Amirkabir Journal of Civil Engineering

Amirkabir J. Civil Eng., 52(9) (2020) 535-538
DOI: 10.22060/ceej.2019.15955.6087

Hydraulic conductivity and uncertainty analysis of between-models and input data by using Bayesian model averaging of artificial intelligence model

Y. Hassanzadeh[1,*] , M. Moazamnia[2], S. Sadeghfam[3], A. Nadiri[4]

[1] Department of Water Engineering, Center of Excellence in Hydroinformatics, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran
[2] PhD student Faculty of Civil Engineering, University of Tabriz
[3] Assistant Professor, Faculty of Engineering, University of Maragheh
[4] Associate Professor, Faculty of Earth Sciences, University of Tabriz

ABSTRACT: The estimation of hydraulic conductivity is one of the most important part of hydrogeological studies which is important in groundwater management. But due to practical, time or cost constraints, direct measurement is difficult. Hence, the using artificial intelligence models with low cost and high performance can be an appropriate alternative for this purpose. Since input data and different training techniques in these models are the most important source of uncertainty, the effect of various sources of uncertainty in output should be considered. In this research a Bayesian Model Averaging (BMA) are developed which includes the model combination of artificial neural network, fuzzy logic and neuro-fuzzy to estimate hydraulic conductivity and uncertainty analysis. In the BMA model, the weight of the models is determined by the Bayesian information criterion (BIC), and the within-model variance, steam from the uncertainty of input data and the between-model variance steam from uncertainty associated with the nature of the artificial intelligence model are calculated. In this study, the developed method has been applied to estimate the hydraulic conductivity in the Urmia aquifer. The results show that although the determination coefficient of BMA is not higher than the determination coefficient of the best model, the output of the BMA is the result of assigning weights that take into account the uncertainty between the models and the input data. Also, the effect of groundwater level variation on estimated hydraulic conductivity from pumpage test up to 2015 was evaluated and the result indicated an insignificant changes in hydraulic conductivity.

Review History:

Received: 2019-03-09
Revised: 2019-04-14
Accepted: 2019-04-16
Available Online: 2019-05-11

Keywords:

Bayesian Model Averaging
Hydraulic Conductivity
Artificial Neural Network
Fuzzy Logic
neuro-fuzzy

1. Introduction

Hydraulic conductivity were estimated through various Artificial Intelligence (AI) methods, such as ANN (Merdun et al, 2006; Sun, 2011; Yao et al, 2015), FL (Ross et al. 2007; Kadkhodaie-Ilkhchi and Amini 2009) and NF (Malki and Baldwin 2002; Hurtado et al. 2009) [1-7]. Although hydraulic conductivity has been estimated by various AI models, limited research are developed which analyze uncertainty associated with artificial intelligence techniques by Bayesian Model Averaging (BMA) method.

BMA transformed into a practical tool since Draper (1995) and Moazamnia et al. (2019), is a strategy to combine Multiple Models (MM) often constructed by perturbing parameters; and to use its capability for assessing inherent uncertainties [8-9]. This paper investigates performances of BMA by combining separate MMs comprising three different AI techniques for predicting hydraulic conductivity of Urmia aquifer. Also, due to the decline in groundwater level in Urmia plain, the effect of the decline on the estimated hydraulic conductivity changes by the Bayesian model are evaluated.

In this research, the effect of uncertainty of input data and AI models for predicting hydraulic conductivity in Urmia plain are investigated by the Bayesian Model Averaging approach. AI models include ANN, SFL, and NF models, which are common models for predicting hydrogeological parameters. After predicting the hydraulic conductivity, the within-model uncertainty and the between-model uncertainty is calculated spatially.

*Corresponding author's email: yhassanzadeh@tabrizu.ac.ir

2. Methodology

The Law of Total Probability: BMA combines n plausible models as expressed by Eq. (1) below as follows, (Draper 1995) [8]:

$$\Pr(\Delta \mid D) = \sum_{p=1}^{n} \Pr(\Delta \mid D, M_p) \Pr(M_p \mid D) \tag{1}$$

where $\Pr(\Delta \mid D)$ is the probability of the prediction of hydraulic conductivity (denoted as Δ) given the measured hydraulic conductivity (denoted as D); $\Pr(\Delta \mid D, M_p)$ is the conditional probability of the predicted quantity given the observed data D and given model (M_p); and $\Pr(M_p \mid D)$ is

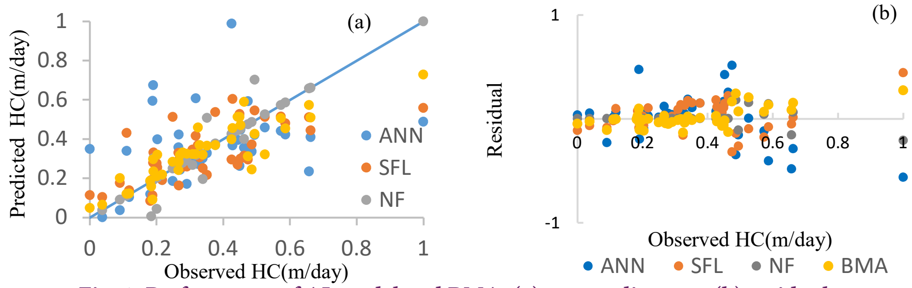

Fig. 1. Performance of AI model and BMA: (a) scatter diagram; (b) residual error

the posterior probability of the model, which are also known as model weight, given the data D, (see Draper 1995; Hoting et al., 1999)[9]. In this study Δ represents predicted hydraulic conductivity, D denotes input data (d, EC, R_t, B) and (M_p) denotes ANN, SFL and NF.

Bayesian Theorem: The Bayes theorem, formulated for BMA, uses n plausible models $\{M_1, M_2, \ldots, M_n\}$, where each array is one representation of the state variable of predicting hydraulic conductivity, and their corresponding measured hydraulic conductivity values at each of the observation well are denoted by D. The theorem is detailed by Berger, 1985 and is expressed as[10]:

$$\Pr(M_p \mid D) = \frac{\Pr(D \mid M_p)\Pr(M_p)}{\sum_{j=1}^{n} \Pr(D \mid M_j)\Pr(M_j)} \tag{2}$$

Where $\Pr(M_p \mid D)$ is the posterior probability, which learns a better estimate from the given data; $\Pr(M_p)$ is a prior model probability for the model M_p, evaluated by expert judgments or estimated e.g. Wöhling et al. (2015); $\Pr(D \mid M_p)$ is marginal likelihood function for model M_p. As per Li and Tsai (2009), marginal likelihood function is approximated by [11-12]:

$$\Pr(D \mid M_p) \approx \exp[-\frac{1}{2} BIC_p] \tag{3}$$

$$BIC_p = Q_p + N \ln 2\pi + m_p \ln N \tag{4}$$

$$Q_p = (\Delta^{cal} - \Delta^{obs}) C_\Delta^{-1} (\Delta^{cal} - \Delta^{obs})^T \tag{5}$$

where N is number of data; m_p is number of model parameters; Q_p is the sum of weighted squared errors expressed by Eq. (5); Δ is predicted hydraulic conductivity and D is measured hydraulic conductivity; C_Δ is the variance matrix of prediction errors using Monte Carlo simulations on model parameters (Li and Tsai 2009) [12]. Eqs. (3), (4) and (5) are replaced in Eq. (2) and are manipulated to derive the following:

$$\Pr(M_p \mid D) = \frac{\exp(-\frac{1}{2}\alpha \Delta BIC_p)}{\sum_{i=1}^{n} \exp(-\frac{1}{2}\alpha \Delta BIC_p)} \tag{6}$$

where $\Delta BIC_p = BIC_p - BIC_{min}$; in which BIC_{min} (Bayesian Information Criteria) is the lowest BIC value among the models; α is a scaling factor used in the variance window.

3. DISCUSSION AND RESULTS

To estimate the hydraulic conductivity using the AI models, the authors used *d* (distance of each estimation point to the origin of the coordinate system), *B* (the thickness of the aquifer), R_t (transverse resistance of the aquifer), and *EC* (the salinity of formation water) as input parameters. The input data uncertainty was considered from Kriging variances, which propagated to the AI model logK output through weights and rules. The uncertainty for input data is considered for *Rt* and *EC*, because the values of *D* and *B* in the location of the hydraulic conductivity measurement are definite. The scatter diagram and residual error of AI models and Bayesian model averaging method is shown in Fig 1. Fig. 1(a) shows that the results of the Bayesian model have less dispersion than other models. In Fig. 1 (b), it is also observed that the residual error diagram of the Bayesian model is lower than other models. The highlights of the overall results presented in this sections are as follows: (i) no single model performs the best in all cases; (ii) performance metrics are useful summaries and together with scatter diagrams they uncover the aspects hidden by performance metrics that the fitted models are hardly perfect.

4. CONCLUSIONS

The results of the three AI models show that there is no single model performing the best but they have convergences and divergences. BMA combines these modelling results into a single model, in which the combined model is a learning from the convergence and divergence of the AI models and as such it performs better than the individual models most of the time but overwhelmingly reduces the scatters in the error residuals.

REFERENCES

[1] Merdun, H., Çınar, Ö., Meral, R., & Apan, M. (2006). Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. *Soil and Tillage Research*, *90*(1-2), 108-116.

[2] Sun, J., Zhao, Z., & Zhang, Y. (2011). Determination of three dimensional hydraulic conductivities using a combined analytical/neural network model. *Tunnelling and Underground Space Technology*, *26*(2), 310-319.

[3] Yao, Y., Zheng, C., Liu, J., Cao, G., Xiao, H., Li, H., & Li, W. (2015). Conceptual and numerical models for groundwater flow in an arid inland river basin. *Hydrological Processes*, *29*(6), 1480-1492.

[4] Ross, J., Ozbek, M., & Pinder, G. F. (2007). Hydraulic conductivity estimation via fuzzy analysis of grain size data. *Mathematical Geology*, *39*(8), 765-780.

[5] Kadkhodaie-Ilkhchi, A., & Amini, A. (2009). A fuzzy logic approach to estimating hydraulic flow units from well log data: A case study from the Ahwaz oilfield, South Iran. *Journal of Petroleum Geology*, *32*(1), 67-78.

[6] Malki, H. A., & Baldwin, J. (2002). A neuro-fuzzy based oil/gas producibility estimation method. In *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)* (Vol. 1, pp. 896-901). IEEE.

[7] Hurtado, N., Aldana, M., & Torres, J. (2009). Comparison between neuro-fuzzy and fractal models for permeability prediction. *Computational Geosciences*, *13*(2), 181-186.

[8] Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society. Series B (Methodological)*, 45-97.

[9] Moazamnia, M., Hassanzadeh, Y., Nadiri, A. A., Khatibi, R., & Sadeghfam, S. (2019). Formulating a strategy to combine artificial intelligence models using Bayesian model averaging to study a distressed aquifer with sparse data availability. *Journal of Hydrology*.

[10] Berger, J. O. (1985). Statistical decision theory and Bayesian inference. *Springer0Verlag (New York)*.

[11] Wöhling, T., Schöniger, A., Gayler, S., & Nowak, W. (2015). Bayesian model averaging to explore the worth of data for soil-plant model selection and prediction. *Water Resources Research*, *51*(4), 2825-2846.

[12] Li, X., & Tsai, F. T. C. (2009). Bayesian model averaging for groundwater head prediction and uncertainty analysis using multimodel and multimethod. *Water resources research*, *45*(9).

HOW TO CITE THIS ARTICLE

Y. Hassanzadeh, M. Moazamnia, S. Sadeghfam, A. Nadirid, Hydraulic conductivity and uncertainty analysis of between-models and input data by using Bayesian model averaging of artificial intelligence model, Amirkabir J. Civil Eng., 52(9) (2020) 535-538.

DOI: 10.22060/ceej.2019.15955.6087