Combined risk prediction in the water environment based on an MS–AR model and Copula theory
Niu Jun-yi, Huang Hu and Chen Na

ABSTRACT
Making a quantitative prediction on the combined risk of the water body is helpful for the objective evaluation of the water environment system’s state of health, and also has important results for the water environment system’s safety management. In this paper, the Markov status switching theory (Markov Switching, MS), Monte Carlo method (Monte Carlo, MC) and Copula theory were used together, to establish a method for the water environment system’s combined risk assessment. This method firstly using MS theory established the water quality time series’ autoregression model (MS–AR); then the MS–AR model and MC method were used to carry out random simulation on the water quality time series; finally, multi-dimensional joint distribution among random simulation results were established by Copula function, and this distribution utilized to make a quantitative analysis of the water environment system’s combined risk. By means of the above combined risk analysis model, the combined risk prediction and correlation analysis of the water quality of the Guohe River bridge section were carried out. The results showed that the total phosphorus (TP) and 5-day biochemical oxygen demand (BOD5) had an important effect on the Guohe River water environment’s state of health, and there was a strong positive correlation between TP and BOD5.

Key words | combined risk prediction, Copula function, random simulation, water quality

INTRODUCTION
Quantitative assessment of the combined risk of the water environment system is helpful for the objective evaluation of the water environment system’s state of health, and has important results for the water environment system’s safety management. Aiming at the root risk existing in the water environment system and its harm to people’s health, researchers have put forward accordingly a health evaluation model (Geng et al. 2006), stochastic model (He & Zeng 2002), grey model (Zhou & Xu 2007), fuzzy model (Li et al. 2007), information entropy model (Jin & Wei 2008), etc., and a water environment risk assessment model (Jin et al. 2008) coupling with other theories, but so far a study on the combined risk assessment of water environment has not been reported. In this article, we will use Markov status switching theory (Markov Switching, MS), Monte Carlo method (Monte Carlo, MC) and Copula theory jointly to establish a method for the water environment system’s combined risk assessment, and provide theoretical support for assessing the water environment system’s state of health.

METHODOLOGY
MS–AR status switching-autoregression model
The Markov status switching model (Hamilton 1989; Krolzig 1997) was applied initially to studying the dependency data in economics, then it obtained wide application in economic and financial areas (Francq & Zakoïan 2001; Liu & Chyi 2006; Liu 2007; Cologni & Manera 2009). It is one kind of analysis method studying time series about changing structural dynamic characteristics, and has good adaptability for time series variables’ structural slowness or sudden changes. Its basic idea is through exterior data to calculate the probability of a system’s interior being an unobservable structure, so as to carry out modeling analysis on the system.

For time series \{x_t\} being of complex structure, we may assume that its evolvement process could be simulated by using \(M\) amounts of a \(p\)-order autoregressive process. Then state variable \(S_t\) can be introduced, \(S_t\) representing the...
structural pattern of a time series at time $t$. It is a one-order Markov process being of state space $J = \{1, 2, \ldots, M\}$, namely the current state is only related to the preceding state, and has no relation with other past states. Here it assumes that model’s residual obeys normal distribution, then the autoregression (AR) model based on Markov status switching can be shown as formula (1):

$$
\begin{align*}
    x_t &= c^S_t + \sum_{k=1}^{p} \varphi_k^S x_{t-k} + \sigma^S \varepsilon_t, \quad S_t \in J \\
    \varepsilon_t &\sim i.i.d. N(0, 1) \\
    \Pr (S_t = j | S_{t-1} = i) &= P_{ij}, \\
    \sum_{i=1}^{M} P_{ij} &= 1, \quad (i, j \in J)
\end{align*}
$$

In the model, the undetermined parameter set is $\Theta = \{c^S_t, \varphi_k^S, \sigma^S, P_{ij} | S_t \in J; k = 1, 2, \ldots, p; i, j \in J\}$, in which $c^S_t$ is the AR model’s intercept with the system under condition $S_t$ at time $t$, $\varphi_k^S$ is the AR coefficient under condition $S_t$, $\sigma^S$ is the model residual’s mean square deviation under condition $S_t$, and $\varepsilon_t$ ($t = 1, 2, \ldots, T$) is the zero mean white noise sequence obeying identical normal distribution. The parameter set $\Theta$ uses maximum likelihood estimation to be calculated, and its detailed calculation courses can be seen in the literature (Hamilton 1989; Niu & Feng 2010). What needs to be pointed out here is that when using the above model to carry out parameter estimation on water quality time series, it should carry out a Gaussian test firstly on the water quality time series. If it cannot satisfy Gaussian requirements, proper transformation should be taken (Liang & Dai 2005; Liu 2007; Niu & Feng 2010) to make data obey normal distribution.

**Random simulation of water quality time series based on MS–AR model**

The random simulation based on Markov status switching-AR model is the foundation of a given model’s initial state, according to the state transition matrix, and with the aid of the MC method to simulate the switching process among each status. MC simulation uses numbers from a probabilistic distribution to generate a series of parameter estimates that follow a distribution and can be analyzed statistically. General input requirements to perform this analysis are the probability density function (i.e., normal), the variance (or standard deviation) and the number of simulations to be performed. Simulations can be performed by correlating, or not, the variables involved in the model. The technique that uses random numbers from a probability density function is termed Random Monte Carlo (RMC) simulation (Carrasco & Chang 2005). If a model’s state is $i$ at time $t$, then the state process using the MC method to simulate its transition from time $t$ to time $t + 1$ is: firstly producing one pseudo-random number $-\text{Rand}$ which is between 0 and 1 and obeys a uniform distribution by random number generator; then comparing $\text{Rand}$ with $p_{ij}$. If $\text{Rand} < p_{ij}$, the model at time $t + 1$ will continue staying at state $i$ and evolving forward; if $\text{Rand} > p_{ij}$, the model will switch to another state, here calculating out the value of $(\text{Rand}-p_{ij})$, and comparing it with $\{p_{ij}, \in J, j \neq i\}$ ($J$ is the model’s state space) separately, if the value of $p_{i,k}$ approaches $(\text{Rand}-p_{ij})$ most, then at time $t + 1$ the model will switch to state $k$ and evolve forward.

**State value’s random simulation**

If a model’s state is $S_t$ at time $t$, then using a random number algorithm it will produce one pseudo-random number $-\text{Rand}$ which obeys normal distribution with mean 0 and standard square deviation $\sigma^S_t$. This is because the parameter estimation process is based on the residual obeying a normal distribution. The water environment system’s quality state variable generally has the minimum possible value, the most possible value and the maximum possible value, therefore the model’s error also has a corresponding value range. Because the model residual obeys normal distribution approximately, here regulating that $\text{Randn}$’s possible value interval is $[\text{Res} - 2\sigma^S_t, \text{Res} + 2\sigma^S_t]$ in which $\text{Res}$ is at state $S_t$. The MS–AR model residual’s mean value of sequence $\{y_t\}$ switched from time series $\{x_t\}$ through Box–Cox transformation and standardization, and $\sigma^S_t$ is residual’s mean square deviation. The $y_t$’s simulation value $z_t$ can be calculated by Equation (2) below:

$$
    z_t = c^S_t + \sum_{k=1}^{p} \varphi_k^S x_{t-k} + \text{Randn}
$$
Combined risk assessment theory of water environment system

Marginal distribution’s establishment

Using a constructed MS–AR model and random simulation method, sequence \( \{x_i\} \)’s simulation sequence \( \{z_n\}_{n=1,2, \ldots, N} \) can be obtained, in which \( N \) is the test number of random simulation. Researches show that with more test numbers, the frequency distribution of \( z \) is more close to sequence \( \{x_i\} \)’s real probability distribution (Jin et al. 2008). Theoretically, \( N \) should take the corresponding test number when simulation sequence \( \{z_n\} \)’s frequency distribution starts to converge. Practically, for water quality time series approximately obeying normal distribution, \( N \) can take the corresponding test number when the mean and median values in the simulation sequence \( \{z_n\} \) corresponding to original sequence \( \{x_i\} \) present frequency stability (Li et al. 2007; Jin et al. 2008). For any sample value \( x_m \) in sequence \( \{x_i\} \), if the number of values in the simulation sequence \( \{Z_n\} \) smaller than \( x_m \) is \( N_m \), then we can consider that the corresponding theoretical cumulative probability of \( x_m \) is:

\[
P(x_m) = \Pr (x_t \leq x_m) = N_m / N \tag{3}
\]

In Equation (3), \( x_t \) can be the water quality cross-section’s certain monitoring index, and also can be the comprehensive assessment index normalized from many monitoring indices through the water quality comprehensive assessment model. \( Pr \) is the probability function of event \( x_t \leq x_m \). After calculating out all the sample value probabilities in sequence \( \{x_i\} \), we can obtain sequence \( \{x_i\} \)’s probability distribution curve.

For certain pollutant’s state variable \( (x_t) \), if the given water environment quality standard’s control limit value is \( x_c \), then this pollutant’s over standard risk can be defined as:

\[
R_c = \Pr (x_t \geq x_c) \tag{4}
\]

In formula (4), \( Pr \) is the probability function of event \( x_t \geq x_c \).

Joint distribution’s establishment

The indices set assessing certain regional water environment state generally consists of several key indices, thus assessing this regional water environment system’s risk should consider the different combined risk formed by each monitoring index’s over standard limit value. This paper will use the Copula function to establish the joint distribution among each key index, and with the aid of this joint distribution, we can quantitatively analyze the over standard combined risk of water quality key indices. Copulas describe the dependence between random variables independently of their univariate marginal distributions. A Copula is a distribution function on the \( n \)-dimensional unit cube with uniformly distributed marginals (Kazianka & Pilz 2010).

1. Four-dimensional Clayton–Copula joint distribution model

\[
C^{(4)}_\theta (u) = (u_1^{-\theta} + u_2^{-\theta} + u_3^{-\theta} + u_4^{-\theta} - 3)^{-1/\theta}, \quad (\theta > 0)
\tag{5}
\]

2. Four-dimensional Frank–Copula joint distribution model

\[
C^{(4)}_\theta (u) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)(e^{-\theta u_3} - 1)(e^{-\theta u_4} - 1)}{(e^{-\theta} - 1)^4}\right)^{1/\theta}, \quad (\theta > 0)
\tag{6}
\]

3. Four-dimensional Gumbel–Copula joint distribution model

\[
C^{(4)}_\theta (u) = \exp \left(-\left(-\ln u_1\right)^\theta + (-\ln u_2)^\theta + (-\ln u_3)^\theta + (-\ln u_4)^\theta \right)^{1/\theta}, \quad (\theta > 1)
\tag{7}
\]

4. Four-dimensional Normal–Copula joint distribution model

\[
C^{(4)}_\theta (u; \rho) = \Phi_4 (\Phi^{-1}(u_1), \Phi^{-1}(u_2), \Phi^{-1}(u_3), \Phi^{-1}(u_4)),
\tag{8}
\]

In the above four equations, (5)–(8), \( u = (u_1, u_2, u_3, u_4) \), and \( u_1, u_2, u_3, u_4 \) are the four marginal distributions in \([0, 1]\). The \( \theta \) in Equations (5)–(7) is the parameter in the functions of Clayton–Copula, Frank–Copula and Gumbel–Copula which describes four variables’ dependency relations. In Equation (8), \( \rho \) is a \( 4 \times 4 \) symmetrical positive definite matrix with elements being 1 in diagonal, \( \Phi_4 \) denotes the standard multivariate normal distribution with correlation coefficient matrix being \( \rho \), and \( \Phi^{-1}(\cdot) \) denotes standard normal distribution’s inverse function. The methods of the above Copula functions’ parameter value estimation, model optimization,
and joint distribution cumulative probability’s calculation can be seen in references Nelson (2006) and Feng et al. (2010).

APPLICATION OF EXAMPLES

MS–AR model’s parameter estimation

Guohe River is the main tributary of Tianjin Yuqiao reservoir. Since the water diversion project from Luanhe River to Tianjin city was completed and water was transited, Guohe River has become an important water delivery channel. The average annual runoff of the Guohe River accounts for about 85% of the total flow of Yuqiao reservoir’s annual inflowing, and the Guohe River bridge section is the key water quality monitoring section of the Haihe River basin, this section’s water quality condition reflects the water quality’s general condition in Tianjin Yuqiao reservoir’s upstream inflowing water. This paper takes the Guohe River bridge section’s monthly average observations of ammonia nitrogen (NH$_3$-N), total phosphorus (TP), 5-day biochemical oxygen demand (BOD$_5$), and dissolved oxygen (DO) as research objects from November 1999 to December 2007, and uses the method system constructed above to analyze Yuqiao reservoir’s water environment combined risk.

Before applying the MS–AR model to simulation of time series, the original monitoring data should be pretreated firstly to make the input data of MS–AR normal. Taking the monitoring data of NH$_3$-N as an example, carry the Box–Cox transform on the original data of NH$_3$-N firstly and then applying the standardized transform as: $y = \frac{x - \text{mean}(x)}{\text{std}(x)}$, $y(t)$ will be a time series obeying standard normal distribution.

Figure 1 is the status of the MS–AR model fitting $y$. The $\varepsilon$ in Figure 2 is the MS–AR model residual of $y(t)$. The $p$-value of the Kolmogorov–Smirnov normal test of $\varepsilon$ is 0.865, that is to say, the hypothesis of $\varepsilon$ obeying normal distribution can be accepted.

Utilizing the above methods to carry on the pretreatment of the monitoring data of NH$_3$-N, TP, BOD$_5$ and DO, we can get four time series obeying normal distribution. And then the parameters of the MS–AR model of the four normal time series was estimated. Table 1 is the MS–AR model parameters’ estimation results of Guohe River bridge section’s NH$_3$-N, TP, BOD$_5$, DO. The $P_{ii}$ in Table 1 represents the self-transferring probability at state $i$. Figure 3 is the residuals’ Q–Q graph of the MS–AR model of Box–Cox transformed data. Finally, to carry the inverse normal transformation and inverse standardization transformation on the fitting values of the above model, the MS–AR model fitted values of original data of NH$_3$-N, TP, BOD$_5$ and DO, as is shown in Figure 4.

Marginal distribution’s establishment

Using the four Markov status switching AR model constructed in previous section, combining the MC method carrying on random simulation on the four monitoring indices of NH$_3$-N, TP, BOD$_5$, DO, four simulation series can be obtained. Firstly, calculating these four simulation series’ empirical probability distributions, and the results obtained can be seen as the above four water quality monitoring indices’ theoretical probability distributions. Then taking these four probability distributions as marginal distributions, and using the Copula function the probability joint distribution of the above four water quality monitoring indices can be established. Table 2 shows statistical eigenvalues of the simulation sequences of NH$_3$-N, TP, BOD$_5$ and DO. Figure 5 is these four monitoring indices’ fitting diagram between simulation frequency distribution curves and empirical probability distribution dots.
Table 1 | The estimated values of MS-AR model’s autoregression coefficient

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>State I</th>
<th>State II</th>
<th>Coefficient</th>
<th>State I</th>
<th>State II</th>
<th>Coefficient</th>
<th>State I</th>
<th>State II</th>
<th>Coefficient</th>
<th>State I</th>
<th>State II</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₃-N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>φ₁ (Std error)</td>
<td>0.514 (0.008)</td>
<td>0.029 (0.003)</td>
<td>c (Std error)</td>
<td>0.436 (0.005)</td>
<td>-0.542 (0.006)</td>
<td>φ₁ (Std error)</td>
<td>0.804 (0.021)</td>
<td>-0.305 (0.005)</td>
<td>c (Std error)</td>
<td>-0.165 (0.006)</td>
<td>0.682 (0.004)</td>
</tr>
<tr>
<td>φ₂ (Std error)</td>
<td>-0.264 (0.007)</td>
<td>0.241 (0.005)</td>
<td>σ (Std error)</td>
<td>0.424 (0.003)</td>
<td>0.834 (0.011)</td>
<td>φ₂ (Std error)</td>
<td>-0.338 (0.003)</td>
<td>0.433 (0.007)</td>
<td>σ (Std error)</td>
<td>0.652 (0.008)</td>
<td>0.703 (0.004)</td>
</tr>
<tr>
<td>φ₃ (Std error)</td>
<td>-0.202 (0.005)</td>
<td>-0.149 (0.001)</td>
<td>Pᵣ (Std error)</td>
<td>0.737 (0.009)</td>
<td>0.742 (0.004)</td>
<td>φ₃ (Std error)</td>
<td>0.047 (0.002)</td>
<td>-0.030 (0.001)</td>
<td>Pᵣ (Std error)</td>
<td>0.838 (0.005)</td>
<td>0.504 (0.007)</td>
</tr>
<tr>
<td>φ₄ (Std error)</td>
<td>0.460 (0.007)</td>
<td>-0.236 (0.009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BOD₅</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>φ₁ (Std error)</td>
<td>0.779 (0.028)</td>
<td>-0.057 (0.001)</td>
<td>φ₂ (Std error)</td>
<td>-0.205 (0.006)</td>
<td>0.168 (0.006)</td>
<td>φ₁ (Std error)</td>
<td>0.788 (0.004)</td>
<td>0.435 (0.009)</td>
<td>c (Std error)</td>
<td>0.128 (0.005)</td>
<td>-0.356 (0.063)</td>
</tr>
<tr>
<td>φ₂ (Std error)</td>
<td>-0.383 (0.004)</td>
<td>-0.250 (0.003)</td>
<td>c (Std error)</td>
<td>0.040 (0.001)</td>
<td>-0.359 (0.005)</td>
<td>φ₂ (Std error)</td>
<td>-0.016 (0.002)</td>
<td>0.055 (0.003)</td>
<td>σ (Std error)</td>
<td>0.615 (0.002)</td>
<td>0.662 (0.033)</td>
</tr>
<tr>
<td>φ₃ (Std error)</td>
<td>-0.006 (0.002)</td>
<td>-0.057 (0.005)</td>
<td>σ (Std error)</td>
<td>0.722 (0.014)</td>
<td>0.722 (0.006)</td>
<td>φ₃ (Std error)</td>
<td>-0.243 (0.001)</td>
<td>0.182 (0.006)</td>
<td>Pᵣ (Std error)</td>
<td>0.938 (0.047)</td>
<td>0.554 (0.039)</td>
</tr>
<tr>
<td>φ₄ (Std error)</td>
<td>0.203 (0.007)</td>
<td>-0.467 (0.002)</td>
<td>Pᵣ (Std error)</td>
<td>0.928 (0.027)</td>
<td>0.650 (0.009)</td>
<td>φ₄ (Std error)</td>
<td>-0.255 (0.004)</td>
<td>0.133 (0.006)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Joint distribution’s establishment

According to Equations (5)–(8) and the simulation frequency obtained from simulated series, the $\theta$, $\rho$, $D$, OLS, and AIC values of various Copula functions that are used to couple the above four water quality monitoring item concentration series’ joint distribution were calculated. Here, $\theta$ and $\rho$ are the undetermined parameters in...
Copula functions, \( D \) (critical values) and the \( p \)-values are the Kolmogorov–Smirnov test result, OLS is the least sum of deviation square, and AIC is the AIC value under Akaike Information Criterion respectively based on difference Copula functions’ joint distribution models. These parameters’ estimation methods and calculation courses can be seen in the literature (Feng et al. 2010). As to Kolmogorov–Smirnov test, taking the significant level \( \alpha = 0.05 \), the fractile corresponding to \( n = 98 \) is 0.13403; and when the value of \( D \) is less than 0.13403, it will pass K–S test. The results of a specific calculation, test and assessment are shown in Table 3. At the same time, this paper selects Normal Copula as the Copula function of the joint distribution model, and this model’s fitting to empirical distribution results is shown in Figures 6 and 7.

Based on the constructed joint distribution model, the combined over standard risks of the four monitoring items of NH\(_3\)-N, TP, BOD\(_5\) and DO in the Guohe River bridge section were studied, and these four monitoring items’ combined over standard risk can be divided into 81 kinds of combination. For example, in which the combined risk parameters’ estimation methods and calculation courses can be seen in the literature (Feng et al. 2010).

<table>
<thead>
<tr>
<th>Copula function type</th>
<th>Parameter</th>
<th>( D ) (p-value)</th>
<th>OLS</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clayton</td>
<td>( \theta ) = 0.126</td>
<td>0.1361 (0.1324)</td>
<td>0.0487</td>
<td>−590.39</td>
</tr>
<tr>
<td>Frank</td>
<td>( \theta ) = 1.113</td>
<td>0.1016 (0.1324)</td>
<td>0.0378</td>
<td>−659.90</td>
</tr>
<tr>
<td>Gumbel</td>
<td>( \theta ) = 1.076</td>
<td>0.1276 (0.0644)</td>
<td>0.0480</td>
<td>−593.00</td>
</tr>
<tr>
<td>Normal Copula</td>
<td>( \rho_3 ), ( \rho_2 ), ( \rho_5 ), ( \rho_4 ), ( \rho_5 ), ( \rho_6 )</td>
<td>0.0664 (0.1324)</td>
<td>0.0212</td>
<td>−742.55</td>
</tr>
</tbody>
</table>

Table 2 | Statistical eigenvalue of four monitoring indices’ simulation sequences in Guohe River bridge section

<table>
<thead>
<tr>
<th>Experiment times</th>
<th>Mean value</th>
<th>Minimum value</th>
<th>Maximum value</th>
<th>Risk estimated value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH(_3)-N</td>
<td>40,000</td>
<td>0.379</td>
<td>0.098</td>
<td>12.293</td>
</tr>
<tr>
<td>TP</td>
<td>40,000</td>
<td>0.146</td>
<td>0.029</td>
<td>6.87</td>
</tr>
<tr>
<td>BOD(_5)</td>
<td>40,000</td>
<td>2.592</td>
<td>0.366</td>
<td>11.054</td>
</tr>
<tr>
<td>DO</td>
<td>20,000</td>
<td>11.280</td>
<td>3.001</td>
<td>45.880</td>
</tr>
</tbody>
</table>

Figure 5 | The fitting schematic diagram between simulation frequency and empirical frequency.
of NH$_3$-N inferior to class III water quality standard, TP inferior to class IV water quality standard, BOD$_5$ inferior to class III water quality standard, and DO inferior to class II water quality standard is:

$$p = \text{prob}(X_1 \geq 1.0, X_2 \geq 0.3, X_3 \geq 4.0, X_4 \leq 6)$$  \hspace{1cm} (9)$$

In Equation (9), 1.0, 0.3, 4.0 and 6 are the standard limited values for NH$_3$-N being class III standard, TP being class IV standard, BOD$_5$ being class III standard and DO being class II standard respectively in the surface water environment quality standard.

Table 4 shows the calculation results on combined risk, in which (II) denotes the probability of the monitoring item’s superior or equal to class II water quality standard, (III) denotes the probability of the monitoring item’s inferior to class II water quality standard but superior or equal to class III standard, and (IV) denotes the probability of the monitoring item’s inferior to class III water quality standard.

From Table 4, the results can be shown:

1. The probability of NH$_3$-N, TP, BOD$_5$ and DO simultaneously reaching class II water quality standard is 32.2%; under the condition of NH$_3$-N, BOD$_5$ and DO all reaching class II water quality standard, the probability of TP inferior to class II water quality standard is 21.08%. This shows that if we control well the TP discharge in the upstream basin, the total qualified ratio of water quality in the Guohe River bridge section will be able to reach 53.28%.

2. For the BOD$_5$ inferior to class III water quality standard, the probability of DO inferior to class II water quality standard is 0. For the BOD$_5$ concentration inferior to class II but superior to class III water quality standard, the probability of DO inferior to class II water quality standard is 0.06%. That is the probability of neither BOD$_5$ nor DO in the Guohe River bridge section reaching class II water quality standard simultaneously is very small, which indicates that there is a strong positive correlation between BOD$_5$ and the DO concentration time series, namely there is negative correlation between their over standard probability. The reason may be that

![Figure 6](image_url) | The fitting effect between empirical frequency and theoretical frequency.

![Figure 7](image_url) | The Q-Q graph of empirical distribution and theoretical distribution.

<table>
<thead>
<tr>
<th>Combined risk probability (%)</th>
<th>BOD$_5$(IV)</th>
<th>BOD$_5$(III)</th>
<th>BOD$_5$(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH$_3$-N(IV)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(IV)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.38</td>
</tr>
<tr>
<td>TP(III)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.09</td>
</tr>
<tr>
<td>TP(II)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>NH$_3$-N(III)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(IV)</td>
<td>0.00</td>
<td>0.00</td>
<td>1.45</td>
</tr>
<tr>
<td>TP(III)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.80</td>
</tr>
<tr>
<td>TP(II)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.27</td>
</tr>
<tr>
<td>NH$_3$-N(II)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP(IV)</td>
<td>0.00</td>
<td>0.00</td>
<td>2.50</td>
</tr>
<tr>
<td>TP(III)</td>
<td>0.00</td>
<td>0.00</td>
<td>3.44</td>
</tr>
<tr>
<td>TP(II)</td>
<td>0.00</td>
<td>0.00</td>
<td>3.30</td>
</tr>
</tbody>
</table>

Table 4 | The over standard combined risk analysis results on NH$_3$-N, TP, BOD$_5$, and DO
for DO inferior to class II water quality standard, the BOD₃ upstream is mostly oxidized, thus making BOD₅ concentration flowing through the Guohe River bridge section small; when DO concentration is high and reaching class II water quality standard, the BOD₃ upstream is not oxidized but gets through the Guohe River bridge section directly, thus resulting in BOD₅ over standard.

3. For the 30.85% probability of BOD₅ inferior to class II water quality standard, the probability of TP inferior to class II water quality standard is 20.07%, which shows that TP over standard has strong positive correlation with BOD₅ over standard in the Guohe River bridge section.

4. In the above four water quality monitoring indices, the main control factors affecting the Guohe River bridge section’s water quality status are TP and BOD₅, and the probabilities of them being inferior to class II water quality standard are high. At the same time, DO inferior to class II water quality standard has a strong correlation with TP and BOD₅ inferior to class II water quality standard in the Guohe River bridge section.

CONCLUSIONS

According to water quality time series change’s structural complexity and variance nonstationarity characteristics, this paper established a water quality time series AR model based on Markov status switching, and novelty combined a status switching AR model with the MC method to carry on water quality time series random simulation, then applied multidimensional joint distribution models based on multivariate Copula functions to four monitoring items’ over standard combined risk assessment. The risk assessment method system constructed in this paper can be used to make quantitative evaluation of the water environment’s combined risk. Through the combined risk assessment model, the correlativity among key pollutants affecting water environment quality can be analyzed, and the key factors affecting water quality condition can be found, so as to provide references for working out the water environment system’s safety management measures.

REFERENCES


First received 11 September 2012; accepted in revised form 17 December 2012