{reg}}/F_{\text{demand initial}})^{0.6}130\%]$	$/system	
	Treat all ablution WW by using microfiltration, activated carbon and UV. Need treatment system, installations, control and piping (Retrofit)	—	$[10{,}000(F_{\text{reg}}/7.27)^{0.6})150\%]+[(499(F_{\text{reg}}/22.71)^{0.6})+(30F_{\text{reg}})+8000(F_{\text{reg}}/F_{\text{demand initial}})^{0.6}150\%]$	$/system	

A.3. Instrumentation and control, C_{IC}

To enable water reuse, pumps and control systems must also be installed. This should include instrumentation cost, installation labor cost and the operating cost for auxiliary equipment such as pumps and motors. For preliminary design, the costs of instrumentation and control may range between 8% and 50% of the total delivered equipment cost depending on the extent of control required (Peters et al., 2003).

Appendix B. MySem cost formula

For MySem case study, the formula for $OC_{\text{base-case}}$, OC_{new} and $CC_{\text{new system}}$ are listed in Tables 12–14. Tables 15 and 16 show the results of calculating the NCI and NAS using MWN and CEMWN methods, respectively, for each WMH option.

Table 21
MWN analysis for SIM

WMH levels	Strategy	FW target (t/day)	WW target (t/day)	NAS ($/yr)	NCI (retrofit) ($)	NCI (grassroots) ($)	TPP (retrofit) (years)	TPP (Grassroots) (years)
Initial	None	29.10	25.63					
Reuse	Reuse	16.46	12.99	2380	8371	6837	3.52	2.87
+ Eliminate	Option 1: eliminate D8 by changing to composting toilet	15.57	13.67	2573	3.8085	30.602	14.80	11.89
+ Reduction	Ablution tap change to laminar tap	8.50	6.60	3905	43.979	33,480	11.26	8.57
+ Outsourcing	Harvest 11.14 t/day of rainwater	2.20	11.45	5091	47.160	36.215	9.26	7.11
+ Regeneration	Regenerate 2.67 t/day (F_{MU}) ablution to 4.2 ppm	0.03	9.27	5500	56.082	43,915	10.20	7.98

Table 22
CEMWN analysis for SIM (Strategy 1)-change toilet flush to dual flush instead

WMH levels	Strategy	FW target (t/day)	WW target (t/day)	NAS ($/yr)	NCI (retrofit) ($)	NCI (grassroots) ($)	TPP (retrofit) (yr)	TPP (Grassroots) (yr)
Initial	None	29.10	25.63					
Reuse	Reuse	16.46	12.99	2380	8371	6837	3.52	2.87
+ Reduction	Ablution tap change to laminar tap	9.39	5.92	3712	14,223	9680	3.83	2.61
	Option 2: reduce D8 to 1.05 t/day by changing to dual flush toilet	9.10	6.15	3776	23,143	12,614	6.13	3.34
+ Outsourcing	Harvest 11.14 t/day of rainwater	2.80	10.99	4962	26,308	15,336	5.30	3.09
+ Regeneration	Regenerate 3.39 t/day (F_{MU}) ablution to 4.2 ppm	0.03	8.22	5483	36,679	24,284	6.69	4.43

Table 23
CEMWN analysis for SIM (Strategy 2): retrofit—use base case toilet flush and reduce regeneration

WMH levels	Strategy	FW target (t/day)	WW target (t/day)	NAS ($/yr)	NCI (retrofit) ($)	TPP (retrofit) (yr)
Initial	None	29.1	25.63			
Reuse	Reuse	16.46	12.99	2380	8371	3.52
+ Reduction	Ablution tap change to laminar tap	9.39	5.92	3712	14 223	3.83
+ Outsourcing	Harvest 11.14 t/day of rainwater	3.09	10.76	4898	17 381	3.55
+ Regeneration	Regenerate 2.891 t/day ablution to 4.2 ppm	0.73	8.40	5343	26 757	5.01

Table 24
CEMWN analysis for SIM (Strategy 2): grassroots—use base case toilet flush and reduce regeneration

WMH levels	Strategy	FW target (t/day)	WW target (t/day)	NAS ($/yr)	NCI (grassroots) ($)	TPP (Grassroots) (yr)
Initial	None	29.1	25.63			
Reuse	Reuse	16.46	12.99	2380	6837	2.87
+ Reduction	Ablution tap change to laminar tap	9.39	5.92	3712	9680	2.61
+ Outsourcing	Harvest 11.14 t/day of rainwater	3.09	10.76	4898	12 396	2.53
+ Regeneration	Regenerate 1.773 t/day ablution to 4.2 ppm	2.77	10.44	4959	14 739	2.97

Appendix C. SIM cost formula

For SIM case study, the formula for $OC_{base-case}$, OC_{new} and $CC_{new\ system}$ are listed in Tables 17–20. Tables 21–24 show the results of calculating the NCI and NAS using MWN and CEMWN methods, respectively, for each WMH option.

Note added in proof

Part of the concepts discussed in this paper have been published in Wan Alwi and Manan (2006) as an R&D note. This paper presents a full version of the completed and tested version of the technique on two case studies involving an industry and an urban system. A complete and more elaborate step-wise description of the CEMWN procedure and SHARPS methodology, which include derivations of cost equation have been presented in this paper.

References

Alva-Argáez, A., Kokossis, A.C., Smith, R., 1998. Wastewater minimization of industrial systems using an integrated approach. Computers and Chemical Engineering 22, 741–744.

Bagajewicz, M., Savelski, M., 2001. On the use of linear models for the design of water utilization systems in process plants with a single contaminant. Transactions of the Institute of Chemical Engineers Part A 79, 600–610.

BBC News (22 March 2002). UN warns of looming water crisis.

Benkö, N., Rév, E., Fonyó, Z., 2000. The use of nonlinear programming to optimal water allocations. Chemical Engineering Communications 178, 67–101.

Burkhard, R., Deletic, A., Craig, A., 2000. Techniques for water and wastewater management: a review of techniques and their integration in planning. Urban Water 2, 197–221.

Castro, P., Matos, H., Fernandes, M.C., Nunes, C.P., 1999. Improvements for mass-exchange networks design. Chemical Engineering Science 54, 1649–1665.

Dhole, V.R., Ramchandani, N., Tainsh, R.A., Wasilewski, M., 1996. Make your process water pay for itself. Chemical Engineering 103, 100–103.

El-Halwagi, M.M., 1997. Pollution Prevention Through Process Integration: Systematic Design Tools. Academic Press, San Diego.

El-Halwagi, M.M., Manousiouthakis, V., 1989. Synthesis of mass exchange network. AIChe Journal 35 (8), 1233–1244.

Feng, X., Chu, K.H., 2004. Cost optimisation of industrial wastewater reuse systems. Transactions of the IChemE, Part B, Process Safety and Environmental Protection 82 (B3), 249–255.

Foo, C.Y., Manan, Z.A., Yunus, R.M., Aziz, R.A., 2004. Synthesis of mass exchange network for batch processes—Part 1: utility targeting. Chemical Engineering Sciences 59, 1009–1026.

Foo, C.Y., Manan, Z.A., Yunus, R.M., Aziz, R.A., 2005. Synthesis of mass exchange network for batch processes—part II: minimum unit target and batch network design. Chemical Engineering Sciences 60 (5), 1349–1362.

Geldermann, J., Treitz, M., Schollenberger, H., Rentz, O., 2005. Plant layout planning based on multi objective pinch analysis. In: The Preeminent International Conference, IFORS 2005, Honolulu, HI, July 11–15, 2005.

Gunaratnam, M., Alva-Argáez, A., A. Kokossis, A., Kim, J.-K., Smith, R., 2005. Automated design of total water systems. Industrial and Engineering Chemistry Research 44, 588–599.

Hallale, N., 2002. A new graphical targeting method for water minimization. Advances in Environmental Research 6 (3), 377–390.

Hallale, N., Fraser, D.M., 1998. Capital cost targets for mass exchange networks. A special case: water minimization. Chemical Engineering Science 53 (2), 293–313.

Huang, C.H., Chang, C.T., Ling, H.C., Chang, C.C., 1999. A mathematical programming model for water usage and treatment network design. Industrial and Engineering Chemistry Research 38, 2666–2679.

Jödicke, G., Fischer, U., Hungerbühler, K., 2001. Wastewater reuse: a new approach to screen for designs with minimal total costs. Computers and Chemical Engineering 25, 203–215.

Koppol, A.P.R., Bagajewicz, M.J., Dericks, B.J., Savelski, M.J., 2003. On zero water discharge solutions in the process industry. Advances in Environmental Research 8, 151–171.

Kuo, W.C.J., Smith, R., 1998. Design of water-using systems involving regeneration. Transactions of the Institute of Chemical Engineers Part B 76, 94–114.

Liu, Y.A., Lucas, B., Mann, J., 2004. Up-to-date tools for water-system optimization. Chemical Engineering Magazine, 30–41.

Manan, Z.A., Wan Alwi, S.R., 2006. Stretching the limits on urban and industrial water savings. Jurutera (01), 24–27.

Manan, Z.A., Tan, Y.L., Foo, D.C.Y., 2004. Targeting the minimum water flowrate using water cascade analysis technique. AIChE Journal 50 (12), 3169–3183.

Manan, Z.A., Wan Alwi, S.R., Ujang, Z., 2006. Water pinch analysis for an urban system: A case study on the Sultan Ismail Mosque at the Universiti Teknologi Malaysia (UTM). Desalination 194, 52–68.

Mariano-Romero, C.E., Alcocer-Yamanaka, V., Morales, E.F., 2005. 'Multiobjective Water Pinch Analysis of the Cuernava City Water Distribution Network. Lecture Notes in Computer Science, Evolutionary Multi-Criterion Optimization, vol. 3410. Springer, Berlin, Heidelberg.

MIMOS Semiconductor (M) Sdn. Bhd. (MySem), 2005. UTM-MIMOS Water Minimization Project (UMWMP). Unpublished report.

Mustow, S., Grey, R., Smerdon, T., Pinney, C., Waggett, R., 1997. Water Conservation: Implications of Using Recycled Greywater and Stored Rainwater in the UK. BSRIA, Bracknell.

Naisby, C., 1997. Greywater recycling and rainwater harvesting. A viable means of domestic water conservation? Unpublished Masters Thesis, Department of Geography and Civil Engineering, University of Leeds.

Olesen, S.G., Polley, G.T., 1997. A simple methodology for the design of water networks handling single contaminants. Transactions of IChemE Part A 75, 420–426.

Peters, M.X., Timmerhaus, K.D., West, R.E., 2003. Plant Design and Economics for Chemical Engineers, fifth ed. McGraw-Hill, New York.

Polley, G.T., Polley, H.L., 2000. Design better water networks. Chemical Engineering Progress 96 (2), 47–52.

Prakash, R., Shenoy, U.V., 2005. Targeting and design of water networks for fixed flowrate and fixed contaminant load operations. Chemical Engineering Science 60 (1), 255–268.

Sayers, D., 1998. A study of domestic greywater recycling. Interim Report, National Water Demand Management Centre, Environment Agency, Worthing, West Sussex.

Smith, R., 1995. Chemical Process Design. McGraw-Hill Inc., New York.

Sorin, M., Bédard, S., 1999. The global pinch point in water reuse networks. Transactions of IChemE Part B 77, 305–308.

Takama, N., Kuriyama, T., Shiroko, K., Umeda, T., 1980. Optimal water allocation in a petroleum refinery. Computers and Chemical Engineering 4, 251–258.

Tan, Y. L., 2005. Development of New Systematic Techniques for Retrofit of Water Network. M.Sc. Thesis, Universiti Teknologi Malaysia, Johor, Malaysia.

The Institution of Chemical Engineers, 2000. Guide to Industrial Water Conservation. Part 1. Water Re-use In Single Contaminant Systems. ESDU 0020.

Ullmer, C., Kunde, N., Lassahn, A., Gruhn, G., Schulz, K., 2005. WADO: water design optimization-methodology and software for the synthesis of process water systems. Journal of Cleaner Production 13 (5), 485–494.

Wan Alwi, S.R., Manan, Z.A., Ujang, Z., 2004. Systematic technique for water minimization in urban water system using water pinch analysis. In: Proceedings of ASIA WATER 2004, 30–31 March 2004.

Wan Alwi, S.R., 2006. A new holistic framework for minimum water network design for urban and industrial sectors. Ph.D. Thesis, Universiti Teknologi Malaysia, Johor, Malaysia (in preparation).

Wan Alwi, S.R., Manan, Z.A., 2006. SHARPS—a new cost-screening technique to attain cost-effective minimum water utilisation network. R&D note. AIChe Journal 52 (11), 3981–3988.

Wang, Y.P., Smith, R., 1994. Wastewater minimization. Chemical Engineering Science 49, 981–1006.

Wang, Y.P., Smith, R., 1995. Wastewater minimization with flowrate constraints. Transactions of IChemE Part A 73, 889–904.

Water-MATRIX©, 2006. A software for water minimization, copyright owned by Universiti Teknologi Malaysia.

Wu, M., Sun, D., Tay, J.H., 2004. Process-to-process recycling of high-purity water for semiconductor wafer backgrinding wastes. Resources, Conservation and Recycling 41, 119–132.

Xu, D., Hu, Y., Hua, B., Wang, X., 2003. Minimization of the flow rate of freshwater and corresponding regenerated water in water-using system with regeneration reuse. Chinese Journal of Chemical Engineering 11 (3), 257–263.

ELSEVIER

Journal of Environmental Management 88 (2008) 253–274

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Flowrate targeting for threshold problems and plant-wide integration for water network synthesis

Dominic Chwan Yee Foo*

School of Chemical and Environmental Engineering, University of Nottingham Malaysia, Broga Road, 43500 Semenyih, Selangor, Malaysia

Received 24 July 2006; received in revised form 22 January 2007; accepted 17 February 2007
Available online 12 April 2007

Abstract

Water reuse/recycle has gained much attention in recent years for environmental sustainability reasons, as well as the rising costs of fresh water and effluent treatment. Process integration techniques for the synthesis of water network have been widely accepted as a promising tool to reduce fresh water and wastewater flowrates via in-plant water reuse/recycle. To date, the focus in this area has been on water network synthesis problems, with little attention dedicated to the rare but realistic cases of so-called threshold problems. In this work, targeting for threshold problems in a water network is addressed using the recently developed numerical tool of water cascade analysis (WCA). Targeting for plant-wide integration is then addressed. By sending water sources across different geographical zones in plant-wide integration, the overall fresh water and wastewater flowrates are reduced simultaneously.

Keywords: Process integration; Process synthesis; Water minimisation; Threshold problems; Targeting; Plant-wide integration

1. Introduction

Environmental sustainability goals, the rising cost of raw material and waste treatment, and increasingly stringent emission regulations are among factors that drive the process industries to shift from conventional end-of-pipe treatment to more sustainable pollution prevention or waste minimisation practises. One of the active areas for waste minimisation activities has been that of in-plant material reuse/recycle. The benefits of implementing material reuse/recycle are two fold. Apart from the reduction of raw material purchases needed for a process, less waste is generated; thus, direct and indirect costs of handling raw material pre-treatment and wastes are also reduced. Recent efforts have seen the advancement of systematic design of material reuse/recycle network for resource conservation, with the most active area being that of water network synthesis.

Wang and Smith (1994) initiated research in the field of water integration by developing a two-stage graphical pinch approach based on the more general mass exchange network synthesis problems (El-Halwagi and Manousiouthakis, 1989). In the first stage, limiting water profile is derived from the limiting water data to locate the minimum fresh water and wastewater flowrates; this step is followed by the detailed network design stage. Flowrate constraints and the integration of regeneration units were considered in their later work (Wang and Smith, 1995a; Kuo and Smith, 1998a). The basic concept underlying these seminal work is that, all water-using processes are modelled as mass transfer operations (more commonly known as *fixed load* problems). However, later works (e.g. Dhole et al., 1996; Hallale, 2002; Manan et al., 2004) showed that not all water-using processes can be handled using the mass transfer-based model. Some water-using processes, such as boilers, cooling towers and reactors cannot be modelled as mass transfer operations. Water network synthesis that involves these units is commonly known as the *fixed flowrate* problems.

The first generation of water source and sink composite curves to handle the fixed flowrate problems was proposed by Dhole et al. (1996); and was later improved by Polley and Polley (2000). A numerical tool that is equivalent to

*Tel.: +60 3 8924 8130; fax: +60 3 8924 8017.
E-mail address: dominic.foo@nottingham.edu.my.

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.02.007

Dhole's composite curves was also developed by Sorin and Bédard (1999). However, Hallale (2002) later pointed out that both of these graphical and numerical tools fail to locate the global pinch point for multiple-pinch problems.

Two promising graphical approaches to locating the right minimum water flowrate targets for the fixed flowrate problem are the water surplus diagram (Hallale, 2002) and the material recovery pinch diagram (El-Halwagi et al., 2003; Prakash and Shenoy, 2005a). The water surplus diagram (Hallale, 2002) is analogous to the grand composite curves in heat exchanger network synthesis (Linnhoff et al., 1982); however, it requires a tedious iterative procedure to implement. To avoid this problem, a non-iterative material recovery pinch diagram was independently developed by El-Halwagi et al. (2003) and Prakash and Shenoy (2005a).

On the other hand, equivalent numerical tools for targeting minimum water flowrates have also been developed. The water cascade analysis (WCA) technique presented by Manan and co-workers (Manan et al., 2004; Foo et al., 2006b) is the tabular equivalent to the water surplus diagram (Hallale, 2002), with the iterative calculation steps having been eliminated. Almutlaq and El-Halwagi (2007) as well as Almutlaq et al. (2005) presented another cascade analysis tool that is based on the composite curves of El-Halwagi et al. (2003). These four tools are by far the most promising techniques for locating the minimum water targets in a water network. While graphical targeting tools provide the conceptual insights for network synthesis, numerical tools are preferred when rapid and accurate answers are needed. More recently, concept of flowrate targeting is extended to locate targets for process changes (Bandyopadhyay, 2006; Wan Alwi and Manan, 2006), regeneration flowrate (Tan and Manan, 2006; Agrawal and Shenoy, 2006; Ng et al., 2006), as well as waste treatment targets after targets for reuse/recycle is located (Kuo and Smith, 1998b; Bandyopadhyay et al., 2006).

Apart from the targeting stage, numerous techniques have also been proposed to design water network that achieve the flowrate targets. This includes the water grid diagram (Wang and Smith, 1994), load table (Olesen and Polley, 1996, 1997), water main method (Kuo and Smith, 1998b; Castro et al., 1999; Feng and Seider, 2001; Wang et al., 2003; Cao et al., 2004; Feng et al., 2005; Zheng et al., 2006b), heuristic procedures (Liu et al., 2004; Liu et al., 2005b), water source diagram (Gomes et al., 2006; Queiroz and Pessoa, 2006) for fixed load problems; as well as source sink mapping diagram (Dunn and Bush, 2001; Dunn and Wenzel, 2001; Parthasarathy and Krishnagopalan, 2001; El-Halwagi, 1997), source sink approach (El-Halwagi, 1997; Vaidyanathan et al., 1998); nearest neighbour algorithm (Prakash and Shenoy, 2005a) as well as the load problem table (Aly et al., 2005) for fixed flowrate problem. Once the flowrate targets are established, the water network is designed to achieve the minimum targets using any of the above-mentioned design tools. Subsequently, a preliminary synthesised network can be evolved to yield simplified network (Prakash and Shenoy, 2005b; Ng and Foo, 2006).

On the other hand, the mathematical optimisation approach for water network synthesis has also received much attention from the research community. Early work in this area was reported by Takama and co-workers (Takama et al., 1980a, b, 1981). Alva-Argáez and co-authors developed the integrated approach combining the insights from water pinch and mathematical programming in handing the fixed load problems (Alva-Argáez et al., 1998, 1999). The combination use of pinch and linear programming (LP) techniques was later presented by Jacob et al. (2002) for the fixed flowrate problems. Optimisation approach based on non-linear programming (NLP) was later presented for fixed load problems (Rossiter and Nath, 1995; Yang et al., 2000; Abebe et al., 2003) and for the fixed flowrate problems (Dunn et al., 2001). Huang et al. (1999) and Benko et al. (1999, 2000) individually developed the mathematical approach to include water treatment in the total water network synthesis. Bagajewicz and co-workers utilised LP and algorithmic procedures for the design of water network, for both single (Bagajewicz and Savelski, 2001; Savelski and Bagajewicz, 2000a, b, 2001; Gómez et al., 2001) and multiple impurities (Bagajewicz et al., 2000; Savelski and Bagajewicz, 2003). More recent works on this area are dominated by the advanced mathematical optimisation approaches, such as fuzzy programming (Tan, 2002; Tan and Cruz, 2004), genetic algorithm (Tsai and Chang, 2001; Li et al., 2003; Prakotpol and Srinophakun, 2004; Shafiei et al., 2004; Lavric et al., 2005), random search optimisation (Poplewski et al., 2002; Jeżowski et al., 2003; Poplewski and Jeżowski, 2005) and particle swarm optimisation (Hul et al., 2007; Luo et al., 2007).

Mathematical optimisation approach serves as a supplementary tool to graphical pinch approach in addressing more complex systems, such as systems with large number of water-using processes (Savelski and Bagajewicz, 2001), multiple impurities (Takama et al., 1980a; Doyle and Smith, 1997; Alva-Argáez et al., 1999; Bagajewicz et al., 2000; Dunn et al., 2001; Teles et al., 2006), mass load uncertainty (Tan et al., 2007; Koppol and Bagajewicz, 2003), capital cost estimation (Alva-Argáez et al., 1998; Jödicke et al., 2001; Feng and Chu, 2004), integration with interception network (Gabriel and El-Halwagi, 2005), as well as water treatment system (Huang et al., 1999; Tsai and Chang, 2001; Karuppiah and Grossmann, 2006) or evaluation of zero discharge possibility (Koppol et al., 2003). A good review of various existing techniques in addressing water network synthesis problem is presented by Bagajewicz (2000).

Another area where water reuse/recycle has seen significant advances in recent years is the water network synthesis for batch processes. As in continuous processes, both graphical pinch approaches (Wang and Smith, 1995b; Foo et al., 2005b; Majozi et al., 2006) and mathematical optimisation approaches (Almató et al., 1997, 1999a, b;

Puigjaner et al., 2000; Kim and Smith, 2004; Chang and Li, 2005; Li and Chang, 2006; Majozi, 2005a, b, 2006; Cheng and Chang, 2006) have been proposed for the synthesis of an optimal batch water network.

The successful applications of water network synthesis in various process industries have been documented for both the water pinch and mathematical optimisation approaches. This includes petrochemical complexes (Liu et al., 2005a; Mann and Liu, 1999), chemical manufacturing (Hall, 1997; El-Halwagi, 1997, 2006; Brouckaert et al., 1999; Tainsh and Rudman, 1999; Gianadda et al., 2001; Mann, 2003; Forstmeier et al., 2005; Ku-Pineda and Tan, 2006; Zheng et al., 2006a; Feng et al., 2006), oil refineries (El-Halwagi et al., 1992; Wang and Smith, 1994; Tainsh and Rudman, 1999), pulp and paper mills (Tripathi, 1996; Tainsh and Rudman, 1999; Bédard et al., 1999; Yang et al., 2000; Parthasarathy and Krishnagopalan, 2001; Andersen, et al., 2002; Jacob et al., 2002; Tan and Manan, 2003; Manan and Tan, 2004; Lovelady et al., 2007; Manan et al., 2007; Foo et al., 2006b; Delgado et al., 2006; Chiang et al., 2006), textile/fabric mills (Jödicke and Hungerbühler, 1999; Jödicke et al., 2001; Wenzel et al., 2002; Ujang et al., 2002; Hamad et al., 2003), thermal power stations (Brouckaert et al., 1999), polymer plants (Tainsh and Rudman, 1999), electroplating processes (Yang et al., 2000; Zhou et al., 2001), fuel production plants (Noureldin and El-Halwagi, 1999; El-Halwagi, 1997), food and beverage plants (Brouckaert and Buckley, 2000; Thevendiraraj et al., 2003), palm oil mills (Chungsiriporn et al., 2006), agrochemical plants (Majozi et al., 2006) and municipal water systems (Manan et al., 2006; Mariano-Romero et al., 2006). Despite the wide spread application of water network synthesis techniques, there remain several areas that yet to be explored by researchers and industrial practitioners. This includes the special case of water network synthesis, i.e. *threshold problems*, as well as the exploration of cross-plant water integration, known as the *plant-wide integration*. These are the subjects of this work.

In this paper, flowrate targeting for water network synthesis with threshold problems is first addressed. The recently developed WCA targeting technique (Manan et al., 2004, Foo et al., 2006b) is used to locate fresh water and wastewater flowrate targets in these networks. Unlike many water networks that normally require fresh water intake and discharges wastewater, water network with threshold problems may either require fresh water without waste generation, or on the other extreme generate wastewater without any fresh water intake.

In the second section of this paper, plant-wide integration is analysed. After water reuse/recycle has been exhausted within a single water network, cross-plant water integration possibility is analysed. By sending wastewater source from one network to the other, overall plant fresh water intake and wastewater generation is reduced simultaneously. WCA targeting technique (Manan et al., 2004, Foo et al., 2006b) is utilised to locate the overall flowrate targets for the plant-wide integration. In the next section, steps to carry out a cascade analysis are briefly reviewed, before the threshold problems and plant-wide integration cases are analysed.

2. Water cascade analysis technique for flowrate targeting

The water cascade table (WCT) in Table 1 summarises how WCA is carried out for flowrate targeting in a water network. The first step in conducting a WCA is to locate the various water sinks and sources at their respective concentration levels. As shown in the first two columns of Table 1, the concentration levels (C_k) are arranged in an ascending order ($k = 1, 2, \ldots, n$), and the flowrates of water sink (F_j) and source (F_i) are summed at their respective concentration level k in columns 3 and 4. Column 5 represents the net flowrate, ($\Sigma_i F_i - \Sigma_j F_j$) between water sources and sinks at each concentration level k; with positive indicating surplus, negative indicating deficit. Next, the net water flowrate surplus/deficit is cascaded

Table 1
WCT for water flowrate targeting

Column								
1	2	3	4	5	6	7	8	9
k	C_k	$\sum_j F_j$	$\sum_i F_i$	$\sum_i F_i - \sum_j F_j$	$F_{C,k}$	Δm_k	Cum.Δm_k	$F_{FW,k}$
					F_{FW}			
k	C_k	$(\sum_j F_j)_1$	$(\sum_i F_i)_1$	$(\sum_i F_i - \sum_j F_j)_1$				
					$F_{C,k}$	Δm_k		
$k+1$	C_{k+1}	$(\sum_j F_j)_{k+1}$	$(\sum_i F_i)_{k+1}$	$(\sum_i F_i - \sum_j F_j)_{k+1}$			Cum. Δm_{k+1}	$F_{FW,k+1}$
					$F_{C,k+1}$	Δm_{k+1}		
$n-2$	C_{n-2}	$(\sum_j F_j)_{n-2}$	$(\sum_i F_i)_{n-2}$	$(\sum_i F_i - \sum_j F_j)_{n-2}$				
					$F_{C,n-2}$	Δm_{n-2}		
$n-1$	C_{n-1}	$(\sum_j F_j)_{n-1}$	$(\sum_i F_i)_{n-1}$	$(\sum_i F_i - \sum_j F_j)_{n-1}$			Cum. Δm_{n-1}	$F_{FW,n-1}$
					$F_{C,n-1} = F_{WW}$	Δm_{n-1}		
n	C_n						Cum. Δm_n	$F_{FW,n}$

down the concentration levels to yield the cumulative surplus/deficit flowrate ($F_{C,\ k}$) in column 6 with an assumed zero fresh water flowrate ($F_{FW} = 0$). This assumed flowrate is to facilitate the search for the minimum water flowrate and will later be replaced once the rigorous fresh water target is located.

Two important parameters to fulfil in water network targeting are the flowrate and impurity load constraints (Foo et al., 2006b). Flowrate constraints are fulfilled once the above-described flowrate cascading is carried out. The next step involves setting up the cumulative impurity load cascade (Cum.Δm) to fulfil the load constraint. Impurity load in column 7 (Δm_k) is obtained via the product of cumulative flowrate ($F_{C,\ k}$) and the concentration difference across two subsequent concentration levels ($C_{k+1}-C_k$). Cascading the impurity load down the concentration levels of column 8 yields the cumulative load (Cum.Δm_k), which is essentially the numerical equivalent of graphical targeting tool of water surplus diagram (Hallale, 2002). A feasible water network is characterised by the presence of only positive values of Cum.Δm in column 8. A negative Cum.Δm_k means the impurity load is transferred from higher to lower concentration level, which is infeasible. In such a case, an interval fresh water flowrate ($F_{FW,\ k}$, column 9) is calculated by dividing Cum.Δm_k by the concentration difference between level k (C_k) and the fresh water concentration (C_{FW}) i.e.,

$$F_{FW,k} = \frac{\text{Cum.}\Delta m_k}{C_k - C_{FW}} \tag{1}$$

The absolute value of the largest negative $F_{FW,\ k}$ will then replace the earlier assumed zero fresh water flowrate in the flowrate targeting (column 6) to obtain a new set of feasible flowrate cascade and hence a feasible load cascade. This new fresh water flowrate represents the minimum fresh water flowrate (F_{FW}) of the network; while the final row in column 6 represents the wastewater flowrate (F_{WW}) generated from the network. The network pinch concentration is the impurity concentration with zero Cum.Δm_k, while the water source that exists at the pinch is called the pinch-causing source. To achieve the minimum water targets, a portion of the pinch-causing source flowrate is allocated to the region above the pinch while the rest to the region below the pinch (these flowrates are found in the $F_{C,\ k}$ column just above and below the pinch concentration). These are termed as the water allocation targets (Manan et al., 2004; Foo et al., 2006a, b). Note that once the flowrate and impurity load cascading are carried out using the minimum fresh water flowrate, final column of Table 1 is omitted. Flowrate targeting for various water network problems using WCA have been presented by many earlier works (e.g. Manan et al., 2004; Foo et al., 2006a, b; Foo, 2006, 2007).

Ng et al. (2006) recently extended the WCA targeting technique to identify individual wastewater streams that emit from a water reuse/recycle network. In the waste targeting procedure, water sinks and sources are firstly segregated to regions above and below the pinch according to their concentration levels. WCA targeting procedure is then carrying out for region below the pinch to identify the individual wastewater streams. Eq. (1) is used to locate the minimum pinch-causing source flowrate ($F_{FW,\ k}$ target is replaced by interval pinch source flowrate $F_{P,\ k}$ target) that will fulfil the flowrate and load constraints in this region (Ng et al., 2006).

In the following sections, WCA targeting tool is firstly used to locate the flowrate targets in the special cases of threshold problems. This will be discussed using several examples.

3. Example 1—Zero discharge network with fresh feed

Table 2 shows the limiting data for Example 1, with three water sinks and sources. Prior to water reuse/recycle, the network requires a fresh water flowrates of 170 ton/h and generates 110 ton/h of wastewater (sum of the individual sink and source flowrates in Table 2). Flowrate targeting can either be carried out using graphical tools such as the composite curves (e.g. El-Halwagi et al., 2003; Prakash and Shenoy, 2005a), water surplus diagram (Hallale, 2002), or numerical techniques such as WCA (Manan et al., 2004; Foo et al., 2006b). In this work, WCA targeting technique is used.

An infeasible WCT is shown in Table 3 in which negative impurity loads are found in impurity concentration level between 50–250 ppm ($k = 3$–5). Following the WCA procedure outlined in previous section, interval fresh water flowrate ($F_{FW,\ k}$) are identified in the final column of Table 3. The largest negative value of $F_{FW,\ k}$, i.e. 59.97 ton/h is used as the minimum fresh water flowrate (F_{FW}) and next to produce a new WCT in Table 4. Even though a pinch is found at the Cum.Δm column (in the last row of Table 4), however, it is evident that the result of the flowrate cascading yields a negative wastewater flowrate (F_{WW}) for the network. This means that the fresh water flowrate of 59.97 ton/h has only fulfilled the impurity load constraints (with none negative Cum.Δm_k values in the final column) but not the flowrate constraint. This flowrate infeasibility can be easily resolved by adding the absolute value of F_{WW} to the F_{FW} used earlier (Table 4) to yield the exact target, i.e. $59.97 + 0.03 = 60.00$ ton/h.

With the revised F_{FW} value, the flowrate cascade yields a network without wastewater, i.e. zero discharge. The revised WCT with revised flowrate cascading is shown in Table 5 while the material recovery pinch diagram (El-Halwagi et al., 2003; Prakash and Shenoy, 2005a) for this problem is shown in Fig. 1. Note that the changes of Δm_k and Cum.Δm_k values in the final two columns of Tables 4 and 5 are insignificant, due to the insignificant change in flowrate cascading ($F_{C,\ k}$).

There are a few special characteristics in Table 5 that are not commonly found in other water network problems, and hence worth mentioning. Firstly, due to the added fresh water flowrate, the pinch concentration is removed

Table 2
Limiting water data for Example 1

Water sink SK_j	Flowrate F_j (ton/h)	Concentration C_j (ppm)	Water source SR_i	Flowrate F_i (ton/h)	Concentration C_i (ppm)
1	50	20	1	20	20
2	20	50	2	50	100
3	100	400	3	40	250

Table 3
Infeasible WCT for Example 1

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)	$F_{FW,k}$ (ton/h)
					0.00			
1	0							
					0.00	0.00		
2	20	50	20	−30			0.00	0.00
					−30.00	−0.90		
3	50	20		−20			−0.90	−18.00
					−50.00	−2.50		
4	100		50	50			−3.40	−34.00
					0.00	0.00		
5	250		40	40			−3.40	−13.60
					40.00	6.00		
6	400	100		−100			2.60	6.50
					−60.00	−59976.00		
7	1000 000						−59973.40	−59.97

Table 4
WCT for Example 1—Infeasible flowrate cascade

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 59.97$		
1	0						
					59.97	1.20	
2	20	50	20	−30			1.20
					29.97	0.90	
3	50	20		−20			2.10
					9.97	0.50	
4	100		50	50			2.60
					59.97	9.00	
5	250		40	40			11.59
					99.97	15.00	
6	400	100		−100			26.59
					$F_{WW} = -0.03$	−26.59	
7	1000 000						0.00 (Pinch)

and no pinch concentration is observed for this case. We shall term the concentration where pinch is firstly located in Table 4 (i.e. 1,000,000 ppm) as the *threshold concentration*, similar to the threshold temperature difference in the heat exchanger network synthesis problems that exhibit similar threshold situation (Linnhoff et al., 1982; Smith, 1995, 2005). Secondly, due to the removal of the pinch concentration, no pinch causing source is found for this case. This is dissimilar to other water network problems in which almost all water networks consist of pinch causing stream(s) at the pinch.

4. Example 2—Network without fresh water feed

Table 6 shows the limiting data of another case study adapted from Jacob et al. (2002). As shown, this water

Table 5
Revised WCT for Example 1 (zero discharge network with fresh water feed)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 60.00$		
1	0						
					60.00	1.20	
2	20	50	20	−30			1.20
					30.00	0.90	
3	50	20		−20			2.10
					10.00	0.50	
4	100		50	50	–		2.60
					60.00	9.00	
5	250		40	40			11.60
					100.00	15.00	
6	400	100		−100			26.60
					$F_{WW} = 0.00$	0.00	
7	1000 000						26.60 (Threshold)

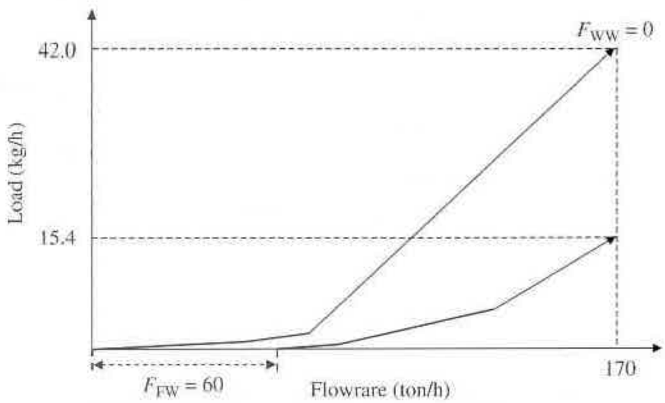


Fig. 1. Material recovery pinch diagram for Example 1.

network has three water sinks and four water sources. Prior to water reuse/recycling, the network requires a fresh water flowrates of 2500 g/min and generates 3200 g/min of wastewater. Flowrate targeting result is shown in WCT in Table 7 and the material recovery pinch diagram (El-Halwagi et al., 2003; Prakash and Shenoy, 2005a) is shown in Fig. 2.

Similar to Example 1, a few special characteristics of Table 7 worth mentioning. Firstly, with the assumed zero fresh water flowrate ($F_{FW} = 0$), none of the Cum.Δm values in the final column is found to be negative. This means that the final column of Table 7 is a feasible load cascade and the determination of $F_{FW,k}$ as illustrated in earlier section is no longer necessary. Hence, the water network does not require any fresh water intake, but will however discharge wastewater at 700 g/min. Secondly, the pinch for this problem lies at the lowest impurity concentration level, i.e. 60 ppm. From Table 7 and Fig. 2, the pinch causing stream with a flowrate of 300 g/min is identified as the water source SR_4. Another special characteristic here is that, the entire flowrate of the pinch causing stream (found in the $F_{C,k}$ column, in interval between 60 and 80 ppm) is sent to the region below the pinch, and none of them to the above pinch region. This situation is uncommon in other water network problems, in which flowrate of the pinch causing stream is often partially allocated to regions above and below the pinch.

All of the above-described characteristic of Examples 1 and 2 are not commonly found in other water network problems studied in any earlier works (Hallale, 2002; El-Halwagi et al., 2003; Prakash and Shenoy, 2005a; Manan et al., 2004; Foo et al., 2006a, b; Almutlaq et al., 2005; Almutlaq and El-Halwagi, 2007). Similar threshold problems were also found in the synthesis of heat exchanger network, in which some networks require only hot utility (cooling duty are fulfilled via process heat recovery) and others require cold utility (heating load is self-sustain) (Linnhoff et al., 1982; Smith, 1995, 2005). Next example will show an industrial case study that exhibits a special case, i.e. combination of the situation found in Examples 1 and 2.

5. Example 3—Special case with no fresh intake and zero discharge

Fig. 3 shows the process flowsheet of an organic chemical production case study from Hall (1997). The limiting data for water demands and sources are given in Table 8. Prior to water reuse/recycling, 40.5 ton/h of fresh water is required and an equivalent amount of wastewater is generated from the network.

Flowrate targeting result is shown in Table 9, where both the fresh water and wastewater requirements are reduced to zero. This is a special case of the previously discussed threshold problems (Examples 1 and 2). Hence, the few special characteristics of Examples 1 and 2 are also found here. As shown in Table 9, a pinch is observed at the lowest concentration of the network, i.e. 22 ppm. Column bottom stream (SK_4) is identified as the pinch causing source at 22 ppm, where the entire source is sent to region below the

Table 6
Limiting water data for Example 2 (Jacob et al., 2002)

Water sink SK_j	Flowrate F_j (g/min)	Concentration C_j (ppm)	Water source SR_i	Flowrate F_i (g/min)	Concentration C_i (ppm)
1	1200	120	1	500	100
2	800	105	2	2000	110
3	500	80	3	400	110
			4	300	60

Table 7
WCT for Example 2 (no fresh feed network)

k	C_k (ppm)	$\Sigma_j F_j$ (g/min)	$\Sigma_i F_i$ (g/min)	$\Sigma_i F_i - \Sigma_j F_j$ (g/min)	$F_{C,k}$ (g/min)	Δm_k (mg/min)	Cum. Δm_k (mg/ min)
					$F_{FW} = 0$		
1	0			0			
					0	0	
2	60		300	300			0 (Pinch)
					300	6	
3	80	−500		−500			6
					−200	−4	
4	100		500	500			2
					300	1.5	
5	105	−800		−800			3.5
					−500	−2.5	
6	110		2400	2400			1
					1900	19	
7	120	−1200		−1200			20
					$F_{WW} = 700$	699916	
8	1000000						699936

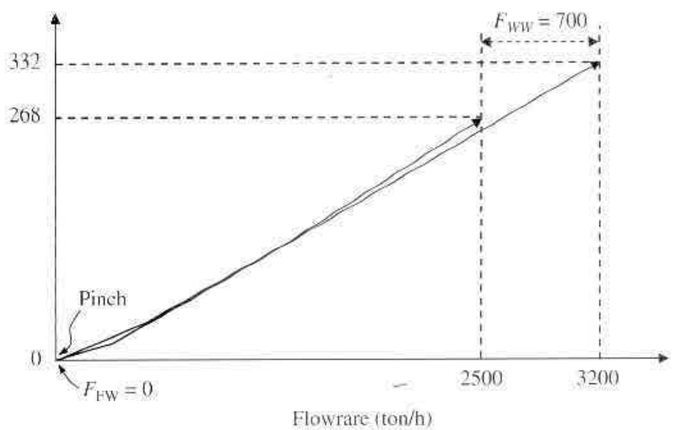


Fig. 2. Material recovery pinch diagram for Example 2.

pinch. This resembles the case of Example 2. On the other hand, threshold concentration is observed at the highest level, i.e. 1,000,000 ppm and no pinch causing source is found, resembles the case of Example 1. The final network configuration to achieve zero fresh water intake and zero discharge is shown in Fig. 4.

Note that there are several earlier works where this case study was reported. In the work of Hall (1997), due to the use of water source and sink composite curves (Dhole et al., 1996) that do not guarantee global optimum, much higher water flowrates were accounted, i.e. 13 ton/h of fresh water and wastewater respectively. Isahak and Amminudin (2004) later suggested adding additional fresh water to satisfy the intake constraint of the wastewater treatment facility. In both cases, a sub-optimal network with much higher water flowrates was synthesised. The true solution for this case is actually a network with zero fresh water intake and zero discharge. However, in practical application, fresh water line and wastewater outlet should be provided to cater for any contingency or process disturbance during operation upset.

6. Plant-wide integration

When water-using processes are located far apart in a plant, geographical constraints should be considered to carry out water reuse/recycle. Often, water-using processes are grouped according to geographical location before water reuse/recycle opportunity is analysed. In another situation, when a few process plants are served by a centralised utility plant, there exists an incentive to consider a plant-wide integration to reduce the overall fresh water and wastewater flowrates that the

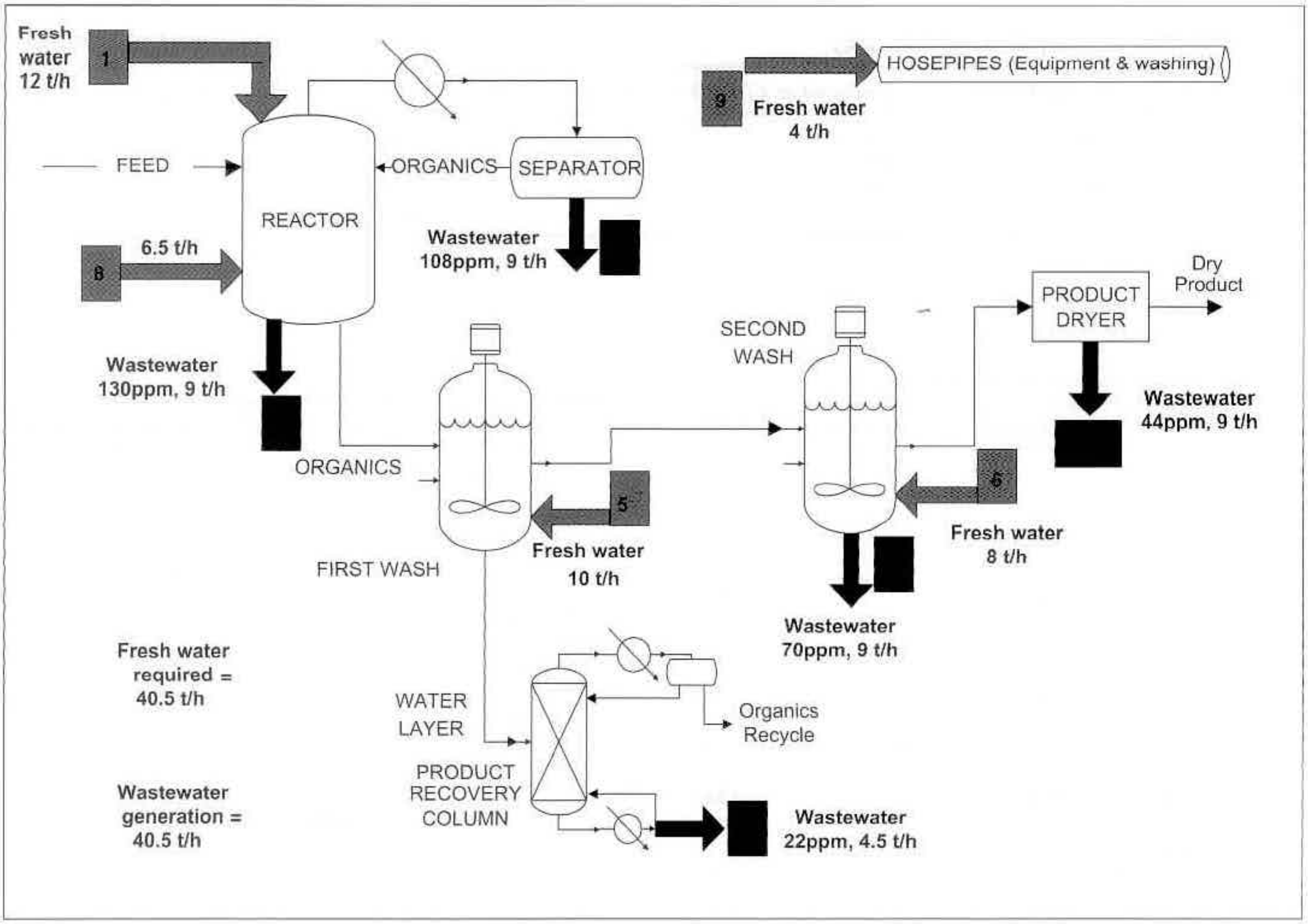


Fig. 3. Process flowsheet for organic chemical production in Example 3 (Hall, 1997).

Table 8
Limiting water data for organic chemical production in Example 3 (Hall, 1997)

Water sinks, SK_j		Flowrate F_j (ton/h)	Concentration C_j (ppm)
j	Stream		
1	Reactor	12	63
5	First wash	10	140
6	Second wash	8	63
8	Stream	6.5	46
9	Hosepipes	4	130
Water sources, SR_i		Flowrate F_i (ton/h)	Concentration C_i (ppm)
i	Stream		
2	Separator	9	108
3	Second wash	9	70
4	Column bottom	4.5	22
7	Reactor discharge	9	130
10	Dryer	9	44

utility plant is serving. Other factors for considering zoning of water-using processes include operability, ease of maintenance and lower piping and instrumentation costs.

The early work on plant-wide integration was found in the area of heat exchanger network synthesis. Ahmad and Hui (1991) termed this as areas of integrity; while Amidpour and Polley (1997) approached the work using their problem decomposition strategy. Olesen and Polley (1996) later extended the problem into water network synthesis. In their work, upon the completion of integration within individual zone/processes, plant-wide integration was carried out to reduce the overall water flowrates of the entire site. Water source in the region below the pinch at one zone is sent to region above the pinch in another zone. Doing this not only will help reducing wastewater discharge in the first zone, but also reducing fresh water flowrate in the second zone. Spriggs et al. (2004) introduced the concept called eco-industrial parks where waste materials are exchanged among different industries within the same industrial zone. Hwang and Chung (1999) had also reported a similar concept on the minimum use of mass separating agents. Besides, similar concept of indirect integration has also been applied to heat and mass integration for batch processes, where heat or mass load were transferred across time intervals to reduce the overall utility targets (Kemp and Deakin, 1989a, b; Foo et al., 2004, 2005a). In this section, flowrate targeting for

Table 9
WCT for Example 3 (organic chemical production)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i$–$\Sigma_j F_j$ (ton/h)	$F_{C,\,k}$ (ton/h)	Δm_k (g/h)	Cum. Δm_k (g/h)
					$F_{FW} = 0$		
1	0			0			
					0	0	
2	22		4.5	4.5			0 (Pinch)
					4.5	99	
3	44		9	9			99
					13.5	27	
4	46	6.5		−6.5			126
					7	119	
5	63	20		−20			245
					−13	−91	
6	70		9	9			154
					−4	−152	
7	108		9	9			2
					5	110	
8	130	4	9	5			112
					10	100	
9	140	10		−10			212
					$F_{WW} = 0$	0	
10	1000 000						212 (Threshold)

plant-wide integration is assessed using WCA technique (Manan et al., 2004: Foo et al., 2006b). It is then extended for threshold problem in the next section.

7. Example 4—Plant-wide integration

Table 10 shows the limiting data for Example 4, taken from Olesen and Polley (1996). As shown, 15 water sinks and sources are categorised in three geographical zones, i.e. Zone A, B and C. Without considering water reuse/recycle, the overall plant consumes 880.15 ton/h of fresh water and generates the same amount of wastewater. By considering water reuse/recycle for individual zones, the total flowrates are reduced 61.4% from 880.15 ton/h to 339.64 ton/h (for both fresh water and wastewater). WCT for targeting in these individual zones are shown in Tables 11–13. Further reduction of total water flowrate is possible by considering plant-wide integration.

From Tables 11–13, it is observed that water network in Zone A has the lowest pinch concentration among three zones, i.e. 100 ppm. Zone C, on the other hand, has a relatively higher pinch concentration of 150 ppm; while the pinch concentration of Zone B is the highest among three zones, i.e. 400 ppm. At this point, it is worth mentioning the significance of the pinch concentration. In the region above the pinch (with concentration lower than the pinch concentration), water is in deficit and fresh water is needed to supplement the need. Below the pinch (with concentration higher than the pinch concentration), water is in excess and wastewater is discharged from this region. Since all three zones are having different pinch concentrations, in principle it is possible to send a "cleaner" water source from regions which have lower pinch concentrations to regions of higher pinch concentrations. Doing this will reduce the overall fresh water and wastewater flowrates of the combined regions. However, it needs to be ensured that the cross-plant water source(s) has a lower concentration than the pinch concentration of the receiving zone, since sending water to region below the pinch will end up in wastewater discharge (Hallale, 2002; Manan et al., 2004). For instance, cross-plant water with Zone B should have a concentration lower than its pinch concentration at 400 ppm.

Due to the existence of three geographical zones, the plant-wide integration problems for Example 4 may further be classified into those that involve only two zones (e.g. Zones A and B) and those that involve all zones (i.e. Zones A, B and C). Both schemes will be analysed in the following section. The first step to carry out such an analysis is to identify the individual wastewater streams that emit in each geographical zone that can be used as the cross-plant water sources. This calls for the use of WCA for waste targeting in a reuse/recycle network (Ng et al., 2006).

Zone A which has the lowest pinch concentration is firstly examined to identify its cross-plant water source(s). From Tables 10 and 11, three water sources are observed to lie in the region below the pinch in Zone A, i.e. SR_1, SR_2 and SR_3. However, not all sources are eligible for plant-wide integration as portion of these sources is needed to fulfil the flowrate requirement of sink SK_5. Following the waste targeting approach of Ng et al. (2006), WCA is carried out for region below the pinch for Zone A (Table 14). As shown, the minimum pinch flowrate (F_P) that fulfils both the flowrate and load constraints for the region below the pinch is determined as 5.71 ton/h. Since 56.67 ton/h of the pinch-causing sources (SR_1 and/or SR_3)

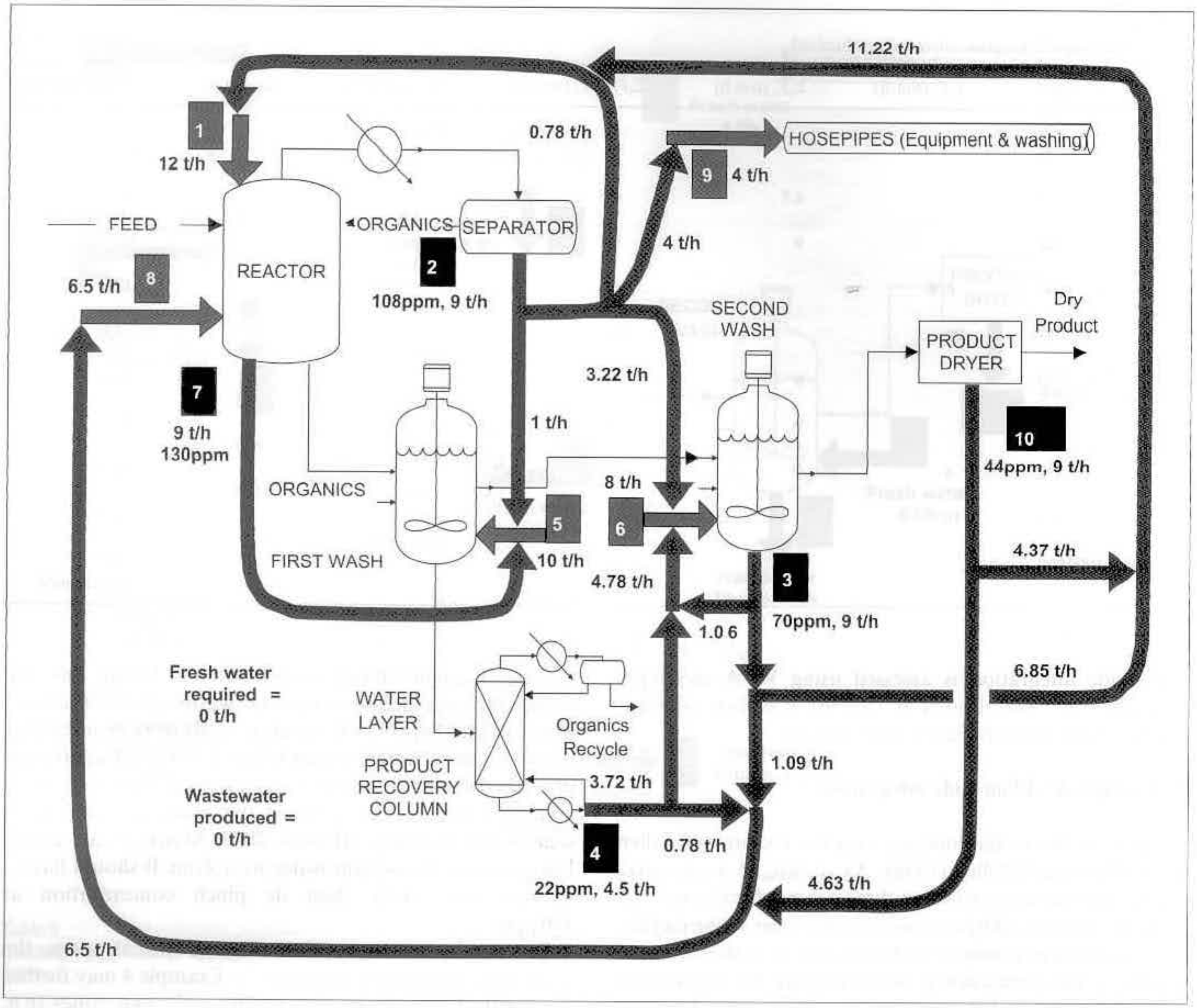


Fig. 4. Water network for Example 3 with water reuse/recycle.

Table 10
Limiting water data for Example 4 (Olesen and Polley, 2006)

Zone	Water sink SK_j	Flowrate F_j (ton/h)	Concentration C_j (ppm)	Water source SR_i	Flowrate F_i (ton/h)	Concentration C_i (ppm)
A	1	20.00	0	1	20.00	100
	2	66.67	50	2	66.67	80
	3	100.00	50	3	100.00	100
	4	41.67	80	4	41.67	800
	5	10.00	400	5	10.00	800
B	6	20.00	0	6	20.00	100
	7	66.67	50	7	66.67	80
	8	15.63	80	8	15.63	400
	9	42.86	100	9	42.86	800
	10	6.67	400	10	6.67	1000
C	11	20.00	0	11	20.00	100
	12	80.00	25	12	80.00	50
	13	50.00	25	13	50.00	125
	14	40.00	50	14	40.00	800
	15	300.00	100	15	300.00	150

Table 11
WCT for Zone A (Example 4)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i – \Sigma_j F_j$ (ton/h)	$F_{C,\,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 98.33$		
1	0	20.00		−20.00			
					78.33	3.92	
2	50	166.67		−166.67			3.92
					−88.33	−2.65	
3	80	41.67	66.67	25.00			1.27
					−63.33	−1.27	
4	100		120.00	120.00			0.00
					56.67	17.00	(Pinch)
5	400	10.00		−10.00			17.00
					46.67	18.67	
6	800		51.67	51.67			35.67
					$F_{WW} = 98.33$	98254.67	
7	1 000 000						98 290.33

Table 12
WCT for Zone B (Example 4)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i – \Sigma_j F_j$ (ton/h)	$F_{C,\,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 54.64$		
1	0	20.00		−20.00			
					34.64	1.73	
2	50	66.67		−66.67			1.73
					−32.02	−0.96	
3	80	15.63	66.67	51.04			0.77
					19.02	0.38	
4	100	42.86	20.00	−22.86			1.15
					−3.84	−1.15	
5	400	6.67	15.63	8.96			0.00
					5.12	2.05	(Pinch)
6	800		42.86	42.86			2.05
					47.98	9.60	
7	1000		6.67	6.67			11.64
					$F_{WW} = 54.64$	54588.21	
8	1000 000						54 599.86

Table 13
WCT for Zone C (Example 4)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i – \Sigma_j F_j$ (ton/h)–	$F_{C,\,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 186.67$		
1	0	20.00		−20.00			
					166.67	4.17	
2	25	130.00		−130.00			4.17
					36.67	0.92	
3	50	40.00	80.00	40.00			5.08
					76.67	3.83	
4	100	300.00	20.00	−280.00			8.92
					−203.33	−5.08	
5	125		50.00	50.00			3.83
					−153.33	−3.83	
6	150		300.00	300.00			0.00
					146.67	95.33	(Pinch)
7	800		40.00	40.00			95.33
					$F_{WW} = 186.67$	186517.33	
8	1000 000						186 612.67

Table 14
Waste targeting for region below the pinch in Zone A (Example 4)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_P = 5.71$		
1	100						
					5.71	1714.29	
2	400	10.00		−10.00			1714.29
					−4.29	−1714.29	
3	800		51.67	51.67			0.00
					$F_W = 47.38$	47343 047.62	
4	1000 000					–	47343 047.62

is allocated to this region in the reuse/recycle network (identified from F_C column just below the pinch concentration of 100 ppm in Table 11), 50.95 ton/h of the pinch-causing source is in excess and will emit as wastewater stream in Zone A. Another wastewater source emits from Zone A is identified from the final column of F_C column in Table 14, i.e. $F_W = 47.38$ ton/h at 800 ppm. Note that summing both the wastewater source flowrates gives the total wastewater flowrate as targeted in Table 11. Hence both the wastewater streams that emit at 100 (SR_3) and 800 ppm (SR_4 and SR_5) are identified as the cross-plant water source candidates for Zone A.

On the other hand, identification of cross-plant water source(s) in the Zone C is relatively straight forward. Since no water sink exist in the region below the pinch, the pinch-causing source flowrate that is allocated to the below pinch region, i.e. 146.67 ton/h (identified from the F_C column just below the pinch concentration of 150 ppm in Table 13), as well as the source at the final concentration level (800 ppm), i.e. 40.00 ton/h, will emit as wastewater streams from this zone. From Table 10, these wastewater streams are identified as SR_4 (800 ppm) and SR_5 (150 ppm) and will serve as the cross-plant water source candidates in Zone C.

7.1. Integration between two geographical zones

Plant-wide integration schemes that involve two zones are firstly examined. Zone B with the highest pinch concentration is firstly considered for the plant-wide integration with Zones A. Due to the pinch concentration of Zone B that is located at 400 ppm, its region above the pinch can only accept a cross-plant water source of lower concentration. This corresponds to 50.95 ton/h wastewater source (SR_3) from Zone A at 100 ppm. Adding this flowrate to Zone B reduces the fresh water flowrate to 45 ton/h (Table 15), consistent with the finding of Olesen and Polley (1996). However, the wastewater flowrate in Zone B has increase drastically to 95.95 ton/h (Table 15). Sending the total available flowrate from Zone A is unnecessary as this leads to bigger cross-plant piping and hence higher piping cost. Hence, one should determine the minimum cross-plant flowrate to be supplied from Zone A. Flowrate targeting for this case resembles the case of multiple fresh water sources targeting (Foo, 2006, 2007). Eq. (1) is used to determine the excessive flowrate of the cross-plant flowrate at each concentration level k, given in the final column of Table 15. Deducting the smallest flowrate in this column (i.e. 38.10 ton/h) from the water source at 100 ppm in Table 14, and recall that Zone B originally consists of a water source at 100 ppm with a flowrate of 20 ton/h, this means that only 12.86 ton/h is required as the cross-plant flowrate from Zone A. Detailed procedure for this multiple sources targeting are referred to the original work (Foo, 2006, 2007). The overall result of the minimum cross-plant flowrate is the reduced wastewater flowrate of 57.86 ton/h in Zone B (Table 16) and a double-pinch network for Zone B. As shown in Table 16, a new limiting pinch concentration has emerged at 80 ppm, in addition to the original pinch concentration at 400 ppm (Table 12), resembling the multiple-pinch problems in water network synthesis (Hallale, 2002; Manan et al., 2004).

On the other hand, if one were to integrate between Zones B and C, the only cross-plant source candidate that is lower than 400 ppm (pinch concentration of Zone B) is the pinch-causing source of Zone C, i.e. SR_{15} with a flowrate of 146.67 ton/h. Following the multiple sources targeting procedure (Foo, 2006, 2007), 10.24 ton/h of this source is required for plant wide integration (Table 17). This leads to the reduced fresh water and wastewater flowrates in Zone C as 48.24 and 58.48 ton/h respectively. Note that a double-pinch network is resulted in Zone B, similar to the earlier scheme where integration is carried out with Zone A. Note also that due to the cross-plant water source from Zone C (150 ppm) is of higher concentration than that in Zone A (100 ppm), hence a higher fresh water and wastewater flowrates are expected in this scheme, as compared to the integration scheme between Zone A and B.

Another plant-wide integration scheme that involves two geographical zones calls for the integration between Zones A and C. Since Zone C has a higher pinch concentration of 150 ppm, wastewater source of Zone A at 100 ppm

Table 15
WCT for Zone B after integration with Zone A (50.95 ton/h from SR_3 at 100 ppm)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)	$F_{FW,k}$ (ton/h)
					$F_{FW} = 45.00$			
1	0	20.00		−20.00				
					25.00	1.25		
2	50	66.67		−66.67			1.25	
					−41.67	−1.25		
3	80	15.63	66.67	51.04			0.00 (Pinch)	
					9.38	0.19		
4	100	42.86	70.95	28.10			0.19	
					37.47	11.24		
5	400	6.67	15.63	8.96			11.43	38.10
					46.43	18.57		
6	800		42.86	42.86			30.00	42.86
					89.29	17.86		
7	1000		6.67	6.67			47.86	53.17
					$F_{WW} = 95.95$	95856.43		
8	1000 000						95 904.29	95.91

Table 16
WCT for Zone B after integration with Zone A (12.86 ton/h from SR_3 at 100 ppm)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 45.00$		
1	0	20.00		−20.00			
					25.00	1.25	
2	50	66.67		−66.67			1.25
					−41.67	−1.25	
3	80	15.63	66.67	51.04			0.00 (Pinch)
					9.38	0.19	
4	100	42.86	32.86	−10.00			0.19
					−0.63	−0.19	
5	400	6.67	15.63	8.96			0.00 (Pinch)
					8.33	3.33	
6	800		42.86	42.86			3.33
					51.19	10.24	
7	1000		6.67	6.67			13.57
					$F_{WW} = 57.86$	57799.29	
8	1000 000						57 812.86

Table 17
WCT for Zone B after integration with Zone C (10.24 ton/h from SR_{15} at 150 ppm)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 48.24$		
1	0	20.00		−20.00			
					28.24	1.41	
2	50	66.67		−66.67			1.41
					−38.42	−1.15	
3	80	15.63	66.67	51.04			0.26
					12.62	0.25	
4	100	42.86	20.00	−22.86			0.51
					−10.24	−0.51	
5	150		10.24	10.30			0.00 (Pinch)
					0.00	0.00	
6	400	6.67	15.63	8.96			0.00 (Pinch)
					8.96	3.58	
7	800		42.86	42.86			3.58
					51.82	10.36	
8	1000		6.67	6.67			13.95
					$F_{WW} = 58.48$	58423.66	
9	1000 000						58 437.61

Table 18
WCT for Zone C after integration with Zone A (50.95 ton/h from SR_3 at 100 ppm)

k	C_k (ppm)	$\Sigma_i F_i$ (ton/h)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 169.68$		
1	0	20.00		−20.00			
					149.68	3.74	
2	25	130.00		−130.00			3.74
					19.68	0.49	
3	50	40.00	80.00	40.00			4.23
					59.68	2.98	
4	100	300.00	70.95	−229.05			7.22
					−169.37	−4.23	
5	125		50.00	50.00			2.98
					−119.37	−2.98	
6	150		300.00	300.00			0.00 (Pinch)
					180.63	117.41	
7	800		40.00	40.00			117.41
					$F_{WW} = 220.63$	220458.41	
8	1000 000						220 575.83

Table 19
WCT for Zone C after integration with Zone A (38.10 ton/h from SR_3 at 100 ppm)

k	C_k (ppm)	$\Sigma_i F_i$ (ton/h)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 173.97$		
1	0	20.00		−20.00			
					153.97	3.85	
2	25	130.00		−130.00			3.85
					23.97	0.60	
3	50	40.00	80.00	40.00			4.45
					63.97	3.20	
4	100	300.00	58.10	−241.90			7.65
					−177.94	−4.45	
5	125		50.00	50.00			3.20
					−127.94	−3.20	
6	150		300.00	300.00			0.00 (Pinch)
					172.06	111.84	
7	800		40.00	40.00			111.84
					$F_{WW} = 212.06$	211 893.84	
8	1000 000						212 005.68

(50.95 ton/h) can be sent to Zone C to reduce its water flowrates. Table 18 shows the result of this scheme. Note that in this scheme, due to the large flowrate deficit in Zone C, the cross-plant flowrate from Zone A does not lead to a double-pinch network in Zone C. Hence the determination of minimum cross-plant flowrate is irrelevant here.

7.2. Integration among all geographical zones

Various plant-wide integration schemes that involve all geographical zones are now analysed. The first option would be to integrate wastewater sources from both Zones A and C with Zone B of the highest pinch concentration. However, targeting result shows that the overall flowrates for this scheme equals that of the integration between Zones A and B. This is the fact that after the integration with Zone A, the resulting network in Zone B has become a multiple pinch problem, with a limiting pinch at 80 ppm and a secondary pinch at 400 ppm (Table 15). Further reduction of fresh water flowrate is only possible if the cross-plant water source from Zone C is available at a concentration lower than 80 ppm (the cleanest possible water source is available at 150 ppm in Zone C). This resembles the typical situation in any multiple pinch problems (Hallale, 2002; Manan et al., 2004; Foo and Manan, 2006).

The second plant-wide integration scheme that involves all zones is the extension of the earlier integration scheme

Table 20
Comparison between different plant-wide integration schemes

Integration schemes	Overall fresh water flowrate (ton/h)	Overall wastewater flowrate (ton/h)	Total cross-plant flowrate (ton/h)	Number of cross-plant interconnections
Integration between two zones				
Zone C → Zone B	333.24	333.24	10.24	1
Zone A → Zone B	330.00	330.00	12.86	1
Zone A → Zone C	322.66	322.66	50.95	1
Integration across all zones				
Zone A → Zone B; Zone C → Zone B	330.00	330.00	19.97	2
Zone A → Zone B; Zone A → Zone C	317.30	317.30	50.95	2
Zone A → Zone C; Zone C → Zone B	316.26	316.26	61.19	2
Without consideration of geographical zones	314.36	314.36	Nil	Nil

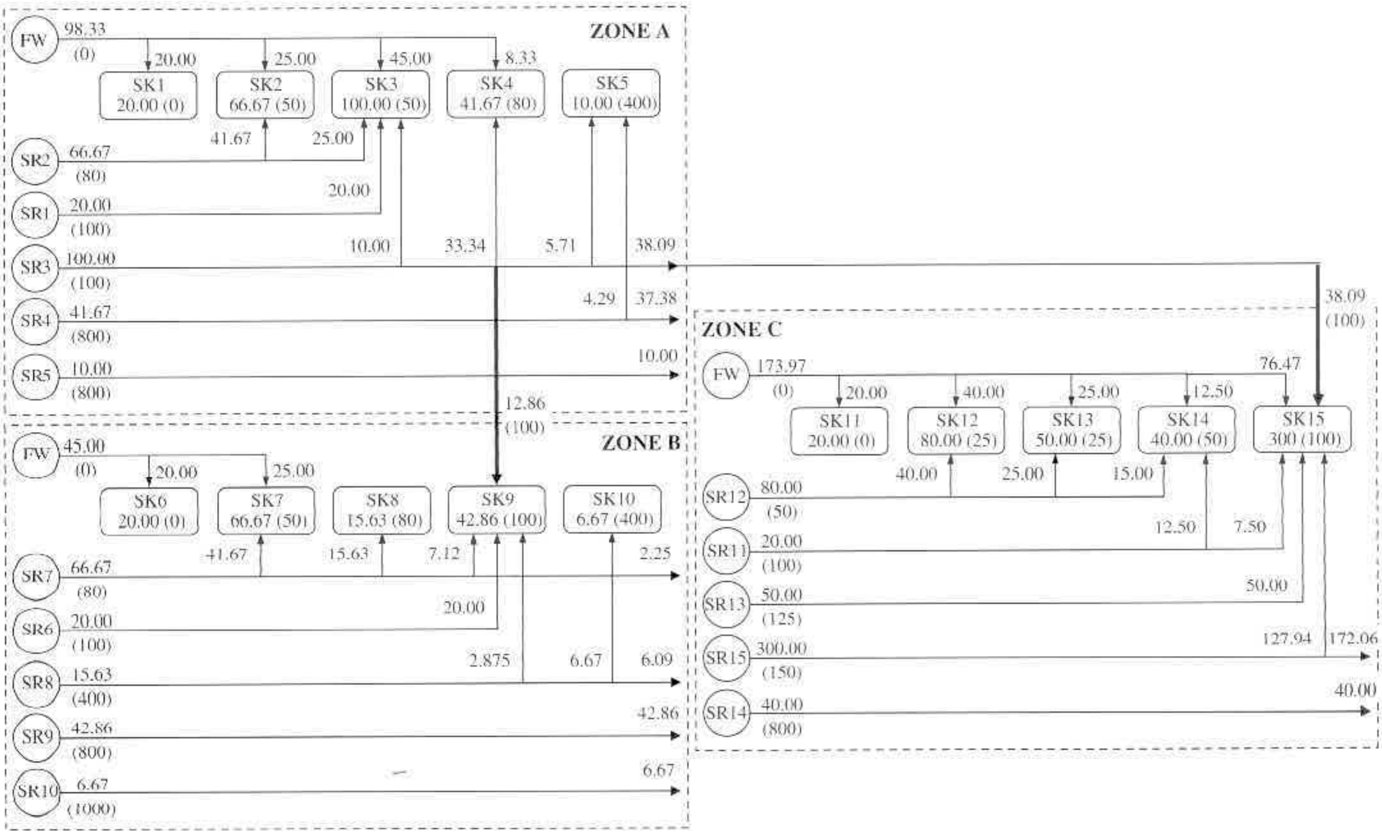


Fig. 5. Water networks for Example 4, with plant-wide integration between Zones A and B (stream in bold). Values indicate stream flowrate (ton/h) and concentration in parenthesis (ppm).

that involves Zones A and B. In this scheme, the leftover wastewater source from Zone A after its integration with Zone B, i.e. 38.10 ton/h (= 50.95–12.86 ton/h) is sent for integration with Zone C. This results in the reduced fresh water and wastewater flowrates in Zone C to 173.97 and 212.06 ton/h, respectively (Table 19).

The last plant-wide integration scheme that worth investigating is the combination of two earlier schemes that involves integration between Zones A and C as well as Zones B and C. In this scheme, the wastewater source from Zone A is sent to Zone C (flowrate targeting reported in Table 18) while the wastewater source from Zone C is sent

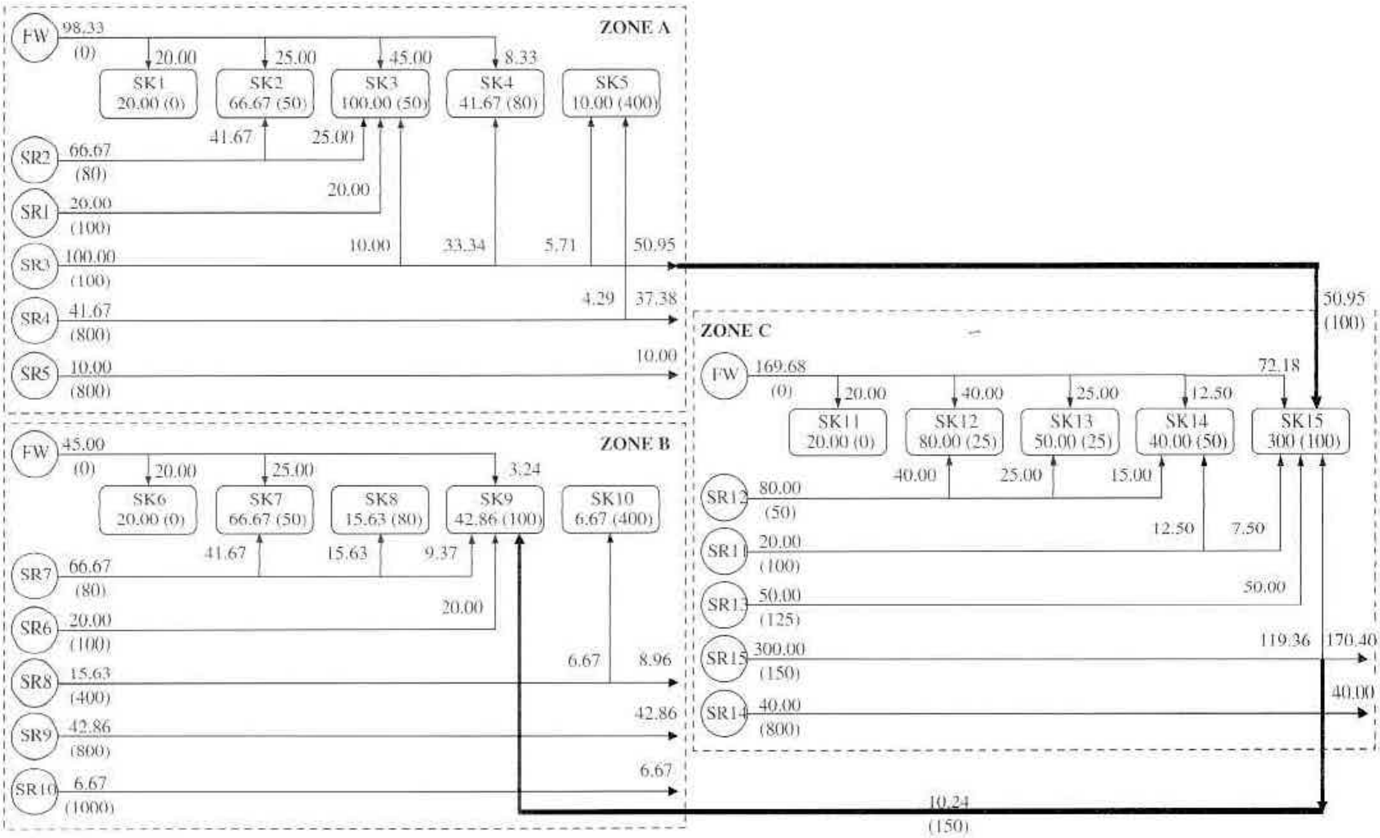


Fig. 6. Water networks for Example 4, with plant-wide integration between Zones A and B (stream in bold). Values indicate stream flowrate (ton/h) and concentration in parenthesis (ppm).

Table 21
WCT without considering geographical zone (Example 4)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 314.36$		
1	0	60.00		−60.00			
					254.36	6.36	
2	25	130.00		−130.00			6.36
					124.36	3.11	
3	50	273.33	80.00	−193.33			9.47
					−68.98	−2.07	
4	80	57.29	133.33	76.04			7.40
					7.06	0.14	
5	100	342.86	160.00	−182.86			7.54
					−175.79	−4.39	
6	125		50.00	50.00			3.14
					−125.79	−3.14	
7	150		300.00	300.00			0.00 (Pinch)
					174.21	43.55	
8	400	16.67	15.63	−1.04			43.55
					173.16	69.27	
9	800		134.52	134.52			112.82
					307.69	61.54	
10	1000		6.67	6.67			174.36
					$F_{WW} = 314.36$	314040.80	
11	1000000						314215.16

to Zone B (flowrate targeting in Table 17). Doing this reduces the fresh water flowrates in Zones B and C as well as the wastewater flowrates in Zones A and C.

Overall results for different plant-wide integration schemes are reported in Table 20, with the total cross-plant flowrates and number of cross-plant interconnections reported in the final two columns. As shown, the final two schemes that involve integration among all zones achieve significant reduction for overall plant fresh water and wastewater flowrates. The overall networks for these schemes are shown in Figs. 5 and 6, designed by the nearest neighbour algorithm of Prakash and Shenoy (2005a). If one were to integrate all water-using processes without considering the geographical zones, the minimum fresh water and wastewater flowrates will be 314.36 ton/h (Table 21). However, this does not appear to be a favourable option considering the additional water saving of less than 1% at the expense of other important potentially significant factors, such as expensive piping costs, reduction of plant operability, etc.

8. Plant-wide integration for threshold problems

Assume that the earlier discussed Examples 1 and 2 are two water networks within the same plant. Conducting a plant-wide integration between the two networks helps to reduce the overall water flowrates. From Tables 5 and 7, it is observed that Example 1 network has a pinch at the highest concentration level (1,000,000 ppm) while Example 2 network has a pinch at 60 ppm. Hence, theoretically all water sources from Example 2 network can be used for plant-wide integration with Example 1 network. However, this is subject to the maximum flowrate available in each water source, as some portion of these sources are integrated with the water sinks in Example 2 network.

Table 22
WCT for Example 2 network after plant-wide integration with Example 1 network

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 0$		
1	0			0			
					0.00	0.00	
2	60		280	280			0.00 (Pinch)
					280.00	5.60	
3	80	−500		−500			5.60
					−220.00	−4.40	
4	100		500	500			1.20
					280.00	1.40	
5	105	−800		−800			2.60
					−520.00	−2.60	
6	110		2386	2386			0.00 (Pinch)
					1866.00	18.66	
7	120	−1200		−1200			18.66
					$F_{WW} = 666.00$	665920.08	
8	1000 000						665 938.74

Table 23
WCT for Example 1 network after plant-wide integration with Example 2 network (source flowrate from Example 2 network are added at 60 and 110 ppm)

k	C_k (ppm)	$\Sigma_j F_j$ (ton/h)	$\Sigma_i F_i$ (ton/h)	$\Sigma_i F_i - \Sigma_j F_j$ (ton/h)	$F_{C,k}$ (ton/h)	Δm_k (kg/h)	Cum. Δm_k (kg/h)
					$F_{FW} = 26.00$		
1	0			0			
					26.00	0.52	
2	20	50	20	−30			0.52
					−4.00	−0.12	
3	50	20		−20			0.40
					−24.00	−0.24	
4	60		20	20			0.16
					−4.00	−0.16	
5	100		50	50			0.00 (Pinch)
					46.00	0.46	
6	110		14	14			0.46
					60.00	8.40	
7	250		40	40			8.86
					100.00	15.00	
8	400	100		−100			23.86
					$F_{WW} = 0.00$	0.00	
9	1000 000			0			23.86

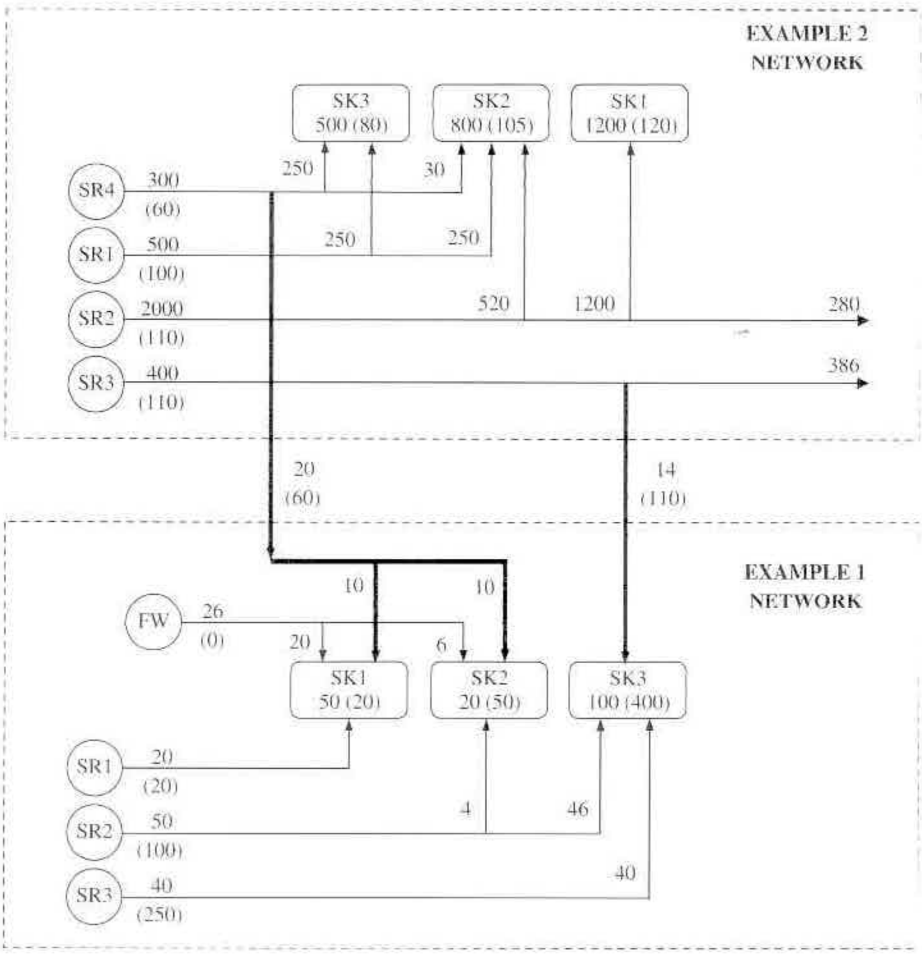


Fig. 7. Water networks for Example 1 and 2, with plant-wide integration (stream in bold). Values indicate stream flowrate (ton/h) and concentration in parenthesis (ppm).

From Table 7, it is observed that four source candidates in Example 2 network can be used for plant-wide integration, i.e. SR_1 (60 ppm), SR_4 (100 ppm), SR_2 and SR_3 (both at 110 ppm). SR_1 with the lowest concentration is the first candidate to be considered for plant-wide integration, as fresh water saving by integrating this source is expected to be higher than using other sources of higher concentration. However, through WCA targeting, only 20 ton/h of this source shall be used for plant-wide integration. Extracting higher flowrate from this source incurs a fresh water penalty in Example 2 network, which defeats the purpose of plant-wide integration (recall that no fresh water is needed originally in this network).

With the extraction of 20 ton/h of water source from SR_1, a double-pinch problem resulted for the Example 2 network. Apart from the original pinch concentration at 60 ppm, a secondary pinch emerges at 110 ppm (Table 22). As a result, SR_4 which lies between the two pinches at 100 ppm cannot be used as a candidate for plant-wide integration (see discussion in Foo et al., 2006b). Hence, two water sources SR_2 or SR_3 at 110 ppm are left as candidate for plant-wide integration. The total available source flowrate is found to be 1866 ton/h (found in intervals 110 and 120 ppm in $F_{C,k}$ column in Table 22). Targeting shows that only 14 ton/h of these sources (either SR_2 or SR_3) is needed for plant-wide integration. Extracting higher flowrate from these sources will not help in reducing the overall fresh water flowrate in Example 1 network, but will just end up as wastewater discharge (recall that Example 1 network is originally a zero discharge network). The WCT for Example 1 network after plant-wide integration is shown in Table 23. As a result, fresh water flowrate of Example 1 network is reduced from 60 to 26 ton/h (Table 23), while wastewater for Example 2 network is reduced from 700 to 666 ton/h (Table 22), corresponds to a reduction of 56.7% and 4.8% respectively. Network design for this case is presented in Fig. 7.

9. Conclusion

Synthesis of water network has thus far focused on the more commonly found water reuse/recycle problems. Little attention has been dedicated to the rare but realistic cases of threshold problems. In this work, targeting for threshold problems in water network has been addressed using the numerical targeting tool of water cascade analysis (WCA) technique. In addition, targeting for plant-wide integration was presented. By sending water sources across different zones via plant-wide integration, overall fresh water and wastewater flowrates are reduced significantly.

Acknowledgement

The financial support from University of Nottingham via Research Committee New Researcher Fund (NRF 3822) is gratefully acknowledged. Feedback of the two reviewers and the editorial help of Dr Raymond Tan (De La Salle University—Manila) are appreciated.

References

Abebe, S., Shang, Z., Kokossis, A., 2003. A two-stage optimisation approach to the design of water-using systems in process plants. European Symposium on Computer-Aided Process Engineering (ESCAPE 13). Elsevier Science, pp. 11–16.

Agrawal, V., Shenoy, U.V., 2006. Unified conceptual approach to targeting and design of water and hydrogen networks. AIChE Journal 52 (3), 1071–1081.

Ahmad, Hui, 1991. Heat recovery between areas of integrity. Computers and Chemical Engineering 15 (12), 809–832.

Almató, M., Sanmart, E., Espuna, A., Puigjaner, L., 1997. Rationalizing the water use in the batch process industry. Computers and Chemical Engineering 21, S971–S976.

Almató, M., Espuña, A., Puigjaner, L., 1999a. Optimisation of water use in batch process industries. Computers and Chemical Engineering 23, 1427–1437.

Almató, M., Sanmatí, E., Espuña, A., Puigjaner, L., 1999b. Economic optimization of the water reuse network in batch process industries. Computers and Chemical Engineering 23, S153–S156.

Almutlaq, A.M., El-Halwagi, M.M., 2007. An algebraic targeting approach to resource conservation via material recycle/reuse. International Journal of Environmental Pollution 29 (1/2/3), 4–18.

Almutlaq, A.M., Kazantzi, V., El-Halwagi, M.M., 2005. An algebraic approach to targeting waste discharge and impure fresh usage via material recycle/reuse networks. Clean Technologies and Environmental Policy 7 (4), 294–305.

Alva-Argáez, A., Kokossis, A.C., Smith, R., 1998. Wastewater minimisation of industrial systems using an integrated approach. Computers and Chemical Engineering 22, S741–S744.

Alva-Argáez, A., Vallianatos, A., Kokossis, A., 1999. A multi-contaminant transhipment model for mass exchange network and wastewater minimisation problems. Computers and Chemical Engineering 23, 1439–1453.

Aly, S., Abeer, S., Awad, M., 2005. A new systematic approach for water network design. Clean Technologies and Environmental Policy 7 (3), 154–161.

Amidpour, M., Polley, G.T., 1997. Application of problem decomposition in process integration. Transactions of the Institute of Chemical Engineers, Part A 75 (A1), 53–63.

Andersen, M., Kristensen, G.H., Wenzel, H., 2002. Tools for evaluation of water reuse scenarios in industry. In: Second International Conference on Industrial Wastewater Recovery and Reuse. Cranfield University, 17 July 2002.

Bagajewicz, M., 2000. A review of recent design procedures for water networks in refineries and process plants. Computers and Chemical Engineering 24, 2093–2113.

Bagajewicz, M., Savelski, M., 2001. On the use of linear models for the design of water utilization systems in process plants with a single contaminant. Transactions of the Institute of Chemical Engineers, Part A 79, 600–610.

Bagajewicz, M., Rivas, M., Savelski, M., 2000. A robust method to obtain optimal and sub-optimal design and retrofit solutions of water utilization systems with multiple contaminants in process plants. Computers and Chemical Engineering 24, 1461–1466.

Bandyopadhyay, S., 2006. Source composite curve for waste reduction. Chemical Engineering Journal 125, 99–110.

Bandyopadhyay, S., Ghanekar, M.D., Pillai, H.K., 2006. Process water management. Industrial and Engineering Chemistry Research 45, 5287–5297.

Bédard, S., Sorin, M., Legault, A., 1999. Reducing the water consumption of industrial processes through process integration. Annual Conference of Association Québécoise de la Maîtrise de l'Énergie.

Benko, N., Rév, E., Szitkai, Z., Fonyó, Z., 1999. Optimal water use and treatment allocation. Computers and Chemical Engineering, S157–S160.

Benko, N., Rév, E., Fonyó, Z., 2000. The use of nonlinear programming to optimal water allocation. Chemical Engineering Communication 178, 67–101.

Brouckaert, C.J., Buckley, C.A., 2000. Water pinch analysis: a strategic tool for water management in the food processing industry. Re-use of Process Water in the Food and Beverage Industries, 17 November.

Brouckaert, C.J., Majozi, T., Gardner, G., Buckley, C.A., 1999. The application of pinch analysis to water and effluent management in the process industry. In: Scott, J.A., Pagan, R.J. (Eds.), Global Competitiveness through Cleaner Production, Proceedings of the 2nd Asia Pacific Cleaner Production Roundtable. 21–23 April 1999. Brisbane, Australia.

Cao, D., Feng, X., Duan, X., 2004. Design of water network with internal mains for multi-containment wastewater regeneration recycle. Transactions of the Institute of Chemical Engineers, Part A 82 (A10), 1331–1336.

Castro, P., Matos, H., Fernandes, M.C., Nunes, C.P., 1999. Improvements for Mass-exchange networks design. Chemical Engineering Science 54, 1649–1665.

Chang, C., Li, B.H., 2005. Design of Discontinuous Water Reuse Systems. 3rd International Symposium on Design, Operation and Control of Chemical Processes (PSE ASIA 2005), Seoul, Korea.

Cheng, K.-F., Chang, C.-T., 2006. Mathematical programming model for designing integrated batch water networks. Paper presented in the 2006 Symposium on Process System Engineering, Taipei, Taiwan.

Chiang, C.L., Hul, S., Ng, D.K.S., Tan, R.R., Foo, D.C.Y., 2006. A linear programming model for Retrofit a Paper Mill Plant. Paper presented in the 2006 Symposium on Process System Engineering, Taipei, Taiwan.

Chungsiriporn, J., Prasertsan, S., Bunyakan, C., 2006. Minimization of water consumption and process optimization of palm oil mills. Clean Technologies and Environmental Policy 8, 151–158.

Delgado, B.E.P.C., Queiroz, E.M., Pessoa, F.L.P., 2006. Application of water source diagram (WSD) procedure in the pulp and paper industry. Paper presented in the Ninth Conference on Process Integration, Modeling and Optimisation for Energy Saving and Pollution Reduction (PRES 2006) in conjunction with 17th International Congress of Chemical and Process Engineering (CHISA), Praha, Czech Republic.

Dhole, V.R., Ramchandani, N., Tainsh, R.A., Wasilewski, M., 1996. Make your process water pay for itself. Chemical Engineering 103, 100–103.

Dunn, R.F., Bush, G.E., 2001. Process integration technology for CLEANER production. Journal of Cleaner Production 9, 1–23.

Doyle, S.J., Smith, R., 1997. Targeting water reuse with multiple contaminants. Transactions of the Institute of Chemical Engineers, Part B 75, 181–189.

Dunn, R., Wenzel, H., 2001. Process integration design method for water conservation and wastewater reduction in industry. Part 1: design for single contaminants. Clean Product and Processes 3, 307–318.

Dunn, R.F., Wenzel, H., Overcash, M.R., 2001. Process integration design method for water conservation and wastewater reduction in industry.

Part 2: design for multiple contaminants. Clean Product and Processes 3, 319–329.

El-Halwagi, M.M., 1997. Pollution Prevention Through Process Integration: Systematic Design Tools. Academic Press, San Diego.

El-Halwagi, M.M., Manousiouthakis, V., 1989. Synthesis of mass exchange networks. AIChE Journal 35 (8), 1233–1244.

El-Halwagi, M.M., El-Halwagi, A.M., Manousiouthakis, V., 1992. Optimal design of dephenolization networks for petroleum-refinery wastes. Transactions of the Institute of Chemical Engineers, Part B 70, 131–139.

El-Halwagi, M.M., Gabriel, F., Harell, D., 2003. Rigorous graphical targeting for resource conservation via material recycle/reuse networks. Industrial and Engineering Chemistry Research 42, 4319–4328.

Feng, X., Chu, K.H., 2004. Cost optimization of industrial wastewater reuse systems. Transactions of the Institute of Chemical Engineers, Part B 82 (B3), 249–255.

Feng, X., Seider, W.D., 2001. New structure and design method for water networks. Industrial and Engineering Chemistry Research 40, 6140–6146.

Feng, X., Shen, R., Zheng, X., 2005. Water network with internal water mains and optimal design methodology. In: Third International Symposium on Design, Operation and Control of Chemical Processes (PSE ASIA 2005), Seoul, Korea.

Feng, X., Wang, N., Chen, E., 2006. Water system integration in a catalyst plant. Transactions of the Institute of Chemical Engineers, Part A 84 (A8), 645–651.

Foo, C.Y., Manan, Z.A., Yunus, R.M., Aziz, R.A., 2004. Synthesis of mass exchange network for batch processes—Part I: utility targeting. Chemical Engineering Science 59 (5), 1009–1026.

Foo, C.Y., Manan, Z.A., Yunus, R.M., Aziz, R.A., 2005a. Synthesis of mass exchange network for batch processes—Part II: Minimum units target and batch network design. Chemical Engineering Science 60 (5), 1349–1362.

Foo, D.C.Y., 2006. Flowrate targeting for multiple fresh water sources using water cascade analysis. Paper presented in 20th Symposium of Malaysian Chemical Engineers (SOMChE 2006).

Foo, D.C.Y., 2007. Water cascade analysis for single and multiple impure fresh water feed. Chemical Engineering Research and Design, in press.

Foo, D.C.Y., Manan, Z.A., 2006. Setting the minimum flowrate targets for utility gases using cascade analysis technique. Industrial and Engineering Chemistry Research 45 (17), 5986–5995.

Foo, D.C.Y., Manan, Z.A., Tan, Y.L., 2005b. Synthesis of maximum water recovery network for batch process systems. Journal of Cleaner Production 13 (15), 1381–1394.

Foo, D.C.Y., Manan, Z.A., El-Halwagi, M.M., 2006a. Correct identification of limiting water data for water network synthesis. Clean Technologies and Environmental Policy 8 (2), 96–104.

Foo, D.C.Y., Manan, Z.A., Tan, Y.L., 2006b. Use cascade analysis to optimize water networks. Chemical Engineering Progress 102 (7), 45–52.

Forstmeier, M., Goers, B., Wozny, G., 2005. Water network optimisation in the process industry—Case study of a liquid detergent plant. Journal of Cleaner Production 13, 495–498.

Gabriel, F.B., El-Halwagi, M.M., 2005. Simultaneous synthesis of waste interception and material reuse networks: problem reformulation for global optimization. Environmental Progress 24 (2), 171–180.

Gianadda, P., Brouckaert, C.J., Sayer, R., Buckley, C.A., 2001. The application of pinch analysis to water, reagen and effluent management in a Chlor-alkali facility. The International Water Association Conference on Water and Wastewater Management for Developing Countries, 29–31 October 2001. Universiti Teknologi Malaysia, Johor Bahru, pp. 743–753.

Gómez, J., Savelski, M., Bagajewicz, M., 2001. On a systematic design procedure for single contaminant water utilization systems in process plant. Chemical Engineering Communication 186, 183–203.

Gomes, J.F.S., Queiroz, E.M., Pessoa, F.L.P., 2006. Design procedure for water/wastewater minimization: single contaminant. Journal of Cleaner Production 15, 474–485.

Hall, S.G., 1997. Water and effluent minimisation. Institution of Chemical Engineers, North Western Branch Papers, No 4, 12 November 1997.

Hallale, N., 2002. A new graphical targeting method for water minimisation. Advances in Environmental Research 6 (3), 377–390.

Hamad, A., Aidan, A., Douboni, M., 2003. Cost-effective wastewater treatment and recycling in mini-plants using mass integration. Clean Technologies and Environmental Policy 4, 246–256.

Huang, C.-H., Chang, C.-T., Ling, H.-C., Chang, C.-C., 1999. A mathematical programming model for water usage and treatment network design. Industrial and Engineering Chemistry Research 38, 2666–2679.

Hul, S., Tan, R.R., Auresenia, J., Fuchino, T., Foo, D.C.Y., 2007. Synthesis of near-optimal topologically-constrained property-based water network using swarm intelligence. Clean Technologies and Environmental Policy 9 (1), 27–36.

Hwang, P.-S., Chung, S.T., 1999. Mass separating agents minimization through process design. Paper presented in the Eighth Asia Pacific Confederation of Chemical Engineering (APCChE 1999), Seoul, Korea.

Isahak, A., Amminudin, K.A., 2004. Wastewater minimisation using water pinch analysis: a pseudo pinch consideration. Paper presented in the 18th Symposium of Malaysian Chemical Engineers (SOMChE 2004).

Jacob, J., Kaipe, H., Couderc, F., Paris, J., 2002. Water network analysis in pulp and paper processes by pinch and linear programming techniques. Chemical Engineering Communication 189 (2), 184–206.

Jeżowski, J., Poplewski, G., Jeżowska, A., 2003. Optimisation of water usage in chemical industry. Environmental Protection Engineering 29 (1), 97–117.

Jödicke, G., Hungerbühler, K., 1999. Reuse of wastewater in a textile dyestuff batch production plant. Recovery, Recycling, Re-integration (R'99 Congress). 1 February 1999.

Jödicke, G., Fisher, U., Hungerbühler, K., 2001. Wastewater reuse: a new approach to screen for designs with minimal total cost. Computers and Chemical Engineering 2, 67–75.

Karuppiah, R., Grossmann, I.E., 2006. Global optimization for the synthesis of integrated water systems in chemical processes. Computers and Chemical Engineering 30, 650–673.

Kemp, I.C., Deakin, A.W., 1989a. The cascade analysis for energy and process integration of batch processes. Part 1: Calculation of energy targets. Transactions of the Institute of Chemical Engineers, Part A 67, 495–509.

Kemp, I.C., Deakin, A.W., 1989b. The cascade analysis for energy and process integration of batch processes. Part 2: Network design and process scheduling. Transactions of the Institute of Chemical Engineers, Part A 67, 510–516.

Kim, J.-K., Smith, R., 2004. Automated design of discontinuous water systems. Transactions of the Institute of Chemical Engineers, Part B 82 (B3), 238–248.

Koppol, A.P.R., Bagajewicz, M.J., 2003. Financial risk management in the design of water utilization systems in process plants. Industrial and Engineering Chemistry Research 42, 5249–5255.

Koppol, A.P.R., Bagajewicz, M.J., Dericks, B.J., Savelski, M.J., 2003. On zero water discharge solutions in the process industry. Advances in Environmental Research 8, 151–171.

Ku-Pineda, V., Tan, R., 2006. Environmental performance optimisation using process water integration and sustainability process index. Journal of Cleaner Production 14, 1586–1592.

Kuo, W.-C.J., Smith, R., 1998a. Design of water-using systems involving regeneration. Transactions of the Institute of Chemical Engineers, Part B 76, 94–114.

Kuo, W.-C.J., Smith, R., 1998b. Designing for the interactions between water-use and effluent treatment. Transactions of the Institute of Chemical Engineers, Part A, 287–301.

Lavric, V., Iancu, P., Plesu, V., 2005. Genetic algorithm optimisation of water consumption and wastewater network topology. Journal of Cleaner Production 13 (15), 1395–1405.

Li, B.H., Chang, C., 2006. A mathematical programming model for discontinuous water-reuse system design. Industrial and Engineering Chemistry Research 45, 5027–5036.

Li, Y., Han, F., Zheng, S., Xiang, S., Tan, X., 2003. An automatic approach to design water utilisation network. In: Chen, B., Westerberg, A.W. (Eds.), Process System Engineering (PSE 2003). Elsevier, Amsterdam, pp. 952–955.

Linnhoff, B., Townsend, D.W., Boland, D., Hewitt, G.F., Thomas, B.E.A., Guy, A.R., Marshall, R.H., 1982. A user Guide on Process Integration for the Efficient use of Energy. IchemE, Rugby.

Liu, Y.A., Lucas, B., Mann, J., 2005a. Up-to-date tools for water-system optimisation. Chemical Engineering 111 (1), 30–41.

Liu, Y.J., Yuan, X.G., Luo, Y.Q., 2005b. Improved concentration interval method for synthesis of water utilization network. Process System Engineering (PSEAsia 2005), Seoul, Korea.

Liu, Z.-Y., Zhang, J.Z., Hu, L., Wu, Z., 2004. Wastewater minimisation using a heuristic procedure. International Journal of Chemical Reactor Engineering 2 (A25), 1–10.

Lovelady, E., El-Halwagi, M.M., Krishnagopalan, G., 2007. An integrated approach to the optimization of water usage and discharge in pulp and paper plants. Journal of Environmental and Pollution 29 (1–3), 274–307.

Luo, Y., Yuan, X., Liu, Y., 2007. An improved PSO algorithm for solving non-convex NLP/MINLP problems with equality constraints. Computers and Chemical Engineering 31, 153–162.

Majozi, T., Brouckaert, C.J., Buckley, C.A., 2006. A graphical technique for wastewater minimisation in batch processes. Journal of Environmental Management 78, 317–329.

Majozi, T., 2005a. Wastewater minimisation using central reusable water storage in batch plants. Computers and Chemical Engineering 29, 1631–1646.

Majozi, T., 2005b. An effective technique for wastewater minimisation in batch processes. Journal of Cleaner Production 13 (15), 1374–1380.

Majozi, T., 2006. Storage design for maximum wastewater reuse in multipurpose batch plants. Industrial and Engineering Chemistry Research 45, 5936–5943.

Manan, Z.A., Tan, Y.L., 2004. Systematic water minimisation in process industry using water cascade analysis technique. Asia Water.

Manan, Z.A., Tan, Y.L., Foo, D.C.Y., 2004. Targeting the minimum water flowrate using water cascade analysis technique. AIChE Journal 50 (12), 3169–3183.

Manan, Z.A., Wan Alwi, S.R., Ujang, Z., 2006. Water pinch analysis for an urban system: a case study on the Sultan Ismail Mosque at the Universiti Teknologi Malaysia (UTM). Desalination 194, 52–68.

Manan, Z.A., Tan, Y.L., Foo, D.C.Y., Tea, S.Y., 2007. Application of water cascade analysis technique for water minimisation in a paper mill plant. International Journal of Environment and Pollution 29 (1–3), 90–103.

Mann, J.G., 2003. Cultural changes and water-asset realignment to support water-reuse projects. Resources, Conservation and Recycling 37, 175–180.

Mann, J.G., Liu, Y.A., 1999. Industrial Water Reuse and Wastewater Minimization. McGraw Hill, New York.

Mariano-Romero, C.E., Alcocer-Yamanaka, V.H., Morales, E.F., 2006. Multi-objective optimization of water-using systems. European Journal of Operational Research, in press.

Ng, D.K.S., Foo, D.C.Y., 2006. Evolution of water network with improved source shift algorithm and water path analysis. Industrial and Engineering Chemistry Research 45, 8095–8104.

Ng, D.K.S., Foo, D.C.Y., Tan, R.R., Tan, Y.L., 2006. Targeting for total water network based on pinch analysis. Paper presented in AIChE Annual Meeting 2006, San Francisco.

Noureldin, M.B., El-Halwagi, M.M., 1999. Interval-based targeting for pollution prevention via mass integration. Computers and Chemical Engineering 23, 1527–1543.

Olesen, S.G., Polley, G.T., 1996. Dealing with plant geography and piping constraints in water network design. Transactions of the Institute of Chemical Engineers, Part B 74, 273–276.

Olesen, S.G., Polley, G.T., 1997. A simple methodology for the design of water networks handling single contaminants. Transactions of the Institute of Chemical Engineers, Part A 75, 420–426.

Parthasarathy, G., Krishnagopalan, G., 2001. Systematic reallocation of aqueous resources using mass integration in a typical pulp mill. Advances in Environmental Research 5, 61–79.

Prakash, R., Shenoy, U.V., 2005a. Targeting and design of water networks for fixed flowrate and fixed contaminant load operations. Chemical Engineering Science 60 (1), 255–268.

Prakash, R., Shenoy, U.V., 2005b. Design and evolution of water networks by source shifts. Chemical Engineering Science 60 (7), 2089–2093.

Prakotpol, D., Srinophakun, T., 2004. GAPinch: genetic algorithm toolbox for water pinch technology. Chemical Engineering and Processing 43 (2), 203–217.

Polley, G.T., Polley, H.L., 2000. Design better water networks. Chemical Engineering Progress 96 (2), 47–52.

Poplewski, G., Jeżowski, J.M., 2005. Stochastic optimization based approach for designing cost optimal water networks. European Symposium on Computer-Aided Process Engineering (ESCAPE 15). 29 May–1 June 2005. Elsevier Science.

Poplewski, G., Jeżowski, J., Jeżowska, A., 2002. Optimal wastewater reuse networks design by adaptive random search optimization. In: 15th International Congress of Chemical and Process Engineering (CHISA2002). 25–29 August 2002. Praha, Czech Republic.

Puigjaner, L., Espuña, A., Almató, M., 2000. A software tool for helping in decision-making about water management in batch process industries. Waste Management 20, 645–649.

Queiroz, E.M., Pessoa, F.L.P., 2006. Water source diagram procedure: wastewater reduction for single contaminant. Paper presented in the Ninth Conference on Process Integration, Modeling and Optimisation for Energy Saving and Pollution Reduction (PRES 2006) in conjunction with 17th International Congress of Chemical and Process Engineering (CHISA), Praha, Czech Republic.

Rossiter, A.P., Nath, R., 1995. Wastewater minimisation using nonlinear programming. In: Rossiter, A.P. (Ed.), Waste Minimisation through Process Design. McGraw Hill, New York.

Savelski, M., Bagajewicz, M., 2000a. Design of water utilization systems in process plants with a single contaminant. Waste Management 20, 659–664.

Savelski, M., Bagajewicz, M., 2000b. On the optimality conditions of water utilization systems in process plants with single contaminants. Chemical Engineering Science 55, 5035–5048.

Savelski, M., Bagajewicz, M., 2001. Algorithmic procedure to design water utilization systems featuring a single contaminant in process plants. Chemical Engineering Science 56, 1897–1911.

Savelski, M., Bagajewicz, M., 2003. On the necessary conditions of optimality of water utilization systems in process plants with multiple contaminants. Chemical Engineering Science 58, 5349–5362.

Shafiei, S., Domenech, S., Koteles, R., Paris, J., 2004. System closure in pulp and paper mills: network analysis by genetic algorithm. Journal of Cleaner Production 12, 131–135.

Smith, R., 1995. Chemical process design. McGraw Hill, New York.

Smith, R., 2005. Chemical process design and integration. John Wiley & Sons, New York.

Sorin, M., Bédard, S., 1999. The global pinch point in water reuse networks. Transactions of the Institute of Chemical Engineers, Part B 77, 305–308.

Spriggs, D., Lowe, E., Watz, J., El-Halwagi, M.M., 2004. Design and development of eco-industrial parks. Paper presented in AIChE Spring Meeting 2004.

Tainsh, R.A., Rudman, A.R., 1999. Practical Techniques and Methods to Develop an Efficient Water Management Strategy. IQPC conference: Water Recycling and Effluent Re-Use. 26–27 April 1999.

Takama, N., Kuriyama, T., Shiroko, K., Umeda, T., 1980a. Optimal water allocation in a petroleum refinery. Computers and Chemical Engineering 4, 251–258.

Takama, N., Kuriyama, T., Shiroko, K., Umeda, T., 1980b. Optimal planning of water allocation in industry. Journal of Chemical Engineering of Japan 13 (6), 478–483.

Takama, N., Kuriyama, T., Shiroko, K., Umeda, T., 1981. On the formulation of optimal water allocation problem by linear programming. Computers and Chemical Engineering 5, 119–121.

Tan, R.R., 2002. Assessing the Sensitivity of Wastewater Reuse Networks to Noisy Loads using Possibility Theory. Chemical Engineering Congress, Manila, Philippines.

Tan, R.R., Cruz, D.E., 2004. Synthesis of robust water reuse networks for single-component source/sink retrofit problems using symmetric fuzzy linear programming. Computers and Chemical Engineering 28, 2547–2551.

Tan, R.R., Foo, D.C.Y., Manan, Z.A., 2007. Assessing the sensitivity of water reuse network to noisy mass loads using Monte Carlo Simulation. Computers and Chemical Engineering, in press.

Tan, Y.L., Manan, Z.A., 2003. Retrofit of Water Network in a Paper Making Process Based on Pinch Analysis Technique. In: 17th Symposium of Malaysian Chemical Engineers (SOMChE 2003), Pulau Pinang, 29–30 December 2003.

Tan, Y.L., Manan, Z.A., 2006. Retrofit of water network with optimization of existing regeneration units. Industrial and Engineering Chemistry Research 45, 7592–7602.

Teles, J., Castro, P., Barbosa-Póvoa, A., Novais, A., 2006. LP-based heuristic procedure for the optimal design of water using networks with multi-contaminants. Paper presented in the Ninth Conference on Process Integration, Modeling and Optimisation for Energy Saving and Pollution Reduction (PRES 2006) in conjunction with 17th International Congress of Chemical and Process Engineering (CHISA), Praha, Czech Republic.

Thevendiraraj, S., Klemeš, J., Paz, D., Aso, G., Cardenas, G.J., 2003. Water and wastewater minimisation study of a citrus plant. Resources, Conservation and Recycling 37, 250–277.

Tripathi, P., 1996. Pinch technology reduces wastewater. Chemical Engineering, 87–90.

Tsai, M.-J., Chang, C.-T., 2001. Water usage and treatment networl design using genetic algorithms. Industrial and Engineering Chemistr Research 40, 4874–4888.

Ujang, Z., Wong, C.L., Manan, Z.A., 2002. Industrial wastewate minimization using water pinch analysis: a case study of an old textil plant. Water Science and Technology 46 (11–12), 77–84.

Vaidyanathan, R., Wilks, T., Manousiouthakis, V., Wilson, S., 1998 Cost-effective design of clean chemical processes by the application o mass integration tools. Hydrocarbon Processing 77 (4), 79–84.

Wan Alwi, S.R., Manan, Z.A., 2006. SHARPS: a new cost-screenin technique to attain cost-effective minimum water network. AIChI Journal 52 (11), 3981–3988.

Wang, Y.P., Smith, R., 1994. Wastewater minimisation. Chemica Engineering Science 49, 981–1006.

Wang, Y.P., Smith, R., 1995a. Wastewater minimization with flowrat constraints. Transactions of the Institute of Chemical Engineers Part A 73, 889–904.

Wang, Y.P., Smith, R., 1995b. Time pinch analysis. Transactions of th Institute of Chemical Engineers, Part A 73, 905–914.

Wang, B., Feng, X., Zhang, Z., 2003. A design methodology for multiple contaminant water networks with single internal water main Computers and Chemical Engineering 27, 903–911.

Wenzel, H., Dunn, R.F., Gottrup, L., Kringelum, J., 2002. Proces integration design methods for water conservation and wastewate reduction in industry. Part 3: experience of industrial application Clean Technologies and Environmental Policy 4, 16–25.

Yang, Y.H., Lou, H.H., Huang, Y.L., 2000. Synthesis of an optima wastewater reuse network. Waste Management 20, 311–319.

Zheng, P., Feng, X., Qian, F., Cao, D., 2006a. Water system integration o a chemical plant. Energy Conversion and Management 47, 2470–2478

Zheng, X., Feng, X., Shen, R., Seider, W., 2006b. Design of optima water-using networks with internal water mains. Industrial an Engineering Chemistry Research 45, 8413–8420.

Zhou, Q., Lou, H.H., Huang, Y.L., 2001. Design of a switchable wate allocation network based on process dynamics. Industrial an Engineering Chemistry Research 40, 4866–4873.

ELSEVIER

Journal of Environmental Management 88 (2008) 275–285

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Issues for small businesses with waste management

Janice Redmond[a,b,*], Elizabeth Walker[b], Calvin Wang[b]

[a] School of Management, Edith Cowan University, 100 Joondalup Drive Joondalup, WA 6027, Australia
[b] Small and Medium Enterprise Research Centre, School of Management, Edith Cowan University, Australia

Received 25 May 2006; received in revised form 12 January 2007; accepted 17 February 2007
Available online 18 April 2007

Abstract

Participation by small and medium enterprise (SME) in corporate social responsibility issues has been found to be lacking. This is a critical issue, as individually SMEs may have little impact on the environment but their collective footprint is significant. The management style and ethical stance of the owner-manager affects business decision making and therefore has a direct impact on the environmental actions of the business. Although adoption of environmental practices to create competitive advantage has been advocated, many businesses see implementation as a cost which cannot be transferred to their customers. After a brief review of pertinent literature this paper reports on an exploratory investigation into the issue. Results show that whereas owner-managers of small enterprises express concern regarding the environment, this does not then translate into better waste management practices.

Keywords: SMEs; Environmental waste management; Attitudes

1. Introduction

Attitudes toward environmental management have been found to be influenced by many internal and external factors such as: company size and resource availability, strategic attitude, sector, and geographic location (Friedman et al., 2000; Gonzales-Benito and Gonzales-Benito, 2006). Small business is defined as an organisation with more than 1 and less than 20 employees and medium business has between 20 and 200 employees (Australian Bureau of Statistics, 2004). For the purposes of this paper these businesses will be referred to collectively as small businesses. These small businesses contribute to waste generation yet they appear not to be committed to waste management practices. As much of the previous research in the field has focussed on large business, business size differences which impact on a small firm's capacity to engage in waste management practices, such as their lower level of resources and flatter management structure may not have been given due attention. This omission may have contributed to the lack of progress made in encouraging small business to engage in waste management practices. Moreover, it is important to business and community sustainability that small businesses are actively engaged both in reducing the waste generated by their business and in implementing practices which appropriately dispose of such waste.

Owner-managers who have the decision-making power in small businesses have expressed interest in the environment yet this interest has not been translated into better waste management practices. As a consequence, the owner-managers of these businesses were surveyed to determine the reasons for the disparity between expressed interest and inaction.

Four factors were explored with the owner-managers: (a) interest in the environment, (b) business impact on the environment, (c) current environmental management practices and (d) awareness of local environmental matters. The first factor was included to confirm that the participant sample was interested in the environment. The other three factors were identified as having the capacity to explain both the disparity between interest and inaction and/or to reveal attitudes which may contribute to a lack of engagement by owner-managers. In particular, the owner-managers' responses to these matters were considered

*Corresponding author. Tel.: +61 08 6304 2153.
E-mail address: j.redmond@ecu.edu.au (J. Redmond).

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.02.006

important indicators of their level of interest and commitment to environmental preservation.

Identifying attitudes that might affect engagement in practices either to reduce or to effectively manage the business impact on the environment is important for both business and community sustainability. In addition, waste volumes were collected to explore the notion that small businesses, as a collective, contribute substantially to the environmental footprint of society. The results of this exploratory study provide a basis for future work to increase small business engagement in effective waste management practices.

2. Background

2.1. Business and environmental management

Globally, businesses are being asked to respond to the concept of corporate social responsibility (Idowu and Towler, 2004; Longenecker et al., 2006; Quazi, 2003) and environmental management has emerged as a very important issue. More specifically, due to its impact on the environment and the community, pressure is mounting on businesses to more effectively manage and reduce their level of waste (Department of Environment, 2005; International Chamber of Commerce, 2001; US EPA, 2003).

Australia currently consumes more resources and "produces more waste than at any time in its history" and "is in the top 10 solid waste generators in the OECD" (Australian Bureau of Statistics, 2003, p. 156). "In 1996–1997, landfills in Australia received 21.2 million tonnes of solid waste, equating to a disposal rate of around 1.146 tonnes per person, or 3.14 kg per person per day" (Australian Bureau of Statistics, 2003, p. 156). The various costs associated with managing this amount of waste exacerbates the issue and supports the need for waste reduction through changes in consumption and behaviour.

The International Chamber of Commerce (ICC) has designed a Business Charter for Sustainable Development as a tool to help companies tackle the challenges and opportunities presented by the environmental issues that emerged during the 1980s and 1990s. The Charter has 16 key principles that provide a basis for good environmental practice. The first of these is the need for businesses to recognise environmental management as being among the highest of corporate priorities and as a key determinant of sustainable development (ICC, 2001).

2.2. Business size as an influence on environmental response

The level of recognition of the importance of environmental management does vary between large and small businesses (Banerjee, 2001; Luetkenhorst, 2004; Sharma et al., 1999). This is partly due to the fact that environmental management, like many corporate social responsibility issues, has primarily been practiced in the "domain of large TNCs [trans-national companies]" (Luetkenhorst, 2004, p. 158).

Gaps between theory and practice of corporate social responsibility in small business are seen as stemming from the narrow research focus on large business. Thompson and Smith (1991) advised that the narrow focus led to assumptions about small business involvement. For example, that small business has less opportunity to exercise social responsibility. Thompson and Smith (1991) suggested that as a consequence of the narrow focus on large business there is a lack of information on how smaller businesses should manage corporate social responsibility tasks and this is a major impediment to progress by small business.

Understanding corporate social responsibility and environmental management practices predominantly from a large business perspective and not from a small business perspective is problematic as small businesses are not scaled down versions of big business (Burns, 1996; Keats and Bracker, 1988). There are major distinctions between businesses of different sizes, primarily in their management practices (Jennings and Beaver, 1997). Worthington and Patton (2005, p. 197) assert that "what drives environmental behaviour of companies is an under-researched and under-developed area of study, particularly in the context of small and medium enterprise (SMEs)".

The level of managerial and financial resources of small business is related to the size of the business and this can have both positive and negative influences on the capacity of the business to implement environmental strategies. Positive influences of business size have been provided by researchers (Condon, 2004; Sarbutts, 2003; Wills, 2003). Condon (2004, p. 57) suggested after working in a series of workshops with a group of SMEs in Australia that small businesses "have a major advantage over larger organisations addressing sustainability issues—their size means that they can react very quickly to changes in the business environment." Moreover, Sarbutts (2003, p. 346) suggested that small business,

> by being flatter [structure] and potentially quicker on their feet and without analysts and shareholders fixated by price/earnings ratios, are better placed than major corporations to take advantage of the fact that society and the media revere qualities such as honesty, integrity and the ability to say sorry...

Another positive for small businesses is that once an environmental strategy is decided upon, because of their relatively small staff numbers, the costs of learning the new routine and renegotiating responsibilities will be less than for a large bureaucratic organisation (Wills, 2003). This does presume however that small businesses are willing to engage in training, which unless proven to be financially beneficial and operationally imperative is not always the case (Storey, 2004; Webster et al., 2005a, b).

Influences of business size which may have negative consequences for implementation of environmental practices have also been highlighted in the literature (Condon, 2004; McKeiver and Gadenne, 2005). McKeiver and Gadenne (2005) found that the number of employees in a business is a driver of both internal and external factors that affect the level of adoption of environmental management. This is in part due to the fact that if the business is very small the sole decision-making responsibility may rest with the owner-manager. Size therefore does matter in many instances and small businesses are economically important from a global perspective, as they are the biggest sector by number in all economies. Small business in Australia, for example, accounts for 97% of all private sector businesses and is recognised as a major contributor to regional economic growth and development (Australian Bureau of Statistics, 2002).

Recent research on small business and environmental management has demonstrated that small business participation is critical for both economic reasons and to achieve sustainable development (Australian Bureau of Statistics, 2002; Condon, 2004; Luetkenhorst, 2004; US EPA, 2003), yet engagement is lacking (Hillary, 2000 cited in McKeiver and Gadenne, 2005). As business involvement in environmental management is a relatively new business issue, literature about small business response to this challenge has only recently emerged.

2.3. *Resources and attitudes as influences on small business response*

While many barriers to small business engagement in environmental practices have been raised by researchers (Condon, 2004; Luetkenhorst, 2004; Simpson et al., 2004; Worthington and Patton, 2005), two issues appear to have the most influence. These are the level of resources available and the prevailing attitudes of the small business owner-managers.

A business resource in this instance refers to undercapitalising, being owned and managed by the same person, and having few employees. In addition, many small businesses are characterised by informality, poor information systems, being operational rather than strategic in their decision making and being time poor (O'Gorman, 2000; Wang et al., 2007; Webster et al., 2005a,b). These characteristics all contribute to the difficulties of implementing good environmental management practices in small businesses, even if the owner-manager attitude is positive.

An example of the impact of limited resources can be found in the recent literature which has investigated formal implementation of environmental management systems. Williams et al. (2000, p. 106) when discussing the implementation of environmental self-regulation systems advised that, "much of this development ... has been in large, well-resourced firms. SMEs have not exhibited the same level of commitment to these new management tools". These results affirm the disparity between the levels of resources available to implement environmental strategies that large business may have compared with small businesses. Williams et al. (2000) also draw attention to the difference in commitment by businesses to environmental practices. The difference in commitment by small business may be due to the fact that in these businesses decisions about the commitment of resources are made by the owner-manager.

The influence of the manager is central to all business operations as "there is a critical relationship between planning in small firms and the strategic-awareness capability of the owner-manager..." (Hannon and Atherton, 1998, p. 114) and "this appears to be strongly influenced by the personal competence of the owner-managers and the type, uncertainty and complexity of the business." Small business owner-managers are known to implement informal rather than formal management practices and to think operationally rather than strategically (Jennings and Beaver, 1997; Upton et al., 2001; Wang et al., 2007). In practice, this means that most small business planning at all levels is at the operational or functional level and involves the day-to-day concerns of running a business and its survival (Longenecker et al., 2006; Shrader et al., 1989; Upton et al., 2001). The majority of small businesses operate on a survival management culture rather than a strategic management culture, making long-term operational changes difficult to action as they require a level of forward thinking. Therefore, as asserted by Luetkenhorst (2004, p. 164) small businesses will be "more dependent on direct economic benefits of CSR-oriented strategies." This need for small business to gain direct benefit from the use of their limited resources is considered a key trigger to encourage environmental action. Indeed, Williamson and Lynch-Wood (2001, p. 431) suggested that a more proactive model is needed after they found that "low commitment levels (i.e., attitudes) [to environmental practice] are supported and reinforced by reactive practices."

Another influence on the use of business resources for environmental management is the owner-manager's ethical stance on the environment. Hornsby et al. (1994, p. 9) noted that among small business owners, "it is apparent that an owner's value system is a critical component of the ethical considerations that surround a business decision." Karp (2003, p. 15) has suggested that, "in an increasingly complex environment, the integrity of the single business leader will matter..."

Small business owner-managers have many diverse demands on their time and finances. This means that having the opportunity and resources to evaluate environmental practice options can be more difficult than it would be for larger businesses (Wills, 2003). To compound this problem owner-managers often lack the managerial skills required to implement practices outside of their technical expertise (Webster et al., 2005b). Therefore, implementation of environmental management strategies may challenge the owner-manager's managerial expertise, ethical stance and their capacity to apply, monitor and evaluate the outcomes of these practices.

While some small businesses do not see environmental issues as important (Condon, 2004), most see them as affecting their business (Simpson et al., 2004), however, this has not always encouraged engagement (Petts et al., 1998; McKeiver and Gadenne, 2005). Another impediment identified by Simpson et al. (2004) was that the owner-managers' view that the costs of managing these matters are not able to be passed onto customers. As a consequence, only a few sought advice about how to reduce the environmental impact of their business. The scant empirical evidence would suggest that whilst owner-managers of small businesses perceive there to be a cost to changing any operational practices to be more "environmental", they are unlikely to make any actual change, thus perpetuating a belief that good environmental practice has a negative effect on the bottom line. Simpson et al. (2004) concluded that the businesses often did not realise the potential of environmental improvements to reduce costs or improve profits. These attitudes, derived from perceived economic implications, contribute to small businesses resistance to reducing the impact of their business (McKeiver and Gadenne, 2005).

In addition, small businesses often have a limited amount of one type of waste and this may contribute to an attitude that suggests that their environmental "footprint" is negligible. It is considered, however, that the collective impact of SMEs on the environment is considerable and could outweigh that of large companies (Hillary, 2000 cited in McKeiver and Gadenne, 2005; Rajendran and Barrett, 2003; Williamson and Lynch-Wood, 2001). When small businesses do not see their business as having a high impact on the environment (McKeiver and Gadenne, 2005) this may further deter them from engaging in environmental activities. Attitudes that support the notions that businesses with small quantities of waste products are not important, or do not need to implement good environmental practices, oppose effective waste reduction and disposal.

When small businesses respond to environmental management issues, they are most often reactive, defensive and frequently limit their response to legislative requirements (Revell and Blackburn, 2004; Worthington and Patton, 2005). Often small businesses do not comply with compulsory regulatory requirements (Rajendran and Barrett, 2003) or even take the time to understand legislative requirements that affect them (Condon, 2004). This said, some small businesses have been shown to have moderate to high "environmentally friendly attitudes" (McKeiver and Gadenne, 2005). Moreover, adoption of environmental practices improves with management-level support (Friedman et al., 2000; Gonzales-Benito and Gonzales-Benito, 2006; Worthington and Patton, 2005) and attitudes to the environment improve with the adoption of environmental management systems (Hillary, 1999 cited in McKeiver and Gadenne, 2005). Two empirical studies provide examples of owner-managers responding to the call to address the waste impact of their businesses.

First, a UK study by Simpson et al. (2004, p. 163) o SMEs found that 73% of respondents "were using or ha used environmental services for waste management." Thi demonstrated that some level of waste management practic had been embedded into their operational strategies. Thi figure was also much higher than for all other environ mental practices (e.g., water management—35%; energ efficiency—49%). In fact, Simpson et al. (2004, p. 166 reported that the majority of the businesses "believed tha 'waste was money' and had a good housekeeping approacl to its management", particularly those in the manufactur ing companies as they appeared to have an interest in th cost savings to be gained from waste reduction.

Second, in a small sample of Australian SMEs McKeiver and Gadenne (2005) examined the variou factors that influenced the implementation of environmen tal management systems and identified both formal (i.e. engaged in a formal certification process) and informa practices (e.g., changing business processes to reduc waste). Formal implementation was found to be enhance by "education, legislation, and awareness" and informa implementation by "...age, customers and employees' (McKeiver and Gadenne, 2005, p. 513). As with othe studies, only some of the businesses had engaged wit 27.1% informal and 11.4% formal implementation o environmental management systems.

The results of both studies from the UK and Australi (Simpson et al., 2004; McKeiver and Gadenne, 2005 respectively) are encouraging and confirm that there i interest in effective waste management among some of th owner-managers of small businesses. However, the result also point to the need for government, employer, employe and community involvement to improve engagement.

Competitive advantage has been promoted as a reaso for encouraging the engagement of small businesses i environmental management practices. However, for thes businesses in the UK attempting to meet the requirement of environmental issues Simpson et al. (2004, p. 156) foun that "few organisations could show that it led to competitive advantage." Therefore, the ability to pre sent the business case for implementing good environmenta practices would seem to be considered critical in influencin small business owner-managers to decide to participate.

2.4. The business case for environmental management practices

It has been suggested that there is no conclusive evidenc that environmental strategies have a positive effect on economic performance for business (Longo et al., 2005; Naffziger et al., 2003). If there was no proof of improved economic performance then small businesses could be excused for not allocating their limited resources to environmental practices as they already have substantial competitive business pressure on them to survive. However, recent research in this area has provided more evidence to support implementation to gain competitive advantage.

Arguments for business to adopt environmental management practices have been made from many different standpoints including: successful market niches and the development of strong consumer loyalty from so-called "green consumers"(Isaak, 1998), operational cost savings, enhanced staff loyalty, improved government relations, innovation and learning, enhanced reputation, and consumer response (Luetkenhorst, 2004). Moreover, Karp (2003, p. 19) proposed that there is sufficient evidence for the business case and stressed that even though there is still debate about whether financial performance can be enhanced "there has been no evidence that social responsibility does not pay back in any form."

Perhaps some of the confusion regarding the potential benefits of adopting environmental strategies may be allayed with education as Simpson et al. (2004, p. 160) have suggested that, as many small businesses do not realise the benefits that can be gained from pursuing environmental improvements, "there is an opportunity for support services, regulators and stakeholders to raise awareness of environmental issues as they relate to businesses, evidencing cost savings and encouraging businesses to become more sustainable." Luetkenhorst (2004, p. 158) goes further by stating:

> At the end of the day, CSR will only prevail and remain an important force if SMEs can be effectively engaged and if CSR can be shown to impact on the development agenda, i.e., first and foremost on enhancing productivity as a long-term determinant of economic growth.

In summary the literature indicates strong support for small businesses to implement effective environmental management practices to gain competitive advantage and sustain both the business and the community. This does however require the allocation of resources by the small business owner-manager to engage in good environmental practices. The literature has shown that one difficulty for small business owner-managers is that they have limited resources and it appears that the assumption is that these resources will be allocated to environmental management only if the business case can be made. While this may be the case, existing attitudes which reduce the importance of the environment may prevent this from occurring. Therefore, it is important to understand the current attitudes of small business owner-managers which affect implementation of environmental management practices.

3. Methodology

3.1. Aim

The aim of this research was to test the following two propositions:

1. That owner-managers of small business would express interest in the environment yet this would not necessarily translate into better waste management practices, as action would be influenced by the owner-managers' attitudes and awareness of the local environment.
2. That the volume of waste generated by small business in one industrial area could be extrapolated to indicate the collective impact on the environment of all small businesses is considerable.

3.2. Research questions and design

As this was an exploratory investigation, the research design incorporated quantitative methodology to collect categorical data from the small business owner-managers. In addition, qualitative data was collected to provide explanation of barriers which might be influencing the owner-managers' decisions to engage in waste management practices. To examine the two research propositions, the main research questions were:

1. What are the small business owner-managers' current attitudes toward the environment?
2. What barriers did owner-managers experience which limited their engagement in waste management practices?
3. What is the approximate volume of waste produced by small businesses in this area in one week?

The owner-managers' interest in the environment, whether they thought their business had an impact on the environment, the current environmental management practices within their business and their awareness of local environmental matters were also explored.

3.3. Survey instrument

Preliminary meetings with the stakeholders led to the development of a survey instrument which was piloted with eight small business operators in the industrial area. After analysing the pilot data further input from professionals working within local environmental management organisations was sought and the survey instrument refined. A decision was made not to include the data collected from the pilot study beyond this point as some survey items differed and therefore the results were not directly comparable.

The final 30 item instrument consisted of a mixture of questions which related to the business (e.g., What is your business?), the environment (e.g., Are you interested in the environment?), waste management (e.g., What type(s) and approximate volume of waste is produced and disposed of during your business operations each week?) and the local environment (e.g., Where do the local stormwater drains flow?). Three answer formats were used in the survey, yes-no, multiple-choice and longer answer format. Prior to conducting the main survey checks of the instrument for both face validity and content validity were made (Cavana et al., 2001).

The survey instrument was taken by hand to the total population of business owner-managers in an urban industrial area (147–8 pilot respondents = 139) over the period November 2005 to February 2006 (breaking for 1 month prior to and after Christmas). This area is one of several small light industrial areas located within the metropolitan area of Perth, Western Australia. The local area is experiencing rapid growth and impacts on the environment from industrial pollution sources have been reported in the light industrial area in recent years. Thus the light industrial area was selected as an appropriate location for this investigation due to its small size and environmental history.

The business owner-managers were given details regarding the purpose of the study and advised that participation was voluntary. This methodology provided an excellent participation rate (86%) as 120 businesses agreed to participate.

3.4. Data analysis

As this was an exploratory study frequencies were calculated from the categorical data collected to investigate the owner-managers' interest, knowledge and practice of environmental management. In addition, the volume of waste generated and costs associated with recycling the waste were also collected. Where we report volume of waste products the data has been reduced as far as possible to provide a meaningful response. For example, steel volumes were collected in many different volume capacities to allow the respondents to provide the data in their own context and therefore included cubic metres, tonnes, tons, and kilograms. The Australian Government national metric conversion tables (Department of Industry, Tourism and Resources, 2006) were utilised to reduce the data to two types of volume (i.e., cubic metres and tons).

Qualitative data was collected to identify the barriers to engagement in waste management practices. The barriers were categorised into internal or external factors to identify the location of the difficulty.

4. Results and discussion

After providing the contextual data for the study, the two propositions that were tested in this research are discussed in relation to their implications for small business engagement in environmental waste management.

4.1. Contextual data

Profiles of the owner-managers who contributed the data and information of their businesses are shown in Tables 1 and 2.

The demographic profile of the business owner and the businesses are consistent with light "industrial" areas in Australia and elsewhere. That is, the types of businesses are still dominated by independent older male owner-operators, with trade related skills, operating at a micro level (i.e., under five employees). This profile needs to be borne in mind when strategies are suggested to deal with environmental issues, as this cohort can be one of the most difficult to engage in regard to behaviour change. It is, however, the very group that collectively is responsible for a high level of industrial waste.

Table 1
Profile of the owner manager

	Gender (%)	Age (%)	Highest education (%)
Male	85		
Female	15		
Under 30		5	
31–40		18	
41–50		30	
51–60		35	
Over 60	–	12	
High school			33
TAFE			13
Trade			34
University			20

4.2. Proposition 1

4.2.1. Small business interest and implementation

It was necessary before testing Proposition 1 to confirm that this sample of small business owner-managers were interested in the environment before an assessment could be made of whether this would then translate into better waste management practices.

The results confirmed that these owner-managers were interested in the environment, with a positive response given by 98% of respondents who rated environmental issues as important or very important. These results provide the basis on which to test the first proposition of the study and they are consistent with other research (McKeiver and Gadenne, 2005; Simpson et al., 2004; Williams et al., 2000).

Owner-managers with a strong commitment to the environment were expected to have measures in place to ensure that they disposed of their waste appropriately. Therefore, their actions toward waste management were probed (e.g., having measures in place to prevent waste water polluting the environment). Waste water was a focus in the study as the industrial area under scrutiny is extremely close to a major waterway and as such, waste water disposal is a critical local environmental issue.

When the owner-managers were asked whether they had prevention measures in place for waste water produced by their business, 58% advised they did not produce waste water. Of those who produced waste water 69% had prevention measures in place and the other 31% did not. The potential for even a small amount of mismanagement by these owner-managers could cause environmental harm

Table 2
Profile of the businesses

	Type business by (%) (ANZSIC code)	Structure (%)	Premises (%)	Employees (%)
Agriculture	3			
Mining	1			
Manufacturing	25			
Motor vehicle services	24			
Other service	12			
Production	3			
Retail trade	17			
Technical services	4			
Transport and storage	4			
Wholesale trade	6			
Independently owner		89		
Subsidiary or branch		8		
Head office		1		
Franchise		2		
Leased			50	
Owned			50	
Full time staff				
1 only				20
2–5				54
6–19				19
20 and over				7

Note: Where totals do not add to 100% no response was given.

as the area lacks deep sewerage. In addition, the majority of the businesses (59%) were unaware of where the stormwater drains flowed. Without measures in place, any polluted water flows directly to the natural waterways.

Together, a lack of prevention measures and a limited awareness of facts about a key contributor to ensuring liquid waste is disposed of correctly, show that the owner-managers' commitment is somewhat at odds with the previous high acknowledgement of the importance of the environment. These findings support the contention stated in Proposition 1 that the owner-managers' interest in the environment would not necessarily translate into better waste management. This environmental rhetoric versus business reality concept is often explained by the barriers that small businesses face in pursuing good environmental management practices.

Part of the reason for this gap between interest and practice may have been identified in another piece of evidence that supported Proposition 1. The owner-managers were concerned about general impacts on the environment (99%) however, their interest was due to the impact on their family (63%), thus making it a personal rather than an economic reason. If their interest in the environment is based on family protection rather than environmental preservation this could account for some of the lack of engagement by these businesses, which is an area for future research.

4.2.2. Small business impact on the environment

Many of the small businesses owner-managers (61%) acknowledged the impact of their business on the environment. Unfortunately, a considerable number of the small business owner-managers did not see their business as having an impact on the environment. This result provides partial support for the statement in Proposition 1 that the owner-managers' attitudes may influence their action. Importantly, the result indicates that there is a need to demonstrate to the small business owner-managers how their business has an impact on the environment.

Table 3
External and internal barriers to recycling

External barriers	Internal barriers
Cannot even give away	Already recycling as much as possible
Cannot recycle	Cost
Shire does not recycle all products	Lack of knowledge
Lack of facilities	Not always viable
Advice of bin supply	Sorting material
Lack of government support	Space
Lack of bins and/or bin space	Time
No notification of Shire recycling	Willingness of staff to comply
Locating a suitable contractor	
Unreliable storage	

4.2.3. Small business barriers to engagement in recycling waste

By gathering qualitative data additional barriers to those identified in the previous literature were revealed. Table 3 illustrates that the owner-managers saw the majority of the barriers to be external (59%), with only 26% identifying an internal barrier to recycling.

The nature of the external barriers calls for both greater support from government and others responsible for environmental management and an increase in co-operation between those organisations and small business. However, it can be argued that the internal barriers are of greater concern because these can be more difficult to change. While some are inherent to small businesses (i.e., time, space) many of these are a matter of perception and/or attitude (doing as much as possible, cost, willingness, lack of knowledge, sorting materials). Further education of owner-manager about the benefits to their business of implementing good environmental practices may result in critical changes being made within the workplace to reduce the internal barriers to participation.

4.3. Proposition 2

4.3.1. Environmental impact of waste by small business

The second proposition to be tested was that the volume of waste generated by small business in this industrial area would indicate that the collective impact on the environment of all small business is considerable. Table 4 provides an overview of the main types and approximate volume of waste produced in the area per week from the 120 small businesses. As can be seen, over 350 cubic metres of waste material and 3500 litres of liquid waste is being generated in this small industrial area each week.

When viewed on an annual basis these weekly amounts convert to over 18 000 cubic metres of waste material and over 184 000 litres of fluid. The majority of the liquid from this area was waste oil which generated approximately 180 000 litres per year from this area. This is equivalent to filling between two and three domestic swimming pools per year (based on a 50 000–70 000 litre pool capacity) or three sea containers (based on 40 ft dry cargo Hi-cube container). "Used oil is a valuable resource. Even though Australians are good at recycling used oil (around 194 million litres was recycled last year), at least 60 million litres of used oil goes missing annually" (Department of Environment and Heritage, 2005, p. 1). In addition, the potential danger is obvious when it is known that "oil is also a pollutant: it takes only one litre of oil to contaminate one million litres of water (which is about half the size of an Olympic swimming pool), and a single automotive oil change produces four to five litres of used oil" (Department of Environment and Heritage, 2006, p. 1). While it is not being suggested that small business owner-managers are responsible for the problems associated with oil used in vehicles, it is important that when they, like others, use oil or service vehicles that the oil is disposed of correctly.

With the potential for multiplication of the generated waste by many other similar small businesses in Australia, these results suggest support for Proposition 2. Unfortunately, as reported by Bremner (2006) it is difficult to make clear conclusions about the collective impact on the environment of all small businesses as there is a lack of waste and recycling data available in Western Australia. Therefore, while more data will need to be collected using other samples, this research provides initial support that Hillary's (1999, cited in McKeiver and Gadenne, 2005) contention that small business operations have a substantial impact on the environment may be correct.

4.3.2. Competitive advantage from waste management

While not a specific aim of the study, the notion that small business will be interested in participating in waste management strategies when they can see a direct benefit was supported by the results of this research. The evidence

Table 4
Type and volume of waste produced by SMEs in one week

Product	Volume			
	Cubic metres	Tons	Units	Litres
Metals				
Steel	26	117		
Cardboard				
Packing material, boxes and paper	171			
Plastics				
Including polystyrene, containers, bags, packing film, and vehicle parts	144			
Wood				
Including pallets, MDF, solid timber, and particle board	12		40	
Batteries			35	
Rubber				
Tyres			420	
Liquids				
Waste oil				3288
Radiator coolant				265

Note: Amounts have been rounded off to nearest full number.

for this was given by some of the owner-managers that provided details on products that were both recycled and were income generators for their business. However, those who did not make a competitive advantage also engaged. An explanation of why the small businesses may have chosen the path they did is provided and then it is suggested that those who are not gaining advantage should be assisted to achieve benefits for their participation.

Strategies to gain benefit from recycling efforts of the businesses varied. Some businesses had a defined procedure for waste disposal and therefore reaped the commensurate financial benefit, while others simply accepted whatever financial benefit they received from removal contractors. More worrying was that other owner-managers paid contractors for the same sort of waste to be removed from the business premises. This provides further evidence that many small businesses work at a survival/operational level rather than planning strategically, and that this can have negative economic consequences for the business. Three examples are given to provide illustration of the results of the different approaches used by the owner-manager.

Steel is one of the most economically valuable products to recycling contractors, yet the benefits reported by the small businesses ranged from a cost to the business for removal, to an income of $A90 per tonne.

Paper when given to recyclers also cost some businesses yet others received $A38 per load. In this industrial area, 171 cubic metres of cardboard and paper products were recycled. According to Visyboard (2006) it is estimated that every "one tonne of recycled paper or cardboard saves: approximately 13 trees, 2.5 barrels of oil, 4100 kW h of electricity, 4 cubic metres of landfill and 31 780 litres of water." In addition, other broad effects on the environment are evident in paper manufacturing processes including the production of significant quantities of waste water (Thompson et al., 2001). Recycling can assist in the reduction of these effects on the environment.

A third example is oil recycling, where some businesses were paying to have oil removed while others were receiving up to $A40 per drum for their waste oil. At the highest resale value given ($A40 per drum), the market recycle value of the collective annual oil production from the small businesses within the industrial area is approximately $A36 000. While it is acknowledged that this is not a great deal of money, when the average annual turnover of a small business in Australia ranges from $A82 000 to 1.08 million (ABS, 2002), loss of any profit should be important for those at the lower end of the scale. Moreover, recovered waste oil can be recycled and approximately 80% of this oil reused. Therefore, this resource is not only of some economic value to small businesses, the recycling efforts are critical particularly in the 21st century when greater pressure is being placed on the availability of these commodities due to growing consumption rates globally.

These figures indicate that some businesses gain competitive advantage and contribute to the community by implementing good waste management practices and some do not. The owner-managers' views also highlighted that they were often unaware of what happened to the products once they were given to the recycling contractors and no longer "their business". Combined, the evidence of different economic return for generated waste and lack of awareness of where waste is disposed of after it leaves the business premises provide opportunities for local environmental organisations to work with the business owner-managers and waste removal contractors to provide the best outcomes for the business, the community and the environment.

4.4. *Summary of results*

The results indicated support for Proposition 1 that their interest in the environment would not necessarily translate into better waste management as a limited number of businesses had implemented prevention measures to ensure that they disposed of their waste appropriately.

It was also established during the testing of Proposition 1 that the owner-managers are more interested in the environment for personal rather than economic reasons. This may confuse the research on environmental practices by small business as the assumption may be that by saying that they are interested in the environment they mean that they are interested in preserving the environment and this may not be the case.

Partial support was provided that the owner-managers' actions were being influenced by their attitudes and awareness of the local environment. This was manifest in three results. First, the majority of owner-managers had an attitude that their business had no impact on the environment. Second, the internal barriers identified by the owner-managers to engagement related to perceptions and/or attitudes. Third, the owner-managers' low awareness of local environmental matters which assist appropriate waste disposal indicated a lack of commitment to the local and global environment. Each of these was considered to be a contributor to inaction. Although only partial support was given here, it is considered that their attitudes and awareness did influence their actions. However, there is insufficient data to measure the link between these variables.

Proposition 2 was supported by the research findings as the total volume of waste generated in this one industrial area supported the view that small businesses as a collective may have a considerable impact on the environment and this impact could outweigh that of large companies.

5. Conclusion

Sustainable business requires an appropriate response to change and in today's society, environmental issues have become critical as it is recognised that the outcomes achieved in environmental management areas such as waste management will have an important impact on the local community.

The importance of both small businesses and sound environmental management to the growth and sustainability of developed countries has been acknowledged and as the environmental fragility becomes more apparent, it is the obligation of everyone to take responsibility. Some inroads have already been made on environmental matters by both large and small businesses, however, more needs to be done to encourage greater commitment by small business to engage with the local community in which they operate.

The literature has shown that small businesses have several advantages over large businesses that could enable them to address the environmental impact of the waste generated from their business. However, it is still not clear how great the gap is between their interest in the environment and their day-to-day practices to reduce the impact from their business operations. There still appears to be an element of rhetoric versus reality among small business owners and little progress has been made about how to bridge this gap.

The attitudes of small business owner-managers have been shown to have an influence on what happens in their business and on the decision to participate (or not) in any waste management practices. The evidence is growing that, by implementing environmental management practices, lasting economic benefits can be made and that these practices can create competitive advantage. Small businesses need to be made aware of this and further evidence will continue to be needed to prove the business case to small business owner-managers. An approach to education of this sector is also needed that will provide the owner-managers with knowledge about how to make appropriate changes in their business to take advantage of the benefits. It is suggested that when the education programs are developed they consider the perspective of the small business owner-manager and emphasise the benefits to be derived from incorporating business environmental strategies.

A snapshot of the current attitudes of owner-managers regarding waste management practices in one industrial area has been provided to illustrate their influence on the implementation of environmental management practices. Increasing engagement by these and other small business owner-managers to reduce waste and implement good environmental management practices is crucial to local and national economic and social outcomes. It is clear that this will need the co-operation of government, businesses and the local community to initiate and sustain the changes necessary to create and maintain a sustainable environment.

References

Australian Bureau of Statistics, 2002. Small Business in Australia 2001 (Electronic version). Australian Bureau of Statistics, Canberra.

Australian Bureau of Statistics, 2003. Australia's Environment: Issues and Trends, Cat. No. 4613.0 (Electronic version). Australian Bureau of Statistics, Canberra.

Australian Bureau of Statistics, 2004. Characteristics of Small Business: Australia, Cat. No. 8127.0 (Electronic version). Australian Bureau of Statistics, Canberra.

Banerjee, S.B., 2001. Managerial perceptions of corporate environmen alism: interpretations from industry and strategic implications fc organizations. Journal of Management Studies 38 (4), 489–513.

Bremner, A., 2006. Is there actually any crisis? Measurement an reporting of the waste management situation in Western Australi: Paper Presented at the Waste & Recycle 2006 Conference, 19–2 September, 2006, Fremantle Western Australia.

Burns, P., 1996. Introduction: the significance of small firms. In: Burns, P Dewhurst, J. (Eds.), Small Business and Entrepreneurship, second ec MacMillan, Basingstoke.

Cavana, R.Y., Delahaye, B.L., Sekaran, U., 2001. Applied Busines Research: Qualitative and Quantitative Methods. Wiley, Australia.

Condon, L., 2004. Sustainability and small to medium sized enterprises– How to engage them. Australian Journal of Environmental Educatio 20 (1), 57–67.

Department of Environment, 2005. Extended producer responsibility policy statement. Retrieved April 27, 2006 from 〈www.zerowastewa com.au/ourwork〉.

Department of Environment and Heritage, 2005. Used oil—it can b recycled. Retrieved April 27, 2006 from 〈www.oilrecycling.gov.au factsheet-2.html〉.

Department of Environment and Heritage, 2006. Used oil—health an environmental impacts. Retrieved April 27, 2006 from 〈ww.oilrecycling gov.au/factsheet-3.html〉.

Department of Industry, Tourism and Resources, 2006. Metric conver sion. Retrieved April 20, 2006 from 〈www.measurement.gov.au index.cfm?event=conversions〉.

Friedman, A.L., Miles, S., Adams, C., 2000. Small and medium-size enterprises and the environment: evaluation of a specific initiativ aimed at all small and medium-sized enterprises. Journal or Sma Business and Enterprise Development 7 (4), 325–342.

Gonzales-Benito, J., Gonzales-Benito, O., 2006. A review of determinan factors of environmental proactivity. Business Strategy and th Environment 15, 87–102.

Hannon, P.D., Atherton, A., 1998. Small firm success and the art c orienteering: the values of plans, planning and strategic awareness i the competitive small firm. Journal of Small Business and Enterpris Development 5 (2), 102–119.

Hornsby, J.S., Kuratko, D.F., Naffziger, D.W., LaFollette, W.R Hodgetts, R.M., 1994. The ethical perceptions of small busines owners: a factor analytic study. Journal of Small Business Manage ment 32 (4), 9–16.

Idowu, S.O., Towler, B.A., 2004. A comparative study of the contents o corporate social responsibility reports of UK companies. Managemen of Environmental Quality: An International Journal 15 (4), 420–437

International Chamber of Commerce, 2001. The Business Charter fo Sustainable Development. Retrieved April 20, 2006 from 〈www old.iccwbo.org/home/environment_and_energy/sdcharter〉.

Isaak, R., 1998. Green Logic: Ecopreneurship, Theory and Ethics Greenleaf, London.

Jennings, P., Beaver, G., 1997. The performance and competitiv advantage of small firms: a management perspective. Internationa Small Business Journal 15 (2), 63–68.

Karp, T., 2003. Social responsible leadership. Foresight 5 (2), 15–23.

Keats, B., Bracker, J.S., 1988. Toward a theory of small firm performance a conceptual model. Amercian Journal of Small Business 12 (4) 41–58.

Longenecker, J.G., Moore, C.W., Petty, J.W., Palich, L.E., 2006. Smal Business Management: An Entrepreneurial Emphasis, 13th ed Thomson, Australia.

Longo, M., Mura, M., Bonoli, A., 2005. Corporate social responsibility and corporate performance: the case of Italian SMEs. Corporat Governance 5 (4), 28–42.

Luetkenhorst, W., 2004. Corporate social responsibility and the develop ment agenda. Intereconomics 39 (3), 157–166.

McKeiver, C., Gadenne, D., 2005. Environmental management systems in small and medium businesses. International Small Business Journal 2: (5), 513–537, 10.1177/0266242605055910.

Naffziger, D.W., Ahmed, N.U., Montagno, R.V., 2003. Perceptions of environmental consciousness in US small businesses: an empirical study. SAM Advanced Management Journal 68 (2), 23–32.

O'Gorman, C., 2000. Strategy and the small firm. In: Carter, S., Jones-Evans, D. (Eds.), Enterprise and Small Business: Principles, Practice and Policy. Prentice Hall, Harlow, England.

Petts, J., Herd, A., O'hEocha, M., 1998. Environmental responsiveness, individuals and organisational learning: SME experience. Journal of Environmental Planning and Management 41 (6), 711–730.

Quazi, A.M., 2003. Identifying the determinants of corporate manager's perceived social obligations. Management Decisions 41 (9), 822–831.

Rajendran, D., Barrett, R., 2003. Managing environmental risk in small business: an agenda for research. Paper Presented at the 16th Annual Small Enterprise Association of Australia and New Zealand (SEAANZ) Conference 28 September—2 October, 2003, Australia.

Revell, A., Blackburn, R., 2004. UK SMEs and their Response to Environmental Issues: Executive Summary, Small Business Research Centre, Kingston University, Surrey.

Sarbutts, N., 2003. Can SMEs 'do' CSR? A practitioner's view of the ways small- and medium-sized enterprises are able to manage reputation through corporate social responsibility. Journal of Communication Management 7 (4), 340–347.

Sharma, S., Pablo, A., Vredenburg, H., 1999. Corporate environmental responsiveness strategies: the importance of issue interpretation and organizational context. The Journal of Applied Behavioural Science 35 (1), 87–108.

Shrader, C.B., Mulford, C.L., Blackburn, V.L., 1989. Strategic and operational planning, uncertainty and performance in small firms. Journal of Small Business Management 27 (4), 45–60.

Simpson, M., Taylor, N., Barker, K., 2004. Environmental responsibility in SMEs: does it deliver competitive advantage. Business Strategy and the Environment 13 (3), 156–171.

Storey, D., 2004. Exploring the link, among small firms, between management training and firm performance: a comparison between the UK and other OECD countries. International Journal of Resource Management 15 (1), 112–130.

Thompson, G., Swain, J., Kay, M., Forster, C.F., 2001. The treatment of pulp and paper mill effluent: a review. Bioresource Technology 77 (3), 275–286.

Thompson, J.K., Smith, H.L., 1991. Social responsibility and small business: suggestions for research. Journal of Small Business Management 29 (1), 30–45.

Upton, N., Teal, E.J., Felan, J.T., 2001. Strategic business planning practices of fast growing family firms. Journal of Small Business Management 39 (1), 60–72.

US EPA, 2003. Unifying EPA's small business activities: a strategy to meet the needs of small businesses. Retrieved April 20, 2006 from ⟨www.smallbiz-enviroweb.org/html/pdf/Stragegy_Final_062703.pdf⟩.

Visyboard, 2006. About Visy: frequently asked questions. Retrieved April 29, 2006 from ⟨http://www.visy.com.au/about/about_faqs.aspx⟩.

Wang, C., Walker, E.A., Redmond, J.L., 2007. Explaining the lack of strategic planning in SMEs: the importance of owner motivation. International Journal of Organisational Behavior 12 (1), 1–16.

Webster, B.J., Walker, E.A., Barrett, R., 2005a. Small business and on-line training: who is willing to participate? New Technologies Work and Employment 20 (3), 248–258.

Webster, B.J., Walker, E.A., Brown, A., 2005b. Small business participation in training activities. Education and Training 47 (8/9), 552–561.

Williams, H., van Hooydonk, A., Dingle, P., Annandale, D., 2000. Developing tailored environmental management systems for small business. Eco-Management and Auditing 7 (3), 106–113.

Williamson, D., Lynch-Wood, G., 2001. A new paradigm for SME environmental practice. The TQM Magazine 13 (6), 424–432.

Wills, I., 2003. Informational barriers to pollution reduction in small business. Economic Society of Australia 22 (2), 84–95.

Worthington, I., Patton, D., 2005. Strategic intent in the management of the green environment within SMEs. Long Range Planning 38, 197–212.

Journal of Environmental Management

Journal of Environmental Management 88 (2008) 286–292

www.elsevier.com/locate/jenvman

Identification of pollution of Tapeng Lagoon from neighbouring rivers using multivariate statistical method

Shao-Wei Liao*, Hwa-Sheng Gau, Wen-Liang Lai, Jen-Jeng Chen, Chang-Gai Lee

Department of Environmental Resources Management, Tajen University, Pingtong, Taiwan 10907, Republic of China

Received 24 May 2006; received in revised form 25 January 2007; accepted 17 February 2007
Available online 4 May 2007

Abstract

This work investigated water samples collected from Tapeng Lagoon and three neighbouring rivers (the Kaoping River, Tungkang River and Lingbeng River) in Taiwan, Republic of China. Canonical discriminant analysis was applied to identify the source of pollution in neighbouring rivers outside Tapeng Lagoon. The two constructed discriminant functions showed a marked contribution to all discriminant variables, and the total nitrogen, algae, dissolved oxygen and total phosphate were combined as the nutrient effect factor. The recognition capacities of the two discriminant functions were 95.6% and 4.4%, respectively. The water quality in the Kaoping River most strongly controlled the water quality in Tapeng Lagoon. Disassembling the oyster frames and fishery boxes had improved the water quality markedly. The methodology and results provide useful information concerning watershed management and may be applicable to other basins with similar properties that are experiencing similar coastal environmental issues.

Keywords: Canonical discriminant analysis; River water quality; Tapeng Lagoon

1. Introduction

Marine environmental resources were an important global issue over the past decade, especially in the field of estuarine water-quality management. In Taiwan, Republic of China, and throughout the rest of the world, end-point treatment of point pollution sources is the best way to improve the quality of water and prevent coastal pollution. Non-point pollution sources that influence the quality of surface water are also serious concerns. Rivers carry municipal and industrial wastewater as well as run-off from agricultural land in vast drainage basins. Therefore, these water bodies are highly vulnerable to pollution (Singh et al., 2005). The quality of surface water within a region is governed by both natural processes (such as precipitation rate, weathering processes and soil erosion) and anthropogenic effects (such as urban, industrial and agricultural activities, and the human exploitation of water resources) (Jarvie et al., 1998). Seasonal variations in precipitation, surface run-off, groundwater flow, interception and abstraction strongly affect river discharge and, consequently the concentrations of pollutants in river water (Vega et al., 1998). The pollution of rivers must therefore be controlled or prevented, and reliable information on the quality of water must be obtained in order to support effective management. Lagoon remediation is both inefficient and costly. Therefore, effective pollution control and water-resource management, especially in lagoons, depend upon identifying the main sources of pollution.

Human activities have gradually destroyed the balance of the ecosystem in Tapeng Lagoon, Taiwan. The seawater exchange rate is not sufficient to eliminate pollutants because the process only occurs through one pocket-shape of hole. Therefore, the quality of the water is becoming worse and eutrophication is occurring. The medium-to-strong pollution levels in the Lingbeng River, Tungkang River and Kaoping River during rising tidal periods, and the sinking of the inner canals during ebbing tidal periods, probably influence the pollutants in Tapeng Lagoon. Water-quality data reveal that at many monitoring points

*Corresponding author. Tel.: +886 8762 4002x361; fax: +886 8762 2397.
E-mail address: swliao@mail.tajen.edu.tw (S.-W. Liao).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.02.010

along the Kaoping River and the Tungkang River, the water quality is poor. In the current work, data obtained after disassembling the oyster frames and fishery boxes in the Tapeng Lagoon 13 observations of 12 water parameters, and in the neighbouring rivers 21 observations of 14 water parameters, were analysed using canonical discriminant analysis (CDA) approaches. These statistics were applied in order to distinguish the effect of neighbouring rivers on the water-quality parameters, and to identify the sources of pollution from neighbouring rivers. The water quality was also analysed both in the rivers and in the lagoon. In addition, CDA was used to construct the functions of the monitoring parameters, identify and estimate the contributions of possible sources to the parameters associated with neighbouring rivers. The aim of the current work was to perform CDA in order to demonstrate its applicability and effectiveness in environmental research. This is the first study in Taiwan that has taken such an approach. The results will be helpful in developing a methodology for use by the government in refining its management programmes.

2. Materials and methods

2.1. Site description

Tapeng Lagoon is a semi-enclosed coastal lagoon that is located at a latitude of around 120°27′ north and a longitude of around 22°26′ east. The lagoon has only one tidal inlet, through which water is exchanged with the coastal currents along the Kaoping coast, which is located on a narrow shelf in southwest of Taiwan. Tapeng Lagoon has a total water area of around 532 ha. The study area was located in a zone with a tropical climate, and was exposed to sunshine throughout the year, except for an average of 28.9 rainy days during the summer. The southwest monsoon influenced the rainfall. The study area was enclosed by the Taiwan Strait to the west, the central mountains to the east, the Kaoping River and the Tungkang River to the north, and the Lingbeng River to the south.

The Lingbeng River, Tungkang River and Kaoping River have total lengths of around 42, 44 and 171 km, before they drain into the Taiwan Strait, with catchments areas of around 336, 472 and 3257 km^2, which are home to around 0.1, 0.5 and 2 million people, respectively. The Lingbeng River has suffered almost no industrial pollution. The Kaoping River has the highest annual run-off in Taiwan (around 8455 million m^3). These rivers are direct recipients of untreated domestic wastewater and effluents from a few (mainly petrochemical) industries located along their courses. They also receive agricultural run-off from its non-point sources throughout their vast catchments areas. Tapeng Lagoon might be affected by these neighbouring rivers. Seasonal variations strongly affect the ecosystem of this area, which is known as the Bay area. Aquaculture, including oyster farming and fish cages, is common in the Bay area, because the water is calm and nutrient-rich. In fact, the quality of the lake water has reportedly worsened and eutrophication has occurred as a result of their pollution (Liao et al., 2003). However, all of the fishery facilities were disassembled when the Tapeng Lagoon National Scenic Area was established in 2003. Therefore, the water quality improved after the oyster frames and fishery boxes used in aquaculture were removed. But the balance between the nature and anthropogenic disturbance has been disrupted.

2.2. Water sampling and analysis

Water samples were collected during July 2004 at 13 sites in Tapeng Lagoon. A grid-sampling method was used to obtain representative measures of the quality of the lagoon water. In addition, 21 sites (seven in each river) were sampled in three neighbouring rivers outside the lagoon during July and August 2004, in order to identify the source that had the most negative effect on the quality of water in Tapeng Lagoon (Fig. 1). The sampling, preservation and transportation of the water samples followed the standard methods of the American Public Health Association (APHA, 1995). The temperature, electrical conductivity (EC) and pH of the water were measured on-site using a mercury thermometer, EC meter (WTW LF 320/Set) and pH meter (WTW pH-320 Set-2), respectively. All other parameters were determined in the laboratory following the standard protocols (APHA, 1995). The samples were analysed in order to evaluate 14 parameters as follows: temperature (T), total alkalinity (T-Alk), concentration of chloride (Cl^-), EC, total concentration of kjeldahl nitrogen (TKN), concentration of phosphate (PO_4^{-2}), pH, concentration of dissolved oxygen (DO), UV-254, total concentration of coliform (T. Coli), number of algae (algae), concentration of chlorophyl-*a* (Chloro-*a*), concentration of sulphate (SO_4^{-2}) and total solids (TS). The quality of the analytical data was ensured by careful standardization, procedural blank measurements and triplicate samples. The laboratory also participated in a national analytical quality-control programme.

2.3. Statistical methods

Environmental monitoring typically generates vast amounts of data, which are difficult to analyse and interpret because the relationships among the variables are complex. Multivariate approaches have been used successfully to support the interpretation of complex field measurements, and to extract meaningful information from such databases (Simeonov et al., 2003a, b; Lambrakis et al., 2004; Singh et al., 2004, 2005; Papatheodorou et al., 2006). CDA can be used to interpret the spatial distribution of bioassemblages with multiple environmental parameters (Cruz-Castillo et al., 1994; Momen and Zehr, 1998; Comber et al., 2005; Liao and Chang, 2005).

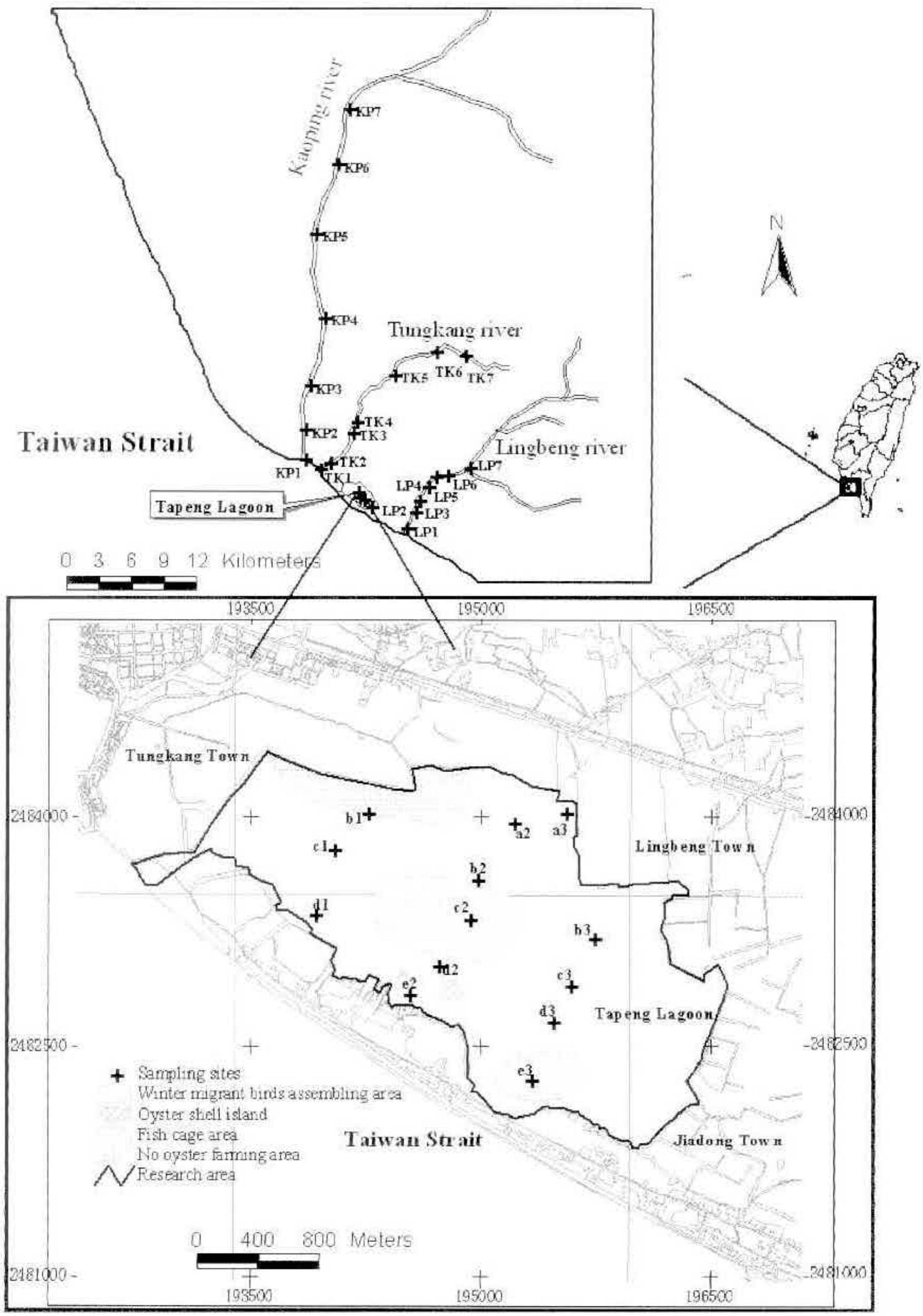


Fig. 1. Water sampling sites in Tapeng Lagoon and neighbouring rivers.

CDA determines how a set of quantitative variables can differentiate among several known classes. It yields linear functions of quantitative variables that maximally separate two or more groups of individuals, while minimizing the variation within groups. This approach distinguishe uncorrelated canonical discriminant functions (CDFs) which are the linear combinations of the original variable that most strongly separate the average values of th groups of observations (Rencher, 1992). The maximum number of CDFs is the smaller of the number of independent variables and the number of class variables minus one. The first CDF, which is denoted as CDF_1, quantifies the maximum possible variation between groups. CDF_2 represents the group differences that are not captured by CDF_1, and is not correlated with CDF_1. Similarly, CDF_3 is not correlated with either CDF_1 or CDF_2, and indicates group differences that are not displayed by CDF_1 or CDF_2, and so on (Sharma, 1996). This implies that each CDF extracts a unique dimension of information from the dataset. The CDA approach also allows the relationships among the groups to be graphically represented by plotting the values of the canonical scores of sample observations. As well as identifying outliers within the data, these plots can also be used to assign a new observation to an existing group. The canonical scores for a new observation are determined for the first two (or first few) CDFs and its position is plotted. The new observation is then assigned to the group whose average value is closest to its position.

Two main statistics explain the characteristics and structure of the CDFs: first, the total canonical structure coefficients (TCSCs), which are the correlation coefficients between the individual variables and the canonical scores (similar to the variable loadings in factor analysis); and second, the total standardized canonical coefficients (TSCCs), which are the multipliers of the standardized independent variables that yield the standardized canonical scores. The validity and use of these two indices have been discussed elsewhere, although controversies still remain (Rencher, 1992; Cruz-Castillo et al., 1994; Huberty, 1994; Matthew et al., 1994; Momen and Zehr, 1998). The TSCCs specify the joint effects of the independent variables of a given CDF, and so are more informative than the TCSCs (Rencher, 1992). However, the TSCCs can be misleading when independent variables are related to each other (Cruz-Castillo et al., 1994). In the current work, the TCSCs were used to interpret the CDFs, because significant correlations were obtained among some of the independent variables (Table 1). The TSCCs were used to yield the CDFs in the present analysis.

In the current work, spatial variations in the river water-quality parameters for Tapeng Lagoon were evaluated using a correlation matrix and the Spearman non-parametric correlation coefficient. Multivariate analysis of the water-quality data was conducted using a CDA approaches. Because the water-quality variables had different measurement units, CDA was applied after the experimental data had been standardized by a z-scale transformation: this approach was intended to prevent any misclassification caused by variations in the dimensionality of the data (Rencher, 1992). Standardization tends to minimize the effects of variation among the variances and to eliminate the effects of different measurement units.

Three neighbouring rivers were selected for evaluating the source of pollution to the lagoon. Stepwise multiple-discriminant analysis was conducted to correlate the river groupings with the physical and chemical variables obtained from the water samples. The 14 parameters (T,

Table 1
Pairwise correlation coefficients of water data in the three neighbouring rivers of Tapeng Lagoon

Discriminate variables	T	Alk	Cl	EC	TKN	TP	UV-254	pH	HPC	DO	Algae	Chloro-a	SO_4^{-2}	TS
T	1.00													
Alk	0.33	1.00												
Cl	−0.10	−0.02	1.00											
EC	−0.07	−0.05	**0.55**	1.00										
TKN	0.38	−0.01	−0.13	−0.20	1.00									
TP	**0.64**	0.28	−0.07	−0.12	**0.77**	1.00								
UV-254	0.03	0.10	−0.04	0.08	0.13	0.22	1.00							
pH	−0.20	**−0.54**	0.09	0.04	**−0.49**	−0.39	−0.23	1.00						
HPC	**0.67**	0.35	−0.12	−0.10	0.39	**0.65**	0.16	−0.13	1.00					
DO	**0.46**	−0.15	0.06	0.16	−0.12	0.06	−0.22	0.38	0.36	1.00				
Algae	−0.29	−0.09	−0.21	−0.03	−0.39	**−0.45**	−0.22	0.11	−0.15	−0.12	1.00			
Chloro-a	**0.54**	**0.45**	−0.06	−0.10	**0.62**	**0.82**	0.10	−0.39	**0.83**	0.08	−0.22	1.00		
SO_4^{-2}	−0.10	0.07	**0.98**	0.42	−0.13	−0.06	−0.04	0.06	−0.08	0.02	−0.22	−0.02	1.00	
TS	−0.08	0.01	0.13	−0.05	−0.23	−0.12	0.02	0.22	−0.08	0.24	−0.12	0.05	0.14	1.00
KP														
Mean	28.64	206.70	516.00	1786.71	0.18	0.00	0.62	7.95	5.42E+07	5.74	191.29	22.86	125.30	984.43
SD	0.41	130.09	1108.07	2726.10	0.08	0.00	0.04	0.30	7.27E+07	1.97	101.55	5.71	49.72	2251.56
TK														
Mean	28.78	148.27	1817.44	1157.86	0.83	0.01	98.37	7.76	2.76E+07	4.46	81.71	22.86	239.85	343.57
SD	0.95	53.71	4622.29	1385.78	0.25	0.01	258.24	0.30	3.87E+07	1.58	35.07	5.71	408.14	595.52
LP														
Mean	30.79	188.50	266.36	1206.00	0.46	0.01	0.84	7.91	7.48E+07	7.03	113.00	20.17	92.64	561.14
SD	1.94	74.21	461.09	1440.98	0.48	0.02	0.09	0.19	1.19E+08	1.13	47.64	34.23	17.84	307.38

Bold values are significant at $P<0.05$; Abbreviations: KP, Kaoping River; TK, Tungkang River; LP, Lingbeng River; SD, standard derivations.

T-Alk, Cl^-, EC, TKN, PO_4^{-2}, pH, DO, UV-254, T. Coli, Algae, Chloro-*a*, SO_4^{-2} and TS) were divided into groups according to the three rivers based on their classification and ordination. The calculation was performed using the STATISTICA package from StatSoft (1996).

3. Results and discussion

CDA was applied to the normalized data matrix (21×14) for July 2004, in order to determine the CDFs that governed the observed responses. The class membership for the three rivers (the Kaoping River, Tungkang River and Lingbeng River) was lower than the number of independent variables (14 water parameters), so two CDFs were determined. These main water-quality parameters were assumed to affect the river classes. A forward-stepwise approach was applied to determine which of the variables could be incorporated into the model (StatSoft, 1996). An *F*-test was conducted to identify the most discriminating variables. The process was terminated when the differences ceased to be significant. Table 2 shows that all of the discriminant variables included was significant according to Wilks' Lambda test. The order of inclusion in the model according to the *F*-test was as follows: TKN, Chloro-*a*, Algae, DO, HPC, EC, UV-254, TP, *T*, pH, Cl^-, TS, SO_4^{-2} and Alk. The factors we considered were the ocean-caused factor, the nutrient factor, the dissolved-oxygen factor, the primary-productivity factor and the environmental-pollution factor. Accordingly, the rivers showed great variation. The Lingbeng River was not polluted by industrial wastewater, while the other rivers were polluted by industrial domestic and agricultural wastewater. The discriminatory capacity of the most important variables in the models followed the following order: TKN, Chloro-*a*, Algae, DO, HPC, EC, UV-254 and TP. TKN and TP are nutrients that influence the large-scale formation of algae, and therefore affect the measurement of Chloro-*a*, Algae and DO. The HPC content indicates the level of pollution. The canonical correlatio coefficient (that is, the square roots of the ratio of th between-group sum of squares to the total sum of square for a given CDF) exceeded 0.8 for both of the CDF (Table 3). The eigenvalue (that is, the ratio of the betweer group sum of squares to the within-group sum of square for a given CDF) also exceeded 1 for both CDFs. The tw CDFs together explained 100% (95.6% and 4.4% respectively) of the variance at the 21 sampling sites. Th values of the discriminant variables were standardized i order to determine the relationships between these var ables and the functions. The standardized CDFs wer obtained using the following formulae:

$$CDF_1 = 1.59T + 0.91Alk - 8.37Cl^- + 2.81EC - 6.80TKN - 1.38TP - 1.09UV\text{-}254 + 1.30\,pH - 3.42HPC + 2.05D\ldots + 2.25Algae + 9.15Chloro\text{-}a + 3.88SO_4^{-2} - 0.56TS;$$

$$CDF_2 = -0.40T - 1.06Alk - 7.83Cl^- + 1.05EC - 0.18TKN - 2.53TP + 1.02UV\text{-}254 + 0.53pH - 3.33HPC + 0.12DC\ldots + 0.06Algae + 6.21Chloro\text{-}a + 7.46SO_4^{-2} - 0.88TS.$$

The three river classes were distinguished using tw CDFs. CDF_1 had the highest canonical coefficient (0.9 and was defined by eight discriminant variables, whos canonical coefficients had high absolute values: Chloro-*a* Cl^-, TKN, SO_4^{-2}, HPC, EC, Algae and DO. Accordingly CDF_1 comprised the caused factors mainly the nutrien factor, the domestic wastewater factor and the ocean caused factor. The TCSCs indicated that the variances c the variables in CDF_1 were as follows (in ascending order TKN, Algae, DO, and TP (Table 3). TKN and TP ar

Table 2
Outcomes of CDA determined by forward stepwise method for the three rivers (CDA was not chosen when Wilks' Lambda<0.0014, or *P*-values >0.0004)

Discriminate variables	Wilks' Lambda	*F*-remove	*P*-level
T (°C)	0.003	2.099	0.218
Alk (mg/L)	0.002	0.714	0.534
Cl (mg/L)	0.002	1.235	0.366
EC (μs/cm)	0.003	2.822	0.151
TKN (mg/L)	0.018	30.127	0.002
TP (mg/L)	0.003	2.405	0.185
UV-254 (cm^{-1})	0.003	2.727	0.158
pH	0.002	1.527	0.304
HPC (CFU/mL)	0.003	3.209	0.127
DO (mg/L)	0.004	4.950	0.065
Algae (unit/mL)	0.007	9.679	0.019
Chloro-*a* (mg/m^3)	0.008	11.541	0.013
SO_4^{-2} (mg/L)	0.002	0.740	0.523
TS (mg/L)	0.002	0.960	0.444

Table 3
Outcomes of the total standardized canonical coefficients (TSCC) and tl total canonical structure coefficients (TCSC) between canonical discrim nant functions (CDF_1, CDF_2) and discriminant variables

Discriminant variables	CDF_1		CDF_2	
	TSCC	TCSC	TSCC	TCS
T (°C)	1.59	0.01	−0.40	**−0.3**
Alk (mg/L)	0.91	0.03	−1.06	−0.0
Cl (mg/L)	**−8.37**	−0.02	**−7.83**	0.0
EC (μs/cm)	**2.81**	0.01	1.05	0.0
TKN (mg/L)	**−6.80**	**−0.08**	−0.18	−0.0
TP (mg/L)	−1.38	**−0.03**	**−2.53**	**0.1**
UV-254 (cm^{-1})	−1.09	−0.03	1.02	0.0
pH	1.30	0.03	0.53	−0.0
HPC (CFU/mL)	**−3.42**	0.02	**−3.33**	−0.0
DO (mg/L)	**2.05**	**0.04**	0.12	**−0.2**
Algae (unit/mL)	**2.25**	**0.06**	0.06	**0.1**
Chloro-*a* (mg/m^3)	**9.15**	−0.00	**6.21**	0.0
SO_4^{-2} (mg/L)	**3.88**	−0.02	**7.46**	0.0
TS (mg/L)	−0.56	0.02	−0.88	0.0
χ^2 test	75.87		21.14	
Canonical coefficient	0.99		0.92	
Eigenvalue	115.65		5.29	
Cumulative variance	0.96		1.00	

Bold numbers are which relatively high absolute values.

nutrients that can influence the production of algae, and thus the DO content. Table 2 shows that the measured TKN values in the Tungkang River and Lingbeng River exceeded that in the Kaoping River. Visual reconnaissance in the field established that some segments of these two rivers were exposed to the effects of livestock breeding and tillage wastewater. The amount of algae was highest in the Kaoping River and the lowest in the Tungkang River. The amount of algae was correlated not only with TKN but also with other potential factors, including water temperature, sunlight intensity and duration, and algal inhibitors. Additionally, the DO was highest in the Lingbeng River and lowest in the Tungkang River. Sampling sites in the Lingbeng River were concentrated in the middle-lower reaches, where duck-breeding farms were densely located, and nutrients and algae were prevalent. These sites therefore had a relatively high DO during the day. TP is also an important factor for algae but was not prevalent in the three rivers. In summary, CDF_1 was identified as the nutrient factor based on the TCSCs. The nutrient factor was able to distinguish among the three rivers and explained 95.6% of the total variance in river water quality.

CDF_2 had a higher canonical correlation of 0.92 and was defined by five discriminant variables, which had canonical coefficients with high absolute values: Cl^-, SO_4^{-2}, Chloro-*a*, HPC and TP (Table 3). CDF_2 comprised the caused factors mainly the nutrient factor, the domestic wastewater factor and the ocean factor. The order of variance of the variables that contributed to CDF_2 was as follows: *T*, DO, Algae and TP. The intensity and duration of sunshine strongly affected the surface water temperature. And the other important factors were nutrient factors. Although the discriminant capacity of CDF_2 was only 4.4%, it was also identified as the nutrient factor. Fig. 2 presents a

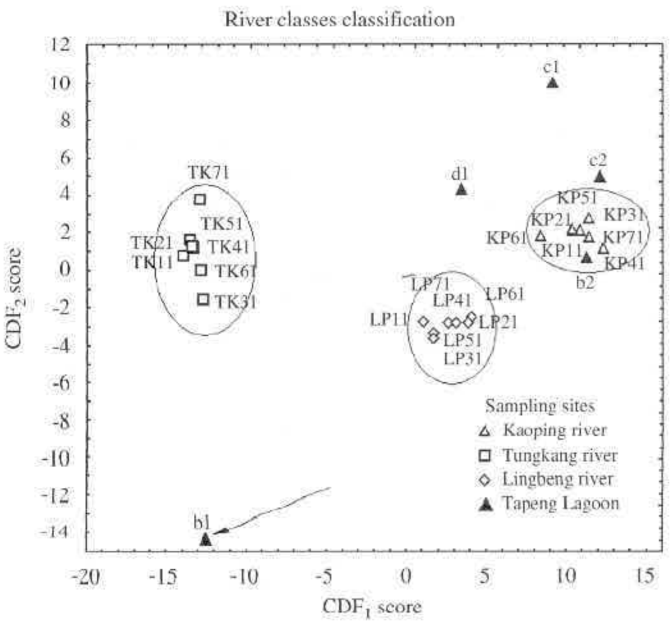


Fig. 2. Scatterplot of the three rivers neighbouring Tapeng Lagoon in these two discriminant functions.

scatterplot of the two CDFs to distinguish among the three rivers. Finally, five sampling sites that were significantly influenced by the ocean tide and located at the entrance of Tapeng Lagoon were selected: b1, c1, d1, b2 and c2. The two constructed CDFs were substituted with standardized water-quality parameters for the five sites in Tapeng Lagoon; this indicated that the quality of water at each site, except b1, was similar to that in the Kaoping River (Fig. 2). Thus, the Kaoping River most strongly affected the water quality in Tapeng Lagoon (on the basis of statistical approach). The reason caused this conclusion was that Kaoping River had the biggest catchments in these three rivers and also status of the ocean flow mechanism in these estuaries.

4. Conclusions

CDA was demonstrated as a useful tool for classifying river types in the present study, which identified physical and chemical properties of water in Tapeng Lagoon. This method was used to establish a feasible conceptual model of the relationships between water characteristics and river distributions. Moreover, it revealed that the Kaoping River water quality had the greatest effect on the Tapeng Lagoon. Furthermore, considering the two constructed linear discriminate functions, the main factor was the nutrient-affected factor. The CDA summarized the important differences among predetermined groups by reducing the dimensionality of the datasets, while accounting for the complex relationships among multiple characteristics. Such insights cannot be gained from univariate statistical techniques. Knowledge of CDA clarifies the variables to which various types of river are most sensitive. Controlling these discriminate variables will enable Tapeng Lagoon to be managed, and a biodiverse habitat to be created and supported. The approaches and results described in this report will help the government to refine the current monitoring programme by selecting the determinants of physical and chemical characteristics of river water samples, which might be applicable to other wetlands with similar properties and environmental problems.

Acknowledgements

This work was sponsored by Tajen University, Taiwan, Republic of China. The authors especially thank our colleagues at the Department of Environmental Resources Management, who helped with sampling and in the laboratory.

References

American Public Health Association/American Water Works Association/Water Environment Federation, 1995. Standard Methods for the Examination of Water and Wastewater, 19th ed. Washington DC, USA.

Comber, A., Fiaher, P., Wadsworth, R., 2005. Comparing statistical and semantic approaches for identifying change from land cover datasets. Journal of Environmental Management 77, 47–55.

Cruz-Castillo, J., Ganeshanandam, S., Mackay, B.R., Lawes, G.S., Woolley, D.J., 1994. Application of canonical discriminant analysis in horticultural research. HortScience 29, 1115–1119.

Huberty, C.J., 1994. Applied Discriminant Analysis. Wiley, New York.

Jarvie, H.P., Whitton, B.A., Neal, C., 1998. Nitrogen and phosphorus in east coast British rivers: speciation, sources and biological significance. Science of the Total Environment 210/211, 79–109.

Lambrakis, N., Antonakos, A., Panagopoulos, G., 2004. The use of multicomponent statistical analysis in hydrogeological environmental research. Water Research 38, 1862–1872.

Liao, S.W., Chang, W.L., 2005. Interpretation and discrimination of marshy wetlands by soil factors in the Kuan-Tu Natural Park, Taiwan. Environmental Monitoring and Assessment 107, 181–202.

Liao, S.W., Hwang, M.S., Lai, W.L., Shue, C.W., 2003. Spatial variation of heavy metal bioaccumulation in *Crassostrea gigas* of the Tapeng Lagoon. Tajen J. 23, 13–29 (in Chinese).

Matthew, C., Lawoko, C.R.O., Korte, C.J., Smith, D., 1994. Application of canonical discriminant analysis, principal component analysis, and canonical correlation analysis as tools for evaluation differences in pasture botanical composition. New Zealand Journal of Agricultural Research 37, 509–520.

Momen, B., Zehr, J.P., 1998. Watershed classification by discrimiant analyses of lakewater-chemistry and terrestrial characteristics. Ecological Applications 8, 497–507.

Papatheodorou, G., Demopoulou, G., Lambrakis, N., 2006. A long term study of temporal hydrochemical data in a shallow lak using multivariate statistical techniques. Ecological Modelling 19: 759–776.

Rencher, A.C., 1992. Interpretation of canonical discriminant function canonical varieties, and principal components. American Statisticia 46, 217–225.

Sharma, S., 1996. Applied Multivariate Techniques. Wiley, New York.

Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsa, D Anthemidis, A., Sofoniou, M., Kouimtzis, Th., 2003a. Assessment c the surface water quality in Northern Greece. Water Research 3' 4119–4124.

Simeonova, P., Simeonov, V., Andreev, G., 2003b. Water quality study c the Struma river basin, Bulgaria (1989–1998). Central Europea Journal of Chemistry 2, 121–136.

Singh, K.P., Malik, A., Mohan, D., Sinha, S., 2004. Multivariat statistical techniques for the evaluation of spatial and tempor: variations in water quality of Gomti River (India)—a case stud: Water Research 38, 3980–3992.

Singh, K.P., Malik, A., Sinha, S., 2005. Water quality assessment an apportionment of pollution sources of Gomti river (India) usin multivariate statistical techniques—a case study. Analytica Chimic Acta 538, 355–374.

StatSoft, 1996. STATISTICA for Windows. StatSoft, Inc, Tulsa OK.

Vega, M., Pardo, R., Barrado, E., Deban, L., 1998. Assessment c seasonal and polluting effects on the quality of river water b exploratory data analysis. Water Research 32 (12), 3581–3592.

ELSEVIER

Journal of Environmental Management 88 (2008) 293–299

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Utilization of washed MSWI fly ash as partial cement substitute with the addition of dithiocarbamic chelate

Xingbao Gao[a,*], Wei Wang[a], Tunmin Ye[a,b], Feng Wang[a], Yuxin Lan[a]

[a]Department of Environmental Science and Engineering, Tsinghua University, Beijing 100084, China
[b]School of Municipal and Environmental Engineering, Harbin Institute of Technology, Harbin 150090, China

Received 28 March 2006; received in revised form 28 January 2007; accepted 17 February 2007
Available online 26 April 2007

Abstract

The management of the big amount of fly ash as hazardous waste from the municipal solid waste incinerator (MSWI) has encountered many problems in China. In this study, a feasibility research on MSWI fly ash utilization as partial cement substitute in cement mortars was therefore carried out. MSWI fly ash was subjected to washing process to reduce its chlorine content (from 10.16% to 1.28%). Consequently, it was used in cement mortars. Ten percent and 20% replacement of cement by washed ash showed acceptable strength properties. In TCLP and 180-day monolithic tests, the mortars with washed ash presented a little stronger heavy metal leachability, but this fell to the blank level (mortar without washed ash) with the addition of 0.25% chelate. Therefore, this method is proposed as an environment-friendly technology to achieve a satisfactory solution for MSWI fly ash management.

Keywords: MSWI fly ash; Heavy metals; Leaching; Compressive strength; Dithiocarbamic chelate

1. Introduction

Incineration has become a preferred option to dispose municipal solid waste (MSW) in big cities (e.g. Shenzhen, Shanghai, Beijing, Tianjin, Guangzhou, Hangzhou) in China. Nowadays, there are about 70 MSW incinerators (MSWI) and the total capacity has reached 33010 t/d (China Statistical Yearbook, 2006). The pollutants generated in the incineration process are trapped in the air pollution control system with production of a big amount of fly ash. MSWI fly ash with concentrated heavy metals (e.g. Cd, Cr, Cu, Ni, Pb) and organochlorine compounds is considered as hazardous waste and requested to be disposed of in hazardous waste landfills after the immobilization of heavy metals. However, most cities have no hazardous waste landfill and although a few cities (e.g. Shenzhen, Shanghai, Tianjin) have such, their limited landfill capacity is not enough to contain the MSWI fly ash. In addition, MSWI fly ash is almost out of control. Hence, finding an alternative way for the disposal of MSWI fly ash has been an essential issue in these cities.

The reuse of industrial waste in basic municipal constructions has been a common practice for a very long time now. For instance, coal fly ash is reused successfully in cement, concrete, and roadbed. It has been shown that without pollutants and chlorine, MSWI fly ash is mineralogically similar to coal fly ash, and hence has pozzolanic properties (Huang and Chu, 2003). MSWI fly ash can also be added into cement mortars or concrete as substitute or inert aggregate (Collivignarelli and Sorlini, 2002). Rémond et al. (2002) had incorporated fly ash into the hydration model. The results showed that the MSWI fly ash essentially affects the hydration of the aluminate phases of the cement by particularly forming Friedel's salt. The reactions between chlorides and aluminates lead in particular to a high sulfate concentration in the interstitial solution of cement pastes, which slows down the transformation of ettringite into monosulfoaluminate. Chlorine concentrations from 5% to 20% in MSWI fly ash exceed

*Corresponding author. Tel.: +86 10 62794483; fax: +86 10 62782910.
E-mail address: gaoxingbao00@mails.tsinghua.edu.cn (X. Gao).

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.02.008

the maximum allowable concentration in most cement mixtures. In the meantime, Mangialardi (2004) found a four-stage washing process to be able to convert the raw MSWI fly ash into a material with improved chemical characteristics for its incorporation into cementitious matrices. As a result, the cementitious mixtures incorporating washed fly ash in place of raw fly ash were found to exhibit better performance characteristics in terms of setting, dimensional stability, compressive strength, and environmental quality. Specifically, it has been demonstrated MSWI fly ash that was subjected to a washing process to reduce its chloride content has about the same property on the concrete 28-day compressive strength as common coal fly ash when the 30% of the cement was replaced by ashes (Bertolini et al., 2004). Moreover, the physical properties of fresh and hardened concrete are not deteriorated (Aubert et al., 2004). Aubert et al. (2006) developed a conventional process based on the washing, phosphation and calcinations of MSWI fly ash to make the utilization of MSWI fly ash in blended cement safe and feasible, and furthermore, a modified process intended to eliminate metallic aluminum and sulfate to improve the quality of the ash.

Another essential issue regarding the reuse of MSWI fly ash is the environmental impact of the pollutants. The removal of organochlorine compounds is generally done by the melting process (Abe et al., 1996) or a low-temperature dechlorination process (Ishida et al., 1998; Wang et al., 2006). The leaching behavior of heavy metals from fly ash is mainly pH dependent due to the solubility product of hydroxide, so the heavy metal leachability would increase in the long-term as pH decreases in the condition of acid rain or organic acid solution. The most common stabilization/solidification (S/S) processes are based on the use of hydraulic binders or chemical agents. The use of cement for hazardous waste S/S disposal has an extensive documented history and well-established technology (Glasser, 1997), but the increase volume of products leads to a high landfill cost. Cement S/S has encountered some difficulties in MSWI fly ash, like the enhancement of leachability of amphoteric heavy metals (Pb and Zn), for instance. Mizutani et al. (2000) compared different chemical agents (a chelating agent, phosphate and ferrite) used for the treatment of MSWI fly ash. The authors concluded that the chelating-treated material showed a high reducing capacity and strong retention for metals over a wide pH range, while the phosphate-treated material showed a significant decrease in availability (especially for Pb), and the ferrite treated material showed an increased physical retention. Moreover, Jiang et al. (2004) investigated heavy metal stabilization in MSWI fly ash by heavy metal chelating agents. The results indicated that the heavy metals in fly ash can be stabilized more effectively by using a kind of dithiocarbamic chelate than by using chemical agents such as sodium sulfide and lime by the formation of coordinated complex. Also, the coordinated complex is stable and then the heavy metal leachability is not pH dependent.

Therefore, the aim of this study is to achieve a sa utilization of MSWI fly ash. Fly ash was subjected to washing process to reduce its chlorine content befo use, while chelate was added to reduce heavy met leachability. We focused on compressive strength proper and heavy metal leachability from mortars to evaluate tl physical properties and environmental impacts of th utilization.

2. Materials

MSWI fly ash used in this study was from the Hangzhc MSW Incineration Plant. The plant uses reciprocatir stoker incinerators and semi-dry lime adsorption and ba filters as air pollution control system. The ash samples we stored in dry and sealed containers (avoiding hydratic and carbonation). Table 1 shows the heavy metal conte of MSWI fly ash. Notably, Pb and Zn were the mo abundant species. The total amount of heavy metals wa about 2.65% of the dry weight of MSWI fly asl Meanwhile, Table 2 shows the results of the toxicit characteristic leaching procedure (TCLP, a standar method to determine waste leaching toxicity by US EPA of MSWI fly ash. The concentration of Cd and Zn inflecte the potential toxicity and inapplicability of considerin MSWI fly ash as normal waste.

The cement used as received was ordinary Portlan cement 32.5R, and the sand was the China standar sand produced by Xiamen ISO Standard Sand Co., Ltc The water used for making and curing cement mortars wa tap water, and for the leaching test was de-ionized wat (pH = 7.67). The chemical agent used to stabilize heav metals was a kind of dithiocarbamic chelate. The chelat was synthesized experimentally through the reaction b different types of polyamine or polyethyleneimin and carbon disulfide in alkaline conditions (Jiang et al 2004). The schematic structure of the chelate and the heav

Table 1
The heavy metal content in MSWI fly ash (g/kg)

Sampling date	Cd	Cr	Cu	Ni	Pb	Zn	Tot
4/16/2005	0.29	0.24	0.74	0.08	5.33	17.61	24.2
4/27/2005	0.35	0.28	0.39	0.09	5.67	19.30	26.0
5/10/2005	0.41	0.24	0.37	0.10	5.44	22.72	29.2
Average values	0.35	0.25	0.50	0.09	5.48	19.88	26.5

Table 2
TCLP test results of MSWI fly ash (mg/L)

Sampling date	Cd	Cr	Cu	Ni	Pb	Zn
4/16/2005	1.12	0.05	0.88	0.36	0.42	42.3
4/27/2005	3.05	0.27	0.92	0.23	0.56	43.8
5/10/2005	2.35	0.42	0.21	0.14	0.43	36.9

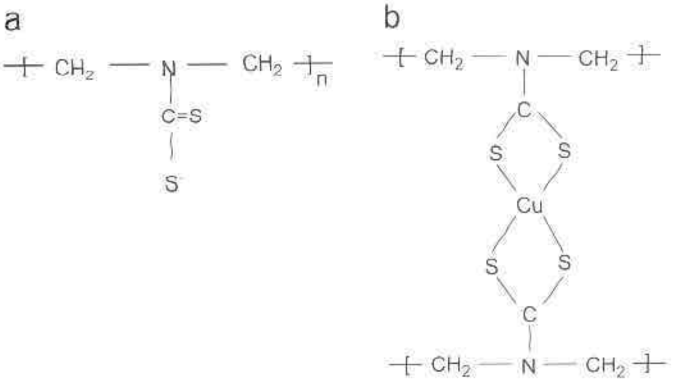


Fig. 1. Schematic structure of (a) the chelate ($n = 2$–30) and (b) the heavy metal coordinated complex (Cu as an example).

metal coordinated complex (Cu as an example) is shown in Fig. 1.

3. Experimental

3.1. Experimental procedures

For the reduction of chlorine in MSWI fly ash, water-washing process was introduced as a pretreatment method. The washing process can also reduce free CaO through hydration that has negative effect in pozzolanic reaction. In this study, we adopted an L/S ratio of 5 and a washing time of 0.5 h. After washing, the ash was dried by frame filter press to reduce the water content to about 20%. The pH of wastewater was adjusted to 6.0–10.5 (may need to add acid) to avoid heavy metal leaching.

The washed ash was then added into cement mortars partially replacing the cement at the ratio of 10%, 20%, and 30%. The cement mortars were made and tested according to the "Test method for the strength of hydraulic cement mortars (Chinese standard method, GB/T17671-1999)." In addition, to avoid the long-term leachability of heavy metals as pH drops (Poon and Chen, 1999), dithiocarbamic chelate was added into the mortars by the production of chelate complex. The recipe of the cement mortars is listed in Table 3.

The cement/sand ratio was 3, while liquid/cement ratio was 0.5. The cement, washed ash, sand and water were stirred and then were added to the matrix models. After which, they were librated on a plain bumper 120 times following GB/T17671-1999. The size of the models was 40 mm × 40 mm × 160 mm. The mortars were cured at 20 ± 2 °C in 100% humidity for 24 h, and then the models were removed, and the mortars were placed in water. Before compressive strength tests, the mortars were taken out of the water and were tested immediately. We could get six values from three tested mortars to obtain the average compressive strength (excluding the highest and the lowest). The test intervals were 3, 7, 14, 28, and 180 days.

Table 3
The recipe of the cement mortars (g)

	Cement	Sand	Washed ash	Water	Chelate
F0	450	1350		225	
F1	405	1350	45	225	
F2	360	1350	90	225	
F3	315	1350	135	225	
F21	360	1350	90	225	0.25%[a]
F22	360	1350	90	225	0.50%[a]

[a]Chelate weight/washed ash weight.

3.2. Analytical methods

3.2.1. Characterization of the ash samples

Major chemical composition analysis was carried out with X-ray fluorescence spectroscopy (XRF, Shimadzu Lab Center XRF-1700, Japan), while the concentrations of the trace metals were determined by a modified digesting ASTM method (Wan et al., 2006). Mineralogical investigation was carried out by X-ray powder diffraction (XRD, Rigaku D/max-rB, Japan). The setting conditions for the XRD were as follows: Cu Kα radiation, 40 keV accelerating voltage, 80 mA current, 15–50° 2θ scanning range, 0.02° step and 6°/min scan speed.

3.2.2. The leaching test

The toxicity of MSWI fly ash samples and crushed hardened mortars was determined by TCLP. For TCLP, the acetic solution of pH 2.88 was chosen for all the samples. The concentrations of heavy metals in leachate were determined by inductively coupled plasma-atomic emission spectrometry (ICP-AES, IRIS Intrepid II, USA).

The monolithic leaching behavior of cement mortars with MSWI fly ash is important to test its feasibility. Therefore, monolithic leaching tests were conducted to assess the leaching behavior of heavy metals from monolithic mortars (Kosson et al., 2002). After cured at 20 ± 2 °C in 100% humidity for 24 h in models, the monolithic mortars were immersed immediately in deionized water, and the container was sealed during the leaching process. The liquid/surface ratio was 10 g/m^2 and the renewal intervals were at 2 h, 8 h, 1, 2, 4, 8, 15, 30, 60 and 180 days. The tests conducted were pH electrode for the pH of the leachate, ion chromatograph for chlorine concentration, and inductively coupled plasma mass spectrometry (ICP-MS, Perkin-Elmer Elan6000, USA) for metal concentration.

4. Results

4.1. The washing process

4.1.1. Change in major chemical composition and mineralogical species during washing

Table 4 shows the major chemical composition of MSWI fly ash and washed ash. MSWI fly ash showed a high

Table 4
The major components of the original MSWI fly ash and washed fly ash (%)

Compound	Al_2O_3	SiO_2	CaO	Fe_2O_3	Na_2O	K_2O	MgO	Cl	F	P_2O_5	SO_3	SnO_2	TiO_2	ZnO	Others
Original ash	6.97	19.81	23.63	4.00	6.68	6.23	3.78	10.16	1.83	2.54	8.74	1.02	1.21	2.79	0.61
Washed ash	9.82	28.46	26.78	5.33	1.56	2.09	5.35	1.28	1.79	3.59	6.63	1.28	1.60	3.65	0.79

quantity of CaO, Cl and SO_3, which would have a negative effect on the mechanical strength of the cement mortars. The washed ash had less chlorine, but the CaO and SO_3 were still very high. According to the chemical composition and physical requirements specified in ASTM C618, the washed ash was similar to Class C coal fly ash. The Portland cement mixtures are rich in washed ash (the total of SiO_2, Al_2O_3 and Fe_2O_3 was about 43.61%). Therefore, the washed ash added into the cement mortars to replace cement partially may have little effect on mechanical strength.

Fig. 2 shows the XRD analysis of original fly ash and washed ash. The NaCl, KCl and CaClOH disappeared as soluble salts, while $Ca(OH)_2$ appeared for the hydration of free CaO in washed ash. $CaCO_3$ increased for the carbonation during the washing process. The formation of new mineral species such as gehlenite ($Ca_2Al_2SiO_7$) and plagioclase ($CaAl_2SiO_8 4H_2O$), which are common in cement, indicates that washed ash is more suitable for reuse in cement.

4.1.2. *Characterization of wastewater*

The pH value of the wastewater was 8.8 ± 0.3, and the concentrations of the heavy metals in wastewater are listed in Table 5. These concentrations of heavy metals were below level II of the "Integrated Wastewater Discharge Standard of China (GB8978-1996)," so the wastewater could be treated in municipal wastewater treatment plant. The transfer ratios of soluble salts from fly ash to water during the washing process are listed in Table 6. Most chloride and some sulfate move to the wastewater. The Na, K, and Ca were almost in wastewater, and these could be used to recycle sodium chloride and potassium chloride by chemical methods.

4.2. *Compressive strength*

Fig. 3 shows the unconfined compressive strength of the cement mortars at 3, 7, 14, 28 and 180 days. The 30% replacement of cement by washed ash in mortars showed deterioration in strength evolution (compressive strength of about 60% of the reference mortars at 28 and 180 curing days), while the 10% and 20% replacement showed much better strength properties. However, the 28-day compressive strength of 30% replacement was 23.4 MPa, which could also be used in some particular conditions. From the results, we set the limit of MSWI fly ash replacement of cement at 20% to ensure the property of the mortars.

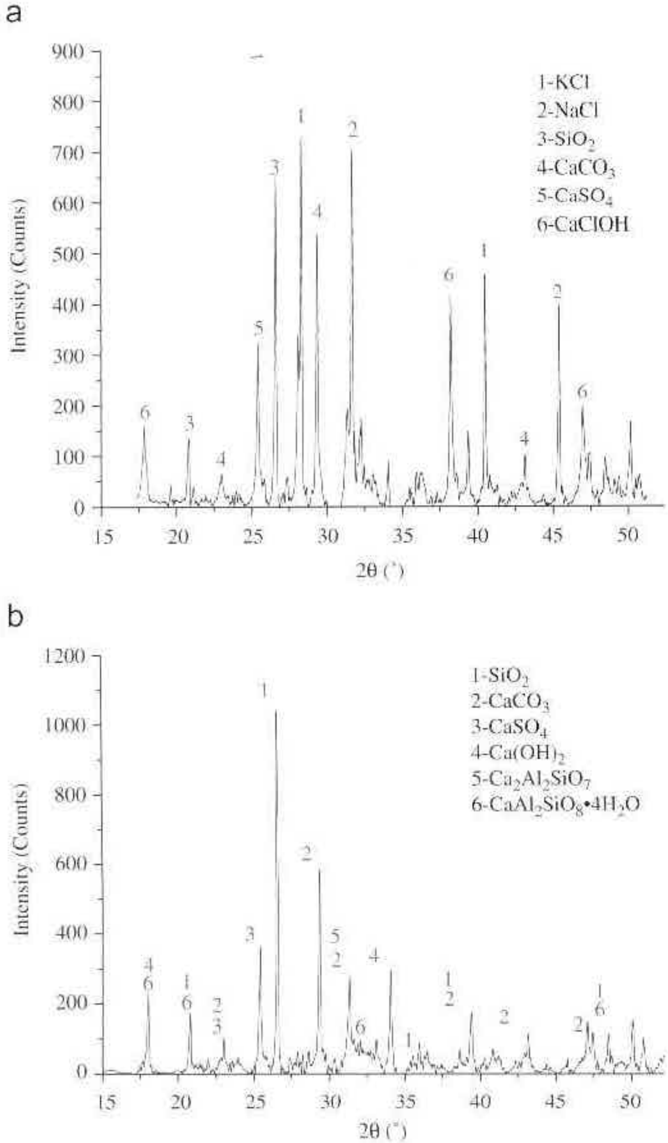


Fig. 2. XRD analysis of (a) original fly ash and (b) washed fly ash.

Table 5
The concentration of heavy metals in the washed water

	Cd	Cr	Cu	Ni	Pb	Zn
Values (mg/L)	0.03	0.19	0.55	0.09	0.65	2.5
Limitation[a]	0.1	1.5	1.0	1.0	1.0	5.0

[a]Level II of "Integrated Wastewater Discharge Standard of Chin (GB8978-1996)".

Table 6
The transfer ratio of soluble salts from fly ash to water during the washing process (L/S = 5.0, $T = 0.5$ h)

	HCO_3^-	SO_4^{2-}	Cl	Ca	Na	K	Loss in mass
Values (%)[a]	0.22	5.24	9.44	3.52	4.03	3.90	30.57[b]
Recovery rate (%)[b]	NA[c]	102.8	101.7	99.7	97.6	98.8	

[a]The ratio is based on fly ash weight.
[b]Calculated value supposing that Si does not move to water.
[c]Not available.

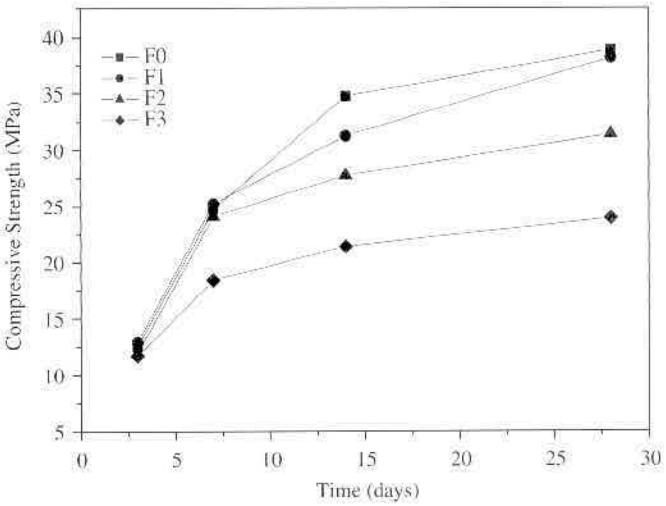


Fig. 3. The unconfined compressive strength as a function of curing time of reference mortars (F0) and mortars with washed MSWI ashes (F1, F2, and F3).

In fact, although the mechanical characteristics were worsened by MSWI fly ash addition, the 180-day compressive strength of mortars with 20% replacement of 42.4 MPa (about 71% of the reference mortars) was enough for use in basic municipal constructions.

In this study, the mortars with chelate were also tested to determine the influence of chelate on the strength evolution at 28 and 180 days. The results showed that chelate had little influence on the strength evolution (the compressive strength of mortars with chelate was about 95.4–98.7% of the reference).

4.3. Leaching test results

4.3.1. TCLP tests for the crushed samples

TCLP tests were carried out for the crushed hardened mortars at 7 and 180 days. Table 7 shows the TCLP tests results for the crushed hardened mortars. The concentration of heavy metals in leachate of F2 (without chelate) was much higher than that of the reference mortars F0 (without MSWI fly ash), but in F21 and F22, the concentration of the heavy metals fell due to the addition of chelate. Generally, the concentration was lower in F22 than F21 and with the curing time growth, the leachability dropped.

Table 7
TCLP tests results of the crushed hardened mortars (mg/L)

Curing days	Mortars	Cd	Cr	Cu	Ni	Pb	Zn
7 days	F0	N.D.	0.09	1.14	0.03	0.06	2.87
	F2	N.D.	0.13	4.68	0.29	0.19	15.39
	F21	N.D.	0.09	1.24	0.09	0.06	5.75
	F22	N.D.	N.D.	0.88	0.12	0.08	4.98
180 days	F0	N.D.	0.11	1.01	N.D.	N.D.	0.57
	F2	N.D.	N.D.	2.57	0.13	0.10	7.36
	F21	N.D.	0.07	0.07	0.06	0.06	1.19
	F22	N.D.	N.D.	0.07	N.D.	N.D.	0.65

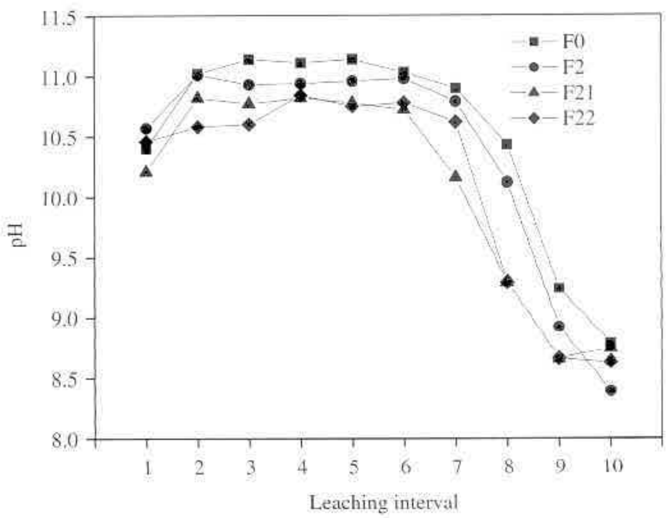


Fig. 4. The change of pH of the leachates in monolithic leaching tests.

4.3.2. Monolithic leaching tests

The change of pH of the leachates is shown in Fig. 4. At the beginning, the pH was about 11 for a long time due to the release of lime from fly ash and cement. After the 10th leachate renewal, the pH fell quickly to 8.5. In the pH range, many heavy metals could not be released because the solubility product constant of heavy metal hydrate was very small. Hence, the cement mortars have the ability to immobilize the heavy metals.

The accumulative release of Cl, Ca, Na, and K in the leachates is shown in Fig. 5. Chlorine is one of the steel corrosion substances (Kayali and Zhu, 2005). The accumulative release of chlorine from mortars with washed

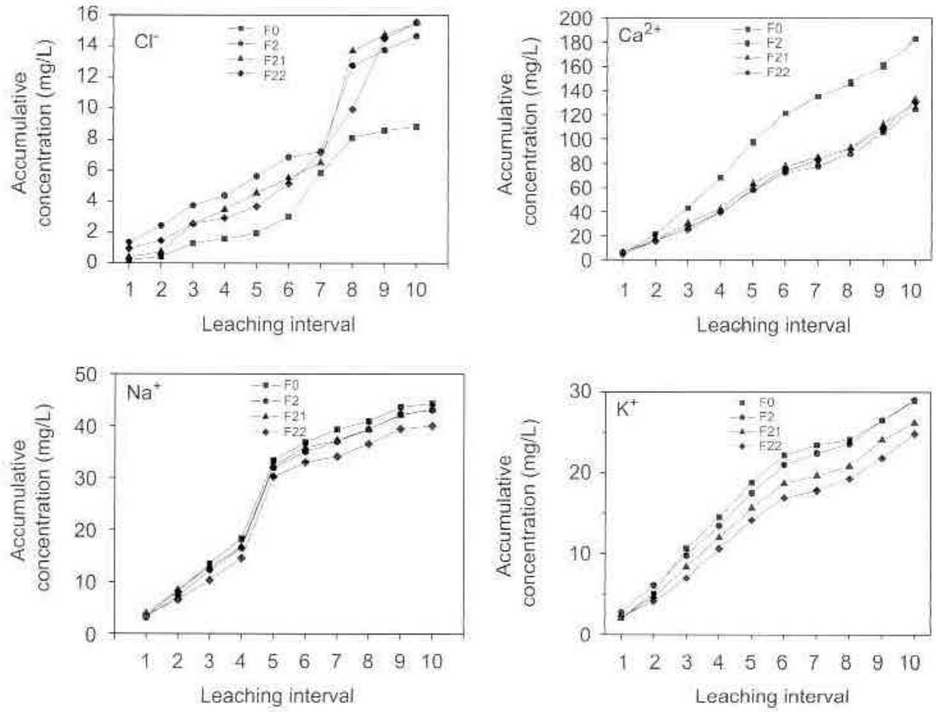


Fig. 5. Accumulative release of Cl, Ca, Na, and K as a function of leaching intervals in monolithic leaching tests.

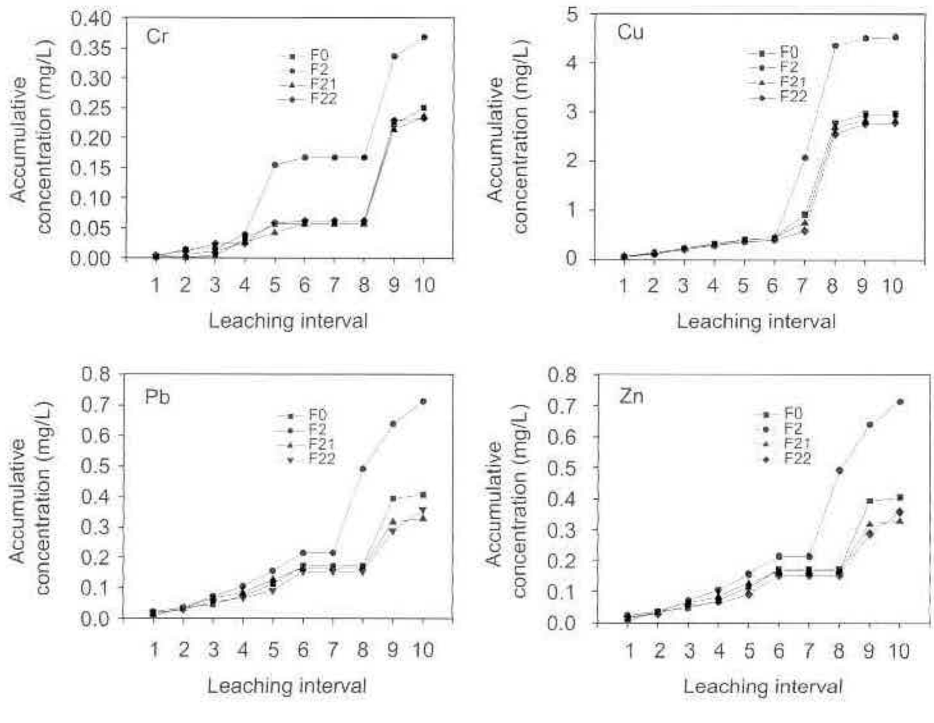


Fig. 6. Accumulative release of heavy metals (Cr, Cu, Pb, and Zn) as a function of leaching intervals in monolithic leaching tests.

MSWI fly ash (F2, F21, and F22) was much higher than that F0. This was caused by the residual chlorine content of about 1.28% in the washed ash. In this case, the chlorine enhanced the steel corrosion speed. To avoid this situation, a multistep washing process which can decrease th chlorine content to 0.1% (UK standard BS EN450 fo coal fly ash use in concrete) shall be adopted in furthe study and practice (Mangialardi, 2004). The accumulativ release of Ca from F2, F21 and F22 was less than the blank mortar (F0) because the soluble Ca in MSWI fly ash was washed out during the washing process. Ca release caused the pH value of F0 to be higher than other mortars in the monolithic leaching tests. The accumulative release of Na and K is similar in all the tested mortars.

The accumulative release of heavy metals from the mortars versus leachate renewal intervals is shown in Fig. 6. It can be seen that the accumulative heavy metal release from F2 was much higher than the reference F0, but the mortars with 0.25% or 0.5% chelate appeared like the reference F0. The leaching behavior of heavy metals from F2 showed that cement could immobilize the heavy metals in the short-term, but in the long-term, the leachability still increased. The 180-day leaching behavior showed that chelate could stabilize the heavy metals in cement mortars in the short- and long-term.

5. Conclusions

MSWI fly ash with a high content of chlorine was found to be unsuitable for utilization in cement mortars. After washing process, the chlorine content reduced to 1.28%, while the washed ash was similar to Class C coal fly ash in chemical composition and could be added into cement mortars as substitute. During the washing process, little heavy metals leached to the water with a pH value of 8.8 ± 0.3 while the soluble salts (chloride and sulfate of Na, K and Ca) mainly move with the water. The formation of new mineral species showed good properties of washed ash.

Ten percent and 20% replacements showed acceptable strength properties and therefore 20% was set as the replacement limit. The long-term compressive strength (180-day) indicated that this method is feasible in basic municipal constructions.

In TCLP and 180-day monolithic leaching tests, the mortars with washed ash presented a little stronger heavy metal leachability, but the leachability fell to the blank level with the addition of chelate. In monolithic leaching tests, the chlorine leachability increased in long-term, thus causing steel corrosion. The strict limit for chlorine content in washed ash at 0.1% was noted, and the washing process needed to be improved.

As shown in this study, the partial replacement of cement by washed MSWI ash in cement mortars with the addition of chelate could be a promising environment-friendly technology in MSWI fly ash management.

Acknowledgments

This research was partially funded by the National Science Foundation of China (NSFC) under Grant no. NSFC 20277024 and the National High Technology Research and Development Program of China (863 Program) under Grant no. 2002AA644010; the authors gratefully acknowledge these supports. The authors would like to appreciate the editors and the reviewers for their comments and suggestions in improving the manuscript. As well, the authors thank Ms. Huaxia Chen for her help in editorial revision.

References

Abe, S.-I., Kambayashi, F., Okada, M., 1996. Ash melting treatment by rotating type surface melting furnace. Waste Management 16, 431–443.

Aubert, J.E., Husson, B., Vaquier, A., 2004. Use of municipal solid waste incineration fly ash in concrete. Cement and Concrete Research 34, 957–963.

Aubert, J.E., Husson, B., Sarramone, N., 2006. Utilization of municipal solid waste incineration (MSWI) fly ash in blended cement Part 1: processing and characterization of MSWI fly ash. Journal of Hazardous Materials 136, 624–631.

Bertolini, L., Carsana, M., Cassago, D., 2004. MSWI ashes as mineral additions in concrete. Cement and Concrete Research 34, 1899–1906.

Collivignarelli, C., Sorlini, S., 2002. Reuse of municipal solid wastes incineration fly ashes in concrete mixtures. Waste Management 22, 909–912.

Glasser, F.P., 1997. Fundamental aspects of cement solidification and stabilization. Journal of Hazardous Materials 52, 151–170.

Huang, W.-J., Chu, S.-C., 2003. A study on the cementlike properties of municipal waste incineration ashes. Cement and Concrete Research 33, 1795–1799.

Ishida, M., Shiji, R., Nie, P., Nakamura, N., Sakai, S.-i., 1998. Full-scale plant study on low temperature thermal dechlorination of PCDDs/PCDFs in fly ash. Chemosphere 37, 2299–2308.

Jiang, J., Wang, J., Xu, X., Wang, W., Deng, Z., Zhang, Y., 2004. Heavy metal stabilization in municipal solid waste incineration fly ash using heavy metal chelating agents. Journal of Hazardous Materials 113, 141–146.

Kayali, O., Zhu, B., 2005. Chloride induced reinforcement corrosion in lightweight aggregate high-strength fly ash concrete. Construction and Building Materials 19, 327–336.

Kosson, D.S., van der Sloot, H.A., Sanchez, F., Garrabrants, A.C., 2002. An integrated framework for evaluating leaching in waste management and utilization of secondary materials. Environmental Engineering and Science 19, 159–204.

Mangialardi, T., 2004. Effects of a washing pre-treatment of municipal solid waste incineration fly ash on the hydration behaviour and properties of ash-Portland cement mixtures. Advances in Cement Research 16, 45–54.

Mizutani, S., van der Sloot, H.A., Sakai, S.-I., 2000. Evaluation of treatment of gas cleaning residues from MSWI with chemical agents. Waste Management 20, 233–240.

Poon, C.S., Chen, Z.Q., 1999. Comparison of the characteristics of flow-through and flow-around leaching tests of solidified heavy metal wastes. Chemosphere 38, 663–680.

Rémond, S., Bentz, D.P., Pimienta, P., 2002. Effects of the incorporation of municipal solid waste incineration fly ash in cement pastes and mortars II: modeling. Cement and Concrete Research 32, 565–576.

Wan, X., Wang, W., Ye, T., Guo, Y., Gao, X., 2006. A study on the chemical and mineralogical characterization of MSWI fly ash using a sequential extraction procedure. Journal of Hazardous Materials 134, 197–201.

Wang, W., Gao, X., Zheng, L., Lan, Y., 2006. Reductive dechlorination of polychlorinated dibenzo-*p*-dioxins and dibenzofurans in MSWI fly ash by sodium hypophosphite. Separation and Purification Technology 52, 186–190.

ELSEVIER

Journal of Environmental Management 88 (2008) 300–306

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Photocatalytic degradation of disperse blue 1 using UV/TiO_2/H_2O_2 process

M. Saquib[a,*], M. Abu Tariq[a], M.M. Haque[a], M. Muneer[b,1]

[a]*Department of Chemistry, Aligarh Muslim University, Aligarh 202002, India*

[b]*Centre for Material Science, Faculty of Science, University of Central Lancashire, Preston PR1 2HE, UK*

Received 7 April 2006; received in revised form 6 February 2007; accepted 10 March 2007

Available online 8 May 2007

Abstract

The photocatalytic degradation of a dye derivative, C.I. disperse blue 1 (**1**), has been investigated under UV light irradiation in the presence of TiO_2 and H_2O_2 under a variety of conditions. The degradation was studied by monitoring the change in substrate concentration employing UV spectroscopic technique as a function of irradiation time. The degradation was studied under different conditions such as different types of TiO_2, reaction pH, catalyst and substrate concentration containing hydrogen peroxide (H_2O_2) besides molecular oxygen in the presence of TiO_2. The degradation of dye was also investigated under sunlight and the efficiency of degradation was compared with that of the artificial light source. The degradation rates were found to be strongly influenced by all the above parameters. The photocatalyst Degussa P25 was found to be more efficient for the degradation of the dye.

Keywords: Photocatalysis; Textile dye; Disperse blue 1; Titanium dioxide; Semiconductor

1. Introduction

Treatment of colored wastewaters from textile or dye industry is a serious problem that attracts the attention of many researches during last decades. A substantial amount of dyestuff is lost during the dyeing process in the textile industry, which poses a major problem for the industry as well as a threat to the environment (Zollinger, 1987). Decolorization of dye effluents has therefore acquired increasing attention. During the past two decades, photocatalytic process involving semiconductor particles under UV light illumination has been shown to be potentially advantageous and useful in the treatment of wastewater pollutants. Earlier studies (Blake, 2001) have shown that a wide range of organic substrates can be completely photomineralized in the presence of TiO_2 and oxygen. The mechanism constituting heterogeneous photocatalytic oxidation processes has been discussed extensively in the literature (Turchi and Ollis, 1990; Mathews and McEvoy, 1992). Briefly, when a semiconductor such as TiO_2 absorbs a photon of energy equal to or greater than its band gap width, an electron may be promoted from the valence band to the conduction band (e^-_{cb}) leaving behind an electron vacancy or "hole" in the valence band (h^+_{vb}). If charge separation is maintained, the electron and hole may migrate to the catalyst surface where they participate in redox reactions with sorbed species. Specially, h^+_{vb} may react with surface-bound H_2O or OH^- to produce the hydroxyl radical and e^-_{cb} is picked up by oxygen to generate superoxide radical anion ($O_2^{\cdot-}$), as indicated in the following equations:

$$TiO_2 + h\upsilon \rightarrow e^-{}_{cb} + h^+{}_{vb}, \tag{1}$$

$$O_2 + e^-{}_{cb} \rightarrow O_2{}^{\cdot-}, \tag{2}$$

$$H_2O + h^+{}_{vb} \rightarrow OH^{\cdot} + H^+. \tag{3}$$

It has been suggested that the hydroxyl radicals ($OH^{\cdot}$) and superoxide radical anions ($O_2^{\cdot-}$) are the primary oxidizing species in the photocatalytic oxidation processes. These oxidative reactions would result in the bleaching of

*Corresponding author. Tel.: +91 571 2703515; fax: +91 571 2702758.
E-mail addresses: saquibazmi2002@gmail.com (M. Saquib), mmuneer@uclan.ac.uk, readermuneer@yahoo.co.in (M. Muneer).

[1]On leave from AMU, Aligarh, India.

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.03.012

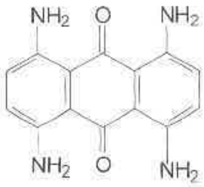
the dye. Alternatively, direct absorption of light by the dye can lead to charge injection from the excited state of the dye to the conduction band of the semiconductor as summarized in the following equations:

$$Dye_{ads} + h\nu \rightarrow Dye_{ads}^{*}, \quad (4)$$

$$Dye_{ads}^{*} + TiO_2 \rightarrow Dye_{ads}^{+} + TiO_2\,(e^-). \quad (5)$$

It has been shown earlier that the heterogeneous photocatalytic oxidation processes can be used for removing the pollutants using $UV/TiO_2/H_2O_2$ and $Solar/TiO_2/H_2O_2$ systems (Bozzi et al., 2004; Sreedhar Reddy and Kotaiah, 2005).

The dye derivative, disperse blue 1 (**1**), has been extensively used in textile industry, leather dyeing, paper printing, photography and as a biological stain. There has not been any report on the photocatalytic degradation of the dye derivative **1**. Therefore we have undertaken a detailed study on the photodegradation of **1** sensitized by TiO_2 containing H_2O_2 in aqueous solution under a variety of conditions, with an aim to determine the optimal degradation condition, which is essential for any application process.



1
disperse blue 1

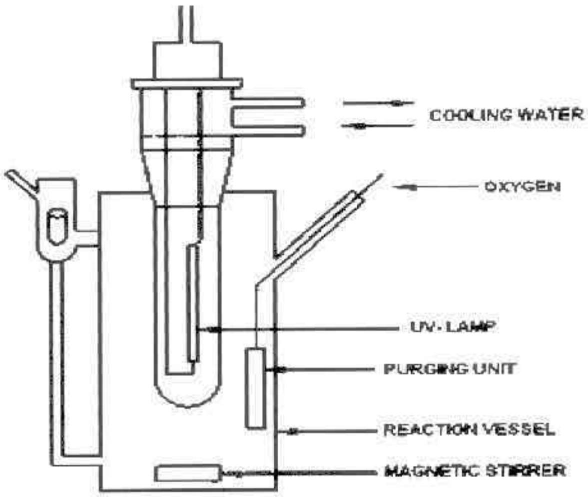
2. Experimental methods

2.1. Reagent and chemicals

Disperse blue 1 (**1**) was obtained from ACROS ORGANICS, USA, and used as such without any further purification. The water employed in all the studies was double distilled. The photocatalyst P25 (Degussa AG) was used for the degradation of dye derivative in most of the experiment. The photocatalyst Hombikat UV 100 and PC500 (Millenium inorganic chemicals) was used for comparative studies. Degussa P25 consists of 75% anatase and 25% rutile with a specific BET-surface area of $50\,m^2\,g^{-1}$ and primary particle size of 20 nm (Bickley et al., 1992). Hombikat UV100 consist of 100% anatase with a specific BET-surface area $>250\,m^2\,g^{-1}$ and primary particle size of 5 nm (Lindner et al., 1997). The photocatalyst PC500 has a BET-surface area of $287\,m^2\,g^{-1}$ with 100% anatase and primary particle size of 5–10 nm (Rauer, 1998). The other chemicals used in this study such as NaOH, HNO_3 and H_2O_2 were obtained from Merck.

2.2. Procedure

Stock solutions of the dye of desired concentration (0.25–1.0 mM) were prepared in double distilled water. An immersion well photochemical reactor made of Pyrex glass equipped with a magnetic stirring bar, water circulating jacket and an opening for supply of air bubbling was used. A schematic diagram of the reactor is shown below.



For irradiation experiment, 250 (mL) solution of the dye of desired concentration was taken into the photoreactor and required amount of photocatalyst was added and the solution was stirred and bubbled with air for at least 15 min in the dark to allow equilibration of the system so that the loss of compound due to adsorption can be taken into account. The pH of the reaction mixture was adjusted by adding a dilute aqueous solution of HNO_3 or NaOH. The zero time reading was obtained from blank solution kept in the dark but otherwise treated similarly to the irradiated solution. The suspensions were continuously purged with air bubbling throughout each experiment. Irradiations were carried out using a 125 W medium pressure mercury lamp (Philips, radiant flux: $1250\,\mu W\,cm^{-2}$ $\lambda_{max} > 254$ nm). Samples (6 mL) were collected before and at regular time intervals during the irradiation. They were centrifuged before analysis.

The sunlight experiments were carried out in order to compare the degradation efficiency of the compound under investigation with that of artificial light. Reactions were carried out in the same photochemical reaction vessel where irradiation was carried out with the artificial light source. Aqueous solution (250 mL) of desired concentration of the model compound containing required amount of photocatalyst was taken and stirred for 10 min in the dark and the solution was then placed on flat platform under sunlight (radiant flux: $450\,\mu W\,cm^{-2}$) with continuous stirring and purging of air. Samples (6 mL) were collected before and at regular time intervals during the illumination and analyzed after centrifugation.

2.3. Analysis

The photocatalytic degradation of the dye derivative was monitored using UV spectroscopic analysis technique (Shimadzu UV–Vis 1601). The double beam spectrophotometer has an in-built tungsten and deuterium lamps, which provide the measurement of optical density (OD) in the range 200–1100 nm (near UV and visible regions). The samples were analyzed using quartz cuvette, as it has zero absorption in the above wavelength regions. The degradation was monitored by measuring the absorbance on a Shimadzu UV–Vis Spectrophotometer (Model 1601). The absorbance of dye (1, 0.25 mM) was followed at λ_{max} 598 nm.

The concentrations of dye derivatives were calculated by standard calibration curve obtained from the absorbance of the dye derivative at different concentrations. The change in absorbance of the dye derivative **1** was followed at their λ_{max} as a function of irradiation time.

3. Results and discussion

3.1. Photocatalysis of aqueous suspensions of dye derivative (1) in the presence of TiO_2

Irradiation of an aqueous suspension of the dye derivatives 1 in the presence of TiO_2 containing H_2O_2 with a Pyrex filtered output of a 125 W medium pressure mercury lamp with constant bubbling of air leads to decrease in absorption intensity as a function of irradiation time. This absorption intensity is used to calculate the concentration using the standard calibration. The change in absorption spectra for the photocatalytic degradation of the dye (absorbance vs wavelength) is shown in Fig. 1.

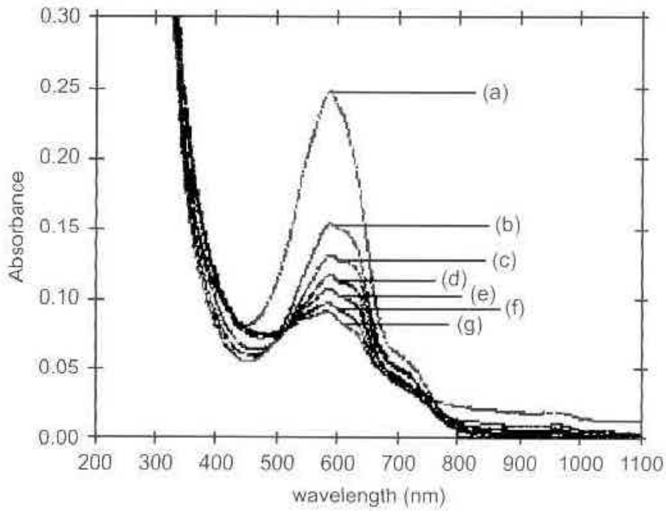


Fig. 1. Change in absorption intensity as a function of time on irradiation of disperse blue 1 (0.25 mM) containing H_2O_2 in the presence of TiO_2 (Degussa P25 1 g l^{-1}) under oxygen. Light source: Pyrex filtered output of 125 W medium pressure mercury lamp. Irradiation time: (a) 0 min, (b) 10 min, (c) 20 min, (d) 30 min, (e) 40 min, (f) 50 min and (g) 60 min.

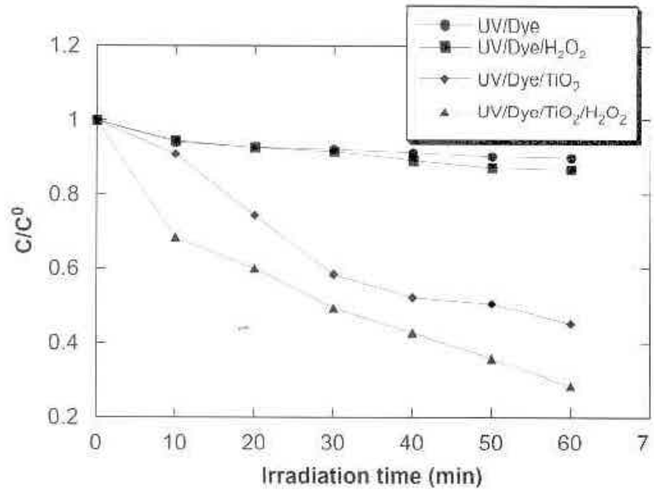


Fig. 2. Comparison of change in concentration as a function of time f irradiation of an aqueous suspension of disperse blue 1 (1) in the presen and absence of TiO_2 containing hydrogen peroxide under UV light sourc Experimental conditions: dye conc.: (0.25 mM), photocatalyst: Deguss P25, 1 g L^{-1}, H_2O_2 (0.1 mL), $V = 250$ mL, immersion well photoreacto 125 W medium pressure mercury lamp, continuous air purging an stirring, irradiation time = 60 min.

Fig. 2 shows the change in concentration as a function c time on irradiation of an aqueous suspension of dy derivative **1** containing TiO_2 in the presence and absence c H_2O_2. It could be seen from the figure that 72% degradation of the compound takes place in the presenc of H_2O_2 and 55% in the absence of hydrogen peroxide whereas in the absence of H_2O_2 the degradation rate wa found slower hence all the experiments were carried out i the presence of hydrogen peroxide.

The curve for the change in substrate concentration as function of irradiation time for the degradation of the dy derivative **1** can be fitted reasonably well by exponentia decay curve suggesting the first order kinetics. For eacl experiment, the degradation rate constant of the dy derivative was calculated from the plot of the natura logarithm of the concentration of the dye derivative as function of irradiation time. The degradation rate for th decomposition of the dye derivative was calculated usin formula given below:

$$-d[C]/dt = kc, \quad (6$$

where C is the concentration of the pollutant and k is th rate constant.

The degradation rate for the decomposition of the dy derivative for the first order reaction was calculated i terms of min^{-1}.

Control experiments were carried out by irradiatin aqueous solution of the dye in the absence of photo catalyst, where no observable loss of the dye derivative wa observed. The zero irradiation time reading was obtaine from blank solutions kept in the dark, but otherwis treated similarly to the irradiated solutions.

3.2. Comparison of different photocatalysts

Titanium dioxide is known to be the semiconductor with the highest photocatalytic activity, non-toxic, relatively inexpensive and stable in aqueous solution. Several reviews have been written, regarding the mechanistic and kinetic details as well as the influence of experimental parameters. It has been demonstrated that degradation by photocatalysis can be more efficient than by other wet-oxidation technique (Weichgrebe and Vogelpohl, 1995). The aim of present study was to determine the best photocatalyst among commercially available different TiO_2 material and find further means to accelerate the efficiency of the photocatalytic process.

Therefore, we have tested the photocatalytic activity of three different commercially available TiO_2 powders (namely Degussa P25, Hombikat UV100 and Millennium Inorganic PC500) on the degradation kinetic of the pollutant under investigation. The degradation rate obtained for the decomposition (decrease in absorption intensity) of **1** in the presence of different types of TiO_2 powders is shown in Fig. 3. It has been observed that the degradation of the pollutant proceeds much more rapidly in the presence of Degussa P25 as compared with other TiO_2 samples under UV light source.

The differences in the photocatalytic activity are likely to be due to differences in the BET-surface, impurities, lattice mismatches or density of hydroxyl groups on the catalyst's surface, since they will affect the adsorption behavior of a pollutant or intermediate molecule and the lifetime and recombination rate of electron–hole pairs. Earlier studies have also shown that Degussa P25 was found to show better activity for the photocatalytic degradation of a large number of organic compounds (Muneer et al., 1999, 2001). Also Lindner et al. (1995) showed that Hombikat UV100 was almost four times more effective than P25 when dichloroacetic acid was used as the model pollutant. Another explanation for the greater photoeffectiveness of mixed phase titania photocatalyst (here Degussa 25) could be due to three factors: (1) the smaller band gap of rutile extends the useful range of photoactivity into the visible region; (2) the stabilization of charge separation by electron transfer from rutile to anatase slows the recombination; (3) the small size of the rutile crystallites facilitates the electron transfer (Hurum et al., 2003).

3.3. Effect of pH

An important parameter in the photocatalytic reactions taking place on the particulate surfaces is the pH of the solution, since it dictates the surface charge properties of the photocatalyst and size of aggregates it forms. Employing Degussa P25 as photocatalyst for the decomposition of model compound in aqueous suspensions of TiO_2 was studied in the pH range between 3 and 11. Fig. 4 shows the degradation rate for the decomposition of the dye as a function of reaction pH.

The interpretation of pH effect on the photocatalytic process is very difficult because of its multiple roles such as electrostatic interactions between the semiconductor surface, solvent molecules, substrate and charged radicals formed during the reaction process. The ionization state of the surface of the photocatalyst can be protonated and deprotonated under acidic and alkaline conditions, respectively, as shown in the following equations:

$$TiOH + H^+ \rightarrow TiOH_2^+, \tag{7}$$

$$TiOH + OH^- \rightarrow TiO^- + H_2O. \tag{8}$$

The point of zero charge (pzc) of the TiO_2 (Degussa P25) is widely reported at pH ~6.25 (Augustynski, 1988). Thus, the TiO_2 surface will remain positively charged in acidic medium (pH < 6.25) and negatively charged in alkaline

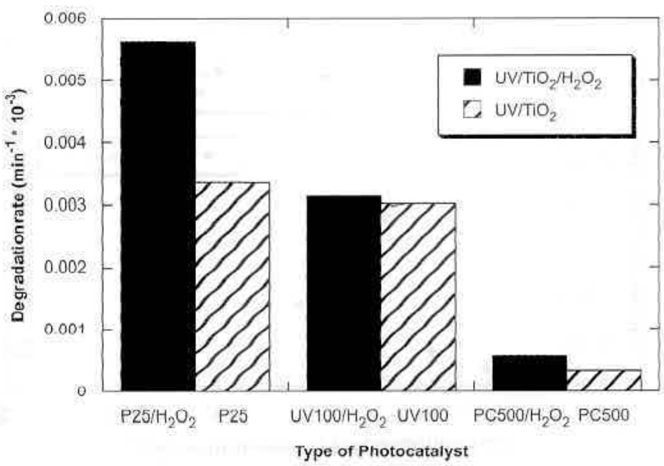


Fig. 3. Comparison of degradation rate for the decomposition of disperse blue 1 (**1**) in the presence of different types of photocatalyst in the presence and absence of hydrogen peroxide. Experimental conditions: dye conc.: 0.25 mM, H_2O_2 (0.1 mL), $V = 250$ mL, photocatalyst: TiO_2 (Degussa P25, Hombikat UV 100 and PC 500 (1 g L^{-1}), irradiation time = 60 min.

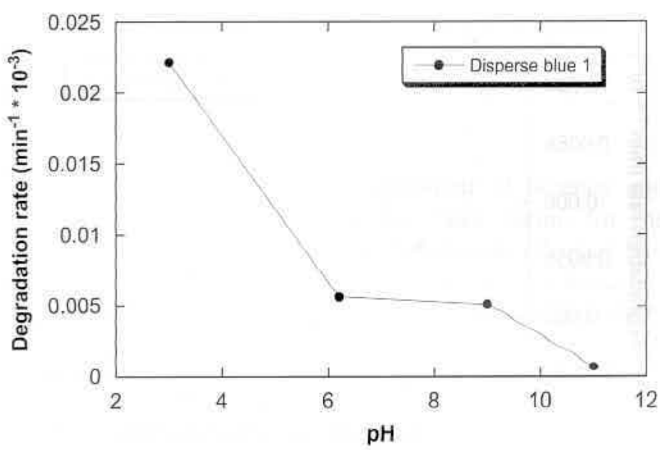


Fig. 4. Influence of pH on the degradation rate for the decomposition of disperse blue 1 (**1**) in the presence of TiO_2 containing hydrogen peroxide. Experimental conditions: dye conc. 0.25 mM, H_2O_2 (0.1 mL), $V = 250$ mL, photocatalyst TiO_2 (Degussa P25, 1 g L^{-1}), reaction pH (3, 6.2, 9 and 11), irradiation time = 60 min.

medium (pH > 6.25). The degradation rate for the dye derivative 1 was found to decrease with increase in pH from 3 to 11 and highest efficiency was obtained at pH 3. In this study it has been shown that the degradation rate for the decomposition of the dye derivatives under investigation is highly influenced by the reaction pH. This may be due to the fact that the structural orientation under this condition is favored for the attack of reactive species.

The adsorption of the dye derivative 1 on the surface of photocatalyst was investigated by stirring the aqueous solution of the dye in the dark for 24 h at different pH values. Analysis of the samples after centrifugation indicates no observable loss of compound.

3.4. Effect of substrate concentration

It is important both from mechanistic and from application point of view to study the dependence of initial substrate concentration on the degradation kinetics of the pollutant. Effect of substrate concentration on the degradation of the dye derivative 1 was studied at different concentrations varying from 0.125 to 1 mM. The degradation rate for the decomposition of dye derivative 1 as a function of substrate concentration is shown in Fig. 5.

It is interesting to note that the degradation rate for the decomposition of dye derivative 1 decreases with the increase in substrate concentration from 0.125 to 0.5 mM. A further increase in substrate concentration leads to increase in the degradation rate.

The decrease in degradation rate with increase in substrate concentration may be due to the fact that as the initial concentrations of the dye increases, the color of the irradiating mixture becomes more and more intense which prevents the penetration of light to the surface of the catalyst. Hence, the generation of relative amount of $OH^{\cdot}$ and $O_2^{\cdot-}$ on the surface of the catalyst do not increase as the intensity of light, irradiation time ar catalyst concentration are constant. Conversely, the concentrations will decrease with increase in concentr tion of the dye as the light photons are largely absorbe and prevented from reaching the catalyst surface by tl dye molecules. Consequently, the degradation efi ciency of the dye decreases as the dye concentratic increases.

3.5. Effect of catalyst concentration

Whether in static, slurry, or dynamic flow reactors, tl initial reaction rates were found to be directly proportion to catalyst concentration, indicating a heterogeneou regime. However, in some cases it was observed that abov a certain concentration, the reaction rate even decrease and becomes independent of the catalyst concentratio This limit depends on the geometry and working cond tions of the photoreactor and for a definite amount of TiC in which all the particles, i.e., the entire surface expose are totally illuminated. When the catalyst concentration very high, after traveling a certain distance on an optic path, turbidity impedes further penetration of light in th reactor. In any given application, this optimum cataly concentration $[(TiO_2)_{OPT}]$ has to be found, in order t avoid excess catalyst and ensure total absorption c efficient photons.

The effect of photocatalyst concentration on th degradation kinetics of the model compound was studie employing different concentrations of Degussa P25 varyin from 0.5 to 4 g L^{-1}.

The effect of varying concentrations of Degussa P2 on the degradation rate for the decomposition of 1 shown in Fig. 6. It was found that degradation rat increases with the increase in catalyst concentration up t 3.0 g L^{-1}, which on further increase leads to decrease i degradation rate.

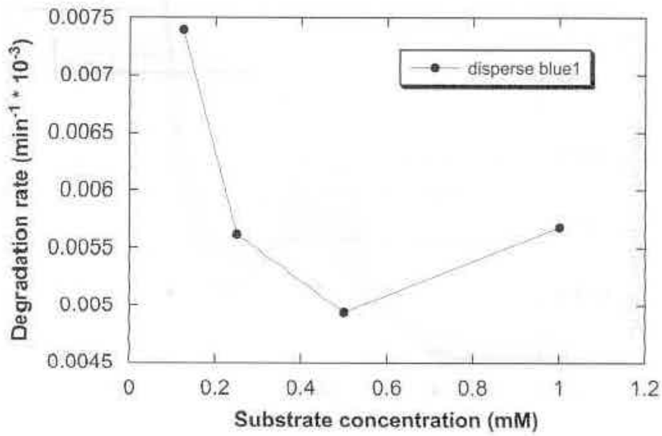


Fig. 5. Influence of substrate concentration on the degradation rate for the decomposition of disperse blue 1 (1) in the presence of TiO_2 containing hydrogen peroxide. Experimental conditions: substrate concentrations (0.25, 0.5, 0.75 and 1 mM), H_2O_2 (0.1 mL) V = 250 mL, photocatalyst: TiO_2 (Degussa P25, 1 g L^{-1}), irradiation time = 60 min.

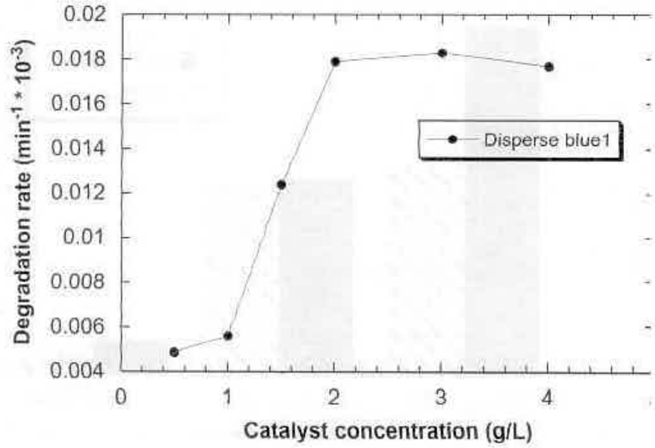


Fig. 6. Influence of catalyst concentration on the degradation rate for th decomposition of disperse blue 1 (1) in the presence of TiO_2 containin hydrogen peroxide. Experimental conditions: dye conc.: 0.25 mM, H_2C (0.1 mL), catalyst concentrations (0.5, 1.0, 1.5, and 2 g L^{-1}), V = 250 ml photocatalyst: TiO_2 (Degussa P25, 1 g L^{-1}), irradiation time = 60 min.

3.6. Effect of hydrogen peroxide

Hydrogen peroxide has been found to enhance the degradation of compound due to more efficient generation of hydroxyl radical and inhibition of electron/hole (e^-/h^+) pair recombination according to the following equations:

$$H_2O_2 + O_2^{\cdot -} \rightarrow {}^{\cdot}OH + OH^- + O_2, \tag{9}$$

$$H_2O_2 + h\nu \rightarrow 2\,{}^{\cdot}OH, \tag{10}$$

$$H_2O_2 + eCB^- \rightarrow {}^{\cdot}OH + OH^-. \tag{11}$$

In order to determine the optimum degradation condition, hydrogen peroxide is to be added for the degradation of model compound under investigation. Fig. 7 shows the degradation rate for decomposition of compound as a function of different concentrations of H_2O_2 containing TiO_2. The rate was found to increase with increase in H_2O_2 concentration from 0.1 to 0.3 mL followed by leveling off.

3.7. Photocatalysis of TiO_2 suspension containing dye derivatives under sunlight

For practical applications of wastewater treatment based on these processes, the utilization of sunlight is preferred. Hence the aqueous suspension of TiO_2 containing dye derivative was exposed to solar radiation. Fig. 8 shows the comparison of change in concentration as a function of irradiation time on illumination of an aqueous suspension of dye derivative in the absence and presence of H_2O_2 containing TiO_2 under sunlight. Blank experiments were carried out under sunlight in the absence of H_2O_2 and TiO_2 where no observable loss of the dye derivative takes place as shown in Fig. 8.

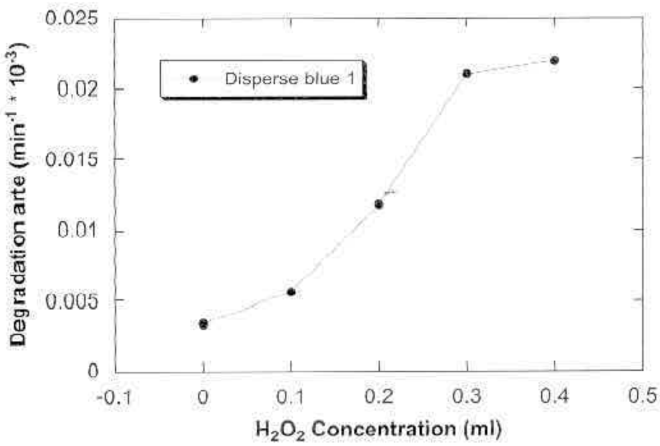


Fig. 7. Comparison of degradation rate for the decomposition of disperse blue 1 (**1**) in the presence of hydrogen peroxide at different concentrations. Experimental conditions: dye conc.: (0.25 mM) photocatalyst: TiO_2 (Degussa P25, $1\,g\,L^{-1}$) $V = 250$ mL, H_2O_2 (0.1, 0.2, 0.3 and 0.4 mL), irradiation time = 60 min.

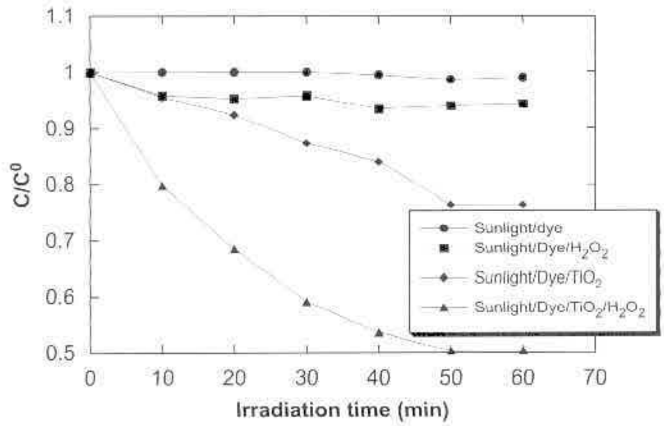


Fig. 8. Comparison of change in concentration as a function of time for irradiation of an aqueous suspension of disperse blue 1 (**1**) in the presence and absence of TiO_2 containing H_2O_2 under sunlight. Experimental conditions: dye conc.: (0.25 mM) H_2O_2 (0.1 mL) photocatalyst: TiO_2 (Degussa P25 $1\,g\,L^{-1}$), $V = 250$ mL, immersion well photoreactor, continuous air purging and stirring, irradiation time = 60 min.

5. Conclusion

TiO_2 can efficiently catalyze the degradation of dye derivatives **1** in the presence of light. The photocatalyst Degussa P25 showed better photocatalytic activity for the degradation of the model compound under investigation. The degradation of dye derivative was found to be slower under sunlight than that of the artificial light source. The addition of electron acceptors such as hydrogen peroxide can enhance the decomposition of model system. The observations of these investigations clearly demonstrate the importance of choosing the optimum degradation parameters to obtain high degradation rate, which is essential for any practical application of photocatalytic oxidation processes. The investigations were conducted at the laboratory scale in order to determine the optimum degradation condition and further studies are require for the practical effluent treatment.

Acknowledgments

Financial support by the Department of Science and Technology (DST), Govt. of India, New Delhi, for the award of Young Scientist to Dr. M. Saquib (Project No. SR/FTP/CS-25/2005) and The Department of Chemistry, Aligarh Muslim University, Aligarh, India, is gratefully acknowledged.

References

Augustynski, J., 1988. In: Structural Bonding. Springer, Berlin, New York, p. 69.

Bickley, R.I., Carreno, T.G., Lees, J.S., Palmisano, L., Tilley, R.J.D., 1992. A spectral investigation of titanium dioxide photocatalysts. Journal of Solid State Chemistry 92, 178.

Blake, D.M., 2001. Bibliography of work on the photocatalytic removal of hazardous compounds from water and air. National Renewal Energy Laboratory, USA pp. 40–253.

Bozzi, A., Dhananjeyan, M., Guasaquillo, I., Parra, S., Pulgarin, C., Weins, C., Kiwi, J., 2004. Evolution of toxicity during melamine photocatalysis with TiO_2 suspensions. Journal of Photochemistry and Photobiology A: Chemistry 162, 179.

Hurum, D.C., Agrios, A.G., Gray, K.A., Rajh, T., Thurnauer, M.C., 2003. Explaining the enhanced photocatalytic activity of Degussa P25 mixed phase TiO_2 using EPR. Journal of Phyical Chemistry B 107, 4545.

Lindner, M., Bahnemann, D.W., Hirthe, B., Griebler, W.D., 1995. Novel TiO_2 powders as highly active photocatalysts. In: Stine, W.R., Tanaka, T., Claridge, D.E. (Eds.), Solar Water Detoxification: Solar Engineering. ASME, New York, p. 339.

Lindner, M., Bahnemann, D., Hirthe, B., Griebler, W.D., 1997. Solar water detoxification: novel TiO_2 powders as highly active photocatalysts. Journal of Solar Energy Engineering 119, 120.

Mathews, RW., McEvoy, S.R., 1992. Photocatalytic degradation of phenol in the presence of near-UV illuminated titanium dioxide. Journal of Photochemistry and Photobiology A: Chemistry 64, 231.

Muneer, M., Theurich, J., Bahnemann, D., 1999. Formation of tox intermediates upon the photocatalytic degradation of the pestici Diuron. Research on Chemical Intermediates 25, 667.

Muneer, M., Theurich, J., Bahnemann, D., 2001. Titanium dioxi mediated photocatalytic degradation of 1,2-diethyl phthalate. Journ of Photochemistry Photobiology A: Chemistry 143, 213.

Rauer, S., 1998. Untersunchung von kommerziell erhaltlichen Titandio iden hinsichtlich ihrer photokatalytischen Aktivtat, Diplomarbe fachhochschule Hannover, Fachbereich Maschinenbau Vertiefu Umwelt-und Verfahrenstechnil, Hannover, Germany.

Sreedhar Reddy, S., Kotaiah, B., 2005. Decolorization of simulated spe reactive dye bath using solar/TiO_2/H_2O_2. International Journal Environmental Science & Technology 2, 245.

Turchi, C.S., Ollis, D.F., 1990. Photocatalytic degradation of organ contaminants: mechanism involving hydroxyl radical attack. Journ of Catalysis 122, 178.

Weichgrebe, D., Vogelpohl, A., 1995. Stratgie zur Auswal geeignet Oxidationsverfahren. 2. Fachtagung naboxidative abwasserbehadlun Clausthal, Germany.

Zollinger, H., 1987. Synthesis, properties and application of organic dy and pigments. In: Eblel, H.F., Brenzinger, C.D. (Eds.), Colo Chemistry, first ed. VCH, Inc., NewYork.

Journal of Environmental Management 88 (2008) 307–317

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Nutrient removal in a pilot and full scale constructed wetland, Putrajaya city, Malaysia

Cheng Hua Sim[a,*], Mohd Kamil Yusoff[a], Brian Shutes[b], Sinn Chye Ho[c], Mashhor Mansor[d]

[a]*Department of Environmental Sciences, Faculty of Environmental Studies, University Putra Malaysia, 43400 UPM-Serdang, Malaysia*
[b]*Biomedical and Environmental Sciences, Middlesex University, Queensway, Enfield EN3 4SA, UK*
[c]*National Oceanography Directorate, Ministry of Science, Technology and Innovation, Block C5, Parcel C, 62662 Putrajaya, Malaysia*
[d]*School of Biological Sciences, University Science Malaysia, 11800 Pulau Pinang, Malaysia*

Received 22 March 2006; received in revised form 1 February 2007; accepted 10 March 2007
Available online 30 April 2007

Abstract

Putrajaya Wetlands in Malaysia, a 200 ha constructed wetland system consisting of 24 cells, was created in 1997–1998 to treat surface runoff caused by development and agricultural activities from an upstream catchment before entering Putrajaya Lake (400 ha). It was designed for stormwater treatment, flood control and amenity use. The water quality improvement performance of a section of the wetland cells is described. The nutrient removal performance was 82.11% for total nitrogen, 70.73% for nitrate–nitrogen and 84.32% for phosphate, respectively, along six wetland cells from Upper North UN6 to UN1 from April to December 2004.

Nutrient removal in pilot scale tank systems, simulating a constructed wetland and planted with examples of common species at Putrajaya, the Common Reed *Phragmites karka* and Tube Sedge *Lepironia articulata*, and the capacity of these species to retain nutrients in above and below-ground plant biomass and substrate is reported. The uptake of nutrients by the Common Reed and Tube Sedge from the pilot tank system was 42.1% TKN; 28.9% P and 17.4% TKN; 26.1% P, respectively.

The nutrient uptake efficiency of the Common Reed was higher in above-ground than in below-ground tissue. The results have implications for plant species selection in the design of constructed wetlands in Malaysia and for optimizing the performance of these systems.

Keywords: Constructed wetland; Nutrient removal; Surface flow; Wetland plant

1. Introduction

Constructed wetlands are man-made systems or engineered wetlands that are designed, built and operated to emulate functions of natural wetlands. They are created from a non-wetland ecosystem or a former terrestrial environment, mainly for the purpose of pollutant removal from wastewater. The constructed wetland treatment system is a cheaper alternative for wastewater treatment using local resources and is an energy-efficient technology. The system utilizes wetland plants and micro-organisms, which are the active agents in the treatment processes (Kadlec and Knight, 1996). Most of the constructed wetland systems are marshes with shallow water regions dominated by emergent marsh plants such as cattails, bulrushes, rushes and reeds. Constructed wetlands offer many multiple use values such as the creation of habitat, water quality improvement, flood control, and production of food and fibres (also termed as constructed aquaculture wetlands).

Constructed wetland systems can potentially tolerate variable volumes of water and varying contaminant levels. The sources include municipal and domestic wastewater, urban surface runoff, agricultural wastewater, industrial effluents and polluted surface waters in rivers and lakes (Sekiranda and Kiwanuka, 1998). The wetland systems can

*Corresponding author. Tel.: +60 12 6349218; fax: +60 3 80617264.
E-mail addresses: michellesim.chenghua@gmail.com (C.H. Sim), mkamil@env.upm.edu.my (M.K. Yusoff), B.Shutes@mdx.ac.uk (B. Shutes), scho@mosti.gov.my (S.C. Ho), mashhor@usm.my (M. Mansor).

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2007.03.011

also be aesthetically pleasing and serve as an attractive destination for tourists and local urban residents. They can also serve as a public attraction or sanctuary for visitors who wish to explore their environmental and educational possibilities. They appeal to different groups varying from engineers to those involved in wastewater facilities as well as environmentalists and people concerned with recreation. These systems also provide a research and training ground for young scientists in this new research and education arena.

1.1. Water quality in Malaysia

Nutrients are the main agricultural pollutants in Malaysia. Drainage ditches, irrigation channels, ponds and other waterways are polluted by agricultural runoff from fertilizer rich land such as vegetable farms, fruits and flower nurseries, golf courses and animal farms. More than 63% of the rivers in Malaysia are classified as moderately to highly polluted. They receive urban runoff polluted with domestic sewage discharges and livestock excreta, as well as from agricultural uses and wastewater from factories. The river waters have high concentrations of biological oxygen demand, nutrients and pathogens, resulting in a risk to public health for bathing and fishing, particularly in areas of poor or impoverished human population and water recreational areas. Constructed wetlands are considered an attractive alternative to achieving good river water quality with significant reduction of pathogens as well as low values of biological oxygen demand and nutrients (Ceballos et al., 2001). Kadlec and Hey (1994) have successfully illustrated the potential of constructed wetlands in controlling non-point source pollution in the watershed and for river water quality improvement. A riparian wetland can be constructed to treat nutrient pollution in rivers in a reasonably cost effective manner (White et al., 1994).

1.2. Putrajaya Wetlands

Putrajaya Wetlands (200 ha), comprising 24 wetland cells, were created in the valley of the Chuau and Bisa Rivers from agricultural lands of oil palm and rubber plantations, within a period of 17.5 months in 1997 and 1998. The wetland system and lake were fully inundated in January 1999. The wetlands were created mainly to restore the polluted Chuau and Bisa river systems caused by agricultural activities in the upstream catchment; and also to play a role in stormwater treatment and flood control. Putrajaya Wetlands are a vegetated horizontal surface flow multi-cell wetland system, designed with different water levels in each of the cells that are separated by a weir (Shutes, 2001). The water flows through these wetland cells and finally discharges into Putrajaya Lake (Fig. 1). The wetland helps to treat non-point source pollutants in order to achieve the national water quality standard Class IIB for recreational use in the lake area.

The wetland cells were excavated to create different water depths of 0.5–3.0 m and filled with topsoil for wetland planting. The bed slope was gradually increased from upstream to downstream to achieve a continuous flow by gravity. Wastewater flows horizontally through the wetland cells and is visible on the surface. The planted marsh area covers 77.7 ha or 39.4% of the total wetland area; the zone of intermittent inundation for swamp tree planting covers 23.7 ha or 12.0%; and the open water covers 76.8 ha or 38.9%. The remaining area comprises weirs, islands and maintenance tracks which cover about 10%.

The wetland cells were planted with 27 types of emergent wetland plants, which play a role in sediment retention and in nutrient and toxicant removal. Another 35 species of herbaceous plants were planted in the zone of intermittent inundation fringing the marsh area as a border for the wetland system, and these species assist in erosion control and bank stabilization. Twenty-two small islands were created in open water areas, which are vegetated and provide isolated habitats for birds and other fauna. These islands also serve to facilitate water filtration, divert water flow to avoid stagnation and help in the flow to the fringes of wetland cells (especially the planting areas). The ornamental ponds were planted with colourful flower species, such as water lilies, to increase their aesthetic value. These wetland plants were sourced locally and they possess high adaptability to conditions of intermittent inundation.

1.3. Field and pilot experimental wetland studies

The existing and newly developed constructed wetland systems in Malaysia are all designed for stormwater or surface water treatment. Nutrient levels in these waters are generally low depending on land use activities in the catchment. The aim of this pilot wetland study was to understand the potential capacity of constructed wetlands for treating water with high nutrient concentrations such as septic effluent and agricultural runoff from nurseries as well as livestock and crop farms. The experimental results are compared with the results of a field study in the full scale Putrajaya Wetlands.

2. Materials and methods

2.1. Field study

The three wetland cells at Upper North arm (UN4–6) located at 02°58.02–02°58.53N; 101°41.94–101°42.08E were selected as the project site (Fig. 1). Water quality was monitored in these three wetland cells for two separate periods, from October 2001 to December 2002 at bimonthly intervals and from April to December 2004 at monthly intervals. The final cell of the UN arm (UN1A) was also monitored with four lateral inlet points S5–S9 at monthly intervals from April to December 2004 (Fig. 1).

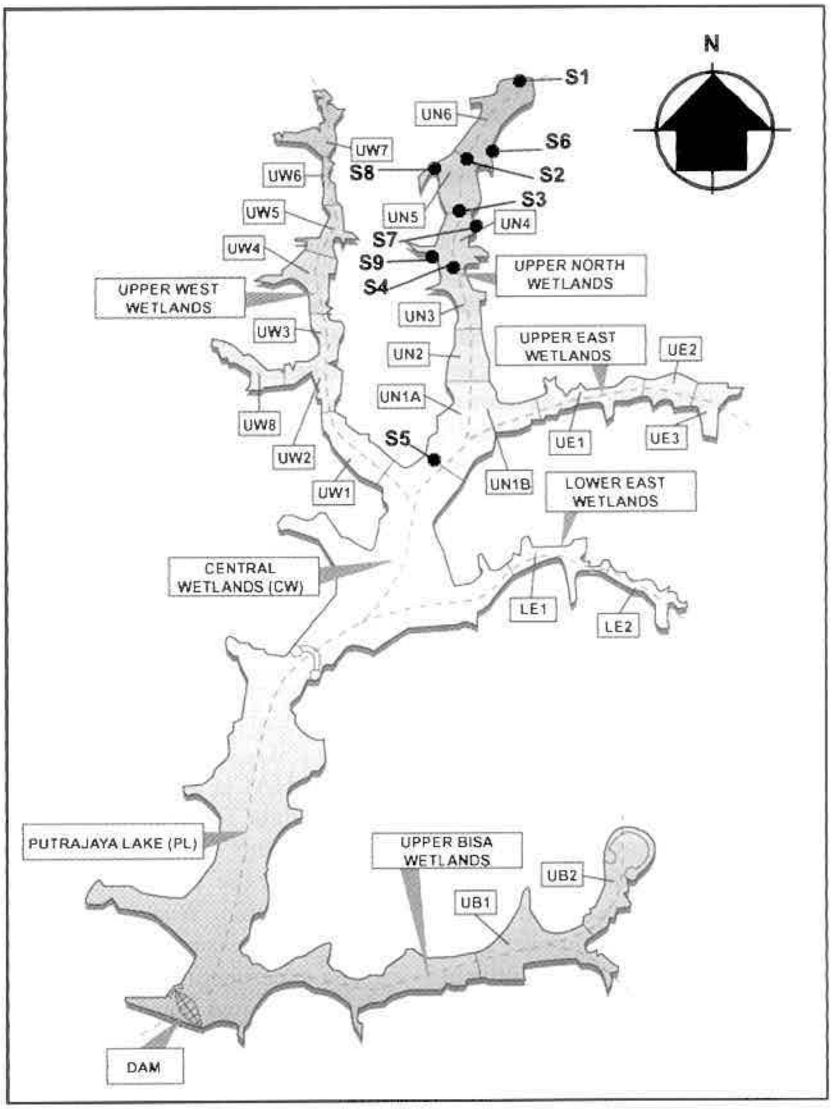


Fig. 1. Location of wetland cells at Putrajaya Wetlands (Ariffin, 1998).

Water samples were collected from outlet points (on the weir, S1–S4) of wetland cells UN4–6 (Fig. 1). The inlet of the wetland cell downstream resembles the outlet of the upstream cell. The water quality parameters measured in situ were dissolved oxygen, pH, temperature, conductivity, turbidity, water transparency and water depth. Water samples were collected and brought back to the laboratory at University Putra Malaysia for analysis. Water samples were analysed for total suspended solids, total dissolved solids and nutrient content using the HACH method (1989). Plant samples were collected from the field for nutrient content analysis.

A broad crested weir was installed at the outlet of each Putrajaya Wetland cell. The flow rate was undetectable by a flow meter, thus for calculating the flow, the basic weir equation (Wiese et al., 1998) was used: $Q = C_{\mathrm{w}}LH^{1.5}$ where Q is the flow ($\mathrm{m^3\,s^{-1}}$), L the length of weir (m), H the depth of flow over the weir (m) and C_{w} the constant factor (1.84 for a broad crested weir).

The water flow was controlled by the inlet chamber and the outlet conduit. The weir crest level was made adjustable using a system of drop boards. Riffle rocks were used to create a pool from which water can be diverted into the downstream wetland.

Nutrient load estimates were also used to assess the nutrient removal efficiency of the wetland. The nutrient removal percentage of nitrate–nitrogen and phosphate along the six wetland cells for the two sampling periods was calculated.

2.2. *Pilot wetland study*

A pilot study was conducted at University Putra Malaysia to determine the nutrient removal efficiency of two wetland plant species. The tank system was constructed simulating the surface flow wetland system in Putrajaya Wetlands. Two native wetland plants, the Common Reed *Phragmites karka* and the Tube Sedge

Lepironia articulata were selected and the experimental period ran from July 2002 to February 2003. These two wetland plants are the dominant marsh plants in the Putrajaya Wetland cells. Eighty-one young stands of the Common Reed and Tube Sedge were planted with their potting materials in three rows at equal intervals into three replicate fibreglass tanks (sized 1.5 m length × 0.7 m width × 0.75 m height), respectively. Each tank had a thickness of 5-mm to avoid leakage and to hold a large capacity. A piping system was installed to these tanks on site, which included inlet feeding pipes, outlet taps and overflow holes (Fig. 2). They were filled largely with soil and some small portions of gravel and sand (in a mixture of soil: gravel: sand at 8:1:1) to avoid clogging and short circuiting, and filled to a depth of 0.6 m for maximum root growth. Gravel sizes used ranged from 0.5 to 4.0 cm in diameter.

The young Tube Sedge plants were allowed to establish for 5 weeks before nutrients were loaded, whereas nutrients were loaded to the Common Reed treatment tanks after 4 weeks. The Common Reed and Tube Sedge tanks were fed continuously for 30 weeks and 8 weeks, respectively, with nutrient solution concentrations of $50.0\,\mathrm{mg\,l^{-1}}$ N and $5.0\,\mathrm{mg\,l^{-1}}$ P. The nutrient solution was prepared by mixing a soluble granule $Ca(NO_3)_2$ (contained 24% N) and powdered commercial manure KH_2PO_4 (contained 50% P). The nutrient application rate to the treatment tanks was 22.5 g N and 2.25 g P per week, which was diluted into 450 l of water and then fed into three replicate treatment tanks. The flow rate of nutrient solution from the supplier tank into the treatment tanks was regulated by taps. Influent flow across the 'wetland' and effluent was collected in a collector tank and manually pumped back using a submersible pump to the supplier tank to be re-circulated to the treatment tanks for three cycles before it wa discarded and replaced with new solution on day 7 and ne cycles repeated.

A 20 mm PVC pipe and a 20 mm valve were used t regulate flow. The inlet feeding pipe in each treatment tan was installed on 5 cm below the surface of the substrat with an approximate length of 0.7 m long and wit perforated holes to ensure an even distribution of surfac flow. The flow rate was regulated to ensure that a unifor flow of $2.0\text{–}7.3 \times 10^{-6}\,\mathrm{m^3\,s^{-1}}$ was achieved. The outl sampling tap was fixed at 5 cm below the substrate surfac An overflow hole was installed at 5 cm above the substrat surface, therefore ensuring a consistent water level of abou 5 cm above the substrate. A further two outlet taps wer installed at 15 and 35 cm from the bottom of the bed fc maintenance purposes, for example, if a clogging probler occurs. Another three replicate tanks, planted with th Common Reed and Tube Sedge but not receiving nutrien supplements, were set up as controls. These control tank were filled with tap water initially and topped up wit rainfall.

The water samples were analysed weekly for chemica oxygen demand (COD), ammoniacal–nitrogen (NH_3–N nitrate (NO_3–N), nitrite (NO_2–N) and phosphate (PO_4–P using a HACH spectrophotometer (HACH, 1989). Plan growth data for the Common Reed was collected fron measurements of leaf and stem length, and shoot numbe whereas for the Tube Sedge, stem length and stem numbe were measured. Weekly visual inspections for pest attac and plant diseases were carried out.

The experiments were continuously monitored until th plants indicated signs of ageing or senescence or wilte when the system became overloaded with nutrients. At th end of each experiment, three replicate plant samples fron

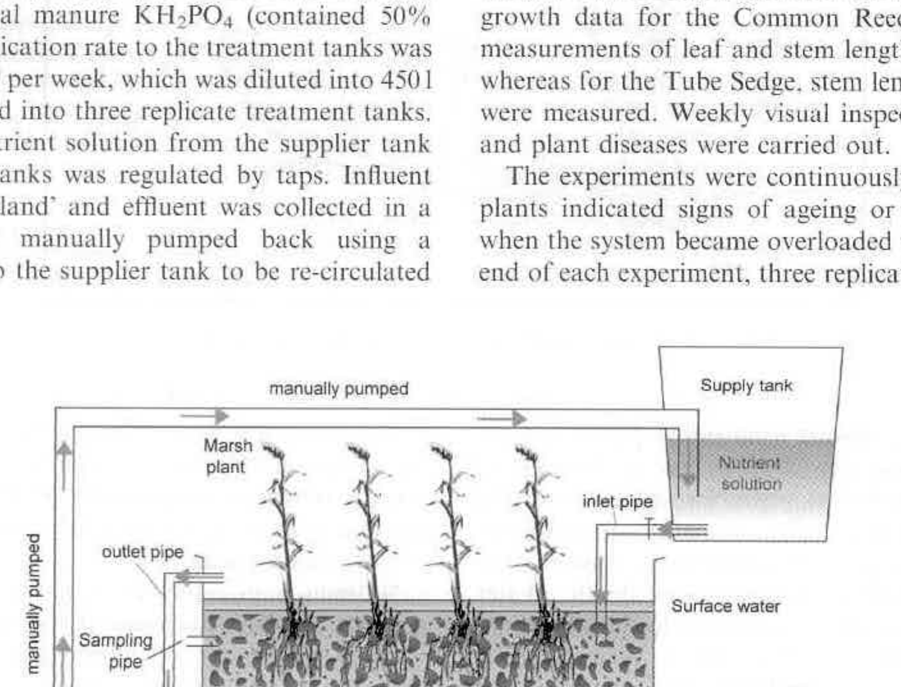


Fig. 2. Schematic diagram of a surface-flow wetland mesocosm at University Putra Malaysia.

the treatment and control tanks were selected and harvested. The plants were cleaned and sorted into leaf, stem and root components. The fresh biomass weights were recorded and above-ground biomass/below-ground biomass ratio and leaf/above-ground stem ratio were determined. The plants were dried in an oven at 70 °C and the water content of the plant biomass was determined. The samples were then ground into powder form for digestion and N and P content was analysed using the Kjeldahl (Blamire, 2003) and drying ash methods (Bureau of Nutritional Sciences Ottawa, 1983). Plant litter from the Common Reed treatment tanks was also collected for nutrient analysis as were the substrate samples.

3. Results

3.1. Field study

The following parameters occasionally exceeded the two water quality standards: the Putrajaya Lake water quality standard and the Class IIB water quality standard of the Department of Environment for primary body contact for water recreational activities; turbidity, ammoniacal–nitrogen, phosphate, total suspended solids and COD (Table 1).

Water depths in the planted sections of the wetland cells ranged from 0.4 to 2.0 m, although the drop boards were not regulated regularly, as the normal water level in the inlet chamber remained in the same range i.e. 28.5–28.8 m in UN6, 27.2–27.4 m in UN5, 26.0–26.3 m in UN4 and 23.8–24.9 m in UN1A. The flow rates were in the range of 1.38–19.78 $m^3 s^{-1}$ in wetland cells UN1–6 throughout the experimental period from April to December 2004.

The nutrient removal performance was 70.73% for nitrate–nitrogen and 84.32% for phosphate, respectively, along six wetland cells from Upper North UN6 to UN1 from April to December 2004 (Tables 2 and 3). The nutrient removal performance detected along wetland cells UN6 to UN4 was 12.96% for nitrate–nitrogen and 33.15% for phosphate from October 2001 to December 2002. The nutrient removal performance was higher from UN6 to UN4 from April to December 2004 at 36.28% for nitrate–nitrogen and 40.42% for phosphate. The variability and lower nutrient removal is due to the intermittent addition of lateral sources of nutrients to the wetland cells between UN6 and UN1 or to low nutrient concentrations. The overall improvement of nitrate–nitrogen and phosphate concentrations in UN6 to UN4 within these two periods is shown in Figs. 3 and 4.

Nutrient accumulation by the Common Reed in the field showed an increase of 25% N and 60% P in below-ground tissue between October 2001 and April 2004 whereas in substrates, an increase of 8.5% P was recorded. In the same period of 30 months, there was a low increase in nutrient accumulation at 2.45 $mg\,kg^{-1}\,day^{-1}$ N and 0.276 $mg\,kg^{-1}\,day^{-1}$ P in the Common Reed plant biomass collected from the field. Thus, harvesting practice is not applied in Putrajaya Wetlands.

3.2. Pilot wetland study

3.2.1. Growth rate and plant biomass harvested

The growth rate and total harvested biomass of both treated wetland plants were higher than those in the control tanks. The fresh biomass per plant of Common Reed was

Table 1
Ranges of water quality data from sampling sites S1–S4 in Putrajaya Wetlands and water quality standards

Water quality parameter	Sampling period		Putrajaya Lake water quality standard	Department of Environment Class IIB standard
	October 2001–December 2002	April–December 2004		
In situ parameters				
pH	5.5–7.4	6.85–7.65	6.5–9	6–9
Conductivity ($\mu S\,cm^{-1}$)	54.8–146.3	98.05–69.4	1000	–
Dissolved oxygen ($mg\,l^{-1}$)	0.78–13.25	2.5–5.02	5–7	5–7
Water transparency (m)	0.069–0.51	0.098–0.51	0.6	–
Water depth (m)	1.01–1.52	1.18–1.56	–	–
Turbidity (Nephelometric Turbidity Unit)	21.7–284.3[a]	18.7–134.2[a]	50	50
Laboratory analysed parameters				
Ammoniacal–nitrogen ($mg\,l^{-1}$)	0.13–0.72[a]	0.21–1.67[a]	0.3	0.3
Nitrate–nitrogen ($mg\,l^{-1}$)	0.07–2.23	0.7–1.78	7	7
Phosphate ($mg\,l^{-1}$)	0.07–0.32[a]	0.05–0.28[a]	0.05	0.2
Total suspended solids ($mg\,l^{-1}$)	10.25–137.5[a]	7.2–73.2[a]	50	50
Total dissolved solids ($mg\,l^{-1}$)	15.62–63.57	41.1–69.7	–	–
Chemical oxygen demand ($mg\,l^{-1}$)	24–48.75[a]	–	25	25
Biological oxygen demand ($mg\,l^{-1}$)	0.38–1.65	–	3	3

[a]Indicates parameter that does not comply to Malaysian water quality standard.

Table 2
Nitrate–nitrogen (NO_3–N) and phosphate (PO_4–P) removal rate from sampling sites S1–S4 in Putrajaya Wetlands from October 2001 to December 2002

	$Q=C_wLH^{1.5}$ ($m^3 s^{-1}$)	NO_3–N ($mg l^{-1}$)	PO_4–P ($mg l^{-1}$)	Loading nitrate–N ($kg day^{-1}$)	Loading PO_4 (kg day^{-1})
S1	16.48±2.47	1.29±0.36	0.18±0.11	1.83	0.26
S2	11.11±2.72	1.79±0.52	0.17±0.09	1.71	0.16
S3	8.48±2.89	1.50±0.43	0.17±0.10	1.1	0.12
S4	11.82±3.19	1.56±0.70	0.17±0.08	1.6	0.17

Sampling points	NO_3 removal rate (%)	PO_4 removal rate (%)
S1–S2	6.4	39.0
S2–S3	36.0	24.3
S3–S4	−45.3	−44.8
S1–S3	40.1	53.8
S1–S4	12.9	33.2

$n=8$. Loading = flow, $Q \times$ nutrient concentration.

Table 3
Nitrate–nitrogen (NO_3–N) and phosphate (PO_4–P) removal rate from sampling sites S1–S5 in Putrajaya Wetlands from April to December 2004

	$Q=C_wLH^{1.5}$ ($m^3 s^{-1}$)	NO_3–N ($mg l^{-1}$)	PO_4–P ($mg l^{-1}$)	Loading nitrate–N ($kg day^{-1}$)	Loading PO_4 ($kg day^{-1}$)
S1	17.22±1.81	1.38±0.67	0.16±0.07	2.0	0.2
S2	5.95±2.47	1.23±0.35	0.17±0.06	0.6	0.1
S3	15.86±2.43	1.26±0.36	0.16±0.11	1.7	0.2
S4	11.83±3.63	1.28±0.45	0.14±0.11	1.3	0.1
S5	7.20±3.97	0.96±0.33	0.06±0.07	0.6	0.1

Sampling points	NO_3 removal rate (%)	PO_4 removal rate (%)
S1–S2	69.2	64.4
S2–S3	−174.9	−152.6
S3–S4	24.7	33.8
S4–S5	54.1	73.7
S1–S3	15.4	10.1
S1–S4	36.3	40.4
S1–S5	70.7	84.3

$n=8$. Loading = flow, $Q \times$ nutrient concentration.

0.58 ± 0.08 and 0.16 ± 0.09 kg, respectively, whereas for Tube Sedge the fresh biomass per plant was 0.31 ± 0.10 and 0.15 ± 0.01 kg, respectively for plant samples harvested from the treatment and control tanks (Table 4 and Fig. 5).

In this 30-week growth period, the treatment plant samples of the Common Reed showed a long period of growth but a shorter maturity stage compared to the control plant samples which demonstrated a longer maturity stage. However, Common Reed treated plants experienced senescence and Tube Sedge treated plants collapsed in a shorter period. Generally a wetland plant experiences senescence after the post-flowering stage. However, no flowering stage was observed in both control and treatment tanks of the Common Reed. The treated

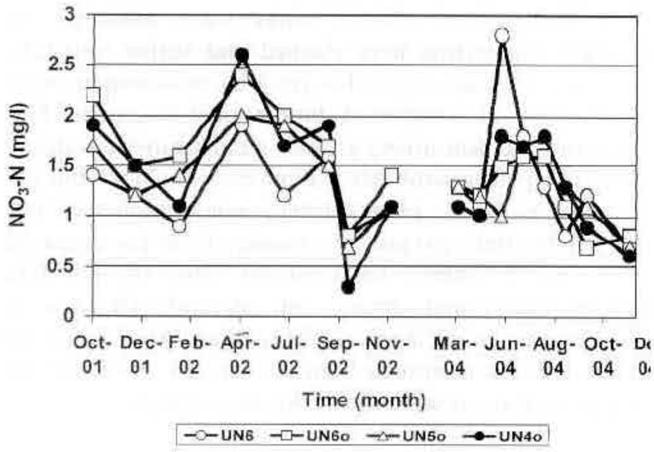


Fig. 3. Nitrate–nitrogen (NO_3–N) concentrations ($mg l^{-1}$) in Upp North cells UN4–6 in 2001–2002 and 2004 (o = outlet).

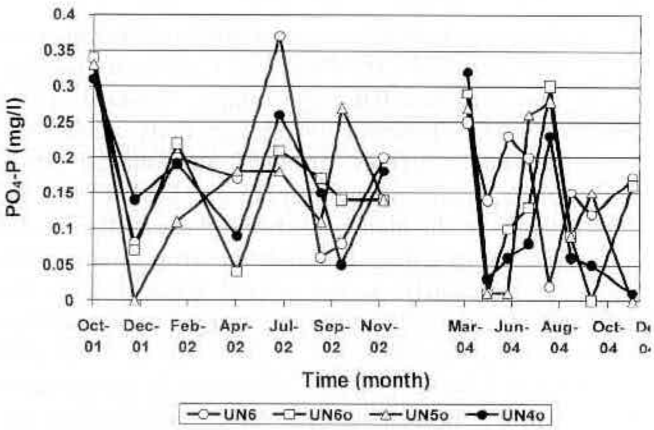


Fig. 4. Phosphate (PO_4–P) concentrations ($mg l^{-1}$) in Upper North ce UN4–6 in 2001–2002 and 2004 (o = outlet).

Tube Sedge stands collapsed after 8 weeks of th experimental period, probably due to nutrient overloa conditions. Wetland plants are able to uptake hig amounts of nutrient under high nutrient load condition until a threshold concentration is achieved.

3.2.2. Nutrient uptake

The nutrient uptake efficiency of the Common Reed wa higher in leaf than in root and above-ground stem. Fo Tube Sedge, the nutrient uptake efficiency was higher i root than in stem (Table 5). Nutrient content in treated lea and stem samples of the Common Reed (25.100 g kg^{-1} N 1.270 g kg^{-1} P and 11.370 g kg^{-1} N; 0.680 g kg^{-1} P) wa higher than those in stem samples of the Tube Sedg (8.430 g kg^{-1} N; 0.870 g kg^{-1} P), except the phospha accumulation in the stem sample was lower in the Commo Reed. Nutrient content in root samples of the Tube Sedg (12.500 g kg^{-1} N; 1.620 g kg^{-1} P) was higher than in th root samples of the Common Reed (12.300 g kg^{-1} N 1.010 g kg^{-1} P). The total net nutrient accumulation in th plant biomass of the Common Reed was 0.076 kg N an 0.005 kg P per tank for a period of 30 weeks and the Tub

Table 4
Fresh and dry biomass of the Common Reed and Tube Sedge in the treatment and control tanks

	Common Reed (210 days)		Tube Sedge (56 days)	
	Treatment plant sample	Control plant sample	Treatment plant sample	Control plant sample
Number of samples	$n = 3$			
Fresh biomass per plant (kg)	0.58±0.08	0.16±0.09	0.31±0.10	0.15±0.01
Total above-ground fresh biomass per tank (kg)	6.21±1.69	1.106±0.155	3.60±0.53	1.87±0.15
Above-ground biomass/below-ground biomass ratio	0.59±0.27	0.47±0.20	4.40±0.85	2.90±1.01
Leaf/above-ground stem ratio	0.28±0.028			
Water content in plant biomass (%)				
Leaf	43.9	43.9		
Above-ground stem	47.3	47.3	58.9	58.9
Below-ground biomass	72.0	72.0	65.1	65.1
Total dry plant biomass (kg)				
Leaf	0.76	0.14		
Above-ground stem	2.56	0.46	1.48	0.77
Below-ground biomass	2.95	0.66	0.29	0.23

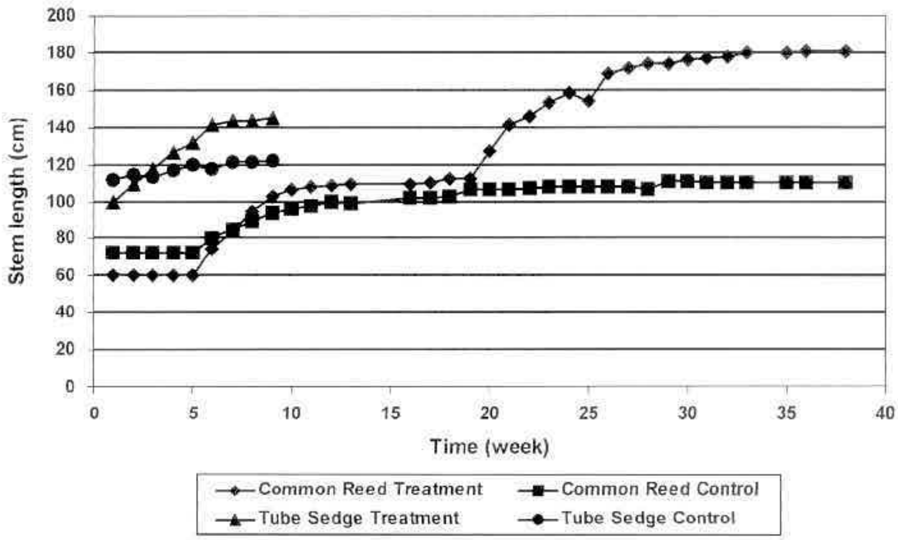


Fig. 5. Growth of Common Reed and Tube Sedge stems in treatment and control tanks.

Sedge showed a lower nutrient content at 0.009 kg N and 0.001 kg P per tank for a period of 8 weeks (Table 6), which is probably partly due to the different time periods. This result showed that the Common Reed has a higher nutrient absorption and storage capacity. Total nitrogen accumulation in the Common Reed treatment plant samples was 0.344 $g\,m^{-2}\,day^{-1}$ compared to 0.155 $g\,m^{-2}\,day^{-1}$ in the Tube Sedge whereas total phosphate accumulation in the Common Reed treatment plant samples was 0.024 $g\,m^{-2}\,day^{-1}$ and 0.023 $g\,m^{-2}\,day^{-1}$ in Tube Sedge (Table 6). Nutrient accumulation of 24.000 $g\,kg^{-1}$ N and 1.200 $g\,kg^{-1}$ P also was recorded in plant litter collected from the Common Reed treatment tanks compared to 16.300 $g\,kg^{-1}$ N; 0.500 $g\,kg^{-1}$ P in the control tanks.

3.2.3. Nutrient removal

The water sample nutrient content analysis showed that the removal efficiency in the Common Reed and Tube Sedge treatment tank systems was 81.3% TN (76.3% nitrate–nitrogen); 32.5% P and 52.6% TN (43.4% nitrate–nitrogen); 24.7% P, respectively, throughout the experimental period from July 2002 to February 2003. It was not possible to continuously monitor the variable flow rate in the outlet of the tanks, therefore the nutrient removal efficiency was determined from nutrient concentrations and not from loadings. The nutrient concentrations improved throughout the experimental period (Fig. 6). Phosphate concentrations showed marked variability in comparison to ammoniacal–nitrogen and nitrate–nitrogen concentrations in the Common Reed treatment tank study.

For the experimental period of 30 weeks, a total of 0.180 kg N and 0.018 kg P was applied to each of the treatment tank of the Common Reed. While for the Tube Sedge, a total of 0.053 kg of N and 0.005 kg of P was applied to each treatment tank for an experimental period of 8 weeks. Nutrient removal through uptake by the

Table 5
Nutrient content in above-ground and below-ground plant biomass of the Common Reed and Tube Sedge in the treatment and control tanks

	Common Reed		Tube Sedge	
	Treatment plant samples	Control plant samples	Treatment plant samples	Control plant samples
N (mg kg^{-1})				
Number of samples	$n = 3$			
Above-ground biomass				
Stem samples	11370 ± 1650	5670 ± 1680	8430 ± 5450	7570 ± 1860
Leaf samples	25100 ± 2340	10830 ± 3100		
Below-ground biomass	12300 ± 1212	6930 ± 2050	12500 ± 4850	4800 ± 2780
P (mg kg^{-1})				
Above-ground biomass				
Stem samples	680 ± 360	330 ± 90	870 ± 40	360 ± 150
Leaf samples	1270 ± 270	700 ± 260		
Below-ground biomass	1010 ± 400	360 ± 50	1620 ± 760	450 ± 180

Table 6
Nutrient uptake of the Common Reed and Tube Sedge in the pilot tank system

	Common Reed (210 days)		Tube Sedge (56 days)	
	N	P	N	P
Number of samples	$n = 3$			
Total net nutrient uptake by plant per tank (kg)	0.076	0.005	0.009	0.001
Total nutrient applied per treatment tank (kg)	0.18	0.018	0.0525	0.0053
% of net plant uptake	42.1	28.9	17.4	26.1
Total net nutrient accumulation rate (g m^{-2} day^{-1})	0.344	0.024	0.155	0.023

Note: Net value was calculated by deduction of nutrient content in control plant samples.

Common Reed and Tube Sedge was 42.1% N; 28.9% P and 17.4% N; 26.1% P, respectively (Table 6). There was no accumulation of either N or P in the substrate in the Common Reed and Tube Sedge tanks compared to the initial substrate samples. However, higher nutrient accumulation was recorded in the substrate of Tube Sedge treatment tanks (0.600 g kg^{-1} N; 0.052 g kg^{-1} P) compared to those in the control tanks (0.500 g kg^{-1} N; 0.049 g kg^{-1} P).

4. Discussion

4.1. Field study

The field water quality results showed that water quality normally improved with flow length along the wetland cells. Although the aspect ratio of each wetland cell UN6–1 was below 4, the minimum recommended by Crites (1994), the combined maximum aspect ratio of UN6–UN1 is 8.19 (Table 7). The water quality improvement was reduced during periods of rainfall when levels of total suspended solids, nitrate and phosphate were highly variable. Malaysia has high rainfall throughout the year, with an annual rainfall of 2300–3000 mm year^{-1}. The field study results showed that nutrient removal efficiency increases with the hydraulic residence time, but decreases during periods of high precipitation. In wet climates, both treatment and hydraulic performance should improve if design strategies that minimize ingress of rainfall can be developed (Davison et al., 2001).

Negative nutrient removal was recorded in UN4 during October 2001 to December 2002, due to the discharge of higher nitrate–nitrogen concentrations from the side inflows to the wetland cell. Furthermore, it is also due to fluctuation in water discharge rates between inlet and outlet of UN4.

High rainfall will cause heavy silt load being washed to the wetland. Siltation can reduce long term wetland performance and the service-life of a wetland considerably (Tanner et al., 2002). Serious siltation occurred in wetland cell UN5 during April to December 2004, caused by discharges of sediment from the side inflows, which impaired the wetland function resulting in no nitrate–nitrogen removal being recorded (Table 3).

About 50–60% of the planted zone of this wetland cell was invaded by terrestrial weeds. The siltation resulted in shallower wetland bed and flow channelization. The open water area has been largely reduced to a small area resulting in a ponding effect. Dredging of silt is recommended to be undertaken to restore the nutrient removal function.

In Putrajaya Wetlands, development and landuse activities in the catchment cause a high pollutant load to

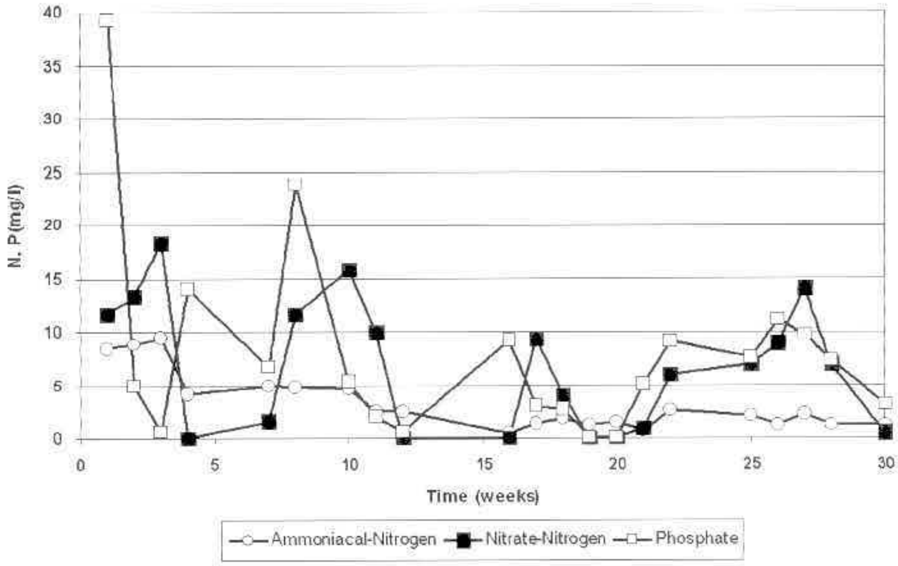


Fig. 6. Nutrient content in treated water samples from the Common Reed treatment tank.

Table 7
Aspect ratio of wetland cells UN1–6

Wetland cell	Wetland cell length (m)	Wetland cell maximum width (m)	Wetland cell minimum width (m)	Maximum aspect ratio	Minimum aspect ratio
UN1	320	517	467	0.69	0.62
UN2	255	367	233	1.09	0.69
UN3	365	173	133	2.74	2.11
UN4	345	200	175	1.97	1.73
UN5	265	260	225	1.18	1.02
UN6	475	250	240	1.98	1.90
UN1–6	2025	293	247	8.19	6.91

be discharged into the wetland through the v-drains and gross pollutant traps. However, a study of inflow from a lateral bank of the wetland showed no impact on water quality improvement except for an increase in total suspended solids levels.

The high evaporation (1300–1400 mm year^{-1}) and evapotranspiration rate (6–16.76 mm day^{-1}) in Malaysia may cause concentrations of parameters in a wetland outflow to be higher than that of the inflow. It is important to control the water discharge rate along the wetland cells to ensure that nutrient removal processes are optimized in the wetlands. Water level control is important to ensure the survival of wetland plants in long periods of inundation.

4.2. Pilot wetland study

The planted wetlands in the pilot study showed higher nutrient removal efficiency than the unplanted wetlands. A clear trend of improved total nitrogen removal and enhanced phosphorus mass removal achieved by planted wetlands compared to unplanted controls was also shown by Tanner et al. (1995). The wetland plants have a large above and below biomass and these sub-surface plant tissues grow horizontally and vertically, and create a large surface area for the uptake of nutrients and ions (Cooper et al., 1996). Each wetland plant species shows differential accumulation and release of N and P and may influence the overall potential of a treatment wetland (Kao et al., 2003). Tropical wetland plants show high biomass production and the values can be used as an indicator to estimate the nutrient uptake capacity of the plants (Greenway and Woolley, 2001). In this study, both the Common Reed and Tube Sedge had a high biomass growth of 0.49–0.66 and 0.195–0.380 kg per plant, respectively (Table 4), and a maximum growth of stem length at 9.30 and 8.37 cm week^{-1}, respectively (Fig. 5). The growth rate of Common Reed and Tube Sedge (above-ground biomass) was higher, at 0.028 and 0.061 kg m^{-2} day^{-1}, respectively, compared to other wetland plant species such as *Typha domingensis*, *Schoenoplectus validus* and *Eleocharis* spp. (0.002–0.006 kg m^{-2} day^{-1}) in a pilot wetland system in Cairns, Australia (Greenway and Woolley, 2001). However, the Common Reed, with an extensive root system and a higher below-ground biomass achieved higher plant uptake in comparison to the Tube Sedge.

Nutrient removal through plant uptake by the Common Reed and Tube Sedge in this pilot study was 42.1% N; 28.9% P and 17.4% N; 26.1% P, respectively. Table 8 shows the nutrient removal efficiencies of studies of plant species in constructed wetlands. Plant uptake by soft-stem bulrush *Schoenoplectus tabernaemontani* and *Baumea articulata* accounted for around 11–26% of the N and 3–29% of the P removal rates (Tanner et al., 1995; Browning and Greenway, 2003). Breen (1990) and Rogers et al. (1991) reported the nitrogen plant uptake of 55% for cattails and 85% for *Schoenoplectus validus*, respectively. A study by Lim et al. (2001) in Malaysia showed that about

Table 8
Nutrient removal efficiencies of studies of plant species in constructed wetlands

Studies	Nutrient removal efficiencies ($kg\,ha^{-1}\,day^{-1}$)		Types of wastewaters	Plant species used	Nutrient plant uptake in plant biomass
	N	P			
Pilot study	3.44	0.24	Nutrient solution	*Phragmites karka*	42.12% N; 28.92% P
	1.56	0.23	Nutrient solution	*Lepironia articulata*	17.43% N; 26.08% P
Headley (2004)			Nursery runoff	*Phragmites australis*	41–54% N; 36–63% P in above-ground biomass, 24–30% N; 36–39% P in below-ground biomass
Toet (2003)	0.53	0.082	Sewage effluent	*Phragmites australis*	37–42% N; 22–40% P
Browning and Greenway (2003)	0.8–7.3			*Baumea articulata, Carex fascicularis, Philydrum lanuginosum and Schoenoplectus mucronatus*	11% N; 3% P
Greenway and Woolley (2001)	0.72–1.93	0.22–0.68	Secondary effluent	*Typha domingensis, Schoenoplectus validus, Eleochrais equisetina, Eleocharis sphacelata*	14.5–80% N, 24–80% P
Kantawanichkul et al. (2001)	11.2		Livestock effluent	*Cyperus flabelliformis*	7–9% N
					0.414–0.491 $g\,N\,m^{-2}\,day^{-1}$ in above-ground biomass 0.197 $g\,N\,m^{-2}\,day^{-1}$ in below-ground biomass
Lim et al. (2001)	4.5		Septic tank effluent	Cattail *Typha sp.*	2.6 $kg\,ha^{-1}\,day^{-1}$ (50% N)
Tanner (2001)	3.0	1.0	Dairy farm wastewaters	Soft-stem bulrush *Schoenoplectus tabernaemontani*	
Okurut (2000)	7.1	0.24	Septic tank effluent	*Cyperus papyrus*	14.95–21.91 $mg\,g^{-1}$ N; 5.61–5.95 $mg\,g^{-1}$ TP in above-ground tissue.
Okurut (2000)	10.4	0.26	Septic tank effluent	*Phragmites australis*	19.96–22.16 $mg\,g^{-1}$ TN; 7.90–10.05 $mg\,g^{-1}$ TP in below-ground biomass
Koottatep and Polprasert (1997)	3.0		Septic tank effluent	Cattail *Typha angustifolia*	43% TN (31% in leaf, 10% in stem, 2% in root)
	1.08	0.229	Sewage effluent	*Typha latifolia*	
Greenway (1997)	3.68	1.997			64.8% N
Tanner et al. (1995)	1.5–14	1.3–3.2	Dairy farm wastewaters	Soft-stem bulrush *Schoenoplectus tabernaemontani*	

50% of the nitrogen was stored in the leaves of cattail plants whereas a maximum rate of total nitrogen accumulation in plant biomass at 80% was recorded by Greenway and Woolley (2001).

One of the major sink for phosphorus in most wetlands is the soil (Kadlec and Knight, 1996). Most of the Phosphorus component may fix within the soil media (Brix, 1987) in the tanks. However, the nutrient stored in the substrate in this pilot study was not significant, probably due to the short experimental period. Higher plant growth was achieved under higher nutrient loading.

A higher nutrient removal of 52.6% N and 24.7% P was achieved by the Tube Sedge pilot tank system at a higher nutrient loading compared to 36.7% N and 12.7% P at a lower nutrient loading in this study. The plant nutrient uptake rate for the Tube Sedge was 14.1% N and 12.0% P compared to 17.4% N; 26.1% P at a higher concentration. A pilot study treating nursery runoff in Australia showed that Nitrogen removal increase with higher loading rates, until a threshold at 2.75 $g\,N\,m^{-2}\,day^{-1}$ (Headley, 2003).

4.3. Conclusion

The nutrient removal efficiencies in the pilot study tanks except for nitrogen removal by the Common Reed were lower than in the section of wetland cells in the field. The overall nutrient removal efficiency in Upper North wetland cells of the Putrajaya Wetlands was satisfactory. The field study also showed that nutrient removal efficiencies improved, as expected, with a longer flow length from UN6 to UN1. However, wetland size is largely predetermined by land availability and land value, although it should be determined by the targeted pollutant removal. The study confirms that the selected plants are suitable for a treatment wetland in a tropical climate. However, long term monitoring and maintenance is crucial to ensure the performance of the wetlands.

Acknowledgements

The authors would like to acknowledge Putrajaya Corporation (Perbadanan Putrajaya) for their kind permission to carry out this research in Putrajaya Wetlands; as well as Wetlands International-Malaysia and the Ford Motor Company Malaysia Conservation and Environmental Grants 2001, and University Putra Malaysia Fundamental Research Grants 2002 for their financial support.

References

Ariffin, S.J., 1998. Putrajaya and Putrajaya Lake development. In: Proceedings of Seminar on Putrajaya Lake Water Quality Management. Selangor Darul Ehsan, pp. 1–30.

Blamire, J., 2003. Quantitative analysis Kjeldahl method. ⟨http://www.brooklyn.cuny.edu/bc/ahp/SDKC/Chem/SD_Kjeldahl-Method.html⟩ Accepted on 13 June 2006.

Breen, P.F., 1990. A mass balance methods for assessing the potential of artificial wetlands. Water Research 24 (6), 689–697.

Brix, H., 1987. Treatment of wastewater in the rhizosphere of wetlands plants—the root-zone method. Water Science and Technology 19 (10), 107–118.

Browning, K., Greenway, M., 2003. Nutrient removal and plant biomass in a subsurface flow constructed wetland in Brisbane, Australia. Water Science and Technology 48 (5), 183–189.

Bureau of Nutritional Sciences Ottawa, 1983. Dry ash method—a rapid method for the determination of sodium and potassium. Health Protection Branch Laboratories, Ottawa, Ont., Canada. ⟨http://www.hc-sc.gc.ca/fn-an/res-rech/analy-meth/chem/reg_determination_sod_pot-reg_dosage_sod_pot_e.html⟩. Accessed on 13 June 2006.

Ceballos, B.S.O., Olliveira, H., Meira, C.M.B.S., Konig, A., Guimaraes, A.O., Souza, J.T., 2001. River water quality improvement by natural and constructed wetland systems in the tropical semi-arid region of Northeastern Brazil. Water Science and Technology 44 (11–12), 599–605.

Cooper, P.F., Job, G.D., Green, M.B., Shutes, R.B.E., 1996. Reed beds and constructed wetlands for wastewater treatment. WRc plc, Swindon, UK, p. 179.

Crites, R.W., 1994. Design criteria and practice for constructed wetland. In: Bavor, H.J., Mitchell, D.S. (Eds.), Wetland System in Water Pollution Control. Water Science and Technology, vol. 29(4), pp. 1–6.

Davison, L., Headley, T., Edmonds, M., 2001. On-site domestic wastewater treatment by reed bed in the moist subtropics. Water Science and Technology 44 (11–12), 353–360.

Greenway, M., 1997. Nutrient content of wetland plants in constructed wetlands receiving municipal effluent in tropical Australia. Water Science and Technology 35 (5), 135–142.

Greenway, M., Woolley, A., 2001. Changes in plant biomass and nutrient removal over 3 years in a constructed wetland in Cairns, Australia. Water Science and Technology 44 (11–12), 303–310.

HACH, 1989. Water Analysis Handbook. HACH Company, Loveland, CO, USA.

Headley, T.D., 2003. A study of the effect of reed beds on the nutrient concentrations in nursery runoff. Dissertation, Ph.D. Thesis, Lismore, NSW, Australia, p. 170.

Headley, T.D., 2004. Removal of nutrients and plant pathogens from recycled nursery runoff using horizontal subsurface flow constructed wetlands. Ph.D. Thesis, School of Environmental Sciences and Management, Southern Cross University, Lismore, NSW, Australia, 170pp.

Kadlec, R.H., Hey, D.L., 1994. Constructed wetlands for river water quality improvement. Water Science and Technology 29 (4), 159–168.

Kadlec, R.H., Knight, R.L., 1996. Treatment Wetlands. CRC Press Inc., Boca Raton, FL, USA.

Kantawanichkul, S., Neamkam, P., Shutes, R.B.E., 2001. Nutrient removal in a combined system: vertical vegetated bed over horizontal flow sand bed. Water Science and Technology 44 (11), 137–142.

Kao, J.T., Titus, J.E., Zhu, W.-X., 2003. Differential nitrogen and phosphorus retention by five wetland plant species. Wetlands 23 (4), 979–987.

Koottatep, T., Polprasert, C., 1997. Role of plant uptake on nitrogen removal in constructed wetlands located in the tropics. Water Science and Technology 36 (12), 1–8.

Lim, P.E., Wong, T.F., Lim, D.V., 2001. Oxygen demand, nitrogen and copper removal by free water surface and subsurface flow constructed wetlands under tropical conditions. Environment International 26, 425–431.

Okurut, T.O., 2000. A pilot study on municipal wastewater treatment using a constructed wetland in Uganda. Dissertation, Ph.D. Thesis, Wageningen University and International Institute of Infrastructural, Hydraulic and Environmental Engineering. Delft, the Netherlands, p. 170.

Rogers, K.H., Breen, P.F., Chick, A.J., 1991. Nitrogen removal in experimental wetland treatment systems—evidence for the role of aquatic plants. Journal of Water Pollution Control Federation 63 (7), 934–941.

Sekiranda, S.B.K., Kiwanuka, S., 1998. A study of nutrient removal efficiency of *Phragmites mauritianus* in experimental reactors in Uganda. Hydrobiologia 364, 83–91.

Shutes, R.B.E., 2001. Artificial wetlands and water quality improvement. Environment International 26 (5–6), 441–447.

Tanner, C.C., 2001. Growth and nutrient dynamics of soft-stem bulrush in constructed wetlands treating nutrient—rich wastewaters. Wetlands Ecology and Management 9, 49–73.

Tanner, C.C., Clayton, J.S., Upsdell, M.P., 1995. Effect of loading rate and planting on treatment of dairy farm wastewaters in constructed wetlands—I. Removal of oxygen demand, suspended solids and faecal coliforms. Water Research 29 (1), 17–26.

Tanner, C.C., Kadlec, R.H., Gibbs, M.M., Sukias, J.P.S., Nguyen, M.T., 2002. Nitrogen processing gradients in subsurface-flow treatment wetlands—influence of wastewater characteristics. Ecological Engineering 18, 499–520.

Toet, S., 2003. A treatment wetland used for polishing tertiary effluent from a sewage treatment plant—performance and processes. Dissertation, Ph.D. Thesis, University Utrecht, the Netherlands, 201pp.

White, G.C., Small, I.C., Bek, P.A., 1994. Carcoar wetland—a wetland system for river nutrient removal. Water Science and Technology 29 (4), 169–176.

Wiese, R., Raft, S., White, G., 1998. Detailed component design. In: Young, R., White, G., Brown, M., Burton, J., Atkins, B. (Eds.), The Constructed Wetlands Manual, vols. 1 & 2. Department of Land and Water Conservation, NSW, Australia, pp. 321–355.

Journal of Environmental Management 88 (2008) 318–325

Journal of Environmental Management

www.elsevier.com/locate/jenvman

An online water quality monitoring and management system developed for the Liming River basin in Daqing, China

Wei Yang, Jun Nan, Dezhi Sun*

School of Municipal and Environmental Engineering, Harbin Institute of Technology, 202 Haihe Road, Nangang District, Harbin 150090, China

Received 15 March 2006; received in revised form 2 February 2007; accepted 10 March 2007
Available online 25 April 2007

Abstract

This paper describes an online water quality monitoring and management system that was developed by combining a chemical oxygen demand sensor with an artificial neural network technology and a virtual instrument technique. The system was used to model the hydrological environment of the Liming River basin in Daqing City, China, in an effort to maintain the water quality in this basin at a level compatible with the status of Daqing City as a scenic resort. Operation of the system during the past 2 years has shown that an optimal allocation of water (including water released from an environmental reservoir to mitigate pollution events) could be achieved for the basin using the information gathered by the system; using mathematic models established for this system, the quantity of water released from the reservoir is adequate to improve the overall water environment. The results demonstrate that the system provides an effective approach to water quality control for environmental protection.

Keywords: Online monitoring; Water environment management; Hydrological environmental modeling; One-dimensional advection–diffusion

1. Introduction

Many urban rivers in China, and particularly scenic rivers, have been polluted by overland runoff from point and non-point sources (Xu et al., 2004; Dong et al., 2004; Deng, 2003). Accidental pollution has often occurred, and sometimes identification of water pollutants and polluters was not possible because water samples could not be obtained in a timely manner (Chai et al., 2004). For example, fish mortality occurred overnight in one incident and was only detected the next morning, after the contaminated water had already disappeared (Bode and Nusch, 1999). In China, online monitoring installations have been constructed for several large rivers, including the Huanghe River (Zhao, 2004), the Huaihe River (Chen et al., 2003), and the Haihe River (Meng, 2002), to provide real-time information to guide environmental protection decision-makers. Although considerable progress has been made in recent years to develop an online water quality monitoring capability, these installations still only complement laboratory testing, which is not yet a fully viable alternative (Drage et al., 1998).

For most medium and small rivers, few of the hydrological stations are well-equipped, and the apparatus that are being used are outdated and cannot satisfy the requirements of detecting and responding to pollution events. Some researchers have investigated integrated water quality models (Richards et al., 1996; Ning et al., 2001; Beck, 2005; Lindenschmidt et al., 2005) and environmental management systems based on hydrologic modeling (Chau et al., 2002; Mujumdar and Saxena, 2004; Zacharias et al., 2005), but these systems are not connected with any online monitoring system. Even in emergency cases of water pollution, no feasible management scheme can be worked out in a timely manner (Thoms and Swirepik, 1998; Rauch and Harremoës, 1999; Huang and Xia, 2001; Quinn, 2003). These problems justify the development of an online water quality monitoring and management system that can provide an early warning of water-pollution events. In recent years, the Chinese government has paid much

*Corresponding author. Tel.: +86 451 86283066; fax: +86 451 86412596.
E-mail address: sdzlab@126.com (D. Sun).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.03.010

attention to monitoring and management of the country's water environment. The online water quality monitoring and management system that has been implemented for the Liming River basin in Daqing is one example of the resulting government-funded programs.

This paper describes the Liming online water quality monitoring and management system, which uses modern data transmission and artificial neural network (ANN) techniques to monitor the river's water environment and hydrological–environmental models to forecast the potential environmental water demand. This combination of techniques allows optimal allocation of water using information acquired from the monitoring system and estimates from the water environment models.

2. Background information on the Liming River basin

Fig. 1 illustrates 37 km of the Liming River in the eastern part of Daqing City of China's Heilongjiang province. It is one of the six major streams in this area that are managed for flood prevention and scenic purposes. In recent years, different sources of contamination have caused deterioration of the water quality in the river and other bodies of water, including lakes and reservoirs: oil-contaminated soil (from which oil is leached into the river by overland runoff and percolation through the soil), domestic sewage, and wastewater produced by oil-extraction plants. Many measures have been taken to improve water quality, including the construction of a wastewater treatment plant for the removal of oil pollutants from surface runoff and accidental oil spills, and the construction of an underground sequencing batch reactor with aerated sludge facilities used to treat domestic sewage concentrated from several geographically distinct locations. These measures have effectively controlled pollutant sources to some degree, but organic pollutants and a lack of clean water in the Liming River are both responsible for poor water quality. To help resolve this problem, an environmental reservoir with a capacity of $0.74 \times 10^8\,m^3$ has been built north of Daqing City to provide a source of clean water. Water can be released from this environmental reservoir to improve water quality in the Liming River as a result of dilution and flushing effects. However, it was necessary to develop an online monitoring and management system for the Liming River basin to coordinate the release of water from seasonal lakes and reservoirs and to assess the assimilative capacity of the river and thus, improve our ability to manage water quality.

3. Water quality monitoring

3.1. Configuration of the water quality monitoring system

In order to provide an early warning when water quality in the river drops below an acceptable level, five monitoring stations were installed along the river, and one central control station was established at the Daqing Flood Prevention Distribution Center (Fig. 1). Water management software was installed on a computer at the Center to monitor such parameters as flow rate, total organic matter, total petroleum hydrocarbons, and total suspended solids. In this paper, we have chosen chemical oxygen demand (COD) as the water quality parameter used to represent total organic matter, since COD data are available from the online monitoring stations.

3.2. Data transmission process

Fig. 2 illustrates the existing signal-transmission network for data from the monitoring stations. The system at each

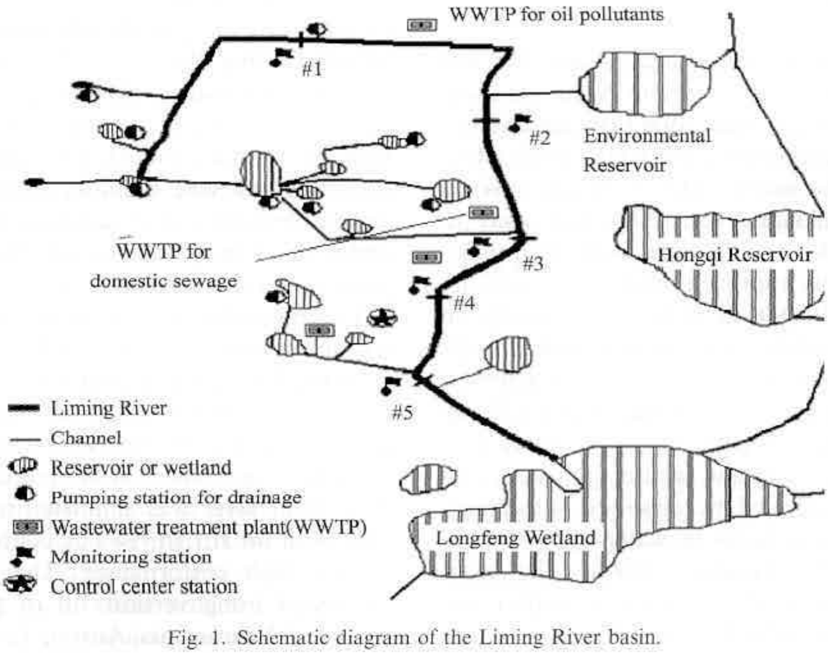


Fig. 1. Schematic diagram of the Liming River basin.

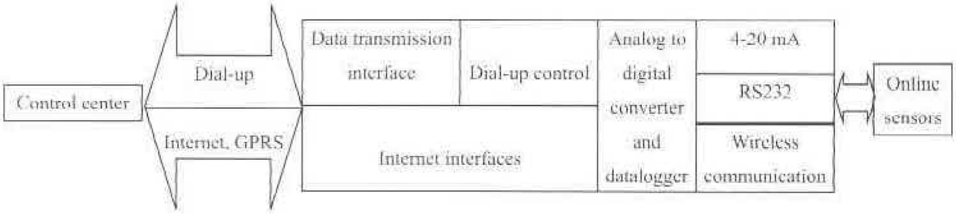


Fig. 2. Schematic diagram of the communication pathways in the monitoring system.

monitoring station includes data links using 4–20-mA power cables, RS232 connections, and wireless communication. Real-time analog signals carried by the power cables are obtained from each monitoring station through a series of water quality sensors. A programmable logic controller is used to convert the analog signals into digital signals, and then dataloggers at each monitoring station read these signals through an RS232 interface. In addition, general packet radio service was adopted; this service relies on retransmission and data integrity protocols to ensure that data packets transmitted by radio do not deteriorate or become lost. This technique can be used to greatly improve the reliability of data transmission (Lindemann and Thummler, 2003). Communication between dataloggers, the monitoring stations, and the control center is mainly carried out by means of the short-message service (SMS) technology complemented by a dial-up connection for use when this service is unavailable. The control center sends out a request to each station every 30 min. The station packages its monitoring data once per 30 s and transfers a compilation of this data to the control center when it receives the request from the control center. In addition, the control center can be connected to the Internet by means of a dial-up connection at any time to publish information and share it with the public.

3.3. Online monitoring using an ANN

The online water quality monitoring system that was developed for the Liming River basin used standard techniques for monitoring flow rate, total suspended solids, and total petroleum hydrocarbons, using instruments that are readily available on the market (Hu and Yang, 2004). A "soft" measurement technique for COD was used to overcome the drawbacks encountered with traditional online instruments. In this approach, multiple sensors are combined to evaluate COD in terms of changes in ultraviolet (UV) and visible (Vis) spectra and in pH. Because almost all organic matter exhibits characteristic absorbance in the range of 215–316 nm (especially in 254 nm), UV–Vis absorbance spectroscopy is widely used to characterize dissolved organic matter in water. In addition, pH, which is affected by dissolved substances, can sensitively indicate variations in water quality (Benjathapanun et al., 1997; Grattan, 1998). The most commonly used computational algorithm, back-propagation, was used in an ANN model to parameterize the non-

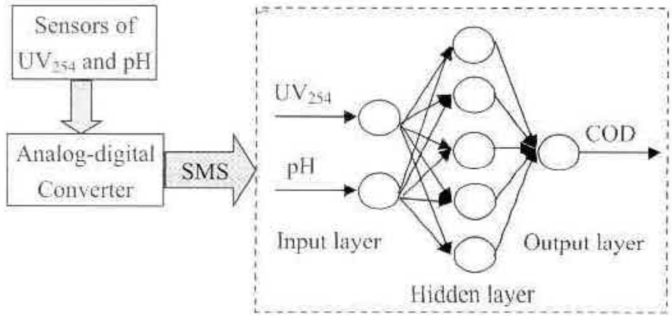


Fig. 3. Illustration of the ANN model used to convert absorbance of UV and Vis light (UV_{254}) and pH into an estimated COD value.

linear relationship between two water quality parameters (UV_{254} and pH) and COD. With UV_{254} and pH used as the inputs for the ANN and COD used as the output, a two-layer feed-forward neural network was created (Fig. 3).

In the process of training, one iteration of this algorithm can be written as follows:

$$x_{k+1} = x_k - \alpha_k g_k, \tag{1}$$

where x_k is a vector array of current weights and biases, x_{k+1} is the value used as the input in the next iteration, g_k is the current gradient, and α_k is the learning rate.

Starting with an initial learning rate ($\alpha_k = 0.1$), an initial momentum constant ($m_c = 0.9$), five hidden neurons, and an error rate of 0.01, the weights and biases are iteratively adjusted using the momentum method to evaluate the network performance (Hill et al., 1993), and the goal is to minimize the mean squared error (MSE) between the network outputs and target outputs during the training process. If the MSE becomes smaller than the training goal and stable at the end of each learning epoch by adjusting α_k and m_c, then the parameter set can be determined and post-processing can be carried out.

This algorithm is realized in a virtual instrument layer. In the process illustrated in Fig. 3, analog signals (data from the UV_{254} and pH sensors) are directly converted into initial digital signals using an analog to digital converter, and then the digital signals are transferred through SMS to the virtual instrument layer to quantify COD. The virtual instrument layer was simulated in hardware (a VXI bus card with an IEEE1394 bus controller) that was selected for its high performance. The required software was developed using version 7.0 of the LabVIEW software (National Instruments, Austin, Texas), which facilitates the

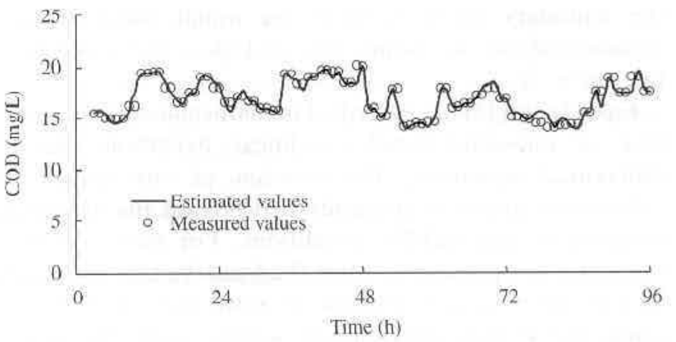


Fig. 4. Results of the training and learning stages for the ANN used to estimate COD (with pH in the range of 4.3–6.2).

development of virtual instruments and produces software that can be run on several types of computers and operating systems without changing the source code (Tanner and White, 1996; Torán et al., 2004). The complete virtual instrument was designed using the G language.

Once the models have been embedded in the computation software, the computers and instruments used for measurement and control are integrated through the virtual instrument layer. During the learning and training stage, data obtained from historical records (2004–2005) provided by the Daqing Flood Prevention Distribution Center was used. During the application stage, Fig. 4 shows that the estimated COD values were in good agreement with the observed values for pH values ranging from 4.3 to 6.2. The calculated values and measured values were fitted using version 11.5 of the SPSS software (SPSS Inc., Chicago, Illinois). The correlation coefficient was 0.924, which suggests that the model was acceptable for application in the Liming River basin and that the determination of COD values using UV_{254} and pH data could be used to rapidly perform online real-time measurements.

4. Hydrological–environmental modeling

4.1. Water quantity submodel for the river

In this paper, submodels for the environmental water requirements and of rainfall-runoff forecasting are included in the overall water quantity model. The environmental water requirements include modeling of the water required for assimilation of polluted river water, for evaporation, and for conservation of groundwater.

4.1.1. Quantity of water required for assimilation of polluted river water

The assimilative capacity of a river is defined as its capacity to "digest" pollution by means of biological activity and physical purification, both of which depend on the uses of the body of water and the quality standards adopted by the management agency (Lee and Wen, 1996). Calculation of the quantity of water needed for assimilation of polluted river water requires calculation of the inverse of assimilative capacity; that is, it represents the minimum quantity of water needed to permit self-purification and dilution of the pollutants, including the quantity of water diverted from other water conservation projects when the river water has been badly polluted. The quantity of water needed for assimilation of polluted water in the Liming River can be calculated as follows:

$$Q_1 = \frac{(q_1 c_1 + Q_0^* C_0)\exp(-kv/x_1) - (Q_0^* + q_1)C_N}{C_N - C\exp(-kt)}, \qquad (2)$$

where Q_0^* is the flow rate from upstream ($m^3\,s^{-1}$), C_0 is the concentration of pollutants from upstream ($mg\,L^{-1}$), q_1 is the flow rate from the pollutant sources ($m^3\,s^{-1}$), c_1 is the concentration of pollutants from the pollutant sources ($mg\,L^{-1}$), Q_1 is the flow rate released from the environmental reservoir ($m^3\,s^{-1}$), C is the concentration of pollutants from the environmental reservoir ($mg\,L^{-1}$), C_N is the standard for water quality ($mg\,L^{-1}$), k is a degradation coefficient (1/day), x_1 is the length of the river (m), and v is the average flow velocity ($m\,day^{-1}$).

4.1.2. Quantity of water needed for evaporation

Evaporative losses are an important part of a river's environmental water demand, especially during the summer. Evaporation of river water decreases the quantity of river water, without greatly affecting the quantity of pollutants in the river. Thus, evaporative losses should be compensated for by water diversion from other bodies of water using the following formula:

$$Q_2 = \begin{cases} 0.1A(E-P), & E > P, \\ 0, & E \leqslant P, \end{cases} \qquad (3)$$

where Q_2 is the water demand created by evaporation ($10^4\,m^3$), A is the average surface area of the water (km^2), P is the monthly rainfall (mm), and E is the monthly evaporation (mm).

4.1.3. Quantity of water needed for conservation of groundwater

Leakage from the river occurs when the water table is lower than the river water, and can be another important environmental water demand. Leakage losses can be calculated using the following equation:

$$Q_3 = k_1 A, \qquad (4)$$

where Q_3 is the annual loss of river water to leakage ($m^3\,yr^{-1}$) and k_1 is the leakage coefficient ($m\,yr^{-1}$).

4.1.4. Calculation of runoff from precipitation

The Soil Conservation Service curve number (CN) runoff-estimation approach (Soil Conservation Service, 1972) was used, with some modifications, to calculate the runoff from precipitation. This method uses the following equation (Smith and Williams, 1980):

$$Q_4 = (P - 0.2S)^2/(P + 0.8S), \qquad (5)$$

where Q_4 is the runoff amount (mm), P is the rainfall depth (mm), and S is the maximum retention estimated for dry-soil antecedent moisture condition I (AMC-I), and can be calculated using the following equation:

$$S = (25\,400/CN) - 254, \tag{6}$$

where CN is the curve number used for the AMC-I soil moisture condition.

To compute the runoff amount from rainfall depth as a function of these initial abstractions and soil water storage, S is estimated from the actual water content in the upper soil layers and from the CN that characterizes the soil and its vegetation or other cover. To further improve accuracy, CN is calibrated from the observed data for wet, average, and dry antecedent soil moisture conditions. When rainfall data is used, the CN parameter is calibrated by combining the analysis of observed runoff hydrographs with the rainfall breakthrough curves for the same runoff events. As long as a suitable CN (here, one that falls within the 90% confidence interval for the calibration data, Bhunya et al., 2003) is obtained, runoff can be forecasted with considerable accuracy.

4.2. Hydrodynamic submodels

The motion of bodies of water in open channels can be described using the Saint-Venant equations, which express the conservation of mass and momentum (Luis and José, 2004). Conservation of mass leads to a continuity equation, which establishes balances between the rate of rise in the water level and the wedge and prism storage components (Singh and Woolhiser, 2002). Conservation of momentum leads to a dynamic equation that establishes balances between inertia, diffusion, gravity, and frictional forces. The governing continuity and momentum equations can therefore be written as

$$\frac{\partial Q}{\partial x} + \frac{\partial A_1}{\partial t} = q, \tag{7}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial(\alpha Q^2/A_1)}{\partial x} + gA_1\frac{\partial h}{\partial x} + \frac{gQ|Q|}{C_2 A_1 R} = 0, \tag{8}$$

$$h(x)|_\zeta = h_1, \tag{9}$$

$$Q(x)|_\zeta = q_1, \tag{10}$$

$$h(t), Q(t)|_{t=0} = h_0, Q_0, \tag{11}$$

where Q is the flow rate ($m^3 s^{-1}$), A_1 is the cross-sectional flow area (m^2), x is the horizontal distance (m), t is the time (s), q is the lateral inflow or outflow (positive for inflow and negative for outflow; $m^3 s^{-1} m^{-1}$), α is the momentum correction coefficient, g is the gravitational acceleration ($m\,s^{-2}$), h is the water surface elevation above datum (m), $R = A/P_w$ is the hydraulic radius (m), P_w is the wetted perimeter (m), C is the de Chezy resistance coefficient, ζ denotes the boundary, h_1 is the water surface elevation above the datum at the boundary (m), q_1 is the flow rate at the boundary ($m^3 s^{-1}$), h_0 is the initial water surfac elevation above the datum (m), and Q_0 is the initial flo rate ($m^3 s^{-1}$).

Eqs. (7) and (8) are described in mathematical terms as pair of one-dimensional non-linear hyperbolic parti differential equations. The solution of any system o differential equations generally depends on the existence uniqueness, and stability conditions. For many applica tions, it is not possible to solve the Saint-Venant equation analytically, but it is possible to solve them numericall using the Preissmann implicit scheme with the mode boundaries represented by flow–time, stage–time, o stage–flow relationships (Crossley and Wright, 1997).

This model can also be used to generate the necessar input data for simulating water quality.

4.3. Water quality submodels

The transport of pollutants is modeled using a finite difference approximation to the one-dimensional advec tion–diffusion equation (Siegel et al., 1997):

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - v\frac{\partial C}{\partial x} - kc + S_0, \tag{12}$$

$$C(x)|_\zeta = c_1, \tag{13}$$

$$C(t)|_{t=0} = c_0, \tag{14}$$

where C is the pollutant concentration ($kg\,m^{-3}$), v is th cross-sectional average flow velocity ($m\,s^{-1}$), D is th diffusion coefficient ($m^2 s^{-1}$), S_0 is the source/sink tern (representing decay, growth, erosion, deposition, and othe processes; $kg\,m^{-1} s^{-1}$), and c_1 and c_0 are the boundary an initial concentrations in the river, respectively ($kg\,m^{-3}$).

Model boundaries are represented by concentration time or concentration–flow relationships. Pollutants ca also be added or removed from any point in the modelin process.

Water quality modeling can cover a wide range of value for the water quality parameters, including the concentra tions of solutes and suspended sediments. All the variable in Eq. (12) represent cross-sectional average quantitie This equation is solved using a novel implicit scheme base on a finite-volume central-difference scheme and a highl accurate Ultimate Quickest scheme (Yang et al., 2002).

4.4. Modeling procedure

The three submodels comprise the hydrological–envir onmental model. The integrated modeling is capable o predicting the water surface elevations, velocity, distribu tion of water quality parameters along the river, the river' assimilative capability, and the quantity of environmenta reservoir water allocated for the Liming River.

The hydrodynamic submodel can provide the water flo data needed for the construction of the water qualit submodel and the water quality submodel can be used t simulate water quality so that the results of water quantity modeling can be verified and optimized in a timely manner. If water quality in the river does not comply with the management agency's water quality standard, the water quantity modeling would be continued until the amount of released water required to produce a satisfactory water quality is achieved.

5. Results and discussion

It is necessary to divert water to improve water quality when pollution levels exceed the limits defined by the management agency. The quantity of water that must be released from the environmental reservoir to maintain an appropriate water quality is determined by the results of the hydrological–environmental modeling. In addition, the decision-making process requires that, in order to reduce the pressure on the clean water resource, the quantity of water released from the reservoir must be as little as possible to bring water quality in the river into compliance with the standard. In the decision-making process, the model is only used to calculate the release of clean water when the river's water quality exceeds the defined limit, and the calculation continues until the water quality complies with the standard: the quantity of water released for dilution of the pollution is thus defined. Fig. 5 shows several examples that illustrate the monitored water quality in 2004 and 2005. Fig. 6 shows an example of a water release schedule based on the online monitoring information. Both figures show that the system can effectively analyze water quality. Fig. 5 shows that the measured water quality remained well below the warning line (i.e., the COD level at which additional water must be released to maintain water quality at an acceptable level) in 2004 and on 5/9/2005, but exceeded the warning line on 30/8/2005, which indicates that water released from the environmental reservoir during the first three periods was higher than the amount required, and that some clean water was wasted. The information provided by the water management system thus supported a decision to reduce the amount of water released, and the required reduction in water flow was thus obtained using hydrological–environmental modeling. Water quality remained acceptable (i.e., COD remained below the warning line) when water flow was reduced by $0.5\,m^3/s$. Soon after this change, the trend line for water quality produced by the simulation remained gentle and the water release rate remained steady, indicating that no new source of water pollution had been detected and that the assimilative capacity of the river remained satisfactory.

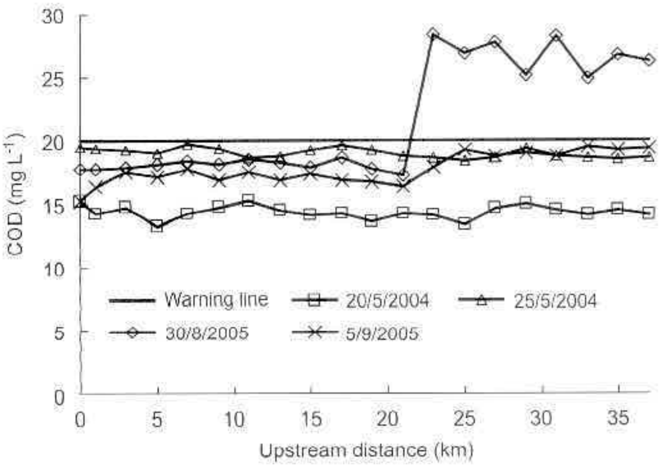


Fig. 5. Four examples of the water quality distributions along the Liming River after the release of water from an upstream environmental reservoir.

However, on 30 August 2005, there was a sharp increase in COD concentration in the river 20 km from the source of the Liming River. This suggests that a pollution incident occurred between monitoring stations #3 and #4 on the river. A survey revealed that this spike in COD was caused by an accident at the wastewater treatment plant in the upper reaches of the river: wastewater flowed directly into the river through a bypass valve. COD concentrations of the outflow from the wastewater treatment plant during the whole day are shown in Fig. 7. The accident lasted for 4 h, during which time raw wastewater flowed directly into the river. The results of the hydrological–environment modeling suggested that water flow should be increased by about

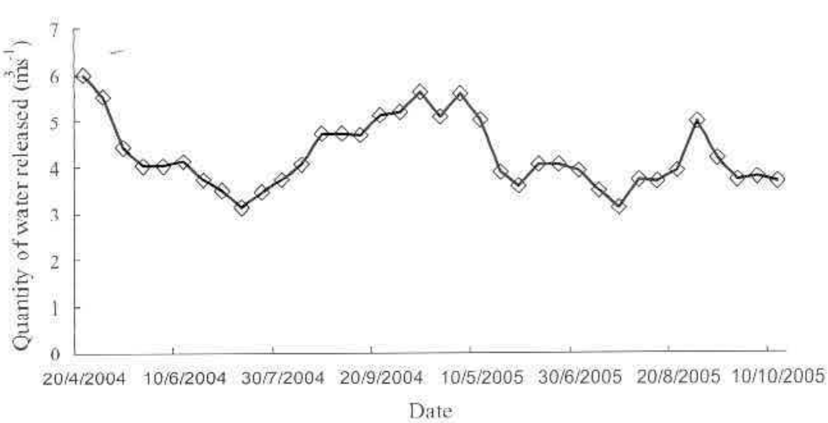


Fig. 6. Water release patterns from the environmental reservoir in 2004 and 2005.

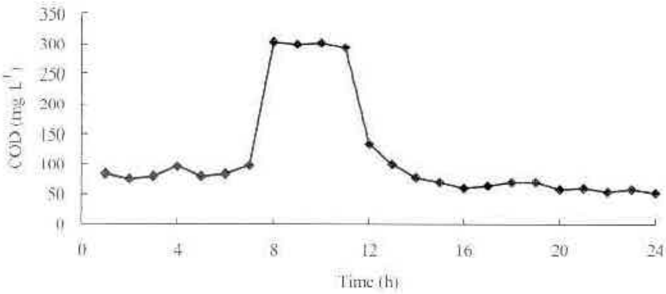


Fig. 7. COD concentrations in the outflow from the wastewater treatment plant on 30 August 2005.

$1.2\,m^3/s$, and water quality in the river was quickly improved by increasing the water flow to this level.

6. Conclusions

This paper describes an online water quality monitoring and management system developed for China's Liming River basin by combining an ANN-based model for sensing COD levels with a virtual instrument technique and using hydrological–environmental modeling to characterize the responses of the river to pollution loads and changes in the flow rate. The goal was to maintain the river's water quality at a level compatible with the status of Daqing as a scenic resort area of China. Using this approach, an optimal environmental water allocation was obtained, permitting a marked improvement in water quality and rapid responses to pollution incidents. The approach used a large quantity of real-time monitoring data as input for the hydrological–environmental model. The results of this approach show that the system provides an effective approach to control water quality for environmental protection purposes.

Acknowledgments

We thank the National 863 High-Tech Research Foundation of China (Grant no. 2003AA601090) for their financial support and the journal's anonymous referees for their valuable comments on an early draft of this manuscript.

References

Beck, M.B., 2005. Vulnerability of water quality in intensively developing urban watersheds. Environmental Modelling and Software 20 (1), 381–400.

Benjathapanun, N., Boyle, W.J.O., Grattan, K.T.V., 1997. Binary encoded 2nd-differential spectrometry using UV Vis spectral data and neural networks in the estimation of species type and concentration. IEEE Proceedings. Science, Measurement and Technology 144 (2), 73–80.

Bhunya, P.K., Mishra, S.K., Berndtsson, R., 2003. Approach to confidence interval estimation for curve numbers. Journal of Hydrologic Engineering 8 (4), 232–233.

Bode, H., Nusch, E.A., 1999. Advanced river quality monitoring in the RUHR Basin. Water Science and Technology 140 (10), 145–152.

Chai, C.G., Li, M., Wang, Y.H., 2004. Problems and suggestions of wat quality monitoring for accidental water pollution events. China Wat Resources 15, 52–53 (in Chinese).

Chau, K.W., Chuntian, C., Li, C.W., 2002. Knowledge manageme system on flow and water quality modelling. Expert Systems wi Applications 22, 321–330.

Chen, Z.D., Chen, D.Q., Li, L., et al., 2003. Dynamic monitoring l remote sensing techniques for the flood in Huaihe River in 2003. Chi Water Resources 12, 26–28 (in Chinese).

Crossley, A.J., Wright, N.G., 1997. Solution of the Saint Venant equatio through the use of Riemann based methods. In: The 27th Congress the International Association for Hydraulic Research, San Francisc CA, pp. 204–209.

Deng, J.M., 2003. The study on contaminated river bioremediatio Environmental Science and Technology S1, 55–57 (in Chinese).

Dong, X.D., Zhou, Q., Zhou, X.D., 2004. Technology and development pollution prevention and treatment on rivers and lakes in Chin Journal of Precious Metallic Geology 1, 27–30 (in Chinese).

Drage, B.E., Upton, J.E., Purvis, M., 1998. On-line monitoring micropollutants in the river Trent (UK) with respect to drinking wat abstraction. Water Science and Technology 38 (11), 123–130.

Grattan, K.T.V., 1998. Water quality monitoring—technological adva tages and potential developments in the use of fibre optic technolog Water Science and Technology 37 (12), 247–253.

Hill, T., Marquez O' connor, M., Remus, W., 1993. Artificial neur network models for forecasting and decision making. Internation Journal of Forecasting 10 (3), 5–15.

Hu, Y., Yang, Z.Y., 2004. A simple chemiluminescence method fo determination of chemical oxygen demand values in water. Talan (Oxford) 63, 521–526.

Huang, G.H., Xia, J., 2001. Barriers to sustainable wate quality management. Journal of Environmental Management 61 (1 1–23.

Lee, C.S., Wen, C.G., 1996. Application of multi-objective programmir to water quality management in a river basin. Journal of Enviro mental Management 47, 11–26.

Lindemann, C., Thummler, A., 2003. Performance analysis of the gener packet radio service. Computer Networks 41, 1–17.

Lindenschmidt, K.E., Hesser, F.B., Rode, M., 2005. Integrating wat quality models in the high level architecture (HLA) environmen Advances in Geosciences 4, 51–56.

Luis, F.G.L., José, S.A.D., 2004. Hydrodynamics and water quali modelling in a regulated river segment: application on the instrea flow definition. Ecological Modeling 173, 197–218.

Meng, X.Z., 2002. The development and application of remote sensin and GIS in Haihe river-basin. Haihe Water Resources 6, 46–48 (i Chinese).

Mujumdar, P.P., Saxena, P., 2004. A stochastic dynamic programmin model for stream water quality management. Sadhana 29 (5), 477–49

Ning, S.K., Chang, N.B., Yang, L., et al., 2001. Assessing pollutio prevention program by QUAL2E simulation analysis for the Kao-Pin River Basin, Taiwan. Journal of Environmental Management 61 (1 61–76.

Quinn, N.W.T., 2003. A decision support system for adaptive real-tim management of seasonal lakes in California. Environmental Modellin and Software 18, 503–511.

Rauch, W., Harremoës, P., 1999. Genetic algorithms in real time contr applied to minimize transient pollution from urban wastewat systems. Water Research 33, 1265–1277.

Richards, K., Sharp, M., Arnold, N., et al., 1996. An integrated approac to modelling hydrology and water quality in glacierized catchment Hydrological Processes 10 (44), 479–508.

Siegel, P., Mose, R., Ackerer, P., et al., 1997. Solution of th advection diffusion equation using a combination of discontinuo and mixed finite. International Journal for Numerical Methods i Fluids 24 (6), 595–613.

Singh, V.P., Woolhiser, D.A., 2002. Mathematical modeling of watershe hydrology. Journal of Hydrologic Engineering 7 (4), 270–292.

Soil Conservation Service, 1972. Estimation of direct runoff from storm rainfall. National Engineering Handbook. Soil Conservation Service. US Government Printing Office, Washington, DC (Chapter 10).

Smith, R.E., Williams, J.R., 1980. Simulation of the surface water hydrology. CREAMS—a field scale model for chemical, runoff and erosion from agriculture management system. USDA-Agricultural Research Service Conservation Research Report No. 26.

Tanner, A.H., White, N.M., 1996. Virtual instrumentation: a solution to the problem of design complexity in intelligent instruments. Measurement and Control 29, 165–171.

Thoms, M.C., Swirepik, J., 1998. Environmental flow management in New South Wales, Australia. In: Harding, J.K., Gilvear, D.G. (Eds.), Hydrology in a Changing Environment: Ecological and Hydrological Interactions. Wiley, Chichester, pp. 281–288.

Torán, F., Ramirez, D., Navarro, A.E., et al., 2004. Design of a virtual instrument for water quality monitoring across the Internet. Sensors and Actuators B76, 281–285.

Xu, W., Yuan, B., Sun, S.Y., et al., 2004. Progress of water pollution remediation techniques in city and town river. Journal of Guangdong University of Technology 21 (4), 85–90 (in Chinese).

Yang, L., Lin, B., Kashefipour, S.M., et al., 2002. Integration of a 1-D river model with object-oriented methodology. Environmental Modelling and Software 17, 693–701.

Zacharias, I., Dimitriou, E., Koussouris, T., 2005. Integrated water management scenarios for wetland protection: application in Trichonis Lake. Environmental Modelling and Software 20 (2), 177–185.

Zhao, W.Z., 2004. The development and application of water quality automatic monitoring in Huanghe River. China Water Resources 7, 27–29 (in Chinese).

Journal of Environmental Management 88 (2008) 326–347

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Environmental accounting as a management tool in the Mediterranean context: The Spanish economy during the last 20 years

Pedro L. Lomas*, Sergio Álvarez[1], Marta Rodríguez, Carlos Montes

Laboratorio de Socio-ecosistemas, Dpto. Ecología, c/Darwin, 2 Despacho C-201, Edificio Biología, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain

Received 2 March 2006; received in revised form 6 February 2007; accepted 10 March 2007
Available online 10 May 2007

Abstract

Although human presence is one of the main characteristics of the Mediterranean identity since ancient times, a false dialectic between conservation and social–economic development has emerged in recent decades. On the one hand, an economic growth policy is taken as the paradigm of social–economic development; on the other hand, there is a multi-scale conservation policy, in which natural protected areas, as patches of preserved nature, are used as one of the main tools to deal with the challenge of sustainability. The Mediterranean Basin is the habitat of many unique species and one of the 25 main biodiversity hotspots in the world, and as a consequence a strong conservation policy has been used to protect environmental values. At the same time, Mediterranean countries are deeply involved in promoting strong economic growth policies, which are not always compatible with environmental ones. In this paper, Spain has been studied as one model of this situation. Due to political reasons, Spanish economic growth and conservationist policies were pursued together during the last 20 years. As a result, Spain owns one of the largest networks of natural protected areas in Western Europe, and at the same time it has experienced one of the strongest periods of economic growths in the European and Mediterranean context during the 1980s and 1990s. An historical series of resource use in five annual periods in the last 20 years of conservation policy, and the effects on the preservation of natural capital have been investigated by means of the eMergy (spelled with an 'm') synthesis approach, which was used to characterize the flow of environmental services supplied by ecosystems, but not in monetary terms. This study shows that Spain is becoming less self-sufficient and more inefficient in resource use, comprehensively measured in eMergy terms. A large part of Spain's economy depends on imported goods and services, and most economic activities are based on tourist services and associated construction, which promotes intensification in the urban use of the territory and more intense environmental impacts and resource use intensification of those countries supplying the raw materials. The consequence is a decoupling of the Spanish economy from local environmental services and the increase of Ecological footprint of Spain, measured by means of eMergy-based indicators. In spite of the increase in number, area and associated budget of the natural protected areas and other conservation measures, the general sustainability of the nation is decreasing.

Keywords: eMergy synthesis; Spain; Sustainability indicators; Conservation policy

1. Introduction

1.1. Background

The recent millennium ecosystem assessment (MEA) Synthesis Report (MA (Milleninum Ecosystem Assessment), 2005) estimates that one third of the planet that has been altered for production purposes. This report shows that 50% of freshwater from rivers and lakes is eventually used by society, and that human activities produce more biologically available nitrogen than all natural cycles combined. Furthermore, this study estimates that 60% of the 24 great global ecosystems are experiencing degradation, and that extinction rates are increasing from 100 to 1000 times over the average estimated for geological time. In addition, they found that up to 20% of known species in many groups are disappearing. These figures are much

*Corresponding author. Tel.: +34914978008; fax: +34914978001.
E-mail address: pedro.lomas@uam.es (P.L. Lomas).

[1]Present address: Subdirección General de Relaciones Internacionales, Secretaría General Técnica, Ministerio de Medio Ambiente, Plaza de San Juan de la Cruz, s/n. 28071 Madrid, Spain.

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.03.009

worse than those calculated before MEA (Postel et al., 1996; Vitousek et al., 1986, 1997; Rojstaczer et al., 2001), and turn the ideas of "natural environment" or "wild nature" as isolated areas without human participation into a fantasy. In this context, the concepts of Noosphere (Verdnasky, 1945), a biogeochemical cycling entity dominated by human influences, and Anthropocene (Crutzen and Stoermer, 2000), a new geological era in which main biophysical processes that control global dynamics (ecosphere) are driven by humankind, emerge.

From this perspective, one paradigmatic case is the Mediterranean Basin, where relationships between humans and other (living and non-living) components of ecosystems can be traced to before Neolithic times (Grove and Rackham, 2003), and where many of the original forests had already been used 7000 years Before the Common Era (BCE) (Makhzoumi and Pungetti, 1999). Naveh and Liberman (1993) suggest that we should only speak about cultural landscapes in the Mediterranean context.

In the Mediterranean Basin, 52% of plants, 30.5% of vertebrates, 25% of mammals, 13% of birds, 61% of reptiles and 52% of amphibians are endemic species. Consequently, it is considered as one of the 25 most significant hotspots of biodiversity, paradoxically located in one of the most densely populated areas of the world (Myers et al., 2000; Cincotta et al., 2000). It is generally accepted that this ancient relationship between humanity and nature, which is based on combined exploitation systems (mainly agriculture, forestry and livestock) that adapt human cycles to natural ones reinforcing ecological processes, has promoted biodiversity and long-term sustainability (González Bernáldez, 1991; Pineda and Montalvo, 1995; Schmitz et al., 2001; de Miguel and Gómez-Sal, 2002; García and Montes, 2003).

However, since the 1950s, increases in mechanized farming, population growth and economic globalization have radically changed ancient agricultural, forestry and pastoral practices. Many socio-economic constraints, like agricultural subsidies (mainly European Union Common Agriculture Policy, CAP), rural–urban migration and abandonment of traditional practices and land have affected the historic agro-ecosystems (Grove and Rackham, 2003; Mulligan et al., 2004). These changes are being accelerated by the growth of commercial relations among countries and their socio-economic consequences.

1.2. A case study Spain as a social–ecological system

Spain could be considered a typical case presenting the characteristic pattern described in the previous Section. Because of its location and its geological history, Spain is a land of natural contrasts, especially lithologic (lime, siliceous and clayey soils) and climatic ones (Mediterranean and continental in the central-southern area, oceanic in the north, areas of dry subtropical climate in the south-eastern Spain, and specific climatic conditions on mountainous areas all over the country), which create a great variety of ecosystems, from deserts to Atlantic forests. Because of its history and location as a bridge between Europe and Africa, in the Mediterranean framework, Spain is also a land of social contrasts. With four different official languages (Catalonian, Galician, Castilian/Spanish and Basque, the latter being a non-Romance language) and many dialects, Spain is organised into 17 regions and two autonomous cities, each of them with its own government and institutional framework. Although it is the fifth most populous country in the European Union (EU), and its population density has grown considerably in the last century (from 36.79 inhabitants/km^2 in 1900 to 81.26 inhabitants/km^2 in 2000), Spain has the fourth lowest population density in the EU, so it may validly be considered a relatively rural country in the European context.

Similar to other European countries, many changes have affected traditional exploitation systems in the last decades in Spain. If we use the historical series of official statistics, there has been a loss of cultivated areas (percentage of total area has changed from 40.15% in 1980 to 35.43% in 2002), and a relative increase of irrigated lands (from 13.76% of total cultivated areas in 1980 to 19.35% in 2000). In contrast, Spain has suffered an increase of 2.07% in the item "other types", which includes infrastructures and cities (Ministerio de Agricultura, Pesca y Alimentación (MAPA (Ministerio de Agricultura and Pesca y Alimentación), 1991, 1998, On-line). In fact, road density reached 0.32 km/km^2 in 2001 (Ministerio de Fomento (MFOM (Ministerio de Fomento On-line-a), and road surface will be doubled by 2020 according to the new Infrastructures Plan 2005–2020 (MFOM (Ministerio de Fomento On-line-b). There has also been an increase of 1.65% in forest area, probably because of replacement of croplands by forests, which has been favoured under the CAP to reduce the so-called European Community's agricultural surplus. In addition, energy use is growing (Ministerio de Economia (MINECO), On-line; IEA (International Energy Agency), 2003; 1997), and as a result greenhouse gas emissions have grown 45% from 1990 to 2004, and Spain's emissions are 25.6% above the Kyoto Protocol agreements for the country (European Environment Agency (EEA (European Environment Agency), 2005; Observatorio de la Sostenibilidad en España (OSE (Observatorio de la Sostenibilidad en España), 2005). Furthermore, Spain has been transformed into a country devoted to the services sector (Tamames and Rueda, 2000), especially tourism and commerce. This sector was responsible for 61.2% of Spanish employment and 64% of the gross added value in 1999 (INE (Instituto Nacional de Estadística)), On-line; MINECO, On-line), although it only involved 31% of the working population and accounted for 45% of the Gross Domestic Product (GDP) in 1960 (Cuadrado, 1999). In contrast, during this time industry and agriculture have declined in importance for the Spanish economy (Cuadrado, 1999; Tamames and Rueda, 2000).

In addition, the Spanish economy has created an enormous pressure on aquatic ecosystems to satisfy demand for water in a country where water is relatively scarce. This pressure is the result of a policy based on satisfying demand instead of controlling it (Arrojo, 2001). Therefore, with a consumption of $530\,m^3$ of water/inhabitant/year (Ministerio de Medio Ambiente (MIMAM (Ministerio de Medio Ambiente), 2000), Spain is one of the highest *per capita* water-consuming in 15 countries of the EU (EU-15). It has the greatest number of dams per inhabitant and per unit area in the world, with more than 1150 large dams (World Commission on Dams (WCD (World Commission on Dams), 2000). Water use has become more intensified, with more than 3 400 000 ha of land under irrigation in recent years (Llamas, 2000). Furthermore, it is estimated that there are more than 75 aquifers that are overexploited or have serious salinization problems, 13 of which have been declared "provisionally overexploited" and two "overexploited" under the Spanish Water Act (MIMAM (Ministerio de Medio Ambiente), 2000). It is estimated that 60% of Spain's wetlands were already lost at the beginning of the 1990s (Casado and Montes, 1991), and 40% of rivers are polluted or seriously polluted (Prats et al., 2000).

Despite the changes and pressures of the last decades, 80–90% of EU-15 vascular plants can be found in Spain, 1500 of which are endemic in a worldwide context, and more than 500 are exclusively shared with Northern Africa. Spain is also the habitat to approximately 50% of the fauna species in EU-15, with more than 7.5% endemic (MIMAM (Ministerio de Medio Ambiente), 1999). Within the EU context, Spain, among other Mediterranean countries, is probably the region with the highest biological diversity.

Due to political and historical reasons, Spain has dealt with most of the processes of intensive industrialization and transformation into a country dedicated to services sector, common to Western European economies, in the last 20 years. In fact, Spain has received considerable EU funding for territorial cohesion, because a great part of its territory has been considered as a priority area to be supported economically within the EU. For these reasons, Spain probably constitutes one of the best laboratories in the Mediterranean world for assessing the effect of the acceleration of societal growth promoted by globalization, its effects on sustainability and the success of different strategies of environmental management adopted by governments.

1.3. Objectives

The main objectives of this paper are (1) to study patterns in the use of environmental goods and services flowing to the Spanish economy, and changes in these patterns over a historical series of 20 years in Spain, (2) to show the changes in patterns of consumption and trade that have promoted economic globalization in Spain in the past two decades, and (3) to study the success of Spanish natural conservation policies and management during the last 20 years, in relation to the preservation of the natural capital that maintains the Spanish economy.

2. Methods

To deal with trends of resource use in the context of these objectives, within the general framework of Ecological Economics (Daly, 1991; Goodland and Daly, 1996; Costanza, 1997; Costanza et al., 1997; Martínez-Alier, 1999), a biophysical valuation of Spain by means of eMergy synthesis (Odum, 1996; Hau and Bakshi, 2004; Brown and Ulgiati, 2004) has been performed for five annual periods (1984, 1989, 1994, 2000 and 2002).

For the purposes of this study, Spain has been considered as a social–ecological system, SES (Berkes and Folke, 1998), comprised of its territories in the Iberian Peninsula and the Balearic Islands, in the Mediterranean Basin (Fig. 1), with a land area of $498\,476\,km^2$ (IGN (Instituto Geográfico Nacional), 1996), which constitutes the second biggest country in terms of area in the EU, after France. Neither the Canary Islands nor the other African territories of Spain have been studied, because of their peculiar characteristics with respect to the rest of the country (distance, singular eMergy flows, different eMergy sources, etc.). To delimit borders, the continental shelf was defined by the area between 300 m of depth, and was estimated by the Spanish Oceanographic Institute (IEO) staff as $74\,037\,km^2$ (Fig. 1).

From the brief description of Spain in the previous section, a flow diagram (Fig. 2) has been drawn to characterize the Spanish SES as a kind of system picture or macroscopic view (Rosnay, 1979; Brown, 2004), allowing us to model interactions between economic and ecological systems in terms of eMergy flows, using energy symbols from Odum (1994). Symbols have been used in accordance with criteria from Odum (1996) to represent the Spanish environmental window or SES. Under these criteria, symbols are placed on the diagram in order of increasing transformity (a measure of quality in eMergy terms, as defined below), and consequently renewable sources and ecological systems are placed on the left and economic flows and components are placed on the right. An aggregated diagram and three-arm diagrams for each year of the study period are respectively presented in Figs. 3 and 4. These diagrams show explicitly the main input-output flows described in the following sections, linking flows from biogeochemical sources (sun, rain, earth cycle, tides, etc.) to those of social–economic processes (industry, commerce, imports, immigration, etc.) in order to provide environmental and social services to Spanish society.

From the flow diagram of Spain presented in Fig. 2 and the summary of this, presented in Fig. 3, most important flows and overview indexes have been calculated for each year studied with the same methodology given by the example shown in Tables 1(a–c), in order to obtain a view of the Spanish social–ecological system dynamics over 20 years (Table 2). According to the usual eMergy evaluation

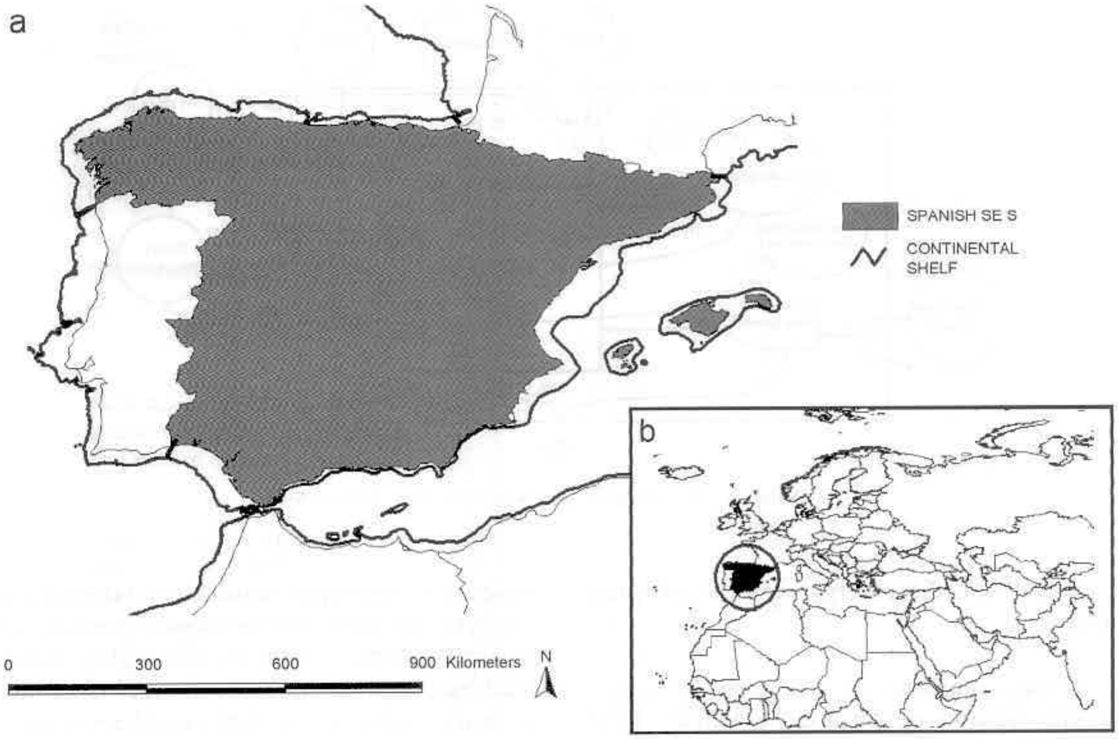


Fig. 1. Spanish social–ecological system in its context: (a) continental Spain and the continental shelf and (b) geographical context within Europe and the Mediterranean Basin.

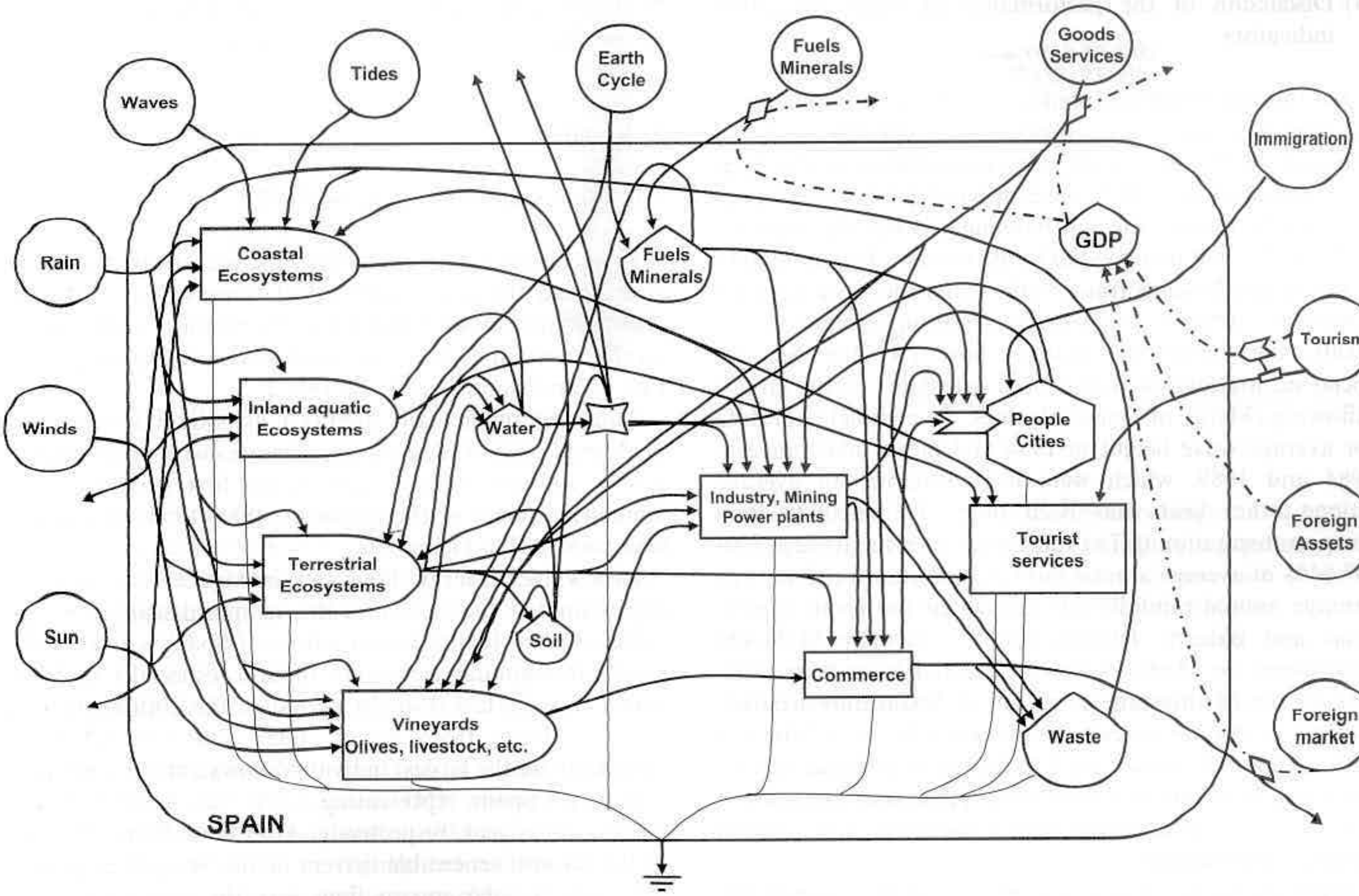


Fig. 2. Emergy flow diagram for Spain.

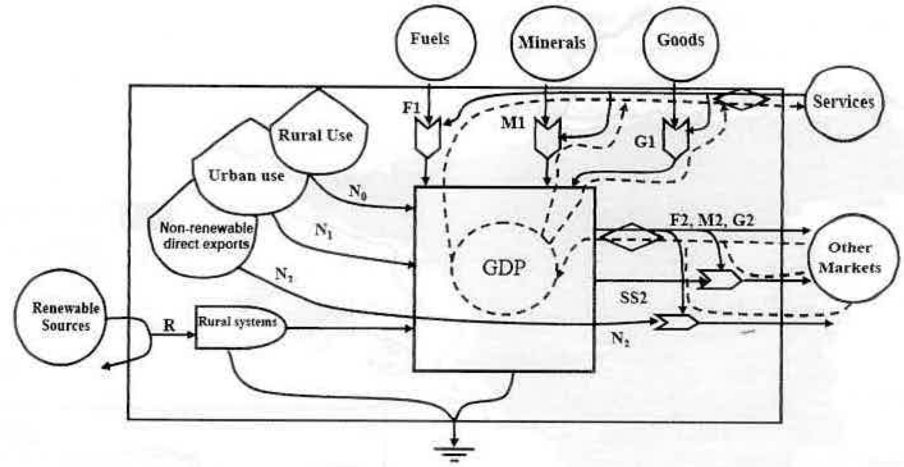


Fig. 3. Summary flow diagram for the main emergy flows in Spain.

procedure (Odum, 1996), the eMergy synthesis of Spain proceeds as follows:

(1) Drawing a flow diagram of Spain as a system (following the method established by Odum, 1996; Tilley and Swank, 2003).
(2) Preparation of an eMergy evaluation table.
(3) Calculation of main flows, storages and unit eMergy values or transformities.
(4) Discussion of the performance of main evaluation indicators.

All the transformities used in this work are updated to the new baseline (total contribution of eMergy to global processes = 15.83E + 24 sej/year) recalculated in the year 2000 (Odum et al., 2000) by multiplying unit eMergy values by 1.68 (the ratio of the new baseline between the past one: 15.83/9.44), as it is suggested by Brown and Ulgiati (2004).

In the new baseline framework, under the most accepted criteria in order to avoid double-counting (Odum et al., 2000), the renewable sources flow (R) for Spanish social–ecological system has been calculated as the largest inflowing eMergy of renewable ones. To complete our data for average wave height in 1984 and mean tidal range in 1984 and 1989, which were not available, an average among other years has been used. To calculate Real Evapotranspiration (ETa) and runoff rate, an average ETa (67.84% of average annual rainfall) and runoff (32.16% of average annual rainfall) rate calculated for Spain (peninsular and Balearic Islands) for 55 years by MIMAM (Ministerio de Medio Ambiente) (2000) have been used. Taking the Mediterranean nature of Spain into account, mature vegetation forests are assumed to have little net gain or loss of topsoil and it has been considered that harvested lands are net soil-losers. Thus, only the erosion rate in cultivated areas has been used to calculate topsoil energy contributions.

GDP at market prices has been used to calculate the eMergy-money ratios; although the use of Gross National Product (GNP) is very widespread. GDP instead of GNP is used to measure the economic activity within Spain regardless of the producer's nationality, following criteria used by Cialani et al. (2005) in order to avoid problems of measuring economic activity in eMergy terms.

In addition, previous eMergy synthesis data for Italy (Cialani et al., 2005) and a worldwide investigation of national economies (Brown, 2003) are used for comparison with other national economies and, especially, with the Mediterranean and European context of Spain (Appendix A and B).

3. Results

3.1. Main sources of the Spanish SES

In accordance with the system picture of Spain (Figs. 2 and 3) and the consequent calculations shown in Tables 1(a–c), main flows introduced to the Spanish social–ecological system (SES) for the studied years are represented by Fig. 5a and summarized in Table 2.

Total eMergy actually used (U), as potential investment in eMergy yield of the country, increases with an average of 3.77% annually with a peak in the first period (7.00% annually), except in the period of 1989–1994, in which it decreases 0.65% (Table 2).

Renewable eMergy flow (R) introduced to Spain is approximately unchanged at this temporal scale (Fig. 5a), although the Mediterranean nature of Spain is clear in the strong interannual variability of rain, especially in 1994, which was the last drought period of the 20th century in Spain. eMergy from waves, tides and rain (chemical potential) are the largest individual flows among renewable sources for Spain, representing 6.09E + 22, 5.66E + 22, and 4.62E + 22 sej/year, respectively. After these flows, the rank of the natural renewable drivers of the Spanish economy, according to solar emergy flow, was: the earth cycle, solar radiation absorbed, and kinetic energy of wind.

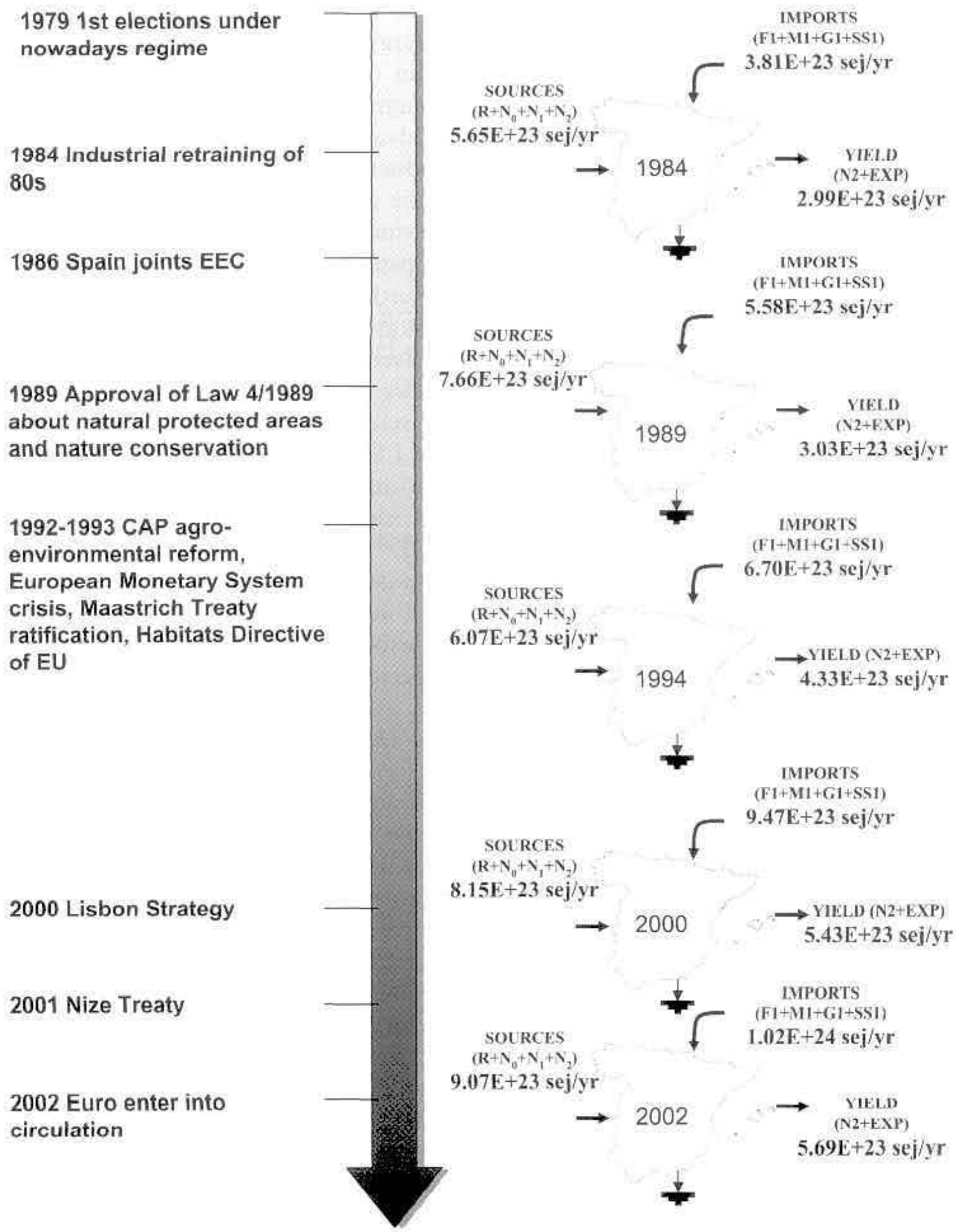


Fig. 4. Summary three arm flow diagram for the main flows in Spain contrasted with some of the main socio-economic events of the decades studied.

Participation of non-renewable eMergy flows from local Spanish sources (N) has increased from 5.09E + 23 sej/year in 1984 to 8.47E + 23 sej/year in 2002, although the annual increase rates have been reduced from 7.72% annually in the period 1984–1989 to 6.02% annually in the period 2000–2002 (Fig. 5a). Construction raw materials, like clay, calcium carbonate, sand and gravel are the largest individual N flows.

Table 1b lists main imported inputs in terms of eMergy flows (IMP) for 2000 in Spain. IMP eMergy flow has increased from 3.81E + 23 sej/year in 1984 to 1.02E + 24 sej/year in 2002. Oil could be highlighted among the most important imported goods and services eMergy flows, as it involves from 21% to 16% of total eMergy used in 1984 and 2002; worth mentioning are also leather and leather products, textiles, mechanical, transport equipment, and the increasing importance of natural gas and coal. In emergy to money terms, imported goods have become the most important item in this category. Among the export eMergy flows (EXP) could be highlighted petroleum-derived products, minerals and mechanical and transport equipment. In emergy to money terms, eMergy flows related to services exported, in general, and particularly tourism, are increasing as a result of the tourism model introduced in Spain beginning the 1960s. As Table 2 shows, eMergy flow of exports has decreased in relation to IMP in the studied period (IMP/EXP index increases from 1.42 in 1984 to 1.87 in 2002).

The percentage of R involved in U decreases 50% (from 7% to 3%) during the 20 years examined (Fig. 5b), so U is increasingly supported by IMP and N, with a growing role of concentrated eMergy sources (N_1) (concentrated against rural eMergy index grows from 13.64 in 1984 to 29.67 in 2002), and the increase of the relative weight of IMP, exceeding N in 1989–1994 period (Fig. 5a). This increase in IMP results in decreased self-sufficiency (fraction of eMergy actually used derived from home sources decreased from 59% to 46%; Fig. 5b) for Spain, except in the last period, where there is a relative increase of N_1 in U, so self-sufficiency is maintained. In this regard, Spanish dependence on imported energy sources is still increasing strongly, and more than 21% of U is imported oil in 1984 and 16% in 2002, so the purchased (non-free) component of the total economy is becoming more important, supporting the growth of the economy (Fig. 5a).

3.2. Some factors of scale to understand eMergy indicators

Taking into account the population factor of scale, the evolution of the potential standard of living in eMergy terms or eMergy use *per capita* shows an increase (Fig. 6a), although the growth rate of this indicator has continuously decreased from 6.69% annually in the 1984–1989 period to 2.68% in the last one, with the exception of the 1994–2000 period, in which it decreases 0.80%.

Taking into account the economic size factor of scale in terms of GDP, the average eMergy which is mobilized per monetary unit or the buying power in the Spanish SES (eMergy to money ratio; EMR) has decreased an average of 2.37% annually throughout the studied period (Fig. 6b), although this indicator experienced an increase of 3.48% annually in the period between 1994 and 2000.

Taking into account the territorial size factor of scale in terms of the Spanish SES area, territorial intensity of the eMergy actually used or Empower Density (flow of eMergy per unit area and time) has increased (except in the recession period in 1989–1994) in absolute terms, at an average annual rate of 3.77% over the whole period, but the annual growth rate within periods has decreased from the 7% annually in 1984–1989 period to 3.23% in 2000–2002 (Fig. 7c). If we consider that the Spanish economy is increasingly dependent on imports and non-renewable sources, territorial intensity of eMergy use depends mainly on the non-renewable fraction of Empower Density (96–97% of total Empower Density is non-renewable; Table 2), especially the imported fraction.

3.3. Interaction of Spanish SES with other systems

If the eMergy flow associated with imports and its significance was taken into proper account, the issue of trade would become crucial for Spain from an eMergy point of view. The most important component in the flow of purchased goods and services is the one for fuels and electricity.

An important aspect of trade is highlighted by the EMergy Exchange Ratio (EER, i.e. the ratio of EMR of Spain to EMRs of trade partner countries or the global economy), which shows the relative advantages and disadvantages for Spain in its international trade of products and resources. The EER for Spain with respect to the global economy has increased from 0.33 to 0.64 in the studied period, with a decline between 1994 and 2000 (Appendix B).

Furthermore, we can use macroeconomic value or eMergy price (emprice) to study the amount of eMergy received per monetary unit invested. As we can see in Tables 1(a–c), the highest values in renewable sources are related to waves and tides, with 1.97E+10 em\$/year and 1.83 em\$/year in the 2000, respectively; in non-renewable indigenous sources, those of calcium carbonate and sand and gravel, with 1.06E+11 em\$/year and 4.43E+10 em\$/year in 2000, respectively; oil and petroleum-derived products in imports, with 9.71E+10 em\$/year; and textiles and mechanical and transport equipment in exports, with 2.36E+10 and 2.62E+10 em\$/year, respectively.

3.4. The appropriation of eMergy by the Spanish SES

To get information about the appropriation of resources by the Spanish system (Raugei et al., 2004), a comparison of U with emergy purchased by the national economy or eMergy yield ratio (EYR; Fig. 7a) has been used. The EYR decreases an average of 0.90% annually (except in the last period 2000–2002, in which it increases 0.52% annually) because Spain shows a growing pattern of energy and matter consumption, which is imported to produce goods and services, and this increase is higher than the growth experienced in the use of N and R.

Regarding the non-renewable and purchased resources used to produce the yield in relation to these renewable sources, the environmental loading ratio (ELR; Fig. 7a) is used to obtain information about economic pressure on ecosystems and their functions as suppliers of environmental services to society. The ELR increases during the whole period of the study (except in the recession period of 1989–1994), especially in the first part of 1984–1989 (with a growth rate of more than 6% annually) and 1994–2000 (with a growth rate of 5.14% annually).

Both indexes can be combined to evaluate the competence of transformation processes (the ability of foreign and national economic investments to exploit local resources or the return on eMergy investment) in relation to the pressure produced on the environment (relative weight of non-renewable and purchased sources in U), which is called the eMergy sustainability index or ESI (Brown and Ulgiati, 1997; Ulgiati and Brown, 1998). Under a local social–ecological perspective, ESI decreases continuously (Fig. 7b), 5.15% annually in the period 1984–1989 and 3.99% in the 1994–2000 period, because the

Table 1a
Emergy flows supporting Spanish social–ecological system in 2000

		Unit	Amount 2000 (unit/year)	Trans. (sej/unit)	Ref. trans.*	Emergy 2000 (sej/year)	Macroeconomic value 2000 (em$/year)
Renewable inputs							
1	Sunlight[a]	J/year	2.55E+21	1.00E+00	0	2.55E+21	8.24E+08
2	Rain (chemical potential)[b]	J/year	1.11E+18	3.06E+04	A	3.40E+22	1.10E+10
3	Rain (geopotential)[c]	J/year	6.91E+17	1.76E+04	A	1.22E+22	3.93E+09
4	Wind kinetic energy[d]	J/year	3.10E+17	2.52E+03	A	7.80E+20	2.52E+08
5	Waves[e]	J/year	1.18E+18	5.14E+04	A	6.09E+22	1.97E+10
6	Tides[f]	J/year	7.66E+17	7.39E+04	A	5.66E+22	1.83E+10
7	Earth cycle[g]	J/year	4.98E+17	1.20E+04	A	5.98E+21	1.93E+09
Indigenous non-renewable inputs							
8	Oil[h]	J/year	9.63E+15	9.06E+04	A	8.72E+20	2.82E+08
9	Coal[i]	J/year	3.21E+17	6.71E+04	A	2.15E+22	6.95E+09
10	Natural gas[j]	J/year	6.20E+15	8.05E+04	A	4.99E+20	1.61E+08
11	Iron[k]	g/year	7.51E+10	1.68E+09	A	1.26E+20	4.08E+07
12	Gold[k]	g/year	4.32E+06	7.39E+14	D	3.19E+21	1.03E+09
13	Silver[k]	g/year	1.15E+08	5.04E+14	D	5.77E+22	1.87E+10
14	Copper[k]	g/year	2.44E+10	3.36E+09	E	8.18E+19	2.65E+07
15	Feldspar[k]	g/year	4.78E+11	1.68E+09	A	8.03E+20	2.60E+08
16	Zinc[k]	g/year	2.02E+11	1.68E+09	A	3.40E+20	1.10E+08
17	Lead[k]	g/year	5.17E+10	1.68E+09	A	8.68E+19	2.81E+07
18	Salt rock[k]	g/year	3.87E+12	1.68E+09	A	6.50E+21	2.10E+09
19	Sulphur[k]	g/year	7.70E+11	1.68E+09	A	1.29E+21	4.18E+08
20	Glauberite y Thernardite[k]	g/year	8.34E+11	1.68E+09	A	1.40E+21	4.53E+08
21	Fluorite[k]	g/year	1.35E+11	1.68E+09	A	2.27E+20	7.33E+07
22	Magnesite[k]	g/year	2.21E+11	1.68E+09	A	3.71E+20	1.20E+08
23	Pumice[k]	g/year	7.62E+11	7.56E+09	A	5.76E+21	1.86E+09
24	Talc[k]	g/year	1.15E+11	1.68E+09	A	1.93E+20	6.22E+07
25	Quartz and silica sand[k]	g/year	6.59E+12	1.68E+09	A	1.11E+22	3.58E+09
26	Calcium carbonate[k]	g/year	1.95E+14	1.68E+09	A	3.28E+23	1.06E+11
27	Potash[k]	g/year	8.70E+11	1.68E+09	A	1.46E+21	4.72E+08
28	Barite[k]	g/year	3.27E+10	1.68E+09	A	5.49E+19	1.77E+07
29	Sand and gravel[k]	g/year	8.17E+13	1.68E+09	A	1.37E+23	4.43E+10
30	Clay[k]	g/year	4.33E+13	1.68E+09	A	7.27E+22	2.35E+10
31	Gypsum[k]	g/year	9.93E+12	1.68E+09	A	1.67E+22	5.39E+09
32	Quartzite[k]	g/year	2.13E+12	1.68E+09	A	3.58E+21	1.16E+09
33	Dolomite[k]	g/year	8.75E+12	1.68E+09	A	1.47E+22	4.75E+09
34	Ophite and porphyry[k]	g/year	4.74E+12	2.44E+09	A	1.15E+22	3.73E+09
35	Serpentine[k]	g/year	7.39E+11	1.68E+09	A	1.24E+21	4.01E+08
36	Marble[k]	g/year	3.66E+12	2.44E+09	A	8.92E+21	2.88E+09
37	Granite[k]	g/year	1.96E+13	8.40E+08	A	1.65E+22	5.32E+09

Table 1a (*continued*)

		Unit	Amount 2000 (unit/year)	Trans. (sej/unit)	Ref. trans.*	Emergy 2000 (sej/year)	Macroeconomic value 2000 (em\$/year)
38	Slate[k]	g/year	2.27E+12	1.68E+09	A	3.82E+21	1.23E−09
39	Net topsoil loss[l]	J/year	1.25E+16	1.05E+05	A	1.31E+21	4.22E−08

*References for transformities (Tables 1(a–c)):

0. Solar transformity is 1 sej J−1 by definition.

A. Odum, H.T., Brown, M.T., Brandt-Williams, S.B. (Eds.), 2000. Handbook of Emergy Evaluation. A Compendium of Data for Emergy Computation [Consulted: 1/6/2005]. Center for Environmental Policy, University of Florida, FL, USA. http://www.emergysystems.org/folios.php.

B. Bargigli, S., Ulgiati, S., 2003. Emergy and life-cycle assessment of steel production in Europe. In: Brown, M.T. (Ed.), Emergy Synthesis, 2: Theory and Application of the Emergy Methodology. Proceedings of the Second Biennial Emergy Conference. Center for Environmental Policy, University of Florida, Gainesville, FL, USA.

C. Brown, M.T. 2003. Resource imperialism: emergy perspectives on sustainability, international trade, and balancing the welfare of nations. In: Ulgiati, S., Brown, M.T., Giampietro, M., Herendeen, R., Mayumi, K. (Eds.), Proceedings of Third Biennial International Workshop. Advances in Energy Studies: Reconsidering the Importance of Energy, Porto Venere, Italy, pp. 135–149.

D. Brown, M.T., Arding, J., 1991. Transformities Working Paper. Center for Wetlands, University of Florida, Gainesville, FL, USA.

E. Brown, M. T., Ulgiati, S., 2004. EMergy analysis and environmental accounting. Encyclopedia of Energy 2, 329–354.

F. This study.

Calculations:

[a]*Sunlight*: Continental area = 498 476 km^2 (IGN, 1996), continental shelf area = 74 037 km^2 (IEO staff), average insolation = 5 704 091.20 kJ/m^2/año (INM (Instituto Nacional de Meteorologia), 2001, 2002, continental albedo = 0.35, shelf continental albedo = 0.2 (Henning, 1989). Energy = (Area)(Average insolation)(1−Albedo) = (498 476 km^2)(10E+6 m^2/km^2)(5 704 091.20 kJ/m^2/year)(10E+3 kJ/J)(1−0.35) + (74 037 km^2)(10E+6 m^2/km^2)(5 704 091.20 kJ/m^2/año)(103 kJ/J)(1−0.2) = 2.55E+21 J/year.

[b]*Rain (chemical potential)*: Continental area = 498 476 km^2 (IGN, 1996), average rain = 0.67 m (INM, 2001, 2002), ETa (67.84% of rainfall; MIMAM, 2000) = 0.45 m, Gibb's free energy of rainfall (G) = 4.94 J/g (Odum, 1996), rain water density = 1.00E+3 kg/m^3, Energy = (Area)(ETa)(G)(Rain water density) = (498 476 km^2)(10E+6 m^2/km^2)(0.45 m)(4.94 J/g)(1.00E+3 kg/m^3)(1.00E+3 g/kg) = 1.11E+18 J/year.

[c]*Rain (geopotential)*: Continental area = 498 476 km^2 (IGN, 1996), average rain = 0.67 m (INM, 2001, 2002), runoff (32.16% of rainfall; MIMAM, 2000) = 0.21 m, average elevation = 660 m, rain water density = 1.00E+3 kg/m^3, g = 9.8 m/s^2, Energy = (Area)(Runoff)(Rain water density)(Average elevation)(g) = (498 476 km^2)(10E+6 m^2/km^2)(0.21 m)(1.00E+3 kg/m^3) (660 m)(9.8 m/s^2) = 6.91E+17 J/year.

[d]*Wind (kinetic energy)*: Wind (kinetic energy) estimated from EU/EWEA (European Commission-DG TREN/European Wind Energy Association), 2003 as an annual technical onshore potential of wind energy.

[e]*Waves*: component of length parallel to front wave = 2340 (our estimate from IGN, 1996), sea water density = 1027 kg/m^3, g = 9.8 m/s^2, wave height (H_s) = 1.29 m (OAPEstado, On-line), average water depth in the breaker zone (h) = 6 m (Odum and Odum, 1983). Energy = (component parallel to front wave)(1/8)(sea water density)(H_s)^2(g × h)1/2 = (2340 km)(1.00E+3 m/km)(1/8)(1027 kg/m^3)(1.29)2(9.8 m/s^2 × 6 m)1/2(3.15E+7 s/year) = 1.18E+18 J/year.

[f]*Tides*: Continental shelf area = 74 037 km^2 (IEO staff), Tides/year = 730 (2/day), mean tidal range = 167.80 cm (OAPEstado, On-line), sea water density = 1027 kg/m^3, g = 9.8 m/s^2, Energy = (Area)(1/2)(Annual tides)(Mean tidal range)2(Sea water density)(g) = (74 037 km^2) (10E+6 m^2/km^2)(1/2)(730)(1.68 m)^2(1027 kg/m^3)(9.8 m/s^2) = 7.66E+17 J/year.

[g]*Earth cycle*: Continental area = 498 476 km^2 (IGN, 1996), heat flow estimated from Odum (1996) = 1.00E+6 J/m^2/year, Energy = (Area)(Heat flow) = (498 476 km^2)(10E+6 m^2/km^2)(1.00E+6 J/m^2/year) = 4.98E+17 J/year.

[h]*Oil*: 2.30E+5 toe (IEA, 2003), Energy = (toe)(1E+7 kcal/toe)(4186 J/kcal) = 9.63E+15 J/year.

[i]*Coal*: 7.66E+6 toe (IEA, 2003). Energy = (toe)(1E+7 kcal/toe)(4186 J/kcal) = 3.21E+17 J/year.

[j]*Natural gas*: 1.48E+5 toe (EUROSTAT, On-line), Energy = (toe)(1E+7 kcal/toe)(4186 J/kcal) = 6.20E+15 J/year.

[k]Iron, gold, silver, copper, feldspar, zinc, lead, salt rock, sulphur, glauberite and thernardite, fluorite, magnesite, pumice, talc, quartz and silica sand, calcium carbonate, potash, barite, sand and gravel, clay, gypsum, quartzite, dolomite, ophite and porphyry, serpentine, marble, granite and slate extraction from IGME (Instituto Geológico y Minero de España).

[l]*Net topsoil loss*: Farmed area = 1.83E+5 km^2 (MAPA, On-line), erosion rate = 150.5 Tm/km^2 (Soto, 1990; Cerdà, 2001), % organic matter in soil = 0.02 (Porta et al., 1994), energy = (Farmed area)(Erosion rate)(% Organic matter in soil)(Energy from organic matter)(4186 J/kcal) = (1.83E+5 km^2)(150.5 Tm/km^2)(1E+6 g/Tm)(0.02)(5 kcal/g)(4186 J/kcal) = 1.25E+16 J/year.

Table 1b
Emergy imports for Spanish social–ecological system in 2000

		Unit	Amount 2000 (unit/year)	Trans. (sej/unit)	Ref. trans*	Emergy 2000 (sej/year)	Macroeconomic value 2000 (em\$/year)
40	Oil and petroleum-derived products[a]	J/year	3.32E+18	9.06E+04	A	3.00E+23	9.43E+10
41	Coal[b]	J/year	5.56E+17	6.71E+04	A	3.73E+22	1.17E+10
42	Natural gas[c]	J/year	6.47E+17	8.05E+04	A	5.21E+22	1.64E+10
43	Electricity[d]	J/year	4.44E+16	3.36E+05	A	1.49E+22	4.68E+09
44	Agriculture and forest products[e]	J/year	4.51E+16	1.75E+05	A	7.88E+21	2.48E+09
45	Livestock and products[f]	J/year	4.85E+15	5.33E+06	A	2.58E+22	8.11E+09
46	Food industry products[g]	g/year	7.64E+12	3.36E+05	A	2.57E+18	8.06E+05
47	Fishery products[h]	J/year	5.44E+15	3.36E+06	A	1.83E+22	5.74E+09
48	Metallic minerals[g]	g/year	1.14E+13	1.68E+09	A	1.92E+22	6.01E+09
49	Non-metallic minerals[g]	g/year	1.20E+13	1.68E+09	A	2.02E+22	6.34E+09
50	Steel and pig iron[g]	g/year	1.79E+13	3.69E+09	B	6.61E+22	2.08E+10
51	Metallic minerals (products/alloys)[g]	g/year	1.27E+12	1.68E+09	A	2.13E+21	6.69E+08
52	Mechanical and transport equipment[g]	g/year	7.03E+12	1.13E+10	A	7.91E+22	2.49E+10
53	Industrial minerals[g]	g/year	1.75E+12	1.68E+09	A	2.94E+21	9.23E+08
54	Leather and products[i]	J/year	4.47E+15	1.44E+07	A	6.46E+22	2.03E+10
55	Textiles[j]	J/year	1.91E+16	6.38E+06	A	1.22E+23	3.83E+10
56	Wood and products[k]	J/year	9.19E+16	5.86E+04	A	5.39E+21	1.69E+09
57	Paper[g]	g/year	5.33E+12	6.55E+09	A	3.49E+22	1.10E+10
58	Chemicals[g]	g/year	1.39E+13	6.38E+08	A	8.89E+21	2.79E+09
59	Rubber[g]	g/year	8.85E+11	7.22E+09	A	6.39E+21	2.01E+09
60	Total goods associated to imports[l]	\$	1.51E+11	1.85E+12	C	2.80E+23	8.80E+10
61	Total services associated to imports (without tourism)[l]	\$	2.60E+10	1.85E+12	C	4.81E+22	1.51E+10
62	Total money associated to tourism services imports[l]	\$	5.51E+09	1.85E+12	C	1.02E+22	3.20E+09

*See footnotes in Table 1a.

[a] *Oil and petroleum-derived products*: 7.92E+7 toe (IEA, 2003), energy = (toe)(1.00E+7 kcal/toe)(4186 J/kcal) = (7.92E+7 toe)(1.00E+7 kcal/toe)(4186 J/kcal) = 3.32E+18 J/year.

[b] *Coal*: 1.33E+7 toe (IEA, 2003), Energy = (toe)(1.00E+7 kcal/toe)(4186 J/kcal) = (1.33E+7 toe) (1.00E+7 kcal/toe)(4186 J/kcal) = 5.56E+17 J/year.

[c] *Natural gas*: 1.55E+7 toe (IEA, 2003), energy = (toe)(1.00E+7 kcal/toe)(4186 J/kcal) = (1.55E+7 toe) (1.00E+7 kcal/toe)(4186 J/kcal) = 6.47E+17 J/year.

[d] *Electricity*: 1.06E+6 toe (IEA, 2003), energy = (toe)(1.00E+7 kcal/toe)(4186 J/kcal) = (1.06E+6 toe) (1.00E+7 kcal/toe)(4186 J/kcal) = 4.44E+16 J/year.

[e] *Agriculture and forest products*: 1.54E+13 g (AEAT, On-line), energy = (1.54E+13 g)(0.20)(3.5 kcal/g)(4186 J/kcal) = 4.51E+16 J/year.

[f] *Livestock and products*: 1.05E+12 g (AEAT, On-line), energy = (1.05E+12 g)(0.22)(5.0 kcal/g)(4186 J/kcal) = 4.85E+15 J/year.

[h] *Fishery products*: 1.18E+12 g (AEAT, On-line), energy = (1.18E+12 g)(0.22)(5.0 kcal/g)(0.22)(5.0 kcal/g)(4186 J/kcal) = 5.44E+15 J/year.

[i] *Leather and products*: 2.83E+11 g (AEAT, On-line), energy = (matter)(15 800 J/g) = (2.83E+11 g) (15 800 J/g) = 4.47E+15 J/year.

[j] *Textiles*: 1.21E+12 g (AEAT, On-line), Energy = (matter)(15 800 J/g) = (1.21E+12 g) (15 800 J/g) = 1.91E+16 J/year.

[k] *Woods and products*: 6.10E+12 g (AEAT, On-line), Energy = (matter)(15 800 J/g) = (6.10E+12 g) (15 800 J/g) = 9.19E+16 J/year.

[g] Food industry products, metallic minerals, non-metallic minerals, steel and pig iron, metallic minerals (products and alloys), mechanical and transport equipment, industrial minerals, paper, chemicals and rubber (AEAT, On-line).

[l] Total goods associated to imports, total services associated to imports (without tourism), total money associated to tourism from BDE (On-line-a).

relative importance of the flow of renewable eMergy sources is reduced.

3.5. Carrying capacity of the Spanish SES

In eMergy terms, carrying capacity may have two main approaches (Fig. 8): a people-based one (more similar to the classical concept of carrying capacity), linked to number of people supported by eMergy used (Odum, 1996; Campbell, 1998), and an area-based one (similar to ecological footprint), associated to support area needed to maintain the standard of living of people (in terms of eMergy use *per capita*) (Brown and Ulgiati, 2001). These two approaches could be applied to both "only renewable" and "developed" scenarios. The renewable scenario would be a lower limit, based only on renewable flows, and the

Table 1c
Emergy exports and selected products for Spanish social-ecological system in 2000

		Unit	Amount 2000 (unit/year)	Trans. (sej/unit)	Ref. Trans.*	Emergy 2000 (sej/year)	Macroeconomic value 2000 (cm$/year)
Exports							
63	Petroleum-derived products[a]	J/year	3.17E+17	9.06E+04	A	2.87E+22	9.02E+09
64	Coal[b]	J/year	2.26E+16	6.71E+04	A	1.52E+21	4.76E+08
65	Electricity[c]	J/year	2.58E+16	3.36E+05	A	8.68E+21	2.73E+09
66	Agriculture and forest products[d]	J/year	3.42E+16	1.75E+05	A	5.98E+21	1.88E+09
67	Livestock and products[e]	J/year	2.97E+15	5.33E+06	A	1.58E+22	4.97E+09
68	Food industry products[f]	g/year	5.35E+12	3.36E+05	A	1.80E+18	5.65E+05
69	Fishery products[g]	J/year	3.07E+15	3.36E+06	A	1.03E+22	3.24E+09
70	Metallic minerals[f]	g/year	9.93E+11	1.68E+09	A	1.67E+21	5.24E+08
71	Non-metallic minerals[f]	g/year	1.27E+13	1.68E+09	A	2.14E+22	6.71E+09
72	Steel and pig iron[f]	g/year	7.50E+12	3.69E+09	B	2.77E+22	8.69E+09
73	Metallic minerals (products/alloys)[f]	g/year	8.80E+11	1.68E+09	A	1.48E+21	4.64E+08
74	Mechanical and transport equipment[f]	g/year	7.21E+12	1.13E+10	A	8.12E+22	2.55E+10
75	Industrial minerals[f]	g/year	8.00E+12	1.68E+09	A	1.34E+22	4.22E+09
76	Leather and products[h]	J/year	2.12E+15	1.44E+07	A	3.06E+22	9.61E+09
77	Textiles[i]	J/year	1.14E+16	6.38E+06	A	7.30E+22	2.29E+10
78	Wood and products[j]	J/year	1.96E+16	5.86E+04	A	1.15E+21	3.61E+08
79	Paper[f]	g/year	2.79E+12	6.55E+09	A	1.83E+22	5.74E+09
80	Chemicals[f]	g/year	9.11E+12	6.38E+08	A	5.82E+21	1.83E+09
81	Rubber[f]	g/year	6.65E+11	7.22E+09	A	4.80E+21	1.51E+09
82	Total goods associated to exports[k]	$	1.17E+11	3.09E+12	F	3.61E+23	—
83	Total services associated to exports (without tourism)[k]	$	2.28E+10	3.09E+12	F	7.06E+22	—
84	Total money associated to tourism services exports[k]	$	3.12E+10	3.09E+12	F	9.66E+22	—
Selected products							
85	Population 2000[l]	Inhabitants	3.99E+7	4.35E+16	F	1.73E+24	—
86	GDP 2000[m]	$	5.62E+11	3.09E+12	F	1.73E+24	—

*See footnotes in Table 1a.

[a] *Petroleum-derived products*: 7.57E+6 toe (IEA, 2003), Energy = (toe)(1.00E+7 kcal/toe)(4186 J/kcal) = (7.57E+6 toe)(1.00E+7 kcal/toe)(4186 J/kcal) = 3.17E+17 J/year.

[b] *Coal*: 5.40E+5 toe (IEA, 2003), Energy = (toe)(1.00E+7 kcal/toe)(4186 J/kcal) = (5.40E+7 toe) (1.00E+7 kcal/toe)(4186 J/kcal) = 2.26E+16 J/year.

[c] *Electricity*: 6.73E+5 toe (IEA, 2003), Energy = (toe)(1.00E+7 kcal/toe)(4186 J/kcal) = (6.73E+5 toe) (1.00E+7 kcal/toe)(4186 J/kcal) = 2.82E+16 J/year.

[d] *Agriculture and forest products*: 1.17E+13 g (AEAT, On-line), Energy = (1.17E+13 g)(0.20)(3.5 kcal/g)(4186 J/kcal) = 3.42E+16 J/year.

[e] *Livestock and products*: 6.46E+11 g (AEAT, On-line), Energy = (6.46E+11 g)(0.22)(5.0 kcal/g)(4186 J/kcal) = 2.97E+15 J/year.

[g] *Fishery products*: 6.67E+11 g (AEAT, On-line), Energy = (6.67E+11 g)(0.22)(5.0 kcal/g)(0.22)(5.0 kcal/g)(4186 J/kcal) = 3.07E+15 J/year.

[h] *Leather and products*: 1.34E+11 g (AEAT, On-line), Energy = (matter)(15 800 J/g) = (1.34E+11 g)(15 800 J/g) = 2.12E+15 J/year.

[i] *Textiles*: 7.24E+11 g (AEAT, On-line), Energy = (matter)(15 800 J/g) = (7.24E+11 g)(15 800 J/g) = 1.14E+16 J/year.

[j] *Woods and products*: 1.30E+12 g (AEAT, On-line), Energy = (matter)(15 800 J/g) = (1.30E+12 g) (15 800 J/g) = 1.96E+16 J/year.

[f] Food industry products, metallic minerals, non-metallic minerals, steel and pig iron, metallic minerals (products and alloys), mechanical and transport equipment, industrial minerals, paper, chemicals and rubber (AEAT, On-line).

[k] Total goods associated to exports, Total services associated to exports (without tourism), Total money associated to tourism from BDE (Banco de España), 2006.

[l] Population from INE (On-line).

[m] GDP from UNSD (On-line).

developed scenario would be an upper limit, based on actual conditions.

If a people-based approach to renewable carrying capacity for Spain is employed, the population that could be supported only with renewable sources shows a decline of 50%, shifting from values of 6% to 3% of the actual population, caused by the relative decrease in the use of renewable eMergy flows in relation to local non-renewable or imported flows.

When using the people-based approach for developed carrying capacity, Spain is considered embedded in the European or the Mediterranean contexts. The Mediterranean context implies the use of the traditional ecological knowledge accumulated during centuries of adaptive

Table 2
Comparison of main emergy indexes and flows for time series emergy synthesis of Spain

No.	Flow/index	Expression	1984	1989	1994	2000	2002	Units
1	Renewable sources used	R	6.09E+22	6.28E+22	5.90E+22	6.09E+23	6.09E+23	Sej/year
2	Non-renewable indigenous sources	N	5.09E+23	7.05E+23	5.49E+23	7.55E+23	8.47E+23	Sej/year
	Dispersed rural sources	N_0	1.47E+21	1.45E+21	1.31E+21	1.31E+21	1.28E+21	Sej/year
	Concentrated used	N_1	4.76E+23	6.84E+23	5.25E+23	7.30E+23	8.22E+23	Sej/year
	Exported without use	N_2	3.11E+22	2.00E+22	2.26E+22	2.46E+22	2.27E+22	Sej/year
3	Imported emergy	IMP	3.81E+23	5.58E+23	6.70E+23	9.47E+23	1.02E+24	Sej/year
	Fuels and electricity	F1	2.19E+23	2.63E+23	3.00E+23	4.05E+23	4.28E+23	Sej/year
	Minerals	M1	4.21E+22	5.48E+22	6.48E+22	1.08E+23	1.16E+23	Sej/year
	Goods (without fuels and electricity)	G1	9.85E+22	2.00E+23	2.68E+23	3.76E+23	4.10E+23	Sej/year
	Services (without tourism)	SS1	2.03E+22	3.52E+22	2.90E+22	4.81E+22	5.68E+22	Sej/year
	Touristic services	P1E3	1.54E+21	5.70E+21	7.63E+21	1.02E+22	1.23E+22	Sej/year
4	Exported emergy	EXP	2.68E+23	2.83E+23	4.10E+23	5.19E+23	5.46E+23	Sej/year
	Fuels and electricity	F2	5.17E+22	4.79E+22	5.03E+22	3.89E+22	3.34E+22	Sej/year
	Minerals	M2	5.80E+22	3.32E+22	4.85E+22	5.22E+22	5.07E+22	Sej/year
	Goods (without fuels and electricity)	G2	6.40E+22	1.11E+23	1.74E+23	2.60E+23	2.80E+23	Sej/year
	Services (without tourism)	SS2	5.11E+22	3.76E+22	8.45E+22	7.06E+22	8.40E+22	Sej/year
	Touristic services	P1E2	4.34E+22	5.39E+22	5.30E+22	9.66E+22	9.78E+22	Sej/year
5	Total emergy available	$R+N+$IMP	9.50E+24	1.33E+24	1.28E+24	1.76E+24	1.93E+24	Sej/year
	Total emergy actually used	$U=R+N+$IMP$-N_2$	9.19E+23	1.31E+24	1.25E+24	1.74E+24	1.91E+24	Sej/year
6	Economic component of emergy used	$U-R$	8.58E+23	1.24E+24	1.20E+24	1.68E+24	1.85E+24	Sej/year
7	Fraction of use derived from indigenous sources	$(R+N_0+N_1)/U$	0.59	0.57	0.47	0.46	0.46	—
8	Fraction of use that is renewable	R/U	0.07	0.05	0.05	0.04	0.03	—
9	Fraction of use that is free	$(R+N_0)/U$	0.07	0.05	0.05	0.04	0.03	—
10	Fraction of use that is imported	IMP/U	0.41	0.43	0.53	0.54	0.54	—
11	Imports minus exports	IMP−EXP	1.13E+23	2.75E+23	2.59E+23	4.28E+23	4.76E+23	Sej/year
12	Imports/exports	IMP/EXP	1.42	1.97	1.63	1.83	1.87	—
13	Ration of concentrated to rural	$(U-R-N_0)/(R+N_0)$	13.74	19.32	19.79	26.95	29.67	—
14	Emergy use *per capita*	U/population	2.40E+16	3.37E+16	3.20E+16	4.35E+16	4.70E+16	Sej/people/year
15	Fraction of use that is electrical	Electrical emergy/U	0.20	0.22	0.23	0.20	0.18	—
16	Fuels *per capita*	Fuels/population	6.77E+15	7.65E+15	8.12E+15	9.58E+15	1.05E+16	Sej/people/year
17	Population		3.83E+07	3.88E+07	3.92E+07	3.99E+07	4.05E+07	people
18a	Empower density	U/Area	1.84E+12	2.62E+12	2.52E+12	3.49E+12	3.83E+12	Sej/m^2/year
18b	Non-renewable empower density	(IMP$+N$)/Area	1.78E+12	2.53E+12	2.44E+12	3.41E+12	3.75E+12	Sej/m^2/year
19	Emergy investment ratio (EIR)	IMP/$(R+N_0+N_1)$	0.71	0.75	1.14	1.20	1.16	—
20	Emergy yield ratio (EYR)	1+1/EIR	2.41	2.34	1.87	1.84	1.87	—
21	Environmental loading ratio (ELR)	$(U-R)/R$	14.10	19.79	20.26	27.55	30.32	—
22	Emergy sustainability index (ESI)	EYR/ELR	0.17	0.12	0.09	0.07	0.06	—
23	Emergy exchange ratio (EER)	EMRg/EMR	0.33	0.56	0.74	0.60	0.64	—
24	Gross DOMESTIC product at market prices (GDP)	GDP	1.64E+11	3.94E+11	5.04E+11	5.62E+11	6.55E+11	$
25	Emergy to money ratio (EMR)	$P_1=U$/GDP	5.62E+12	3.31E+12	2.49E+12	3.09E+12	2.91E+12	Sej/$
26	Global emergy to money ratio	P_2	1.85E+12	1.85E+12	1.85E+12	1.85E+12	1.85E+12	Sej/$
27	Renewable carrying capacity	$(R/U)\times$ population	2.45E+06	1.87E+06	1.84E+06	1.40E+06	1.29E+06	People
28	Developed carrying capacity at European standard of living	ESL $(R/U)\times$ population	6.09E+07	4.48E+07	4.43E+07	3.36E+07	3.11E+07	People
29	Developed carrying capacity at Mediterranean standard of living	MSL $(R/U)\times$ population	3.41E+07	2.51E+07	2.48E+07	1.88E+07	1.74E+07	People
30	Renewable support area (SA$_{(r)}$)	(IMP$+N$)/REmpD$_{(r)}$	7.28E+12	1.00E+13	1.03E+13	1.39E+13	1.53E+13	m^2
31	Synchronal support area at European standard of living (SSA$_E$)	R^*/E_REmpD$_{(r)}$	1.87E+11	2.66E+11	2.56E+11	3.58E+11	3.93E+11	m^2
32	Synchronal support area at Mediterranean standard of living (SSA$_M$)	R^*/M_REmpD$_{(r)}$	4.42E+11	6.27E+11	6.05E+11	8.45E+11	9.28E+11	m^2

Details about ESL (European Standard of Living) and MSL (Mediterranean Standard of Living) are contained in Appendix C.

R^* is the required amount of renewable emergy necessary to lower the ELR of the country to that of the region. ($R^*=(\mathrm{IMP}+N)/\mathrm{ELR}(r)$, where ELR($r$) is the environmental loading ratio of the region, calculated in Appendix C).

EMRg is the global Emergy to money ratio from Brown (2003) and EMR is the Emergy to money ratio calculated in this paper for Spain.

E_REmpD(r) and M_R EmpD(r) are the renewable empower density for European and Mediterranean context, respectively, and are calculated as renewable emergy used of the region (Europe or Mediterranean Basin)/total area of the region (data in Appendix C).

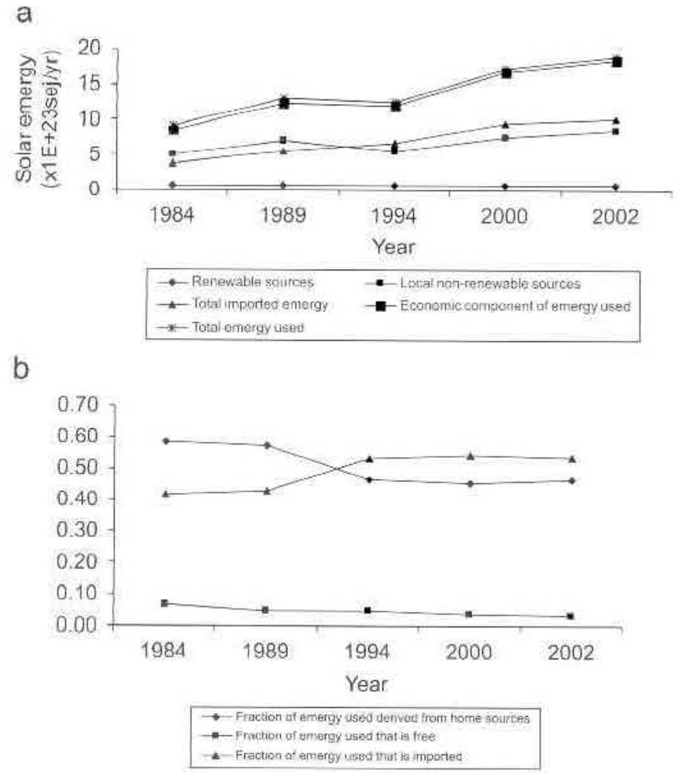


Fig. 5. Main emergy flows in the Spanish social–ecological system: (a) total emergy actually used and its components and (b) main relations among components.

learning to couple the Mediterranean natural perturbation regime with the human activities, although possibilities of economic growth are limited by the connections between the economy and the flow of environmental services that Mediterranean ecosystems supply to Spain. The European context implies the increasing use of international trade to supply goods and services for the national economy, so there is a growing disconnection between local flow of environmental services and the national economy. Potential possibilities for growth are higher for the European context but it means a disconnection between the use of energy and materials and the supply of local environmental services and a loss of resilience as a result. To estimate a benchmark standard of living (ratio of the total eMergy actually used to the renewable one) for those two regions, we have used data contained in Brown (2003) and Cialani et al. (2005) for 14 European and five Mediterranean countries in the 1990s, which are summarized in Appendix C.

If we assume that the Western European standard of living is the correct one, we use the developed carrying capacity with the European Standard of Living (ESL) as the upper limit. The developed carrying capacity at the European standard of living is above present Spanish population in the periods from 1984 to 1994. This means that there was a margin for growth that has been exceeded in the period 1994–2000, in which the

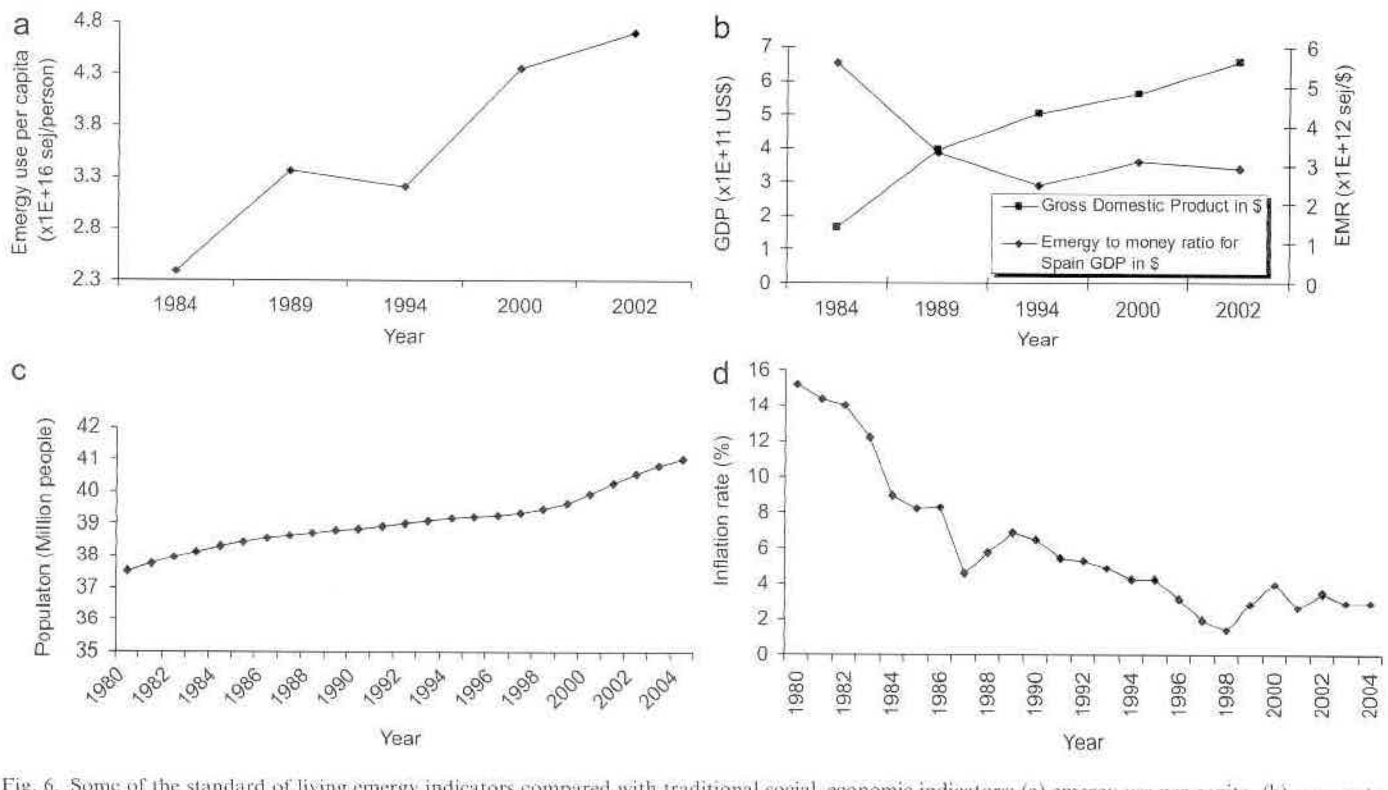


Fig. 6. Some of the standard of living emergy indicators compared with traditional social–economic indicators: (a) emergy use per capita, (b) emergy to money ratio compared with GDP, (c) population patterns in Spain (1980–2004) and (d) inflation rate patterns in Spain (1980–2004).

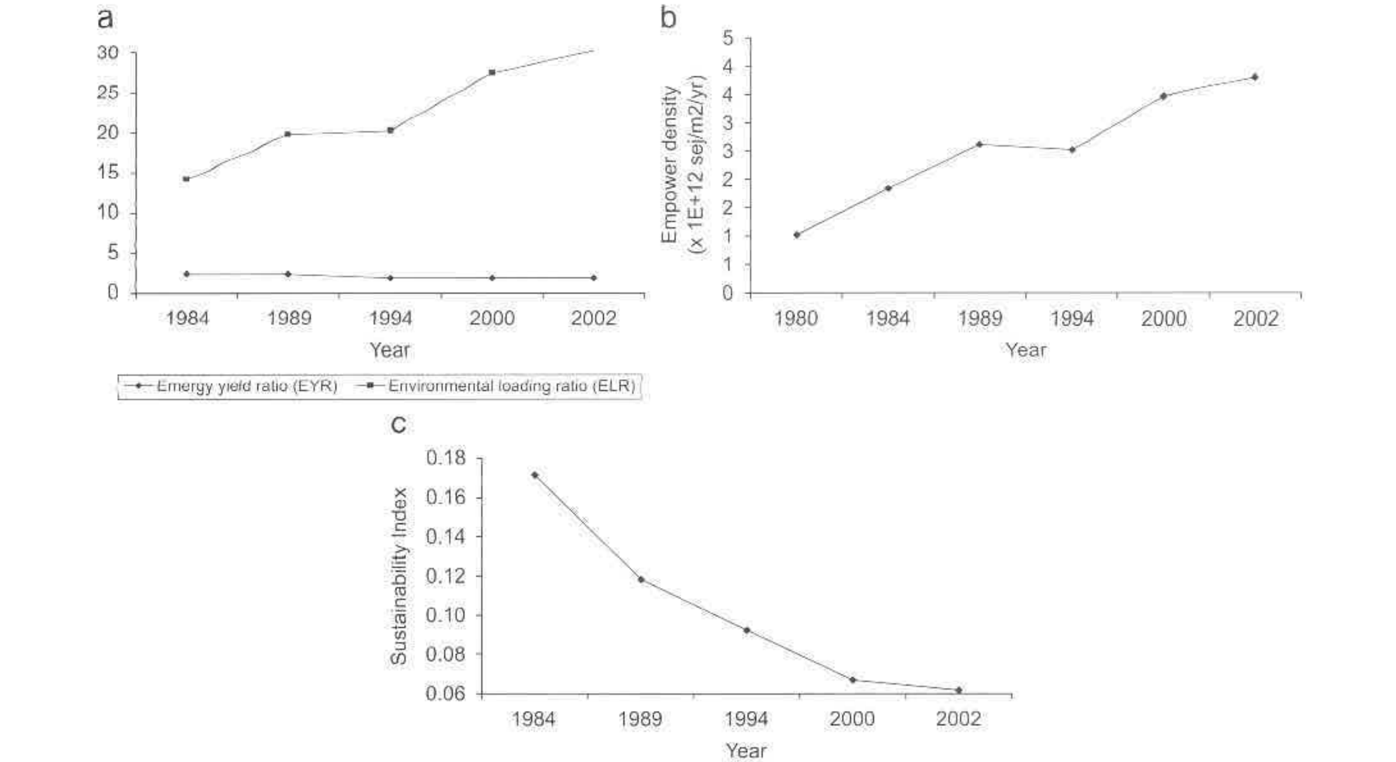


Fig. 7. Emergy indicators for Spain 1984–2002: (a) emergy yield ratio and environmental loading ratio, (b) sustainability index and (c) empower density.

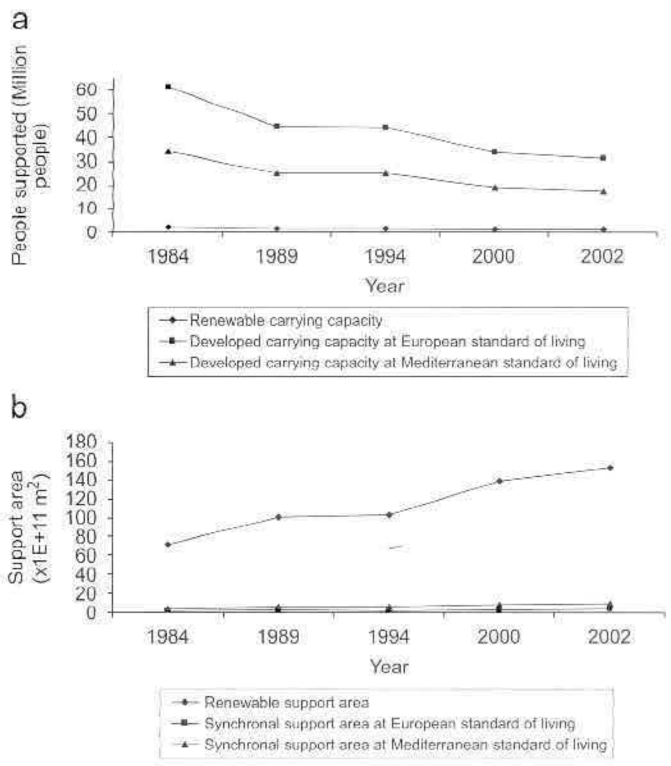


Fig. 8. Evolution of carrying capacity in emergy terms for Spain (1984–2002): (a) people supported and (b) support area.

population supported was 84% of actual population in 2000. In the case of the Mediterranean standard of living, we use the developed carrying capacity at the Mediterranean Standard of Living (MSL) as the upper limit. We have to take into account that it has been calculated with data from only five countries, which was all the available data in the literature for the Mediterranean Basin. The number of people supported decreases from 88% (1984) to 43% (2002) of the respective actual population for these years. It means that the Spanish standard of living had already exceeded the Mediterranean one in the middle 1980s, so if Spain wants to maintain the Mediterranean way of life, which it has experienced so far, Spain has to reduce its eMergy consumption *per capita*.

If the support area-based approach is used for renewable carrying capacity, the area that we would have to use in order to maintain the Spanish standard of living with only renewable sources is 15 (1984) to 30 (2002) times the actual area of the country. This illustrates a doubling of the renewable ecological footprint due to the strong growth experienced in recent decades.

To estimate the support area-based approach for developed carrying capacity (sometimes called synchronal support area), we have to calculate ELR (Appendix C the reference region (Mediterranean or European). We use the same sources that the previous people-based approach to calculate Mediterranean and European regional ELR for 14 European and five Mediterranean countries in the 1990s, summarized in Appendix C. If the standard of living of the European region is employed as a reference, our territorial margin, in terms of the area of Spain that remains after using our territory to reach European ELR, is decreasing, because synchronal support area has increased from 37% of actual area in 1984 to 78% in 2002. In contrast, if the Mediterranean Basin is employed as a reference regional ELR, the increase in support area needed to equal our actual ELR to the Mediterranean regional one was growing from the 88% of the actual area in 1984 to 185% in 2002.

4. Discussion

4.1. Patterns in the supply of environmental goods and services flows to the Spanish economy and its changes

Spanish eMergy use, U, has similar values to other western countries (Appendix A). Compared with population size, U significantly increased over the investigated two decades. As a result, the Spanish standard of living in terms of resources use has increased (eMergy *per capita*). Spain ranks within the group of highly industrialised countries, although still below the average level of several European countries (Appendix A). The particular decrease in eMergy use *per capita* that took place in the 1989–1994 period may have been affected by the recession of the European Monetary System from 1992 to 1993, which caused the peseta (the Spanish currency at that time) to leave the Exchange Rate Mechanism of the European Monetary System in 1992. Three devaluations of the peseta took place between 1992 and 1993 (Gadea, 2000), causing strong disturbances in the energy and materials required and in economic growth levels within the nation. This indicator coincides with the increase of energy use and total material requirements emphasized in the first sustainability report for Spain (OSE (Observatorio de la Sostenibilidad en España), 2005).

This situation could be considered as a case of inflation in eMergy terms, therefore, more money circulating for the same eMergy. However, we have to take into account that the study has been done during some of the years involved in increasing monetary inflation periods for Spain (1987–1989, 1998–2000 and 2001–2002), despite the decreasing inflation rates during the last two decades (Fig. 6d). It will be necessary to study more years to avoid accounting for only inflation-peak years.

As a consequence, there has been an intensification of transformation activity on the territory (empower density) and an increase of pressure or stress on ecosystems due to production (environmental loading ratio). Spain reaches ELR values close to those of the USA or Switzerland for 1999. This process has been supplied by flows of matter and energy mainly based on imports of external energy memory, thus, Spain has become less autonomous (self-sufficient), especially with regards to fuels. As a consequence, there has been a loss in the potential contribution of local eMergy sources to the main economy (eMergy yield ratio), because growing amounts of resources have to be imported to support the growing Spanish standard of living. In the international context, the Spanish EYR is within the range of EYR for European and other western countries. The result is that the Spanish ESI decreases because of the growing pressure on ecosystems derived from the intensification of the economy related to its high dependency on external eMergy sources, added to the relative low contribution of local eMergy sources to production. The ESI change rate has to be emphasized, especially in the mid-1990s. In the international context, there are many countries which show higher ESI indexes than Spain (Appendix B), but because of different causes. There are countries which have an extremely low value of eMergy use *per capita* with a high use of locally available renewable resources, which sometimes could mean potential wealth not adequately used, and others with a low value of eMergy *use per capita*, but with a high use of non-renewable sources. This could be the case of countries like Bolivia, Kenia, India, etc. Spain shows the patterns of a western country, with a small ESI derived from its high IMP and N flows, but still with higher values than most of the European countries, as a result of the relatively late incorporation into the European Union (Fig. 4) economic and consumption patterns.

4.2. Patterns of trade in the context of economic globalization

As we have seen, the IMP has become by far the most significant eMergy flow in the Spanish economy, and so trade is a crucial aspect to the study of Spain as a social–ecological system. In classic and environmental economic assessments of trade employed to support decision-making, the predominant approaches are monetary ones, with a user-side value approach. In these approaches, economic policy is reduced to the balance of payments, and value is measured by what is considered to be the best indicator of utility: price. In this sense, in the official statistics on foreign trade for Spain for every year studied, the countries mainly involved in trade exchanges with Spain in monetary terms are those of the European Union and the Organisation for Economic Co-operation and Development (OECD), whose economies are mainly based on manufacture exports. As a result, Spain could be considered within the group of countries reaching a kind of dematerialization, growing without an increase in matter and energy use, but with some problems in the balance of payments.

On the contrary, in a donor-side value approach, as provided by eMergy Synthesis, the concept of value is related to the work done by nature to produce environmental goods and services that support the economy. In eMergy synthesis, value is related to the energy memory of these environmental goods and services. And, in this case, buying power is not estimated by price but by the EMR or eMergy potentially bought by one monetary unit. Therefore, the origin of the main imported eMergy flows for Spain is the oil and natural gas extracting countries (Nigeria, Algeria and some Middle Eastern countries), whose economies are mainly based on raw materials exports. As a result, Spain could be considered as a net importer of raw materials, with a high increase in the use of energy and matter, promoting a kind of false dematerialization by moving the environmental loading required by its growth to countries that supply raw materials (Muradian et al., 2002; Ramos Martín, 2001, 2003; Cañellas et al., 2004; Carpintero, 2005).

A comparison of EMRs (or buying power in eMergy terms) for different countries to the global EMR (Appendix B) or EER shows that there are differences in the relative buying power of different countries, so in the commerce trade with raw materials exporting countries, Spain commerce with an eMergy advantage in these product exchanges. In these terms, there is a natural decapitalization of supplier countries, promoted by the organization of trade, international division of labour, and economies of scale related to the export of primary sources by developing countries and the import by western countries. In this context, Spain, like other industrialised countries, is promoting natural decapitalization and poverty in the supplier countries with trading disadvantages in eMergy terms (those which have an EER smaller than our EER). This is another example of what Brown (2003) calls resource imperialism. On the other hand, Spanish EER is below the value of most western countries (Appendix B), and therefore many of them have eMergy advantages in trade relations with Spain.

Thus, the greatest part of the pressure, in terms of non-renewable stocks of resource depletion or exploitation, is transferred to the exporting countries (they have to use their own resources and processes to satisfy Spanish demand). These resources are used to exploit and develop the importing country beyond the possibilities that a renewable economy would provide Spain, promoting a decoupling of the Spanish national economy from the flow of local environmental goods and services (natural capital), and the limits that this imposes on the local growth of the importing country.

4.3. Decoupling between national flow of environmental services and the Spanish economy

The Mediterranean standard of living has supported an agricultural way of life for more than eight millennia. This fact might intuitively be interpreted as a measure of the sustainability of this way of life (Butzer, 2005). In the last 40 years, many economies, especially in the northern part of the Mediterranean Basin have become disconnected from this ancient way of life: that is, disconnected from the goods and services that their territories supply. In the present, the standard of living of these countries is mainly supported by imported flows of goods. As we have seen from the eMergy indicators, the strong growth of the Spanish standard of living (eMergy *per capita*) has been mainly supported by imports of primary resources (high content in eMergy and a low monetary value), promoting a disconnection between the original flow of environmental services and the requirements of the Spanish economy. How important is this decoupling? Or to what extent is Spain exploiting its system over its endogenous possibilities?

It seems clear that the Spanish endorsement of the European economic community (EEC) Treaty in 1986 entailed great social–economic changes. It is probable that previous patterns of strong growth in the 1960s were accelerated, and, as is shown by standard of living, carrying capacity and footprint eMergy indicators, Spain left the Mediterranean standard of living to adopt a Western European one. This disconnection becomes evident from the middle 1980s, but its growth rate is especially strong after the middle 1990s. In this sense, both carrying capacity measures show that in the mid-1980s Spain disconnects definitively from its Mediterranean way of life to adopt an European one.

To deal with the challenge of natural capital decapitalization (strong use of N, high dependency on imports, high pressure on environmental systems, low efficiency in the yield, etc.), different Spanish governments invested a great amount of money in conservationist programmes. In fact, Spain ranks as the third country in the EU in terms of the money spent on environmental protection measures, with an average of 0.8% of GDP and 108 € *per capita* (EUROSTAT On-line-a, b). The natural protected areas in Spain will be considered a good measure of conservation policies, in terms of area and money spent during the past 20 years. Creation of a natural protected areas policy has been developed since the 1980s (Morillo and Gómez-Campo, 2000), supported mainly by international and European legislation. The Conservation of Nature-Wild Flora and Fauna Act of 1989 created different types of natural protected areas to preserve some parts of the country outside of the general economic process of growth and land transformation, and it is the real starting point of the natural protected areas declaration in Spain (Fig. 9). In 2003, there were already 950 protected areas in 38 different protection categories embracing more than 9% of the country's surface (EUROPARC-España, 2004).

eMergy indicators illustrate that conservation policies are not successful enough in terms of preservation of natural capital to enhance sustainability. It has been shown that the intensity of use of the territory has grown and that carrying capacity is strongly decreasing, so the Spanish ecological footprint, in eMergy terms, is increasing too. In this Mediterranean context, natural protected areas cannot be managed as islands inside the territory in which they are embedded, since a full set of biophysical, socio-economic and historical–cultural aspects are shared by both sides of the fence (García and Montes, 2003). In fact, other indicators, like the Natural Capital index (NCI) illustrate that Spain has a great quantity of "natural

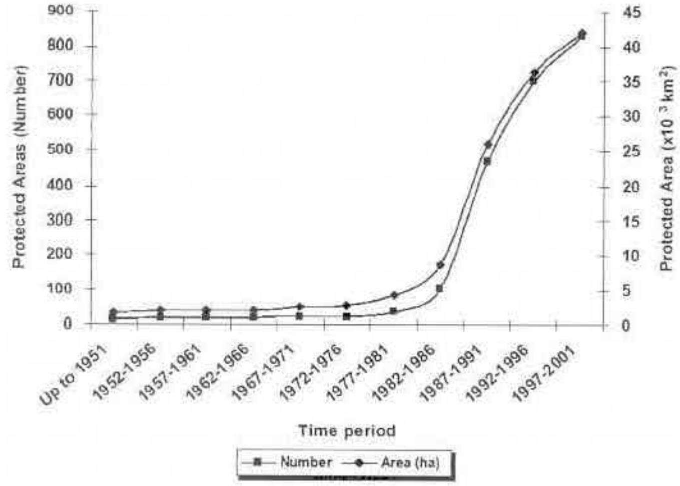


Fig. 9. Evolution of number and area of natural protected areas in Spain during the last 50 years (*Data source*: EUROPARC-España, 2003.)

areas" but that their quality (ratio between current state of the ecosystem and the defined baseline state) is low (Ten Brink, 2000). In fact, there are countries with a low quantity of natural areas but with a high quality, so their NCI is higher than Spain. This is the case with regard to another Mediterranean country, Greece (De Groot et al., 2003).

It would be interesting to have historical series to study previous periods and compare evolution in the last 15 years with the past decades prior to the entry of Spain in the EU. However, these eMergy indicators confirm patterns suggested by other studies of ecological footprint in Spain. Carpintero (2005) has studied the economic metabolism of Spain, and estimates the ecological footprint changes from 2 ha/inhabitant (1955) to 5 ha/inhabitant (2000), which is more than three times the total area of Spain, including the marine portion. The World Wildlife Fund (WWF) (WWF/WCMC-UNEP, 2004) estimates the ecological footprint to be 4.8 ha/inhabitant, so that there would be an ecological deficit of 2.9 ha/inhabitant.

Although Spain has only been studied between 1984 and 2002, it has to be underlined that patterns obtained are confirmed by the partial indicators of the OSE (Observatorio de la Sostenibilidad en España) (2005): so far from recovering a Mediterranean way of life connected to the flow of goods and services of its own territory, sustainability indicators are getting worse and deepening in the "growth without limits" model.

5. Conclusions

Despite ancient transformations of its territory, Spain began the first part of the 1980s with one of the best-preserved natural heritages in the Mediterranean and European area. From a socio-economic point of view, the 1980s starts with a political transition and with the economy in a growth period, without strong pressure on ecosystems and with a productive system that was still extensive in many cases.

In this paper, an historical series of eMergy indicators, instead of traditional monetary ones, has been studied in Spain for five different years to determine the balance and evolution of social–ecological dynamics (trends of resource use) during the last two decades. It can be inferred from the use of these indicators that Spain has suffered a global backward movement in sustainability, with increased intensity in the second part of the 1990s. eMergy indicators stress the magnitude and speed of the changes that the Spanish economy faced in the last two decades, as well as its strong dependence on imported resources. Other eMergy indicators estimate the consequences that those changes have had on the territory, in terms of natural capital decapitalization and the increasing need to spend money to substitute for the free environmental services formerly supplied by the lost of the past Mediterranean way of life to adopt a Western European one.

The sustainable use of resources in the Mediterranean Basin has been accomplished as a consequence of human and ecological resilience (Butzer, 2005). The Mediterranean nature of most of Spain produces highly-resilient ecosystems, because their ecosystems obtain their stability by adjusting their dynamics to couple with climatic local perturbation regime (García and Montes, 2003). Mediterranean way of life has been characterized by the reproduction of these patterns (management of fire, water, etc.) in a smaller scale to avoid great perturbations (wild fires, flooding, etc.). Today, the Mediterranean standard of living is endangered, and there is an effort to preserve some of it characteristics. In this sense, although the Mediterranean nature of the Spanish social–ecological system guarantees a high level of ecological resilience *sensu* Holling (1973), management policies, distant from Mediterranean traditional management that was its identity in the past, are not succeeding in preserving the flow of environmental goods and services that supports our economy. As we have seen in the results of this eMergy synthesis figures, Spain is still in the reversible phase of its economic evolution: in other words, it is more endangered than irreversibly degraded.

A transition to a global and coherent landscape management that overcomes the current dichotomy between territories exclusively managed for conservation and those exclusively dedicated to production is needed. In a Mediterranean context, this goal would be achieved by a landscape management proposal in which natural protected areas contributed to the preservation of a heterogeneous mosaic of traditional uses, in which different ecosystems in many states of maturity that changed with time would be combined and complemented (Burel and Baudry, 1995, 1999; Farina, 1997; González Bernáldez, 1991, 1992). Also, a real integration of conservation practices and the sustainable use of biological diversity with other sectoral or cross sectoral activities, plans and programs that have and impact upon them, is desirable.

Acknowledgements

For discussion and ideas we would like to thank Sergio Ulgiati, from University of Naples "Parthenope", Mark T. Brown, from the Florida University, and Fernando Valladares, from Environmental Sciences Centre of the Spanish National Research Council in Madrid. We would also thank the Spanish Meteorology Institute, the Spanish Oceanographic Institute and the Spanish Port Organism staffs for their help with some geophysical variables used in the synthesis. Also, we would like to thank Phil Mason and Daniel Welsch for their English language edition. Finally, we would like to thank two anonymous reviewers for their critical comments about the earlier draft paper. Financial support for this research activity has been made available thanks to the Andalusian Regional Government Department of Environment (Project NET413308/1).

Appendix A

Main emergy flows supporting national economies for Spain and other selected countries, arranged by EMR are given in Table A1.

Table A1
Main emergy flows supporting national economies for Spain and other selected countries, arranged by EMR[a]

Country	U (E+20 sej/year)	Renewable (E+20 sej/year)	Non-renewable (E+20 sej/year)	Population (E+6 inhabitants)	GDP (E+9 US$/year)	Emergy per capita (E+15 sej/inhab)	EMR (E+12 sej/US$)
Nicaragua 1994	816.06	720.00	90.00	4.51	1.40	18.09	58.29
Zambia 1997	1250.00	1030.00	220.00	8.96	2.50	13.94	50.00
Morocco 1994	976.21	380.00	600.00	28.55	8.28	3.42	11.79
Argentina 1994	4520.00	1940.00	2580.00	35.66	54.80	12.67	8.25
Kenia 1999	765.60	370.00	390.00	29.35	10.24	2.61	7.48
India 1999	26 210.60	6750.00	19 440.00	442.00	442.00	2.62	5.93
Spain 1984	**9190.00**	**609.00**	**5090.00**	**38.28**	**164.00**	**24.00**	**5.62**
Syria	790.00	90.00	700.00	15.02	17.00	5.25	4.64
Italy 1984	16 100.00	2030.00	5040.00	56.64	390.00	28.40	4.12
Canada 1999	23 359.05	7800.00	15 550.00	30.49	598.95	76.56	3.90
Spain 1989	**13 100.00**	**628.00**	**7050.00**	**38.79**	**394.00**	**33.70**	**3.31**
Saudi Arabia 1994	7953.00	2580.00	5370.00	22.03	241.00	36.11	3.30
Italy 2000	37 900.00	2030.00	7430.00	57.84	1210.86	65.50	3.13
Spain 2000	**17 400.00**	**609.00**	**7550.00**	**39.93**	**562.00**	**43.50**	**3.09**
Brazil 1995	17 880.00	6870.00	8830.00	167.20	600.00	10.71	2.98
Italy 2002	34 700.00	2030.00	5850.00	57.30	1176.27	60.50	2.95
Spain 2002	**19 100.00**	**609.00**	**8470.00**	**40.55**	**655.00**	**47.00**	**2.91**
Spain 1994	**12 500.00**	**590.00**	**5490.00**	**39.17**	**504.00**	**32.00**	**2.49**
Italy 1989	21 300.00	2030.00	6000.00	56.70	866.00	37.50	2.45
Bolivia 1997	195.20	180.00	10.00	8.04	8.00	2.43	2.44
Italy 1995	25 900.00	2030.00	8020.00	57.33	1070.00	45.10	2.41
South Africa 1999	9270.00	2400.00	6860.00	43.20	412.00	21.44	2.25
Uruguay 1995	308.70	200.00	110.00	3.22	14.70	9.56	2.10
Italy 1991	23 200.00	2030.00	8430.00	56.76	1150.00	40.90	2.02
Global economy 1999	**510 350.00**	**158 600.00**	**343 700.00**	**5900.00**	**27 100.00**	**8.52**	**1.85**
Netherlands 1994	6789.30	250.00	6550.00	15.67	371.00	43.40	1.83
USA 1999	90 100.00	8380.00	81 620.00	266.56	8500.00	33.76	1.75
Denmark 1997	1786.40	20.00	1760.00	5.35	123.20	33.27	1.45
Switzerland 1999	2538.00	280.00	2260.00	7.20	270.00	35.25	0.94
Ireland 1994	469.56	70.00	400.00	3.67	54.60	12.74	0.86
Japan 1999	36 000.00	1330.00	34 600.00	126.97	4500.00	28.30	0.80
Germany 1995	15 257.77	220.00	15 100.00	83.03	2090.10	18.43	0.73

[a]Data source for selected countries, Brown (2003), for Italy Cialani etal., (2005), and this study for Spain.

Appendix B

Some of the main emergy indicators for Spain and other selected countries, arranged by ESI are given in Table B1.

Table B1
Some of the main emergy indicators for Spain and other selected countries, arranged by ESI[a]

Country	U (E+20 sej/year)	EER [EMR_{ge}/EMR_i][b]	ELR [$(U-R)/R$]	EYR [$U/(N_0+N_1+IMP)$]	ESI [EYR/ELR]
Bolivia 1997	195.20	0.76	1.07	15.00	14.00
Nicaragua 1994	816.06	0.03	1.14	8.33	7.33
Zambia 1997	1 250.00	0.04	1.21	5.68	4.68
Uruguay 1995	308.70	0.88	1.57	2.75	1.75
Kenia 1999	765.60	0.25	2.05	1.95	0.95
Brazil 1995	17 880.00	0.62	2.61	2.03	0.78
Argentina 1994	4520.00	0.22	2.33	1.75	0.75
Canada 1999	23 359.05	0.47	1.99	1.50	0.75
Global economy 1999	**510 350.00**	—	**2.17**	**1.46**	**0.70**
Morocco 1994	976.21	0.16	2.56	1.64	0.64
Saudi Arabia 1994	7953.00	0.56	3.08	1.48	0.48
India 1999	26 210.60	0.31	3.88	1.35	0.35
South Africa 1999	9270.00	0.82	3.85	1.35	0.35
Italy 1984	16 100.00	0.45	6.91	1.78	0.26
Spain 1984	**9190.00**	**0.33**	**14.10**	**2.41**	**0.17**
Italy 1991	23 200.00	0.92	10.46	1.76	0.17
Italy 1989	21 300.00	0.76	9.47	1.61	0.17
Ireland 1994	469.56	2.16	6.87	1.17	0.17
Italy 1995	25 900.00	0.77	11.72	1.59	0.14
Spain 1989	**13 100.00**	**0.56**	**19.79**	**2.34**	**0.12**
Syria	790.00	0.40	9.20	1.12	0.12
Switzerland 1999	2538.00	1.97	9.10	1.12	0.12
USA 1999	90 100.00	1.28	10.74	1.01	0.10
Spain 1994	**12 500.00**	**0.74**	**20.26**	**1.87**	**0.09**
Italy 2002	34 700.00	0.63	16.13	1.29	0.08
Italy 2000	37 900.00	0.59	17.65	1.33	0.08
Spain 2000	**17 400.00**	**0.60**	**27.55**	**1.84**	**0.07**
Spain 2002	**19 100.00**	**0.64**	**30.32**	**1.87**	**0.06**
Japan 1999	36 000.00	2.32	27.06	1.04	0.04
Netherlands 1994	6789.30	1.01	27.20	1.04	0.04
Germany 1995	15 257.77	2.53	69.55	1.01	0.01
Denmark 1997	1786.40	1.75	89.00	1.01	0.01

[a] *Data source*: For selected countries Brown (2003), for Italy Cialani et al. (2005), and this study for Spain.
[b] EMRge = EMR of global economy; EMRi = EMR of the country.

Appendix C

Calculations of the average standard of living (ESL and MSL) and regional ELR to be used in carrying capacity and support area for selected European and Mediterranean Basin countries are given in Table C1.

Table C1
Calculations of the average standard of living (ESL and MSL) and regional ELR to be used in carrying capacity and support area for selected European and Mediterranean Basin countries

European Countries	Total emergy actually used (U_E) (sej/year)	Renewable emergy used (R_E) (sej/year)	Area (m^2)	Analysis year	R_E/U_E	ESL = U_E/R_E	ELR
Spain	1.25E+24	5.90E+22	4.98E+11	1994	0.05	20.26	19.67
Italy	2.26E+24	1.21E+23	3.01E+11	1995	0.08	12.73	11.72
Czech Republic	1.55E+23	5.60E+22	7.90E+10	1998	0.36	2.77	2.57
Finland	1.20E+23	2.70E+22	3.38E+11	1994	0.23	4.44	4.44
Ireland	4.70E+22	7.00E+21	8.40E+10	1994	0.15	6.71	6.87
Portugal	1.76E+23	1.70E+22	9.20E+10	1995	0.10	10.35	10.35

Table C1 (*continued*)

European Countries	Total emergy actually used (U_E) (sej/year)	Renewable emergy used (R_E) (sej/year)	Area (m^2)	Analysis year	R_E/U_E	ESL = U_E/R_E	ELR
Slovakia	6.70E+22	6.00E+21	4.90E+10	1994	0.09	11.17	11.75
France	1.32E+24	8.30E+22	5.91E+11	1999	0.06	15.90	15.92
Netherlands	6.80E+23	2.50E+22	4.10E+10	1994	0.04	27.20	27.20
England	2.82E+24	8.30E+22	1.30E+11	1999	0.03	33.95	34.05
Germany	1.53E+24	2.20E+22	3.57E+11	1995	0.01	69.55	69.55
Austria	2.59E+23	1.60E+22	8.39E+10	1997	0.06	16.19	16.19
Switzerland	2.54E+23	2.80E+22	4.13E+10	1999	0.11	9.07	9.10
Denmark	1.78E+23	2.00E+21	4.31E+10	1997	0.01	89.00	89.00
		Total R_E = 5.52E+23	Total area of European countries used = 2.73E+12		Average =	ESL = 24.02	ELR(*r*) = 23.50
					SD =	25.23	25.41
Mediterranean basin countries	Total emergy actually used (U_M) (sej/year)	Renewable emergy used (R_M) (sej/year)	Area (m^2)	Analysis year	R_M/U_M	MSL = U_M/R_M	ELR
Spain	1.25E+24	5.90E+22	4.98E+11	1994	0.05	20.72	19.67
Italy	2.26E+24	1.21E+23	3.01E+11	2000	0.05	18.68	17.65
France	1.32E+24	8.30E+22	5.91E+11	1999	0.06	15.90	15.92
Morocco	9.80E+22	3.80E+22	4.44E+11	1994	0.39	2.58 2	2.56
Syria	7.90E+22	9.00E+21	1.85E+11	1997	0.11	8.78	9.20
		Total R_M = 3.10E+23	Total area of Mediterranean countries used = 2.02E+12		Average =	MSL = 13.44	ELR(*r*) = 13.12
					SD =	7.65	7.18

Data source: Brown (2003) and Cialani et al. (2005).

References

AEAT (Agencia Estatal de Administración Tributaria) [On-line]. Información estadistica sobre el comercio exterior de España-Diciembre 2000 [Consulted: 2/2/2006]. Ministerio de Hacienda. Departamento de Aduanas e II.EE., Madrid. ⟨http://www.aeat.es/aduanas/descarga/est1200.pdf⟩.

Arrojo, P (Co-ord), 2001. El Plan Hidrológico Nacional a debate. Colección Nueva Cultura del Agua, 8. Bakeaz/Fundación Nueva Cultura del Agua, Bilbao.

BDE (Banco de España) [On-line]. Balanza de Pagos y posición de inversión internacional frente a otros residentes en la Zona del Euro y al resto del mundo. Boletín Estadístico del Banco de España [Consulted: 2/2/2006]. Banco de España, Madrid. ⟨http://www.bde.es/infoest/a1603.pdf⟩.

Berkes, F., Folke, C. (Eds.), 1998. Linking Social and Ecological Systems. Cambridge University Press, Cambridge.

Brown, M.T., Ulgiati, S., 1997. Emergy-based indices and ratios to evaluate sustainability monitoring economies and technology toward environmentally sound innovation. Ecological Engineering 9, 51–69.

Brown, M.T., Ulgiati, S., 2001. eMergy measures of carrying capacity to evaluate economic investments. Population and Environment 22 (5), 471–501.

Brown, M.T., Ulgiati, S., 2004. eMergy analysis and environmental accounting. Encyclopedia of Energy 2, 329–354.

Brown, M.T., 2003. Resource imperialism: eMergy perspectives on sustainability, international trade, and balancing the welfare of nations. In: Ulgiati, S., Brown, M.T., Giampietro, M., Herendeen, R., Mayumi, K. (Eds.), Proceedings of Third Biennial International Workshop. Advances in Energy Studies: Reconsidering the Importance of Energy. Porto Venere, Italy, pp. 135–149.

Brown, M.T., 2004. A picture is worth a thousand words: energy systems language and simulation. Ecological Modelling 178 (1–2), 83–100.

Burel, F., Baudry, J., 1995. Species biodiversity in changing agricultural landscapes: a case study in the Pays d'Auge, France. Agriculture Ecosystems & Environment 55, 193–200.

Burel, F., Baudry, J., 1999. Ecologie du paysage. Concepts, methods et applications. Edition Tec et Doc, París, France.

Butzer, K.W., 2005. Environmental history in the Mediterranean world: cross-disciplinary investigation of cause- and effect for degradation and soil erosion. Journal of Archaeological Science 32, 1773–1800.

Campbell, D., 1998. eMergy analysis of human carrying capacity and regional sustainability: an example using the state of Maine. Environmental Monitoring and Assessment 51, 531–569.

Cañellas, S., Citalic, A., Puig, I., Russi, D., Sendra, C., Sojo, A., 2004. Material flow accounting of Spain. Journal of Global Environmental Issues 4 (4), 229–241.

Carpintero, O., 2005. El metabolismo de la economía española. Recursos naturales y huella ecológica (1955–2000). Fundación César Manrique. Islas Canarias (España).

Casado, S., Montes, C., 1991. Estado de conservación de los humedales peninsulares españoles. Quercus 66, 18–26.

Cerdá, A., 2001. La erosión del suelo y sus tasas en España. [Consulted: 2/2/2006]. Ecosistemas, 3, Año X., Asociación Española de Ecología Terrestre. ⟨http://www.aeet.org/ecosistemas/013/revisiones1.htm⟩.

Cialani, C., Russi, D., Ulgiati, S., 2005. Investigating a 20-year national economic dynamics by means of eMergy-based indicators. In: Brown, M.T., Bardi, E., Tilley, D., Ulgiati, S. (Eds.), eMergy Synthesis 3: Theory and Applications of the eMergy Methodology. Proceedings of the Third Biennial eMergy Conference. Center for Environmental Policy, University of Florida, Gainesville, FL, USA, pp. 325–334.

Cincotta, R.P., Wisnewski, J., Engelman, R., 2000. Human population in the biodiversity hotspots. Nature 404, 990–992.

Costanza, R., 1997. La Economía Ecológica de la sostenibilidad Invertir en capital natural. In: Goodland, R., Daly, H., El Serafy, S., von Droste, B. (Eds.). Medio Ambiente y desarrollo sostenible. Más allá del Informe Brundtland, Editorial Trotta, Serie Medio Ambiente, Madrid, pp. 103–114.

Costanza, R., Cumberland, J., Daly, H., Goodland, R., Norgaard, R., 1997. An Introduction to Ecological Economics. St. Lucie Press, FL, USA.

Crutzen, P.J., Stoermer, E.F., 2000. The Anthropocene IGBP Newsletter, 41. Royal Swedish Academy of Sciences, Stockholm, Sweden.

Cuadrado, J.R., 1999. Sector servicios: una visión de conjunto. In: García Delgado, J.L. (Dir.). Economía ante el s. XXI. Ed. Espasa Calpe, España, Madrid.

Daly, H.E., 1991. Ecological Economics and Sustainable Development: From Concept to Policy. The World Bank, Washington, DC, USA.

De Groot, R., Van der Perk, J., Chiesura, A., van Vliet, A., 2003. Importance and threat as determining factors for criticality of natural capital. Ecological Economics 44, 187–204.

De Miguel, J.M., Gómez-Sal, A., 2002. Diversidad y funcionalidad de los paisajes agrarios tradicionales. In: Montalvo, J. (Coords.). La diversidad Biológica de España. Pearson Education S.A., Madrid, pp. 273–284.

EEA (European Environment Agency), 2005. The European Environment—State and Outlook 2005 [Consulted: 2/2/2006]. Copenhagen, Denmark. 〈http://reports.eea.eu.int/state_of_environment_report_2005_1/en/tab_content_RLR〉.

EU/EWEA (European Commission-DG TREN/European Wind Energy Association), 2003. Wind Energy-the facts. An analysis of wind energy in the EU25 [Consulted: 2/2/2006]. European Wind Energy Association. 〈http://www.ewea.org/index.php?id = 91〉.

EUROPARC-España, 2004. Anuario EUROPARC-España del estado de los espacios naturales protegidos 2003. Fundación Fernando González Bernáldez, Madrid, 123pp.

EUROSTAT [On-line-a]. Energy Yearly Statistics: Data, 2001 [Consulted: 2/2/2006]. 〈http://epp.eurostat.cec.eu.int/cache/ITY_OFFPUB/KS-CN-03-001-3A/EN/KS-CN-03-001-3A-EN.PDF〉.

EUROSTAT [On-line-b]. Environmental protection expenditure in Europe by public sector and specialised producers 1995–2002 [Consulted: 2/2/2006]. 〈http://epp.eurostat.cec.eu.int/cache/ITY_OFFPUB/KS-NQ-05-010/EN/KS-NQ-05-010-EN.PDF〉.

Farina, A., 1997. Landscape structure and breeding bird distribution in a sub Mediterranean agro-ecosystem. Landscape Ecology 12, 365–378.

Gadea, M.D., 2000. La peseta en la cultura de la estabilidad, 1989–1999. In: García Delgado, J.L., Serrano Sanz, J.M. (Dirs.), Del real al euro. Una historia de la peseta. Colección Estudios Económicos, vol. 21. Servicio de Estudios de La Caixa, Barcelona, pp. 169–190.

García, R., Montes, C., 2003. (Eds.). Vínculos en el paisaje mediterráneo. El papel de los espacios protegidos en el contexto territorial/Linkages in the Mediterranean Landscape. The Role of Protected Areas in the Territorial Context. Junta de Andalucía, Sevilla, 216pp.

González Bernáldez, F., 1991. Diversidad biológica, gestión de ecosistemas y nuevas políticas agrarias. In: Pineda, F.D., Casado, M.A., de Miguel, J.M., Montalvo, J. (Eds.), Biological Diversity/Diversidad Biológica, Fundación Areces, WWF/ADENA, SCOPE, Madrid, pp. 23–31.

González Bernáldez, F., 1992. Ecological consequences of the abandonment of traditional land use systemes in central Spain. Options Mediteranéennes Series Seminars 15, 23–29.

Goodland, R., Daly, H., 1996. Environmental sustainability: universal and non-negotiable. Ecological Applications 6 (4), 1002–1017.

Grove, A.T., Rackham, O., 2003. The Nature of Mediterranean Europe: An Ecological History, second ed. Yale University Press, London, UK.

Hau, J.L., Bakshi, B.R., 2004. Promise and problems of eMergy analysis. Ecological Modelling 178 (1), 215–225.

Henning, D., 1989. Atlas of the Surface Heat Balance of the Continents. Gebruder Borntraeger, Berlin, Germany.

Holling, C.S., 1973. Resilience and stability of ecological systems. Annual Review of Ecology and Systematics 4, 1–23.

IEA (International Energy Agency), 2003. Energy Balances of OECD Countries 2000–2001. Paris, France.

IGME (Instituto Geológico y Minero de España) [On-line]. Panorama minero España 1993–2002. [Consulted: 2/2/2006]. 〈http://www.igme.es/internet/RecursosMinerales/indexc.htm〉.

IGN (Instituto Geográfico Nacional), 1996. Atlas Nacional de España Tomo I. Centro Nacional de Información Cartográfica. Ministerio de Fomento, Madrid.

INE (Instituto Nacional de Estadística) [On-line], 2006. Padrón municipal Cifras oficiales de población [Consulted: 2/2/2006]. 〈http://www.ine.es/inebase/cgi/um?M = %2Ft20%2Fe260&O = inebase&N = &L〉.

INM (Instituto Nacional de Meteorología), 2001. Calendario Meteorológico 2001. Serie Monografías. Secretaría General Técnica, Madrid.

INM (Instituto Nacional de Meteorología), 2002. Calendario Meteorológico 2002. Serie Monografías. Secretaría General Técnica, Madrid.

Llamas, R., 2000. Análisis de los sistemas hidráulicos. Recursos, demandas, balances. [Consulted: 2/2/2006]. In: Jornadas Colegio de Ingenieros de Caminos, Canales y Puertos, Madrid, 〈http://www.unizar.es/fnca/docu1.php〉.

MA (Millenimum Ecosystem Assessment), 2005. Ecosystems and Human Well-being: Synthesis. Island Press, Washington, DC, USA.

Makhzoumi, J., Pungetti, G., 1999. Ecological Landscape Design and Planning: The Mediterranean Context. E and FN Spon, London, UK.

MAPA (Ministerio de Agricultura, Pesca y Alimentación), 1991. Anuario de estadísticas agrarias—1990. Secretaría General Técnica, Madrid.

MAPA (Ministerio de Agricultura, Pesca y Alimentación) [On-line]. Anuario de Estadística Agroalimentaria—2003 [Consulted: 2/2/2006]. 〈http://www.mapya.es/es/estadistica/pags/anuario/Anu_03/indice.asp〉.

MAPA (Ministerio de Agricultura, Pesca y Alimentación), 1998. Anuario de Estadística Agraria—1997, Madrid.

Martínez-Alier, J., 1999. Introducción a la Economía Ecológica. Editorial, Rubes, Barcelona.

MFOM (Ministerio de Fomento) [On-line-a]. Plan Estratégico de Infraestructuras y Transporte. Horizonte 2005–2020 [Consulted: 2/2/2006]. Secretaría General Técnica, Madrid. 〈http://www.fomento.es/MFOM/LANG_CASTELLANO/INFORMACION_MFOM/PEIT2/050711-1.htm〉.

MFOM (Ministerio de Fomento) [On-line-b]. Anuario Estadístico 2004. [Consulta: 2/2/2006]. Secretaría General Técnica, Madrid. 〈http://www.fomento.es/MFOM/LANG_CASTELLANO/INFORMACION_MFOM/INFORMACION_ESTADISTICA/Anuari2004/anuario.htm〉.

MIMAM (Ministerio de Medio Ambiente), 1999. Estrategia español para la conservación y el uso sostenible de la diversidad biológica. Secretaría General de Medio Ambiente. Dirección General de Conservación de la Naturaleza, Madrid.

MIMAM (Ministerio de Medio Ambiente), 2000. Libro blanco del agua en España. Dirección General de Obras Hidráulicas. Secretaría General Técnica, Madrid.

Morillo, C., Gómez-Campo, C., 2000. Conservation in Spain, 1980–2000. Biological Conservation 95, 165–174.

Mulligan, M., Burke, S., Ramos, C., 2004. Climate change, land-use change and the "Desertification" of Mediterranean Europe. In: Mazzoleni, S., di Pasquale, G., Mulligan, M., di Martino, P., Rego, F. (Eds.). Recent Dynamics of the Mediterranean Vegetation and Landscape. Wiley, Chichester, pp. 259–279.

Muradian, R., O'Connor, M., Martínez-Alier, J., 2002. Embodied pollution in trade: estimating the "environmental load displacement" of industrialised countries. Ecological Economics 41, 51–67.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.

Naveh, Z., Liberman, A., 1993. Landscape Ecology. Theory and Application. Springer, New York, USA.

Odum, H.T., Odum, E.C., 1983. Energy Analysis Overview of Nations Working Paper, WP-83-82. International Institute for Applied System Analysis, Laxenburg, Austria.

Odum, H.T., 1994. Ecological and General Systems: An Introduction to Systems Ecology, revised ed. University Press of Colorado, Niwot, CO, USA.

Odum, H.T., 1996. Environmental Accounting: eMergy and Decision Making. Wiley, New York, USA.

Odum, H.T., Brown, M.T., Brandt-Williams, S., 2000. Folio #1. Introduction and Global Budget. Handbook of eMergy Evaluation. A Compendium of Data for eMergy Computation issued in a Series of Folios. Center for Environmental Policy, Gainesville, FL, USA.

OSE (Observatorio de la Sostenibilidad en España), 2005. Sostenibilidad en España 2005, Informe de primavera. Observatorio de la Sostenibilidad en España. Alcalá de Henares, España.

Pineda, F.D., Montalvo, J., 1995. Biological diversity in dehesa systems. In: Gilmour, D. (Ed.), Biological Diversity Outside Protected Areas. Overview of Traditional Agroecosystems, IUCN, Forest Conservation Programme, Gland Suiza, pp. 107–122.

Porta, J., López-Acebedo, M., Roquero, C., 1994. Edafologia para la agricultura y el medio ambiente. Mundi-Prensa, Madrid.

Postel, S.L., Daily, G.C., Ehrlich, P.R., 1996. Human appropriation of renewable fresh water. Science 271, 785–788.

Prats, N., Munné, A., Rieradevall, M., Bonada, N., 2000. La determinación del estado ecológico de los sistemas acuáticos en España. In: Fabra, A., Barreira, A. (Eds.), La aplicación de la Directiva Marco del Agua en España. Instituto Internacional de Derecho y Medio Ambiente, Barcelona, pp. 47–81.

OAPEstado [On-line]. Base de datos oceanográficos [Consulted: 2/2/2006]. Ministerio de Fomento, Madrid, ⟨http://www.puertos.es/index2.jsp?langId=1&catId=1014806377970&pageId=1037009598954⟩.

Ramos Martin, J., 2001. Non-linearity in Energy Metabolism of Spain: Attractor Points for the Development of Energy Intensity. In: Ulgiati, S., Brown, M.T., Giampietro, M., Herendeen, R.A., Mayumi, K. (Eds.), Proceedings of the Second Biennial International Workshop in Advances in Energy Studies. Exploring Supplies, Constraints and Strategies, Padova, Italy, pp. 535–542.

Ramos Martin, J., 2003. Intensidad energética de la economia española: una perspectiva integrada. Economia Industrial 351, 59–72.

Raugei, M., Bargigli, S., Ulgiati, S., 2004. eMergy "Yield" Ratio: problems and misapplications. In: Brown, M.T., Bardi, E., Tilley, D., Ulgiati, S. (Eds.), eMergy Synthesis 3: Theory and Applications of the eMergy Methodology. Proceedings of the Third Biennial eMergy Conference. Center for Environmental Policy, University of Florida, Gainesville, FL, USA, pp. 105–124.

Rojstaczer, S., Sterling, S.M., Moore, N.J., 2001. Human appropriation of photosynthesis products. Science 294, 2549–2552.

Rosnay, J., 1979. The Macroscope: A New World Scientific System. Harper & Row, Publishers Inc., New York, USA.

Schmitz, M.F., de Aranzabal, I., Rescia, A., Pineda, F.D., 2001. Implications of socioeconomic changes in Mediterranean cultural landscapes. In: Mander, U., Printsmann, A., Palang, H. (Eds.), Development of European Landscapes, vol. 92. Publicaciones Instituti Geographici Universitatis, Tartuensis, pp. 788–793.

Soto, D., 1990. Aproximación a la medida de la erosión y medios para reducir ésta en la España peninsular. Ecologia 1, 169–196.

Tamames, R., Rueda, A., 2000. Estructura económica de España, 24th ed. Alianza Editorial, Madrid.

Ten Brink, B., 2000. Biodiversity indicators for the OECD environmental outlook and strategy, RIVM Rapport, 402001014, Globo Report Series, 25.

Tilley, D.R., Swank, W.T., 2003. EMERGY-based environmental systems assessment of a multi-purpose temperate mixed-forest watershed of the southern Appalachian Mountains, USA. Journal of Environmental Management 69, 213–227.

Ulgiati, S., Brown, M.T., 1998. Monitoring patterns of sustainability in natural and man-made ecosystems. Ecological Modelling 108, 23–36.

UNSD (United Nations Statistical Division) [On-line]. GDPmp for Spain. National Accounts Main Aggregates Database. Basic Data Selection [Consulted: 2/2/2006]. ⟨http://unstats.un.org/unsd/snaama/selection-basicFast.asp⟩.

Verdnasky, V.I., 1945. The biosphere and the noosphere. American Scientist 33, 1–12.

Vitousek, P., Ehrlich, P., Ehrlich, A., Matson, P., 1986. Human appropriation of the products of photosynthesis. Bioscience 36 (6), 368–373.

Vitousek, P.M., Mooney, H.A., Lubchenco, J., Melillo, J.M., 1997. Human domination of Earth's ecosystems. Science 277 (5325), 494–499.

WCD (World Commission on Dams), 2000. Dams and Development. A new framework for Decision Making: The Report of the World Commission on Dams. Earthscan, London, UK.

WWF/WCMC-UNEP (World Wildlife Fund/World Conservation Monitoring Centre—United Nations Environmental Program), 2004. Informe Planeta Vivo 2004. [Consulted: 2/2/2006]. Gland, Switzerland, ⟨http://www.wwf.es/planeta_vivo04.php⟩.

ELSEVIER

Journal of Environmental Management 88 (2008) 348–359

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Baseline assessment for environmental services payments from satellite imagery: A case study from Costa Rica and Mexico

M. Kalacska[a], G.A. Sanchez-Azofeifa[a,*], B. Rivard[a], J.C. Calvo-Alvarado[b], M. Quesada[c]

[a]Earth and Atmospheric Sciences Department, University of Alberta, Edmonton, Alta., Canada T6G 2E3

[b]School of Forestry Engineering, Costa Rican Technology Institute (ITCR), P.O. Box 159-7050, Cartago, Costa Rica

[c]Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México, Apartado Postal 27-3 (Xangari), 58089 Morelia, México

Received 16 February 2006; received in revised form 2 February 2007; accepted 10 March 2007
Available online 1 May 2007

Abstract

In this study we evaluate the accuracy of four global and regional forest cover assessments (MODIS, IGBP, GLC2000, PROARCA) as tools for baseline estimation. We conduct this research at the national scale for Costa Rica and for two tropical dry forest study sites in Costa Rica (Santa Rosa) and Mexico (Chamela-Cuixmala). We found that at the national level, the total forest cover accuracy of the four land cover maps was inflated due to an overestimation of forest in areas with an evergreen canopy. However, the four maps greatly underestimated the extent of the deciduous forest (dry forest); an ecosystem that faces high deforestation pressure and poses complications to the mapping of its extent from remotely sensed data. For the tropical dry forest sites, all maps have low forest cover accuracies (mean for Santa Rosa: 27%; mean for Chamela-Cuixmala: 56%). This has implications for policy implementation.

Keywords: Carbon; Chamela-Cuximala; Costa Rica; Forest cover; Global land cover; Mexico; Payments for environmental services; Santa Rosa; Tropical dry forest; Tropical wet forest

1. Introduction

The implications of global climate change due to an increasing build-up of anthropogenic greenhouse gasses (GHG) such as carbon dioxide (CO_2) in the atmosphere have become internationally important issues (Choi, 2005; Gerlagh and van der Zwann, 2006; IPCCC, 2005; Lenton, 2006; Litynski et al., 2006) The Kyoto Protocol was proposed as potential means to mitigate climate change (United Nations, 1997). Each country listed in Annex B of the Protocol is allocated an "assigned amount of emission reduction units" (AAUs) that correspond to the nations' allowable GHG emissions (Choi, 2005; United Nations, 1997). Annex B countries include developed countries and countries with transitional economies. The AAUs are a country's baseline emission minus the percentage of emission reductions that are required by the Protocol (Choi, 2005). While most countries are required to reduce their emissions, a few such as Iceland are permitted an increase (United Nations, 1997).

Specific mechanisms are permitted under the Kyoto Protocol to help parties attain their GHG emission reduction targets: joint implementation (JI) and the Clean Development Mechanism (CDM)(Choi, 2005; Olschewski and Benitez, 2005; Pfaff et al., 2000; UNFCCC, 2003b). The JI and CDM mechanisms are project-based where an Annex B country can earn emission reduction credits by participating in a project that helps another country reduce its GHG emissions (e.g. reforestation). The difference between the JI and CDM is that JI projects are between Annex B countries whereas in CDM projects Annex B countries receive certified emission reduction units (CERs) by collaborating with non-Annex B countries (Choi, 2005; Olschewski and Benitez, 2005; UNFCCC, 2003b). Annex B countries may engage in either JI or CDM projects to help another Annex B or non-Annex B country and use the credits for their own compliance (Subak, 2000). Anthropogenic land-based activities—"land-use, land-use change and forestry" (LULUCF) are a means of achieving

*Corresponding author. Tel.: +1 780 492 1822; fax: +1 780 492 2030.
E-mail address: arturo.sanchez@ualberta.ca (G.A. Sanchez-Azofeifa).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.03.015

compliance to the Protocol (UNFCCC, 2001b, 2003a, b). While a part of the CDM, LULUCF activities are limited to afforestation and reforestation projects. Also, during the first commitment period, the credits resulting from LULUCF projects are capped at 1% of the base year emissions times five (Choi, 2005; UNFCCC, 2003b). Nevertheless, such activities do form part of the initiatives taken by Annex B countries. An example of a recent project in the tropics is the investment of US$2 million by the Government of Norway into Costa Rica's Private Forestry Project in exchange for 230 kt of carbon offsets (US$10/t/C) (Subak, 2000).

To evaluate the effectiveness of mitigation projects in the forestry sector and the clear definition of solid baselines, three fundamental questions must be addressed: *what is the initial extent of the forests?, what type of forest is there (primary, secondary)? what is the rate of change of the forest extent?* Estimates of payments for environmental services are greatly dependent upon the differences between the baseline and mitigation scenarios and on deforestation rates before and after the implementation of a project; the greater the difference the greater the estimate of carbon sequestration/value of environmental services (Busch et al., 2000). Therefore, it is imperative that the initial state and extent of the forest (baseline determination) be characterized as accurately as possible, a problem that is not trivial due to the lack of standardized mechanisms that can provide accurate information for different types of forests especially in the tropics (Foster et al., 2002; Subak, 2000).

While CDM projects can take place in any ecosystem, in this study we focus on tropical forests because of their importance in carbon sequestration, their general overall threatened status and the complications that arise when mapping their extent from remotely sensed data (Sánchez-Azofeifa et al., 2001). To accurately assess land cover changes in tropical environments at reasonable costs requires remote sensing technology (Subak, 2000). However, an initial problem in accurate estimation of forest cover is one of nomenclature (Jung et al., 2006). While there are numerous definitions of what is a forest (see (ITTO, 2002) for an in depth review), there is no general consensus (Jung et al., 2006). In addition, many definitions are biased towards mature wet or rain forests, neglecting seasonally deciduous forests and stages of vegetation succession (i.e. secondary forests) (Sánchez-Azofeifa et al., 2005a, b). Attempts have been made to consolidate definitions during the Marrakesh Accords (UNFCCC, 2001a, 2003a). Based on those accords "forest" is defined as "*a minimum area of land of* 0.05–1.0 *ha with tree crown cover, or equivalent stocking level, of more than* 10–30% *and containing trees with the potential to reach a minimum height of* 2–5 *m at maturity*". This is also the definition adopted by the Eleventh Conference of the Parties when discussing the implementation of CERs (UNFCCC, 2003b). In addition, stands temporarily below the thresholds but which are expected to grow or revert to forest are also included in the forest category (UNFCCC, 2001a, 2003b). However, the definition of "forest" adopted by any one country is optional within the stated minimum levels defined by the Marrakesh Accords. In this study we adopt the Marrakesh Accord's definition of "forest" and more precisely the following refinement describing the environment where dry forests are found: an *area with a vegetation type dominated by deciduous trees located in an area with a mean temperature* $>25\,^{\circ}C$*, a total annual precipitation range of* 700–2000 *mm and three or more dry months* (*precipitation* <100 *mm*) (Sánchez-Azofeifa et al., 2005b).

Methods for using remote sensing to monitor and detect tropical deforestation in the humid tropics have been successfully developed, tested and applied ((Sánchez-Azofeifa et al., 2001; Skole and Tucker, 1993; Stone and Lefebvre, 1998; Zhang et al., 2005) among many others), providing important information on the extent of tropical evergreen forests. Tropical dry deciduous forests (T-df), however, have received less attention and thus the development of remote sensing methods for quantifying the extent of the T-df has been neglected in comparison to wet/rain forests (Arroyo-Mora et al., 2005b). Significant errors have resulted from mapping the extent of the tropical dry forest from satellite images because the cloud free images are most easily acquired during the dry season when an increased percentage of the canopy is leafless, lacking the spectral signature of green leaf biomass (Arroyo-Mora et al., 2005a; Asner, 1998; Kalacska et al., 2007). This property of the canopy induces the misinterpretation of forested areas in the image for pastures or areas with dispersed trees (Pfaff et al., 2000).

Our study investigates the implications of using various global remote sensing derived land cover classification data sets of forest cover as a baseline scenario at the national level for Costa Rica. We use Costa Rica as an example for several reasons. Firstly, it has a multitude of forest types (from dry forest to rain forest) and also has a considerable history in payment of environmental services. In addition, a substantial database of remotely sensed imagery both satellite and airborne have been acquired for the entire country, including multi and hyperspectral images and aerial photograph that cannot be matched by any other tropical or sub-tropical country. Significant efforts have been made by the Costa Rican Centre for High Technology (CENAT- El Centro Nacional de Alta Tecnologia) and its National Airborne and Remote Sensing Program (El Programa Nacional de Investigaciones Aerotransportadas y Sensores Remotos) to ensure continual collection of remotely sensed data for the country. Finally, substantial efforts have been in the country through the national ministry of environment (MINAE), NGOs such as FONAFIFO and several provincial agencies such as FUNDECOR to manage and inventory the nation's forests. Costa Rica has the largest proportion of national territory in protected areas (national parks, absolute reserves, biological reserves) of any Central American country. Subsequently, we focus specifically on the dry forest in two protected areas located in Costa Rica and

Mexico and assess how the different interpretations of the global land cover classifications affect the quantification of the value associated with environmental services in the dry forest. In terms the tropical dry forest in Central America, Costa Rica and Mexico have the largest most thoroughly inventoried areas remaining, as well as under protection in terms of biological reserves or national parks.

2. Methods

2.1. Study areas

Our study is conducted at two levels: first, at the national level for Costa Rica (total area of 51 000 km^2) and second at a specific ecosystem level (tropical dry forest) for two study sites. Because of the central mountain range, the country encompasses numerous physiognomically different tropical forest types (Holdridge et al., 1971). However, to facilitate the interpretation of the results and to be consistent with forest types reliably detectable from imagery we refer to two distinct forest types: predominantly evergreen (e.g. tropical wet) and predominantly dry season deciduous (e.g. tropical dry) forest types. Subsequently, in a more detailed analysis, the two tropical dry forest areas we examine in greater detail are the Santa Rosa National Park in Costa Rica (10°48′53″N, 85°36′54″W) and the Chamela-Cuixmala Biosphere Reserve in Mexico (19°22′–19°39′N, 104°56′–105°10′W).

Santa Rosa National Park is composed of a mixture of secondary forest in various stages of regeneration (Arroyo-Mora et al., 2005a; Janzen, 2000; Kalácska et al., 2004). We refer to four stages of succession in Santa Rosa: pasture SR-P, early SR-E, intermediate SR-I and late SR-L (Table 1). The total study area for Santa Rosa is 500 km^2. The Chamela-Cuixmala Biosphere Reserve is comprised of approximately 126 km^2 of forest, the majority of which has been undisturbed for hundreds of years (Maass et al., 1982). We refer to four physiognomically different forest types in and around the biological station. Upper ridge-top and slope (CH-U) and Lower Riparian (CH-L) forest classes are mature undisturbed forests, intermediate (CH-I) is a secondary forest stage and early (CH-E) is the first stage of regeneration populated entirely by low *Acacia* sp. bushes (Table 1) (Kalácska et al., 2005). We included a buffer of 30 km around the station for the analyses, in order to include land cover types other than forest in the study area. The total study area (including the buffer zone) for Chamela-Cuixmala is 2465 km^2. In Santa Rosa there are 6 months of little to no precipitation (December–May) and a total yearly precipitation that is highly variable (915–2558 mm)(Janzen, 1993). In Chamela, the precipitation ranges from 374 to 897 mm with 80% falling between July and October (Maass et al., 1982). Drought deciduousness is the general leaf phenological response to the dry season (Gentry, 1995; Lobo et al., 2003). Gentry (1995) estimates that approximately 40–60% of the species in Santa Rosa are deciduous compared to over 80% in Chamela.

2.2. Total aboveground biomass

Of the many direct methods currently available to estimate total above ground biomass, allometric equations are probably one of the most broadly used (Brown, 2002). Site-specific regression equations have been developed to estimate plant biomass from values of diameter at breast height (DBH), height and wood specific gravity (Maass et al., 2002). In other studies, general regression equations have been developed for specific forest types from DBH. Carbon is assumed to be approximately 50% of the biomass (Brown, 2002). The most complete estimates of carbon include not only standing live vegetation but also dead wood, root biomass and soil carbon. From forest structure data we calculated total live aboveground biomass from Brown (1997) for stems above 2.5 cm DBH:

$$\mathrm{ATB} = \exp(-1.996 + 2.32(\ln D)), \tag{1}$$

where biomass is expressed in kilograms of dry mass and D is DBH in centimetres. Root biomass was estimated from

Table 1
Forest structure characteristics and allometric carbon (C) content from live total aboveground and root biomass for all stages of forest in Santa Rosa and Chamela

Forest stage[a]	Canopy height (m)	Basal area (m^2/ha)	Stem density (No./0.1 ha)	Species density (No./0.1 ha)	No. strata	C (tons/ha)
SR-P	2 ± 0.5	0 ± 0.0	0 ± 0	1 ± 0	0	—
SR-E	7.5 ± 2.2	11.7 ± 5.4	112 ± 64	15 ± 7	1	31.8 ± 1.8
SR-I	10.3 ± 3.4	21.4 ± 6.8	130 ± 35	29 ± 5	1	60.9 ± 2.5
SR-L	15.0 ± 2.2	30.1 ± 6.5	107 ± 42	29 ± 7	2	88.9 ± 2.0
CH-E	1.3 ± 0.4	0 ± 0.0	0 ± 0	1 ± 0	0	—
CH-I	11.0 ± 2.2	10.4 ± 4.7	146 ± 103	8 ± 6	1	22.4 ± 2.3
CH-U	8.9 ± 1.8	13.2 ± 2.5	181 ± 5	31 ± 12	1	29.5 ± 3.0
CH-L	21.1 ± 1.0	25.8 ± 2.6	124 ± 10	31 ± 4	2	72.6 ± 3.4

[a]Forest structure data for Santa Rosa are from Kalacska et al. (2004) and Arroyo-Mora et al. (2005) and from Kalacska et al. (2005.) for Chamela. Census includes only woody stems with a DBH $\geqslant$ 5 cm.Canopy height for SR-P reflects height of African grass *Hyparrhenia rufa* (Jaragua) during the wet season. For CH-E it reflects the height of the *Acacia* sp. bushes.

Cairns et al. (1997):

$$RB = \exp(-1.0587 + 0.8836(\ln ATB)), \quad (2)$$

where ATB is aboveground tree biomass (Mg/ha). In all our estimates we include only live ATB and root biomass because we do not have data regarding dead wood or soil carbon for the study sites. In total, 26 plots of 20×50 m were sampled in Santa Rosa where stems $\geqslant 5$ cm DBH were identified and measured (Arroyo-Mora et al., 2005a; Kalácska et al., 2004). Thirteen plots of the same dimensions were included in the census in Chamela-Cuixmala (Kalácska et al., 2005) (Table 1).

2.3. Total forest cover assessment

For Costa Rica a supervised classification map (CR2000) was produced by a combination of 14 Landsat Thematic Mapper 5 and 7 images acquired in 1997 and 2000 using NASA pathfinder methodologies with a minimum mapping unit of 0.03 km^2 (Sánchez-Azofeifa et al., 2001; Zhang et al., 2005). The overall accuracy of the CR2000 data set (forest/non-forest) was estimated to be 90–92%. For the accuracy assessment, a total of 800 control points for forest with a minimum area of 0.03 km^2 were chosen and assessed on the ground (Sánchez-Azofeifa et al., 2001). The evaluation of the accuracy of the forest cover data set with the extensive collection of ground control points provide greater confidence in this data set in comparison to continental or global algorithms. This data set, resampled to 1 km^2 resolution (in order to have the same pixel size as the land cover maps), was used as the control in all analyses for Costa Rica.

For the Costa Rican national level baseline and the dry forest study sites' baseline determination analysis we examine three published and readily available global land cover maps created from different sensors: the Global Land Cover, 2000 (GLC2000) generated by the Canada Centre for Remote Sensing using SPOTVEG imagery (22 classes) using regionally defined classifications (Latifovic and Olthof, 2004), IGBP from the International Geosphere Biosphere Program created with AHVRR imagery (17 classes) (Loveland et al., 2000) and MODIS Land Cover data produced by Boston University (17 classes)(Muchoney et al., 2000). In addition, we include a regional land cover map for Central America produced by the Center for International Earth Science Information Network (CIESIN) at Columbia University under the Proyecto Ambiental Regional de Centroamerica (PROARCA). This regional land cover map was created from AVHRR imagery (17 classes)(Central American Commission on Environment and Development, 1998). All the above data sets have a spatial resolution of 1 km^2. PROARCA data are not available for the dry forest site in Chamela-Cuixmala Mexico.

Comparison of all coarse resolution maps was performed using the Costa Rica 2000 (CR2000) database. Accuracy at the national level and the specific dry forest study sites was estimated from sets of randomly generated control points extracted from the CR2000 data set. As the control data for Chamela-Cuixmala Mexico we used a supervised classification map (MX2002) generated using ASTER imagery (resampled to 1 km^2 resolution) for the year 2002 (Sanchez-Azofeifa and Quesada, unpublished).

3. Results

3.1. Forest cover estimation at the national level for Costa Rica

The overall forest cover estimates for Costa Rica from the different land cover maps are shown in Fig. 1. In comparison to the CR2000 data, each land cover map underestimates the actual forest cover for the predominantly deciduous (dry) ecosystem (Fig. 1f). The area demarcated as "dry" in Fig. 1f is predominantly deciduous or contains trees that are deciduous during the dry season (periods of little to no rainfall). This area encompasses 14% of Costa Rica (7140 km^2). The forest extent of the evergreen vegetation (Fig. 1f) however, in general, compares well between the land cover maps and the CR2000 data set except for the overestimation of the forest in the northeastern and central sectors of the country most likely due to the inclusion of palm, coffee, pineapple, yucca and other plantations in the forest cover class.

The overall forest accuracies for GLC2000 (74%) and MODIS (76%) (Table 2) reflect the high prevalence and accuracy of "forest" for the majority of the evergreen forest areas in the country (Fig. 1) (e.g. 79%—GLC2000 and 88%—MODIS) (Table 2). However, the low non-forest accuracies (e.g. 32% and 16% for GLC2000 and MODIS, respectively) for all land cover maps indicate an overestimation of the forest in the evergreen forest areas (Table 2). The poor accuracies of all data sets for the predominantly deciduous forest areas (36% mean accuracy for all data sets) indicate a severe underestimation of the deciduous forest (Table 2).

Table 3 indicates the total forest area from the land cover maps in comparison to the CR2000 data set. The total forest area estimated for Costa Rica by Mayaux et al. (1998) is also included in Table 3. Every land cover map derived from coarse resolution sensors used in this study overestimates the extent of the forest for Costa Rica by as much as 16 000 km^2 with the exception of Mayaux et al. (1998) who underestimate the forest cover by 9777 km^2. PROARCA is the closest in the overall estimate for forest area for the country with 27 792 km^2 which is only 4565 km^2 more than the CR2000 data set. The MODIS land cover data set is the farthest from CR2000 at 39 409 km^2.

3.2. Tropical dry forest sites

The total forest area from the CR2000 data set for Santa Rosa is 276.6 km^2. For Chamela-Cuixmala the MX2002 data set reveals 1925.6 km^2 of forest as control. The range

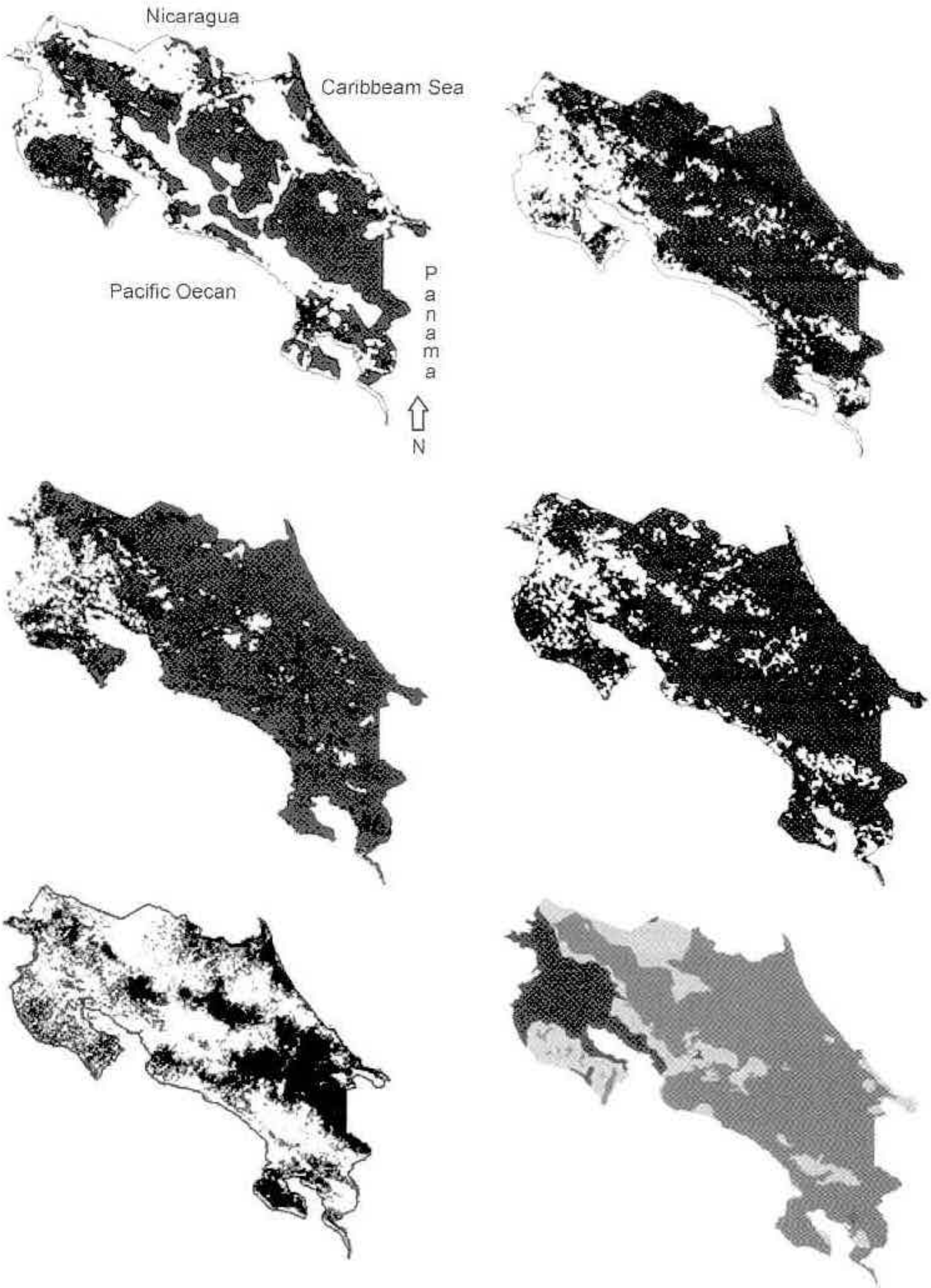


Fig. 1. Total forest cover extracted for Costa Rica from the various global/regional land cover maps. (a) PROARCA (b) IGBP (c) MODIS (d) GLC2000 (e) CR2000 (f) areas comprised of a predominantly dry season deciduous canopy (dark grey tone) and areas with a predominantly evergreen canopy (medium grey tone) in Costa Rica. Deciduous and evergreen areas were determined based on the Holdridge life zone database for Costa Rica.

of land cover classes found in the deciduous study areas based on the four global/regional land cover data sets is illustrated in Table 4. The consensus for the dominant class in Santa Rosa from the land cover maps is "cropland" or "agriculture", with the forest classes being minimal in comparison. However, from the CR2000 data for Santa Rosa, the dominant class is "forest" with an actual coverage of 55%. The actual forest cover for Santa Rosa has also been reported by an independent study from Arroyo-Mora et al. (2005b). Every land cover map derived from coarse resolution sensors underestimated the total forest in Santa Rosa by 130.6 km^2(MODIS) to 196.8 km^2 (PROARCA) in comparison to CR2000. The highest accuracy is from both the PROARCA and MODIS data sets for forest at 34% (non-forest accuracies of 92% and 48%, respectively, Table 2). The lowest accuracy is from GLC2000 at 16% for forest (78% non-forest).

Table 2
Forest and non-forest accuracies (%) from the national level analysis for Costa Rica and the two tropical dry forest sites (Santa Rosa and Chamela-Cuixmala)

Analysis level and thematic class	GLC2000	MODIS	IGBP	PROARCA	Mean
Costa Rica predominantly deciduous	56	35	18	36	36
Costa Rica predominantly evergreen	79	88	70	73	78
Costa Rica overall forest	74	76	59	64	68
Costa Rica overall non-forest	32	16	32	59	35
Santa Rosa forest	16	34	24	34	27
Santa Rosa non-forest	78	48	72	92	73
Chamela-Cuixmala Forest	41	66	61	—	56
Chamela-Cuixmala non-forest	42	37	38	—	39

Table 3
Total forest cover for Costa Rica for each data set

Land cover map	Total forest area (km^2)
MODIS	39 409
GLC2000	36 961
IGBP	32 648
PROARCA	27 792
CR2000	*23 227*
Mayaux et al. (1998)	13 450

CR2000 control data set is shown in italics.

For the Chamela-Cuixmala region the dominant class from the land cover maps is cropland (GLC2000), mixed forest (MODIS) or evergreen broadleaf forest (IGBP) (Table 4). From the MX2002 database, the dominant class is forest with an extent of 78% (1925.6 km^2). With all forest classes combined, for total extent, MODIS was very close at 79% (1950.3 km^2) followed by IGBP at 70% (1728 km^2) and GLC2000 at 46% (1135.6 km^2). The highest accuracy for forest cover was from MODIS at 66% (non-forest accuracy 37%) indicating that while the amount of forest is close to the MX2002 data set, the precision (actual location) of the forest is incorrect (Table 2). The lowest forest accuracy was from the GLC2002 data set for forest (41%) with a non-forest accuracy of 42%. For all the land cover maps forest classes were assigned based on descriptions as well as nomenclature. For example, the classes such as "woody savannah" because of the description were also included in the forest classes along with those that were labelled "forest". Improved land cover classifications for all forest types could be expected with standardization and broader descriptions of the land cover classes (Jung et al., 2006).

3.3. Environmental services payments: forecasting carbon sequestration

Arroyo-Mora et al. (2005b) found a rate of change of +4.91% per year in forest cover for the period of 1986–2000 in a larger dry forest area encompassing the Santa Rosa study area. Assuming the same constant rate of change in forest cover for the 2000–2010 period, Table 5 and Fig. 2 illustrate the forecasted total forest area for each land cover map taking the results from this study as the year 2000 baseline for each. The CR2000 data set shows a total increase in forest cover by 2010 of 170 km^2 followed most closely by MODIS at 89.8 km^2 and with the greatest difference, PROARCA at 49 km^2. With the exception of PROARCA, the land cover maps produced a relatively similar forecast in total forest area. With the assumption that the ratio of forest stage (i.e. early: 22%, intermediate: 47%, late: 31%) found in the area by Arroyo-Mora et al. (2005a) remains relatively constant over that time period, and the values of Mg C/ha/stage from Table 1 are used the carbon gains forecasted by each data set are shown in Table 5. The greatest carbon gain is from the CR2000 data set with a total of 1 074 691 Mg C followed by MODIS with a total of 567 107 Mg C and with the greatest difference, PROARCA at 310 207 Mg C. If successional stages are disregarded and an average value of 91.32 Mg C/ha is used (Kauffman et al., unpublished), 1 553 401 Mg C are the forecasted gain from CR2000 in comparison to 819 720 Mg C (MODIS) or 448 386 Mg C (PROARCA) (Table 5). As can be seen in Fig. 2, the rate of change of the forest cover is much lower for all land cover maps in comparison to CR2000 and each year the difference is compounded. An unprecedented rate of change would be needed by models incorporating any of the land cover maps to reach the same final forecasted value of total forest area as shown by CR2000. From the projected increase in forest cover (2000–2010), the estimated value of carbon sequestration from the CR2000 data set is $US 500 109 ($US 29.4 ha^{-1}) with a range of $US 248 354–746 762 ($US 14.6–43.9 ha^{-1}) (Fig. 3). The MODIS land cover map is the closest in its projection with a projected value of carbon sequestration in 2010 at $US 263 904 and a range of $US 131 055–394 061 (Fig. 3).

4. Discussion and conclusion

While this study is not meant to cast an overly negative impression of the value of coarse resolution remotely

Table 4
Forest and non-forest classes from the land cover maps for Santa Rosa and Chamela

Location	Land cover map	No. forest classes (extent)	Dominant class (extent)	2nd dominant class (extent)	Other classes
Santa Rosa	GLC2000	6 (26%)	Cropland (70.1%)	Tropical broadleaved evergreen forest—open canopy (19%)	Grassland, water
Santa Rosa	IGBP	5 (14.3%)	Cropland (51.2%)	Cropland—natural vegetation mosaic (14.7%)	Savannah, grassland
Santa Rosa	MODIS	5 (35%)	Cropland (64.8%)	Woody savannah (2.3%)	Shrubland, savannah, grassland, cropland/natural veg. Mosaic
Santa Rosa	PROARCA	7 (21.9%)	Agriculture (72%)	Tropical broadleaf deciduous woodland (7.2%)	Tropical perennial graminoid grassland, forest-woodland-agriculture complex, urban-veg. complex, agriculture and urban-industrial
Chamela	GLC2000	6 (46.4%)	Cropland (50.6%)	Tropical broadleaved evergreen forest—closed canopy (28.8%)	Grassland, water, consolidated rock with sparse vegetation
Chamela	IGBP	5 (70.1%)	Evergreen broadleaved forest (26.7%)	Cropland (25%)	Savannah, grassland
Chamela	MODIS	5 (79.1%)	Mixed forest (29%)	Evergreen broadleaf forest (28.3%)	Shrubland, savannah, grassland, cropland, cropland/natural veg.mosaic

PROARCA is only available for Santa Rosa.

Table 5
Initial and 10 yr forecasted total forest cover and carbon gain (assuming a +4.91% increase in forest cover per year) for the Santa Rosa dry forest study area (total area 500 km^2)

	CR2000	GLC2000	IGBP	MODIS	PROARCA
Initial forest cover (km^2)	276.6	128.3	137.6	146.0	79.8
10 yr forecasted forest cover (km^2)	446.7	207.2	222.2	235.7	128.9
Change in forest cover (km^2)	170.1	78.9	84.6	89.8	49.1
Total carbon gain Mg C (allometric with stages from Table 1)	1 074 691	498 492	534 626	567 107	310 207
Total carbon gain Mg C (average 91.32 Mg C/ha)	1 553 401	720 540	772 769	819 720	448 386

sensed data used to create forest cover maps as tools for the estimation of payments of environmental services and carbon sequestration baselines, it does intend to illustrate that caution must be used when selecting the appropriate derived products for regional planning exercises and environmental monitoring assessments. The intended use varies among data sets and is generally based on the scale of the product and the amount of effort placed on in situ validation. The participation of local agencies and the consideration for region-specific complications is also imperative to the quality of the final products. Imagery, preferably medium resolution (i.e. 30 m), should be acquired for each study site/region projected for environmental services payments. The subsequent classification of the imagery to establish a baseline would then be specific to the area and a validation should also be conducted with the unique characteristics of the ecosystem in mind. Used with caution, medium and high resolution remotely sensed data are powerful tools for the study of land use–ecosystem interactions, for providing information to decision makers, and the regional application of integrated models.

Previous studies have examined the inconsistencies between forest cover data sets at national levels (e.g. Jung et al., 2006; Kerr et al., 2001; Kleinn et al., 2002; Mayaux et al., 1998). Inconsistencies are important to recognize because numerous problems with the accuracy estimations are due to under-estimation, over-estimation and general misclassification of forest and other land cover types (Jung et al., 2006). However, we believe there is a need to move beyond examining accuracies and to begin examining their implications that in many cases can be greater than differences in a simple measure of how accurate the forest extent is in comparison with other classification efforts. For example, the risk of using unreliable forest cover estimates for establishing baselines include forcing trading to increase global net emissions whereby any mitigation projects would be both misdirected and inefficient (Kerr et al., 2002).

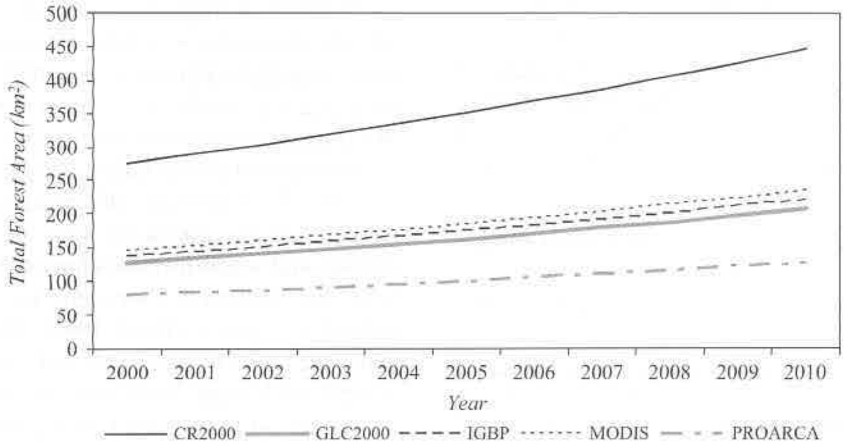


Fig. 2. Ten-year (2000–2010) forecasted total forest area (km^2) assuming a constant $+4.91\,km^2/yr$ rate of change.

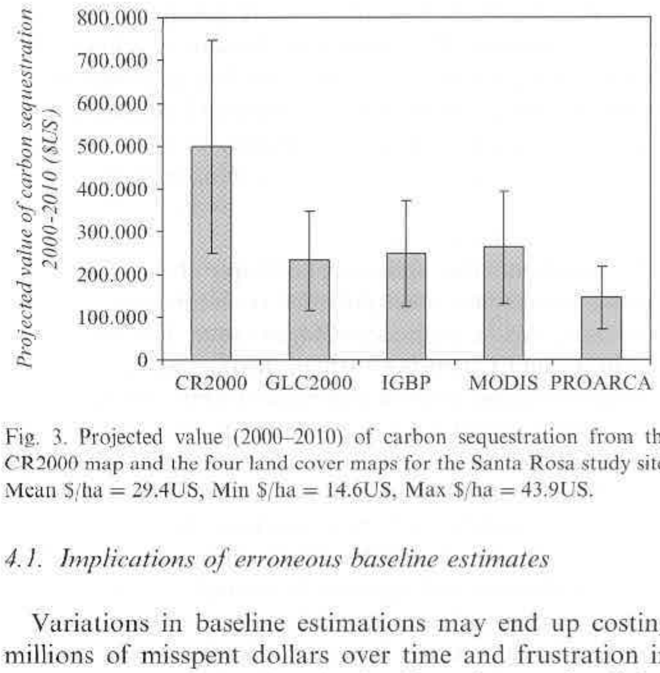


Fig. 3. Projected value (2000–2010) of carbon sequestration from the CR2000 map and the four land cover maps for the Santa Rosa study site. Mean $/ha = 29.4US, Min $/ha = 14.6US, Max $/ha = 43.9US.

4.1. Implications of erroneous baseline estimates

Variations in baseline estimations may end up costing millions of misspent dollars over time and frustration in both the governmental and scientific spheres of political action. As indicated in Table 5, the relatively large underestimation of the baseline by the land cover maps for the Costa Rican example leads to compounded errors in forecasted carbon sequestration and it is unrealistic and erroneous to assume that if the rate of change is correct that the initial baselines are not important. For a project to be eligible for Emission Reduction Units (ERUs) under the United Nation's Framework Convention on Climate Change (UNFCCC) in the LULUCF category, it has to show a successful accumulation of sequestered carbon(UNFCCC, 2003b). One ERU is equal to 1 metric tonne of carbon dioxide equivalent (UNFCCC, 2003a). And the only way a project can claim ERUs or credits through the CDM, is if it can show sequestration of carbon above the baseline scenario. Any erroneous estimates of either the initial forest cover or change (i.e. deforestation rate) would lead to diverse and unrealistic values for the carbon stocks.

4.2. Consideration for the tropical dry forest

For the tropical dry forest specifically, the forest identified from wet season images is a mixture of both deciduous and semi-evergreen species as well as pasture lands with enough green herbaceous biomass to produce a spectral signature comparable to trees resulting in the non-forested areas to possibly be mistaken for forest (Fig. 4). Estimates of carbon would therefore be an overestimation of the baseline in these areas. In comparison, the forest area readily extracted from dry season images are most likely located in areas where the microclimate enables them to retain their foliage along with areas including species that practice inverse phenology. Kalácska et al. (2004, 2005) have shown that the areas in Santa Rosa that predominantly retain partial foliage in the dry season are found in the late successional stage. The spectral signature of dry woody matter does not resemble that of green vegetation (Asner, 1998; Kalacska et al., 2007) (Fig. 4) and therefore, these areas are likely to be missed by automated algorithms. In Chamela-Cuixmala, the Riparian forest (CH-L) which also retains its foliage in the dry season has a different species composition (i.e. more semi-evergreen species), different microclimate and different biomass (Quesada, unpublished observation) from other forest stages in the area. Yet, the other late/mature stage in Chamela-Cuixmala (CH-U) is almost entirely deciduous in the dry season reinforcing the complication that when the majority of the trees are without foliage, the spectral signature is comprised of a mixture from soil, leaf litter, rock, bark, etc., rather than predominantly green leaves (Asner, 1998) (Fig. 4).

In general, a higher accuracy for the non-forest classes compared to the forest classes can be seen for every classification and land cover map. For Santa Rosa the reason for the large discrepancy in accuracy for the two classes with the land cover maps is that the majority of the area is classified as non-forest. Therefore, the chance that a "forest" control point will fall into a pixel classified as

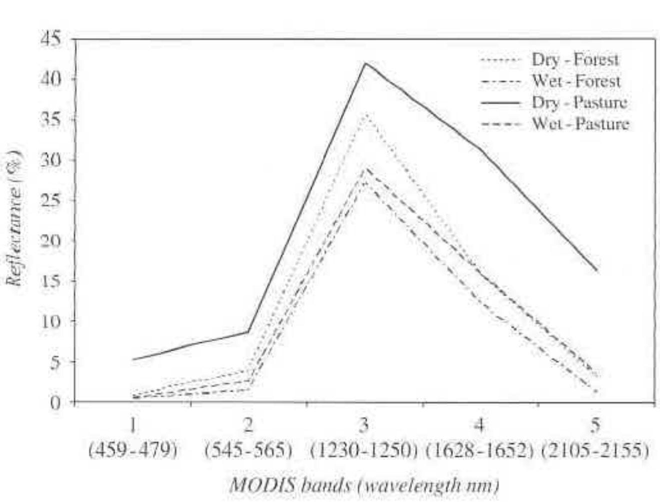


Fig. 4. Spectra of pasture and forest in the dry and wet seasons. Hyperion hyperspectral spectra were averaged to the spectral bands of the MODIS sensor.

forest is much less than the chance a "non-forest" control point will fall into a pixel classified as non-forest. For Chamela-Cuixmala, the location of the forest is incorrect (i.e. low accuracy) while the extent (i.e. total area) is generally close to what can be found on the ground. Thus, the application of inappropriate classification techniques will also result in large discrepancies. Methods must be flexible and may not all be used on an operational mapping project without extensive ground truth information. In addition, calibration or validation of large-scale maps without consideration of the ecological characteristics specific to each environment may contribute to the errors. These precautions must be taken in order to assure the most reliable baseline scenarios for both environmental services payments and carbon sequestration.

A common assumption is that at the spatial resolution of most global land cover maps ($1\,km^2$) the majority of the pixels is not homogenous in the land cover class they represent and therefore, under and/or over estimation of various classes is accepted. For the dry deciduous forest in Costa Rica Arroyo-Mora et al. (2005b) found that the mean patch size of forest was $1.07\,km^2$. In addition the forest patches are in general comparable in size or in some cases larger than the agriculture patches (i.e. dominant class in the land cover maps) (Table 4). These results indicate that although there may be many "mixed $1\,km^2$ pixels" there are still a sufficient number of "primarily forest pixels" in order for the deciduous forest to be present and included on land cover maps.

4.2.1. Deforestation pressure and implications for conservation

Based on examinations of socioeconomic uncertainty it has been stated that conservation research should focus in the wet forest life zones (Kerr et al., 2004). In Costa Rica with the collapse of the beef industry (thus changing the country's economy to cash crops such as pineapple and heart of palm that are grown in wet forest zones) and the short utility of tropical wet forest soils for agriculture, there are significant deforestation pressures in the wet forest life zones (Sánchez-Azofeifa et al., 2001; Subak, 2000). However, we argue that similar if not greater pressures also exist in the dry forest, compounded by the fact that they are practically non-existent in global land cover classifications and thus are not in the forefront of conservation policies(Sánchez-Azofeifa et al., 2005a, b). Due to low biotic and abiotic stresses and a comfortable climate, the dry forest has always been the preferred ecosystem for human settlement and animal husbandry (Ewel, 1999). The dry forest is globally extensive (42% of tropical forests are dry) but because of its appeal to human settlement it is also among the least protected (Murphy and Lugo, 1986). In Mesoamerica less than 1% has official conservation status, and only 2% is in patches large enough to attract the attention of conservation organizations (Janzen, 1988). Pfaff and Sánchez-Azofeifa (2004) illustrate large areas in the dry deciduous forest in Costa Rica where the pressure of deforestation is as high is in areas of wet forest. This along with the need to acknowledge their existence/location and the phenological complication accounting for the problems associated with estimating their true extent should make them a priority in global environmental services payments and carbon mitigation projects.

4.2.2. Cost of deforestation

In addition, the following example from Chamela-Cuixmala illustrates other potential problems that can arise from using various estimates of forest cover. The land tenure system around Chamela-Cuixmala favours subsistence and commercial crops, tourism and cattle grazing. Presently, the land is most valued for tourism rather than any other land use, including forest (Maass et al., 2005). However, the short-term return of such land uses does not make up for the long-term cost associated with these practices. For example, the clearing of the forests could result in either a scenario where certain pollinators will have to be brought to the area for various crops (which are of considerable value) or a scenario that could result in the loss of hundreds of thousands of dollars worth of crops (Maass et al., 2005). Both cases would result in very expensive endeavours compared to leaving the forest intact and utilizing the various services it could provide. If the maps being used by decision makers do not show the true extent of the forest, possible mechanisms for its protection cannot be considered.

However, in order to determine deforestation pressure (Pfaff and Sánchez-Azofeifa, 2004) or localize deforestation hot spots (Van Laake and Sánchez-Azofeifa, 2004), spatially reliable estimates of deforestation are needed: the basic requirement for which are accurate baseline forest cover maps from which to begin modelling.

4.3. Utility of remote sensing

A final broad question that must be considered concerns both the utility and facility of using remotely sensed data for payment of environmental services projects in general.

Rosenqvist et al. (2003) review the possible functions of remote sensing technology as part of decision support systems for the Kyoto Protocol. Two specific points from their review require special consideration for the use of remotely sensed data. First, the definition of "forest" from the Marrakesh Accords (UNFCCC, 2001a, 2003a) as referred to earlier and the subsequent standardization of ground control point verification (Subak, 2000). Second, based on the Bonn Agreements all forest and afforestation/reforestation/deforestation activities are defined based on land use rather than land cover (UNFCCC, 2001b). The implications of these are such that an area of cleared land that is expected to return to forest will still be counted as forest under the Kyoto protocol and will not count as deforestation (Rosenqvist et al., 2003). In addition, only direct-human-induced afforestation/reforestation/deforestation events will be considered (UNFCCC, 2003a, b). And, for reforestation specifically the land must have been cleared for a minimum of 10 years prior to human-induced reforestation (UNFCCC, 2003a). Thus, once a reliable land cover map of forest and non-forest areas is produced it must further be subject to additional in situ verification for land use classification.

4.4. Implications and questions to be addressed

We have shown that depending on which study is used, the estimations of environmental services payments, carbon content and the accuracy of the forest cover will vary accordingly. Until questions regarding nomenclature and types of forest classes are resolved, even for the simplest questions of "how much forest is there?" and "where is the forest?", discrepancies between various studies and problems with the estimations of payments of environmental services will persist. These discrepancies may end up costing hundreds of millions of dollars in erroneous payments and unsuccessful carbon mitigation projects as well as the irrevocable loss of biodiversity. In order to rectify the discrepancies, more rigorous methods including a greater emphasis on the collection of ground control data are required. In addition, a standardized description of the "forest" class which takes into account the heterogeneity and deciduousness of the dry forests as well as every class included in a land cover analysis would reduce the uncertainty associated with the current land cover classifications. Because, as shown, some large-scale global land cover maps are inherently unrealistic when examined closely at the ecosystem or country scales.

The implications of our study present a need for looking beyond simple accuracy assessments of these products and examining them in broader contexts such as environmental services payments. Fassnacht et al. (2006) identify the importance of understanding the limitations and caveats associated with using products created from remotely sensed data. The four key issues they identify are differences in direct and indirect models, the difference between class-based and continuous mapping models, scale and accuracy assessment. Similar issues illustrated in this study strengthen the need for a stronger awareness about how maps created from imagery should be used and the technical limitations associated with such data. Nevertheless, there is no doubt that remote sensing is a powerful tool for policy makers when used appropriately.

Acknowledgements

We would like to thank Dr. Joseph Tosi Jr. for providing life zone data for Costa Rica. In addition, we thank the generous financial support of the Tinker Foundation and the Inter-American Institute for Global Change Research. We would also like to acknowledge the Carbon Sequestration Working Group at the National Center for Ecological Analysis and Synthesis (NCEAS), University of Santa Barbara, California for discussions leading to the development of this study. We would also like to thank Dr. Frédéric Achard for comments and suggestions on an earlier version of the manuscript.

References

Arroyo-Mora, J.P., Sánchez-Azofeifa, G.A., Kalácska, M., Calvo-Alvarado, J., C.Janzen, D.H., 2005a. Secondary forest detection in a Neotropical dry forest using Landsat 7 ETM+ and IKONOS imagery. Biotropica 37, 497–507.

Arroyo-Mora, J.P., Sánchez-Azofeifa, G.A., Rivard, B., Calvo, J.C., Janzen, D.H., 2005b. Dynamics in landscape structure and composition for the Chorotega region, Costa Rica from 1960 to 2000. Agriculture, Ecosystems and Environment 106, 27–39.

Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sensing of Environment 64, 234–253.

Brown, S., 1997. Estimating biomass and biomass change of tropical forests: a primer. FAO Forestry Paper No. 143, Rome.

Brown, S., 2002. Measuring, monitoring and verification of carbon benefits for forest-based projects. Philosophical Transactions of the Royal Society of London A 360, 1669–1683.

Busch, C.B., Sathaye, J.A., Sánchez-Azofeifa, G.A., 2000. Estimating the greehouse gas benefits of forestry projects: a Costa Rican case study. Ernest Orlando Lawrence Berkely National Laboratory, Berkely, CA.

Cairns, M.A., Brown, S., Helmer, E.H., Baumgardner, G.A., 1997. Root biomass allocation in the world's upland forests. Oecologia 111, 1–11.

Central American Commission on Environment and Development, US Agency for International Development, International Resources Group, L., The Nature ConservancyWinrock International, 1998. Central American Vegetation/Land Cover Classification and Conservation Status (PROARCA). CIESIN, Palisades NY 〈http://sedac.ciesin.columbia.edu/conservation/proarca.html〉.

Choi, I., 2005. Global climate change and the use of economic approaches: The ideal design features of domestic greenhouse gas emissions trading with an analysis of the European Union's CO_2 Emissions Trading Directive and the Climate Stewardship Act. Natural Resources Journal 45, 865–952.

Ewel, J.J., 1999. Natural systems as models for the design of sustainable systems of land use. Agroforestry Systems 45, 1–21.

Fassnacht, K.S., Cohen, W.B., Spies, T.A., 2006. Key issues in making and using satellite-based maps in ecology: a primer. Forest Ecology and Management 222 (1–3), 167–181.

Foster, J. R., Kingdon, C. C., Townsend, P. A., 2002. Predicting tropical forest carbon from EO-1 hyperspectral imagery in Noel Kempff Mercado National Park, Bolivia. IEEE 3108–3110.

Gentry, A. H., 1995.Diversity of floristic composition of Neotropical dry forests. In: Bullock, S. H., Mooney, H. A., Medina, E.(Eds.), Seasonally Dry Tropical Forests. Cambridge University Press, Cambridge, UK.

Gerlagh, R., van der Zwann, B., 2006. Options and instruments for a deep cut in CO_2 emissions: carbon dioxide capture or renewables, taxes or subsidies? The Energy Journal 27, 25–48.

Holdridge, L.R., Grenke, W.C., Hatheway, W.H., Liang, T., Tosi Jr., J., 1971. Forest Environments in Tropical Life Zones: A Pilot Study. Pergamon Press, New York.

IPCCC, 2005. IPCCC special report on carbon dioxide capture and storage. Intergovernmental Panel on Climate Change.

ITTO, 2002. ITTO Guidelines for the restoration, management and rehabilitation of degraded and secondary tropical forests. ITTO Policy Development Series, vol. 13. Yokohama, Japan.

Janzen, D.H., 1988. Guanacaste National Park: tropical ecological and biocultural restoration. In: Cairns, J.J. (Ed.), Rehabilitating Damaged Ecosystems. CRC Press, Boca Raton, FL, pp. 143–192.

Janzen, D.H., 1993. Caterpillar seasonality in a Costa Rican dry forest. In: Stamp, N.E., Casey, T.M. (Eds.), Caterpillars: Ecological and Evolutionary Constraints on Foraging. Chapman Hill, New York, pp. 448–477.

Janzen, D.H., 2000. Costa Rica's Area de Conservacion Guanacaste: a long march to survival through non-damaging biodevelopment. Biodiversity 1, 7–20.

Jung, M., Henkel, K., Herold, M., Churkina, G., 2006. Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sensing of Environment 101, 534–553.

Kalácska, M., Sánchez-Azofeifa, G.A., Calvo-Alvarado, J.C., Rivard, B., Quesada, M., 2005. Effects of season and successional stage on leaf area index and spectral vegetation indices in three Mesoamerican tropical dry forests. Biotropica 37, 486–496.

Kalácska, M., Sánchez-Azofeifa, G.A., Calvo-Alvarado, J.C., Rivard, B., Quesada, M., Janzen, D.H., 2004. Speceis composition, similarity and diversity in three successional stages of tropical dry forest. Forest Ecology and Management 200, 227–247.

Kalacska, M., Sánchez-Azofeifa, G.A., Rivard, B., Caelli, T., White, H.P., Calvo-Alvarado, J.C., 2007. Ecological fingerprinting of ecosystem succession: estimating secondary tropical dry forest structure and diversity using imaging spectroscopy. Remote Sensing of Environment 108 (1), 82–96.

Kerr, S., Hendy, J., Liu, S.Pfaff, A. S. P., 2004. Tropical forest protection, uncertainty, and the environmental integrity of carbon mitigation policies. EconWPA. 〈http://ideas.repec.org/p/wpa/wuwpot/0411001.html〉.

Kerr, S., Pfaff, A. S. P.Sánchez-Azofeifa, G. A., 2002. The dynamics of deforestation: evidence from Costa Rica. MOTU Economic and Public Policy Research.

Kerr, S., Pfaff, A. S. P., Sánchez-Azofeifa, G. A.Boscolo, M., 2001. How can carbon sequestration in tropical forests be rewarded? Evidence from Costa Rica in Panayotou, T. (Ed.), Central America, Project, Environment: Conservation and Competitiveness. Harvard Institute for International Development.

Kleinn, C., Corrales, L., Morales, D., 2002. Forest area in Costa Rica: a comparative study of tropical forest cover estimates over time. Environmental Monitoring Assessment 73, 17–40.

Latifovic, R., Olthof, I., 2004. Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data. Remote Sensing of Environment 90, 153–165.

Lenton, T.M., 2006. Climate change to the end of the millennium. Climatic Change 76, 7–29.

Litynski, J.T., Klara, S.M., McIlvried, H.G., Srivastava, R.D., 2006. An overview of terrestrial sequestration of carbon dioxide: The United States Department of Energy's fossil energy R&D program. Climatic Change 74, 81–95.

Lobo, J.A., Quesada, M., Stoner, K.E., Fuchs, E.J., Herrerias-Diego, Y., Rojas-Sandoval, J., Saborio-Rodriguez, G., 2003. Factors affecting phenological patterns of Bombacaceous trees in seasonal forests in Costa Rica and Mexico. American Journal of Botany 90, 1054–1063.

Loveland, T.R., Reed, B.C., Brown, J.F., Olhen, D.O., Zhu, Z., Yang, L.W.M.J., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing 21, 1303–1330.

Maass, J. M., Balvanera, P., Castillo, A., Daily, G. C., Mooney, H. A., Ehrlich, P., Quesada, M., Miranda, A., Jaramillo, V. J., García-Oliva, F., Martinez-Yrizar, A., Cotler, H., López-Blanco, J., Pérez-Jiménez, A., Búrquez, A., Tinoco, C., Ceballos, G., Barrraza, L., Ayala, R.Sarukhán, J., 2005. Ecosystem services of tropical dry forests: insights from long-term ecological and social research on the Pacific Coast of Mexico. Ecology and Society. 〈http://www.ecologyandsociety.org/vol10/iss1/art17〉.

Maass, J. M., Martinez-Yrizar, A., 2001. NPP Tropical Forest: Chamela, Mexico, 1982–1995. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee.

Maass, J.M., Martínez-Yrizar, A., Patino, C., Sarukhán, J., 2002. Distribution and annual net accumulation of above-ground dead phytomass and its influence on throughfall quality in a Mexican tropical deciduous forest ecosystem. Journal of Tropical Ecology 18, 821–834.

Mayaux, P., Achard, P., Malingreau, J.-P., 1998. Global tropical forest area measurements derived from coarse resolution satellite imagery: a comparison with other approaches. Environmental Conservation 25, 37–52.

Muchoney, D., Borak, J., Chi, H., Friedly, M., Gopal, S., Hodges, J., Morrow, N., Strahler, A., 2000. Application of the MODIS global supervised classification model to vegetation and land cover mapping of Central America. International Journal of Remote Sensing 21, 1115–1138.

Murphy, P., G.Lugo, A.E., 1986. Structure and biomass of a subtropical dry forest in Puerto Rico. Biotropica 18, 89–96.

Olschewski, R., Benitez, P.C., 2005. Seconday forests as carbon sinks? The economic impact of accounting methods on reforestation projects in the tropics. Ecological Economics 55, 380–394.

Pfaff, A.S.P., Kerr, S., Hughes, R.F., Liu, S., Sánchez-Azofeifa, G.A., Schimel, D., Tosi Jr., J., Watson, V., 2000. The Kyoto Protocol and payments for tropical forests: An interdisciplinary method for estimating carbon-offset supply and increasing the feasibility of a carbon market under the CDM. Ecological Economics 35, 203–221.

Pfaff, A.S.P., Sánchez-Azofeifa, G.A., 2004. Deforestation pressure and biological reserve planning: a conceptual approach and an illustrative application for Costa Rica. Resource Energy Economics 26, 237–254.

Rosenqvist, A., Milne, A., Lucas, R., Imhoff, M., Dobson, C., 2003. A review of remote sensing technology in support of the Kyoto Protocal. Environmental Science Policy 6, 441–455.

Sánchez-Azofeifa, G.A., Harriss, R.C., Skole, D.L., 2001. Deforestation in Costa Rica: a quantitative analysis using remote sensing imagery. Biotropica 33, 378–384.

Sánchez-Azofeifa, G.A., Kalácska, M., Quesada, M., Calvo, J.C., Nassar, J., Rodriguez, J.P., 2005a. Need for integrated research for a sustainable future in tropical dry forests. Conservation Biology 19, 285–286.

Sánchez-Azofeifa, G.A., Quesada, M., Rodriguez, J.P., Nassar, J., Stoner, K.E., Castillo, A., Garvin, T., Zent, E.L., Calvo, J.C., Kalácska, M., Fajardo, L., Gamon, J., Cueves-Reyes, P., 2005b. Research priorities for Neotropical dry forests. Biotropica 37, 477–485.

Skole, D.L., Tucker, C., 1993. Tropical deforestation and habitat fragmentation in the Amazon—satellite data from 1978 to 1988. Science 260, 1909.

Stone, T.A., Lefebvre, P., 1998. Using multi-temporal satellite data to evaluate selective logging in Pará, Brazil. International Journal of Remote Sensing 19, 2516–2517.

Subak, S., 2000. Forest protection and reforestation in Costa Rica: evaluation of a Clean Development Mechanism prototype. Environmental Management 26, 283–297.

UNFCCC, 2001a. The Marrakesh accords and the Marrakesh declaration. In: Proceedings of the Seventh Conference of the Parties (COP 7). FCCC/CP/2001/13/Add.1 (Annex, Section A).

UNFCCC, 2001b. Report of the Conference of the Parties of the Second Part of its Sixth Session held at Bonn from 16 to 27 of July 2001 (COP 6, Part 2). FCCC/CP/2001/5 (Annex, Section VII).

UNFCCC, 2003a. Methodological issues. Land use, land-use change and forestry: definitions and modalities for including afforestation and reforestation activities under article 12 of the Kyoto Protocol. United Nations office in Geneva, Geneva.

UNFCCC, 2003b. Report of the Subsidiary Body for Scientific and Technological Advice on its eighteenth session, held at Bonn, from 4 to 13 June 2003. Addendum. Methodological issues. Land use, land-use change and forestry: definitions and modalities for including afforestation and reforestation activities under Article 12 of the Kyoto Protocol (Agenda item 4 (d)). Modalities and procedures for afforestation and reforestation project activities under the clean development mechanism (in the first commitment period). United Nations Office at Geneva, Geneva.

United Nations, 1997. Kyoto Protocol to the United Nations Framework Convention on Climate Change, Dec. 10, 1997 in Change, U. N. F. C. o. C., editor. U.N. Doc. FCC/CP/1997/L.7/Add.1.

Van Laake, P., Sánchez-Azofeifa, G.A., 2004. Focus on deforestation: zooming in on hot spots in highly fragmented ecosystems in Costa Rica. Agriculture, Ecosystems and Environment 102, 3–15.

Zhang, Q., Devers, D., Desch, A., Justice, C.O., Townshend, J., 2005. Mapping tropical deforestation in Central Africa. Environmental Monitoring Assessment 101, 69–83.

Journal of Environmental Management 88 (2008) 360–371

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Social Sciences and landscape analysis: Opportunities for the improvement of conservation policy design

Ismael Vaccaro[a], Karma Norman[b,*]

[a]Department of Anthropology and McGill School of Environment, McGill University, Room 717, Leacock Building, 855 Sherbrooke Street West, Montreal, Que., Canada H3A 2T7

[b]Socioeconomics Program, Northwest Fisheries Science Center, NMFS/NOAA, 2725 Montlake Blvd. East Seattle, WA 98112-2097, USA

Received 12 February 2006; received in revised form 1 February 2007; accepted 10 March 2007
Available online 6 June 2007

Abstract

This article develops a methodology aimed at generating a systematic social diagnosis of social and natural landscapes. The analytical process is divided into six easily replicable and causatively connected steps. The goal is two-fold: first, to present the inextricable connections between physical landscapes and the communities that occupy them. And second, to provide a fundamental tool to public policy designers that should simultaneously improve social acceptability of conservation policies and policy efficiency and effectiveness. Finally, this methodology is consciously heterogeneous from a theoretical perspective. This article puts together, in fruitful dialogue, contributions from varying places on the social theory spectrum: from political economy to poststructural theory.

Keywords: Landscape; Conservation; Policy; Social analysis

1. Introduction

During the last 150 years, the protection of nature has evolved from an elitist *pass tempo* to a widely shared social priority linked to local development. Although colonialism, expropriation and exclusion have been historical descriptors closely associated with the conservation effort, the rate of environmental degradation recorded during the 20th century leaves little room for doubt regarding the need to protect the global environment from the global economy (Saberwal and Rangajaran, 2003; Wilshusen et al., 2002).

The social sciences in general have a long tradition of public policy analysis. In the late 1980s, the application of an analysis rooted in political economy to ecological conflicts resulted in the emergence of political ecology (Blaikie, 1985; Bryant and Bailey, 1997). The goal of this multidisciplinary theoretical movement was to understand the sociocultural context of environmental conflicts, paying special attention to the political conflicts associated with natural resource management. Political ecology resists a reductionist analysis of environmental conflict focused only on proximate causes. Following the political economy tradition, political ecologists implemented multiscalar analyses in order to identify relevant but remote actors and causes (Paulson and Gezon, 2005). Cognitive and poststructural approaches that focused on the subjective accounts of environmental conflicts were integrated as part of political ecology's heterogeneous theoretical and methodological toolkit (Neumann, 2005; Robbins, 2004). It was only a matter of time before work in political ecology approached the design and consequences of conservation policies (Guha, 2000; Neumann, 1998).

Notwithstanding a few exceptions (Peña, 2003; Stocks, 2003), political ecology and the social sciences in general have tended to remain on the fence, offering an arsenal of analytical tools to critically analyze the effects of conservation policies, but not participating in their design and management. The critical work has proven useful to empower subaltern groups that had been abused by national or international interest groups (Brosius and Russell, 2003; Peluso, 1992). It has also played a major role in land restitution processes (Maluleke and Steenkamp,

*Corresponding author. Tel.: +1 206 302 2418; fax: +1 206 860 3475.
E-mail addresses: Ismael.vaccaro@mcgill.ca (I. Vaccaro), Karma.norman@noaa.gov (K. Norman).

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.03.007

1998; Russell, 2005). In general, anthropologists, geographers and other social scientists appear as potential mediators, interlocutors or translators between actors engaged in active conflicts (Minnegal, 2005).

The goal of this article, drawing on the multiple theoretical traditions of the social sciences, is to develop a systematic methodology that offers a social diagnosis of any given landscape. Trained as anthropologists, the authors of this article rely predominantly on knowledge and references produced from within their discipline. The thrust of the argument presented here, however, as with political ecology itself, draws on citations, case studies, knowledge and concepts from the political sciences, sociology, history, and geography, in addition to anthropology. We are proposing a six step analysis that should provide valuable baseline data fundamental to the design of ecologically and socially successful conservation policies. This methodology is consciously generalist and needs a context-specific adaptation, but it points out basic social variables that need to be considered when trying to understand (and protect) a landscape.

2. Landscape and uncertainty

Landscape is a fundamental concept within any sociocultural analysis that seeks to relate society and nature (Balee, 1998). A landscape is not simply a topographic configuration sustaining a particular combination of ecosystems. Nor should the analysis of a landscape consider the space as a mere abstraction dependent upon social processes. A landscape is the combination of biophysical reality and the human uses, reconstructions, and representations of it. From the ranks of historical ecology the landscape has been "defined as the material manifestation of the relation between humans and the environment" (Crumley, 1994, p. 6). Human agency, however, cannot be understood with analyses that solely scrutinize flows of matter and energy. The differential cultural perception of landscape and natural resources is fundamental to understand the aforementioned flows. "Landscape refers to the organization of land and nature at the scale of human experience" (Greenough and Tsing, 2003, p. 14). Decision making depends on cultural values. Or as Hirsch puts it "There is not one absolute landscape here, but a series of related, if contradictory, moments—perspectives—which cohere in what can be recognized as a singular form: landscape as a cultural process" (Hirsch and O'Hanlon, 1995, p. 23). Here, we will not suggest that the analysis of a singular landscape cannot be achieved due to the impossibility of accessing all its possible representations. On the contrary, we are in fact proposing a method that offers data capable of informing conservation policies while simultaneously sensing the multiple social perspectives of the ecosystem configuration that we might otherwise refer to as a "landscape."

In previous decades, the complexity and scarce predictability of social systems had separated 'the social' realm from the scientific world (Snow, 1959; Hollingshead, 1940). A methodological abyss separated the unpredictable social landscape from the tightly structured and assessable biological landscape. However, during the previous two decades of ecological thinking, concepts such as patch dynamics, disturbance, stochasticity, and non-linear response have overcome or replaced static and homeostatic models. This intellectual process has resulted in the introduction of unpredictability and uncertainty into the understanding of ecological systems (Moran, 1984; Scoones, 1999; Winterhalder, 1994). The acceptance of this perspective has necessary consequences for the design and management of conservation policies (Zimmerer, 2000).

Contemporary ecology has also developed a theoretical apparatus capable of inserting human beings into ecological equations. Presently, the expression 'anthropogenic landscape' refers not only to an extremely degraded or non-natural landscape such as an urban center or landfill, but also to species-rich, patchy environments with desired biodiversity values (Posey, 1984). Human beings have been accepted into the biological paradigm (Abel and Stepp, 2003; Botkin, 1990).

Within anthropology, for example, this sort of revolutionary merger occurred forty years ago when people like Steward (1955) and Rappaport (1968) attempted to integrate environment and society as an interactive whole. Their work, albeit limited to a materialist perspective, anticipated an integrated conceptualization of humans and nature. The subsequent development of ethnobiology at the heels of cognitive anthropology, and the adoption of poststructural and hermeneutic precepts included and emphasized the role of culture in the discussions of the interactions between nature and society, until then monopolized by more materialistic models.

3. Why conservation policies?

There are several elements that drive us to study conservation policies. Some of them are applied, and some are theoretical. As mentioned earlier, in societies worldwide there is a widespread sense of environmental crisis. In this Malthusian perception of crisis, the scarce resource is 'pristine nature' and considerable resources are devoted to its management. Most conservation policies, however, do not happen in a social void. Most pockets of 'pristine nature' have been historically inhabited and managed by all sorts of human communities. In the name of conservation, or on behalf of global or national ecological heritage, innumerable acts of injustice and forced expropriation have been perpetrated. On the other hand, many have pointed out the emancipatory potential for minorities that conservation policies present (Stevens, 1997; Wilshusen et al., 2002). Protected areas managed by or with local groups are tools of local development and of political recognition of otherwise blurred ownership rights. The involvement of social scientists, particularly those interested in microsocial analysis, in the design and management of protected areas may have positive outcomes at several levels. These objective facts play into two interesting fields of reflection.

First, we acknowledge at the outset, that satisfactory conservation goals can be achieved by developments other than the setting-aside of a "protected area." Indeed, it is with the consideration of social variables that policymakers often redesign or abandon protected areas as a means to conservation ends. In theory, the establishment of a protected area is a political event designed to impact biological variables. The consequences of the transformations fostered by such a process reverberate across the entirety of the area's social fabric. The creation of a protected area entails a redefinition of administrative and jurisdictional regimes extant in the area. These changes, traditionally promoted by external entities such as states, reorganize access to and control of the natural resources of the area: they are territorialization policies (Braun, 2002; Hannah, 2000; Peluso and Vandergeest, 2001). In general these jurisdictional reorganizations divert control and access to resources to external regulatory entities. This control affects not only ownership but also management. Uses and productive practices fall under strict control and often prohibition by the new public managers (Moore, 1998; Neumann, 1998).

Moreover, the rationale informing these new managerial regulations, the protection of landscapes or marine areas as national parks in particular, is often built upon urban values (Williams et al., 2002; Lee and Field, 2005). The guiding principle is not about production of resources for direct consumption by a local population. A protected area's management structure is organized around protection and promotion of natural values. Ecosystem integrity and biodiversity preservation have become goals in themselves with profound ideological and functional ramifications. The area is also protected for indirect consumption by visitors (Duncan and Duncan, 2003; Vaccaro and Beltran, 2007). The managers of conservation policies additionally have the responsibility to decide which kind of environment they want to protect. In other words, they are actually in the practice of environmental engineering, which may result in radical alterations of local ecological conditions. Notable examples include the reintroduction of charismatic predator species such as the wolf in Yellowstone, and the brown bear in the Pyrenees. These initiatives do not necessarily coincide with the goals and views of the rancher populations that surround or live inside these areas or the timber or mining corporations that would like to extract profit from them.

The entire process therefore is not merely a political takeover, but it is also one which transforms the economic life of the area as well as the collective identity of the affected communities (Sivaramakrishnan and Vaccaro, 2006). Conservation policies, until now, could not be typically characterized as locally created and managed. On the contrary, they are designed, funded, implemented and managed by institutions, public or private, sited and controlled from urban centers (Brosius and Russell, 2003).

From an applied perspective, microsocial analysis can contribute to solving some of the fundamental conflicts associated with the implementation of conservation policies. The fact that these policies are implemented over a social stage often results in conflicts between park managers and locals, or unsuccessful conservation due to a lack of managerial awareness of the local uses and ecological knowledge.

As was previously stated, this piece is a conscious effort to offer new tools for conservation policy design. Policy success, however, is a concept that needs to be qualified. Such success will be necessarily connected to biodiversity preservation and environmental justice associated with some form of local development. There is no single recipe to define success, but innumerable experiences have taught us that a biodiversity preservation which results in further degradation of the social circumstances for already marginal populations—politically, socially, and economically marginal peoples—will collapse or at minimum suffer extreme social pressures.

The next two sections are devoted to the development of a systematic approach explaining the potential of the social sciences as an intellectual means through which practitioners can better understand and inform conservation policy design.

4. Environment and use: anthropogenesis

In 2001, Ismael Vaccaro, one of the authors, resided for three weeks in the Aggtelekk National Park in northern Hungary. The park was created to protect a network of spectacular caves and an unusual combination of ecosystems covering the high hills that connect Hungary and Slovakia. One day, while walking across the forested areas with one of the rangers, the author and a park official intersected a tractor being driven by park staff to the upper meadows. When questioned about the purpose of this trip, the ranger explained that they were going to plow the upper meadows. The goal was to keep the forest at bay and foster the conditions for the survival of an endemic species of grass. Upon further questioning regarding the species they were planting, the ranger said that they were planting nothing in particular, but simply plowing.

The anecdote lends itself to a discussion of the interactions between environment and society through productive practices. The region, densely populated until recent times, contains heavily humanized ecosystems. This area, as with most of rural Europe, has suffered a process of severe depopulation. The removal of a "key species" like human beings has resulted in a "natural" alteration in the ecosystem. The park is attempting to conserve and protect an environment that, although perceived and explained as natural, is clearly anthropogenic. In order to do so, the park has to mimic the traditional uses that were once implemented by its now vanished local population.

The disappearance of humans from many protected environments globally, due to exclusion or unforced migration patterns, has triggered ecological processes with myriad consequences. In the Spanish Pyrenees the satisfaction provided by recovering forests has been tainted by a preoccupation with the state of these forests. Without rural populations to work and manage them, the likelihood of devastating wildfires has increased exponentially (Vaccaro, 2005).

Anthropogenic landscapes contain biophysical features manipulated or generated through human agency. In many cases, places deemed native, wild, undisturbed, and in need of preservation, are actually the products of long and deep histories of human impact and manipulation. Here we are writing, of course, of the production of nature (Castree, 2000). Anthropogenesis is, thus, a key concept that allows us to bridge the divide between nature and culture.

Examples of ecosystems that depend on specific traditional managerial practices are too numerous to name here. Examples of suppressions of these practices by modern, scientifically oriented managers are equally numerous. Fire, as a suppressed managerial tool and as a biodiversity enhancer, has attracted much attention all over the world (Kull, 2004; Langston, 1995; Lewis, 1989; Mathews, 2003; Pyne, 1997). Another illustrative example is ranching, the use and management of the territory and the resources that it generates, its ecology, and the inhabitation patterns associated with it (Ensminger, 1992; McCabe, 2004; Peters, 1994; Sayre, 2002).

The effective protection of a number of ecosystems depends on the effective assessment of the impacts and managerial characteristics of the traditional practices historically implemented on them. Microsocial analysis situates us in a privileged theoretical and methodological position to assess this interaction and to design policies capable of articulating social background and environmental values (Aswani and Hamilton, 2004; Sepez and Lazrus, 2005).

Anthropogenic landscapes and features may be associated with systems focused on long-term localized use as well as sudden extractive practices. The former, often related to subsistence practices, have low impact indexes that often translate into significant levels of sustainability. In general, these practices channel productivity to local households and maintain local welfare. The latter, associated with the delocalized current global economic system and with mass production directed towards a market economy, is, more often than not, controlled by outsiders not necessarily concerned with the long-term viability of the system as a whole. In other words, anthropogenesis can be related to all sorts of productive systems and it may have very different consequences in terms of the ecology and social fabric of an area.

5. Six steps towards landscape analysis: methodology

In this paper we want to formalize a simple six step method that should allow for a basic understanding of the social fabric of most landscapes, as well as the social background's connections with its concomitant ecosystem. This systematization of elements historically well known to anthropology and other disciplines is an attempt to convince social scientists and conservation professionals alike of the policy design potential harbored within social analysis.

As a heuristic device we will use a multilayered approach that mimics the explanatory structure of geographic information systems (GIS). In other words, each of these variables can be seen as a layer of social meaning attributed to the territory. The connection of all elements tells us about the fundamental social elements relevant to the management of natural resources in any given landscape. Several of these variables are also easy to map, thereby showing their distribution over the territory in question. The correlation of this social diagnosis with the structure of the area's ecosystem should shed light on the potential causality behind this structure: on the potential anthropogenic qualities of the examined landscape. The understanding of the relationships between local practices and ecosystem structure is fundamental for the conservation of the latter. The acknowledgement of the rightful importance of local managerial practices should also improve local acceptance of the policy and have a positive impact on local development.

5.1. Demographic patterns

In order to understand the social characteristics of any given territory we must both assess the population in quantitative and qualitative terms, and examine the local demographic context of the inhabited area. Use of—and pressures on—the natural resources of an area is necessarily linked to the density and distribution of the human populations to which the resources are connected.

Settlement patterns, and the demographics of the settled, are fundamental variables that will allow for a basic understanding of a given territory. Local inhabitants, for example, may be concentrated in villages, or dispersed in homesteads, and these differing densities represent very different social realities. Additionally, a population may exhibit seasonal spatial behaviors (nomadic, transhumant, or sedentary) and researchers would need to thoroughly examine these temporal patterns.

A historical approach should also unveil patterns associated with migratory movements: depopulation, repopulation, transfers from mountains to valleys, inland areas to coasts, or rural to urban areas. There are strong connections between population distribution and the economic and political systems in place. Demographic fluctuations are not isolated data but reflections of larger processes at play.

Peters (1994) has described significant changes in demographic aggregations in Botswana, and succeeds in linking them to ecological, economic and political changes associated with the creation of water boreholes as a consequence of development projects. She points to increasing levels of concentration and sedentarization related to year round water availability. Thus, this example is one of many that demonstrates that natural resource sustainability and availability is affected by such population shifts and vice versa.

In the valley of Lillet, in the central Catalan Pyrenees, high levels of dispersed populations spread across the mountain ranges, visible in the 19th century, declined until almost disappearing while the valleys below were filled with factories and mines. The global oil crisis of the 1970s set the conditions that led to the dismantling of these factories and mines. Consequently, the relatively large towns of the valleys entered into a process of acute depopulation. The demographic records from the period show a migratory movement towards the cities of the lowlands. The landscape has been deeply affected by this succession of demographic changes. The agricultural landscape of the 19th century was replaced by empty slopes ravaged by the needs of the mines and the factories. During the last 30 years forests have been in a clear process of expansion due to the lack of human pressure. The understanding of the landscape thus, requires a deep understanding of the historical demographic patterns.

Fairhead and Leach (1996), in their classic piece on Guinea Conakry, demonstrated the connections between concentrated settlements and forest expansion. Over time, the physical overlap of villages and forest clusters was striking, and this relationship between demographic distributions and ecological attributes was at the base of their study.

Similarly, in coastal areas, population levels and ethnic distributions within communities are linked to the ways in which these communities will accept or integrate environmental policies affecting adjacent marine areas, including marine protected areas (Christie et al., 2003).

As with the mountainous areas of Spain, coastal areas on the United States Pacific coast have faced major demographic shifts characterized by depopulation. Initially, resident Native American communities, which were involved in particular management and extractive pursuits as well as ownership regimes, were nearly wiped out by introduced diseases and subsequent violent conflicts with Euro-American settlers. The resultant depopulation paved the way for a Euro-American repopulation based on mining and novel extractive uses that shifted seascapes and adjacent landscapes away from a salmon economy (Beckham, 1971). A more recent change reflects depopulation in some of the same areas due to the decline of natural resource industries, such as forestry and fishing. Rather than face rapid depopulation, some communities have worked to develop demographic shifts which reflect a community in transition from a natural resource-dependent entity to one dependent on a tourist or retiree economy. Such a community, visible in its demographic analysis, may be more likely than a fishing-dependent community to accept prohibitions on use and extraction from an adjacent marine or terrestrial protected area.

5.2. Property regimes

A second significant level of analysis is centered on property regimes. Understanding the ownership of land and natural resources is fundamental to understanding the uses of natural resources in the territory, and the actors involved in their utilization. These uses ultimately have significant consequences with respect to the general configuration of the landscape.

Ownership is a sanctioned, hence legitimized, social acknowledgement of the rights of an individual or group of individuals over something. It is a contract—a social relationship—between individuals (Hann, 2003). The rights to possess and dispose, to exclude others and manage resources, are not universal. Certainly different cultural contexts have offered varying means of defining and enacting possession.

An accurate social analysis of the landscape will necessarily include an understanding of the property regimes in place. The establishment of a conservation policy should account for and articulate with previous structures of ownership. Even the conservation effort itself requires a deep understanding of property regimes. If an area contains an especially significant habitat, it is of fundamental importance to know who has owned this area and how it has been managed in the past. Obviously, extant management may have been at the root of this ecologically valuable habitat. Simple expropriation and total exclusion may very well result in the destruction of valuable habitat. So, protection may easily require a symbiosis between modern means of protection and traditional and prior uses (Johnson and Nelson, 2004).

The understanding of property regimes, however, is not limited to identifying the absolute owner of a given piece of land. Ownership is, at its essence, a bundle of rights (Ostrom et al., 2002). Although in modern terms we tend to focus on absolute ownership, this upper level of possession can, and usually is, divided into several tenancy levels. In other words, absolute ownership does not always coincide with managerial jurisdiction, or diverse levels of usufruct. An owner can delegate, cede, or lease, the right to manage, and the right to use a particular resource. These different levels may be exercised by a single individual or institutions, or they may be shared by different actors (Hann, 2003; Verdery and Humphrey, 2004). The understanding and management of a social and ecological landscape requires an assessment of ownership. Do the people who work and live in the land own it? Or are they landless peasants paying rent to powerful absentee owners? Were the managerial decisions that contributed to the formation of the current landscape taken locally or by external powers?

Legal ownership however, does not exclude *de facto* use of territory and resources. Illegal uses or occupations are common and are often associated with conflicts between systems of legitimacy. 'Poaching' and 'trespassing' are concepts associated with the aforementioned distinction that is so relevant to conservation policies (Gibson, 1999).

Although each specific case will require a context specific analysis, ownership regimes can be divided into five general types: (1) territory and resources can be owned by the state, (2) territory and resources may exist as common property, (3) as private or (4) corporate property, or they may exist as (5) 'open access' (Ostrom et al., 2002). Each of these ownership types implies different actors with different managerial approaches, as well as different levels of legitimacy, and potential levels of excludability.

When pursuing the analysis of property regimes it is important to understand how categories of ownership are locally defined. For instance, in many modern Western societies, seascapes and maritime resources have usually been considered and treated as open access. Although the temptation to universalize this open access character of coastal waters is strong, several researches have shown that in non-Western geographical and cultural locales more strict ownership regimes have traditionally been applied to the sea (Aswani, 1999; Johannes, 1991; Norman, 2007). The same caveat applies to other areas, including forested areas and graze lands that have otherwise been understood as common property (McKean, 1982; Netting, 1981).

The analysis of property regimes requires a dual methodological approach that combines archival research with ethnohistory. The work on property and cadastral registers should reveal the modern and legal, state-sanctioned, ownership structure of an area. This, however, is not sufficient in an examination of the complex structures that manage a given territory. The recollection and analysis of ethnohistorical accounts should unveil the local rules and perceptions of ownership which, as mentioned earlier, will not necessarily coincide with the contemporary legal registers.

5.3. *Managerial institutions*

Associated directly with property regimes are managerial institutions, the political structures designed to organize and regulate agency related to owned resources (Agrawal, 2002). A discussion on managerial institutions speaks to the widespread failure of many conservation areas to generate acceptance and retain legitimacy in the eyes of local populations.

Property theorists have pointed out that particular institutional attributes, the strength of monitoring structures for example, are key in making predictions about the capacity of an institution to purposefully manage resources, whether the desired end is conservation or equitable distribution or some other goal (Ostrom, 1990). Further, in the history of environmental management, management institutions imposed on a community from a distance are often met with resistance by local groups, and "top-down management" in general has been widely critiqued (Acheson, 2000; Ostrom, 1990; Sharp, 1998). In many cases, "top-down management" for conservation ends may be understood as another extension of state-building and state maintenance processes (Scott, 1998; Sivaramakrishnan, 1999; Vaccaro, 2005).

For this reason, efforts aimed at designing "community-based management" or "co-management", or other variants on this policy theme, have evolved in multiple natural resource management settings globally (Stevens, 1997; Usher, 1995). In these efforts, local institutions are enlisted in a larger effort to protect a resource or area.

Nevertheless, in some instances, local institutions are not culturally designed or equipped for novel conservation demands. In remote marine-dependent parts of Indonesia, the attempt to give ritual *sasi*—a temporary, chiefly ban on harvest from marine and reef areas—a broader policy meaning and enforcement capacity led to local social unrest and political disruptions not anticipated by external managers (Zerner, 1994). This dramatic failure, among others, suggests that rigorous institutional research, on the local and larger scales both, would need to precede any policy design.

In the analysis of managerial institutions, researchers first need to systematically inventory the variety of institutions present. Management institutions may be formal or informal, culturally traditional in nature, or highly bureaucratized. They may be initiated and maintained on multiple government levels, including local entities (e.g. tribal councils), regional bodies (e.g. multi-state management commissions), federal agencies (e.g. the U.S. Department of the Interior) and even multi-national institutions (as with trans-border protected areas including the Torres Strait Protected Zone or the continent of Antarctica). Resource management institutions may also be public or private in their structure and design. Microsocial analysis facilitates the identification of entities in each of these categories, allowing for the analysis and consideration of local institutions frequently overlooked in policy design.

After the institutions have been surveyed, research into both the institutional history and the political and cultural interplay between the institutions needs to be conducted. Institutional attributes can frequently be linked to managerial success, and researchers would need to characterize the institutions according to a tested list of variables, including the presence of conflict-resolution mechanisms, rule congruence and clarity of membership or coercive rights. Common property theory provides a rich literature in this regard (Ostrom, 1990).

5.4. *Productive practices*

Once we know about demographic distributions, property regimes and institutions with managerial jurisdiction over the area we seek to protect, we necessarily need to focus on productive practices. The three previous methodological steps provide, so to speak, a framework to understand the relationship between society and ecology. An analysis of demography, ownership regimes and institutions, identifies actors, their distribution across the landscape, and the jurisdictions and institutions with which they interact. The analysis of productive practices—the uses of territory and natural resources at its most basic—looks toward agency, and the physical interaction between actors and the environment (Steward, 1955; Butzer, 1982).

The possibilities for this economic interaction are multiple and as diverse as any possible combination of ecology and society. The physical production of goods could include hunting, gathering, fishing, a range of agricultural activities (from swidden practices to industrial agriculture), ranching, massive industries of extraction and transformation, and tourism and ecosystem services in general.

The traditional approach in conservation policies has been to exclude and prohibit productive practices inside the boundaries of a protected area (Wilshusen et al., 2002). This perspective ignores two important factors that may have a key impact on the success of a protected area. First, to singularly prohibit all productive practices in an area may suddenly suppress important or key income sources for the population living inside or around the protected area (Haenn, 2005). This development will likely result in resistance to and a boycott of the protective effort by the affected locals (Guha, 2000). Second, it is probable that important ecological features of the protected area may depend for their existence on specific local uses of resources (Lewis, 1989). The prohibition of these uses, then, will result in the loss of ecological richness. So, the effective management of a protected area requires an intimate knowledge of the internal dynamics of this landscape, including the interactions between nature and society through productive practices.

The assessment of the productive practices of a society within a territory results from the analysis of the economic life of its inhabitants. The connection of these practices with ownership regimes and local managerial institutions will open the door to a discussion of the political economy within an area. Classic ethnographic collection of data, including time allocation and income surveys, will allow for more ready identification and description of productive activities. The time allocated to each activity as well as the relative importance of the income it produces would tell researchers about the economic and social relevance of each activity.

The localization of these practices within a territory, taking advantage of the knowledge previously developed with respect to ownership, should allow for the identification of their ecological impacts: does a practice degrade the ecological integrity of the area? Is it a biodiversity-enhancing practice? Is the role of the activity in the local ecology indeterminate?

In any case, the point is that without a clear understanding of what has been or what is being done with the landscape, and without local political integration, managerial decisions lack a clear basis for support. Furthermore, traditional managerial practices are often informed by a deep historical knowledge of the local ecological conditions of the area. Conservation policies that ignore the local cultures of nature are wasting a precious resource that could prove essential to a better design of the policy itself (Aswani, 1999; Stevens, 1997).

A combination of anthropological and ecological research has demonstrated, for instance, that traditional horticulture, with its multilayered and diversified species composition, includes the fundamentals of sound agroforestry. Horticultural gardens, so often criticized because of their backwardness, foster biodiversity by multiplying the ecological niches available in a given area, and by the implementation of well-grounded fallow cycles (Altieri, 1995). Such analysis certainly does not apply to the monocrop plantations of contemporary industrial agriculture.

5.5. Cultures of nature

The previous four levels of analysis offer a nearly complete picture of the social characteristics of any landscape: demography, property regimes, managerial institutions and productive practices. However, a materialistic analysis of flows and structures must be complemented with a more hermeneutic approach to landscape management. This approach entails, first, an identification of the main actors living and involved with the management of a specific landscape. The second step implies an assessment of the specific perspectives on nature and resources which each of the involved actors holds. Although a given landscape is a unique and tangible biophysical reality, the way in which this tangible reality is perceived and culturally constructed is fundamental in understanding the way multiple actors interact with the environment. It is important to examine the cultures of nature present and active in the area (Lowe, 2004). Landscapes are far from homogeneous cultural constructions no matter where they exist. As Raffles writes:

> Locality is both embodied and narrated and is, as a consequence, often highly mobile: places travel with the people through whom they are constituted.... A locality emerges in complex ways through the multiple practices of numerous individuals in the midst of various situated projects. Its location is never secure and is always in need of reaffirmation and redefinition .... There is a clear tension here between fragmentation and proliferation of a unitary concept of locality, and the fact that locality is always ... shared (Raffles, 1999, pp. 324–328)

In the mountains of Spain, forest engineers see the forest as a productive unit to be managed in order to extract the maximum sustainable yield. Biologists look at it as a group of ecosystems that sustain valuable species in need of protection. Farmers, on the other hand, perceive the forest as the limit of their world. To them, the forest is a wall that separates them from the beasts. The landscape cannot be understood without the *agency* that each one of these groups exerts over it. And by agency we refer to the capacity of human individuals and societies to shape their worlds with intentionality. Agency, however, cannot be analyzed without understanding the cultural backgrounds, values, and goals associated with it (Knight, 2006; Theodossopoulus, 2003). Depending upon whether the goal is the production of short term benefits on a mass scale, self-sufficiency, or conservation, resources would be managed differently.

Furthermore, agency is informed by the specific knowledge that each group of actors has developed about it. It is of utmost importance to understand how knowledge is produced and transmitted. Successful conservation needs to incorporate information on how resources have been, and are, managed by all the social actors present in the area designated for protection (Hunn et al., 2003). In addition, knowledge is not an innocuous element. Its accepted validity and its legitimacy will have fundamental political consequences (Foucault, 1991). In other words, actors who succeed in imposing a perspective of nature will succeed in imposing a concomitant managerial perspective. Presently, the vast majority of the world is under some national jurisdiction, and the form of knowledge accepted and managed by state bureaucracies and technocracies is modern science.

Most current conservation policies are sustained by modern actors to which science is the main cognitive and managerial tool. In general, parks and reserves are public policies sustained by states and multilateral organizations, or private initiatives funded by non-governmental organizations. In both cases science plays a major role in their decision-making processes. Discourses about the need for explicit conservation were or are often originating in urban settings (Smith and Wishnie, 2000).

The analysis of human agency also has to include an analysis of the political power held by each of the actors involved. The creation of protected areas is a political process that requires a scrutiny of the values, goals and methods used by the social contenders (O'Neill, 1996). Discourse analysis is a primary tool to dissect the texts and statements produced by the social actors (Brosius, 1997; Nazarea, 1999).

Scientists do not share a single, unified point of view, independent of their disciplinary tradition, institutional context, and professional goals. Disciplinary traditions and professional goals do matter. The analysis of the cultures of nature interacting in a landscape will have to take into account the different perspectives of forestry, mining, road, and hydraulic engineers, as well as biologists, climatologists, geologists, geographers, and economists, among others. Moreover, an analysis will not be complete unless we associate each one of the actors with the institutions that pays his or her salary. Forest engineers working for the department of agriculture may generate completely different conclusions from forest engineers working for the department of the environment, even when both use precisely the same methodological tools (Guha, 2000; Moore, 1998; Sivaramakrishnan, 1999; Sodikoff, 2005). While the institutions in which scientists are situated may shape the science generated, scientists are nevertheless shaped to some degree by the environment as it is encountered and observed by them. Field scientists and technical managers operate on a basic environmental substrate regardless of their institutional pedigrees (Ingold, 2000; Mathews, 2003).

The analysis of the scientific knowledge generated about a landscape and its natural resources does not exhaust all the potential venues of reliable knowledge about a place. Through the years, ethnobiologists have been capable of identifying extremely detailed systematizations of knowledge about local environments developed by "traditional" communities that existed without access to modern science (Berlin et al., 1974; Conklin, 1975; Gonzalez, 2001).

The design and management of conservation policies has to comprehensively examine local cultures of nature. These cultures of nature are an important source of knowledge generated with years of careful observation and experience. And, equally important, consideration of local cultures of nature incorporates local perspectives into the design of these policies. Locals become shareholders in the policy with greater interest in the success of the policy itself.

The analysis of the cultures of nature present and active in the area cannot, however, be relegated to locally contextualized descriptions of nature. Such an analysis requires the inclusion of systems of values developed by each community and associated with the management and use of natural resources—a local moral economy if you will (Polanyi, 1944; Netting, 1981; Thompson, 1966). Managerial practices are not developed in a moral vacuum. They serve the socially accepted goals of the community that generated them. This fact confers legitimacy on them. Any attempt to regulate a landscape has to address, or understand, the moral economies articulating collective agency in the area. The imposition of a particular set of rules that disregards others is likely to be resisted (Guha, 2000; Neumann, 1998). The opponents of a conservation area more often than not are the inhabitants of the area affected and are quickly identified as poachers, trespassers, or bandits, and subsequently prosecuted (Gibson, 1999; Jacoby, 2001). Attention to moral economies will benefit management, particularly with regard to considerations of ancestral fishing, hunting, ranching and gathering rights, which are significant, for example, in many of the national parks of Alaska (Hunn et al., 2003). Enclosure or privatization of resources held and managed in common may result in the violation of historical collective rights and awaken organized resistance (Vaccaro, 2007).

As mentioned in the previous section, the triad of knowledge, values systems, and productive practices have considerable impacts on collective and individual senses of identity. Human beings become social individuals through a culturally contextualized process of socialization. Most conservation policies assume that the reduction of resources available to local communities as a result of the enclosure measures will be compensated for by the emergence of new sources of revenues associated with eco-tourism. The transition from agriculture, fishing, ranching, or timber harvesting activities to tour guides or hostel manager is not an easy one, from the perspective of infrastructure, but also from an intellectual, cultural and emotional perspective.

5.6. *Anthropogenic features of the landscape and ecosystem correlation*

Conservation policy design, traditionally, has only accounted for a series of natural values that seem to be worth protecting. In general, there has been little or no consideration of the social backgrounds in which these natural values have thrived or survived. Thus far, we hope to have demonstrated the limitations of an approach that ignores variables that may have a fundamental effect on the values that we are trying to protect.

The combination of the previous five steps of landscape social analysis with the ecological map of the area should provide even more information. The comparison of species distributions and concentrations with property regimes, oral history, and uses of the territory could unveil fundamental causal relationships between the social and the biological elements, and elucidate the connections between human inhabitation and ecosystem configuration. The understanding of these connections is a *sine qua non* precondition for the design of a just and successful conservation policy.

During past decades the term "anthropogenesis" has been associated with features of the landscape that show high levels of ecological degradation brought on by human intervention. Paradigmatic examples could be open-air mining operations, landfills, dumpsites and so on. Historical ecology, however, asserts that the history of every landscape cannot be written without paying close attention to the impact of human habitation within the area (Crumley, 1994; Balee and Erikson, 2006). Roosevelt (1989) and Carneiro (1995), for instance, describe the evolution of Amazonia's ecological structure and processes, linking it unquestionably to its different historical phases of human habitation.

Globally omnipresent discussions about concepts such as how to define or deal with invasive, beneficial, endemic, or ornamental species, benefit from an accurate understanding of the study of economic, social and ecological fluxes (Crosby, 1986; Grove, 1995). The productive history of an area with its connections to timber extraction (Haenn, 2005), plantation economy (Stiffler and Moberg, 2003), fisheries extraction (Acheson, 1988), or endemic warfare (Wiessner, 1998) may hold the key to understanding the current ecological situation of its territory. The ecological distribution of the area and the distribution and predictability of its basic resources may, in return, offer also clues to understanding the patterns of human occupation and use (Dyson-Hudson and Smith, 1978).

In the eastern Pyrenees, an area traditionally devoted to timber extraction, the forests have been dominated by pines for centuries. In these areas, oak and beech are actually the naturally occurring species. The depopulation of the area over the last 30 years has resulted in a slow recovery of oak and beech. These forests, at least in their species composition, are clearly anthropogenic.

In the western Torres Strait between Australia and Papua New Guinea, small uninhabited islands are frequently home to multiple coconut trees. While coconut palms, as a species, are readily distributed by natural forces (e.g. wind and waves), an ethnohistorical analysis suggests that on some of the islands, coconut groves were the products of anthropogenic propagation by ancestors who maintained property rights at a great distance from home villages and encampments, including on uninhabited islands (Norman, 2007).

In the aforementioned West African example brought to the fore by Fairhead and Leach (1996), 100 years of colonization, followed by independence, combined with science and development, had produced an environmental discourse that considered the forest a remnant of history, whereas the savanna was a man-made assertion into a declining primeval forest. Their research, drawing on local knowledge and aerial photograph analysis, suggested exactly the opposite. The savanna was not the result of aggressive local managerial practices, and the forest was the result of the collective effort of each village. Villages had created forest cover as an anthropogenic barrier between the houses and the beasts, keeping away, as well, the wildfires from the human inhabitants of the savanna.

In sum, the potential anthropogenic origin of ecological features must not be ignored. The implementation of the proposed five steps of social analysis and correlation with traditional ecological mapping can become a powerful tool to understand complex landscape dynamics and help on its successful management.

6. Conclusions

Conservation policies are much more than ecological engineering projects, and their ultimate viability does not solely depend on thorough assessments of ecosystemic distributions. Conservation policies are social projects on several levels. First, the environmental structure of any given area is not the exclusive result of non-human biophysical interactions. Anthropogenesis—human agency over the landscape—is a major factor in the shaping of ecosystems. Second, conservation policies are, in their inception and in their consequences, political projects and social processes. They are political in their very inception because they are created and implemented by institutions regulating public issues associated with natural resources management. The legal process of conservation area declaration is in itself also political. They are political in their consequences because the management of natural resources includes decisions about access to said resources. Political and economic power affects—and is affected by—these managerial decisions. Conservation policies are exclusionary in nature because they set jurisdictional boundaries.

The goal of this article is twofold. First, the article seeks to link anthropology and related social sciences including geography, economics, political science and sociology, to ecology and conservation. Secondly, it introduces a systematic methodology that shows the potential contributions that social sciences can offer to landscape analysis and, hence, conservation policy design, analysis and management.

This method, as discussed above, includes the analysis of demographic patterns, property regimes, managerial institutions, productive practices, cultures of nature, and their correlations with anthropogenic features of the landscape. We do not claim that this is a comprehensive way of analyzing the social and ecological variables of a landscape, but it offers a first approach at a set of data relating ecology to society. It identifies a set of fundamental social variables that need to be taken into account by researchers and managers in order to fully understand a landscape in its social, but also in its ecological composition. Such data could prove to be the key to significant improvements in any given conservation policy. The generality of the variables discussed allows for, once locally contextualized of course, widespread applicability in most contexts around the world.

Accepting the significance of the social background of a landscape while designing a conservation policy will improve its internal coherence and functionality; will accentuate its perceived legitimacy; and as a consequence, promote local acceptance. The process of social data gathering and negotiation results in the *de facto* incorporation of local knowledge, local institutions and local individuals into the creation and management of the conservation policy. In many of the legal texts guiding the creation of protected areas and the use of natural resources, considerations of significant social elements have now been explicitly advocated. Local participation has become a key element that managerial institutions all over the world, with more or less success, are trying to explore and incorporate (Macinko and Bromley, 2002; Environment Australia, 2002; Russell and Harshbarger, 2003). Such incorporation translates into local participation and local empowerment, both of which are much needed in a political process that otherwise results in the frequent exclusion of locals from historically accessible resources.

References

Abel, T., Stepp, J.R., 2003. A new systems ecology for anthropology. Conservation Ecology 7 (3), 12.

Acheson, J., 1988. The Lobster Gangs of Maine. University Press of New England, Hanover, NH.

Acheson, J., 2000. Varieties of Institutional Failure. Presented at the Keynote Address at Constituting the Commons: Crafting Sustainable Commons in the New Millenium, the Eighth Conference of the International Association for the Study of Common Property, Bloomington, Indiana, USA, May 31–June 4.

Agrawal, A., 2002. Common Resources and Institutional Sustainability. In: Ostrom et al. (Eds.), The Drama of the Commons, 41-86 National Science Foundation, Washington, DC.

Altieri, M.A., 1995. Agroecology: The Science of Sustainable Agriculture. Westview Press, Boulder.

Aswani, S., 1999. Common property models of sea tenure: a case study from the Roviana and Vonavona Lagoons, New Georgia, Solomon Islands. Human Ecology 27 (3), 417–453.

Aswani, S., Hamilton, R., 2004. Integrating indigenous ecological knowledge and customary sea tenure with marine and social science for conservation of bumphead parrotfish (*Bolbometopon muricatum*) in the Roviana Lagoon, Solomon Islands. Environmental Conservation 31 (1), 69–83.

Balee, W. (Ed.), 1998. Advances in Historical Ecology. Columbia University Press, New York.

Balee, W., Erikson, C., 2006. Time and Complexity in Historical Ecology: Studies in the Neotropical Lowlands. Columbia University Press.

Beckham, S.D., 1971. Requiem for a People: The Rogue Indians and the Frontiersmen. University of Oklahoma Press, Norman.

Berlin, B., Breedlove, D.E., Raven, P.H., 1974. Principles of Tzeltal plant classification: an introduction to the botanical ethnography of a Mayan-speaking people of highland, Chiapas. Academic Press, New York.

Blaikie, P., 1985. The Political Economy of Soil Erosion in Developing Countries. Longman, London.

Botkin, D., 1990. Discordant Harmonies: A New Ecology for the Twenty-first Century. Oxford University Press, New York.

Braun, B., 2002. The Intemperate Forest: Nature, Culture and Power in Canada's West Coast. Minnesota University Press, Minneapolis.

Brosius, P., 1997. Endangered forest, endangered people: environmentalist representations of indigenous knowledge. Human Ecology 25 (1), 47–69.

Brosius, P., Russell, D., 2003. Conservation from above: an anthropological perspective on transboundary protected areas and ecoregional planning. Journal of Sustainable Forestry 17 (1/2), 39–65.

Bryant, R., Bailey, S. (Eds.), 1997. Third World Political Ecology. Routledge, London.

Butzer, K., 1982. Archaeology as Human Ecology: Method and Theory for a Contextual Approach. Cambridge University Press, Cambridge.

Carneiro, R.J., 1995. The history of ecological interpretations of Amazonia: does Roosevelt have it right? In: Sponsel, L. (Ed.), Indigenous Peoples and the Future of Amazonia: An Ecological Anthropology of an Endangered World. University of Arizona Press, Tucson, AZ.

Castree, N., 2000. The production of nature. In: Sheppard, E., Barnes, T.J. (Eds.), A Companion to Economic Geography. Blackwell, Oxford.

Christie, P., Pollnac, R.B., Pomeroy, R., 2003. The Search for Integrated Coastal Management Sustainability: Results from a Multidisciplinary Examination in the Philippines and Indonesia. Presentation for the American Fisheries Society Annual Meeting. Quebec City.

Conklin, H., 1975. Hanunoo Agriculture: A Report on an Integral System of Shifting Cultivation in the Philippines. Elliot's Books, Northford, Conn.

Crosby, A., 1986. Ecological Imperialism: The Biological Expansion of Europe, 900–1900. Cambridge University Press, Cambridge.

Crumley, C., 1994. Historical Ecology: Cultural Knowledge and Changing Landscapes. School of American Research Press, Santa Fe, NM.

Duncan, J., Duncan, N., 2003. Landscapes of Privilege: Aesthetics and Affluence in An American Suburb. Routledge, London.

Dyson-Hudson, R., Smith, E.A., 1978. Human Territoriality: An Ecological Assessment. American Anthropologist 80, 21–41.

Ensminger, J., 1992. Making a Market: The Institutional Transformation of An African Society. Cambridge University Press, Cambridge.

Environment Australia. 2002. Indigenous Land Management, The Land Needs Its People, Indigenous Protected Areas. Natural Heritage Trust, Department of Environment and Heritage, Canberra.

Fairhead, J., Leach, M., 1996. Misreading the African Landscape: Society and Ecology in a Forest-savanna Mosaic. Cambridge University Press, Cambridge.

Foucault, M., 1991. The Foucault Effect: Studies on Governmentality. University of Chicago Press, Chicago.

Gibson, C., 1999. Politicians and Poachers: The Political Economy of Wildlife Policy in Africa. Cambridge University Press, Cambridge.

Gonzalez, R.J., 2001. Zapotec Science: Farming and Food in the Northern Sierra of Oaxaca. University of Texas Press, Austin.

Greenough, P., Tsing, A. (Eds.), 2003. Nature in The Global South: Environmental Projects in South and Southeast Asia. Duke University Press, Durham.

Grove, R., 1995. Green Imperialism: Colonial Expansion, Tropical Island Edens and The Origins of Environmentalism 1600–1860. Cambridge University Press, Cambridge.

Guha, R., 2000. The Unquiet Woods: Ecological Change and Peasant Resistance in the Himalaya. University of California Press, Berkeley, CA.

Haenn, N., 2005. Fields of Power, Forest of Discontent: Culture, Conservation and the State in Mexico. The University of Arizona Press, Tucson.

Hann, C. (Ed.), 2003. The Postsocialist Agrarian Question: Property Relations and the Rural Condition. Transaction Publishers, Piscataway, NJ.

Hannah, M., 2000. Governmentality and the Mastery of Territory in Nineteenth Century America. Cambridge University Press, Cambridge.

Hirsch, E., O'Hanlon, M. (Eds.), 1995. The Anthropology of Landscape: Perspectives on Place and Space. Clarendon Press, Oxford.

Hollingshead, A.B., 1940. Human ecology and human society. Ecological Monographs 10, 354–366.

Hunn, E., Johnson, D., Russell, P., Thornton, T., 2003. Huna Tlingit Traditional Environmental Knowledge, Conservation, and the Management of a "Wilderness" Park. Current Anthropology 44 (supplement), s79–s103.

Ingold, T., 2000. The Perception of the Environment: Essays on Livelihood, Dwelling and Skill. Routledge, London.

Jacoby, K., 2001. Crimes Against Nature: Squatters, Poachers, Thieves and the Hidden History of American Conservation. University of California Press, Berkeley.

Johannes, R.E., 1991. Words of the Lagoon: Fishing and Marine Lore in the Palau District of Micronesia. University of California Press, Berkeley.

Johnson, K., Nelson, K., 2004. Common Property and Conservation: the potential for effective communal forest management within a National Park in Mexico. Human Ecology 32 (6), 703–733.

Knight, J., 2006. Waiting for the Wolves in Japan: An Study of People–wildlife Relations. University of Hawai'i Press, Honolulu.

Kull, C., 2004. Isle of Fire Chicago: The Political Ecology of Landscape Burning in Madagascar. University of Chicago Press, Chicago.

Langston, N., 1995. Forest Dreams, Forest Nightmares: the Paradox of Old Growth in the Inland West. University of Washington Press, Seattle.

Lee, R., Field, D., 2005. Communities and Forests: Where People Meet the Land. Oregon State University Press, Corvallis.

Lewis, H., 1989. Ecological and Technological Knowledge of Fire: aborigines versus park rangers in Northern Australia. American Anthropologist 91 (4), 940–961.

Lowe, C., 2004. Making the monkey: the rise, fall, and re-emergence of 'M. togeanus' in Indonesians' conservation biology. Cultural Anthropology 19 (4), 491–516.

Macinko, S., Bromley, D.W., 2002. Who Owns America's Fisheries? Island Press, Covelo, CA.

Maluleke, L., Steenkamp, C., 1998. The Makuleke Experience. In: Workshop on Land Claims on Conservation Land: 3–4 September 1998. IUCN South Africa Country Office, Pretoria.

Mathews, A., 2003. Suppressing fire and memory: environmental degradation and political restoration in the Sierra Juarez of Oaxaca, 1887–2001. Environmental History 8, 77–108.

McCabe, T., 2004. Cattle Brings us to Our Enemies: Turkana, Ecology, Politics, and Raiding in a Disequilibrium System. University of Michigan Press, Ann Harbor.

McKean, M.A., 1982. The Japanese Experience with Scarcity: Management of Traditional Common Lands. Environmental Review 6, 63–88.

Minnegal, M., (Ed.), 2005. Sustainable environments, sustainable communities: potential dialogues between anthropologists, scientists and managers. School of Anthropology, Geography & Environmental Studies. University of Melbourne, Melbourne.

Moore, D., 1998. Clear waters and muddied histories: environmental history and the politics of community in Zimbabwe's Eastern Highlands. Journal of Southern African Studies 24, 377–404.

Moran, E. (Ed.), 1984. The Ecosystem Concept in Anthropology. Westview Press, Boulder.

Nazarea, V. (Ed.), 1999. Ethnoecology: Situation Knowledge/Local Lives. University of Arizona Press, Tucson.

Netting, RM., 1981. Balancing on an Alp: Ecological Change and Continuity in a Swiss Mountain Community. Cambridge University Press, New York.

Neumann, R., 1998. Imposing Wilderness: Struggles Over Livelihood and Nature Preservation in Africa. University of California Press, Berkeley.

Neumann, R., 2005. Making Political Ecology. Hodder Arnold, London.

Norman, K., 2007. Grasping Adubad: Badulgal Knowledge, Management, Tenure and Harvest within the Marine Environment of the Torres Strait. Ph.D., dissertation. University of Washington, Seattle.

O'Neill, K., 1996. The International Politics of National Parks. Human Ecology 24 (4), 521–539.

Ostrom, E., 1990. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press, Cambridge.

Ostrom, E., Dietz, T., Dolsak, N., Stern, P.C., Stonich, S., Weber, E.U. (Eds.), 2002. The Drama of the Commons. National Science Foundation, Washington DC.

Paulson, S., Gezon, L., 2005. Political Ecology Across Spaces, Scales, and Social Groups. Rutgers University Press, New Brunswick.

Peluso, N., 1992. Rich Forests, Poor People: Resource Control and Resistance in Java. University of California Press, Berkeley.

Peluso, N., Vandergeest, P., 2001. Genealogies of the Political Forest and Customary Rights in Indonesia, Malaysia and Thailand. Journal of Asian Studies 61, 761–812.

Peña, D., 2003. Community Acequias in Colorado's Rio Culebra Watershed: A Customary Commons in the Domain of Prior Appropriation. University of Colorado Law Review 74, 101–195 (Co-authored with Greg Hicks, July).

Peters, P., 1994. Dividing de Commons: Politics, Policy and Culture in Botswana. University Press of Virginia, Charlottesville.

Polanyi, K., 1944. The Great Transformation. Farrar and Rinehart, Inc, New York.

Posey, D., 1984. A Preliminary Report on Diversified Management of Tropical Forest by the Kayapo Indians of the Brazilian Amazon. Advances in Economics Botany 1, 112–126.

Pyne, S., 1997. World Fire: The Culture of Fire on Earth. University of Washington Press, Seattle.

Raffles, H., 1999. Local theory: nature and the making of an Amazonian place. Current Anthropology 14, 323–360.

Rappaport, R., 1968. Pigs for the Ancestors: Ritual in the Ecology of a New Guinea People. Yale University Press, New Haven.

Robbins, P., 2004. Political Ecology. Blackwell, Oxford.

Roosevelt, A., 1989. Resource management in Amazonia before the conquest: beyond ethnographic projection. Advances in Economic Botany 7, 30–62.

Russell, D., Harshbarger, C., 2003. Groundwork for Community-Based Conservation: Strategies for Social Research. Altamira Press, Walnut Creek, CA.

Russell, P., 2005. Recognizing Aboriginal Title: The Mabo Case and Indigenous Resistance to English-Settler Colonialism. University of Toronto Press, Toronto.

Saberwal, V., Rangajaran, M., 2003. Battles over Nature: Science and The Politics of Conservation. Permanent Black, Delhi.

Sayre, N., 2002. Ranching, Endangered Species, and Urbanization in the Southwest. The University of Arizona Press, Tucson.

Scoones, I., 1999. New Ecology and the Social Sciences: What prospects for a fruitful engagement? Annual Review of Anthropology 28, 479–507.

Scott, J.C., 1998. Seeing Like A State: How Certain Schemes to Improve the Human Condition Have Failed. Yale University Press, New Haven.

Sepez, J., Lazrus, H. (Eds.), 2005. Traditional Ecological Knowledge in Federal natural resource management agencies. Special Issue. Practicing Anthropology 27 (1).

Sharp, N., 1998. Reimagining sea space: from Grotius to Mabo. In: Peterson, N., Rigsby, B. (Eds.), Customary Marine Tenure in Australia. Oceania Publications, University of Sydney, Sydney.

Sivaramakrishnan, K., 1999. Modern Forests: Statemaking and Environmental Change in Colonial Eastern India. Cambridge University Press, New Delhi.

Sivaramakrishnan, K., Vaccaro, I., 2006. Postindustrial natures: hypermobility and place attachments. Journal of Social Anthropology 14 (3), 301–317.

Smith, E.A., Wishnie, M., 2000. Conservation and Subsistence in Small-Scale Societies. Annual Review of Anthropology 29, 493–524.

Snow, C.P., 1959. The Two Cultures and the Scientific Revolution. Cambridge University Press, Cambridge.

Sodikoff, G., 2005. Forced and Forest Labor Regimes in Colonial Madagascar, 1926–1936. Ethnohistory 52 (2), 407–435.

Stevens, S. (Ed.), 1997. Conservation Through Cultural Survival: Indigenous Peoples and Protected Areas. Island Press, Washington, DC.

Steward, J., 1955. Theory of Culture Change: The Methodology of Multilineal Evolution. University of Illinois Press, Urbana.

Stiffler, S., Moberg, M. (Eds.), 2003. Banana Wars: Power, Production and History in the Americas. Duke University Press, Durham.

Stocks, A., 2003. Mapping Dreams in Nicaragua's Bosawas Biosphere Reserve. Human Organization 62 (4), 65–78.

Theodossopoulus, D., 2003. Troubles with Turtles: Cultural Understandings of the Environment on a Greek Island. Berghahn Books, New York.

Thompson, E.P., 1966. The Making of the English Working Class. Vintage Books, New York.

Usher, P.J., 1995. Co-management of Natural Resources: Some Aspects of the Canadian Experience. In: Peterson, D.L., Johnson, D.R. (Eds.), Human Ecology and Climate Change: People and Resources in the Far North. Taylor and Francis, Bristol, PA.

Vaccaro, I., 2005. Property Mosaic and State-making: Governmentality, Expropriation and Conservation in the Pyrenees. Journal of Ecological Anthropology 9, 4–19.

Vaccaro, I., 2007. Sovereignty, collective ingenuity and moral economies: the confluence of transnational trends, states and local strategies in the Pyrenees. Environment and History 13 (1), 25–46.

Vaccaro, I., Beltran, O., 2007. Consuming space, nature and culture: patrimonial discussions in the hyper-modern era. Journal of Tourism Geographies 9 (3).

Verdery, K., Humphrey, C. (Eds.), 2004. Property in Question: Value Transformation in the Global Economy. Berg, Oxford.

Wiessner, P., 1998. Historical Vines: Enga Networks of Exchange, Ritual and Warfare in Papua New Guinea. Smithsonian Institute Press, Washington, DC.

Williams, J., Walker, G., Hatton, T., 2002. Dryland Salinization: a challenge for water and land management in the Australian landscape. In: Haygarth, P.M., Davis, S.C. (Eds.), Agriculture, Hydrology and Water Quality. CAB International, London.

Wilshusen, P., Brechin, S.R., Fortwangler, C., West, P.C., 2002. Reinventing a Square Wheel: A Critique of a Resurgent "Protection Paradigm" in International Biodiversity Conservation. Society & Natural Resources: An International Journal 15, 17–40.

Winterhalder, B., 1994. Concepts in Historical Ecology: The View from Evolutionary Theory. In: Crumley, C. Historical Ecology, School of American Research Press, Santa Fe, NM, pp. 17–42.

Zerner, C., 1994. Tracking sasi: the transformation of a central Moluccan reef management institution in Indonesia. In: White, T., Hale, L.Z., Renard, Y., Cortesi, L. (Eds.), Collaborative and Community Based Management of Coral Reefs: Lessons from Experience. Kumarian Press, West Hartford, CT.

Zimmerer, K., 2000. The Reworking of Conservation Geographies: non-equilibrium landscapes and nature-society hybrids. Annals of the Association of American Geographers 90 (2), 356–369.

ELSEVIER

Journal of Environmental Management 88 (2008) 372–381

Journal of Environmental Management

www.elsevier.com/locate/jenvman

The economic dimensions of integrating flood management and agri-environment through washland creation: A case from Somerset, England

J. Morris[a,*], A.P. Bailey[b,1], C.S. Lawson[b], P.B. Leeds-Harrison[c], D. Alsop[d], R. Vivash[d]

[a]Department of Natural Resources, Building 42a, Cranfield University, Cranfield, Bedford MK43 0AL, UK
[b]Department of Agriculture, The University of Reading, Earley Gate, Reading RG6 6AR, UK
[c]National Soil Resources Institute, Cranfield University, Silsoe, Bedfordshire MK45 4DT, UK
[d]River Restoration Centre, Silsoe Campus, Silsoe, Bedfordshire MK45 4DT, UK

Received 9 August 2006; received in revised form 29 January 2007; accepted 10 March 2007
Available online 27 April 2007

Abstract

In many river floodplains in the UK, there has been a long history of flood defence, land reclamation and water regime management for farming. In recent years, however, changing European and national policies with respect to farming, environment and flood management are encouraging a re-appraisal of land use in rural areas. In particular, there is scope to develop, through the use of appropriate promotional mechanisms, washland areas, which will simultaneously accommodate winter inundation, support extensive farming methods, deliver environmental benefits, and do this in a way which can underpin the rural economy. This paper explores the likely economic impacts of the development of flood storage and washland creation. In doing so, consideration is given to feasibility of this type of development, the environmental implications for a variety of habitats and species, and the financial and institutional mechanisms required to achieve implementation.

Keywords: Flood defence; Land use; Washlands; Agri-environment

1. Introduction

Flood defence for farmland has for many years been an important element of Britain's production-oriented agricultural policy. Many floodplain areas benefited from publicly funded flood defence and land drainage schemes, which reduced crop damage and facilitated a change to more intensive farming systems. In recent years, however, the limits of these floodplains have been demonstrated by fluvial floods during the winter months (English Nature, 2001a). Further, the move towards decoupling under the MacSharry and Agenda 2000 reforms of the EU Common Agricultural Policy (CAP), and the more recent introduction of the single payment scheme (Defra, 2004a, b) and cross-compliance (Defra, 2004c), has meant that current policy emphasis is directed towards environmental enhancement and diversity of economic activity, with a diversion of funds away from support for farm outputs.

This has encouraged a re-appraisal of land use in rural areas encompassing farming, environment and flood management, including an examination of the scope for the positive creation of flood storage facilities (Morris et al., 2004a). These could provide relief to areas presently subject to unacceptable flooding, reduce the need for expensive flood defence measures elsewhere in the catchment, help the management of scarce freshwater resources, provide wildlife and amenity benefits and, through credits for flood storage and extensive farming methods, provide alternative sources of income to land managers.

Washlands, which are flood storage areas used during times of high flow to reduce flooding in other parts of the

*Corresponding author. Tel.: +44 1525 863309; fax: +44 1525 863344.
E-mail addresses: j.morris@cranfield.ac.uk (J. Morris), a.p.bailey@reading.ac.uk (A.P. Bailey).
[1]Tel.: +44 118 3786270; fax: +44 118 9352421.

0301-4797/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2007.03.023

catchment (English Nature, 2001a), are one mechanism for flood defence and management (Environment Agency, 2002). For the purpose here, a washland is defined as an area of the natural floodplain that is allowed to flood or is deliberately flooded by a river or stream for flood management purposes, simultaneously providing potential for a wetland habitat (Morris et al., 2002, 2004b). Wetlands, as defined by the Ramsar Convention (Article 1.1), encompass a wider variety of habitats. In line with the Ramsar definition, Barbier et al. (1997) define five broad wetland systems, of which riverine—land periodically inundated by river over-topping—is one and closest to the washland concept.

In addition, to their potential contribution to commitments in the Ramsar Convention, washlands may also have a role in the context of the Water Framework Directive (WFD) and 1994 UK Biodiversity Action Plan (BAP) targets. The WFD (Directive 2000/60/EC) has a requirement for an integrated approach to water management at the catchment level. This may present opportunities to improve the management of flood risk through washland creation. Washlands also provide real opportunities to enhance biodiversity and thus contribute to meeting UK BAP targets (Detr, 1995). Floodplains contain several important habitats, including grazing marshes, fens and reedbeds, and the landscape, wildlife and/or historic interest can be of national, if not international, importance (Joyce and Wade, 1998; Wilson et al., 2004).

The environmental importance of floodplains within the UK has long been recognised with the establishment of a number of floodplain environmentally sensitive areas (ESAs) under agri-environment policy (English Nature, 2000; Defra, 2002). In addition, the recent review of UK strategy for the management of flood risk, aptly entitled Making Space for Water (Defra, 2004d), identifies a clear role for land management in general and washlands, which integrate habitat and flood management in particular. Finally, the launch of environmental stewardship (Defra, 2005a, b), replacing existing schemes such as countryside stewardship (Defra, 2003a) and ESAs, includes objectives, primarily under higher level stewardship, for flood management, specifically, 'to provide additional flood water storage and flood defence through the restoration and recreation of wetland habitat for other objectives' (Defra, 2003b).

The remainder of this paper, using a study of the mid and lower Parrett catchment within the Somerset Levels and Moors ESA in south west England (Morris et al., 2002), explores how public funds might be used more effectively to improve flood risk management through the appropriate use of agricultural land in ways, which reduce the adverse effects of unwanted flooding and simultaneously exploit the beneficial opportunities that managed storage of flood waters would bring. The feasibility of washland creation and potential environmental benefits are considered, before focusing on the impacts of washland creation and the financial and institutional mechanisms required to achieve implementation. In this respect, the paper provides an example of the opportunities that exist for 'joining-up' policy regimes and funding mechanisms regarding farm income support, nature conservation, flood risk and water resources management, especially during this period of considerable policy reform.

2. Feasibility of washland creation

Catchments can be classified into a number of zones, which vary in terms of topography, hydraulic characteristics and potential contribution to flood storage management. Within this there may be a number of options including temporary storage and managed evacuation of water in the lower levels and holding back potential flood waters in the middle catchment.

The suitability of potential sites for washland creation depends on a large number of factors: technical, economic, environmental and social. Hydraulic potential should be the initial selection criterion, followed by other criteria which reflect opportunity for environmental enhancement and likely social and economic impacts.

In the Parrett catchment, criteria for screening site selection for storage were developed and applied (Morris et al., 2002, 2004b). These were hydraulic suitability (ease of filling, evacuation and containment) existing flooding regimes, opportunity for environmental enhancement, suitability of land use, and site constraints such as that imposed by settlements and infrastructure.

3. Opportunities for environmental enhancement in floodplain areas

Just as commercial agriculture requires suitable water regimes, so do environmental and ecological characteristics and processes. The water regime requirements of features of the natural environment can be defined in terms of inundation and groundwater levels and these very amongst species and habitats during the course of the year.

The main conservation objectives in the Somerset Levels and Moors concern wintering wildfowl, breeding waders, rare aquatic invertebrates and diverse aquatic plant communities, species-rich lowland wet grassland features, and the wider wetland. These objectives are pursued through the designation of Special Protected Area status, Ramsar and SSSI sites, and the Natural Area Biodiversity Action Plan (English Nature, 2001b). In that these objectives require management of water regimes, with respect to both flooding and groundwater levels, they can be met through judicious management of flood storage areas and washlands.

Fig. 1 illustrates the variation in water regime requirements, measured in terms of depth of the water table level from the surface, for selected environmental characteristics during the calendar year. The gap in the diagram which runs through the year shows the minimum and maximum

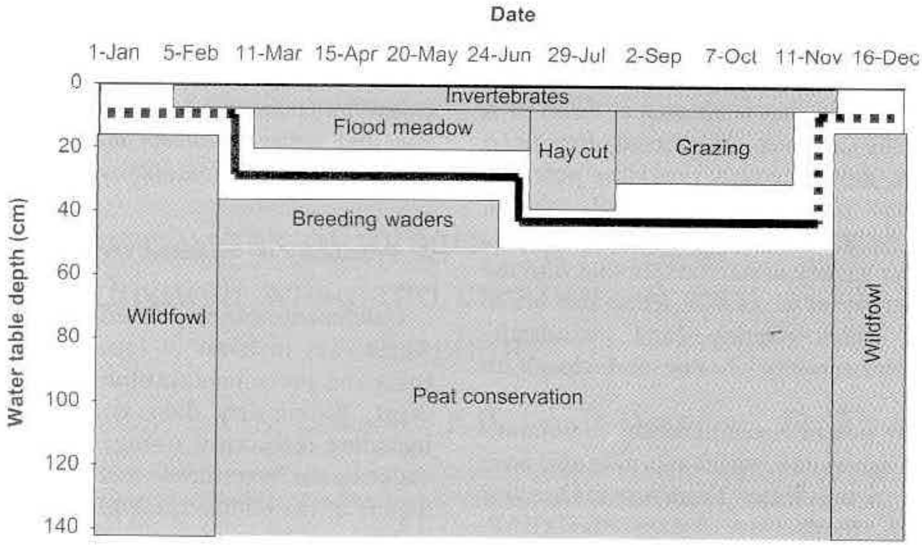


Fig. 1. Water regime requirements for environmental characteristics.

heights of water table levels which would satisfy the water regime needs of specific characteristics.

Wintering wildfowl require controlled water levels over critical minimum areas during the winter period (December–February inclusive), and a mix of splash (up to 10 cm deep), shallow (10–30 cm deep) and deep (30–75 cm) flooding. It is recommended that each of these areas should be at least 20 ha ideally in proximity of each other (English Nature, 2001b). If this was the case, a total area of 60 ha would be required. Breeding waders require water table levels, which are at least within 35 cm of the surface during the March to end June period inclusive. More generally, areas should be flood free after mid March, except for some surface pools and water tables at around 20 cm of the surface in early spring. It is recommended that managed units should be about 50 ha in size, with a mix of field water level conditions.

Fig. 1 also shows that there is general tolerance to flooding during the winter months. Indeed, wintering wildfowl are attracted by this condition. Of course, other fauna, such as small mammals, would need to be able to take refuge on higher ground during flood periods.

Further, there is a wide range of different plant communities found within the floodplain areas which vary greatly in their response to water regime. They can be divided into three broad groups according to their species characteristics: species-rich communities, washland communities (inundation grassland, swamps) and species-poor agricultural communities.[2]

Species-rich communities cannot tolerate inundation by surface water for more than a few days during the spring and summer (March–September), and they require water tables that are at least 20 cm below the surface during the mid-March to end-June period inclusive. Several of the plant communities have specific water regime requirements. For example, *Cynosurus cristatus–Caltha palustris* (MG8) flood pasture requires a constant water table throughout the summer, usually within the top 50 cm of the ground surface. *Alopecurus pratensis–Sanguisorba officinalis* (MG4) hay meadow can tolerate a deeper water table (over a metre in depth) in the summer. These plant communities are of high conservation value, which have developed under traditionally low-input agricultural management.

With respect to the two other broad groups, washland plant communities are tolerant of extended periods of flooding with regular inundation of surface water. They can provide a valuable grazing source for wildfowl and summer grazing for stock. The agricultural communities can be divided into two groups. Improved grasslands, which tend to be dominated by one or two grass species such as *Lolium perenne*. These grasslands are rarely flooded and support intensive grazing and silage. Wet meadows and rush pastures are considered unproductive agricultural land as they are found on permanently moist soil, they can provide valuable habitat for breeding waders, which require tussocky grasslands and rushes as well as soft soil for feeding.

Table 1 summarises the tolerance of the main types of grassland communities to flood regimes during winter and summer periods, and thereby the water management criteria that must be satisfied to secure these communities. The table also shows the equivalent values for dry matter (DM) and utilisable metabolisable energy (UME) production from the various grassland types. These values indicate the relative capacity of the grassland types to support livestock production, whether by forage conservation or grazing. These estimates are used in the economic

[2] This is based upon the British National Vegetation Classification. For further information refer to the *British Plant Communities* series edited by J.S. Rodwell (1991).

Table 1
Class of plant community, flood regime requirements and grassland productivity

Class[a] (and Tier reference)	Winter flooding regime October–February	Summer flooding regime March–September[c]	Relative dry matter yield	Relative utilisable metabolisable energy
Species-rich communities MG4/5/7C/8/M22/24 (Tier 3 target)	Periodic flooding 1–2-week duration	MG4 rarely inundated in summer	0.25–0.61	0.19–0.4
		MG5 rarely inundated throughout the year MG7C/8 constant water table between 20 and 50 cm depth M22/24 soils wet/moist throughout year		
Inundation grassland MG13/OV28 (Tier 4 washland)	Tolerant to extended periods of flooding	Will tolerate splash flooding in March and April	0.47	0.45
Swamps dominated by grasses/sedges S5/6/7/22 (Tier 3 tendency)[b] but limited information for these communities	Tolerant to extended periods of flooding	Tolerant to water table to the soil surface	0.95	0.3
Agriculturally improved MG6/7	Periodic flooding 1–2-week duration	Rarely flooded	1.0	1.0
Wet grassland MG9/10 (may previously been improved then abandoned) (as per Tier 2)	Periodic flooding 1–2-week duration	Occasionally flooded always moist soils, main species that develop are dominated by *Deschampsia cespitosa* (Tufted hair grass) and *Juncus effusus* (Soft rush) which are unpalatable	0.63	0.37

[a]Based upon the British National Vegetation Classification, see footnote [1].
[b]Dry matter yields are unknown for these communities. Some species such as *Glyceria maxima*, Reed swamp grass (which is unlikely to develop on the peat soils of the Somerset level) can provide very productive grazing. Relative dry matter yield 0.94–1.71. Based on figures from Westlake (1966).
[c]Difference in summer flooding period between flood scenarios and vegetation requirements. DM and UME data based on figures from Tallowin and Jefferson (1999).

analysis of grassland systems under different flood storage options.

Fig. 1 illustrates two main points. First, water regime requirements and tolerances vary between species and habitats through the year. For example, although flooding and water logging in winter suit visiting wildfowl, excessive flooding in spring is detrimental to breeding waders, invertebrates, small mammals and some plant species. Second, it is possible to manage water tables during the year (for example, along the ranges shown by the gap in the diagram) in order to deliver multiple environmental and farming objectives. This is the essence of water level management.

4. Impacts of washland creation on farming

Flood defence for agriculture, as for most land-engaging activities, refers to acceptable levels of flooding above and below the surface of the ground. Acceptable levels of flooding depend on the types of farming activities and practices. Generally, the more intensive is the system of production, the greater the need is for flood defence and land drainage, where the latter involves the evacuation of excess water and the control of field water levels to support commercial farming. Arable systems, involving cereal, protein, vegetable and root crops, are more sensitive to waterlogging and flooding than grassland systems, which can tolerate wetter conditions, especially in winter. In some floodplain and adjacent lowland areas, such as the Fens of East Anglia, flood defence and land drainage works carried out over many years, financed by a mixture of public and private funds, support some of the most intensive arable farming systems in the UK. In floodplain areas where standards of flood protection and land drainage are relatively low, most land is down to grassland, supporting an agricultural economy based on dairy, cattle and sheep production. The greater the incidence of flooding and waterlogging, the less intensive is the livestock system, with summer grazing of wetgrass being the most extensive form.

In the Somerset Levels and Moors, the floodplain areas correspond closely to the ESA. The voluntary scheme pays farmers annual amounts in return for the adoption of agreed prescribed practices, which are classified in 'tiers' of compliance (Table 2). Designated in 1987, it now covers over 29,000 ha, of which almost 18,000 (60% of the total eligible area) are subject to a total of over 1000 agreements.[3] Farmers receive annual payments of about

[3]From 2005 ESAs have been replaced by Environmental Stewardship (ES). Current agreement holders will continue to receive payments under

Table 2
ESA payments: Somerset levels and moors

Tiers and supplements	Annual payment rate (2001–2004) £/ha
Tier 1 permanent grassland	£125
Tier 1A extensive permanent grassland	£200
Tier 2 wet permanent grassland	£225
Tier 3 permanent grass raised water level areas	£430
Buffer strip supplement	£110/ha equivalent
All year penning supplement on peat soils	£18
Raised water level area supplement	£80
Public access tier	£170

Source: www.Defra.gov.uk.

£125/ha to retain permanent grassland, and between £200/ha and £430/ha to maintain wet grassland. The higher rate applies for permanently raised field water levels. These payments represent a mixture of compensation and incentive. At the time of the study, one area of concern was the potential for the design of a similar payment regime for washland creation/flood storage.[4]

Regional farm business data (University of Exeter, 1999, 2000, 2001) for the south west of England, within which the Somerset Levels and Moors ESA lies, suggests that farms within the floodplain and ESA areas tend to be smaller and less intensive, and therefore generally have lower average farm incomes than their regional average. Although very good agricultural performance is possible, many farms operate at lower levels of intensity than either their potential or the regional average. Most of the areas are used for grass conservation and/or stock grazing. The latter mainly involves dry milk cows, beef cattle and sheep. In the drier areas, over-wintering of sheep provides a useful income source.

The alleviation of flooding in those areas worst affected will reduce damage costs and increase output and profitability, other things being equal. Conversely, an increase in winter flooding in the newly created washlands would impose restrictions on farming, which, in the absence of incentive payments, would reduce the income and profitability for farmers. The extent to which this occurs depends on the degree of change in flooding and waterlogging, and the extent to which existing land use is sensitive to this. For example, increasing the extent of flooding, particularly in the spring, on agriculturally productive grassland would result in the development of flood tolerant vegetation such as rush-pasture, inundation grassland or swamps depending on the degree of flooding. Although these communities can provide summer grazing, changes in agricultural practices will be necessary as a result of the change in flood water distribution. Further, flooding depth and duration has obvious impacts for the timing of critical field operations and access. These apply to arable operations such as crop establishment and the grassland management operations such as forage harvesting and turnout dates of grazing animals. Disruption and delay will have an impact on revenues and costs.

The method needed to assess the financial and economic impacts of changes in drainage conditions and flood risk associated with the adoption of flood storage options is summarised in Fig. 2 (Morris and Hess, 1987; Dunderdale and Morris, 1997a, b; Penning-Rowsell et al., 2003). First, field drainage conditions should be identified as these determine the physical productivity of farming activity, whether for grassland (energy Mj/ha) or arable crops (t/ha). Energy from grass converts into a potential livestock carrying capacity (livestock units per ha) and, depending on the type and mix of livestock in the floodplain, into financial returns £/ha. Financial returns from arable production (£/ha) reflect the type and mix of crops in the flood plain. Second, therefore, estimates of the financial returns from each enterprise are needed (see University of Exeter, 2001). These are net of relevant production expenses (such as seeds, fertilisers, veterinary expenses and machinery operating costs). At the time, they also included receipts from government schemes such as the Integrated Administration and Control System (IACS) area payments and Beef and Sheep Premia (Defra, 2003c, 2004c)[5] and ESA membership. Third, flood damage costs (£/ha/year), which vary according to flood risk and land use, will need to be identified and deducted from financial returns to give an overall estimate of financial performance (£/ha/year) for each major land use type.

In order to assess the impact of adopting flood storage options, two broad scenarios were identified for the assessment of changes in flood regimes, which were distinguished in terms of their severity and impacts, as follows:

Damage and recovery scenario: relatively small changes in annual flood risk, which may include infrequent long duration events, but not to the degree that results in changes in agricultural land use. Examples include damage

(*footnote continued*)
this scheme until their agreement expires. They may then wish to transfer to ES. Payments rates under ES are similar to those found under ESAs. Similar payments to Tier 1 can be found under Entry Level Stewardship (ELS), Tier 2 and 3 payments under Higher Level Stewardship (HLS). Additional supplements also remain for raised water levels (HLS) and access (HLS), the latter at a reduced rate.

[4]Under the Higher Level of Environmental Stewardship (Defra, 2005b) a supplement (inundation grassland supplement) of £85/ha is now available in areas of river floodplain, currently protected by flood defence banks, which are made available for additional inundation by floodwater.

[5]IACS payments have now been replaced by the single farm payment and the single payment scheme, decoupled from levels of production. In England, the payment takes a hybrid approach involving an historic element (90% in 2005) and regional element (10% in 2005), the latter gradually replacing the former over a 7-year timescale. It has been suggested that farmers exclude these payments from their enterprise gross margin calculations and farm income, although in reality this may not occur in the immediate future.

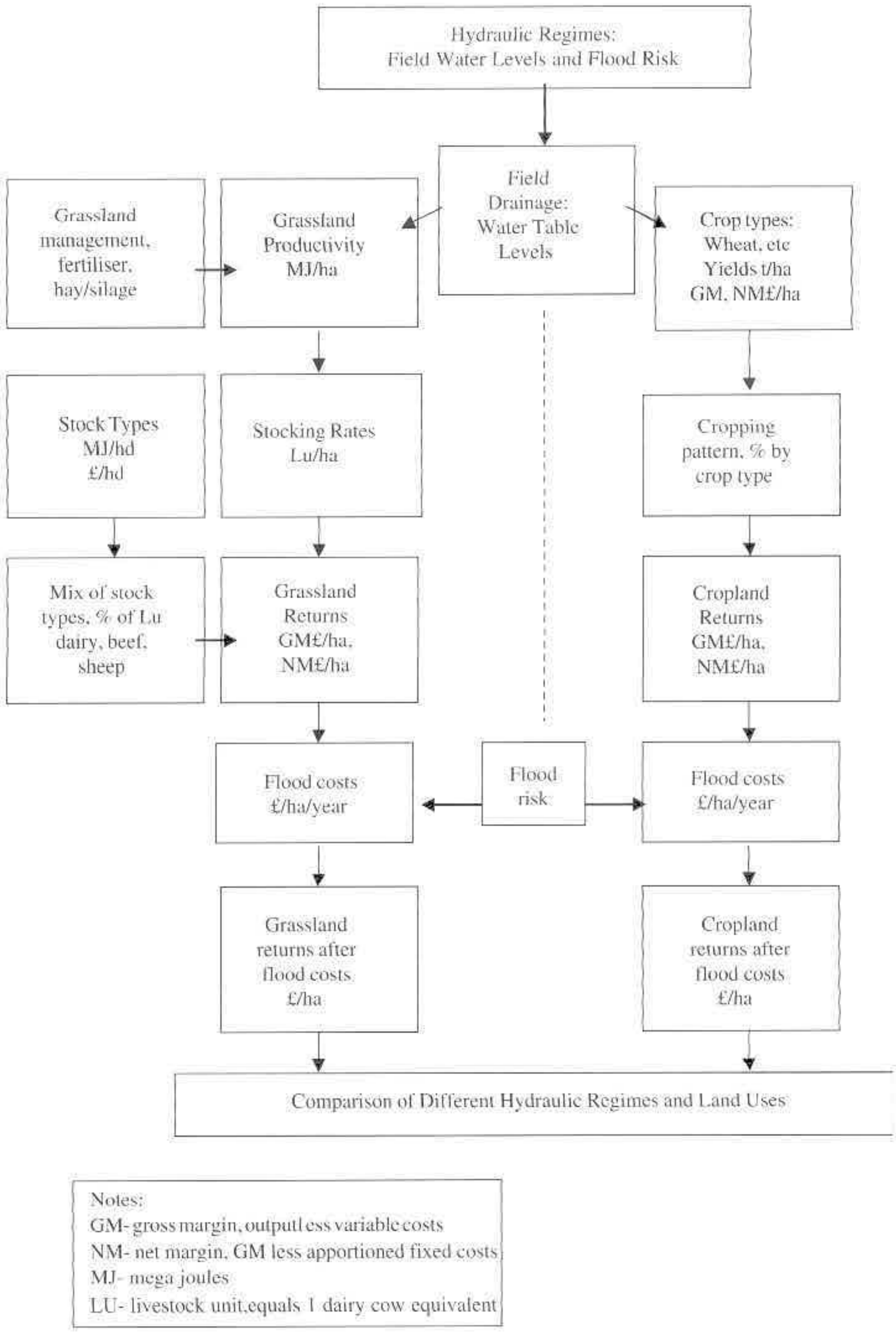


Fig. 2. Approach to financial and economic appraisal of washland options.

to the yield of grass or cereals, in some cases requiring reseeding of grass or winter cereals.

Land use change scenario: significant change in flood risk, which results in a shift in land use and farming practice, for example a shift from arable to grassland, or from intensive to extensive grassland.

With respect to damage and recovery, Fig. 3 contains estimated flood damage costs (2003/03 prices) on improved grassland according to the duration (in weeks) and depth (75 mm to over 750 mm) of flooding during the winter (October–March). Short duration flooding of about 1–2 weeks in mid winter has little impact, with costs of around £15/ha, including clean-up costs. Long duration floods of 2 months or so are likely to kill 'improved' ryegrass varieties and require reseeding at a cost of about £200/ha. Repeated relatively short duration flooding would have a similar impact. Persistent long duration flooding of more than 2 months would encourage a switch to a lower intensity land use with an estimated cost of about £170/ha to £220/ha.

Fig. 4 shows the impact of flooding on ESA Tier 1 type grassland, which is subject to limits on fertiliser use (max 75 kg N/ha) and hence livestock carrying capacity. Flood

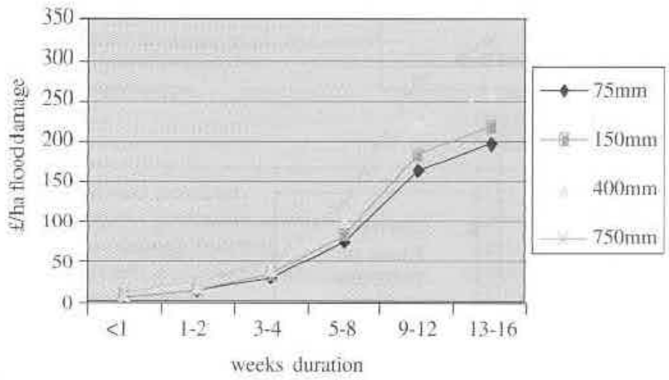


Fig. 3. Winter flood damage costs (£/ha) on improved grassland by duration and depth of flooding.

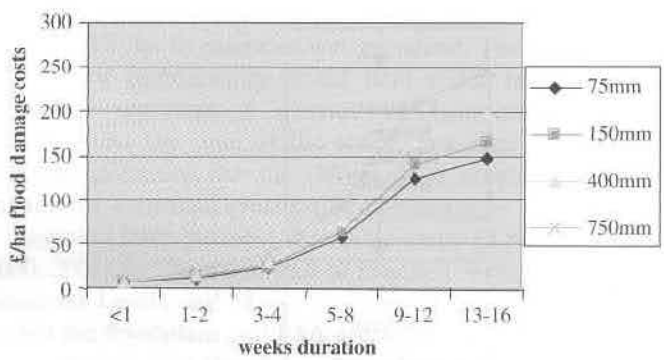


Fig. 4. Winter flood damage costs (£/ha) on Tier 1 grassland by duration and depth of flooding.

costs are proportionately lower than for improved grass, but the same principles concerning the duration of flooding apply.

For arable crops, it is assumed that winter flooding of more than a few days would destroy the crop and require reseeding with a lower yielding spring cereal, if feasible. Damage costs could be about £450 to £500/ha.

Figs. 3 and 4 also indicate the value of flood alleviation relief that might be obtained where there are reductions in flood risk due to controlled washland flooding elsewhere.

With respect to land use options, standards of flood defence have a major influence on agricultural practices, not only as a consequence of surface flooding but also, and often more critically, as a consequence of waterlogging of soils. In many respects, the land use in the Parrett catchment reflects these drainage circumstances, modified by ESA prescriptions where these apply. A change in flood risk associated with the adoption of flood storage options could involve a change in land use, for example from intensive to extensive grassland.

A 'washland package' was identified comprising flood tolerant grass species and related grassland management offering potential advantage over the current Tier 3 arrangements, which require permanent raised water levels. Financial indicators include gross margins (output less direct costs such as fertiliser and feed), and different definitions of net margins according to whether semi-fixed costs are charged (such as some direct labour and machinery operating costs) or full fixed costs (including full labour costs and depreciation on machinery and buildings). In the longer term, the net margin is a better indicator of performance.

By way of example, Table 3 shows the income losses (excluding ESA payments) associated with a switch from Tier 1 Grassland (the dominant grassland system) and arable systems to an extensive washland system. For the overall dairy and livestock mix in the Parrett catchment, the reduction in gross margin is about £260/ha, and in net margin (after full fixed costs) about £90/ha. The impact of a switch from cereal-based arable cropping to washland is also shown.

At present ESA rates (with Tier 1 at £125/ha), annual payments for a washland option would probably need to be about £300/ha/year (in 2001 values) to attract farmer interest. An analysis of whole farm budgets for dairy farms and beef farms in the catchment showed that a payment of about £300/ha/year would maintain the viability of these farms in a washland environment. Site specific environmental enhancements would need to be identified and built into the washland prescriptions. It envisaged that the washland option would specify grassland management requirements such as grass sward composition, grazing/cutting regimes, and zero chemical nitrogen.

An economic analysis of washland creation (MAFF, 1999; Penning-Rowsell et al., 2003) shows that there appears to be advantage from the viewpoint of the national economy of reducing the intensity of farm production in washland areas of the Levels and Moors. That is, some types of farming are unprofitable in the absence of subsidies paid per ha of crop or head of livestock. As a consequence the nation is 'better off' if these hidden losses are avoided, although there are important implications, at least in the short term, for local incomes and economy.[6]

Furthermore, it can be argued that agri-environment payments to farmers associated with reduced intensification and environmental protection are indicative of society's willingness to pay for environmental goods and services. Registering these as a benefit of washland development increases the economic value of the washland option. Given the opportunity to achieve economic and environmental benefits through washland creation, and through targeted support to help sustain incomes to the farming community, it would appear in the public interest to redirect funding, both from agricultural support and flood defence for agriculture, into flood storage and washland creation. The introduction of the single payment scheme and an inundation grassland supplement in higher

[6]In this situation, decoupling support to farm incomes from commodity prices is a more economically efficient way of supporting the rural economy than supporting commodity prices. Many farmers in the area reportedly reduced stock numbers following the introduction of the Single Payment Regime because they were no longer profitable.

Table 3
Reduction in the value of financial indicators associated with a switch from Tier 1 type grassland or arable to extensive grazing on washland (excluding ESA payments, but including area payments on arable)

£/ha/year reductions 2001 values	Dairy	Beef	Beef and sheep	'Average' catchment dairy and livestock	Arable
Gross Margin (excl. forage costs)[a]	410	280	260	350	300
Gross margin (incl. forage costs)[a]	330	190	170	260	300
Net margin after semi-fixed costs[b]	215	140	100	170	150
Net margins after full fixed costs[c]	150	80	0	90	90

[a]Forage casts equate to agrochemical and silage/haymaking costs.
[b]Direct labour and machinery operating costs only.
[c]Including labour costs, machinery operating and depreciation, and housing/building costs for stock. Adapted from Farm Business Survey Data from the University of Exeter.

level environmental stewardship has initiated this process, but there exists further potential for more extensive flood storage and washland creation options.

5. Administrative options for washlands

There are a number of alternative forms of management and administration for washland creation and operation (for examples, see English Nature, 2001a). These include land purchase, easements on flooding, management agreements supported by annual payments and leaseback partnership arrangements. For the Parrett catchment study, these were screened against the criteria of effectiveness, efficiency, fairness, risks and whether they had a good chance of meeting the overall objective of wise use of floodplains (Table 4). It was shown (Morris et al., 2002) that the suitability of these options vary according to the purposes to be achieved, the need to provide long-term robust solutions, and linked to these, the preferred link between the farming community and the management of the land.

All of the aforementioned approaches are potentially feasible. Land purchase, annual payments and leaseback have their particular advantages, disadvantages and risks, and suit the interest of different groups. Easements may be appropriate to accommodate modest increases in flood risk, but probably not for regular deep flooding and where there is a wish to achieve environmental enhancement.

The diversity of circumstance and practice within floodplain areas suggests that a diversity of approaches to washland administrative arrangements will be needed. A mosaic of land tenure arrangements may be acceptable provided this can deliver the scale, integration and reliability of service required.

Further, institutional and administrative arrangements for flood storage and washland creation should reflect the land management and funding mechanisms. Given the multiple objectives to be served and organisations involved in land management and flood defence, a multi-agency approach is required to promote washlands as a future option for floodplain management. In the UK, the flood storage facility could be managed by the Environment Agency in collaboration with Defra and locally administered Internal Drainage Boards. Statutory and voluntary conservation organisations collaborating with land managers could help manage habitats. If it was decided to progress washland creation through an annual payment regime, it would make sense to administer this through existing mechanisms such as those operated by Defra or Natural England.

Finally, in addition to the possibility of land purchase, easement costs and/or annual payment, a meaningful estimate of costs needs to include design, supervision, engineering works, and operation and maintenance costs for the washland. These costs, reflecting the capital and operating costs of 'engineering' washland projects, are very site specific and will vary according to flood characteristics and impacts. However, it is likely they will replace some costs currently committed to conventional flood defence.

6. Conclusion

Creating facilities in the floodplain of a river catchment, which will store water and create washland, and which will potentially meet all or some of the multiple objectives of flood and water resources management, environmental enhancement, and support the rural farming community, is feasible. However, there is both synergy and conflict of interest in washlands amongst flood storage, environment and farming objectives. Different sites are likely to have different priorities and management systems. Accordingly, prescriptions for flood facilities, environmental and farming management will require local definition.

Further, the storage of winter floodwater and washland creation on farmland will result in income losses to farmers, and would mean a switch from arable to grass, and from improved grassland to extensive systems. The extent of loss of net income (revenues less costs) associated with a switch in cropping and extensification will depend on whether farmers can achieve savings in costs to offset reductions in revenues.

Set against this, is the scope to design a washland flood storage package, involving land purchase, easement, annual payments, or partnership/leaseback. Given the

Table 4
Administrative options for washlands

Option	Strengths	Weaknesses
Land purchase and transfer of ownership to authority or trust	Good chance of delivering flood storage and environmental objectives	Risk of reduced ties to farming community
		Problems of attracting and negotiating tenants
	Efficient up-front funding Funded under capital budgets	Difficult to arrange purchases in large blocks
	Provide exit route for some farmers	Relies on voluntary participation, unless made compulsory
Easement: one-off payment to compensate future flood risk	Focus on flood defence aspects	Less suited to significant changes in flood risk
	Suited to compensating risk of infrequent flood events	Less suited to delivering environmental enhancement
	Attractive to flood defence agency: one-off capital payment	
Annual payments to compensate for income loss and/or environmental enhancement	Potential to deliver range of objectives: social, economic and environmental	Inflate land prices, encourage subletting
	Maintain farmer and community links with land	Mixed success of ESA schemes
		Participation dependent on 'incentives'
	Farmer familiarity with payment mechanism	Expensive, dependent on 'revenue' budgets
	Integrate with ESA scheme	Create dependency
	Can be adjusted over time according to objectives/circumstances	
Lease-back: transfer of ownership or control to authority or trust with tenancies to previous owners	Ability to focus on scheme objectives	Administratively and legally complicated to establish
	Partnership approach	Reluctance to transfer assets, until scheme proven
	Farmers/community engaged in implementation	Requires clear community of interest amongst participants
	Diverse 'partner' funding sources	

diversity of circumstance and practice across the UK, a range of approaches to washland tenure arrangements is suggested, provided this can deliver the scale, integration and reliability of service required. Payments should reflect the different levels of risk. Where the owner maintains occupation, payments should also reflect the timing, duration and depth of flooding during the season.

At present, however, there may be a number of barriers to the adoption of washland options by farmers, especially relating to the viability and practicality of 'new' washland farming practices; the perceived adequacy and predictability of existing agri-environment payments, and insufficient targeting for flood storage and washland options; and concern about the potential irreversibility of wetland creation from an agricultural perspective.

It is important that ways are found to overcome these barriers, as at current levels of government support, there appears to be an economic advantage in moving to extensive washland farming systems. Furthermore, given the potential to achieve additional social, economic and environmental benefit, it would appear in the public interest to redirect funding into flood storage and washland creation as part of an integrated and sustainable approach to flood management. The current reappraisal of agricultural and rural policy objectives and options provides astounding opportunities to do this.

Acknowledgements

This paper is based on research commissioned by the EU Wise Use of Floodplains Project. The authors are grateful for their financial support. Nevertheless, the views expressed here and conclusions reached are solely the responsibility of the authors.

References

Barbier, E.B., Acreman, M., Knowler, D., 1997. Economic Valuation of Wetlands: A Guide for Policy Makers and Planners. The Ramsar Library. Ramsar Convention Bureau, Gland, Switzerland.

Defra, 2002. Environmentally Sensitive Areas: Explanatory Notes. Defra, London.

Defra, 2003a. The Countryside Stewardship Scheme: Information and How to Apply. Defra, London.

Defra, 2003b. Review of Agri-Environment Schemes: Third Public Consultation. Environmental Stewardship Higher Level, Annex A and Annex D. Conservation Management Division. Defra, London.

Defra, 2003c. IACS The Integrated Administration and Control System. IACS 1: 2003 Edition. Rural Payments Agency, Defra, London.

Defra, 2004a. Single Payment Scheme: Information for Farmers and Growers in England. July 2004 Update. Rural Payments Agency, Defra, London.

Defra, 2004b. Single Payment Scheme: Information for Farmers and Growers in England. November 2004 Update. Rural Payments Agency, Defra, London.

Defra, 2004c. Single Payment Scheme: Cross Compliance Handbook for England. 2005 Edition. Rural Payments Agency, Defra, London.

Defra, 2004d. Making Space for Water: Developing a Government Strategy for Flood and Coastal Erosion Risk Management in England. July 2004. Defra, London.

Defra, 2004e. IACS The Integrated Administration and Control System, IACS 1A: (Supp 2004). 2004 Supplement to 2003 Edition. Rural Payments Agency, Defra, London.

Defra, 2005a. Environmental Stewardship: Entry Level Stewardship Handbook. Terms and Conditions and How to Apply. Rural Development Service, Defra, London.

Defra, 2005b. Environmental Stewardship: Higher Level Stewardship Handbook. Terms and Conditions and How to Apply. Rural Development Service, Defra, London.

DETR, 1995. Biodiversity: The UK Steering Group Report, vol. 2. HMSO, London, UK.Dunderdale, J.A.L., Morris, J., 1997. Agricultural impacts of river maintenance, a method of assessment. Journal of Agricultural Engineering Research 68, 317–327.

Dunderdale, J.A.L., Morris, J., 1997b. The benefit and cost analysis of river maintenance. Journal of the Institute of Water and Environmental Management 11 (6), 423–430.

English Nature, 2000. The effectiveness of the floodplain ESA schemes in the maintenance and enhancement of biodiversity. No. 364 English Nature Research Report prepared for EN by Risk Policy Analysts, Peterborough.

English Nature, 2001a. Sustainable flood defence: the case for washlands. No. 406 Research Report prepared for EN by Risk Policy Analysts, Peterborough.

English Nature, 2001b. Somerset Levels & Moors SPA/Ramsar/SSSI: An Introduction to its Importance. Taunton, Somerset.

Environment Agency, 2002. Agriculture and Natural Resources: Benefits, Costs and Potential Solutions. Environment Agency, Bristol (May).

Joyce, C.B., Wade, P.M., 1998. European Wet Grasslands: Biodiversity, Management and Restoration. Wiley, Chichester, UK.

MAFF, 1999. Flood Defence Project Appraisal Guidance (FCDPAG3), Economic Appraisal. Ministry of Agriculture Fisheries and Food, London.

Morris, J., Vivash, R., Alsop, D., Lawson, C., Leeds-Harrison, P., Bailey, A., 2002. Economic basis and practicalities of washland creation on the Somerset moors and levels. Report to EU Wise Use of Flood Plains Project, Cranfield University at Silsoe and River Restoration Centre, Silsoe.

Morris, J., Hess, T.M., Gowing, D.J. Leeds-Harrison, P.B., Bannister, N., Wade, M. and Vivash, R.M., 2004a. Integrated washland management for flood defence and biodiversity. Report to Department for Environment, Food and Rural Affairs & English Nature. Cranfield University at Silsoe, Bedfordshire, UK, March.

Morris, J., Bailey, A., Alsop, D., Vivash, R., Lawson, C.S., Leeds-Harrison, P.B., 2004b. Integrating flood management and agri-environment through washland creation in the UK. Farm Management 12 (1), 1–16.

Morris, J., Hess, T.M., 1987. Agricultural flood alleviation benefit assessment: a case study. Journal of Agricultural Economics 38 (2), 402–412.

Penning-Rowsell, E., Johnson, C., Tunstall, S., Tapsell, S., Morris, J., Chatterton, J., Coker, A., Green, C., 2003. The Benefits of Flood and Coastal Defence, Techniques and Data 2002. Flood Hazard Research Centre, Middlesex University.

Rodwell, J.S. (Ed.), 1991 et seq. British Plant Communities, vol. 1–5. Cambridge University Press, Cambridge (Published on behalf of the Joint Nature Conservation Committee with a research team of C.D. Pigott, D.A. Ratcliffe, A.J.C. Malloch, H.J.B. Birks, M.C.F. Proctor, D.W. Shimwell, J.P. Huntley, E. Radford, M.J. Wigginton and P. Wilkins).

Tallowin, J.R.B., Jefferson, R.G., 1999. Hay production from lowland semi-natural grasslands: a review of implications for ruminant livestock systems. Grass and Forage Science 54, 99–115.

Westlake, D.F., 1966. Biomass and Productivity of *Glyceria maxima*. I. Seasonal changes in Biomass. Journal of Ecology 54, 745–753.

Wilson, A.M., Ausden, M., Milsom, T.P., 2004. Changes in breeding wader populations on lowland wet grassland in England and Wales: cause and potential solutions. Ibis 146 (Suppl. 2), 32–40.

University of Exeter, 1999. Farm Management Handbook. Agricultural Economics Unit, University of Exeter.

University of Exeter, 2000. Farm Management Handbook. Agricultural Economics Unit, University of Exeter.

University of Exeter, 2001. Farm Management Handbook. Agricultural Economics Unit, University of Exeter.

Continued from outside back cover

Journal of
Environmental Management

July 2008 Volume 88 Number 2

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(200807)88:2;1-F