SIMULATION METHODS OF SPRING DISCHARGE VARIATIONS IN KARST AREA

CHEN XI, LIU CHUANJI

Key Lab of Water Resources Development, Hohai University, Nanjing 210098, PR China

HU Zhongming, WANG Zhonggui

Anyang Water Resources Bureau, Henan, 45500, PR China

Spring discharge variations are closely related with rise and decline of groundwater table, which depends on rainfall, groundwater withdrawal and irrigated water from Hongqi canal in the Xiao-nan-hai spring catchment. From the 80s of the last century, the spring discharge declines rapidly due to increase of groundwater withdrawal for agriculture and industry utilization. The mean annual discharge reduces from 8.06 m$^3$/s in the 70s to 4.48 m$^3$/s in the 90s of the last century. For planning water utilization scheme to cease spring discharge decline, we should quantitatively measure spring discharge variations related with their influences. On the basis of analyzing spring discharge variation and its influences in the Xiao-nan-hai spring catchment, a BP neural network model in a structure of three dimension inputs (annual rainfall, groundwater mining and division water), one output (spring discharge) and two hidden neurons, and a time series model were established. Comparison of two models shows that the BP neural network model can be used to better simulate spring discharge in both calibration and validation periods because it reveals relationship of spring discharge and its influences. However, errors of simulated spring discharges by the time series model are much larger in the validation period than that in the calibration period, which indicate that time series model is not suitable in prediction of spring discharge affected by both natural and artificial conditions in the region.

INTRODUCTION

There are 3.44 million km$^2$ karst area in China. The total water resources in the karst area are approximately 2039 billion cubic meter per year, 23.4 percent of total groundwater in China. 91.6 percent of karst water is in the southern China and 9.4 percent in the northern China which concentrates in a small region with larger spring discharges and is very important to water resources utilization [1]. However, over-exploration of water resources in the kast regions has induced lower groundwater tables and thus greatly decrease of spring discharges, resulting in ecologic and environmental disasters. Therefore, analysis of influences to spring discharges and simulation of their variations are very important to water resources protection and planning in the karst areas.

Groundwater estimation and evaluation and simulation of spring discharges in the karst region could be based on a numerical method. Thus, plenty of observation data of groundwater tables and hydrogeological data from drilling holes should be available or
explored because of tremendously non-homogeneous and anisotropy aquifer in the karst regions. Alternatively, simulation of spring discharges could be based on non-structure methods with a less data requirement, such as a neural network and time series analysis. Spring discharge amount is usually dependent on regional groundwater table, precipitation recharge, groundwater withdrawal and irrigation. Therefore, if these data series are available, a neural network model could be used to simulate the spring discharge variations and their relationship with precipitation, groundwater withdrawal and streamflow in the canal. A time series model could be used for spring discharge simulation based on statistical characteristics of spring discharge series itself. These two models are able to reveal natural and artificial influences to spring discharge and to predict spring discharge variations as well.

Artificial neural networks are able to simulate mankind neural structures and functions and are suitable for identification and reflection of fuzzy information and nonlinear relationship [2]. Sun et al. [3] have succeed in application of this method in spring discharge simulation in Shanxi Jingchi.

Time series analysis has been widely used in engineering, meteorology, hydrology, earthquake [4]. Qian Jiazhong et al. [5] used one dimensional unstationary time series model to simulate precipitation variations for groundwater evaluation in Xuzhou.

In this study, an artificial neural work model and a time series analysis model are used to simulate spring discharge variations in the Xiao Nanhai catchment of Anyang city, Henan province of China. Comparison of simulated results between these two models is made for revealing pre-conditions of the model suitability for karst spring discharge predictions.

HYDROGEOLOGICAL CONDITIONS

Xiao-nan-hai spring catchment is situated in the transitional area between upheaval zone along the Taiheng Mountain and the Northern China plain sink zone with a big fracture of Tangxi Mountain in the east generally stretching to sinking blocks, structurally discontinuous from the west to the east. Limestone is exposed on the surface in the large area of lower hill area except the Lingzhou basin that karst water system is covered by surface soil. Rich fractures, shallow holes, sinks, and caves in the carbonate rock areas produce many water movement passages and large storage rooms. They join together forming two main runoff concentration zones. Groundwater flow out in the Heng River valley which is intercepted by the rich runoff zone. The outlet of Nan-hai Spring is situated at Heng River valley with a low ground surface elevation from 131 m to 135 m. The whole spring controlling region is 934.6 km², consisting recharging zone in the west mountain region and Linzhou basin, groundwater flowing zone in the middle lower mountain region, and discharging zone in the east nearby Xiao-nan-hai spring. Spring discharge from Xiao-nan-hai spring is a primary water source supplying water utilization for Anyang city. Therefore the spring is called as “life spring” in the region.

Annual mean precipitation is 574mm and annual mean potential evaporation is 942
Heng River stretches from the north to the south in the middle of the study region and a well-known artificial canal in the world - Hongqi canal is in the northern high mountain.

Spring discharge variations are closely related with groundwater rise and decline, which depend on rainfall, groundwater withdrawal and irrigated water from Hongqi canal.

Groundwater rises from September till the highest between October and December, two or three months lag behind the largest precipitation in a year. Spring discharge variations behave with similar patterns. From the 80s of the last century, the spring discharge declines rapidly due to increase of groundwater withdrawal for agriculture and industry utilization. The mean annual discharge reduces from 8.06m$^3$/s in the 70s to 4.48 m$^3$/s in the 90s of the 21th century (Fig. 1). Table 1 shows that spring discharge declines correspond to different increase amount of groundwater withdrawal.

Infiltration of canal water and irrigation water from Hongqi canal and Heng River are main sources of groundwater recharges. Influences of evaporation to the spring discharge can be neglect due to large depth to groundwater in the region.

Decrease of spring discharge has affected sustainable utilization of water resources, and existence of the “life spring”, which could seriously affect economic development in the region. Therefore, analysis of spring discharge variations and their influences for scheduling reduces of groundwater withdrawal is urgent for water resources utilization and environmental protection.

**SIMULATION MODELS FOR SPRING DISCHARGE VARIATIONS**

Two simulation models, artificial neural network model and time series model, are applied for spring discharge simulation. Based on annual mean discharge, precipitation, groundwater withdrawal and division water from Hongqi canal from 1971 to 2000, a BP network model is trained or calibrated and validated. Unstationary time series model can be used to simulate spring discharge variations because the spring discharge series is usually supposed to be an unstationary process with both natural and artificial influences. Compatibilities of these two models for spring discharge simulation and prediction are estimated on the basis of errors of simulated results and observed data from 1971 to 2000.

**BP neural network model**

The Back Propagation (BP) neural network is a multilayered, feedforward and by far the most extensively utilized due to its well-studied theory. There are one or more layers of neural cells between input layer and output layer. These cells have no direct communications to the external, but the change of their states can influence the relations between input and output. The BP neural network approximates the non-linear relationship between the input and the output by adjusting the weight values internally instead of giving the function expression explicitly. Further, the BP neural network can
be generalized for the input that is not included in the training patterns even in the noise-contained environment.

Figure 1. Variations of precipitation, groundwater withdrawal, Hongqi canal discharge, and spring discharge

Table 1. Amount of spring discharge, groundwater withdrawal and division water from Hongqi Canal in three periods

<table>
<thead>
<tr>
<th>Period</th>
<th>Discharge (m³/s)</th>
<th>Precipitation (mm)</th>
<th>Groundwater withdrawal (10⁴ m³/a)</th>
<th>Division water from Hongqi canal (10⁴ m³/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971-1976</td>
<td>8.06</td>
<td>720.35</td>
<td>1467.98</td>
<td>27850.0</td>
</tr>
<tr>
<td>1977-1989</td>
<td>5.3</td>
<td>560.49</td>
<td>2817.10</td>
<td>19607.7</td>
</tr>
<tr>
<td>1990-2000</td>
<td>4.48</td>
<td>559.55</td>
<td>6614.25</td>
<td>6000.0</td>
</tr>
</tbody>
</table>

The number of hidden layers and the number of neurons in one hidden layer are not straightforward. No rules are available to determine the number exactly. Flood [6] suggests that two hidden layers provide the greater flexibility necessary to model complex-shaped solution surfaces, and are thus recommended as a starting point when developing a layered feedforward network of sigmoidal neurons.

A dilemma exists when determining the number of hidden neurons [7]. One technique starts by training a relatively large network that is later reduced in size by removing the hidden neurons that do not significantly contribute to the solution. Another approach is the radial-Gaussian system developed by Flood, which adds hidden neurons
to the network, in a sequential manner, training each on the error left over from its predecessors. The number of hidden neurons required to achieve a given level of accuracy is thus determined automatically during training. Moreover, Hajela (1991) suggested the number of hidden-layer neurons in one hidden layer network to be between the average of the input and output layer neurons and the sum of these two-layer neurons.

The output of study problem is spring discharge and its influence factors are precipitation, groundwater mining and irrigation water from Hongqi canal division. Based on the trial-error method, network model is developed with analysissitus structure of three inputs, two layers and ten hidden-layer neurons.

All inputs $x_i$ are multiplied with a corresponding input weight $\omega_i$ and summed:

$$u_j = \sum \omega_i x_i - \theta_j$$

where $\theta_j$ is threshold. Normally, the threshold is named as bias, which is estimated as a trainable weight for an additional input signal attached to each node having a constant input value of $x = 1$. This sum is passed through a non-linearity, commonly a sigmoid function:

$$f(x) = \frac{1}{1+\exp(-x)}$$

The neuron output from the hidden layer is then calculated as:

$$y_j = f(u_j)$$

Usually all inputs and outputs to the network have been normalized to an internal representation between 0 and 1.

$$T = T_{\text{min}} + \frac{T_{\text{max}} - T_{\text{min}}}{X - X_{\text{min}}} (X_{\text{max}} - X_{\text{min}})$$

and the output is then rescaled to match the range of the training data.

$$X = X_{\text{min}} + \frac{X_{\text{max}} - X_{\text{min}}}{T_{\text{max}} - T_{\text{min}}} (T - T_{\text{min}})$$

where $X$ is observation data, $X_{\text{max}}$ and $X_{\text{min}}$ are maximum and minimum of the data series, respectively. $T$ is the normalized data, $T_{\text{max}}$ and $T_{\text{min}}$ are maximum and minimum of the normalized data series, respectively.

A major concern in the development of the neural network is to determine an appropriate set of weights. The actual output value is then compared to the desired output, and an error signal is computed for each output node. These error signals are transmitted backwards from the output layer to each node in the intermediate layer that contributes directly to the output. However, each unit in the intermediate layer receives only a portion of the total error signal, based roughly on the relative contribution the unit made to the original output. This process repeats, layer by layer, until each node in the network has received an error signal that describes its relative contribution to the total error. Based on the error signal received, connection weights are then updated by each unit to cause the network to converge toward a state that allows all the training patterns to be encoded.

The whole observation data series are divided into two sets: training data set from
1971 to 1992, and validation data set from 1993 to 2000. Model training results are shown in Fig 2. Simulated spring discharges $Q_s$ generally match the observed values $Q_o$ well. The average relative error of spring discharge is 6.5%, with maximum 20.8% and minimum 0.2%. Same input weights are used in validations of spring discharge from 1993 to 2000. Validation results are shown in Table 3. The average relative error of spring discharges is 6.5%, with maximum 17.8% and minimum 1.5%.

**Time series analysis model**

Most unstationary time series patterns can be described in terms of three basic components: trend $M_t$, periodic fluctuations $T_t$, and random $W_t$.

$$X_t = M_t + T_t + W_t$$ (6)

Trend of the series can be adequately approximated by a polynomial function.

$$M_t = c_1 + c_2 t + c_3 t^2 + c_4 t^3 + c_5 t^4$$ (7)

where $c_i (i=1,2,3,4,5)$ is coefficient to be determined by regression analysis method based on observation data (Table 2).

Periodic fluctuations $T_t$ can be approximated by following function.

$$T_t = \sum_{i=1}^{n} [a_i \cos(2\pi f_i t) + b_i \sin(2\pi f_i t)]$$ (8)

If there are $n$ data points in the series, then there will be $n/2-1$ cosine functions and $n/2-1$ sine functions. $a_i$ and $b_i$ are coefficient, $f_i$ is frequency of the periodic fluctuations.

After removing the trend and the periodic fluctuations from spring discharge series, remaining component can be regarded as stationary series and can be approximated by autoregressive process AP(2).

$$W_t = \phi_1 W(t-1) + \phi_2 W(t-2) + \epsilon$$ (9)

where $\epsilon_t$ is one dimensional white noise, $E \epsilon_t = 0$. $\phi_i (i=1,2)$ is autoregressive model parameters.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$a_i$</th>
<th>$b_i$</th>
<th>$f_i$</th>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$c_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.6768</td>
<td>-0.0488</td>
<td>0</td>
<td>0.98</td>
<td>8.392</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.1725</td>
<td>0.5071</td>
<td>0.32</td>
<td>-0.04</td>
<td>-0.3097</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.3341</td>
<td>-0.1409</td>
<td>0.15</td>
<td></td>
<td></td>
<td>-0.0016</td>
</tr>
<tr>
<td>4</td>
<td>-0.0054</td>
<td>-0.531</td>
<td>0.31</td>
<td></td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.06</td>
<td>-0.0409</td>
<td>0.01</td>
<td></td>
<td></td>
<td>-0.0002</td>
</tr>
<tr>
<td>6</td>
<td>0.1080</td>
<td>-0.0488</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.0089</td>
<td>-0.0768</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-0.1306</td>
<td>-0.2379</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.2757</td>
<td>-0.1048</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.0883</td>
<td>-0.613</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In this study, spring discharge series from 1971-1992 is used to determine model parameters by least squares estimation (Table 2). Simulated spring series \( Q_s \) is compared with the observed series \( Q_o \) as shown in Figure 2. Average relative error is 1.6 percent, with maximum 26.7 percent and minimum 0.5 percent. Another spring discharge set of eight years from 1993 to 2000 is used for validation. If same model parameters are used, simulation results in Table 3 indicate that errors between simulated and observed values become large. Average relative error is 17.3 percent; with the maximum 31.5 percent and the minimum 1.8 percent.

![Comparison between observed and simulated spring discharges](image)

**Figure 2.** Comparison between observed and simulated spring discharges

**Table 3.** Validation results of BP neural network and time series model

<table>
<thead>
<tr>
<th>Year</th>
<th>( Q_o ) (m³/s)</th>
<th>( Q_s ) (m³/s)</th>
<th>Absolute error (m³/s)</th>
<th>Relative error (%)</th>
<th>( Q_o ) (m³/s)</th>
<th>Absolute error (m³/s)</th>
<th>Relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>4.41</td>
<td>4.10</td>
<td>0.31</td>
<td>7.0</td>
<td>4.33</td>
<td>0.08</td>
<td>1.8</td>
</tr>
<tr>
<td>1994</td>
<td>4.99</td>
<td>4.10</td>
<td>0.89</td>
<td>17.8</td>
<td>3.75</td>
<td>1.24</td>
<td>24.8</td>
</tr>
<tr>
<td>1995</td>
<td>4.44</td>
<td>4.36</td>
<td>0.08</td>
<td>1.8</td>
<td>4.96</td>
<td>0.52</td>
<td>11.7</td>
</tr>
<tr>
<td>1996</td>
<td>5.28</td>
<td>4.42</td>
<td>0.86</td>
<td>16.3</td>
<td>4.11</td>
<td>1.17</td>
<td>22.2</td>
</tr>
<tr>
<td>1997</td>
<td>4.01</td>
<td>4.07</td>
<td>0.06</td>
<td>1.5</td>
<td>2.84</td>
<td>1.17</td>
<td>29.2</td>
</tr>
<tr>
<td>1998</td>
<td>4.71</td>
<td>4.62</td>
<td>0.09</td>
<td>1.9</td>
<td>4.26</td>
<td>0.45</td>
<td>9.6</td>
</tr>
<tr>
<td>1999</td>
<td>4.02</td>
<td>3.87</td>
<td>0.15</td>
<td>3.7</td>
<td>4.33</td>
<td>0.31</td>
<td>7.7</td>
</tr>
<tr>
<td>2000</td>
<td>4.1</td>
<td>4.17</td>
<td>0.07</td>
<td>1.7</td>
<td>2.81</td>
<td>1.29</td>
<td>31.5</td>
</tr>
<tr>
<td>Mean</td>
<td>4.5</td>
<td>4.22</td>
<td>0.28</td>
<td>6.5</td>
<td>3.93</td>
<td>0.57</td>
<td>17.3</td>
</tr>
</tbody>
</table>
CONCLUSIONS AND DISCUSSIONS

Comparison of the simulated and observed results indicates that two models can simulate spring discharge well in calibration periods, but errors between observed series and predicted series by the time series model are much larger than those by the neural network model in the validation period. Because the neural network model establishes relationship between spring discharge and its influences, spring discharge prediction is reliable if variations of its influences are known. The time series model is developed only based on characteristics of spring discharge series: its trend, periodic fluctuation and random. The model can be used to repeat historic series well. However, trend and periodic fluctuation related with human-regulated events may be changed in an inconsistent and unrepeatable way. Study results show that groundwater withdrawal has been reduced since protection measures have been taken from the middle of 90s. Therefore, decrease trend of spring discharge become less than that during 1971~1992. If the trend from 1971 to 1992 is still used for prediction, the simulated spring discharge in the middle of 90s must be much less that of observed value. Therefore, capabilities of model predictions rely on robust of model physical structures and representative of history data series.

REFERENCES