

Eutrophication Hazard Evaluation Using Copula-Cloud

Dengfeng Liu, Dong Wang*, Yuankun Wang

Key Laboratory of Surficial Geochemistry, Ministry of Education; Department of Hydro-sciences, School of Earth Sciences and Engineering, Nanjing University Nanjing, 210023, China

Received 1 July 2015

Accepted 25 January 2016

Abstract

A copula-cloud (C-C) coupled framework is proposed with respect to the uncertainty and multi-criteria mapping in water environmental hazard evaluation, with introducing copulas and cloud model theory. In the C-C framework, cloud models are firstly generated with given risk criteria which quantify the concept of risk from each dimension, before that copulas are estimated with the dependence of observed water quality data. The multi-dimensional copula-cloud model can be finally constructed with connecting all clouds by copulas, which can model the randomness, fuzziness and dependence of each risk criterion in assessment. The C-C technique have been applied on cases of representative lakes or reservoirs in China. Results in comparison with relevant proposed methods verify the completeness and effectiveness of the technique. As a beneficial exploration on clouds and copulas, the proposed C-C technique provide an innovative approach in relevant risk assessment.

Keywords: Risk assessment, Copula, Cloud model, Eutrophication hazard; Water environment

COPULA-云在富营养化风险评价中的应用

刘登峰, 王栋*, 王远坤

表生地球化学教育部重点实验室, 南京大学地球科学与工程学院水科学系, 南京, 210046

摘要: 针对现有水环境风险评价模型中概念不确定性和多指标映射问题, 引入云模型和 copula 函数, 提出了一种耦合 copula 和云的风险评价模型(C-C)。在 C-C 方法框架下, 首先根据富营养化指标量化体系通过云发生器生成相应云模型, 再由各类指标下实测数据的相依特性构建 copula 多元映射模型, 最后由 copula 联接各个云得到多维 copula-云评价模型。该模型能够同时刻画富营养化风险评价中的模糊性、随机性以及各风险分量间的相依性。将 C-C 模型应用于我国代表性湖库的富营养化风险评价中, 并将评价结果与多种现有方法进行对比。分析表明 C-C 模型考虑因素全面, 评价结果准确可靠, 是云模型和 copula 多元分析理论在风险评价领域的有益探索。

关键词: 风险评价; copula; 云模型; 富营养化

1. 引言

在环境问题日益严重的大背景下, 以水体富营养化为代表的水环境风险日渐显现, 严重威胁了居民的生产生活安全。OECD 将富营养化定义为“水体营养盐增加引起的一系列征兆变化, 其中藻类和大型植物生产力的增加、水质恶化和其他征兆变化破坏了水的利用”^[1]。对以富营养化为代表的水环境风险的科学评价是控制水体污染, 保证区域水环境安全的重要前提。

富营养化风险评价, 即通过既定风险指标体系确定水体的富营养化状态属性, 是一个涉及定性定量转化的多变量综合决策过程。近年来的研究从不同角度入手, 提出的评价模型也各有特点。如基于 Weber-Fechner 定律、突变模型、集对分析与可变模糊集等方法^[2-5]。丁昊和王栋考虑了评价过程中模糊-随机性二元共存的特点, 引入云模型理论用于富营养化评价^[6]; 刘登峰等进一步总结了富营养化评价的 3 个重要问题, 并以此提出了一种熵-云耦合评价模型^[7]。已有的混合熵权方法虽有利于使权重分配更趋合理, 但也无法排除其存在更大不确定的可能性。而权重本质上作为多指标映射的一种线性实现

* Corresponding Author: wangdong@nju.edu.cn

方式，在原理上具有一定局限性，亟待进一步研究完善。

Copula 函数是一类连接多元随机变量的边缘分布以得到其联合分布的联结函数，由于该函数形式灵活且构造简单，因此被逐渐广泛应用于水资源多元分析领域^[8-10]。Copula 函数能够捕捉和刻画多元变量间的相依结构^[11-12]，实质上是一种基于变量相依特性的多元映射方法。张翔等利用 copula 函数构建了水质水量的多维联合分布^[13]；陈晶等进一步探讨了 copula 函数在水环境评价中的合理性和可靠性^[14]。本文针对现有云评价模型中结构的复杂性以及定权问题，简化了云评价模型的构造，从各指标间的相依性入手引入 copula 函数，提出了一种耦合 copula 和云的风险评价模型 (C-C)。首先根据风险指标体系通过云发生器生成若干云模型，再通过各类指标下实测数据的相依性构造 copula 函数，最后由 copula 函数联接云得到多维 copula-云模型。将 C-C 模型应用于我国代表性湖库的富营养化风险评价中，并将评价结果与多种现有方法对比，以验证模型的可靠性。

2. Copula-云风险评价模型(C-C)

2.1. 云模型

云模型是李德毅提出的一种定性概念与其定量表示之间的不确定性转换模型^[15]。云模型已被证明具有应用的普适性，并开始应用于风险评价领域^[16]。此外，云模型在富营养化评价中的应用原理见文献^[6-7]，此处不作赘述。

文献^[7]所提熵-云模型采用一个指标下每个等级对应一个云模型的研建方式，即一个富营养化熵-云评价模型由 30 组云模型构成，这种模型结构略显复杂，不仅增加了建模的难度，也增加了评价的工作量。本文提出了一种简化的云模型研建方式的方法，即通过半云发生器^[17]生成一个云模型以表征一个指标下所有等级的风险程度。在该建模思路下只需要针对每个评价指标建立 5 组云模型，这样较大程度上减少了原有方法的工作量，是对原有云评价方法的改进。

2.1.1 参数确定

云模型用参数组 (Ex, En, He) 表达某一个指标下的风险程度。各参数由式(1)得到^[7-8]：

$$\begin{cases} Ex = (B_{\min} + B_{\max}) / 2 \\ En = (B_{\max} - B_{\min}) / 6 \\ He = k \cdot En \end{cases} \quad (1)$$

B_{\max}, B_{\min} 为描述某一风险等级定量区间的上下确界； k 取 0.1^[7]。

参考文献^[18]建议的指标体系确定各指标下的云模型参数（见表 1）。

表 1 各风险评价指标下云模型参数 (Ex, En, He) 表

风险指标	云模型参数		
	Ex	En	He
Chl-a	75	25.00	2.50
TP	250	83.33	8.33
TN	2500	833.33	83.33
COD	14	4.67	0.93
SD	0	5.00	1.00

2.1.2 模型生成

对于本例中的各风险指标变量，采用半云发生器生成其相应云模型：

① 生成以 En 为期望值、 He 为标准差的一个正态随机数 En' ；

② 生成以 Ex 为期望值、 En' 的绝对值为标准差的一个正态随机数 x 。 x 称为富营养化风险论域内的一个云滴；

③ 计算，令

$$u = U(x) = \begin{cases} \frac{-(x-Ex)^2}{2(En')^2}, & x < En \text{ (上升云)} \\ 1, & x \geq En \end{cases} \quad (2)$$

$$\text{或 } u = U(x) = \begin{cases} \frac{-(x-Ex)^2}{2(En')^2}, & x > En \text{ (下降云)} \\ 1, & x \leq En \end{cases} \quad (3)$$

u 为 x 隶属于富营养化概念的确度，以量化富营养化风险程度；

④ 重复(1)-(3)，直至产生 N 个云滴为止。

2.2. 多维 copula-云模型

Copula 函数 $C(\cdot, \dots, \cdot)$ 是定义在 $[0, 1]^d$ 上的具有零基面且 d 维递增的分布函数。根据 Sklar 定理^[21]，令 $H(\cdot, \dots, \cdot)$ 为具有边缘分布 $U_1(\cdot), \dots, U_d(\cdot)$ 的 d 维联合分布函数，那么存在一个 copula 函数 $C(\cdot, \dots, \cdot)$ ，满足：

$$H(\cdot, \dots, \cdot) = C(U_1, \dots, U_d) \quad (4)$$

若 $U_1(\cdot), \dots, U_d(\cdot)$ 连续, 则 $C(\cdot, \dots, \cdot)$ 唯一确定。

常用的 copula 函数包括多元 Gaussian copula、多元 t-copula、Archimedean 型 copula、极值 copula 等, 关于 copula 理论的详细介绍参见 [22-23], 此处不做赘述。在各类 copula 函数中, 多元 Gaussian copula 因其能够描述多元正态变量间的相依性, 且在变量维数大于 3 时效果优良, 故广泛应用于各类 Gaussian 过程或近似 Gaussian 过程的多元建模中 [24]。令 ρ 为对角元素为 1 的 d 维对称正定矩阵, 且 Φ_ρ 是以 ρ 为相关系数矩阵的标准正态分布, 则相应 d 维 Gaussian copula 函数为:

$$C(U_1, \dots, U_d; \rho) = \Phi_\rho[\Phi_\rho^{-1}(U_1), \dots, \Phi_\rho^{-1}(U_d)] \quad (5)$$

本文采用 Gaussian copula 构造多维 copula-云模型, 从而实现多指标到风险等级的映射。富营养化指标体系下包含 5 个评价指标, 因此维数 $d=5$ 。首先通过半云发生器生成 5 个指标维度上的云模型 $U_1(\cdot), \dots, U_5(\cdot)$, 再计算水质数据的相关系数矩阵 ρ 并得到 Gaussian copula 函数 C , 最终通过 C 联结 5 个风险指标下的云得到 5 维 copula-云评价模型 $C(U_1, \dots, U_5)$ 。

2.3. C-C 模型评价

C-C 建模的基本框架是: 首先根据既定富营养化风险指标体系生成各个指标下云模型, 再通过各实测水质数据的相依性构造 Gaussian copula 函数, 最后由 copula 函数联接各个云得到多维 copula-云模型

C-C 模型构建与评价的具体实现步骤如下:

步骤1 确定富营养化指标体系。本文参考文献 [19] 中建议的富营养化指标体系, 选取叶绿素 a (Chl-a)、总磷 (TP)、总氮 (TN)、化学需氧量 (COD)、透明度 (SD) 作为 5 个风险评价指标, 并将各指标按富营养化风险由低到高划分为 6 个等级 (I, II, III, IV, V, VI)。

步骤2 生成各风险指标下的云模型。根据云模型参数计算公式 (式 (1)), 由富营养化等级限值确定云模型参数 (Ex, En, He) (表1), 由半云发生器 (式 (2)、(3)) 得到各指标下的半云模型 $U_{Chl-a}(\cdot)$ 、 $U_{TP}(\cdot)$ 、 $U_{TN}(\cdot)$ 、 $U_{COD}(\cdot)$ 、 $U_{SD}(\cdot)$ 。

步骤3 构建多维 copula-云模型。根据水质数据的相关系数矩阵得到相应 Gaussian copula 函数 C

(式 (5)), 联结各指标下的云模型得到多维 copula-云模型 $C(U_{Chl-a}, U_{TP}, U_{TN}, U_{COD}, U_{SD})$ 。

步骤4 富营养化风险评价。将水体水质实测值 $X = (x_1, \dots, x_5)$ 代入多维 copula-云模型得到确定度 $CE = C[U_{Chl-a}(x_1), U_{TP}(x_2), U_{TN}(x_3), U_{COD}(x_4), U_{SD}(x_5)]$ 。通过比对 CE 与各等级所对应的标准限值确定最终富营养化风险等级 L_{CE} , 评价结束。

3. 实例应用

3.1. 评价结果

以我国 12 处代表性湖库为例 [19], 验证 C-C 富营养化风险评价模型的适用性。将水质实测值带入各风险指标维度上的云模型, 得到湖库在各风险指标维度上的确定度分布 (见图 3)。

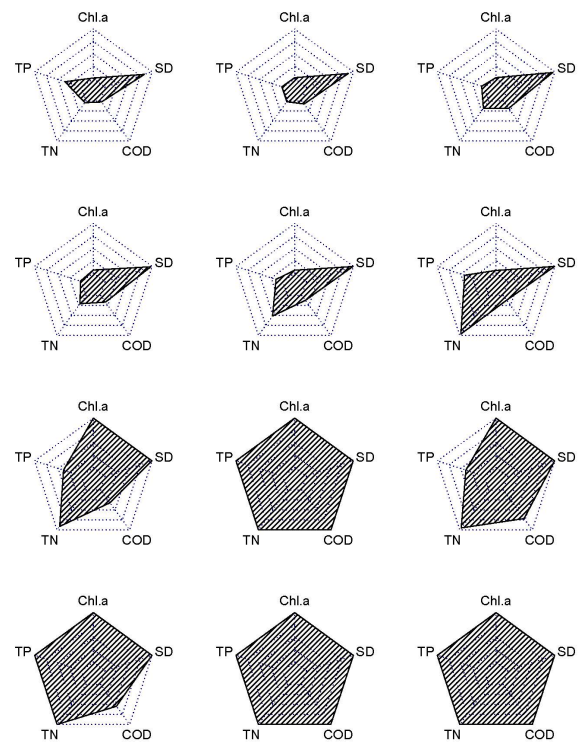


图 3 湖库各风险指标确定度雷达图 (左上至右下依此: 洱海, 博斯腾湖, 于桥水库, 磁湖, 巢湖, 甘棠湖, 蘑菇湖, 西湖, 玄武湖, 墨水湖, 东山湖)

运用 C-C 模型进行评价时, 首先将各风险指标的等级限值代入多维 copula-云模型, 得到不同风险等级下确定度的标准限值; 其次, 将各水体实测

值代入多维 copula-云模型得到富营养化确定度 CE, 结果详见表 2。

表 2 各湖库富营养化确定度及风险等级

湖库	$U_{C_{chl-a}}$	U_{TP}	U_{TN}	U_{COD}	U_{SD}	CE	L_{CE}
邛海	0.0123	0.3546	0.0431	0.0266	0.8373	0.0123	III
洱海	0.0184	0.0229	0.0207	0.0751	0.8912	0.0184	III
博斯腾湖	0.0196	0.0561	0.1850	0.1845	0.9583	0.0196	IV
于桥水库	0.0629	0.0270	0.2066	0.1663	0.9734	0.0629	V
磁湖	0.0582	0.1476	0.5150	0.1205	0.9916	0.0582	V
巢湖	0.0538	0.4184	0.9626	0.1174	0.9985	0.0538	V
甘棠湖	1.0000	0.3859	0.9109	0.3205	0.9974	1.0000	V
蘑菇湖	1.0000	1.0000	1.0000	1.0000	0.9952	1.0000	VI
西湖	1.0000	0.3923	0.9489	0.7153	0.9973	1.0000	V
玄武湖	1.0000	1.0000	1.0000	0.5452	0.9981	1.0000	VI
墨水湖	1.0000	1.0000	1.0000	0.9963	0.9996	1.0000	VI
东山湖	1.0000	1.0000	1.0000	1.0000	0.9986	1.0000	VI

(注: 各等级所对应确定度标准限值为I[0, 0.0026]; II[0.0026, 0.0029]; III[0.0029, 0.0063]; IV[0.0063, 0.0162]; V[0.0162, 0.6570]; VI[0.6570, 1.0000])

3.2. 结果对比

为进一步验证评价结果的有效性, 选取单一指标法、文献[19]、云模型法^[7]等现有评价方法进行对比, 各方法的评价结果见表 3。

表 3 水体富营养化风险评价结果对比

湖库	单一指标	文献[19]	云模型法	C-C 模型
邛海	V	III	IV	III
洱海	IV	III	III	III
博斯腾湖	V	III	IV	IV
于桥水库	V	IV	IV	V
磁湖	V	V	V	V
巢湖	VI	V	V	V
西湖	VI	V	V	V
甘棠湖	VI	V	V	V
蘑菇湖	VI	V	VI	VI
玄武湖	VI	VI	VI	VI
墨水湖	VI	VI	VI	VI
东山湖	VI	VI	VI	VI

由表 3, 将 C-C 模型评价结果与几种参照方法进行对比分析, 在所有的 12 个评价实例中, C-C 模型的评价等级均是各种方法结果的众数。这表明 C-C 模型最能代表各方法的评价结果, 即评价结果具有较强的代表性。此外, 评价等级反映出水体富营养化风险的高低。由表中可得, 虽然在某些评价例中各方法的评价结果有所出入, 但整体上能够反映出各湖库水体的富营养化情况, 即由上至下富营养化

风险逐渐加剧, 这与实际水质景观和用水情况相一致^[19]。

4. 结论

本文针对风险评价中的不确定性(随机性、模糊性)以及多指标映射问题, 引入云模型和 copula 多元分析理论, 提出了一种简化的 C-HC 富营养化评价模型。首先, 云模型刻画了富营养化评价中的模糊性和随机性, 半云发生器的引入使得一个指标下的富营养化概念只需要一个云模型与之对应, 简化了云模型的研建; 其次, copula 函数刻画了不同评价指标间的相依特性, 通过 copula 函数连接各指标下的云模型即可得到多维 copula-云评价模型。实例证明 C-HC 模型评价结果直观有效, 该模型亦可应用于其他相关风险评价领域。

Copula 函数理论与云模型的耦合是一个崭新的研究方向, 本文以 Gaussian copula 为例探讨了 copula-云耦合建模在水体富营养化风险评价中的适用性。对于 copula 函数类型、云模型映射模式等方面的研究值得进一步开展。

致谢

资助项目: 教育部博士点基金(20120091110026, 20100091120059)、江苏省教育厅青蓝工程。

参考文献

1. Eutrophication of waters: Monitoring, assessment and control. Organization for Economic Co-operation and Development, 1982.
2. Zhang B., Liu J. L. Evaluation method for lake eutrophication influence and public satisfaction. *Advances in Water Science*, 2009, 20(5): 695-700.
张宝, 刘静玲. 湖泊富营养化影响与公众满意度评价方法. *水科学进展*, 2009, 20(5): 695-700.
3. Li Z. Y., Guo C., Wang J. Y., et al. Generalized expression of potential function in catastrophe model and its application in eutrophication evaluation. *Advances in Water Science*, 2010, 21(1): 101-106.
李祚泳, 郭淳, 汪嘉杨, 等. 突变模型势函数的一般表示式及用于富营养化评价. *水科学进展*, 2010, 21(1): 101-106.
4. Li W. J., Qiu L., Chen X. N., et al. Assessment model for river ecology health based on Set Pair Analysis and Variable Fuzzy Set. *Journal of Hydraulic Engineering*. 2011, 42(7): 775-782.
李文君, 邱林, 陈晓楠, 等. 基于集对分析与可变模糊集的河流生态健康评价模型. *水利学报*, 2011, 42(7): 775-782.
5. Yin X., Li R. Z., Yang J. W., et al., Eutrophication evaluation model for lake and reservoirs based on extended-blind number. *Acta Scientiae Circumstantiae*. 2014, 34(4): 1045-1053.
尹星, 李如忠, 杨继伟等. 基于延拓盲数的湖库水体富营养化评价模型. *环境科学学报*, 2014, 34(4): 1045-1053.
6. Ding H., Wang D. The evaluation method of water eutrophication based on cloud model. *Acta Scientiae Circumstantiae*. 2013, 33(1): 251-257.
丁昊, 王栋. 基于云模型的水体富营养化程度评价方法. *环境科学学报*, 2013, 33(1): 251-257.
7. Liu D. F., Wang D., Ding H., et al. Eutrophication assessment by entropy-cloud model. *Journal of Hydraulic Engineering*, 2014, 45(10): 89-98.
刘登峰, 王栋, 丁昊, 等. 水体富营养化评价的熵-云耦合模型. *水利学报*, 2014, 45(10): 89-98.
8. Favre A C, El Adlouni S, Perreault L, et al. Multivariate hydrological frequency analysis using copulas. *Water resources research*, 2004, 40(1).
9. Genest C, Favre A C, Béliveau J, et al. Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resources Research*, 2007, 43(9).
10. Salvadori G, De Michele C. Multivariate multiparameter extreme value models and return periods: A copula approach. *Water resources research*, 2010, 46(10).
11. Vandenberghe S, Verhoest N E C, De Baets B. Fitting bivariate copulas to the dependence structure between storm characteristics: A detailed analysis based on 105 year 10 min rainfall. *Water resources research*, 2010, 46(1).
12. Hao Z, Singh V P. Modeling multisite streamflow dependence with maximum entropy copula. *Water Resources Research*, 2013, 49(10): 7139-7143.
13. Zhang X., Ran Q. X., Xia J., et al., Jointed distribution function of water quality and water quantity based on Copula. *Journal of Hydraulic Engineering*. 2011, 42(4): 483-489.
张翔, 冉啟香, 夏军, 等. 基于 Copula 函数的水量水质联合分布函数. *水利学报*, 2011, 42(4): 483-489.
14. Chen J., Wang W. S., Li Y. Q., Copula assessment method and its application for eutrophication of lake water quality assessment. *Journal of Sichuan University (Engineering Science Edition)*, 2011, 1.
陈晶, 王文圣, 李跃清. Copula 评价法及其在湖泊水质富营养化评价中的应用. *四川大学学报 (工程科学版)*, 2011, 1.
15. Li D, Liu C, Gan W. A new cognitive model: cloud model. *International Journal of Intelligent Systems*, 2009, 24(3): 357-375.
16. Liu D, Wang D, Wu J, et al. A risk assessment method based on RBF artificial neural network-cloud model for urban water hazard. *Journal of Intelligent and Fuzzy Systems*, 2013, 27(5): 2409-2416.
17. Li D.Y., Du Y. *Artificial Intelligence with Uncertainty*. Chapman & Hall/CRC, 2007.
18. Shu J. H. The assessment methodology for eutrophication level of lakes in China. *Environmental Pollution & Control*, 1990, 12(5): 2-7.
舒金华. 我国湖泊富营养化程度评价方法的探讨. *环境污染与防治*, 1990, 12(5): 2-7.
19. Liu D. F., Wang D., Ding H., et al. Gamma-cloud for eutrophication evaluation. *Journal of China Hydrology*. 2014, 36(1):8-11.
刘登峰, 王栋, 丁昊, 等. 一种基于 Gamma-云的富营养化评价方法及其应用. *水文*, 2014, 34(5): 45-50.
20. Sklar A. Fonctions de repartition à n dimensions et leurs marges. *Publication de l'Institut de Statistique de l'Université de Paris*, 1959, 8:229-231.
21. Joe, H., *Dependence Modeling with Copulas*. CRC Monographs on Statistics & Applied Probability. Chapman & Hall, London, 2014.
22. Nelsen, R. B., *An introduction to copulas*. Springer, 1999.
23. Renard B, Lang M. Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. *Advances in Water Resources*, 2007, 30(4): 897-912.