

DOI:10.3876/j.issn.1000-1980.2016.04.015

基于 Copula 熵理论的大坝渗流统计模型因子优选

李小奇^{1,2}, 郑东健^{1,2}, 鞠宜朋³

(1.河海大学水文水资源与水利工程科学国家重点实验室,江苏 南京 210098;

2.河海大学水资源高效利用与工程安全国家工程研究中心,江苏 南京 210098;

3.中国电子科技集团公司第十五研究所,北京 100083)

摘要: 针对大坝渗流统计模型需要考虑较多的前期项,造成参选的因子数量较大,进而导致常规方法的建模误差较高的问题,研究引入 Copula 熵理论,利用 Copula 熵和偏互信息(partial mutual information, PMI)相结合的方法,对输入因子的选取进行优化。针对 Copula 熵的求取, Copula 函数采用 Gumbel 函数,分布采用柯西分布代替正态分布,并引入 Hamble 准则来精确选取因子。将该方法在糯扎渡大坝渗流监测中进行应用,并与常规的因子选择方法进行对比分析,结果表明,采用基于 Copula 熵的因子优化选取方法的渗流统计模型具有更好的预测效果。

关键词: 大坝安全监控; 渗流统计模型; Copula 熵; 输入因子; 偏互信息

中图分类号: TV64; TV698

文献标志码: A

文章编号: 1000-1980(2016)04-0370-07

Input factor optimization study of dam seepage statistical model based on copula entropy theory

LI Xiaoqi^{1, 2}, ZHENG Dongjian^{1, 2}, JU Yipeng³

(1. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China;

2. National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Hohai University, Nanjing 210098, China;

3. The Fifteenth Research Institute, China Electronic Technology Group Corporation, Beijing 100083, China)

Abstract: In order to avoid the conventional method's requirement of selecting large quantities of input factors as well as its large errors in development of a dam seepage statistical model, problems caused by the need for many items to be considered in the earlier stage of model development, the copula entropy theory combined with partial mutual information was used to optimize the input factor selection. To obtain the copula entropy, the Gumbel function was used as the copula function, the Cauchy distribution was used to replace the normal distribution, and the Hamble criterion was used to select the input factors accurately. This approach was applied to seepage detection for the Nuozhadu Dam. Comparison of the present results with those obtained from the conventional factor selection approach shows that the seepage statistical model for optimizing input factor selection based on the copula entropy has a better prediction effect.

Key words: dam safety monitoring; seepage statistical model; copula entropy; input factor; partial mutual information

随着大坝安全监测对预测模型的精度要求越来越高,影响模型精度的预报因子选取日益重要。目前,渗流统计模型预报因子的选择最常用的是线性相关系数^[1]、主成分分析、灰色关联度分析等方法。表1列举了渗流统计模型因子的几种主要选择方法,并对其优劣进行了简要分析。

收稿日期: 2015-11-29

基金项目: 国家自然科学基金(51279052); 水文水资源与水利工程科学国家重点实验室项目(20145028312)

作者简介: 李小奇(1986—),男,山东青州人,博士研究生,主要从事大坝安全监控及健康诊断研究。E-mail: lxq0920@gmail.com

表 1 渗流统计模型因子选择的几种主要方法
Table 1 Several methods for input factor selection of seepage statistical model

筛选方法	基本原理	参数	优势	劣势
先验知识	利用经验知识对因子进行优选		方便快捷	过于依赖经验
简单线性相关	依据简单的变量线性相关进行因子挑选	相关系数	简便准确	不能反映变量间的非线性关系
逐步回归	根据因变量的显著程度依次引入回归方程	显著性 F 检验	计算量小	只能识别一个回归因子集合
主成分分析	将相关的多变量进行降维处理,选取贡献率高的不相关指标作为主成分	特征值	无回归计算,避免了因子相关性影响	因子共线性较差时指标计算精度不高
关联度分析	计算因子间的灰色关联度,根据关联度大小选择建模因子	灰关联系数	避免了原始信息量不足的影响	没有明确的评判指标来引入因子
模糊隶属度	计算因子的隶属度值,根据一定评判指标选取建模因子	关系矩阵	避免了因子认知不足的影响	隶属度函数选择和评判指标不明确
AIC 准则	选择使 AIC 信息量值最小的因子组合	残差平方和	原理简单,计算便捷	因子组合数量过多时,计算繁琐

先验知识方法最为简便,但误差也最大。逐步回归分析中,由 F 检验剔除的因子有可能组成最佳因子集,并且在因子集进行联合分析时只能识别一个因子集合^[2]。主成分分析法对相关变量的要求较高,当变量的共线性较差时,降维处理得到的综合指标精度也较低^[3]。利用关联度进行因子选择时,只能将因子按照关联度值进行排序,没有具体引入指标^[2]。模糊隶属度函数选择对样本分布的依赖性较大,隶属度函数选择对结果精度的影响较高^[4]。AIC 信息量准则法中,计算量与因子组合数量成正比,当因子组合过多时工作量巨大^[5]。鉴于以上不足,需要提出一种新的理论来满足大坝安全监控中对渗流统计模型因子选择的需求。

目前,一种可供选择的方法是互信息(mutual information, MI)法以及 MI 改进算法^[6]。赵铜铁钢等^[7]利用 MI 法选择了神经网络模型的输入因子。卢迪等^[8]利用互信息法对 BP 网络模型的中长期径流预报因子进行了筛选。Sharma^[9]改进了 MI 法,提出偏互信息(partial mutual information, PMI)的概念,并得到了广泛应用。Kan 等^[10]利用 PMI 法对集合神经网络(ENN)的径流预报因子进行了优化选取。Yuan 等^[11]利用 PMI 法对时间序列模型的输入因子进行了选取。应用和研究证明了 PMI 法可以有效提高模型因子选择的精度。然而,PMI 法仍有缺陷:(a)水位、位移、沉降等均为连续变量,而 PMI 主要基于离散数据;(b)对多元分布变量的边缘和联合概率密度进行估计时,概率密度函数难以求得^[6-7]。

针对上述不足,本文提出了 Copula 熵和 PMI 相结合的方法。在已有的大坝渗流统计模型基础上,通过专家经验确定备选因子集,采用 Copula 函数和信息熵结合的 Copula 熵法计算 PMI 值,然后基于 PMI 的选入准则,对大坝渗流影响量的实测数据进行计算,确定最终的输入因子集。最后,将此方法在糯扎渡大坝的渗流监测中进行了应用和对比。

1 基本原理

1.1 渗流统计模型方程

渗流量的大小反映了混凝土坝坝体和地基帷幕以及土石坝防渗体的防渗效果,是评价大坝运行安全的重要依据。大坝渗流主要受上下游水位、降雨和温度的影响,模型方程表示为^[12]

$$Q = Q_{SH} + Q_{XH} + Q_R + Q_P + Q_\theta \quad (1)$$

式中: Q ——渗流量实测值; Q_{SH} ——上游水位分量; Q_{XH} ——下游水位分量; Q_R ——温度分量; Q_P ——降雨分量; Q_θ ——时效分量。

通常 Q_{SH} 和 Q_P 对渗流量都有滞后效应,因此需要考虑滞后天数的影响。常规模型中,混凝土坝坝基渗漏量主要与 H_{SH} 、 H_{SH}^2 和 H_{XH} 有关,土石坝渗流主要与 H_{SH} 和 H_{XH} 有关。另外,土石坝由于坝体材料随温度变化较小,渗流计算通常不考虑 Q_R 分量的作用;混凝土坝的温度以周期项作为因子。

1.2 Copula 函数

1959 年,Sklar 提出可以将一个联合分布分解为 n 个边缘分布和一个 Copula 函数,这个 Copula 函数描述了变量间的相关性^[13]。Copula 函数实际上就是一类将联合分布函数与它们各自的边缘分布函数连接在一

起的函数,自变量的定义域为 $[0,1]$ 。以公式形式表述为

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (2)$$

式中: F —— n 维变量的联合累积分布函数; $x_1, \dots, x_n$ —— n 个随机变量; F_i —— 各变量的边缘分布函数。

如果 $F_1 \dots F_n$ 是连续的,那么 C 是唯一的;反之 C 不唯一。

在多变量的研究中, Archimedean 型 Copula 函数具有模型构造方便和计算简单的优点,因此应用广泛^[14]。本研究采用 Archimedean 型最常用的 Gumbel Copula 函数来进行计算,它主要描述例如某变量剧增时,效应量会发生同样剧增情形时的正相关特性。为了便于计算,本文一律取二维变量的情形,它的二维联合分布函数和密度函数如下:

$$C(u, v) = \exp \left\{ - \left[(-\ln u)^\theta + (-\ln v)^\theta \right]^{\frac{1}{\theta}} \right\} \quad (3)$$

$$c(u, v) = C(u, v) \frac{\left[(-\ln u)(-\ln v) \right]^{\theta-1}}{uv} \left[(-\ln u) + (-\ln v)^\theta \right]^{\frac{2}{\theta}-2} \left\{ (\theta-1) \left[(-\ln u)(\ln v) \right]^{\theta-1} + 1 \right\} \quad (4)$$

式中: u —— 变量 x 的边缘分布 $F(x)$; v —— 变量 y 的边缘分布 $F(y)$; θ —— Gumbel Copula 函数的参数,在 $[1, +\infty)$ 内取值。

1.3 Copula 熵

熵理论是 1948 年 Shannon 将玻尔兹曼熵的概念引入信息论中,用来度量随机事件的不确定性或信息量的方法^[15]。大坝监测信息多为连续随机变量, n 维随机变量的联合熵可以表示为

$$H(x_1, \dots, x_n) = - \int_0^\infty \dots \int_0^\infty f(x_1, \dots, x_n) \lg [f(x_1, \dots, x_n)] dx_1 \dots dx_n \quad (5)$$

式中: $H(x)$ —— n 维连续性随机变量的熵函数; $f(x)$ —— 变量 x 的概率密度函数,可用分布函数的偏微分来表示。

若式(3)中的 $u = F(x)$, $v = F(y)$, 则 Copula 熵的表达式为

$$H_c(u, v) = - \int_0^1 \int_0^1 c(u, v) \lg c(u, v) du dv \quad (6)$$

对 Copula 熵的求解,转化为求变量 x 和 y 的边缘分布 F 和 Copula 联合分布的参数 θ 的过程。 θ 可以通过 Kendall 系数来求取^[16],它与 Kendall 系数 τ 的关系为

$$\theta = \frac{1}{1 - \tau} \quad (7)$$

当 $\theta = 1$ 时, u, v 相互独立;当 $\theta \rightarrow \infty$ 时, u, v 趋于完全相关。

本文选取柯西分布函数作为随机变量 ξ 的边缘分布函数,与正态分布相比,柯西分布产生的随机点具有较广泛的覆盖范围,能真实地反映随机向量的分布情况,同时能较快摆脱局部最优值的现象^[17]。柯西分布的概率密度函数和分布函数分别为

$$f(x; \mu, \gamma) = \frac{1}{\pi} \left(\frac{\gamma}{x - \mu + \gamma^2} \right) \quad (8)$$

$$F(x; \mu, \gamma) = \frac{1}{\pi} \arctan \frac{x - \mu}{\gamma} + \frac{1}{2} \quad (9)$$

式中: μ —— 分布峰值位置的位置参数; γ —— 最大值 1/2 处的 1/2 宽度的尺度参数。

对 μ 采用中位数作为估计量,对 γ 采用分位数作为估计量^[18]。假设求得 μ 的中位数为 $\mu_{n/2}$, 则 $\gamma = \mu_{3n/4} - \mu_{n/2}$ 。

求得单变量的边缘分布函数 $F(x)$ 的参数 μ, γ 后,代入式(3)的联合分布函数 $C(u, v)$,再根据求得的 θ ,即可得到 Gumbel Copula 分布函数和密度函数。然后将联合分布密度函数带入式(6),求得 Copula 熵的值。

1.4 Copula 熵的 PMI 计算

偏互信息(PMI)是 Sharma 提出的,它度量了在消除多余变量影响的条件下,某两变量之间的相关性。

通过计算某两变量的 C_{PMI} 值,设定一定的过滤标准,即可得到筛选后的输入因子集合。以二维情形计算,假设两变量为 x, y , 由偏互信息的定义可知:

$$C_{\text{PMI}} = H(x') + H(y') - H(x', y') = \int_0^{\infty} \int_0^{\infty} f(x', y') \ln \frac{f(x', y')}{f(x')f(y')} dx' dy' \quad (10)$$

其中

$$x' = x - E[x|Z_2] \quad y' = y - E[y|Z_2]$$

式中: x' —— x 的残余信息; y' —— y 的残余信息; Z_2 —— 二级备选因子集合。

式(10)中条件期望 $E[x|Z_2]$ 的求取可以采用式(11)^[9], $E[y|Z_2]$ 的求取方式相同。

$$E[x_i|Z_{2i}] = \frac{1}{N} \sum_{j=1}^N \alpha_j \left[x_i + (Z_{2i} - Z_{2j})^{\xi} S_{zz}^{-1} S_{xz} \right] \quad (11)$$

式中: N —— 样本数; α_j —— 平滑参数,本文取高斯核密度估计的窗宽进行计算^[9]; ξ —— 高斯分布的参数; S_{zz} —— Z_2 中变量 Z_{2i} 样本的协方差; S_{xz} —— 变量 x 和 Z_2 中变量 Z_{2i} 样本的互协方差。

根据多维联合熵的定义,多维变量的联合熵可以用 n 个变量的边缘熵的和与 Copula 函数熵的和来表达。由于 $du = dx f(x)$, $dv = dy f(y)$, $f(x, y) = c(u, v) f(x) f(y)$, 二维联合熵的表达式可以写为

$$H(X, Y) = - \int_0^{\infty} \int_0^{\infty} f(x, y) \ln f(x, y) dx dy = H_c(u, v) + H(X) + H(Y) \quad (12)$$

对于本文二维变量(x', y')的情形,式(12)可以写为

$$H(x', y') = H(x') + H(y') + H_c(u', v') \quad (13)$$

式(13)代入式(10)可知, $C_{\text{PMI}} = -H_c(u', v')$, 即偏互信息为 Copula 函数的负熵, u' 为变量 x' 的分布函数, v' 为变量 y' 的分布函数。通过 Copula 熵计算 PMI 值的过程如图 1 所示。

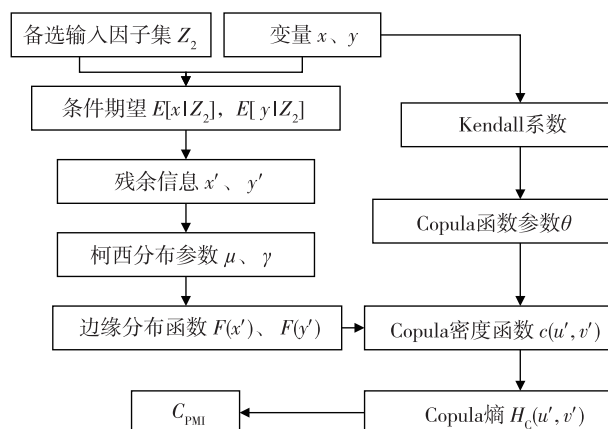


图 1 C_{PMI} 值计算过程

Fig. 1 Calculation process of C_{PMI}

1.5 因子选入准则

PMI 算法需要给定一个标准,来判定 C_{PMI} 多大时将变量 x 纳入模型的输入因子集。本文采用 Fernando 和 May 等推荐的 Hampel 检验作为算法的停止准则^[19-20],其表达式为

$$H_j = \frac{d_j}{1.4826 d_j^{(50)}} \quad d_j = |C_{\text{PMI}} - C_{\text{PMI}}^{(50)}| \quad (14)$$

式中: H_j —— Hampel 距离; 1.4826 —— 归一化因子,使得 H_j 等于数据序列的标准差 σ ; $d_j^{(50)}$ —— d_j 的中位数; $C_{\text{PMI}}^{(50)}$ —— 一组数据中 C_{PMI} 值的中位数。

根据标准差的 3σ 准则,当 $H_j > 3$ 时,因子被选入。

2 实例分析

2.1 概况

土石坝渗流一直是坝体安全监测的关键指标,本研究以糯扎渡黏土心墙堆石坝为例。糯扎渡大坝是世界同类型中第三高坝,位于中国云南省澜沧江上,最大坝高 261.5 m,坝顶长度 627.87 m,坝顶宽

度 18 m,水库正常蓄水位 812 m。坝体于 2008 年 11 月开始填筑,2012 年 12 月 21 日填筑到坝顶。蓄水初期水位变化较大,加上坝体填筑过程中的数据缺失以及外部环境的改变,对坝体安全监测模型的因子选择造成了障碍。因此,考虑利用 Copula 熵和 PMI 法相结合的理论进行计算。心墙渗流是土石坝渗漏监测的重点,这里以蓄水期最关键的心墙中下部 617m 高程的 DB-C-P-14~DB-C-P-17 渗压计测点为例,取 2011-11-30—2013-01-18 时段内的数据进行分析。为了叙述方便,以 DB-C-P-14 测点的计算过程为例进行演示。

2.2 备选因子

通过资料分析和知识经验,考虑温度对坝体材料的影响,引入上游水位 H_{SH} 、下游水位 H_{XH} 、温度 T 、降雨 P 和时效 θ 等作为备选因子来进行筛选,具体结果见表 2。

表 2 二级输入因子集 Z_2 统计表
Table 2 Statistical table of second-level input factor set Z_2

序号	变量符号	参数值	分量名称	Pearson 相关系数	Kendall 相关系数	是否选入 Z_2
1	$\sin \varepsilon$	$\alpha = \frac{2\pi i t}{365}$	温度分量	-0.2478	-0.2351	是
2	$\cos \varepsilon$			0.1635	0.2351	是
3	$\sin^2 \varepsilon$			-0.1977	-0.2185	是
4	$\cos^2 \varepsilon$			0.1553	0.2185	是
5	H_{SH}^1	$H_{SH}^i = (H_{SH})^i$	上游水位	0.8048	0.7531	是
6	H_{SH}^2			0.7876	0.7531	是
7	H_{SH}^3			0.7805	0.7531	是
8	$\bar{H}_{3SH}$	$\bar{H}_{iSH} = \frac{1}{i} \sum_{i=3}^5 H_{iSH}$	上游水位前期项	0.7990	0.7506	是
9	$\bar{H}_{4SH}$			0.7996	0.7503	是
10	$\bar{H}_{5SH}$			0.7998	0.7495	是
11	H_{XH}		下游水位	0.6836	0.5814	是
12	P	$\bar{P}_i = \frac{1}{i} \sum_{i=3}^5 P_i$	当日降雨量	0.0635	0.0438	否
13	$\bar{P}_3$		前期降雨量	0.0536	0.0358	否
14	$\bar{P}_4$			0.0427	0.0256	否
15	$\bar{P}_5$			0.0359	0.0244	否
16	η	$\eta = \frac{d_i - d_0}{100}$	时效分量	-0.4463	-0.2216	是
17	η^2			-0.3704	-0.2313	是
18	$\ln \eta$			0.5175	0.2365	是

2.3 因子优选

经过定性分析和相关系数的显著性分析,得到备选因子集合 Z_2 。运用 Copula 熵计算 Z_2 影响因子的 C_{PMI} 值和 Hample 距离,得到最终的输入因子集合 Z_3 。假设效应量为渗流量 S ,这里只对温度分量因子的筛选进行示例,其他分量筛选方式相同。首先选出温度分量中相关性最高的因子,然后将温度分量的预备因子作为一个集合进行分析。取相关性最高的 $\sin \alpha$ 为分析因子, $\{\cos \alpha, \sin^2 \alpha, \cos^2 \alpha\}$ 作为待选因子进行 Hample 检验,得到温度分量的最终输入因子集 $Z_{3R}(S) = \{\sin \alpha, \cos \alpha\}$ 。通过相同方式的 Hample 检验,最终得到了渗流统计模型的输入因子集合 $Z_3(S) = \{\sin \alpha, \cos \alpha, H_{SH}^1, H_{SH}^2, H_{SH}^3, \theta, \ln \theta\}$ 。

2.4 结果对比

为了显示方法的优越性,将利用经验知识、相关性和 Copula 熵得到的输入因子集合进行了对比分析和计算,模型采用逐步回归法,结果如表 3 所示。针对 2011-11-30—2013-01-18 数据进行分析,以 2011-11-30—2013-01-01 作为拟合区间,2013-01-01—2013-01-18 为预测区间,得到拟合结果如图 4 所示,典型日期预测结果如表 4 所示。由图表对比可知,采用 Copula 熵法输入因子的渗流统计模型拟合值更贴近实测值,拟合精度和预测精度都更高。

表 3 DB-C-P-14~DB-C-P-17 测点拟合结果对比
Table 3 Comparison of fitted results at monitoring points DB-C-P-14 through DB-C-P-17

方法	输入因子集	复相关系数 R			
		DB-C-P-14	DB-C-P-15	DB-C-P-16	DB-C-P-17
经验知识	$\sin \alpha, \cos \alpha, H_{SH}, P, \bar{P}_3, \bar{P}_4, \bar{P}_5, \theta, \ln \theta$	0.975	0.978	0.982	0.992
相关性	$H_{SH}, H_{SH}^2, H_{SH}^3, \bar{H}_{3SH}, \bar{H}_{4SH}, \bar{H}_{5SH}, \theta, \theta^2, \ln \theta$	0.912	0.938	0.964	0.982
Copula 熵	$\sin \alpha, \cos \alpha, H_{SH}, H_{SH}^2, H_{SH}^3, \theta, \ln \theta$	0.991	0.991	0.993	0.995

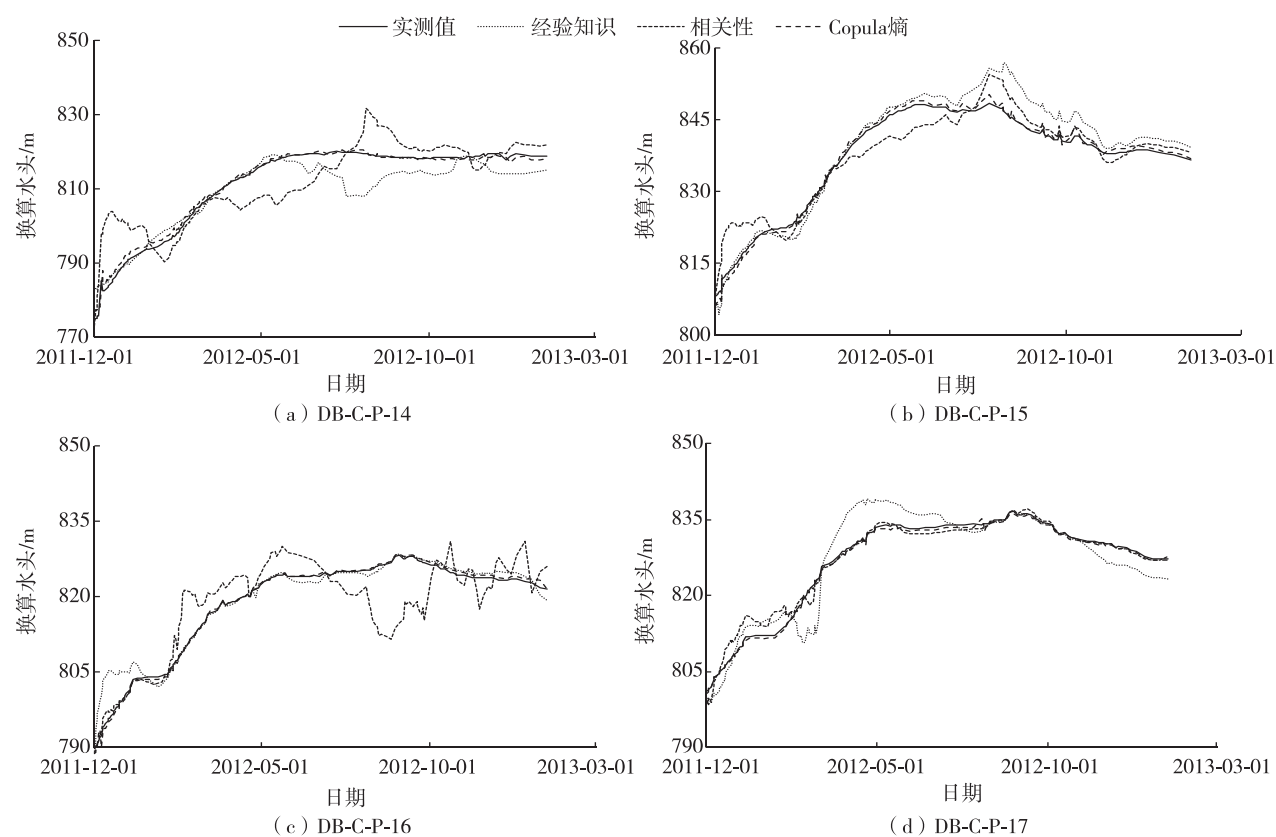


图 2 模型拟合曲线对比

Fig. 2 Comparison of fitted curves by present model for different monitoring points

表 4 DB-C-P-14~DB-C-P-17 测点水头预测结果对比

测点	2013-01-05 水头预测结果				2013-01-10 水头预测结果			
	实测值	经验知识	相关性	Copula 熵	实测值	经验知识	相关性	Copula 熵
DB-C-P-14	818.8	821.8	812.6	820.4	818.9	821.1	813.8	820.2
DB-C-P-15	837.6	840.8	831.2	836.2	837.2	842.4	833.2	836.2
DB-C-P-16	822.7	824.7	817.9	823.1	822.8	825.1	818.6	823.6
DB-C-P-17	826.0	823.2	820.6	824.8	825.6	823.0	820.4	825.8

3 结 论

- a. 依据大坝监测样本,采用 Archimedean 型 Copula 函数进行计算,为了计算简便,采用二维的 Gumbel Copula 函数。利用 Kendall 秩相关系数,求解了 Gumbel Copula 函数的相关参数。
- b. 对于变量的边缘分布,采用柯西分布代替正态分布,并利用中位数和分位数的方法,对柯西分布参数进行求取,从而进一步精确了 Gumbel Copula 联合分布函数。
- c. 引入 PMI 方法,推导得到 Copula 函数的负熵等于偏互信息值。通过 PMI 的计算,并引入 Hample 检验作为停止准则,建立了因子引入的标准和具体步骤。
- d. 建立基于 Copula 熵的因子选择流程,并根据优选后的因子建立了渗流统计模型。通过与常规方法对比,证明新的因子优化方法建立的预测模型效果更好。

参考文献:

[1] 陈西江, 鲁铁定, 谭成芳. 大坝位移 BP 网络模型影响因子的优选[J]. 江西科学, 2010, 28(1):72-76. (CHEN Xijiang, LU Tieding, TAN Chengfang. Optimization for influence factors of dam displacement base on BP neural network model [J].

- Jiangxi Science, 2010, 28(1):72-76. (in Chinese))
- [2] 姚远, 李姝昱, 张博. 逐步回归-PLS 模型在大坝位移监控中的应用[J]. 水电能源科学, 2011, 29(4):81-82. (YAO Yuan, LI Shuyu, ZHANG Bo. Application of stepwise regression-PLS model to dam displacement monitoring [J]. Water Resources & Power, 2011, 29(4):81-82. (in Chinese))
- [3] YU Hong, WU Zhongru, BAO Tengfei, et al. Multivariate analysis in dam monitoring data with PCA [J]. Science China Technological Sciences, 2010, 53(4):1088-1097.
- [4] 王季方, 卢正鼎. 模糊控制中隶属度函数的确定方法[J]. 河南科学, 2000, 18(4):348-351. (WANG Jifang, LU Zhengding. The determine method of membership function in fuzzy control [J]. Henan Science, 2000, 18(4):348-351. (in Chinese))
- [5] HU S. Akaike information criterion [M]. North Carolina State: Center for Research in Scientific Computation, 2007.
- [6] CHEN Lu, GUO Shenglian. Copula entropy and its application in hydrological correlation analysis [J]. Journal of Water Resources Research, 2013, 2(2):103-108.
- [7] 赵钢铁, 杨大文. 神经网络径流预报模型中基于互信息的预报因子选择方法[J]. 水力发电学报, 2011, 30(1):24-30. (ZHAO Tongtiegang, YANG Dawen. Mutual information-based input variable selection method for runoff-forecasting neural network model [J]. Journal of Hydroelectric Engineering, 2011, 30(1):24-30. (in Chinese))
- [8] 卢迪, 周惠成. 基于互信息量与 BP 神经网络的中长期径流预报方法研究[J]. 水文, 2014, 4(4):8-14. (LU Di, ZHOU Huicheng. Medium and long-term runoff forecasting based on mutual information and BP neural network [J]. Journal of China Hydrology, 2014, 4(4):8-14. (in Chinese))
- [9] SHARMA A. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management (part 1): a strategy for system predictor identification [J]. Journal of Hydrology, 2000, 239(1):232-239.
- [10] KAN Guangyuan, YAO Cheng, LI Qiaoling, et al. Improving event-based rainfall-runoff simulation using an ensemble artificial neural network based hybrid data-driven model [J]. Stochastic Environmental Research & Risk Assessment, 2015, 29(5):1-26.
- [11] YUAN Conggui, ZHANG Xinzheng, XU Shuqiong. Partial mutual information for input selection of time series prediction[C]// Northeastern University. Proceedings of the 2011 Chinese Control and Decision Conference. New York: IEEE, 2011:2010-2014.
- [12] 吴中如. 大坝与坝基安全监控理论和方法及其应用[J]. 江苏科技信息, 2005(12):1-6. (WU Zhongru. Dam and dam foundation safety monitoring theory and method and its application [J]. Jiangsu Science and Technology Information, 2005(12):1-6. (in Chinese))
- [13] SKLAR M. Fonctions de répartition à n dimensions et leurs marges [M]. Paris: Université Paris, 1959.
- [14] 赵婷. Copula 理论及其在金融分析上的应用[D]. 长沙:湖南大学, 2011.
- [15] GRAY R M. Entropy and information theory [M]. New York: Springer Science & Business Media, 2011.
- [16] 郭生练, 闫宝伟, 肖义, 等. Copula 函数在多变量水文分析计算中的应用及研究进展[J]. 水文, 2008, 28(3):1-7. (GUO Shenglian, YAN Baowei, XIAO Yi, et al. Multivariate hydrological analysis and estimation [J]. Journal of China Hydrology, 2008, 28(3):1-7. (in Chinese))
- [17] 赵慧. 基于 K-S 检验 copula 分布估计算法中边缘分布的研究[D]. 太原:太原科技大学, 2013.
- [18] 吴庆波, 李再兴, 景平. 一元 Cauchy 分布族中两参数的分位数估计及其性质[J]. 廊坊师范学院学报(自然科学版), 2010, 10(1):8-9. (WU Qingbo, LI Zaixing, JING Ping. Quantile estimators of the two parameters in the univariate Cauchy distribution and their property [J]. Journal of Langfang Teachers College (Natural Sciences), 2010, 10(1):8-9. (in Chinese))
- [19] FERNANDO T, MAIER H R, DANDY G C. Selection of input variables for data driven models: an average shifted histogram partial mutual information estimator approach [J]. Journal of Hydrology, 2009, 367(3):165-176.
- [20] MAY R J, MAIER H R, DANDY G C, et al. Non-linear variable selection for artificial neural networks using partial mutual information [J]. Environmental Modeling & Software, 2008, 23:1312-1326.