



# Um sistema de monitorização contínua da probabilidade de secas baseado em cópulas, CDPMS

João Dehon PONTES FILHO<sup>1</sup>, Maria Manuela PORTELA<sup>2</sup>, Ticiania Marinho de Carvalho STUDART<sup>3</sup>,  
Francisco Assis SOUZA FILHO<sup>4</sup>

1. Universidade Federal do Ceará, Brasil, dehon@alu.ufc.br

2. Instituto Superior Técnico/Universidade de Lisboa (IST/UL), CERIS - Investigação e Inovação em Engenharia Civil  
maria.manuela.portela@tecnico.ulisboa.pt

3. Universidade Federal do Ceará, ticiania@ufc.br

4. Universidade Federal do Ceará, assis@ufc.br

## RESUMO

Apresenta-se um sistema de monitorização em tempo real da probabilidade de secas baseado em cópulas, **CDPMS** (do inglês, *Continuous Drought Probability Monitoring System*). Para o efeito recorreu-se ao conceito de superfície limiar de precipitação relativa a um dado intervalo de tempo e intensidade de seca, no pressuposto de que se pretendia conhecer a probabilidade de ocorrência de uma seca com aquela intensidade no final do intervalo de tempo em causa, a partir do progressivo conhecimento da precipitação ao longo do mesmo. Os resultados da aplicação do **CDPMS** a Portugal Continental evidenciam muito bom desempenho o qual advém, por um lado, do facto de a precipitação ocorrer fundamentalmente no período de outubro a março pelo que um défice de precipitação em tal semestre é fortemente indicativo da plausibilidade de seca, e, por outro lado, da capacidade de as cópulas modelarem estruturas de dependência complexas, como as que ocorrem entre precipitações em diferentes intervalos de tempo. O **CDPMS** constitui uma ferramenta que, uma vez implementada, se pretendeu de fácil manuseamento, destinada a monitorizar a precipitação e consequentemente a probabilidade da ocorrência de secas, permitindo ao utilizador antecipar medidas de mitigação e de adaptação ou mesmo emitir alertas, desde que devida e institucionalmente capacitado para o efeito.

**Palavras-Chave:** seca; SPI; superfície limiar de precipitação, cópulas; probabilidade multivariadas

## INTRODUÇÃO. OBJETIVO E DADOS DE BASE

As secas são causadas por precipitações persistentemente inferiores às médias sendo que à medida que aumenta a duração do período de tempo com défice de precipitação assim se manifestam pela diminuição progressiva da água armazenada no solo, dos recursos hídricos naturais superficiais e subterrâneos e dos volumes de água armazenados em albufeiras, com impactes progressivamente mais gravosos nos ecossistemas, na economia, na saúde pública e na própria segurança nacional. É o desastre natural que maiores custos acarreta à escala mundial. Contudo, é difícil perceber o seu início dado que se desenvolve gradualmente, sendo que frequentemente quando se identifica uma situação de seca já é tarde para implementar medidas de mitigação dos seus efeitos. Afigurou-se, assim, que seria importante desenvolver um sistema que permitisse com alguma facilidade antecipar medidas face à plausibilidade de, a breve prazo, se declarar uma situação seca. Para o efeito recorreu-se ao Índice de Precipitação Padronizada, SPI (*Standardized Precipitation Index*), designadamente, através do conceito de superfícies de limiares de precipitação, desenvolvido por Portela *et al.*, 2012, e à modelação bivariada com base em cópulas. Tal sistema de monitorização contínua da probabilidade de secas baseado em cópulas foi identificado pelo acrónimo decorrente da sua designação em Inglês, ou seja, por **CDPMS - Continuous Drought Probability Monitoring System** (Pontes Filho *et al.*, 2019).

As superfícies de limiares de precipitação fornecem os valores mínimos da precipitação acumulada num dado intervalo de tempo, pré-fixado consoante o objetivo em vista, para que não ocorra uma seca, com um dado grau de intensidade, de seca moderada, a seca severa e, por fim a seca extrema (Portela *et al.*, 2012). Uma vez que numa dada localização se obtenham as anteriores precipitações, o **CDPMS** fornece, a partir do primeiro mês do período em causa, a estimativa da probabilidade de ocorrência no final desse mesmo período de uma seca com

dada intensidade, sendo que tal estimativa é progressivamente atualizada à medida que as precipitações nos sucessivos meses do período vão sendo conhecidas.

Em termos matemáticos o problema relativamente ao qual o **CDPMS** pretendeu dar resposta pode ser formulado da seguinte forma: seja  $N$  o número de meses em estudo a que se refere a superfície limiar de precipitação relativa a dada intensidade de seca;  $R_N^*$  a precipitação mínima fornecida por essa superfície para uma dada localização;  $R_n$  a precipitação observada nos  $n$  meses iniciais do período ( $1 \leq n \leq N$ );  $R_{N-n}$  a precipitação nos restante ( $N-n$ ) meses do período (desconhecida); e  $R_N$  a precipitação total no período (igualmente desconhecida). Ocorrerá uma seca no final do período de  $N$  meses se a precipitação já registada  $R_n$  adicionada da precipitação  $R_{N-n}$  ainda por ocorrer ficar abaixo do limiar de precipitação indicativo de seca com a intensidade considerada,  $R_N^*$ , ou seja, se:

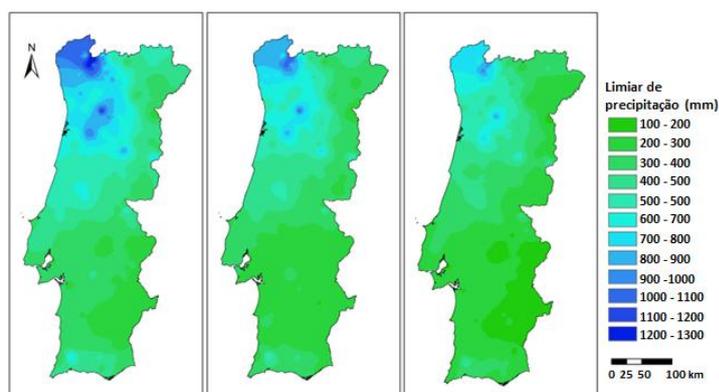
$$R_n + R_{(N-n)} \leq R_N^* \Leftrightarrow R_N \leq R_N^*$$

A probabilidade de ocorrer uma seca no final do período de  $N$  meses uma vez que seja conhecida a precipitação nos  $n$  meses iniciais desse período é dada por:

$$P(R_N \leq R_N^* | R_n)$$

ou seja, está-se perante um problema de probabilidade condicional, o qual pode ser resolvido por recurso a cópulas. A aplicação de cópulas em hidrologia tem merecido uma atenção progressiva, mencionando-se, a título de exemplo relacionado com secas, a análise de frequência combinado diferentes características das secas (intensidade, duração, magnitude e distribuição espacial) (Shiau, 2006, Ayantobo *et al.*, 2018).

Na aplicação do anterior modelo a Portugal Continental selecionou-se o semestre chuvoso, de outubro a março, dada a elevada contribuição da precipitação durante o mesmo para o total anual (cerca de 70%, nas regiões mais húmidas do País, e perto de 80%, nas mais secas), o que faz com que o défice de precipitação em tal semestre seja um forte indicativo da plausibilidade de ocorrência de secas. Na **Figura 1** apresentam-se as correspondentes superfícies de limiares de precipitação obtidas por inversão do SPI a essa mesma escala temporal ( $SPI6_{out-mar}$ ). Sempre que num dado ponto a precipitação registada no período em causa for inferior à fornecida por uma dessas superfícies assiste-se a uma seca com a intensidade definida por essa figura.



**Fig. 1.** Superfícies de limiares de precipitação obtidas por inversão do SPI entre outubro e março ( $SPI6_{out-mar}$ ) da esquerda para a direita, para seca moderada (-0.84), severa (-1.28) e extrema (-1.65). Tais superfícies fornecem em cada ponto o valor de  $R_N^*$  (adaptada de Portela *et al.*, 2012).

De modo a permitir a monitorização contínua da probabilidade de ocorrência de seca em todo o País selecionaram-se os 45 postos udométricos a que se refere a **Figura 2**. Os registos de precipitações mensais nos anteriores postos foram obtidos através do SNIRH. As falhas de precipitação foram preenchidas através de modelos de regressão linear simples, sendo que cada falha utiliza uma equação de regressão específica (Portela *et al.*, 2012). Foi assim possível dispor de 100 anos, entre 1918/19 e 2017/2018 de precipitações mensais na rede de postos da **Figura 2**.

## 2. MODELOS

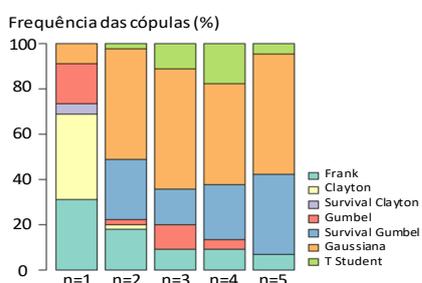
O modelo bivariado aplicado em cada posto udométrico é detalhadamente descrito em Pontes Filho *et al.*, 2019. Menciona-se apenas que, para cada mês do semestre chuvoso ( $1 \leq n \leq 5$  e  $N = 6$ ) num dado posto, tal modelo foi construído tendo por base a amostra emparelhada ( $R_n, R_N$ ) constituída pela precipitação acumulada entre outubro e esse mês,  $R_n$ , e a correspondente precipitação no semestre  $R_N$ . A dimensão dessa amostra é igual à do período que serviu de base ao estudo, ou seja, 100 anos. Tendo em vista a descrição da estrutura de

dependência entre as precipitações em causa por meios de cópulas consideraram-se diferentes famílias candidatas de cópulas, designadamente de Gumbel, Frank, Clayton, Gaussiana e T-Student. Os parâmetros dessas famílias foram estimados pelo método da máxima verosimilhança (*maximum likelihood estimation*, MLE), concretamente, pelo método da máxima pseudo-verosimilhança (*maximum pseudo-likelihood method*, MPL). A seleção de cópulas recorreu ao Critério de Informação de Akaike (*Akaike Information Criteria*, AIC). Anota-se que para cada posto da **Figura 2** e para cada intervalo de tempo  $n$  se aplicou uma cópula específica. Na **Figura 3** sintetiza-se a frequência de utilização das diferentes cópulas consideradas.



Nome (código)	ID	Lat (°)	Long (°)	Nome (código)	ID	Lat (°)	Long (°)
Merufe (01G/03)	RG01	42.0180	-8.3890	Bemposta (17I/02)	RG24	39.3490	-8.1410
Travancas (03N/01)	RG02	41.8280	-7.3056	Alter do Chão (18L/01)	RG25	39.2182	-7.6844
Leonte (03I/03)	RG03	41.7650	-8.1470	Pragança (18C/01)	RG26	39.1990	-9.0640
Soutelo (Chaves) (03L/02)	RG04	41.7530	-7.5348	Pavia (20I/01)	RG27	38.8965	-8.0136
Campo de Viboras (04R/03)	RG05	41.5240	-6.5580	Caia (Monte Caldeiras) (20O/02)	RG28	38.8873	-7.0898
Cabeceiras de Basto (04J/06)	RG06	41.5127	-7.9792	Santo Estevão (20E/02)	RG29	38.8600	-8.7460
Sta Marta da Montanha (04K/02)	RG07	41.5008	-7.7460	Estremoz (20L/01)	RG30	38.8416	-7.6159
Folgares (06N/01)	RG08	41.3032	-7.2828	Colares (Sarrazola) (21A/01)	RG31	38.8020	-9.4570
Carviçais (06P/02)	RG09	41.1790	-6.8900	Évora-Monte (21K/02)	RG32	38.7690	-7.7161
Moncorvo (06O/04)	RG10	41.1650	-7.0510	São Manços (23K/01)	RG33	38.4605	-7.7505
Adorigo (07L/01)	RG11	41.1460	-7.6070	Bar. de Pego do Altar (23G/01)	RG34	38.4196	-8.3952
Pindelo dos Milagres (09J/02)	RG12	40.8060	-7.9630	Amieira (24L/01)	RG35	38.2793	-7.5605
Freixedas (09O/02)	RG13	40.6880	-7.1630	Barrancos (25P/01)	RG36	38.1321	-7.0013
Gouveia (11L/01)	RG14	40.4940	-7.5930	Santa Vitória (26I/01)	RG37	37.9645	-8.0227
Santo Varão (12F/02)	RG15	40.1840	-8.6020	Serpa (26L/01)	RG38	37.9426	-7.6038
Góis (13I/01)	RG16	40.1568	-8.1133	Relíquias (27G/01)	RG39	37.7030	-8.4825
Soure (13F/01)	RG17	40.0521	-8.6250	Castro Verde (27I/01)	RG40	37.6976	-8.0933
Penha Garcia (13O/01)	RG18	40.0420	-7.0180	Mértola (28L/01)	RG41	37.6371	-7.6619
Alvaizere (15G/01)	RG19	39.8270	-8.3810	Rosário (Almodóvar) (28I/02)	RG42	37.6020	-8.0810
Ladoeiro (14N/02)	RG20	39.8269	-7.2660	Barragem de Mira (28G/01)	RG43	37.5101	-8.4433
Nisa (16L/03)	RG21	39.5160	-7.6690	Sta Catarina (Tavira) (31K/01)	RG44	37.1487	-7.7847
Castelo de Vide (17M/01)	RG22	39.4116	-7.4525	Valverde (31E/03)	RG45	37.0820	-8.7180
Pernes (17F/01)	RG23	39.3910	-8.6630				

**Fig. 2.** Localização e identificação dos 45 postos udométricos utilizados no estudo (coordenadas geográficas no sistema WGS84).



**Fig. 3.** Cópulas selecionadas no conjunto dos 45 postos da **Figura 2**, em função do período inicial de meses do semestre húmido com precipitação conhecida ( $n=1$ , refere-se apenas ao mês de outubro,  $n=2$  aos meses de outubro e novembro e assim sucessivamente, até  $n=5$  que se refere aos meses de outubro a fevereiro) (adaptada de Pontes Filho *et al.*, 2019).

O cálculo das probabilidades condicionadas utilizou a seguinte equação

$$H(X \leq x | Y = y) = C_{\theta}(U_1 | U_2 = u_2) = \frac{\partial}{\partial u_2} C_{\theta}(u_1, u_2)$$

em que  $X$  e  $Y$  designam as duas variáveis aleatórias em presença (a precipitação acumulada entre outubro e esse mês e a correspondente precipitação no semestre de outubro a março) com distribuições marginais  $F_x(x)$  e  $F_y(y)$ , sendo  $U_1 = F_x(x)$  e  $U_2 = F_y(y)$  e  $u_1$  e  $u_2$  valores específicos.  $C$  designa a cópula com parâmetros  $\theta$  que combina as distribuições marginais numa função de distribuição conjunta. Apenas foram consideradas cópulas bivariadas, como decorre da anterior equação.

### 3. RESULTADOS

Menciona-se que, antecedendo a utilização do **CDPMS**, houve que estabelecer e validar o sistema conforme descrito em Pontes Filho *et al.*, 2019. No presente artigo exemplifica-se apenas a sua aplicação à monitorização em tempo real da probabilidade de ocorrência de seca. Uma vez que tal aplicação requer a atualização permanente das precipitações mensais e que se pretendeu produzir resultados para todo o território nacional, selecionou-se, a título de exemplo, o semestre húmido de outubro de 2012 a março de 2013, com o objetivo de avaliar a probabilidade de seca moderada neste último mês à medida que a precipitação entre outubro de 2012

( $n=1$ ) e fevereiro de 2013 ( $n=5$ ) ia sendo progressivamente conhecida. Os resultados obtidos constam da **Figura 4**. As probabilidades de ocorrência de seca moderada em março foram representadas por meio de superfícies obtidas mediante aplicação do método IDW (*inverse distance weighting*) com coeficiente de ponderação de 2 às probabilidades obtidas em cada um dos 45 postos da **Figura 2**. No último mapa identificam-se os postos que, de facto, registaram ou não seca moderada em março de 2013.

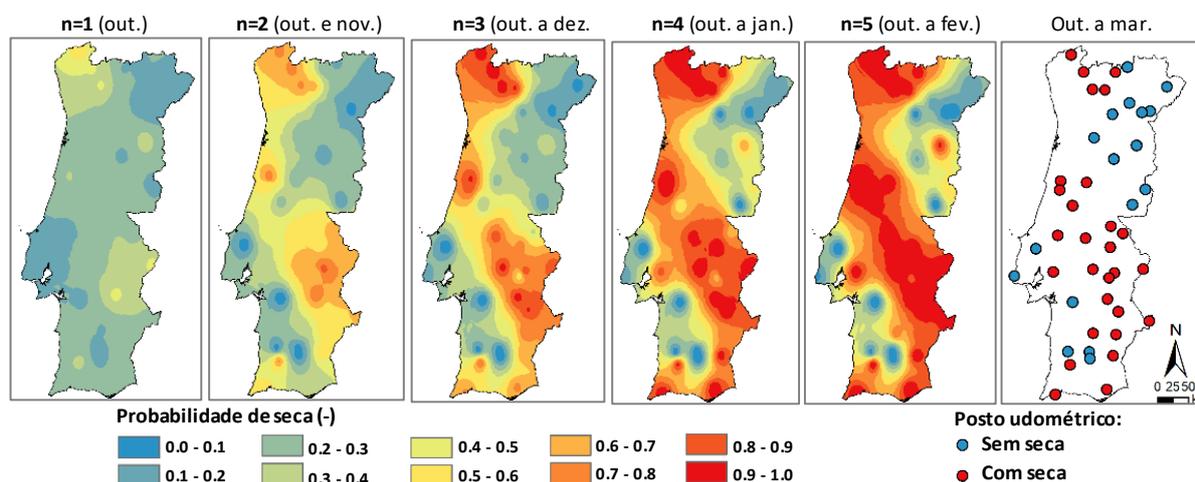


Fig. 4. Aplicação do sistema de monitorização contínua da probabilidade de secas baseado em cópulas, **CDPMS**, à probabilidade de ocorrência de seca moderada em março 2013 à medida que a precipitação mensal entre out./2012 e fev./2013 vai sendo conhecida. O último mapa identifica os postos que efetivamente registaram ou não seca moderada no final do período de outubro a março (adaptada de Pontes Filho *et al.*, 2019).

#### 4. CONCLUSÃO

A **Figura 4** evidencia o bom ajuste entre previsões de ocorrência de seca moderada e localizações que efetivamente experienciaram seca, ou seja, o bom desempenho do **CDPMS**. A figura mostra ainda que, à medida que são incorporadas as precipitações nos sucessivos meses do semestre húmido, se destacam regiões relativamente às quais a probabilidade de seca é francamente não negligenciável, justificando a emissão de alertas, com alguma cautela, em dezembro de 2013, já com alguma certeza, em janeiro de 2014 e definitivamente, em fevereiro deste último ano. Um tal sistema de aviso prévio pode contribuir para alertar autoridades mais diretamente relacionadas com a gestão dos recursos hídricos e com a proteção civil, permitindo a antecipação de medidas e ações gradualmente ajustáveis face à plausibilidade do evento. Contudo importa ter presente que a expectativa de seca pode não se confirmar caso ocorram precipitações inesperadas, sem que isso comprometa a mais-valia do sistema de aviso prévio.

Anota-se que, não obstante o **CDPMS** se ter revelado capaz de antecipar a ocorrência de secas, não permitiu distinguir com igual fiabilidade a intensidade das secas que possivelmente se avizinham, julga-se que em consequência de as diferenças entre limiares de precipitação relativos às diferentes intensidades de seca serem muito pequenas e poderem ser facilmente compensadas por acontecimentos pluviosos pontuais que tornam mais complexa a descrição da estrutura de dependência entre precipitações a diferentes escalas temporais. Por fim regista-se que o **CDPMS** pode ser aplicado a regiões ou localizações específicas, requerendo em qualquer caso uma rede de monitorização da precipitação capaz de fornecer registos praticamente em tempo real.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- Ayantobo, O.O.; Li, Y.; Song, S.; Javed, T.; Yao, N. Probabilistic modelling of drought events in China via 2-dimensional joint copula. *J. Hydrol.* 2018, 559, 373–391.
- Pontes Filho, J.D., Portela, M.M., Studart, T.M.C., Souza Filho, F.A. A continuous drought risk monitoring system, **CDPMS**, based on copulas. *Water*, 2019, 11(9), 1925, <https://doi.org/10.3390/w11091925>.



Portela, M.M, Santos, J.F., Naghettini, M., Matos, J.P., Silva, A.T. Superfícies de limiares de precipitação para identificação de secas em Portugal continental: uma aplicação complementar do Índice de Precipitação Padronizada, SPI. Recursos Hídricos, 2012, 33(2):5-23. <https://doi.org/10.5894/rh33n2-1>.

Shiau, J.T. Fitting drought duration and severity with two-dimensional copulas. Water Resour. Manag. 2006, 20, 795–815.