Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

ISSN: 0972-6268	Vol. 13	No. 2	pp. 413-416	2014

Original Research Paper

Time Series Analysis of Water Quality in Hanjiang River

Juan Zhang and Changjun Zhu
College of Urban Construction, Hebei University of Engineering, Handan, 056038, China

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com
Received: 19-5-2013
Accepted: 10-9-2013

Key Words:
Time series analysis
Exponential smoothing
Water quality
Seasonal decomposition

ABSTRACT

According to characteristics of time series in Hanjiang river, exponential smoothing and seasonal decomposition are adopted to predict the water quality in Hanjiang river. The accuracy of prediction will directly influence whether we can work out a reasonable plan and management measures. The paper introduces the principle of exponential smoothing method and seasonal decomposition, and the process is detailed using SPSS. The results indicate that the predictions are reasonable and reliable, and the methods are fitting for the time series analysis.

INTRODUCTION

Time series analysis is a kind of statistical method on dynamic data processing. The method is based on the random process theory and mathematical statistical methods to study the rule of statistics of random data sequence for solving practical problems. The methods include general statistical analysis (such as correlation analysis, spectrum analysis and so on), statistical modelling and inference, and on the time series analysis of optimal prediction, control and filtering content. Classic statistical analysis assumes that the data sequence is independent, and the time series analysis is focused on the study of data dependencies. The latter is actually on random time series analysis of discrete indicators, which is based on the observed time series analysis data to build mathematical model with parameters estimation theory and method according to the curve fitting. Hu Junhua et al. (2006) forecasted the hydrological level using SPSS. Fu Xinzhong et al. (2009) predicted medium and long-term runoff using ARIMA-ANN model. Ruzhong et al. (2002) predicted the river water quality based on grey dynamic model group. Xiaofang Rui & Linli Wang (2000) studied the flood prediction of flood routing method with forecast period. Guoru Huang & Xiaofang Rui (2003) predicted flood routing using radial basis function-neural network model. Xingming Zhu et al. (2005) predicted flood water level using artificial neural network model.

It is generally that the curve fitting and parameters estimation methods (such as nonlinear least squares) are adopted. Time series analysis can be applied in national economy macroscopical control, regional comprehensive development planning, business management, market forecasting, weather forecasting, hydrological forecasting, earthquake prediction, disaster prediction of crop diseases and insect pests, environment pollution control, ecological balance, astronomy and oceanography, etc.

Time series analysis is a group of sequence of numbers according to the time sequence. Time series analysis uses the sequence which is treated by mathematical statistics to predict the future. Time series analysis is one of the quantitative prediction methods. The basic principle can be expressed by two sides. One is the recognition of continuity, and the other is to consider randoms. Time series' prediction generally can be reflected three kinds of practical change rules: trend, periodic variation and stochastic change.

STUDY AREA

Hanjiang river is the longest tributary of the Yangtze river, which is 1577 kilometres long. Before 1959, the area of the Hanjiang river basin was 174300 square kilometres which is the largest basin among the Yangtze river basin; after 1959, the area was reduced to 159000 square kilometres, which has become the second tributary. Hanjiang river rises in Hanzhong city, Shanxi province, which has the best water quality in the middle region of China. Danjiang reservoir upstream the Hanjiang river is the water source of middle route of south-to-north water transfer project. Upstream is up the Danjiangkou, which is a narrow valley and 925km long. The middle reaches between Danjiangkou and Zhongxiang, which is a wide valley and sandy beach, 270kmm long; downstream is between Zhongxiang to Hankou, 382km long, flowing through the Jianghan plain, undulating gradually narrowing and flowing into Yangtze

river at Hankou Dragon King temple.

Hanjiang basin is involved in 78 counties, 20 cities of 6 provinces. Upper valley of basin is the boundaries of Qinling, Waifang mountain, funiu mountain and the Yellow river. Northeast is the boundary of Funiu mountain, Tongbai mountain and Huaihe river. Southwest to Daba mountain and Jing mountain and Jialing river, Zhanghe river, southeast is Jianghan plain, no obvious natural watershed boundaries.

Hanjiang river basin belongs to subtropical monsoon zone, the climate is mild and humid, annual precipitation is 873mm, water is plentiful but the distribution is uneven in a year. Runoff from May to October accounted for about 75% of the whole year, large inter-annual changes, which is the largest change in river among the tributaries of the Yangtze river.

EXPONENTIAL SMOOTHING ALGORITHM

According to the different smooth times, exponential smoothing method is divided into single exponential smoothing method, two exponential smoothing method and three exponential smoothing method. But their basic ideas are as follows: Predictive value was previously observed values weighted. As the different data give different right, new data are given the larger weight, the older to the smaller weight.

Single exponential smoothing method: Assuming time series $y_1, y_2, \cdots\cdots, y_t$, the single exponential smooting formula is $S_t^{(1)} = \alpha y_t + (1-\alpha) S_{t-1}^{(1)}$ where the single exponential smoothing values in t[th] cycle, α is weighted coefficient, $0 < \alpha < 1$.

In order to understand the essence of the exponential smoothing, the above formulas were launched, the formula can be got as follows.

$$S_t^{(1)} = \alpha \sum_{j=0}^{t-1} (1-\alpha)^j y_{t-j} + (1-\alpha)^t S_0^{(1)}$$

$0 < \alpha < 1$, when $t \to \infty, (1-\alpha)^t \to 0$, So

$$S_t^{(1)} = \alpha \sum_{j=0}^{t-1} (1-\alpha)^j y_{t-j}$$

This shows that the $S_t^{(1)}$ is weighted average of time sequence $y_1, y_2, \cdots\cdots, y_t$, weighted coefficient is $\alpha, \alpha(1-\alpha), \alpha(1-\alpha)^2, \cdots$, which is based on a geometric attenuation:

$$\alpha \sum_{j=0}^{t-1} (1-\alpha)^j = 1$$

Because the weighted coefficient comply with exponential rule, and also has a smoothing function, it is called exponential smoothing.

To predict the value with smoothing predicting is single exponential smoothing method. The prediction model can be as follows:

$$\hat{y}_t = S_t^{(1)} = \alpha y_t + (1-\alpha) \hat{y}_t$$

An exponential smoothing value in t period is as the predictive value in t+1 period.

Two exponential smoothing method: When time series has no apparent tendency changes, single exponential smoothing in t period can directly predict the values in t+1 period. But when the time sequence occurs in a linear trend, obvious lag error is existed when single exponential smoothing method is used to predict. Therefore, it needs to be revised. Correction method is to make two exponential smoothing based on the single exponential smoothing. Using hysteresis deviation rule to find the development direction and trend, then the linear trend prediction model is built. It is known as the two exponential smoothing method.

Assuming single exponential smoothing, the computing formula of two exponential smoothing is as follows:

$$S_t^{(2)} = \alpha S_t^{(1)} + (1-\alpha) S_{t-1}^{(2)}$$

If time sequence begins with a linear trend from a period, it has line trend changes in the next period, the linear model as follows can predict:

$$\hat{y}_{t+T} = a_t + b_t T \quad T = 1, 2, \cdots$$

Where t is current period; T is the period number from current period to predict period; $\hat{y}_{t+T}$ is the forecast value in t+ T period, a_t is intercept; b_t is slop.

$$a_t = 2S_t^{(1)} - S_t^{(2)} \qquad b_t = \frac{\alpha}{1-\alpha}(S_t^{(1)} - S_t^{(2)})$$

Seasonal decomposition model: Changes of time sequence are affected by long-term trend, seasonal variation and irregular changes.

Long-term trend factors: Long-term trend factors (T) reflect the development direction in a long time; it can be going up or going down or a smooth trend in a long time.

Seasonal variation factor (S): Seasonal variation factor (S) is a fixed length and amplitude fluctuation affected by seasonal changes.

Periodic fluctuant element: Periodic fluctuant element, also known as cyclic variation factor, is affected by various factors which influence the formation of the upper and lower volatile fluctuations.

Irregular changes factors: The irregular change is also

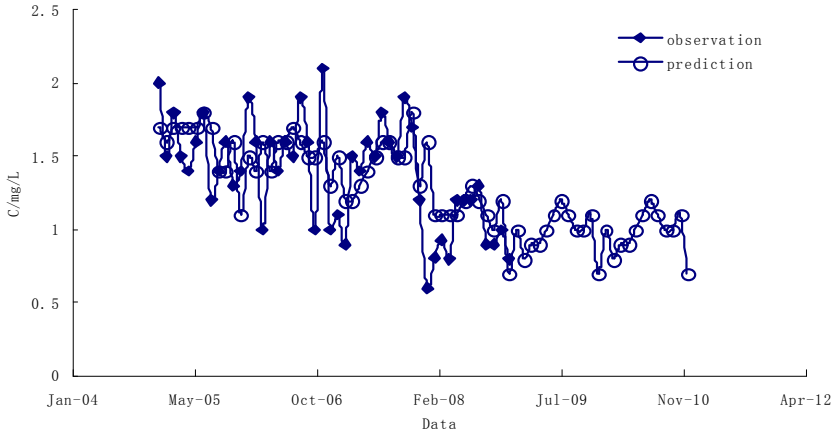

Fig. 1: Comparison of observation and prediction of dissolved oxygen.

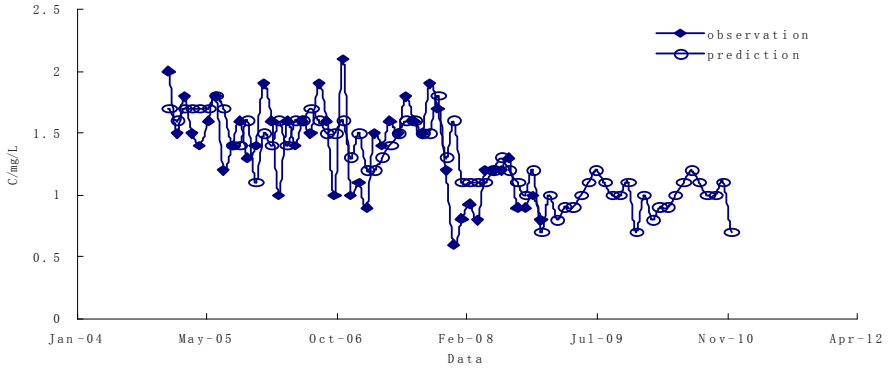

Fig. 2: Comparison of observation and prediction of potassium permanganate index.

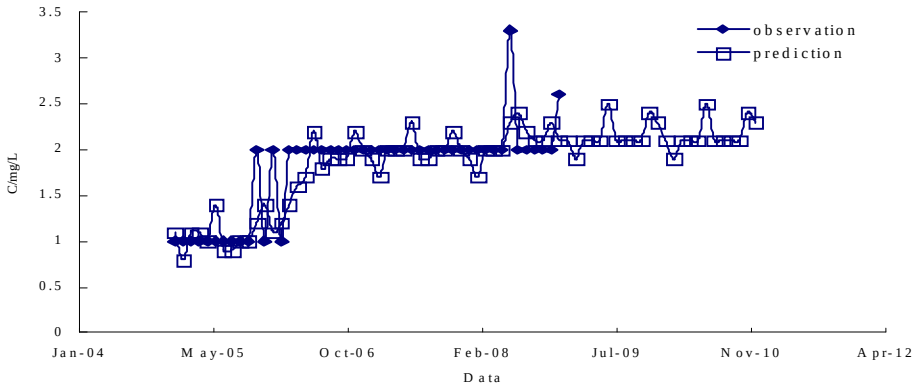

Fig. 3: Comparison of observation and prediction of BOD_5.

known as random variable. It is affected by various causal factors affecting the irregular fluctuant.

Time sequence of *y* can be expressed as a function of the above four factors.

$$y_t = f(T_t, S_t, C_t, I_t)$$

There are many methods for time series decomposition, the commonly employed methods are additive model and multiplicative model.

Additive model: $y_t = T_t + S_t + C_t + I_t$

Multiplicative model: $y_t = T_t \times S_t \times C_t \times I_t$

The fundamental steps of multiplicative decomposition model are as follows:

1. Using moving average method to eliminate the long-term trend and periodicity to get series TC, then seasonal index S can be got by monthly average method.

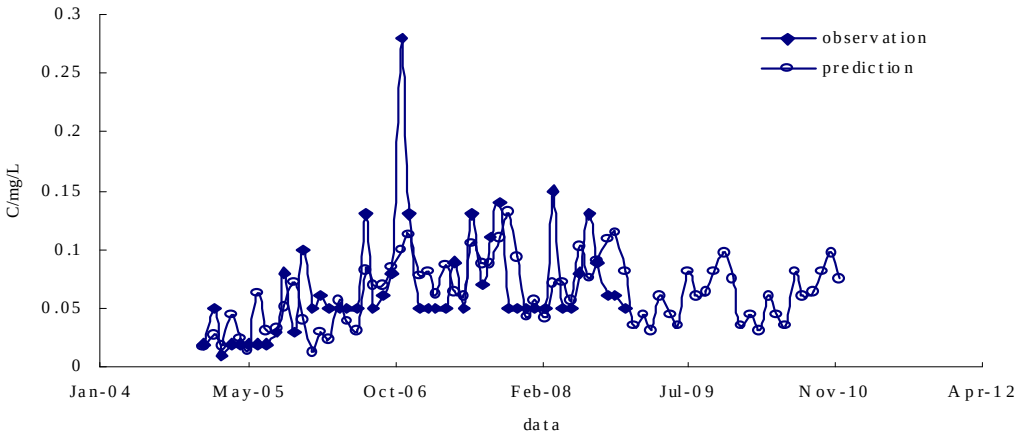

Fig. 4: Comparison of observation and prediction of NH_4-NH_3.

2. Scatter diagram, choose the suitable curve to get the long-term trends T.
3. Calculation of cycle factor C. The periodic fluctuant element C can be got by the TC of sequence divided by T.
4. T, S, C decomposed from series, the remaining is the irregular change, i.e.

$$I = \frac{Y}{TSC}$$

Analysis of every section in Hanjiang basin: The process of data treatment is as follows:

Import the data into SPSS.

Data pretreatment: Click the button "data'→defined date", the time of data can be defined, the initial value is 1997-1, then click"OK" button, the data variables can be created.

Creating data series: Click exchange creating time series, selecting function, developing function parameters, the choice to be processed, after definition, click "OK" then the data variables are generated. Prediction using exponential smoothing

Seen from the Figs. 1 to 4, the simulated results can meet the demands of the hydrology prediction.

CONCLUSION

According to the analysis, it is found that the exponential smoothing model is used to analyse the runoff in Hanjiang river. The results indicate that the method can be regarded as a kind of simple and accurate analysis method.

1. ARIMA model is a kind of simple and accurate analysis method, which can be applied to predict the runoff.
2. The results evaluated by ARIMA can provide references for environment departments.

ACKNOWLEDGMENTS

This study was supported by Natural Science Foundation of Hebei Province (E2012402013), Open Foundation of State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering (2011491511), the Program for Handan Science and Technology Research and Development (1123109066-4), the National Basic Research Program of China (2010CB951101), the National Natural Science Foundation of China (Grant No. 40830639, 50879016, 50979022 and 50679018), the Program for Changjiang Scholars and Innovative Research Team in University (IRT0717) and the Special Fund of State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering (1069-50985512, 1069-50986312).

REFERENCES

Hu Junhua, Zhang Qingbo and Bao Leiming 2006. Application of SPSS statistical software in hydrological forecasting. Northwest Hydropower, 13(2): 3-4.

Fu Xinzhong, Feng Lihua and Chen Wenchen 2009. Application of ARIMA-ANN model in the prediction of medium and long-term runoff. Journal of Water Resources & Water Engineering, 20(5).

Li, Ruzhong, Wang, Jiaquan and Qian, Jiazhong 2002. Prediction of river water quality based on grey dynamic model group. Bulletin of Soil and Water Conservation. 22(4): 10-12.

Xiaofang Rui and Linli Wang 2000. A study of flood routing method with forecast period. Advances in Water Science, 11(3): 291-295.

Guoru Huang, Xiaofang Rui, 2003. Radial basis function-neural network model for channel flood routing, Journal of Hohai University (Natural Sciences), 31(6): 621-625.

Xingming Zhu, Changna Lu, Ruyun Wang and Jingyi Bai 2005. Artificial neural network model for flood water level forecasting. Journal of Hydraulic Engineering, 36(7): 806-811.