Revista da Estatística UFOP, Vol III(3), 2014, ISSN 2237-8111
Edição especial: 59ª Reunião Anual da Regional Brasileira da Sociedade Internacional de Biometria - RBRAS
Trabalhos apresentados na versão poster.

Modelagem da duração e intensidade de secas utilizando uma cópula arquimediana

Jailson de Araujo Rodrigues [1 2]
Ana Paula Coelho Madeira Silva [3]
Jaime dos Santos Filho[1]
Ângela Lima da Silva [1]

1 Introdução

As secas são fenômenos climáticos recorrentes que resultam da persistência de valores de precipitação abaixo da média, afetando diversas regiões do planeta e provocando impactos negativos nas atividades agrícolas, industriais e urbana. Contrariamente aos demais desastres naturais que geralmente atuam de forma rápida e com resultados imediatos, a seca é um desastre natural que afeta maior número de pessoas, com longos períodos de duração e intensidade. Esses dois fatores correlacionados são importantes elementos que caracterizam as secas e geralmente são modelados pelas distribuições exponencial e gama, respectivamente (SHIAU; MORRADES, 2009). O fato de serem modelados por diferentes distribuições, inviabiliza a aplicação das distribuições bivariadas usuais na análise conjunta dessas variáveis.

As cópulas inicialmente desenvolvidas por Sklar (1959), são transformações que são utilizadas para construir distribuições multivariadas a partir de distribuições marginais selecionadas. A utilização de cópulas na análise de fenômenos de secas é algo relativamente novo. No entanto, podemos destacar os seguintes trabalhos: Kao e Govindaraju (2010), Shiau (2006) e Shiau et al. (2007).

O objetivo deste trabalho é utilizar uma cópula para construir uma distribuição bivariada para modelagem de secas. Em particular, vamos obter a distribuição conjunta da duração e intensidade de seca ocorridas em Cascavel no Paraná. A série histórica analisada corresponde as secas ocorridas num período de 30 anos.

2 Materiais e métodos

2.1 Índice Padronizado de Precipitação (SPI)

Na literatura, é possível encontrar vários índices que permitem o monitoramento de secas. O mais utilizado no Brasil é o Índice Padronizado de Precipitação (SPI) desenvolvido por McKee

[1]DEPEN-IFBA. e-mail: *jailsondearaujo@yahoo.com.br*
[2]Agradecimento ao IFBA pelo apoio financeiro.
[3]CSL - UFSJ.

et al. (1993). Esse índice quantifica o déficit ou o excesso de precipitação em diferentes escalas de tempo. Essa característica torna o SPI uma valiosa ferramenta para todos os estudos de disponibilidade hídrica. A escala temporal mais analisada é a mensal (SPI-1 mês). O evento seca começa quando o SPI torna-se negativo e termina quando este volta a apresentar valores positivos. Dentro de sua escala, magnitudes menores ou iguais a -2 indicam seca extrema. Na Tabela 1 é apresentada essa classificação.

Tabela 1. Classificação da severidade da seca.

Valores do SPI	Categoria de seca
$0,00$ a $-0,99$	Seca ligeira
$-1,00$ a $-1,49$	Seca moderada
$-1,50$ a $-1,99$	Seca severa
$\leq -2,00$	Seca extrema

Considerando que a seca ocorre quando o valor do SPI é menor que zero, a duração da seca é a somatoria dos meses consecutivos em que o SPI é negativo. Dessa forma, se X representa a duração da seca, a intensidade da seca, denotada por Y, é a soma do valores do SPI no período de seca. Por conveniência, assumiremos a intensidade da seca como uma grandeza positiva e assim, pode-se escrever:

$$Y = -\sum_{i=0}^{X} SPI_i \tag{1}$$

Os dados de SPI explorados neste trabalho foram coletados na estação meteorológica com coordenadas geográficas $-24,93333$ de latitude, $-53,43333$ de longitude e 760m de altitude, localizada na cidade de Cascavel no Estado do Paraná. A base de dados corresponde a uma série histórica de medições mensais do SPI no período de janeiro de 1976 até dezembro de 2005.

2.2 Cópulas

O teorema de Sklar (1959) estabelece que se $F(x)$ e $F(y)$ são funções de distribuição univariadas e $C(u,v)$ é uma cópula, então a função:

$$F_{X,Y}(x,y) = C(F(x), F(y)) \tag{2}$$

é uma função de distribuição bivariada com distribuições marginais $F(x)$ e $F(y)$. Reciprocamente, se $F_{X,Y}(x,y)$ é uma função de distribuição bivariada com distribuições marginais $F(x)$ e $F(y)$, então, existe uma cópula $C(u,v)$, $0 \leq u,v \leq 1$ tal que (2) vale. Além disso, se $F(x)$ e $F(y)$ são contínuas, C é única e tem representação:

$$C(u,v) = F\left(F_X^{-1}(u), F_Y^{-1}(v)\right) \tag{3}$$

Sob a hipótese que as distribuições marginais são contínuas com funções densidade de probabilidade $f_X(x)$ e $f_Y(y)$, a função densidade de probabilidade conjunta pode ser escrita como:

$$f_{X,Y}(x,y) = c\left(F(x), F(y)\right) f_X(x) f_Y(y) \tag{4}$$

em que $c(u,v)$ é a densidade de C, dada por:

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}. \tag{5}$$

Para a construção da distribuição bivariada será utilizada a cópula arquimediana de Clayton dada por:

$$C(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}, \ \ \theta \geq 0 \tag{6}$$

com função densidade de probabilidade,

$$c(u,v) = (\theta + 1)\left(u^{-\theta} + v^{-\theta} - 1\right)^{-(1/\theta)-2} (uv)^{-\theta-1} \tag{7}$$

Considerando que a duração (X) e a intensidade (Y) da seca tem distribuições exponencial e gama, respectivamente, com funções densidade de probailidade apresentadas a seguir:

$$f(x) = \lambda \exp(-\lambda x) \tag{8}$$

$$f_Y(y) = \frac{\beta^{\alpha} y^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta y) \tag{9}$$

sendo que λ é um parâmetro de escala da distribuição exponencial, $\alpha > 0$ e $\beta > 0$ são, respectivamente, os parâmetros de forma e escala da distribuição gama. Além disso, $\Gamma(\cdot)$ representa a função gama,

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} \exp(-t)\, dt. \tag{10}$$

A distribuição bivariada proposta para análise do duração e da intensidade de secas pode ser expressa como:

$$F_{X,Y}(x,y) = \left(F_X(x)^{-\theta} + F_Y(y)^{-\theta} - 1\right)^{-1/\theta} \tag{11}$$

sendo que $F_X(x)$ e $F_Y(y)$ são, respectivamente, as funções de distribuições acumuladas exponencial e gama dadas por:

$$F_X(x) = 1 - \exp(-\lambda x) \tag{12}$$

e

$$F_Y(y) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \int_0^{y} t^{\alpha-1} \exp(-\beta t) dt \tag{13}$$

Para o ajuste do modelo (11) será utilizado o Método de Inferência pelas Marginais, ver

Joe (1997) e Shiau et al. (2007). Esse método é dividido em duas etapas: primeiramente, são ajustadas as distribuições marginais pelo Método da Máxima Verossimilhança e em seguida estima-se o parâmetro de dependência da cópula.

3 Resultados e discussão

O coeficiente de correlação de Pearson entre a duração da seca e sua intensidade é $0,832$, comprovando a assertiva da existência de correlação entre essas duas grandezas. As estimativas obtidas pelo Método de Inferência pelas Marginais no ajuste da distribuição bivariada (11) foram respectivamente: $\widehat{\lambda} = 0,503$, $\widehat{\alpha} = 1,073$, $\widehat{\beta} = 0,640$ e $\widehat{\theta} = 0,753$. Dessa forma, a distribuição conjunta para a duração e intensidade de secas é dada por:

$$F_{X,Y}(x,y) = \left(F_X(x)^{-0,753} + F_Y(y)^{-0,753} - 1\right)^{-1,328} \tag{14}$$

sendo que $F_X(x)$ e $F_Y(y)$ são as funções de distribuições acumuladas da duração e intensidade da seca apresentadas a seguir:

$$F_X(x) = 1 - \exp(-0,503) \tag{15}$$

$$F_Y(y) = \frac{0,640^{1,073}}{\Gamma(1,073)} \int_0^y t^{0,073} \exp(-0,640t)dt \tag{16}$$

Os gráficos de probabilidade para o ajuste da distribuição da variável X e da variável Y são exibidos na Figura 1, indicando a qualidade de ajuste das distribuições marginais.

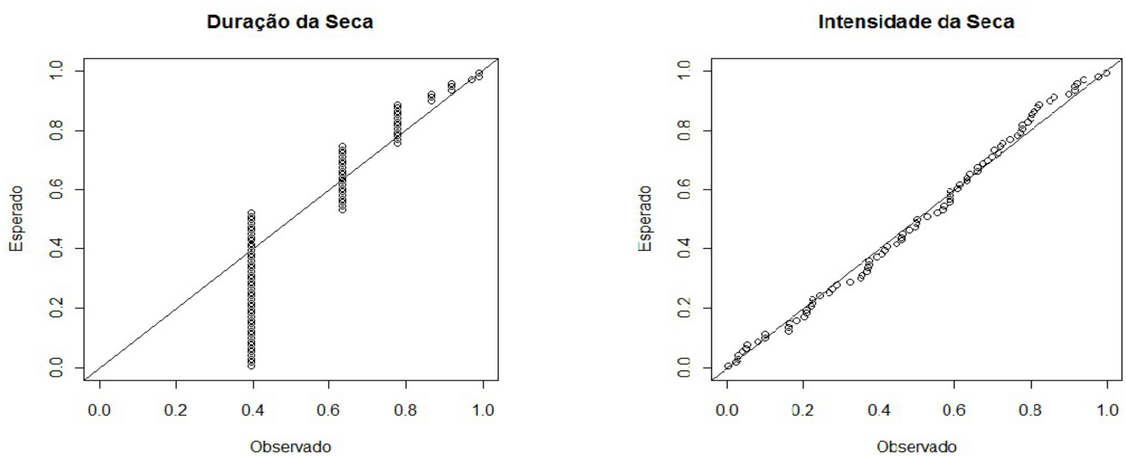

Figura 1: Gráficos de probabilidade para a duração da seca (X) e sua intendidade (Y) para a cidade de Cascavel.

4 Conclusões

As secas são um componente crítico no planejamento do uso racional dos recursos hídricos. Dessa forma, entender os elementos que caracterizam as secas é essencial para que se possa determinar medidas para mitigar os problemas relacionados com esse fenômeno. Nesse trabalho, analisamos a estrutura de dependência entre a duração e a intensidade de secas ocorridas em Cascavel. Para isso foi construída uma distribuição bivariada a partir da cópula arquimediana de Clayton. A qualidade de ajuste do modelo com relação aos dados de seca analisados foi verificada utilizando gráficos de probabilidade. Os resultados obtidos indicam que a distribuição construída pode ser empregada como uma alternativa viável na análise de secas ocorridas em Cascavel, PR.

Referências

[1] KAO, S. C.; GOVINDARAJU, R. S. **A copula-based joint deficit index for droughts** J. Hydrology, v. 380, p. 121-134, 2010.

[2] MCKEE, T. B.; DOESKEN, N. J.; KLEIST, J. **Drought monitoring with multiple time scales**. 9th Conference on Applied Climatology, Preprints. American Meteorological Society,Boston, p. 233-236, 1995.

[3] SHIAU, J. T. **Fitting drougth duration and severity with two-dimensional copulas**. Water Resources Management, v. 20, p. 795-815, 2006.

[4] SHIAU, J. T.; FENG, S.; NADARAJAH, S. **Assessment of hydrological droughts for the Yellow River, China, using copulas**. Hydrological Processes, v. 21, p. 2157-2163, 2007.

[5] SHIAU, J. T.; MODARRES, R. **Copula-based drought severity-duration-frequency analysis in Iran**. Meteorological Applications, v. 16, p. 481-489, 2009.

[6] SKLAR, K. **Fonctions de repartition à n dimensions et leura margens**. Publications de l'Institut de Statistique de L'Université de Paris, v. 8, p. 229-231, 1959.