

Received 24.05.2016
Reviewed 16.08.2016
Accepted 04.10.2016A – study design
B – data collection
C – statistical analysis
D – data interpretation
E – manuscript preparation
F – literature search

Assessing the suitability of extreme learning machines (ELM) for groundwater level prediction

Basant YADAV¹⁾ ABCDEF ✉, Sudheer CH²⁾ ABCDEF,
Shashi MATHUR¹⁾ ABCDEF, Jan ADAMOWSKI³⁾ ABCDEF

¹⁾ Indian Institute of Technology, Department of Civil Engineering, Hauz Khas, New Delhi 110 016, India; e-mail: basant1488@gmail.com

²⁾ Ministry of Environment, Forest and Climate Change; e-mail: sudheer108@gmail.com

³⁾ McGill University, Faculty of Agricultural and Environmental Sciences, Department of Bioresource Engineering, Quebec, Canada, H9X 3V9; e-mail: jan.adamowski@mcgill.ca

For citation: Yadav B., Ch S., Mathur S., Adamowski J. 2017. Assessing the suitability of extreme learning machines (ELM) for groundwater level prediction. *Journal of Water and Land Development*. No. 32 p. 103–112. DOI: 10.1515/jwld-2017-0012.

Abstract

Fluctuation of groundwater levels around the world is an important theme in hydrological research. Rising water demand, faulty irrigation practices, mismanagement of soil and uncontrolled exploitation of aquifers are some of the reasons why groundwater levels are fluctuating. In order to effectively manage groundwater resources, it is important to have accurate readings and forecasts of groundwater levels. Due to the uncertain and complex nature of groundwater systems, the development of soft computing techniques (data-driven models) in the field of hydrology has significant potential. This study employs two soft computing techniques, namely, extreme learning machine (ELM) and support vector machine (SVM) to forecast groundwater levels at two observation wells located in Canada. A monthly data set of eight years from 2006 to 2014 consisting of both hydrological and meteorological parameters (rainfall, temperature, evapotranspiration and groundwater level) was used for the comparative study of the models. These variables were used in various combinations for univariate and multivariate analysis of the models. The study demonstrates that the proposed ELM model has better forecasting ability compared to the SVM model for monthly groundwater level forecasting.

Key words: *extreme learning machine (ELM), forecasting, groundwater level, support vector machine (SVM), water resource management*

INTRODUCTION

Groundwater is an important source of freshwater worldwide and has a variety of domestic, industrial, and agricultural applications. It plays an integral role in supporting various types of ecosystems, particularly in arid climates [LI *et al.* 2014; PÉREZ- MARTIN *et al.* 2014; TODD, MAYS 2005]. This finite resource is often overexploited, resulting in dangerously low levels of groundwater around the world. As such, it is becoming increasingly important to improve the man-

agement of groundwater. MOHANTY *et al.* [2013] suggest that an overall groundwater management strategy depends on various factors, including availability of accurate information, financial support, policy framing and implementation. However, another key component of groundwater management is being able to forecast groundwater levels with a high degree of accuracy. The precise prediction of groundwater levels can help policymakers resolve the best approach to groundwater management problems [EMAMGHOLIZADEH *et al.* 2014].

Modelling of groundwater levels depends on various hydro-meteorological parameters; therefore, physically-based models have often been the preferred method to capture the complex and dynamic hydrogeological phenomena. However, insufficient amounts of accurate hydrogeological data as well as a lack of accuracy during data collection and pre-processing are common issues that arise when using physically-based models, particularly in underdeveloped and developing countries. This study investigates the suitability of extreme learning machines to forecast groundwater levels.

PREVIOUS RESEARCH

In the past decade, various studies have investigated the advantages and disadvantages of conceptual-based models. Subsequently, their performance has been compared with data-driven models such as artificial neural networks (ANN) [DALIAKOPOULOS 2005; MASKEY *et al.* 2000; MOHAMMADI 2008; MOHANTY *et al.* 2013]. A comparative study suggests that conceptual-based models require many parameters for calibration and have large computation times. In practice, however, poor model performance and associated uncertainties of such models can be attributed to data collection cost and time as well as inaccessibility of sites [KIM *et al.* 2005]. Data-driven models are able to develop interrelationships between input-output variables and generalize complex phenomena with high accuracy and minimal computation time [PLATT 1999]. ADAMOWSKI and CHAN [2011] brought attention to the relative importance of numerical models (for capturing the complex underlying processes) and data-driven models (for accurate prediction with a limited data set) when modelling a particular site.

Recently, the use of artificial intelligence (AI) approaches such as genetic programming (GP), artificial neural networks (ANN), and support vector machines (SVM) for use in water resources problems has gained popularity. GP has been successfully used in many water management problems and researchers have concluded that GP simulation equations decrease computational effort by using common simulation packages that can yield results with acceptable accuracy [FALLAH-MEHDIPOUR *et al.* 2013; 2014; RABUNAL *et al.* 2007; SAVIC *et al.* 1999]. ANNs have been successfully used in many fields and are capable of describing highly nonlinear and complex hydrological processes. The ANN approach to groundwater level prediction has been successfully applied by many researchers [ALMASRI, KALUARACHCHI 2005; BANERJEE *et al.* 2009; EMAMGHOLIZADEH *et al.* 2014; MOHANTY *et al.* 2013; SETHI *et al.* 2010; SREEKANTH *et al.* 2009; YOON *et al.* 2007].

The basis of support vector machines was developed by VAPNIK [1995]. The SVM approach overcomes common problems associated with ANNs (local minimum and the network over fitting) as it is based on structural risk minimization (SRM) instead

of the empirical risk minimization (ERM) concept of ANNs. This new concept in SVM has two unique features, including excellent generalization capability and sparse representation, resulting in a superior forecasting model when compared to traditional data-driven models. Furthermore, the SVM solution is always unique and global as its implementation requires the solution of a convex quadratic constrained optimization problem [SCHÖLKOPF, SMOLA 2002]. The application of SVM in water resources problems is recent and has been found to be more efficient than the traditional soft computing techniques [BEHZAD *et al.* 2009; SAFAVI, ESMIKHANI 2013; SURYANARAYANA *et al.* 2014; TRIPATHI *et al.* 2006]. Although SVM has been used successfully in many fields of research, its output depends on the choice of employing a suitable kernel function and its parameters. The hyper parameters of SVM are heuristic and generally selected by a trial and error process which can be time consuming.

HUANG *et al.* [2004] proposed a data-driven algorithm for a single-layer feed forward neural network (SLFN) known as extreme learning machine (ELM) which significantly reduces the computational time required for training a neural network. Studies using ELM have yielded very fast learning times with good generalization performance due to the fact that the ELM simplifies the entire training process [HUANG *et al.* 2004; LIANG *et al.* 2006; MOHAMMADI *et al.* 2015a, b; SHAMSHIRBAND *et al.* 2015a, b]. A recent study by HUANG *et al.* [2015] further showed that the ELM algorithm overcomes the problem of slow learning speed associated with traditional methods such as the back-propagation method and yields a better performance due to its ability to obtain the smallest training error and norm of weight. Thus, the ELM algorithm has gained popularity in various scientific fields such as the forecasting of coal mine water inrush [ZHAO *et al.* 2013], non-stationary time series prediction [WANG, HAN 2014], estimation of monsoon rainfall [ACHARYA *et al.* 2014], estimation of wind speed distribution [SHAMSHIRBAND *et al.* 2015a], and sales forecasting [SUN *et al.* 2008]. However, few studies exist that explore the predictive ability of ELM for groundwater level forecasting. NURHAYATI *et al.* [2013] conducted a study to test the use of ELM for forecasting groundwater levels on tidal lowlands in Indonesia. In this study, backward propagation ANN (BPANN) was used for validation, and the results showed that the training result and the groundwater prediction using ELM yielded better results than BPANN methods.

The focus of the present study is to explore the suitability of the ELM method to forecast monthly groundwater levels. The performance of ELM is further compared and validated using SVM, which has been found by many researchers to be superior to other methods such as ANN GP and neuro-fuzzy models [BEHZAD *et al.* 2009; HE *et al.* 2014; PRADHAN 2013; SAFAVI, ESMIKHANI 2013; SURYANARAYANA *et al.*

2014; YOON *et al.* 2011]. The goal of this comparative study is to test the use of ELM and SVM models to forecast groundwater levels at two on-site well locations with a one-month lead time.

To achieve the aforementioned goals, consideration is given to the following four different types of input parameters: mean monthly groundwater level (h), mean monthly temperature (T), monthly evapotranspiration (ET) and monthly precipitation (P). The performance assessment is conducted via three broadly utilized statistical indicators to attain reliable results.

METHODS

SUPPORT VECTOR MACHINE (SVM)

VAPNIK [1995] proposed a kernel-based algorithm as an SVM based on the Vapnik–Chervonenkis (VC) theory, which is now considered to be one of the leading techniques for pattern classification and function approximation. SVM can be used for applications in diverse scientific fields due to the fact that it has good generalization ability, is less prone to overfitting, and it allows for simultaneous minimization of estimation errors. SVMs use kernel functions that make the original inputs linearly separable in a mapped high dimensional feature space [QU, ZUO 2010].

However, the main advantage of SVM is that it not only has the strength of ANN, but at the same time it can overcome some of the widely discussed deficiencies associated with ANN [ASCE... 2000]. In SVM, the threshold of the network and the final set of optimal weights can be interpreted, unlike ANN where the threshold of the network and the final set of optimal weight cannot be interpreted. The network training in ANN is time-consuming and a multi-dimensional input structure provides more tunable parameters, unlike SVM, where an optimal structure is the solution of the quadratic optimization problem and the number of tunable parameters does not increase with the size of input variables. DIBIKE *et al.* [2001] suggested that SVM is more efficient than ANN in dealing with a multi-dimensional input.

The SVM equations are formulated as per Vapnik's theory, that if $\{(I_1, T_1), \dots, (I_N, T_N)\}$ are assumed as the given training data sets, where $I_k \in \mathcal{R}^n$ refers to the space of input variable, $T_k \in \mathcal{R}$ refers to the space of target value and N represents the length of the training data. The linear regression of SVM is estimated by solving [VAPNIK 1995] Eq. (1):

$$\text{minimize } \frac{1}{2} \|W\|^2 + C \sum_{k=0}^N (\xi + \xi^*) \quad (1)$$

Subject to

$$\begin{aligned} T_k - \langle w, I_k \rangle - b &\leq \varepsilon_k + \xi \\ \langle w, I_k \rangle + b - T_k &\leq \varepsilon_k + \xi^* \\ \xi_k, \xi_k^* &\geq 0, \quad k = 1, \dots, N \end{aligned}$$

where w denotes weight vector, b is a bias, C represents the regularization constant, ε is the error tolerance range of the function, and ξ, ξ^* are the slack variables. The schematic representation of SVM is depicted in Figure 1.

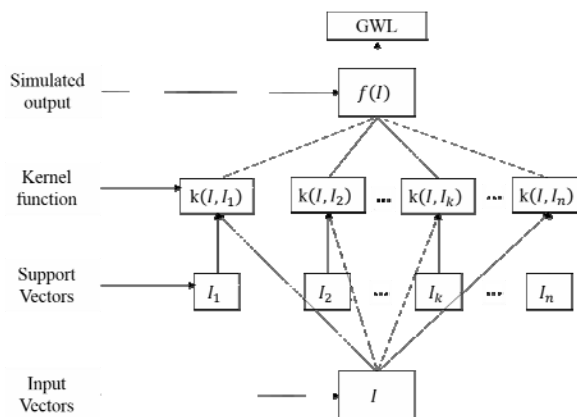


Fig. 1. Schematic representation of the process followed in SVM; source: own elaboration

EXTREME LEARNING MACHINE (ELM)

To overcome the deficiencies of traditional soft computing techniques, HUANG *et al.* [2004] proposed a simple three-layer structure algorithm, ELM. In the ELM structure the input weight (connection between input layer and hidden layer) and the bias values (in the hidden layer) are randomly generated. ELM analytically calculates the output weight matrix between hidden layers and output layers through a simple generalized inverse operation of the hidden layer output matrix. Generally, ELM has interpolation capability and universal approximation capability [HUANG *et al.* 2006], making ELM a promising time series prediction tool. Mathematically, ELM can be formulated as a function with L hidden nodes and N training samples, as follows [HUANG *et al.* 2006]:

$$\sum_{i=1}^L w_i g(W_{in(i)}, b_i, x_j) = \sum_{i=1}^L w_i g(W_{in(i)} \cdot x_j + b_i) = y_j \quad (2)$$

$$j = 1, \dots, N$$

where $x_j \in \mathcal{R}^n$ is the input vector, $W_{in(i)} \in \mathcal{R}^n$ is the input weight vector, $W_{in(i)} \cdot x_j$ represents the inner product of $W_{in(i)}$ and x_j , $b_i \in \mathcal{R}^n$ represents the bias of the i th hidden node, $g(\cdot)$ denotes the approximation function (sigmoid), $w_i \in \mathcal{R}^n$ is the output weight matrix and $y_j \in \mathcal{R}$ denotes the simulated output of ELM. In the ELM algorithm, the input weight and bias are randomly chosen at the initial stage.

Equation (2) can be written compactly in the form [HUANG *et al.* 2006]:

$$Aw = Y$$

$$A = \begin{bmatrix} g(W_{in(1)}, b_1, x_1) & \dots & g(W_{in(L)}, b_L, x_1) \\ \vdots & \dots & \vdots \\ g(W_{in(1)}, b_1, x_N) & \dots & g(W_{in(L)}, b_L, x_N) \end{bmatrix}_{N \times L}$$

where, $Y = [y_1, \dots, y_N]^T$, and $w = [w_1, w_2, \dots, w_L]^T$. Matrix A is called the hidden layer output matrix of ELM [HUANG *et al.* 2011]; the i th column of A is the i th hidden node's output vector with respect to inputs x_1, x_2, \dots, x_N and the j th row of A is the output vector of the hidden layer with respect to input x_j . If the ELM model with L hidden nodes is able to learn these N training samples with no residuals, w can be predicted, such that [HUANG *et al.* 2011]:

$$\sum_{i=1}^L w_i g(W_{in(i)} \cdot x_j + b_i) = t_j \quad j = 1, \dots, N \quad (3)$$

where t_j represents the target output. Eq. (3) can be further expressed as:

$$Aw = T \quad (4)$$

where $T = [t_1, \dots, t_N]^T$, is the target vector. The random selection of $W_{in(i)}$ and b_i convert Eq. (4) into a linear parameter system such that the minimum norm least squares solution of the linear parameter system can further be written as [HUANG *et al.* 2011]:

$$w = A^\dagger T$$

where A^\dagger is the Moore–Penrose generalized inverse of the hidden layer output matrix A . In practice, A^\dagger is calculated using the singular value decomposition (SVD), and then the non-zero singular values construct the output weights. However, when L and N are large, the computation complexity of the SVD decomposition impacts the learning speed of ELM immensely. The architecture of ELM is given in Figure 2.

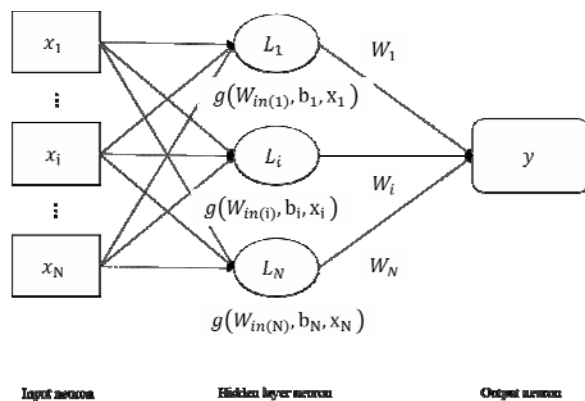


Fig. 2. Schematic representation of the process followed in ELM; source: own elaboration

STUDY AREA

The locations selected for this study are in British Columbia, Canada (Fig. 3). In the study area, approximately 1 million people are estimated to use groundwater supplied by hundreds of aquifers across the region. The study site is highly hydro-climatically complex, and hydrological parameters like precipitation, temperature, humidity are closely related to physical changes in the Pacific Ocean (i.e. currents, salinity, etc.). In Canada, precipitation varies from $4000 \text{ mm}\cdot\text{y}^{-1}$ (in the west) to $200 \text{ mm}\cdot\text{y}^{-1}$ in the rain

shadow regions of North America [MOORE, MCKENDRY 1996]. The coastal areas, like the selected study site, have a mild climate with summer daytime temperatures around 20°C and mild winters. The central regions in Canada experience a comparatively hot summer (July temperatures are 30°C or more) and a colder winter. The northern region is the coldest part with severe snowy winter and a short summer.

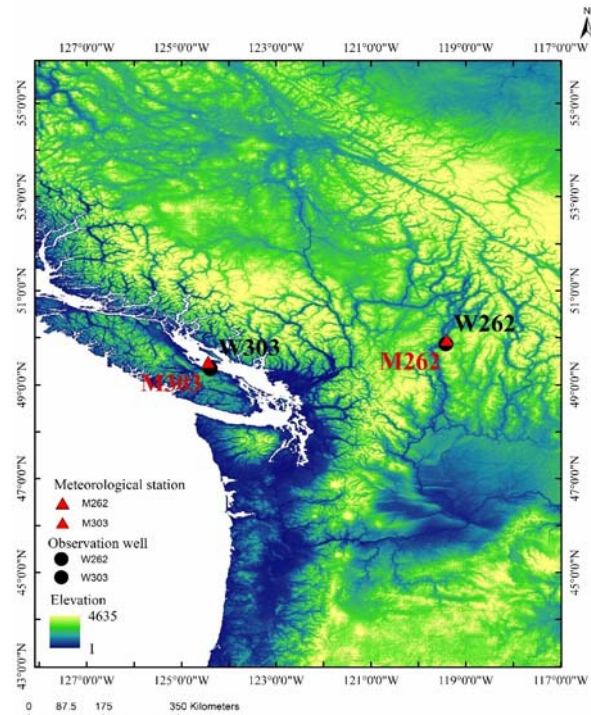


Fig. 3. Study area with selected well (W262 and W303) and meteorological stations (M262 and M303); source: own elaboration

Climate change also has an impact on various hydrological variables such as temperature, precipitation, evapotranspiration, and ultimately, runoff [ALLEN *et al.* 2014]. Changes to these variables also have a large impact on the groundwater recharge system. Moreover, at many sites, significant inter-annual variability in water levels is expected to have occurred due to natural and/or anthropogenic causes which can threaten water supplies for domestic consumption, hydroelectric generation and irrigation.

The well network in British Columbia was established for monitoring of groundwater levels over time in areas of high human use. The information on groundwater level fluctuations helps in understanding the aquifer properties, patterns of human use and climatic conditions of the area [GURDAK *et al.* 2009]. The well network in the study area has 119 groundwater level observation wells of which 78% have a stable or increasing water level, 9% have a large rate of decline (greater than $10 \text{ cm}\cdot\text{y}^{-1}$), and 13% show a moderate rate of decline (3 to $10 \text{ cm}\cdot\text{y}^{-1}$) [Ministry of Environment undated].

In this study, two sites which show a decline rate greater than $10 \text{ cm}\cdot\text{y}^{-1}$ in the water level were se-

lected. These wells are established in areas of human settlement, and collect data automatically to provide an understanding of how specific aquifers are replenished and how ongoing use is affecting water availability. The selected wells are part of an area where well density is very high and which represents a significant part of the aquifer. Well 262 covers the western part of the aquifer and well 303 covers the eastern part. The input parameters used in this study are monthly total rainfall (P), mean monthly temperature (T), monthly evapotranspiration (ET) and monthly average groundwater level (h). The information on precipitation and temperature was obtained from the Government of Canada [undated], whereas the evapotranspiration data was obtained from a non-profit association [Farmwest undated]. Furthermore, the data of groundwater levels from January 2006 to December 2014 were obtained online from Environment Canada's database.

MODEL DEVELOPMENT

PARAMETRIC ANALYSIS

In most of the previous studies conducted, a combination of input parameters was selected at random and the final parameters were obtained after a number of simulations. This study analyses the impact of individual input parameters on the accuracy of groundwater level forecasts. The methodology is based on the monthly lead time as it encompasses the long-term trend, which is fundamental to the resolution of many complex problems regarding the availability and sustainability of groundwater. Extensive amounts of monthly data are required to assess many different properties including: the effect of climate variability, monitoring and management of regional aquifer development, effects of groundwater withdrawals and other hydrologic stresses on groundwater availability, land subsidence, changes in groundwater quality, and surface water-groundwater interaction.

The ELM models are developed by considering groundwater levels for the current month (h_t) and previous month (h_{t-1}) while the other meteorological parameters (P_t , T_t , ET_t) are taken only for the current month. In the case of SVM, the input structure consists of groundwater levels for the current month (h_t) and previous month (h_{t-1}) and meteorological parameters (P_t , T_t , ET_t) for the current month as well as the values (P_{t-1} , T_{t-1} , ET_{t-1}) of the previous month.

DEVELOPMENT OF ELM MODELS

The ELM models for groundwater level prediction for both wells (W262 and W303) are developed using the MATLAB R2014a program which is used for both a univariate and multivariate analysis. The data set from January 2006 to December 2012 is used for training the model, while the data set from January 2013 to December 2014 is used for testing the devel-

oped ELM. The monthly data for all input variables are normalized in the range of $(-1, +1)$, keeping target values in the original form for ELM. In the case of the univariate analysis, only the past groundwater level data (h) is used as an input to the developed model, while in the multivariate analysis four variables (P , T , ET , and h) in various combinations are used as input data. In the development phase, the past data of each variable is considered from time lag t to $t-4$ month (current month to four previous months) on the basis of the developed correlation between the original and lagged data. The best ELM model is the model with an optimal number of hidden neurons (50 for both the wells) with highest values of statistical indicators among the various model input combinations and is developed for both the observation wells.

DEVELOPMENT OF SVM MODELS

LIBSVM toolbox [CHANG, LIN 2011] with MATLAB R2014a program is used for the development of the SVM models to forecast groundwater levels. The accuracy of the SVM model and its generalization ability depends on the selection of suitable kernels and its parameters. In this study, several SVM models for different input structures are developed and for each model, the kernel function (Radial Basis Function) is kept constant, while the model parameters (regularization constant, C ; insensitive loss function, ε ; and parameter of radial basis function, γ) are estimated by a trial and error method. The parameters C , ε , and γ for the best SVM model are estimated as 1.67, 0.38, 0.86 (for W262) and 1.44, 0.38, 0.018 (for W303), respectively. The developed SVM models utilized all input parameters (P , T , ET , and h) in different combinations with a time lag starting from the current month (t) to the previous four months ($t-4$). The best SVM model is the model with suitable kernel function, optimized model parameters (C , ε and γ), and the highest values of statistical indicators among the various model input combinations.

MODEL PERFORMANCE EVALUATION

The performance of the developed models is evaluated using several widely used statistical tests. The coefficient of determination (R^2) indicates the degree of correlation between two variables. It is a measure of how well the model is able to develop a relationship between observed and predicted variables. R^2 is given as:

$$\left(\frac{\sum_{i=1}^n (h_{oi} - \bar{h}_o)(h_{pi} - \bar{h}_p)}{\sqrt{\sum_{i=1}^n (h_{oi} - \bar{h}_o)^2} \sqrt{\sum_{i=1}^n (h_{pi} - \bar{h}_p)^2}} \right)^2 \quad (5)$$

Root mean square error ($RMSE$) indicates model performance by measuring the difference between observed and predicted values. $RMSE$ is given as:

$$\sqrt{\frac{\sum_{i=1}^n (h_{oi} - h_{pi})^2}{n}} \quad (6)$$

Nash–Sutcliffe efficiency criterion (NS) is used to assess the predictive ability of hydrological models. It is a normalized measure that compares the mean square error generated by the model to the variance of the target output. NS is given as:

$$\left(1 - \frac{\sum_{i=1}^n (h_{pi} - h_{oi})^2}{\sum_{i=1}^n (h_{oi} - \bar{h}_o)^2}\right) \quad (7)$$

where in all three equations, n is the number of data, h is the groundwater level (m) and the subscripts o and p denote the observed and predicted values, respectively.

RESULTS AND DISCUSSION

In this research work, an ELM model is employed to predict monthly groundwater level. To demonstrate the merit of the developed ELM model, its prediction performance is evaluated by providing comparisons with the SVM model. To assess the performance of the developed ELM and SVM models, a statistical comparison between the predicted and measured values of groundwater levels is performed using the statistical indicators: coefficient of determination (R^2), root mean square error ($RMSE$) and Nash–Sutcliffe efficiency criterion (NS).

The developed ELM and SVM models are evaluated by using different combinations of more correlated input variables (Table 1). For both study sites (W262 and W303), ELM models are more accurate in predicting the groundwater levels up to one month ahead compared to the SVM models. In the case of W262, the best ELM model for one month ahead prediction (h_{t+1}) has input variables of $\{h_t, h_{t-1}\}$ where h_{t+1} is the average groundwater level one month ahead, h_t is the groundwater level of the current month, and h_{t-1} is the groundwater level of the previous month. However, in the case of W303, the best ELM model is achieved when the current month temperature (T_t) is also included. The best SVM model for W262 is obtained when the average groundwater level (h) and monthly precipitation (P) are considered, and the input set is given as $\{h_t, h_{t-1}, P_t, P_{t-1}\}$, where P_t is the total precipitation of the current month and P_{t-1} is the total precipitation of the previous month. For W303, the best SVM model is achieved when the average temperature of the current month, T_t and average temperature of the previous month, T_{t-1} are included. Thus, the input set for the models is given as $\{h_t, h_{t-1}, P_t, P_{t-1}, T_t, T_{t-1}\}$. Moreover, for W303, groundwater levels were significantly impacted by seasonal variability in temperature as heavy snowfall during the winter lead to increased groundwater recharge as the snow melted.

Table 1. Performance of ELM and SVM with different hydrological and meteorological inputs for selected wells (W262 and W303)

Input set	R^2		$RMSE$		NS	
	W262	W303	W262	W303	W262	W303
ELM						
$\{h_t, h_{t-1}\}$	0.901	0.755	0.053	0.681	0.885	0.742
$\{h_t, h_{t-1}, P_t\}$	0.916	0.828	0.072	0.648	0.794	0.767
$\{h_t, h_{t-1}, T_t\}$	0.879	0.862	0.072	0.502	0.796	0.860
$\{h_t, h_{t-1}, ET_t\}$	0.882	0.800	0.085	0.734	0.715	0.700
SVM						
$\{h_t, h_{t-1}\}$	0.898	0.731	0.088	0.714	0.688	0.717
$\{h_t, h_{t-1}, P_t, P_{t-1}\}$	0.845	0.726	0.074	0.708	0.782	0.721
$\{h_t, h_{t-1}, P_t, P_{t-1}, T_t, T_{t-1}\}$	0.725	0.788	0.092	0.637	0.661	0.775
$\{h_t, h_{t-1}, ET_t, ET_{t-1}\}$	0.706	0.728	0.089	0.715	0.684	0.716

Explanations: R^2 = determination coefficient, $RMSE$ = root mean square error, NS = Nash–Sutcliffe efficiency criterion, h_t = groundwater level for the current month, h_{t-1} = groundwater level for the previous month, P_t = precipitation for the current month, P_{t-1} = precipitation for the previous month, T_t = temperature for the current month, T_{t-1} = temperature for the previous month, ET_t = evapotranspiration for the current month, ET_{t-1} = evapotranspiration for the previous month.

Source: own study.

The statistical analysis for both sites is presented in Table 2. The best ELM models for the W262 and W303 sites have testing $RMSE$ values of 0.053 m and 0.502 m, respectively, and are superior to the best SVM models with values of 0.074 m and 0.637 m, respectively. Lower $RMSE$ values indicate that the best ELM model can successfully generalize the complex system close to the observed data. This argument is further supported by the general statistics given in Table 3.

Table 2. Performance of best ELM and SVM models for groundwater level prediction for well 262 and well 303 for a one-month lead period

Model	Well number	R^2	$RMSE$	NS
ELM	W262	0.901	0.053	0.885
	W303	0.862	0.502	0.860
SVM	W262	0.845	0.074	0.782
	W303	0.788	0.637	0.775

Explanations: R^2 , $RMSE$, NS as under Table 1.

Source: own study.

The best ELM model for W262 and W303 has R^2 values in the testing phase of 0.901 and 0.862, respectively, and simulates better than the best SVM model with R^2 values of 0.845 and 0.788 respectively. Similarly, the best ELM model for W262 and W303 has a Nash efficiency of 0.885 and 0.860, respectively, and performs better than the best SVM model, which has a Nash efficiency of 0.782 and 0.775, respectively. The higher R^2 and Nash efficiency values indicate that the ELM model has a greater ability to predict and generalize complex nonlinear systems when compared to the best SVM model for both wells.

Table 3. General statistics of the observed and simulated groundwater level for well 262 and 303

Well number	Statistics	Observed time series	ELM model output	SVM model output
262	min	56.83	56.82	56.80
	max	57.33	57.31	57.17
	mean	57.02	57.02	56.99
	median	56.99	57.03	56.98
	mode	56.86	56.85	56.80
	SD	0.1642	0.1578	0.1225
303	range	0.500	0.494	0.378
	min	21.77	21.49	22.01
	max	26.82	26.10	26.34
	mean	24.35	24.29	24.27
	median	24.32	24.42	24.37
	mode	24.52	21.49	22.01
	SD	1.372	1.246	1.231
range	5.050	4.608	4.328	

Source: own study.

Likewise, Table 3 presents the comparison between the prediction ability of ELM and SVM for W262 and W303. The best SVM model overestimates the target values for both wells, while the best ELM model simulates the target output closer to the observed groundwater level at both locations.

To graphically analyse the capability of the developed ELM and SVM models for the two wells, the predicted values are plotted against the measured data. Figure 4 compares observed and predicted groundwater levels in the testing phase at W262 for the best ELM and SVM models. Figure 5 makes the same comparison for the W303 site. The best SVM model

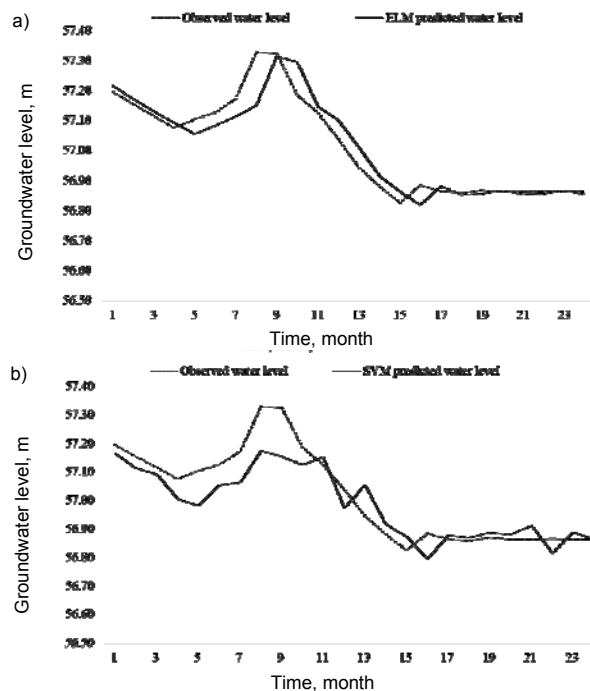


Fig. 4. The predicted groundwater level curve by best ELM model (a) and SVM model (b) and observed groundwater level curves for one-month ahead prediction at well 262 during testing period; source: own study

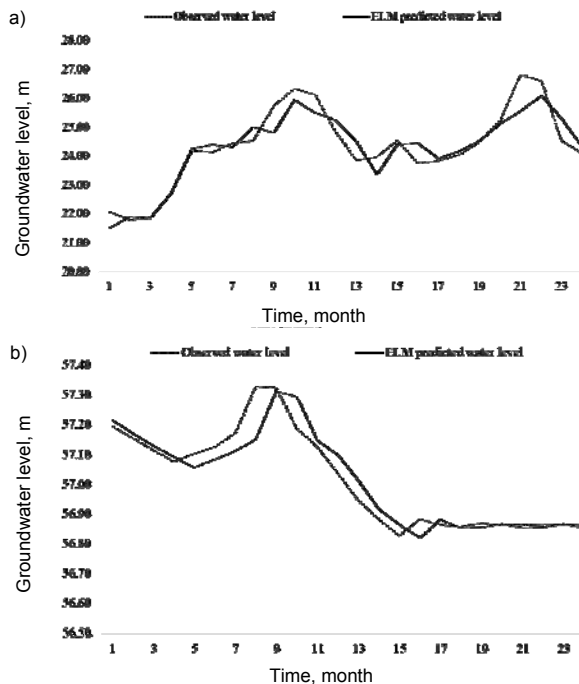


Fig. 5. The predicted groundwater level curve by best ELM model (a) and SVM model (b) and observed groundwater level curves for one-month ahead prediction at well 303 during testing period; source: own study

overestimates the target values for both wells, while the best ELM model simulates the target output closer to the observed groundwater level at both locations.

Further, the line of best fit of predicted groundwater level by ELM and SVM against the measured data for the testing phase are illustrated, respectively, in Figures 6 and 7 for W262 and W303. The plots clearly show that SVM overestimates/underestimates many points for both the wells and ELM predicts the groundwater level with a lower degree of overestimation/underestimation. It is also observed that the ELM model is able to capture the underlying dynamics of groundwater level variations and predicts the lower and higher groundwater level values closely to observed values. In the case of SVM, lower values of groundwater level are estimated accurately, however the higher groundwater level values are overestimated significantly.

The overestimation/underestimation of the target output by the ELM/SVM can be explained by the fact that the simulation ability of data-based models directly depends on input data. The selected input structure may not be providing the sufficient information to the model and hence the prediction accuracy is low. The magnification of error in the predicted groundwater level values can also be attributed to the structural shortcomings of the data-based models. SVM performance depends on the proper selection of hyper parameters, which are selected via a trial and error method. Likewise, the activation function (sigmoid) in ELM can have a somewhat unstable performance if the random numbers are not selected properly.

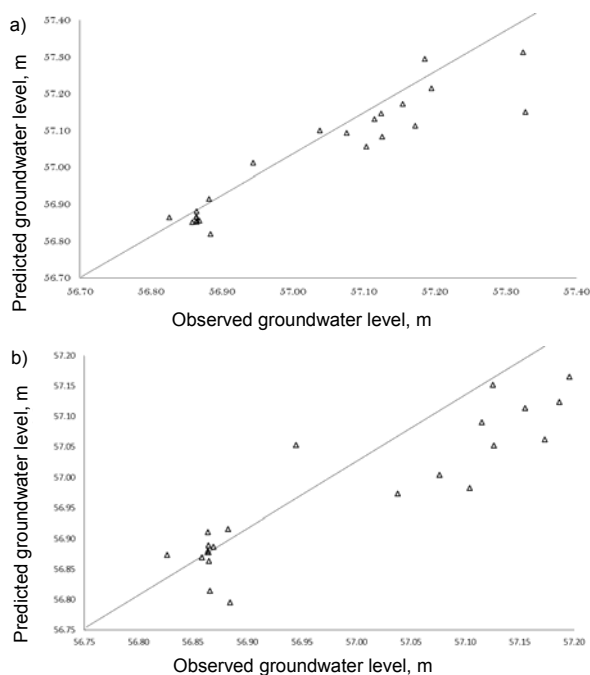


Fig. 6. Scatter plot of W262 comparing observed and simulated target values by best ELM model (a) and SVM model (b) in testing phase for one-month lead time; source: own study

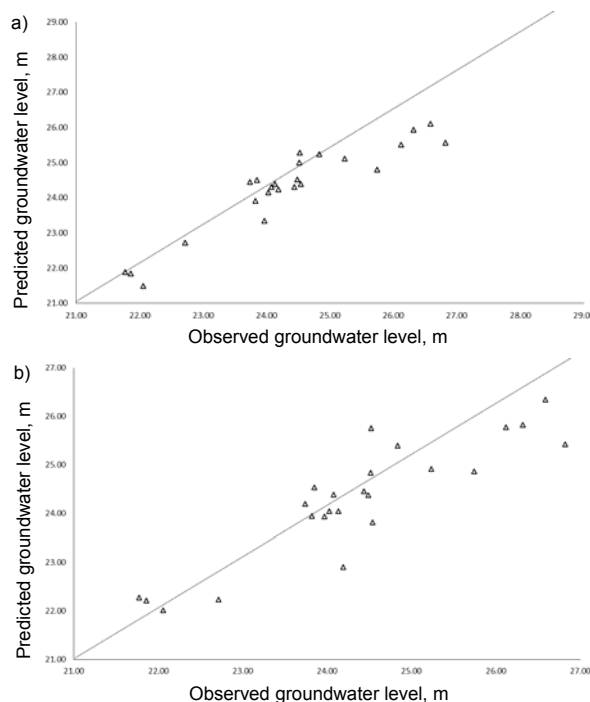


Fig. 7. Scatter plot of W303 comparing observed and simulated target values by best ELM model (a) and SVM model (b) in testing phase for one-month lead time; source: own study

CONCLUSION

This study evaluated the ability of a new extreme learning machine method to predict monthly groundwater levels. Two observation wells (W262 and

W303) located in British Columbia, Canada were studied. Hydrological and meteorological parameters for both sites were used in various input structures for the comparison study between SVM and ELM soft computing methods. The parametric analysis indicated that past groundwater level data are very important for accurate predictions, however, the selection of other parameters depends on the location of the study site. The numerical and graphical analyses indicate that the ELM algorithm is capable of more precise monthly forecasts of groundwater levels and displays its superiority over SVM models. Future research efforts should be directed towards using ELM models to predict groundwater levels at different lead times and at different geographical locations with varying physical characteristics. Forecasting groundwater levels with greater lead times will help policymakers to achieve better groundwater management planning.

REFERENCES

- ACHARYA N., SHRIVASTAVA N.A., PANIGRAHI B.K., MOHANTY U.C. 2014. Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine. *Climate Dynamics*. Vol. 43. Iss. 5–6 p. 1303–1310.
- ADAMOWSKI J., CHAN H.F. 2011. A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology*. Vol. 407. Iss. 1 p. 28–40.
- ALLEN D.M., STAHL K., WHITFIELD P.H., MOORE R.D. 2014. Trends in groundwater levels in British Columbia. *Canadian Water Resources Journal*. Vol. 39. Iss. 1 p. 15–31. DOI: 10.1080/07011784.2014.885677.
- ALMASRI M.N., KALUARACHCHI J.J. 2005. Modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrate loading and recharge data. *Environmental Modelling and Software*. Vol. 20. Iss. 7 p. 851–871.
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology 2000. Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of Hydrologic Engineering*. Vol. 5. Iss. 2 p. 115–123.
- BANERJEE P.R., PRASAD K., SINGH V.S. 2009. Forecasting of groundwater level in hard rock region using artificial neural network. *Environmental Geology*. Vol. 58. Iss. 6 p. 1239–1246.
- BEHZAD M., ASGHARI K., COPPOLA Jr E.A. 2009. Comparative study of SVMs and ANNs in aquifer water level prediction. *Journal of Computing in Civil Engineering*. Vol. 24. Iss. 5 p. 408–413.
- CHANG C.C., LIN C.J. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*. Vol. 2. Iss. 3. No. 27.
- DALIAKOPOULOS I., COULIBALYA P., TSANI I.K. 2005. Groundwater level forecasting using artificial neural network. *Journal of Hydrology*. Vol. 309. Iss. 1–4 p. 229–240.
- DIBIKE Y.B., VELICKOV S., SOLOMATINE D., ABBOTT M.B. 2001. Model induction with support vector machines: introduction and applications. *Journal of Computing in Civil Engineering*. Vol. 15 p. 208–216.
- EMAMGHOLIZADEH S., MOSLEMI K., KARAMI G. 2014. Prediction the groundwater level of Bastam Plain (Iran) by artificial neural network (ANN) and adaptive neuro-

- fuzzy inference system (ANFIS). *Water Resources Management*. Vol. 28. Iss. 15 p. 5433–5446.
- FALLAH-MEHDIPOUR E., HADDAD O.B., MARIÑO M.A. 2013. Prediction and simulation of monthly groundwater levels by genetic programming. *Journal of Hydro-environment Research*. Vol. 7. Iss. 4 p. 253–260.
- FALLAH-MEHDIPOUR E., HADDAD O.B., MARIÑO M.A. 2014. Genetic programming in groundwater modeling. *Journal of Hydrologic Engineering*. Vol. 19. Iss. 12 04014031.
- Farmwest undated. Evapotranspiration [online]. [Access 20.03.2016]. Available at: www.farmwest.com
- Government of Canada undated. Weather information [online]. [Access 20.03.2016]. Available at: www.weather.gc.ca
- GURDAK J.S., HANSON R.T., GREEN T.R. 2009. Effects of climate variability and change on groundwater resources of the United States. *United States Geological Survey. Fact Sheet 2009–3074*.
- HE Z., WEN X., LIU H., DU J. 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *Journal of Hydrology*. Vol. 509 p. 379–386.
- HUANG G., HUANG G.B., SONG S., YOU K. 2015. Trends in extreme learning machines: A review. *Neural Network*. Vol. 61 p. 32–48.
- HUANG G.B., WANG D.H., LAN Y. 2011. Extreme learning machines: A survey. *International Journal of Machine Learning and Cybernetics*. Vol. 2. Iss. 2 p. 107–122.
- HUANG G.B., ZHU Q.Y., SIEW C.K. 2004. Extreme learning machine: A new learning scheme of feedforward neural networks. In: *Proceedings of the international joint conference on neural networks*. 25–29 July 2004. Budapest, Hungary. Vol. 2 p. 985–990.
- HUANG G.B., ZHU Q.Y., SIEW C.K. 2006. Extreme learning machine: Theory and applications. *Neurocomputing*. Vol. 70 p. 489–501.
- KIM S.J., HYUN Y., LEE K.K. 2005. Time series modeling for evaluation of groundwater discharge rates into an urban subway system. *Geoscience Journal*. Vol. 9. Iss. 1 p. 15–22.
- LI F., QIAO J., ZHAO Y., ZHANG W. 2014. Risk assessment of groundwater and its application. Part II. Using a groundwater risk maps to determine control levels of the groundwater. *Water Resources Management*. Vol. 28. Iss. 13 p. 4875–4893.
- LIANG N.Y., HUANG G.B., RONG H.J., SARATCHANDRAN P., SUNDARARAJAN N. 2006. A fast and accurate on-line sequential learning algorithm for feedforward networks. *IEEE Transactions Neural Networks*. No. 17 p. 1411–1423.
- MASKEY S., DIBIKE Y.B., JONOSKI A., SOLOMATINE D. 2000. Groundwater model approximation with artificial neural network for selecting optimal pumping strategy for plume removal. In: *Proceedings Workshop 2nd Joint Artificial Intelligence in Civil Engineering Applications*. Eds. O. Schleider, A. Zijderveld. March 2000. Cottbus, Germany p. 67–80.
- Ministry of Environment undated. [online]. [Access 20.03.2016]. Available at: www.env.gov.bc.ca
- MOHAMMADI K. 2008. Groundwater table estimation using MODFLOW and artificial neural networks. *Water Science and Technology Library*. Vol. 68. Iss. 2 p. 127–138.
- MOHAMMADI K., SHAMSHIRBAND S., MOTAMEDI S., PETKOVIĆ D., HASHIM R., GOCIC M. 2015a. Extreme learning machine based prediction of daily dew point temperature. *Computers and Electronics in Agriculture*. Vol. 117 p. 214–225.
- MOHAMMADI K., SHAMSHIRBAND S., YEE P.L., PETKOVIĆ D., ZAMANI M., CH S. 2015b. Predicting the wind power density based upon extreme learning machine. *Energy*. Vol. 86 p. 232–239.
- MOHANTY S., JHA M.K., KUMAR A., PANDA D.K. 2013. Comparative evaluation of numerical model and artificial neural network for simulating groundwater flow in Kathajodi–Surua Inter-basin of Odisha, India. *Journal of Hydrology*. Vol. 495 p. 38–51.
- MOORE R.D., MCKENDRY I.G. 1996. Spring snowpack anomaly patterns and winter climatic variability, British Columbia, Canada. *Water Resources Research*. Vol. 32. Iss. 3 p. 623–632.
- NURHAYATI, SOEKARNO I., HADIHARAJA I.K., CAHYONO M. 2013. The prediction of groundwater level on tidal lowlands reclamation using extreme learning machine. *Journal of Theoretical and Applied Information Technology*. Vol. 56. Iss. 1 p. 75–84.
- PÉREZ-MARTÍN M.A., ESTRELA T., ANDREU J., FERRER J. 2014. Modeling water resources and river-aquifer interaction in the Júcar River Basin, Spain. *Water Resources Management*. Vol. 28. Iss. 12 p. 4337–4358.
- PLATT J.C. 1999. Fast training of support vector machines using sequential minimal optimization. In: *Advances in kernel methods: Support vector learning*. Eds. B. Schölkopf, Ch.J.C. Burges, A.J. Smola. Cambridge. MIT Press p. 185–208.
- PRADHAN B. 2013. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Computers and Geosciences*. Vol. 51 p. 350–365.
- QU J., ZUO M.J. 2010. Support vector machine based data processing algorithm for wear degree classification of slurry pump systems. *Measurement*. Vol. 43. Iss. 6 p. 781–791.
- RABUNAL J.R., PUERTAS J., SUAREZ J., RIVERO D. 2007. Determination of the unit hydrograph of a typical urban basin genetic programming and artificial neural networks. *Hydrological Processes*. Vol. 21. Iss. 4 p. 476–485.
- SAFAVI H.R., ESMIKHANI M. 2013. Conjunctive use of surface water and groundwater: Application of support vector machines (SVMs) and genetic algorithms. *Water Resources Management*. Vol. 27. Iss. 7 p. 2623–2644.
- SAVIC D.A., WALTERS G.A., DAVIDSON J.W. 1999. A genetic programming approach to rainfall-runoff modeling. *Water Resources Management*. Vol. 13. Iss. 3 p. 219–231.
- SCHÖLKOPF B., SMOLA A.J. 2002. *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. Cambridge. MIT Press. ISBN 9780262194754 pp. 648.
- SETHI R.R., KUMAR A., SHARMA S.P., VERMA H.C. 2010. Prediction of water table depth in a hard rock basin by using artificial neural network. *International Journal of Water Resources and Environmental Engineering*. Vol. 2. Iss. 4 p. 95–102.
- SHAMSHIRBAND S., MOHAMMADI K., TONG C.W., PETKOVIĆ D., PORCU E., MOSTAFAEIPOUR A., CH S., SEDAGHAT A. 2015a. Application of extreme learning machine for estimation of wind speed distribution. *Climate Dynamics*. Vol. 46. Iss. 5 p. 1893–1907.
- SHAMSHIRBAND S., MOHAMMADI K., YEE L., PETKOVIĆ D., MOSTAFAEIPOUR A. 2015b. A comparative evaluation

- for identifying the suitability of extreme learning machine to predict horizontal global solar radiation. *Renewable and Sustainable Energy Reviews*. Vol. 52 p. 1031–1042.
- SREEKANTH P., GEETHANJALI D.N., SREEDEVI P.D., AHMED S., KUMAR N.R., JAYANTHI P.D.K. 2009. Forecasting groundwater level using artificial neural networks. *Current Science*. Vol. 96. Iss. 7 p. 933–939.
- SUN Z.L., CHOI T.M., AU K.F., YU Y. 2008. Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Systems*. Vol. 46. Iss. 1 p. 411–419.
- SURYANARAYANA C., SUDHEER C., MAHAMMOOD V., PANIGRAHI B.K. 2014. An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India. *Neurocomputing*. Vol. 145 p. 324–335.
- TODD D.K., MAYS L.W. 2005. *Groundwater hydrology*. Third Revision. Danvers, MA. John Wiley and Sons Inc. ISBN 9780471059370 pp. 636.
- TRIPATHI S., SRINIVAS V.V., NANJUNDIAH R.S. 2006. Down-scaling of precipitation for climate change scenarios: A support vector machine approach. *Journal of Hydrology*. Vol. 330. Iss. 3 p. 621–640.
- VAPNIK V.N. 1995. *The nature of statistical learning theory*. New York, USA. Springer-Verlag pp. 314.
- WANG X., HAN M. 2014. Online sequential extreme learning machine with kernels for non-stationary time series prediction. *Neurocomputing*. Vol. 145 p. 90–97.
- YOON H., HYUN Y., LEE K.K. 2007. Forecasting solute breakthrough curves through the unsaturated zone using artificial neural networks. *Journal of Hydrology*. Vol. 335 p. 68–77.
- YOON H., JUN S.C., HYUN Y., BAE G.O., LEE K.K. 2011. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*. Vol. 396. Iss. 1 p. 128–138.
- ZHAO Z., LI P., XU X. 2013. Forecasting model of coal mine water inrush based on extreme learning machine. *Applied Mathematics and Information Sciences*. Vol. 7 p. 1243–1250.

Basant YADAV, Sudheer CH, Shashi MATHUR, Jan ADAMOWSKI

Ocena zdolności ekstremalnych maszyn uczących (ELM) do przewidywania poziomu wód gruntowych

STRESZCZENIE

Na całym świecie fluktuacje poziomów wód gruntowych stanowią ważny temat badań hydrologicznych. Rosnące potrzeby wodne, błędne praktyki irygacyjne, niewłaściwa gospodarka glebowa i niekontrolowana eksploatacja poziomów wodonośnych są powodami, dla których poziom wód gruntowych podlega fluktuacjom. Dla skutecznego zarządzania zasobami wód gruntowych istotne jest dysponowanie dokładnymi zapiskami i zdolność prognozowania poziomu tych wód. Rozwój technik komputerowych (modele wykorzystujące dane) w dziedzinie hydrologii ma istotny potencjał z powodu niepewnego i złożonego charakteru systemów wód gruntowych. W prezentowanych badaniach wykorzystano dwie techniki komputerowe: maszynę uczenia ekstremalnego (ELM) i maszynę wektorów nośnych (SVM – ang. support vector machine) do przewidywania poziomów wód gruntowych w dwóch studzienkach obserwacyjnych w Kanadzie. Do porównawczych badań modeli wykorzystano zestaw danych miesięcznych z ośmiu lat (2006–2014), składający się z danych hydrologicznych i meteorologicznych (opady, temperatura, ewapotranspiracja, poziom wody). Wymienione zmienne zastosowano w rozmaitych kombinacjach do jedno- i wieloparametrycznej analizy modeli. Wyniki dowodzą, że model ELM ma lepsze zdolności przewidywania miesięcznych poziomów wód gruntowych w porównaniu z modelem SVM.

Słowa kluczowe: *maszyna uczenia ekstremalnego (ELM), maszyna wektorów nośnych (SVM), poziom wód gruntowych, prognozowanie, zarządzanie zasobami wodnymi*