Research on the analysis of Yellow River water and sand monitoring data based on ARIMA model

Yan Liu, Xue Bai*, Qian Shi, Sijia Liu

Gansu Finance and Trade Vocational College, Gansu, 730300, China

*Corresponding author: bxue94668@outlook.com

Abstract. The article uses hydrological data from a hydropothe paperr station in Xiaolangdi Reservoir spanning six years. the paper use mutability analysis, seasonal decomposition method, and ARIMA model to reveal hydrological change trends and cyclic characteristics. the paper also use the CUSUM algorithm to identify the water-sand flux mutation points. The study's findings demonstrate that ARIMA(1,2,1) can be used to predict trends in total water flux and total sand content. It also reveals that there is significant seasonality and periodicity in the water-sand fluxes. These results offer a solid scientific foundation for decision-making and have substantial implications for the management and control of the region's water resources.

Keywords: Yellow River sand, ARIMA model, CUSUM algorithm.

1. Introduction

The Yellow River is one of the important rivers in China, and the change of its water and sand flux is affected by many factors, such as population increase, economic development and climate change. The study of the changing law of water and sand flux in the Yellow River can help to understand the hydrological characteristics of the Yellow River Basin, the current status and trend of water resources utilization, and provide a scientific basis for water resources allocation. In addition, the study can help predict future water supply and demand, provide reference for economic and social development planning and decision-making, and understand the impact of climate change on hydrological processes, providing scientific support for climate adaptation and ecological protection.

Recent studies have shown that changes in water and sand fluxes in the Yellow River Basin are influenced by a variety of factors, including climate change, human activities and watershed management. The research of Guo the papern[1] and other scholars found that the gradual decrease of water-sand flux in Jialing River was affected by the combination of human activities, climate change and topography.Liu Xin [2] and other scholars studied the influence of channel morphology changes on water-sand processes in the lower reaches of the Yellow River, and found that there are obvious trends in channel morphology parameters on different time scales. Cui Anhua [3] and other scholars explored the index problem of water and sand regulation in the Ningmeng section of the Yellow River, and came up with a set of water and sand regulation indexes applicable to this section of the river. Liu Xiaoyan [4] mainly thought about the water-sand situation and countermeasures of the Yellow River, and believed that measures such as strengthening river management, improving water resources utilization efficiency, and strengthening ecological protection are the key to solving the water-sand problem of the Yellow River. The study of Wang Junjie[5] and other scholars shothe paperd that the change of water-sand flux in the Yellow River has seasonal and interannual variations, and is also affected by human activities, climate change and topography of the basin. Cheng Yifei[6] et al. effectively modeled the process of flood evolution, which can provide an important basis for flood management of the Yellow River. Feng Shaojie[7] et al. emphasized that it is crucial to understand and quantify the impacts of various natural and anthropogenic factors on water systems in order to make more effective water management decisions. Chen Peng[8] et al. investigated and analyzed the characteristics of water and sand changes in the Yangtze River source area over the past decade. It helps to understand the effects of soil erosion and flood management strategies in the Yangtze River source area. Zhao Yang[9] et al. shothe paperd that soil and water conservation measures can significantly improve water and sand conditions in the watershed.

Hothe paperver, what is the pattern of change in the time series of total water flux and total sand discharge? As the paperll as the use of CUSUM algorithm to analyze the sudden change of water and sand fluxes related research is less. In this paper, through the water level, water flow and monitoring data of a hydropothe paperr station on the Yellow River downstream of Xiaolangdi Reservoir in the past 6 years, the paper study the change rule of total water flow and total sand discharge(The data are derived from http://www.mcm.edu.cn/). It also analyzes the seasonality, periodicity and mutability of the water and sand fluxes to analyze the hydrological trend more comprehensively.

2. Time series analysis and forecasting methods

In order to investigate the time series characteristics of total water flow and total sand discharge, this paper adopts ARIMA model to predict the changes of total water flow and sand discharge. The modeling algorithm is as follows:

ARIMA (p, d, q) model[10] : ARIMA model is a commonly used method for forecasting and modeling based on time series data.ARIMA full name is Autoregressive Integrated Moving Average, which can be regarded as a combination of Autoregressive (AR) and Moving Average (MA) models with the introduction of differencing (I), which is used to deal with the non-stationarity of time series data. Its specific process is:

Define a specific time series to abstract the timeline:

$$DT_t = \{D_t, D_{t-1}, \cdots, D_{,t-k}\},\ k > 0 \tag{1}$$

whereD_t is the time sample at moment t

Differential smooth processing, ARIMA model requires that the sample series under study is a smooth series, i.e., it is from the trend graph of the sample did not find a clear trend before the study, that is, the mean and variance of the sample under study do not change significantly, this paper first use d = 1 for the difference until the series under study is a smooth series:

$$\nabla DT_t^d = D_t^d - D_{t-1}^d \tag{2}$$

The final smooth difference sequence is obtained as:

$$\nabla DT_t^d = \left\{\nabla D_t^d, \nabla D_{t-1}^d, \cdots, \nabla D_{t-k}^d\right\} \tag{3}$$

where d is the number of differences

The p, q values the paperre determined by calculating the autocorrelation function as the paperll as the partial autocorrelation function, and the most suitable ARIMA (p, d, q) model was accurately determined according to BIC (Bayesian)

$$BIC = -2\ln(L) + k\ln \tag{4}$$

where L is the likelihood function and k is the number of parameters. The final) ARIMA (p, d, q) model is derived by LB (Liung-Box) test.

The model predictions, which the paperre∇DT_t^d brought into the ARIMA (p, d, q) model and the inverse variance score was performed to obtain the predictions:

$$T_i = \{T_1, T_2, \cdots T_i, \cdots, T_m\}(1 \le i \le m) \tag{5}$$

whereT_i is the predicted time series, and$m$ is the number of time periods predicted, in this study for the prediction of the total water flow and total sand discharge in the last 6 years, so the number of time periods$m = 6$.

3. Research on influencing factors and trend prediction of water and sand changes in the Yellow River

3.1. Time scale analysis of total water flow and total sand discharge

In order to further analyze the change rule of the time series of the total water flow and the total sand discharge of this hydropothe paperr station, as well as to identify the appropriate ARIMA (p, d, q), the hydrological changes of the chosen hydropothe paperr station are studied in this paper using the ARIMA model. The ADF test is conducted in this work for both the total water flow and the total sand discharge, respectively.

Table 1. Total Water Flow ADF Inspection Table

ADF Inspection Form							
variant	difference in order	t	P	AIC	threshold value		
					1%	5%	10%
Annual flood (m^3)	0	-5.398	0.000***	185.768	-7.355	-4.474	-3.127
	1	-2.721	0.070*	197.392	-7.355	-4.474	-3.127
	2	-5.367	0.000***	147.194	-10.417	-5.778	-3.392

Table 1 showed the ADF test table for total water flow. Based on the results of Table 1, the analysis was performed for the total water flow variable with difference orders of 0, 1 and 2. In the case of difference order of 0, the original hypothesis assumes that the variable is non-stationary, and the p-value is less than 0.01 at an ADF statistic of -5.398, so the original hypothesis can be rejected. In terms of the AIC value obtained, moreover, it is 185.768, which is lower than the critical value, indicating that the model is well fitted. Yet, at a difference order of 1, the ADF statistic of -2.721, with a p-value of about 0.07, fails to reject the original hypothesis, and although the first-order differenced series may be smooth, the corresponding AIC value of 197.392 is higher than the critical value, suggesting a relatively poor model fit. Finally, with a difference order of 2, the ADF statistic is -5.367, whose p-value is less than 0.01, which rejects the original hypothesis and indicates that the series after second-order differencing is smooth, while the corresponding AIC value is 147.194, which is lower than the critical value, suggesting that the model is well-fitted. In conclusion, according to the statistical indicators and information criterion, for the selection of d in the ARIMA (p, d, q) model of total water flow, this paper chooses the model with a difference order of 2 to obtain a smooth series.

Table 2. Total sediment discharge ADF inspection table

ADF Inspection Form							
variant	difference in order	t	P	AIC	threshold value		
					1%	5%	10%
Annual total sediment discharge (m^3)	0	-1.484	0.541	211.474	-6.045	-3.929	-2.987
	1	-1.603	0.482	220.5	-7.355	-4.474	-3.127
	2	-17.915	0.000***	154.589	-10.417	-5.778	-3.392

Table 2 showed the ADF test table for total sand discharge . Based on the results of Table 2, the following results were obtained after analyzing the ADF test for the total sand discharge variable with difference orders of 0, 1 and 2. In the case of difference order of 0, the original hypothesis cannot be rejected, which indicates that the variable may be smooth in its original state and the model is well fitted. Whereas, in the case of a difference order of 1, the series after first order differencing may be smooth, but the model fit is relatively poor. Finally, in the case of a difference order of 2, the series after second-order differencing is recognized as smooth and the model is well fitted. The combination of statistical indicators and information criterion, for the selection of d in the ARIMA (p, d, q) model

for total sand discharge, it is recommended to choose a model with a difference order of 2 to obtain a smooth series.

To further analyze the total water flow and total annual sand discharge their respective ARIMA (p, d, q) model values of q, p. In this paper, the ACF and PACF plots of total water flow and total sand discharge are plotted. From Fig. 1 and Fig. 2, the paper can get the time series of total water flow and total sand discharge, whose p, q values are 1. Combined with the previous judgment of its differential order, finally this paper determines that the optimal ARIMA model of total water flow and total sand discharge is the ARIMA(1, 2, 1) model.

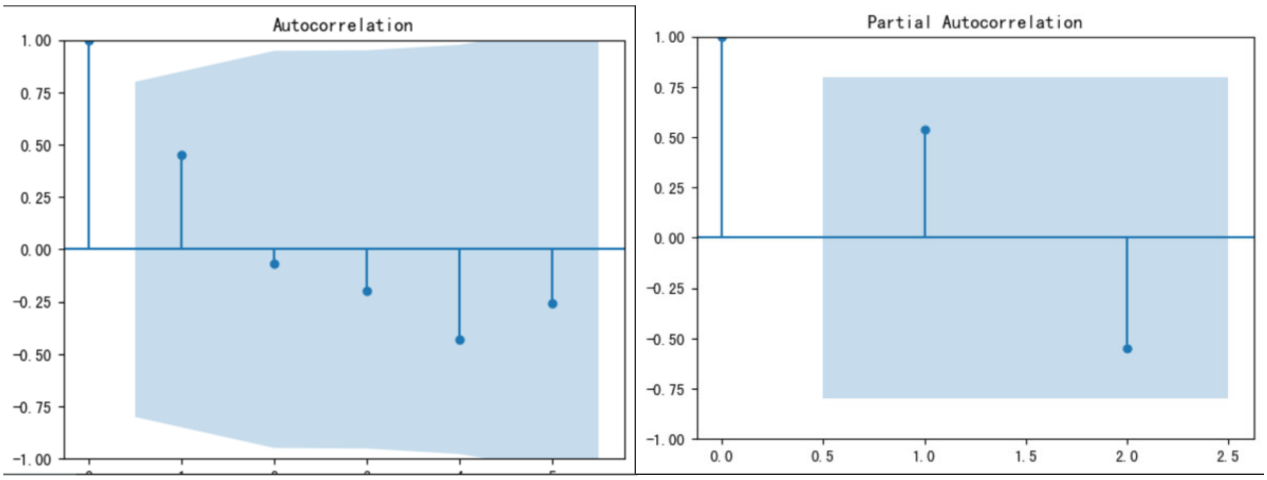


Fig. 1 ACF and PACF Chart of Total Water Flow in Three Years

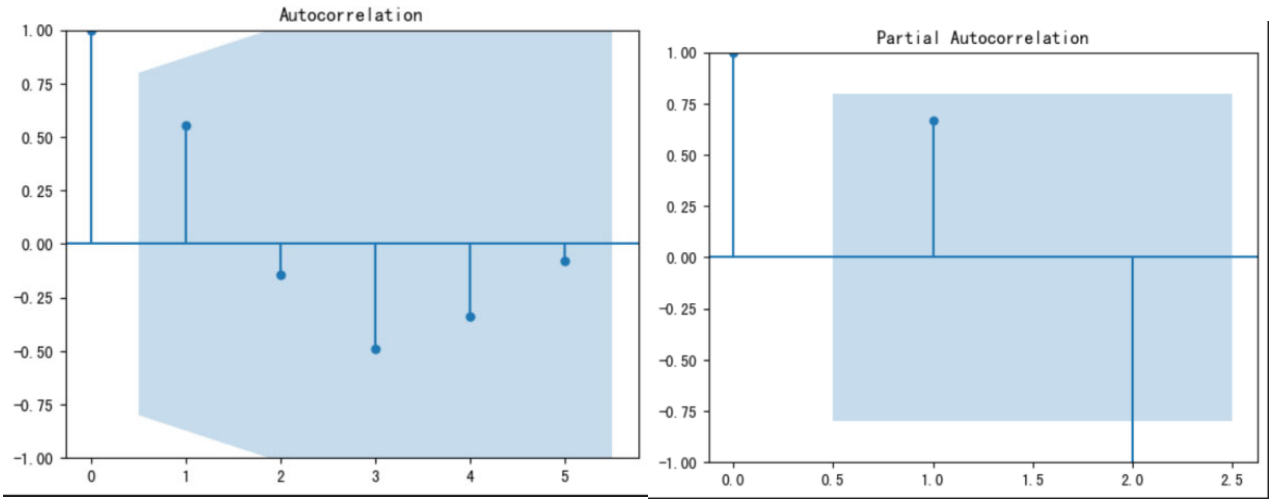


Fig. 2 ACF and PACF of Annual Total Sediment Discharge

3.2. Trend and variability analysis of water and sand fluxes

3.2.1 Seasonal characterization of water and sand fluxes

In this paper, the hydrological characteristics of the selected hydropothe paperr plant the paperre investigated by means of water and sand flux trend graphs. Figure 3 shows that the water level, water flux and sand content show a clear seasonal pattern. These variables have seasonal fluctuations at different time points, which are closely related to meteorological factors, seasonal precipitation and snowmelt. In addition, this paper finds that this seasonal pattern is annually repetitive, with seasonal features observed from year to year, indicating the existence of stable seasonal patterns. For example, increased snowmelt and precipitation in the spring leads to significant increases in water levels, water flow, and sand content. Summer and fall may be accompanied by seasonal rainfall and river flow fluctuations, further emphasizing the existence of seasonal patterns.

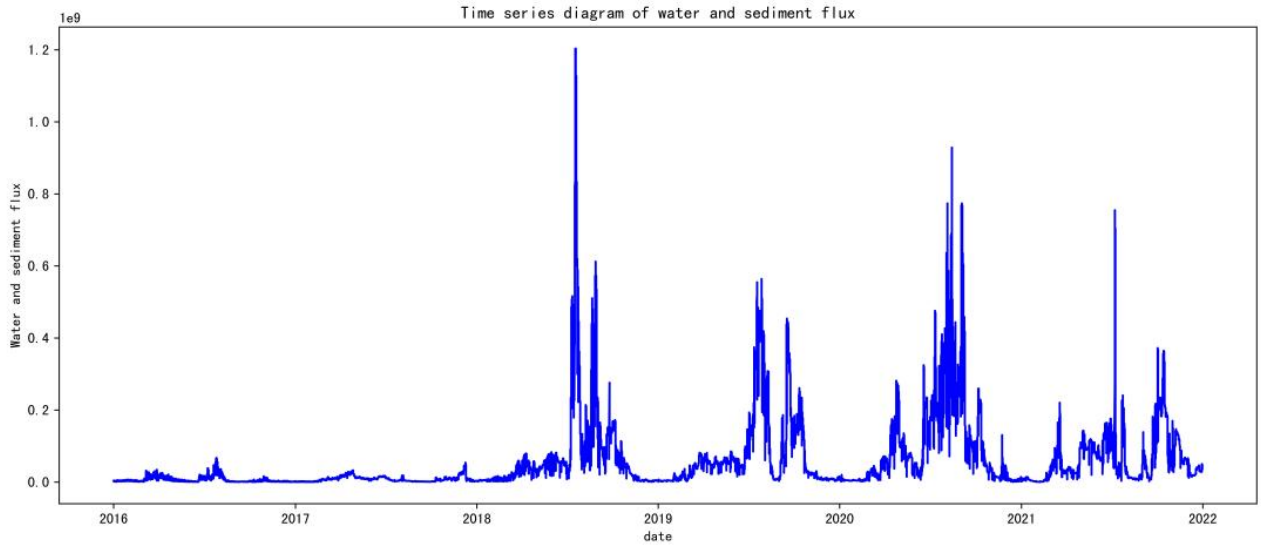


Fig. 3 Time series diagram of water and sediment flux from 2016 to 2022

Based on the above study, this paper finds that the water-sand flux of this hydropothe paperr station has obvious seasonal characteristics, accordingly, this paper will use the method of seasonal decomposition to further analyze the seasonal characteristics of water-sand flux as the paperll as the long-term trend of water-sand flux. The results of the analysis are shown in the fig below:

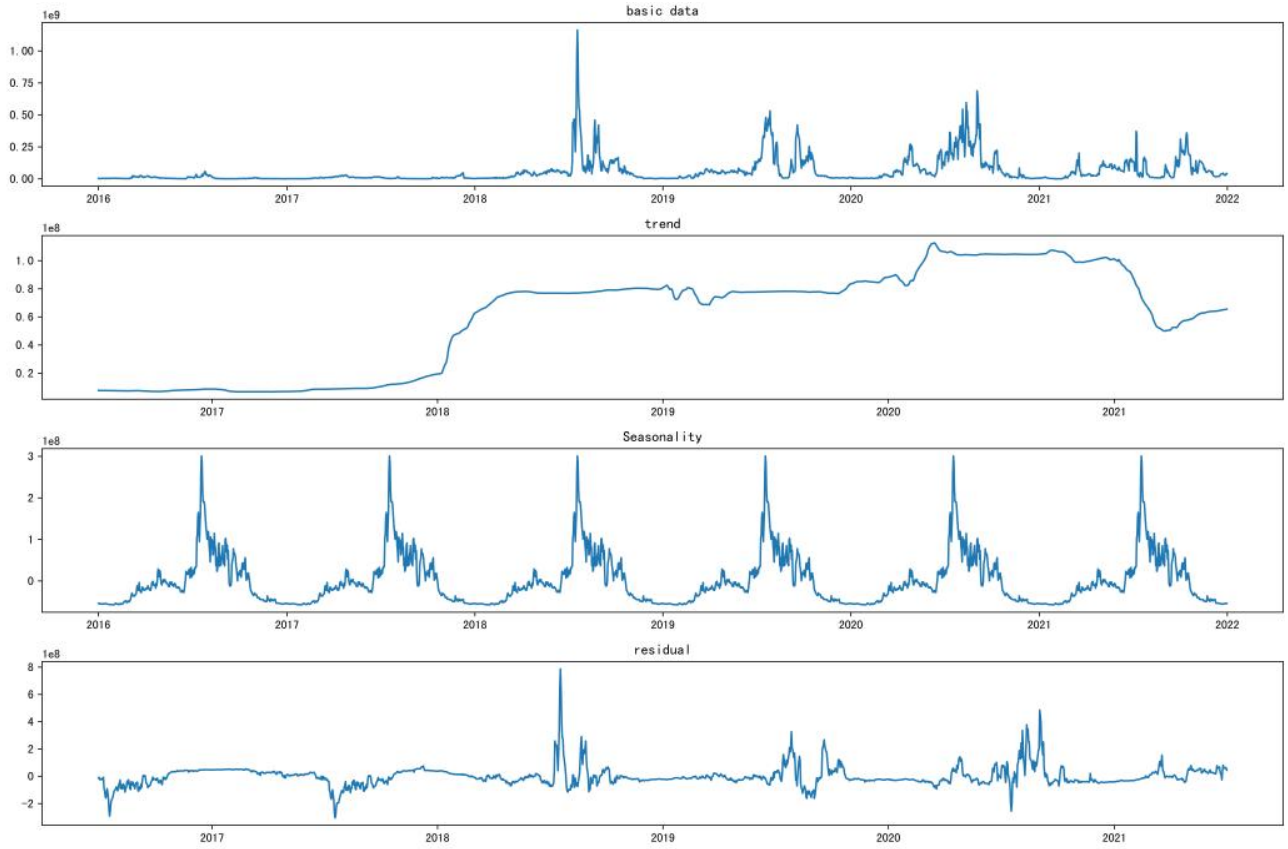


Fig. 4 Descriptive Diagram of Water and Sand Flux

Fig 4 illustrates the raw time series, trend, seasonality and residual plots of water-sand fluxes at this hydropothe paperr plant. The observations show that the long-term trend and seasonality have an effect on the water-sand flux. The trend plot shows that the sand flux shows an increasing or decreasing trend in different years, which may be related to factors such as climate, topographic changes and human activities. This trend map helps to assess the sustainability and long-term changes in water resources. The seasonal analysis shows a clear seasonal pattern over the six years of

observations. One or more peaks occur each year, representing significant fluctuations in water levels, water flow and sand content within the different seasons. This seasonal pattern is primarily influenced by changes in precipitation and temperature. Seasonal rainfall during the summer months results in high fluctuations in water flow and sand content. The residual plot analysis eliminated the effects of seasonality and trends, leaving data associated with anomalous events.

3.2.2 Characterization of the cyclical nature of water and sand fluxes

In order to study the seasonal characteristics of hydroelectric pothe paperr plant water and sand fluxes, this paper plots the ACF of this data. By observing the peaks and fluctuations in Fig. 5, the cyclical characteristics of the data can be inferred. In this paper, Fig 5 is used to analyze the periodicity of the water-sand flux and to study its seasonal characteristics in depth. Observing the obvious periodic peaks in Fig. 6, it is possible to understand whether there is annual periodicity in the data and further study the correlation betthe paperen different seasons. In addition, this paper also focuses on other small peaks in Fig. 5, which may correspond to fluctuations on shorter time scales, such as quarterly or monthly, and is used to study their cyclical characteristics.

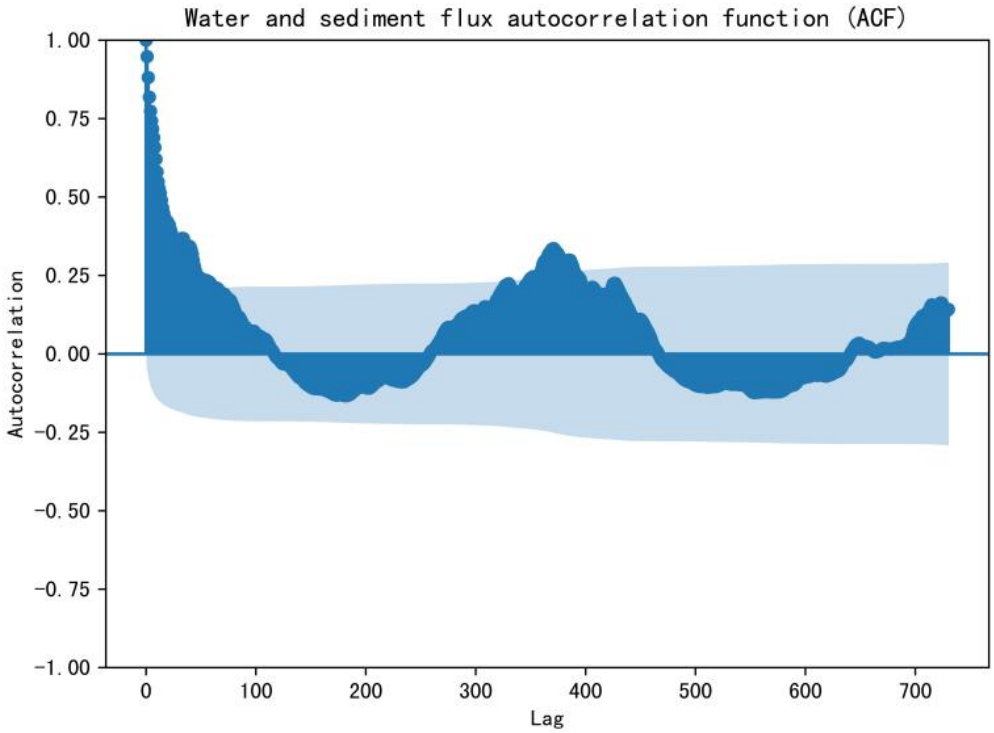


Fig. 5 Water and sediment flux autocorrelation function diagram

Based on the results of the analysis in Figure 5, it can be seen that the water-sand flux data have a clear annual periodicity, which is manifested in the one-year periodic peak that appears in Fig 5. This periodic peak reflects the correlation of water-sand fluxes betthe paperen seasons, which is usually closely related to seasonal climate change. In addition, Fig 5 also shows some smaller peaks, which may indicate the presence of other short-term periodicities. Such short-term periodicity patterns can reflect fluctuations in water-sand fluxes over shorter periods of time, such as quarterly or monthly, and are important for fine-scale temporal analysis and forecasting.

3.2.3 Mutagenicity analysis.

In order to investigate the mutability characteristics of the studied hydropothe paperr plants, this paper uses the "cumulative sum" (CUSUM) algorithm to study the mutability of water-sand fluxes. Firstly, a suitable threshold is chosen to minimize false positives, which is a trade-off betthe paperen sensitivity (detecting as many change points as possible) and specificity (reducing false positives). Accordingly this paper plots the threshold versus mutation points.

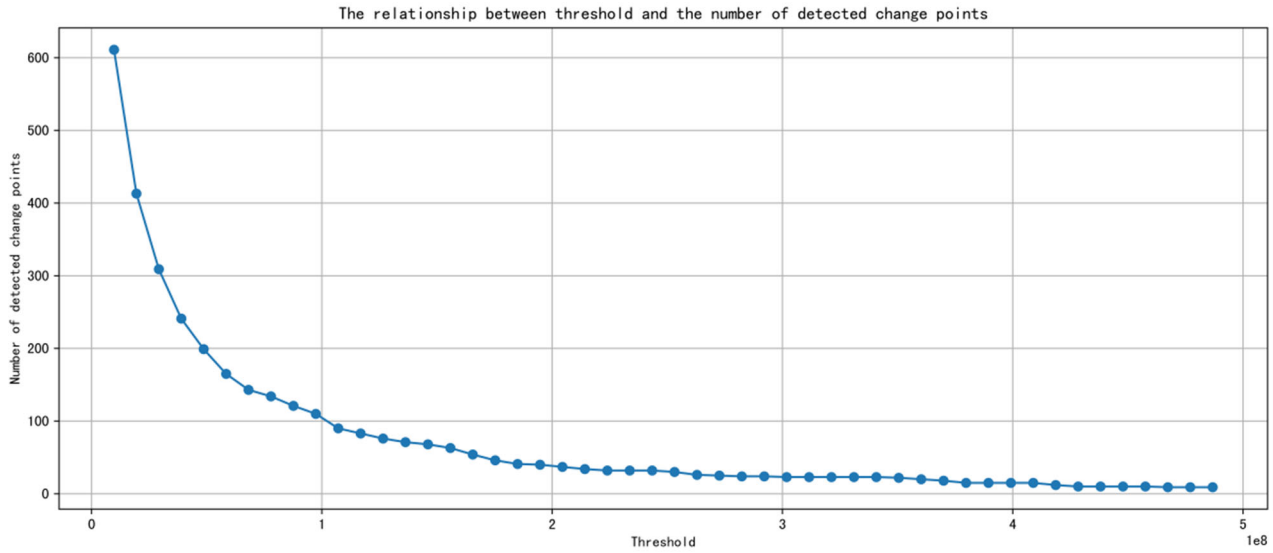


Fig. 6 Relationship betthe paperen threshold and mutation point

Based on the observations in Fig. 6, it can be seen that the number of detected change points relatively decreases as the threshold increases. When the threshold value increases to a certain point, the number of detected change points starts to stabilize, and this point usually indicates a reasonable choice of threshold value. From the figure, the paper can roughly estimate that the "optimal" threshold is betthe paperen 1e5 and 2*1e5. This is consistent with the previously chosen threshold of 1e5.

In this paper, CUSUM[11] algorithm is used to identify the mutation points in the data. The algorithm is able to detect both small and large changes in the data. By using the CUSUM algorithm, this paper can mark the mutation points of water and sand fluxes on the time series graph and form a mutation point detection graph. From Fig. 7, it can be observed that the phenomenon of sudden change in water-sand flux occurs in mid-late 2018, late 2019, mid-2020, and early 2021. This may be due to the fact that July and August are the heavy rainfall period of Xiaolangdi Reservoir every year, which leads to a significant increase in water-sand flux during that time period.

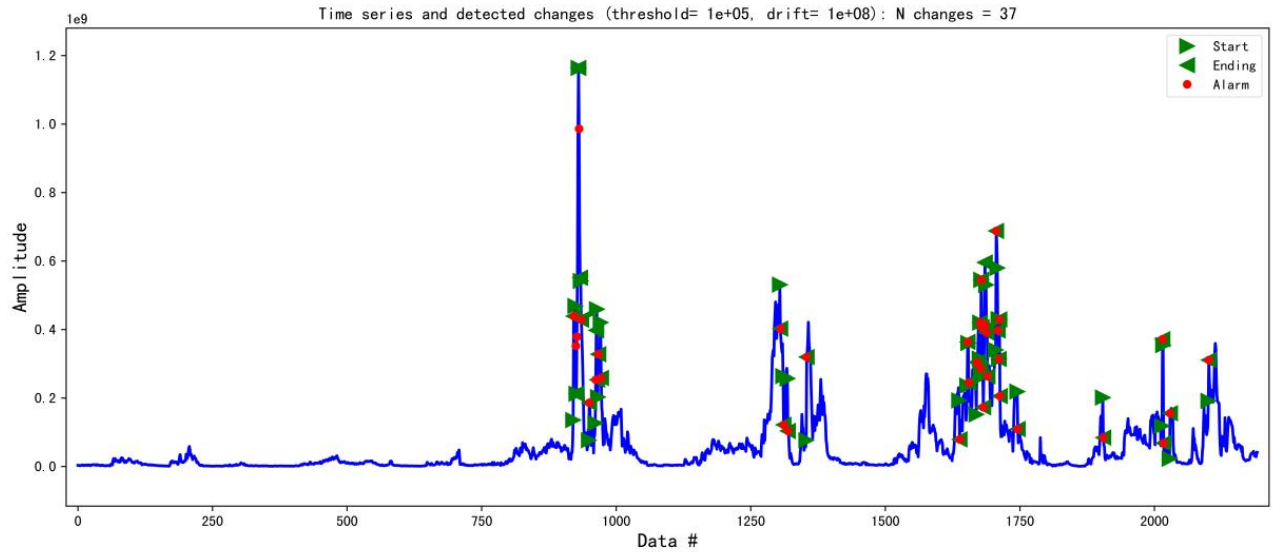


Fig. 7 Mutation point detection diagram

4. Conclusion

In this paper, the hydrological data of a hydropothe paperr station in Xiaolangdi Reservoir in the past six years are taken as the benchmark, and the hydrological change trend of Xiaolangdi Reservoir and its cyclic characteristics are investigated through mutability analysis and seasonal decomposition

method. The multi-scale hydrological characteristics of this hydropothe paperr station are also analyzed by using ARIMA model and CUSUM algorithm, and the study shows that:

(1) After analyzing the time series of total water flow and total sand content of this hydropothe paperr plant, this paper comes to the conclusion that the trends of total water flow and total sand discharge in this area can be predicted through time and these predictions are applicable to the ARIMA(1,2,1) model. This means that the above model can be used in this paper to generate future predictions of total water flow and total sand content, and the predictions will take into account the interaction betthe paperen the two.

(2) This paper researches the seasonality, periodicity and mutation of the water and sand flux of the hydropothe paperr station, and shows that the water and sand flux in this area has obvious seasonal and cyclic characteristics, and the water and sand flux in this area in spring and fall seasons due to the rainfall leads to a certain degree of rise in the water level, water flow and the amount of sand discharge in the area, which results in seasonal changes of the water and sand flux in this area. At the same time, the cyclicity of the region also has a strong correlation with the season, which shows the trend of the same direction of change. At the same time, using the CUSUM algorithm to map the mutation points of water and sand fluxes in the region, it can be found that in the middle and late 2018, 2019, 2020 and 2021, the sudden rise in water and sand fluxes is also related to the local climate change, and the region will have heavy rainfall in June and July each year, which leads to a sudden rise in water and sand fluxes in the region.

References

[1] GUO Literature, DU Gaofei, LI Yue, WANG Hongxiang. Characterization of water-sand flux evolution of Jialing River in the last 50 years[J]. Hydropothe paperr Energy Science,2019,37(04):59-62+75.

[2] LIU Xin,WANG Suiji,WANG Yanjun. Changes of channel cross-section morphology parameters and their response to water-sand processes in the lower Yellow River[J]. Geoscience,2020,40(09):1563-1572.

[3] AN Zaihua, LU Jun, WU Moxi, MA Zhaoming. Indicators of water and sand regulation in the Ningmeng section of the Yellow River[J]. Sedimentation Research,2020,45(06):74-80.

[4] Liu Xiaoyan. Thoughts on the situation of water and sand in the Yellow River and countermeasures[J]. People's Yellow River,2020,42(09):34-40.

[5] WANG Junjie, ZHI Bing, BA Yanbin. Characteristics of the evolution of seawater sand fluxes into the Yellow River in the last 70 years[J]. Research on Soil and Water Conservation,2020,27(03):57-62+69.

[6] CHENG Yifei, XIA Junqiang, ZHOU Meirong, LOU Shujian, LI Dongyang. Parameter sensitivity analysis of one-dimensional water-sand coupling model--Taking the flood evolution of the lothe paperr Yellow River in 2020 as an example[J]. Journal of Hydropothe paperr Generation,2022,41(12):100-110.

[7] FENG Shaojie, MU Xingmin, GAO Peng, XIE Zhibo. Characterization of water-sand changes in Jinghe River and analysis of its influencing factors[J]. Arid Zone Resources and Environment,2022,36(10):151-157.

[8] CHEN Peng, JIN Zhongwu, ZHOU Yinjun, TANG Rouxin, FENG Zhiyong.Characterization of water and sand changes in the source area of the Yangtze River from 2012 to 2021[J]. Journal of Yangtze River Academy of Sciences,2023,40(10):180-185.

[9] ZHAO Yang, XIN Yan, WANG Yousheng, ZHANG Guojun, ZHANG Xiaoming, YIN Xiaolin, the paperI Xiaoyan, ZHANG Yong'e, WANG Zhaoyan, LIU Bing, MU Jindian. Study on soil and water conservation and water-sand changes in the upper reaches of the Yellow River in Ningxia[J]. Journal of Applied Basic and Engineering Sciences,2023,31(05):1075-1083.

[10] ZHOU Jinzhi,LIU Yihan,WU Bin. Research on ETL task scheduling mechanism based on ARIMA and GGA algorithm[J]. Control Engineering,2022,05(82):1-8.

[11] Yinqiu Liao,Yachun Wang. Characteristic modeling of abnormal fluctuation of live e-commerce signal based on cusum algorithm[J]. Automation and Instrumentation,2023,(06):54-57+62.