A STOCHASTIC APPROACH TO STANDARDIZED PRECIPITATION INDEX CLASS TRANSITIONS

Isabel Abreu
Professora Auxiliar - Faculdade de Ciência e Tecnologia (UFP)
iabreu@ufp.edu.pt

Maria João Guerreiro
Professora Auxiliar - Faculdade de Ciência e Tecnologia (UFP)
mariajoao@ufp.edu.pt

Teresa Lajinha
Professora Auxiliar - Faculdade de Ciência e Tecnologia (UFP)
tlajinha@ufp.edu.pt

ABSTRACT
Floods and droughts are natural phenomena of difficult prevision. Awareness, mitigation and water resource management, depend upon timely information on the beginning of a dry or wet period and on their extension in time and space. This study predicts SPI steady state class probabilities in the Douro region, Portugal, through an empirical and Markov chain approach, expected residence time in each class of severity, expected first passage time, recurrence time and class prediction in a short-term basis. It was verified that all these variables are useful for water resources management.

KEYWORDS
SPI, Markov chain, floods and droughts, water resources management.

RESUMO
Cheias e secas são fenómenos naturais de difícil previsão. O conhecimento, mitigação e gestão dos recursos hídricos depende de informação sobre o início dos períodos secos ou húmidos e a sua extensão no tempo e no espaço. Este estudo prevê as probabilidades das classes do índice SPI, na região do Douro, Portugal, através das séries de Markov e de um método empírico, tempo de residência esperada em cada classe, tempo médio de transição entre classes pela primeira vez, tempo de recorrência e previsão de classes a curto prazo. Verificou-se que todas as variáveis são convenientes para a gestão dos recursos hídricos.

PALAVRAS-CHAVE
SPI, cadeia de Markov, cheias e secas, gestão de recursos hídricos.
Floods and droughts are climatic events that occur at variable time frequencies in many areas of the world (Seiler et al., 2002). Several regions in Portugal are flood-prone areas, while other regions of the country are frequently threatened by dry periods. Drought, as well as floods, may be produced by natural causes or may be induced by human activities (Paulo et al., 2005). Drought is often defined as a temporary water scarcity situation due to a precipitation deficit. On the other hand, a flood can be produced by rainfall excess over a period of time. These conjugated phenomena are part of the normal behaviour of any climate, but they can have severe impacts on regional and national economics, having also social and environmental consequences.

Because of its slow development and its difficulty to be detected, drought is possibly one of the more complex natural occurrences (Morid et al., 2007). Floods are also difficult to predict, because they depend not only on precipitation volume and intensity, but also on the characteristics of the soil that affect drainage capacity. This capacity can be overloaded under abnormal abundant rainfall conditions.

Drought and flood awareness and mitigation, as well as water resource management plans depend upon timely information on the beginning of a dry or wet period and on their extension in time and space. The kind of information needed to deal with drought events may be obtained through continuous monitoring, which is normally performed using drought indices (Paulo et al., 2005). One of such indices is the Standardized Precipitation Index (SPI) (McKee et al., 1993) and some authors suggest that it can also be used as an indicator of the progress of soil saturation conditions conducting to floods (Seiler et al., 2002).

The SPI was defined as the number of standard deviations that the experimental cumulative precipitation (monthly amount) at given time scales (normally one, three or six months, or one or two years) would deviate from the long-term mean. The cumulative precipitation reports to a particular month; that is, a SPI-3 in April represents the SPI value for the cumulative precipitation of the actual and previous 2 months to April (February, March, and April), a SPI-6, the SPI value for the cumulative precipitation of the actual and previous 5 months, and so on.

Shorter or longer time scales may reproduce response delays to precipitation anomalies (Paulo et al., 2005). Since the cumulative rainfall may not be normally distributed, the data is approximately transformed to a normal standardize distribution (Ntale and Gan, 2003), so that a SPI equal to zero implies that the corresponding monthly amount represents 50% of the cumulative fitted distribution (McKee et al., 1993). SPIs ranging from −1 to +1 express a mild pluviometric regime and values out of this range represent relevant deviations from the average rainfall amount (Lana et al., 2001).

An important characteristic of the SPI is that it can monitor dry and wet periods over an extensive variety of time scales, assessing the precipitation effects on different water resource components, like groundwater, reservoir storage, soil moisture and stream flow (Morid et al., 2006). In practice, a monthly precipitation time series is ‘flattened’ using a moving window of width equal to the number of months desired, e.g. a 6-months SPI in July would be computed using the cumulative rainfall over February to July. Ntale and Gan (2003) point that Edwards and McKee (1997) selected a 3 month SPI for a short-term drought index, a 12 month SPI for an intermediate-term drought index, and a 48 month SPI for a long-term drought index. Labedzki (2007) stated that the 1-3 month SPI reflects better the agricultural drought than the 6 month SPI.

The SPI demonstrated to be a tool that should be used operationally as part of a state, regional, or national drought watch system in the United States. During the 1996 drought in the USA, the SPI detected the onset of the drought at least 1 month in advance of the Palmer Drought Severity Index (Hayes et al., 1999), in addition, it satisfactorily explains the development of conditions leading up to the three main flood events to occur in the southern Cordoba Province in Argentina during a period of 25 years (Seiler et al., 2002).

The hazard and catastrophic nature of droughts and the existence of a simple index to compute their severity and time extension are important conditions to develop prediction tools, as probabilistic ones, which may support opportune implementation of vigilance and mitigation measures (Paulo et al., 2005).

Studies developed by Fernandez and Salas (1999) have provided analytical formulations for estimating return periods (defined as the average number of trials required to the first occurrence of a critical event) of drought events with duration greater than or equal to a critical value for both time independent and Markov time dependent series (Bonaccorso et al., 2003).

Markov chains are used in climatology to model rainy and drought behaviour or evolution from wet to dry episodes (Lana and Burgueño, 1998). A Markov chain approach, using both the homogeneous and non-homogeneous formulations, was used for several locations in Alentejo, southern Portugal, to characterize the stochastic nature of droughts and to predict the evolution from a class of severity to another (Paulo et al., 2005; Paulo et al., 2007).

The Nash-Sutcliffe efficiency (NSE) is a normalized statistic which evaluates the goodness of fit of a model, computing the relative magnitude of the residual variance to the measured data variance. The optimum value is 1.0, meaning that modelled data fits a 1:1 line when compared to measured data. A value between 0.0 and 1.0 indicates that the model is within an acceptable level of performance whereas, values smaller than 0.0, indicate that the mean observed value is a better estimate than the model (Nash-Sutcliffe, 1970).

This paper aims to characterize dry and wet periods in the region of Douro, northern Portugal, and to predict SPI severity class transitions. The parameters analyzed in this study include: a) state probabilities, which represent the probabilities of occurrence of the various SPI classes using both empirical analysis and Markov chain; b) expected residence time in each class of severity, which is the average time the system stays under the same class; c) expected first passage time, which is the time it takes for the system to shift from a particular class to another; d) recurrence time, which is the time it takes for the system to come back to the same class; e) state prediction in a short-term basis, which is the most plausible class 1–3 months in advance. This is especially important for water resources management under extreme conditions (extreme dry or extreme wet conditions).

2. METHODS OF ANALYSIS

This study was conducted using monthly precipitation data series available in the meteorological network of Douro river basin (Figure 1). The raingauges are distributed within the Portuguese limits of this network.
This network has free public access through the INAG website [www.inag.pt]. The stations were chosen based not only on the length of record, but also on guaranteeing an overlapping period in all of them. All stations have monthly precipitation data from at least 1958 to 1992, which was the time range used for the analysis. Recent records are not available for all of the stations, and therefore were not included in this study (Table 1).

### Table 1 - Characteristics of raingauges

<table>
<thead>
<tr>
<th>CODE</th>
<th>NAME</th>
<th>LATITUDE (ºN)</th>
<th>LONGITUDE (ºW)</th>
<th>ALTITUDE (m)</th>
<th>BEGINNING YEAR OF OPERATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>08P/02G</td>
<td>ESCALHÃO</td>
<td>40º 56’</td>
<td>-7º 4’</td>
<td>33</td>
<td>615</td>
</tr>
<tr>
<td>03K/04UG</td>
<td>FIRVIDAS</td>
<td>41º 47’</td>
<td>-8º 16’</td>
<td>37</td>
<td>935</td>
</tr>
<tr>
<td>06Q/01UG</td>
<td>FORNOS DE LAGOAÇA</td>
<td>41º 10’</td>
<td>-7º 14’</td>
<td>34</td>
<td>697</td>
</tr>
<tr>
<td>07J/05UG</td>
<td>GRALHEIRA</td>
<td>41º 0’</td>
<td>-8º 1’</td>
<td>49</td>
<td>1104</td>
</tr>
<tr>
<td>05K/02UG</td>
<td>LAMAS DE ALVADIA</td>
<td>41º 27’</td>
<td>-8º 14’</td>
<td>34</td>
<td>964</td>
</tr>
<tr>
<td>05K/03UG</td>
<td>LAMAS OLO</td>
<td>41º 22’</td>
<td>-8º 12’</td>
<td>14</td>
<td>984</td>
</tr>
<tr>
<td>04K/03UG</td>
<td>LIJA DO ALVÃO</td>
<td>41º 30’</td>
<td>-8º 18’</td>
<td>34</td>
<td>939</td>
</tr>
<tr>
<td>06I/02UG</td>
<td>MARCO DE CANAVEZES</td>
<td>41º 10’</td>
<td>-9º 50’</td>
<td>54</td>
<td>215</td>
</tr>
<tr>
<td>08I/01UG</td>
<td>MEZIO (PAIVA)</td>
<td>40º 59’</td>
<td>-8º 6’</td>
<td>36</td>
<td>611</td>
</tr>
<tr>
<td>05L/02C</td>
<td>MINAS DE JALES</td>
<td>41º 27’</td>
<td>-8º 24’</td>
<td>36</td>
<td>853</td>
</tr>
<tr>
<td>05Q/03UG</td>
<td>MOGADOURO</td>
<td>41º 28’</td>
<td>-7º 16’</td>
<td>48</td>
<td>537</td>
</tr>
<tr>
<td>02P/01C</td>
<td>MOIMENTA DA RIA</td>
<td>41º 56’</td>
<td>-7º 1’</td>
<td>22</td>
<td>837</td>
</tr>
<tr>
<td>08I/01UG</td>
<td>MOSTEIRO DE CABREL</td>
<td>40º 56’</td>
<td>-9º 14’</td>
<td>39</td>
<td>189</td>
</tr>
</tbody>
</table>

The standardized precipitation index (SPI) for each month was calculated using the previous 3-months (SPI-3) and 6-months (SPI-6) precipitation data, for a short-term and intermediate-term analysis for management purposes, respectively. Report to Guerreiro et al. (2007) for the detailed methodology of computation of SPI values.

A SPI value equal to zero means that there are no deviations between the precipitation amount computed to that month and the mean precipitation computed to that month in the analysed time period. Positive values of the SPI indicate an excess rainfall relative to the mean and negative values of SPI indicate a lack of precipitation relative to the mean value. Therefore, dry periods are characterized by negative SPI values whilst wet periods are expressed by positive ones. The SPI values are grouped in eight classes, as suggested by Lloyd-Hughes and Saunders (2002), from extreme drought (SPI < -2.0) to extremely wet (SPI ≥ 2.0) as shown in Table 2.

### Table 2 - Drought classification by SPI value and corresponding event probabilities (Lloyd-Hughes and Saunders, 2002)

<table>
<thead>
<tr>
<th>Class</th>
<th>SPI value</th>
<th>Category</th>
<th>Probability %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.00 or more</td>
<td>Extremely wet</td>
<td>2.3</td>
</tr>
<tr>
<td>2</td>
<td>1.5 to 1.99</td>
<td>Severely wet</td>
<td>4.4</td>
</tr>
<tr>
<td>3</td>
<td>1.00 to 1.49</td>
<td>Moderately wet</td>
<td>5.2</td>
</tr>
<tr>
<td>4</td>
<td>0 to 0.99</td>
<td>Mildly wet</td>
<td>13.1</td>
</tr>
<tr>
<td>5</td>
<td>0 to -0.99</td>
<td>Mild drought</td>
<td>13.1</td>
</tr>
<tr>
<td>6</td>
<td>-1.00 to -1.49</td>
<td>Moderate drought</td>
<td>9.2</td>
</tr>
<tr>
<td>7</td>
<td>-1.50 to -1.99</td>
<td>Severe drought</td>
<td>4.4</td>
</tr>
<tr>
<td>8</td>
<td>-2.00 or less</td>
<td>Extreme drought</td>
<td>2.3</td>
</tr>
</tbody>
</table>

The long-term probabilities (steady-state) of each class were computed both using the available data for each raingauge station, and using a Markov chain approach. Given an initial state and class transition probabilities, the long term (steady-state) probabilities using Markov-chain approach were calculated.

The Markovian process is characterized by a set of states, which were the eight SPI classes (Table 2) in this case, and by the transition probability, Pij, between those eight states. This probability means that the process will be in state j at the next time point, t+1, knowing that at the present time point, t, it is in state i. Class transition probabilities were calculated based on the available data for each raingauge.

With a finite space of states, the transition probability matrix, P=P{Xt+1 = j|Xt = i}, is estimated from the sample by computing the relative frequency that a SPI value shifts from each state i to each state j. Hence, the number of i x j elements of the probability transition matrix depends on the number of states (Paulo et al., 2005, 2007).

The homogeneous formulation has some limitations namely, it assumes the transition probabilities to be independent from the starting month, unlike the non-homogeneous formulation, where the transition probabilities are dependent on the initial month, however, requiring a large bulk of data. Having this in mind, the homogeneous formulation was therefore adopted.
Provided the transition probability matrix $P$ is independent of a particular time point $t$, the Markov chain is time-homogeneous, and so the $n$-step transition probability was computed by $P^n$, where $P$ is the transition matrix.

The Nash-Sutcliffe efficiency test was performed on the empirical and Markovian class steady state probabilities in order to verify the goodness of fit.

Expected Residence Time (ERT) was calculated based on the average time the system stays in a particular class. The Expected First Passage Time (EFPT) was calculated based on the average time it takes for the system to change to a particular class, including the actual class. In this last case, it is called Recurrence Time (RT).

Given a class, the probability of transition to a particular class in 1-month, 2-months and 3-months period, was calculated. In order to estimate the most probable class 1–3 months ahead of the present state, the higher probability value of all class transitions was identified.

3. RESULTS AND DISCUSSION

Results from 13 rain gauges distributed over the Portuguese side of the Douro basin, reveal that 73% of precipitation occurs during the humid semester (October to March), and 12% occurs between June and September, meaning that April and May are the wettest months of the dry semester (Figure 2). The same pattern is observed in other regions of Portugal, e.g. Paulo et al. (2005). However, in the wet semester, the Douro basin precipitation values are about twice as high as the precipitation values in the Alentejo region, but in the dry semester, they are about four times as high, suggesting that the dry semester is drier in the Alentejo than in the Douro.

A similar drought SPI value in the Douro basin and in the Alentejo region will mean different precipitation values, even though both (and all) regions will need to face restrictions on water resources allocations and adequate implementation of measures to cope with a drought.

The results observed from the steady state class probability analysis indicate that the Markov-chain analysis is an adequate model for its calculation (Figure 3a and Figure 3b). The Nash-Sutcliffe efficiency test between the empirical and Markovian steady state class probabilities had an outcome of 0.9998, being close to the optimum value (1.0), confirming its adequacy.
is a slightly better adjustment between these values and those calculated using SPI-3, implying that a 3 month SPI may be more adequate to forecast SPI class transitions (Figure 3b).

All raingauge stations show a similar pattern for expected residence time, being mild conditions (Classes 4 and 5) more persistent than any other at both time scales, 3-months and 6-months (Figures 4a and 4b). However, the expected residence times are different at the two time scales, being higher for the 6-months than for the 3-months evaluation. A higher scatter is also observed at the 6-months than at the 3-months for the various raingauge stations.

For water management purposes, in case of a drought event, the 6-months expected residence time value might be preferred over the 3-months, since it indicates a higher persistence for the same class. This means that if, in case of a mild to severe drought, the water resources’ manager should take into account the 6-months expected residence time, for it will be on the safe side.

The calculated expected residence time is approximately 2.0 and 2.7 months for mild conditions, whereas the other classes have lower values of approximately 1.3 and 1.7 months, for SPI-3 and SPI-6 calculations, respectively. As expected, these results imply that extreme conditions do not last as long as do mild conditions, it is predicted that the residence time in an extreme condition does not overcome 2 months. The water resources’ manager may take this information into account when planning for protection measures to deal with the extreme conditions.

The expected first passage time (in months) is lower for the mild conditions (Classes 4 and 5), not only from one to itself, but also from all other classes to classes 4. This is evident in Figures 5a and 5b, where the larger values are at extreme classes 1 and 8. This is valuable information for a water resources manager, meaning that, in average, it takes a much longer time for an extreme episode (dry or wet) to occur.
Recurrence time for both SPI-3 and SPI-6 show the same behaviour (Figure 6). Extreme drought, for example, has an average recurrence time of approximately four to seven years, whereas extremely wet has an average recurrence time of approximately six to ten years, for SPI-3 and SPI-6, respectively.

Finally, the most probable state one, two and three months from the actual state, using SPI-3, are class 4 and class 5, representing the mild conditions (wet and dry). It is evident from the example presented in Table 3a that dry conditions tend to follow dry conditions, and wet conditions tend to follow wet conditions one month from the actual state. As Markov transition probability matrices show a strong diagonal trend, recent SPI conditions have a tendency to be reproduced in the short-term. There is a tendency to go to mild conditions two and three months from the actual state, which is more evident when leaving a dry condition. Wet conditions are more persistent over time.

On the other hand, SPI-6 results highlight the persistence in maintaining the same class from the actual state, especially for the wet states, as shown in Table 3b. These results are consistent with the analysis of the expected residence time, which indicates a value of approximately 3 months for the mild conditions and approximately 2 months for the other classes (Figure 4b). The same analogy applies to SPI-3.
4. CONCLUSIONS

Steady state probabilities for each drought class may be adequately modelled through Markov chain analysis, even though the results are conditioned by the probabilistic nature of the SPI calculation.

There is a good fit between the empirical data and Markovian solutions, both for SPI-3 and SPI-6, as it was confirmed by Nash-Sutcliffe efficiency test. In general, the expected residence time in each class is about 1 to 2 months for the states of more concern and 2 to 3 months for mild conditions.

In most cases, the expected time to reach mild conditions from any initial drought/wet state is about 2 months considering SPI-3 and is at least 3 months when using SPI-6. In addition, in the studied region, dry spells tend to restore faster than wet ones.

For water management purposes, in case of a drought event, SPI-6 might be preferred over the SPI-3, since the first indicates a higher persistence for the same class, being a better guarantee of less water scarcity, when adequately managed.

BIBLIOGRAPHY

INAG Home Page. [Em linha]. Disponível em www.inag.pt. [Consultado em 12/03/2008]