ORIGINAL ARTICLE

Variogram-based approach for selection of grid size in groundwater modeling

Zargham Mohammadi

Abstract In this paper a geostatistical approach based on the variogram concept is used for decisions about the size of a grid network applicable to a groundwater model. One of the important properties of the variogram function is the range of influence, which is interpreted as a measure of similarity and correlation distance between spatial phenomena. Taking the concept of range into account, several available spatial variables of the under study aquifer were used to plot the variogram in different directions. The study area is an unconfined aquifer (Boushkan Plain in southwest Iran) with an average thickness of 50 m and area of 100 km^2. The variables used for computing the variogram include aquifer thickness (D), groundwater pumping rate (Q) and electrical resistivity (R) of the aquifer material. The range of influence was estimated to be 2,000, 2,500, and 2,000 m for D, Q and R, respectively. Comparisons of statistical parameters of spatial variables over a grid size ranging from 1,000 to 5,000 m were done to confirm the proposed grid size according to variogram plots. The bounded area of each cell could be considered as homogenous media according to the spatial variation of the variables used. The optimum size of the grid network was selected according to the minimum of the variance and coefficient of variation over each cell size. Results suggest a 2,500 × 2,500 m grid size for the modeling process in the studied aquifer. The results emphasize the role of variogram function in selection of grid size for the groundwater model.

Z. Mohammadi (✉)
Department of Earth Sciences, Shiraz University,
71454 Shiraz, Iran
e-mail: zmohammadi@shirazu.ac.ir

Introduction

Modeling groundwater system is a common practice for understanding and management of an aquifer. Generally, the modeling process is initiated by the development of a conceptual model followed by the development of a mathematical model. Discretization of the groundwater system by grid network helps constant hydrodynamic properties of fluid and porous media (i.e., homogeneous media). The size of the grid network is conventionally selected with regard to the objective of the model, available data, elements of conceptual model, and modeler experience. The development and application of a groundwater model is a common practice for management of groundwater resources. There are several well-known text books for describing groundwater modeling (e.g., Anderson and Woessner 1992; Bear and Verruijt 1987). Numerical modeling of an aquifer is increasingly used as a management tool. The behavior of a groundwater system during the past several years and present conditions are considered for the prediction of future behavior in response to probable hydrogeological stresses. In a numerical model, the continuous domain of groundwater system is replaced by a discretized grid network. Many numerical models of a groundwater aquifer e.g., MODFLOW (McDonald 1984), assign a finite-difference network over the study area and the governing groundwater flow equation is solved at the network nodes. Determination of the grid orientation and the cell size are critical to the modeling and calibration of a groundwater flow model. As recommended by Anderson and Woessner (1992), for minimizing the water balance

error, the grid network should be aligned parallel to the features that control groundwater flow direction. The appropriate grid size is a function of the curvature in the water table, the variability in aquifer properties, the limitation in computer memory, and the size of study area (Anderson and Woessner 1992). However, variable grid size may be considered over the study aquifer for accurate modeling of aquifer response to hydrogeological stresses such as pumping, artificial and natural recharge, and pollution source. The need for local grid refinement and related methods have recently been emphasized by some individuals (e.g., Mehl and Hill 2002, 2004; Dickinson et al. 2007). Generally, the physical parameters (e.g., hydraulic conductivity, aquifer thickness, porosity, and specific storage) and hydrogeological stresses (e.g., pumping rate, recharge, evaporation, and variation of water table elevation) are assumed to be constant over each cell or over a zone consisting of several cells. Therefore, the spatial variability of hydrodynamic parameters could be hypothesized as one of constraint for selection of cell size.

This paper describes a variogram-based technique for selection of grid network in groundwater modeling. Several conventional data, e.g., aquifer electrical resistivity, pumping rate of production wells, and aquifer thickness are used in the introduced approach. This approach is applied in an alluvial unconfined aquifer, the Boushkan Plain located in southwest Iran, and the results discussed. This research attempts to answer the question: how we can select the grid size for groundwater modeling?

Methodology

The variogram has been extensively used to measure the spatial variability of spatial data in different fields of earth sciences such as petroleum industry, mining activities, environmental science, hydrology, and climatology. Geostatistical tools are being extensively applied for the spatial and/or temporal modeling of regionalized variables. Regionalized variables (i.e., $Z(u)$, $u \in$ study area) were considered as a realization of regionalized function. The spatial variability of a regionalized variable could be evaluated by the variogram. Given two locations, u and $u + h$, inside the field of a regionalized variable $Z(u)$, the variogram is a measure of one-half of the mean square error between $Z(u)$ and $Z(u + h)$ (Rouhani and Hall 1988). Theoretically, the variogram ($\gamma(h)$) is the expected squared difference between two data values separated by a distance between h according to Gringarten and Deutsch (2001):

$$2\gamma(h) = \text{Var}[Z(u) \cdot Z(u + h)] = E\{[Z(u) \cdot Z(u + h)]^2\} \quad (1)$$

The behavior of variogram versus lag distance can be directly used for comparative purpose and for sampling design (Theodossiou and Latinopoulos 2006; Modis and Papaodysseus 2006; Rouhani and Hall 1988). The spatial sampling strategies were summarized by Cox et al. (1997). Evaluation and optimization of a groundwater sampling network were addressed in several technical papers (Bueso et al. 1999; Olea 1995; Stevick et al. 2005). The degree of dissimilarity of the attribute could be evidenced by increasing the rate of the variogram (Chiles 1999). The range is defined as a distance where the attribute values are correlated. The range gives an exact sense to the conventional concept of area of influence of a sample (Chiles 1999).

Example application

A case study is presented here to demonstrate the approach described in this paper. The study area, Boushkan aquifer, is located in southwest Iran (Fig. 1). The Boushkan aquifer is an unconfined alluvial aquifer with a thickness 5 m to more than 120 m. The study area is one of the most productive agriculture activities in southwest Iran. Uncontrolled exploitation of groundwater and over pumping caused more than 13 m of drawdown of the water table during the years 1999–2009 (Bousheher Regional Water Company 2010). Accordingly, several technical studies such as surface geophysical surveys, completion of a piezometric network, hydrochemical analysis, and prevention of new production wells were implemented to achieve hydrogeological balance in the Boushkan aquifer. Hydraulic conductivity was determined based on four pumping tests. The results of the pumping tests show that hydraulic conductivity ranges from 2.1 to 25 m/day (Mohammadi and Nassimi 2010). Despite the importance of hydraulic conductivity in selection of grid size for groundwater modeling, the spatial distribution of this variable in the study area was insufficient for computing a variogram and for grid size selection according to the approach described in this study. It is emphasized that a full analysis must include the spatial distribution of all hydrodynamic variables in the study aquifer such as hydraulic conductivity (Fig. 2).
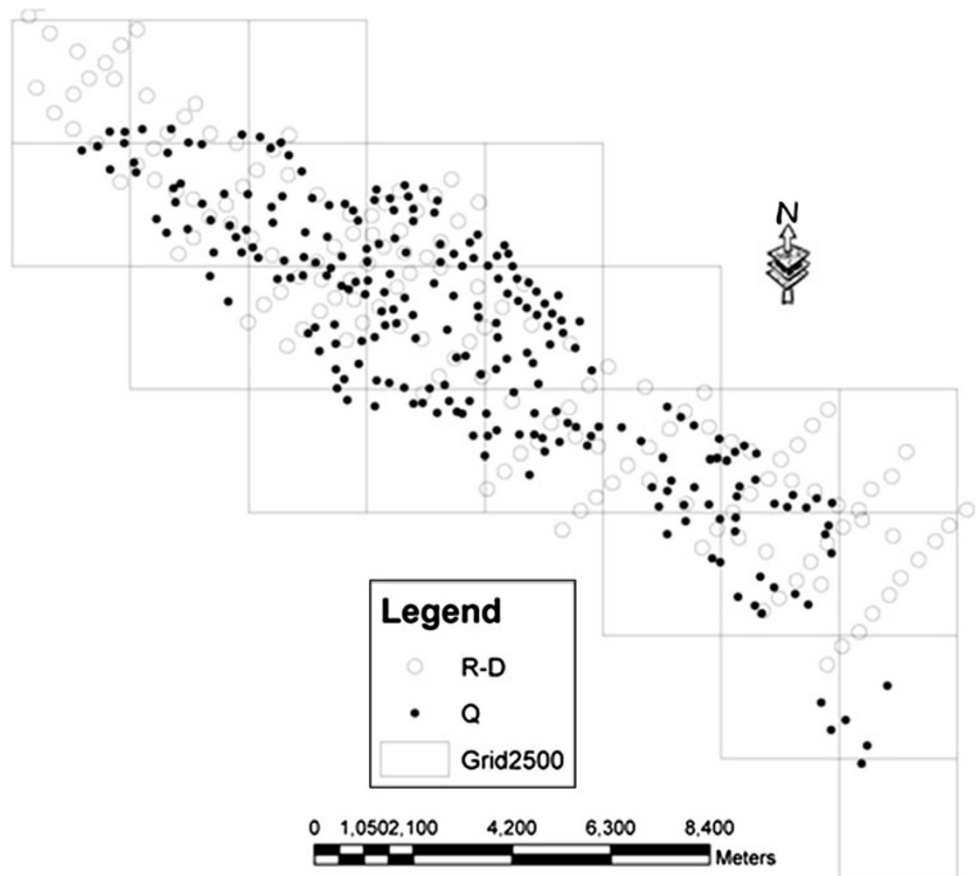
Available data in this study include (1) electrical resistivity (R) of aquifer sediment at 156 points, (2) aquifer thickness (D) at 156 points, and (3) average annual pumping volume (Q) from 252 pumping wells (Fig. 2). The electrical resistivity of the aquifer material was estimated by geo-electrical method using the Schlumberger pattern. Thickness of the aquifer and topography of the aquifer bottom were determined based on the sharp variation of the R value (at 156 points) between the aquifer material and lower impervious layer (Bousheher Regional Water Company 2010). The estimated aquifer thickness was calibrated

Fig. 1 The study area

with data obtained from several borings, including 11 exploration wells and 17 observation wells drilled to bed rock. Aquifer thickness estimates based on resistivity data were significantly improved as a result. The pumping rate was measured using a volumetric gauge in the production wells (Bousheher Regional Water Company 2010).

Results and discussion

Statistical characteristics of the variables are given in Table 1. Coefficient of variation (CV) for the variable is computed from:

$$CV = \frac{\sigma}{m} \tag{2}$$

where σ and m are standard deviation and mean, respectively. CV is a dimensionless statistical ratio that represents a comparison of variability over a limited domain free from scale effect.

The aquifer thickness shows data that exhibit a relatively low coefficient of variation of about 0.4 with an average of 54 m. The average annual pumping rate in the study area was about 3.3 l/s (or 103,743 m^3/year) with a coefficient of variation of 0.6. However, the electrical resistivity had relatively a high coefficient of variation of 1.1 with an average of 1,396 Ω/m. Despite a relatively low variation in aquifer thickness and annual pumping rate in the study area, resistivity exhibited a higher variation.

Variograms are used for developing a spatial dependence model of the variables (Fig. 3). Table 2 presents the characteristics of the variograms. The variogram range is estimated from 2,000 m to 2,500 m for Q, R and D,

Fig. 2 Assumed grid network and location of data points

Table 1 Summary statistics of the variables

Variable	Number of points	Average	Minimum	Maximum	SD	CV
Q (m³/year)	230	103,743	6,832	518,400	61,318	0.6
R (Ω/m)	156	1,396	0	8,540	1,507	1.1
D (m)	153	54	6	94	18	0.4

Table 2 Characteristics of variogram of the variables

Variable	Range	Sill	Nugget effect	Model
Q	1,800	0.8	0	Exponential
R	2,000	1	0	Exponential
D	2,500	0.9	0	Exponential
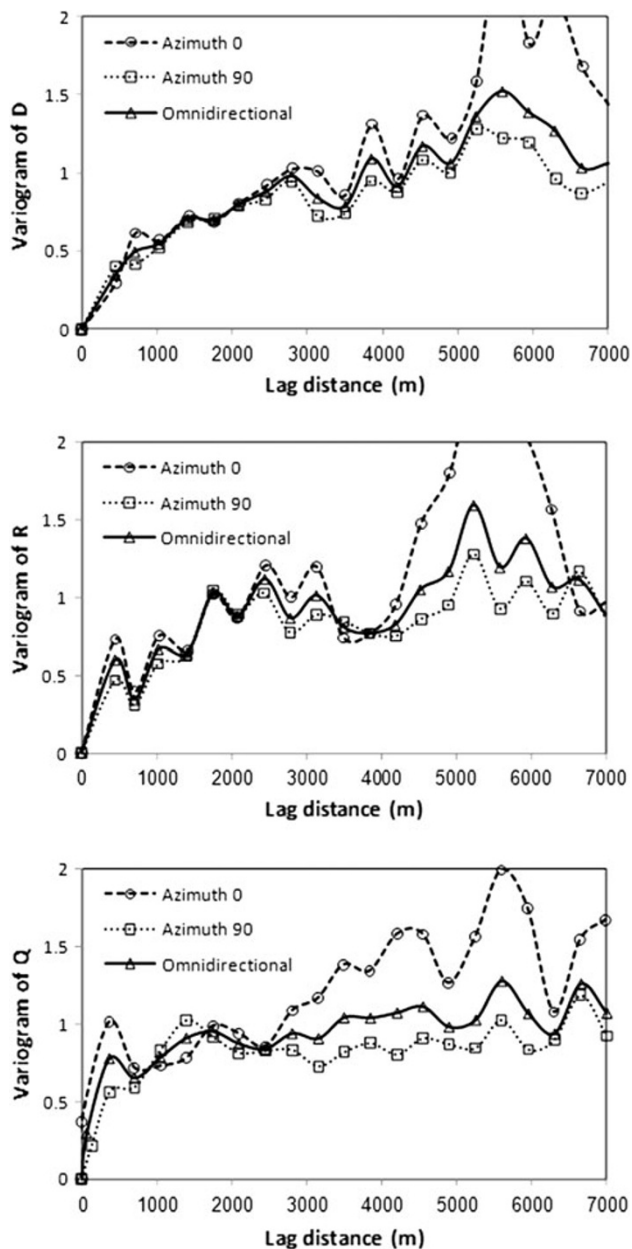
Fig. 3 Standardized variogram of Q, R and D in different directions

respectively. Conceptually, similarity between each variable at distances less than the range of influence could be replaced with a homogeneous cell in a groundwater model. Although the range of influence is sensitive to the density

of data points, a sense of similarity and homogeneity of each variable from the available data set was obtained.

Since, the spatial dependence (i.e., range in Table 2) is not the same for all variables used, a statistical test is proposed for confirming the homogeneous grid size. Assuming different grid networks (from 1,000 to 5,000 m) it is possible to compare the characteristics of data points in each cell for each grid size. Different grid networks were overlaid on to the data points and then the data points in each grid were selected for computation of statistical parameters of the variables. The mean and standard deviation of the data points of the variables (Q, R and D) in each cell are used for computation of the CV in each cell based on Eq. (2). These computations were repeated for all cell sizes ranging from 1,000 to 5,000 m. The values for each variable over a homogenous grid size (i.e., each cell in grid network) typically showed low variation. Therefore, the grid network that exhibited a low CV associated with each cells (Fig. 4) could be selected as the optimum grid size for a homogenous domain in the groundwater model. The degree of variability of the variables (R, D and Q) is compared using CV versus grid size in Fig. 4. CV was computed in each cell (e.g., 25 cells in grid network of 2,500 m) for eight grid sizes including 1,000, 1,500, 2,000, 2,500, 3,000, 3,500, 4,000 and 5,000 m. Then a box plot of CV was developed for each grid size. Relatively, small dimension of the box plot of CV means that the data points of the variable tend to have same values in each cell (Fig. 4). The results showed relatively high variability of CV for small cell sizes (Fig. 4). It seems that a higher variability of CV in a small cell size is caused by small scale variation of the variables due to grain size distribution, soil compaction, and variation in soil packing. Alternatively, large-scale variation of the variables, such as changes in sediment facies or/and structural and geological variations could effect on the variation of CV in higher cell sizes. As shown in Fig. 4, fluctuations in CV of the variables are smooth for grid sizes larger than 2,500 m, which was proposed as the range of influence in the variograms earlier in Fig. 3. Although the grid size was selected based on the analysis of three available variables in the study area, it seems that optimum grid size should be developed according to all hydrodynamic variables such as hydraulic conductivity.
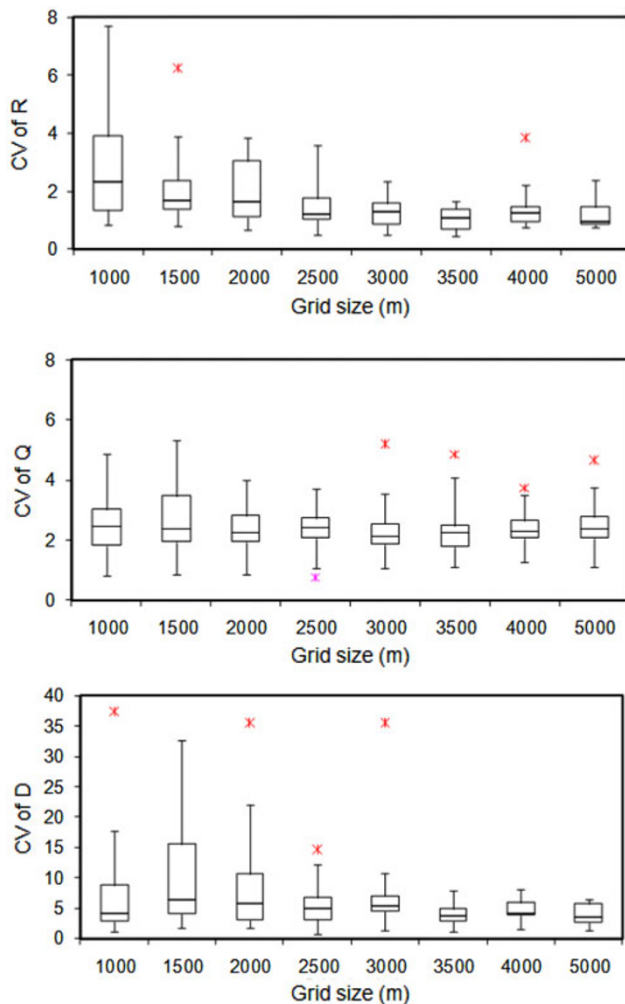
Fig. 4 Box plot of coefficient of variation in different grid sizes

Conclusion

This paper has shown that variogram can be used for selecting grid size in groundwater model. The range of influence in the variogram concept is used to delineate a homogeneous domain (i.e., grid size) with maximum similarity of the variables. The approach is applicable in most instances when primary hydrogeological data such as aquifer thickness, pumping rate and electrical resistivity of the aquifer are available. This approach can be applied as a preliminary analysis of the data that will later be used in a groundwater model. The variogram-based approach provides possible spatial anisotropies of the data and range of similarity between measured values of the data. The approach could be applied as one of the steps that have to be followed to select the optimum grid size of an aquifer model. The approach requires a greater density of data points at some locations for all variables to provide more reliable results due to scale dependence of the range. A

comprehensive analysis based on the proposed approach can include the spatial distribution of the hydraulic conductivity for improving the results.

References

Anderson MP, Woessner WW (1992) Applied groundwater modeling: simulation of flow and advective transport. Academic Press, San Diego

Bear J, Verruijt A (1987) Modeling groundwater flow and pollution theory and application of transport in porous media. D Reidal Publishing, Dordrecht

Bousheher Regional Water Company (2010) Hydrogeology of Boushkan aquifer (in Farsi) (unpublished)

Bueso MC, Angulo JM, Cruz-Sanjulian J, Garcia-Arostegui JL (1999) Optimal spatial sampling design in a multivariate framework. Math Geol 31(5):507–525

Chiles J-P, Delfiner P (1999) Geostatistics: modelling spatial uncertainty. Wiley Interscience, New York, p 695

Cox DD, Cox LH, Ensor KB (1997) Spatial sampling and the environment: some issues and directions. Environ Ecol Stat 4:219–233

Dickinson JE, James SC, Mehl S, Hill MC, Leake SA, Zyvoloski GA, Faunt CC, Eddebbarh A (2007) A new ghost-node method for linking different models and initial investigations of heterogeneity and nonmatching grid. Adv Water Resour 30:1722–1736

Gringarten E, Deutsch CV (2001) Variogram interpretation and modeling. Math Geol 33(4):507–534

McDonald MC, Harbaugh AW (1984) A modular three-dimensional finite difference groundwater flow model. US Geological Survey, technique of water resources investigations (book 6), Washington

Mehl S, Hill MC (2002) Evaluation of a local grid refinement method for steady-state block-centered finite-difference groundwater models. Dev Water Sci 47:367–374

Mehl S, Hill MC (2004) Three-dimensional local grid refinement for block-centered finite-difference groundwater models using iteratively coupled shared nods: a new method of interpolation and analysis of errors. Adv Water Resour 27:899–912

Modis K, Papaodysseus K (2006) Theoretical estimation of the critical sampling size for homogeneous ore bodies with small nugget effect. Math Geol 38(8):489–501

Mohammadi Z, Nassimi A (2010) Comparison of the pumping test results using different methods in the unconfined aquifer. Appl Adv Geol (Shahid Chamran University, Iran) 2(1):8–21 (in Farsi)

Olea RA (1995) Fundamentals of semivariogram estimation, modeling, and usage. In: Yarus JM, Chambers RL (eds) Stochastic modeling and geostatistics: principles, methods, and case

studies. American Association of Petroleum Geologists, Tulsa, pp 27–36

Rouhani S, Hall TJ (1988) Geostatistical schemes for groundwater sampling. J Hydrol 103:85–102

Stevick E, Pohll G, Huntington J (2005) Locating new production wells using a probabilistic-based groundwater model. J Hydrol 303(1–4):231–246

Theodossiou N, Latinopoulos P (2006) Evaluation and optimization of groundwater observation networks using the Kriging methodology. Environ Model Softw 21:991–1000