Environ. Sci. Technol. **2009,** *43,* 6662–6668

Determining the Probability of Arsenic in Groundwater Using a Parsimonious Model

JIN-JING LEE,[†] CHENG-SHIN JANG,[‡] CHEN-WUING LIU,[*,§] CHING-PING LIANG,[∥] AND SHENG-WEI WANG[§]

Environmental Professionals Training Institute, Environmental Protection Administration, Executive Yuan, Taoyuan, Taiwan 320, ROC, Department of Leisure and Recreation Management, Kainan University, Luzhu, Taoyuan, Taiwan 338, ROC, Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan, 106, ROC, and Department of Environmental Engineering and Science, Fooyin University, Kaoshiung, Taiwan, 831, ROC

Received February 19, 2009. Revised manuscript received May 15, 2009. Accepted July 17, 2009.

Spatial distributions of groundwater quality are commonly heterogeneous, varying with depths and locations, which is important in assessing the health and ecological risks. Owing to time and cost constraints, it is not practical or economical to measure arsenic everywhere. A predictive model is necessary to estimate the distribution of a specific pollutant in groundwater. This study developed a logistic regression (LR) model to predict the residential well water quality in the Lanyang plain. Six hydrochemical parameters, pH, NO_3^--N, NO_2^--N, NH_4^+-N, Fe, and Mn, and a regional variable (binary type) were used to evaluate the probability of arsenic concentrations exceeding 10 μg/L in groundwater. The developed parsimonious LR model indicates that four parameters in the Lanyang plain aquifer, (pH, NH_4^+, Fe(aq), and a component to account for regional heterogeneity) can accurately predict probability of arsenic concentration $\geq 10 \mu$g/L in groundwater. These parameters provide an explanation for release of arsenic by reductive dissolution of As-rich FeOOH in NH_4^+ containing groundwater. A comparison of LR and indicator kriging (IK) show similar results in modeling the distributions of arsenic. LR can be applied to assess the probability of groundwater arsenic at sampled sites without arsenic concentration data apriori. However, arsenic sampling is still needed and required in arsenic-assessment stages in other areas, and the need for long-term monitoring and maintenance is not precluded.

Introduction

Arsenic has been identified as a major risk factor for blackfoot disease (BFD) on the southwestern coast of Taiwan (*1*). The main source of arsenic exposure for the general population is the ingestion of drinking water with high levels of arsenic. Ingestion of arsenic compounds in well water is also associated with age-adjusted mortality from diabetes (*2*), hypertension and cerebrovascular disease (*3*), and cancers of the lung, liver, bladder, and kidney in Taiwan (*4*, *5*). Arsenic is widely distributed in the Earth's crust and can concentrate in pyrite, hydrous Fe oxides, and sulfide compounds and can also be adsorbed by Fe and Mn after the weathering of sulfide minerals in sedimentary environments (*6*). Arsenic can be released into groundwater through redox processes such as the oxidation of arsenic-bearing pyrite minerals and through the reductive dissolution of arsenic-rich Fe-oxides (*7*); these processes depend on pH and redox conditions in the aquifer (*8*) and may be an important geochemical interface controlling the speciation and the mobility of many elements along the flow direction (*9*). The distribution and transport of arsenic in groundwater are associated with the complex nature of arsenic chemistry and the geochemical variability of groundwater according to the hydrogeological setting, which can be examined to determine the occurrence of arsenic (*10*).

The estimated spatial distribution of contaminated groundwater quality is important in assessing the health and ecological risks. However, spatial distributions of groundwater quality are commonly heterogeneous, varying with depths and locations. This is problematic because only limited field data can be obtained owing to time and cost constraints, and sparsely measured data contain considerable uncertainty. A model is needed because it is not practical or economical to measure arsenic everywhere. Generally, the nonparametric geostatistical method, indicator kriging (IK), is widely used to model the spatial variability and distribution of field data with the uncertainty. The values estimated by IK represent the probability that the concentration of the specific element does not exceed a prespecified threshold (*11*). IK has been applied to estimate the probability distribution of heavy metal pollution in fields, to delineate hazardous areas (*12*, *13*), and to evaluate the extent of arsenic-polluted aquifers (*14*). However, arsenic testing in a laboratory is a complicated and costly process. A model using relevant hydrochemical parameters to evaluate the probability of arsenic exceeding a particular concentration in groundwater is necessary in assessing health risk. Multivariate statistical techniques are commonly used to examine relationships among several variables. Logistic regression (LR) is one of the most frequently used multivariate statistical methods for exploring environmental and epidemiological issues and predicting risks to human health and ecology. Logistic regression differs from classical linear regression in which the modeled response is the probability of being in a category, rather than the observed quantity of a response variable (*15*). The LR model has been applied to predict the probability of arsenic exceeding 5 μg/L in groundwaters (*16*), to identify the important variables that markedly increase the probability of NO_3-N contamination (*17*), and to explain a compound associated with pesticide pollution levels in groundwater (*18*).

Because arsenic in groundwater is closely correlated with other parameters, the objective of this study is to develop an alternative approach using a multivariate method (LR) to predict probability of specific arsenic concentrations exceeding Taiwan's EPA limit in Lianyang's residential wells based on groundwater NO_3-N, NO_2-N, NH_4^+-N, pH, Fe(aq), and Mn(aq), as well as a variable accounting for regional variability between northern and southern portions of the aquifer plain. The utility of the LR model is that it can be applied to assess the potential exposure risk of arsenic to individuals without measuring arsenic concentration directly in order to locate risk areas and to source safe drinking wells.

---

[*] Corresponding author phone: +886-2-2362-6480; fax: +886-2-2363-9557; e-mail: lcw@gwater.agec.ntu.edu.tw.

[†] Environmental Protection Administration.

[‡] Kainan University.

[§] National Taiwan University.

[∥] Fooyin University.

10.1021/es900540s CCC: $40.75

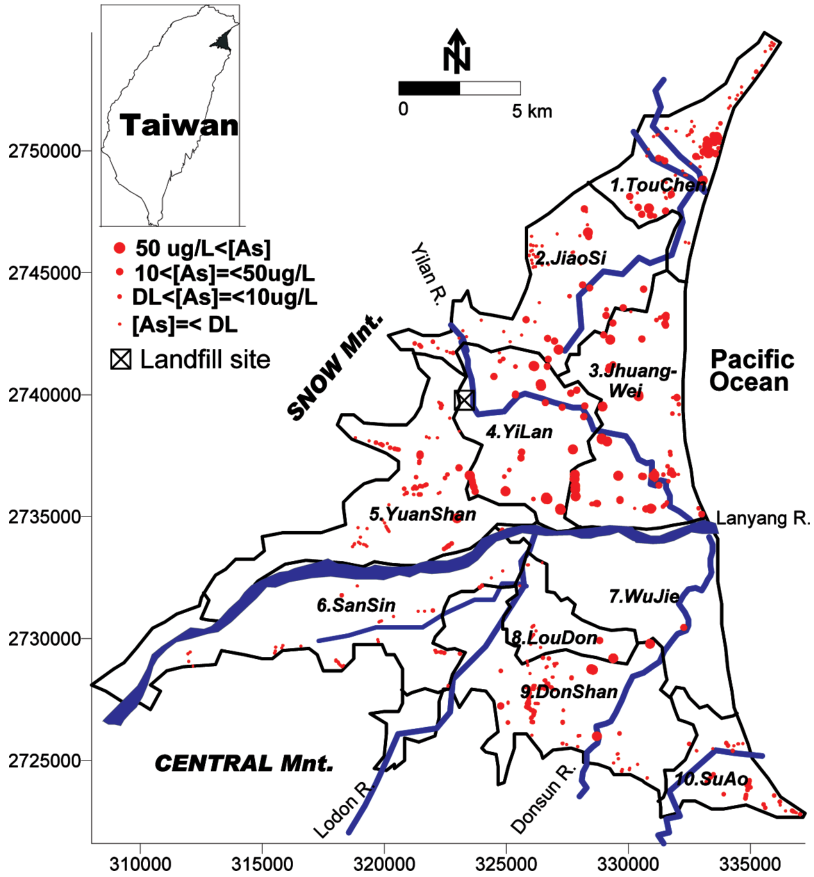

FIGURE 1. Study area and spatial distributions of measured arsenic concentrations.

Materials and Methods

Study Area. The Lanyang plain, which is approximately 400 km^2 (about 18.6% of total county area) and is located in Yilan County in northeastern Taiwan, consists of ten townships and is the home of 97% of the population (about 451,000) in Yilan County (Figure 1). The study area is the alluvial fan of the Lanyang River. The groundwater flows from west to east. The western parts of the plain near by the mountains form the main recharging area of the groundwater, and natural recharge is the source of groundwater (*19*). Although the major river in the Lanyang plain is the Lanyang River, several small rivers also flow through the area. In the northern portion, the Yllan River is the main tributary and receives water mainly from Snow Mountain. In the southern portion, the Lodon and Donsun Rivers are the main tributaries and receive water mainly from Central Mountain. The varied sources of groundwater recharge have caused different hydrogeochemical environments in the northern and southern portions of the Lanyang plain. The surface layer is covered by sediments from the Quaternary period, including silt, sand, and clay, and is partitioned into proximal, mid, and distal areas (*20*). The sediments contain abundant Fe and Mn oxides (*21*).

Arsenic levels in the well waters differ markedly between regions. The main source of exposure of residents to inorganic arsenic is through drinking water obtained from the wells (*22*). Arsenic in groundwater mainly results from geogenic release from sedimentary formations of the marine deposits formed during the Quaternary period (*23*). Although a high percentage of households have access to tap water, groundwater is still used as a popular source of drinking water by many residents in the study area. Arsenic concentrations that exceed the current Taiwan Environmental Protection Administration (Taiwan EPA) limit of 10 μg/L are out of compliance with WHO provisions. The arsenic concentration in some well water is up to 600 μg/L or higher, and most residential wells are less than 40 m deep (*24*). Arsenite and arsenate contents represented 87% and 5.8% of the total arsenic content, respectively (*25*). Long-term exposure to inorganic arsenic from drinking water had a significant dose–response relationship with various internal cancer in the Lanyang plain, but no blackfoot disease (BFD) case has been found here, indicating the geochemical characteristics of arsenic-affected groundwater in the Lanyang plain may be different from the BFD areas in the southwestern Taiwan region (*22, 24*).

Aquaculture and agriculture are the primary economic sectors of local residents since 1970. Fishponds, which are a mix of groundwater and seawater, are mostly situated in the northern region and near the Lanyang River in the southern region, near the coast (*26*). Croplands, such as rice paddies, are dispersed throughout this study area, where irrigation water has been supplied by the Yilan irrigation association for over 80 years. Factories are located on the southern region of the Lanyang plain; a large amount of groundwater has been extracted to supply the need of agriculture, aquaculture, and industry.

Well Water Samples. This study used data collected by the Environmental Protection Bureau (YiLan EPB, Taiwan) of YiLan County during 1997–1999 (*27*). A total number of 928 well water samples were collected in the study area. Figure 1 presents the spatial distributions of arsenic concentrations in the collected samples. All groundwater samples were collected from household wells. The well water was run through a pumping tube for at least 10 min before sampling to ensure that the water sampled was representative of the aquifer water chemistry. Seven water quality parameters were analyzed, including concentrations of arsenic, ammonia (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-), iron (Fe), manganese (Mn), and pH. Sample pH was measured in the field using

TABLE 1. Logistic Regression Classification for Calibration, Validation, and a Combined Data Set in Different Threshold Values

data set (number of observations)	threshold value	classification criteria, for 50% probability cut point: total correct responses	model sensitivity	model specificity	model goodness-of-fit test (HL)
calibration (600)	1 μg/L	81.5	73.9	88.3	<0.001
calibration (600)	5 μg/L	85.8	62.4	93.6	<0.001
calibration (600)	10 μg/L	87.7	54.1	95.3	0.54
validation (328)	10 μg/L	88.1	64.6	92.1	

a pH probe. Other items were analyzed in the laboratory. The procedures for analyzing arsenic concentrations followed the APHA Method 3500-AsB. Water samples were filtered with 0.45 μm glass microfiber filter papers and acidified with HNO_3 (Merck ultrapure grade) to pH< 2. A graphite furnace atomic absorption spectrometer (GFAAS, Perkin-Elmer Model 2100), a hydride generation system (HG, MHS-10), and an automatic sampler (AS-70) were adopted in arsenic analysis, with the detection limit of 0.9 μg/L. A GFAAS was used to determine Fe and Mn concentrations; the Nesslerization method, colorimetric method, and ultraviolet spectrophotometric screening method were used for determining concentrations of NH_4^+-N, NO_2^--N, and NO_3^--N, respectively. Strict quality control procedures, such as reagent blank analysis, field blank analysis, duplicated test, and check test, were executed during the field sampling (*27*). Table S1 (see the Supporting Information) presents a statistical overview of the data set. The average arsenic concentration was 11.82 ± 45.07 μg/L, with the maximum of 772 μg/L. Arsenic concentrations in approximately 82% of the 928 samples were below the Taiwan EPA limit of 10 μg/L, and 49.5% was below the instrumental detection limit of 0.9 μg/L. This study adopted a half of the instrumental detection limit to represent the value below the instrumental detection limit (*28*).

Logistic Regression. Logistic regression is widely applied to analyze nondetects because a threshold value is specified to define the response categories (*29*). The model takes the following form:

$$P[y = 1|x] = \frac{e^{b_0+b_1x_1+b_2x_2+...+b_kx_k}}{1 + e^{b_0+b_1x_1+b_2x_2+...+b_kx_k}} \qquad (1)$$

where *P* is probability of observing an event, *y* is an indicator variable, x_1, x_2,..., and x_k are explanatory variables, and b_1, b_2,..., and b_k are regression coefficients in eq 1. This study used the LR model to predict the probability that arsenic concentrations in wells exceeded the current Taiwan EPA limit for arsenic in drinking water (10 μg/L); $y = 1$ denotes a measurement exceeding 10 μg/L, and $y = 0$ denotes a measurement less than 10 μg/L. Other thresholds, including 5 and 1 μg/L, were also tested; however, these models were less correct and had lower goodness-of-fit than 10 μg/L (Table 1). Explanatory variables were pH and NO_3^--N, NO_2^--N, NH_4^+-N, Fe, and Mn concentrations, and a binary-variable region defined the northern and southern regions of Lanyang River as 0 and 1, respectively. The binary variable is included in our logistical model to account for local environmental differences, which, in our study area, come from differences in recharge from Snow and Central Mountains. The SAS system software (*30*) with stepwise selection was applied in this study, and selective evaluation of variables was followed. The Hosmer-Lemeshow (HL, *p* value) goodness-of-fit test

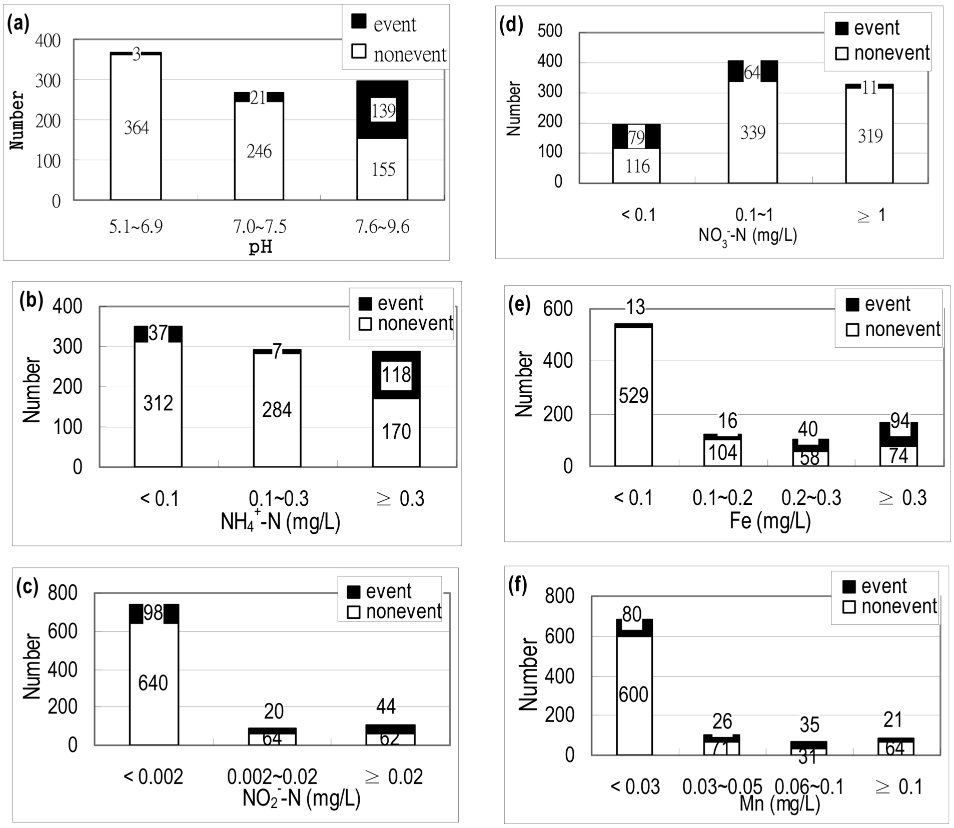

FIGURE 2. Frequency distribution of explanatory variables and occurrence of events (event: As > 10 μg/L; nonevent: As < 10 μg/L).

TABLE 2. Explanatory Variables in the Final Logistic-Regression Model

variable	estimated coefficient (b)	Wald *p*-value	odds ratio (exp(b))	95% LCL	95% UCL
intercept	−21.60	<0.0001			
pH	2.42	<0.0001	11.25	5.53	22.89
NH_4^+-N (mg/L)	0.95	<0.0001	2.59	1.75	3.84
Fe (mg/L)	1.96	<0.0001	7.08	2.92	17.18
region (0 vs 1)	1.21	0.0012	11.34	2.62	49.12

was used to examine the fit of values to the model, and the Wald statistic was utilized to test individual model parameters, with a significance level of $p < 0.05$. Model discrimination is represented by the *c* statistic (*29*). The overall rates of correct classification, sensitivity (rate of predicting true positive measurements), and specificity (rate of predicting true negative measurements) were computed. A predicted probability of 0.5 was used to delineate contaminated and uncontaminated samples and calculate the overall rates of correct classification, sensitivity, and specificity. Thus, a value of zero was assigned when the prediction P' $(Y_j = 1)$ was less than 0.5, and a value of 1 was assigned when the prediction P' $(Y_j = 1)$ equaled or exceeded 0.5. Calibration data of 600 samples (65%) were used to develop the initial model, and the remaining data (328 samples) were used for model validation. Samples used for calibration and validation were selected by random processes.

Results and Discussion

Logistic Regression. Figure 2 presents the relationships between explanatory variables and event occurrence. Event frequency was high when pH, NH_4^+, and Fe concentrations were high. Menard (*31*) indicated that high levels of collinearity may occur and pose problems for the model when correlation coefficients are higher than 0.8. Absolute values of the rank-based correlation coefficients (Spearman's rho) among independent variables were <0.55 (see Table S2 in the Supporting Information), suggesting that weak correlations are presented.

The final LR model consisted of the following four explanatory variables: pH, NH_4^+-N concentrations, Fe concentrations, and region (0 vs 1) (Table 2). An odds ratio of more than 1 indicates that the odds of detecting arsenic concentrations ≥10 μg/L increase when the explanatory variable increases (*31*). The presence of arsenic ≥10 μg/L in a well was associated with an 11-fold greater odds for the pH value, a 2-fold greater odds for the NH_4^+-N concentrations, a 7-fold greater odds for the Fe concentrations, and an 11-fold greater odds for region (0 vs 1). The *p*-value of HL was 0.43, and the observed and average predicted probabilities were well-correlated, indicating that the model fit was acceptable. The rate of total correct classification for the model was 87.7%, with 95.3% of arsenic concentrations less than 10 μg/L (specificity) and 54.1% of arsenic concentrations exceeding 10 μg/L (sensitivity) correctly classified. The *c* statistic was 0.94 for the LR model, indicating excellent

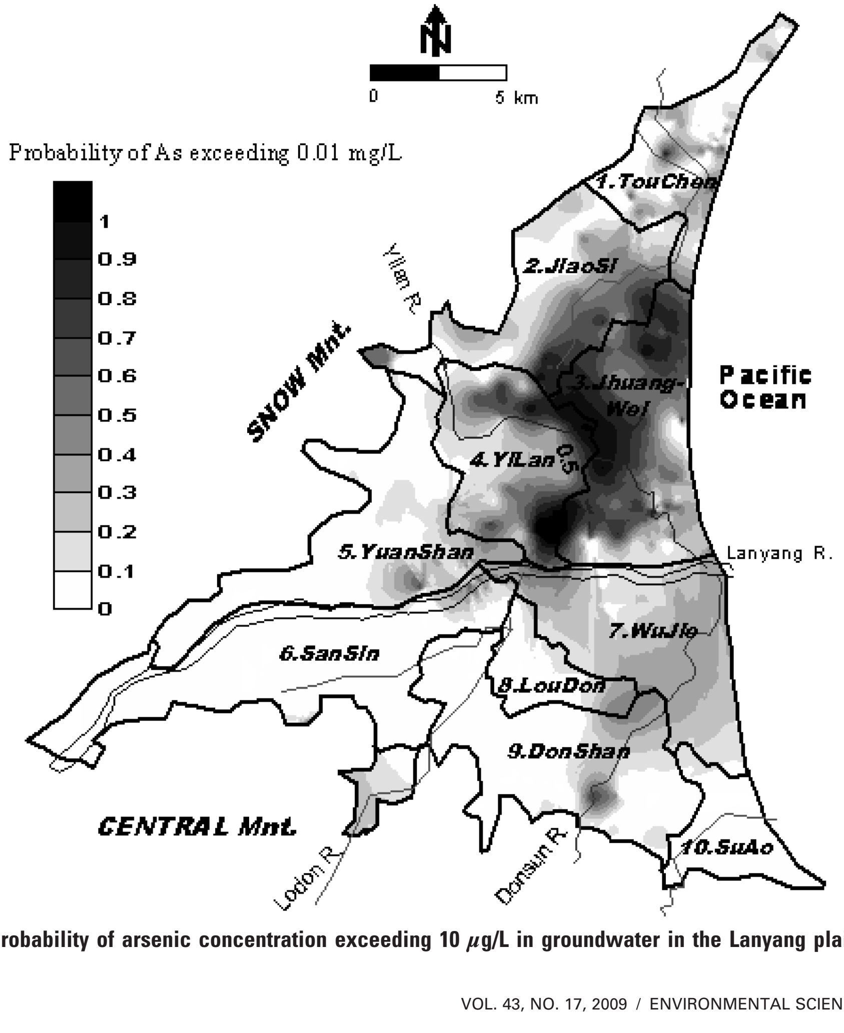

FIGURE 3. Predicted probability of arsenic concentration exceeding 10 μg/L in groundwater in the Lanyang plain.

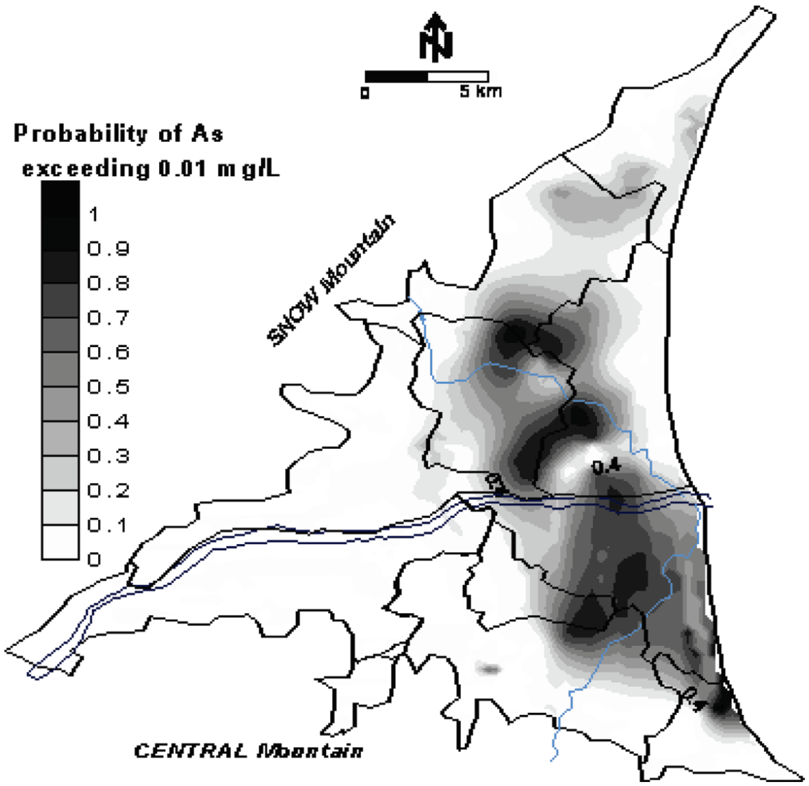

FIGURE 4. Predicted probability of arsenic concentration exceeding 10 μg/L in groundwater in the Lanyang plain by IK.

discrimination for the model (*29*). The proposed LR model was validated using an independent data set of 328 wells. The rate of total correct classification for the model is 88.1%, with 92.1% of arsenic concentrations less than 10 μg/L (specificity) and 64.6% of arsenic concentrations exceeding 10 μg/L (sensitivity). The observed and average predicted probabilities are reasonably well correlated. Thus, the proposed LR model correctly predicted events and nonevents according to the validation procedure. Our Supporting Information describes the validation of LR model.

How Selected Variables Can Account for Arsenic in Lanyang Plain's Groundwater. The correlation coefficients of arsenic concentration with pH, NH_4^+-N, and Fe were 0.59, 0.41, and 0.58, respectively, indicating that they were moderately corrected (Table S2 in the Supporting Information). Microbial degradation of sedimentary organic matter is regarded as the biogeochemical factor reductive dissolution of arsenic-rich FeOOH in anaerobic groundwater, causing the sorbed arsenic species to behave differently in conditions of differing pH (*32*). An increase in pH likely influenced their mobilization as well as changes in Hartree energy, E_h.The ratio of arsenite and arsenate to total arsenic concentration is controlled by pH, redox conditions in groundwater, and adsorption/desorption processes of metal hydroxides (*33*). The event frequency for pH ≥7 (28%) was higher than that of pH < 7 (0.8%) (Figure 2a) and may reduce affinity of As onto solid phases. The circumneutral values (median 7.1, range 5.1−9.6) in groundwater play a role in enhancing desorption of arsenic species adsorbed onto minerals (*10*).

Microbial oxidation of natural organic matter also generates high NH_4^+ concentrations in groundwater (*8*). Groundwater in mid and distal fan regions of Lanyang plain may be mixed with old and fresh groundwater; old groundwater may be attributed to mineralization of organic matter in sediment, and new groundwater may result from leaching from the ground above (*19*). Additionally, the landfill site is a traditional dumping site which was operated from 1980 to 1999, where no impermeable linear was placed underneath. The anaerobic pollutants easily infiltrate from the landfill site to groundwater and create a more local reducing environment (*34*). The presence of the NH_4^+ (media 1.6 mg/L, range <0.07−19.97 mg/L) reflects the moderately reducing condition indicating a dissimilatory reduction of NO_3-N may occur (*35*), in which arsenic and Fe ions can be released by the reductive dissolution of arsenic-bearing Fe-oxyhydroxides (*36*).

Fe oxides are the most likely sorbent of arsenic in aquifers and have been closely linked to the behavior of arsenic in soil and sediment (*37*). Chen et al. (*25*) investigated geochemical characteristics of porewater and alluded that arsenic was released to groundwater through reductive dissolution of Fe-oxides in the midfan area and was closely associated with mineralization of organic carbon in the distal fan area. The range of Fe concentrations is <0.09−3.48 mg/L (median <0.09 mg/L), and the event frequency at Fe concentrations ≥0.1 mg/L (39%) was higher than that at Fe concentrations <0.1 mg/L (2%) (Figure 2e). Arsenic can be released into groundwater via the reductive dissolution of Fe oxyhydroxides (*10*) or mobilized from an Fe oxide surface through desorption of arsenic species under neutral pH conditions (*7*).

The final LR model does not incorporate NO_2^--N, NO_3^--N, and Mn. The croplands (such as rice paddies) are dispersed throughout this study area, and high nitrate concentration in groundwater can result from surface soil leaching. Both NO_3^--N and NO_2^--N concentrations were not statistically significant and did not improve the model fit, reflecting that developed land and population density are not directly associated with current levels of arsenic in groundwater. Most events (arsenic concentrations exceeding 10 μg/L) occur with low NO_3^--N concentrations (Figure 2d). Additionally, a negative correlation existed between arsenic and NO_3^--N concentrations (Table S2 in the Supporting Information), indicating that arsenic may be sorbed onto aquifer material in the presence of nitrate (*38*). Most NO_2^--N concentrations in groundwater (79.4%) are less than the detection limit of 0.03 mg/L, and the remaining NO_2^--N concentrations are very low (Figure 2c). Nitrite is an unstable intermediate species and is considered to be evidence of ongoing denitrification.

Compared to Fe oxides, the fraction of arsenic adsorbed by Mn oxides is considerably low (*39*). Additionally, Mn(IV) is reduced at a higher redox potential level than Fe. Some of the arsenic can be readsorbed on Fe oxyhydroxide while arsenic is mobilized from Mn oxides (*39*). Most Mn concentrations in groundwater (73.2%) are less than the detection limit (0.03 mg/L), and the remaining Mn concentrations are very low (Figure 2f). Moreover, the Mn concentrations did not correlate with current levels of arsenic in groundwater (Table S2 in the Supporting Information).

In this study, the use of a few key geochemical process parameters in the LR model can yield acceptable prediction results in assessing the probability of arsenic in groundwater, because the study area is a middle size plain (-400km^2) with a flat surface topography where the surface parameters are less important. For a large-size region where local heterogeneities are not considered, the use of surface parameters can provide reasonable prediction results (*40*). However, to obtain more refined prediction results in a large-size region, surface, geological, and anthropogenic sources of arsenic, geochemical processes, and hydrogeological and land-use parameters need to be included in the LR model (*16*).

Comparison of IK and LR. Figure 3 presents the model-based predicted probabilities for arsenic concentrations exceeding 10 μg/L in groundwater. The result is output from the LR model with model coefficients in Table 2 and is interpolated throughout the study area. The probability map reveals the spatial variability of elevated arsenic concentrations in the Lanyang plain. Three townships, Jiao-Si, Jhuang-Wei, and Yi-Lan, have the highest probabilities. Don-Shan, Tou-Chen, and Wu-Jei have moderate probabilities. The anaerobic pollutants were infiltrated from the landfill site to groundwater and transported to the downstream region, Jiao-Si, Yi-Lan, and Jhung-Wei, along with the flow direction of groundwater, creating a reducing environment, and arsenic was mobilized through reductive dissolution of Fe oxyhydroxide (*34*). Additionally, reductive dissolution of Fe oxyhydroxide coupled with mineralization of organic matter in sediment can result in arsenic mobilized in groundwater of the southern region in the study area, Don-Shan, Tou-Chen, and Wu-Jei (*23, 34*).

The rate of total correct classification for the calibration data was 86%, with 95.7% of arsenic concentrations less than 10 μg/L (specificity) and 40% of arsenic concentrations exceeding 10 μg/L (sensitivity) correctly classified by indicator kriging (see the Supporting Information). Additionally, the rate of total correct classification for the validation data is 87.2%, with 92% of arsenic concentrations less than 10 μg/L (specificity) and 60% of arsenic concentrations exceeding 10 μg/L (sensitivity).

The classification correct rates of LR were similar with those obtained by IK, but sensitivity of LR (54%) was better than that of IK. The probability of exceeding 50% located at Don-Shan, Lou-Don, Su-Ao, and Wu-Jei predicted by IK was higher than that predicted by LR (Figure 4). The primary cause is that the LR model was based on the correlations among three hydrochemical parameters and one regional variable, rather than the monovariable-spatial distribution of arsenic. The LR model efficiently predicted the occurrence probability of arsenic concentration exceeding 10 μg/L using these four parameters. Moreover, relationships between occurrence probabilities of As $\geq$ 10 μg/L and variability of these hydrochemical variable concentrations are also explored. In the future, it is possible that the provisional guidelines may become more strict in some countries while our continual improvement on the analytical precision in measuring arsenic and the As provision in the LR model might be changed accordingly.

Both LR and IK can be used to predict the probability of As concentration exceeding 10 μg/L; however, LR is applied to estimate the probability of As concentration exceeding 10 μg/L when no As is present while the LR model is being developed, and it is necessary to further investigate the relationship between As and other hydrochemical parameters. The IK can handle highly variant phenomena without having to trim off important high-value data and provide risk-qualified estimates to a specific pollutant. Additionally, the pollutant concentration can be estimated through the conditional cumulative distribution function (*ccdf*) of each site. However, the uncertainty of IK estimations is based on the spatial variation of observed data at sampled locations only, and the *ccdf* provides one measure to a single location. Thus, geostatistics is generally used to estimate spatial distribution of pollutants for unsampled sites, rather than to estimate point pollution. The LR is another applicable method to estimate the occurrence probability of pollutants at each sample without directly measuring pollutants everywhere. The parsimonious nature of the developed LR model can be used as a screening model and applied to other arsenic-prone regions where the local governments have limited funding and technology for mitigating As (*40, 41*).

Acknowledgments

The authors would like to thank The Environmental Protection Bureau of the YiLan County of the Republic of China, Taiwan, for its assistance. Professor Hann-Chyuan Chiang of National I-Lan University, is appreciated for providing the well sample data. This work is supported by the National Science Council of the Republic of China under Contract No. NSC 096-2628-B-002-021-MY3.

Supporting Information Available

Univariate statistical overview and correlation analysis of measured parameters of the data set (Tables S1 and S2), the detail validation of LR model, and indicator kriging estimation. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) Tseng, W. P. Effects and dose-response relationships of skin cancer and blackfoot disease with arsenic. *Environ. Health Perspect.* **1977**, *19*, 109–119.

(2) Lai, M. S.; Hsueh, Y. M.; Chen, C. J.; Shyu, M. P.; Chen, S. Y.; Kuo, T. L.; Wu, M. M.; Tai, T. Y. Ingested inorganic arsenic and prevalence of diabetes mellitus. *Am. J. Epidemiol.* **1994**, *139*, 484–492.

(3) Chen, C. J.; Hsueh, Y. M.; Lai, M. S.; Shyu, M. P.; Chen, S. Y.; Wu, M. M.; Kuo, T. L.; Tai, T. Y. Increased prevalence of hypertension and long-term arsenic exposure. *Hypertension* **1995**, *25*, 53–60.

(4) Wu, M. M.; Kuo, T. L.; Hwang, Y. H.; Chen, C. J. Dose-response relation between arsenic well water and mortality from cancers and vascular diseases. *Am. J. Epidemiol.* **1989**, *130*, 1123–132.

(5) Chen, C. J.; Wang, C. J. Ecological correlation between arsenic level in well water and age-adjusted mortality from malignant neoplasma. *Cancer Res.* **1990**, *50*, 5470–5474.

(6) Nickson, R. T.; McArthur, J. M.; Ravenscroft, P.; Burgess, W. G.; Ahmed, K. M. Mechanism of arsenic release to groundwater, Bangladesh and West Bengal. *Appl. Geochem.* **2000**, *15*, 403–413.

(7) Anawar, H. M.; Akai, J.; Komaki, K.; Terao, H.; Takahashi, T.; Ishzuak, T.; Safiulah, S.; Kato, K. Geochemical occurrence of arsenic in groundwater of Bangladesh: Source and mobilization process. *J. Geochem. Explor.* **2003**, *77*, 109–131.

(8) Smedley, P. L.; Kinniburgh, D. G. A review of the source, behaviour and distribution of arsenic in natural waters. *Appl. Geochem.* **2002**, *17*, 1093–1103.

(9) Smedley, P. L.; Kinniburgh, D. G.; Macdonald, D. M. J.; Nicolli, H. B.; Barros, A. J.; Tullio, J. O.; Pearce, J. M.; Alonso, M. S. Arsenic associations in sediments from the loess aquifer of La Pampa, Argentina. *Appl. Geochem.* **2005**, *20*, 989–1016.

(10) Lipfert, G.; Reeve, A. S.; Sidle, W. C.; Marvinney, R. Geochemical patterns of arsenic-enriched ground water in fracture, crystalline bedrock, Northport, Maine, USA. *Appl. Geochem.* **2006**, *21*, 528–545.

(11) Smith, J. L.; Halvorson, J. J.; Papendick, R. I. Using multiple-variable indicator kriging for evaluating soil quality. *Soil Sci. Soc. Am. J.* **1993**, *57*, 743–749.

(12) Juang, K. W.; Lee, D. Y. Simple indicator kriging for estimating the probability of incorrectly delineating hazardous areas in a contaminated site. *Environ. Sci. Technol.* **1998**, *32*, 2487–2493.

(13) van Meirvenne, M.; Goovaerts, P. Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold. *Geoderma* **2001**, *102*, 75–100.

(14) Liu, C. W.; Jang, C. S.; Liao, C. M. Evaluation of arsenic contamination potential using indicator kriging in the Yun-Lin aquifer (Taiwan). *Sci. Total Environ.* **2004**, *321*, 173–188.

(15) Nolan, B. T.; Hitt, K. J.; Ruddy, B. C. Probability of nitrate contamination of recently recharged groundwater in the conterminous United States. *Environ. Sci. Technol.* **2002**, *36*, 2138–2145.

(16) Ayotte, J. D.; Nolan, B. T.; Nuckols, J. R.; Cantor, K. P.; Robinson, G. R., Jr.; Baris, D.; Hayes, L.; Karagas, M.; Bress, W.; Silverman, D. T.; Lubin, J. H. Modeling the probability of arsenic in groundwater in New England as a tool for exposure assessment. *Environ. Sci. Technol.* **2006**, *40*, 3578–3585.

(17) Aelion, C. M.; Conte, B. C. Susceptibility of residential wells to VOC and nitrate contamination. *Environ. Sci. Technol.* **2004**, *38*, 1648–1653.

(18) Worrall, F. A molecular topology approach to predicting pesticide pollution of groundwater. *Environ. Sci. Technol.* **2001**, *35*, 2282–2287.

(19) Peng, T. R. *Environmental Isotopic Study (^{13}C, ^{18}O, ^{14}C, D, T) on Meteroric Water and Groundwater in I-Lan Area*. Doctoral Dissertation, Institute of Geology, National Taiwan University, Taiwan, 1995 (in Chinese).

(20) Chen, W. S. *Analysis of Sediments and Sedimentary Environments in Stratigraphic Correlation of the Lanyang Plain*. Taiwan Central Geological Survey Report, 2000 (in Chinese).

(21) Lee, J. Y. *The Geochemical Analysis of Sediments*. Taiwan Central Geological Survey Report, 2000 (in Chinese).

(22) Chiou, H. Y.; Chiou, S. T.; Hsu, Y. H.; Chou, Y. L.; Tseng, C. H.; Wei, M. L.; Chen, C. J. Incidence of transitional cell carcinoma and arsenic in drinking water: A follow-up study of 8120 residents in an arseniasis-endemic area in northeastern Taiwan. *Am. J. Epidemiol.* **2001**, *153*, 411–418.

(23) Chen, I. J. *Geochemical Characteristics of Porewater and Sediments from Chung-Hsing, Wu-Jie and Long-De of I-Lan Plain, Taiwan*. Master Thesis, Institute of Geology, National Taiwan University, Taiwan, 2001.

(24) Chiou, H. Y.; Huang, W. I.; Su, C. L. Dose-response relationship between prevalence of cerebrovascular disease and ingested inorganic arsenic. *Stroke* **1997**, *28*, 1717–1723.

(25) Chen, S. L.; Yeh, S. J.; Lin, T. H. Trace element concentration and arsenic speciation in the well water of a Taiwan area with endemic blackfoot disease. *Biol. Trace Elem. Res.* **1995**, *48* (3), 263–274.

(26) Taiwan FACOA. *Fisheries Year Book of Taiwan (ROC)*. Fisheries Agency, Council of Agriculture, Taiwan (ROC), 2004 (in Chinese).

(27) YiLan EPB, Taiwan. *Survey of Arsenic Contents of Drinking Water (Surface Water and Groundwater) in YiLan County*. Environmental Protection Bureau of YiLan County, Taiwan (ROC), 1997−1999 Report.

(28) Gaus, I.; Kinniburgh, D. G.; Talbot, J. C.; Webster, R. Geostatistical analysis of arsenic concentration in groundwater in Bangladesh using disjunctive kriging. *Environ. Geol.* **2003**, *44*, 939–948.

(29) Hosmer, D. W.; Lemeshow, S. *Applied Logistic Regression*, 2nd ed.; John Wiley and Sons: New York, 2000.

(30) SAS Institute. *SAS/STAT User' Guide, Version 8*. SAS Institute Inc: Cary, NC, 1999.

(31) Menard, S. *Applied Logistic Regression Analysis*, 2nd ed., Sage Publication: Thousand Oaks, CA, 2002.

(32) Dixit, S.; Hering, J. G. Comparison of arsenic(V) and arsenic(III) sorption onto iron oxides minerals: Implications for arsenic mobility. *Environ. Sci. Technol.* **2003**, *37*, 4182–4189.

(33) Peters, S. C.; Blum, J. D.; Klaue, B.; Karagas, M. R. Arsenic occurrence in New Hampshire drinking water. *Environ. Sci. Technol.* **1999**, *33*, 1328–1333.

(34) Lee, J. J.; Jang, C. S.; Wang, S. W.; Liang, C. P.; Liu, C. W. Delineation of spatial redox zones using discriminant analysis and geochemical modeling in arsenic-affected alluvial aquifers. *Hydrol. Proces.* **2007**, DOI: 10.1002/hyp.6884.

(35) Smith, R. L.; Howes, B. L.; Duff, J. H. 1991. Denitrification in a sand and gravel aquifer. *Appl. Environ. Microbiol.* **1991**, *54*, 1071–1078.

(36) Stüben, D.; Berner, Z.; Chandrasekharam, D.; Karmakar, J. Arsenic enrichment in groundwater of West Bengal, India: geochemical evidence for mobilization of As under reducing conditions. *Appl. Geochem.* **2003**, *18*, 1417–1434.

(37) Pierce, M. L.; Moore, C. B. Adsorption of arsenite and arsenate on amorphous iron hydroxides. *Water Res.* **1982**, *16*, 1247–1253.

(38) Oremland, R. S.; Stoltz, J. F. Arsenic, microbes, and contaminated aquifers. *Trends Microbiol.* **2005**, *13* (2), 45–49.

(39) Norra, S.; Berner, Z. A.; Agarwala, P.; Wagner, F.; Chandrasekharam, D.; Stüben, D. Impact of irrigation with As rich groundwater on soil and crops: A geochemical case study in West Bengal Delta Plain, India. *Appl. Geochem.* **2005**, *20*, 1890–1906.

(40) Winkel, L.; Berg, M.; Amini, M.; Hug, S. J.; Johnson, C. A. Predicting groundwater arsenic contamination in Southeast Asia from surface parameters. *Nat. Geosci.* **2008**, *1*, 536–542.

(41) Berg, M.; Stengel, C.; Trang, P. T. K.; Viet, P. H.; Sampson, M. L.; Leng, M.; Samreth, S.; Fredericks, D. Magnitude of arsenic pollution in Mekong and Red River Deltas−Cambodia and Vietnam. *Sci. Total Environ.* **2007**, *372*, 413–425.

ES900540S