Identification of long-term trends and seasonality in high-frequency water quality data from the Yangtze River basin, China

Weili Duan¹,²*, Bin He², Yaning Chen¹, Shan Zou¹,³, Yi Wang¹, Daniel Nover⁴, Wen Chen², Guishan Yang²

¹ State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences (CAS), Urumqi, Xinjiang, China, ² Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China, ³ Hengyang Normal University, Hengyang, China, ⁴ Department of Engineering, University of California–Merced, United States of America

* duan.scut.cn@gmail.com

Abstract

Comprehensive understanding of the long-term trends and seasonality of water quality is important for controlling water pollution. This study focuses on spatio-temporal distributions, long-term trends, and seasonality of water quality in the Yangtze River basin using a combination of the seasonal Mann-Kendall test and time-series decomposition. The used weekly water quality data were from 17 environmental stations for the period January 2004 to December 2015. Results show gradual improvement in water quality during this period in the Yangtze River basin and greater improvement in the Uppermost Yangtze River basin. The larger cities, with high GDP and population density, experienced relatively higher pollution levels due to discharge of industrial and household wastewater. There are higher pollution levels in Xiang and Gan River basins, as indicated by higher NH₄-N and COD_Mn concentrations measured at the stations within these basins. Significant trends in water quality were identified for the 2004–2015 period. Operations of the three Gorges Reservoir (TGR) enhanced pH fluctuations and possibly attenuated COD_Mn, and NH₄-N transporta-
tion. Finally, seasonal cycles of varying strength were detected for time-series of pollutants in river discharge. Seasonal patterns in pH indicate that maxima appear in winter, and minima in summer, with the opposite true for COD_Mn. Accurate understanding of long-term trends and seasonality are necessary goals of water quality monitoring system efforts and the analysis methods described here provide essential information for effectively controlling water pollution.

Introduction

Since 1979, China has experienced sustained and rapid economic growth with a concurrent and growing problem of environmental pollution [1–3]. With rapid urban development and accelerating industrialization, the demand for water is increasing, and meanwhile water quality is declining because of pollution. All these changes have made water security a major contemporary
challenge [4, 5]. According to the “2015 China Environmental Bulletin” in June 2016 [6], the proportion of water quality stations belonging to the best water quality category in seven major river basins (the Yangtze River, the Yellow River, the Pearl River, the Huai River, the Hai River, the Songhua River and the Liao River) in 2014 is 64.5%. There is no obvious change from 2014 to 2016, and much work will be necessary to achieve the 70% target (the ratio of the best water quality category in seven major river basins) for water quality by 2020. Pollution control and water quality restoration in these river basins are therefore key for future environmental efforts.

In order to make rivers and lakes clean and prevent water from being polluted, a surface water quality monitoring system was built by China’s Ministry of Environmental Protection in 2004. This system is similar to the National Water-Quality Assessment (NAWQA) Program in the United States [7, 8], the Harmonized Monitoring Scheme (HMS) in Britain [9, 10], and the National Monitoring and Assessment Program (NOVA) in Denmark [11, 12]. These systems can obtain reliable water quality data, which is the first step to estimate and control the water pollution. Meanwhile, based on these monitoring data, trend analysis (e.g., linear trend test and Mann–Kendall trend test [13]), load estimation methods (e.g., LOADEST[14]), and evaluation models (e.g., SPARROW[15], SWAT[16]) for water quality evaluation have been gradually developed and improved [17], and all these methods have been used successfully to assess pollution at a basin scale [18]. However, using a combination of several methods, few academic studies have employed these datasets to improve water quality assessments and regional water resources management by analyzing numerous parameters over various spatio-temporal scales in China [19, 20].

The Yangtze River plays a significant role in the history, culture, and economy of China, and it discharges large amounts of sediment and freshwater into the East China Sea, thus influencing the sea’s biogeochemical cycles. The aquatic environment of the Yangtze has been degraded drastically during recent decades [21, 22]. Current research on Yangtze River water quality monitoring focuses on the evaluation of water quality for human health, regional water pollution characteristics [23, 24] and related modeling exercises [25]. However, based on weekly data, little research has been done to analyze and understand seasonal water quality trends on the Yangtze River over ten years.

The objectives of this study are therefore (1) to analyze spatial and temporal distributions of water quality data collected from January 2004 to December 2015 at 17 monitoring stations in the Yangtze River basin, (2) to calculate trends and observe seasonality, and (3) to explore the relationship between water quality and river discharge, using a combination of several trend analysis methods including the seasonal Mann-Kendall test and time-series decomposition.

Data and methods

Study area

The Yangtze River, originates from Tanggula Mountain (elev. 6621 m), and flows 6,397 km eastward before emptying into the Eastern China Sea (ESC). Its drainage basin lies between 91˚E and 122˚E and 25˚N and 35˚N, covering one fifth of the whole nation with an area of about 1.8 x10^6 km^2 (Fig 1). The river is divided into an uppermost section (from the source to Yibin, Sichuan Province, with a length of 3,499 km), an upper section (from Yibin to Yichang, Hubei Province, with a length of 1,030 km), a middle section (from Yichang to Hukou, Jiangxi Province, with a length of 950 km) and a lower section (from Hukou to Shanghai, with a length of 938 km). The main river meanders through diverse landforms from west to east. The whole basin has complicated hydroclimatic conditions because of the East and South Asian monsoon activities, which brings seasonal discharge from wet (May to October) to dry (December to April) conditions. The mean annual precipitation varies from 300 to 500 mm in the western
region to 1,600–1,900 mm in the southeastern region, and the annual mean temperature in the southern and northern parts of the middle and lower Yangtze River basin is 19 and 15°C, respectively [26]. The Yangtze River basin supplies about 50% of the country’s runoff and 40% of Chinese freshwater resources [27].

Datasets
A total of 17 environmental monitoring stations were selected for analysis in this study, in which 7 stations (Number 1, 5, 7, 11, 13, 15 and 16) are located along the trunk stream, 2 stations (8 and 9) were setup for monitoring the Danjiangkou Reservoir, and the rest were built to assess the influence of major tributaries (see Fig 1 and Table 1). Weekly water quality data (including pH, CODMn, NH4-N, and DO) at these 17 stations were obtained and processed from a surface water quality monitoring system. The monitoring system covers seven river systems including the Yangtze, the Yellow, the Pearl, the Huai, the Hai, the Songhua and the Liao Rivers. So far, 148 monitoring stations have been completely established, containing 122 stations for monitoring rivers and 26 stations for monitoring lakes. All these stations test five

Fig 1. Geological map of the Yangtze River basin and studying stations.

https://doi.org/10.1371/journal.pone.0188889.g001
water quality parameters including pH, total organic carbon (TOC), chemical oxygen demand (COD\text{Mn}), ammonia-nitrogen (NH\text{4}-N), and dissolved oxygen (DO), every four hours. These data are processed by the government and used to generate weekly data and then the weekly data are published through the Data Center of China’s Ministry of Environmental Protection. So only weekly pH, COD\text{Mn}, NH\text{4}-N, and DO were selected and because generally the water quality data are not made public in China. Data were collected at all stations for the period 2004–2015, except for three stations (Qingfengxia, Xingang and Chuchuo) which were only available for 2006 to 2015. Weekly water grades are also described according to environmental quality standards for surface waters in China (GB3838-2002, see S1 Table) (Grade I–V level means that water is "Excellent", "Good", "Satisfactory", "Bad", and "Very bad" respectively).

Data related to river discharge, land use, gross domestic product (GDP), population density, and precipitation in the Yangtze River basin were also collected in order to analyze causes of spatiotemporal changes of water quality. Two hydrological stations (Yichang and Hukou), located in the junction zone of the upper, the middle and the lower Yangtze River, were chosen to obtain daily runoff data (Fig 1). The Yichang hydrological station is next to the Nanjingguan environmental station, and the Hukou hydrological station is next to the Hexishuichang environmental station. Land use data from 2010 were extracted from Landsat TM/ETM images with a spatial resolution of 1 km × 1 km at the national scale [28] and land use types were categorized as forest, wetland, desert, agriculture, grass, settlement, and others. Population and GDP for 2010 were provided by the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC), with 1km × 1km resolution. Average annual precipitation from 2000 to 2009 came from the China meteorological data center.

Methods

Data pre-processing and summary statistics. After the data extraction process, weekly water quality data and daily runoff data were categorized and transformed into an Excel-
format database. MATLAB (R2010a) was used to identify and assess data quality [29]. Data were then analyzed using descriptive statistics to quantitatively describe the main features of the dataset and time-series analysis (such as box plots) was performed on weekly data (the ‘weekly’ data come from processing of samples taken at 4 hourly time-steps) to allow a direct comparison between the hydrochemical information at different stations. Summary statistics including mean, median, and standard deviation were calculated for each pollutant. Box plots were used to evaluate changes in water quality and ArcGIS 10.0 was applied to display the spatial and temporal features of water quality at different stations. In addition, the Pearson’s correlation coefficient [30] was used to evaluate the relationship between water quality and river discharge.

To further understand water quality variability for the whole Yangtze River basin, the full period 2004–2015 was divided into three sub-periods (2004–2007, 2008–2011, and 2012–2015), which can obtain more characters for the changes of water quality. We first checked the trend for the whole period and then the probability distribution functions (PDFs) were calculated for regional mean (the average value of water quality at 17 stations) pH, COD$_{Mn}$, NH$_4$-N, and DO concentrations for all three sub-periods. A two-tailed Kolmogorov-Smirnov test was applied to assess whether the probabilities for different time periods are significantly different.

**Seasonal Mann–Kendall trend analysis.** The seasonal Mann–Kendall test (SMK) proposed by [31] was employed to detect monotonic trends of weekly water quality data. It is a nonparametric test and is used to detect potential trend change points in water quality trends [32–35]. Let $X = (X_1,X_2,\ldots,X_n)^T$ be a time series of independent water quality observations, and $X_i = (X_{i1},X_{i2},\ldots,X_{ij})$. Here, $n$ is the number of years and $j$ is the number of weekly data for each year and so for weekly ‘seasons’, the first week data are compared only with the first weeks data of every year, the second week data only with the second weeks data of every year, and so on. The null hypothesis ($H_0$) is that there are no monotonic trends in time. The statistic for the $g$th season is:

$$S_g = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\text{sgn}(X_{ig} - X_{jg}), g = 1, 2, \ldots, m$$

According to Hirsch et al. (1982), the SMK statistic, $\hat{S}$, for the entire series is calculated according to

$$\hat{S} = \sum_{g=1}^{m}S_g$$

For detailed treatment of this analysis, see [31]. Here, the significance level $p$ is selected at 0.05 and 0.10 with corresponding $Z$ statistics of 1.96 and 1.65, respectively. A positive value of $Z$ indicates an ’upward trend’ and a negative value of $Z$ indicates a ’downward trend’.

**Time-series decomposition.** Analysis of long-term trends and seasonal variability was carried out using the dynamic harmonic regression (DHR) technique, extensively described in [36]. Compared to a modified Clausius–Clapeyron equation, a multiple linear regression that includes both a linear and a harmonic dependence on time, and digital filtration (DF), DHR better fits the data and captures both the seasonal variations of the pollutant concentrations and the smaller scale interannual variations in the long term trends [37]. This method has been employed successfully in many examples with non-stationary environmental time-series analysis [38–40]. DHR decomposes an observed time series into component parts such as trend and seasonality [41]

$$y_t = T_t + C_t + S_t + e_t$$

where $y_t$ is the water quality time series, $T_t$ is a longer-term trend or low frequency
component), $C_t$ is a sustained cyclical component with a period separate from the seasonal component (e.g., a diurnal cycle), $S_t$ is the seasonal component (e.g., annual water quality seasonality), and $e_t$ is the residuals for water quality predications.

The trend ($T_t$) was extracted by the integrated random walk (IRW) model, which is a special case of the generalized random walk (GRW) [42]. The seasonal components ($S_t$) were defined and calculated as follows:

$$S_t = \sum_{i}^{N/2} [a_i \cos(\omega_i t) + b_i \sin(\omega_i t)]$$

$$\omega_i = \frac{2\pi \cdot i}{N} \quad i = 1, 2, \ldots, \left[\frac{N}{2}\right]$$

where $\omega_i$ values are the fundamental and harmonic frequencies associated with the periodicity in the observed water quality series chosen by reference to the spectral properties. For instance, the period 52 corresponds to a weekly sampling data in an annual cycle. The generalized random walk was applied to model the phase and amplitude parameters, which were estimated recursively using the Kalman filter and a fixed interval smoother. The squared correlation coefficient between the calculated seasonal component and detrended water quality data was used to determine significance of seasonality; the significance of the trend was decided on the basis of the squared correlation coefficient between the calculated trend and deseasonalized water quality data. Finally, the coefficient of determination ($R^2$) was applied to evaluate the performance of the model.

$$R^2 = 1 - \frac{\text{var}(\text{residuals})}{\text{var}(\text{observations})}$$

where $\text{var}$ is the variance estimate of the analyzed weekly water quality series.

Here, considering the regional representativeness of stations in different river segments and the distance between hydrological stations and environmental stations, two environmental stations (Nanjinguan and Hexishuichang) and two hydrological stations (Yichang and Hukou) that represent a large fraction of the total data were selected to present and discuss the DHR analysis. The runoff at Yichang station was highly impacted by the TGD, but it may also reflect the changes in annual runoff. The runoff data are daily data. Before discussing the correlations between runoff and water quality, we calculated the monthly annual runoff and water quality.

**Results**

**Overall water quality and temporal/spatial distribution**

Analysis shows that overall water quality improved from 2004 to 2015 (Figs 2 and 3), consistent with the results from [19]. Fig 2 shows changes in different grades of river water at 17 stations in the Yangtze River basin from 2004 to 2015 and shows a marked increase in Grade II water, occupying 66.7% of all water quality categories in 2015, and a slight increase in Grade I water, and a slight decrease in water of Grades III-IV. S1 Fig indicates overall change in fraction of Grade I-III water, Grade IV-V water, and Grade V+ water. Proportion of Grade I-III water has increased from 2004 to 2015, with 97.3% in 2015, suggesting gradual improvement in water quality in this time period.

The univariate statistics of the water quality parameters are presented in Table 2, and the distribution of pH, DO, COD$_{Mn}$, and NH$_4$-N from 2004 to 2015 at 17 stations are shown in the boxplots in S2 Fig. Generally, as the figure shows, the median pH and DO concentration at Station 1 and 6 are generally higher relative to the other stations, while the median COD$_{Mn}$
and NH$_3$-N concentrations are lower, revealing that the water quality in the Uppermost Yangtze River basin is better than in other areas. Except for Station 1, we see increasing concentrations of median DO and pH from Station 2 to Station 7, but decreasing concentrations of median COD$_{Mn}$ and NH$_4$-N concentrations. Station 2 (Minjiang Bride) had the highest median COD$_{Mn}$ concentration (3.47 mg/l), the lowest median DO concentration (6.47 mg/l), and relative highly NH$_4$-N concentration. We also found that the water was generally below Grade III during this period, consistent with previous work demonstrating that the middle and lower Min River was seriously polluted during 2003–2008 [43]. Station 14 (Chuchuo) had the highest median NH$_4$-N concentration (0.97 mg/l) and the lowest median pH (6.94), followed by station 10 (Xingang), revealing high ammonia nitrogen pollution in the Xiang River and Gan River.

Fig 3 shows the weekly probability distribution functions (PDFs) for regional mean concentrations (the average value of water quality at 17 stations) between 2004 and 2015 for the three time periods: 2004–2007 (black curve), 2008–2011 (green curve), and 2012–2015 (red curve). According to the changes can be inferred through the shift and shape of the curves between three time periods from the Figure, we can see that there has been a noticeable decrease in COD$_{Mn}$ and NH$_3$-N over the period (the PDF curve shifted from right (black curve) to left (red curve)), suggesting improvement overall in Yangtze River Basin water quality, especially over the 2012–2015 period; a tendency towards smaller pH-values was also found in recent years.

Overall water quality improvement is shown based on an evaluation of 17 environmental stations, which differs slightly from the State of Environment Report of China [44] indicating Grade I-III water increased slightly from 88.1% in 2014 to 89.4% in 2015, and Grade V+ water remained flat at 3.1%. However, the report agreed in that it showed overall improvement in the situation of the Yangtze River in recent years.
Fig 3. Weekly probability distribution functions for regional mean concentrations in the Yangtze River basin, between 2004 and 2015 for the three time periods: 2004–2007 (black curve), 2008–2011 (green curve), and 2012–2015 (red curve). (a) pH, (b) COD$_{Mn}$ (mg/L), (c) NH$_3$-N (mg/L), and (d) DO (mg/L).

https://doi.org/10.1371/journal.pone.0188889.g003

Table 2. Univariate statistics of the water quality parameters.

| Station | pH    | DO    | COD$_{Mn}$ | NH$_3$-N |
|---------|-------|-------|------------|----------|
|         | Mean  | SD    | Mean       | SD       | Mean  | SD    |
| 1       | 8.21  | 0.33  | 8.48       | 1.41     | 2.29  | 2.52  |
| 2       | 7.57  | 0.28  | 6.50       | 1.94     | 3.47  | 1.11  |
| 3       | 7.74  | 0.46  | 8.29       | 1.64     | 2.49  | 1.25  |
| 4       | 7.74  | 0.35  | 7.86       | 1.43     | 2.84  | 1.16  |
| 5       | 7.83  | 0.41  | 8.56       | 1.26     | 1.98  | 0.86  |
| 6       | 8.02  | 0.51  | 8.88       | 1.55     | 1.49  | 0.70  |
| 7       | 7.90  | 0.35  | 8.17       | 1.13     | 2.16  | 0.73  |
| 8       | 7.83  | 0.37  | 8.84       | 1.66     | 1.99  | 0.96  |
| 9       | 7.99  | 0.35  | 8.93       | 1.37     | 1.97  | 0.47  |
| 10      | 7.28  | 0.45  | 7.21       | 2.02     | 2.44  | 0.85  |
| 11      | 7.90  | 0.36  | 9.05       | 1.69     | 2.57  | 0.94  |
| 12      | 7.81  | 0.38  | 7.91       | 1.86     | 2.69  | 0.97  |
| 13      | 7.51  | 0.27  | 7.76       | 1.35     | 2.42  | 0.59  |
| 14      | 6.94  | 0.49  | 6.80       | 1.79     | 2.36  | 0.91  |
| 15      | 7.52  | 0.21  | 7.77       | 1.19     | 2.44  | 0.43  |
| 16      | 7.85  | 0.31  | 7.43       | 1.61     | 2.29  | 0.71  |
| 17      | 7.64  | 0.36  | 8.35       | 1.75     | 2.55  | 1.28  |

https://doi.org/10.1371/journal.pone.0188889.t002
Seasonal Mann Kendall trends

Fig 4 and Table 3 show seasonal trends (the value of the test statistic $Z$) at 17 stations in the Yangtze River basin from 2004 to 2015. pH tended to decrease at 11 stations (approximately 65%, 8 stations with significance at 95% confidence), which are mainly located on the upper and middle of the Yangtze River basin (Fig 4A). Total 10 stations (among them, 9 stations with significance at 95% confidence) exhibited a positive trend in COD$_{Mn}$ concentration, mainly occurring on major tributaries (e.g., station 10 on Xiang River, station 14 on Gan River) and the lower reach of the trunk stream (Fig 4B). For NH$_4$-N concentration, about half of the stations showed decreasing trends, most of which are distributed on the middle and lower reaches of the trunk stream (Fig 4C). Increasing seasonal trends were found at 12 stations for DO concentration (Fig 4D).

Fig 5 shows the Mann-Kendall trends (the value of $Z$) for 52 weeks at 17 stations in the Yangtze River basin, between 2004 and 2015, showing increasing long-term trends. From the figure, we can see that decreasing trends were found in COD$_{Mn}$ and NH$_4$-N concentrations at station 7 and station 11 over 52 weeks, while increasing trends in four water quality parameters at station 6 over 52 weeks, in NH$_3$-N concentration at station 16 and in COD$_{Mn}$ and DO concentrations at station 10 and station 14 over the time.

### Table 3. Seasonal trends (the value of $Z$) of stations with positive or negative trends for water quality parameters in the Yangtze River basin during 2004–2015.

| ID  | Range in z scores | Showing positive trend | Showing significant positive trend | Showing negative trend | Showing significant negative trend |
|-----|-------------------|------------------------|------------------------------------|------------------------|-----------------------------------|
| pH  | -10.3–11.6        | 6                      | 5                                  | 11                     | 8                                 |
| COD$_{Mn}$ | -12.9–11.3    | 10                     | 9                                  | 7                      | 6                                 |
| NH$_4$-N | -10.8–15.3      | 8                      | 6                                  | 9                      | 5                                 |
| DO  | -3.1–17.6         | 12                     | 11                                 | 5                      | 3                                 |

https://doi.org/10.1371/journal.pone.0188889.t003

Time-series decomposition results

**Long-term trends.** Results of the DRH analysis, conducted based on two representative environmental stations, are illustrated in Fig 6. Significant long-term trends at these two
environmental stations were identified for pH, DO, COD$_{Mn}$, and NH$_3$-N. The R$^2$ values for them ranged from 0.81 for COD$_{Mn}$ at Nanjinguan station to 0.94 for DO at Hexishuichang station, indicating good performance. Additionally, pH showed very clear seasonal variation at both stations, but the periodic disturbance at Nanjinguan station was more frequent (Fig 6A and 6B). It is possible that the operation of the three Gorges Reservoir (TGR) enhanced pH fluctuations at Nanjinguan station. For example, Wang et al. (2015) showed that the operation of releasing bottom water for power generation decreased the pH downstream from the dam because of the change of [CO$_2$]/[CO$_3^{2-}$] ratios [45]. Fig 6A shows that pH increased from 7.8 in January 2004 to 8.4 in July 2007 with a strong seasonal signal, then declined to 7.2 around January 2009 and then reached a new high in August 2011 and finally declined. At Hexishuichang station, pH increased from 7 in January 2004 to 8.3 in July 2008 and then decreased until June 2013 and finally remained at the same level.

As shown in Fig 6C, significant long-term declines were found in stream water COD$_{Mn}$ concentrations at Nanjingguan station, the value of which was about 2.5 mg/l in January 2004 and finally decreased to1.5 mg/l in December 2015. One reason might be the attenuation of the TGR, with nutrient and sediment removal rates in reservoirs significantly affecting transport in river systems [46]. Moreover, the impact of the TGR might be unstable because the conditions in the TGR have changed (e.g., operating rules or water supply) during the time period. The establishment of the TGR has three engineering phases including 1992–1997, 1998–2003 and 2003–2009. Also, operating rules have been adjusted for different flooding situations. By comparison, COD$_{Mn}$ at Hexishuichang station initially decreased to around 1.8 mg/l in February 2007, and then experienced two peaks (August 2010 and June 2012), and finally plateaued with only slight variation over the remaining years (Fig 6D).

Clear seasonal variability was found in NH$_4$-N concentrations at both stations, and both of them finally decreased from January 2004 to December 2015 (Fig 6E and 6F). NH$_4$-N at
Nanjinguan station experienced one peak (in June 2010) and then exhibited only slight variability from November 2013, while Hexhishuichang showed three peaks (two big peaks in August 2007 and July 2011, and a smaller one in 2014). At Nanjinguan station, the long-term decreasing trend for COD$_{Mm}$ generally parallels the trend in NH$_{4}$-N, with a significant Pearson’s Correlation coefficient between the COD$_{Mm}$ and NH$_{4}$-N trends (Pearson’s r = 0.62, p<0.01).
Fig 6G shows that DO concentrations at Nanjingguan station decreased slightly from January 2004 to June 2010, then increased dramatically around July 2012, and finally decreased in December 2015. In comparison, after three step-wise increases, DO concentrations at Hexishuichang station initially increased and peaked in May 2010, and then decreased to 6 mg/l in July 2014, and finally increased to 10 mg/l in December 2015.

Discharge at Yichang station initially decreased by February 2007, and then varied only slightly over the remaining period of record (Fig 6I). There are no similar overall trends for pH, DO, COD$_{Mn}$, or NH$_4$-N concentrations at this station. However, during the period from January 2007- February 2014, a similar decreasing trend was found for DO, COD$_{Mn}$, and NH$_4$-N. A slight variability over January 2004—February 2011 and a marked decline from March 2011 were found at Hukou station (Fig 6J), and similarly there was no overall trend in pH, DO, COD$_{Mn}$, or NH$_4$-N concentrations.

**Seasonality.** Initial review of weekly time-series plots (Fig 6) indicated a seasonal pattern for pollutants such as pH. Through DHR analysis, seasonal cycles of varying strength were extracted for all pollutant time-series (Table 4 and S3 Fig). Table 4 shows that the strongest seasonal cycle was identified for NH$_4$-N at Nanjinguan station (50.00%), while the weakest was for pH at Nanjinguan station (9.55%). Seasonal patterns in pH had small phases at both stations, with concentration maxima in winter, and minima in summer, which is consistent with previous work demonstrating that pH in June–August was slightly lower than in January–March in the middle Yangtze River [47]. The result was similar with other study sites. For example, Zeb et al. indicated that pH in winter was higher than in summer in Khyber Pakhtunkhawa (KPK) province of Pakistan [48].

COD$_{Mn}$ at both stations exhibited a seasonal cycle with variability in amplitude throughout the study period. Concentration peaks occurred in summer and minima in winter at Nanjinguan station, and peaks in spring and minima in the end of summer and start of autumn at Hexishuichang station. Generally, NH$_4$-N and DO exhibited similar seasonal cycles at both stations. The crest and trough of DO reflected higher values in winter and lower values in summer, which agrees with results from [49]. DO concentration tended to track river temperature, with high concentrations in winter and early spring (times of lower water temperature), and low concentrations in summer and fall (times of higher water temperature).

Similar and clear seasonal cycles of varying strength were found for discharge at both stations (S3E and S3F Fig), with discharge maxima in summer, and minima in winter, reflecting seasonal precipitation (wet condition from May to October and dry condition from December...
to April) influenced by both East and South Asian monsoon activities exhibited in the Yangtze River basin. Discharge patterns mirror pH in seasonality, although the amplitude of discharge was variable, especially at Hukou station.

Discussions

Rapid economic development and increasing population have brought significant changes in land cover and placed stress on water quantity and quality within the Yangtze River basin. Spatial distribution of mean concentration at each station, with GDP of 2010, population of 2010, mean precipitation between 2001 and 2010, gross domestic product (GDP), and land use of 2010, are shown in Fig 7. The Figure shows that GDP and population is concentrated in big cities such as the Chengdu-Chongqing Economic Zone along the Upper Yangtze River, the Wuhan-Changsha-Nanchang Economic Zone along the Middle Yangtze River, and the Yangtze River Delta Economic Zone (e.g., Shanghai, Hangzhou, and Nanjing) along the Lower Yangtze River. Water quality in these area (e.g., stations 2–4, 10, 14, and 17) was relatively poor because of pollutants such as NH$_4^+$-N and COD$_{Mn}$ associated with industrial activities and sanitary sewage [50, 51]. Pollutants in industrial and household waste-water discharges were high in these cities (Fig 8 and S4 Fig), especially in Shanghai, Hangzhou, Chongqing and Chengdu. Moreover, the household waste-water discharged increased between 2011 and 2013 as people migrated to cities from the country [52]. Regional high pollution may be a consequence of drainage from municipal sewage, agricultural wastewater, and livestock production facilities. Water quality is strongly affected by land-use through changes of pollutant input and the water cycle. On the one hand, with the growth of agricultural and settlement land, point source pollution and nonpoint source pollution are likely to increase because of incremental sewage treatment plants and additional use of agricultural fertilizer, like other countries [53]. On the other hand, land use may accelerate pollutants transport by changing the form of runoff. For example, most rain that falls on a parking lot runs off immediately [54], often draining into storm sewers with pollutants that transport it to a stream or ditch without filtration.

Fig 7. Spatial distributions of mean concentrations at 17 stations, with GDP of 2010 (Chinese yuan/km$^2$), population of 2010 (person/km$^2$), mean precipitation between 2001 and 2010 (mm), and land use of 2010. (a) pH, (b) COD$_{Mn}$ (mg/L), (c) NH$_4^+$-N (mg/L), and (d) DO (mg/L).

https://doi.org/10.1371/journal.pone.0188889.g007
In addition, industrial waste-water discharge was nearly unchanged in Nanchang and Changsha between 2005 and 2013, but industrial pollutants and ammonia nitrogen discharge have increased, especially in the last three years (Fig 8). These pollutants entered the Xiang and Gan Rivers, possibly explaining the high concentration and increasing trends at stations 10 and 14. Yan et al. indicated that agricultural activities, industrial and domestic wastewater have caused increasing loads of nitrogen to be discharged into the Yangtze River between 1970 and 2003 [55]. In general, water quality at stations on northern tributaries (including 6, 8 and 9) were better than southern tributaries (e.g., station 10 on the Xiang River and station 14 on the Gan River), and NH$_4^+$-N at stations on the trunk stream were relatively low. Meanwhile, heavy metals and nitrogen are the major pollutants in the lower Xiang [56, 57] and Gan Rivers [58]. Although these changes are likely to contribute to deteriorating water quality, enormous effort has been made in recent decades in China including an ever-improving legal system and the popularity of sewage treatment plants, leading to improving water quality.
Finally, Fig 4 and Table 3 show that there has been a measurable decrease in pollutants before flowing into dams and reservoirs and after flowing from them. For example, water quality after the dams is improving than before the dams; mean NH$_4$-N at stations 10 and 14 were 0.53 mg/l and 0.65 mg/l respectively, but dramatically decreased to 0.27 mg/l and 0.20 mg/l at stations 11 and 15 after flowing from Dongting and Poyang Lakes. Possibly, high runoff from other tributaries with low pollutant concentration may dilute outflows from lakes and reservoirs and dams are positively influencing the retention of pollutants along the river. Additionally, attenuation of pollutant concentrations due to lakes is another well-established explanation for the data [59], especially under some of appropriate planning and contamination monitoring policies.

Conclusions

This study characterizes the spatio-temporal distribution, long-term trend, and seasonality of water quality in the Yangtze River basin using statistic methods and time-series decomposition. The used dataset were weekly water quality data (pH, COD, NH$_4$-N, and DO) from 17 environmental stations for the period January 2004 through December 2015. The weekly water quality data allowed analysis of long-term trends and provided new insights into the changing amplitude and phase of the seasonality of the pollutants within the Yangtze River basin. Water quality gradually improved during this time period in the Yangtze River basin, but regional differences are still obvious. For example, high ammonia nitrogen pollution can be seen in the Xiang and Gan River basins because of high pollutants from industrial activities and sanitary sewage around these rivers. In addition, significant seasonal trends were identified in weekly pH, DO, COD$_{Mn}$, and NH$_4$-N concentration over the 2004–2015 period, and seasonal cycles of varying strength were extracted for pollutant time-series, suggesting the seasonal cycles of water quality in the Yangtze River basin. All these results could be helpful for fully understanding the seasonal trends of long-term water quality in the Yangtze River basin, which can provide essential information for effectively controlling water pollution and managing water resources.

Supporting information

S1 Fig. Changes of number of different grades of river water based on 17 stations in the Yangtze River basin from 2004 to 2015. (TIF)

S2 Fig. Boxplots illustrating distribution of water quality at 17 stations from 2004 to 2015 in the Yangtze River basin. Red lines are the mediums. (a) pH, (b) COD$_{Mn}$ (mg/L), (c) NH$_3$-N (mg/L), and (d) DO (mg/L). (TIF)

S3 Fig. Seasonal cycles extracted from the weekly pH, DO, COD$_{Mn}$, NH$_4$-N at environmental stations 7 (left) and 13 (right). The last two sub-graphs are for seasonal cycles of river discharges at Yichang and Hukou respectively. (TIF)

S4 Fig. Main pollutant emission in household waste water in main cities from 2005 to 2013. (TIF)

S1 Table. Environmental quality standards for surface water(mg/L) (GB3838 2002a). (DOCX)
Acknowledgments

This study was sponsored by the National Natural Science Foundation of China (No. 41471460 and 41501552), the Natural Science Foundation of Jiangsu Province (Grants No BK20161612), the Talent Introduction Project of Nanjing Institute of Geography and Limnology, CAS (NIGLAS2015QD09), and “One Hundred Talents Program” of the Chinese Academy of Sciences. We also wish to acknowledge Dr. C. James Taylor for his help.

Author Contributions

Data curation: Shan Zou.
Formal analysis: Weili Duan.
Funding acquisition: Weili Duan, Bin He.
Investigation: Weili Duan.
Methodology: Weili Duan, Yaning Chen, Shan Zou.
Supervision: Bin He.
Validation: Weili Duan.
Writing – original draft: Weili Duan.
Writing – review & editing: Weili Duan, Yaning Chen, Yi Wang, Daniel Nover, Wen Chen, Guishan Yang.

References

1. Liu J, Diamond J. China’s environment in a globalizing world. Nature. 2005; 435(7046):1179–86. https://doi.org/10.1038/4351179a PMID: 15988514
2. Zhang X, Chen C, Ding J, Hou A, Li Y, Niu Z, et al. The 2007 water crisis in Wuxi, China: analysis of the origin. J Hazard Mater. 2010; 182(1):130–5.
3. Xu Z, Zhang X, Xie J, Yuan G, Tang X, Sun X, et al. Total nitrogen concentrations in surface water of typical agro- and forest ecosystems in China, 2004–2009. PloS one. 2014; 9(3):e92850. https://doi.org/10.1371/journal.pone.0092850 PMID: 24667701
4. Piao S, Ciais P, Huang Y, Shen Z, Peng S, Li J, et al. The impacts of climate change on water resources and agriculture in China. Nature. 2010; 467(7311):43–51. https://doi.org/10.1038/nature09364 PMID: 20811450
5. Yin Z, Walcott S, Kaplan B, Cao J, Lin W, Chen M, et al. An analysis of the relationship between spatial patterns of water quality and urban development in Shanghai, China. Computers, Comput Environ Urban Syst. 2005; 29(2):197–221.
6. China’s Ministry of Environmental Protection. 2015 China Environmental Bulletin.; 2016. p. 2014–5.
7. Gilliom RJ, Alley WM, Gurtz ME. Design of National Water-Quality Assessment Program. United States Government Print, Office; 1995.
8. Rosen MR, Lapham WW. Introduction to the US Geological Survey National Water-Quality Assessment (NAWQA) of ground-water quality trends and comparison to other national programs. J Environ Qual. 2008; 37(5_Supplement):190.
9. Morvan X, Saby N, Arrouays D, Le Bas C, Jones R, Verheijen F, et al. Soil monitoring in Europe: a review of existing systems and requirements for harmonisation. Sci Total Environ. 2008; 391(1):1–12. https://doi.org/10.1016/j.scitotenv.2007.10.046 PMID: 18063012
10. Hurley MA, Currie JE, Gough J, Butterwick C. A framework for the analysis of harmonised monitoring scheme data for England and Wales. Environmetrics. 1996; 7(4):379–90.
11. Kronvang B, Jeppesen E, Conley DJ, Sandergaard M, Larsen SE, Ovesen NB, et al. Nutrient pressures and ecological responses to nutrient loading reductions in Danish streams, lakes and coastal waters. J Hydrol. 2005; 304(1):274–88.
12. Conley DJ, Markager S, Andersen J, Ellermann T, Svendsen LM. Coastal eutrophication and the Danish national aquatic monitoring and assessment program. Estuar Coast. 2002; 25(4):848–61.
13. Hirsch RM, Slack JR, Smith RA. Techniques of trend analysis for monthly water quality data. Water Resour Res. 1982; 18(1):107–21.
14. Runkel RL, Crawford CG, Cohn TA, US GS. Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. US Department of the Interior, US Geological Survey; 2004.
15. Schwarz GE, Hoos AB, Alexander RB, Smith RA. The SPARROW surface water-quality model: theory, application and user documentation. US geological survey techniques and methods report, book. 2006; 6(10):248.
16. Arnold JG, Moriasi DN, Gassman PW, Abbaspour KC, White MJ, Srinivasan R, et al. SWAT: Model use, calibration, and validation. T ASABE. 2012; 55(4):1491–508.
17. Shrestha S, Kazama F, Newham LT. A framework for estimating pollutant export coefficients from long-term in-stream water quality monitoring data. Environ Modell Softw. 2008; 23(2):182–94.
18. Duan W, He B, Nover D, Yang G, Chen W, Meng H, et al. Water quality assessment and pollution source identification of the Eastern Poyang Lake basin using multivariate statistical methods. Sustainability. 2016; 8(2):133.
19. Sun J, Chen Y, Zhang Z, Wang P, Song X, Wei X, et al. The spatio-temporal variations of surface water quality in China during the "Eleventh Five-Year Plan". Environ Monit Assess. 2015; 187(3):1–12.
20. Qiao Y, Feng J, Liu X, Wang W, Zhang P, Zhu L. Surface water pH variations and trends in China from 2004 to 2014. Environ Monit Assess. 2016; 188(7):1–13.
21. Müller B, Berg M, Yao ZP, Zhang XF, Wang D, Pfluger A. How polluted is the Yangtze river? Water quality downstream from the Three Gorges Dam. Sci Total Environ. 2008; 402(2–3):232–47. https://doi.org/10.1016/j.scitotenv.2008.04.049 PMID: 18554685
22. Li M, Xu K, Watanabe M, Chen Z. Long-term variations in dissolved silicate, nitrogen, and phosphorus flux from the Yangtze River into the East China Sea and impacts on estuarine ecosystem. Estuar Coast Shelf Sci. 2007; 71(1–2):3–12.
23. Wu B, Zhao D, Zhang Y, Zhang X, Cheng S. Multivariate statistical study of organic pollutants in Nanjing reach of Yangtze River. J Hazard Mater. 2009; 169(1):1093–8.
24. Ye C, Li S, Zhang Y, Zhang Q. Assessing soil heavy metal pollution in the water-level-fluctuation zone of the Three Gorges Reservoir, China. J Hazard Mater. 2011; 191(1):366–72.
25. Ding X, Shen Z, Hong Q, Yang Z, Wu X, Liu R. Development and test of the export coefficient model in the upper reach of the Yangtze River. J Hydrol. 2010; 383(3):233–44.
26. Zhang Q, Xu C, Zhang Z, Chen YD, Liu C, Lin H. Spatial and temporal variability of precipitation maxima during 1960–2005 in the Yangtze River basin and possible association with large-scale circulation. J Hydrol. 2008; 353(3):215–27.
27. Hollert H. Processes and environmental quality in the Yangtze River system. Environ Sci Pollut R. 2013; 20(10):6904–6.
28. Liu J, Zhang Z, Xu X, Kuang W, Zhou W, Zhang S, et al. Spatial patterns and driving forces of land use change in China during the early 21st century. J Geogr Sci. 2010; 20(4):483–94.
29. MathWorks. MATLAB version 7.10.0 (R2010a). Natick, Massachusetts: The MathWorks. 2010.
30. Ahlgren P, Jarneving B, Rousseau R. Requirements for a cocitation similarity measure, with special reference to Pearson’s correlation coefficient. J Am Soc Inf Sci Tec. 2003; 54(6):550–60.
31. Hirsch RM, Slack JR. A nonparametric trend test for seasonal data with serial dependence. Water Resour Res. 1984; 20(6):727–32.
32. Bouza-Deaño R, Ternero-Rodríguez M, Fernández-Espinosa AJ. Trend study and assessment of surface water quality in the Ebro River (Spain). J Hydrol. 2008; 361(3–4):227–39.
33. Chiang H. Spatial analysis of water quality trends in the Han River basin, South Korea. Water Res. 2008; 42(13):3285–304. https://doi.org/10.1016/j.watres.2008.04.006 PMID: 18490047
34. Zhang Y, Cabilio P, Nadeem K. Improved Seasonal Mann-Kendall Tests for Trend Analysis in Water Resources Time Series. Advances in Time Series Methods and Applications : Springer; 2016. p. 215–29.
35. Helsel DR, Frans LM. Regional Kendall test for trend. Environ Sci Technol. 2006; 40(13):4066–73. PMID: 16856718
36. Taylor CJ, Pedregal DJ, Young PC, Tych W. Environmental time series analysis and forecasting with the Captain toolbox. Environ Modell Softw. 2007; 22(8):797–814.
37. Venier M, Hung H, Tych W, Hites RA. Temporal trends of persistent organic pollutants: A comparison of different time series models. Environ Sci Technol. 2012; 46(7):3928–34. https://doi.org/10.1021/es204527q PMID: 22432692
38. Halliday SJ, Wade AJ, Skeffington RA, Neale C, Reynolds B, Rowland P, et al. An analysis of long-term trends, seasonality and short-term dynamics in water quality data from Plynlimon, Wales. Sci Total Environ. 2012; 434:186–200. https://doi.org/10.1016/j.scitotenv.2011.10.052 PMID: 22119034

39. Haddad M, Taibi H, Arezki SMM. On the recent global mean sea level changes: Trend extraction and El Niño’s impact. Cr Geosci. 2013; 345(4):167–75.

40. Minaudo C, Meybeck M, Moatar F, Gassama N, Curie F. Eutrophication mitigation in rivers: 30 years of trends in spatial and seasonal patterns of biogeochemistry of the Loire River (1980–2012). Biogeoosciences. 2015; 12(8):2549–63.

41. Young PC, Pedregal DJ, Tych W. Dynamic harmonic regression. J Forecasting. 1999; 18(6):369–94.

42. Pedregal DJ, Trapero JR. Electricity prices forecasting by automatic dynamic harmonic regression models. Energ Convers Manage. 2007; 48(5):1710–9.

43. Zhao J, Fu G, Lei K, Li Y. Multivariate analysis of surface water quality in the Three Gorges area of China and implications for water management. J Environ Sci. 2011; 23(9):1460–71.

44. China’s Ministry of Environmental Protection. 2015 State of Environment Report, Department/Division, 2016: http://www.zhb.gov.cn/gkml/hbb/qt/201606/t20160602_353078.htm, (Chinese, accessed January 20, 2017).

45. Wang B, Liu C, Wang F, Liu X, Wang Z. A decrease in pH downstream from the hydroelectric dam in relation to the carbon biogeochemical cycle. Environ Earth Sci. 2015; 73(9):5299–306.

46. Brakebill JW, Ator SW, Schwarz GE. Sources of Suspended-Sediment Flux in Streams of the Chesapeake Bay Watershed: A Regional Application of the SPARROW Model I. J Am Water Resour Assoc. 2010; 46(4):757–76.

47. Li X, Gan Y, Zhou A, Liu Y. Relationship between water discharge and sulfate sources of the Yangtze River inferred from seasonal variations of sulfur and oxygen isotopic compositions. J Geochem Explor. 2015; 153:30–9.

48. Zeb BS, Malik AH, Waseem A, Mahmood Q. Water quality assessment of Siran river, Pakistan. Int. J Phys Sci. 2011; 6(34):7789–98.

49. Zhao J, Fu G, Lei K. Water Quality Assessment and Apportionment of Pollution Sources of Selected Pollutants in the Min Jiang, a Headwater Tributary of The Yangtze River. Hydrof Current Res. 2015; 6: 211. https://doi.org/10.4172/2157-7587.1000211

50. Wang J, Da L, Song K, Li B. Temporal variations of surface water quality in urban, suburban and rural areas during rapid urbanization in Shanghai, China. Environ Pollut. 2008; 152(2):387–93. https://doi.org/10.1016/j.envpol.2007.06.050 PMID: 17681654

51. Zhang Z, Tao F, Du J, Shi P, Yu D, Meng Y, et al. Surface water quality and its control in a river with intensive human impacts—a case study of the Xiangjiang River, China. J Environ Manage. 2010; 91(12):2483–90. https://doi.org/10.1016/j.jenvman.2010.07.002 PMID: 20674147

52. He G, Mol AP, Lu Y. Wasted cities in urbanizing China. Environ Dev. 2016; 18:2–13.

53. Luo P, He B, Takara K, Razafindrabeh BH, Nover D, Yamashiki Y. Spatiotemporal trend analysis of recent river water quality conditions in Japan. J Environ Monitor. 2011; 13(10):2819–29.

54. Duan W, He B, Nover D, Fan J, Yang G, Chen W, et al. Floods and associated socioeconomic damages in China over the last century. Nat Hazards. 2016; 82(1):401–13.

55. Yan W, Mayorga E, Li X, Seitzinger SP, Bouwman AF. Increasing anthropogenic nitrogen inputs and riverine DIN exports from the Changjiang River basin under changing human pressures. Global Biogeochem Cy. 2010; 24(4):1134–8.

56. Zhang Q, Li Z, Zeng G, Li J, Fang Y, Yuan Q, et al. Assessment of surface water quality using multivariate statistical techniques in red soil hilly region: a case study of Xiangjiang watershed, China. Environ Monit Assess. 2009; 152(1–4):123–31. https://doi.org/10.1007/s10661-008-0301-y PMID: 18535918

57. Wang L, Guo Z, Xiao X, Chen T, Liao X, Song J, et al. Heavy metal pollution of soils and vegetables in the midstream and downstream of the Xiangjiang River, Hunan Province. J Geogr Sci. 2008; 18(3):353–62.

58. Zhang X, Huang GH, Nie X. Possibilistic stochastic water management model for agricultural nonpoint source pollution. J Water Resour Plan Manage. 2010; 137(1):101–12.

59. Duan W, He B, Takara K, Luo P, Nover D, Hu M. Modeling suspended sediment sources and transport in the Ishikari River basin, Japan, using SPARROW. Hydrol Earth Syst Sc. 2015; 19(3):1293–306.