Journal of Environmental Management 182 (2016) 308–321

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Groundwater level prediction using a SOM-aided stepwise cluster inference model

Jing-Cheng Han [a], Yuefei Huang [a, b, *], Zhong Li [c], Chunhong Zhao [a], Guanhui Cheng [d], Pengfei Huang [a]

[a] State Key Laboratory of Hydroscience and Engineering, Dept. of Hydraulic Engineering, Tsinghua University, Beijing 100084, China
[b] State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining 810016, China
[c] Department of Civil Engineering, McMaster University, Hamilton, ON L8S 4L7, Canada
[d] Institute for Energy, Environment and Sustainable Communities, University of Regina, Regina, Saskatchewan S4S 0A2, Canada

ARTICLE INFO

Article history:
Received 11 January 2016
Received in revised form
17 June 2016
Accepted 22 July 2016

Keywords:
Groundwater level modeling
Uncertainty
SOM
Stepwise cluster inference
Autoregressive error model
Hexi Corridor

ABSTRACT

Accurate groundwater level (GWL) prediction can contribute to sustaining reliable water supply to domestic, agricultural and industrial uses as well as ecological services, especially in arid and semi-arid areas. In this paper, a regional GWL modeling framework was first presented through coupling both spatial and temporal clustering techniques. Specifically, the self-organizing map (SOM) was applied to identify spatially homogeneous clusters of GWL piezometers, while GWL time series forecasting was performed through developing a stepwise cluster multisite inference model with various predictors including climate conditions, well extractions, surface runoffs, reservoir operations and GWL measurements at previous steps. The proposed modeling approach was then demonstrated by a case of an arid irrigation district in the western Hexi Corridor, northwest China. Spatial clustering analysis identified 6 regionally representative central piezometers out of 30, for which sensitivity and uncertainty analysis were carried out regarding GWL predictions. As the stepwise cluster tree provided uncertain predictions, we added an AR(1) error model to the mean prediction to forecast GWL 1 month ahead. Model performance indicators suggest that the modeling system is a useful tool to aid decision-making for informed groundwater resource management in arid areas, and would have a great potential to extend its applications to more areas or regions in the future.

1. Introduction

Groundwater resource is commonly the most important water resource in semi-arid and arid areas that are often subject to water shortage. It plays a fundamental role in supplying clean and safe water to competing uses for domestic, industrial and agricultural sectors, and increasing attentions are also paid to its significance for ecological integrity. However, groundwater aquifer systems always feature complexity, high nonlinearity, being multi-scale and random as a result of the frequent interactions between surface water and groundwater as well as acute human disturbance (Nourani et al., 2015). Thus, effective modeling techniques would be required for providing efficient ground water management strategies. As for dynamic groundwater level (GWL) prediction, physical-based or conceptual models represent the hydrological variables and physical processes in real-world systems (Han et al., 2015), but they have practical limitations in terms of prediction accuracy as a result of unavoidable discrepancies between the model and the real-world system (Adamowski and Chan, 2011; Nourani et al., 2015; Salas et al., 1990). As far as increasingly scarce water resources accompanying with expanding population growth are concerned, improvements and innovations in groundwater predictions become critical. Hence, such black box or data driven models as Artificial Neural Networks (ANNs) were found to be widely employed by hydrogeologists (Chen et al., 2010; Coppola et al., 2005; Izady et al., 2013; Mohanty et al., 2010; Tapoglou et al., 2014; Yoon et al., 2011; Zahmatkesh et al., 2015).

Although various data-driven models were developed to predict GWL fluctuations, there are no consistent agreements on how to select an appropriate model with high efficiency in a real case

* Corresponding author. State Key Laboratory of Hydroscience and Engineering, Dept. of Hydraulic Engineering, Tsinghua University, Beijing 100084, China.
E-mail addresses: hanjc2012@gmail.com (J.-C. Han), yuefeihuang@tsinghua.edu.cn (Y. Huang).

http://dx.doi.org/10.1016/j.jenvman.2016.07.069
0301-4797/

(Coulibaly et al., 2001). Considering the "multiple inputs-multiple outputs" structure of regional GWL prediction models, a promising approach would be the Stepwise Cluster Analysis (SCA), which has been widely used for flow prediction recently by virtue of its ability to represent the nonlinear and complex relationships between various inputs and response variables (Fan et al., 2015; Huang et al., 2006; Li et al., 2015b). Moreover, it proved to be an effective and promising method for air-quality prediction and pilot-scale groundwater simulation (Huang et al., 2006; Qin et al., 2007; Sun et al., 2009). However, to perform GWL predictions at a regional level is still a hard work, considering the complexities of specific hydrogeological conditions and interactions between groundwater and surface water as well as climatic factors (Adamowski and Chan, 2011; Dash et al., 2010; Nourani et al., 2011). Predictor selection and parameter setting during the training phase would also lead to variations in the model performance with respect to the reliability and robustness of simulations, and optimal model configurations would be a key point of generating reliable simulations when SCA is applied. In order to deal with such issue, a stepwise cluster multisite inference model based on SCA specifically for accurately predicting regional GWL fluctuations would be indispensable.

Generally, data-driven models use statistical techniques instead of numerical simulation to relate the system response to various inputs, which are termed as predictors. As such, these models are able to "learn" system behavior of interest through exploring the patterns of representative data. Accordingly, the data quality for both predictors and training samples would impose varying influences on the model performance, and the introduction of irrelevant and redundant information might mislead the knowledge discovery process during the training phase and further yield unreliable predictions (Lábó, 2012). To tackle such dilemma, on one hand, it was recommended that pre-processing is recommended on the raw data to achieve accurate forecasting as conducted by many studies (Chen et al., 2010, 2011; Moosavi et al., 2013; Nourani et al., 2015). On the other hand, prevalent post-processing procedures were also adopted to correct the predictions (Li et al., 2015a; Morawietz et al., 2011). Accordingly, auxiliary procedures based on pre-processing or post-processing would be favorable alternatives to obtain more accurate GWL predictions.

Regional GWL observations are usually comprised of GWL time series for many piezometers, and a clustering technique may be preferred as a spatial data pre-processing tool to help to identify the regional characteristics dependent upon several representative observations instead of all observations involved. As such, the reduction of dimension would support efficient and informed decisions when black box models are used, although specific loss of GWL information might occur. Hence, intensive modeling efforts would be made based on these representative sites such as centroids for the obtained homogenous clusters. As an unsupervised machine learning technique, the self-organizing map (SOM) operates to reduce dimensions of high-dimensional data, and it could reveal the complex, nonlinear, and statistical relationships between high-dimensional data items on a low-dimensional display so as to allow optimal clusters to be determined (Chen et al., 2010; Kalteh et al., 2008; Kohonen, 1997; Nourani et al., 2015; Yang et al., 2012). Accordingly, the dimensionality of input variables as well as the resulting model complexity would be decreased (Hsu and Li, 2010; Hsu et al., 2002; Kalteh et al., 2008; Nourani et al., 2013, 2015).

Thus, a promising approach to achieve accurate and efficient regional GWL predictions would combine both spatial clustering method and data-driven model in association with pre/post-processing procedures. With multisite representative GWL observations being considered at the same time, it would then lead to a SOM-aided stepwise cluster multisite inference model for GWL prediction. In this study, this model would be developed to predict GWL a month into the future, and this paper is arranged as follows. The study area for an arid irrigation district is first introduced in the western Hexi Corridor of China. The proposed methodology is then presented with respect to modeling framework, SOM, stepwise cluster multisite inference and autoregressive error models as well as model performance indicators. Afterwards, the obtained results are presented in association with discussions, which are subsequently followed by summarized conclusions.

2. Study area and data

The study area is located in the Shule River watershed of China, which covers an area of approximately 160,000 km^2, and drained by such major inland rivers as Shule River, Dang River and Shiyou River (Fig. 1). Moreover, the Shule River is one of the three longest inland rivers in the Hexi Corridor, extending a length of 670 km. As depicted in Fig. 1, most part of the Shule River watershed is situated in Gansu Province, while almost all source waters are formed in the mountains of Qinghai Province. The whole watershed bears a semi-arid to arid climate, and according to the observations for Dunhuang, Yumen and Guazhou station during 1961–2010, annual mean temperature in the plain region is 7.2–9.7 °C together with annual precipitation ranging from 40.1 to 66 mm and annual evaporation up to 2755 mm (E20 pan evaporation). Therefore, the whole watershed is very vulnerable to water shortage, and the extent to which water demands across the plain region could be satisfied is highly dependent on the runoff contribution from source water areas. Although groundwater plays a vital role in sustaining the oasis and conserving the eco-environment, its pumping extraction is not effectively controlled and managed in this area, which is resided by almost 200,000 populations. Hence, careful management of water resources has been emphasized in order to guarantee a coordinated development between economy, society and eco-environment, and the water-saving policies and water-efficient techniques are always highlighted by local government. However, few models were developed for supporting water resource management, and no simulation modeling tools are really applied for GWL prediction as a result of a lack of hydrogeological investigations. Besides, complex interactions between groundwater and surface water further complicate the groundwater flow modeling procedures. Therefore, accurate predictions for GWL fluctuations through developing data-driven models like SCA would be imperative to enhance efficient water utilization as well as ecological conservation for this area.

In this study, the area of interest focuses on the middle reach of the Shule River. As we can see in Figs. 1 and 2, through construction of irrigation canals, the Changma reservoir serves as the main water supply to three irrigation districts, including Changma, Shuangta and Huahai. Besides, two more dams of Shuangta and Chijinxia were built to manage water resources for irrigation in the Shuangta and Huahai irrigation district. The average runoff for the Shule River that is drained annually into the Changma reservoir is 9.48×10^8 m^3 with apparent variations in monthly runoff. A large alluvial fan was formed after the river goes through the Changma Valley, and three main irrigation canals have been constructed along the river channel to divert water in a synergistic manner to three irrigation districts. The Changma irrigation district is selected herein to conduct modeling simulations for GWL prediction, taking its social-economic conditions and water consumptions into consideration. This entire irrigation district is located on the Yumen Basin, and has an area of 460 km^2, which is the largest in the Shule River watershed.

Fig. 3 presents the locations of 30 GWL observation wells

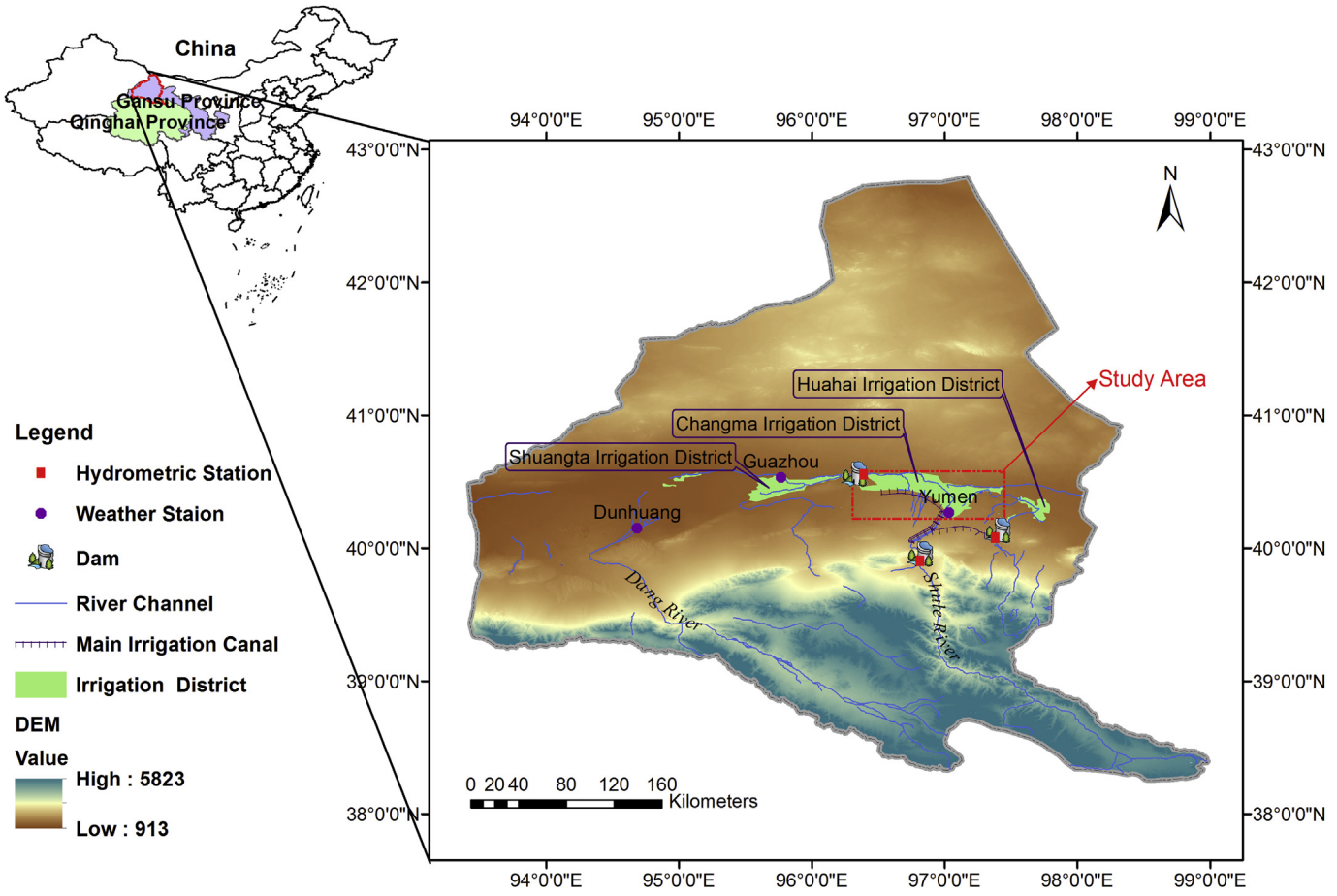

Fig. 1. Location of the study area.

(piezometers) across the Yumen Basin, and each well is also notated. To perform simulations, local data was collected and prepared as carefully as possible, and thorough examinations on these data were also carried out. Specifically, weather observations for Dunhuang, Yumen and Guazhou station during 1961–2010 were acquired from the China Meteorological Data Sharing Service System at a monthly step, consisting of precipitation, mean air temperature, monthly maximum/minimum daily average temperature, relative humidity, sunshine ratio and hours. Monthly hydrologic records based on daily observations from 1998 to 2010 such as monthly mean discharges at Changma and Yumen station, monthly mean inflow into Changma and Chijinxia Reservoir as well as monthly mean outflow from Changma Reservoir were provided by the Bureau of Shule River Water Resources Management. In addition to the GWL elevations for 30 wells across the Yumen Basin, the Bureau kept a clear record of groundwater withdrawal information on five delineated subareas from 1998 to 2010.

3. Methodology

Fig. 4 illustrates the flowchart of proposed forecasting framework. Generally, SOM is used as a spatial clustering method to group the GWL piezometers into several clusters, the number of which is determined with the help of non-hierarchical K-means classification method (Tarsitano, 2003). While the GWL elevations for central piezometers in the future are predicted through developing such temporal clustering method as stepwise cluster inference model. At the first step, input pattern analysis is undertaken on the GWL time series through establishing a self-organizing map. Then, the non-hierarchical K-means classification method is applied to identify central piezometers for representing the regional GWL fluctuations. To accurately predict the GWL for central piezometers at a lead time, candidate predictors such as weather, hydrology, antecedent GWL observations are prepared as well as to be screened based on experience and trial and error method. Afterwards, a stepwise cluster tree will be built through cutting and merging operation based on the training data sets. According to updated inputs, the GWL simulation would be then performed through the cluster tree. Thus, a forecasting system is developed for GWL prediction, and post-modeling analysis would be carried out accordingly. Details on the principles of modeling procedures will be further elaborated as follows.

3.1. Self organizing map (SOM)

SOM, also referred to Kohonen network (Kohonen, 1995), implements projection of an orderly mapping of a high-dimensional distribution onto a regular low-dimensional grid, and it takes advantages of visualizing high-dimensional data using discrete lattices (Nourani et al., 2015). Thus, a map of 1 or 2 dimensions produced by SOM is usually favored to illustrate the similarities of the data, and accordingly, similar data sets can be further grouped together to result in various categories with similar input patterns (Kohonen, 1997).

In general, the SOM network consists of two layers, an input layer and a Kohonen layer (SOM layer). The n-dimensional input layer is fully projected to the SOM layer through weight vector w, and a two dimensional SOM layer is the form that is applied in most

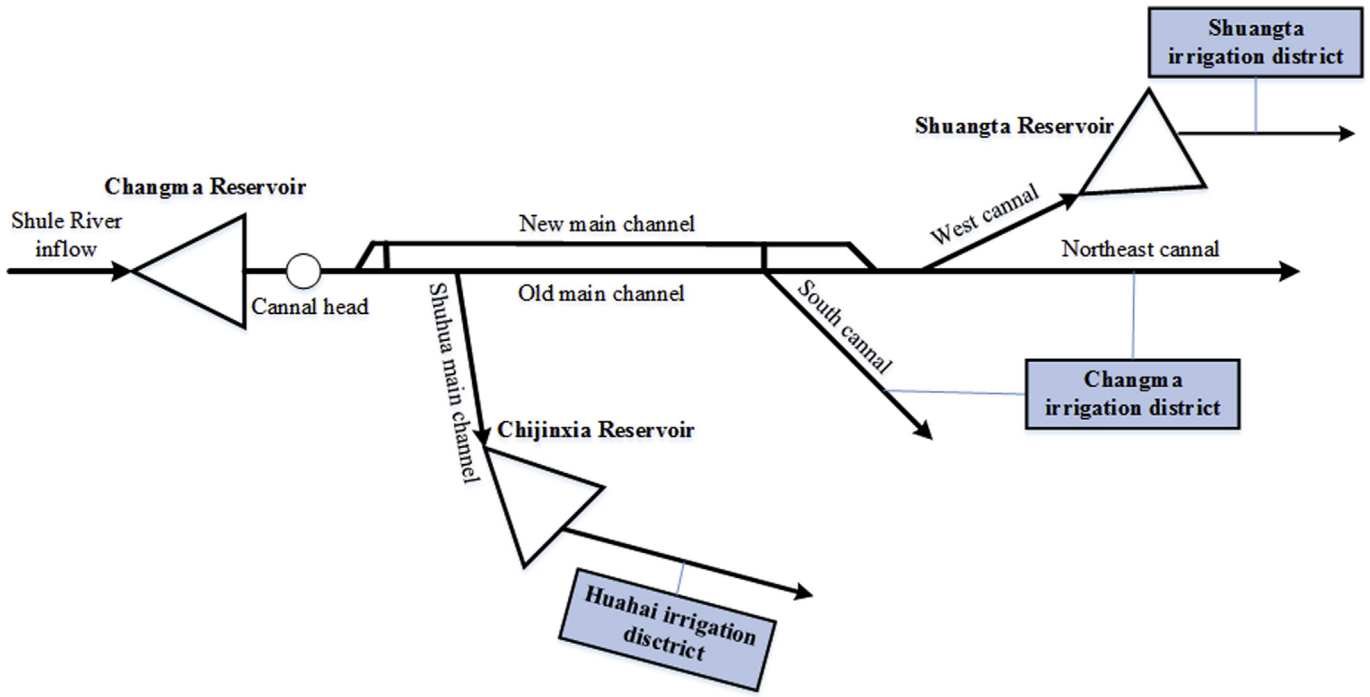

Fig. 2. Schematic representation of the water resource system in the study area.

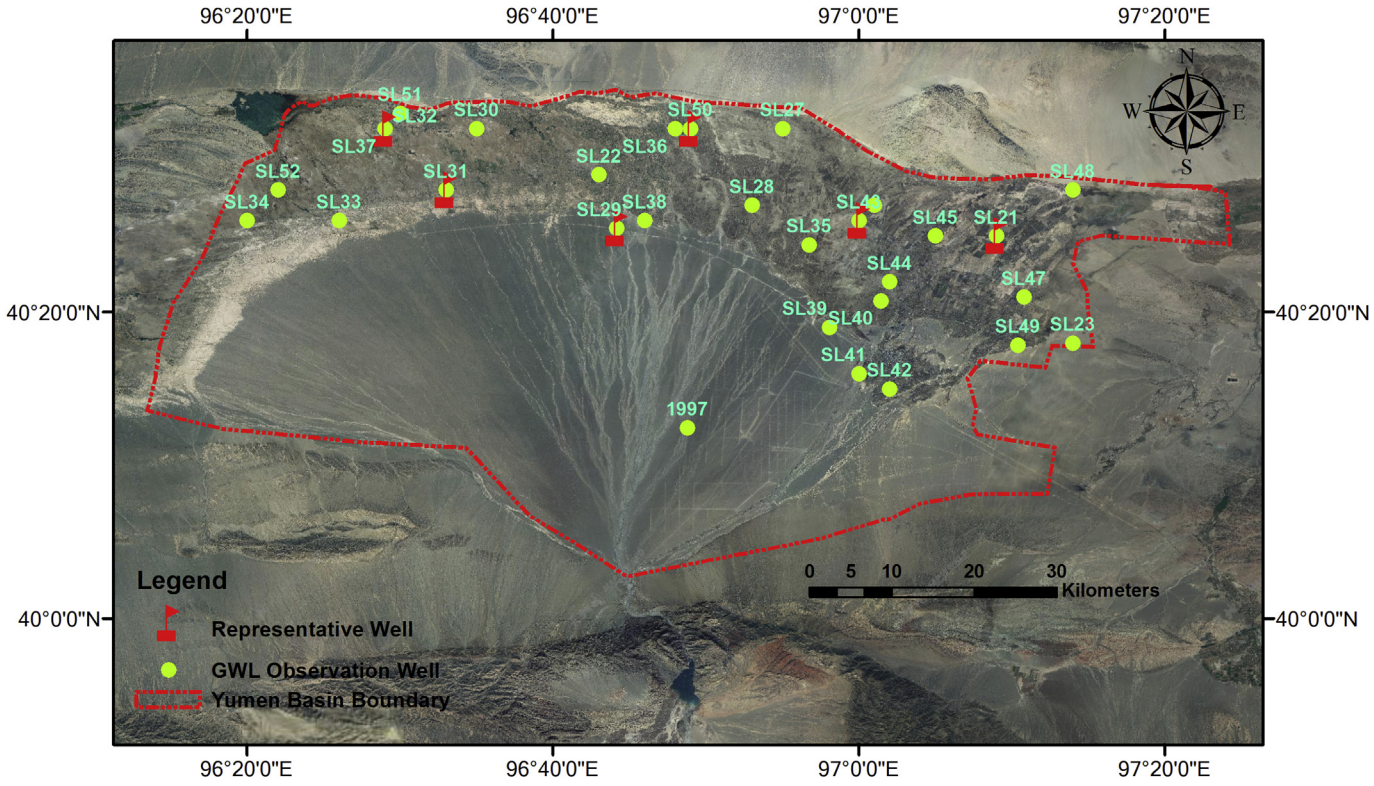

Fig. 3. Distribution of groundwater level observation wells across the Yumen Basin.

common applications. To apply SOM successfully, following procedures would be required (Kalteh et al., 2008).

i. Normalization. Normalization ensures that all variables have equal importance in the formation of SOM by transforming all the variables to the range of e.g. 0–1. In this study, twice normalization on groundwater level time series was carried out in order to avoid the varying effects due to different geographic locations and time steps. The GWL observations were firstly normalized for each observation well to eliminate the effects of

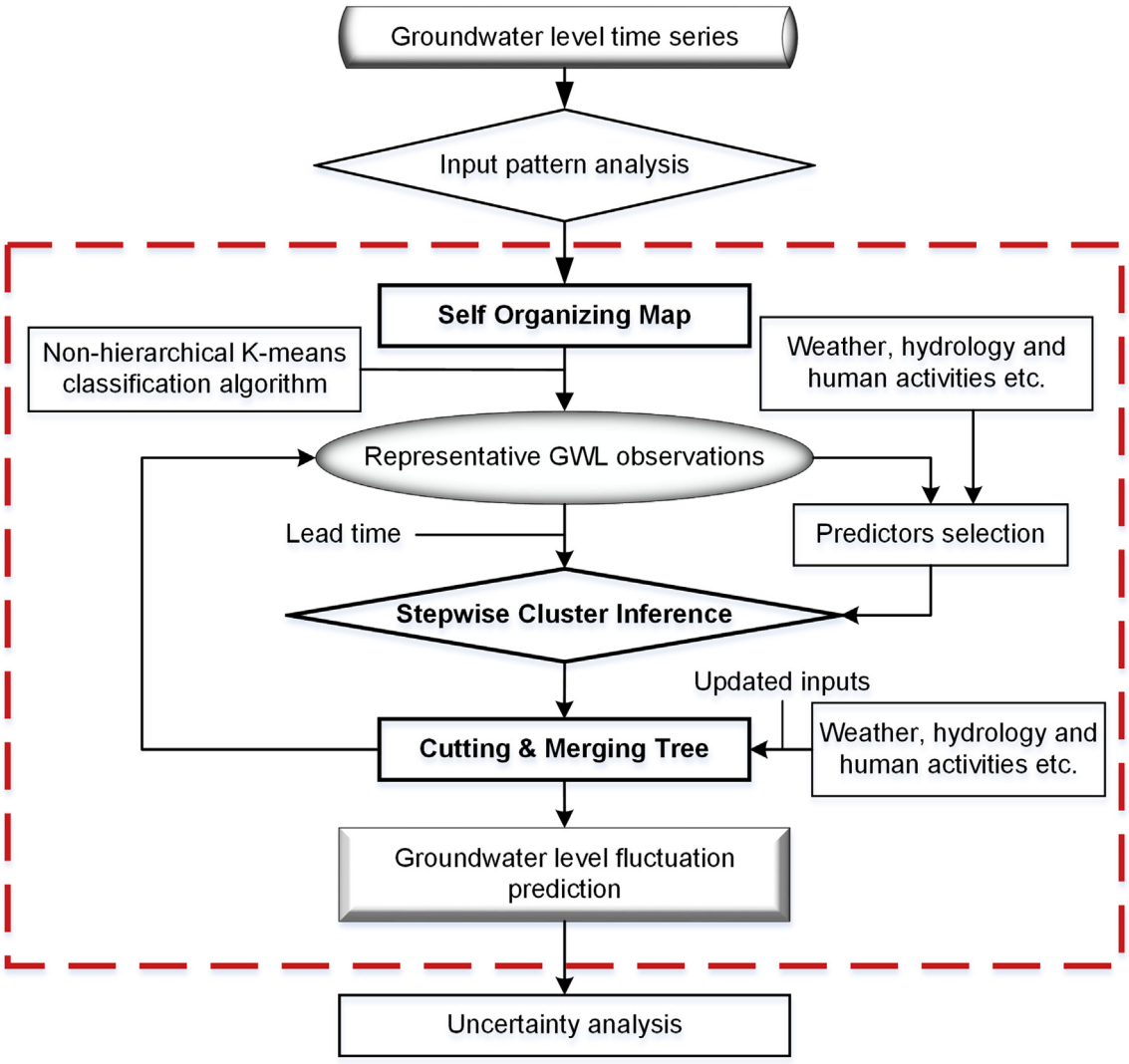

Fig. 4. Flowchart of SOM-aided stepwise cluster multisite inference model.

GWL elevations as a result of geographic locations, and then normalized at each time step to consider varying effects of the GWL distinctions at different time steps.

$$GWL_{t,i} = \frac{\left(GWL_t^i - GWL_{\min}^i\right)}{GWL_{\max}^i - GWL_{\min}^i} \quad (1)$$

$$GWL_{t,i}^* = \frac{\left(GWL_{t,i} - GWL_{t,\min}\right)}{GWL_{t,\max} - GWL_{t,\min}} \quad (2)$$

in which, i denotes the order of observation well; t is the time; GWL_t^i is the observed groundwater level for the ith observation well at time t; $GWL_{t,i}$ is the firstly normalized groundwater level for the ith observation well at time t; $GWL_{t,i}^*$ is the secondly normalized GWL for the ith observation well at time t; $GWL_{\min}^i$ and $GWL_{\max}^i$ are the minimum and maximum observed GWL at the ith observation well, respectively; $GWL_{t,\min}$ and $GWL_{t,\max}$ are the minimum and maximum groundwater levels for all observations at time t, respectively.

ii. Training. After normalization, a sample input vector from the data matrix is selected to iterative training procedures to form the SOM. The random weights are initially assigned, and the Euclidean distance between the input vector and weight w is then computed as follows.

$$\|y - w_j\| = \sqrt{\sum_{t=1}^{T}(y_t - w_{j,t})} \quad (3)$$

in which, j stands for the jth neuron. Through comparing each of the SOM neuron weight vectors with the present input pattern, a type of competitive and unsupervised learning is applied. The neuron with the closest match is called winner neuron or Best Matching Unit (BMU). Then, the weight vector of the BMU and its topologically neighboring neurons are updated by changing the weights at each training iteration to further reduce the distance between the weights and the input vector (Nourani et al., 2015). The most commonly used neighborhood function for updating the weights is the Gaussian (Kohonen, 1997). This training process is repeated until convergence. In this study, the SOM training was accomplished based on SOM Toolbox 2.0 (http://www.cis.hut.fi/somtoolbox/). Moreover, the Kohonen map has been chosen as a

hexagonal lattices with number of nodes determined according to Astel et al. (2007) and Yang et al. (2012).

iii. Application of the trained SOM map. Once the training is finished, post-processing can proceed based on the resulting SOM. Herein, it is employed to identify the spatial clusters of GWL fluctuations, and locate the representative piezometers. In order to achieve this, the non-hierarchical K-means classification algorithm is used to help determine the number of optimal clusters for the GWL piezometers, in which the David-Bouldin index is chosen as criterion (Davies and Bouldin, 1979). Then, the SOM lattices is subject to a second-level clustering into a given number of clusters, and subsequently, the Euclidean distance is calculated for each cluster to find the representative GWL well (central piezometer), which is closest to the centroid.

3.2. Stepwise cluster multisite inference model

According to the principle of SCA (Huang et al., 2006), it makes an effort to build a classification tree in the sense of probability based on a series of cutting or mergence processes under given statistical criteria (Sun et al., 2009). To obtain reliable and accurate GWL predictions, both empirical design and careful configurations would be very necessary when SCA is considered. Thus, a stepwise cluster multisite inference model based on SCA specifically for regional GWL prediction would be necessary. In accordance with SCA, the primary objective of proposed inference model is also to characterize the nonlinear relationships between multiple predictors and multisite GWL observations through generating clustering trees and classifying the observation samples into various groups (Li et al., 2015b).

The training samples should be prepared in advance to construct a cluster tree for prediction. Firstly, the training set is formed by combining all the predictors and predictands of n samples, which are then cut and merged progressively until convergence or criterion being violated. The essence of training process is actually to cut the training set into two and to merge two sets into one, step by step, and thus to result in a clustering tree. Accordingly, the training set is divided into many smaller sets, and n samples are finally classified. The clustering criterion used is the F test based on Wilks' likelihood-ratio criterion (Wilks, 1963). Through the cut-merge loop, the training procedure is completed when all hypotheses of further cutting or merging are rejected, and the cluster that cannot be cut or merged is called an end node. More details on the operation could refer to previous literature (Huang, 1992; Li et al., 2015b; Sun et al., 2009).

After the training is done, the derived cluster tree can be employed for prediction, where the chosen predictors are used progressively as the cutting criteria. When a new sample is being input, thus, the predictor values against the criteria would help reach an end node along a cutting and merging route. Moreover, the predictand value at each end node can be expressed as:

$$y_p = y_p^{\varsigma} \pm R_p^{\varsigma} \tag{4}$$

where y_p^{ς} is the derived mean value of predictand y_p, and R_p^{ς} stands for the radius of y_p for end node ς, calculated respectively as follows.

$$y_p^{\varsigma} = \frac{1}{n_{\varsigma}} \sum_{k=1}^{n_{\varsigma}} y_{p,k}^{\varsigma} \tag{5}$$

$$R_p^{\varsigma} = \left[\max\left(y_{p,k}^{\varsigma}\right) - \min\left(y_{p,k}^{\varsigma}\right)\right] \Big/ 2 \tag{6}$$

Correspondingly, the predictions at end node are not precise results, and various percentile values can also be obtained.

To perform GWL predictions for the study area of interest, many empirical variables are used as the candidate input predictors for the inference model based on the available data. Due to the GWL observations collected at the beginning of each month, the inference model is trained to predict the GWL elevations 30 days into the future at the central piezometers. As shown in Table 1, 32 candidate input predictors at a monthly step are considered as well as antecedent GWL observations for central piezometers. Needless to say, interactions among input predictors would lead to significant influence on the predictions. To address such dilemma, the trial and error method coupled with empirical evidence are adopted to further determine input predictors as well as to configure the models, in which the Nash-Sutcliffe efficiency is chosen as the criteria.

3.3. Autoregressive (AR) error model

Model errors always reflect the limitations of models in mimicking the hydrological processes as well as inaccuracies in data used to drive the models (Li et al., 2015a), so post-processors or updating procedures are often applied to correct forecasts based on the latest available observations and their difference from simulations. As a popular approach for updating forecasts, autoregressive (AR) error models investigate the “memory” of errors in hydrological simulations and use a simple linear regression function to estimate errors in a forecast period based on the known errors at previous time steps (Brockwell and Davis, 1996). Then, forecasts are updated according to the estimated errors. Let y_t and y_t^M represent the observed and simulated GWL, respectively. A lag-1 AR (AR(1)) model can be formulated as:

$$\varepsilon_t = y_t - y_t^M \tag{7}$$

$$\varepsilon_t = a \cdot \varepsilon_{t-1} + b \tag{8}$$

in which, ε_t represents the residual of y_t and y_t^M; a and b denote the coefficients of linear regression function.

The AR model is considered to update the GWL predictions, and the accuracy is expected to increase through incorporating the AR error model with the stepwise cluster inference model. To allow for hydraulic characteristics in the irrigation area and the lead time for GWL prediction, the AR(1) model would be applied in this study.

3.4. Model performance indicators

The data was split into two series of samples for training and validation, corresponding to 1998–2005 and 2006–2010, respectively. Comparisons of training and validation simulations against observations were then carried out, respectively. Moreover, model performance was evaluated both visually using the hydrographs/regressive chart and statistically for its general “goodness of fit” in terms of Nash-Sutcliffe Efficiency (*NSE*), Correlation coefficient (*R*), Deviation of Volume (*DV*) and Root Mean Square Error (*RMSE*).

In order to analyze the simulation uncertainty, the upper and lower bounds of predictions as well as the 50 Percent Prediction Uncertainty (50PPU) would be visualized in the GWL hydrographs (Han et al., 2014a). Besides, two indicators are used to evaluate the uncertainty, i.e., Average Relative Interval Length (*ARIL*) and the percent of observations bracketed by the prediction band (*PCI*) as follows (Jin et al., 2010).

Table 1
List of candidate input predictors.

Label	Parameter	Label	Parameter	Label	Parameter
X1	Monthly precipitation, DH*	X12	Monthly sunshine hours, YM	X23	Monthly average runoff, Yumen hydrometric Station
X2	Monthly average daily air temperature, DH	X13	Monthly sunshine ratio, YM	X24	Monthly inflow into Changma Reservoir
X3	Monthly minimum daily air temperature, DH	X14	Monthly average relative humidity, YM	X25	Monthly outflow from Changma Reservoir
X4	Monthly maximum daily air temperature, DH	X15	Monthly precipitation, GZ*	X26	Monthly inflow into Chijinxia Reservoir
X5	Monthly sunshine hours, DH	X16	Monthly average daily air temperature, GZ	X27	Groundwater extraction in south of the irrigation district
X6	Monthly sunshine ratio, DH	X17	Monthly minimum daily air temperature, GZ	X28	Groundwater extraction in east of the irrigation district
X7	Monthly average relative humidity, DH	X18	Monthly maximum daily air temperature, GZ	X29	Groundwater extraction in north of the irrigation district
X8	Monthly precipitation, YM*	X19	Monthly sunshine hours, GZ	X30	Groundwater extraction in west of the irrigation district
X9	Monthly average daily air temperature, YM	X20	Monthly sunshine ratio, GZ	X31	Groundwater extraction in the main canal area
X10	Monthly minimum daily air temperature, YM	X21	Monthly average relative humidity, GZ	X32	Total groundwater extraction in the irrigation district
X11	Monthly maximum daily air temperature, YM	X22	Monthly runoff, Changma Station	X33~	Antecedent GWL observations at representative wells

Note: DH, YM and GZ stand for Dunhuang, Yumen and Guazhou Station.

$$ARIL = \frac{1}{n}\sum \frac{Limit_{Upper,t} - Limit_{Lower,t}}{R_{obs,t}} \quad (9)$$

where $Limit_{Upper,t}$ and $Limit_{Lower,t}$ are the upper and lower boundary values of the predictand; n is the number of time steps; $R_{obs,t}$ is the observed GWL elevation.

$$PCI = \frac{NQ_{in}}{n} \times 100\% \quad (10)$$

where NQ_{in} is the number of observations which are contained in the prediction interval bracketed by the upper and lower bounds.

4. Results & discussions

4.1. Results of SOM

For the SOM output layer, hexagonal discrete lattices are usually preferred for visualization (Kalteh et al., 2008). In order to accomplish spatial clustering of GWL piezometers, a lattice with 6 × 5 hexagons was utilized to illustrate the similarities of GWL observations for 30 wells in the study area (Fig. 3). In this study, a 30 × 156 matrix (i.e., 156 normalized GWL records for each piezometer from 1998 to 2010) comprised the high dimensional inputs into the SOM training process. Fig. 5(a) shows the neighbor weight distances obtained by the two dimensional SOM output layer. The dark hexagons represent the BMU neurons with one connecting input vector at least, indicating that a GWL well or more wells would be clustered as shown in Fig. 5(b). In addition, the colors in the regions denote the relative distances between adjacent neurons. Thus, a total of nineteen BMUs or clusters are first generated based on thirty series of the GWL observations, and Fig. 5(b) presents the resulting SOM with neurons marked by the observation wells. Therefore, it can be found that several GWL piezometers are directly grouped onto a hexagonal neuron, indicating high similarities between them. According to the obtained neighbor weight distance, the smaller it is, the more likely the adjacent neurons can be clustered. Hence, further clustering of the SOM neurons could be undertaken based on the neighbor weight distances, but the appropriate number of clusters should be necessarily designated herein.

With the help of non-hierarchical K-means classification algorithm, the Davies-Bouldin index was calculated to determine the optimum number of clusters, which was set in the range of [2,19] according to the SOM results. A lower value of Davies-Bouldin index indicates that the clustering is better, so the best clustering scheme essentially minimizes the Davies–Bouldin index (Davies and Bouldin, 1979). To achieve robust simulations, 1000 runs of non-hierarchical K-means clustering operation were implemented simultaneously. As a result, 1000 clustering schemes were ultimately derived. Accordingly, frequencies for the resulting number of clusters were computed as illustrated in Fig. 6, and hence, the optimum number of clusters would be 6 for the GWL piezometers.

The GWL piezometers could be then classified into the given number of clusters as shown in Table 2. Regarding the specific locations of piezometers, the clusters seem to have no direct relationship with the direction of main stream flow (Fig. 3), which has been being altered for several years due to construction of irrigation canals and increasing human activities in the irrigation district, especially for unregulated pumping extractions. Nevertheless, interactions between groundwater and surface water including springs cannot be neglected as well. In order to identify predominant piezometers, which can best represent the GWL fluctuation pattern of the study area, the Euclidean distance was adopted to identify the central piezometer for each cluster. As a consequence, six central piezometers were then obtained for reflecting regional GWL fluctuations (Table 2). Therefore, emphasis should be imposed on them to understand the relationships between groundwater resources with economic and social behaviors and ecological processes for reliable and effective groundwater management of the plain. Also, these six central piezometers were focused to demonstrate stepwise cluster GWL predictions.

4.2. GWL prediction

With respect to six central piezometers, antecedent GWL observations for SL29, SL37, SL43, SL31, SL21 and SL50 were also assigned as candidate predictors X33~X38, respectively. Thus, a total of 38 predictors were used as candidates for the inputs into stepwise cluster inference model. In addition to predictors, two independent parameters should be set before the training of stepwise cluster inference model. The first parameter is the minimum number of samples ($N_{\min}$) for the end node, and the second is the level of significance for the F-test. Through trial and test in conjunction with empirical evidence, appropriate set of candidate predictors was ultimately selected as inputs for stepwise cluster inference. Hence, six stepwise cluster inference models were established for the central piezometers, and six cluster trees were built accordingly. Fig. 7 illustrates an example of the resulting cluster tree for well SL50. As presented in the figure, there are totally 22 end nodes, which are highlighted with blue background, and under each end node, the values of predictand are expressed as mean ± radius as described above. As for nodes which are split into two nodes at lower layer, the cutting criteria is also notated as to which predictor is applied. For instance, Node #1 is cut into Nodes #2 (left branch) and #3 (right branch) by comparing the normalized value of X4 with 0.92,927. When the normalized value is less than 0.92,927, the left branch will be taken, and otherwise, the right

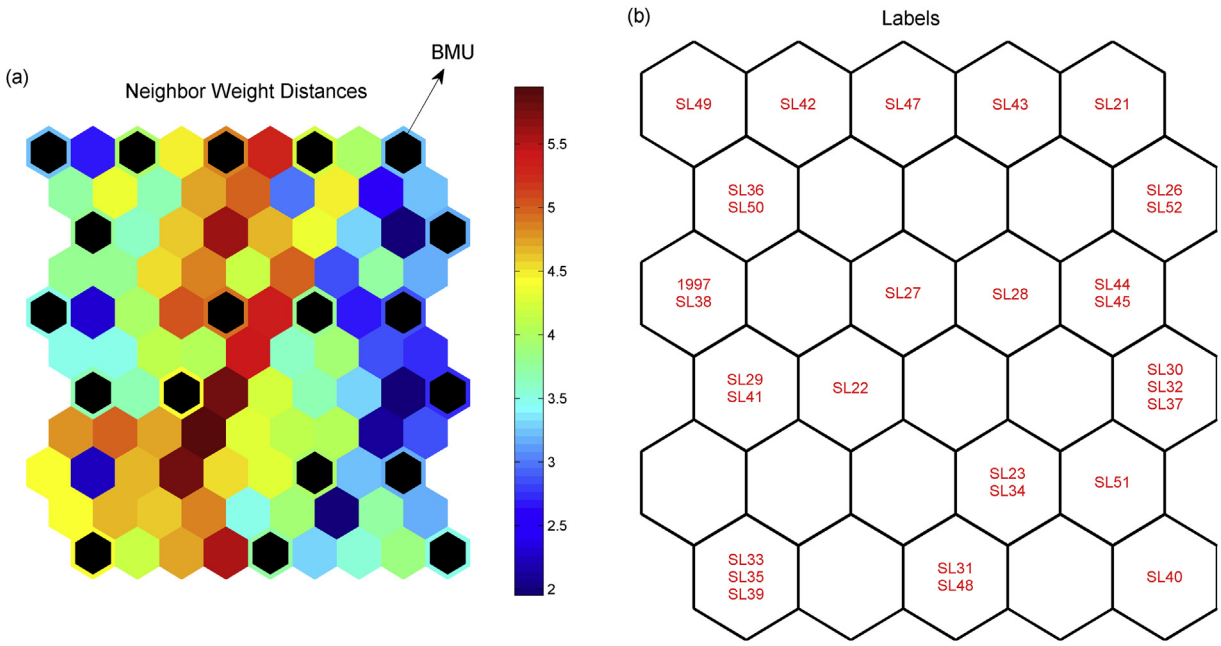

Fig. 5. 2-Dimensional SOM analysis of GWL data. (a) SOM neighbor weight distance; (b) projected observation wells.

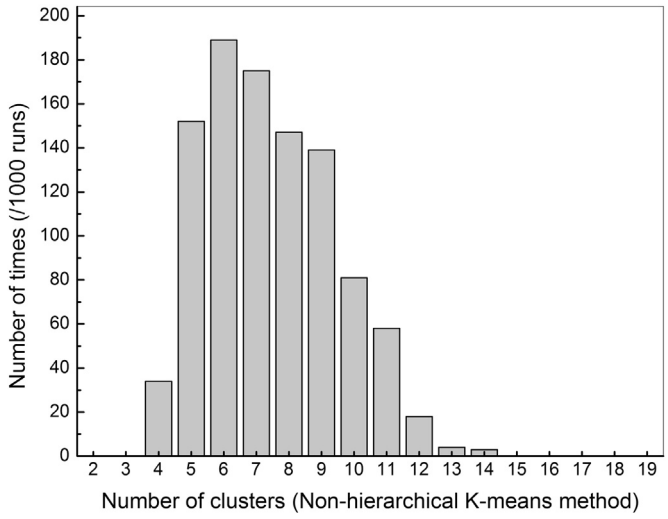

Fig. 6. Optimal clusters based on K-means clustering algorithm (1000 runs).

Table 2
The resulting classification of GWL observations and corresponding central piezometer for each cluster.

Cluster	GWL piezometers						Central piezometers
1	SL29	SL41	SL22	SL33	SL35	SL39	SL29
2	SL28	SL44	SL45	SL30	SL32	SL37	SL37
3	SL47	SL43	SL27				SL43
4	SL23	SL34	SL51	SL31	SL48	SL40	SL31
5	SL21	SL26	SL52				SL21
6	SL49	SL42	SL36	SL50	1997	SL38	SL50

branch is taken until the end node.

Table 3 shows the final settings of stepwise cluster inference models and the total number of end nodes for the resulting cluster trees as well as the behavioral predictors, which are defined herein as the predictors playing a part in GWL forecasting according to the cluster tree. The minimum number of samples for clusters at each end node is required to be greater than 2, but it was found that no better predictions could be ensured to produce the maximum end nodes through keeping a minimum number as 3. Especially for SL31, of which the determined minimum number is 6, only 11 clusters/end nodes were generated for prediction. Hence, on the one hand, it is difficult to accurately capture the fluctuation pattern of GWL for SL31 compared to other piezometers, indicating high uncertainty for the GWL predictions. On the other hand, it further reveals that the simulation efficiency was significantly subject to the observation data since the proposed model is data-driven. Consequently, it would reflect that high uncertainties exist in the GWL observations due to complex interactions among multiple influencing factors. Nevertheless, the proposed method still demonstrates its superiority relative to predictions only using the mean value of observations because the best model performance was obtained at a minimum number of not greater than 6 rather than only one cluster for the predictand. As discussed above, to increase the level of significance for the *F*-test would help lead to more cutting, and as a result, more clusters/end nodes for predictions would be probable. Besides, merging operations might be reduced, which would also contribute to more end nodes for prediction. Based on the results as shown in Table 3, increasing stepwise clusters for GWL prediction would be not positive for improving the accuracy of predictions, and again, there exists an optimal level of clusters for GWL time series prediction. When it comes to the behavioral predictors, the resulting input structures for central piezometers differed a lot from each other, suggesting distinct GWL fluctuation patterns for six clusters of GWL wells. However, it should be worth emphasizing that the predictors are subject to screening before they are used during training. Therefore, the interactions among candidate predictors still need to be discussed according to the implications of predictors on GWL predictions.

Fig. 8 indicates the simulation results of GWL time series for six central piezometers, which are demonstrated for both the training and validation period. As shown in the figure, five statistics for the predictand (the minimum, maximum, mean, 25th and 75th percentile values of predictions) were obtained against the observations, which are denoted by the red dots. The mean predictions

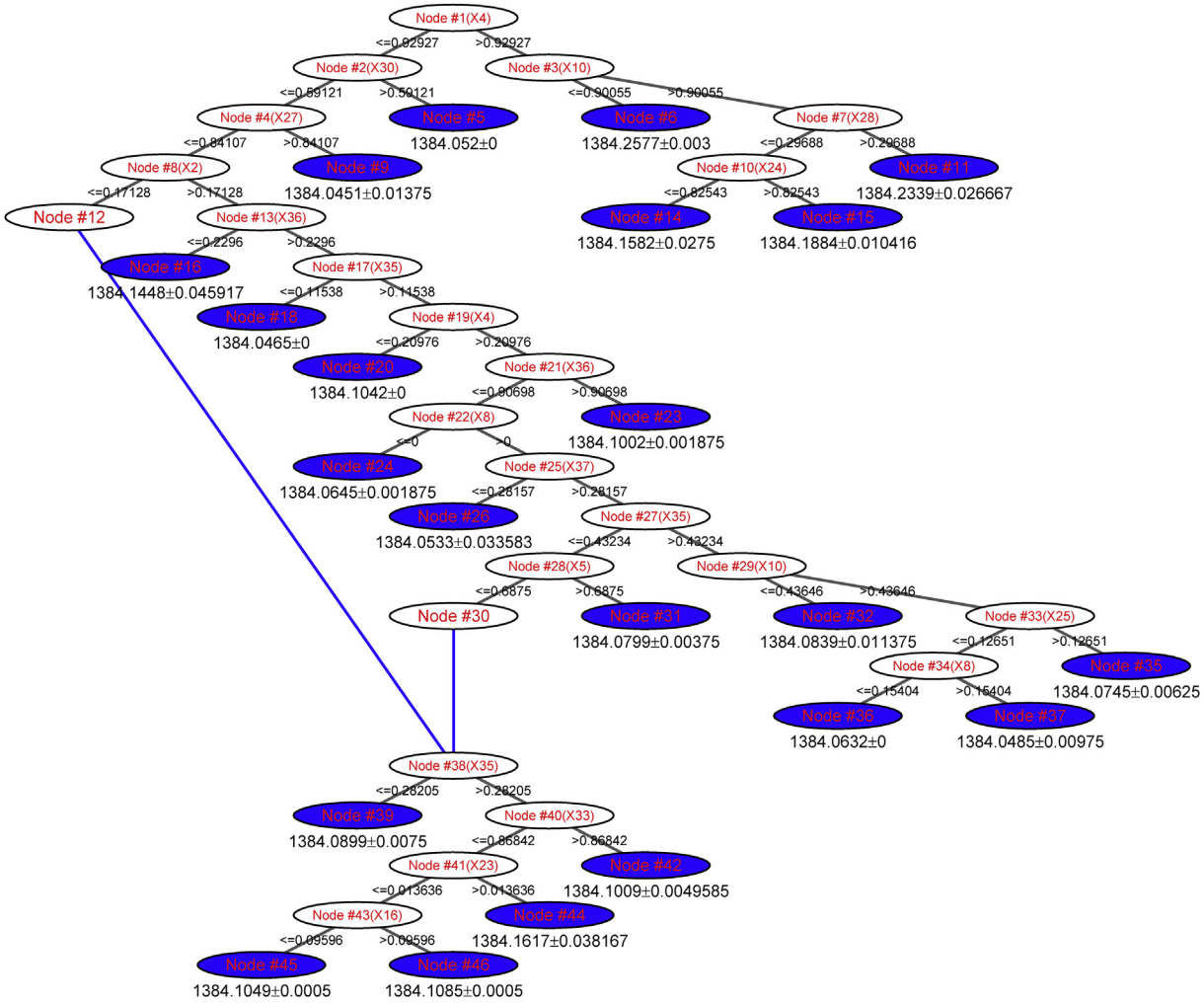

Fig. 7. The resulting stepwise cluster tree (take SL50 for example).

Table 3
Results of stepwise cluster tree.

Piezometer	N_{min}	Significance level	End node	Behavioral predictors
SL21	3	0.01	15	X33, X28, X23, X31, X20, X29, X4, X15, X7, X17, X26, X20
SL29	4	0.01	14	X34, X33, X2, X27, X26, X6, X7
SL31	6	0.01	11	X31, X35, X36, X18, X3, X19, X23, X6, X9, X25, X38, X13
SL37	5	0.1	14	X2, X36, X4, X34, X21, X7, X19, X28, X20
SL43	4	0.05	17	X29, X38, X31, X26, X27, X33, X36, X24, X25, X28, X6, X19, X2, X7
SL50	3	0.05	22	X4, X30, X10, X27, X28, X2, X24, X36, X35, X8, X37, X5, X25, X33, X23, X16

and corrected predictions with the AR(1) model are depicted as triangle and fork symbols, respectively. In addition, the maximum and minimum values were described as the upper and lower bounds, which are shown using two red lines in the figure. Besides, the grey interval band of 50PPU was used to reflect the intervals between 25th and 75th percentiles.

According to the GWL hydrograph, the stepwise cluster inference model accurately reproduced the rising and falling tendencies of GWL elevations. However, the range of GWL fluctuations is highly dependent on the observation sites, and the training results confirms a good match between observations and simulations when the mean value (green triangle) was chosen to predict the GWL elevations. Thus, the mean value would be a reliable alternative for GWL prediction in comparison with the other statistics. By contrast, it seems that no comparable simulations were produced for the validation period relative to the training, and moreover, predictions are relatively poor for SL31. It might be attributed to complex hydrogeological conditions and varying human activities, which were not reflected sufficiently in the model (Han et al., 2014b). Besides, well pumping and groundwater extraction was not effectively considered, and unauthorized wells have also been found to be digged and operated by the farmers in this region. Such influences on hydrologic regime cannot be accurately represented in the model due to lack of reliable records and data availability. Especially for piezometers with small range of GWL fluctuations such as SL50, GWL has a range of fluctuations not greater than 0.3 m for the training period, while the maximum amplitude of GWL fluctuations for the validation period is nearly as high as 0.5 m.

Some observations are even beyond the upper/lower bounds, signifying inconsistent hydrologic regime of GWL. The identical phenomenon would also be found for SL43, SL37 and SL31. Hence, it is too difficult to accurately represent the varying pattern of GWL fluctuations without sufficient consideration of inconsistent GWL changes. Nevertheless, the primary changing trend of GWL could still be captured at a satisfying level through applying the mean value for prediction based on the evaluation indicators in Table 4. In order to improve the predictions, the AR(1) error model was then integrated with the inference model to correct the predictions, and the corresponding results were represented by the asterisk symbols in Fig. 8. As for the simulation results, the newly obtained predictions apparently approach closer to the observations, especially for the validation period. Consequently, such post-processing procedure as the AR(1) model would help to increase the modeling efficiency of GWL prediction. In addition to the illustration of GWL hydrograph, the model performance were further evaluated for both training and validation. As shown in Table 4, the predictions based on the mean values corrected with the AR(1) model were compared with those using the mean value. According to the recommendations of Moriasi et al. (2007), the mean values corrected with the AR(1) model for the simulation error would result in "good" GWL predictions for all the central piezometers. By contrast, estimation of the GWL elevation using only the mean value is relatively poor, although most of them for the training period are very good. For instance, the obtained *NSE* values for the GWL prediction of well SL31 for training and validation are 0.917 and −0.29, respectively. While after AR correction, the values become 0.862 and 0.767. In the meanwhile, the value of RMSE for validation decreases from 0.206 to 0.087 m. Such issue, on one hand, could be

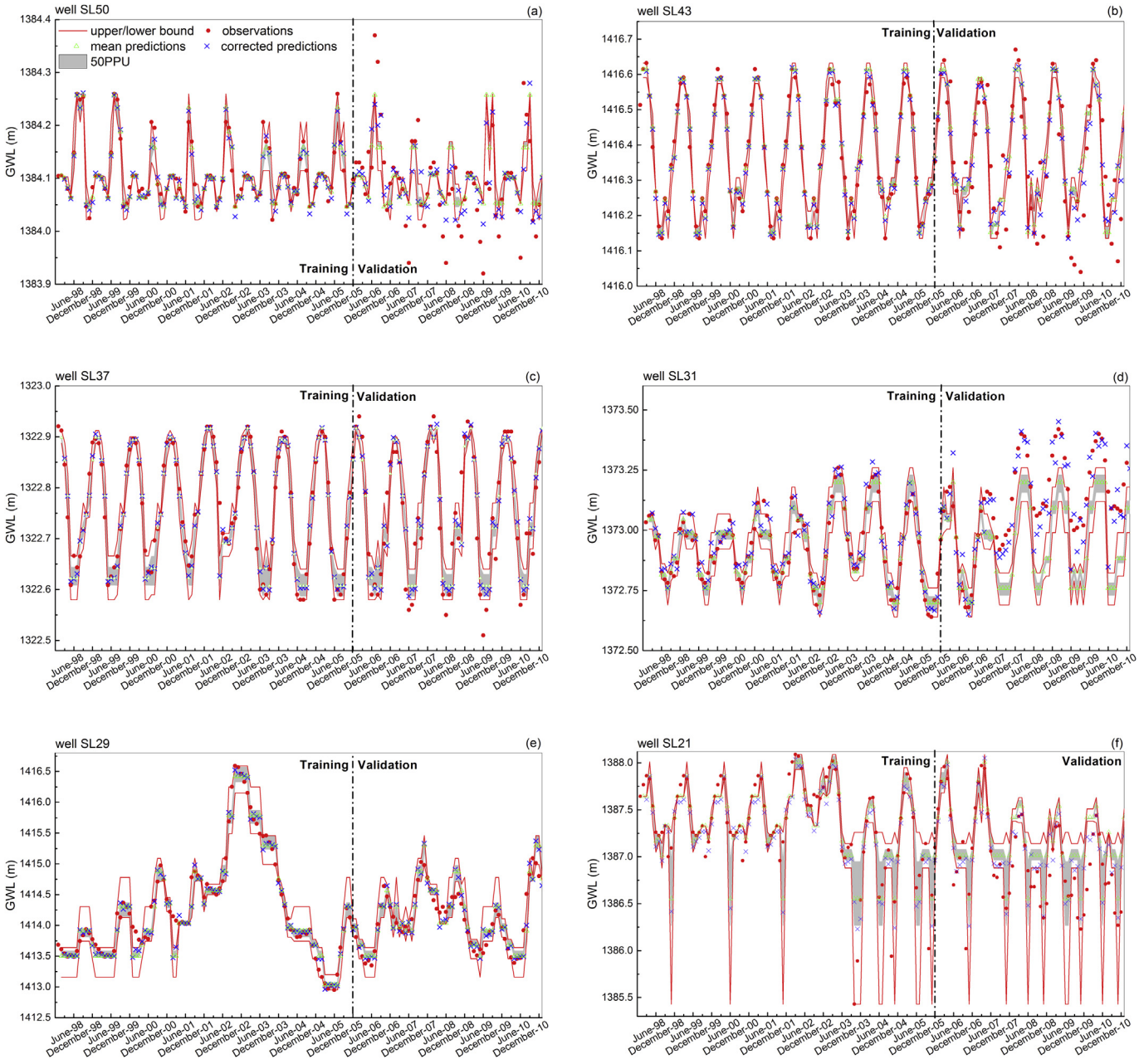

Fig. 8. GWL simulation results through stepwise cluster inference for six representative piezometers.

attributed to the limits of assigned training samples, which might be inadequate for demonstrating future changes of GWL elevations. On the other hand, inconsistent hydrologic regime would lead to such deficiency as previously discussed. Besides, short-term transient pumping changes, which were not captured in the groundwater extraction records, would also cause apparent predictive errors. Regardless of slight or even negative improvements on GWL predictions for SL29, SL37 and SL43, it could still be generalized that more reliable and robust predictions were obtained using the corrected mean values based on the evaluation of model performance. Therefore, incorporation of the AR(1) model would contribute to better GWL predictions into the future.

Fig. 9 illustrates predictions using the mean values as well as the corrected values with the AR(1) model against the observed, and corresponding fitting lines are also plotted for the entire simulation period. Besides, a solid black line of 45° is added for reference. The closer the fitted line gets to 45°, the better the prediction would perform. Thus, the GWL prediction using the corrected mean values represented as the red dash line seems to be more reliable than the dark dash line. Moreover, significant improvements on GWL predictions would be easily observed for SL21 and SL31. As for other piezometers, general contribution of the AR(1) error model to GWL prediction was also positive although it seems to be negligible.

4.3. Uncertainty analysis

Such indicators as *ARIL* and *PCI* were used to analyze the simulation uncertainty of the stepwise cluster inference model. According to the obtained results as indicated in Table 5, GWL predictions for SL21 and SL29 show high uncertainty relative to other piezometers. It is mainly caused by remarkable changes of GWL elevations during the simulation period. Nonetheless, the values of *ARIL* are not greater than 0.47 m. It is interesting that the *ARIL* value decreases generally from east to west as delineated by SL21-SL37 and SL43-SL50 (Fig. 3), which is consistent with the direction of surface water flow. In addition, the GWL fluctuations for SL43 and SL50 in the north were also found to be lower than those for piezometers in the southwest. However, special attention should be paid to GWL changes of well SL21, which has a distinct difference from its physically neighboring piezometer of SL43. It might be caused by complex interactions between groundwater and surface water, i.e. canal construction and irrigation operation, which has dramatically altered the natural flow regime. By contrast, the predictions for SL29, situating at edge of the alluvial fan, also present prominent uncertainty. Such results would be attributed to frequent exchanges between surface water and groundwater, which often overflows as springs.

PCI is always used as an indicator to evaluate the reliability of model. From Table 5, the *PCI* value varies from well to well, and notable differences could also be observed for training and validation. In comparison with satisfying *PCI* levels for the training period, the obtained *PCI* values for the validation period seem to be relatively poor, especially for SL50 and SL31. Thus, other than SL29 and SL37, the modeling reliability for predicting the GWL into the future is quite susceptible to inconsistent fluctuations of GWL due to such changes as land cover and groundwater usage. Correspondingly, prediction uncertainty in the future should be carefully addressed as a result of changing environmental conditions. It would further suggest that post-processing procedures for accurate GWL predictions would be crucial. Besides, it is worth mentioning that hydrologic predictions using such black box models were highly dependent on historic records. As such, it would result in modeling deficiencies when varying and different conditions occurred compared to those during calibration. Accordingly, promising alternatives would be recommended to assimilate both the advantages of black box models and process-based models to make useful predictions.

In the process of stepwise cluster inference, the appropriate set of predictors plays an important part in improving the model's efficiency for both steps of training and validation. As aforementioned, 38 candidate predictors were prepared for selection as inputs to the stepwise cluster inference model. When all the candidate predictors were used as inputs, inappropriate predictors might take the roles of behavioral predictors, and lead to unreasonable inference. To get rid of adverse effects of noisy predictors, therefore, screening process was firstly required to select approximate predictors. For instance, to predict GWL fluctuations for SL21, predictors such as antecedent GWL observations for other piezometers should be removed. If not, predictions for SL21 would become relatively poor due to the introduction of these predictors as cutting criteria. It is consistent with the conclusions of Chen et al. (2010), who found no better performance of six-site model was achieved than that of four-site model, and too much information could not help improve the generalization ability of the GWL prediction model. Thus, an experience based trial and error method was applied to help determine behavioral predictors in this study (Table 3). For each cutting operation of stepwise cluster training, only one predictor would be determined as the best criteria for splitting the clusters into two sub-clusters. Once the introduced noisy predictor might be chosen as behavioral, modeling efficiency or reliability would diminish because of insufficient accounting of interactions among predictors. According to the resulting cluster trees for GWL elevation predictions using only behavioral predictors and all 38 candidate predictors as inputs, it was found that the modeling efficiency decreased to various extents. Taking the *F*-test for samples based on the predictand values into consideration, both the minimum number of sample for end node and the level of significance have significant influences on the model performance. Li et al. (2015b) investigated the sensitivity of a daily flow prediction model on such parameters, and found that both of them have significant impacts on the shaping of the cluster tree, and should be carefully set to obtain accurate simulations, although various

Table 4
Evaluation of stepwise cluster inference model performance for both training (*C*) and test (*V*).

Prediction	Indicator	Sl21		SL29		SL31		SL37		SL43		SL50	
		C	*V*	*C*	*V*	*C*	*V*	*C*	*V*	*C*	*V*	*C*	*V*
Mean value corrected by AR(1)	*NSE*	0.688	0.617	0.958	0.813	0.862	0.767	0.944	0.853	0.976	0.662	0.917	0.523
	R	0.840	0.813	0.980	0.917	0.939	0.904	0.972	0.924	0.988	0.815	0.959	0.726
	DV (%)	−0.0045	0.007	−0.0015	0.0024	0.0011	−0.0017	0	0.00012	−0.00037	0.00058	−0.00021	0.00032
	RMSE	0.277	0.290	0.174	0.204	0.0539	0.0878	0.0253	0.0486	0.024	0.111	0.0166	0.0600
Mean value	*NSE*	0.714	0.449	0.956	0.805	0.917	−0.29	0.942	0.849	0.976	0.636	0.929	0.417
	R	0.846	0.798	0.978	0.910	0.958	0.786	0.971	0.922	0.988	0.803	0.964	0.657
	DV (%)	−0.00012	0.015	−0.0020	0.0029	−0.0003	−0.012	−0.00016	0	0	0.0012	−0.0001	0.0007
	RMSE	0.265	0.348	0.178	0.209	0.0419	0.206	0.0259	0.049	0.0236	0.115	0.0154	0.066

Note: *C* and *V* stand for training and validation, respectively.

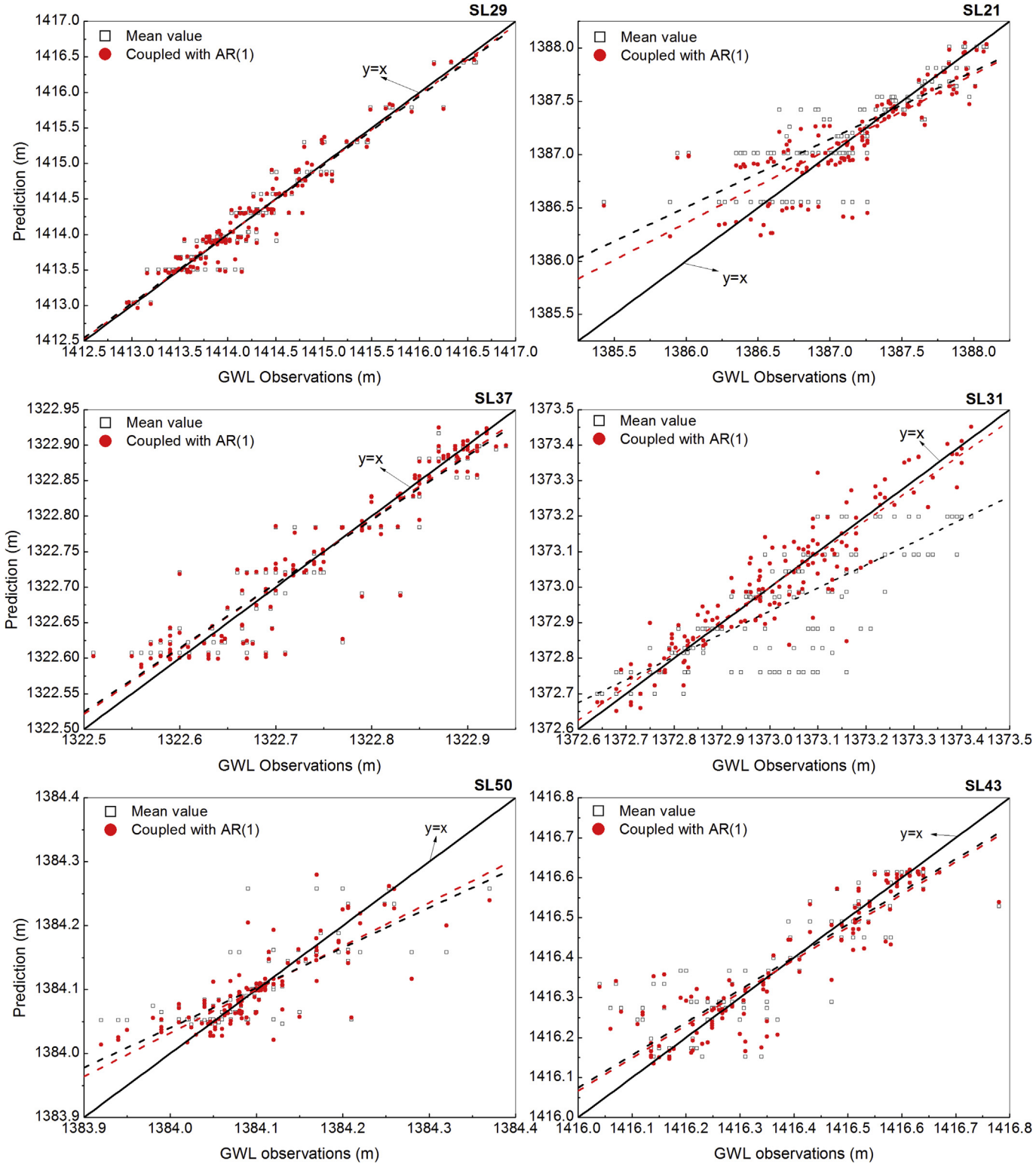

Fig. 9. GWL predictions against observations using mean value coupled or not with AR(1) model.

combinations might produce the same cluster tree for predictions due to parameter uncertainty.

5. Summary

In this study, a SOM-aided stepwise cluster multi-site inference model was developed to make regional GWL predictions a month ahead. To represent various patterns of GWL fluctuations, a self-organizing map coupled with non-hierarchical K-means classification method was applied to identify the optimal clusters for GWL observation wells as well as the central piezometers. According to the available data, there were totally 38 candidate predictors used

Table 5
Uncertainty analysis of the GWL simulation for the training (*C*) and test (*V*) as well as the entire simulation (*S*) period.

Piezometer	*ARIL*			*PCI*		
	C	*V*	*S*	*C*	*V*	*S*
SL21	0.3451	0.4612	0.3897	0.6042	0.4333	0.5355
SL29	0.4250	0.3855	0.4098	0.7812	0.6167	0.7226
SL31	0.1231	0.1343	0.1274	0.8021	0.11,667	0.5355
SL37	0.05621	0.05935	0.05742	0.6875	0.4333	0.5935
SL43	0.04469	0.04850	0.04615	0.4271	0.2333	0.3484
SL50	0.02505	0.02266	0.02413	0.5938	0.08333	0.40

to predict the GWL. Based on the obtained predictions, several conclusions would be generalized as follows.

(1) Six central piezometers were identified with varying GWL fluctuations in the Changma irrigation district. Then, six corresponding stepwise cluster inference trees were constructed to infer the range of GWL predictions through comparing the predictor values with the cutting criteria. It was found that the mean value of predictions would be employed as a reasonable alternative for GWL prediction, while incorporation of the AR(1) error model would further improve the predictions, rendering the predictions more accurate and reliable. Due to its simplicity and robustness, the proposed model would be a useful tool to predict the regional GWL fluctuations within a multisite forecasting framework. Thus, such accurate prediction capability would help develop appropriate wellfield operation policies while mitigating environmental impacts such as aquifer overdraft, streamflow depletion and downstream ecological flow based on reservoir operation. As a future work, the resulting inference modeling structure would be further extended to predictions for other piezometers. Using interpolation and estimation methods, in addition, sufficient GWL forecasting coverage across the study area of interest would be able to eventually support informed regional water resources decision-making.
(2) Various combinations of candidate predictors constitute the behavioral predictors for six kinds of GWL fluctuation pattern. However, predictors' screening procedure should be first implemented to eliminate noisy predictors, which would lead to irrational cutting in the cluster tree with regard to the chosen criteria. Consequently, more predictors being introduced would not ensure more reliable simulations. To determine behavioral predictors, such advanced correlation analysis methods as mutual information theory might be a feasible alternative, and further works will be advised on such issues in hydrological application. In addition to predictor selection, care should also be given to configurations of significance level and the minimum number for end node, which would impose significant influence on the structure of cluster tree as well as the resulting predictions.
(3) As a result of background changes that are hard to account for, inconsistent hydrological characteristics would make accurate predictions even harder based only on historical records. Although an AR(1) error model was applied to timely correct the predictions, background changes such as land cover changes due to human activities might not be effectively reflected by modeling efforts, and moreover, it is still difficult to capture many extreme values according to the validation results. In this regard, alternatives to improve model performance would require more understandings of the GWL time series, and hence, more deliberate preprocessing and post-processing procedures would be helpful and highly recommended in the future. Also, the prediction accuracy could further be improved through gaining more insights into groundwater systems as provided by data-driven and physical-based models.

Acknowledgement

This study was financially supported by the National Basic Research Program of China (973 Program 2013CB036402), National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2013BAB05B03), Research and Development Special Fund for Public Welfare Industry of the Ministry of Water Research in China (201501028) and China Postdoctoral Science Foundation (2015M571048). Also, the authors wish to acknowledge both the editors and reviewers for their valuable comments and suggestions on the improvement of the manuscript.

References

Adamowski, J., Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. J. Hydrol. 407, 28–40.
Astel, A., Tsakovski, S., Barbieri, P., Simeonov, V., 2007. Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. Water Res. 41, 4566–4578.
Brockwell, P.J., Davis, R.A., 1996. Introduction to Time Series and Forecasting. Springer, New York.
Chen, L., Chen, C., Pan, Y., 2010. Groundwater level prediction using SOM-RBFN multisite model. J. Hydrol. Eng. 15, 624–631.
Chen, L.H., Chen, C.T., Lin, D.W., 2011. Application of integrated back-propagation network and self-organizing map for groundwater level forecasting. J. Water Resour. Plan. Manag. 137, 352–365.
Coppola, E.A., Rana, A.J., Poulton, M.M., Szidarovszky, F., Uhl, V.W., 2005. A neural network model for predicting aquifer water level elevations. Ground Water 43, 231–241.
Coulibaly, P., Anctil, F., Aravena, R., Bobée, B., 2001. Artificial neural network modeling of water table depth fluctuations. Water Resour. Res. 37, 885–896.
Dash, N., Panda, S., Remesan, R., Sahoo, N., 2010. Hybrid neural modeling for groundwater level prediction. Neural Comput. Applic 19, 1251–1263.
Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. 1, 224–227.
Fan, Y.R., Huang, W., Huang, G.H., Li, Z., Li, Y.P., Wang, X.Q., Cheng, G.H., Jin, L., 2015. A stepwise-cluster forecasting approach for monthly streamflows based on climate teleconnections. Stoch. Environ. Res. Risk Assess. 29, 1557–1569.
Han, J.-C., Huang, G.-H., Zhang, H., Li, Z., Li, Y.-P., 2014a. Bayesian uncertainty analysis in hydrological modeling associated with watershed subdivision level: a case study of SLURP model applied to the Xiangxi River watershed, China. Stoch. Environ. Res. Risk Assess. 28, 973–989.
Han, J.-C., Huang, G., Huang, Y., Zhang, H., Li, Z., Chen, Q., 2015. Chance-constrained overland flow modeling for improving conceptual distributed hydrologic simulations based on scaling representation of sub-daily rainfall variability. Sci. Total Environ. 524–525, 8–22.
Han, J.-C., Huang, G., Zhang, H., Li, Z., Li, Y., 2014b. Heterogeneous precipitation and streamflow trends in the Xiangxi river watershed, 1961–2010. J. Hydrol. Eng. 19, 1247–1258.
Hsu, K.-C., Li, S.-T., 2010. Clustering spatial–temporal precipitation data using wavelet transform and self-organizing map neural network. Adv. Water Resour. 33, 190–200.
Hsu, K.-l., Gupta, H.V., Gao, X., Sorooshian, S., Imam, B., 2002. Self-organizing linear output map (SOLO): an artificial neural network suitable for hydrologic modeling and analysis. Water Resour. Res. 38, 38–31-38-17.
Huang, G., 1992. A stepwise cluster analysis method for predicting air quality in an urban environment. Atmos. Environ. Part B. Urban Atmos. 26, 349–357.
Huang, G.H., Huang, Y.F., Wang, G.Q., Xiao, H.N., 2006. Development of a forecasting system for supporting remediation design and process control based on NAPL-biodegradation simulation and stepwise-cluster analysis. Water Resour. Res. 42 n/a-n/a.
Izady, A., Davary, K., Alizadeh, A., Nia, A.M., Ziaei, A.N., Hasheminia, S.M., 2013. Application of NN-ARX model to predict groundwater levels in the Neishaboor plain, Iran. Water Resour. Manage 27, 4773–4794.
Jin, X., Xu, C.-Y., Zhang, Q., Singh, V.P., 2010. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. J. Hydrol. 383, 147–155.
Kalteh, A.M., Hjorth, P., Berndtsson, R., 2008. Review of the self-organizing map (SOM) approach in water resources: analysis, modelling and application. Environ. Model. Softw. 23, 835–845.

Kohonen, T., 1995. Self-organizing Maps. Springer, Berlin, Heidelberg, New York.
Kohonen, T., 1997. Self-organizing Maps. Springer-Verlag, Berlin, Heidelberg.
Lábó, E., 2012. Validation studies of precipitation estimates from different satellite sensors over Hungary – analysis of new satellite-derived rain rate products for hydrological purposes. J. Hydrol. 468–469, 173–187.
Li, M., Wang, Q.J., Bennett, J.C., Robertson, D.E., 2015a. A strategy to overcome adverse effects of autoregressive updating of streamflow forecasts. Hydrol. Earth Syst. Sci. 19, 1–15.
Li, Z., Huang, G.H., Han, J.-C., Wang, X., Fan, Y., Cheng, G., Zhang, H., Huang, W., 2015b. Development of a stepwise-clustered hydrological inference model. J. Hydrol. Eng. 20 (10), 04015008.
Mohanty, S., Jha, M., Kumar, A., Sudheer, K.P., 2010. Artificial neural network modeling for groundwater level forecasting in a river island of eastern India. Water Resour. Manage 24, 1845–1865.
Moosavi, V., Vafakhah, M., Shirmohammadi, B., Behnia, N., 2013. A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. Water Resour. Manage 27, 1301–1321.
Morawietz, M., Xu, C.Y., Gottschalk, L., 2011. Reliability of autoregressive error models as post-processors for probabilistic streamflow forecasts. Adv. Geosci. 29, 109–118.
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50, 885–900.
Nourani, V., Alami, M.T., Vousoughi, F.D., 2015. Wavelet-entropy data pre-processing approach for ANN-based groundwater level modeling. J. Hydrol. 524, 255–269.
Nourani, V., Baghanam, A.H., Adamowski, J., Gebremichael, M., 2013. Using self-organizing maps and wavelet transforms for space–time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling. J. Hydrol. 476, 228–243.
Nourani, V., Ejlali, R.G., Alami, M.T., 2011. Spatiotemporal groundwater level forecasting in coastal aquifers by hybrid artificial neural network-geostatistics model: a case study. Environ. Eng. Sci. 28, 217–228.
Qin, X.S., Huang, G.H., Chakma, A., 2007. A stepwise-inference-based optimization system for supporting remediation of petroleum-contaminated sites. Water Air Soil Pollut. 185, 349–368.
Salas, J.D., Delleur, J.W., Yevjevich, V., Lane, W.L., 1990. Applied Modeling of Hydrological Time Series. Water Resources Publications, Denver.
Sun, W., Huang, G.H., Zeng, G., Qin, X., Sun, X., 2009. A stepwise-cluster microbial biomass inference model in food waste composting. Waste Manag. 29, 2956–2968.
Tapoglou, E., Karatzas, G.P., Trichakis, I.C., Varouchakis, E.A., 2014. A spatio-temporal hybrid neural network-Kriging model for groundwater level simulation. J. Hydrol. 519 (Part D), 3193–3203.
Tarsitano, A., 2003. A computational study of several relocation methods for k-means algorithms. Pattern Recognit. 36, 2955–2966.
Wilks, S.S., 1963. Statistical Inference in Geology. Wiley, New York.
Yang, Y., Wang, C., Guo, H., Sheng, H., Zhou, F., 2012. An integrated SOM-based multivariate approach for spatio-temporal patterns identification and source apportionment of pollution in complex river network. Environ. Pollut. 168, 71–79.
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, 128–138.
Zahmatkesh, Z., Karamouz, M., Nazif, S., 2015. Uncertainty based modeling of rainfall-runoff: combined differential evolution adaptive Metropolis (DREAM) and K-means clustering. Adv. Water Resour. 83, 405–420.