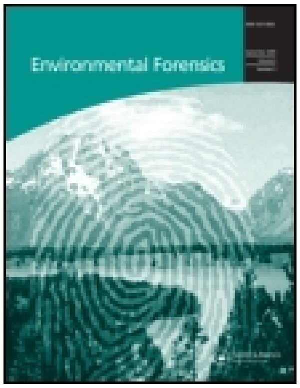

Environmental Forensics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uenf20

Applying Spatial Data Mining for Watershed Site Selection to Perform Field Sampling

Yen-Jen Lee [a] , Cheng-Mau Wu [b c] , Andy I-Fu Shen [d] , Kuo Chyang Chang [e] , Jet-Chau Wen [f] & Ginn-Yein Chen [g]

[a] Graduate School of Engineering Science and Technology, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan, ROC

[b] Research Center for Soil & Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan

[c] Department of Landscape Architecture, National Chiayi University, Chiayi, Taiwan, ROC

[d] Environmental Protection Administration, Executive Yuan/Department of Water Quality Protection, Chung-Hwa Road, Taipei, Taiwan, ROC

[e] Water Resources Agency, Ministry of Economic Affairs/Planning Division, Taipei, Taiwan, ROC

[f] Department and Graduate School of Safety Health and Environment Engineering, Research Center for Soil & Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan, ROC

[g] Department of Civil and Water Resources Engineering, National Chiayi University, Chiayi, Taiwan, ROC

Version of record first published: 21 Aug 2012

Environmental Forensics, 13:238–247, 2012

ISSN: 1527-5922 print / 1527-5930 online
DOI: 10.1080/15275922.2012.702329

Applying Spatial Data Mining for Watershed Site Selection to Perform Field Sampling

Yen-Jen Lee,[1] Cheng-Mau Wu,[2] Andy I-Fu Shen,[3] Kuo Chyang Chang,[4] Jet-Chau Wen,[5] and Ginn-Yein Chen[6]

[1]*Graduate School of Engineering Science and Technology, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan, ROC*
[2]*Research Center for Soil & Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan, and Department of Landscape Architecture, National Chiayi University, Chiayi, Taiwan, ROC*
[3]*Environmental Protection Administration, Executive Yuan/Department of Water Quality Protection, Chung-Hwa Road, Taipei, Taiwan, ROC*
[4]*Water Resources Agency, Ministry of Economic Affairs/Planning Division, Taipei, Taiwan, ROC*
[5]*Department and Graduate School of Safety Health and Environment Engineering, Research Center for Soil & Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan, ROC*
[6]*Department of Civil and Water Resources Engineering, National Chiayi University, Chiayi, Taiwan, ROC*

This study presents a systematical site selection framework for a wide-range survey. Spatial data mining was used to delimit three classes of pollution potential areas. Results show that pollution samples highly correlate with the classified potentially contaminated areas. For example, most of the soil pollution and water pollution appear under the high potentially contaminated area (HPCA). In contrast, under the low potentially contaminated area (LPCA), no pollution was discovered. These findings shows that the proposed approach is fairly reliable and can be applied to wide-range of areas within a river watershed to determine site selections for performing field sampling.

Keywords: spatial data mining, geographical information system (GIS), site selection, field sampling

Illegal dumping of industrial wastes has recently become one of the most serious social and environmental problems (Simons, 1989), especially with polluted soil and groundwater created by illegal dumping of industrial wastes from the past decades. These wastes not only pollute soil and groundwater but also cause health problems (Lakshmikantha, 2006). Due to these ramifications, many pollution-producing factories and illegal dumping sites were closed down and these lands were redeveloped into residential areas or public facilities. In other words, could it be that these sites were redeveloped with contaminants still existing in the subsurface? If so, could it be possible that these contaminants or toxic wastes may become exposed later after endangering the health of residents or accidentally become uncovered through erosion or some other means? The urgency of such environmental problems in recent years has been generally acknowledged. More and more effort is, therefore, being put into working out realistic solutions to such problems. Researchers are currently seeking to characterize waste disposal sites which include the detection of the location and extent of contamination patches in areas as small as landfill sites (Calvo et al., 2005). In such context, the integrated use of geophysical methods and chemical analysis was implemented for the evaluation and characterization of contaminated sites (Orlando and Marchesi, 2001).

Since the late 1970s, Taiwan has become more industrialized and prosperous, but its capability to treat waste has not kept pace with this advancement. In particular, some smaller cities, factories, and farms often discharge wastes into the rivers or accumulate them along the riverbanks with little or no treatment. Rivers in southern Taiwan are affected by various contaminants, including domestic wastes, animal husbandry wastes, agricultural pesticides, dioxins, polychlorinated biphenyls (PCBs), and heavy metals (Sun et al., 2009). The majority of research in southern Taiwan has focused on the Era-Jiin River (referred to as *Erren River* herein) as the most visibly polluted river (Hung and Shy, 1995). Many toxic wastes were dumped along its riverbanks. Electronic waste recyclers and metal smelters accounted for approximately 80% of all illegal dumping

Address correspondence to Jet-Chau Wen, Department and Graduate School of Safety Health and Environment Engineering, Research Center for Soil & Water Resources and Natural Disaster Prevention, National Yunlin University of Science and Technology, 123, Section 3, University Road, Douliou, Yunlin 64002, Taiwan, ROC. E-mail: wenjc@yuntech.edu.tw

activity along the Erren River, especially downstream from the conjunction of Kang-Wei-Kuo Creek to the estuary. Heavy metals, such as lead, cadmium and tin, from circuit boards were haphazardly discarded and great quantities of industrial wastes were found which created a lot of serious environmental problems. Restoration of the Erren River has been ongoing since 2001. The Environmental Protection Agency of Taiwan (EPA Taiwan, 2004) has spent $50 to $60 million (Taiwan dollars) to clean up sites along the river. However, funding for the clean-up effort has been difficult to secure. In 2007, huge amounts of electronic waste (e-waste), which included stripped electronic circuit boards, plastic-coated metals, and unknown composites, were found on both sides of the riverbanks during a riverbank construction project along a 3-kilometer stretch downstream to the estuary (Industrial Technology Research Institute, ITRI, 2009). This incident created some concern to nearby residents and attracted great attention from environmental groups. The Taiwan EPA then began to investigate other locations along the riverbanks or watershed to see whether they were contaminated. Taiwan EPA authorities were eager to know the pollution status of the Erren River watershed in order to propose a management and remediation plan. The main objective of field sampling in this case was to provide information on the environmental status of the Erren River Basin. However, due to the time and cost concerns, field sampling could not be conducted thoroughly over the whole basin. Representative sites then were selected to be assessed.

In regard to site selection, there are many research studies on mapping, selecting, and monitoring waste disposal sites using variety of methods. Muttiah et al. (1996) used geographical information system (GIS)-based simulated annealing to select waste disposal sites. Hokkanen and Salminen (1997) used multicriteria decision analysis to choose a solid waste management system. Lahdelma et al. (2000) proposed various multicriteria methods aimed at supporting such complex planning and decision processes by providing a framework for collecting, storing, and processing all relevant information. Soupios et al. (2007) used an integrated method, which combined geophysical methods and chemical analysis for characterization and management of landfills. Calvo et al. (2005) developed a method to carry out environmental diagnosis of the landfill sites. Simsek et al. (2006) proposed a solid waste disposal site selection procedure based on using a groundwater vulnerability mapping.

With the support of advanced technologies such as GIS, airborne, and satellite remote sensing, and global positioning system (GPS), better spatial information can be obtained (Sharifi and Retsios, 2004; Simsek et al., 2006). This procedure involves a pre-survey work, acquisition of aerial photography, development of an appropriate classification system to be used in the inventory process, survey of public information, airphoto analysis and georeferencing, site prioritization, and finally, implementation of a monitoring program. The technique was described as being "effective" in performing comprehensive inventories of waste disposal sites over county-sized areas.

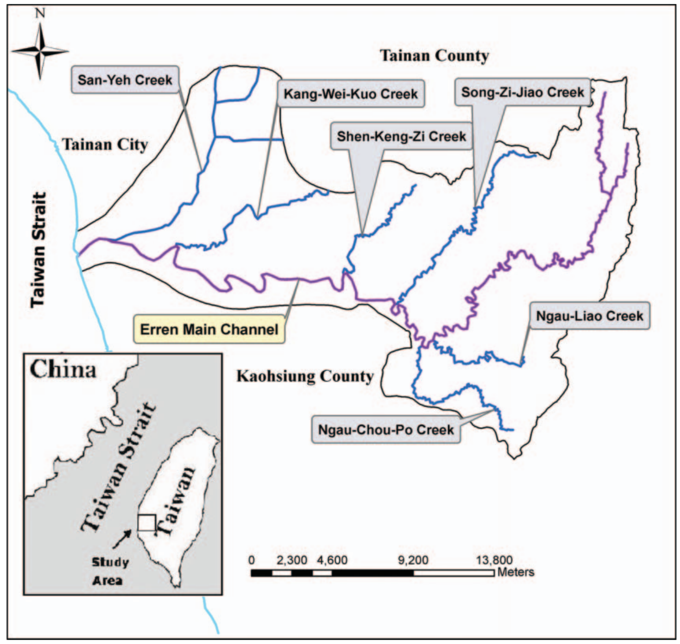

Figure 1. Map of the Erren River Basin (Taiwan). (color figure available online.)

This study presents an improved framework of selecting sampling sites of potentially contaminated areas to perform sampling and surveying for contaminants. In order to know how serious pollution (including pollutants and their extent) is along a river, a comprehensive survey is needed. However, it is time consuming and costly to conduct comprehensive field sampling within a river basin. Therefore, representative site selections within a whole river basin for soil and groundwater sampling are always conducted through various strategies for different purposes. The significant locations have to be chosen and a systematical procedure has to be established.

In this study, the Erren River Basin (Figure 1) was selected as an example to conduct a framework of site selection for further sampling and inspection. An integrating approach of 3S—meaning GIS, remote sensing (RS), and GPS—and spatial data mining for potentially polluted sites is presented. Spatial data mining was used to classify potentially polluted areas (PPAs) into three groups based on information gathered on the nature and distribution of industrial operations, historical pollution events, illegal dumping sites, reports and interviews of local residents. GIS was used to create maps of PPAs by overlaying the available historical background data. Suspected spots were then located through aerial-photo interpretation of the PPAs (Wu et al., 2010). Finally, field checks were carried out to inspect the suspected spots located by aerial-photo interpretation, and the candidate sites were then chosen for soil sampling.

Procedure of Site Selection

The procedure of site selection is performed according to the schematic of forensic concept as shown in Figure 2. The

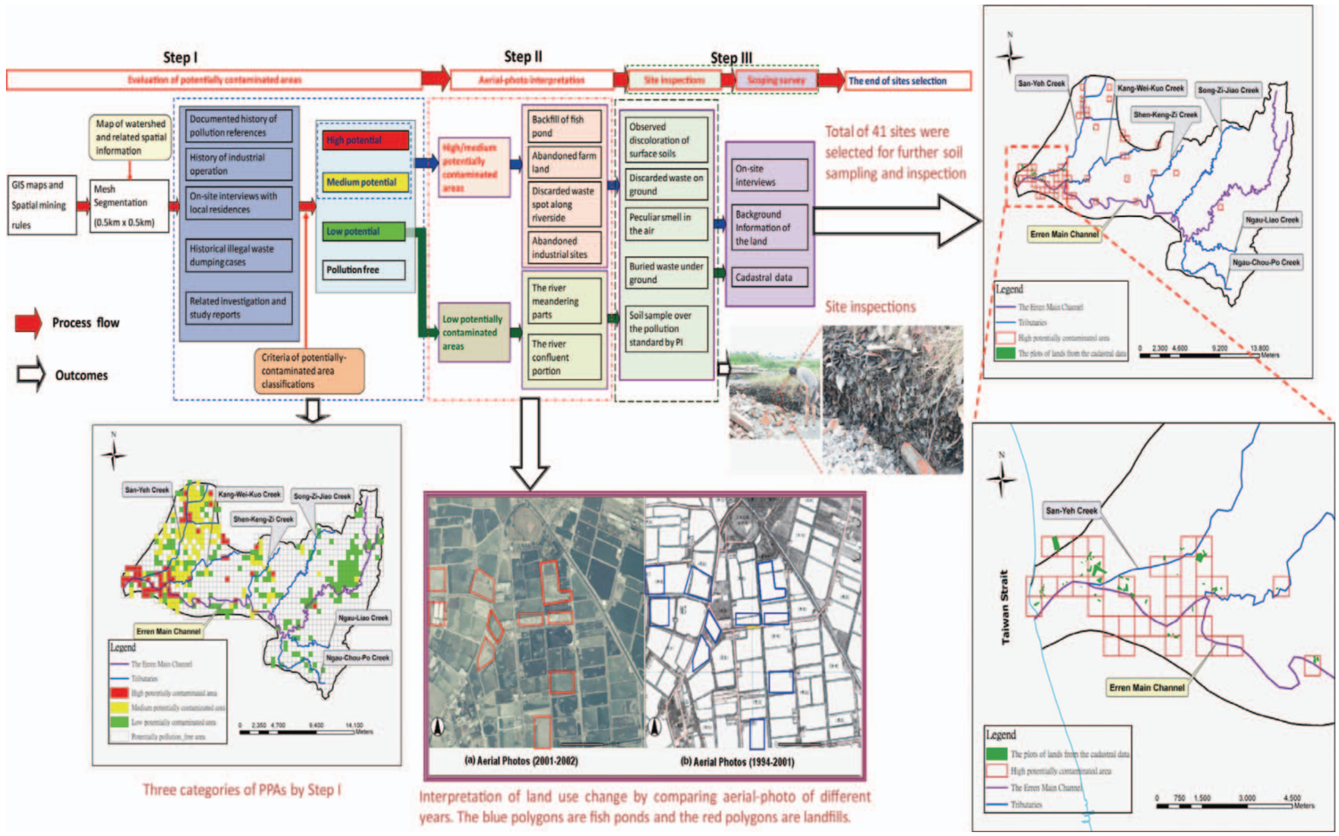

Figure 2. Schematic illustration of forensic concept. (color figure available online.)

systematical procedure includes three assessment steps, namely Step I, Step II, and Step III; each builds upon the previous step with more intensive information (Figure 3). Step I assessment is evaluation of potentially contaminated areas, which provides a basis for further investigation by Step II. Step II involves aerial-photo interpretation for high, medium, and low PPAs that are located in Step I according to historical background data. Step III is conducted after the suspicion of contamination are confirmed in Step II. Step III assessment includes two tasks: site inspections and scoping surveys.

Step I: Evaluation of Potentially Contaminated Areas

In Step I, GIS was used to create maps of PPAs by overlaying the available historical background data. The data set used included: 1) history of industrial operation, 2) historical illegal waste dumping cases, 3) related investigation and study reports regarding the historical background, 4) on-site interviews with local residences, 5) digital maps of study area, and 6) GIS shapefiles (industrial area distribution, rivers, roads, etc.). According to classification rules of potentially contaminated areas, four classes of areas were established, which are high potentially contaminated, medium potentially contaminated, low potentially contaminated, and potentially pollution-free areas (Figure 4).

Grid segments were delimited before carrying out classifications for potentially contaminated areas. The grid of 0.5 km × 0.5 km was made within a 339-square kilometer area

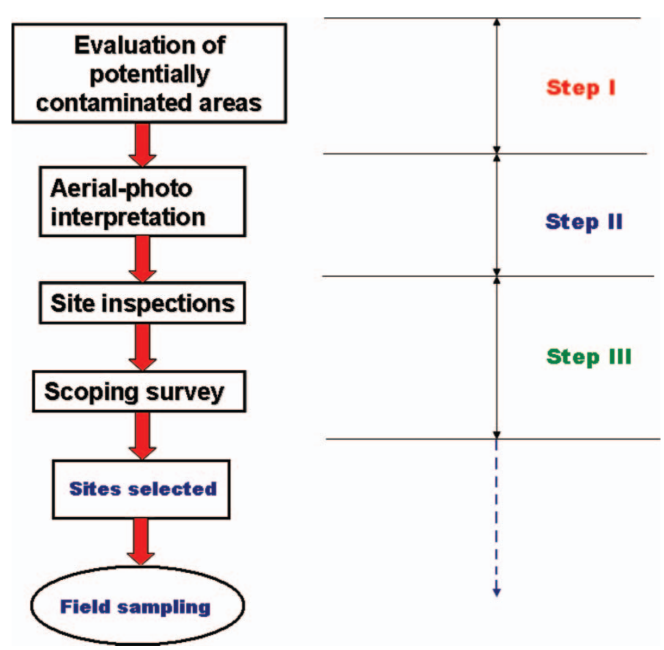

Figure 3. Procedure of site selections for performing further field sampling. (color figure available online.)

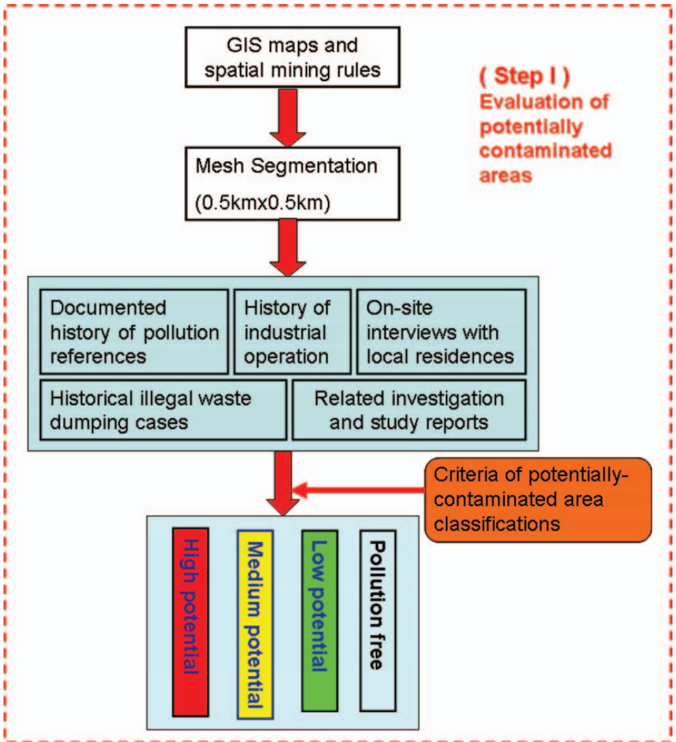

Figure 4. Flowchart for evaluation of potentially contaminated areas (Step I). (color figure available online.)

of the Erren River watershed (Figure 5). Selection of the grid size 0.5 km × 0.5 km was based on the criteria of a mesoscale sampling-area survey for soil pollution of crop farms by the EPA of Taiwan (2004).

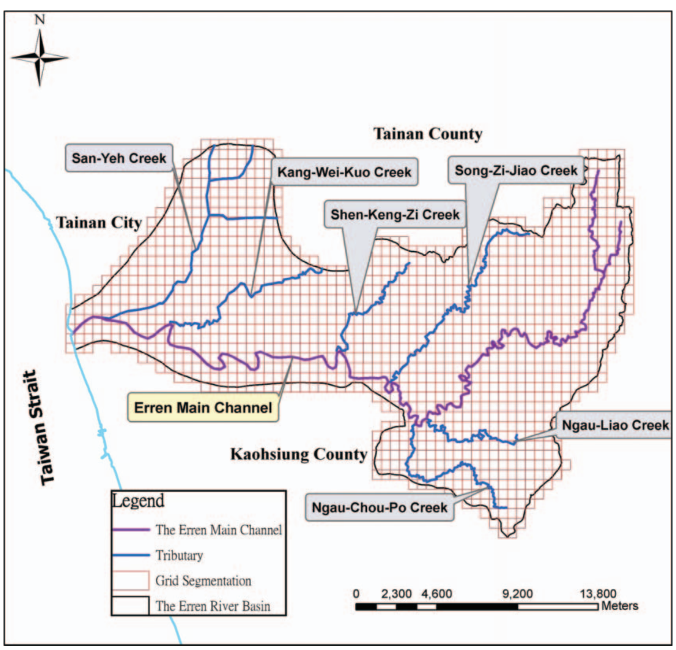

Figure 5. Grid segmentation of Erren River (Step I). (color figure available online.)

The classification criteria of potentially contaminated areas are:

1. *High potentially contaminated area (HPCA)*: The HPCA was defined by the following facts uncovered within a grid: (a) historical sample results of soil (or water) pollution over the standard regulation level (EPA Taiwan, 2000), (b) existing illegal waste dump or previous contaminated sites listed by local EPA for follow-up, (c) a suspicious site reported by environmental groups or claimed by residents, and (d) occurrence of any historical pollution-incident with any high pollution industry as stipulated by the EPA regulations (EPA Taiwan, 2000).
2. *Medium potentially contaminated area (MPCA)*: The MPCA was defined by one of the following facts uncovered within a grid: (a) a site where there was documented releases of contaminants into the environment, or (b) previous site of a high pollution industry as stipulated by the EPA regulations (EPA Taiwan, 2000).
3. *Low potentially contaminated area (LPCA)*: The LPCA was defined as a site in a grid as never having a high pollution industry as stipulated by the EPA regulations (EPA Taiwan, 2000) and never having any of the other criteria mentioned for H/MPCAs.
4. *Potentially pollution-free area (PPFA)*: A PPFA was exclusive of all the other three classes and the land was utilized under normal agricultural operations.

Step II: Aerial-Photo Interpretation

The flowchart of Step II: aerial interpretation and Step III: site inspection and scoping survey is presented in

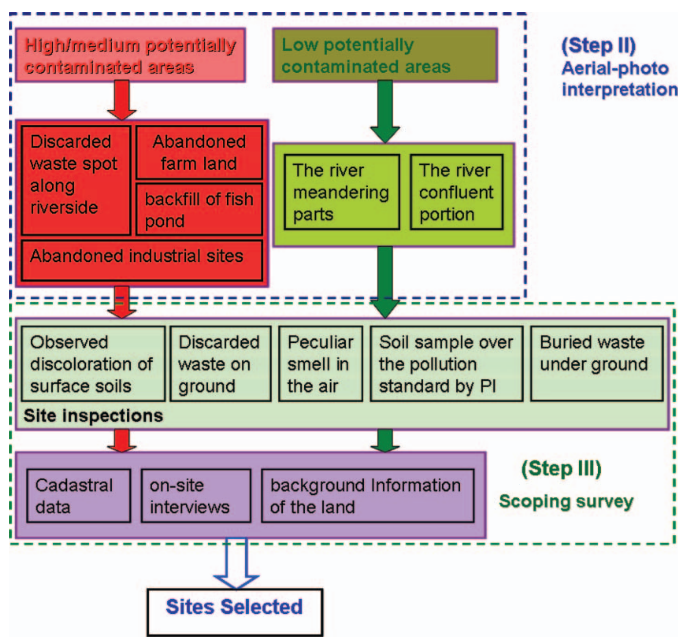

Figure 6. Flowchart for site selection of Step II and Step III. (color figure available online.)

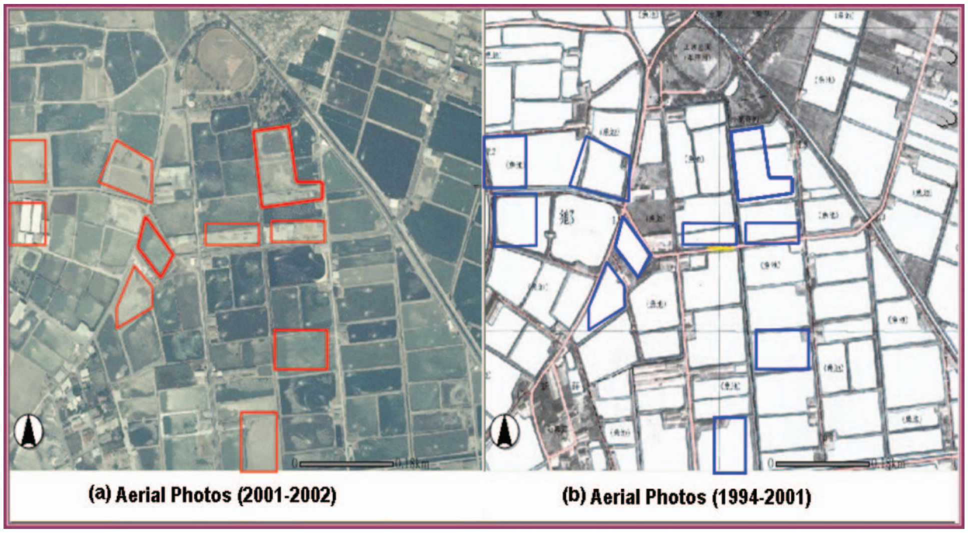

Figure 7. Map of the land use change by comparing aerial-photos of different years at the same locations: (a) the red polygons representing filled in fishponds (years 2001–2002) in comparison with (b) the blue polygons representing existing fishponds (years 1994–2001). (color figure available online.)

Figure 6. In Step II, the four classes of areas determined by the GIS evaluation were further identified by aerial-photo interpretation.

Owing to land utilization and historical background, the H/MPCAs and the LPCAs were distributed over different portions of the watershed. The aerial-photos of suspicious locations were examined at different targets. The screening conditions for the H/MPCAs were: (1) fish ponds converted to landfills; (2) discarded waste spots along riversides; (3) abandoned farmlands; and (4) abandoned industrial sites. In other words, any one of these conditions was chosen preferentially to be probed for contaminants in the H/MPCAs. All these factors are obviously related to pollutant sources, and professional judgments were considered as well.

In contrast, for LPCAs, the river meandering parts and the river confluent portions were examined. By comparing aerial photos of different years, land use showed evidence of changes (e.g., utilization, land cover, or aberration) and the suspected spots were then located on aerial-photos (Figure 7), and then Step III for the site selection procedure was performed. As can be seen in Figure 7, land use of 2001–2002 changed from that of 1994–2001, which was traced back to fishponds

Figure 8. Site inspection photo images: (a) Electronic waste covered by soil of riverbank. (b) E-waste from recycling operations included ink recovery, burning of plastic-coated metals, plastic recovery, solder collection, and gold extraction. (color figure available online.)

Table 1. *Main pollution sources of the Erren River*

Sources	Animal Husbandry	Industry	Domestic	Seepage
Description	Swine, duck, chicken, and fish farming	Heavy metal refinery, electrical processing, acid washing	Sewage and garbage	Non-point source
Percentage (%)	51	38	10	<1

Source: (Simons, 1989; Hung and Shy, 1995; Sun et al., 2009).

that were filled in and covered up. In Figures 7(a) and 7(b) at the same locations, the red polygons representing filled in fishponds (years 2001–2002) are compared with the blue polygons representing existing fishponds (years 1994–2001), respectively.

Step III: Site Inspections and Scoping Surveys

Two tasks were included in this step, namely site inspections and scoping surveys.

Site inspections

Site inspections and field checks were carried out to examine the suspected spots located by aerial-photo interpretation. The suspected spots were further checked by site investigators. The location of each spot was targeted and verified using GPS, a camera, the field-portable X-ray fluorescence (XRF), photo ionization detector (PID), flame ionization detector (FID), and site lists. Physical appearance checks were performed, which included (1) observed discoloration of surface soils, (2) discarded waste on ground, (3) peculiar smell in the air, (4) soil samples over the pollution standard detected by primary inspection (PI) via XRF, PID, and FID, and (5) uncovered waste buried underground.

To carry out the PI process, the field-portable XRF tests were conducted (Figure 8) for quick site checks and pollutant screening (Wu et al., 2012). The field XRF was used to rapidly pre-screen samples directly on site samples and to obtain the optimal utility from the laboratory sampling effort.

Scoping surveys

Scoping surveys were carried out to decide the candidate sites, which were then chosen for further soil samplings. To complete a scoping survey, three data requirements were needed: (1) background information of the land, (2) on-site interviews with landowners, and (3) cadastral data.

The digital cadastral data were obtained from the National Land Surveying and mapping Center (NLSC) of the Interior Department. With the screening conditions in Step II and by overlapping the layer of digital cadastral maps with the grids of the HPCA map created by GIS, the boundaries of selected sites were determined. Furthermore, on-site interviews were conducted with the landowners, which is different from on-site interviews in Step I (with local residents).

Study Area

In this study, the Erren River was selected as an example to conduct a framework of site selection. With approximately a 339-square kilometer drainage area, the Erren River is approximately 62.5 kilometers in length and the bed slope is approximately 1:142. It flows through Tainan County and Kaohsiung County, past Tainan City, and runs into the Taiwan Strait. The average annual rainfall of the Erren River basin is 1909.9 mm and the mean annual runoff is about 498.88×10^6 m^3. Six tributaries that flow into the Erren River from upstream to downstream are Ngau-Liao Creek, Ngau-Chou–Po Creek, Song-Zi-Jiao Creek, Shen-Keng-Zi Creek, Kang-Wei-Kuo Creek, and San-Yeh Creek (Figure 1). The major pollution sources of the Erren River are summarized in Table 1. Animal wastes, mainly from swine farming, and to a lesser extent, duck, chicken, and fish farming, were the main pollution sources along these rivers, with untreated animal wastes frequently seeping directly into the water. Industrial pollution was also a significant source of contamination. The heavy metal refinery industry along the Erren River was a major source of pollution and PCBs and dioxins were also discharged directly into the river. The proposed framework prioritized the sites into four classes of potentially contaminated areas (Figure 4). The worst sites (high potential sites) were given more attention.

Field Sampling

After site selection was performed by the proposed procedure, the candidate sites were chosen. In this paper, forty-one selected sites were chosen and field sampling was carried out on these sites. Based on the collected information of each site, both judgmental sampling and random sampling were taken individually. Three types of samples were collected during field sampling, which include soil, groundwater, and wastes. Each individual sampling method was based on the Taiwan EPA standard (2000, 2006) as shown in Table 2.

Direct push technology (DPT) with dual-tube was applied for soil sampling. DPT was conducted from the surface to groundwater level. The topsoil with a depth of 0– 15 cm and the subsoil with a depth of 15–30 cm were collected. Then, soil samples were collected per 1.2 m in depth until the groundwater level was reached. Soil samples were rapidly pre-screened using primary inspection (PI) by XRF for heavy metals and by PID and FID for volatile organic compounds (VOCs). Both PID and FID methods are applicable for VOCs. The PID is a site-screening

Table 2. Standards for field sampling and laboratory experiments

Object	Taiwan EPA*	Related Standards	Description
Soil	NIEA S102.61B	ISO11466.2	Soil sampling method.
	1. NIEA S103.61C	U.S. EPA SW – 846	1. General regulations and test methods for soil samples.
	2. NIEA M103.01C	ISO / DIS 11464	2. Laboratory experiment methods for heavy metals (ICP-AES)
	3. NIEA S701.60C	ASTM D7691 – 11 US EPA,SW - 846, Method 4030	3. Laboratory experiment methods for petroleum hydrocarbons
Ground water	NIEA W106.50C	ASTM, D-6001-96 ASTM, D6282-98	Groundwater sampling method.
	NIEA W102.51C	U.S. EPA/240 / B–01/002	General regulations and test methods for water and wastewater.
	NIEA W785.54B	US EPA Method 524.2	Laboratory experiment methods for purgeable organic compounds in water
Wastes	1. NIEA R118.02B	JIS K0060	Waste sampling method.
	2. NIEA M152.01C	U.S. EPA., Method 5000	Sample preparation for volatile organic compounds.
	NIEA R101.02C	U.S. EPA SW-846 ASTM D4547-98	General regulations and test methods for evaluating solid waste.

*By Environmental Analysis Laboratory, EPA Taiwan. Available at: http://www.niea.gov.tw/analysis/method/m_t.asp.

test used for examining polynuclear aromatic hydrocarbons, and FID is used for methane or butane types of chemical compositions.

When pollution was discovered during PI, then groundwater was sampled. In the case where waste was discovered whether on the surface or under the surface at any point, samples were then collected for further toxicity characteristic leaching procedure (TCLP) testing. These collected soil, groundwater, and waste samples underwent lab-testing to confirm the pre-screening results. Soil samples were analyzed for possible metal ion concentrations or major pollutants which were anticipated based on the industrial activity in the surrounding area. A total of 117 soil samples were screened on the field using FID and portable XRF (Wu et al., 2012). Groundwater samples collected were analyzed for various parameters, using groundwater standards as the reference.

To justify the site selection framework, sixty top ranking soil samples that revealed soil pollutants by pre-screening of the selected sites were delivered to the laboratory for further inductively coupled plasma atomic emission spectrometer (ICP-AES) analysis. Water quality analysis and TCLP were conducted on both groundwater samples and waste samples individually.

Results and Discussion

Results of Site Selection

Figure 9 shows the selected site distributions obtained by Step I. According to Step I of the site selection procedure, the HPCAs (red grids) are concentrated along the downstream parts from the conjunction of Kang-Wei-Kuo Creek to the estuary. Most of the MPCAs (yellow grids) are distributed along San-Yeh Creek. The LPCAs (green grids) are scattered within the Erren River Basin with most concentrations along the upper stream part of Erren River. The PPFAs (white grids) are excluded of all the above three classes and the land use involved normal agricultural operations.

Based on the results of Step I, the second step was carried out. Figure 7 presents the selected spots obtained by Step II, the aerial-photo interpretation and comparisons. Interpretation of land use change was performed by comparing aerial photos from different years. Step III, site inspections and scoping surveys, was conducted after Step II. Figure 10 presents the site distributions after site inspection and scoping surveys for the selected spots in the previous step. The cadastral data were used to delimit the boundaries of the selected sites. As shown in Figure 11, the plots of land (green polygons) are marked on the figure. For further field sampling, 41 sites were selected. Site inspections were carried out and the information of the landowner and historical operation on

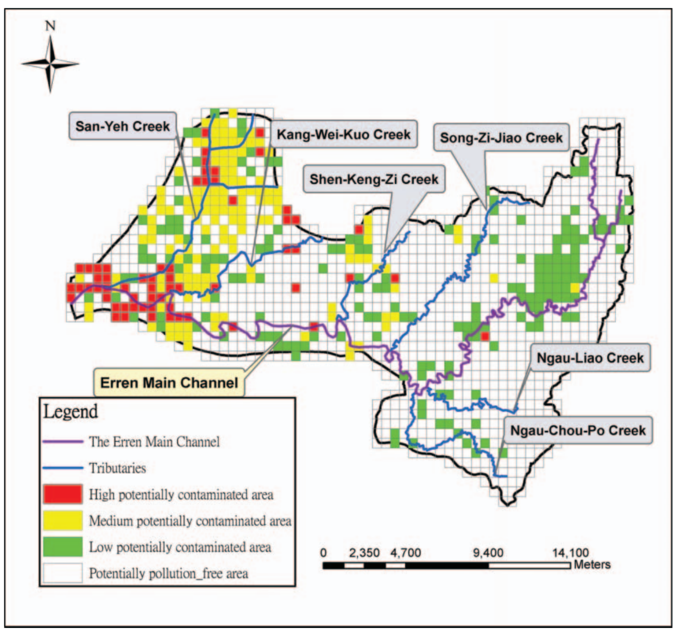

Figure 9. Map of the selected sites distribution (Step I). (color figure available online.)

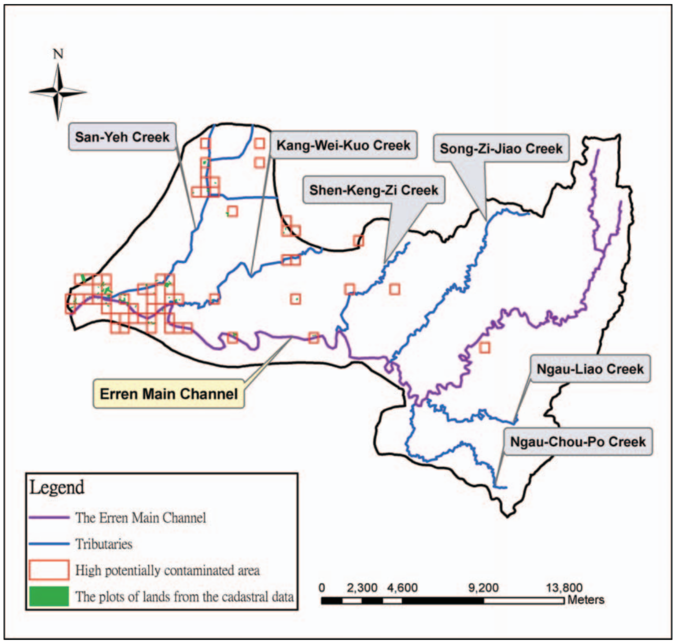

Figure 10. Map of the selected sites distribution by aerial-photo interpretation (Step II). (color figure available online.)

the land was collected before field sampling. Soil samples collected in the field were delivered to the laboratory for further analysis by the inductively coupled plasma atomic emission spectrometer (ICP-AES), and relative laboratory experiments were carried out for groundwater samples and waste samples simultaneously.

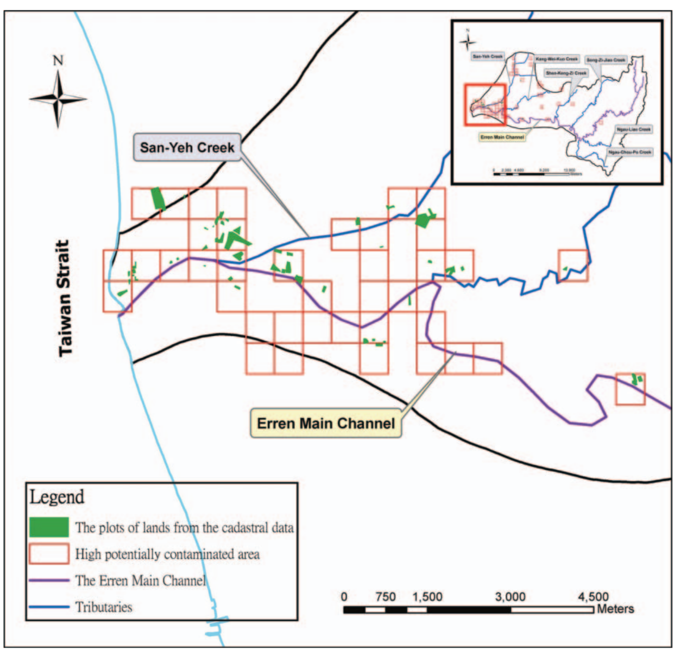

Figure 11. Map of the distribution of selected sites determined by site inspection and scoping surveys (Step III). (color figure available online.)

Results with Field Soil Sampling Surveys

A total of 117 soil samples were pre-screened in the field using XRF, PID, and FID screening tests. From those samples, 60 top-ranking soil samples that revealed soil pollutants when pre-screened were selected for further confirmation analytical laboratory analysis. The confirmation analysis used ICP-AES according to the soil contamination standard for heavy metals

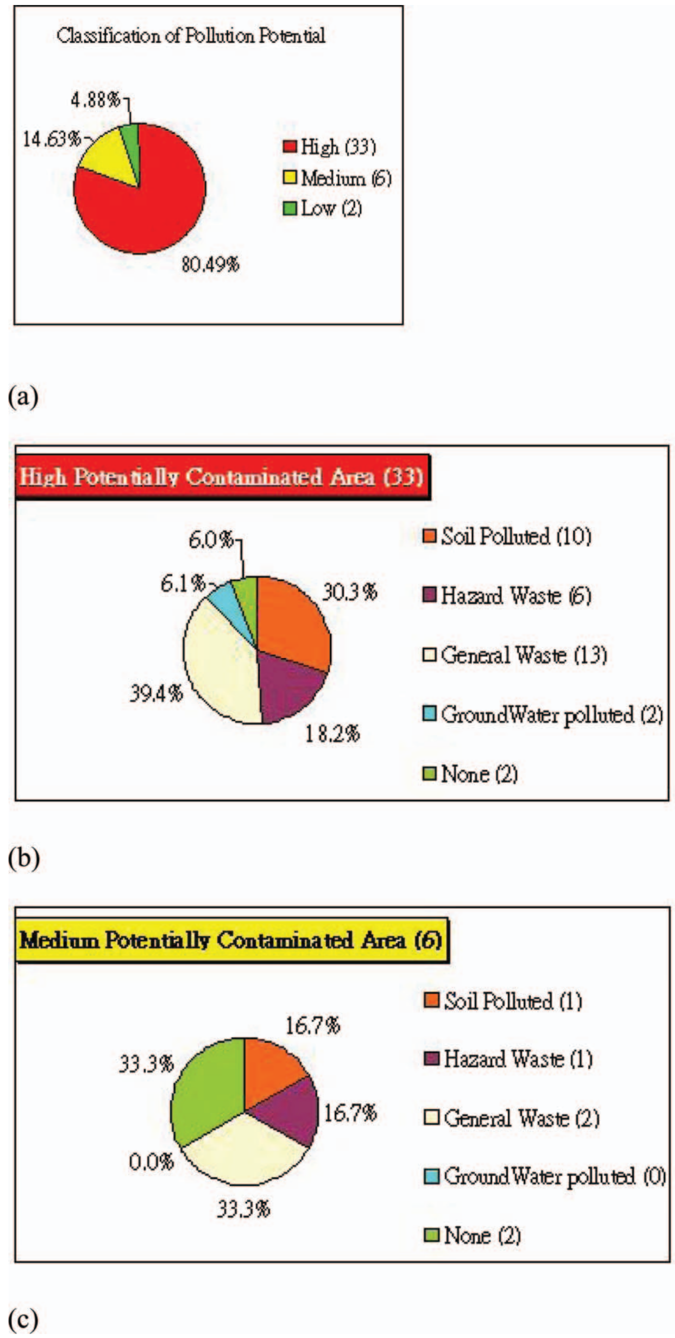

Figure 12. Pie charts of site selection results. (a) Percentage of the selected sites in different classifications. (b) Percentage of various contaminants in HPCAs. (c) Percentage of various contaminants in MPCAs. (color figure available online.)

Table 3. *Soil contamination standard for heavy metals (EPA, Taiwan, 2000, 2006)*

Elements Concentration	Ni mg/kg	Cu mg/kg	Zn mg/kg	Pb mg/kg	Cd mg/kg	Cr mg/kg	Hg mg/kg	As mg/kg
Minimum Detection Limit‡ (MDL)	1.35	0.72	6.62	1	0.1	1.37	0.037	0.499
Pollution Threshold Limit to be controlled‡ (PTL)	200	400(200§)	2000(600)	2000(500)	20(5)	250	20(5)	60
Monitored Threshold Limit‡ (MTL)	130	220(120)	1000(260)	1000(300)	10(2.5)	175	10(2)	30

‡ Taiwan EPA Standard (EPA Taiwan, 2000, 2006); § Limit for farmland

(Table 3) and organic compounds (Table 4) (EPA Taiwan, 2000, 2006). The ICP-AES analysis results were then used to confirm whether or not the site selection procedure was correct. Table 5 shows the results of the numbers over the pollution threshold limit (PTL) to be controlled (in Table 3) and were obtained from XRF and ICP-AES individually (Wu et al., 2012).

The following discussions were made according to 41 selected sites conducted by ICP-AES experiments. Table 6 shows the result of the site selection procedures and Figure 12a shows the percentages regarding those results. Among the 33 HPCAs (Figure 12b), there are 10 (30.3%) selected sites with soil contamination, two (6.1%) with groundwater pollution, and 19 (57.6%) waste sites, which include six (18.2%) hazard waste plus 13 (39.4%) general waste; within the six MPCAs (Figure 12c), there is one (16.7%) selected site with soil contamination, none with ground pollution, and three (50.0%) (16.7% plus 33.3%) waste dumping sites which include one (16.7%) with hazardous waste and two (33.3%) without; and no pollution is found in the LPCAs. On the other hand, the risk of soil pollution in the HPCAs is nearly two times of the MPCAs.

Table 4. *Soil contamination standard for organic compounds (EPA Taiwan, 2000, 2006)*

Item	Description	Soil pollution control standards (mg/kg)
001	Benzene	5
002	Carbon tetrachloride	5
003	Chloroform	100
004	1,2-Dichloroethane	8
005	(cis-1.2-Dichloroethylene)	7
006	(trans-1,2-Dichloroethylene)	50
007	1,2-Dichloropropane	0.5
008	1,2-Dichlorobenzene	100
009	1.3-Dichlorobenzene	100
010	3,3′-Dichlorobenzidine	2
011	Ethylbenzene	250
012	Hexachlorobenzene	500
013	Pentachlorophenol	200
014	Tetrachloroethylen	10
015	Toluene	500
016	TPH (Total Petroleum Hydrocarbons)	1000
017	Trichloroethylene	60
018	2,4,5-Trichlorophenol	350
019	2,4,5-Trichlorophenol	40
020	Vinyl chloride	10

Table 5. *The numbers over the pollution threshold limit (PTL) obtained from x-ray fluorescence (XRF) and inductively coupled plasma atomic emission spectrometer (ICP-AES) samples*

Element XRF	No. over PTL†	ICP-AES No over PTL	RP‡ (%)
Pb	7	6	85.71
Zn	15	12	80.00
Ni	20	10	50.00
Cu	48	17	35.42
As	12	3	25.00
Cr	30	5	16.67
Cd	52	3	5.77
Hg	43	1	2.33

†Pollution Threshold Limit; ‡ Relative Proximity: Number of detected samples over the PTL in the field of the ICP-AES results divided by the number of detected XRF results over the PTL.

Furthermore, the risk of groundwater pollution in the HPCAs is much greater than the MPCAs and LPCAs. Among the 41 selected sites, eleven were soil-polluted sites and two groundwater polluted sites with heavy metals, and twenty-two dumping sites (seven of them were hazardous industrial waste sites).

Summary

A total of 41 sites were selected in which 33 sites were HPCAs (80.49%), six sites were MPCAs (14.63%), and two sites were

Table 6. *Site selection results, examined by an inductively coupled plasma atomic emission spectrometer (ICP-AES)*

Classification of Pollution Potential	HPCA	MPCA	LPCA	Subtotal no. (Pcs.)
Selected sites (Pcs.)	33 (80.49%)	6 (14.63%)	2 (4.88%)	41 (100%)
Soil pollution (no.)	10	1	0	11
	30.3%	16.7%	0	26.8%
Illegal dumping				
Hazardous waste (no.)	6	1	0	7
	18.2%	16.7%	0	17%
General waste (no.)	13	2	0	15
	39.4%	33.3%	0	36.6%
Groundwater	2	0	0	2
pollution (no.)	6.1%	0	0	4.9%

LPCAs (4.88%) as shown in Figure 12a. As can be seen from Table 6, Figure 12b, and Figure 12c, pollution samples highly correlate with the classified potentially contaminated areas. For instance, most of the soil pollution and water pollution appear under the HPCA category. In contrast, under the LPCA category, no pollution appears.

This systematical site selection framework is different from conventional field sampling where the specific polluted site of the latter is already delimited; the only matter of concern is how to conduct a rational site sampling method to collect representative samples. The proposed approach was applied to a wide-range area—a river watershed for site selection. The consequent field sampling was carried out according to the chosen sites.

Conclusions

Just as a prosecutor works on evidence collection, so it is with our proposed systematical site selection framework. The more background information obtained, the more accurate a contaminant site is targeted. Philosophy of the proposed framework also matches the forensic process. It is inevitably a problem where sufficient information is needed at different stages (Stage I to III) in order to carry out proper decisions. This proposed solution provides an effective framework when used together with the tools of data mining, which can reduce the burden of cumbersome and time-consuming problems. This framework is also fairly reliable when applied to a wide-range survey. It is useful for field sampling and also for local environmental protection departments of county governments for continuing environmental protection. Integrating 3S and spatial data mining for site inspection of pollutants will make a great contribution to future environmental contaminant detection.

Acknowledgement

The authors would like to acknowledge the Environmental Protection Agency (EPA) of Taiwan for its support under grant EPA-98-GA12-03-A108, the National Science Foundation of Taiwan for grant NSC 99-2221-E-224-028-, and the Water Resources Agency, Ministry of Economic Affairs, of Taiwan for grant WRA-0990029.

References

Calvo, F., Moreno, B., Zamorano, M., and Szanto, M. 2005. Environmental diagnosis methodology for municipal waste landfills. *Waste Management* 25:768–779.

Environmental Protection Agency of Taiwan. 2000. Soil and Groundwater Remediation Act. In *Compilation of laws and regulations of Soil and Groundwater Remediation*, pp. 1–22.

Environmental Protection Agency of Taiwan. 2004. *Soil and Groundwater Pollution Remediation Annual Report*, pp. 2–4.

Environmental Protection Agency of Taiwan. 2006. Appendix 2: Soil pollution grading table. In *Regulations Governing the Preliminary Assessment of Soil and Groundwater Pollution Control Sites*. Available at: :http://law.epa.gov.tw/en/laws/278076101.html#att02

Hokkanen, J., and Salminen, P. 1997. Theory and methodology. Choosing a solid waste management system using multicriteria decision analysis. *European Journal of Operational Research* 98:19–36.

Hung, J. J., and Shy, C. P. 1995. Speciation of dissolved selenium in the Kaoping and Erhjen (Erren) Rivers and Estuaries. *Southwestern Taiwan Estuaries* 18:234–240.

Industrial Technology Research Institute (ITRI). 2009. *Report on the Erren River Remediation and Reclamation Along Wei-Ziu-Nei section of Hu-Nei Township in Kaohsiung County*, pp. 1–15 [in Chinese].

Lahdelma, R., Salminen, P., and Hokkanen, J. 2000. Using multicriteria methods in environmental planning and management. *Environmental Management* 26(6):595–605.

Lakshmikantha, H. 2006. Report on waste sites around Bangalore. *Waste Management* 26:640–650.

Muttiah, R. S., Engel, B. A., and Jones, D. D. 1996. Waste disposal site selection using GIS-based simulated annealing. *Computers & Geosciences* 22(9):1013–1017.

Orlando, L., and Marchesi, E. 2001. Georadar as a tool to identify and characterise solid waste dump deposits. *Journal of Applied Geophysics* 48:163–174.

Sharifi, M. A., and Retsios, V. 2004. Site selection for waste disposal through spatial multiple criteria decision analysis. *Journal of Telecommunications and Information Technology* 3:1–11.

Simons Li Consulting Co. 1989. *Pollution Control Plan of Era-Jiin River, Drainage Basin*. Taipei: Taiwan: Simons Li Consulting Co.

Simsek, C., Kincal, C., and Gunduz, O. 2006. A solid waste disposal site selection procedure based on groundwater vulnerability mapping. *Environmental Geology* 49:620–633.

Soupios, P., Papadopoulos, N., Papadopoulos, I., et al. 2007. Application of integrated methods in mapping waste disposal areas. *Environmental Geology* 53:661–675.

Sun, P. L., Hawkins, W. E., Overstreet, R. M., and Brown-Peterson, N. J. 2009. Morphological deformities as biomarkers in fish from contaminated rivers in Taiwan. *International Journal of Environmental Research and Public Health* 6:2307–2331.

Wu, C.-M., Chen, G.-Y., and Wen, J.-C. 2010. Integration of 3S and spatial information for site inspection of pollutants. In *Proceedings of the H21A-165, 2010 Western Pacific Geophysics Meeting*, 22–25 June 2010 (Taipei International Convention Center, Taipei, Taiwan).

Wu, C.-M., Tsai, H.-T., Yang, K.-H., and Wen, J.-C. 2012. How reliable is XRF measurements for different metals in soil contamination? *Environmental Forensics* 13:110–121.