

# Archimedean Copulas 在降水变异概率 分布特征研究中的应用<sup>\*</sup>

## ——以新疆福海县为例

李剑锋<sup>1,2</sup>, 张强<sup>1,2</sup>, 陈晓宏<sup>1,2</sup>

(1. 中山大学 水资源与环境系, 广东 广州 510275; 2. 中山大学 华南地区水循环与水安全广东省普通高校重点实验室, 广东 广州 510275)

**摘要:** 探究二维情况下变异不同步, 也就是一个变量发生变异, 而另一个变量不发生变异的情况, 对 Copula 联合分布的影响。在气温变化的环境下, 降水过程发生了显著变异, 而年降水量和年总降水天数的联合分布对长期旱涝情况有重要意义。以 Copula 函数为基础, 计算新疆福海站年降水量和年总降水天数联合分布, 用传统变异分析方法进行变异分析。针对年降水量发生变异, 而年总降水天数没有发生变异的情况, 探讨其 Copula 联合分布的影响。计算结果表明, 年降水变异对年降水量和年总降水天数 Kendall 相关系数  $\tau$  影响不大, 但对 Copula 种类的选择影响较大。

**关键词:** Copula; 变异分析; 年降水量; 年降水天数; 新疆福海县

**中图分类号:** P426.6    **文献标志码:** A    **文章编号:** 1000-811X(2011)03-0001-07

当前以气温上升为主要特征的气候变化显著改变了水循环时空演变特征<sup>[1-2]</sup>, 其主要表现是降水过程发生了显著变异, 而降水的时空变异是导致地表水文水资源时空分布变化的关键因素<sup>[3-4]</sup>, 因而研究降水概率变化特征对区域水资源演变的科学理解与水资源管理具有重要的科学与现实意义。降水过程是包括许多统计要素特征, 如降水量、降水历时与降水强度等, 而传统意义上的单边量分布函数只能描述降水某一要素的特征变化, 实际上, 对某两个或者三个要素的相关结构或者联合分布情况的了解, 对于降水过程的全面理解具有重要意义。国际上已开展了许多相关的研究工作, 如 Yue 采用 Box-Cox 转换方法把两变量正态化, 再用两变量正态模型描述降雨特征量的联合分布<sup>[5]</sup>。另外, Yue 采用二维 Gumbel 逻辑模型分布对降水频率进行分析<sup>[6]</sup>。以往的研究中, 随机变量之间或多或少都需要一些假设。各个降水随

机变量之间需要有相同的边缘分布, 或者假定随机变量符合正态分布, 或者随机变量被转换成符合正态分布。但实际上, 这些假设在某种程度上具有一定的局限性, 主要问题在于降水变量可能服从不同的概率分布, 另外, 区域不同, 降水的概率分布特征也可能不同。

Copula 联合分布计算函数则有效地避免了上述局限性, 因而逐步在水文气象领域获得应用, 虽然当前的 Copula 函数在水文气象领域的应用与研究尚处于初级阶段。Salvadori 等以 Copula 为基础对频率分析的理论方面进行研究<sup>[7]</sup>。Faver 等提出运用 Copula 对极值进行建模的方法<sup>[8]</sup>。在水文研究方面, De Michele 等使用 Copula 研究洪峰与洪量的二维极值分布<sup>[9]</sup>。Zhang 等使用二维 Copula 对降雨资料和洪水资料进行频率分析<sup>[10]</sup>。运用 Copula 能够简单快捷地构造出降水资料的联合分布。但是当前气候变化情景下, 降水时空分布与

\* 收稿日期: 2010-12-11

基金项目: 新疆自治区科技攻关项目(200931105); 中山大学理工科青年教师重点培育计划项目(2009-37000-1132381); 中山大学“百人计划”人才引进项目(37000-3171315); 南京地理与湖泊研究所前沿领域创新研究项目(CXNIGLAS200814; 08SL141001-28005)

作者简介: 李剑锋(1987-), 男, 广东广州人, 硕士研究生, 主要从事生态水文研究、极端降水及旱涝灾害研究。

E-mail: jianfengli.sysu@gmail.com

通讯作者: 张强(1974-), 男, 山东沂水人, 博士, 教授, 博士生导师, 主要从事流域气象水文学研究、旱涝灾害机理、流域地表水文过程及其对气候变化的响应机制与机理以及流域生态需水等领域的研究工作。E-mail: zhangq68@mail.sysu.edu.cn

降水过程发生显著变异,在多降水变量情况下,不同变量的变异情况可能不同,这种变异的同步在运用 Copula 函数分析时会对分析结果造成怎样的影响?基于此,本文探讨两随机变量变异不同步情况下,对 Copula 函数分析结果的影响。

年降水量和年总降水天数联合分布对于研究一个区域的长期旱涝情况有重要意义。由于气候变化原因,年降水量和年总降水天数均可能发生变异,或者其中一个变量发生变异。本文以 Copula 函数为基础,对年总降水量与年总降水日数进行变异分析,探讨两变量变异不同步的情况下,对 Copula 函数分析结果的影响。

## 1 变异点分析方法

### 1.1 单变异点检验方法

Reeves 等对世界气候组织气候指导手册推荐的 8 种主要变异点分析方法进行对比<sup>[11]</sup>,结果表明标准正态齐性检验(Standard normal homogeneity test, SNH test)和非参数标准正态齐性检验(Non-parametric SNH test)对变异点的位置判断比较准确,但当变异点处于序列两端附近时,这两种方法的判断效果较低。Reeves 对标准正态齐性检验作了改进,其实质是两样本的  $t$  检验。Wilcoxon 秩和检验和 Mann-Whitney 检验是非参数标准正态齐性检验。本文基于 Mann-Whitney U 检验法<sup>[12-13]</sup>和由 Reeves 改进后的  $t$  检验<sup>[14]</sup>进行变异点分析。

### 1.2 多变异点分析方法

Hawkins 于 2001 年提出一种连续假设检验算法<sup>[15]</sup>。该方法把多变异点计算分为两步:拆分计算和整合计算。拆分计算后进行整合计算,检验前一步计算的变异点在整合计算过程是否会失去其显著性。

本文根据研究需要,对 Hawkins 连续假设检验算法改进。在整合计算步骤中,子序列分为 1 到  $C_2$ ,  $C_1 + 1$  到  $C_3 \cdots \cdots C_{k-2}$  到  $n$ ,对子序列逐个检验,可能出现三种检验结果:接受  $H_0$ 、拒绝  $H_0$  且变异点没变、拒绝  $H_0$  且变异点改变。根据以上情况,本文的多变异点分析方法如下:(1)拆分计算。拆分计算中,子序列分为 1 到  $C_1$ ,  $C_1 + 1$  到  $C_2 \cdots \cdots C_{k-1} + 1$  到  $n$ ,如果任意一个子序列的检验结果拒绝  $H_0$ ,那么就把这个子序列拆分,确定出这个子序列的变异点, $k$  增加;(2)整合计算。子序列分为 1 到  $C_2$ ,  $C_1 + 1$  到  $C_3 \cdots \cdots C_{k-2}$  到  $n$ 。任意

子序列的检验结果接受  $H_0$ ,子序列中没有变异点,则删除该子序列中原来的变异点,该子序列整合成功, $k$  减小。如检验 1 到  $C_2$  子序列,检验结果接受  $H_0$ ,则删除原来的变异点  $C_1$ 。检验结果拒绝  $H_0$  且子序列中变异点没有改变,则保留该子序列中原来的变异点。如检验 1 到  $C_2$  子序列,检验结果拒绝  $H_0$ ,且变异点仍然是  $C_1$ ,则保留变异点  $C_1$ 。任意子序列的检验结果拒绝  $H_0$  且子序列中变异点发生改变,则以新的变异点替换原来的变异点,如检测 1 到  $C_2$  子序列,检验结果拒绝  $H_0$ ,变异点为  $C'_1$ ,则变异点更改为  $C'_1$ ,子序列分为 1 到  $C_2$ ,  $C'_1 + 1$  到  $C_3 \cdots \cdots C_{k-2} + 1$  到  $n$ ,继续进行整合计算;(3)拆分计算后进行整合计算,检验前一步计算的变异点在整合计算过程是否会失去其显著性或发生改变。

本文采用  $t$  检验和 Mann-Whitney U 检验的连续假设检验方法。采用这两种方法的连续假设检验结果进行比较,如果点  $c$  都被检验为变异点,或者变异点差异在两个时段内时,才确定其为变异点。

## 2 Copula 函数

Copula 函数可以用不同边缘分布的变量来构造联合分布。以二维随机变量为例。假设二维随机变量  $X$  和  $Y$ ,它们的边缘分布函数是  $F(x) = P[X \leq x]$  和  $G(y) = P[Y \leq y]$ ,则它们的联合分布为  $H(x, y) = P[X \leq x, Y \leq y]$ 。则存在 Copula 函数  $C$  使得

$$H(x, y) = C(F(x), G(y))。 \quad (1)$$

Copula 函数  $C$  本质上是边缘分布为  $F(x)$  和  $G(y)$  的随机变量  $X, Y$  的二元联合分布函数。同理 Copula 可推广至二维以上的多维联合分布<sup>[16]</sup>。

Copula 函数主要类型有 Archimedean Copulas, Elliptical Copulas, Marshall-Olkin Copulas<sup>[17]</sup>。

### 2.1 Archimedean Copulas 介绍

Archimedean Copulas, 是一类非常重要的 Copula, 具有构造方便、容易使用的特点,在各种领域都得到广泛应用。Archimedean Copulas 具有以下几种类型: Gumbel-Hougaard Copula、Ali-Mikhail-Haq Copula、Clayton Copula、Frank Copula<sup>[18]</sup>。

假设随机变量  $X, Y$ , 它们的边缘分布函数是  $F(x)$  和  $G(y)$ 。令  $U = F(x)$  和  $V = G(y)$ , 那么  $U \in [0, 1], V \in [0, 1]$ 。 $u, v$  是  $U$  和  $V$  的某个值。

令  $\varphi$  是连续严格递减的函数, 且  $\varphi(1) = 0$ 。定义  $\varphi^{[-1]}$  是  $\varphi$  的伪逆函数。若  $\varphi(0) = \infty$ , 则  $\varphi^{[-1]} = \varphi^{-1}$ 。则形式如下的 C 函数为 Archimedean Copula。

$$C(u, v) = \varphi^{[-1]}[\varphi(u) + \varphi(v)]。 \quad (2)$$

对于每一种 Archimedean Copula, 都有一个参数  $\theta$  隐含在 Copula 生产元  $\varphi$  中<sup>[17]</sup>。

## 2.2 经验 Copula

Deheuvels 提出经验 Copula<sup>[19]</sup>, 其中比较简单的是二维情况。 $\{(x_k, y_k)\}_{k=1}^n$  是长度为  $n$  的连续两变量分布样本。则经验 Copula  $C_e$  计算公式如下

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{x_k < x_i \text{ 且 } y_k < y_j \text{ 的 } (x, y) \text{ 的个数}}{n}, \quad (3)$$

式中:  $x_i$  和  $y_j$ ,  $1 \leq i, j \leq n$ , 表示从样本中来的次序统计<sup>[20]</sup>。

## 2.3 AIC 准则法

Akaike 提出 Akaike information criterion (AIC 准则法)<sup>[21]</sup>, 本文运用该法检验各种 Copula 分布的拟合程度。AIC 包括两个部分: 函数拟合的偏差和参数数量带来的不可靠性。AIC 的计算过程如下

$$AIC = n \log(RSS/n) + 2m, \quad (4)$$

式中:  $n$  是样本个数;  $m$  是参数个数;  $RSS$  是残差平方和。AIC 值越小, 说明函数拟合程度越好。

## 2.4 Copula 函数非参数估计

Genest 和 Rivest (Genest, et al, 1993) 提出 Copula 函数的非参数估计方法, 本文根据该非参数估计方法计算 Copula 函数。方法如下:

(1) 根据下式估算随机样本的 Kendall 系数  $\tau$

$$\tau = \frac{\sum_{i < j} \text{sign}[(x_i - x_j)(y_i - y_j)]}{\frac{n(n-1)}{2}}, \quad (5)$$

式中:  $n$  是随机样本的观测值个数;  $i, j = 1, 2, \dots, n$ ; 如果  $x_i \leq x_j$  且  $y_i \leq y_j$ , 则  $\text{sign} = 1$ , 否则  $\text{sign} = -1$ 。

(2) 根据以上介绍的 4 种 Copula 函数 Kendall 系数  $\tau$  与  $\theta$  的关系式计算  $\theta$ 。

(3) 由  $\theta$  得到各种类型 Copula 的生成元, 也就得到各种类型的 Copula 函数。

(4) 判断最合适的 Copula 类型。判别过程如下:

①对于 Archimedean Copula 函数  $C(x, y)$ , 定义参数估计值  $K(t) = P[C(x, y) \leq t]$ ,  $t \in (0, 1)$ , 计算公式如下

$$K(t) = t - \frac{\varphi(t)}{\varphi'(t)}。 \quad (6)$$

②计算  $K(t)$  的非参数估计  $K_N(t)$ 。先计算  $t_i$ 。

$$t_i = \frac{x_j < x_i \text{ 且 } y_j < y_i \text{ 的 } (x_j, y_j) \text{ 个数}}{n-1}, \quad i = 1, \dots, n, \quad (7)$$

然后计算  $K_N(t)$ 。

$$K_N(t) = \frac{t_i < t \text{ 的个数}}{n}, \quad i = 1, \dots, n。 \quad (8)$$

③根据参数估计值  $K(t)$  和非参数估计值  $K_N(t)$  点绘 Q-Q 关系图。检验  $K(t)$  和  $K_N(t)$  是否相等, 如果图上的点落在 45° 线上, 表明  $K(t)$  和  $K_N(t)$  相等。选择图上点和 45° 线拟合程度最好的 Copula 类型。

## 3 降水资料

福海县位于新疆维吾尔自治区北部的阿勒泰地区, 准噶尔盆地北缘、阿勒泰山南麓。福海站地理坐标为 87°28'E, 47°7'N。降水量资料序列是从 1957 年 11 月至 2009 年 1 月的日降水量, 资料长度为 52 年, 且无缺测数据, 故而用于本文研究。本文研究的变量有年降水量  $V$  与年降水总天数  $D$ 。

## 4 年降水资料变异点分析

对福海站 52 年资料进行年降水量变异点分析 (表 1)。福海站共有 2 个变异点: 1973 年和 1982 年, 由此将该站降水量序列分为: 1958-1973 年; 1974-1982 年; 1983-2008 年, 共三段。1973 年变异后, 年降水量均值从 1 257 mm 减少至 845 mm; 1982 年变异后, 年降水量均值增加到 1 338 mm。

表 1 福海站降水资料年降水量变异分析结果

子序列	子序列长度	年降水量均值/mm
1958-1973 年	16	1 257
1974-1982 年	9	845
1983-2008 年	26	1 338

福海站年降水天数变异分析结果如表 2 所示。可见, 年降水天数与年降水量变异情况并不一致。另外, 年降水天数变异点在序列前端, 有可能是由变异检验方法自身的一些缺陷引起的: 如果检验变异点在端点, 其检验结果可信度较差<sup>[11]</sup>。考虑到 1958-1960 年子序列样本容量较小, 分析计算精度差, 结果不稳定, 且有可能是由于检验方法问题的错误变异点。故在本文计算中, 不考虑年降水天数变异。

表 2 福海站降水资料年降水天数变异分析结果

子序列	子序列长度	年降水天数均值/d
1958 - 1960 年	3	98
1961 - 2008 年	48	76

## 5 不考虑水文变异的 Copula 函数

本节不考虑降水变异,把整个降水序列当作一个整体进行计算分析。用式(5)计算年降水量和年降水天数的 Kendall 相关系数  $\tau = 0.49$ ,所以年降水量与年降水天数具有一定的正相关关系,以年为尺度,降水天数越多,年降水量也可能越大(图 1)。

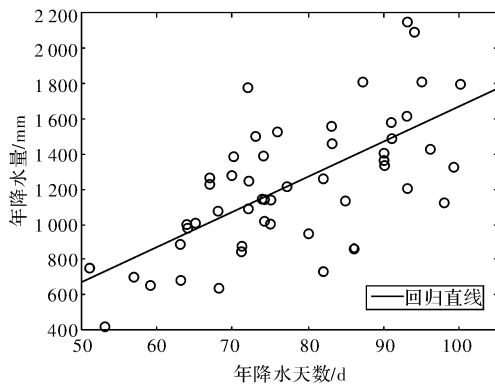


图 1 福海站年降水量与年降水天数散点图

我们把年降水量序列和年降水天数序列应用到 Gumbel-Hougaard Copula、Clayton Copula、Frank Copula,求出拟合程度最好的 Copula。年降水量和年降水天数的 Kendall 相关系数  $\tau = 0.49$ ,然后求出各类型 Copula  $\theta$  值。各种类型 Copula 的  $\theta$  值如表 3 所示。

表 3  $\tau = 0.49$  时,各类型 Copula 的  $\theta$  值和 AIC 值

	Gumbel-Hougaard	Clayton	Frank
$\theta$	1.96	1.92	5.55
AIC	-21 098	-20 872	-21 336

根据以上  $\theta$  值,得到年降水量和年降水天数各类型 Copula 的理论分布。各类型 Copula 函数的 AIC 值如表 3 所示。Frank Copula 的 AIC 值最小, Gumbel-Hougaard Copula 次之, Clayton Copula 的 AIC 值最大,因此, Frank Copula 拟合程度最好。根据式(6)至式(8),计算得参数估计值  $K(t)$  和非参数估计值  $K_N(t)$ ,再点绘 Q-Q 关系如图 2 所示。Gumbel-Hougaard Copula 上尾部分拟合程度较好,但下尾拟合程度较差; Clayton Copula 下尾部分拟合程度较好,但上尾拟合程度较差; Frank Copula 中间拟合程度较好,上尾部分拟合效果处

于 Gumbel-Hougaard Copula 之下, Clayton Copula 之上。同样,下尾部分拟合效果也是在另外两个类型 Copula 之间,综合来看, Frank Copula 整体拟合程度较好。福海站 1958 - 2008 年年降水量  $V$  和年降水总天数  $D$  的 Frank Copula 图像如图 3 所示。

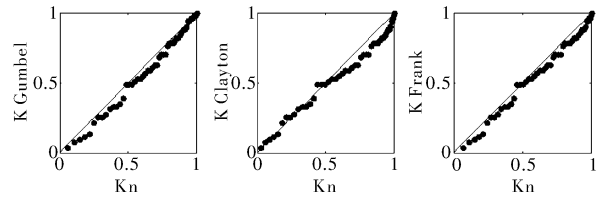


图 2 年降水量和年降水天数 Q-Q 图

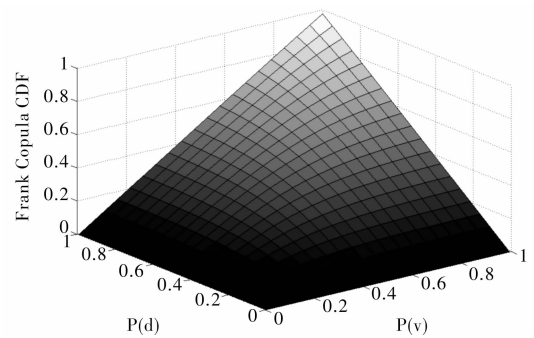


图 3 年降水量和年降水天数 Frank Copula 函数

## 6 考虑变异的 Copula 函数

由变异分析结果知,福海站年降水量发生两次变异,而年降水天数不考虑变异。所以考虑变异后,福海站降水资料分成 3 段: 1958 - 1973 年、1974 - 1982 年、1983 - 2008 年。通过式(5)计算各子序列年降水量  $V$  和年降水天数  $D$  的 Kendall 相关系数  $\tau$ ,然后求出各类型 Copula  $\theta$  值。各子序列 Kendall 相关系数  $\tau$  和各类型 Copula  $\theta$  值如表 4 所示。各子序列年降水量与年降水天数散点图如图 4 所示。

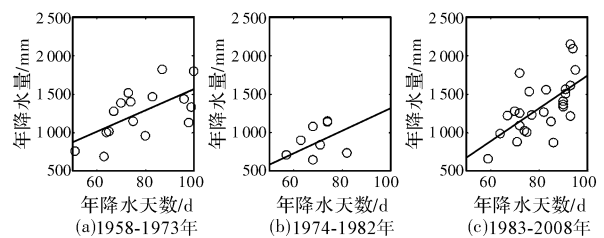


图 4 各子序列年降水量与年降水天数散点图

表 4 各子序列 Kendall 相关系数  $\tau$  和各类型 Copula  $\theta$  值

子序列	$\tau$	Gumbel-Hougaard	Clayton	Frank
1958 - 1973	0.47	1.88	1.75	5.16
1974 - 1982	0.46	1.84	1.69	5.01
1983 - 2008	0.50	1.99	1.98	5.68

根据以上  $\theta$  值, 得到各子序列年降水量和年降水天数各类型 Copula 的理论分布。各子序列 Copula 函数的 AIC 值如表 5 所示。1958 – 1973 年, Clayton Copula 的 AIC 值最小, Frank Copula 次之, Gumbel-Hougaard Copula 最大, 该序列 Clayton Copula 拟合程度最好。1974 – 1982 年, Clayton Copula 的 AIC 值最小, Frank Copula 次之, Gumbel-Hougaard Copula 最大, 但三者 AIC 都相差不大, 该序列 Clayton Copula 拟合程度最好。1983 – 2008 年, Gumbel-Hougaard Copula 的 AIC 值最小, Frank Copula 次之, Clayton Copula 最大, 该序列 Gumbel-Hougaard Copula 拟合程度最好。根据式 (6) 至式 (8), 计算得各子序列参数估计值 和非参数估计值, 再点绘 Q – Q 关系如图 5 所示。

表 5 各子序列 Copula 函数的 AIC 值

子序列	Gumbel-Hougaard	Clayton	Frank
1958 – 1973	-1 767	-1 903	-1 795
1974 – 1982	-491	-504	-497
1983 – 2008	-5 141	-4 671	-4 943

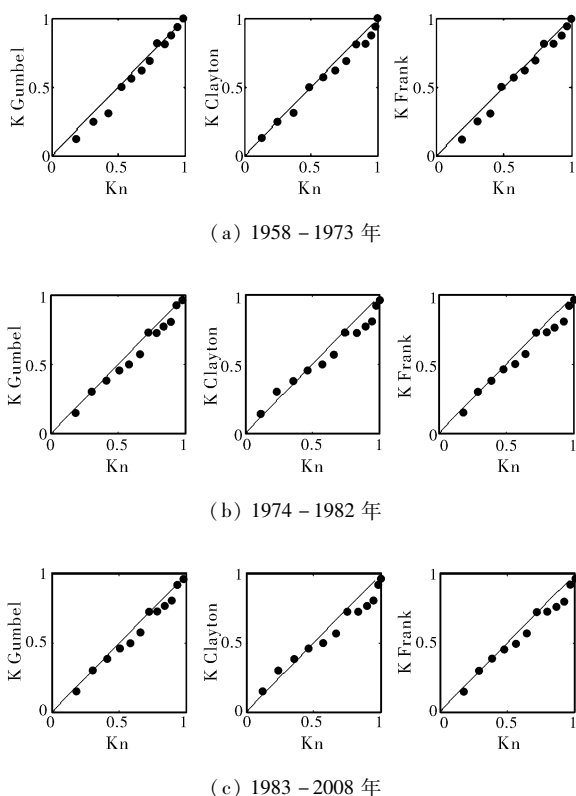


图 5 各子序列年降水量和年降水天数 Q – Q 图

1958 – 1973 年, Gumbel-Hougaard Copula 中部及上尾部分拟合较好; Clayton Copula 中部及下尾部分拟合较好; Frank Copula 中部拟合较好。从 Q – Q 关系图, 三者对不同部位拟合程度各有优劣, 可根据实际情况选取。Clayton Copula 的 AIC 值最

小, 本文选取 Clayton Copula 作为 1958 – 1973 年年降水量和年降水天数 Copula。运用同样的分析方法可判断出, 对 1974 – 1982 年降水量和年降水天数 Clayton Copula 拟合优度为最佳。1983 – 2008 年降水量和年降水天数以 Gumbel-Hougaard Copula 拟合优度最佳。各子序列拟合程度最好的 Copula 函数如图 6 所示。1958 – 1973 年与 1974 – 1982 年 Kendall 相关系数  $\tau$  相近, 并且都是 Clayton Copula, 所以年降水量和年降水天数 Copula 函数非常接近。1983 – 2008 年 Kendall 相关系数  $\tau = 0.50$ , 服从 Gumbel-Hougaard Copula。

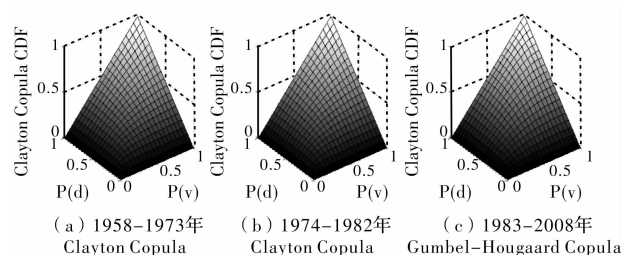


图 6 各子序列拟合程度最好的 Copula 函数

## 7 考虑变异与不考虑变异 Copula 对比

根据 Genest 和 Rivest 的 Copula 非参数估计方法, 确定某一序列年降水量和年降水天数 Copula, 只需要计算该序列年降水量和年降水天数的 Kendall 相关系数  $\tau$  和确定使用那种类型的 Copula 函数。

### 7.1 考虑变异与不考虑变异 Kendall 相关系数 $\tau$

不考虑变异, 1958 – 2008 年年降水量和年降水天数 Kendall 相关系数  $\tau = 0.49$ 。考虑变异后, 1958 – 1973 年  $\tau = 0.47$ , 1974 – 1982 年  $\tau = 0.46$ , 1983 – 2008 年  $\tau = 0.50$ 。各子序列 Kendall 相关系数  $\tau$  与整个序列  $\tau$  相差不大。考虑变异对年降水量和年降水天数 Kendall 相关系数影响不大。

### 7.2 考虑变异与不考虑变异 Copula 函数的选择

本文同时使用 AIC 准则法和 Genest 和 Rivest 的 Copula 非参数估计方法选择拟合程度最好的 Copula 函数。AIC 准则法以函数拟合的偏差和参数数量带来的不可靠性为基础构造 AIC 值, AIC 值最小的 Copula 函数拟合程度最好。AIC 准则法直观快捷, 是判断 Copula 整体拟合程度的良好指标, 但不能具体地判别不同部位的拟合情况。Copula 非参数估计方法通过 Q – Q 图, 能够详细地考察不同部位理论频率与经验频率的关系, 方便地判别

Copula 函数在不同部位的拟合情况,但对于整体拟合情况的判别具有一定的主观性。AIC 准则法和 Copula 非参数估计方法优缺点刚好互补,同时运用这两种方法,能够根据实际需要选择 Copula 函数。

不考虑变异时,本文选择 Frank Copula 作为年降水量和年降水天数 Copula。在该序列, Frank Copula AIC 值最小,整体拟合情况最好。从 Q-Q 图看出, Frank Copula 中部拟合程度最好,上尾和下尾检验效果在 3 种 Copula 中处于中等水平。本文研究整体的联合分布情况,所以选择 Frank Copula。在实际工作中,可根据 AIC 和 Q-Q 图选择满足需要的 Copula。考虑变异后,1958-1973 年和 1974-1982 年均是 Clayton Copula 拟合程度最好,1983-2008 年是 Gumbel-Hougaard Copula 拟合程度最好。1958-1973 年和 1974-1982 年 Kendall 相关系数  $\tau$  变化不大,而且 Copula 种类相同,该次变异并没有导致年降水量和年降水天数联合分布太大的变化。1974-1982 年和 1983-2008 年 Kendall 相关系数  $\tau$  变化不大,但最优 Copula 函数不同,年降水量和年降水天数联合分布变化较大。

考虑变异和不考虑变异,得到不同的最优 Copula 函数类型。考虑变异的子序列的最优 Copula 类型也不相同。变异对 Kendall 相关系数  $\tau$  影响较小,但对 Copula 函数选择影响较大。

## 8 结论与讨论

针对年降水量与年降水天数变异不同步的情况,即年降水量发生变异,而年降水天数没有发生变异。采用传统变异分析方法进行分析,通过考虑变异和不考虑变异的 Copula 对比,和考虑变异不同子序列的 Copula 对比得出以下结论:①考虑变异和不考虑变异,以及考虑变异不同子序列的 Kendall 相关系数  $\tau$  变化不大。②考虑变异和不考虑变异,以及考虑变异不同子序列拟合程度最好的 Copula 种类可能不同,即对 Copula 函数选择影响较大。

本文采用  $t$  检验和 Mann-Whitney U 检验的连续假设检验方法对福海站年降水量和年降水天数序列进行多变异点分析。在考虑变异的多变量联合分布中,各变量的变异情况可能不同,分析计算需要综合考虑各个变量的变异情况。目前,变异分析方法只能够应用于单变量,还没有公认的多变量变异分析方法。本文例子中,年降水天数变

异发生在序列前端,有可能是由于分析方法自身缺陷而产生的错误变异点,观察年降水天数图也没有发现显著变异,故本文认为年降水天数没有发生变异,只需要考虑年降水量变异情况。计算表明,变异前后 Kendall 相关系数  $\tau$  没有发生太大变化。年降水量与年降水天数之间有比较稳定的物理关系,一般来说,年降水天数越长,降水时间长,年降水量可能越大。年降水量即使发生变异,年降水量与年降水天数的这种物理关系仍存在,这可能是 Kendall 相关系数  $\tau$  变化不大的原因。

Dias 等<sup>[22-23]</sup>以 Csörgö 等<sup>[24]</sup>的方法为基础,在金融方面应用多变量 Copula 变异分析方法。该法构造 Copula 参数  $\theta$  的统计检验量,当统计检验量超出给定概率的阈值,则认为  $\theta$  发生显著变异,即 Copula 发生变异。可以把这个思路应用于气象和水文分析。

## 参考文献:

- [1] Alan D Z, Justin S, Edwin P M, et al. Detection of intensification in global-and continental-scale hydrological cycles: Temporal scale of evaluation [J]. *Journal of Climate*, 2003, 16: 535-547.
- [2] Zhang Qiang, Xu Chongyu, Zhang Zengxin, et al. Changes of atmospheric water vapor budget in the Pearl River basin and possible implications for hydrological cycle[J]. *Theoretical and Applied Climatology*, 2010, DOI 10.1007/s00704-010-0257-z.
- [3] Zhang Qiang, Xu Chongyu, Chen Yongqin David, et al. Spatial assessment of hydrologic alteration across the Pearl River Delta, China, and possible underlying causes[J]. *Hydrological Processes*, 2009, 23: 1565-1574.
- [4] Zhang Qiang, Xu Chongyu, Singh V P. Multi-scale variability of sediment load and streamflow of the Lower Yangtze River basin: possible causes and implications [J]. *Journal of Hydrology*, 2009, 368: 96-104.
- [5] Yue S. Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls[J]. *Hydroscience Journal*, 2009, 45(2): 315-326.
- [6] Yue S. The Gumbel logistic model for representing a multivariate storm event[J]. *Advances in Water Resources*, 2000, 24(2): 315-326.
- [7] Salvadori G, Michele C. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events[J]. *Water Resources Research*, 2004, 40(12), DOI: 10.1029/2004WR003133.
- [8] Favre A C, Adlouni E, Perreault L, et al. Multivariate hydrological frequency analysis using copulas[J]. *Water Resources Research*, 2004, 40, DOI: 10.1029/2003WR002456.
- [9] De M, Salvadori C, Canossi M, et al. Bivariate statistical approach to check adequacy of dam spillway [J]. *Journal of*

- Hydrologic Engineering, 2005, 10(1): 50–57.
- [10] Zhang L, Singh V P. Bivariate flood frequency analysis using the copula method [J]. Journal of Hydrologic Engineering, 2006, 11(2): 150–164.
- [11] Reeves J, Chen Jian, Wang Xiaolan, et al. A Review and comparison of changepoint detection techniques for climate data [J]. American Meteorological Society, 2007, 46: 900–915.
- [12] Seigel S. Nonparametric statistics for the behavioral sciences [J]. Journal of Nervous & Mental Disease, 1957, 125(3): 497.
- [13] Mann H B., Whitney D R. On a test of whether one of two random variables is stochastically larger than the other [J]. The Annals of Mathematical Statistics, 1947, 18(1): 50–60.
- [14] Alexandersson H. A homogeneity test applied to precipitation data [J]. J Climatol, 1986, 6: 661–675.
- [15] Hawkins D M. Point estimation of the parameters of a piecewise regression model [J]. Appl Statist, 1976, 25: 51–57.
- [16] Nelsen RB. 2006. An introduction to copulas [M]. USA: Springer Science, 2006: 269.
- [17] Embrechts P, Lindskog F, McNeil A. Modelling dependence with copulas and applications to risk management [M]. Switzerland, 2001.
- [18] Genest C, Rivest L P. Statistical inference procedures for Bivariate Archimedean Copulas [J]. American Statistical Association, 1993, 88(423): 1034–1043.
- [19] Deheuvels P. La fonction de dépendance empirique et ses propriétés [J]. Un test non paramétrique d'indépendance Acad Roy Bel Bull Cl Sci, 1979, 65(5): 274–292.
- [20] 肖义. 基于 Copula 函数的多变量水文分析计算方法及应用研究 [D]. 武汉: 武汉大学, 2007: 131.
- [21] Akaike H. A new look at the statistical model identification [J]. IEEE Transactions on automatic control, 1974, 19(6): 716–723.
- [22] Dias A, Embrechts P. Change-point analysis for dependence structures in finance and insurance [M]. Sociedade Portuguesa de Estatística, 2002: 69–86.
- [23] Dias A, Embrechts P. Dynamic copula models for multivariate high-frequency data in finance [M]. ETH Zurich, 2004: 42.
- [24] Csörgő M, Horváth L. Limit theorems in change-point analysis [M]. Canada: John Wiley & Sons, 1997: 409.

## Application of Archimedean Copulas in Research on Probability

### Distribution Characteristics of Precipitation Variation

#### —Taking Fuhai County in Xinjiang as an Example

Li Jianfeng<sup>1, 2</sup>, Zhang Qiang<sup>1, 2</sup> and Chen Xiaohong<sup>1, 2</sup>

(1. Department of Water Resources and Environment, Sun Yat-sen University, Guangzhou 510275, China;

2. Key Laboratory of Water Cycle and Water Security in Southern China of Guangdong High Education Institute, Sun Yat-sen University, Guangzhou 510275, China)

**Abstract:** The influences of two-dimensional non-synchronous variation, that a variable changes but the other does not change simultaneously, on joint distribution of Copula are discussed. In the environment of temperature changes, significant variability occurs in the precipitation process, and joint distribution of annual precipitation and annual total precipitation is important to the long-term drought conditions. Based on Copula function, joint distribution of annual precipitation and annual total precipitation of Fuhai County station in Xinjiang is calculated, and variation analysis is done by traditional variation analysis method. In the case that precipitation significantly changed but precipitation days didn't, the influence of precipitation changes on copula joint distribution is discussed. Research results indicate that Kendall's  $\tau$  does not change much on the effect of precipitation changes, but the best fitting copula is affected much.

**Key words:** Copula; variation analysis; annual precipitation; annual precipitation days; Fuhai County of Xinjiang