Engineering Geology 207 (2016) 56–65

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo

Spatial estimation of the thickness of low permeability topsoil materials by using a combined ordinary-indicator kriging approach with multiple thresholds

Cheng-Shin Jang [a], Shih-Kai Chen [b,*], Ya-Ting Cheng [b]

[a] Department of Leisure and Recreation Management, Kainan University, Luzhu, Taoyuan 338, Taiwan
[b] Department of Civil Engineering, National Taipei University of Technology, Taipei 106, Taiwan

ARTICLE INFO

Article history:
Received 21 October 2015
Received in revised form 7 April 2016
Accepted 10 April 2016
Available online 16 April 2016

Keywords:
Groundwater recharge
Permeability
Ordinary–indicator kriging
Probability
Topsoil

ABSTRACT

Accurately estimating a spatial pattern of low permeability topsoil materials, such as clay, silt, and mud, is critical because these materials are a natural barrier to groundwater recharge and flow in subsurface engineering applications. This study determined the spatial distributions and thicknesses of low permeability topsoil materials in the Choushui River alluvial fan, Taiwan by using a combined ordinary–indicator kriging (OIK) approach with multiple thresholds. The thicknesses were first estimated using ordinary kriging (OK). To reduce the overestimation of low values and underestimation of high values, indicator kriging (IK) was then adopted to probabilistically categorize the thickness of the materials. The maximum occurrence probability among the categories was selected for determining the most suitable thickness category of low permeability topsoil materials. Finally, the thicknesses of the materials estimated using OK were amended according to the upper and lower limits of the most suitable category determined using multi-threshold IK. The research results reveal that the combined OIK approach with multiple thresholds can reduce the overestimation of low values and underestimation of high values for various topsoil thicknesses and characterize a reliable hydrogeological pattern in topsoil. The approach facilitates evaluating groundwater recharge and constructing a numerical model of coupled surface water and groundwater flow.

1. Introduction

Groundwater is one of the most fundamental water resources in Taiwan and is typically extracted for various purposes, such as drinking, irrigation, and aquaculture (Jang et al., 2012, 2013a). However, groundwater overexploitation frequently occurs in many Taiwan aquifers, resulting in a substantial drop in groundwater levels, seawater intrusion, and land subsidence (Ting et al., 1998). Regarding the safe yield of an aquifer, reasonable and sustainable use of groundwater depends on aquifer recharge originating from precipitation infiltration (Zhou, 2009). Several natural factors, such as low permeability subsoil materials, can affect the capacity for aquifer recharge (Jang et al., 2013b; Evans and Hartemink, 2014). In general, vertical infiltration of surface water into the subsoil is more difficult when a higher amount of low permeability soil materials, such as clay, silt, and mud, are present in groundwater surfaces. The presence of thicker low permeability soil materials leads to slower percolation of soil water into groundwater surfaces. Furthermore, before numerical models of coupled surface water and groundwater flow, such as WASH123D (Yeh et al., 2006) and SWR (Hughes et al., 2012), are used to quantify the recharge effect for a specified hydrological event (Huang and Yeh, 2012; Hughes et al., 2014), a correct hydrogeological setting must be determined. The distributions and thicknesses of low permeability topsoil materials are critical for groundwater recharge because they play a major role in determining whether surface water (stream-river network and overland regime) infiltrates subsurface media or forms runoff. Consequently, to conserve groundwater resources, the distribution of low permeability topsoil materials must be accurately determined for assessing and managing regional water resources.

Distributions of low permeability topsoil materials vary spatially. However, only limited data can be surveyed in situ because of time and cost constraints. Determining a spatial distribution by using limited survey data is commonly uncertain. Geostatistics is a spatial interpolation technique that models the spatial variability and distributions of surveyed data with uncertainty. Ordinary kriging (OK) is a weighted-average estimator that is based on neighboring observation data. However, OK frequently overestimates low values and underestimates high values, resulting in smoothing the spatial variability of variables (Marinoni, 2003). A nonparametric geostatistical approach, indicator kriging (IK), relies on no specified distributions of spatial variables and, compared with OK, yields an estimate that is less sensitive to outliers through binary transformation of data. At an unsampled location, the values estimated using IK signify the occurrence probability of spatial variables not exceeding a certain threshold. The kriging technique

* Corresponding author.
E-mail address: chensk@ntut.edu.tw (S.-K. Chen).

http://dx.doi.org/10.1016/j.enggeo.2016.04.008

has been frequently applied to spatially estimate the depth or thickness of soil materials (Bourennane et al., 2000; Marinoni, 2003; Penížek and Borůvka, 2006; Kuriakose et al., 2009) and characterize soil texture types (Pohlmann et al., 2000; Guadagnini et al., 2002; Trevisani and Fabbri, 2010; Jang et al., 2013b; Ryu et al., 2013). In addition, estimates obtained using a combined ordinary–indicator kriging (OIK) approach proposed by Marinoni (2003) substantially reduced smoothing effects on zero-thickness areas with low hydraulic conductivity (K) compared with estimates obtained using OK. Furthermore, Marinoni emphasized that accurately estimating a zero thickness is critical for geological and geotechnical applications because the areas with low K can be regarded as a natural barrier to groundwater flow. However, Marinoni only used a single estimation threshold to conduct the zero-thickness problem of low K soil. The single estimation threshold in IK could give rise to incorrect estimates and loss of information (Goovaerts, 1997). Therefore, this study attempted to extend the method of Marinoni (2003) and adopted multiple thresholds in the OIK approach to comprehensively categorize various thicknesses of low permeability topsoil materials.

This study determined the spatial distributions and thicknesses of low permeability topsoil materials in the Choushui River alluvial fan, Taiwan by using the combined OIK approach with multiple thresholds. First, the thicknesses of low permeability topsoil materials were analyzed using OK. To reduce the overestimation of low values and underestimation of high values, IK was then adopted to probabilistically categorize these thicknesses. A maximum occurrence probability among the categories was selected for determining the most suitable category of the thickness of low permeability topsoil materials. Finally, the thickness of low permeability topsoil materials estimated using OK was amended according to the upper and lower limits of the most suitable category determined using multi-threshold IK.

2. Material and methods

2.1. Hydrogeology of the study area

The Choushui River alluvial fan is located in western Taiwan and is enclosed by the Wu River to the north, Peikang River to the south, Taiwan Strait to the west, and Dulliu Hill and Baguah Mountain to the east (Fig. 1). This alluvial fan occupies an area of approximately 2500 km^2, and the hydrogeological condition can principally be partitioned into proximal-, mid-, and distal-fan zones. The Quaternary unconsolidated sediment underlying this alluvial fan is rich in groundwater. The eastern proximal-fan and foothills are a major source of natural groundwater recharge. The mean annual precipitation was approximately 1590 mm in the study area according to climate data for 2000–2014 (Taiwan Central Weather Bureau, 2014).

Subsurface stratigraphic analyses of a depth of approximately 300 m were conducted between 1992 and 1998 to examine the hydrogeological features of the alluvial fan, including 72 boreholes. Because the focus was on evaluating the capacity of topsoil infiltration, this study considered soil materials only in shallow aquifers. In general, low permeability topsoil materials, such as clay, silt, and mud, can serve as a natural barrier to groundwater recharge and flow (Taiwan Central Geological Survey (CGS), 1999). Low permeability topsoil materials were primarily present in the distal- and mid-fan regions, but the amount of these materials in the proximal-fan region was low.

2.2. Data preparation for thicknesses of low permeability topsoil materials

This study used data on soil materials in 72 boreholes reported by the Taiwan CGS (1999) and considered the soil materials above groundwater levels in shallow aquifers as topsoil materials. The thickness of low permeability topsoil materials was calculated according to elevation. A thickness in a borehole of zero indicated the absence of low permeability topsoil materials. Fig. 2 shows the observed distribution and histogram of the thicknesses of low permeability topsoil materials in the boreholes. Low permeability topsoil materials were thick in the distal- and mid-fan zones. Particularly, the thickness of low permeability topsoil materials of more than 19.58 m occurred in the western foothills of the Dulliu Hill and Baguah Mountain and the Peikang River mouth because of marked river sedimentation. Zero thicknesses of low permeability topsoil materials were present in the coastal, southern, and proximal-fan regions. Table 1 lists statistics on the thickness of these materials, which ranged from 0 to 27.6 m, with the average being 7.48 m and standard deviation being 7.19 m. Approximately 21.12% of the low permeability topsoil materials were absent in the boreholes.

2.3. Geostatistical approaches

2.3.1. Variogram analysis

A variogram can characterize the spatial variability of random variables between two locations. In practice, an experimental variogram, $\gamma(\mathbf{h})$, is used to compute pairs of data separated by a vector, $\mathbf{h}$ (Isaaks and Srivastava, 1989).

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

where $N(\mathbf{h})$ is the number of data pairs, $z(\mathbf{u}_i)$ represents the random variable, and $\mathbf{h}$ denotes the lag distance. The experimental variogram is then fitted by a theoretical model, which is a spherical, exponential, or Gaussian model, with the nugget effect (c_0), sill (c), and range (a). When the variogram value in a theoretical model levels out at a certain lag distance, the beginning distance is the range (Fig. 3). The nugget effect represents the variogram value greater than zero at zero lag distance and is attributed to measurement errors or spatial sources of variation at distances smaller than the sampling interval. When the variogram model reaches the range, the variogram value minus the nugget effect is called the sill (Isaaks and Srivastava, 1989). The functions of the spherical, exponential, and Gaussian models are defined as follows (Isaaks and Srivastava, 1989).

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

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

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

The shape in the neighborhood of the origin is the primary difference among the three models. The spherical, exponential, and Gaussian variogram models typically exhibit linear, fast, and slow increases, respectively, in a short lag distance. This study adopted a least-squares method for fitting the variogram model and the parameters to obtain a best-fit model with the lowest fitting errors. Variograms can be calculated in different directions to gauge the anisotropy of spatial variability. This anisotropic model generally includes geometric and zonal anisotropies (Deutsch, 2002). The geometric anisotropy occurs when the range varies with the direction of the variogram for the constant sill. The zonal anisotropy occurs when both the range and sill vary with the direction of the variogram. This study considered only the geometric anisotropy.

2.3.2. Ordinary kriging

OK is the most fundamental geostatistical method for modeling spatial distributions of random variables. Moreover, OK is a linear weighted