Joint occurrence of water quality indexes in relation to river streamflow in the heavily polluted Huai River Basin, China

Shaofei Wu, Xiang Zhang and Dunxian She

ABSTRACT

Effective joint management of water quantity and quality of rivers requires a good understanding of the interrelationships between these variables. This study aims to quantitatively evaluate the impact of river streamflow on the joint occurrence of water quality variables in the heavily polluted Huai River Basin (HRB). Using data from three representative stations, joint distributions were set up for two key water quality indexes, COD_{Mn} and NH_3-N, for three different streamflow scenarios: low, medium and high streamflows, and the joint probabilities of different water quality combinations were calculated. The results showed that there was a significant negative relationship between streamflow and the water quality index, while the water quality indexes were significantly positively correlated. In general, the higher the river streamflow, the lower the joint probability of water quality variables under the three scenarios in each station. It is concluded that, in both the main stream and the tributary, high streamflow scenarios do help to improve the joint probability of water quality combination under a higher water quality grade, compared with a decrease under a lower quality grade. This study is expected to provide scientific references for water quality management and implementation of water pollution prevention in the HRB.

Key words | copula functions, different scenarios of river streamflow, Huai River Basin (HRB), relationship between water quantity and quality

Shaofei Wu
Xiang Zhang (corresponding author)
Dunxian She
State Key Laboratory of Water Resources and Hydropower Engineering Science,
Wuhan University,
Wuhan 430072,
China
and
Hubei Provincial Collaborative Innovation Center for Water Resources Security,
Wuhan 430072,
China
E-mail: *zhangxiang@whu.edu.cn*

Shaofei Wu
Jiangxi Provincial Institute of Water Science,
Nanchang 330029,
China

INTRODUCTION

Managing water resources for the maintenance of ecosystems and human benefits in developed watersheds with regulated flows usually requires consideration of both water quantity and quality. Approaches to water quality management include source control, ensuring effluents remain within tolerable limits, encouraging pollutant degradation or settling, and diluting contaminants through management of discharge. While all these approaches can be effective, in rivers where source control is problematic and the risk of exceeding permitted or desirable limits of contaminant concentration is high, joint regulation of water quality and water quantity is, and will continue to be, an important decision-making technology (Zhang *et al.* 2010).

The surface water quality at any point on a river reflects both natural sources and processes and anthropogenic activities (Hesse & Krysanova 2016). The distribution of water quality is naturally determined by soil characteristics, soil erosion, precipitation, and biological and physical catchment processes (Shi *et al.* 2017), and the influence of these factors varies with geology, topography and season (Liu *et al.* 2016; Shi *et al.* 2016). Relevant anthropogenic activities include construction and operation of reservoirs and dams, agriculture, urbanization, and industrial effluents (Shin

doi: 10.2166/ws.2017.046

et al. 2013). Domestic and industrial wastewater contributes a relatively constant source of pollution, whereas surface streamflow is a seasonal phenomenon, affected both by climate condition and the operation of river regulation infrastructure (Singh et al. 2004).

Numerous studies have examined the effect of river streamflow on water quality (Wright & Worrall 2001; Mosley et al. 2012), most of which were based on a single variable (often using regression analysis) or qualitative assessment of multiple variables (usually using principle component analysis, cluster analysis, or factor analysis). However, much less attention has been paid to quantitative assessment of the concomitant changes in multiple water quality variables in association with discharge variation. The contribution of this paper is to present a quantitative evaluation of the influence of streamflow on the joint occurrence of multiple interdependent water quality variables. We approached the problem by using copulas to construct the joint distribution function of water quality variables under different flow scenarios. The copula function is capable of interpreting the structural relationships among related variables. It can be used to construct the joint distribution with random marginal distribution, without assuming that the variables are independent, normally distributed or have the same marginal distribution (Nelsen 2006).

The objective of this paper is to assess and quantify the joint occurrence of key water quality indexes under different streamflow scenarios. This problem, and the methods and results of the investigation described here, are relevant to any river where water quality management is a recognized issue. We illustrate the approach by considering two locally important water quality indexes, concentrations of COD_{Mn} (permanganate index) and NH_3-N (ammonia nitrogen), at three hydrological stations in the middle Huai River Basin (HRB), China. Over the past 40 years, the HRB, like many other rivers in China, has experienced marked declines in water quality due to social and economic development. It was reported that China has experienced roughly 1,000 water pollution events every year between 2001 and 2004, which seriously restricts the availability of water for industrial, agricultural and domestic usage. According to the Chinese Academy of Engineering, Ministry of Environmental Protection of PRC (2011), the economic losses from water pollution were calculated to be more than 280 billion Chinese yuan. Moreover, the water pollution also has significant impacts on human health, especially in rural areas, where about 300 million people lack access to clean drinking water. The excess cancer mortality attributable to water pollution in rural China in 2003 was 6.2 deaths/10,000 population (World Bank 2007). Hence, a series of effective measures including joint regulation of water quality and quantity were conducted by the Chinese government. In this situation, a scientific understanding of how river streamflow will affect joint water quality indexes will provide a foundation for rational scheduling of water resources and policy making.

River streamflow has been regulated basin-wide since the late 1980s to meet irrigation demand (Xia et al. 2011). Climate, water resource development, land use, land cover and human activities all affect water quality. Although these factors are in a state of continuous change, in the HRB, the major historical changes in the driving forces that have shaped the current relationships between water quality and discharge can be considered to have stabilized. This assumption allowed for the focus of this paper to be on the influence of streamflow on water quality. More distinct evidence supporting this assumption is included in the study area and data section.

In this study, we first provide a brief description of copulas. Then, we describe the application of copulas to the case of COD_{Mn} and NH_3-N concentrations at three middle HRB hydrological stations. Construction of seven copulas from the Archimedean family, meta-elliptical family and the Plackett family (Nelsen 2006) is explained, and model performance assessed. Finally, the impact of streamflow on the joint occurrence of water quality indexes is discussed.

METHODOLOGY

Copula distribution function

Previous studies indicated that using the copula function, various marginal distribution structures can be selected according to the situation to establish flexible multivariate distributions and to describe non-linear and asymmetric relationships among various variables. A more detailed treatment of copulas and their properties can be found in Joe (1997) and Nelsen (2006).

Copulas can be classified into several families (Nelsen 2006), including the Archimedean copulas, meta-elliptical copulas, quadratic copulas with cubic form, extreme value copulas and other copulas. In this study, only the G-H, Clayton, Frank, and AMH copula from the Archimedean family, normal and Student's t copula from the meta-elliptical family and the Plackett copula are employed to model joint dependence of water quality variables. The expressions for the cumulative distribution function (CDF) and associated parameter space of bivariate copulas are presented in Table 1, where Φ represents the CDF of the standard normal distribution, u and v represent the two marginal CDFs, θ and ρ are the copula parameters, and t_v is the standard t distribution with v degrees of freedom.

Modelling joint occurrence of water quality variables under different streamflow scenarios

The joint probability between the variables of water quality indexes, i.e., COD_{Mn} and NH_3-N, can be established according to the criteria of water quality standards based on the Chinese standard grades (China State EPA 2002; Yuan et al. 2015). Let x_0 and y_0 represent the mass concentration (mg/L) of water quality COD_{Mn} and NH_3-N, respectively; then, the following three combinations of joint probability of water quality indexes are considered:

Case 1: Both water quality indexes are less than a given threshold, i.e., $P1 = P(X \leq x_0 \wedge Y \leq y_0)$

Case 2: At least one water quality index exceeds a given threshold, i.e., $P2 = P(X \geq x_0 \vee Y \geq y_0)$

Case 3: Both water quality indexes exceed a given threshold, i.e., $P3 = P(X \geq x_0 \wedge Y \geq y_0)$

Marginal distributions of water quality variables

Water quality variables can be modeled using both two- and three-parameter distributions such as normal, log-normal (Chin 2012) distributions. As the normal density function fails in representing the skewed densities that have been observed in many water quality datasets due to their symmetrical and bell-shaped characteristics, in this study, five widely used statistical models with three parameters in hydrological research, i.e., Pearson type III (P-III), Log-normal (LN3), generalized Pareto distribution (GPA), generalized logistic distribution (GLO) and generalized extreme value distribution (GEV), were selected as the candidate distributions to model the marginal distribution of the water quality variables in order to model the skewed pattern of the water quality variables. The CDF of these models can be found in Hosking (1990).

Table 1 | Expression of CDF, associated parameters and parameter range of applied copulas

Copulas	C(u, v)	θ/ρ
G-H	$\exp\{-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}\}$	$\theta \geq 0$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\theta > 0$
Frank	$-\frac{1}{\theta}\ln[1 + (e^{-\theta} - 1)^{-1}(e^{-\theta u} - 1)(e^{-\theta v} - 1)]$	$\theta \neq 0$
AMH	$\frac{uv}{1 - \theta(1-u)(1-v)}$	$-1 < \theta < 1$
Normal	$\int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) dxdy$	$-1 < \rho < 1$
Student's t	$\int_{-\infty}^{t_v^{-1}(u)}\int_{-\infty}^{t_v^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \left[1 + \frac{x^2 - 2\rho xy + y^2}{v(1-\rho^2)}\right]^{-(v+2)/2} dxdy$	$-1 < \rho < 1$
Plackett	$-\frac{1}{2(\theta-1)}\left[s - \sqrt{s^2 - 4uv\theta(\theta-1)}\right]$	$\theta \geq 0$

Note: $s = 1 + (\theta - 1)(u + v)$.

CASE STUDY AREA AND DATA

The HRB, located in eastern China between the Yangtze River Basin to the south and the Yellow River Basin to the north (Figure 1), contains more than 5,700 reservoirs and 5,000 dams (also called weirs, sluices or floodgates) to control river streamflow. The total storage capacity of these water projects is 303 billion cubic metres, accounting for 51% of the annual runoff of the whole basin. The construction of these water projects reached its peak during 1970 to 1980 (Zhang *et al.* 2010). During the dry season, the numerous sluices upstream of Bengbu pond water of generally poor quality. A significant rainfall event requires opening of the sluice gates for flood management, which can result in the sudden release of polluted water to the river downstream, potentially causing social, economic and environmental harm. A number of serious incidents of this type have occurred since the first one was recorded in 1994, and the management of such incidents remains a high priority issue (Bai & Shi 2006; Zhang *et al.* 2010). Thus, understanding the impact of different scenarios of streamflow on water quality events is of central importance to the overall management of joint water quantity and quality in the HRB (Zhang *et al.* 2010).

In this study, monthly streamflow data and the corresponding two key water quality indexes, the concentrations of COD_{Mn} and NH_3-N, for the period of 1985–2005, were collected at three monitoring stations, Bengbu and Lutaizi on the main stream, and Jieshou on the Shaying River tributary (Figure 1). Data after 1985 were selected (after the peak construction of water projects in the river basin), a time during which there was a relatively stable relationship between the discharge and water quality variables. Only months with observations of river streamflow and water quality COD_{Mn} and NH_3-N concentrations were included. For this river system, the target for water quality is Grade III, according to standard GB 3838-2002 Environmental Quality Standards for Surface Water (China State EPA 2002). We selected the above two water quality indexes because: (1) the increasing chemical oxygen demand (COD)

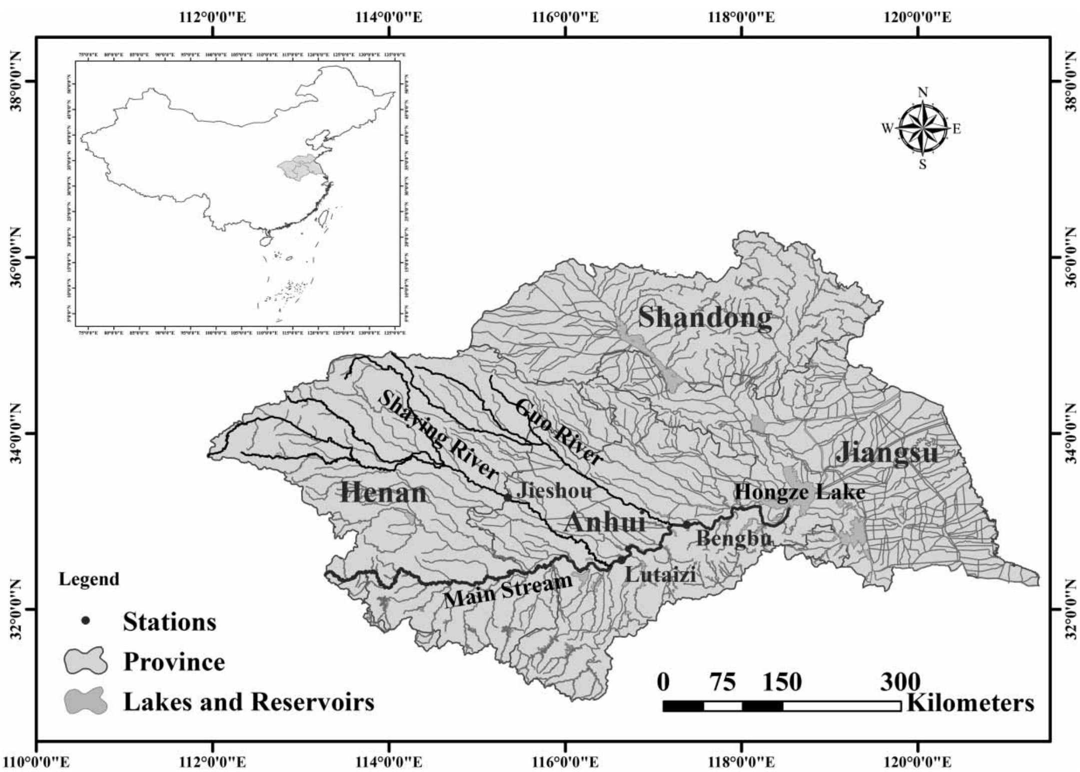

Figure 1 | Location of the HRB.

and NH_3-N discharges from industrial pollutants, agricultural lands and household sewage plants have exceeded the assimilation capacity of the aquatic environment and have become the greatest priority in China due to the negative effects of these oxygen-consuming substances on aquatic organisms (He *et al.* 2015; Yuan *et al.* 2015; Meng *et al.* 2016); (2) further, these indexes are also treated as the two and only major water pollutants in watershed management of the total pollutant emission control by the Action Plan for Prevention and Control of Water Pollution in China (Ministry of Environmental Protection of PRC 2015), and thus are more widely and frequently monitored in most rivers in China. Note that the collection of long-term water quality and hydrological observations at the same monitoring stations has been, and still is, a major challenge, particularly in developing countries like China, which has few environmental data sets that can be effectively shared (He *et al.* 2015). Therefore, owing to water quality accessibility, and more importantly, to closely meet the actual demands of current watershed management in the HRB, many researchers have focused on these two water quality indexes (Zhang *et al.* 2010; Zhai *et al.* 2014; Meng *et al.* 2016).

RESULTS AND DISCUSSION

Classification of streamflow scenarios

The selected water quality variables COD_{Mn} and NH_3-N were both significantly negatively correlated with streamflow and positively correlated with each other (Table 2), which means that it is possible to quantify the impact of streamflow on the joint occurrences of these water quality variables. Each river monthly streamflow series was divided into three classes based on the 37.5% and 62.5% quantiles (Liu & Zheng 2002), which were termed scenarios. Streamflow observations lower than the 37.5% quantile were named scenario 1 (S1), and observations greater than the 62.5% quantile were named scenario 3 (S3), while those between the 37.5% and 62.5% quantiles were named scenario 2 (S2). This allowed establishment of bivariate distributions of the corresponding observations of COD_{Mn} and NH_3-N concentrations. Classification of the streamflow data resulted in three scenarios that were markedly different in terms of descriptive statistics (Table 3). For each streamflow scenario, the corresponding water quality observations were selected, which allowed quantitative assessment of the influence of river streamflow (as defined by low, medium and high streamflow classes) on the joint occurrence of the water quality indexes.

Performance of the marginal distributions

We first calculated the skewness and kurtosis of the water quality variables pertaining to the three scenarios for all stations; the results showed that the skewness of both water quality variables ranged from 0.259 to 1.442, while the kurtosis ranged between 0.586 and 3.621, which indicated that the selected water quality datasets were highly skewed. Furthermore, the Shapiro–Wilk test, a tool for demonstrating the normality of the dataset, showed that both of the water quality variables for all scenarios for the three stations violated the assumption of normality. Therefore, we used five three-parameter statistical models, introduced in the methodology section, to model the marginal distribution in order to model the skewed pattern of the water quality variables.

Parameters of the marginal distributions were estimated using the L-moment method, and the performance of the probability functions was tested using the Kolmogorov–Smirnov

Table 2 | Correlation between streamflow and water quality indexes for the three stations

	Streamflow and COD_{Mn}		Streamflow and NH_3-N		COD_{Mn} and NH_3-N	
Station	Spearman's ρ	Kendall's τ	Spearman's ρ	Kendall's τ	Spearman's ρ	Kendall's τ
Jieshou	−0.407	−0.285	−0.471	−0.335	0.654	0.483
Lutaizi	−0.262	−0.180	−0.361	−0.243	0.611	0.438
Bengbu	−0.302	−0.213	−0.428	−0.304	0.493	0.351

All values of ρ and τ are highly significant, with $p < 0.001$.

Table 3 | Characteristics of the three subsets (scenarios) of the monthly streamflow time series of the three stations

Station	Scenario	Minimum, (m^3/s)	Mean, (m^3/s)	Maximum, (m^3/s)	Standard deviation	Skewness
Jieshou	S1	0.04	3.51	10.90	0.905	0.546
	S2	11.19	21.39	41.41	0.348	0.819
	S3	42.61	175.30	1,380.00	1.091	3.588
Lutaizi	S1	49.44	109.30	157.40	0.264	−0.243
	S2	161.10	243.10	367.40	0.238	0.367
	S3	401.60	1,274.0	5,749.00	0.857	1.895
Bengbu	S1	11.11	92.39	163.00	0.475	−0.247
	S2	167.00	274.84	478.01	0.340	0.715
	S3	538.00	1,685.47	6,850.32	0.772	1.649

(K-S) statistic under the 95% significance level and root-mean-square error (RMSE) criterion. For the low streamflow scenario S1 at Jieshou (Table 4), the lower values of D_n compared with $D_{n,\,0.95}$ indicated that all of the candidate distributions could be used to simulate the observed water quality series. However, the RMSE statistics suggested that the P-III and GPA were the best-fit distribution for COD_{Mn} and NH_3-N, respectively, given that the lower the RMSE, the better the modelling result (Table 4 and Figure 2). On the other hand, for the other two scenarios at Jieshou, the KS statistic also indicated that all distributions were suitable for the water quality series, while GEV and P-III were determined as the optimal distributions for COD_{Mn} and NH_3-N, respectively, in both cases (Table 4).

Joint dependence of water quality variables using copulas

Seven copulas were employed to identify and describe the joint distribution for the three scenarios for all stations. The parameter of each copula was estimated using Kendall's

Table 4 | Performance of the probability distributions in modelling marginal water quality indexes for the three river streamflow scenarios of the three stations

				Parameter			K-S test		
Station	Scenario	Water quality index	Distribution	Shape	Scale	Location	D_n	$D_{n,0.95}$	RMSE
Jieshou	S1	COD_{Mn}	P-III	0.817	0.0531	2.374	0.105	0.189	0.038
		NH_3-N	GPA	22.345	0.192	−1.106	0.072		0.025
	S2	COD_{Mn}	GEV	5.287	−0.367	8.162	0.066	0.150	0.027
		NH_3-N	P-III	0.888	0.068	−0.786	0.079		0.032
	S3	COD_{Mn}	GEV	2.225	−0.255	5.377	0.052	0.148	0.015
		NH_3-N	P-III	0.672	0.144	−0.077	0.087		0.039
Lutaizi	S1	COD_{Mn}	GEV	4.630	−0.535	1.407	0.086	0.177	0.033
		NH_3-N	GPA	1.158	−0.305	0.034	0.076		0.024
	S2	COD_{Mn}	GLO	5.113	−0.205	1.093	0.064	0.173	0.023
		NH_3-N	GPA	1.656	0.159	−0.087	0.062		0.022
	S3	COD_{Mn}	GLO	0.760	−0.227	4.530	0.079	0.156	0.028
		NH_3-N	P-III	0.576	0.888	0.018	0.069		0.029
Bengbu	S1	COD_{Mn}	LN3	1.408	0.536	1.601	0.045	0.164	0.019
		NH_3-N	P-III	1.049	0.399	−0.112	0.065		0.022
	S2	COD_{Mn}	GLO	0.708	−0.256	4.842	0.058	0.161	0.019
		NH_3-N	GPA	1.606	0.003	−0.162	0.11		0.042
	S3	COD_{Mn}	GLO	0.612	−0.166	4.432	0.067	0.150	0.026
		NH_3-N	LN3	−1.034	1.169	−0.054	0.10		0.039

Only the best-fit distributions for each scenario are presented.

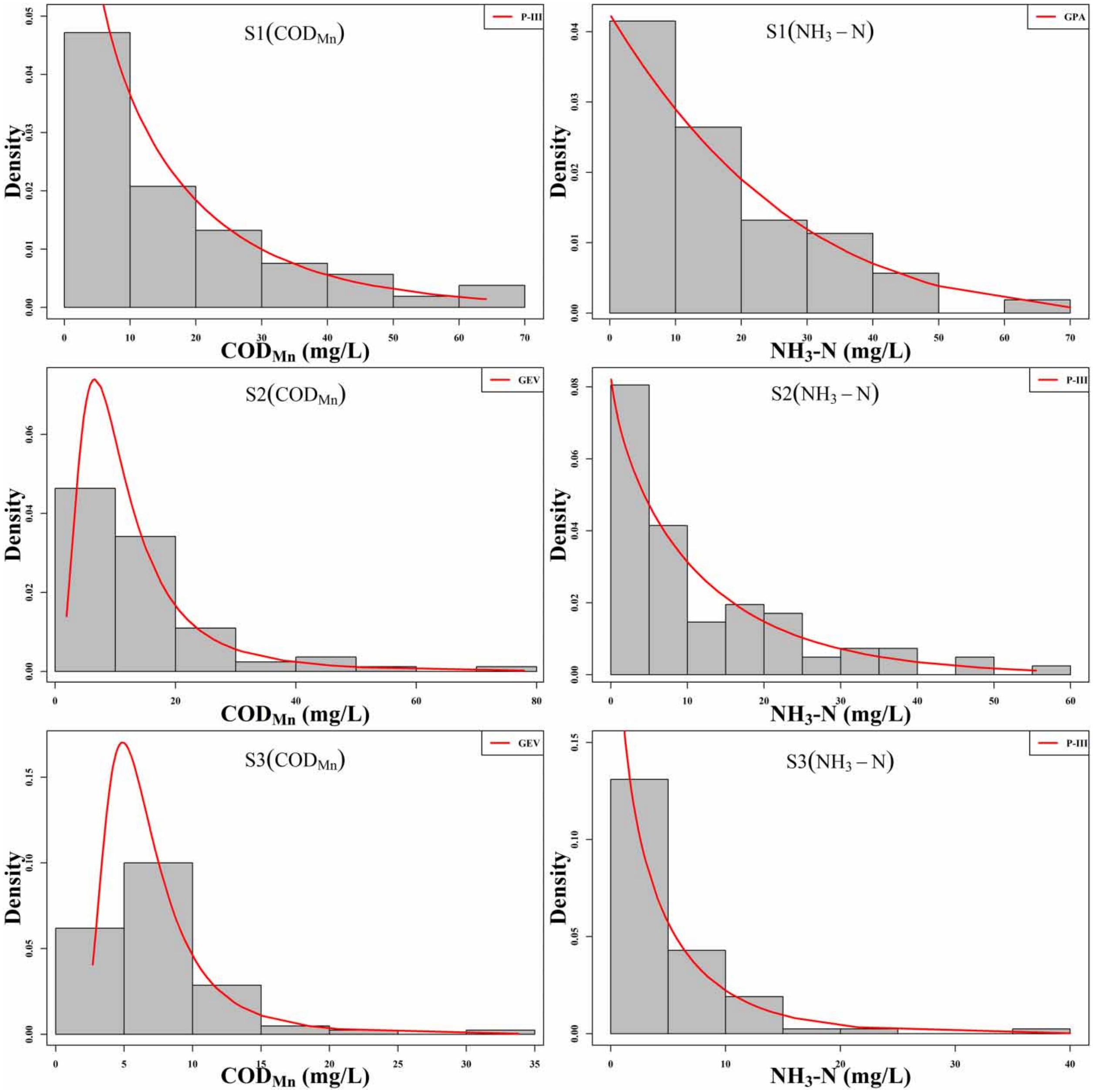

Figure 2 | Best-fit marginal distributions of probability density functions for the two water quality indexes for the three river streamflow scenarios at Jieshou station.

correlation method or the pseudo-likelihood method (Table 5). For scenario S1 at Lutaizi, all three goodness-of-fit tests suggested that the G-H copula was better than the other copulas (Table 5). For scenario S2, although the bias values of the Frank copula were slightly smaller than those of the G-H copula, the RMSE and AIC statistics of the G-H copulas were much smaller those of the Frank copula; therefore, the G-H copula was deemed the best-fit copula for scenario S2, and for the same reason, the Frank copula was selected as the best-fit copula for scenario S3 at Lutaizi station. The same procedure was followed to determine the best-fit copulas for the scenarios of the other two stations (Table 5). Overall, the G-H and Frank copulas performed best in all cases. Q-Q plots of joint distributions were used to visualize the degree of fit for the chosen best-fit copula functions (e.g. Figure 3).

Table 5 | Estimated parameters of various copulas and goodness-of-fit test results; only the best-fit copulas for each scenario are presented

			Parameter(s)		Goodness-of-fit test		
Station	**Scenario**	**Copulas**	**Kendall's τ**	**$\theta/\rho(v)$**	**RMSE**	**AIC**	**Bias**
Jieshou	S1	Frank	0.450	4.900	0.0479	−321.061	14.021
	S2	Frank	0.449	4.874	0.0283	−583.009	15.296
	S3	Frank	0.428	4.565	0.0292	−592.494	23.405
Lutaizi	S1	G-H	0.489	1.955	0.0227	−445.455	9.282
	S2	G-H	0.479	1.918	0.0224	−470.192	13.116
	S3	Frank	0.318	3.122	0.0229	−572.941	12.316
Bengbu	S1	G-H	0.506	2.025	0.0195	−542.582	6.024
	S2	G-H	0.279	1.386	0.0231	−533.998	24.634
	S3	Frank	0.114	1.032	0.0340	−553.433	37.981

Joint occurrence of water quality variables in relation to river streamflow

The joint probability of water quality was visualized using isoline plots. For example, for Bengbu station (Figure 4), considering the joint probability combination P1 that the COD_{Mn} concentration is less than 6.0 mg/L and the concentration of NH_3-N is less than 1.0 mg/L, meeting the Chinese Grade III standard, is approximately 30% for the lowest streamflow scenario S1, while for S2 it is approximately 45%, and for the highest streamflow scenario S3 it is approximately 75%. Consideration of the joint probabilities of all three water quality combinations among the three streamflow scenarios at Bengbu in relation to the

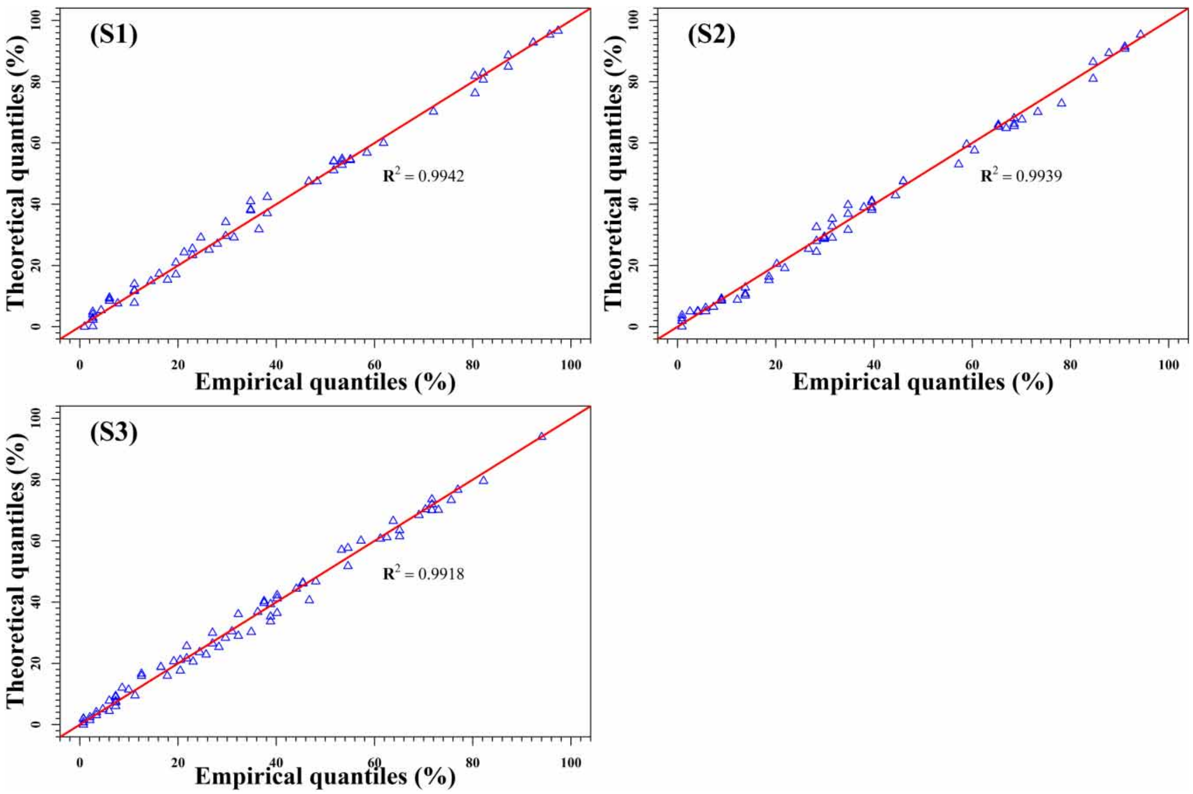

Figure 3 | Q-Q plots for Lutaizi station using G-H copulas under the three river streamflow scenarios.

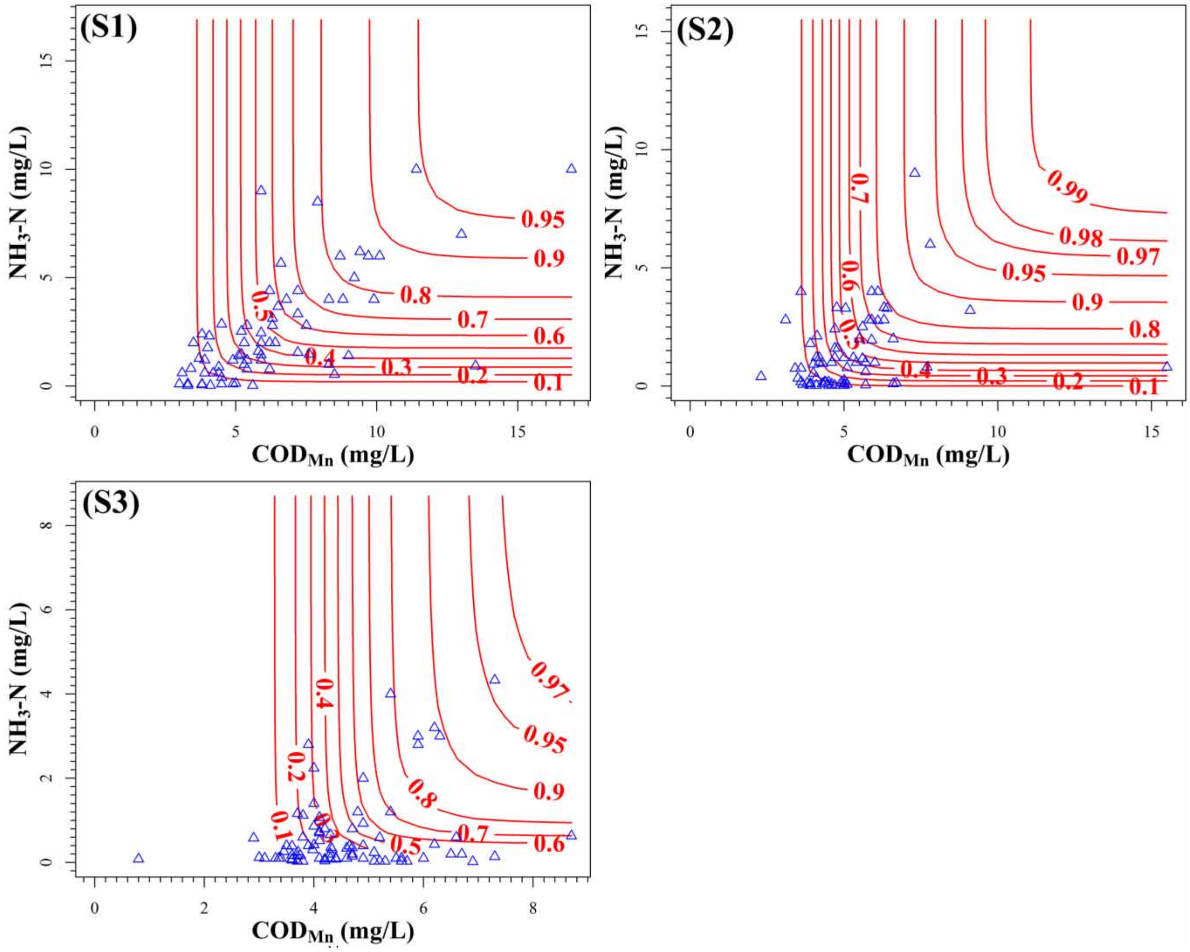

Figure 4 | Isolines of joint probability P1 that both water quality indexes will be less than a given threshold among the three streamflow scenarios for Bengbu station. Historical water quality observations for each scenario are also plotted.

Chinese standards for water quality Grades II to V (Grade I was not met at this station) provided a comprehensive summary of the likelihood of meeting these standards in relation to streamflow (Figure 5). Combination P1 showed increasing probability of worsening water quality grade because this combination involved both water quality indexes being less than the relevant threshold. In contrast, combinations P2 and P3 showed decreasing probability of worsening water quality grade because they involved one or both of the water quality indexes exceeding the relevant threshold. The joint probabilities were derived for P2 and P3 combinations among the three scenarios for all stations (see Figure 6 for the plot of the S2 and S3 results). Note that the sum of the probability of P1 and P2 is equal to 1, which is in accordance with the complementary relationship between these two water quality combinations, so the results for P1 were omitted.

Discussion

For Bengbu station, water quality combination P1 indicated that water quality Grade II would be achieved for both water quality indexes in less than 25% of months (Figure 5). Lowering the expectation to Grade III or worse increased the probability that the standard would be met, although the probability was strongly conditioned by river streamflow. Streamflow scenario S1 was always associated with the lowest probability, and streamflow scenario S3 was always associated with the highest probability. The patterns of joint probabilities of water quality combinations P2 and P3 were

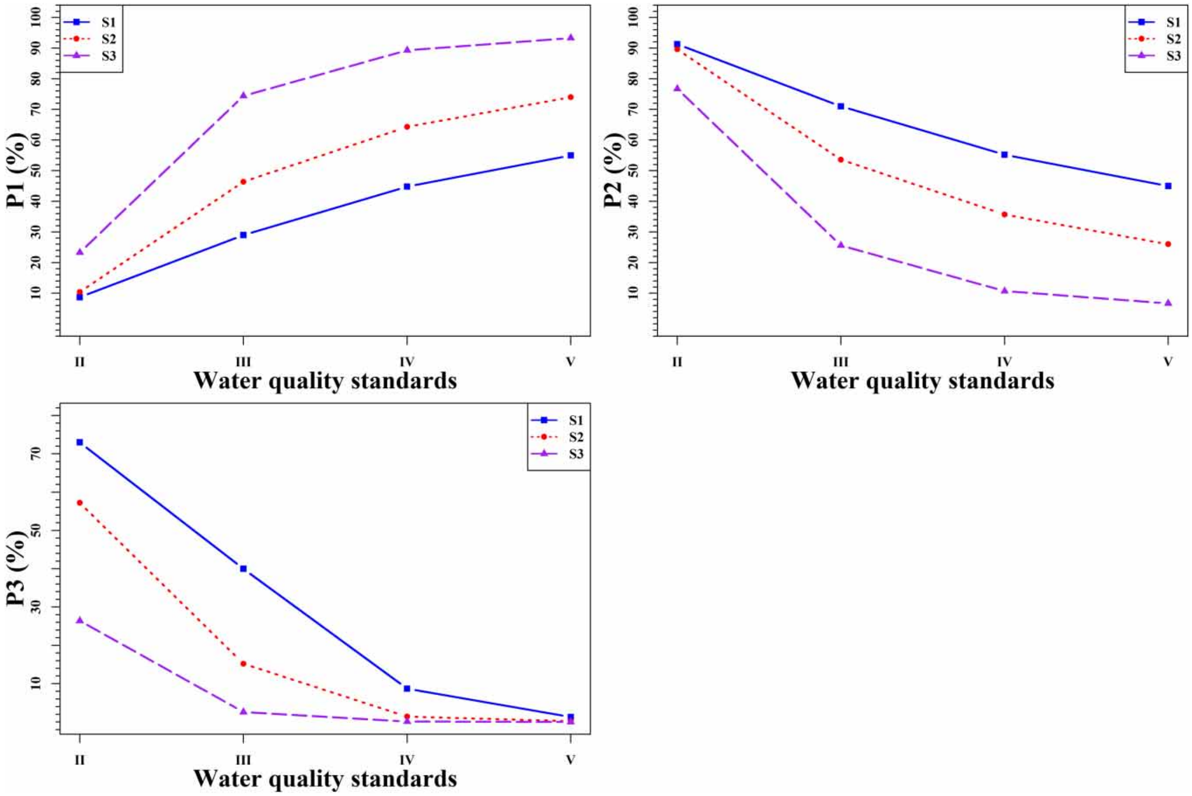

Figure 5 | Joint probabilities of the three water quality index combinations for the three river streamflow scenarios at Bengbu station.

the same under the three streamflow scenarios, with probabilities for S3 (high streamflow scenario) less than those for S2 (medium streamflow scenario), and those for S2 less than those for S1 (low streamflow scenario) (Figure 5). Streamflow scenario S2 had a much lower probability than S1 for combinations P2 and P3 under the water quality Grade III, while the higher streamflow scenario S3 significantly reduced the risk of jointly exceeding the water quality standards for combination P2 of water quality Grade III and combination P3 of water quality Grade II (Figure 5).

The current water quality management target for the HRB is Grade III, both in the main stream and in the tributary Shaying River. Taking Bengbu as a case and assuming this standard, the probability of P2, when at least one of the water quality indexes exceeds Grade III, was approximately 70% for the low streamflow scenario S1. For the medium streamflow scenario S2, the joint probability decreased to 54%, but this could still be considered an unacceptably high risk. For the high streamflow scenario S3, the joint probability was 26%, which suggests that the exceeding of water quality standards for at least one index was a problem irrespective of the flow conditions. The probabilities for P3, when both water quality indexes exceed the standard, were 40%, 15% and 3% for the streamflow scenarios S1, S2 and S3, respectively. These probabilities were lower than those for P2, but it is normal practice in China to consider the standard not met if only one index fails to meet the standard. This is based on the assumption that exceeding the standard by only one contaminant is sufficient to render water quality impaired to the standard met by that contaminant.

A comparison of water quality combinations of P2 and P3 among the three stations is displayed in Figure 6. Among the stations, Jieshou was consistently at the highest risk of poor water quality conditions under any combination of water quality index and streamflow scenario. This is explained by the generally low streamflow in Jieshou (Table 3) relative to the quantity of pollution directed to the river.

Under the streamflow scenario S1, the probability of combinations P2 and P3 was much higher at Bengbu compared with Lutaizi, indicating that Bengbu was exposed to a

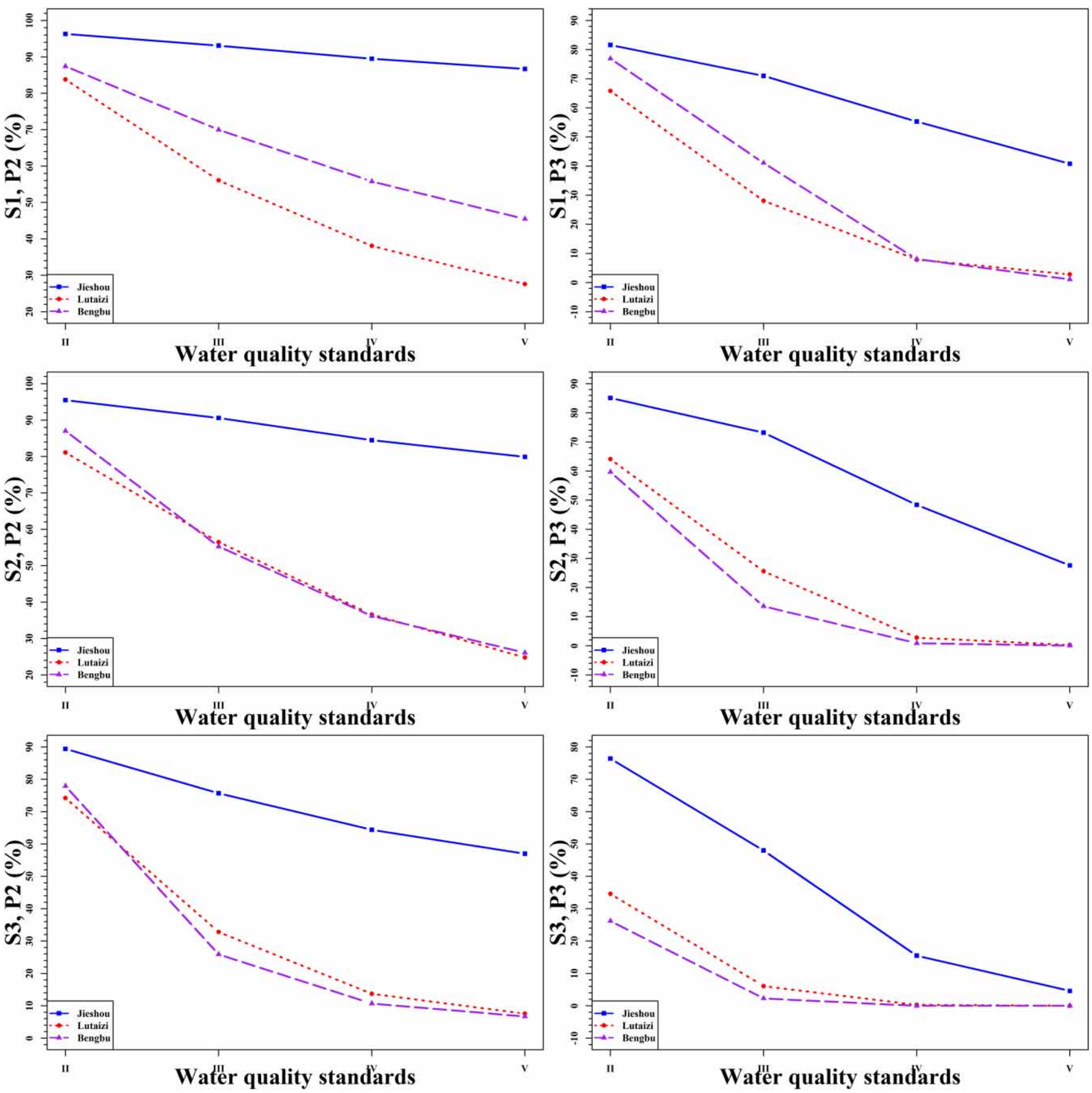

Figure 6 | Comparison of the joint probabilities of water quality comparisons P2 and P3 under the three river streamflow scenarios for all stations.

greater risk of water pollution during times of low streamflow. However, with the higher streamflow scenarios S2 and S3, the risk at Bengbu was lower than that at Lutaizi. This is likely explained by the relatively constant contribution of industrial, domestic, and agriculture wastewater to the HRB from three large urban areas, Fengtai county, Huainan city and Huaiyuan county, located between Lutaizi and Bengbu. During times of low streamflow, there is little dilution and assimilation of pollutants. The mean streamflow at Bengbu under scenario S1 was slightly lower than at Lutaizi, while streamflow scenarios S2 and S3 indicated higher flows at Bengbu than at Lutaizi. This suggests that under high streamflow conditions, the inflow of relatively higher quality storm runoff water between Lutaizi and Bengbu provided the opportunity for dilution and assimilation processes to improve water quality in the HRB main stream.

A more detailed illustration to investigate the impact of river streamflow on the water quality combined probability is demonstrated in Figure 7, where the joint probability of water quality combinations within each water quality grade is displayed. It can be concluded that high streamflow scenarios do help to improve the joint probability of water quality combinations within Grade II and III, while decreasing those within Grade V or inferior V.

CONCLUSIONS

In this study, a quantitative evaluation of the influence of river streamflow on the joint occurrence of multiple interdependent water quality variables was presented. Seven copulas were tested for their suitability to establish a joint distribution between water quality indexes under three streamflow scenarios (representing low, medium and high streamflow conditions) at three monitoring stations in the heavily polluted HRB. The joint occurrence of two key water quality indexes, COD_{Mn} and NH_3-N concentrations, was examined in relation to streamflow. Using best-fit copula models, the joint probabilities of different combinations of the water quality indexes for each streamflow scenario were established. The main conclusions from this study are as follows:

(1) At all monitoring stations, there was a significant negative correlation between streamflow and COD_{Mn} and NH_3-N concentrations. This correlation provided the possibility of using copulas to characterize the joint occurrence of water quality indexes in relation to river streamflow. It was found that in all cases, G-H and Frank copulas performed well in modelling the joint distributions of water quality indexes.

(2) In both the main stream of the HRB (Lutaizi and Bengbu) and the tributary Shaying River (Jieshou), the low streamflow scenario had very high joint probabilities of exceeding the water quality standards. For the higher streamflow scenario, the corresponding probabilities decreased, which is consistent with the negative correlation between streamflow and water quality indexes.

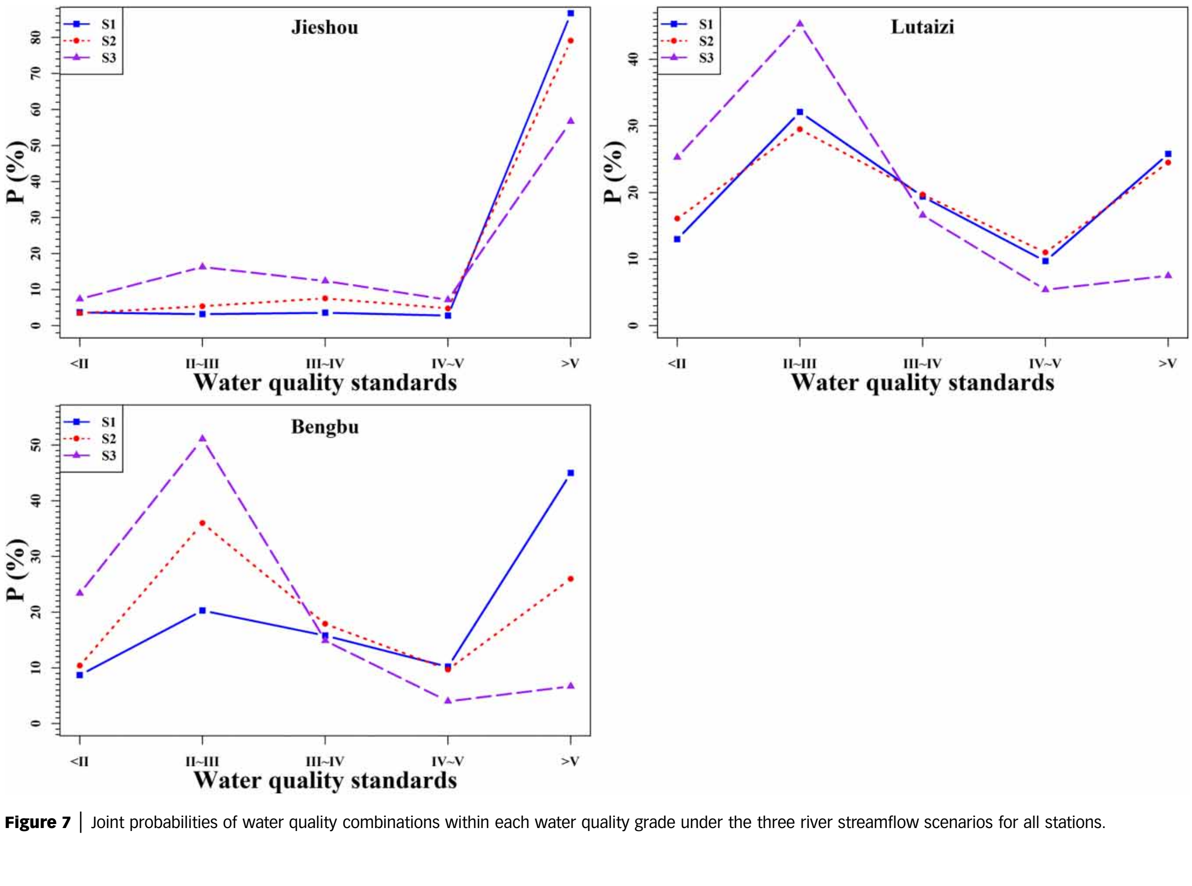

Figure 7 | Joint probabilities of water quality combinations within each water quality grade under the three river streamflow scenarios for all stations.

(3) Under the current water quality management targets for the HRB, for low streamflow and medium streamflow conditions, the probability of P2 at Lutaizi and Bengbu exceeded 50%, and this probability exceeded 90% at Jieshou. This result quantified the considerable risk of exceeding the pollution standards under low and medium streamflow conditions. Under the high streamflow condition, the probability P2 was lower but still at a higher than desirable level. At Jieshou, this probability exceeded 75%, and at Lutaizi and Bengbu, it was approximately 30%.
(4) Among all stations, both in the main stream and the tributary, high streamflow scenarios helped to improve the joint probability of water quality combinations within higher water quality grades (most often Grade II and III) but decreased that within lower grades (most often Grade V or inferior V).

This paper analyzed the impact of river streamflow on the joint pattern of water quality variables in the HRB using copulas. All of the results can be used as a scientific basis of sustainable water resource management and water pollution improvement in the HRB and provide a foundation for water quantity–quality joint operation. However, it is well known that different copula functions do have different abilities in capturing the tail dependence structure (Nelsen 2006). For example, the Clayton copula is more suitable for modelling the dependence between series with lower-tail dependence, while the G-H copula has a stronger ability in characterizing the dependence between series with upper-tail dependence. More detailed consideration of the impact of different types of copula functions in terms of different tail dependence structures between the hydrological time series should be included in subsequent studies. However, due to the limited water quality observations and to closely meet the actual demands of current watershed management in the HRB, this paper only selected two major water quality indexes. In future, more representative water quality indexes, such as total phosphorus and total nitrogen, should be taken into consideration. In addition, the uncertainty analyses should also be strengthened in future studies.

ACKNOWLEDGEMENTS

This present study is financially supported by the National Natural Science Foundation of China (51279143, 51369011), the Major Science and Technology Program for Water Pollution Control and Treatment (2014ZX07204-006), the National Key Research and Development Program (during the 13th Five-Year Plan), Ministry of Science and Technology, PRC (Grant No. 2016YFC0401301), the Science and Technology Projects of the Jiangxi Provincial Department of Water Resources (KT201302) and the Open Fund of Poyang Lake Water Resources, Water Ecology and Environment Research Center, MWR (KFJJ201404).

REFERENCES

Bai, X. & Shi, P. 2006 Pollution control in China's Huai River basin, what lessons for sustainability? *Environment: Science and Policy for Sustainable Development* **48** (7), 22–38.
Chin, D. A. 2012 *Water-Quality Engineering in Natural Systems: Fate and Transport Processes in the Water Environment*, 2nd edn. John Wiley & Sons, Hoboken, NJ, USA.
China State EPA 2002 *China's National Environmental Quality Standards for Surface Water (GB 3838-2002)*. Environmental Science Press, Beijing, China.
He, T., Lu, Y., Cui, Y., Luo, Y., Wang, M., Meng, W., Zhang, K. & Zhao, F. 2015 Detecting gradual and abrupt changes in water quality time series in response to regional payment programs for watershed services in an agricultural area. *Journal of Hydrology* **525**, 457–471.
Hesse, C. & Krysanova, V. 2016 Modeling climate and management change impacts on water quality and in-stream processes in the Elbe River Basin. *Water* **8** (2), article 40.
Hosking, J. R. M. 1990 L-moment: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society: Series B (Methodological)* **52** (1), 105–124.
Joe, H. 1997 *Multivariate Models and Dependence Concept*. Chapman & Hall, London, UK.
Liu, C. M. & Zheng, H. X. 2002 South-to-north water transfer schemes for China. *International Journal of Water Resources Development* **18** (3), 453–471.
Liu, J., Zheng, X., Xia, J., Wu, S., She, D. & Zou, L. 2016 Characterizing and explaining spatio-temporal variation of water quality in a highly disturbed river by multi-statistical techniques. *SpringerPlus* **5** (1), article 1171.
Meng, Y., Zhang, X., She, D., Wang, J. & Wu, S. 2016 The spatial and temporal variation of water use efficiency in the Huai River Basin using a comprehensive indicator. *Water Science & Technology: Water Supply* **17** (1), 229–237.

Ministry of Environmental Protection of PRC 2015 *Action Plan for Prevention and Control of Water Pollution in China*. People's Publishing House, Beijing, China.

Mosley, L. M., Zammit, B., Leyden, E., Heneker, T. M., Hipsey, M. R., Skinner, D. & Aldridge, K. T. 2012 The impact of extreme low flows on the water quality of the Lower Murray River and Lakes (South Australia). *Water Resources Management* **26** (13), 3923–3946.

Nelsen, R. B. 2006 *An Introduction to Copulas*, 2nd edn. Springer, New York, USA.

Shi, W., Xia, J. & Zhang, X. 2016 Influences of anthropogenic activities and topography on water quality in the highly regulated Huai River Basin, China. *Environmental Science & Pollution Research International* **23** (21), 21460–21474.

Shi, Y., Xu, G., Wang, Y., Engel, B. A., Peng, H., Zhang, W., Cheng, M. & Dai, M. 2017 Modelling hydrology and water quality processes in the Pengxi River basin of the Three Gorges Reservoir using the soil and water assessment tool. *Agricultural Water Management* **182**, 24–38.

Shin, J. Y., Artigas, F., Hobble, C. & Lee, Y.-S. 2013 Assessment of anthropogenic influences on surface water quality in urban estuary, northern New Jersey: multivariate approach. *Environmental Monitoring & Assessment* **185**, 2777–2794.

Singh, K. P., Malik, A., Mohan, D. & Sinha, S. 2004 Multivariate statistical techniques for the evaluation of spatial and temporal variation in water quality of Gomti River (India) – a case study. *Water Research* **38** (18), 3980–3992.

World Bank 2007 *Cost of Pollution in China: Economic Estimates of Physical Damages*. World Bank, Washington, DC, USA.

Wright, J. & Worrall, F. 2001 The effects of river flow on water quality in estuarine impoundments. *Physics & Chemistry of the Earth: Part B (Hydrology, Oceans & Atmosphere)* **26** (9), 741–746.

Xia, J., Zhang, Y.-Y., Zhan, C. S. & Ye, A. Z. 2011 Water quality management in China: the case of the Huai River Basin. *International Journal of Water Resources Development* **27** (1), 167–180.

Yuan, D., Zhang, Y., Liu, J. & Gao, S. 2015 Water quantity and quality joint-operation modeling of dams and floodgates in Huai River Basin, China. *Journal of Water Resources Planning & Management* **141** (9), 04015005.

Zhai, X. Y., Xia, J. & Zhang, Y. Y. 2014 Water quality variation in the highly disturbed Huai River Basin, China from 1994 to 2005 by multi-statistical analyses. *Science of the Total Environment* **496** (1), 594–606.

Zhang, Y., Xia, J., Liang, T. & Shao, Q. 2010 Impact of water projects on river flow regimes and water quality in Huai River Basin. *Water Resource Management* **24**, 889–908.

First received 29 September 2016; accepted in revised form 9 March 2017. Available online 13 April 2017