Journal of Hydrology 548 (2017) 569–587

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

# Conjunction of radial basis function interpolator and artificial intelligence models for time-space modeling of contaminant transport in porous media

Vahid Nourani [a,b,*], Shahram Mousavi [c], Dominika Dabrowska [d], Fahreddin Sadikoglu [e]

[a] Dept. of Water Resources Engineering, Faculty of Civil Eng., Univ. of Tabriz, Iran
[b] Department of Civil Engineering, Near East University, Nicosia, North Cyprus, Mersin 10, Turkey
[c] Young Researchers and Elite Club, Miyaneh Branch, Islamic Azad University, Miyaneh, Iran
[d] Dept. of Hydrogeology and Engineering Geology, Faculty of Earth Sci., Univ. of Silesia, Poland
[e] Engineering Faculty, Near East University, Nicosia, North Cyprus, Mersin 10, Turkey

ARTICLE INFO

*Article history:*
Received 12 November 2016
Received in revised form 8 March 2017
Accepted 18 March 2017
Available online 21 March 2017
This manuscript was handled by G. Syme, Editor-in-Chief, with the assistance of Hazi Mohammad Azamathulla, Associate Editor

*Keywords:*
Groundwater
Contaminant transport
Artificial intelligence
Meshless method
Imperialist competitive algorithm

ABSTRACT

As an innovation, both black box and physical-based models were incorporated into simulating groundwater flow and contaminant transport. Time series of groundwater level (GL) and chloride concentration (CC) observed at different piezometers of study plain were firstly de-noised by the wavelet-based de-noising approach. The effect of de-noised data on the performance of artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) was evaluated. Wavelet transform coherence was employed for spatial clustering of piezometers. Then for each cluster, ANN and ANFIS models were trained to predict GL and CC values. Finally, considering the predicted water heads of piezometers as interior conditions, the radial basis function as a meshless method which solves partial differential equations of GFCT, was used to estimate GL and CC values at any point within the plain where there is not any piezometer. Results indicated that efficiency of ANFIS based spatiotemporal model was more than ANN based model up to 13%.



## 1. Introduction

Groundwater provides a crucial resource of water for drinking, agriculture and industrial purposes. So, of the most important environmental issues are management and conservation of groundwater sources from different contaminants. When groundwater is contaminated, removal of contaminants and the restoration of quality may be slow and sometimes, impractical. It can be harmful for human health, the ecosystem and can result in water shortage. Thus, simulation of contaminant transport can be an important task in hydro-environmental studies and consequently, it is necessary to develop the robust models which can determine the location and amount of pollution. For modeling groundwater flow and contaminant transport (GFCT), several computational methods, namely, finite difference method, finite volume method, finite element method, and boundary element method have been applied for numerical solution of governing physical-based partial differential equation (PDE) (Bear and Cheng, 2010). Recently, meshless or meshfree technique as an alternative method has been used to solve PDEs of the GFCT in porous media. In a meshless technique, a collection of scattered nodes is employed instead of generated mesh over the problem domain. Meshless techniques include the smooth-particle hydrodynamics, kernel method, moving least squares method, the element-free Galerkin method, partition of unity method, local Petrov-Galerkin method, radial point interpolation method and the method of radial basis functions (RBFs). Each method has its own advantages for particular problems (Dehghan and Shirzadi, 2015a,b). Among various meshless methods, the RBF-based methods (e.g., Kansa's collocation method) with many general applications, have been more popular numerical schemes in compared to other meshless methods because: a) There is no need for boundary and domain discretizations; b) There is no need for integration over boundary and domain; c) In some instances, they converge exponentially to provide smooth solutions; d) RBFs are of great importance in solving complex high dimensional problems due to their dependence on the Euclidean distance between points as univariate functions;

* Corresponding author at: Dept. of Water Resources Engineering, Faculty of Civil Eng., Univ. of Tabriz, Iran.
*E-mail addresses:* vnourani@yahoo.com (V. Nourani), mousavi481@gmail.com (S. Mousavi), d.dabrowska@us.edu.pl (D. Dabrowska).

http://dx.doi.org/10.1016/j.jhydrol.2017.03.036
0022-1694/

e) Further information including interior conditions, boundary conditions etc. can be added or deleted at any step of the modeling; f) And finally, their coding and implementation is simple (Nourani and Babakhani, 2013).

Multiquadric (MQ) RBF approximation technique was first used by Hardy (1971, 1990) for the scattered geographical data interpolation and then Kansa (1990) adopted Hardy's MQ-RBF technique to solve PDEs on irregular domains. Franke and Schaback (1997) illustrated that MQ-RBF has the best performance compared to other examined schemes. Also, a comparison between finite element method and RBF methods was conducted by Li et al. (2003). Several studies have investigated the capability of meshless methods in the field of GFCT modeling (e.g., Boztosun and Charafi, 2002; Boztosun et al., 2002; Li et al., 2003; Herrera et al., 2009; Alhuri et al., 2011; Li and Mao, 2011; Meenal and Eldho, 2012; Swathi and Eldho, 2013; Dehghan and Shirzadi, 2015a).

Although the physical-based numerical techniques are widely used for temporal and/or spatial modeling of hydro-environmental systems, some real-world conditions such as anisotropy and heterogeneity can have meaningful impacts on GFCT and restrict the usefulness of such methods. As a result, these methods may be replaced by data-driven or black box techniques when there is no sufficient field data set and output accuracy is preferred over physical perception of the phenomenon. The uncertainty and complexity of the groundwater process have caused data-driven models such as artificial neural networks (ANNs) and adaptive neuro-fuzzy inference system (ANFIS) are widely used by hydrogeologists. Several studies have been performed to examine the susceptibility of artificial intelligence (AI) models for GFCT modeling (e.g., Coulibaly et al., 2001; Mohamed and Hawas, 2004; Singh et al., 2004; Singh and Datta, 2007; Nourani et al., 2008; Li and Tsai, 2009; Bashi-Azghadi et al., 2010; Foddis et al., 2013; Shiri et al., 2013; Taormina and Chau, 2014; Foddis et al., 2015; Nourani et al., 2015).

Uncertainty involved in the model inputs and field parameters like hydraulic conductivity of the soil, dispersion coefficient, their temporal variations and unknown boundary conditions restrict the GFCT model efficiency. In this research, both ANN and ANFIS as AI-based black box models and MQ-RBF as a physical-based PDE solving technique were incorporated accurately into simulating GFCT. On the other hand, since the data quality and involved noise in data have significant impact on the efficiency of any data-driven model, application of a data de-noising technique can increase the capability of GFCT modeling. Recent studies have shown that the performance of models applied for exploration and forecasting of stochastic or deterministic systems can be enhanced using de-noised data. Li and Mao (2011) studied the effect of injected artificial noise to the groundwater quality data for identifying contaminant sources. Nourani et al. (2014b) indicated that the intensity of the noises which effect on the results of AI models relay on both nature and level of the noise. The wavelet threshold-based signal de-noising approach which identifies the localized features of non-stationary signals in both time and frequency domains, is a potential de-noising filter with regard to other de-noising techniques (e.g., Wiener and Kalman filters which is only capable of dealing with linear natural systems; ensemble Kalman filter which is suitable only for Gaussian error and just propagates the first two moments of error and has limited effectiveness for highly nonlinear uncertainty evolutions) (Nourani et al., 2014a). In this regard, Cannas et al. (2006) and Nourani et al. (2009) applied multi-resolution analysis based on wavelet transform to improve the efficiency of ANN-based river flow forecasting models.

The proposed hybrid model in this study contains several features which distinguish it from previously proposed and published GFCT modeling tools. The GFCT modeling in an infinite aquifer can be limited due to the uncertainty of data and conceptualization of model parameters, non-linear nature of the phenomenon, noise involved in the observed data, insufficient data related to the domain boundaries. In order to overcome the mentioned issues of GFCT modeling, a hybrid artificial intelligence-meshless model linked to a threshold-based wavelet de-noising method is suggested in this paper. In the proposed method, in order to predict groundwater level (GL) and chloride concentration (CC) one time step ahead, an AI-based black box non-linear model (i.e., ANN or ANFIS) is used to tackle the non-linear inherent of temporal variation of process. Also, the cross-wavelet coherence is linked to the modeling structure to improve the capability of the AI model by selecting most dominant inputs. On the other hand, linear form of well-known physical-based Richards' and advection-dispersion PDEs are respectively used to identify spatial variations of the GLs and CCs. The PDEs are solved in couple by MQ-RBF technique which is a well-suited method for a problem with uncertain boundary conditions but with known interior data, observed at different points (piezometers, in this study).

## 2. Materials and methods

### 2.1. Proposed hybrid AI-meshless model

In this research, in order to predict spatiotemporal GL and contaminant concentration in porous media, AI models (i.e., ANN and ANFIS) were employed for temporal forecasting and MQ-RBF method as a meshless model was used for spatial prediction of Richards' and advection-dispersion PDEs (i.e., Eqs. (5) and (7)). In other words, temporal and spatial terms of these equations were solved via AI models and meshless method, respectively. Thus, both black box and physical-based models were incorporated into simulating GFCT. Nomenclatures were used in proposed methodology, are presented in Table 1.

The proposed hybrid AI-meshless approach for GFCT modeling includes four distinguished stages (Fig. 1). In the first stage of modeling, observed GLs and CCs time series of all piezometers are de-noised using wavelet de-noising approach. De-noised time series were used as inputs of the AI models into investigate influence of noisy data in the results of modeling. In the second stage, in order to forecast the GL and CC one time step ahead, an AI (i.e., ANN or ANFIS) model is trained and verified for each piezometer considering river discharge (runoff), rainfall, water level of the lake, CC in river and GLs with different lags as potential input data set. In this stage, WTC was employed to cluster piezometers and identify correlation among hydrological parameters and therefore to determine input parameters of the AI model. In the third stage, considering the predicted GLs obtained from the second stage as interior conditions, the MQ-RBF model as a meshless technique was employed to solve spatial terms of PDE of groundwater flow (i.e., Eq. (5)). In MQ-RBF model, the inputs are GL, UTMx (i.e., x-coordinate of collocation points), UTMy (i.e., y- coordinate of collocation points) and shape coefficient. In this stage, one time step ahead GL is determined at any interested point of plain where no piezometer exists to observe GL. Also, the velocity in any desired point is calculated using the Darcy's law and the optimum value of RBF's shape coefficient ($c_{s1}$) (for groundwater flow equation) is determined at any time step using imperialist competitive algorithm. Finally, using the values of velocity and CC values of piezometers predicted in the second stage, the CC value for the next time step is computed at any desired point using MQ-RBF based solution of contaminant equation (Eq. (7)). In MQ-RBF model, the inputs are GL, CC, UTMx, UTMy, hydraulic conductivity, dispersivity and shape coefficient. In this stage, imperialist competitive algorithm is employed as an optimization tool to find optimum values of hydraulic conductivity, longitudinal dispersivity