

Previsão de secas na primavera em Portugal Continental com base em indicadores climáticos de larga escala

Spring drought forecasting in mainland Portugal based on large-scale climatic indices

Santos, J.F.^a, Portela, M.M.^b, Pulido-Calvo, I.^c

^aDpto. Engenharia, ESTIG, Instituto Politécnico de Beja, Rua Afonso III, 7800-050, Portugal. E-mail: joaof.santos@ipbeja.pt

^bDpto. Engenharia Civil, Instituto Superior Técnico, Universidade de Lisboa, Portugal, Avda. Rovisco Pais, 1049-001 Lisboa, Portugal. E-mail: maria.manuela.portela@tecnico.ulisboa.pt

^cDpto. Ciencias Agroforestales, Escuela Técnica Superior de Ingeniería, Campus La Rábida, Universidad de Huelva, 21819 Palos de la Frontera (Huelva), Spain. E-mail: ipulido@uhu.es

Recibido: 08/09/2015

Aceptado: 21/10/2015

Publicado: 28/10/2015

RESUMO

O sucesso de uma estratégia de mitigação dos efeitos da seca passa pela implementação de um sistema de monitorização e previsão eficaz, capaz de identificar os eventos de seca e de seguir a sua evolução espaço-temporal. Neste artigo demonstra-se a eficiência de redes neuronais artificiais na previsão, para Portugal, do índice de precipitação padronizada, SPI, relativo à primavera. A validação dos modelos recorreu ao *hindcasting*, designando-se, por tal, a técnica através da qual um dado modelo é testado mediante a sua aplicação a períodos temporais históricos, com comparação dos resultados obtidos com as respectivas observações. O índice SPI foi calculado à escala temporal de 6 meses tendo o *hindcast* utilizado como indicadores climáticos a oscilação do Atlântico Norte e temperaturas da superfície do mar. O estudo evidenciou a mais valia da inclusão dos anteriores preditores externos no modelo de previsão. Elaboraram-se, ainda, mapas de probabilidade de ocorrência de seca os quais constituem importantes ferramentas no planeamento integrado e na gestão de recursos hídricos.

Palavras chave | Redes neuronais artificiais; *Hindcasting*; SPI; NAO; SST; SPI.

ABSTRACT

The success of a strategy of mitigation of the effects of the droughts requires the implementation of an effective monitoring and forecasting system, able to identify drought events and follow their spatiotemporal evolution. This article demonstrates the capability of the artificial neural networks in predicting the spring standardized precipitation index, SPI, for Portugal. The validation of the models used the hindcasting, which is a technique by which a given model is tested through its application to historical data followed by the comparison of the results thus achieved with the data. The SPI index was calculated at the timescale of six months and the climate indices used as external predictors in the hindcasting were the North Atlantic Oscillation and temperatures of the sea surface. The study showed the added value of the inclusion of previous predictors in the model. Maps of the probabilities of the drought occurrences which may be very important for integrated planning and management of water resources were also developed.

Key words | Artificial neural networks; *Hindcasting*; SPI; NAO; SST; SPI.

INTRODUÇÃO

A previsão sazonal de condições hidro-climatológicas tornou-se recentemente um assunto científico relevante (Trigo *et al.*, 2004; Gámiz-Fortis *et al.*, 2008a,b; Ionita *et al.*, 2008; Gámiz-Fortis *et al.*, 2010). A seca, quando comparada com outros eventos extremos mais recorrentes, como as cheias e os tornados, é reconhecida como um fenómeno que possui impactos igualmente adversos, mas que podem ser mitigados de modo efectivo (Di Mauro *et al.*, 2008). Com efeito, a ocorrência de uma seca e as suas consequências podem requerer bastante tempo até serem percebidas pelos sistemas socioeconómicos, pois tendem a evoluir temporalmente de forma lenta. De acordo com Rossi (2003), um sistema de monitorização e de previsão eficaz, capaz de prontamente identificar a ocorrência de uma seca e de seguir a sua evolução espaço-temporal, constitui um dos principais pré-requisitos para a implementação de uma estratégia de mitigação das suas consequências.

Das inúmeras variáveis climáticas que podem ser utilizadas para analisar secas, a precipitação é reconhecida como o indicador chave de detecção e caracterização do fenómeno e da sua persistência (Lloyd-Hughes, 2002), sendo também a variável essencial na maioria dos modelos hidrológicos (Santos *et al.*, 2005). Para o clima Mediterrâneo que prevalece em Portugal, a precipitação de Inverno é um dos principais factores do balanço hídrico, especialmente tendo em conta o facto de tal clima não fazer esperar ausência de precipitação naquela estação do ano. Deste modo, a precipitação ou a ausência de precipitação no Inverno tem um papel fundamental, tanto no prevenir, como no indiciar situações de seca com consequente afectação dos sistemas naturais e socioeconómicos.

A oscilação do Atlântico Norte (NAO) possui importantes padrões de teleconexão com diferentes variáveis climáticas e, consequentemente, com o fenómeno de seca em toda a Europa, sendo vários os estudos sobre o assunto (Qian *et al.*, 2000a,b, Hurrell *et al.*, 2001, Trigo *et al.*, 2002, e, mais recentemente, López-Moreno *et al.*, 2007, e Vicente-Serrano *et al.*, 2011). De acordo com Ionita *et al.* (2012) as correlações temporais significativas entre padrões de seca e a NAO sugerem a utilização na Europa deste indicador climático como potencial predictor dos padrões de seca a escalas temporais inter-anuais. López-Moreno e Vicente-Serrano (2008) também analisaram o impacto da NAO de Inverno na ocorrência de seca na Europa e Trigo *et al.* (2006) demonstraram que a magnitude das anomalias médias de SPI é notavelmente diferente para as fases positivas e negativas da NAO, sendo o sinal do SPI contrário ao da NAO.

Ionita *et al.* (2012) exploraram igualmente correlações entre a temperatura da superfície do mar (SST) e os padrões de seca na Europa. A escalas temporais inter-anuais, esses autores identificaram padrões de variabilidade de seca fortemente correlacionados com padrões de SST de anos anteriores, sendo que o diferimento temporal entre as ocorrências de seca e as anomalias de SST pode fornecer informação valiosa com vista à previsão daquele fenómeno. Também Hoerling e Kumar (2003) demonstraram a importante relação existente entre temperaturas da superfície do mar, por um lado frias e persistentes no Pacífico (zona Tropical-Este) e, por outro lado, quentes no oceano Índico-Pacífico (zona Tropical-Oeste) com a seca de 1998-2003 que afectou os USA, mas também a Europa do Sul e o Sudoeste Asiático.

A escolha da escala temporal a adoptar na análise de determinadas condições climáticas tem uma considerável influência na caracterização da disponibilidade de água e, consequentemente, nas condições de seca. A escalas temporais sazonais, Gámiz-Fortis *et al.* (2010) mostraram que anomalias nas condições atmosféricas e oceânicas encontram-se muitas vezes associadas a variações sazonais de precipitação e temperatura (ex., Trigo *et al.*, 2004; López-Moreno *et al.*, 2007; López-Moreno e Vicente-Serrano, 2008) podendo, como tal, despoletar variações no escoamento e consequentemente no armazenamento de água em albufeiras. De acordo com Gámiz-Fortis *et al.* (2010), o desempenho das previsões de médio a longo prazo encontra-se associado à introdução de predictores que representam componentes de baixa frequência do sistema climático, como a cobertura de neve, o teor de água no solo, a SST e os principais padrões de circulação atmosférica e oceânica como sejam o ENSO (*El Niño–Southern Oscillation*) ou a NAO na Europa.

Não tendo sido identificadas contribuições relativas ao impacto de indicadores de circulação atmosférica e de condições oceânicas (ex., NAO, ENSO, SOI, campos de SST, etc.) na variabilidade espacial e temporal das secas em Portugal, resume-se um estudo que teve por objectivo efectuar o *hindcast* do fenómeno da seca no território de Portugal Continental fazendo intervir alguns daqueles indicadores (Santos *et al.*, 2014). O *hindcasting* é uma técnica através da qual um dado modelo é testado mediante a sua aplicação a períodos temporais históricos, ou seja, passados, com apreciação do seu desempenho por comparação dos resultados a

que conduz com observações relativas àqueles períodos. Devido às características climáticas da região oeste da Península Ibérica, é particularmente importante averiguar se o *hindcasting* no que respeita a ocorrências de secas é suficientemente preciso quando efectuado tendo por base a previsão de secas para o período primaveril. De facto, o regime de precipitação na região é caracterizado por um pico primário no outono e um pico secundário na primavera e pela ausência de precipitação no verão (Martín *et al.*, 2004), o que significa que, uma vez que as condições de seca se estabeleçam no início da primavera, é previsível que se prolonguem pelo menos até Outubro (início, em Portugal Continental, do período com chuva e do ano hidrológico). De acordo com Lloyd-Hughes (2002), a primavera é a estação do ano mais previsível no que diz respeito à precipitação, e consequentemente, às ocorrências de seca.

A identificação e caracterização das secas recorreram ao índice de precipitação padronizada, SPI (McKee *et al.*, 1993), que tem sido largamente utilizado no âmbito da investigação sobre o fenómeno em Portugal (Santos *et al.*, 2010, 2011, 2013, 2014). Alguns autores têm proposto métodos para prever ou para avaliar a provável evolução temporal do SPI (Cordery e McCall, 2000; Bordi *et al.*, 2005; Cancelliere *et al.*, 2005, 2007). No entanto, apesar de tais esforços, tentar prever o início e o fim de uma seca continua a ser uma tarefa bastante difícil. Na investigação que se resume, o modelo de *hindcasting* de secas utilizou redes neuronais artificiais (ANN), à semelhança de autores como Mishra e Desai (2006), Kim e Valdés (2003), e mais recentemente, Dastorani e Afkhami (2011). No essencial, tentou-se dar resposta às seguintes três questões: i) a consideração de preditores externos, como a NAO e a SST relativos a meses anteriores (estação de Inverno), pode levar a uma melhoria da capacidade preditiva dos modelos?; ii) em que medida o aumento do horizonte temporal de previsão afecta o desempenho dos modelos?; iii) é possível associar os resultados de previsão obtidos a mapas de probabilidade de ocorrência de seca?.

Constituem contribuições relevantes da investigação efectuada a análise do impacto de preditores de larga-escala (SST) e de padrões de circulação atmosférica (NAO) na variabilidade espacial e temporal das secas em Portugal Continental e a avaliação quantitativa, com base num novo conceito de mapas de probabilidade de eventos de seca, da capacidade dos modelos efectuarem o *hindcasting* de secas tendo por base aqueles preditores.

MÉTODOS

Área de estudo e dados

A investigação centrou-se em Portugal Continental, território, em termos climáticos, sob a influência das zonas Atlântica e Mediterrânea. A precipitação média anual no País varia entre mais de 3000 mm, na região noroeste, a menos de 400 mm, na região sul, sendo o relevo um dos factores mais determinantes na distribuição espacial da precipitação. À medida que a disponibilidade de água diminui de norte para sul o regime hidrológico (ao longo do ano e entre sucessivos anos) torna-se progressivamente mais irregular (Portela e Quintela, 2006), facto que sugere que a previsão de secas possa contribuir muito significativamente para a gestão de recursos hídricos, nomeadamente nas regiões semi-áridas do centro e sul do País.

Santos *et al.* (2010) procederam à caracterização do padrão espacial de secas em Portugal mediante recurso ao índice SPI, calculado a partir de 94 anos de registos de precipitação mensal (de Outubro de 1910 a Setembro de 2004) nos 144 postos udométricos esquematicamente localizados na Figura 1. Tendo por base uma análise de *clusters*, aqueles autores demonstraram que em Portugal se podem considerar as três regiões com padrões distintos de secas esquematizadas na Figura 1 e identificadas por norte (*cluster 1* ou CL1), centro (*cluster 2* ou CL2) e sul (*cluster 3* ou CL3). O número de postos udométricos incluídos em cada região é de 56, 53 e 35, respectivamente.

Tendo por base modelos de ANN aplicados a cada posto de cada uma das regiões da Figura 1, procedeu-se ao *hindcasting* de secas na mesma para os meses primaveris, para o que foram incorporados na previsão diferentes preditores.

A informação relativa às secas consistiu na quantificação do défice de precipitação à escala temporal de 6 meses, através do cálculo das séries temporais de SPI a essa escala (SPI6). Os SPI objecto de previsão foram os primaveris, referentes, concretamente,

aos períodos de seis meses terminando em Abril, Maio e Junho, $SPI6_{Abril}$, $SPI6_{Maio}$ e $SPI6_{Junho}$, respectivamente, por se prever que possibilitem a antevisão do estabelecimento de condições de seca nas condições climáticas ocorrentes em Portugal Continental.

A escolha da escala temporal de seis meses atendeu ao facto de se admitir que essa escala é a mais adequada para avaliar a seca agrícola, sendo que, em Portugal, de entre as actividades económicas relevantes que são particularmente afectadas pela ocorrência de seca figura precisamente a agricultura. De acordo com Mo *et al.* (2009), o SPI6 pode ser muito eficaz na monitorização da variabilidade da precipitação em períodos distintos e na identificação das respectivas tendências de médio prazo. Segundo aqueles autores o SPI6 consegue ser mais sensível na detecção de desvios da precipitação do que outros índices (ex., o índice de Palmer).

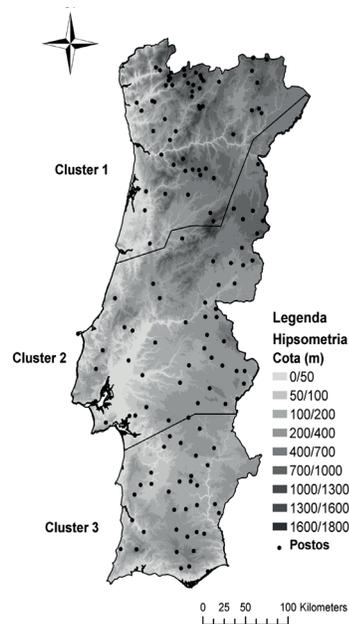


Figura 1 | Padrões espaciais de seca em Portugal continental com base no SPI. Três regiões identificadas por análise de *clusters*. Os pontos representam a localização dos 144 postos utilizados na análise (adaptada de Santos *et al.*, 2010).

Os predictores utilizados foram as SSTs globais recolhidas do *Extended Reconstructed SST v3b dataset*, desenvolvidas pelo *National Oceanic and Atmospheric Administration's National Climatic Data Center* (Xue *et al.*, 2003; Smith *et al.*, 2008) e obtidas a partir do endereço electrónico do *Royal Netherlands Meteorological Institute* (KNMI). A cobertura espacial utilizada para os campos de SST foi calculada com base no Oceano Mediterrâneo (37°E-7°W; 31°N-47°N), SST1, e em duas das áreas previamente identificadas no estudo de Gámiz-Fortis *et al.* (2010), designadamente, no sudoeste do Oceano Atlântico (35°W-25°W; 15°S-10°S), SST2, e no Oceano Atlântico Norte (45°W-25°W; 38°N-42°N), SST3, conforme se representa na Figura 2. As séries mensais de SST foram obtidas através da média em cada mês das células existentes em cada uma das anteriores regiões (cada célula têm 2°×2°). No final e para cada região, gerou-se a série de SST relativo à estação do Inverno através da adimensionalização da correspondente média mensal da SST de Janeiro, Fevereiro e Março, SST (JFM), pela média e desvio-padrão para o período 1910/11-2003/04.

Gámiz Fortis *et al.* (2008a,b) mostraram que o padrão temporal das SSTs de Outono geograficamente representativo das regiões Atlânticas destacadas na Figura 2 possui correlação estatisticamente significativa com as séries temporais de escoamento no Inverno em três importantes bacias hidrográficas Portuguesas, facto que determinou a adopção no presente estudo da SST naquelas regiões como variável predictora, mas agora referente a SPIs em meses primaveris.

Uma vez que a NAO é um dos principais modos de circulação atmosférica que determina o clima na Europa (Hurrell *et al.*, 2003), optou-se por também incluir este índice como predictor, avaliando se o seu comportamento no Inverno poderia ser utilizado para prever seca na subsequente estação primaveril. A NAO é registada ao longo de todo o ano, apesar de também apresentar uma importante variabilidade sazonal, o que levou vários autores a estudar o efeito da NAO sobre as secas da Península Ibérica (Martín-Vide e Fernández, 2001, Trigo *et al.*, 2004, e Xoplaki *et al.*, 2004).

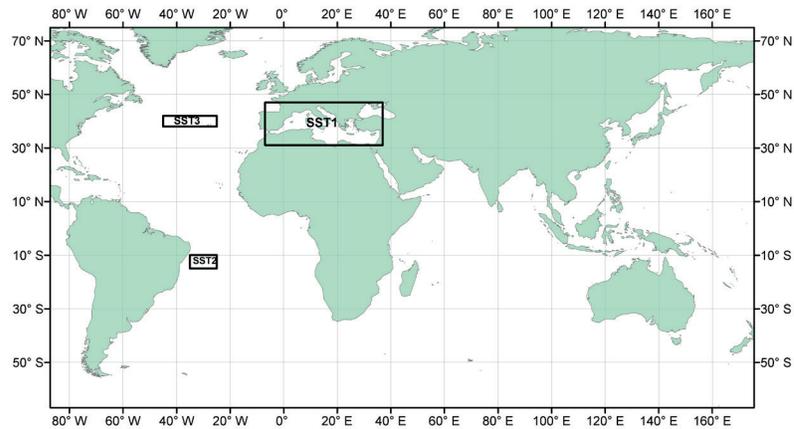


Figura 2 | Localização das zonas SST1, SST2 e SST3 referentes às temperaturas da superfície do mar Mediterrâneo, do Atlântico Sul e do Atlântico Norte, respectivamente.

A NAO possui uma grande intensidade e cobertura espacial, especialmente durante o Inverno (Hurrell *et al.*, 2003), sendo reconhecido que a sua fase positiva se encontra ligada a secas no Sul da Europa e a aumentos de precipitação na Europa do Norte (Trigo *et al.*, 2002; Hurrell *et al.*, 2003; and Silva *et al.*, 2012). O padrão oposto ocorre durante as fases negativas, o que, especificamente para Portugal Continental, conduz a cheias mais frequentes (Silva *et al.*, 2010). Como tal, no presente estudo, o cálculo do índice NAO considerou, como habitualmente, os meses de Inverno de Dezembro, Janeiro e Fevereiro (DJF), mas também o mês de Março (M) de forma a ter em conta condições ocorrentes no início da primavera, as quais são muito importante em Portugal, designadamente, no que respeita à precipitação: NAO (DJFM). O anterior índice foi seleccionado para o período de Outubro de 1910 a Setembro de 2004, como definido por Jones *et al.* (1997). Para o Inverno, a diferença de pressão normalizada ao nível do mar entre Gibraltar e Reiquiavique (Islândia) é um forte indicador do sinal da NAO (<http://www.cru.uea.ac.uk/cru/data/vinther/nao1821.txt>; Vinther *et al.*, 2003), pelo que a NAO foi calculada no presente estudo com base na diferença de pressão entre aquelas duas localizações. Jones *et al.* (1997) verificaram que a escolha da estação de Gibraltar (em detrimento da estação de Lisboa ou da de Ponta Delgada) é a que melhor representa a parte sul do dipolo da NAO. Reforçam tal facto as conclusões de Hurrell (1995) e Jones *et al.* (2003), que indicam que, comparativamente à consideração da estação dos Açores para cálculo da NAO, a escolha das anteriores estações permite uma melhor correlação entre aquele índice e as séries de precipitação europeias.

Redes neuronais – Procedimento Geral

A previsão de fenómenos de seca é fundamental na mitigação de vários tipos de impactos sobre os sistemas de recursos hídricos. Tradicionalmente, têm sido utilizados modelos estatísticos lineares para previsão de secas hidrológicas com base em métodos de análise de séries temporais (ex., regressão linear simples ou múltipla, ARMA, etc.). Em alternativa aos anteriores modelos lineares que, além do mais, assumem os dados como sendo estacionários e que têm capacidade limitada para captar não-estacionariedades e não-linearidades nas séries hidrológicas (Yevjevich, 1972) e, mais especificamente, nos indicadores de seca (Kim e Valdés, 2003), utilizou-se uma abordagem mais robusta, nomeadamente, as redes neuronais artificiais (ANNs).

Vários autores têm confirmado o potencial das ANN quando aplicadas a diferentes tipos de processos e de variáveis hidrológicas. Dawson e Wilby (1998) demonstraram a capacidade das ANNs no campo da modelação dos processos de precipitação-escoamento e da previsão de cheias. Kim e Valdés (2003) desenvolveram uma metodologia de previsão baseada na conjugação da transformada de Fourier e ANNs para efectuar a previsão do PDSI (*Palmer Drought Severity Index*) para vários horizontes temporais. Mishra e Desai (2006) também avaliaram a grande capacidade das ANNs para modelar e efectuar previsões a partir de séries temporais não lineares e não estacionárias, tendo testado e comparado modelos lineares estocásticos (ARIMA/SARIMA)

com diferentes tipos de ANNs na previsão de secas utilizando séries de SPI. Os resultados demonstraram as ANNs suplantam os modelos estocásticos lineares.

ASCE (2000a,b) contém, porventura, um dos estudos mais completos sobre a aplicação de ANNs em hidrologia. Sob a forma de dois artigos, são comparados diferentes modelos e abordagens, sendo referido que as ANNs são ferramentas robustas para modelar quaisquer processos hidrológicos não lineares, tais como processos de precipitação-escoamento, escoamentos, precipitações (e logo, os indicadores de seca deduzidos dessas precipitações), águas subterrâneas, e simulações da qualidade da água. Após treino apropriado, as ANNs são capazes de gerar resultados muito satisfatórios para muitos dos problemas de previsão em hidrologia (Jain e Kumar, 2007).

As redes neuronais artificiais, ANNs, são modelos inspirados na arquitectura neuronal do cérebro humano. Consistem em redes formadas por diferentes camadas ligadas entre si, a que se dá o nome de perceptrões multicamada (MLP) (Rumelhart *et al.* 1986; Senthil-Kumar *et al.* 2005). Estas estruturas permitem o reconhecimento de padrões de ligação entre as variáveis intervenientes num dado sistema a modelar (Rohwer *et al.* in Michie *et al.*, 1994), identificando os “mapeamentos” óptimos entre as “entradas” e as “saídas” desse sistema. No âmbito da previsão de variáveis hidrológicas, Mutlu *et al.* (2008) obtiveram melhores resultados com este tipo de arquitectura de ANN relativamente a outras arquitecturas.

Quanto ao modo de funcionamento as estruturas mais utilizadas e reconhecidas são as designadas *multilayer feed forward networks* constituídas por redes neuronais que, por assim dizer, são unicamente “alimentadas para a frente”, ou seja, as funções implementadas conduzem a “saídas” que dependam exclusivamente das “entradas”, não existindo estado interno, isto é, não possuindo memória. A Figura 3a) exemplifica a arquitectura de uma rede neuronal com duas camadas escondidas e múltiplos nós ou perceptrões. Os elementos que constituem as camadas são designados por nós ou perceptrões. Na Figura 3a), existem g , n , m e s nós na camada de entrada ou de *input*, na primeira e segunda camada escondidas e na camada de saída ou *output*, respectivamente. Os parâmetros associados às diferentes ligações constituem os pesos, sendo que a informação apenas se transfere de camadas antecedentes para camadas subsequentes. A Figura 3b) esquematiza o funcionamento típico de um nó. Cada nó j recebe sinais de entrada a partir de cada nó i da camada antecedente. Associado a um sinal de entrada (X_i) encontra-se um peso (W_{ji}). O sinal de entrada efectivo (I_j) para o nó j é a soma ponderada de todos os sinais de entrada, de acordo com:

$$I_j = \sum_{i=1}^g x_i W_{ji} \quad (1)$$

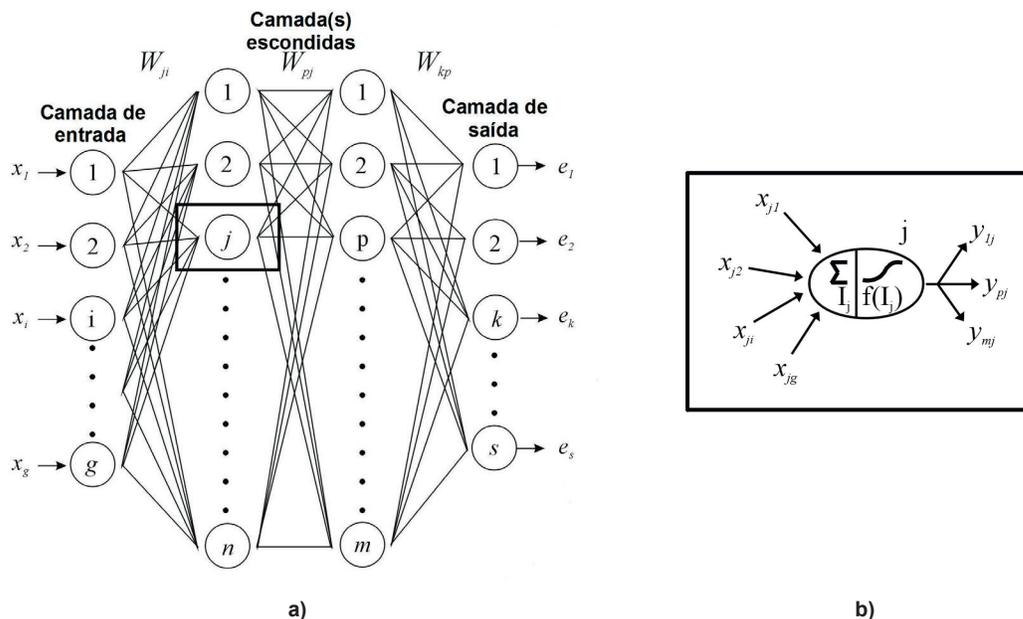


Figura 3 | a) Rede neuronal alimentada para a frente (*multilayer feed forward network*) com duas camadas escondidas. Variáveis de entrada: x_1, \dots, x_g e variáveis de saída: e_1, \dots, e_s ; b) Nó característico de uma rede neuronal (adaptada de Pulido–Calvo e Portela, 2007).

O sinal de entrada efectivo I_j passa através de uma função de activação, também chamada de função de transferência, de forma a produzir o sinal de saída (y_j) do nó j , sendo essa função a responsável pela relação entre os *inputs* e os *outputs* nos nós (Mishra e Desai, 2006). No estudo que se apresenta foi considerada a função linear ($y_i = I_j$) nas camadas de entrada e de saída, sendo a função de activação entre camadas escondidas a seguinte função sigmoideal não linear:

$$y_i = f(I_j) = \sigma(I_j) = \frac{1}{1 + e^{-I_j}} \quad (2)$$

onde I_j pode variar no intervalo $(-\infty, \infty)$, e y_j apresenta valores entre 0 e 1.

Para a estimativa do conjunto de pesos é levado a cabo um processo correctivo-repetitivo, também designado por fase de aprendizagem ou de treino da rede neuronal, no qual se dá forma às interconexões entre os diferentes perceptrones. A adequação dessas interconexões é efectuada recorrendo a uma técnica de convergência do erro de modo a que se produzam *outputs* tão próximos quanto possível dos valores esperados para os mesmos.

De entre os vários métodos de aprendizagem (calibração e validação) das ANNs, no presente estudo foi utilizado o algoritmo de cálculo de retro-propagação do erro (Pulido-Calvo e Portela, 2007). A aprendizagem por retro-propagação do erro consiste num algoritmo iterativo desenvolvido para resolver sistemas de equações lineares aplicado a redes neuronais com três ou mais camadas e que implementa o sistema de cálculo sucessivo das derivadas parciais numa direcção contrária à de propagação normal da informação através da rede (Rumelhart *et al.*, 1986).

O esquema de modelação adoptado consistiu na calibração e validação individual das ANNs para cada posto udométrico de cada sub-região (num total de 144 modelos de ANNs para as três regiões da Figura 1) em que a informação a ser *hindcasted* foi o SPI6 finalizando nos meses de Abril, Maio e Junho. Previamente à calibração de cada ANN, as séries temporais de SPI6 em cada posto foram divididas em dois subconjuntos: [TS], referente ao subconjunto (subset) de validação, e [CSS] compreendendo os subconjuntos de calibração ou [CS] e de dados seleccionados ou [SS]. O subconjunto [CSS] inclui 75% dos dados em cada posto udométrico da Figura 1 (ou seja, os dados em 68 anos) e o subconjunto [TS] os restantes 25% (22 anos). Em cada posto foram deixados 4 anos de registos (de 2001 a 2004) de forma a correr os modelos individuais, com o objectivo de produzir mapas de SPI6 para esses anos.

O sobre ajustamento do modelo ocorre quando o erro referente ao subconjunto [CSS] é muito reduzido mas aumenta ao serem apresentados novos dados à rede. Tal circunstância acontece porque a rede “memorizou” os exemplos de treino, mas não aprendeu a generalizar para novas situações. Assim, como forma de evitar o sobre ajustamento dos modelos, foi efectuada uma monitorização, no final de cada corrida (*epoch*), através do cálculo do somatório dos erros quadráticos para ambos os subconjuntos [CS] e [SS]. Quando este somatório começa a aumentar deve parar-se o treino e os pesos da ANN respeitantes à corrida que deu origem aos erros mínimos devem ser testados com o subconjunto [TS]. Este último passo é denominado de *external validation phase*. Embora Iyer e Rhinehart (1999) recomendem repetir este processo pelo menos 30 vezes para cada modelo, no presente estudo o mesmo foi repetido 100 vezes para cada posto udométrico por forma a escolher a melhor ANN aplicável a esse posto.

A selecção das variáveis de entrada é um aspecto fundamental na modelação com ANNs (Bowden *et al.*, 2005; May *et al.*, 2008; Fernando *et al.*, 2009). Para as fases de calibração e validação o objectivo é o de seleccionar variáveis de entrada que expliquem a maior percentagem de variância nos dados originais. Assim, avaliaram-se previamente as relações lineares entre as possíveis variáveis de entrada e os dados originais através de mapas de correlação linear, sendo que algumas destas variáveis de entrada foram padronizadas antes de serem incorporadas nos modelos de ANNs, como sugerido por vários autores (Mishra e Desai, 2006; Ochoa Rivera *et al.*, 2002, 2007), com o objectivo de tentar melhorar a eficiência da aprendizagem e o desempenho geral da rede neuronal.

A precipitação acumulada até Março em cada um dos 144 postos udométricos representados na Figura 1 foi incluída como principal predictor para cada SPI6 primaveril tendo cada série sido previamente normalizada. Para o *hindcasting* do SPI6_{Abril}, SPI6_{Mai} e SPI6_{Junho} consideraram-se como variáveis de entrada as precipitações acumuladas normalizadas de Novembro a Março (NDJFM), Dezembro a Março (DJFM) e Janeiro a Março (JFM), respectivamente. Consideraram-se também como variáveis de entrada nas redes os predictores externos adoptados para cada região, nomeadamente, as SSTs (JFM) previamente normalizadas e

o índice NAO (DJFM), em si mesmo já normalizado. Os outputs das ANNs foram os valores do SPI6 para os meses da primavera e para cada um dos 144 postos udométricos.

Foram testadas ANNs com uma e duas camadas escondidas, tendo, em cada caso, sido ensaiados entre 5 e 50 neurónios (5, 6, 7, ..., 50). Para alcançar um dos objectivos do estudo, foram testados 100 modelos de ANN em cada posto udométrico e para cada SPI6 dos meses da primavera a ser previsto. Para a validação procedeu-se ao cálculo de medidas de erro por forma a identificar o melhor modelo em cada posto udométrico e cada série temporal.

Mapas de correlação espacial entre os SPI6 primaveris e os potenciais predictores

No *hindcasting* sazonal, o objectivo é o de prever a distribuição espacial e temporal das condições de seca com alguns meses de avanço. Apesar de os complexos sistemas de dinâmica atmosférica não permitirem a previsão exacta, consegue-se antever algumas das características estatísticas e principais padrões temporais das secas.

De forma a entender como as SSTs (JFM) e a NAO (DJFM), adoptados como predictores externos, melhor descrevem o campo de SPI6 da primavera, foram obtidos mapas de correlação linear para Portugal Continental entre cada uma das séries de SPI6 da primavera ($SPI6_{Abril}$, $SPI6_{Maio}$ e $SPI6_{Junho}$) e cada um daqueles predictores externos (SST1, SST2, SST3 e a NAO). Esta análise forneceu uma base para a identificação, de entre as variáveis disponíveis, das a adoptar como variáveis de entrada. Este processo de selecção é um dos principais passos no desenvolvimento de um modelo de *hindcasting* satisfatório, na medida em que as variáveis de entrada determinam a estrutura do modelo e afectam os seus resultados. De acordo com autores como Bowden *et al.* (2005), May *et al.* (2008), Fernando *et al.* (2009) e Pulido-Calvo *et al.* (2012) este é também um aspecto essencial antes de treinar e validar uma ANN.

Medidas de erro

Para seleccionar a melhor ANN em cada posto udométrico e para cada SPI6 a ser utilizada como modelo de *hindcasting*, foram aplicadas medidas de erro no processo de validação externa, nomeadamente, a raiz quadrada do desvio quadrático médio (RMSE), o coeficiente de correlação (r), os coeficientes de eficiência (E1 e E2) e o índice de persistência (PI). As expressões de cálculo para o efeito aplicadas são as que constam de Kitanidis e Bras (1980), Legates e McCabe Jr. (1999) e Pulido-Calvo e Portela (2007).

Incerteza dos modelos de ANN

Por forma a avaliar a verosimilhança das previsões decorrentes dos melhores modelos de ANN, em cada posto e para cada mês de SPI6 da primavera, foram gerados mapas de probabilidade baseados na distribuição dos resíduos daqueles modelos. Os resíduos são dados pela diferença entre os valores de SPI6, por um lado, previstos ou *hindcasted*, e, por outro lado, observados. Em cada posto udométrico e para cada série temporal os resíduos deverão ser uma variável aleatória aproximadamente com distribuição normal (ou seja, com coeficiente de assimetria não significativamente diferente de zero) e com média nula (Demyanov *et al.*, 1998). De forma a testar a normalidade de cada série temporal de resíduos aplicou-se à mesma o teste de Snedecor e Cochran (Snedecor e Cochran, 1989).

Mapas de probabilidade de seca

A criação de mapas de probabilidade atendeu ao pressuposto de normalidade dos resíduos em cada posto udométrico e para cada SPI6 da primavera. Para cada série prevista, $SPI6_m$ a probabilidade de não excedência com respeito a um dado limiar específico de categoria de seca, Tr , é dado por $\Phi(Z)$, onde Φ representa a função de distribuição normal acumulada e Z , o valor da variável normal padrão dada por:

$$z = \frac{(Tr - SPI6_m) - \bar{X}_{res}}{S_{res}} \quad (2)$$

onde \bar{X}_{res} e S_{res} são a média e o desvio-padrão da correspondente série de resíduos.

Para cada $SPI6_m$, $\Phi(Z)$ fornece a probabilidade de ter um $SPI6_m$ menor ou igual a Tr , $P(SPI6_m \leq Tr)$, de acordo com as características estatísticas dos resíduos. O valor do limiar de categoria de seca, Tr , foi estabelecido com base nas categorias de seca propostas por Agnew (2000), apresentadas na Tabela 1.

Tabla 1 | Categorias de seca em concordância com os valores de SPI (Agnew 2000).

Probabilidade de não excedência	SPI	Categoria de seca
0.60	>-0.84 e <0.84	Normal
0.20	<-0.84	Seca moderada
0.10	<-1.28	Seca severa
0.05	<-1.65	Seca extrema

RESULTADOS E DISCUSSÃO

Seleção das variáveis de entrada

A abordagem desenvolvida para efectuar o *hindcasting* dos meses de SPI6 da primavera em Portugal Continental considerou a utilização, para além de precipitações acumuladas, de dados relativos a indicadores climáticos de grande-escala, designadamente, o índice NAO(DJFM) e as temperaturas da superfície do mar, SST1 (JFM), SST2 (JFM) e SST3 (JFM). Foram, assim, obtidos mapas respeitantes ao coeficiente de correlação no período de Outubro de 1910 a Setembro de 2004 entre as séries de SPI6 dos meses da primavera ($SPI6_{Abril}$, $SPI6_{Maio}$ e $SPI6_{Junho}$) e a NAO (DJFM) e os SSTs (JFM). Uma vez que os padrões de correlação espacial que, deste modo, se verificou ocorrerem são semelhantes para os três SPI6 da primavera, apenas se apresentam os resultados para $SPI6_{Abril}$ – Figura 4.

A Figura 4 mostra que a correlação é forte a moderada na maioria dos casos, com alguns valores estatisticamente significativos. De acordo com Yevjevich (1972), para amostras como a dimensão das utilizadas (94 anos), correlações de 0.19 são consideradas significativas para um nível de confiança de 5%.

A figura indica que as séries da NAO (DJFM) se correlacionam negativamente com os SPI6 da primavera, tendo os melhores valores de correlação sido obtidos nas regiões centro e sul (*cluster 2* e *3*). A anomalia SST1 (JFM) apresenta correlação positiva na região norte (*cluster 1*), com valores máximos de 0.5. As correlações entre SST2 (JFM) e qualquer das séries relativas aos SPI6 da primavera ($SPI6_{Abril}$, $SPI6_{Maio}$ e $SPI6_{Junho}$) foram sempre muito fracas e com pouco significado estatístico pelo que se excluiu o predictor SST2 (JFM) da análise. Para SST3 (JFM), foram obtidas correlações negativas significativas, designadamente, na região costeira a sul, mas também com valores aceitáveis para toda a região sul.

De acordo com os anteriores resultados adoptaram-se como potenciais predictores externos do SPI6 finalizando em Abril, Maio e Junho a NAO (DJFM), a SST1 (JFM) e a SST3 (JFM). Atendendo aos valores do coeficiente de correlação, a NAO (DJFM) foi adoptada para todo o País, enquanto a SST1 (JFM) apenas foi considerada no norte e a SST3 (JFM), no sul. Utilizando as séries temporais de tais predictores como *input* para previsão foram estabelecidos os correspondentes modelos de ANN multicamada.

Uma vez que a precipitação é o indicador mais importante para monitorizar e prever secas sendo, simultaneamente, a única variável interveniente no cálculo de SPI6, avaliou-se também a sua capacidade como potencial predictor de seca. As séries temporais de precipitação acumulada adoptadas como variáveis de entrada para a previsão dos SPI6 da primavera foram as consideradas no próprio procedimento de cálculo do SPI, embora finalizando sempre em Março, uma vez que as precipitações neste mês podem desempenhar um papel antecipatório importante no modelo de previsão de secas. Assim, a ANN a aplicar em cada posto udométrico para previsão de $SPI6_{Abril}$, $SPI6_{Maio}$ e $SPI6_{Junho}$ considerou como variável de entrada também a precipitação acumulada normalizada,

respectivamente, de Novembro a Março (NDJFM), de Dezembro a Março (DJFM) e de Janeiro a Março (JFM). Foram também estas as precipitações acumuladas que exibiram maiores correlações com aqueles SPI6.

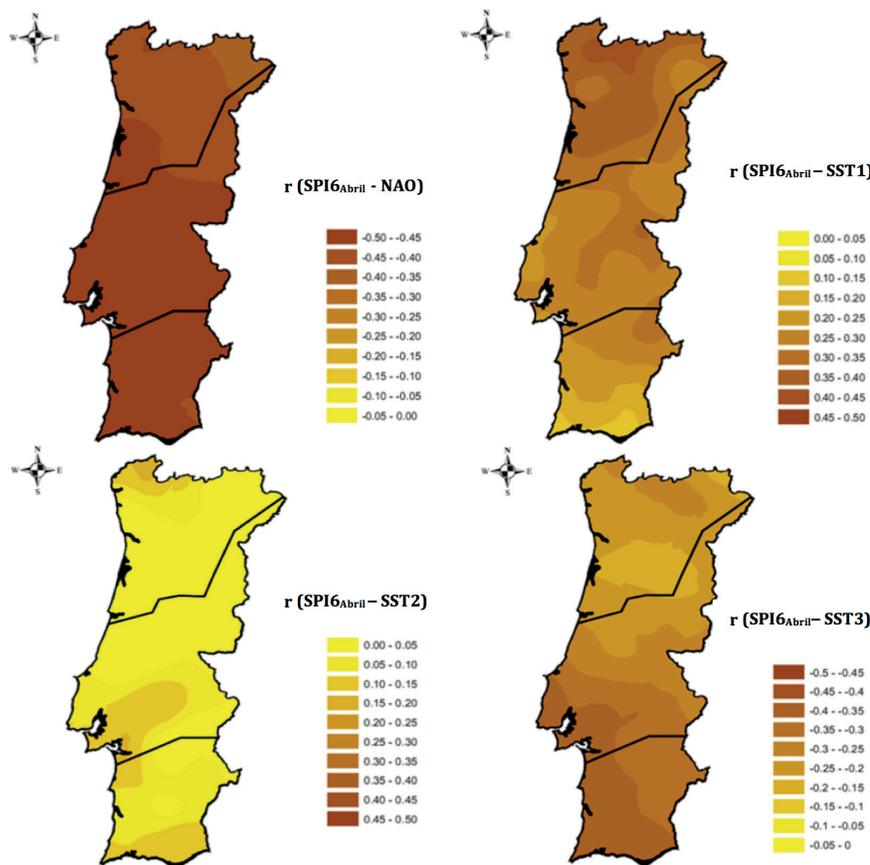


Figura 4 | Mapas da correlação linear, r , entre as séries de SPI6 primaveris ($SPI6_{Abril}$, $SPI6_{Maio}$ e $SPI6_{Junho}$) e a NAO (DJFM) e os SSTs (JFM). Período de Outubro de 1910 a Setembro de 2004 (adaptado de Santos *et al.*, 2014).

Validação dos modelos de ANNs: comparação de mapas de seca primaveril observada e prevista

Para as séries temporais escolhidas como dados de entrada das redes neuronais, apresenta-se na Tabela 2 as médias das medidas de erro para a melhor ANN apurada para cada um dos postos udométricos das três regiões apresentadas na Figura 1. Na referida tabela é descrita a arquitectura genérica das redes seleccionadas em cada caso – ex., 3: 3 – (1 a 9) – 1: 1: rede neuronal com três nós na camada de entrada, de um a nove nós na primeira camada escondida e um nó na camada de saída; notação: g, n, s . Menciona-se que os melhores resultados para cada posto foram sempre obtidos para redes com apenas uma camada escondida.

Os anteriores resultados são particularmente bons, especialmente no que respeita à previsão do $SPI6_{Abril}$ em que consistentemente se obtiveram as melhores medidas de erro, designadamente, os menores valores de RMSE e valores das restantes medidas elevados e mesmo muito próximos de 1.

Em termos gerais, para incrementos crescentes (relativamente ao último mês –Março– presente nos dados) do horizonte temporal de previsão (um, dois e três meses, respectivamente, para $SPI6_{Abril}$, $SPI6_{Maio}$ e $SPI6_{Junho}$), as ANNs prevêm as séries de SPI6 de forma progressivamente menos favorável, facto bem aparente quando se comparam os resultados da Tabela 2 sequencialmente para $SPI6_{Abril}$, $SPI6_{Maio}$ e $SPI6_{Junho}$. Com efeito, as médias das medidas de erro para o $SPI6_{Junho}$ são piores do que as obtidas para os outros dois meses anteriores da primavera ($SPI6_{Abril}$ e $SPI6_{Maio}$). Ainda assim, as medidas de erro para o $SPI6_{Junho}$ são estatisticamente aceitáveis (RMSE < 0.55; $r > 0.78$; $E1 > 0.37$; $E2 > 0.60$; $PI > 0.77$).

Tabla 2 | Média das medidas de erro para a melhor ANN em cada região e para cada SPI6 primaveril.

Arquitectura da ANN	Região	Média das medidas de erro				
		RMSE	r	E_1	E_2	PI
SPI6_{Abril}						
3: 3 – (1 a 9) – 1: 1	CL1	0.292	0.960	0.749	0.915	0.961
2: 2 – (1 a 9) – 1: 1	CL2	0.210	0.977	0.811	0.953	0.975
3: 3 – (1 a 9) – 1: 1	CL3	0.250	0.975	0.786	0.946	0.972
SPI6_{Maió}						
3: 3 – (1 a 9) – 1: 1	CL1	0.418	0.898	0.550	0.789	0.900
2: 2 – (1 a 9) – 1: 1	CL2	0.411	0.900	0.576	0.802	0.908
3: 3 – (1 a 9) – 1: 1	CL3	0.415	0.924	0.642	0.849	0.921
SPI6_{Junho}						
3: 3 – (1 a 9) – 1: 1	CL1	0.460	0.828	0.434	0.667	0.828
2: 2 – (1 a 9) – 1: 1	CL2	0.517	0.780	0.367	0.596	0.791
3: 3 – (1 a 9) – 1: 1	CL3	0.548	0.796	0.393	0.623	0.772

Os resultados alcançados indicaram uma média dos resíduos médios (\bar{X}_{res}) para as melhores ANNs de cada região praticamente igual a 0. Da análise da distribuição dos resíduos concluiu-se poder ocorrer uma ligeira assimetria, embora num número reduzido de postos, aspectos que, globalmente, se consideraram consistentes com o pressuposto de normalidade.

Na Figura 5 apresentam-se os mapas dos valores previstos de SPI6 para os meses da primavera de 2002 e 2004 fornecidos pela melhor ANN em cada posto udométrico. De entre o período de 4 anos não utilizado no estabelecimento dos modelos (2001 a 2004, como antes referido), seleccionaram-se aqueles anos para exemplificar os resultados obtidos uma vez que neles ocorreram os mais gravosos eventos de seca severa. Com efeito, 2002 foi um ano cuja seca primaveril afectou a costa norte, com uma clara recuperação, em termos de área afectada e intensidade, entre Abril e Junho. Em 2004, um ano com uma seca primaveril generalizada afectando as áreas interiores, ocorreu o inverso: a intensidade e extensão espacial de área em seca aumentaram entre Abril a Junho. Em 2002 as áreas mais afectadas foram a costa norte e uma região interior perto da fronteira com Espanha, enquanto em 2004 a seca se iniciou na região noroeste de Portugal, tendo-se expandido progressivamente para oeste e sul.

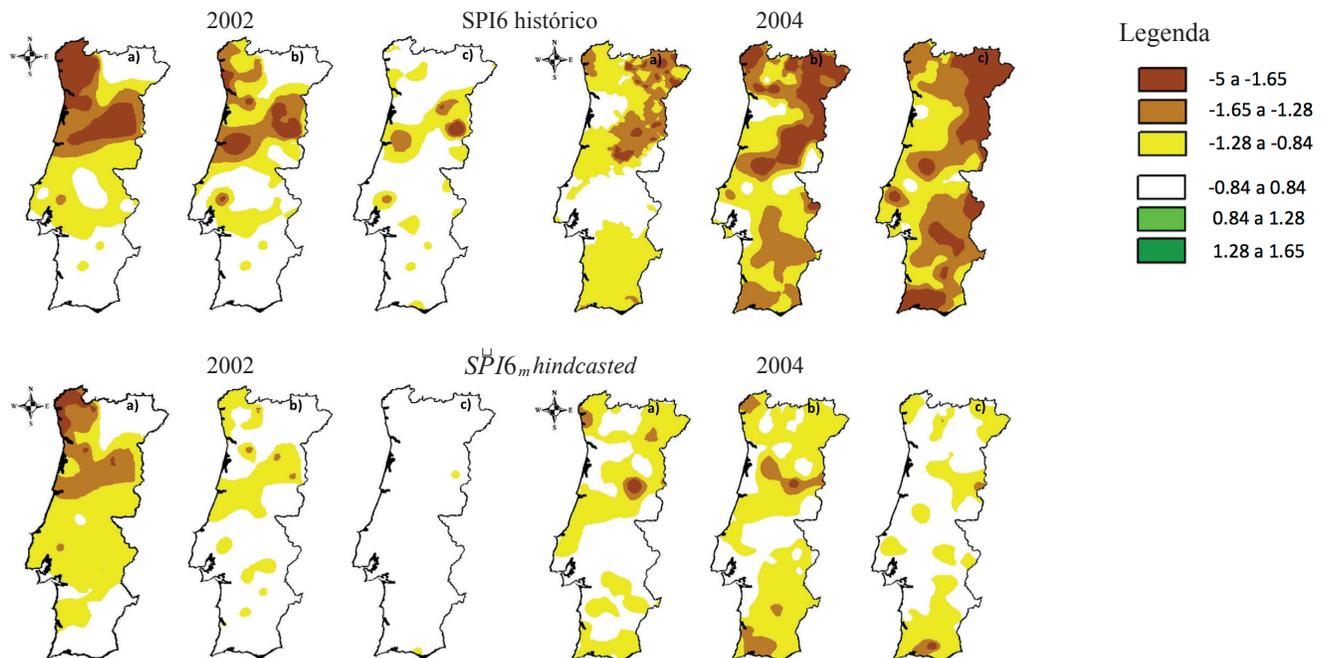


Figura 5 | Mapas de SPI6 histórico e *hindcasted* para os meses primaveris de: a) Abril, SPI6_{Abril}; b) Maio, SPI6_{Maió}; e c) Junho, SPI6_{Junho} de 2002 (esquerda) e 2004 (direita) (ver a Tabela 1 para as categorias de seca) (adaptado de Santos *et al.*, 2014).

Para Abril de 2002 as estimativas da melhor ANN produziram um padrão semelhante ao decorrente das observações. O desempenho é, contudo, significativamente mais fraco em Abril de 2004. É relevante notar que especialmente no ano de 2004, as capacidades de *hindcast* dos modelos são progressivamente menos boas para 2 e 3 meses de horizonte temporal de previsão (Maio e Junho, respectivamente). A Figura 5 indica claramente que as ANNs subestimaram as áreas em seca extrema não sendo capazes de reproduzir a extensão espacial dos eventos históricos. O menor desempenho dos modelos de ANN, na primavera de 2004, pode estar relacionado com o facto de este ter sido um ano de seca extrema, reconhecendo-se que o desempenho das ANN não é o melhor na modelação de valores extremos, designadamente, se esses valores não se encontram representados nas séries temporais utilizadas na calibração. Dawson e Wilby (1998) e Campolo *et al.* (1999) sugerem que a subestimação nos modelos de ANN pode ser atribuída precisamente a uma falta de informação fornecida aos modelos. Autores como Karunanithi *et al.* (1994), sugerem que a inclusão de um maior número de padrões de valores extremos nos dados de treino pode minorar o problema.

Mapas de probabilidade de seca: distribuição espacial e temporal das secas primaveris

A abordagem de previsão de secas desenvolvida incluiu uma componente probabilística, destinada a avaliar quantitativamente a incerteza dos modelos, de modo a permitir a adopção mais fundamentada de estratégias de mitigação do fenómeno. Para tal e tendo por base os valores previstos pelos modelos de ANN para os meses da primavera de 2002 e 2004, obtiveram-se os mapas de probabilidade apresentados na Figura 6.

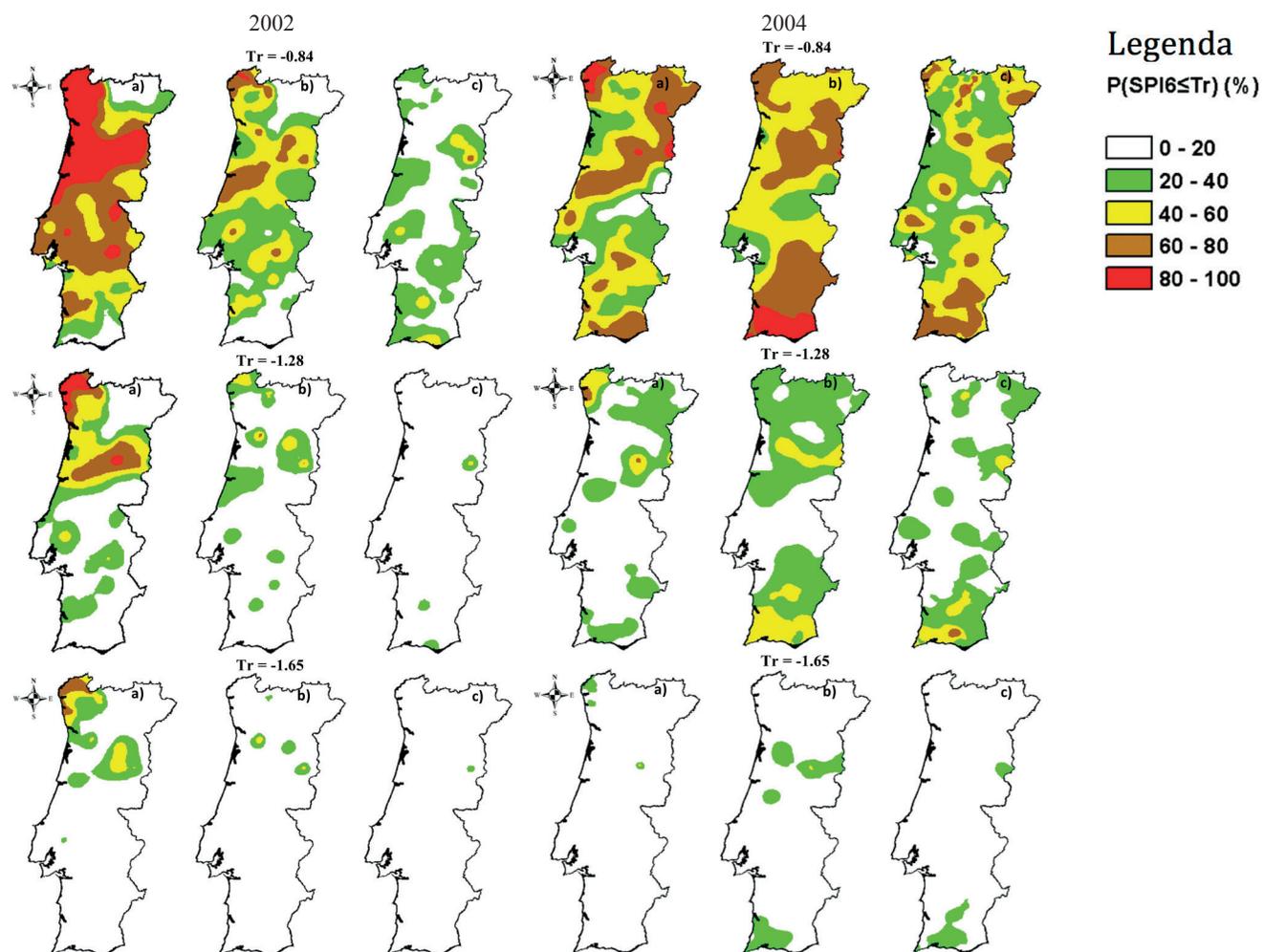


Figura 6 | Mapas de probabilidade para diferentes limiares de seca, Tr , baseados nos valores previstos de SPI6 para os meses primaveris de: a) Abril, $SPI6_{Abril}$; b) Maio, $SPI6_{Maio}$; e c) Junho, $SPI6_{Junho}$ de 2002 (esquerda) e 2004 (direita) (adaptado de Santos *et al.*, 2014).

Cada mapa da Figura 6 caracteriza espacialmente a probabilidade de ocorrer uma seca de intensidade igual ou superior a um dado limiar. Os limiares de seca, Tr , considerados foram -0.84 , -1.28 e -1.65 (categorias de seca moderada, severa e extrema, respectivamente, conforme Tabela 1).

Para os anos de 2002 e 2004 e para o limiar de -0.84 , os mapas obtidos com base nas estimativas para SPI6Abril indicam regiões onde a probabilidade de ter uma seca com SPI menor ou igual àquele limiar é muito alta, o que é consistente com o que realmente ocorreu (mapas históricos da Figura 6). Apesar de a capacidade predictiva diminuir à medida que aumenta o horizonte temporal de previsão (SPI6Maio e SPI6Junho), especialmente para a primavera de 2002, os modelos continuam a indicar seca nas zonas em que a mesma realmente ocorreu. Por outro lado, à medida que o limiar diminui, portanto, no sentido de secas mais severas, as probabilidades de ocorrer seca com intensidade igual ou superior é menor. Não obstante, os modelos conseguem sempre identificar seca nas regiões onde estas de facto se desenvolveram.

Mapas de probabilidade de seca, como os exemplificados na Figura 6, podem servir de ferramenta fundamental na previsão das ocorrências e, conseqüentemente, apoiar, inclusivamente por regiões, o planeamento e gestão integrada de recursos hídricos.

CONCLUSÕES

Um dos aspectos mais importantes no desenvolvimento de um modelo de previsão de secas é a selecção das variáveis de entrada, as quais determinam a estrutura do modelo e afectam os seus resultados. O modelo de previsão (*hindcasting*) de secas que se apresenta para os SPI6 relativos aos meses da primavera em Portugal Continental foi baseado em redes neuronais artificiais (ANN) e em informação relativa a indicadores climáticos de larga escala (índice NAO) e a temperaturas da superfície do mar (SST1 e SST3). A influência da NAO do Inverno nos valores de SPI6 relativos à primavera foi inicialmente avaliada por recurso a uma análise de correlação para cada uma das áreas previamente identificadas para o País como sendo homogéneas no que respeita a secas (norte, centro e sul), tendo sido observadas correlações significativas para o nível de significância de 5%. A NAO tem sido largamente utilizada para prever ocorrências de seca com elevado grau de fiabilidade (Cordery e McCall, 2000; Rodwell, 2003). No entanto, como alguns autores referem (López-Moreno e Vicente-Serrano, 2008), os modelos de previsão devem incluir o efeito sinérgico de diferentes variáveis de entrada e não de uma única variável. Tais recomendações são consistentes com os resultados que se obtiveram, os quais confirmam a capacidade predictiva de secas das SSTs.

A sensibilidade demonstrada pelos resultados a diferentes desfasamentos temporais entre os predictores (NAO e SSTs) e variável a ser prevista (SPIs primaveris) evidencia a necessidade de aprofundar a investigação relativa ao efeito que tais desfasamentos exercem sobre a capacidade de previsão dos modelos. De entre os autores a que, no anterior contexto, se atendeu na concepção do modelo implementado menciona-se o estudo de Trigo *et al.*, (2004), o qual identificou um intervalo temporal de 2-3 meses na resposta dos escoamentos mensais às fases negativa e positiva da NAO em bacias hidrográficas atlânticas da Península Ibérica. A literatura indica aspectos comuns a todos os modelos de previsão, nomeadamente o facto de a eficiência de previsão para escalas temporais de SPI superiores ser crescentemente melhor do que para escalas temporais de SPI inferiores e de a consideração de diferentes zonas hidrológicas não influenciar os resultados finais. Tais aspectos necessitam, contudo, de ser clarificados, especialmente no âmbito da utilização de modelos de previsão para secas.

Anota-se que a abordagem que se apresentou segue algumas das recomendações de Kurnik (2009), nomeadamente, de que a previsão de seca deve ser efectuada numa óptica probabilística e de que é necessário uma metodologia de previsão de secas integrada, utilizando indicadores hidrológicos (ex. humidade do solo, escoamentos, precipitação, etc..) e indicadores meteorológicos e climáticos (ex. SPI, PDSI, etc..).

Os gestores de recursos hídricos e as actividades agrícolas podem beneficiar fortemente de um sistema de previsão capaz de fornecer uma estimativa probabilística do evento com alguns meses de avanço, sendo que a primavera é a estação do ano mais sensível em termos de detecção do seu início. Séries temporais longas de SPI são componentes essenciais para um modelo de previsão fiável, na medida em que permitem que os modelos de redes neuronais possam aprender com o passado.

A desvantagem da utilização de modelos estatísticos para a previsão é a possível natureza não estacionária das correlações históricas entre as variáveis globais (ex. SSTs e NAO) e o clima regional (ex. SPI6) em consequência da mudança climática em

curso. Importa, assim, desenvolver investigação que permita melhorar o conhecimento de tais relações históricas e respectivos padrões, necessariamente numa óptica não-estacionária.

REFERÊNCIAS

- Agnew, C.T. (2000). Using the SPI to identify drought. *Drought Network News*, 12, 6-12.
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000a). Artificial neural networks in hydrology. I. Preliminary concepts. *Journal of Hydrologic Engineering*, 5(2), 115-123. doi:10.1061/(ASCE)1084-0699(2000)5:2(115)
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000b). Artificial neural networks in hydrology. II. Hydrologic applications. *Journal of Hydrologic Engineering*, 5(2), 124-137. doi:10.1061/(ASCE)1084-0699(2000)5:2(124)
- Bordi, I., Fraedrich, K., Petitta, M., Sutera, A. (2005). Methods for predicting drought occurrences. In *Proceedings of the 6th International Conference of the European Water Resources Association*, Menton, France.
- Bowden, G.J., Dandy, G.C., Maier, H.R. (2005). Input determination for neural network models in water resources applications. Part 1—background and methodology. *Journal of Hydrology*, 301(1-4), 75-92. doi:10.1016/j.jhydrol.2004.06.021
- Campolo, M., Andreusi, P., Soldati, A. (1999). River flood forecasting with a neural network model. *Water Resources Research*, 35(4), 1191-1197. doi:10.1029/1998WR900086
- Cancelliere, A., Di Mauro, G., Bonaccorso, B., Rossi, G. (2005). Stochastic forecasting of Standardized Precipitation Index. In *Proceedings of XXXI IAHR Congress Water Engineering for the future: Choice and Challenges*, Seoul, Korea, 3252-3260.
- Cancelliere, A., Di Mauro, G., Bonaccorso, B., Rossi, G. (2007). Drought forecasting using the Standardized Precipitation Index. *Water Resources Management*, 21(5), 801-819. doi:10.1007/s11269-006-9062-y
- Cordery, I., McCall, M. (2000). A model for forecasting drought from teleconnections. *Water Resources Research*, 36(3), 763-768. doi:10.1029/1999WR900318
- Dastorani, M.T., Afkhami, H. (2011). Application of artificial neural networks on drought prediction in Yazd (Central Iran). *Desert*, 16, 39-48.
- Dawson, D.W., Wilby, R. (1998). An artificial neural network approach to precipitation-runoff modeling. *Hydrological Sciences Journal*, 43(1), 47-66. doi:10.1080/02626669809492102
- Demyanov, V., Kanevsky, M., Chernov, S., Savelieva, E., Timonin, V. (1998). Neural network residual kriging application for climatic data. *Journal of Geographic Information and Decision Analysis*, 2(2), 215-232.
- Di Mauro, G., Bonaccorso, G.B., Cancelliere, A., Rossi, G. (2008). Use of NAO index to improve drought forecasting in the Mediterranean area: Application to Sicily region. *Options Méditerranéennes*. Série A: Séminaires Méditerranéens, No. 80.
- Fernando, M.K.G., Maier, H.R., Dandy, G.C. (2009). Selection of input variables for data driven models: An average shifted histogram partial mutual information estimator approach. *Journal of Hydrology*, 367(3-4), 165-176. doi:10.1016/j.jhydrol.2008.10.019
- Gámiz-Fortis, S., Esteban-Parra, M.J., Trigo, R.M., Castro-Díez, Y. (2010). Potential predictability of Iberian river flow based on its relationship with previous winter global SST. *Journal of Hydrology*, 385, 143-149. doi:10.1016/j.jhydrol.2010.02.010
- Gámiz-Fortis, S., Pozo-Vázquez, D., Trigo, R.M., Castro-Díez, Y. (2008a). Quantifying the predictability of winter river flow in Iberia. Part I: Interannual predictability. *Journal of Climate*, 21, 2484-2502. doi:10.1175/2007JCLI1774.1
- Gámiz-Fortis, S., Pozo-Vázquez, D., Trigo, R.M., Castro-Díez, Y. (2008b). Quantifying the predictability of winter river flow in Iberia. Part II: Seasonal predictability. *Journal of Climate*, 21, 2503-2518. doi:10.1175/2007JCLI1775.1

- Hoerling, M., Kumar, A. (2003). The perfect ocean for drought. *Science*, 299(5607), 691-694. *Geophysical Research Abstracts*, 12, EGU2010-8454, EGU General Assembly 2010, Viena, Austria. doi:10.1126/science.1079053
- Hurrell, J.W. (1995). Decadal trends in North Atlantic Oscillation: regional temperatures and precipitation. *Science*, 269(5224), 676-679. doi:10.1126/science.269.5224.676
- Hurrell, J.W., Kushnir, Y., Visbeck, M. (2001). The North Atlantic Oscillation. *Science*, 291(5504), 603-605. doi:10.1126/science.1058761
- Hurrell, J.W., Kushnir, Y., Ottersen, G., Visbeck, M. (2003). The North Atlantic Oscillation: climatic significance and environmental impact. *Geophysical Monograph Series*, 134, American Geophysical Union, Washington, DC, USA.
- Ionita, M., Lhomann, G., Rimbu, N. (2008). Prediction of spring Elbe discharge based on stable teleconnections with winter global temperature and precipitation. *Journal of Climate*, 21(23), 6215-6226. doi:10.1175/2008JCLI2248.1
- Ionita, M., Lohmann, G., Rimbu, N., Chelcea, S., Dima, M. (2012). Interannual to decadal summer drought variability over Europe and its relationship to global sea surface temperature. *Climate Dynamics*, 38(1), 363-377. doi:10.1007/s00382-011-1028-y
- Iyer, M.S., Rhinehart, R.R. (1999). A method to determine the required number of neural-network training repetitions. *IEEE Transactions on Neural Networks*, 10(2), 427-432. doi:10.1109/72.750573
- Jain, A., Kumar, A.M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, 7(2), 585-592. doi:10.1016/j.asoc.2006.03.002
- Jones, P.D., Jonsson, T., Wheeler, D. (1997). Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland. *International Journal of Climatology*, 17(13), 1433-1450. doi:10.1002/(SICI)1097-0088(19971115)17:13<1433::AID-JOC203>3.0.CO;2-P
- Jones, P.D., Osborn, T.J., Briffa, K.R. (2003). Pressure-based measures of the North Atlantic oscillation (NAO): a comparison and an assessment of changes in the strength of the NAO and in its influence on surface climate parameters in The North Atlantic Oscillation: climate significance and environmental impact. *Geophysics Monogram* 134, 51-62, American Geophysical Union.
- Karunanithi, N., Grenney, W.J., Whitely, D., Bovee, K. (1994). Neural networks for river flow prediction. *Journal of Computing Civil Engineering*, 8(2), 201-219. doi:10.1061/(ASCE)0887-3801(1994)8:2(201)
- Kim T. e Juan B. Valdés, (2003). Nonlinear Model for Drought Forecasting Based on a Conjunction of Wavelet Transforms and Neural Networks. *Journal of Hydrologic Engineering*, 8(6), 319-328. doi:10.1061/(ASCE)1084-0699(2003)8:6(319)
- Kitanidis, P.K., Bras, R.L. (1980). Real time forecasting with a conceptual hydrological model. 2. Applications and results. *Water Resources Research*, 16(6), 1034-1044. doi:10.1029/WR016i006p01034
- Kurnik, B. (2009). DESERT Action JRC, Drought forecasting methods. Ljubljana on 24 September 2009 – 1st DMCSEE – JRC Workshop on Drought Monitoring.
- Legates, D.R., McCabe Jr., G.J. (1999). Evaluating the use of ‘goodness-of-fit’ measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1), 233-241. doi:10.1029/1998WR900018
- Lloyd-Hughes, B. (2002). *The long range predictability of European drought*. PhD Thesis, Department of Space and Climate Physics, University of London, University College London, UK.
- López-Moreno, J.I., Vicente-Serrano, S.M. (2008). Extreme phases of the wintertime North Atlantic Oscillation and drought occurrence over Europe: a multi-temporal-scale approach. *Journal of Climate*, 21(6), 1220-1243. doi:10.1175/2007JCLI1739.1
- López-Moreno, J.I., Beguería, S., Vicente-Serrano, S.M., García-Ruiz, J.M. (2007). The influence of the NAO on water resources in central Iberia: precipitation, streamflow anomalies and reservoir management strategies. *Water Resources Research*, 43, W09411, doi:10.1029/2007WR005864

- Martín, M.L., Luna, M.Y., Morata, A., Valero, F. (2004). North Atlantic teleconnection patterns of low-frequency variability and their links with springtime precipitation in the western Mediterranean. *International Journal of Climatology*, 24(2), 213-230. doi:10.1002/joc.993
- Martín-Vide, J., Fernández, D. (2001). El índice NAO y la precipitación mensual en la España peninsular. *Investigaciones Geográficas*, 26, 41-58. doi:10.14198/INGEO2001.26.07
- May, R.J., Maier, H.R., Dandy, G.C., Fernando, T.M.K.G. (2008). Non-linear variable selection for artificial neural networks using partial mutual information. *Environmental Modelling and Software*, 23(10-11), 1312-1326. doi:10.1016/j.envsoft.2008.03.007
- McKee, T.B., Doesken, N.J., Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference on Applied Climatology*. American Meteorological Society, Boston, USA, 179-184.
- Mishra, A.K., Desai, V.R. (2006). Drought forecasting using feed-forward recursive neural network. *Ecological Modelling*, 198(1-2), 127-138. doi:10.1016/j.ecolmodel.2006.04.017
- Mo, K.C., Jae-Kyung, E., Schemm, E., Yoo, S.-H. (2009). Influence of ENSO and the Atlantic multi-decadal Oscillation on drought over the United States. *Journal of Climate*, 22, 5962-5982. doi:10.1175/2009JCLI2966.1
- Mutlu, E., Chaubey, I., Hexmoor, H., Bajwa, S.G. (2008). Comparison of artificial neural network models for hydrologic predictions at multiple gauging stations in an agricultural watershed. *Hydrological Processes*, 22(26), 5097-5106. doi:10.1002/hyp.7136
- Michie, D., Spiegelhalter, D.J., Taylor, C.C. (1994). *Machine learning, neural and statistical classification*. Project StatLog, Department of Statistics, University of Leeds, UK.
- Ochoa-Rivera, J.C., García-Bartual, R., Andreu, J. (2002). Multivariate synthetic streamflow generation using a hybrid model based on artificial neural networks. *Journal of Hydrology and Earth System Sciences*, 6(4), 641-654. doi:10.5194/hess-6-641-2002
- Ochoa-Rivera, J.C., García-Bartual, R., Andreu, J. (2007). Influence of Inflows Modeling on Management Simulation of Water Resources System. *Journal of Water Resources Planning and Management*, ASCE, 133(2), 106-116. doi:10.1061/(ASCE)0733-9496(2007)133:2(106)
- Portela, M.M., Quintela, A.C. (2006). Estimación em Portugal Continental de escoamento e de capacidades úteis de albufeiras de regularização na ausência de informação. *Recursos Hídricos*, 27(2), 7-18.
- Pulido-Calvo, I., Portela, M.M. (2007). Application of neural approaches to one-step daily flow forecasting in Portuguese watersheds. *Journal of Hydrology*, 332(1-2), 1-15. doi:10.1016/j.jhydrol.2006.06.015
- Pulido-Calvo, I., Gutiérrez-Estrada, J.C., Savic, D. (2012). Heuristic modelling of the water resources management in the Guadalquivir River Basin, Southern Spain. *Water Resources Management*, 26(1), 185-209. doi:10.1007/s11269-011-9912-0
- Qian, B., Corte-Real, J.M., Xu, H. (2000a). Is the North Atlantic Oscillation the most important atmospheric pattern for precipitation in Europe? *Journal of Geophysical Research*, 105(D9), 901-910. doi:10.1029/2000JD900102
- Qian, B., Xu, H., Corte-Real, J.M. (2000b). Spatial-temporal structures of the quasi-periodic oscillations in precipitation over Europe. *International Journal of Climatology*, 20(13), 1583-1598. doi:10.1002/1097-0088(20001115)20:13<1583::AID-JOC560>3.0.CO;2-Y
- Rodwell, M.J. (2003). On the predictability of the North Atlantic climate. The North Atlantic Oscillation: climate significance and environmental impact, *Geophysical Monograph*, 134, 173-192, Amer. Geophys. Union. doi:10.1029/134GM08
- Rossi, G. (2003). Requisites for a drought watch system. In: G. Rossi et al. (eds), *Tools for Drought Mitigation in Mediterranean Regions*, pp. 147-157. Kluwer Academic Publishing: Dordrecht. doi:10.1007/978-94-010-0129-8_9
- Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536. doi:10.1038/323533a0

- Santos, J.A., Corte-Real, J., Leite, S.M. (2005). Weather regimes and their connection to the winter precipitation in Portugal. *International Journal of Climatology*, 25(1), 33-50. doi:10.1002/joc.1101
- Santos, J.F., Portela, M.M., Pulido-Calvo, I. (2011). Regional frequency analysis of droughts in Portugal. *Water Resources Management*, 25(14), 3537-3558. doi:10.1007/s11269-011-9869-z
- Santos, J.F., Portela, M.M., Pulido-Calvo, I. (2013). Dimensionality reduction in drought modelling. *Hydrological Processes*, 27(10), 1399-1410. doi:10.1002/hyp.9300
- Santos, J.F., Portela, M.M., Pulido-Calvo, I., (2014). Spring drought prediction based on winter NAO and global SST in Portugal, *Hydrological Processes*, 28(3), 1009-1024. doi:10.1002/hyp.9641
- Santos, J.F., Pulido-Calvo, I., Portela, M.M. (2010). Spatial and temporal variability of droughts in Portugal. *Water Resources Research*, 46(3). DOI: 10.1029/2009WR008071. doi:10.1029/2009WR008071
- Senthil-Kumar, A.R., Sudheer, K.P., Jain, S.K., Agarwal, P.K. (2005). Rainfall-runoff modelling using artificial neural networks: comparison of network types. *Hydrological Processes*, 19(6), 1277-1291. doi:10.1002/hyp.5581
- Silva, A.T., Portela, M.M., Naghettini, M. (2012), Nonstationarities in the occurrence rates of flood events in Portuguese watersheds. *Journal of Hydrology and Earth System Sciences*, 16, 241-254. doi:10.5194/hess-16-241-2012
- Smith, T.M., Reynolds, R.W., Peterson, T.C. Lawrimore, J. (2008). Improvements to NOAA's Historical Merged Land-Ocean Surface Temperature Analysis (1880-2006). *Journal of Climate*, 21, 2283-2296. doi:10.1175/2007JCLI2100.1
- Snedecor, G.W., Cochran, W.G. (1989). *Statistical methods*, Ames, Iowa State University Press (8th edition), Iowa, USA.
- Trigo, R.M., Osborn, T.J., Corte-Real, J.M. (2002). The North Atlantic Oscillation influence on Europe. Climate impacts and associated physical mechanisms. *Climate Research*, 20, 9-17. doi:10.3354/cr020009
- Trigo, R.M., Pozo-Vázquez, D., Osborn, T.J., Castro-Díez, Y., Gámiz-Fortis, S., Esteban-Parra, M.J. (2004). North Atlantic Oscillation influence on precipitation, river flow and water resources in the Iberian Peninsula. *International Journal of Climatology*, 24(8), 925-944. doi:10.1002/joc.1048
- Trigo, R., Xoplaki, E., Zorita, E., Luterbacher, J., Krichak, S.O., Alpert, P., Jacobeit, J., Sáenz, J., Fernández, J., González-Rouco, F., García-Herrera, R., Rodo, X., Brunetti, M., Nanni, T., Maugeri, M., Trkes, M., Gimeno, L., Ribera, P., Brunet, M., Trigo, I.F., Crepon, M., Mariotti, A. (2006). Relations between Variability in the Mediterranean region and mid-latitude variability. In: *Mediterranean Climate Variability*, edited by: Lionello P., Malanotte-Rizzoli P., e R. Boscolo. Amsterdam, Elsevier, 179-226. doi:10.1016/s1571-9197(06)80006-6
- Vicente-Serrano, S.M., López-Moreno, J.I., Lorenzo-Lacruz, J., El Kenawy, A., Azorin-Molina, C., Morán-Tejeda, E., Pasho, E., Zabalza, J., Beguería, S., Angulo-Martínez, M. (2011). The NAO impact on droughts in the Mediterranean region. In: Vicente-Serrano S.M. e Trigo R. (Eds.), *Hydrological, socioeconomic and ecological impacts of the North Atlantic Oscillation in the Mediterranean region*. Advances in Global Research (AGLO) series, Springer-Verlag. doi:10.1007/978-94-007-1372-7_3
- Vinther, B.M., Andersen, K.K., Hansen, A.W., Schmith, T., Jones, P.D. (2003). Improving the Gibraltar/Reykjavik NAO Index. *Geophysical Research Letters*, 30(23), 2222. doi:10.1029/2003GL018220
- Xoplaki E., González-Rouco J.F., Luterbacher J. e H. Wanner, (2004). Wet season Mediterranean precipitation variability: influence of large-scale dynamics and predictability. *Climate Dynamiques* 23, 63-78.
- Xue, Y., Smith, T.M., Reynolds, R.W. (2003). Interdecadal changes of 30-yr SST normals during 1871-2000. *Journal of Climate*, 16, 1601-1612. doi:10.1175/1520-0442-16.10.1601
- Yevjevich, V. (1972). *Stochastic Processes in Hydrology*. Water Resources Publications, Fort Collins, Co.

