Physics and Chemistry of the Earth xxx (2015) 1–10

Contents lists available at ScienceDirect

Physics and Chemistry of the Earth

journal homepage: www.elsevier.com/locate/pce

SPI drought class prediction using log-linear models applied to wet and dry seasons

Elsa E. Moreira [a, b, *]

[a] Center of Mathematics and Applications, Faculty of Sciences and Technology, Nova University of Lisbon, 2829-516, Caparica, Portugal
[b] Research Center for Landscape, Environment, Agriculture and Food, Institute of Agronomy, University of Lisbon, 1349-017, Lisbon, Portugal

A R T I C L E I N F O

Article history:
Received 24 March 2015
Received in revised form
13 August 2015
Accepted 28 October 2015
Available online xxx

Keywords:
Quasi-association log-linear models
Drought class transitions
Odds
Confidence intervals

A B S T R A C T

A log-linear modelling for 3-dimensional contingency tables was used with categorical time series of SPI drought class transitions for prediction of monthly drought severity. Standardized Precipitation Index (SPI) time series in 12- and 6-month time scales were computed for 10 precipitation time series relative to GPCC datasets with 2.5° spatial resolution located over Portugal and with 112 years length (1902–2014). The aim was modelling two-month step class transitions for the wet and dry seasons of the year and then obtain probability ratios – Odds – as well as their respective confidence intervals to estimate how probable a transition is compared to another. The prediction results produced by the modelling applied to wet and dry season separately, for the 6- and the 12-month SPI time scale, were compared with the results produced by the same modelling without the split, using skill scores computed for the entire time series length. Results point to good prediction performances ranging from 70 to 80% in the percentage of corrects (PC) and 50–70% in the Heidke skill score (HSS), with the highest scores obtained when the modelling is applied to the SPI12. The adding up of the wet and dry seasons introduced in the modelling brought improvements in the predictions, of about 0.9–4% in the PC and 1.3–6.8% in the HSS, being the highest improvements obtained in the SPI6 application.

1. Introduction

The hazard and disaster nature of droughts makes it important to develop prediction tools, including probabilistic ones, which may support early warning for timely implementation of preparedness and mitigation measures (Wilhite et al., 2000). Droughts have a slow initiation and they are usually only recognized when the drought is already established. They are of long duration, and usually affect large areas. Forecasting of when a drought is likely to begin or to come to an end is extremely difficult. An adequate period between the release of the prediction and the actual onset of the predicted drought hazard makes it possible for decision and policy makers to timely implement policies and measures to mitigate the effects of drought. (Nichols et al., 2005). The short time drought predictions (from 1 to 3 months) are important for warning farmers about the probable initiation or establishment of a drought, about its continuation or its probable termination in a few months. Short time drought predictions may also be used to alert water managers and decision or policy makers about the need to enforce appropriate preparedness measures before a drought is effectively installed, or to prepare for a post-drought period.

Drought indices are numerical indicators for assessing drought severity, producing also categorical information Examples are the Standardized Precipitation Index (SPI), the Palmer Drought Severity Index (PDSI) and the MedPDSI (McKee, 1993; Palmer, 1965; Pereira et al., 2006; Paulo et al., 2012). The SPI is widely used for the identification of drought events and to evaluate their severity through well-defined drought and wet classes. It may be computed on shorter or longer time scales, which reflect different lags in the response of the water cycle to precipitation anomalies (Steinemann et al., 2005). In addition, due to the standardization, it allows a reliable comparison between different locations and climates (Mishra and Singh, 2010). The stochastic properties of the SPI time series have been explored for a while by the author for analyzing and predicting droughts (Moreira et al., 2008, 2012).

Atmospheric circulation patterns governing wet and dry rainfall regimes in Europe are the NAO (North Atlantic oscillation) and AO (Artic oscillation), which correlate with precipitation time series.

* Center of Mathematics and Applications, Faculty of Sciences and Technology, Nova University of Lisbon, 2829-516, Caparica, Portugal.
E-mail address: efnm@fct.unl.pt.

http://dx.doi.org/10.1016/j.pce.2015.10.019
1474-7065/

Pires and Perdigão (2007) showed high levels of correlation between the NAO and SPI, introducing the NAO as an interesting tool for drought predictions. Combining the stochastic properties of the SPI with weather regimes indices is a challenge of current research for predicting droughts.

Many statistical and non-statistical techniques as well as the combination of both have been proposed to forecast droughts. Examples are the drought forecasting based on the ARIMA models (Han et al., 2010; Mishra and Desai, 2009) and neural networks (Bierkens and Van Beek, 2009; Mishra and Desai, 2006). Hybrid models that combine two techniques are also applied by Mishra et al. (2007), which uses a hybrid stochastic and neural network model and by Kim and Valdes (2003) that conjugate wavelet transforms and neural networks for a nonlinear model. Other techniques for long term drought forecasting are for instance the model combination of wavelet and fuzzy logic model used in Ozger et al. (2012) and the adaptive neuro-fuzzy inference used in Bacanli et al. (2009). Farokhnia et al. (2011) also used data mining and ANFIS techniques for drought forecast. In Mishra and Singh (2011) a review of different methodologies used so far for drought modelling is presented.

Approaches of drought forecasting (through the indices of SPI, PDSI, etc.) common in hydrological and engineering applications use the stochastic, space-temporal properties of the continuous or of the categorized time-series. In particular approaches that use drought indices and hydro-meteorological variables as atmospheric-oceanic indices (e.g. NAO, AO, ENSO) have been suggested for drought prediction at monthly and seasonally scale. An example is the one presented by Morid et al. (2007), which uses artificial neural networks and time series of drought indices additionally forced the NAO index. In a recent work by Bonaccorso et al. (2015), probabilistic models for short and middle term forecasting of SPI drought class transition including information provided by the NAO index, were used. In Ribeiro and Pires (2015), hybrid techniques that combine the statistics and the dynamics applied to SPI time series are used for seasonal drought predictability.

Regarding the forecasting of drought class transition probabilities, the Markov models have been widely applied. For instance, a non-homogeneous Markov chain model has been applied by Lohani and Loganathan (1997) and Lohani et al. (1998) to drought classes identified by means of the Palmer index. Paulo and Pereira (2007) used homogeneous Markov models to predict drought class transitions probabilities of SPI at 12-month time scale. Besides Markov chains, other models have been employed to predict drought class transition probabilities. 3-D log-linear models, which are one step further advanced than the Markov chains are used by Moreira et al. (2008) to predict drought class transitions one and two months ahead, given the drought class for the last two months. Kavalieratou et al. (2012) also used the same approach for short-term forecast.

A probabilistic approach was also used in Cancelliere et al. (2007), which analytically derived the approximate probabilities for drought class transitions by assuming a multivariate normal distribution for the SPI time series and Bonaccorso et al. (2015) used probabilistic models that result from evaluating conditional probability of future SPI classes with respect to current SPI and NAO classes.

Despite all that efforts, forecasting when a drought is likely to begin or to come to an end is still a difficult task. The log-linear modelling of drought class transitions for monthly drought prediction developed by Moreira et al. (2008) is an innovative approach with potential to be improved and was to be later combined with other prediction tools of different nature, including weather regime transitions. The technique of log-linear modelling (Agresti, 1990) is based upon the computation of 3-dimension contingency tables of drought class transitions counts, corresponding two-time step transitions (the drought class at month $t-1$, t and $t+1$) obtained from SPI time series. The log-linear model for 3 categories fits all the transitions frequencies between drought classes and then, ratios of expected frequencies (odds) giving the most probable transition for the next month, are computed as well as their confidence intervals. This approach also allows predicting with a leading time of two or more months.

In the current paper, the work developed by Moreira et al. (2008) is continued; the modelling is now applied to SPI12 and SPI6 and to wet and dry season's months separately in order to obtain improvements in the predictions produced by the previous modelling, which already shown its potentialities. Moreover, differently from past studies, long precipitation time series with more than 100 years are used. The used of long records has clear advantages in model fitting; allows better estimates for the transition probabilities, resulting in better predictions. Also in this work, it is presented an evaluation of the model performance using skill scores (Wilks, 2006; Jolliffe and Stephenson, 2003), unlike the former work where the predictions were only empirically evaluated. The predictions produced by both modelling, with and without the split into wet and dry seasons are compared in order to see if a real improvement exists.

2. Data, drought classes and weather seasons

Input data to this study consists of SPI monthly values in a 12-month (SPI12) and 6-month (SPI6) time scale, computed from the precipitation GPCC dataset with 2.5° spatial resolution and with 112 years length (1902–2014), resulting 10 grid points over mainland Portugal (Fig. 1). The GPCC dataset was used in this study because the time series of monthly precipitation observations available for locations in Portugal are not updated since 2006 and it is important in this kind of modelling to have as many and recent observations as possible in order to be able to fit well the model and obtain good estimates for the model parameters.

The GPCC Data is a gauge-based gridded monthly precipitation dataset for the global land surface, available in 2.5°, 1° and 0.5° spatial resolutions. The dataset is based on both non real-time and real-time stations (Schneider et al., 2010). GPCC monthly precipitation analysis products are based on anomalies from climatological normals at the stations, or from GPCC high-resolution gridded climatology where no station normal is available. The GPCC precipitation climatology consists of normals collected by WMO, delivered by the countries to GPCC, or calculated from time series of monthly data available in the GPCC data base (for details see the GPCC annual reports at http://gpcc.dwd.de). Raziei et al. (2011) for instance assessed the spatial and temporal variability of drought over Iran using the GPCC dataset and found satisfactory agreement with observations.

The SPI with 12-month time scale, as well as larger time scales, identifies anomalous dry and wet periods of relatively long duration and relates well with impacts of drought on the hydrologic regimes and water resources of a region (Vicente-Serrano, 2006), or the effects of rain fluctuations over short intervals (Mishra and Singh, 2010). Shorter time scales of less than 6 months are more useful to detect agricultural droughts; while longer ones, larger than 24 months, may be useful to consider impacts on groundwater resources. For the Portuguese conditions, where a dry summer period of near 6 months occurs, droughts impacting the hydrologic regime are better assessed when using the SPI with 12-month time scale (Paulo and Pereira, 2006; Santos et al., 2010).

Former studies (Moreira et al., 2006, 2008; 2012) showed that the use of the SPI with a 12-month time scale tend to reflect more the persistence of the droughts instead of the change, thus

Please cite this article in press as: Moreira, E.E., SPI drought class prediction using log-linear models applied to wet and dry seasons, Physics and Chemistry of the Earth (2015), http://dx.doi.org/10.1016/j.pce.2015.10.019

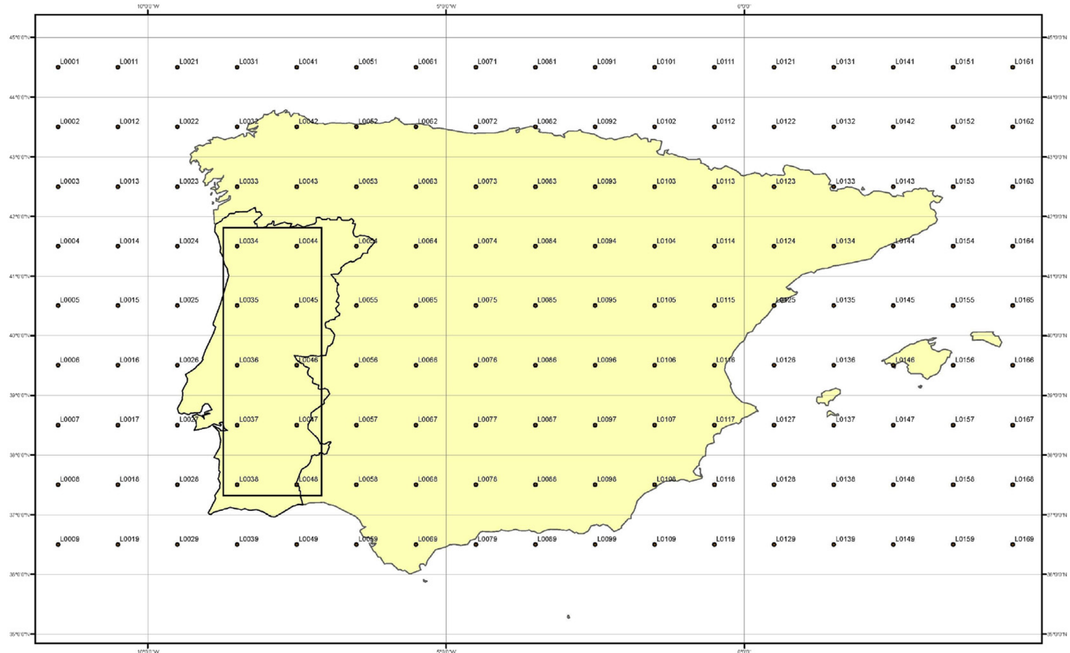

Fig. 1. Selected grid locations in Portugal with a resolution of 2.5 × 2.5° in longitude and latitude.

evidencing the self-perpetuating characteristic of the droughts (Moreira et al., 2008). In order to obtain a little more sensitivity to change of drought classes, SPI time series with 6-month time scale were also used in this study.

The analysis of the prediction results produced by the modelling developed in Moreira et al. (2008), show that the disagreements between the model and the observations, often refer to conditions when a decrease or increase of the drought class category breaks with the drought class established in the preceding two months, which correspond to an increase or decrease in rainfall for the months under prediction. This many times occurs with the beginning or during a new rainy season after several months in severe drought. Soon, it was thought that the introduction of a new category in the contingency tables representing the season of the year could bring some improvement in the predictions. The precipitation in Portugal occurs predominantly during the autumn and winter, and spring and summer are in general dry. As a result two main seasons were define: the wet one, including the months from October to March and the dry one, including the months from April to September.

From the SPI time series, categorical time series of monthly drought classes were then computed based on Table 1. The severity of drought classes adopted are defined in Table 1, which is a modification of those proposed by McKee et al. (1993, 1995) by grouping the severe and extremely severe drought classes. This modification was done for modelling purposes since transitions referring to the extremely severe drought classes are much less frequent than for other classes; thus, a possible bias is avoided since too many zeros in the contingency tables may cause problems in the fitting.

Table 1
Drought class classification of SPI (modified from McKee et al. (1993)).

Code	Drought classes	SPI values
1	Non-drought	SPI ≥ 0
2	Near normal	−1 < SPI < 0
3	Moderate	−1.5 < SPI ≤ −1
4	Severe/Extreme	SPI ≤ −1.5

3. Methods

For the modelling, the number of two-step monthly transitions between any drought class was counted for the wet and dry seasons to form two 3-dimensional (4 × 4 × 4) contingency tables with N = 64 cells each one. The two contingency tables have 3 categories: the drought class at month $t-1$, t and $t+1$ with 4 levels (drought class 1,2,3,4) each one. An example of these contingency tables is presented in Tables 3 and 4. If the season at month t is the wet one then the transition is counted for Table 3 (wet), else it is counted for Table 4 (dry).

The observed frequency, n_{ijk},i,j,k = 1,…,4, reported in the contingency tables, consists of the number of times that in a month a drought class i was followed by a drought class j in the next month and then by a drought class k in the month after that (two step transitions).

3.1. 3-Dimensions log-linear models

In the previous study by Moreira et al. (2008), quasi-association (QA) log-linear models for 3 dimension (3D) contingency tables were fitted (Agresti, 1990). This type of log-linear models, after testing several other models types were selected as the best ones to fit this kind of contingency tables. In this work, QA log-linear models for 4 categories were tested to fit the 4D contingency tables gathering the two 3D contingency tables for wet and dry

Table 2
Contingency table for the prediction of 4 drought classes for computing the Heidke Skill Score.

Drought classes	Observed				Marginal total
Predicted	1	2	3	4	
1	p_{11}	p_{12}	p_{13}	p_{14}	$p'_1 = \Sigma p_{1k}$
2	p_{21}	p_{22}	p_{23}	p_{21}	$p'_2 = \Sigma p_{2k}$
3	p_{31}	p_{32}	p_{33}	p_{34}	$p'_3 = \Sigma p_{3k}$
4	p_{41}	p_{42}	p_{43}	p_{44}	$p'_4 = \Sigma p_{4k}$
Marginal total	$p_1 = \Sigma p_{i1}$	$p_2 = \Sigma p_{i2}$	$p_3 = \Sigma p_{i3}$	$p_4 = \Sigma p_{i4}$	100%

Please cite this article in press as: Moreira, E.E., SPI drought class prediction using log-linear models applied to wet and dry seasons, Physics and Chemistry of the Earth (2015), http://dx.doi.org/10.1016/j.pce.2015.10.019

Table 3
Three-dimensional contingency table for two consecutive transitions between drought classes of SPI with 6-months time-scale computed for the location L0034 in the northwest of Portugal (see Fig. 1).

Drought class month t − 1	Drought class month t				Drought class month t				Drought class month t				Drought class month t			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Wet season	Drought class month t + 1				Drought class month t + 1				Drought class month t + 1				Drought class month t + 1			
	1				2				3				4			
1	223	18	1	0	38	33	1	0	4	4	1	1	0	1	2	0
2	49	35	1	0	19	78	7	2	1	19	11	3	0	3	9	2
3	4	2	1	0	1	20	8	3	0	0	5	3	0	0	5	9
4	2	2	1	1	0	7	4	3	0	1	5	5	0	0	1	13
Dry season	Drought class month t + 1				Drought class month t + 1				Drought class month t + 1				Drought class month t + 1			
	1				2				3				4			
1	233	22	0	0	44	28	1	0	1	9	1	0	1	4	0	0
2	34	35	1	0	17	98	7	0	1	12	4	0	0	4	6	5
3	0	5	0	0	0	16	7	2	0	1	3	3	0	0	4	8
4	0	1	0	1	0	4	9	5	0	0	3	9	0	0	1	22

Table 4
Three-dimensional contingency table for two consecutive transitions between drought classes of SPI with 12-months time-scale computed for the location L0034 in the northwest of Portugal (see Fig. 1).

Drought class month t − 1	Drought class month t				Drought class month t				Drought class month t				Drought class month t			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Wet season	Drought class month t + 1				Drought class month t + 1				Drought class month t + 1				Drought class month t + 1			
	1				2				3				4			
1	249	13	0	0	39	34	0	0	0	2	1	0	0	1	0	0
2	37	23	0	0	9	111	8	2	0	14	11	1	0	2	4	1
3	3	3	0	0	0	17	6	0	0	1	10	5	0	0	8	11
4	1	0	2	1	0	4	4	3	0	2	5	8	0	0	3	13
Dry season	Drought class month t + 1				Drought class month t + 1				Drought class month t + 1				Drought class month t + 1			
	1				2				3				4			
1	312	0	0	0	10	13	0	0	1	0	0	0	0	0	0	0
2	8	19	0	0	5	175	4	0	0	13	5	0	0	1	0	0
3	0	1	1	0	0	5	5	1	0	0	31	1	0	0	6	6
4	0	0	0	0	0	0	0	0	0	0	4	6	0	0	3	36

season, where the forth category is the season (wet or dry). However the results were worse than considering two separate models for dry and wet seasons. In consequence, two simplest 3D log-linear models instead of a more complex one were considered a best option.

Denoting by m_{ijk} the mean value $E(n_{ijk})$ of n_{ijk}, $i,j,k = 1,\ldots,4$, also called expected frequency, the QA log-linear model is described by

$$\log m_{ijk} = \lambda + \lambda_i^a + \lambda_j^b + \lambda_k^c + \beta ij + \alpha ik + \eta jk + \tau ijk + \delta_{1i} I(i=j) + \delta_{2i} I(i=k) + \delta_{3j} I(j=k) + \delta_{4i} I(i=j=k) \quad (1)$$

where λ is the constant parameter also designated by grand mean; λ_i^a is the effect of i level of category A (drought class at month t − 1), $i = 1,\ldots,4$; λ_j^b is the effect of j level of category B (drought class at month t) $j = 1,\ldots,4$; λ_k^c is the effect of k level of category C (drought class at month t + 1) $k = 1,\ldots,4$; β, α, η, τ are the linear association parameters between categories; $\delta_{1,i}$, $\delta_{2,i}$, $\delta_{4,i}$ are parameters associated with the i-th diagonal element of category A; $\delta_{3,j}$ is associated with the j-th diagonal element of category B; and *I(condition)* takes the value 1 when the condition holds and the value 0 otherwise. The expected frequencies m_{ijk} represent the expected number of two successive transitions between drought classes i, j, and k in two consecutive months during the study period.

The QA log-linear models allow linear-by-linear association of the main diagonal of the contingency tables and are adequate to fit tables where the number of levels per category is the same, with ordered categories, resulting from a pairwise comparison of dependent samples, which is the case (Agresti, 1990). In adjusting these models, it is assumed that the n_{ijk}, $i,j,k = 1,\ldots,4$ are values taken by independent Poisson distributed variables and the parameters estimators $\widehat{\lambda}$, $\widehat{\lambda}_i^a$, $\widehat{\lambda}_j^b$, $\widehat{\lambda}_k^c$, $\widehat{\beta}$, $\widehat{\delta}_h$ and $\widehat{m}_{hj}$, $h,j = 1,\ldots,4$, obtained using the maximum likelihood method, are asymptotically normally distributed (Agresti, 1990). The assumption of independency of the n_{ijk}, $i,j,k = 1,\ldots,4$ could be considered because transitions between drought classes in successive months mainly depend on the amount of precipitation occurring in those months, not on transitions in previous months (Paulo and Pereira, 2007).

Not all the parameters in the model are linearly independent because of the constraint

$$\sum_{i=1}^{4} \lambda_i^a = \sum_{j=1}^{4} \lambda_j^b = \sum_{k=1}^{4} \lambda_k^c \quad (2)$$

is required in this kind of modelling in order to make the parameters identifiable (Agresti, 1990). As a result, it was taken $\lambda_1^a = \lambda_1^b = \lambda_1^c = 0$, which simplifies the model as for previous studies in the same area (Moreira et al., 2006, 2008, 2012).

To ease the computations, from now on, matrix notation will be used. The linearly independent parameters in the model are 30. $(\lambda, \lambda_2^a, \lambda_3^a, \lambda_4^a, \lambda_2^b, \lambda_3^b, \lambda_4^b, \lambda_2^c, \lambda_3^c, \lambda_4^c, \alpha, \beta, \eta, \tau, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \delta_{21}, \delta_{22}, \delta_{23}, \delta_{24}, \delta_{31}, \delta_{32}, \delta_{33}, \delta_{34}, \delta_{41}, \delta_{42}, \delta_{43}, \delta_{44})$ and they constitute the parameter vector $\boldsymbol{\theta}$, where, for instance, $\theta_2 = \lambda_2^a$. The corresponding maximum likelihood estimators of the parameters constitute the vector $\widehat{\theta}$. Moreover, $\mathbf{n}$, $\mathbf{m}$ are, respectively, the vectors of observed frequencies, expected frequencies and logarithms of the expected frequencies, all ordered according to the index $s = 16i + 4j + k - 20$. This ordering is required because the QA log-linear models have to be rewritten in a matrix notation for computational purposes. The model matrix $\mathbf{X}$, containing known constants is derived from eq. (1) and is a 64x30 matrix. Let designate the 64 components row vectors of $\mathbf{X}$ by $\mathbf{x}_s$ with $s = 1,\ldots,64$. This matrix $\mathbf{X}$ is the same for all contingency tables because it relates with the QA log-linear model and not depends on the data set. The QA log-linear model in matrix notation is then

$$\log \mathbf{m} = \mathbf{X}\boldsymbol{\theta} \tag{3}$$

Considering that a rather long time span was used, it may be assumed that the vector $\widehat{\theta}$ of estimates has a normal distribution with mean value $\boldsymbol{\theta}$ and with the variance-covariance matrix

$$\mathbf{COV} = (\mathbf{X}^T\mathbf{D}(\widehat{\mathrm{m}})\mathbf{X})^{-1} \tag{4}$$

where $\mathbf{D}(\widehat{\mathbf{m}})$ is the diagonal matrix and (-1) represent the inverse of the matrix, whose diagonal principal elements are the adjusted expected frequencies (Agresti, 1990). Moreover, the vector $\widehat{\boldsymbol{\theta}}$ is independent from the residual deviance

$$G^2 = 2\sum_i \sum_j \sum_k n_{ijk} Log\left(\frac{n_{ijk}}{m_{ijk}}\right) \tag{5}$$

obtained when adjusting a log-linear model to a contingency table. G^2 is asymptotically distributed as a central Chi-Square with 34 degrees of freedom, since there are 64 cells in the contingency tables and 30 linearly independent parameters to be adjusted. In the expression of G^2 the frequencies are also ordered according to the index s.

To validate the adjustment of the model, the Chi-Square test with statistic G^2 may be used (Nelder, 1974; Agresti, 1990). The null hypothesis that the model fits well data is not rejected for those models having a residual deviance not exceeding the Chi-Square quantile for a probability 1-$\alpha = 0.95$ and the correspondent degrees of freedom. In other words, all the models presenting a test p-value exceeding the chosen significance level of $\alpha = 0.05$ are considered well fitted. In Table 5 are presented the models residual deviances and correspondent test p-values.

3.2. Odds and respective confidence intervals

Odds are ratios of expected frequencies that indicate the proportion between the probabilities of occurrence for two different events and assume values from 0 to $+\infty$. Here, odds represent the number of times that it is more, less, or equally probable the occurrence of a drought class transition over another and are defined as

$$Odds_{kl|ij} = \frac{m_{ijk}}{m_{ijl}}, k \neq l, \text{ and } i, j, k, l = 1, \ldots, 4 \tag{6}$$

meaning that, one month from now, it is "$Odds_{kl|ij}$" times more, less, or equally (according with the value obtain for the odds) probable that a specific location will be in class k instead of class l, given that at month t (present) is in class j, and at month $t-1$ (past) was in class i.

Considering the Eq. (3), the logarithms of the expected frequencies are

$$\widehat{z}_S = \mathbf{x}_s^T \widehat{\boldsymbol{\theta}}, s = 1, \ldots, 64 \tag{7}$$

which are also normal distributed with variance

$$V(z_S) = \mathbf{x}_s^T \left(\mathbf{X}^T\mathbf{D}(\widehat{\mathbf{m}})\mathbf{X}\right)\mathbf{x}_s, s = 1, \ldots, 64. \tag{8}$$

For large samples, $Odds_{kl|ij}$ have asymptotic normal distribution and the logarithmic transform $Log\ Odds_{kl|ij} = Log\ m_{ijk} - Log\ m_{ijl}$ converges more rapidly to a normal distribution. Using the ordering of the expected frequencies by index s = 1,…,64, we obtain $Log\ Odds_{kl|ij} = \widehat{z}_{S1} - \widehat{z}_{S2}$, where s1 and s2 correspond to the class transitions ijk and ijl respectively. Then the variance for the log odds is given by

$$V\left(\log Odds_{kl|ij}\right) = (\mathbf{x}_{s1} - \mathbf{x}_{s2})^T \left(\mathbf{X}^T\mathbf{D}(\widehat{\mathbf{m}})\mathbf{X}\right)(\mathbf{x}_{s1} - \mathbf{x}_{s2}). \tag{9}$$

For the Poisson sampling, an estimator of the asymptotic standard error is given by $\sqrt{Var(L\ og\ Odds_{kl|ij})}$ (Agresti, 1990), thus the asymptotic confidence intervals for the $Log\ Odds_{kl|ij}$, associated with a probability $1-\alpha$, are

Table 5
Residual deviances and p-values of the fitted models (SPI6 and SPI12 time series) using the Chi-square test with 5% significance level and 34 degrees of freedom.

Location	Season	SPI16		SPI12	
		Residual deviance	p-value	Residual deviance	p-value
L0034	Wet	17.17	0.9927	12.26	0.9998
	Dry	23.26	0.9173	18.47	0.9861
L0035	Wet	30.84	0.6229	23.52	0.9112
	Dry	48.14	0.0546	26.27	0.8253
L0036	Wet	44.75	0.1025	17.69	0.9905
	dry	28.38	0.7388	16.21	09958
L0037	Wet	48.10	0.0551	20.54	0.9666
	Dry	45.50	0.0898	16.87	0.9938
L0038	Wet	35.10	0.4157	15.64	0.9970
	Dry	44.70	0.1036	23.44	0.9133
L0044	Wet	22.72	0.9297	32.36	0.5481
	Dry	25.60	0.8492	34.50	0.4438
L0045	Wet	38.07	0.2891	22.64	0.9315
	Dry	24.39	0.8877	14.18	0.9989
L0046	Wet	44.23	0.1124	32.14	0.5591
	Dry	35.45	0.3994	24.06	0.8968
L0047	Wet	35.16	0.4128	31.35	0.5978
	Dry	47.80	0.0584	23.47	0.9125
L0048	Wet	46.12	0.0802	32.86	0.5233
	Dry	41.65	0.1721	21.66	0.9501

Please cite this article in press as: Moreira, E.E., SPI drought class prediction using log-linear models applied to wet and dry seasons, Physics and Chemistry of the Earth (2015), http://dx.doi.org/10.1016/j.pce.2015.10.019

$$\left[Log\ Odds_{kl|ij} - q_{1-\alpha/2}\sqrt{V\left(LogOdds_{kl|ij}\right)}, Log\ Odds_{kl|ij} + q_{1-\alpha/2}\sqrt{V\left(LogOdds_{kl|ij}\right)}\right] \quad (10)$$

where $q_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of a standard normal variable. The estimates of the odds and corresponding asymptotic confidence intervals are then obtained by exponentiation of the respective intervals borders for the logarithm of the odds.

The confidence intervals of the odds reflect the sampling variability of the observed drought transitions internal to each time series. Odds confidence intervals, besides reflecting this variability, indicate also if a given odds is significantly different from 1. Considering a level of significance of 5%, if the confidence interval for an odds includes the value 1, then there is 95% probability that the odds in fact equals 1, meaning that the drought transition from class *i* to class *j* to class *k* and the drought transition from class *i* to class *j* to class *l*, are not significantly different. Otherwise, if the confidence interval not includes 1, then there is also 95% probability that the odds is in fact larger or less than 1, depending on the value that it takes, meaning that the first transition is significantly more or less probable than the second. However, if the confidence interval of a given odds is too large, the reliability of the prediction is small.

In order to obtain the most probable class transition for the month $t+1$, the odds for the three closest one-step transitions from the drought class at month t are computed as well as their confidence intervals. Depending on the values obtained, one of the three is chosen as the most probable. For instance, if the drought classes at month $t-1$ and t are equal to 2 and 3 respectively, then$Odds_{32|23}$, $Odds_{34|23}$ and $Odds_{24|23}$ will be computed. If the values and confidence intervals obtain for those odds are for instance $Odds_{32|23} = 1.72[0.82, 3.63]$, $Odds_{34|23} = 4.87[2.42, 9.82]$ and $Odds_{24|23} = 2.83[1.23, 6.52]$, then class 2 or 3 will be equally probable, since the confidence interval for $Odds_{32|23}$ includes the value 1 and the transition 2 → 3→3 is more probable than 2 → 3 → 4 as well as, transition 2 → 3 → 2 is more probable than 2 → 3 → 4.

As said before, the modelling described in section 3.1 was applied to the months of the wet and dry season separately; as a result two different expected frequencies are obtained: m^1_{ijk} from modelling the wet season and m^2_{ijk} from modelling the dry season. When obtaining the most probable transition for the month *t + 1*, the season is evaluated at month *t* in order to use $Odds^1_{kl|ij} = \frac{m^1_{ijk}}{m^1_{ijl}}$ if the season is wet, or $Odds^2_{kl|ij} = \frac{m^2_{ijk}}{m^2_{ijl}}$ if the season is dry.

3.3. Model performance

In order to evaluate the model performance skill scores are used (Wilks, 2006; Jolliffe and Stephenson, 2003). A skill score measures the forecast accuracy relative to some set of control or a reference forecast. It essentially answers the question: is my forecast better or worse than the reference or control forecast?

The Heidke Skill Score (HSS) is one of the most use skill score, because it is relatively easy to compute and also because the standard forecast, chance, is relatively easy to beat. Other standard scores are possible, such as persistence or climatology, but these require additional information to compute, in the form of a separate contingency table. The HSS measures the fractional improvement of the forecast over a random prediction. Like most skill scores, it is normalized by the total range of possible improvement over the standard, which means HSS can safely be compared on different datasets. The range of the HSS is $-\infty$–1. Negative values indicate that the chance forecast is better, 0 means no skill, and a perfect forecast obtains a HSS of 1. In our case the computation of the HSS involves building the contingency table presented in Table 2, which is used in HSS defined as follow:

$$HSS = \left(\sum_{i=1}^{4} p_{ii} - \sum_{i=1}^{4} p_i p'_i\right) \Big/ \left(1 - \sum_{i=1}^{4} p_i p'_i\right) \quad (11)$$

where p_{ii} is the proportion of predictions that agree with the observations for class *i* and p_{ik} is the proportion of predictions that not agree with the observations for entry class (*i*,*k*), with $i \neq k$ and p_i, p'_i are the marginal totals in Table 2.

From Table 2, we also easily obtain the measure that gives us the total number of agreements, called the proportion of corrects which simply is

$$PC = \sum_{i=1}^{4} p_{ii}. \quad (12)$$

4. Results and discussion

As for the previous work (Moreira et al., 2008), both the contingency tables for wet and dry seasons (Tables 3 and 4) show that the highest values occur for the transitions that imply the maintenance of the precedent drought classes, indicating the mentioned self-perpetuating characteristic trend of droughts, as well as their slow initiation and dissipation. This characteristic is less strong as expected for the SPI6 than for the SPI12. The transitions to a different drought class occur more frequently for the SPI6, since it responds quickly to increases or decreases in the precipitation.

To illustrate, it is presented in Table 6 for 7 for the 4 of 10 locations (L0035, L0036, L0045 and L0048), the comparison between the SPI6 drought class categories calculated from observed rainfall data and the predictions produced by the log-linear modelling, computed through the odds estimates. The selected period concerns October 2011 to June 2013, which gives a picture of a drought initiation, development and dissipation. For each site are presented the observed SPI6 drought classes at months $t-1$ and t and the observed and predicted SPI6 drought classes at month $t+1$, using first the log-linear model applied to the wet and dry seasons (5th column in each location of Table 6) and the modelling without the split into wet and dry season (6th column in each location of Table 6). The modelling with the split into wet and dry seasons is named model* producing predictions*. When two or three drought classes are equally probable then in the predicted drought class it appears for instance "1 or 2" or "2 or 3 or 4", meaning that probabilities for transitions into the classes 1 or 2, or 2 or 3 or 4 are not significantly different respectively. The cells in Table 6 highlighted in grey are those when the prediction does not match the observation. The same content are presented in Table 7 for the SPI12.

Results in Tables 6 and 7, as in the previous study by Moreira et al. (2008), show that disagreements between the predicted and the observed often refer to conditions when a decrease or increase of the drought class category breaks with the drought class established in the preceding two months, which correspond to an increase or decrease in rainfall for the months under prevision. These situations are very difficult to predict with a modelling of this type, purely probabilistic, that learns mainly from the past transitions and do not account for predictive variables representing the present weather, like for instance global atmospheric circulation patterns. However, with the log-linear model applied to the wet and dry seasons, some of these situations could be predict, namely

Please cite this article in press as: Moreira, E.E., SPI drought class prediction using log-linear models applied to wet and dry seasons, Physics and Chemistry of the Earth (2015), http://dx.doi.org/10.1016/j.pce.2015.10.019

Table 6
Comparison between observed and predicted SPI6 drought classes of four locations, for the period October 2011–June 2013. The column of "predicted*" uses the modelling with the split into wet and dry season, while just "predicted" uses the modelling without the split.

L0035 date	Drought class at		Drought class at month t + 1			L0036 date	Drought class at		Drought class at month t + 1		
	t − 1	t	Real	Predicted*	Predicted		t − 1	t	Real	Predicted*	Predicted
Oct-11	3	3	2	2 or 3 or 4	2 or 3 or 4	Oct-11	2	2	1	2	2
Nov-11	3	2	2	2	2	Nov-11	2	1	2	1	1
Dec-11	2	2	3	2	2	Dec-11	1	2	2	2	2
Jan-12	2	3	4	2 or 3 or 4	2 or 3 or 4	Jan-12	2	2	3	2	2
Feb-12	3	4	4	4	4	Feb-12	2	3	3	3	3
Mar-12	4	4	4	4	4	Mar-12	3	3	3	2 or 3 or 4	2
Apr-12	4	4	4	4	4	Apr-12	3	3	4	2	2
May-12	4	4	4	4	4	May-12	3	4	4	2 or 3 or 4	2 or 3 or 4
Jun-12	4	4	3	4	4	Jun-12	4	4	3	4	4
Jul-12	4	3	2	2 or 3 or 4	2	Jul-12	4	3	2	2 or 3 or 4	2 or 3 or 4
Aug-12	3	2	1	1 or 2	2	Aug-12	3	2	1	2	2
Sep-12	2	1	1	1	1	Sep-12	2	1	1	1	1
Oct-12	1	1	2	1	1	Oct-12	1	1	1	1	1
Nov-12	1	2	2	2	2	Nov-12	1	1	1	1	1
Dec-12	2	2	1	2	2	Dec-12	1	1	1	1	1
Jan-13	2	1	2	1	1	Jan-13	1	1	1	1	1
Feb-13	1	2	1	2	2	Feb-13	1	1	1	1	1
Mar-13	2	1	1	1	1	Mar-13	1	1	1	1	1
Apr-13	1	1	1	1	1	Apr-13	1	1	1	1	1
May-13	1	1	1	1	1	May-13	1	1	1	1	1
Jun-13	1	1	1	1	1	Jun-13	1	1	1	1	1

L0045 date	Drought class at		Drought class at month t + 1			L0048 date	Drought class at		Drought class at month t + 1		
	t − 1	t	Real	Predicted*	Predicted		t − 1	t	Real	Predicted*	Predicted
Oct-11	3	4	2	2 or 3 or 4	4	Oct-11	1	1	1	1	1
Nov-11	4	2	3	2	2	Nov-11	1	1	2	1	1
Dec-11	2	3	4	2 or 3 or 4	2 or 3 or 4	Dec-11	1	2	2	2	2
Jan-12	3	4	4	2 or 3 or 4	4	Jan-12	2	2	3	2	2
Feb-12	4	4	4	4	4	Feb-12	2	3	3	3	2 or 3
Mar-12	4	4	4	4	4	Mar-12	3	3	3	3	2 or 3
Apr-12	4	4	4	3 or 4	4	Apr-12	3	3	4	2 or 3	2 or 3
May-12	4	4	4	3 or 4	4	May-12	3	4	4	2 or 3 or 4	4
Jun-12	4	4	3	3 or 4	4	Jun-12	4	4	4	4	4
Jul-12	4	3	2	2 or 3 or 4	2	Jul-12	4	4	2	4	4
Aug-12	3	2	1	2	2	Aug-12	4	2	2	2	2
Sep-12	2	1	2	1	1	Sep-12	2	2	2	2	2
Oct-12	1	2	1	1 or 2	2	Oct-12	2	2	1	2	2
Nov-12	2	1	1	1	1	Nov-12	2	1	1	1	1
Dec-12	1	1	1	1	1	Dec-12	1	1	1	1	1
Jan-13	1	1	1	1	1	Jan-13	1	1	2	1	1
Feb-13	1	1	1	1	1	Feb-13	1	2	1	2	2
Mar-13	1	1	1	1	1	Mar-13	2	1	1	1	1
Apr-13	1	1	1	1	1	Apr-13	1	1	1	1	1
May-13	1	1	1	1	1	May-13	1	1	1	1	1
Jun-13	1	1	1	1	1	Jun-13	1	1	1	1	1

those when after some months in drought, with the beginning of the wet season in October a decrease of the drought class is expected. Also when after some months in non-drought, during the dry season it is expectable an increase of the drought class.

From the comparison between the observations and the predictions produced by the two different modelling for the 10 locations (just 4 shown in Tables 6 and 7), the modelling applied to wet and dry season seems to bring a slightly increase in the number of agreements, but mainly apparently due to the increase of the cases when two or three drought class transitions are equally probable. For the SPI6 and the selected period (Table 6), in a few cases the log-linear modelling applied to wet and dry season can predict the correct class when there is a break with the installed drought class in the previous two months. Regarding the SPI12, the number of disagreements for the selected period and the locations used to illustrate, are the same in both modelling. However, from the observation of the entire period of the time series, we can find several cases where the modelling with the split into wet and dry season can predict correctly a change of the drought class in opposition with the modelling without the split.

When looking to the SPI12 results (Table 7), it is noticeable that the number of disagreements decrease compared with the results for the SPI6 (Table 6), using both the modelling, with the split into wet and dry season and without the split. The increase in the agreements when using SPI12 is natural because the number of transitions between different drought classes decreases since as said before, the SPI12 is built to respond slowly to changes in the precipitation than the SPI6. Given that the log-linear modelling, as other stochastic approaches like Markov Chains (Paulo and Pereira, 2007), fails to predict many of these changes of classes, a decline in the model performance is justified when applied to the SPI6.

In order to have a true picture of the model performance applied to wet and dry seasons compared with the modelling without the split, the proportion of corrects (PC) and the Heidke skill score (HSS) were computed for the entire period of the time series, for the SPI6 (Table 8) and the SPI12 (Table 9). Results presented in

Please cite this article in press as: Moreira, E.E., SPI drought class prediction using log-linear models applied to wet and dry seasons, Physics and Chemistry of the Earth (2015), http://dx.doi.org/10.1016/j.pce.2015.10.019

Table 7
Comparison between observed and predicted SPI12 drought classes of four locations, for the period October 2011–June 2013. The column of "predicted*" uses the modelling with the split into wet and dry season, while just "predicted" uses the modelling without the split.

L0035 date	Drought class at		Drought class at month t + 1		
	t − 1	t	Real	Predicted*	Predicted
Oct-11	2	2	2	2	2
Nov-11	2	2	3	2	2
Dec-11	2	3	4	3	3
Jan-12	3	4	4	2 or 3 or 4	3 or 4
Feb-12	4	4	4	4	4
Mar-12	4	4	4	4	4
Apr-12	4	4	4	4	4
May-12	4	4	4	4	4
Jun-12	4	4	4	4	4
Jul-12	4	4	4	4	4
Aug-12	4	4	4	4	4
Sep-12	4	4	3	4	4
Oct-12	4	3	4	2 or 3 or 4	2 or 3 or 4
Nov-12	3	4	4	2 or 3 or 4	3 or 4
Dec-12	4	4	2	4	4
Jan-13	4	2	2	2	2
Feb-13	2	2	1	2	2
Mar-13	2	1	1	1	1
Apr-13	1	1	1	1	1
May-13	1	1	1	1	1
Jun-13	1	1	1	1	1

L0036 date	Drought class at		Drought class at month t + 1		
	t − 1	t	Real	Predicted*	Predicted
Oct-11	1	1	1	1	1
Nov-11	1	1	2	1	1
Dec-11	1	2	2	2	2
Jan-12	2	2	3	2	2
Feb-12	2	3	3	2 or 3	3
Mar-12	3	3	4	2	2 or 3
Apr-12	3	4	4	4	3 or 4
May-12	4	4	3	4	4
Jun-12	4	3	3	2 or 3 or 4	3
Jul-12	3	3	3	3	2 or 3
Aug-12	3	3	3	3	2 or 3
Sep-12	3	3	3	3	2 or 3
Oct-12	3	3	3	2	2 or 3
Nov-12	3	3	3	2	2 or 3
Dec-12	3	3	2	2	2 or 3
Jan-13	3	2	2	2	2
Feb-13	2	2	1	2	2
Mar-13	2	1	1	1	1
Apr-13	1	1	1	1	1
May-13	1	1	1	1	1
Jun-13	1	1	1	1	1

L0045 date	Drought class at		Drought class at month t + 1		
	t − 1	t	Real	Predicted*	Predicted
Oct-11	2	2	2	2	2
Nov-11	2	2	3	2	2
Dec-11	2	3	4	2 or 3 or 4	3
Jan-12	3	4	4	3 or 4	3 or 4
Feb-12	4	4	4	4	4
Mar-12	4	4	4	4	4
Apr-12	4	4	4	4	4
May-12	4	4	4	4	4
Jun-12	4	4	4	4	4
Jul-12	4	4	4	4	4
Aug-12	4	4	4	4	4
Sep-12	4	4	4	4	4
Oct-12	4	4	4	4	4
Nov-12	4	4	3	4	4
Dec-12	4	3	2	2 or 3 or 4	2 or 3 or 4
Jan-13	3	2	2	2	2
Feb-13	2	2	1	2	2
Mar-13	2	1	1	1	1
Apr-13	1	1	1	1	1
May-13	1	1	1	1	1
Jun-13	1	1	1	1	1

L0048 date	Drought class at		Drought class at month t + 1		
	t − 1	t	Real	Predicted*	Predicted
Oct-11	1	1	1	1	1
Nov-11	1	1	1	1	1
Dec-11	1	1	1	1	1
Jan-12	1	1	1	1	1
Feb-12	1	2	2	1	1
Mar-12	2	2	2	2	2
Apr-12	2	2	3	2	2
May-12	2	3	3	2 or 3	2 or 3
Jun-12	3	3	3	3	3
Jul-12	3	3	3	3	3
Aug-12	3	3	3	3	3
Sep-12	3	3	3	3	3
Oct-12	3	3	3	2 or 3 or 4	3
Nov-12	3	2	2	2 or 3 or 4	3
Dec-12	2	2	2	2	2
Jan-13	2	2	2	2	2
Feb-13	2	1	1	2	2
Mar-13	1	1	1	1	1
Apr-13	1	1	1	1	1
May-13	1	1	1	1	1
Jun-13	1	1	1	1	1

Table 8
Results obtained from the SPI6, for the proportion of corrects (PC) and the Heidke skill score (HSS), for modelling with the split into wet and dry season (model*) and modelling without the split (model) and the difference between both.

SPI6	PC			HSS		
	Model*	Model	Difference	Model*	Model	Difference
L0034	76.26%	72.17%	4.09%	60.83%	53.99%	6.84%
L0035	77.68%	74.18%	3.50%	63.06%	57.41%	5.64%
L0036	73.14%	70.83%	2.31%	56.01%	52.08%	3.93%
L0037	73.96%	72.62%	1.34%	56.86%	54.74%	2.12%
L0038	76.48%	74.10%	2.38%	61.23%	57.51%	3.72%
L0044	74.40%	70.91%	3.50%	58.08%	52.49%	5.60%
L0045	77.46%	73.74%	3.72%	63.02%	56.97%	6.05%
L0046	75.97%	73.29%	2.68%	60.39%	55.80%	4.59%
L0047	72.99%	71.95%	1.04%	55.47%	53.69%	1.78%
L0048	74.70%	73.81%	0.89%	58.49%	57.22%	1.27%

Table 9
Results obtained from the SPI12, for the proportion of corrects (PC) and the Heidke skill score (HSS), for modelling with the split into of wet and dry season (model*) and modelling without the split (model) and the difference between both.

SPI12	PC			HSS		
	Model*	Model	Difference	Model*	Model	Difference
L0034	86.24%	83.11%	3.12%	77.57%	72.62%	4.95%
L0035	82.59%	82.96%	−0.37%	71.48%	72.06%	−0.58%
L0036	83.33%	84.15%	−0.82%	73.06%	74.56%	−1.50%
L0037	83.48%	82.14%	1.34%	73.67%	71.56%	2.11%
L0038	85.71%	83.63%	2.08%	76.84%	73.52%	3.32%
L0044	83.71%	82.22%	1.49%	73.22%	70.89%	2.33%
L0045	84.45%	83.78%	0.67%	74.57%	73.52%	1.05%
L0046	82.44%	83.26%	−0.82%	71.53%	72.77%	−1.24%
L0047	84.60%	83.18%	1.42%	75.33%	73.10%	2.23%
L0048	85.49%	83.18%	2.31%	76.38%	72.56%	3.83%

Please cite this article in press as: Moreira, E.E., SPI drought class prediction using log-linear models applied to wet and dry seasons, Physics and Chemistry of the Earth (2015), http://dx.doi.org/10.1016/j.pce.2015.10.019

Table 8 for the SPI6 confirm the slightly improvement in the predictions brought by the modelling with the split into wet and dry seasons when compared with the modelling without the split. The PC presents an improvement ranging from 0.9% to 4%, while the improvement in the HSS varies between 1.3% and 6.8%, depending on the location. The locations in the north and center of Portugal have the highest improvements and in general are also the ones that present the highest skill scores. For the SPI12 (Table 9), three of the locations do not present improvement in the predictions when the modelling is applied separately to wet and dry season, and for the other locations the improvement in the skill scores is a bit lower than compared with the SPI6, ranging from 0.7% to 3.1% for the PC and 2.1%–5% for the HSS. Unlike the SPI6, no trend regarding the skill scores is detected for the locations distribution in Portugal.

From these results we can infer that when applied to the SPI12, the modelling applied to the wet and dry seasons, globally does not produce a real improvement in the predictions. Yet, when applied to the SPI6, a real improvement in the predictions occur in all the locations, being however quite modest in some of those. Nevertheless, the modelling has an overall good performance, since it attains values of 70–77% in the PC and 50–63% for the HSS when applied to SPI6, and even better as expected for the SPI12, attaining values of 82–86% in the PC and 70–77% for the HSS.

In conclusion, we can say that the log-linear modelling performs well in predicting the drought class with 1 month ahead knowing the drought classes in the two previous months, although it fails many times in predicting the transitions to a drought class different than the drought class that was in the previous two months. Nevertheless it captures very well the maintenance of the drought class. Splitting the modelling into wet and dry seasons, some situations of class change can be predicted, namely those when after some months in drought, with the beginning of the wet season there is a decrease of the drought class or when after some months in non-drought, during the dry season there is an increase in the drought class. The improvement in the predictions is globally smaller that our expectations and just effective when using the SPI6, which is more suitable to detect agriculture droughts.

With the monthly prediction produced by the log-linear modelling of SPI drought class transitions, the stakeholders affected by drought like farmers and the energy producing sector using hydrological resources, can better adopt a risk management approach to droughts, through the developing of policies and measures to mitigate drought.

5. Conclusions

The log-linear modelling of SPI drought class transitions shows to be a good tool to produce drought class predictions in a short term. Knowing the drought class values for two precedent months it is possible to make reliable predictions for drought class in the following month, since the modelling obtained performances in the range of 70–86% for the percentage of corrects and 50–77% for the Heidke skill score. The cases where the predicted and the actual drought classes did not match occur often when there is a break with the installed drought class in the previous two months, which correspond to a significant increase or decrease in rainfall for the months under prediction.

The adding up of the wet and dry seasons introduced in the modelling brought modest improvements in the predictions, of about 0.9–4% in the percentage of corrects and 1.3–6.8% in the Heidke skill score. However a real improvement exists, mainly when using the SPI6 the modelling applied to the wet and dry seasons. In several cases when there is a break with class installed in the previous months, the modelling with the split into wet and dry seasons could predict more correct transitions than the modelling without the split.

The approach is therefore appropriate to strengthen the usefulness of drought monitoring and related information to water managers and users, supporting their decisions on drought mitigation measures. Future research will consider the introduction of other stochastic and weather regime tools, like the atmospheric circulation patterns NAO and AO, into the log-linear modelling aiming at a larger improvement in the predictions.

Acknowledgements

This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações)" and the project PTDC/GEO-MET/3476/2012 – Predictability assessment and hybridization of seasonal drought forecasts in Western Europe.

References

Agresti, A., 1990. Categorical Data Analysis. J. Wiley & Sons, New York.
Bacanli, U., Firat, M., Dikbas, F., 2009. Adaptive neuro-fuzzy inference system for drought forecasting. Stochast. Environ. Res. Risk Assess. 23, 1143–1154.
Bierkens, M., Van Beek, L., 2009. Seasonal predictability of European discharge: NAO and hydrological response time. J. Hydrometeorol. 10 (4), 953–968.
Bonaccorso, B., Cancelliere, A., Rossi, G., 2015. Probabilistic forecasting of drought class transitions in Sicily (Italy) using standardized precipitation Index and North Atlantic oscillation. J. Hydrol. 526, 136–150.
Cancelliere, A., Di Mauro, G., Bonaccorso, B., Rossi, G., 2007. Drought forecasting using the standardized precipitation index. Water Resour. Manage 21 (5), 801–819.
Farokhnia, A., Morid, S., Byun, H.R., 2011. Application of global SST and SLP data for drought forecasting on Tehran plain using data mining and ANFIS techniques. Theor. Appl. Climatol. 104, 71–81.
Han, P., Wang, P., Zhang, S., Zhu, D., 2010. Drought forecasting based on the remote sensing data using ARIMA models. Math. Comput. Model. 51 (11–12), 1398–1403.
Jolliffe, I.T., Stephenson, D.B., 2003. Forecast Verification: a Practitioner's Guide in Atmospheric Science. Wiley, Hoboken, NJ.
Kavalieratou, S., Karpouzos, D.K., Babajimopoulos, C., 2012. Drought analysis and short-term forecast in the Aison River Basin (Greece). Nat. Hazard. Earth Syst. Sci. 12 (5), 1561–1572.
Kim, T., Valdes, J.B., 2003. Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. J. Hydrol. Eng. 8 (6), 319–328.
Lohani, V.K., Loganathan, G.V., 1997. An early warning system for drought management using the Palmer drought index. J. Am. Water Resour. Assoc. 33, 1375–1386.
Lohani, V.K., Loganathan, G.V., Mostaghimi, S., 1998. Long-term analysis and short-term forecasting of dry spells by Palmer drought severity index. Nord. Hydrol. 29 (1), 21–40.
McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. In: 8th Conference on Applied Climatology. Am. Meteor. Soc., Boston, pp. 179–184.
McKee, T.B., Doesken, N.J., Kleist, J., 1995. Drought monitoring with multiple time scales. In: 9th Conference on Applied Climatology. Am. Meteor. Soc., Boston, pp. 233–236.
Mishra, A., Desai, V., 2006. Drought forecasting using feed-forward recursive neural network. Ecol. Model 198, 127–138.
Mishra, A., Desai, V., Singh, V., 2007. Drought forecasting using a hybrid stochastic and neural network model. J. Hydrol. Eng. 12, 626–638.
Mishra, A.K., Desai, V., 2009. Drought forecasting using stochastic models. Stoch. Environ. Res. Risk Assess. 19, 326–339.
Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. J. Hydrol. 391, 202–216.
Mishra, A.K., Singh, V.P., 2011. Drought modeling – a review. J. Hydrol. 403 (1–2), 157–175.
Moreira, E.E., Paulo, A.A., Pereira, L.S., Mexia, J.T., 2006. Analysis of SPI drought class transitions using log-linear models. J. Hydrol. 331, 349–359.
Moreira, E.E., Coelho, C.A., Paulo, A.A., Pereira, L.S., Mexia, J.T., 2008. SPI-based drought category prediction using log-linear models. J. Hydrol. 354, 116–130.
Moreira, E.E., Mexia, J.T., Pereira, L.S., 2012. Are drought occurrence and severity aggravating? A study on SPI drought class transitions using log-linear models and ANOVA-like inference. Hydrol. Earth Syst. Sci. 16, 3011–3028.
Morid, S., Smakhtin, V., Bagherzadeh, K., 2007. Drought forecasting using artificial neural networks and time series of drought indices source. Int. J. Climatol. 27 (15), 2103–2211.
Nelder, J.A., 1974. Log-linear models for contingency tables: a generalization of classical least squares. Appl. Stat. 23, 323–329.

Nichols, N., Coughlan, M.J., Monnik, K., 2005. The challenge of climate prediction in mitigating drought impacts. In: Wilhite, D.A. (Ed.), Drought and Water Crisis. Science, Technology, and Management Issues. Taylor & Francis, Bocca Raton, pp. 33–51.

Ozger, M., Mishra, A., Singh, V., 2012. Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. J. Hydrometeorol. 13, 284–297.

Palmer, W.C., 1965. Meteorologic Drought. US Department of Commerce. Weather Bureau, p. 58. Research Paper No. 45.

Paulo, A.A., Pereira, L.S., 2006. Drought concepts and characterization: comparing drought indices applied at local and regional scales. Water Int. 31 (1), 37–49.

Paulo, A.A., Pereira, L.S., 2007. Prediction of SPI drought class transitions using Markov chains. Water Resour. Manag. http://dx.doi.org/10.1007/s11269-006-9129-9.

Paulo, A.A., Rosa, R.D., Pereira, L.S., 2012. Climate trends and behavior of drought indices based on precipitation and evapotranspiration in Portugal. Nat. Hazards Earth Syst. Sci. 12, 1481–1491.

Pereira, L.S., Rosa, R.D., Paulo, A.A., 2006. Testing a modification of the Palmer drought severity index for Mediterranean environments. In: Rossi, G., Vega, T., Bonaccorso, B. (Eds.), Methods and Tools for Drought Analysis and Management. Springer, Dordrecht, pp. 149–167.

Pires, C.A., Perdigão, R.A.P., 2007. Non-Gaussianity and asymmetry of the winter monthly precipitation estimation from the NAO. Mon. Weather Rev. 135 (2), 430–448.

Raziei, T., Bordi, I., Pereira, L.S., 2011. An application of GPCC and NCEP/NCAR datasets for drought variability analysis in Iran. Water Resour. Manage 25, 1075–1086.

Ribeiro, A., Pires, C.A., 2015. Seasonal drought predictability in Portugal using statistical-dynamical techniques. Jornal Phys. Chem. Earth. Special issue of the EcoHCC14, Tomar, Portugal (2015).

Santos, J.F., Pulido-Calvo, I., Portela, M., 2010. Spatial and temporal variability of droughts in Portugal. Water Resour. Res. 46, W03503.

Schneider, U., Becker, A., Meyer-Christoffer, A., Ziese, M., Rudolf, B., 2010. Global Precipitation Analysis Products of the GPCC. Global Precipitation Climatology Centre (GPCC), DWD, Internet Publication. http://www.dwd.de.

Steinemann, A.C., Hayes, M.J., Cavalcanti, L.F.N., 2005. Drought indicators and triggers. In: Wilhite, D.A. (Ed.), Drought and Water Crises. CRC Press, Boca Raton, pp. 71–92.

Vicente-Serrano, S.M., 2006. Differences in spatial patterns of drought on different time scales: an analysis of the Iberian Peninsula. Water Resour. Manage 20, 37–60.

Wilhite, D.A., Sivakumar, M.V.K., Wood, D.A., 2000. In: Early Warning Systems for Drought Preparedness and Drought Management. (Proceedings of an Expert Group Meeting. Lisbon). World Meteorological Organization, Geneva. http://drought.unl.edu/monitor/EWS/EWS_WMO.html.

Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences, second ed. Academic Press/Elsevier, New York.

Please cite this article in press as: Moreira, E.E., SPI drought class prediction using log-linear models applied to wet and dry seasons, Physics and Chemistry of the Earth (2015), http://dx.doi.org/10.1016/j.pce.2015.10.019