Computers & Geosciences 90 (2016) 144–155

Contents lists available at ScienceDirect

Computers & Geosciences

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

Case study

A method to improve the stability and accuracy of ANN- and SVM-based time series models for long-term groundwater level predictions

Heesung Yoon [a], Yunjung Hyun [b,*], Kyoochul Ha [a], Kang-Kun Lee [c], Gyoo-Bum Kim [d]

[a] Geologic Environment Division, Korea Institute of Geoscience and Mineral Resources, Republic of Korea
[b] Korea Environment Institute, 370 Sicheong-daero, Sejong, Republic of Korea
[c] School of Earth and Environmental Sciences, Seoul National University, Republic of Korea
[d] Daejeon University, Republic of Korea

ARTICLE INFO

Article history:
Received 24 January 2014
Received in revised form
16 October 2015
Accepted 1 March 2016
Available online 3 March 2016

Keywords:
Groundwater level
Artificial neural network
Support vector machine
Recursive prediction
Weighted error function

ABSTRACT

The prediction of long-term groundwater level fluctuations is necessary to effectively manage groundwater resources and to assess the effects of changes in rainfall patterns on groundwater resources. In the present study, a weighted error function approach was utilised to improve the performance of artificial neural network (ANN)- and support vector machine (SVM)-based recursive prediction models for the long-term prediction of groundwater levels in response to rainfall. The developed time series models were applied to groundwater level data from 5 groundwater-monitoring stations in South Korea. The results demonstrated that the weighted error function approach can improve the stability and accuracy of recursive prediction models, especially for ANN models. The comparison of the model performance showed that the recursive prediction performance of the SVM was superior to the performance of the ANN in this case study.

1. Introduction

Groundwater is a valuable water resource for various human activities. Therefore, groundwater should be systematically managed for its sustainable use. Rainfall is the primary driver of groundwater level (GWL) fluctuations, and recent climate change studies predict changes in rainfall patterns (IPCC (International Panel on Climate Change) (2007); NIMR, 2009). Predictions of long-term groundwater level fluctuations are necessary to effectively manage groundwater resources and to assess the effects of rainfall pattern changes on groundwater resources.

Physics-based numerical models have generally been developed and applied to predict GWL fluctuations. These models establish mathematical governing equations based on physical concepts, and the equations are solved using numerical techniques for given domains (Rai and Singh, 1995; Knotters and Bierkens, 2000; Park and Parker, 2008). Physics-based numerical models are robust and powerful tools for simulating spatial and temporal variations in GWLs. However, this approach requires a wide variety and quantity of information regarding the physical properties of the domain and boundary conditions, a lack of which can cause poor model performance or can increase the model uncertainty through equifinality (Bardossy, 2007; Pollacco et al., 2008).

Recently, as the quantity and quality of automatic monitoring system data have steadily increased and improved, studies developing time series models for hydrologic problems using data-based learning algorithms, such as artificial neural networks (ANNs) and support vector machines (SVMs), have increased. The number of studies applying ANN-based time series models has considerably increased since the 1990s, mainly to address surface water problems, including modelling of rainfall (French et al., 1992; Kuligowski and Barros, 1998; Hung et al., 2009), stream flow (Karunanithi et al., 1994; Zealand et al., 1999; Campolo et al., 1999; Hu et al., 2005; Akhtar et al., 2009), and water quality (Maier and Dandy, 1996; Lek et al., 1999; Dogan et al., 2009). The application of ANNs for predicting GWL fluctuations has been emerging since the late 1990s (Coulibaly et al., 2001; Coppola et al., 2005; Dialiakopoulos et al., 2005; Giustolisi and Simeone, 2006; Krishna et al., 2008; Mohanty et al., 2010). SVMs, which are relatively new data-based learning algorithms and were introduced by Vapnik (1995), have emerged as an alternative method in ANN-dominated hydrologic research fields. Most SVM applications have been focused on surface water problems (Dibike et al., 2001; Liong and Sivapragasam, 2002; Asefa et al., 2005, 2006; Yu et al., 2006; Khalil et al., 2006; Khan and Coulibaly, 2006); however, the applicability of SVMs has been recently extended to GWL predictions (Asefa et al., 2004; Gill et al., 2007; Behzad et al., 2010). Yoon et al. (2011) applied ANN and SVM to the prediction of GWL fluctuations with

* Corresponding author.
E-mail address: yjhyun@kei.re.kr (Y. Hyun).

http://dx.doi.org/10.1016/j.cageo.2016.03.002
0098-3004/

direct prediction strategy. They concluded that the SVM model performance was better than ANN in their case studies; however, the prediction error increased significantly with lead times, which hinders a long term prediction of GWL fluctuations.

Long-term predictions have been a challenge in the field of time series model development. Several different methodologies have been developed to address this issue; the primary methods are direct prediction and recursive prediction (Ji et al., 2005; Herrera et al., 2007; Sorjamaa et al., 2007). The direct prediction strategy always uses real past measurement data for a target variable as inputs. Thus, there is no prediction error accumulation as the time increases, which can improve model performance. The limitation of the direct prediction strategy is that it uses different models for each prediction horizon, which can considerably increase the computational load. In contrast, the recursive prediction strategy repeatedly uses the same model with previous prediction values as inputs to estimate the next prediction. However, the prediction error at each time step can be cumulative and can deteriorate the model performance as the time increases.

In this study, ANN- and SVM-based recursive prediction time series models were developed for the long-term prediction of GWLs in response to rainfall; a weighted error function approach was utilised to improve the stability and accuracy of the time series models. The developed models were applied to daily GWL data from 5 groundwater stations in South Korea to evaluate the effects of the weighted error function on the model performances.

The paper is organized as follows: Section 2 describes the development of the time series model. Prediction strategies for the long-term prediction of the GWL are discussed, the weighted error function approach is proposed, and the structure of the developed time series models is described. Section 3 describes study site and time series data. Application results of the proposed method are presented in Section 4, and conclusions are drawn in Section 5.

2. Time series model development

For the establishment of the time series models, ANN structure trained by back-propagation algorithm (BPA) (Rumelhart, 1984) and SVM by sequential minimal optimization (SMO) algorithm (Platt, 1999) were employed. Theoretical backgrounds of ANN and SVM are described in Appendix A and B, respectively. The model building process of the ANN and the SVM consists of training and calibration stages. In the training stage, weights and biases values are updated through the learning algorithms for given model parameter sets. The best model parameter sets for the ANN and the SVM are selected in the calibration stage. The performances of the selected ANN and SVM models are examined in the validation stage. For this modelling procedure all the data are divided into three parts corresponding to the training, calibration, and validation stages.

2.1. Prediction strategy

Three types of strategy can be taken into consideration for the prediction of GWL fluctuations using time series models: prediction strategy using only rainfall data as input components, direct and recursive prediction strategies using rainfall and GWL data as input components.

The prediction strategy using only rainfall data as input components is expressed as

$$\hat{g}_{t+h} = F_h{}^{Pa}(\mathbf{x}), \quad \mathbf{x} = \{p_{t-a+1}, p_{t-a+2}, \dots, p_t\}, \tag{1}$$

where F_h^{Pa} is the time series estimator based on rainfall inputs for a prediction horizon h, $\hat{g}$ is a predicted GWL value, p is the rainfall amount, and a is the lag time of rainfall inputs, which determines the size of the input structure. This type of model does not consider the autoregressive property embedded in the GWL time series data. Thus, the model cannot properly estimate the groundwater recession behaviour. Moreover, the number of input components should be greater than the maximum number of days without rain. Otherwise, the predicted values will be identical after the number of continuously dry days exceeds h. This type of prediction strategy, therefore, is not appropriate for the purpose of the present study.

The direct prediction strategy can be expressed as

$$\hat{g}_{t+h} = F_h{}^{PaGb}(\mathbf{x}),$$
$$\mathbf{x} = \{p_{t-a+1}, p_{t-a+2}, \dots, p_t, g_{t-b+1}, g_{t-b+2}, \dots, g_t\}, \tag{2}$$

where F_h^{PaGb} is the time series estimator based on rainfall and GWL inputs for a prediction horizon h, g is a GWL value and a and b are lag times for rainfall and GWL inputs, respectively. The direct prediction strategy uses previously measured GWL data as inputs. Therefore, the errors in the predicted values do not accumulate in the next prediction, which can increase the model accuracy (Sorjamaa et al., 2007). However, in the direct prediction strategy, different models are required for every prediction horizon, which causes computational burden. Moreover, the model performances tend to decrease with the prediction horizon, which inhibits long-term predictions. The direct prediction strategy, therefore, is not adequate for the long-term prediction of GWL.

The recursive prediction strategy, which uses a one-step ahead direct prediction model, can be expressed as

$$\hat{g}_{t+1} = F_1{}^{PaGb}(\mathbf{x}),$$
$$\mathbf{x} = \{p_{t-a+1}, p_{t-a+2}, \dots, p_t, \hat{g}_{t-b+1}, \hat{g}_{t-b+2}, \dots, \hat{g}_t\}, \tag{3}$$

where $F_1{}^{PaGb}$ is the time series estimator with a prediction horizon of one time step. The recursive prediction strategy appears adequate for the long-term prediction problem addressed in this study because it repeatedly uses a given time series model. However, the error from previous prediction horizons can be accumulated and transmitted to further horizons, which can cause substantial deterioration in the model performance (Herrera et al., 2007); namely, even if an excellent one-step ahead direct prediction model is constructed, it is possible that the model is not adequate for recursive prediction and will result in increasing errors as the number of prediction horizons increases. Some preliminary experiments with this method have shown that this phenomenon is more frequent and substantial for the ANN modelling approach. The ANN, based on the empirical risk minimization (ERM), is more vulnerable to overfitting and susceptible to considering noise and error in the data as a pattern in comparison with the SVM based on the structural risk minimization (SRM) (Cimen and Kisi, 2009; Deng et al., 2011; Malekmohamadi et al., 2011). It is possible that this feature of the ANN intensifies the deterioration of the recursive prediction model performance.

To cope with the problem in the recursive prediction strategy, it is necessary and important to select an adequate one-step ahead direct prediction model that is capable of understanding the overall relationship between rainfall and GWL fluctuations.

2.2. A weighted error function approach

In this study, a simple method using a weighted error function is suggested to improve the stability and accuracy of the recursive prediction model. Few studies have utilised weighted error functions for the recursive prediction using time series models based on machine learning techniques such as ANN and SVM. Jung and Kwon (2013) proposed ANN models trained by weighted error

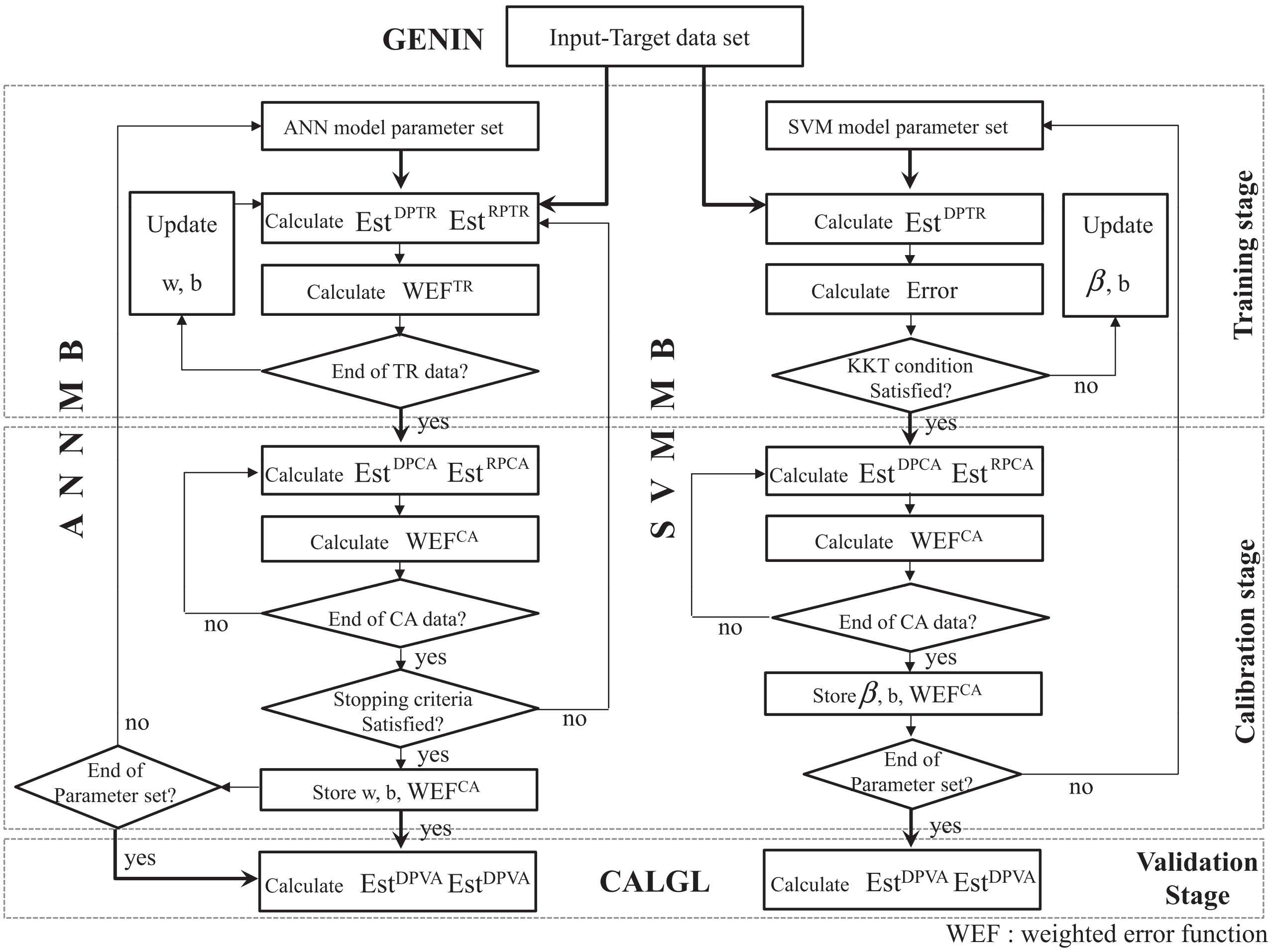

Fig. 1. A schematic diagram of the model structure used in this study: (a) ANN and (b) SVM.

functions of which weights are determined from the frequency of wind speed and the power performance curve. Their method is focused on the wind energy potential estimation. Thus it can be hardly applied to the prediction of the GWL. To the best of our knowledge, there has been no research on improving the stability and accuracy of the recursive prediction model based on machine learning algorithms for the long-term prediction of the GWL.

The conventional model parameter selection rule for the ANN and the SVM is to locate a parameter set that minimizes the sum of the squared errors (SSE) in the calibration stage:

$$\text{minimize} \quad E^{DPCA} = \sum_{i=1}^{N_{CA}} \left(\text{Obs}_i^{CA} - \text{Est}_i^{DPCA}\right)^2, \tag{4}$$

where N_{CA} is the number of data in the calibration stage, E^{DPCA} is the SSE between observed (Obs^{CA}) and directly predicted (Est^{DPCA}) values in the calibration stage. In this study, for the construction of the weighted error function, the recursive prediction as well as the direct prediction is conducted at the calibration stage. And the errors due to the direct and recursive predictions are combined using a weighting factor to construct the weighted error function:

$$\text{minimize} \quad w^{CA}\ E^{DPCA} + (1-w^{CA})\ E^{RPCA}, \quad \left(0 \le W^{CA} \le 1\right), \tag{5}$$

$$E^{RPCA} = \sum_{i=1}^{N_{CA}} \left(\text{Obs}_i^{CA} - \text{Est}_i^{RPCA}\right)^2, \tag{6}$$

where w^{CA} is a weighting factor of the calibration stage and E^{RPCA} is the SSE between observed and recursively predicted (Est^{RPCA}) values in the calibration stage. The proposed method using the weighted error function (Eq. (5)) can reduce the possibility that a poorly performing recursive prediction model is constructed, though its corresponding direct prediction model is excellent. As mentioned in Section 2.1, because the ANN based on the ERM can be more susceptible to the deterioration of the model performance for the recursive prediction, the error function of the ANN in the training stage is modified as follows:

$$\text{minimize} \quad w^{TR}\ E^{DPTR} + (1-w^{TR})\ E^{RPTR}, \quad \left(0 \le W^{TR} \le 1\right), \tag{7}$$

$$E^{DPTR} = \sum_{i=1}^{N_{TR}} \left(\text{Obs}_i^{TR} - \text{Est}_i^{DPTR}\right)^2, \tag{8}$$

$$E^{RPTR} = \sum_{i=1}^{N_{TR}} \left(\text{Obs}_i^{TR} - \text{Est}_i^{RPTR}\right)^2, \tag{9}$$

where w^{TR} is a weighting factor of the training stage, E^{DPTR} and E^{RPTR} are the SSEs between observed (Obs^{TR}), and directly (Est^{DPTR}) and recursively (Est^{RPTR}) predicted values, respectively, in the training stage. The selection process of the weighting factors will be described in Section 4.2.

2.3. Structure of the time series models

The time series model of this study consists of three parts: a preprocessor (GENIN) for input data setting, model building

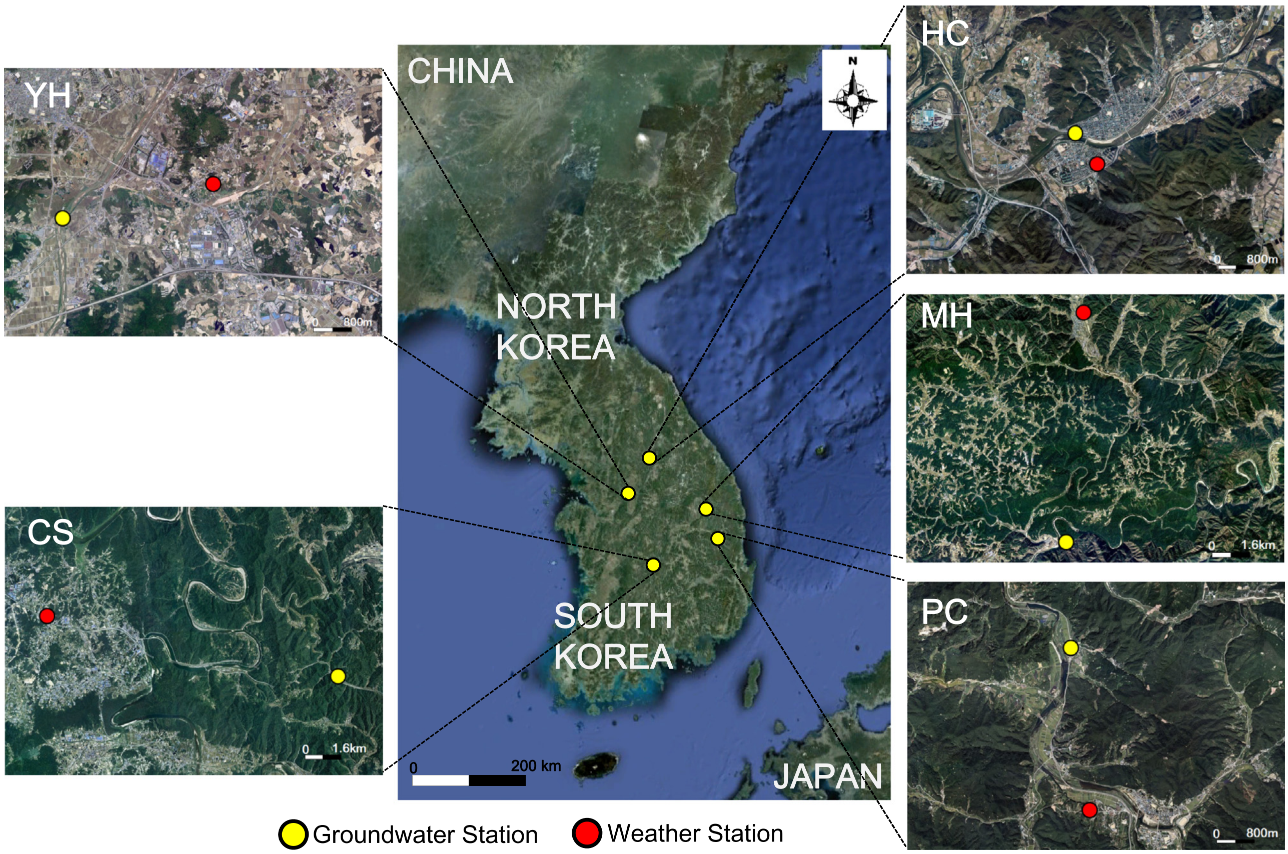

Fig. 2. Location of the selected NGMN stations.

Table 1
Selected NGMS stations and temporal extent of the data (year).

Station	Total data	Training	Testing	Validation
HC	2003–2008	2003–2004	2005–2006	2007–2008
MH	2003–2008	2003–2004	2005–2006	2007–2008
YH	2003–2008	2003–2004	2005–2006	2007–2008
PC	2005–2008	2005	2006	2007–2008
CS	2003–2008	2003–2004	2005–2006	2007–2008

modules, and a post-processor (CALGL) for the estimation of the GWL. GENIN generates input data set in accordance with the structure of the input layers for model building modules and a corresponding target data set from the time series data. The model building modules include two independent model builders: ANN-based model builder (ANNMB) and SVM-based model builder (SVMMB). CALGL, using ANNMB and SVMMB with selected model parameter sets, calculates three types of prediction results for the validation of the model performances: one-step ahead direct prediction, recursive prediction with and without weighted error function. A flowchart of the overall modelling process is shown in Fig. 1.

2.4. Model performance criteria

Four performance criteria were used to evaluate and compare the results of the ANN and SVM models: the mean error (ME), the mean absolute percentage error (MAPE), the root-mean-square error (RMSE), and the correlation coefficient (CORR). The ME

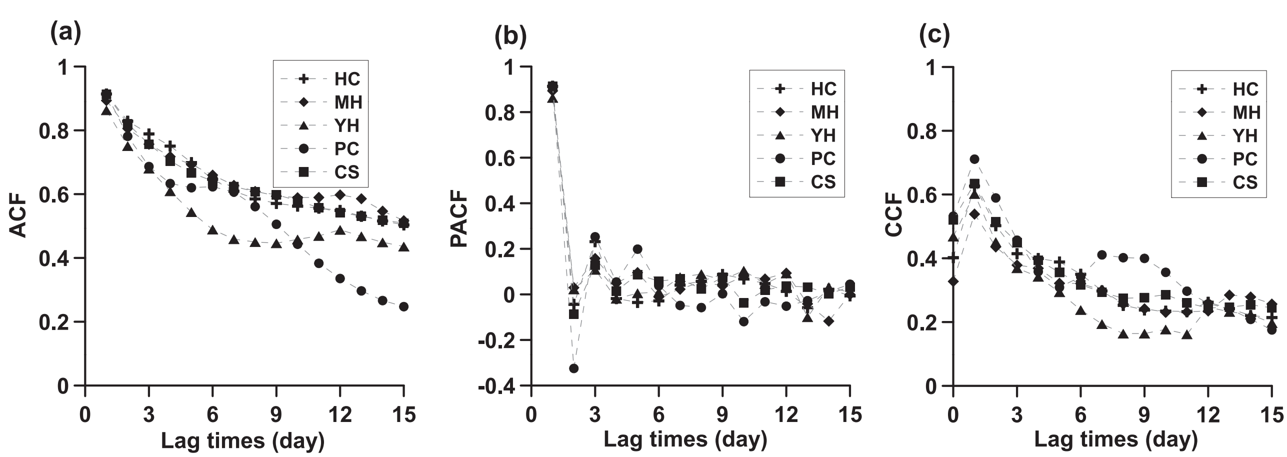

Fig. 3. Statistical analyses for input vector selection: (a) ACF, (b) PACF, and (c) CCF.

Table 2
Selected input structures and model parameters.

Station	Input structure (Rain – GWL)	ANN model parameters				SVM model parameters			
		w^{TR}–w^{CA}	HN	LR	MM	w^{CA}	C	ε	σ
HC	3–3	0.6–0.8	7	0.0015	0.1	0.4	7.5	0.0975	2.9
MH	2–5	0.0–0.3	8	0.005	0.2	0.4	7.0	0.08	2.0
YH	2–3	0.7–0.5	20	0.005	0.4	0.5	9.0	0.1	3.1
PC	3–5	0.0–0.4	18	0.002	0.1	0.8	13.0	0.095	3.0
CS	3–3	0.1–0.8	6	0.003	0.3	0.7	9.5	0.1	3.2

Table 3
The direct prediction results for the 5 NGMN stations.

Station	Model	ME ($\times 10^{-3}$m)	MAPE (%)	RMSE (m)	CORR
HC	ANN	−1.28	0.033	0.077	0.978
	SVM	1.23	0.031	0.072	0.980
MH	**ANN**	8.38	0.018	0.103	0.919
	SVM	−0.056	0.017	0.110	0.911
YH	**ANN**	−13.9	0.081	0.075	0.953
	SVM	−11.6	0.078	0.076	0.951
PC	ANN	−0.198	0.019	0.062	0.962
	SVM	−2.17	0.016	0.050	0.973
CS	**ANN**	−4.56	0.023	0.063	0.930
	SVM	−4.92	0.024	0.063	0.930

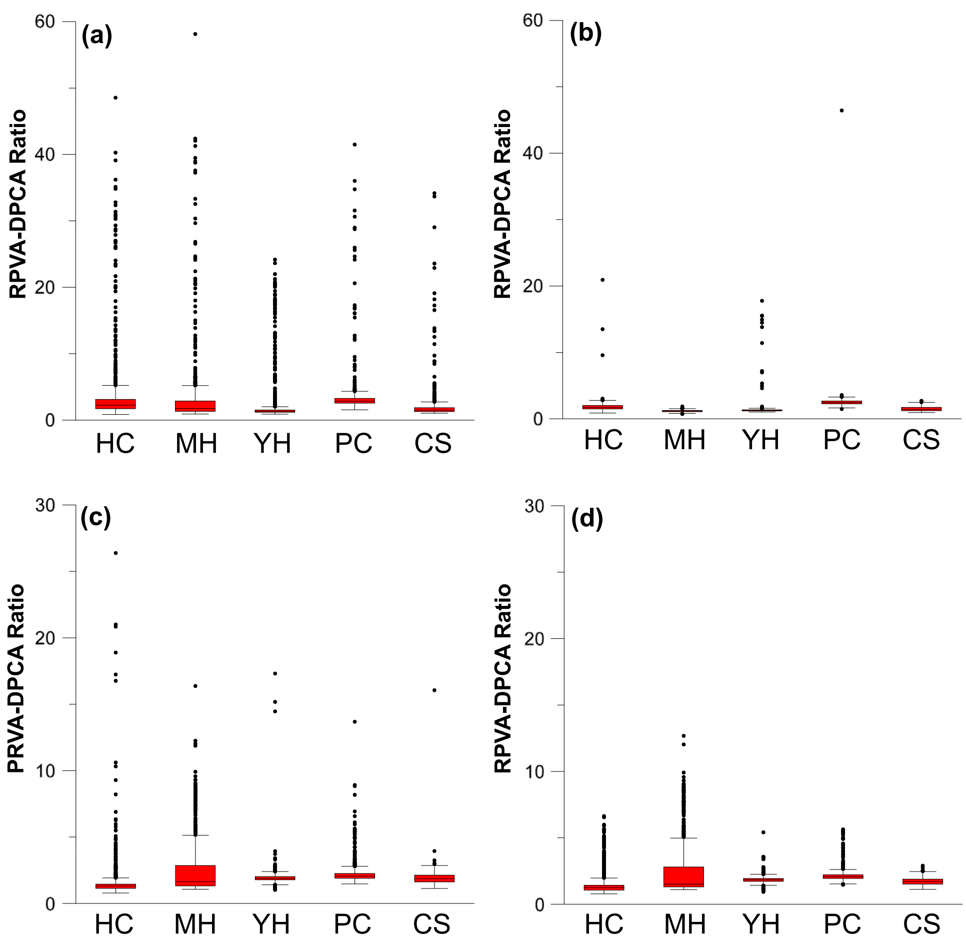

Fig. 4. Comparisons of the RP–TS ratio: (a) ANN without weighting factors, (b) ANN with weighting factors, (c) SVM without weighting factors, and (d) SVM with weighting factors.

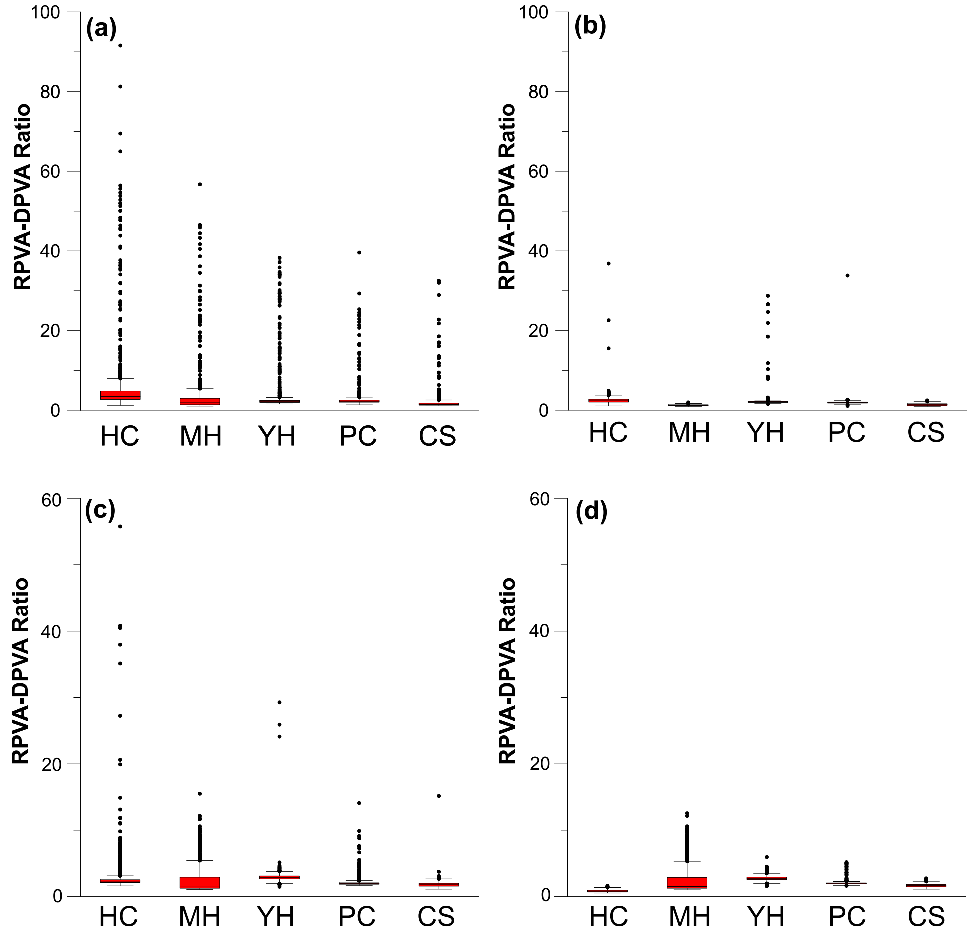

Fig. 5. Comparisons of the RP–DP ratio: (a) ANN without weighting factors, (b) ANN with weighting factors, (c) SVM without weighting factors, and (d) SVM with weighting factors.

evaluates the bias in the total errors; the MAPE represents the relative magnitude of the errors; the RMSE reflects the average magnitude of the errors, assigning higher weights to large errors using quadratic scoring; and the CORR represents the strength and direction of a linear relationship between the target and output values. The performance criteria are mathematically expressed as follows:

$$\mathrm{ME} = \frac{1}{N_{\mathrm{VA}}}\sum_{i=1}^{N_{\mathrm{VA}}}\left(\mathrm{Obs}_i^{\mathrm{VA}} - \mathrm{Est}_i^{\mathrm{VA}}\right), \tag{10}$$

$$\mathrm{MAPE}(\%) = 100 \times \frac{1}{N_{\mathrm{VA}}}\left(\sum_{i=1}^{N_{\mathrm{VA}}}\frac{\left|\mathrm{Obs}_i^{\mathrm{VA}} - \mathrm{Est}_i^{\mathrm{VA}}\right|}{\mathrm{Obs}_i}\right), \tag{11}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_{\mathrm{VA}}}\sum_{i=1}^{N_{\mathrm{VA}}}\left(\mathrm{Obs}_i^{\mathrm{VA}} - \mathrm{Est}_i^{\mathrm{VA}}\right)^2}, \tag{12}$$

$$\mathrm{CORR} = \frac{\frac{1}{N_{\mathrm{VA}}}\sum_{i=1}^{N_{\mathrm{VA}}}\left(\mathrm{Obs}_i^{\mathrm{VA}} - \overline{\mathrm{Obs}_i^{\mathrm{VA}}}\right)\left(\mathrm{Est}_i^{\mathrm{VA}} - \overline{\mathrm{Est}_i^{\mathrm{VA}}}\right)}{\sqrt{\frac{1}{N_{\mathrm{VA}}}\sum_{i=1}^{N_{\mathrm{VA}}}\left(\mathrm{Obs}_i^{\mathrm{VA}} - \overline{\mathrm{Obs}_i^{\mathrm{VA}}}\right)^2}\sqrt{\frac{1}{N_{\mathrm{VA}}}\sum_{i=1}^{N_{\mathrm{VA}}}\left(\mathrm{Est}_i^{\mathrm{VA}} - \overline{\mathrm{Est}_i^{\mathrm{VA}}}\right)^2}}, \tag{13}$$

where N_{VA} is the number of data points in the validation stage, $\mathrm{Obs}^{\mathrm{VA}}$ is an observed value, $\mathrm{Est}^{\mathrm{VA}}$ is an estimated value from ANNMB or SVMMB module, and $\overline{\mathrm{Obs}^{\mathrm{VA}}}$ and $\overline{\mathrm{Est}^{\mathrm{VA}}}$ are the mean values of $\mathrm{Obs}^{\mathrm{VA}}$ and $\mathrm{Est}^{\mathrm{VA}}$, respectively.

3. Study sites and data descriptions

In this study, the aforementioned time series models were applied to forecast daily averaged GWL data from National Groundwater Monitoring Network (NGMN) stations in South Korea. Five NGMN stations were selected for the model validation. For testing the applicability of the suggested methodology, the primary condition for selecting stations was set to be a rapid and accurate response in GWLs to rainfall events, ensuring that the daily total rainfall data of weather stations near the NGMN stations can be assumed as the natural primary input variable. Maximum correlation coefficients from cross-correlation analyses between rainfall and GWLs ranged from 0.54 to 0.71. The temporal extent, continuity, and quality of the data were also considered when selecting stations. The selected NGMN stations were Hongcheon (HC), Myeongho (MH), Yulhyeon (YH), Pacheon (PC), and Cheongseong (CS) (Fig. 2).

The time series data from each station were divided into three sets for the model training, testing, and validation stages. Table 1 describes the temporal extent of the data in the 3 stages for each station. The data for each stage (X) were standardised using the minimum ($X_{\min}^{\mathrm{TR}}$) and maximum ($X_{\max}^{\mathrm{TR}}$) values of the training data as follows:

$$\text{standardized } X = \frac{X - X_{\min}^{\mathrm{TR}}}{X_{\max}^{\mathrm{TR}} - X_{\min}^{\mathrm{TR}}} \tag{14}$$

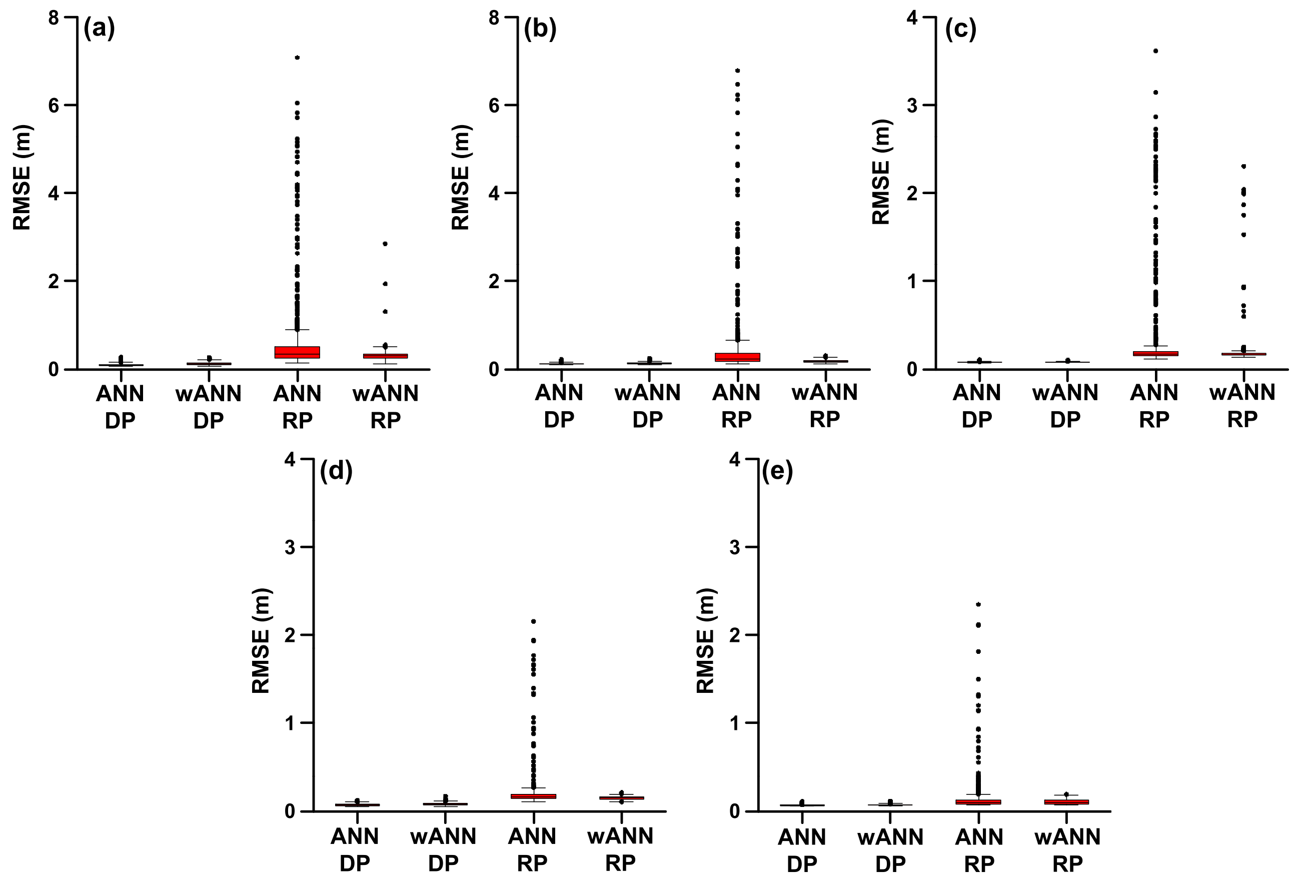

Fig. 6. Comparisons of the RMSEs for the ANNs: (a) HC, (b) MH, (c) YH, (d) PC, and (e) CS. Here, DP and RP denote direct and recursive prediction, respectively; wANN denotes the ANN model that uses weighting factors.

Following the training stage, the calculated standardised GWLs were retransformed using Eq. (14) to evaluate the model performance.

4. Results and discussion

4.1. Input vector selection

Determining an appropriate set of model inputs is an important step in applying learning algorithms to water resources. There have been several studies related to the input structure design using common trial-and-error methods, statistical approaches, and other various optimisation techniques (Coulibaly et al., 2000; Sudheer et al., 2002; Bowden et al., 2005; Nayak et al., 2006). The present study employed a statistical method for determining model input structures using autocorrelation functions (ACFs) and partial autocorrelation functions (PACFs) of GWL data, and cross-correlation functions (CCFs) between rainfall and GWL data. The ACF, PACF, and CCF statistics of the rainfall and GWL data from the 5 NGMN stations are presented in Fig. 3. The gradual decaying pattern of the ACF suggests the presence of a dominant auto-regressive process (Fig. 3a); the partial autocorrelation values with a lag time greater than 3 or 5 days were not significant (Fig. 3b). The CCF shows that the highest correlation value occurred at a lag time of 1 day, which implies a rapid response of rainfall and GWLs (Fig. 3c). In this study, the lag time at which the cross-correlation was greater than 0.4 was chosen as the number of rainfall input components for all NGMN stations. The selected input structures of the 5 NGMN stations are summarised at the second column of Table 2. Input data sets for each NGMN stations were prepared using GENIN module with the information of the selected input structures.

4.2. Model selection

The model parameters in this study are the number of hidden nodes (HN), the momentum (MM), and the learning rate (α) for the ANN (Appendix A). Moreover, the positive trade-off parameter (C), the tolerance of the loss function (ε), and the kernel function parameter (σ) are used for the SVM (Appendix B). The trial-and-error method was employed for selecting the weighting factors and model parameters, which were allowed to vary as follows: w^{TR}, w^{CA}[0.0, 1.0], HN∈[2, 20], $\alpha \in$ [0.0005, 0.005], MM∈[0.0, 0.9], C∈[6.0, 14.0], $\varepsilon \in$ [0.07, 0.16], and $\sigma \in$ [2.0, 4.0]. The parameter set for each model was selected among 1000 combination sets of parameters. The initial distribution of weights and biases is also an important factor for the BPA-based ANN model building process when considering the local minimum problem. In this study, 100 random sets were explored for each combination of model parameters to select the best initial distribution of weights and biases for ANN models. Table 2 summarises the selected weighting factors and model parameter sets for the 5 NGMN stations.

4.3. Comparison of model performances

The one-step ahead direct prediction results and the recursive prediction results were systematically compared with the station observations. The model performance criteria using direct prediction with the selected model parameters for the 5 NGMN stations are described in Table 3. The MEs ranged from -1.39×10^{-2} m to 8.38×10^{-3} m, which implies that the

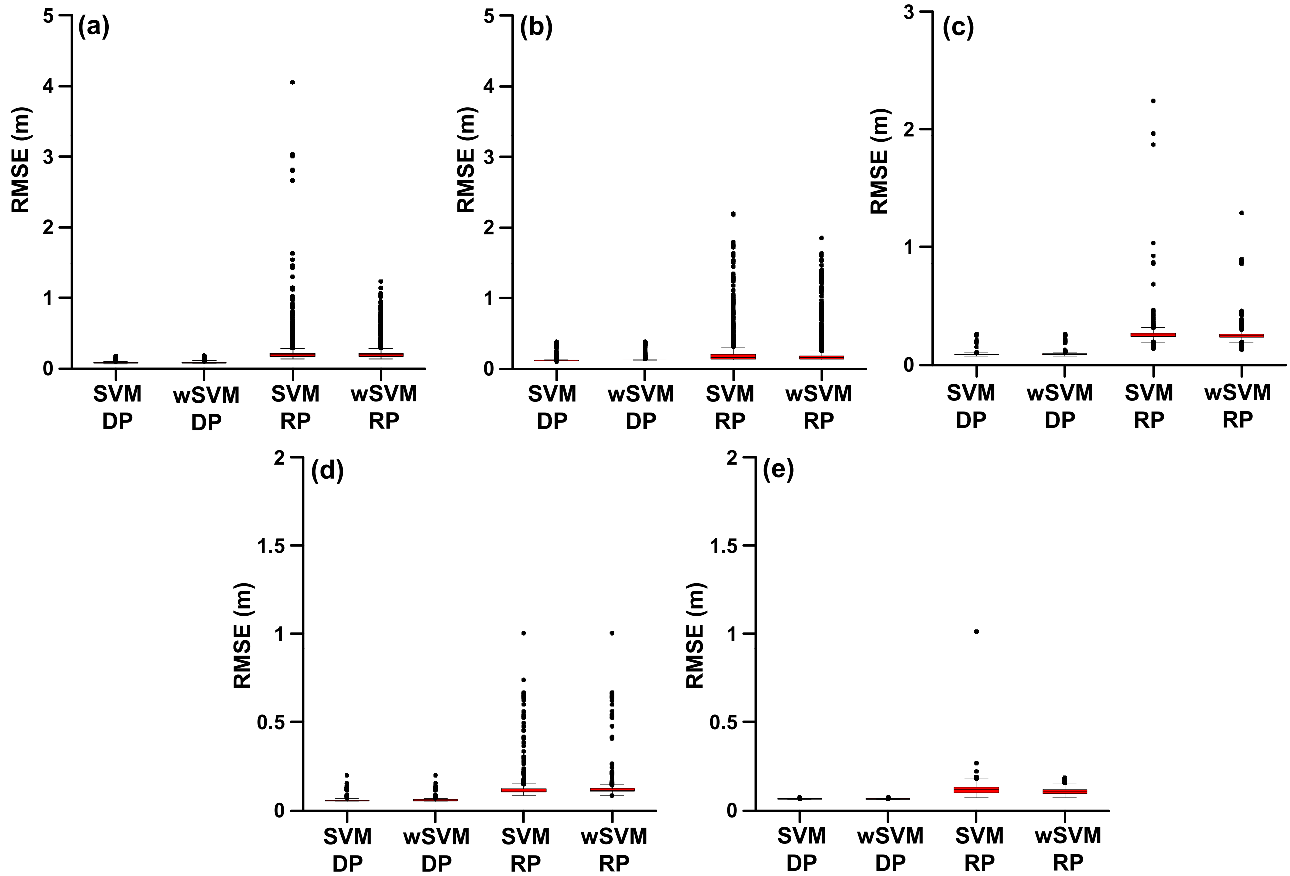

Fig. 7. Comparisons of the RMSEs for the SVMs: (a) HC, (b) MH, (c) YH, (d) PC, and (e) CS. Here, DP and RP denote direct and recursive prediction, respectively; wSVM denotes the SVM model that uses weighting factors.

modelling results were not substantially biased. The MAPEs and RMSEs were all less than 0.081% and 0.11 m, respectively; the CORRs were all greater than 0.90, which indicates that the models were well designed for predicting GWLs at these study sites. Based on the RMSEs, the performance of the ANN model was superior to that of the SVM model for the MH and YH stations. On the contrary, SVM performed better than ANN for the HC and PC stations; the models performed similarly for CS.

To evaluate the effects of the weighting factors on the model performances, the RMSEs at calibration and validation stages were calculated for each parameter set of both the non-weighted (weighting factor values equal to 1.0) and weighted schemes. And an RPVA-DPCA ratio and an RPVA-DPVA ratio were defined as follows and calculated for every model parameter set:

$$\text{RPVA}-\text{DPCA ratio} = \frac{\text{RMSE of recursive prediction in Validation stage}}{\text{RMSE of direct prediction in Calibration stage}}, \tag{15}$$

$$\text{RPVA}-\text{DPVA ratio} = \frac{\text{RMSE of recursive prediction in Validation stage}}{\text{RMSE of direct prediction in Validation stage}}, \tag{16}$$

The model selection rule of this study is to minimize the error of the calibration stage and the recursive prediction model utilizes the selected one-step ahead direct prediction model, therefore, the range of the RPVA-DPCA ratio values indicates the extent of the possibility that inadequate models for the recursive prediction are selected. Thus, as the RPVA-DPCA ratio value and its distribution decrease, the probability that a recursive prediction model with high model performance is selected increases. The RPVA-DPVA ratio value indicates the extent of the consistency between the direct prediction and recursive prediction. Thus, as the RPVA-DPVA ratio value approaches one and its distribution decreases, the probability that a recursive prediction model with high consistency is selected increases. The results of the calculations show that the RPVA-DPCA ratio and RPVA-DPVA ratio distributions for the recursive prediction models without the weighting factor varied substantially, especially for ANN models; however, the ranges decreased when using weighting factors (Figs. 4 and 5), indicating that the weighted error function approach can enhance the stability of the recursive prediction models.

Figs. 6 and 7 describe the RMSE distributions for each station and model; the statistics are provided in Table 4. The wANN and wSVM denote the ANN and SVM model constructed using the weighted error function, DP denotes direct prediction, and RP denotes recursive prediction. For the one-step ahead direct prediction, the median, mean, and standard deviation of the RMSEs with the weighted error function were similar or slightly greater than those without the weighted error function. However, for the recursive prediction, the range and statistics of the RMSEs were substantially smaller when using the weighted error function, especially for the ANN model, indicating that the weighed error function approach can improve the accuracy of recursive prediction models.

Table 5 shows the model performance criteria for recursive GWL predictions with the selected model parameters for the 5 NGMN stations. The poor model performance problem discussed in Section 2.1 is not present; the models performed reasonably well: the MEs ranged from -1.06×10^{-1} m to 5.36×10^{-2} m, the MAPEs ranged from 0.037% to 0.273%, the RMSEs ranged from 0.072 m to 0.159 m, and the CORRs ranged from 0.866 to 0.939.

Table 4
RMSE statistics (m) for the model parameter sets. Stdev. denotes the standard deviation.

Station	Model	Min	Max	Median	Mean	Stdev.
HC	ANN-DP	0.065	0.274	0.093	0.101	0.028
	wANN-DP	0.071	0.266	0.120	0.125	0.034
	SVM-DP	0.069	0.178	0.083	0.085	0.009
	wSVM-DP	0.071	0.183	0.085	0.089	0.015
	ANN-RP	0.136	7.078	0.344	0.534	0.725
	wANN-RP	0.127	2.851	0.299	0.303	0.119
	SVM-RP	0.135	4.051	0.194	0.247	0.263
	wSVM-RP	0.133	1.230	0.189	0.238	0.154
MH	ANN-DP	0.103	0.222	0.121	0.125	0.015
	wANN-DP	0.101	0.244	0.132	0.136	0.020
	SVM-DP	0.103	0.377	0.118	0.124	0.030
	wSVM-DP	0.108	0.377	0.120	0.127	0.036
	ANN-RP	0.124	6.781	0.235	0.414	0.699
	wANN-RP	0.115	0.307	0.171	0.174	0.030
	SVM-RP	0.119	2.192	0.156	0.247	0.260
	wSVM-RP	0.119	1.849	0.152	0.234	0.247
YH	ANN-DP	0.074	0.107	0.079	0.079	0.004
	wANN-DP	0.075	0.103	0.080	0.081	0.004
	SVM-DP	0.075	0.259	0.087	0.088	0.014
	wSVM-DP	0.074	0.258	0.088	0.090	0.016
	ANN-RP	0.119	3.615	0.171	0.314	0.481
	wANN-RP	0.136	2.305	0.164	0.187	0.164
	SVM-RP	0.136	2.242	0.253	0.260	0.103
	wSVM-RP	0.127	1.288	0.246	0.248	0.065
PC	ANN-DP	0.051	0.124	0.071	0.074	0.012
	wANN-DP	0.052	0.169	0.078	0.079	0.013
	SVM-DP	0.048	0.199	0.057	0.058	0.009
	wSVM-DP	0.047	0.199	0.058	0.058	0.009
	ANN-RP	0.103	2.156	0.165	0.196	0.185
	wANN-RP	0.104	0.210	0.149	0.150	0.016
	SVM-RP	0.086	1.006	0.111	0.122	0.062
	wSVM-RP	0.084	1.006	0.114	0.121	0.055
CS	ANN-DP	0.062	0.109	0.065	0.068	0.006
	wANN-DP	0.063	0.113	0.068	0.072	0.008
	SVM-DP	0.062	0.074	0.065	0.065	0.001
	wSVM-DP	0.062	0.075	0.065	0.065	0.001
	ANN-RP	0.069	2.346	0.095	0.131	0.171
	wANN-RP	0.069	0.191	0.095	0.104	0.027
	SVM-RP	0.070	1.013	0.117	0.119	0.036
	wSVM-RP	0.070	0.185	0.108	0.107	0.017

Table 5
The recursive prediction results for the 5 NGMN stations.

Station	**Model**	**ME ($\times 10^{-3}$m)**	**MAPE (%)**	**RMSE (m)**	**CORR**
HC	ANN	14.9	0.086	0.140	0.928
	SVM	−1.35	0.078	0.132	0.939
MH	ANN	53.6	0.042	0.133	0.885
	SVM	−7.40	0.037	0.124	0.883
YH	ANN	−106.0	0.273	0.159	0.872
	SVM	−90.8	0.251	0.149	0.873
PC	ANN	−0.498	0.061	0.131	0.866
	SVM	−2.31	0.043	0.095	0.911
CS	**ANN**	−0.027	0.038	0.072	0.916
	SVM	−0.043	0.047	0.083	0.907

Among the recursive prediction results, the MEs, MAPEs and RMSEs were largest at YH. The YH station is surrounded by many paddy fields where groundwater extraction for rice planting is high and is concentrated from May to June. Unnatural decreases in GWL during this season (Fig. 8(c)) may cause overestimation in the models and deterioration in the model performance. The GWL prediction results from the selected models for the 5 NGMN stations are shown in Fig. 8. The results of one-step ahead direct prediction and recursive prediction with the weighted error function approach are similar, which indicates that the developed recursive prediction models are suitable for the long-term prediction of the GWL. The model performance tended to decrease at peak values; however, the rainfall-GWL response patterns were adequately simulated, including groundwater recession. Based on the RMSEs, the performance of the ANN model was better than the SVM model at CS, whereas SVM was better at HC, MH, YH, and PC. In the present case study, contrary to the direct prediction, the recursive prediction performance of the SVMs was superior to the ANNs.

There have been several comparative studies of the ANN and the SVM for the prediction of GWL fluctuations. Gill et al. (2007) evaluated the effect of missing data on the prediction of the GWL, Behzad et al. (2010) and Yoon et al. (2011) compared the model performances under variable pumping and weather conditions and in a coastal aquifer, respectively. They concluded that the SVM model outperformed the ANN in accuracy and stability and the reason of the outperformance comes from higher generalization ability of the SVM based on the SRM principle, which corresponds to the result of the present study. Most of the previous studies that utilized the ANN or the SVM as time series models for the prediction of the GWL employed the direct prediction strategy. Thus the long-term prediction was limited. The main contribution of this study lies in suggesting a simple and improved method for the long-term prediction of the GWL. However, the current study has a limitation in using the rainfall as a sole exogenous input, whereas other previous researches considered various inputs such as temperature, evapotranspiration, and pumping rates as well as the rainfall.

5. Summary and conclusions

In the present study, a weighted error function approach was utilised to improve the performance of ANN- and SVM-based recursive prediction models for the long-term prediction of GWLs in response to rainfall. The developed time series models were applied to the GWL data from 5 NGMN stations in South Korea. The results demonstrated that the weighted error function approach can improve the stability and accuracy of recursive prediction models, particularly for ANNs. The weighted error function approach reduced the possibility that a poorly performing recursive prediction model is selected because it considers direct and recursive prediction errors simultaneously for the selection of the best model parameter set at the model calibration stage.

There was no superiority between the ANN or SVM models for the direct prediction of GWL fluctuations. However, the SVM model outperformed the ANN for recursive predictions at 4 of the 5 stations according to the RMSEs. The SVM is based on the SRM principle thus theoretically its generalization ability is higher than the ANN based on the ERM principle. This inherent feature probably enables the SVM to capture the rainfall-GWL relationship more effectively and to be superior to the ANN for recursive predictions.

In the field of hydrology, there have been a number of researches on predicting variables related to water resources using time series models; however, the researches on improving the performance of the long-term time series prediction are relatively scarce. In the present research, we designed a simple method, the weighting factor approach, to improve the performance of time series models based on ANN and SVM, and successfully verify the applicability to the long-term prediction of the GWL. The developed time series models with the weighted error function and the overall application results could be useful for the evaluation of the effects of rainfall pattern changes on groundwater resources and for the effective management of groundwater resources. Further study should include the consideration of meteorological and hydrological variables besides the rainfall as inputs and application to various types of groundwater level data.

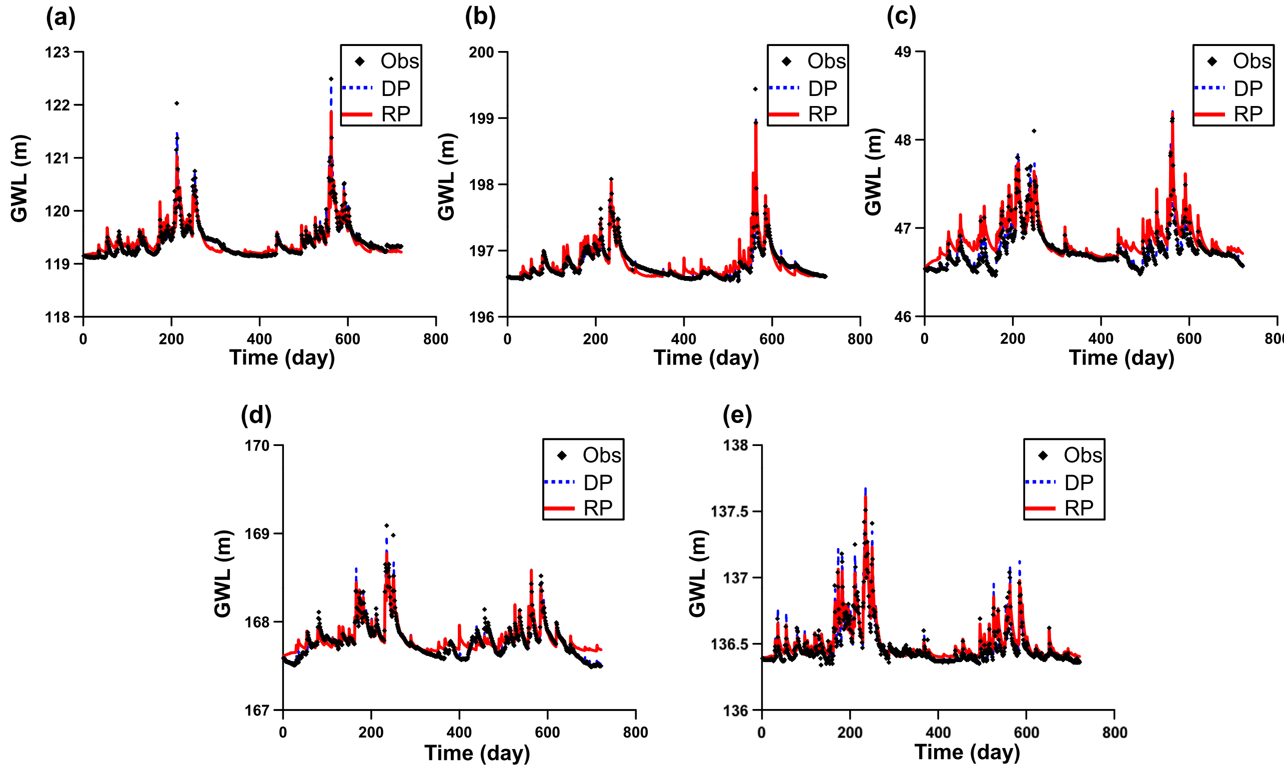

Fig. 8. Direct and recursive prediction results for the 5 stations: (a) HC, (b) MH, (c) YH, (d) PC, and (e) CS.

Acknowledgements

This work was supported by the research project of "Advanced Technology for Groundwater Development and Application in Riversides (Geowater+)" in "Water Resources Management Program (code 11 Technology Innovation C05)" of the MOLIT and the KAIA in Korea and Basic Science Research Program through the National Research Foundation of Korea(NRF) by the Ministry of Science, ICT & Future Planning (NRF-2013R1A1A3013721).

Appendix A. Artificial neural network (ANN)

An ANN is a flexible mathematical framework with parallel-distributed information processing that is patterned after the learning process of neurons in the human brain. A multilayer perceptron network (MLPN) (Rosenblatt (1962)), the most common ANN structure, was used for developing GWL time series models in this study. The designed MLPN consists of a single input and hidden and output layers; log-sigmoid and linear activation functions are used for the hidden and output layers, respectively. The mathematical expressions of the log-sigmoid activation function and the feedforward process of the MLPN are, respectively, described by

$$f(s) = \frac{1}{1 + e^{-s}}, \tag{A. 1}$$

$$y_j = f\left(\sum_{i=1}^{L} w_{ij}x_i + b_j\right), \tag{A. 2}$$

where the subscripts i and j represent nodes in the previous and present layers, respectively, x and y denote nodal values, f denotes the activation function of the present layer, w and b denote weights and bias values, respectively, and L is the number of nodes in the previous layer.

The back-propagation algorithm (BPA) with momentum (Rumelhart and McClelland, 1986) was employed to train the ANN. The weight and bias for the given dataset were updated using the error function between the observed and estimated values. A limitation of the BPA is that it searches for a local minimum in the error surfaces. The momentum term helps the ANN to avoid being captured in local minima by diminishing drastic weight changes over the iterative training process (Rumelhart and McClelland, 1986). The error function and the weight-updating rule of the ANN by the BPA can be expressed as follows:

$$E^m = \sum_{k=1}^{N} \left(\mathrm{Obs}_k^m - \mathrm{Est}_k^m\right)^2, \tag{A. 3}$$

$$w^{m+1} - w^m = \beta\left(w^m - w^{m-1}\right) + (1 - \beta)\ \alpha\ \left(-\frac{\partial E^m}{\partial w^m}\right), \tag{A. 4}$$

where m denotes the mth feedforward process or iteration; N denotes the number of data points in the training stage; Obs, Est, and E are the observed, estimated, and error values, respectively; and α and β denote the learning rate and momentum values, respectively.

Appendix B. Support vector machine (SVM)

An SVM is based on structural risk minimisation (SRM) (Vapnik, 1995) rather than the empirical risk minimisation (ERM) of ANNs, which enhances the generalisation of the SVM by simultaneously minimising both empirical error and model complexity.

The SVM maps input vectors into high-dimensional feature space through a nonlinear mapping function where the SVM constructs an optimal hyperplane with maximum margins. Given a set of N training samples $\{\mathbf{x}_k, \ y_k\}_{k=1}^N, \quad \mathbf{x} \in R^m, \ \ y \in R$, where $\mathbf{x}$ is an input vector of m components and y is an output value, the mathematical form of an SVM estimator (f) is as follows:

$$f(\mathbf{x}) = \mathbf{w}\cdot\varphi(\mathbf{x}) + b \tag{B. 1}$$

where $\mathbf{w}$ denotes a weight vector, b is a bias value, and φ is the nonlinear mapping function. Based on the SRM theorem (Vapnik, 1995), Eq. (B. 1) can be solved by the following convex optimisation problem:

$$\begin{aligned} &\underset{\mathbf{w},b,\xi,\xi^*}{\text{minimize}} \quad \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{k=1}^{N}\left(\xi_k + \xi_k^*\right) \\ &\text{subject to} \quad \begin{cases} y_k - \mathbf{w}^T\varphi(\mathbf{x}_k) - b \le \varepsilon + \xi_k \\ \mathbf{w}^T\varphi(\mathbf{x}_k) + b - y_k \le \varepsilon + \xi_k^* \\ \xi_k, \ \ \xi_k^* \ge 0 \end{cases} \quad k = 1, \ 2, \ \cdots, \ N \end{aligned} \tag{B. 2}$$

where C is a positive trade-off parameter that determines the degree of error in the training stage and ξ and ξ^* are slack variables that penalise training errors by Vapnik's ε-insensitivity loss function. In the optimisation equation, the term $\frac{1}{2}\|\mathbf{w}\|^2$ improves the generalisation of the SVM by regulating the degree of model complexity. Moreover, $C\sum_{k=1}^{N}(\xi_k + \xi_k^*)$ controls the degree of empirical risk. Eq. (B. 2) is usually reformulated into a dual form using Lagrangian multipliers (α, α^*) by imposing an optimality condition in which derivatives with respect to the primary variables (i.e., $w, \ \ b, \ \ \xi$ and ξ^*) should vanish:

$$\begin{aligned} &\underset{\alpha, \ \alpha^*}{\text{maximize}} \quad \begin{cases} -\dfrac{1}{2}\displaystyle\sum_{k, \ l=1}^{N}\left(\alpha_k - \alpha_k^*\right)\left(\alpha_l - \alpha_l^*\right) K(\mathbf{x}_k, \ \ \mathbf{x}_l) \\ -\varepsilon\displaystyle\sum_{k=1}^{N}\left(\alpha_k + \alpha_k^*\right) + \displaystyle\sum_{k=1}^{N} y_k\left(\alpha_k - \alpha_k^*\right) \end{cases} \\ &\text{subject to} \quad \begin{cases} \displaystyle\sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right) \\ 0 \le \alpha_k, \ \ \alpha_k^* \le C \end{cases} \\ &\text{to obtain} \quad f(\mathbf{x}) = \sum_{k=1}^{n}\left(\alpha_k - \alpha_k^*\right) K(\mathbf{x}, \ \ \mathbf{x}_k) + b \end{aligned} \tag{B. 3}$$

where n is the number of support vectors and K is a kernel function defined by an inner product of the nonlinear transfer functions. Here, a radial basis function with parameter σ, which is most commonly used, was employed for the kernel function:

$$K(\mathbf{x}_k, \mathbf{x}_l) = \exp\left(-\frac{\|\mathbf{x}_k - \mathbf{x}_l\|^2}{2\sigma^2}\right) \tag{B. 4}$$

The architecture of a SVM is not designed *a priori*. The input vectors that have nonzero Lagrangian multipliers under the Karush–Kuhn–Tucker condition support the structure of an SVM and are called support vectors. Training an SVM involves selecting the support vectors and optimising the weight and bias values. To train the SVM, the sequential minimal optimisation (SMO) algorithm (Platt, 1999) was employed in this study, which selects two α values at a time and sequentially and analytically solves the optimisation problem for the selected parameters (Platt, 1999; Schölkopf and Smola, 2002).

References

Akhtar, M.K., Corzo, G.A., van Andel, S.J., Jonoski, A., 2009. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. Hydrol. Earth Syst. Sci. 13 (9), 1607–1618.

Asefa, T., Kemblowski, M., Lall, U., Urroz, G., 2005. Support vector machines for nonlinear state space reconstruction: Application to the Great Salt Lake time series. Water Resour. Res. 41 (12), 1–10.

Asefa, T., Kemblowski, M., McKee, M., Khalil, A., 2006. Multi-time scale stream flow predictions: the support vector machines approach. J. Hydrol. 318 (1-4), 7–16.

Asefa, T., Kemblowski, M.W., Urroz, G., McKee, M., Khalil, A., 2004. Support vectors-based groundwater head observation networks design. Water Resour. Res. 40 (11), W1150901–W1150914.

Bardossy, A., 2007. Calibration of hydrological model parameters for ungauged catchments. Hydrol. Earth Syst. Sci. 11 (2), 703–710.

Behzad, M., Asghari, K., Coppola Jr., E.A., 2010. Comparative study of SVMs and ANNs in aquifer water level prediction. J. Comput. Civil Eng. 24 (5), 408–413.

Bowden, G.J., Dandy, G.C., Maier, D.H., 2005. Input determination for artificial neural network models in water resources applications. Part 1 – background and methodology. J. Hydrol. 301 (1–4), 75–92.

Campolo, M., Andeussi, P., Soldati, A., 1999. River flood forecasting with a neural network model. Water Resour. Res. 35 (4), 1191–1197.

Cimen, M., Kisi, O., 2009. Comparison of two different data-driven techniques in modeling lake level fluctuations in Turkey. J. Hydrol. 378 (3–4), 253–262.

Coppola, E., Rana, A.J., Poulton, M.M., Szidarovszky, F., Uhl, V.V., 2005. A neural network model for predicting aquifer water level elevations. Ground Water 43 (2), 231–241.

Coulibaly, P., Anctil, F., Bobee, B., 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. J. Hydrol. 230 (3–4), 244–257.

Coulibaly, P., Anctil, F., Aravena, R., Bobee, B., 2001. Artificial neural network modeling of water table depth fluctuations. Water Resour. Res. 37 (4), 885–896.

Deng, J., Chen, X., Du, Z., Zhang, Y., 2011. Soil water simulation and prediction using stochastic models based on LS-SVM for red soil region of China. Water Resour. Manag. 25 (11), 2823–2836.

Dialiakopoulos, I.N., Coulibaly, P., Tsanis, I.K., 2005. Groundwater level forecasting using artificial neural networks. J. Hydrol. 309 (1–4), 229–240.

Dibike, Y.B., Velickov, S., Solomatine, D., Abbott, M.B., 2001. Model induction with support vector machines: introduction and applications. J. Comput. Civil Eng. 15 (3), 208–216.

Dogan, E., Sengorur, B., Koklu, R., 2009. Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. J. Environ. Manag. 90 (2), 1229–1235.

French, M.N., Krajewski, W.F., Cuykendall, R.R., 1992. Rainfall forecasting in space and time using a neural network. J Hydrol. 137 (1–4), 1–31.

Gill, M.K., Asefa, T., Kaheil, Y., McKee, M., 2007. Effect of missing data on performance of learning algorithms for hydrologic prediction: implication to and imputation technique. Water Resour. Res. 43 (7), W07416.

Giustolisi, O., Simeone, V., 2006. Optimal design of artificial neural networks by a multi-objective strategy: groundwater level predictions. Hydrol. Sci. J. 51 (3), 502–523.

Herrera, L.J., Pomares, H., Rojas, I., Guillen, A., Prieto, A., Valenzuela, O., 2007. Recursive prediction for long term time series forecasting using advanced models. Neurocomputing 70 (16–18), 2870–2880.

Hu, T.S., Lam, K.C., NG, S.T., 2005. A modified neural network for improving river flow prediction. Hydrol. Sci. J. 50 (2), 299–318.

Hung, N.Q., Babel, M.S., Weesakul, S., Tripathi, N.K., 2009. An artificial neural network model for rainfall forecasting in Bangkok, Thailand. Hydrol. Earth Syst. Sci. 13 (8), 1413–1425.

IPCC (International Panel on Climate Change), 2007. Climate change, 2007: The physical science basis. Contributions of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom, New York, NY, USA.

Ji, Y., Hao, J., Reyhani, N., Lendasse, A., 2005. Direct and recursive prediction of time series using mutual information selection. Lect. Notes Comput. Sci. 3512, 1010–1017.

Jung, S., Kwon, S., 2013. Weighted error functions in artificial neural networks for improved wind energy potential estimation. Appl. Energ 111, 778–790.

Karunanithi, N., Grenney, W.J., Whitley, D., Bovee, K., 1994. Neural networks for river flow prediction. J. Comput. Civil Eng. – ASCE 8 (2), 201–220.

Khalil, A.F., McKee, M., Kemblowski, M., Asefa, T., Bastidas, L., 2006. Multiobjective analysis of chaotic dynamic systems with sparse learning machines. Adv. Water Resour. 29 (1), 72–88.

Khan, M.S., Coulibaly, P., 2006. Application of support vector machine in lake water level prediction. J. Hydrol. Eng. 11 (3), 199–205.

Knotters, M., Bierkens, M.F.P., 2000. Physical basis of time series models for water table depths. Water Resour. Res. 36 (1), 181–188.

Krishna, B., Satyaji Rao, Y.R., Vijaya, T., 2008. Modelling groundwater levels in an urban coastal aquifer using artificial neural networks. Hydrol. Process. 22 (8), 1180–1188.

Kuligowski, R.J., Barros, A.P., 1998. Using artificial neural networks to estimate missing rainfall data. J. Am. Water Resour. Assoc. 34 (6), 1437–1447.

Lek, S., Guiresse, M., Giraudel, J.-L., 1999. Prediction stream nitrogen concentration from watershed features using neural networks. Water Res. 33 (16), 3469–3478.

Liong, S.Y., Sivapragasam, C., 2002. Flood stage forecasting with support vector machines. J. Am. Water Resour. Assoc. 38 (1), 173–186.

Maier, H.R., Dandy, G.C., 1996. The use of artificial neural networks for the prediction of water quality parameters. Water Resour. Res. 32 (4), 1013–1022.

Malekmohamadi, I., Bazargan-Lari, M.R., Kerachian, R., Nikoo, M.R., Fallahnia, M.,

2011. Evaluating the efficacy of SVMs, BNs, ANNs and ANFIS in wave height prediction. Ocean. Eng. 38 (2–3), 487–497.
Mohanty, S., Jha, M.K., Kumar, A., Sudheer, K.P., 2010. Artificial neural network modeling for groundwater level forecasting in a river island of eastern India. Water Resour. Manag. 24 (9), 1845–1865.
Nayak, P.C., Satyaji Rao, Y.R., Sudheer, K.P., 2006. Groundwater level forecasting in a shallow aquifer using artificial neural network approach. Water Resour. Manag. 20 (1), 77–90.
NIMR, 2009. Research outcomes on understanding climate change II – climate change in the Korean Peninsula. Dongjin Press, Seoul, Korea.
Park, E., Parker, J.C., 2008. A simple model for water table fluctuations in response to precipitation. J. Hydrol. 356 (3–4), 344–349.
Platt, J.C., 1999. Fast training of support vector machines using sequential minimal optimization. In: Schölkopf, B., Burges, C.J.C., Smolar, A.J. (Eds.), Advances in Kernel Methods—Support Vector Learning. MIT Press, Cambridge, Massachusetts, USA.
Pollacco, J.A.P., Ugalde, J.M.S., Angulo-Jaramillo, R., Braud, I., Saugier, B., 2008. A linking test to reduce the number of hydraulic parameters necessary to simulate groundwater recharge in unsaturated zone. Adv. Water Resour. 31 (2), 355–369.
Rai, S.N., Singh, R.N., 1995. Two-dimensional modeling of water table fluctuation in response to localized transient recharge. J. Hydrol. 167 (1–4), 167–174.
Rosenblatt, F., 1962. Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, Washington DC, USA.
Rumelhart, D.E., McClelland, J.L., 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition The PDP Research Group. MIT Press, Cambridge, Massachusetts, USA.
Schölkopf, B., Smola, A.J., 2002. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge, Massachusetts, USA.
Sorjamaa, A., Hao, J., Reyhni, N., Ji, Y., Lendasse, A., 2007. Methodology for long-term prediction of time series. Neurocomputing 70 (16–18), 2861–2869.
Sudheer, K.P., Gosain, A.K., Ramasastri, K.S., 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. Hydrol. Process. 16 (6), 1325–1330.
Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer-Verlag, New York, USA.
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. J. Hydrol. 396 (1–2), 128–138.
Yu, P.S., Chen, S.T., Chang, I.F., 2006. Support vector regression for real-time flood stage forecasting. J. Hydrol. 328 (3–4), 704–716.
Zealand, C.M., Burn, D.H., Simonovic, S.P., 1999. Short-term streamflow forecasting using artificial neural networks. J. Hydrol. 214 (1–4), 32–48.