Journal of Hydrology 519 (2014) 3193–3203

Contents lists available at ScienceDirect

Journal of Hydrology

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

A spatio-temporal hybrid neural network-Kriging model for groundwater level simulation

Evdokia Tapoglou [a], George P. Karatzas [a,*], Ioannis C. Trichakis [b], Emmanouil A. Varouchakis [a]

[a] School of Environmental Engineering, Technical University of Crete, Polytechneioupolis, Chania 7310, Greece

[b] European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES), Water Resources Unit, Via Enrico Fermi 2749, 21027 Ispra (VA), Italy

ARTICLE INFO

Article history:
Received 10 June 2014
Received in revised form 2 October 2014
Accepted 14 October 2014
Available online 22 October 2014
This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Stephen Worthington, Associate Editor

Keywords:
Artificial Neural Networks
Kriging
Groundwater hydraulic head simulation
Spatial and temporal simulation

SUMMARY

Artificial Neural Networks (ANNs) and Kriging have both been used for hydraulic head simulation. In this study, the two methodologies were combined in order to simulate the spatial and temporal distribution of hydraulic head in a study area. In order to achieve that, a fuzzy logic inference system can also be used. Different ANN architectures and variogram models were tested, together with the use or not of a fuzzy logic system. The developed algorithm was implemented and applied for predicting, spatially and temporally, the hydraulic head in an area located in Bavaria, Germany. The performance of the algorithm was evaluated using leave one out cross validation and various performance indicators were derived. The best results were achieved by using ANNs with two hidden layers, with the use of the fuzzy logic system and by utilizing the power-law variogram. The results obtained from this procedure can be characterized as favorable, since the RMSE of the method is in the order of magnitude of 10^{-2} m. Therefore this method can be used successfully in aquifers where geological characteristics are obscure, but a variety of other, easily accessible data, such as meteorological data can be easily found.

1. Introduction

Data driven models have been used in the last decades in hydrology with many of the applications focusing specifically on groundwater hydrology. These modeling techniques use mathematical equations to describe the physical system and necessitate easily measurable data time series. They are particularly useful for overcoming model development and calibration limitations and their effectiveness depends almost entirely on how closely they represent the real-world system, which can be highly complex with vast changes in properties. Artificial Neural Networks (ANNs) belong to the category of data driven models which treat the system as a black box. Their objective is to find a relationship between input and output data without using any physical data, such as hydraulic conductivity and porosity (ASCE, 2000). The performance and the capabilities of a wide range of data driven models in hydrology, including ANNs, are presented by Elshorbagy et al. (2010a,b).

In this study, ANNs are combined with the Kriging interpolation technique in order to predict the hydraulic head change in a region, both temporally and spatially. More specifically, ANNs are used for the temporal prediction and Kriging for the spatial one. The two methods are combined by using a fuzzy logic system.

1.1. Artificial Neural Networks

Artificial Neural Networks provide an alternative to conventional numerical modeling techniques which are often limited by strict assumptions of normality, linearity and variable independence. Because of their ability to derive meaning from complicated or imprecise data, ANNs can be used to extract patterns and detect trends (Haykin, 1999).

Due to their capacity to describe nonlinear relationships, ANNs are widely used in hydrology, both for surface water and groundwater (Gupta et al., 1997; Hsu et al., 1995; Rajurkar et al., 2004; Rogers and Dowla, 1994). A comprehensive review of the applications of ANNs in hydrology has been presented by Maier and Dandy (2000) and Maier et al. (2010). Various researchers (Lin and Chen, 2006; Samani et al., 2007) used ANNs to identify the parameters of an aquifer and examined the efficiency of networks using Multilayer Perceptron (MLP) and Radial Basis Function (RBF) neural networks. Applications where different parameters are estimated have been presented throughout the years, including the temporal prediction of the hydraulic head in a well, the spatial prediction of aquifer parameters or the existence of pollution (Banerjee et al., 2011; Samani et al., 2007). Such models have also

* Corresponding author.
E-mail address: karatzas@mred.tuc.gr (G.P. Karatzas).

http://dx.doi.org/10.1016/j.jhydrol.2014.10.040
0022-1694/© 2014 Elsevier B.V. All rights reserved.

been successfully used to predict salinity (Maier and Dandy, 1996) and rainfall-runoff processes (Sudheer et al., 2002). A review of the use of ANNs in modeling saltwater intrusion was conducted by Singh (2014).

Several applications of ANNs have focused on the prediction of the water level in wells using different input parameters usually directly related to the aquatic equilibrium, such as temperature, rainfall and water level in neighboring wells (Coppola et al., 2005b; Lallahem et al., 2005; Nayak et al., 2006). Coppola et al. (2005b) developed ANN models that accurately predicted transient groundwater levels in response to variable weather and pumping conditions and extended this work to water quality for an upconing problem in a coastal aquifer (Coppola et al., 2005a). ANNs have also been applied for the prediction of the hydraulic head in wells located in karstic aquifers (Trichakis et al., 2009).

Mohanty et al. (2013) compared the performance of ANNs to that of a commercial, numerical modeling software (MODFLOW) in simulating the groundwater flow in Odisha Interbasin, India. Different error indicators were used and the ANN model provided better predictions for the groundwater level over short time horizons.

The most commonly used ANN training method is back propagation, while there are also models that use optimization methodologies for ANN training (Coppola et al., 2005b, 2007). These alternative methodologies include the use of conjugate gradient methodology (Towsey et al., 1995), Genetic Algorithms (GAs) (Rogers et al., 1995), and Differential Evolution (Trichakis et al., 2009). This body of research collectively demonstrates that ANNs may serve as efficient and accurate models for simulating groundwater systems and can be used for developing effective management and protection strategies.

The uncertainty of ANN predictions was also examined in various studies (Kasiviswanathan and Sudheer, 2013; Trichakis et al., 2011b) in order to evaluate the reproductivity of the results and the influence of the stochastic nature of ANNs on the results.

Finally, ANN models have been used for the spatial definition of aquifer parameters (Rizzo and Dougherty, 1994), where the coordinates of observation wells were used as an input to the neural network and the class of hydraulic conductivity was the output. In this way the hydraulic conductivity range was determined. In other studies, similar methods were used to predict groundwater salinity (Banerjee et al., 2011) and rainfall distribution (Luk et al., 2000).

In ANNs, information processing is performed at different layers, dividing network nodes into three categories, input nodes, output nodes and hidden nodes. Input nodes do not perform calculations, while output and hidden nodes are computational nodes. The output node values are the final numerical outputs of the network. The computational nodes initially multiply every input (x_i) with the corresponding synaptic weight (w_i) and sum up the results. The sum is introduced as argument to the activation function and the resulting value is the output value of the node for the current inputs and weights. This value is then either fed to the next hidden layer or to the output of a neural network (ASCE, 2000).

Various types of activation functions can be used in a neural network (e.g., linear, threshold, sigmoid). The sigmoid function is by far the most common form of activation function used in ANNs, due to its ability to describe nonlinear relationships, in this case a natural process (Nayak et al., 2006; Trichakis et al., 2011a).

In the present study, a sigmoid function is used in a feed forward Artificial Neural Network which simulates the water budget, as described by Eq. (1).

$$\Delta S = I - O + P - \mathrm{EPT} \pm Q \tag{1}$$

where ΔS is the change in storage of the aquifer, I the inflow to the water basin, O the outflow from the water basin, P the precipitation, EPT the evapotranspiration and Q the pumping/recharge rate.

1.2. Kriging

Various techniques of spatial interpolation have been presented throughout the years, as described by Li and Heap (2008). However, most of them are somehow interconnected and lay their basis on similar principles. Spatial interpolation models can be derived from these techniques and can be classified into two categories: mechanical/deterministic and statistical/probability-based.

Mechanical models use arbitrary or empirical model parameters. They include techniques such as Thiessen polygons, Splines and Inverse Distance Interpolation. They do not provide for error estimation and usually have no strict assumptions about the variability of the features (Hengl, 2007).

In statistical/probability models, model parameters are estimated in an objective way, according to probability theory. They include four main groups, namely Kriging, environmental correlation, Bayesian-based models and hybrid models. Their main advantage is that an estimate of the prediction error is calculated along with the predicted values.

The performance of the two categories of spatial interpolation has been compared in various studies; for hydrological data, in particular, probability methods perform overall better (Varouchakis and Hristopulos, 2013b).

Kriging, which falls in the category of statistical/probability models, was selected for application in this study due to its close relation to natural sciences (Hengl, 2007). It is an exact approximation method which also provides an indicator of the accuracy of the estimated value by calculating the error variance (Delhomme, 1978; Theodossiou and Latinopoulos, 2006).

Kriging is a technique first published by Krige (1951), while Matheron (1963) derived the formulas and basically established the whole field of linear geostatistics (Cressie, 1990). The interpolation technique of Kriging has been applied mostly for the spatial estimation of variables rather than the temporal, even though applications on spatio-temporal estimations are also common in recent literature (Snepvangers et al., 2003; Ta′any et al., 2009). By taking into consideration the spatial distribution of data, Kriging overcomes problems that other interpolation methods, such as the nearest neighbor method, the distance weighted method and polynomial interpolation, may have. Basic concepts of the Kriging technique and its application to natural phenomena have been reviewed by the ASCE Task Committee on Geostatistical Techniques in Geohydrology (1990a, 1990b).

As mentioned before, Kriging has been used in hydrology and more specifically in groundwater applications. Delhomme (1978) first presented a series of case studies where Kriging was used in automatic contouring, in measurement network design and other applications. Aboufirassi and Mariño (1983) and Kumar (2006), amongst others, used various Kriging algorithms to estimate the groundwater level in different areas. The most commonly used Kriging methods are the Ordinary (Yang et al., 2008) and the Universal (Aboufirassi and Mariño, 1983). Other more accurate Kriging methods such as Kriging with external drift are also widely used. Auxiliary variables, such as digital elevation models or other geographical and physical parameters, can also be used in order to improve the accuracy of the estimations (Desbarats et al., 2002). Different methods of Kriging are compared in Marinoni (2003).

Kriging assumes that the closer the input data are, the more positively correlated the estimation errors are. The estimated value of parameter z at point x_k, $\bar{z}(x_k)$, is determined as the linear combination of N parameter samples $z(x_i)$ at nearby sampling points x_i:

$$\bar{z}(x_k) = \sum_{i=1}^{N} \lambda_i z(x_i) \tag{2}$$