

Time Series Analysis of Phosphorus Concentration in Chaohe River in Beijing

Weiwei Zhang, Hong Li

Institute of System Comprehensive Development, Beijing Academy of Agriculture and Forest Science, Beijing, 100097, China

Abstract—Taking the phosphorus monitoring data of two sites MS1 and MS2 in the period 1986-2003 in Chaohe River lying Beijing Miyun reservoir stream watershed as a case, time domain analysis method, Fourier and Wavelet analysis were adopted to explore and compare the periodic patterns and temporal pattern characteristics of the two sites. The results showed that the time domain analysis method was not adopted due to there were no serial correlation and heteroscedasticity in each site series. The periodic patterns of two sites were discovered using Fourier analysis. The site MS1 had two periodic patterns of two years and four years while the site MS2 had a period of two and a half years. The temporal pattern characteristics at different scales were obtained through wavelet analysis, which were at moderate and small scales for the site MS1, while at small scale for the site MS2. The Fourier and wavelet analysis method can both be used in the study of surface water quality temporal change pattern, the first is a coarse method and the latter is a more detailed method for analyzing surface water quality temporal pattern characteristics.

Keywords—time domain analysis; fourier; wavelet analysis; phosphorus; water quality; chaohe river

I. INTRODUCTION

In the context of rapid industrialisation and urbanization, discharge of the effluents from industrial, municipal and agricultural activities into water environment has brought surface water pollution problems. In agricultural watersheds, many nutrients such as nitrogen and phosphorus are usually from agricultural practices, and transport and transformation of pollutants are influenced by local hydrological processes. Temporal and spatial evolution of these pollutants can help develop and illustrate transport characteristics in relation to various events, such as agricultural activities and hydrological processes.

Time series analysis in relation to temporal and spatial pattern of water quality includes multivariate statistical techniques, time domain analysis and frequency domain analysis. Multivariate statistical techniques, taking a hypothesis that the data is independent in time series, ignore the temporal and spatial relation [1-3]. Through describing the autocorrelation structure in time series, time domain analysis can explore the dynamic evolution characteristics of time series data. Since easily used, this method was commonly applied to the analysis of water quality time series [4-6]. Frequency domain analysis includes Fourier analysis and wavelet analysis. Fourier analysis, that

transforms time series values from temporal domain to the frequency domain, partitions the total variance into different temporal scales defined by frequencies (i.e., high frequencies=small scale, and low frequencies=large scale), and thus helps identify the dominant temporal scales. Wavelet analysis partitions the variation of time series signals into two domains, time location and scale (frequency or location) with both stationary and non-stationary datasets. Thus, we can zoom in or out detailed variations occurring at a specific temporal scale and time location. Wavelet analysis is a common method for signal and image processing, recently has been successfully applied in some hydrological, climate and oceanographic time series studies [7-9]. However, there is little study yet using frequency domain analysis in water quality signals. How to combine time domain and frequency domain analysis tools to analyze temporal pattern of water quality signals over different temporal scales and further find various processes affecting water quality temporal evolution at different scales, is very important for ensuring watershed water quality security.

II. MATERIALS AND METHODS

A. Watershed description and water quality monitoring

The study watershed, called Chaohe River watershed, is located in Miyun reservoir upriver in the northwest of Beijing, China. There are two water quality monitoring site (MS1 and MS2) in the study area. Elevation within the study area ranges from 5 to 2250 above mean sea level. Mean annual precipitation in the area is about 566 mm, about 60% of which is in July and August. Annual average runoff is about $1.8 \times 10^9 \text{ m}^3$ but has decreased to $1.3 \times 10^9 \text{ m}^3$ at the end of the last century due to climate and land use/cover changes. Rainfall and runoff provide a driving for watershed pollution. The land uses within the study area are agriculture, forest, and rural-residential. Many studies have showed that agricultural non-point source pollution was the major influence factor of Miyun reservoir watershed water quality.

The phosphorus monitoring data of two monitoring sites for the years of 1986-2003 in study area were obtained from Beijing Water Authority. Water samples were collected at May and October from sampling sites during the period 1980-2003, and analyzed by environmental quality standard elementary item analysis method for surface water. The average of May and October monitoring data annually was the annual phosphorus concentration.

B. Time domain analysis method

Through making time series models, time domain analysis can obtain dynamic change of time series signal. Thus, this method is simple and easily used. However, if there are no significant serial autocorrelation and heteroscedasticity in time series, it is unable to make time series models. Therefore, preliminary tests including serial autocorrelation and heteroscedasticity should be performed to determine whether using time series models. The sample autocorrelation function (ACF) was used to examine the serial autocorrelation qualitatively. The above-mentioned qualitative checks for serial autocorrelation are quantified using the Ljung-Box-Pierce Q-test. Also formal hypothesis test, such as Engle's test was used for the presence of heteroscedasticity (i.e. ARCH effects). All the operations were implemented in the software of Matlab.

C. Fourier analysis

Fourier analysis, which breaks down a signal into constituent sinusoids of different frequencies, is as a mathematical technique for transforming our view of the signal from time-based to frequency-based. The fast Fourier transform (FFT) is used to transform phosphorus concentration time series into a frequency-domain signal [10]. The power spectral density, a measurement of the energy at various frequencies, is calculated using the complex conjugate, and a plot of power versus frequency is a periodogram. The scale of frequency is somewhat inconvenient, so power versus period is plotted for convenience (where period=1/frequency).

D. Wavelet analysis

Wavelet analysis was developed from the Fourier transforms. It can investigate detailed temporal patterns from both frequency and time domains through adjusting time and frequency signals. The mother wavelet function can be considered as a series of adjustable moving windows to detect local temporal patterns. For each chosen scale and time location, Fourier transforms are conducted in wavelet analysis. Some detailed introduction and comparison of wavelet analysis were discussed by Torrence [11]. The continuous wavelet transform (CWT) is defined as the sum over all time of the signal multiplied by scaled, shifted versions of the wavelet function ψ :

$$C(a,b) = \frac{1}{\sqrt{a}} \int f(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (1)$$

Where C is the wavelet coefficient, which are a function of scale and position, a and b are scale and position, respectively, f(t) is the time series, and ψ^* is the complex conjugate of wavelet function ψ . The Morlet wavelet function was chosen for wavelet transform in this study.

$$\psi(t) = e^{-i\omega_0 t} e^{(-t^2/2)} \quad (2)$$

Where ω_0 is non-dimensional frequency ($\omega_0=6$), t is time.

The Morlet wavelet is a complex wavelet, thus allowing us to detect both location-dependent amplitude and phase

for different frequencies exhibited in time series. Also, the Morlet wavelet provides a good balance between time and frequency localization [12]. Therefore, the Morlet wavelet has been used for hydrological signals, such as rainfall and runoff analysis. These studies showed that Morlet wavelet can describe well the shape of hydrological signals, and is appropriate for hydrological signals analysis. Because the Morlet wavelet is a complex wavelet, the wavelet transform $C(a,b)$ is also complex. The transform can then be divided into the real part and the imaginary part, or amplitude, $|C(a,b)|$, and phase, $\tan^{-1}\{\text{Im}[C(a,b)]/\text{Re}[C(a,b)]\}$. Consequently, one can also defined the wavelet spectral power at different scale (a) and position (b).

$$P_w(a,b) = |C(a,b)|^2 \quad (3)$$

III. RESULTS

A. Time domain analysis

Fig.1 displayed that the sample autocorrelation (ACF) of the phosphorus concentration time series after extracting the sample mean and the upper and lower standard deviation confidence bounds (at the 95% confidence level) for MS1 and MS2. From these figures for the ACF, there is no significant serial correlation at the corresponding element of lags. Moreover, there is periodic change from the ACF.

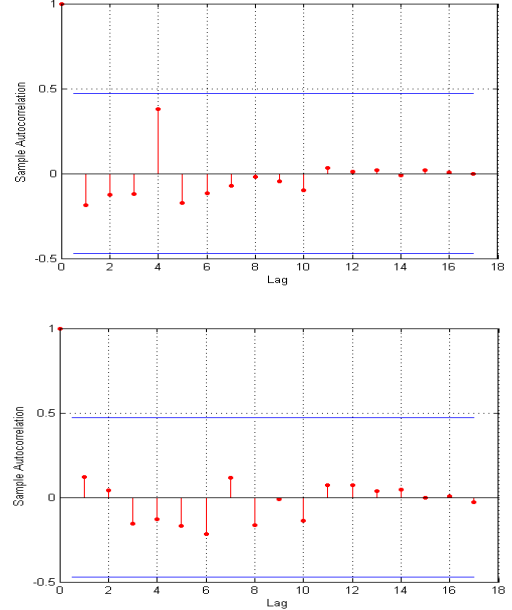


Figure 1. The ACF of TP time series at the two sites MS1 (left) and MS2 (right) (confidence interval 95%)

Quantifying the correlation using the function in Matlab returns a Boolean output. $H = 0$ implies that no significant correlation exists, while $H = 1$ means that significant correlation exists. Table 1 showed that there is no significant serial correlation when tested from 0 up to 17 lags of the ACF at the 0.05 level of significance.

Perform Engle's ARCH test using the function in Matlab also returns a Boolean output. $H = 0$ indicates acceptance of the null hypothesis that no ARCH effects exist; that is, there is homoscedasticity at the corresponding element of lags. H

$= 1$ indicates rejection of the null hypothesis. Thus, Table 2 showed that there is no heteroscedasticity when tested from 0 up to 16 lags at the 0.05 level of significance.

TABLE I. SERIAL AUTOCORRELATION TEST OF TP TIME SERIES

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
MS1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MS2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

TABLE II. HETEROSCEDASTICITY TEST OF TP TIME SERIES

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
MS1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MS2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

B. Fourier analysis

There were two peak values with a length of about 2 and 4 years in the plot of power versus period at MS1 (Fig.2 left). And the two peak values were significant at 90% confidence level, which indicated that the phosphorus of MS1 had two annual temporal patterns with 2 and 4 years. Three peak values were found in the plot of power versus period at MS2 (Fig.2 right). Only the peak values 2.5 were significant at 90% confidence level, which indicated that the phosphorus of MS2 had an annual temporal pattern with 2.5 years.

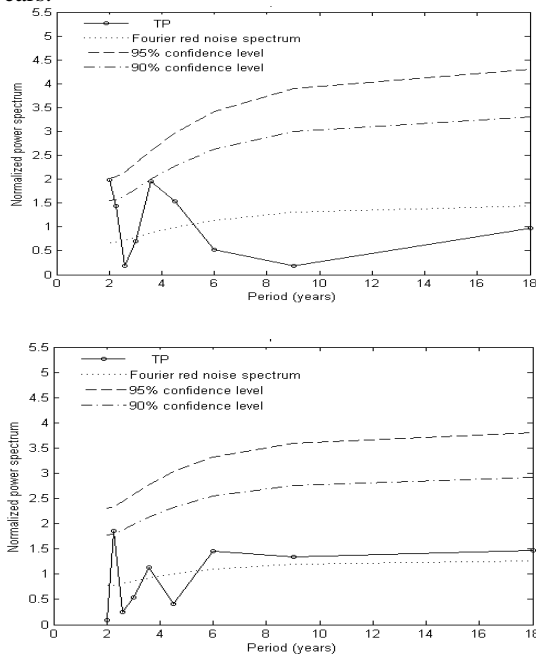


Figure 2. Periodograms of TP for the two sites MS1 (left) and MS2 (right)

The 2 and 4 years periodic patterns of phosphorus concentration were detected for MS1 yearly data collected from 1986 to 2003. This period may correspond to two dominant courses affecting temporal pattern of phosphorus pollutant in surface water quality. Under the dominant courses, phosphorus had two periodic patterns of 2 and 4 years at MS1.

So although the period of phosphorus concentration was detected using Fourier analysis, the Fourier analysis has an obvious shortcoming. In transforming phosphorus concentration series into the frequency domain, time information is lost. Through Fourier transform of phosphorus concentration series, it is impossible to find the change time of phosphorus. Classical Fourier spectral analysis has been considered as a useful method to detect stable periodic characteristics. However, it is a coarse method to research detail periodic pattern together with time information in water quality signals.

C. Wavelet analysis

Fig.3 and 4 showed the local wavelet power spectrum and wavelet coefficient of phosphorus for monitoring sites MS1 and MS2, respectively. In the figures, the thick black contour designates the 5% significance level against red noise and the thin solid lines like 'U' shape indicate the Cone of Influence.

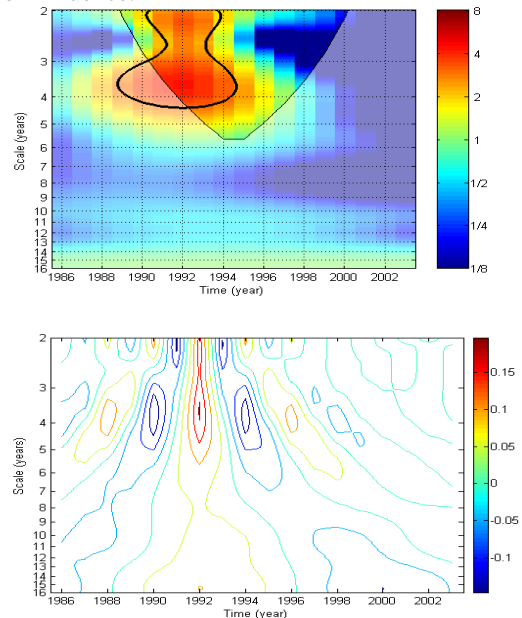


Figure 3. The local wavelet power spectrum (left), wavelet coefficient (right) of TP for MS1

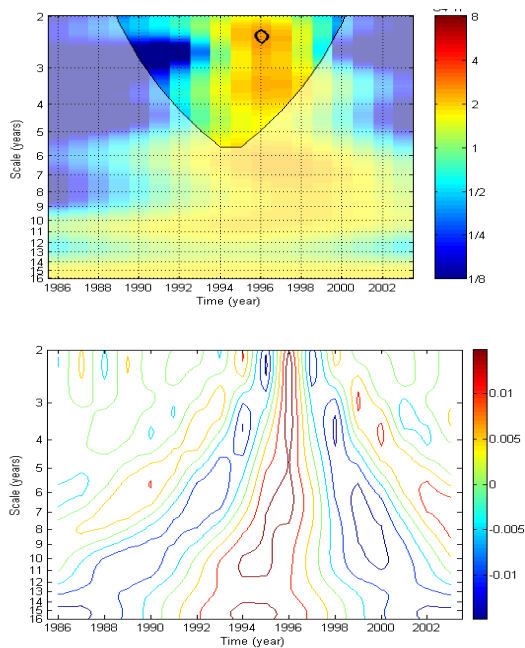


Figure 4. The local wavelet power spectrum (left), wavelet coefficient (right) of TP for MS2

The phosphorus series at site MS1 has high power at the scale of 2-4 years in the period from 1986-2003, which correspond to the small-scale random variations. Some of the variations are statistically significantly different from that of red noise. Therein, there are high variances at the scale of 3-4 years in the periods 1988-1996, while there are high variances at the scale of 2-3 years in the periods 1990-1994.

For the phosphorus at site MS2, there are high variations at small scales around 2 to 4 years in the period from 1994-1998. There are moderate variations at moderate scales after 4 years in the all monitoring period. However, some of these variations are not statistically significant as indicated without thick black contour lines.

The temporal pattern of phosphorus concentrations at different scales were detected through continuous wavelet transform, which might correspond to different dominant courses affecting temporal pattern of phosphorus pollutant in surface water quality. When the scales have changed, the dominant courses affecting surface water quality would also change.

IV. CONCLUSION

In this paper, time domain analysis, Fourier and Wavelet analysis were adopted to explore and compare the periodic patterns and temporal pattern characteristics of phosphorus monitoring data with the period 1986-2003 in surface water, taking two sites MS1 and MS2 in Chaohe River lying Miyun reservoir stream watershed as a case. The time domain analysis method was not adopted due to there were no serial correlation and heteroscedasticity in each site series. The periodic patterns of two sites were

discovered using Fourier analysis. The site MS1 had two periodic patterns of two years and four years while the site MS2 had a period of two and a half years. The temporal pattern characteristics at different scales were obtained through wavelet analysis, which were at moderate and small scales for the site MS1, while at small scale for the site MS2.

The wavelet analysis of water quality signals in this study showed advantages of detecting detailed temporal pattern compared to the classic Fourier analysis. In the future study, wavelet analysis can be used in the study of surface water quality temporal change pattern, which benefit temporal transport processes analysis of water quality pollutants and water quality control.

V. ACKNOWLEDGEMENTS

The authors are grateful to the Beijing Municipal Bureau of Water Management for their water sampling and analysis, and we would like to thank all reviewers and editors for their valuable comments and suggestions during the review process.

REFERENCES

- [1] Shrestha S. & Kazama F., Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environmental Modelling and Software*, 22, pp. 464-475, 2007.
- [2] Zhou F., Liu Y. & Guo H. C., Application of multivariate statistical methods to the water quality assessment of the watercourses in the northwestern New Territories, Hong Kong. *Environmental Monitoring and Assessment*, 132, pp. 1-13, 2007.
- [3] Singh K. P., Malik A., Mohan D. & Sinha S., Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India): A case study. *Water Research*, 38, pp. 3980-3992, 2004.
- [4] Worrall F. & Burt T. P., Time series analysis of long-term river dissolved organic carbon records. *Hydrological Processes*, 18, pp. 893-911, 2004.
- [5] Mohammadi K., Eslami H. R. & Kahawita R., Parameter estimation of an ARMA model for river flow forecasting using goal programming. *Journal of Hydrology*, 331, pp. 293-299, 2006.
- [6] Koutroumanidis T., Sylaios G., Zafeiriou E. & Tsihrintzis V. A., Genetic modeling for the optimal forecasting of hydrologic time-series: Application in Nestos River. *Journal of Hydrology*, 368, pp. 156-164, 2009.
- [7] Kang S. J. & Lin H., Wavelet analysis of hydrological and water quality signals in an agricultural watershed. *Journal of Hydrology*, 338, pp. 1-14, 2007.
- [8] Anctil F. & Coulibaly P., Wavelet analysis of the interannual variability in southern Quebec streamflow. *Journal of Climate*, 17, pp. 163-173, 2004.
- [9] Zhang Q., Chen G. Y., Xu C. Y., Xu Y. P., Liu C. L. & Jiang T., Periodic characters of sediment load and runoff changes in the Yangtze River basin in the past 40 years, China. *Advances in Water Science*, 20, pp. 80-85, 2009.
- [10] Si B. C., Spatial scaling analyses of soil physical properties: A review of spectral and wavelet methods. *Vadose Zone Journal*, 7, pp. 547-562, 2008.
- [11] Torrence C. & Compo G. P., A practical guide to wavelet analysis. *Bulletin of American Meteorological Society*, 79, pp. 61-78, 1998.
- [12] Labat D., Ronchail J. & Guyot J. L., Recent advances in wavelet analysis: Part 2. Amazon, Parana, Orinoco and Congo discharges time scale variability. *Journal of Hydrology*, 314, pp. 289-311, 2005