Optimal estimates and uncertainty of low-permeability soil fractions in a complicated aquifer system

Shih-Kai Chen[1], Cheng-Shin Jang [2, *], and Ya-Hsuan Wang[1]

[1] Department of Civil Engineering, National Taipei University of Technology, Taipei City, Taiwan
[2] Department of Leisure and Recreation Management, Kainan University, Luzhu, Taoyuan City, Taiwan

ABSTRACT

Spatial distributions of low-permeability soils, such as clay, silt, and mud, are important for establishing hydrogeological maps and can affect groundwater flow and recharge and contaminant transport in soil solutes. This study adopted indicator kriging (IK) to spatially characterize the optimal estimates and uncertainty of the low-permeability soil fraction in the Choushui River alluvial fan aquifers in Taiwan. First, IK was used to analyze the occurrence probabilities of low-permeability soil fractions according to several thresholds and to establish the conditional cumulative distribution function (CCDF) using a linear interpolation model. Then, median estimates and E-type estimates of the low-permeability soil fractions in aquifers were determined based on the CCDF. Finally, the integration of the conditional variance and interquartile range was employed to assess the local uncertainty of IK estimates. The analysis results indicated that the median estimates were more reliable than the E-type estimates and capable of modeling the absence of low-permeability soils in aquifers. Moreover, high estimated low-permeability soil fractions frequently gave rise to the high levels of local uncertainty. The study results are useful for modifying hydrogeological system maps and establishing numerical simulations of groundwater models.

1. INTRODUCTION

Establishing a hydrogeological system requires detailed spatial information on the types and distributions of subsurface soils. Generally, gravels and sands with high permeability and excellent groundwater storage are regarded as aquifers, whereas clay, silt, and mud with low permeability and poor groundwater storage serve as aquitards (Gerber et al. 2001; Liu et al. 2006; Cheng and Chen 2007). Low-permeability soils in aquifers, such as clay, silt, and mud, restrict groundwater flow and recharge (Marinoni 2003) and strongly affect contaminant transport in soil solutes (Johnson et al. 1989). Before numerical groundwater models, such as MODFLOW (Harbaugh and McDonald 1996), MT3D (Zheng 1996), and 3DFEMWATER (Yeh 1987), are used to simulate groundwater flow and pollutant transport, the spatial distribution of low-permeability soils must be comprehensively established (Ting et al. 1998;

Jang et al. 2016a). In addition, owing to the interactions between marine and continental sediments, many alluvial sedimentary environments have complicated hydrogeological patterns (Pham and Tsai 2017), and many thin layers of low-permeability soils frequently appear in an aquifer. A low-permeability soil fraction in an aquifer can be adopted to characterize soil permeability and water movement (Delbari et al. 2011; Shahriari et al. 2019) and explore the hydrogeological system of complicated sedimentary environments. Therefore, determining a sound spatial distribution of the low-permeability soil fraction in an aquifer is critical for exploring hydrogeological characteristics and groundwater resource management.

Distributions of low-permeability soils in aquifers typically exhibit high levels of spatial variability due to soil heterogeneity. However, borehole data on soil textures are typically limited, resulting that determining the spatial distribution of low-permeability soils in aquifers is rather difficult according to finite observations. Geostatistics is a common spatial analysis tool that conducts the spatial

---

* Corresponding author
  E-mail: csjang@mail.knu.edu.tw

distributions of observations with uncertainty (Isaaks and Srivastava 1989). Ordinary kriging (OK) is a method of linearly weighted moving averages based on neighboring observations. However, OK, which typically overestimates low observations and underestimates high observations, gives rise to unreal soil layer distributions owing to decreasing spatial variability when OK is used to model soil distributions (Marinoni 2003). A distribution-free geostatistical technique, indicator kriging (IK), is lightly impacted by outliers and produces a robust estimate through the binary transformation of observations. At an unobserved site, the estimated values using IK denote the occurrence probability of variables that do not surpass a certain cutoff (Yeh et al. 2020). The kriging technique has been frequently applied to spatially estimate the thickness or fractions of different soil textures (D'Or et al. 2001; Marinoni 2003; Delbari et al. 2011; Chen et al. 2013; Jang et al. 2013, 2016b; Park and Jang 2014) and the presence or absence of certain soil types (Bastante et al. 2008; Jang et al. 2013; Pham and Tsai 2017). Although having an advantage in estimates of extreme values, the IK technique could cause the loss of information when a spatial variable was transformed into an indicator value according to a specified cutoff (Goovaerts 1997). Jang et al. (2016b) used a combined ordinary-indicator kriging approach with multiple thresholds to determine the spatial distributions and thicknesses of low permeability topsoil materials in the Choushui River alluvial fan, Taiwan, according to maximal probability density functions among different categories. However, Jang et al. (2016b) only focused on low permeability topsoil materials and did not conduct low-permeability soil fractions of all aquifers. Additionally, kriging estimates at unknown sites typically have local uncertainty. The assessment of local uncertainty is important for evaluating risks involved in decision-making processes. Two uncertainty measurement methods-conditional variance (CV) and interquartile range (IQR), which were similar to the kriging variance obtained by OK, can be used to model the uncertainty of IK estimates (Goovaerts 2001; Goovaerts et al. 2005). Accordingly, this study used uncertainty measurement methods based on IK estimates to propagate the uncertainty of the estimated low-permeability soil fraction and to facilitate establishing a correct hydrogeological map and conducting calibration procedures for groundwater model simulations. To compare with Jang et al. (2016b), the novelty of this study is the combination of kriging estimates and uncertainty measurements for the low-permeability soil fractions in aquifers, which is useful for establishing hydrogeological maps and groundwater model simulations.

This study adopted IK to spatially model the optimal estimates and uncertainty of the low-permeability soil fraction in the aquifers of the Choushui River alluvial fan in Taiwan. First, IK was used to quantify the occurrence probabilities of the low-permeability soil fraction according to several thresholds and establish the conditional cumulative distribution function (CCDF) using a linear interpolation model. Then, median estimates and E-type estimates of low-permeability soil fractions in aquifers were determined based on the CCDF. Finally, the integration of CV and IQR was employed to assess the local uncertainty of the IK estimates.

2. MATERIALS AND METHODS

2.1 Hydrogeology of the Study Area

The Choushui River alluvial fan, which covers an area of about 2300 km², is encircled by the Wu River to the north, the Peikang River to the south, the Taiwan Strait to the west, and the Dulliu Hill and Baguah Mountain to the east (Fig. 1). The average annual precipitation recorded between 2010 and 2019 was approximately 1615 mm in this area (Taiwan Central Weather Bureau 2020). This alluvial fan is mainly composed of Quaternary unconsolidated deposits with abundant groundwater. Seventy-two geological boreholes were established at 300 m depths between 1992 and 1998, and the types and depths of subsurface soils were recorded. The hydrogeology in this alluvial fan was generally divided into proximal-, mid-, and distal-fan areas. Because of the interactions among marine and continental deposits, eight overlapping sequences, including four aquifers and four aquitards, were located in the distal- and mid-fan areas (Taiwan CGS 1999). The four aquifers are named "aquifer 1", "aquifer 2", "aquifer 3", and "aquifer 4" from top to bottom in the distal- and mid-fan regions (Fig. 2) (Liu et al. 2006). The eastern proximal-fan areas are an unconfined aquifer for natural groundwater recharge. According to the delineation of the hydrogeological system proposed by the Taiwan CGS (1999), the distal- and mid-fan aquifers frequently exhibit a markedly large fraction of low-permeability soils, indicating a complicated hydrogeological pattern.

2.2 Data on the Low-Permeability Soil Fractions in Aquifers

This study used data on soil textures in 72 boreholes and the hydrogeological system map of the Choushui River alluvial fan reported by the Taiwan CGS (1999). Because this study focused on quantifying low-permeability soils in aquifers, the summation of the length of clay, silt, and mud in an aquifer was divided by the thickness of the aquifer in a borehole to obtain the low-permeability soil fractions. Figure 3 shows the spatial distribution of the low-permeability soil fractions in aquifers. A high low-permeability soil fraction frequently occurred in the distal- and mid-fan aquifers, while a low low-permeability soil fraction was present in the proximal-fan aquifers. The low-permeability soil fractions in aquifers were examined using a Kolmogorov-Smirnov normality test, indicating that they were not

normally distributed (p < 0.001). Thus, this study adopted a distribution-free geostatistical technique, IK, to characterize the low-permeability soil fractions in aquifers. Table 1 lists the statistics regarding the low-permeability soil fractions in aquifers. The low-permeability soil fractions were larger in aquifer 1 than in the remaining aquifers. The soil fractions had greater skewness in aquifers 3 and 4 than in aquifers 1 and 2. Table 2 reports proportions of observed fractions of low-permeability soil in aquifers for various classes. The

classes of 0.1 - 0.3, 0.1 - 0.2, 0 - 0.2, and 0 were predominant in aquifers 1, 2, 3, and 4, respectively, for the proportions of low-permeability soil fractions.

2.3 Indicator-Based Geostatistical Technique

2.3.1 Variogram Analysis

An experimental variogram, $\gamma(\mathbf{h})$, can quantify the spatial variability of variables between two sites separated by a
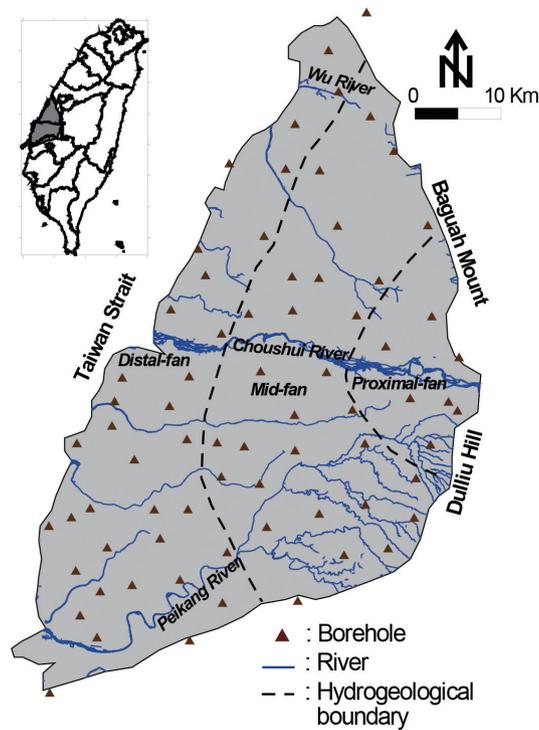
Fig. 1. Study area.
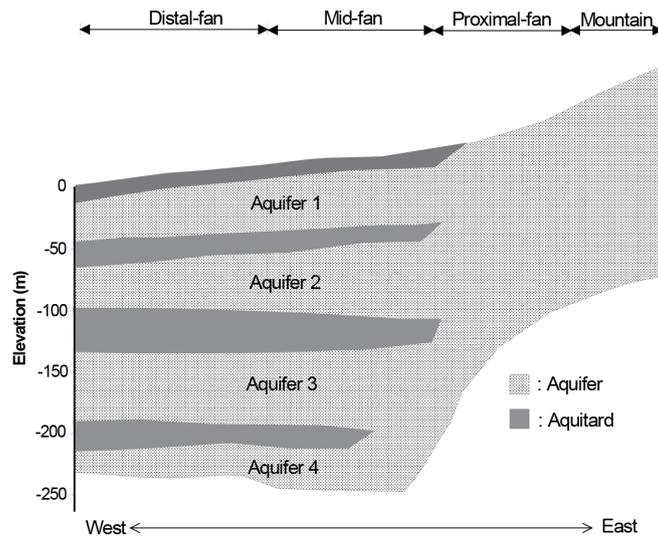
Fig. 2. Conceptual hydrogeological profile.
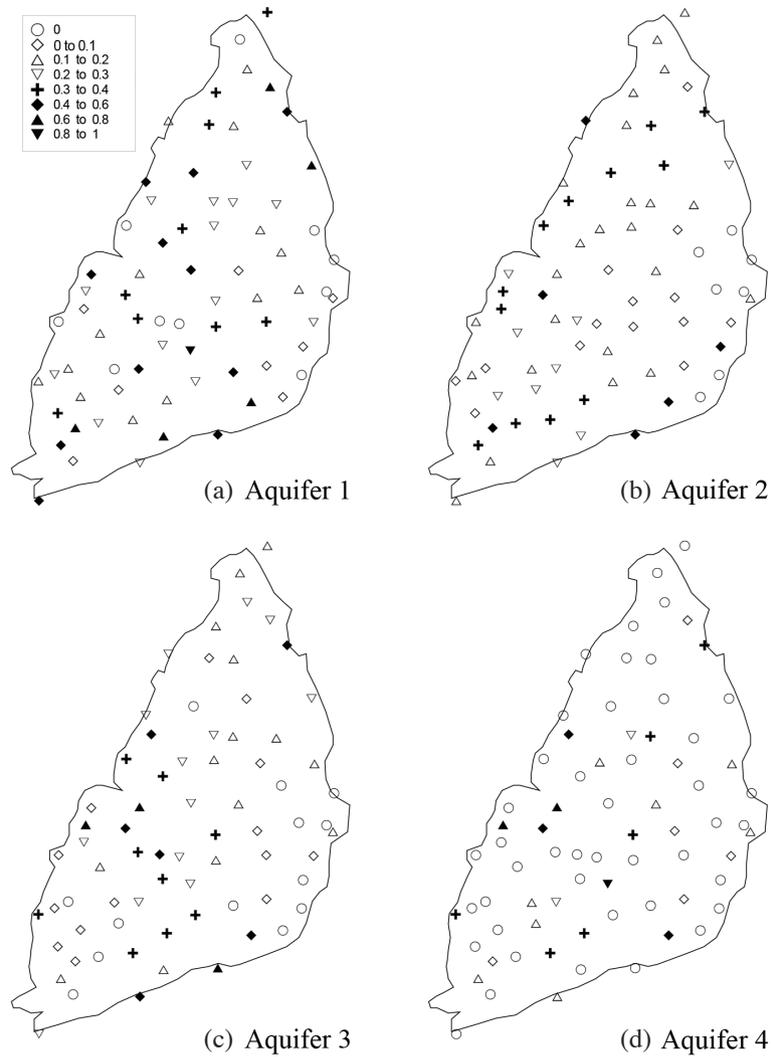
Fig. 3. Observed fractions of low-permeability soils in aquifers.

Table 1. Statistics regarding low-permeability soil fractions in aquifers.

Statistics		Aquifer 1	Aquifer 2	Aquifer 3	Aquifer 4
Average		0.260	0.193	0.200	0.110
Standard deviation		0.203	0.141	0.179	0.192
Skewness		0.795	0.694	1.098	2.085
Maximum		0.865	0.571	0.780	0.865
Zero-fraction percentile		13.8th	9.7th	17.9th	60.3rd
Percentiles	10th	0	0.031	0	0
	20th	0.069	0.074	0.033	0
	30th	0.128	0.105	0.076	0
	40th	0.175	0.131	0.124	0
	50th	0.241	0.158	0.164	0
	60th	0.283	0.189	0.220	0
	70th	0.334	0.262	0.286	0.122
	80th	0.429	0.331	0.311	0.221
	90th	0.564	0.383	0.438	0.389

Table 2. Proportions (%) of observed fractions of low-permeability soil in aquifers for different classes.

Fraction classes	Aquifer 1	Aquifer 2	Aquifer 3	Aquifer 4
0	13.9	9.7	18.1	61.1
0 - 0.1	11.1	19.4	19.4	8.3
0.1 - 0.2	19.4	33.3	19.4	11.1
0.2 - 0.3	19.4	12.5	18.1	2.8
0.3 - 0.4	12.5	16.7	12.5	8.3
0.4 - 0.5	11.1	5.6	5.6	1.4
0.5 - 0.6	4.2	2.8	2.8	2.8
0.6 - 0.7	4.2	0.0	1.4	1.4
> 0.7	4.2	0.0	2.8	2.8

vector $\mathbf{h}$ (Isaaks and Srivastava 1989).

$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \left\{ \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{u}_i + \mathbf{h}) - z(\mathbf{u}_i)]^2 \right\} \qquad (1)$$

where $N(\mathbf{h})$ is the number of pairs, $\mathbf{u}_i$ is the vector of location coordinates, $z(\mathbf{u}_i)$ and $z(\mathbf{u}_i + \mathbf{h})$ are the spatial variables, and $\mathbf{h}$ is the lag distance. The experimental variogram is then fitted by a theoretical model, such as spherical, exponential, or Gaussian models with the nugget effect (c_0), sill (c), and range (a). The least-squares method with the smallest fitting error is used to determine an optimal theoretical variogram model. The mathematical formulas of the spherical, exponential, and Gaussian models are expressed as follows (Isaaks and Srivastava 1989).

Spherical model:
$$\begin{cases} \gamma(\mathbf{h}) = c_0 + c\left[1.5\left(\frac{\mathbf{h}}{a}\right) - 0.5\left(\frac{\mathbf{h}}{a}\right)^3\right], \ \mathbf{h} \leq a \\ \gamma(\mathbf{h}) = c_0 + c, \ \mathbf{h} > a \end{cases} \qquad (2)$$

Exponential model:
$$\gamma(\mathbf{h}) = c_0 + c\left[1 - \exp\left(-\frac{3\mathbf{h}}{a}\right)\right] \qquad (3)$$

Gaussian model:
$$\gamma(\mathbf{h}) = c_0 + c\left\{1 - \exp\left[-\left(\frac{3\mathbf{h}}{a}\right)^2\right]\right\} \qquad (4)$$

2.3.2 Indicator Kriging

IK is a distribution-free geostatistical technique, and the estimate signifies the occurrence probability not exceeding a given cutoff z_k at a certain site $\mathbf{u}$ (Goovaerts 1997). Before the IK algorithm, the spatial variable $Z(\mathbf{u})$ is first transformed into a binary response (i.e., indicator variable), which is defined as follows (Goovaerts 1997).

$$I(\mathbf{u}; z_k) = \begin{cases} 1, & \text{if } Z(\mathbf{u}) \leq z_k, \ k = 1, 2, ..., m \\ 0, & \text{otherwise} \end{cases} \qquad (5)$$

The expected value of $I(\mathbf{u}; z_k)$ subject to n neighboring data can be written as follows (Goovaerts 1997).

$$E[I(\mathbf{u}; z_k \,|\, (n))] = Prob\{Z(\mathbf{u}) \leq z_k \,|\, (n)\} = F(\mathbf{u}; z_k \,|\, (n)) \qquad (6)$$

where $F[\mathbf{u}; z_k/(n)]$ is the CCDF for $Z(\mathbf{u})$ not exceeding z_k. This study employed the GAMV and IK3D codes in GSLIB (Deutsch and Journel 1997) for analyzing the experimental variogram and IK, respectively.

2.3.3 Optimal Estimates

Goovaerts (1997) suggested several methods of optimal estimates for IK when a unique estimate was selected from the range of the corresponding z-values in the CCDF. This study adopted the median estimate and E-type estimate to determine low-permeability soil fractions in aquifers. The median estimate, $z_{ME}^*(\mathbf{u})$, is the corresponding fraction of the 0.5 probability in the CCDF at a site $\mathbf{u}$ and is expressed as follows (Goovaerts 1997):

$$z_{ME}^*(\mathbf{u}) = F^{-1}(\mathbf{u}; 0.5 \,|\, (n)) \qquad (7)$$

The E-type estimate, $z_{EE}^*(\mathbf{u})$, is the expected value of the CCDF at a site $\mathbf{u}$ and is defined as (Goovaerts 1997)

$$z_{EE}^*(\mathbf{u}) = \sum_{k=1}^{K+1} \frac{(z_{k-1} + z_k)}{2} \cdot [F(\mathbf{u}; z_k \,|\, (n)) - F(\mathbf{u}; z_{k-1} \,|\, (n))] \qquad (8)$$

2.3.4 Measure of Local Uncertainty

The methods of CV and IQR are typically used to

measure the local uncertainty of IK estimates (Goovaerts 1997). CV, $\sigma^2_{CV}(\mathbf{u})$, quantifies the spread of the conditional probability density function (CPDF) around the expected value, $z^*_{EE}(\mathbf{u})$. $\sigma^2_{CV}(\mathbf{u})$ is defined as follows (Goovaerts 1997).

$$\sigma^2_{CV}(\mathbf{u}) = \sum_{k=1}^{K+1} [\bar{z}_k - z^*_{EE}(\mathbf{u})]^2 \cdot [F(\mathbf{u}; z_k | (n)) - F(\mathbf{u}; z_{k-1} | (n))] \quad (9)$$

where the K cutoffs are adopted to classify spatial variables, $\bar{z}_k$ is the average of the class (z_{k-1}, z_k), and $F[\mathbf{u}; z_k|(n)]$ and $F[\mathbf{u}; z_{k-1}|(n)]$ are the CCDF at the classes z_k and z_{k-1}, respectively. To simplify the computation of $z^*_{EE}(\mathbf{u})$, a linear model is adopted to interpolate the CCDFs. The IQR($\mathbf{u}$) is the difference between the reverse estimates of the CCDF at the probabilities of 0.75 and 0.25 and can be expressed as follows (Goovaerts 1997).

$$IQR(\mathbf{u}) = F^{-1}(\mathbf{u}; 0.75 | (n)) - F^{-1}(\mathbf{u}; 0.25 | (n)) \quad (10)$$

where $F^{-1}[\mathbf{u}; 0.75|(n)]$ and $F^{-1}[\mathbf{u}; 0.25|(n)]$ are the reverse estimates of the CCDF at the probabilities of 0.75 and 0.25, respectively.

2.3.5 Cross-Validation

In a leave-one-out cross-validation procedure, each observation is first removed and then reestimated from the remaining observations using OK (Isaaks and Srivastava 1989). This study used the kriging mean error (KME) and mean square standard error (MSSE) to evaluate the suitability of the determined spatial variogram structure and the performance of the kriging estimator. The two parameters are defined as (Chilès and Delfiner 1999; Deutsch 2002)

$$KME = \frac{1}{n} \sum_{i=1}^{n} [I^*(\mathbf{u}_i) - I(\mathbf{u}_i)] \quad (11)$$

$$MSSE = \frac{1}{n} \sum_{i=1}^{n} \frac{[I^*(\mathbf{u}_i) - I(\mathbf{u}_i)]^2}{\sigma^2_k(\mathbf{u}_i)} \quad (12)$$

where I*($\mathbf{u}_i$) and I($\mathbf{u}_i$) are the estimated and observed values of the indicator variables, respectively, at the ith site; n is the number of observations; and $\sigma_k(\mathbf{u}_i)$ and $\sigma^2_k(\mathbf{u}_i)$ are the kriging standard deviation and variance, respectively, at the ith site. The KME and MSSE, which are close to nil and unity, respectively, indicate that the fitting model and parameters of the determined theoretical variograms are excellent, and the kriging estimates are robust. In addition, the MSSE can be accepted by the range of $\pm 3(2/n)^{0.5}$ (i.e., the range of 0.5 - 1.5 for 72 borehole data) (Chilès and Delfiner 1999).

3. RESULTS

3.1 Variogram Analyses of the Low-Permeability Soil Fractions

This study considered the zero fraction of low-permeability soils as the first threshold and adopted the 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 90th percentiles of the observed low-permeability soil fractions in each aquifer to establish other thresholds (shown in Table 1) when the corresponding fractions of the percentiles were not zero. However, the 10th and 20th percentiles of the low-permeability soil fractions in aquifers 2 and 3, respectively, were considerably close to the zero-fraction percentile. Thus, the thresholds of the 10th and 20th percentiles in aquifers 2 and 3, respectively, were ignored. The numbers of the thresholds in aquifers 1, 2, 3, and 4 were 9, 9, 8, and 4, respectively. In addition, Goovaerts (1997) suggested that a reasonable number of thresholds ranges from 5 to 15 for establishing the CCDF using IK. In this study, the number of the thresholds was within a reasonable range in aquifers 1, 2, and 3 and below the reasonable range in aquifer 4 due to the limitation of observations.

This study employed a 5000 m lag to analyze the experimental omni-variogram for low-permeability soil fractions (Fig. 4). The number of pairs ranged from 98 to 339 for the experimental variograms. The exponential and spherical theoretical models provided the optimal fit for the variograms in aquifers 1 and 4 and aquifers 2 and 3, respectively. The fitted nugget effect, sill, and range were respectively 0.04 - 0.15, 0.05 - 0.11, and 20000 - 27000 m in aquifer 1, 0.02 - 0.11, 0.06 - 0.16, and 15000 - 45000 m in aquifer 2, 0.05 - 0.12, 0.06 - 0.21, and 22000 - 35000 m in aquifer 3, and 0.04 - 0.15, 0.06 - 0.11, and 20000 - 23000 m in aquifer 4 (Table 3). The nugget effect-to-sill proportions (c_0/c) of ≤ 0.25, 0.25 - 0.75, and > 0.75 signified strong, moderate, and weak spatial correlations (Cambardella et al. 1994), respectively. The c_0/c values ranged from 0.63 to 1.44 in aquifer 1, 0.18 to 1.11 in aquifer 2, 0.33 to 1.09 in aquifer 3, and 0.67 to 1.50 in aquifer 4 (Table 3). The analysis results revealed that moderate and weak spatial correlations were present for most indicator variables of the low-permeability soil fractions.

3.2 CCDF Establishment of the Low-Permeability Soil Fractions

This study utilized IK to establish the CCDF for low-permeability soil fractions in aquifers based on various thresholds. A grid of 55 × 81 cells with a 1000 m spacing was assigned to each aquifer. However, 2225 active cells were found within each aquifer of this study area.

Figure 5 displays the CCDFs at various thresholds in the four aquifers. Generally, a higher probability in the CCDF represents a lower fraction of low-permeability soils. The high probabilities in the CCDF mainly occurred in the

Fig. 4. The experimental and fitted theoretical variograms of indicator variables for the low-permeability soil fractions in aquifers.

Fig. 4. (Continued)

3rd threshold, aquifer 3

4th threshold, aquifer 3

5th threshold, aquifer 3

6th threshold, aquifer 3

7th threshold, aquifer 3

8th threshold, aquifer 3

1st threshold, aquifer 4

2nd threshold, aquifer 4

3rd threshold, aquifer 4

4th threshold, aquifer 4
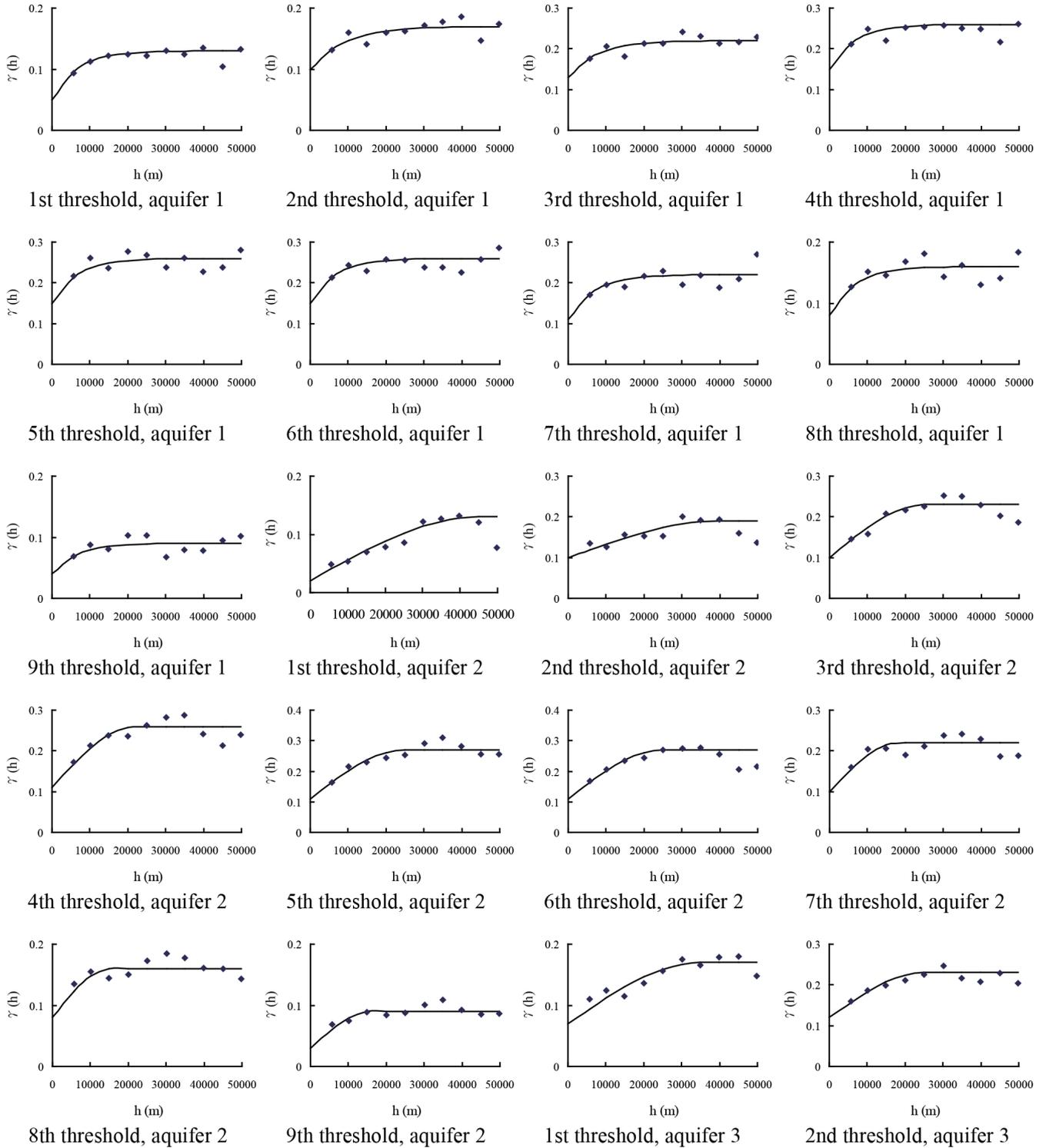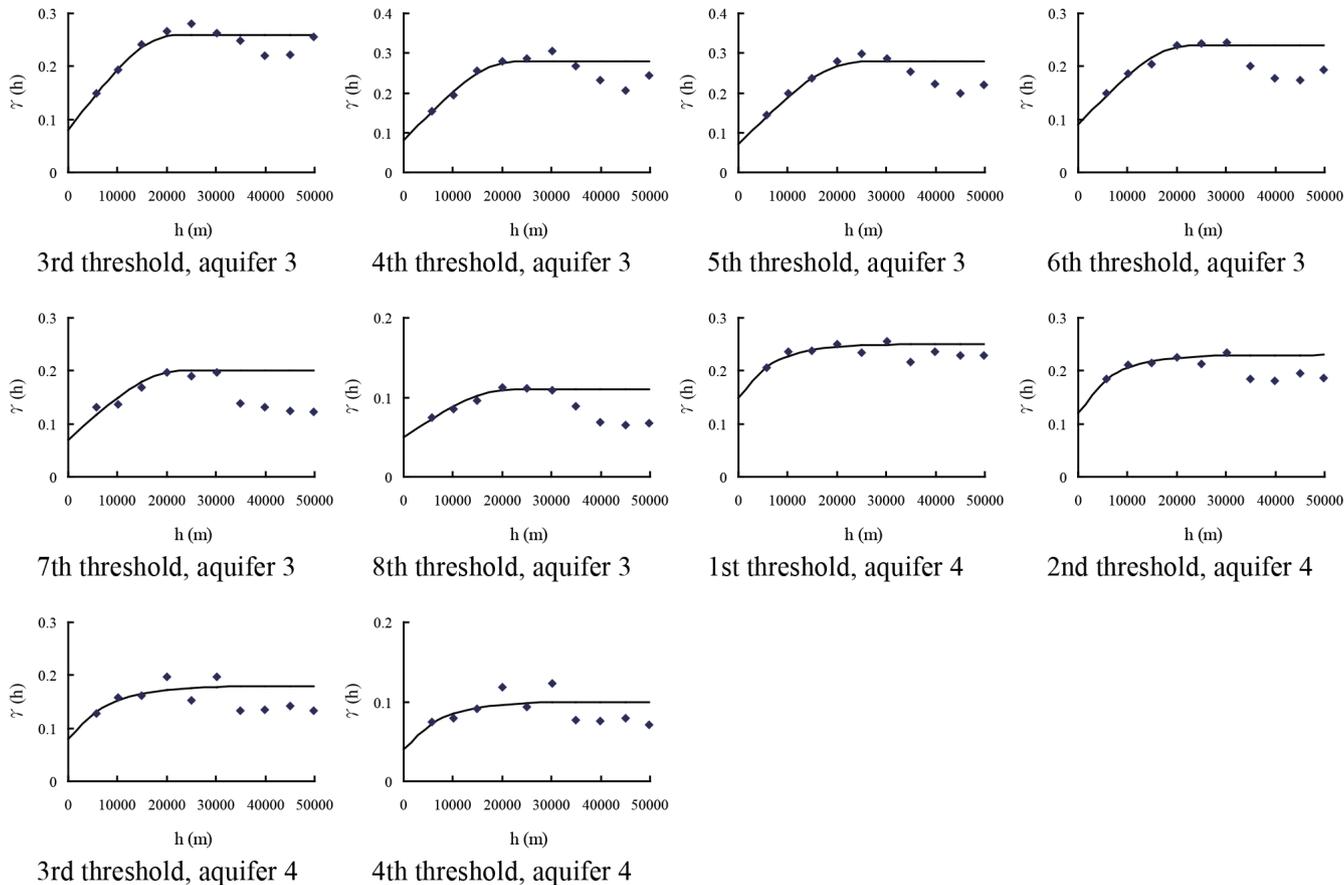
Table 3. Fitted theoretical variograms of indicator variables for the low-permeability soil fractions in aquifers and the cross-validation parameters.

Aquifers	Thresholds (percentiles)	Fractions	Model	c_0	c	a (m)	c_0/c	KME	MSSE
1	1st (13.8th)	0.000	E[a]	0.05	0.08	20000	0.63	-0.015	1.207
	2nd (20th)	0.069	E	0.10	0.07	27000	1.43	-0.015	0.911
	3rd (30th)	0.128	E	0.13	0.09	23000	1.44	-0.025	1.004
	4th (40th)	0.175	E	0.15	0.11	20000	1.36	-0.004	1.022
	5th (50th)	0.241	E	0.15	0.11	20000	1.36	-0.013	1.126
	6th (60th)	0.283	E	0.15	0.11	20000	1.36	0.005	1.170
	7th (70th)	0.334	E	0.11	0.11	20000	1.00	0.008	1.174
	8th (80th)	0.429	E	0.08	0.08	20000	1.00	0.023	1.323
	9th (90th)	0.564	E	0.04	0.05	20000	0.80	0.006	1.415
2	1st (9.7th)	0.000	S[b]	0.02	0.11	45000	0.18	0.002	1.118
	2nd (20th)	0.074	S	0.10	0.09	40000	1.11	0.006	0.911
	3rd (30th)	0.105	S	0.10	0.13	25000	0.77	0.003	0.973
	4th (40th)	0.131	S	0.11	0.15	22000	0.73	0.007	1.159
	5th (50th)	0.158	S	0.11	0.16	25000	0.69	0.012	1.211
	6th (60th)	0.189	S	0.11	0.16	25000	0.69	0.003	1.114
	7th (70th)	0.262	S	0.10	0.12	18000	0.83	-0.011	1.075
	8th (80th)	0.331	S	0.08	0.08	15000	1.00	0.000	1.148
	9th (90th)	0.383	S	0.03	0.06	16000	0.50	0.006	1.357

Note: [a]E: exponential model; [b]S: spherical model.

Table 3. (Continued)

Aquifers	Thresholds (percentiles)	Fractions	Model	c_0	c	a (m)	c_0/c	KME	MSSE
3	1st (17.9th)	0.000	S	0.07	0.10	35000	0.70	0.007	1.351
	2nd (30th)	0.076	S	0.12	0.11	25000	1.09	0.020	1.182
	3rd (40th)	0.124	S	0.08	0.18	22000	0.44	0.024	1.197
	4th (50th)	0.164	S	0.08	0.20	23000	0.40	0.019	1.113
	5th (60th)	0.220	S	0.07	0.21	25000	0.33	0.017	1.173
	6th (70th)	0.286	S	0.09	0.15	23000	0.60	0.015	1.131
	7th (80th)	0.311	S	0.07	0.13	23000	0.54	0.012	1.257
	8th (90th)	0.438	S	0.05	0.06	22000	0.83	0.006	1.152
4	1st (60.3rd)	0.000	E	0.15	0.10	20000	1.50	-0.006	1.353
	2nd (70th)	0.122	E	0.12	0.11	20000	1.09	0.000	1.338
	3rd (80th)	0.221	E	0.08	0.10	23000	0.80	-0.005	1.313
	4st (90th)	0.389	E	0.04	0.06	22000	0.67	-0.001	1.304

Fig. 5. Estimated CCDF at different thresholds in aquifers.

1st threshold, aquifer 1　　2nd threshold, aquifer 1　　3rd threshold, aquifer 1　　4th threshold, aquifer 1

5th threshold, aquifer 1　　6th threshold, aquifer 1　　7th threshold, aquifer 1　　8th threshold, aquifer 1

9th threshold, aquifer 1　　1st threshold, aquifer 2　　2nd threshold, aquifer 2　　3rd threshold, aquifer 2

4th threshold, aquifer 2　　5th threshold, aquifer 2　　6th threshold, aquifer 2　　7th threshold, aquifer 2

Fig. 5. (Continued)

8th threshold, aquifer 2   9th threshold, aquifer 2   1st threshold, aquifer 3   2nd threshold, aquifer 3

3rd threshold, aquifer 3   4th threshold, aquifer 3   5th threshold, aquifer 3   6th threshold, aquifer 3

7th threshold, aquifer 3   8th threshold, aquifer 3   1st threshold, aquifer 4   2nd threshold, aquifer 4

3rd threshold, aquifer 4   4th threshold, aquifer 4

proximal-fan region and the southwestern and southeastern parts of aquifer 1, the proximal- and mid-fan regions and the northern, partially southwestern coastal and southeastern parts of aquifer 2, and the proximal-fan region and the partially northern, southwestern coastal, and southeastern parts of aquifers 3 and 4. In addition, the cross-validation results indicated that the KME and MSSE ranged from -0.025 to 0.024 and from 0.911 to 1.415 (Table 3), respectively, demonstrating that the fitting theoretical variograms and parameters are suitable and the IK estimates are excellent.

3.3 Optimal Estimates for Low-Permeability Soil Fractions

This study used a linear interpolation model within thresholds and the upper bound, unity, to establish CCDFs.

The corresponding fractions of the 0.5 probability in the CCDF were obtained for each cell for median estimates. Figure 6 shows the median estimates for the low-permeability soil fractions in the four aquifers. To compute an expected value, the CPDFs within classes at each cell were determined in advance. The E-type estimate was the summation of the average of a class multiplying the CPDF according to Eq. (8). Figure 7 displays the E-type estimates for the low-permeability soil fractions in the four aquifers. The E-type estimates smoothed the spatial variability, while the median estimates exhibited a higher level of spatial heterogeneity than the E-type estimates. According to the median estimates, the high low-permeability soil fractions primarily occurred in the mid-fan, southwestern, and northern regions of aquifer 1, the western distal-fan and northern and southern mid-fan regions of aquifers 2 and 3, and the western
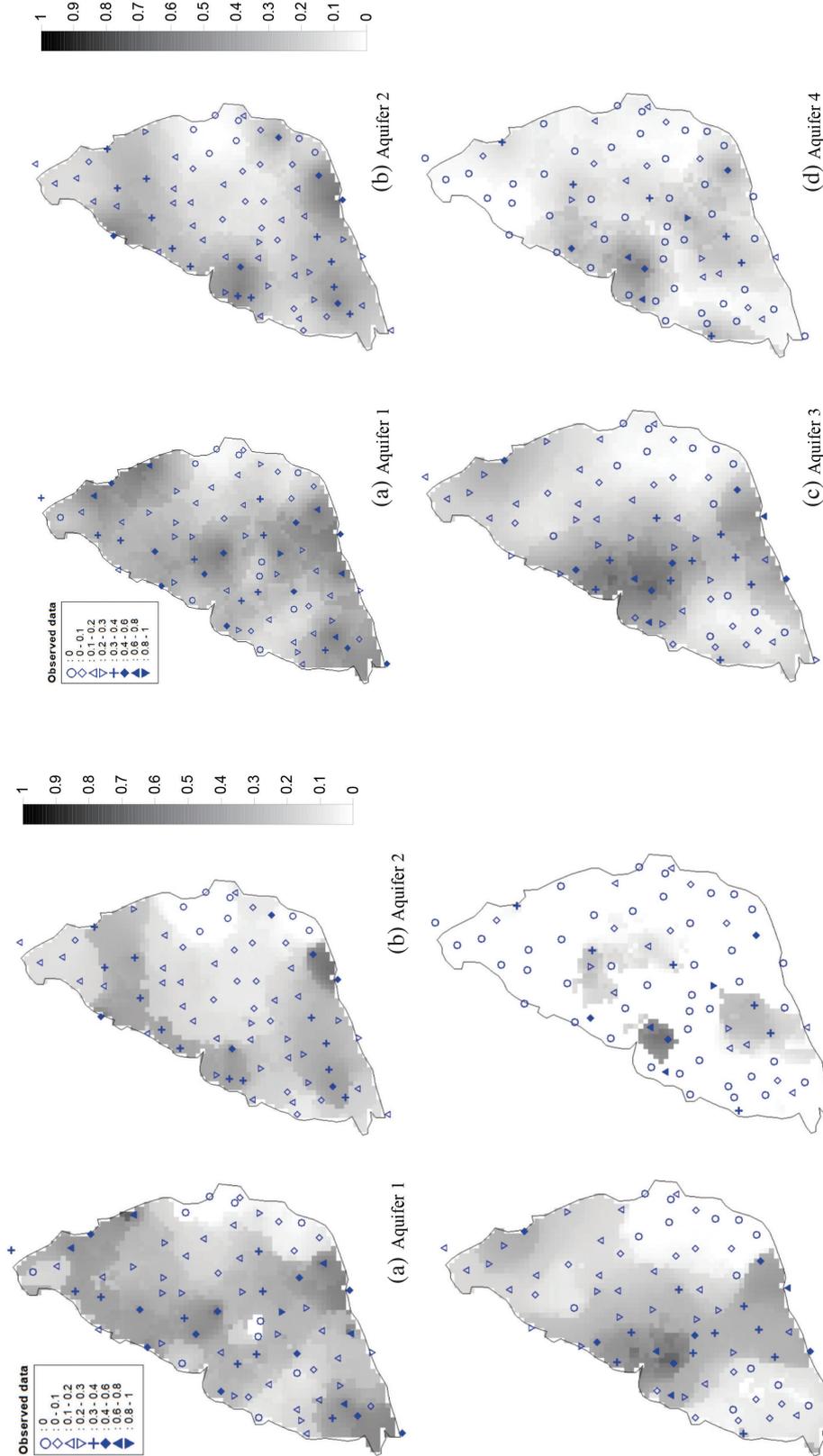
Fig. 7. E-type estimates for the low-permeability soil fractions in the four aquifers.

Fig. 6. Median estimates for low-permeability soil fractions in the four aquifers.

distal-fan regions of aquifer 4. In addition, Table 4 reports the areal proportions of different fraction classes for the median estimates and E-type estimates. The fraction classes of 0.2 - 0.3, 0.1 - 0.2, 0.1 - 0.3, and 0 were dominant in aquifers 1, 2, 3, and 4, respectively, according to the median estimates. The fraction classes of 0.2 - 0.3, 0.1 - 0.2, 0.1 - 0.2, and 0 - 0.1 prevailed in aquifers 1, 2, 3, and 4, respectively, according to the E-type estimates.

3.4 Local Uncertainty of the IK Estimates

This study adopted CV and IQR to measure the local uncertainty of IK estimates conditional to the information available. Five levels of CV and IQR distributions were employed to characterize the uncertainty. The first level, which was below the 20th percentile of the CV and IQR distributions, was very low; the second level, which was the 20th to 40th percentiles of the CV and IQR distributions, was low; the third level, which was the 40th to 60th percentiles of the CV and IQR distributions, was medium; the fourth level, which was the 60th to 80th percentiles of the CV and IQR distributions, was high; and the fifth level, which was above the 80th percentile of the CV and IQR distributions, was very high.

Figures 8 and 9 display the levels of local uncertainty using CV and IQR, respectively. To acquire an integrated result, this study adopted a maximum level of local uncertainty from CV and IQR for each cell (Chen et al. 2019). Figure 10 shows the five levels of integrated local uncertainty. The high and very high levels of local uncertainty were mainly distributed in the mid- and distal-fan aquifers with high estimated low-permeability soil fractions.

4. DISCUSSION

This study considered the zero fraction of low-permeability soils in aquifers as the first threshold. The median estimates and E-type estimates based on CCDF were used to spatially model low-permeability soil fractions in aquifers. Although the absence of low-permeability soils commonly occurred in the aquifers, particularly in aquifer 4, the E-type estimates could not characterize the zero fraction of low-permeability soils in aquifers (see Tables 2 and 4), resulting in incorrect estimated results. Thus, the median estimates were capable of determining the absence of low-permeability soils in aquifers and effectively reduced the smooth estimate effect. However, median estimates easily underestimated sparse high values surrounding numerous low values, particularly in aquifer 4. Because the presence or absence of low-permeability soils in aquifers strongly affects groundwater flow and recharge (Marinoni 2003; Jang et al. 2013, 2016b), the framework of the median estimates based on IK with multiple thresholds can facilitate establishing the spatial characteristics of low-permeability soil fractions in aquifers for simulation models of groundwater flow and pollutant transport. In addition, although the upper tail means of CCDF for CV were more sensitive than those for IQR (Goovaerts 1997), the within-class linear interpolation model used to establish CCDF reduced the difference between CV and IQR, resulting that the spatial patterns of CV were similar to those of IQR (Figs. 8 and 9). Moreover, high local uncertainty typically occurred in high values for CV and IQR.

Aquifers generally have a relatively low low-permeability soil fraction. However, the presence of high low-permeability soil fractions in aquifers indicates a complicated hydrogeological system, easily resulting in incorrect hydrogeological delineation, particularly for high levels of local uncertainty. The Taiwan CGS can modify the hydrogeological system map in the Choushui River alluvial fan based on the research results. Furthermore, low-permeability soils affect the velocity and direction of groundwater flow (Liu et al. 2004) and are a barrier to groundwater flow and contaminant transport in soil solutes (Marinoni 2003; Jang and Liu 2005). Correctly modeling the spatial distribution and uncertainty of low-permeability soils

Table 4. Areal percentages (%) of estimated fraction classes for the median estimates and E-type estimates. (a) Median estimates; (b) E-type estimates.

(a)

Fraction class	Aquifer 1	Aquifer 2	Aquifer 3	Aquifer 4
0	2.8	7.4	8.8	76.0
0 - 0.1	7.4	17.1	23.1	9.9
0.1 - 0.2	20.0	38.1	26.2	7.6
0.2 - 0.3	36.1	20.0	26.2	4.6
0.3 - 0.4	25.6	15.4	10.9	0.1
0.4 - 0.5	7.7	1.1	3.8	1.4
0.5 - 0.6	0.3	0.9	1.0	0.4
0.6 - 0.7	0.1	0.0	0.0	0.0

(b)

Fraction class	Aquifer 1	Aquifer 2	Aquifer 3	Aquifer 4
0	0.0	0.0	0.0	2.8
0 - 0.1	4.8	13.6	23.3	48.7
0.1 - 0.2	18.6	35.9	33.5	31.8
0.2 - 0.3	32.4	32.9	21.1	13.3
0.3 - 0.4	29.7	12.4	13.4	2.7
0.4 - 0.5	14.1	4.8	5.9	0.7
0.5 - 0.6	0.4	0.4	2.8	0.0
0.6 - 0.7	0.0	0.0	0.0	0.0

Fig. 9. Spatial distributions of IQR in the four aquifers based on five levels of local uncertainty.
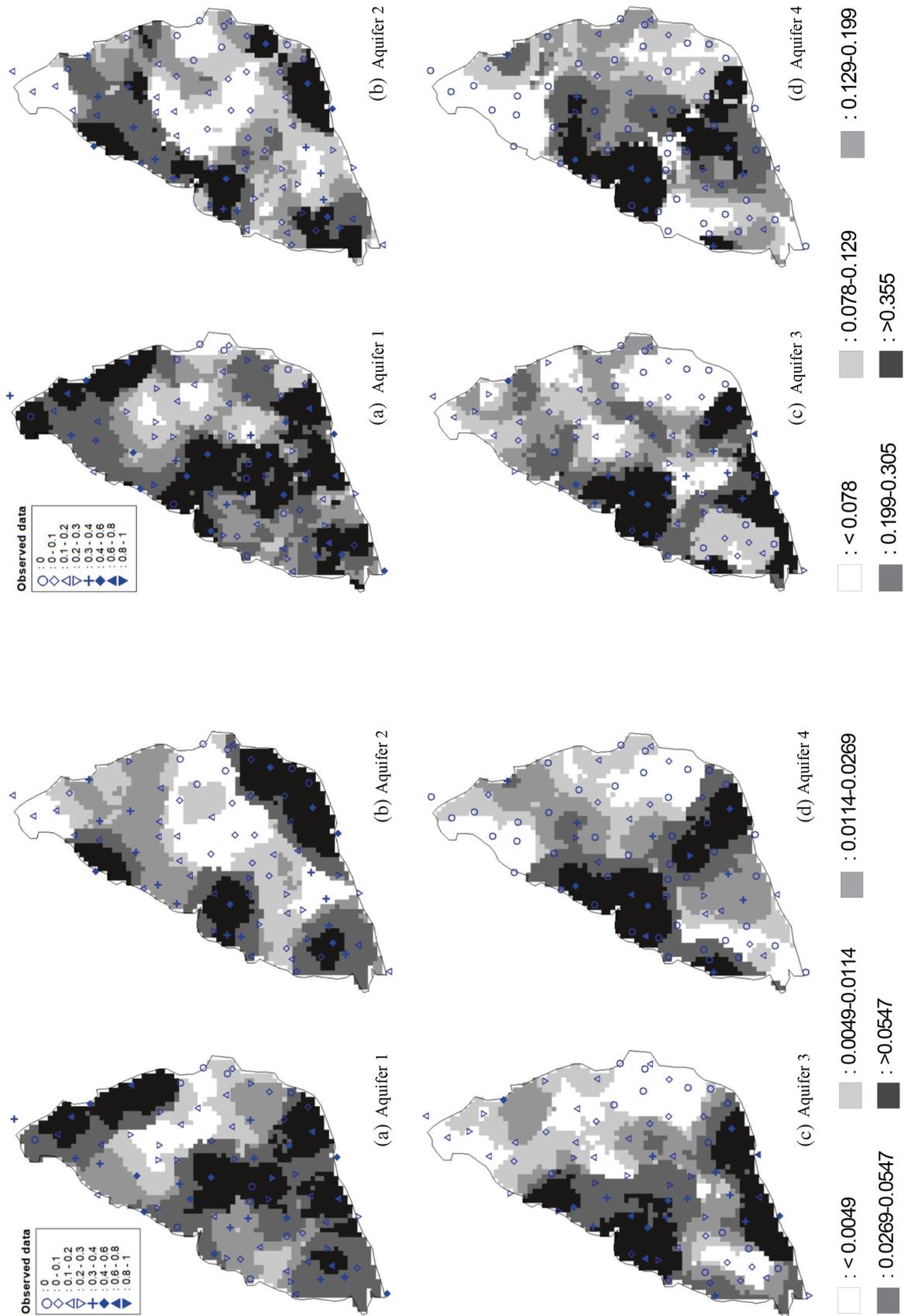
Fig. 8. Spatial distributions of CV in the four aquifers based on five levels of local uncertainty.
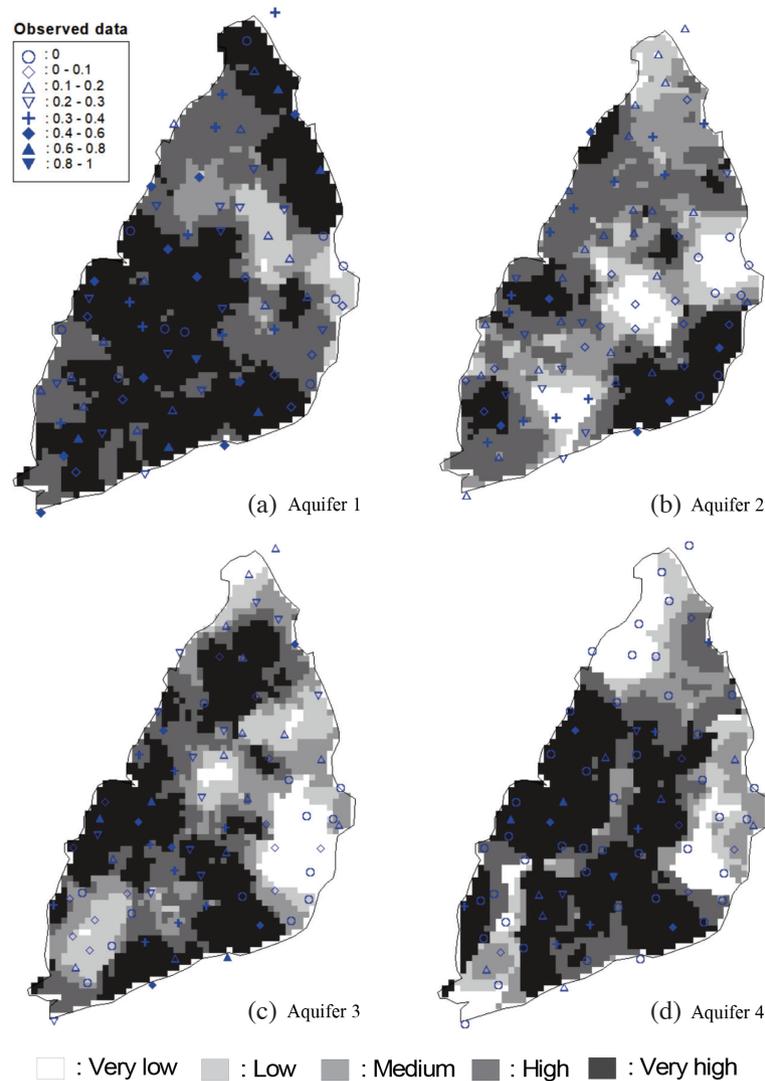
Fig. 10. Spatial distributions of integrated local uncertainty in the four aquifers based on five levels.

in aquifers can facilitate the establishment of numerical simulation models of groundwater flow and contamination transport. Moreover, when high low-permeability soil fractions in aquifers have high levels of local uncertainty at certain sites, researchers should pay more attention to conducting calibration procedures for groundwater model simulations.

5. CONCLUSION

This study adopted IK to determine the optimal estimates and uncertainty of the low-permeability soil fractions in the aquifers of the Choushui River alluvial fan. The median estimates and E-type estimates based on multi-threshold IK gave us an excellent insight into modelled spatial distributions and uncertainty of the low-permeability soil fractions in complicated aquifer systems. The analysis results revealed that the median estimates were more reliable than

the E-type estimates and capable of characterizing the absence of low-permeability soils in aquifers. The high levels of local uncertainty were principally situated in the mid- and distal-fan aquifers with high estimated low-permeability soil fractions. To avoid incorrect hydrogeological delineation, researchers should pay more attention to problems associated with high low-permeability soil fractions in aquifers with high levels of local uncertainty. Furthermore, the study results can provide government administrators with detailed information on modifying the hydrogeological system map in the Choushui River alluvial fan and groundwater researchers with a reference for numerical simulation models of groundwater flow and contaminant transport.

REFERENCES

Bastante, F. G., C. Ordóñez, J. Taboada, and J. M. Matías, 2008: Comparison of indicator kriging, conditional indicator simulation and multiple-point statistics used to model slate deposits. *Eng. Geol.*, **98**, 50-59, doi: 10.1016/j.enggeo.2008.01.006. [Link]

Cambardella, C. A., T. B. Moorman, J. M. Novak, T. B. Parkin, D. L. Karlen, R. F. Turco, and A. E. Konopka, 1994: Field-scale variability of soil properties in central Iowa soils. *Soil Sci. Soc. Am. J.*, **58**, 1501-1511, doi: 10.2136/sssaj1994.03615995005800050033x. [Link]

Chen, S.-K., C.-S. Jang, and Y.-H. Peng, 2013: Developing a probability-based model of aquifer vulnerability in an agricultural region. *J. Hydrol.*, **486**, 494-504, doi: 10.1016/j.jhydrol.2013.02.019. [Link]

Chen, S.-K., C.-S. Jang, and C.-Y. Chou, 2019: Assessment of spatiotemporal variations in river water quality for sustainable environmental and recreational management in the highly urbanized Danshui River basin. *Environ. Monit. Assess.*, **191**, doi: 10.1007/s10661-019-7246-1. [Link]

Cheng, C. and X. Chen, 2007: Evaluation of methods for determination of hydraulic properties in an aquifer–aquitard system hydrologically connected to a river. *Hydrogeol. J.*, **15**, 669-678, doi: 10.1007/s10040-006-0135-z. [Link]

Chilès, J. P. and P. Delfiner, 1999: Geostatistics: Modeling Spatial Uncertainty, John Wiley & Sons Inc., New York, 283-287.

Delbari, M., P. Afrasiab, and W. Loiskandl, 2011: Geostatistical analysis of soil texture fractions on the field scale. *Soil Water Res.*, **6**, 173-189, doi: 10.17221/9/2010-SWR. [Link]

Deutsch, C. V., 2002: Geostatistical Reservoir Modeling, Oxford University Press, New York, 124-152.

Deutsch, C. V. and A. G. Journel, 1997: GSLIB: Geostatistical Software Library and User's Guide, 2nd Edition, Oxford University Press, New York, 369 pp.

D'Or, D., P. Bogaert, and G. Christakos, 2001: Application of the BME approach to soil texture mapping. *Stoch. Environ. Res. Risk Assess.*, **15**, 87-100, doi: 10.1007/s004770000057. [Link]

Gerber, R. E., J. I. Boyce, and K. W. Howard, 2001: Evaluation of heterogeneity and field-scale groundwater flow regime in a leaky till aquitard. *Hydrogeol. J.*, **9**, 60-78, doi: 10.1007/s100400000115. [Link]

Goovaerts, P., 1997: Geostatistics for Natural Resources Evaluation, Oxford University Press, New York, 259-368.

Goovaerts, P., 2001: Geostatistical modelling of uncertainty in soil science. *Geoderma*, **103**, 3-26, doi: 10.1016/S0016-7061(01)00067-2. [Link]

Goovaerts, P., G. AvRuskin, J. Meliker, M. Slotnick, G.

Jacquez, and J. Nriagu, 2005: Geostatistical modeling of the spatial variability of arsenic in groundwater of southeast Michigan. *Water Resour. Res.*, **41**, doi: 10.1029/2004WR003705. [Link]

Harbaugh, A. W. and M. G. McDonald, 1996: User's documentation for MODFLOW-96, an update to the U.S. Geological Survey modular finite-difference groundwater flow model. Open-File Report 96-485, Branch of Information Services, U.S. Geological Survey, USA, 56 pp, doi: 10.3133/ofr96485. [Link]

Isaaks, E. H. and R. M. Srivastava, 1989: An Introduction to Applied Geostatistics, Oxford University Press, New York, 278-322.

Jang, C.-S. and C.-W. Liu, 2005: Contamination potential of nitrogen compounds in the heterogeneous aquifers of the Choushui River alluvial fan, Taiwan. *J. Contam. Hydrol.*, **79**, 135-155, doi: 10.1016/j.jconhyd.2005.07.001. [Link]

Jang, C.-S., S.-K. Chen, and Y.-M. Kuo, 2013: Applying indicator-based geostatistical approaches to determine potential zones of groundwater recharge based on borehole data. *CATENA*, **101**, 178-187, doi: 10.1016/j.catena.2012.09.003. [Link]

Jang, C.-S., C.-F. Chen, C.-P. Liang, and J.-S. Chen, 2016a: Combining groundwater quality analysis and a numerical flow simulation for spatially establishing utilization strategies for groundwater and surface water in the Pingtung Plain. *J. Hydrol.*, **533**, 541-556, doi: 10.1016/j.jhydrol.2015.12.023. [Link]

Jang, C.-S., S.-K. Chen, and Y.-T. Cheng, 2016b: Spatial estimation of the thickness of low permeability topsoil materials by using a combined ordinary-indicator kriging approach with multiple thresholds. *Eng. Geol.*, **207**, 56-65, doi: 10.1016/j.enggeo.2016.04.008. [Link]

Johnson, R. L., J. A. Cherry, and J. F. Pankow, 1989: Diffusive contaminant transport in natural clay: a field example and implications for clay-lined waste disposal sites. *Environ. Sci. Technol.*, **23**, 340-349, doi: 10.1021/es00180a012. [Link]

Liu, C.-W., S.-K. Chen, and C.-S. Jang, 2004: Modelling water infiltration in cracked paddy field soil. *Hydrol. Process.*, **18**, 2503-2513, doi: 10.1002/hyp.1478. [Link]

Liu, C.-W., S.-W. Wang, C.-S. Jang, and K.-H. Lin, 2006: Occurrence of arsenic in ground water in the Choushui river alluvial fan, Taiwan. *J. Environ. Qual.*, **35**, 68-75, doi: 10.2134/jeq2005.0129. [Link]

Marinoni, O., 2003: Improving geological models using a combined ordinary–indicator kriging approach. *Eng. Geol.*, **69**, 37-45, doi: 10.1016/S0013-7952(02)00246-6. [Link]

Park, N.-W. and D.-H. Jang, 2014: Comparison of geostatistical kriging algorithms for intertidal surface sediment facies mapping with grain size data. *Sci. World*

J., **2014**, 145824, doi: 10.1155/2014/145824. [Link]

Pham, H. V. and F. T.-C. Tsai, 2017: Modeling complex aquifer systems: a case study in Baton Rouge, Louisiana (USA). *Hydrogeol. J*., **25**, 601-615, doi: 10.1007/s10040-016-1532-6. [Link]

Shahriari, M., M. Delbari, P. Afrasiab, and M. R. Pahlavan-Rad, 2019: Predicting regional spatial distribution of soil texture in floodplains using remote sensing data: A case of southeastern Iran. *CATENA*, **182**, 104149, doi: 10.1016/j.catena.2019.104149. [Link]

Taiwan Central Geological Survey (Taiwan CGS), 1999: Project of Groundwater Monitoring Network in Taiwan during First Stage: Research Report of Choushui River Alluvial Fan, Water Resources Bureau, Taiwan, 383 pp.

Taiwan Central Weather Bureau, 2020: Internet Application of Climate Data, Central Weather Bureau, Ministry of Transportation & Communications, Executive Yuan, Taiwan. Available at http://e-service.cwb.gov.tw/wdps/index_net.jsp.

Ting, C.-S., Y. Zhou, J. J. de Vries, and I. Simmers, 1998: Development of a preliminary ground water flow model for water resources management in the Pingtung Plain, Taiwan. *Ground Water*, **36**, 20-36, doi: 10.1111/j.1745-6584.1998.tb01062.x. [Link]

Yeh, G. T., 1987: 3DFEMWATER: A three-dimensional finite element model of water flow through saturated-unsaturated media. No. ORNL-6386, Oak Ridge National Lab., TN, USA.

Yeh, H.-C., Y.-C. Chen, and C. Wei, 2020: Mapping dust storm $PM_{2.5}$ pollution risk using indicator kriging in northern Taiwan. *Terr. Atmos. Ocean. Sci*., **31**, 313-323, doi: 10.3319/TAO.2019.11.07.01. [Link]

Zheng, C., 1996: MT3D96 Version DoD_1.5: A modular three-dimensional transport model. The Hydrogeology Group, University of Alabama, USA.