

# 基于 D 藤 Copula 的河湖水系连通伴生水污染 风险识别建模方法研究\*

张晨<sup>1)†</sup> 于若兰<sup>1)</sup> 杨蕊<sup>1,2)</sup> 高学平<sup>1)</sup> 孙嘉琦<sup>1)</sup>

(1)天津大学水利工程仿真与安全国家重点实验室, 300350, 天津;

2)南京水利科学研究院水文水资源与水利工程科学国家重点实验室, 210029, 江苏南京)

**摘要** 河湖水系连通作为提高水资源利用率、优化水资源配置的有效举措,在水资源保护与开发利用中越来越受重视.河湖水系连通工程在促进河湖生态环境改善、缓解水资源分配不均的同时,也伴随着诸多不确定性因素.针对河湖水系连通可能带来的伴生风险,基于 D 藤 Copula 函数在风险种类与指示因子之间建立辨识体系,根据指示因子的变化情况识别可能产生的风险.本研究以南四湖连通为例,应用水污染风险识别模型识别出  $\text{NH}_3\text{-N}$ 、 $\text{TN}$ 、 $\text{Chla}$  等物质为工程运行后水环境联合风险的关键因子,建议重点关注  $\text{TN}$  指标.计算结果表明,该建模方法能够实现水质风险识别,可以为河湖水系连通伴生风险识别管控工作提供技术支持.

**关键词** 河湖水系连通;风险识别;水质;Copula

**中图分类号** X52; X824

**DOI:** 10.12202/j.0476-0301.2020103

河湖水系是水资源的载体、生态环境的重要组成部分、经济社会发展的基础支撑.近年来,随着社会经济的发展,人类社会生产活动对河湖水系的影响不断加剧,河湖水系连通性减弱、水资源与水环境承载能力不足、水安全风险加大等问题不断出现,与水相关的生态与环境问题日益突出.针对我国所亟待解决的一系列水资源问题与挑战,水利部提出了河湖水系连通战略<sup>[1]</sup>.2010年9月30日,国务院加快水利建设专题会议上着重指出:“抓紧建设一批水资源配置工程、重点水源工程和河湖水系连通工程,以提高水资源时空调控和城乡供水保障水平”<sup>[2]</sup>.实施河湖水系连通举措,能够降低水旱灾害风险,修复和改善水生态环境,保障水资源可持续利用.

维持河湖水系连通,可以有效地改善河湖生态环境,优化区域水资源配置,但河湖连通也会伴随一系列潜在风险,例如洪旱灾害转移、水资源分布不均、水沙物质输移、水环境污染、水生态物质输入等<sup>[3]</sup>.当前,在水环境风险分析方面,CSPPC模型<sup>[4]</sup>、模糊随机风险率模型<sup>[5]</sup>、水质管理模拟-优化耦合模型<sup>[6]</sup>等初步探讨水质超标风险率;二维模糊蒙特卡洛法<sup>[7]</sup>及 WaterRAT 软件分析法<sup>[8]</sup>能够实现水环境风险评估.但上述方法大多未考虑到各风险要素间

的相关关系,不能很好地解决实际风险中多因子共同作用的问题. Copula 函数放宽了对传统多变量模型中相同分布族和线性关系的假设,并且在考虑边缘分布和多变量依存关系建模方面很有优势,为多随机变量联合分布建模提供了理想的方法<sup>[9-10]</sup>. D 藤适用于多变量间关系相对独立的情形,具有更高的灵活度.基于 D 藤 Copula 函数构建多变量作用下的水污染风险识别模型,比传统模型能直接构造各项关键指标的联合分布,打破单变量拟合的局限性<sup>[11]</sup>,有效解决实际工程面临的多元风险识别管控问题.

本文以识别河湖水系连通伴生水环境污染风险为目标,结合水环境污染指标评价准则,采用 D 藤 Copula 函数构建多影响因素的水污染风险识别模型,进行联合分布概率密度分析,将建立起的风险识别模型应用于实例识别出关键风险因子,为实现河湖水系连通伴生水污染风险管控工作提出可供参考的方案.

## 1 风险识别理论基础

风险识别是水环境风险管理项目的第一步,目的是在风险事件发生之前,科学系统地识别对项目有显著影响的因子,以便继续进行风险分析和估计.水环境风险识别是指在风险事故发生之前,运用各

\* 国家重点研发计划资助项目(2018YFC0407203);国家自然科学基金资助项目(51679160)

† 通信作者:张晨(1981—),男,博士,教授,研究方向:水力调控及环境效应研究. E-mail: emil@tju.edu.cn

收稿日期: 2020-03-06

种方法体系判断可能出现的各种水环境风险,对风险事故发生潜在因子进行鉴别的过程<sup>[12]</sup>.在风险识别阶段,采取科学、有效的方法和途径对系统风险因子进行全面的分析与识别,不仅能够提高对风险因子客观的认识,也为风险分析工作提供科学的理论依据.做好水环境风险识别这一基础工作,对整个河湖水系连通的风险管理与控制、实现水环境的健康可持续发展具有重要意义.

风险一般是指遭受损害、不利或毁灭的可能性<sup>[12]</sup>,基于概率和期望值来推测结果的可能事件和严重程度<sup>[13]</sup>.在风险理论研究,德克斯坦等<sup>[14]</sup>利用系统  $A$  中 2 个相互作用力荷载  $L$  与阻尼  $R$  的关系来定义系统的风险:

$$P_F = P(A) = P(L > R), \quad (1)$$

式中:  $A$  为系统失事事件;  $L > R$  代表系统荷载超过阻尼;  $P_F$  为系统荷载大于阻尼的概率,即可代表系统风险.

对于本研究的河湖水系连通水环境系统而言,“荷载”可以定义为污染物指标的浓度,“阻尼”为规定范围内可容纳污染物的临界值.由此,河湖水系连通伴生水污染风险可被定义为河湖水系连通后,受系统各种不确定性污染源的影响,水系水质超过一定标准值的可能.

## 2 Copula 函数理论基础

**2.1 定义** Copula 函数是定义在  $[0, 1]^n$  区间上的多维联合分布函数,它能够将多个随机变量的边际分布连接起来构造联合分布.令  $F(X_1, X_2, \dots, X_n)$  为在  $(0, 1)$  上均匀分布的联合分布函数,根据 Sklar 定理<sup>[15]</sup> 可作如下表述:

$$F(X_1, X_2, \dots, X_n) = C_\theta(F_1(x), F_2(x), \dots, F_n(x)), \quad (2)$$

式中  $C_\theta(F_1(x), F_2(x), \dots, F_n(x))$  为 Copula 函数,  $\theta$  为 Copula 参数,  $F(x)$  为各随机变量的边际分布函数.若  $F_1(x), F_2(x), \dots, F_n(x)$  连续,则 Copula 函数唯一且确定<sup>[16]</sup>.

**2.2 类型** Copula 函数种类繁多,应用较广泛的 Copula 函数类型为椭圆 Copula 族函数和阿基米德 Copula 族函数.椭圆 Copula 族函数是指具备椭圆分布的 Copula 函数族,由于其具备多元正态分布的多个属性,因此实际应用十分广泛,包括正态 Copula 函数和 Student t-Copula 函数.阿基米德 Copula 族函数在水文领域中较为常见,它构造简单且种类繁多,能够在水文多变量分析中建立各种类型变量间的相关关系,根据生成元函数的不同,其又可分为 Clayton Copula、Gumbel Copula、Frank Copula 和 Joe Copula 等<sup>[17]</sup>.

**2.3 参数估计** Copula 参数估计常使用非参数估计法、适线法及极大似然估计法<sup>[18]</sup>.除此之外, Kendall 秩相关系数  $\tau$  也很重要,它不仅能够描述 2 个随机变量间的正负线性相关关系,也能刻画非线性相关关系.通过  $\tau$  与 Copula 参数之间的相关关系也能够用来估计参数值.

**2.4 拟合优度检验** 为检测指定的分布模型是否能够拟合变量的实际分布,通常在参数估计之后采用适当的拟合优度准则进行比选. K-S 检验、Q-Q 图检验、 $\chi^2$  检验、赤池信息准则 (AIC) 与贝叶斯信息准则 (BIC) 在 Copula 函数检验中应用广泛.本文选用 AIC 法检验单一变量分布:

$$C'_{AI} = -2F_{LK} + 2k, \quad (3)$$

$$C'_{BI} = -2F_{LK} + \ln N \cdot k, \quad (4)$$

式中  $F_{LK}$  为拟合模型的最大对数似然,  $k$  为拟合参数值,  $N$  为观测值.  $C'_{AI}$  与  $C'_{BI}$  值越小,拟合分布模型适合度越高.

## 3 基于 D 藤 Copula 函数的连通伴生水污染风险识别模型

传统 Copula 模型的应用多为二维变量模型,但由于河湖水系连通工程系统庞大,所涉及的风险因子众多,且风险因子之间大多存在相关关系,因此需要建立多元变量的联合概率分布. Joe<sup>[19]</sup> 提出将多变量联合分布化解为一系列独立的 Pair-Copula 的方法, Bedford 等<sup>[20]</sup> 在此基础上提出了藤式 Copula 的方法.本文使用 D 藤 Copula 函数构建多变量水污染风险识别模型,模型主要包括基础数据处理、构建藤结构和伴生风险因子识别 3 部分.水污染风险识别模型逻辑见图 1.

**3.1 基础数据处理** 计算初始观测序列特征值,并选用 6 种常用于水文变量的分布(正态分布、指数分布、logistic 分布、对数正态分布、伽马分布和威布尔分布)来分别对风险因子初始观测序列进行拟合,根据经验分布曲线与计算分布曲线的拟合程度及拟合优度检验准则 AIC 值选取最优边缘分布函数.随机变量边缘分布的经验频率计算式为

$$P(K \leq k) = \frac{K - 0.44}{N + 0.12}. \quad (5)$$

之后应对选用的最优边缘分布做概率积分变换,将序列转化为  $[0, 1]$  区间上均匀分布的概率值,以满足 Copula 函数应用条件.在此可利用 K-S 检验来判断转化后的序列值是否满足  $[0, 1]$  上均匀分布,当

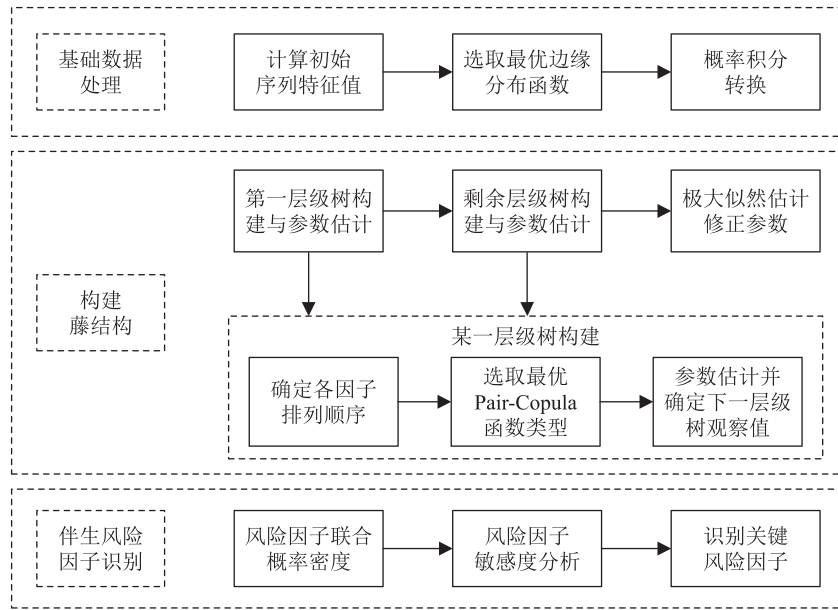


图 1 基于 D 藤 Copula 的水污染风险识别模型逻辑

K-S 检验值伴随概率 > 0.05 时, 可认为序列的边缘分布拟合恰当。

**3.2 构建藤结构** 基于 D 藤 Copula 模型能够有效解决 Copula 函数只能分析二维数据之间相关性的不足. C 藤与 D 藤在高维变量相依结构分析中具有很大优势, 是目前最常见的 2 种 Copula 藤结构. C 藤适用于拟合多变量中具有一个关键变量的情形, 当变量间存在一定主次关系时, 适合采用 C 藤进行刻画<sup>[21]</sup>. D 藤适用于变量间关系相对独立的情况, Czado<sup>[22]</sup> 的研究证明, D 藤能够更为准确地描述变量间的对称和非对称关系, 刻画多变量间的非线性关联结构. 本文选取 D 藤结构对变量进行构建, 四维情况下的 D 藤 Copula 结构如图 2 所示<sup>[23]</sup>.

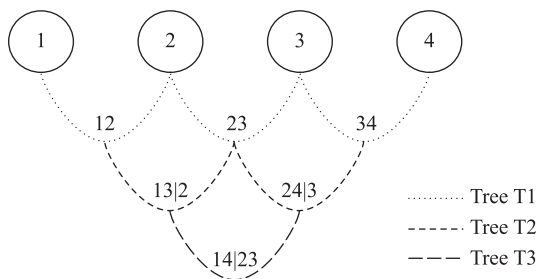


图 2 D 藤 Copula 四维变量结构

利用极大似然估计的方法对 D 藤 Copula 进行参数估计. D 藤 Copula 的  $n$  维联合密度表达式为

$$F(x_1, x_2, \dots, x_n) = \prod_{k=1}^n f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} C_{i, i+j|1, \dots, i+j-1} \{F(x_i|x_{i+1}, \dots, x_{i+j-1}), F(x_{i+j}|x_{i+1}, \dots, x_{i+j-1})\}, \quad (6)$$

式中下标  $j$  表示树的层级遍历, 下标  $i$  表示每棵树中节点的遍历. 其中, 条件分布函数  $F(x|v)$  可由  $h$  函数计算得到:

$$h(x|v, \theta) := F(x|v) = \frac{\partial C_{xv|v_{-j}}(F(x|v_{-j}), F(v_j|v_{-j})|\theta)}{\partial F(v_j|v_{-j})}, \quad (7)$$

式中  $v$  表示一个  $d$  维向量,  $v_j$  表示向量  $v$  的任意一个元素,  $v_{-j}$  表示去除元素  $v_j$  后得到的  $d-1$  维向量. 由此可得出  $n$  维联合分布的对数极大似然函数为:

$$l_{Dv}(\theta_{Dv}|u) = \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} \sum_{t=1}^T \lg(C_{i, i+j|1, \dots, i+j-1}(F(x_{i,t}|x_{i+1,t}, \dots, x_{i+j-1,t}), F(x_{i+j,t}|x_{i+1,t}, \dots, x_{i+j-1,t}))), \quad (8)$$

参数估计的每层极大似然估计方法为: 首先, 从原始数据进行极大似然估计, 得第 1 层级树中各个 Pair-Copula 参数值; 然后将第 1 层树估计出的 Copula 参数代入到第 2 层树中, 利用  $h$  函数计算其条件分布函数, 由此计算出第 2 层级树中的观测值; 再利用第 2 层级树的观测值, 采用极大似然估计法估计第 2 层树的 Copula 参数值. 重复以上步骤, 可得剩余层级树的 Copula 参数值.

以四维变量 D 藤 Copula 函数为例, 构建四维变量 D 藤结构.

1) 根据各变量间的观察值计算出两两变量间的 Kendall 秩相关系数  $\tau$ . 利用 Dissmann 等提出的方法, 以 Kendall 秩系数的绝对值为权重, 采用最大生成树算法计算得出 D 藤 Copula 第 1 层级树的顺序<sup>[24]</sup>.

2) 确定第 1 层级树中各变量之间的 Pair-Copula

函数类型. 尽管阿基米德型 Copula 函数应用最为广泛, 但是大量文献证明, Twan、BB 族等 Copula 函数更适用于降水、洪水等气象水文要素的多变量分析<sup>[25]</sup>, 因此本文采用 7 种 Copula 函数(Gaussian Copula、Student't Copula、Clayton Copula、Gumbel- hougard Copula、Frank Copula、Joe Copula、BB 族 Copula)作拟合对象, 根据 AIC 标准选取最适合的 Pair-Copula 函数, 并利用极大似然估计进行参数估计.

3) 根据 2) 得出的第 1 层级树参数值, 利用式 (7)  $h$  函数计算得到第 1 层级树观测值.

4) 重复以上步骤, 确定剩余层级树藤结构及参数值.

5) 最后利用极大似然估计修正 D 藤 Copula 参数.

**3.3 伴生风险因子识别** 完成藤结构的构建后, 根据 D 藤  $n$  维联合密度表达式, 可计算出  $n$  种风险因子作用下水环境的联合风险概率密度, 定量改变某一因子的输入值, 观察整体风险概率的变化情况, 并根据该结果对各风险因子作敏感度分析. 通过对比河湖连通前后风险联合分布概率密度的变化, 确定河湖水系连通对水环境的影响, 识别河湖水系连通关键风险因子, 完成水污染风险识别.

## 4 模型应用实例

选择沂沭泗水系中南四湖流域进行河湖连通伴生水污染风险因子识别与敏感度分析的研究. 南四湖水环境变量众多, 风险因子间结构复杂, 南水北调东线工程全面通水后, 必将对南四湖产生影响. 本文选用南四湖南阳站 2008—2014 年的氨氮 ( $\text{NH}_3\text{-N}$ )、总氮 (TN)、总磷 (TP) 和叶绿素  $a$  (Chla) 4 个代表性观测指标建立联合风险概率模型. 南四湖连通工程通水时间为 2013 年 11 月, 现将调研数据以 2013 年 11 月为界, 分为 2008 年 1 月—2013 年 10 月和 2013 年 11 月—2014 年 12 月这 2 个时间段, 分别代表连通前、连通后 2 个阶段.

**4.1 基础数据处理** 采用 3.1 中的 6 种分布分别对连通前、后风险因子初始观测序列进行拟合, 并根据 AIC 标准对各分布拟合结果进行比较, 连通前工况下的结果如表 1 所示.

表 1 连通前各因子不同分布类型对应的 AIC 值

风险物质	正态	指数	logistic	对数正态	伽马	威布尔
$\text{NH}_3\text{-N}$	31.60	39.39	27.42	-10.51	-0.72	8.69
TN	266.35	240.43	251.14	212.58	220.69	227.31
TP	-210.13	-243.77	-247.90	-282.59	-270.78	-258.96
Chla	568.46	554.27	559.92	521.77	530.09	536.74

对于各因子, AIC 值越小, 拟合分布模型适合度越高. 由表 1 结果可知, 对于连通前的情况,  $\text{NH}_3\text{-N}$ 、TN、TP 及 Chla 均适用于对数正态分布. 使用同样方法对连通后各因子分布拟合结果进行比较, 可知连通后  $\text{NH}_3\text{-N}$  适用于伽马分布, TN 及 TP 适用于对数正态分布, Chla 适用于威布尔分布.

利用所选用的边缘分布函数对序列进行概率积分转换, 使新序列满足  $[0, 1]$  区间上均匀分布. 对新序列进行 K-S 检验并将结果记录于表 2.

表 2 K-S 检验结果

风险物质	K-S 统计量伴随概率	
	连通前	连通后
$\text{NH}_3\text{-N}$	0.68	0.57
TN	0.15	0.71
TP	0.07	0.41
Chla	0.94	0.31

由表 2 可见, 连通前后各风险因子 K-S 统计量伴随概率值均  $> 0.05$ , 可以认为新序列在  $[0, 1]$  上均匀分布, 可用于 Copula 函数构建.

**4.2 D 藤 Copula 函数构建** 首先建立连通前工况下的藤结构, 对 D 藤 Copula 函数第 1 层级树进行构建和参数估计, 将 4 种变量间的 Kendall 秩相关系数  $\tau$  记录于表 3.

表 3 4 种变量间的 Kendall 秩相关系数  $\tau$

风险物质	$\text{NH}_3\text{-N}$	TN	TP	Chla
$\text{NH}_3\text{-N}$	1.00	0.04	-0.01	-0.22
TN	0.04	1.00	0.10	0.06
TP	-0.01	0.10	1.00	0.06
Chla	-0.22	0.06	0.06	1.00

由表 3 结果, 得到第 1 层级树顺序为 TN、TP、Chla、 $\text{NH}_3\text{-N}$ . 从 3.2 节 7 种备选 Pair-Copula 类型中, 根据拟合优度检验准则选取两两变量间最适宜 Pair-Copula 类型, 并利用  $\tau$  对第 1 层级树进行参数估计, 将每对变量函数类型及参数估计最终结果列于表 4.

表 4 Tree 1 Pair-Copula 类型及参数估计

变量	Best Copula	参数估计
(TN, TP)	Frank Copula	0.948
(TP, Chla)	Gaussian Copula	0.099
(Chla, $\text{NH}_3\text{-N}$ )	Gaussian Copula	-0.337

根据第 1 层级树 Pair-Copula 类型和参数估计结果, 可以利用  $h$  函数生成第 2 层级树的各变量观察

值. 利用同种方法计算得第 2 层级树每对变量函数类型及参数估计结果如表 5 所示.

表 5 Tree 2 Pair-Copula 类型及参数估计

变量	Best Copula	参数估计
(TN, Chla) TP	Gaussian Copula	-0.008
(TP, NH <sub>3</sub> -N) Chla	Frank Copula	-0.162

由第 2 层级树的构建结果, 得出第 3 层级树两变量因子 (TN, Chla) |TP 与 (TP, NH<sub>3</sub>-N) |Chla 之间的  $\tau$  为 -0.004, 最优 Pair-Copula 类型为 Gaussian-Copula, 参数估计结果为 -0.007. 通过拟合优度检验, 利用极大似然估计修正 D 藤 Copula 参数, 修正结果见表 6.

表 6 修正后的 D 藤 Copula 参数

Tree	各层级树 Pair-Copula	参数初始值	参数终值
1	C(TN, TP)	0.948	0.994
	C(TP, Chla)	0.099	0.177
	C(Chla, NH <sub>3</sub> -N)	-0.337	-0.315
2	C(TN, Chla) TP	-0.008	0.074
	C(TP, NH <sub>3</sub> -N) Chla	-0.162	0.006
3	C((TN, NH <sub>3</sub> -N) (TP, Chla))	-0.007	0.997

根据 3 列层级树构建顺序、修正后各 Pair-Copula 类型及参数终值, 可得四元风险因子 D 藤 Copula 模型, 其结构如图 3 所示.

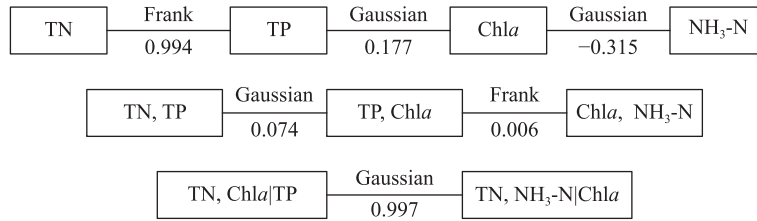


图 3 四元风险因子 D 藤 Copula 模型结构

重复以上步骤, 建立连通后工况下的 D 藤 Copula 函数. 至此, 连通前后 Copula 函数模型均构建完成, 可进行风险因子识别.

4.3 水系连通风险因子识别 以中国地表水环境标准规定的水的分类标准及我国湖泊各指标对应的营养状态分级标准, 结合建立起的连通前后 D 藤 Copula

模型, 分别输入 4 种风险因子均值作为基准值, 改变某一确定因子的质量浓度值, 以此观察模型输出量的变化, 进而得到连通前后 4 种变量作用下整体风险变化情况(图 4).

由图 4 可知, 从总体上来说, 各风险因子浓度变化引发的风险在水系连通后得到明显改善, 说明河

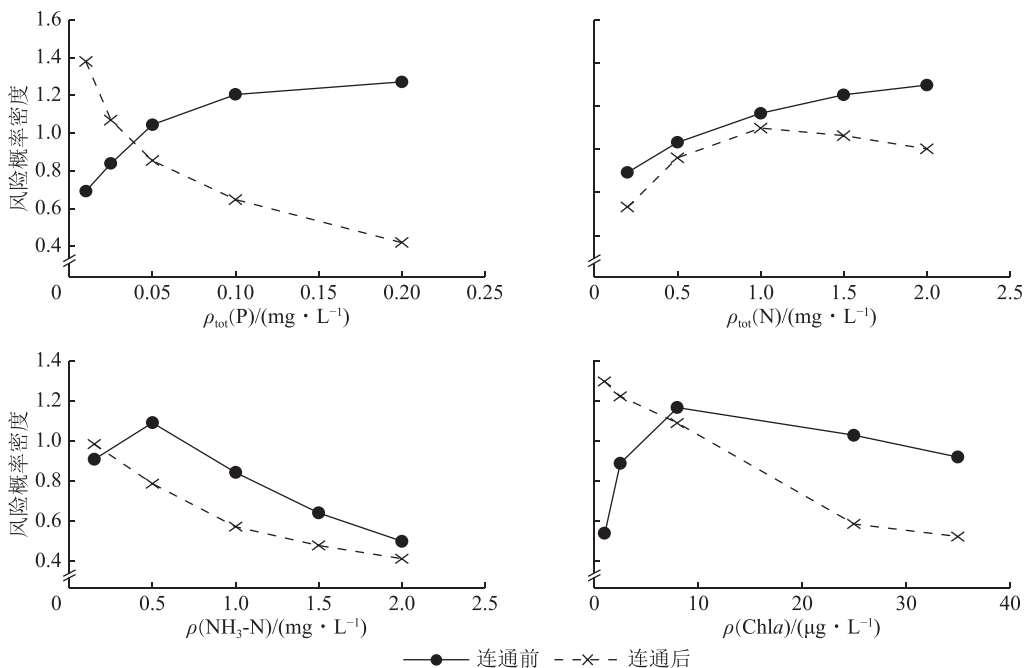


图 4 连通前后各风险因子概率密度变化曲线

湖连通工程效果良好。综合考虑,南水北调东线工程通水后, TN 风险概率密度呈现略微上升的趋势,且相较于其他指标, TN 风险概率密度较大,因此确定 TN 为河湖连通引起的伴生污染风险因子,应在风险管控中重点关注 TN 指标。对各风险因子作敏感度分析, TP 风险概率密度明显降低且敏感度变化最大,可排除 TP 作为关键风险因子;  $\text{NH}_3\text{-N}$ 、TN、Chla 概率密度值前后相近,且较运行前减小幅度不大,因此可认为  $\text{NH}_3\text{-N}$ 、TN、Chla 为工程运行后水环境联合风险的关键因子。

## 5 结论

本文基于 Copula 理论,利用 D 藤 Copula 函数建立起考虑多变量作用下的水环境污染风险识别模型。该模型能够根据样本数据的差异为初始观测序列拟合不同类型的边缘分布,再通过变量间相关关系选择不同类型的 Pair-Copula 函数,体现不同变量因子间的结构差异性,构建多元因子作用下的水污染风险识别模型。结合南四湖研究区域,运用所提出的河湖连通伴生水污染风险识别模型,将  $\text{NH}_3\text{-N}$ 、TN、Chla 识别为工程运行后水环境联合风险的关键因子,以便进一步实现对该流域河湖连通伴生风险的管控,为河湖水系伴生风险识别管控工作提供科学依据和技术支持。

## 6 参考文献

- [1] 李宗礼,李原园,王中根,等. 河湖水系连通研究:概念框架[J]. 自然资源学报,2011,26(3): 513
- [2] 夏军,高扬,左其亭,等. 河湖水系连通特征及其利弊[J]. 地理科学进展,2012,31(1): 26
- [3] 庞博,徐宗学. 河湖水系连通战略研究:关键技术[J]. 长江流域资源与环境,2015,24(增刊1): 146
- [4] 陈小红,涂新军. 水质超标风险率的CSPPC模型[J]. 水利学报,1999,30(12): 1
- [5] 李黎武,施周. 基于模糊事件概率理论的水质风险率计算方法[J]. 水利学报,2007,38(4): 417
- [6] BURN D H. Water-quality management through combined simulation-optimization approach[J]. Journal of Environmental Engineering, 1989, 115(5): 1011
- [7] KENTEL E, ARAL M. 2D Monte Carlo versus ZD fuzzy Monte Carlo health risk assessment[J]. Stochastic Environmental Research and Risk Assessment, 2005, 19(1): 86
- [8] MCINTYRE N R, WHEATER H S. A tool for risk-based management of surface water quality[J]. Environmental Modelling and Software, 2004, 19(12): 1131
- [9] FAN Y, HUANG K, HUANG G H, et al. A factorial Bayesian Copula framework for partitioning uncertainties in multivariate risk inference[J]. Environmental Research, 2020, 183: 109215
- [10] LIU Y R, LI Y P, Ma Y, et al. Development of a Bayesian-copula-based frequency analysis method for hydrological risk assessment-the Naryn River in Central Asia[J]. Journal of Hydrology, 2020, 580: 124349
- [11] 孙妍,王秀茹,王铭浩,等. 渠村引黄口流域水沙情势变化和丰枯遭遇分析[J]. 北京师范大学学报(自然科学版),2019,55(4): 489
- [12] 孙滔滔,赵鑫,尹魁浩,等. 水环境风险源识别和评估研究进展综述[J]. 中国水利,2018(15): 41
- [13] 陈浩,徐宗学,班春广,等. 基于Copula函数的深圳河流域降雨潮位组合风险分析[J]. 北京师范大学学报(自然科学版),2020,56(2): 307
- [14] 德克斯坦 L,勃兰特 E J. 水资源工程可靠性与风险[M]. 北京:水利水电出版社,1993
- [15] SKLAR A. Fonctions de repartition an dimensions et leurs marges[J]. Publication de l'Institut de Statistique de l'Universite de Paris, 1959, 8: 229
- [16] 张翔,冉啟香,夏军,等. 基于Copula函数的水量水质联合分布函数[J]. 水利学报,2011,42(4): 483
- [17] ANNE C, SALAHEDDINE A, LUC P. Multivariate hydrological frequency analysis using copulas[J]. Water Resources Research, 2004, 40(1): 1
- [18] 邱小霞,刘次华,吴娟. Copula函数中参数极大似然估计的性质[J]. 经济数学,2008,25(2): 210
- [19] JOE H. Parametric families of Multivariate distributions with given margins[J]. Journal of Multivariate Analysis, 1993, 46(2): 262
- [20] BEDFORD T, COOK R M. Probability density decomposition for conditionally dependent random variables modeled by vines[J]. Annals of Mathematics and Artificial Intelligence, 2001, 32: 245
- [21] 杨蕊,吴时强,高学平,等. 基于Vine Copula函数的河湖连通水环境多因子联合风险识别研究[J]. 水利学报,2020,51(5): 606
- [22] CZADO C. Pair-copula constructions of multivariate copulas[J]. Copula Theory and Its Applications, 2010, 198(4): 93
- [23] LI Q, WANG H, ZENG Y, et al. An efficient d-vine copula-based coupling uncertainty analysis for variable-stiffness composites[J]. Composite Structures, 2019, 219(1): 221
- [24] DISSMANN J, BRECHMANN E C, CZADO C, et al. Selecting and estimating regular vine copula and application to financial returns[J]. Computational Statistics and Data

Analysis, 2013, 59: 52

[25] 姚蕊, 孙鹏, 张强, 等. 基于Copula淮河流域水文干旱频

率分析及影响研究[J]. 北京师范大学学报(自然科学版), 2019, 55(6): 755

## A D-vine Copula-based model for identification of associated water pollution risk in interconnected river system network

ZHANG Chen<sup>1)†</sup> YU Ruolan<sup>1)</sup> YANG Rui<sup>1,2)</sup> GAO Xueping<sup>1)</sup> SUN Jiaqi<sup>1)</sup>

(1)State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, 300350, Tianjin, China;

2)State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering,

Nanjing Hydraulic Research Institute, 210029, Nanjing, Jiangsu, China)

**Abstract** In water resource conservation and development, interconnected river system network (IRSN) has become an effective measure to increase water resources efficiency and allocation. IRSN improves heterogeneous distribution of water resources and freshwater ecological restoration. However, uncertainties due to natural and anthropogenic forces still exist. Therefore a D-vine Copula-based model was developed to identify risk factors in IRSN. IRSN of Nansi Lake was utilized to identify risk of water quality indicators before and after the connection project. Three water quality indicators ( $\text{NH}_3\text{-N}$ , TN and  $\text{Chl}a$ ) were considered primary risk factors. Water resource managers are suggested to pay more attention to TN. The D-vine Copula-based model provides useful reference and technical assistance in IRSN projects.

**Keywords** interconnected river system network; risk identification; water quality; Copula

【责任编辑: 刘先勤】