

Drought analysis in Slovakia: regionalization, frequency analysis and precipitation thresholds

M. M. Portela¹, M. Zelenáková², J. F. Santos³,
P. Purcz², A. T. Silva¹ & H. Hlavatá⁴

¹*Instituto Superior Técnico, IST, Universidade de Lisboa, Portugal*

²*Technical University of Košice, Slovakia*

³*Instituto Politécnico de Beja, ESTIG, Portugal*

⁴*Slovak Hydrometeorological Institute, Slovakia*

Abstract

The paper presents a drought characterization for Slovakia based on the standardized precipitation index at the time scale of three months (SPI3) applied to the precipitation records at 491 Slovakian rain gauges over a considerable span (33 years). After a drought spatial regionalization, using Principal Component Analysis (PCA), the kernel occurrence rate estimation method coupled with bootstrap confidence band was applied to characterize the yearly drought occurrence rates in each one of the regions given by the PCA, aiming at identifying trends in the frequency of the droughts. The study also includes examples of surfaces of precipitation thresholds that can be easily and reliably utilized to recognize and monitor the drought occurrences at the early stages of their development. Those surfaces were obtained by reverting to the original precipitation field the values of the SPI that represent drought limits.

Keywords: drought, standardized precipitation index, principal component analysis, kernel occurrence rate estimator, precipitation threshold surfaces.

1 Introduction, study area, precipitation data and drought index

Droughts are generally associated with the persistence of low precipitation, soil moisture and water availability relative to the normal levels in a designated area.



Although there is no universally accepted definition for drought [1] defines it as “a sustained and regionally extensive occurrence of below average natural water availability”. Different from other extreme events, like floods and earthquakes, droughts remain a less visible natural risk, whose impacts are not systematically recorded. Droughts are among the most complex and least understood natural hazards, affecting more people than any other one. They are also recurrent hazards particularly in areas with pronounced natural hydrological temporal variability.

The objective of the present study was to provide a comprehensive characterization of the drought occurrences in the entire Slovakia based on the monthly precipitations, from January 1981 to December 2013 (33 years), in the 491 rain gauges evenly distributed over the country schematically located in Fig. 1 over a map of the mean annual precipitation.

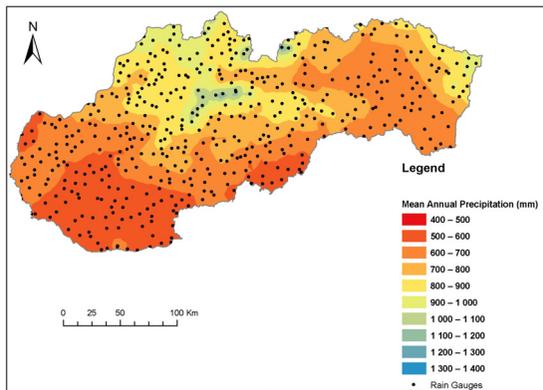


Figure 1: Location on the 491 rain gauges over a map of mean annual precipitation in Slovakia.

From the total number of months of $491 \times 33 \times 12 = 194,436$, there were gaps in 3248 (i.e. in approx. 1.67% of the months) that were filled by linear regression analysis. For that purpose and for each gap in a given rain gauge, R_1 , a near rain gauge, R_2 , was identified provided that: i) R_1 and R_2 had at least 10 years of simultaneous records in the month where the gap occurred; ii) the correlation coefficient, cc , between the two monthly historical series was the highest from all the possible nearest stations that verify condition i) and necessarily higher than 0.7 ($cc \geq 0.7$). This procedure allowed filling all the gaps, except in three months that were filled with the respective monthly average.

The mean annual precipitation map of Fig. 1 was obtained by applying the IDW (Inverse Distance Weighting) spatial interpolation technique with exponent 2 to the mean annual precipitation in the 491 rain gauges that supported the study. Authors such as [2, 3] have successfully applied the IDW technique on precipitation mapping for Slovakia. Only the historical series, i.e. the existing precipitation records prior to the gap filling, were considered to compute the mean annual precipitations.

The droughts in the study region were assessed via one of the most popular and common drought indexes [4], the Standardized Precipitation Index (SPI), developed by [5]. This index was designed to quantify the precipitation deficit at different time scales (from 1 to 24 months), which reflect the impact of droughts on the different types of reservoirs of fresh water at the watershed level. The SPI remaps the precipitation records into a standardized probability distribution function so that an index of zero indicates the median precipitation amount, while a negative index stands for drought conditions and a positive one, for wet conditions [6]. A comprehensive description of the calculation and of the advantages of the SPI index can be found in [7–11]. For the adopted time scale of 3 months, the computation of the SPI index utilized the Pearson type III probability distribution function applied to the series of cumulative precipitation in consecutive periods of 3 months (SPI3). The parameters of the distribution model were estimated by the L-moments method. Values of SPI3 lower than -1.65 represent extreme droughts, between the previous value and -1.28, severe droughts, and between this last value and -0.84, moderate droughts [12]. The time scale of three months (SPI3) was utilized not only as an example but mainly because it reflects short- and medium-term soil moisture conditions and provides a seasonal estimation of precipitation, which is quite important in primary agricultural regions such as Slovakia. In fact, approximately half of the country is under agricultural production [13]. While the short SPI time scales between 3 and 6 months may be relevant for agricultural users, hydrologists or water managers may be more interested in SPI values between 12 and 24 months. For the neighbouring country of Hungary [14] concluded that agricultural drought was best replicated by the SPI on a scale of 2–3 months.

Based on the values of SPI3 in the 491 rain gauges a drought spatial regionalization was done, by applying Principal Component Analysis (PCA). For each region thus identified the yearly drought occurrence rates were characterized through a new approach, the kernel occurrence rate estimation method (KORE) coupled with bootstrap confidence band. Such characterization allowed ascertaining the temporal frequency of the regional droughts. The paper also includes the presentation of surfaces of precipitation thresholds that can be used to identify the droughts at the early stages of their development.

2 Drought spatial patterns

The identification of spatial patterns of droughts in the SPI3 field utilized principal component analysis (PCA), as previously mentioned. PCA is a regionalization technique that can be used to identify homogenous groups of variables that experienced similar drought (or wet) conditions during a study period [4, 15, 16] and for which can be ascribed a physical meaning [17].

Authors such as [18–23] define the PCA method as a technique that allows decomposing the multisite data set of a given variable (e.g. the SPI field) into univariate representations of that variable. In that way, the original intercorrelated variables can be reduced to a small number of new linearly uncorrelated ones that explain most of the total variance [15, 16].



Considering k variables in a given time period i , $X_{i,1}$, $X_{i,2}$, ..., $X_{i,k}$, k principle components, PCs, are produced for the same time period, $Y_{i,1}$, $Y_{i,2}$, ..., $Y_{i,k}$, using linear combinations of the first ones, according to:

$$\begin{cases} Y_{i,1} = a_{11} X_{i,1} + a_{12} X_{i,2} + \dots + a_{1k} X_{i,k} \\ Y_{i,2} = a_{21} X_{i,1} + a_{22} X_{i,2} + \dots + a_{2k} X_{i,k} \\ \dots \\ Y_{i,k} = a_{k1} X_{i,1} + a_{k2} X_{i,2} + \dots + a_{kk} X_{i,k} \end{cases} \quad (1)$$

In the applications developed the variables $X_{i,k}$ refer to the SPI3 series, k is equal to the number of rain gauges considered in the analysis ($k=491$) and i represents the length of SPI3 series in each rain gauge ($i=33 \times 12 - 2 = 394$).

In the previous combinations the Y values or component scores (PC scores) are orthogonal and uncorrelated variables, such that $Y_{i,1}$ explains most of the variance, $Y_{i,2}$ the reminiscent amount of variance, and so on. The coefficients of the linear combinations are called 'loadings' and represent the weights of the original variables in the PCs.

PCs extraction can be based on variance/covariance or correlation matrix of data with $\{a_{11}, a_{12}, \dots, a_{1k}\}$ being the first eigenvector and $\{a_{k1}, a_{k2}, \dots, a_{kk}\}$ the eigenvector of k order. In the present study the Pearson correlation matrix was considered for PCs extraction.

Finally the amount of variance explained by the first PC is called the first eigenvalue, δ_1 , the second is δ_2 , so that $\delta_1 \geq \delta_2 \geq \delta_3 \geq \dots \geq \delta_k$, since each eigenvalue represents the fraction of the total variance in the original data explained by each component [25]. In the present study, the results of PCs were evaluated by analysing the eigenvalues (scree plot), the correlations between PCs and the original variables (factor loadings) and the percentage of the variance explained. To achieve more stable spatial patterns, a rotation of the principal components with the Varimax procedure was applied. This procedure provides a clearer division between components, since the rotation simplifies the spatial structure by isolating regions with similar temporal variations [20]. The patterns defined in this way are referred as rotated principal components, RPCs.

The extracted principle components, either unrotated, PCs, or rotated, RPCs, can be approximately considered representations of the same variable measured in the same units as the SPI3 from which they were derived, which was the assumption of the drought regionalization approach applied in this study.

3 Drought spatial patterns

The analysis of changes in the temporal occurrences of droughts attempts to answer the question: regardless of the severity of the drought, i.e. the precipitation deficit, how has the distribution of the occurrence of droughts

changed over time? To tackle this question, a kernel occurrence rate estimator (KORE) may be applied to a historical series of drought occurrences with the aim of estimating how the mean number of drought periods in a year, λ , changes over time, that is, to characterize $\lambda(t)$ [26–28]. The KORE analysis used the methodology detailed by [28] also applied to the RPCs: a Gaussian kernel was used, combined with pseudodata generation to reduce boundary estimation bias. The analysis focused on the occurrence of severe or worse droughts, that is, the occurrence of SPI3 values lower than -1.28. In order to analyse the uncertainty associated with the KORE estimates, a pointwise 90% bootstrap confidence band was constructed.

4 Results

4.1 Drought index and drought spatial patterns

As mentioned, the PCA was applied to the SPI series for the time scale of 3 months, SPI3, computed for the 491 rain gauges. In order to obtain more localized patterns a number of principal components were retained for Varimax rotation. Their selection was based on the interpretation of the scree plot [29], on the mapping of the factor loadings (raw data) and on the amount of variance explained in the original data.

The scree plot of Fig. 2 shows that the line stops descending and levels out approximately on the fourth PC, which means that three to four principal components should be retained.

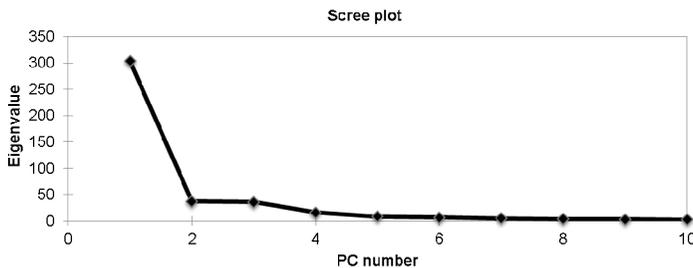


Figure 2: First ten eigenvalues resulting from the PCA applied to the SPI computed at 3 months' time-scale, SPI3.

Based on the variance explained by each component, Table 1 shows the first ten PCs retained (F1 to F10). It is clear that the first four components explain about 80% of the total variance in the original SPI3 series. It is also clear that from factors 4 to 5, i.e. F4 to F5, only a small amount of variance explained is added (approximately 1%). By mapping the factor loadings, four leading components were also suggested, since they fully cover the study area and do not overlap, being representatives of the Northern, Southern, Eastern and Western

regions. Based on the previous results, the number of components retained to rotate was the four main patterns, F1 to F4. The new distribution of the variance explained, which was maximized with Varimax rotation, is also represented in Table 1.

Table 1: Percentage of partial and cumulative variance explained by the unrotated and rotated components of PCA, F1 to F10, extracted from the SPI3 field.

PCA	SPI3			
	Partial variance explained (%)		Cumulative variance explained (%)	
	Unrotated	Rotated	Unrotated	Rotated
F1	61.9	22.1	61.9	22.1
F2	7.6	19.7	69.5	41.8
F3	7.3	15.0	76.8	56.8
F4	3.2	23.3	80.1	80.1
F5	1.7		81.8	
F6	1.3		83.1	
F7	1.0		84.0	
F8	0.7		84.7	
F9	0.6		85.4	
F10	0.6		85.9	

The four leading rotated components, RPCs (F1 to F4), of SPI3 that do not overlap, are spatially identified in Fig. 3, which represents the mapping of the factor loadings (correlations between the RPCs and the SPI3 data field). As previously mentioned the spatial interpolation method used was the Inverse Distance Weighting available on Arcgis version 10.1 (<http://www.esri.com/software/arcgis/arcgis-for-desktop>). Four sub-regions of drought variability characterized by high positive rotated loadings values greater than 0.6 (statistically significant), covering the entire study area, were identified. Also represented in Fig. 3 are the results of the KORE frequency estimator applied to each one of the RPCs time series (F1 to F4). As noted above only severe or worse droughts, represented by SPI3 values lower than -1.28, were considered in the KORE analysis.

Fig. 3 shows that, for the SPI3 time scale, the first rotated component (F1) highlights an area located in the Western part of Slovakia and it explains nearly 22% of the total variance (Table 1). The second rotated component (F2) explains an area in the Central Northern part of Slovakia (20% of total variance), the third (F3) in the Eastern part of the country (15% of total variance), and the fourth component (F4) explains an area in Central Southern part (23% of total variance). All the rotated components relate mostly positive with the original SPI3 series.

The four rotated components mean that the variation measured by the SPI3 among the drought/wet conditions across the entire study region at that time scale can be explained adequately by four components, rather than 491 rain gauges, which means a clear dimensionality reduction of the SPI3 field.

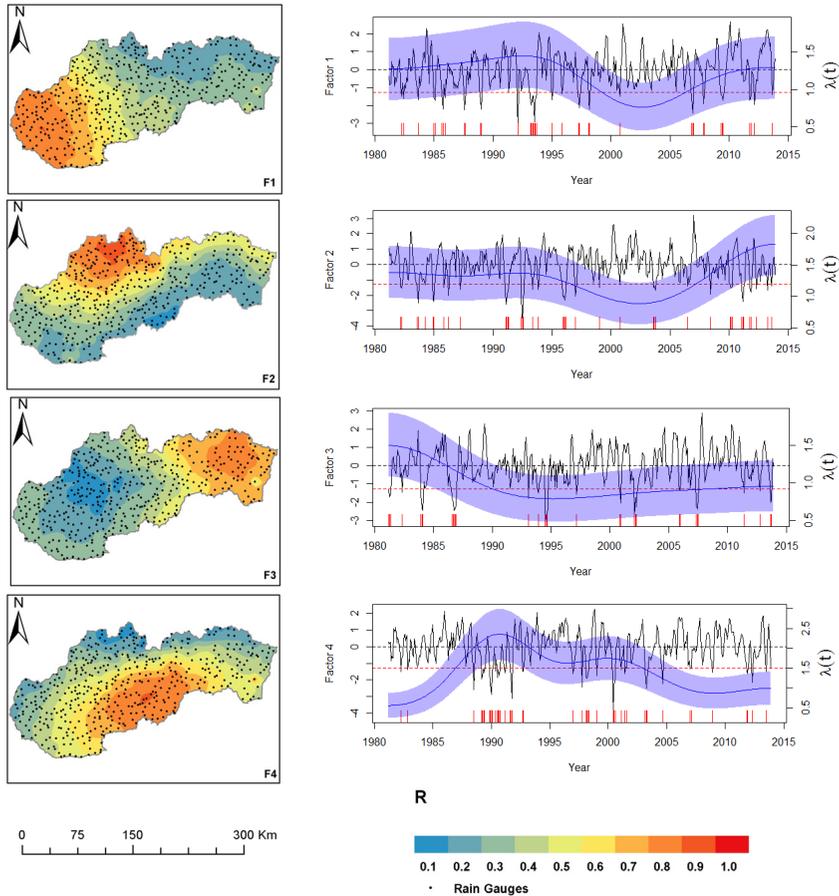


Figure 3: Spatial correlation maps between each RPC and the at-site SPI3 series (left) and time-dependent occurrence rates of severe drought of the 4 RPCs of SPI3 (right). The vertical red ticks indicate the points in time when drought events occurred.

Regarding the frequency of the drought occurrences, Fig. 3 shows that the different regions denote different frequencies of severe droughts although with some similarity between Western and Central North regions, in one hand, and between Eastern and Central South regions, in the other hand.

In fact, the western half of Slovakia experienced an increase in the drought occurrences from about 2002 on. Such increase is much more pronounced in the Central North region, where the highest rate of annual droughts occurs at the end of the analysed period, than in the Western region, where the highest rate of droughts was achieved around 1994, being slightly higher than the rates towards the present. In the eastern part of Slovakia the drought frequency seems to be decreasing towards the present or, at least, it is not as high as it was in the past. Except for Central Northern region, all the other regions experienced more

frequent droughts in the past, which somehow disagrees with the expected effect of the climate change.

5 Precipitation surfaces for drought recognition

Despite the widespread use and the advantages of SPI compared with other drought indices, the interpretation of the values associated to SPI and drought monitoring based on those values are not easy to accomplish, especially because they involve standardized values that are difficult to relate with the precipitation from which mathematical manipulation they result.

Therefore, an additional calculation was developed that gives the SPI values that represent drought thresholds back to the precipitation field, thus facilitating an adequate interpretation of the meaning of such index and quite easily and reliably identifying the drought episodes [28, 30–32]. As a result, monitoring can be operationalized as can the subsequent actions that need to be undertaken.

For that purpose and for all the 491 rain gauges of Fig. 1, the monthly and the cumulative precipitation in 3 consecutive months were estimated for different values of SPI3. The previous estimation required the inversion of the SPI calculation procedure through the use of a set of widely tested computational subroutines [33] that were incorporated in an algorithm developed by the authors.

Fig. 4 exemplifies the results given by the approach for severe drought, i.e. for $SPI3 = -1.28$. Each map shows the spatial distribution of the cumulative precipitation in periods of three consecutive months. Obviously, other drought categories and time scales can be considered.

If the precipitation registered in a given location and period falls below the value given by one of the maps for that location and period, then a severe drought is occurring. Therefore, each map represents the surface of the minimum 3-months precipitation below which a severe drought episode is recognized – surfaces cumulative precipitation thresholds for recognition of severe droughts.

The precipitation surfaces were achieved by applying the IDW (Inverse Distance Weighting) spatial interpolation technique with exponent two to the precipitation thresholds obtained by the inversion of the SPI3 in the 491 rain gauges that supported the study (Fig. 1). The schematic localization of the gauges was also included in Fig. 4.

Fig. 4 shows that the spatial patterns of the threshold precipitations are poorly differentiated across the maps. However, the smooth transition between successive maps is easily understandable because each two consecutive maps always include a common period of analysis of two months.

It is interesting to note that the greater water availability suggested by Fig. 1 for the Central North Slovakia results in slightly higher precipitation thresholds for that region. This is a consequence of the drought index utilized that recognizes a drought not by the reduced values of the precipitation themselves but by the deviations of the precipitation from average conditions. Accordingly, the precipitation thresholds for drought recognition are higher in regions with more water availability.

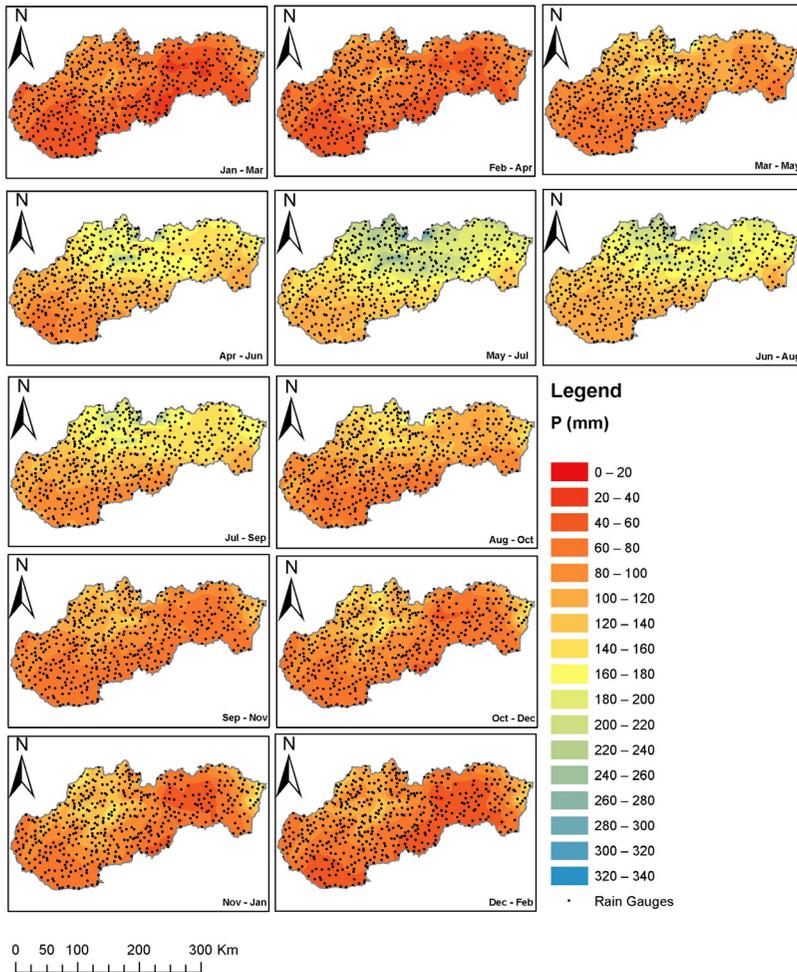


Figure 4: Inversion of SPI₃=-1.28. Precipitation in 3 consecutive months (identified in each map) corresponding to the severe drought threshold.

Figures like the one exemplified can be obtained for other time scales and drought categories. By comparing the registered precipitation in a given location and period with the precipitation thresholds for the same location, either for that period or for a wider one, obtained by including following months, it is possible to conclude not only if a drought is occurring in that location but also how much it should rain in the next month/months in order to avoid drought conditions. By assigning probabilities to these amounts of rain it is possible to estimate the probabilities of recovering from a drought in the next n months. Based on this information different decisions can be taken regarding the exploitation, for example, of water supply systems.

6 Conclusion

This paper presents a comprehensive drought characterization for the entire Slovakia based on equally comprehensive monthly precipitation data. Its results allowed identifying four regions that can be considered homogeneous in terms of droughts and which denote particular behaviour with regard to the temporal frequency of the droughts. A new application of the SPI is proposed aiming at recognizing the drought occurrences in the early stages of their development and to understand the probability of recovering from droughts thus contributing to the sustainable management of the water resources, especially when based on artificial reservoirs.

Acknowledgements

The research was financed by the Fundação para a Ciência e Tecnologia, FCT, by through the project Proc. 441.00 Eslováquia (Transnational Cooperation. Scientific Cooperation Agreement between Portugal and Slovakia) – SRDA SK-PT-0001-012. The authors thank the Slovak Hydrometeorological Institute for providing the precipitation data.

References

- [1] Tallaksen, L.M., Van Lanen, H.A., Hydrological drought: processes and estimation methods for streamflow and groundwater (Vol. 48), Elsevier Science Limited, 2004.
- [2] Szolgay, J., Parajka, J., Kohnová S., Hlavčová K., Comparison of mapping approaches of design annual maximum daily precipitation, Atmospheric Research Journal 92, 289–307, 2009.
- [3] Kohnová, S., Gaál, L., Parajka, J., Szolgay, J. and Hlavčová, K., Mapping Precipitation Quantiles by means of pooling approaches at ungauged sites in Slovakia, article published on behalf of the Slovak Grant Agency VEGA under Project No. 1/0103/10 and the Slovak Research and Development Agency under Contract No. APVV-0443-07, 2010.
- [4] Santos, J.F., Pulido-Calvo, I., Portela, M.M., Spatial and temporal variability of droughts in Portugal, Water Resources Research, 46, W03503, doi: 10.1029/2009WR008071, 2010.
- [5] McKee, T.B., Doesken, N.J., Kleist, J., The relationship of drought frequency and duration to time scales, Proceedings of the 8th Conference on Applied Climatology, American Meteorology Society, 179–184, 1993.
- [6] Santos, J.F., Portela, M.M., Pulido-Calvo I., Regional frequency analysis of droughts in Portugal, Water Resources Management, 25(14), 3537–3558, doi: 10.1007/s11269-0119869-z, 2011.
- [7] Edwards, D.C., McKee, T.B., Characteristics of 20th century drought in the United States at multiple time scales, Climatology Report 97–2, Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, 1997.



- [8] Guttman, N.B., Comparing the Palmer Drought Index and the standardized precipitation index, *Journal of the American Water Resources Association (JAWRA)* 34, 113–121, 1998.
- [9] Guttman, N.B., Accepting the Standardized Precipitation Index: A calculation algorithm, *Journal of the American Water Resources Association (JAWRA)* 35 (2), 311–322, 1999.
- [10] Hayes, M., Wilhite, D.A., Svoboda, M., Vanyarkho, O., Monitoring the 1996 drought using the standardized precipitation index, *Bulletin of the American Meteorological Society*, 80, 429–438, 1999.
- [11] Lloyd-Hughes, B., Saunders, M.A., European drought climatology and prediction using the Standardized Precipitation Index (SPI), 8.11, 13th Conference on Applied Meteorology, 2002.
- [12] Agnew, C.T., Using the SPI to identify drought, *Drought Network News*, 12, 6–12, 2000.
- [13] FAO Database 2006 (website <http://apps.fao.org/>).
- [14] Szalai S, Szinell C., Comparison of two drought indices for drought monitoring in Hungary – a case study. In *Drought and Drought Mitigation in Europe*, Vogt JV, Somma F (eds). Kluwer: Dordrecht; 161–166, 2000.
- [15] Bonaccorso B., Bordi I., Cancelliere A., Rossi G., Sutera A., Spatial variability of drought: An analysis of the SPI in Sicily, *Water Resources Management*, 17, 273–296, doi: 10.1023/A:1024716530289, 2003.
- [16] Vicente-Serrano, S.M., González-Hidalgo, J.S., Luis, M., Raventós, J., Drought patterns in the Mediterranean area: The Valencia region (eastern Spain), *Climate Research*, 26, 5–15, doi 10.3354/cr026005, 2004.
- [17] Ehrendorfer, M., A regionalization of Austria's precipitation climate using principal component analysis, *International Journal of Climatology*, 7: 71–89, doi: 10.1002/joc.3370070107, 1987.
- [18] Tipping, M.E., Bishop, C.M., Probabilistic principal component analysis, *Journal of the Royal Statistical Society, Series B*, 61(3), 611–622, doi 10.1111/1467-9868.00196, 1999.
- [19] Jolliffe, I. T. (Ed.), *Principal Component Analysis*, 2nd ed., 502 pp., Springer, New York, 2002.
- [20] Kahya, E., Demirel, M.C., Beg, O.A., Hydrologic homogeneous regions using monthly streamflow in Turkey, *Earth Sciences Research Journal*, 12(2), 181–193, 2008.
- [21] Kahya, E., Kalayc, S., Piechota, T.C., Streamflow regionalization: Case study of Turkey, *Journal of Hydrologic Engineering*, 13(4), 205–214, doi: 10.1061/(ASCE)1084-0699(2008)13:4(205), 2008.
- [22] Westra, S., Brown, C., Lall, U., Sharma, A., Modeling multivariable hydrological series: Principal component analysis or independent component analysis? *Water Resources Research*, Vol. 43, W06429, doi: 10.1029/2006WR005617, 2007.
- [23] Singh, P.K., Kumar, V., Purohit, R.C., Kothari, M., Dashora, P.K., Application of principal component analysis in grouping geomorphic parameters for hydrologic modeling, *Water Resources Management*, 23, 325–339, doi: 10.1007/s11269-008-9277-1, 2009.



- [24] Rencher, A.C. (Ed.), *Multivariate statistical inference and applications*, John Wiley, New York, 1998.
- [25] Bordi, I., Sutera, A., Fifty years of precipitation: Some spatially remote teleconnections, *Water Resources Management*, 15: 247–280, doi 10.1023/A:1013353822381, 2001.
- [26] Mudelsee, M., Borngen, M., Tetzlaff, G., Grunewald, U., No upward trends in the occurrence of extreme floods in central Europe. *Nature* 425: 166–169, 2003.
- [27] Silva, A.T., Portela, M.M., Naghettini, M., Nonstationarities in the occurrence rates of flood events in Portuguese watersheds. *Hydrology and Earth System Sciences*, 16(1): 241–254, 2012.
- [28] Portela, M. M., Santos, J. F., Silva, A. T., Benitez, J. B., Frank, C., Reichert, J. M., Drought analysis in southern Paraguay, Brazil and northern Argentina: regionalization, occurrence rate and precipitation thresholds. *Hydrology Research*, doi: 10.2166/nh.2014.074, 2014.
- [29] Bryant, F. B., Yarnold, P. R., Principal-components analysis and confirmatory factor analysis, in L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics*, Washington, DC: American Psychological Association, 1995.
- [30] Portela, M.M., Santos, J.F., Naghettini, M., Matos, J.P., Silva, A.T., Superfícies de limiares de precipitação para identificação de secas em Portugal continental: uma aplicação complementar do Índice de Precipitação Padronizada, SPI, *Recursos Hídricos*, 33(2): 5–23, Associação Portuguesa dos Recursos Hídricos, APRH), Lisboa, doi 10.5894/rh33n2-1 (<http://www.aprh.pt/rh/v33n2.html>), 2012.
- [31] Santos, J.F.; Portela, M.M.; Naghettini, M.; Matos, J.P.; Silva, A.T., Precipitation thresholds for drought recognition: a complementary use of the SPI. 7th International Conference on River Basin Management including all aspects of Hydrology, Ecology, Environmental Management, Flood Plains and Wetlands, RBM13, 3–14, Wessex Institute, New Forest, UK, ISBN:978-1-84564-712-4, ISSN:1746-448X, 2013.
- [32] Portela, M.M., Silva, A.T., Santos, J.F., Benitez, J.B., Frank, C., Reicher, J.M., Chapter 3. Analysis of Temporal Variability of Droughts in Southern Paraguay and Northern Argentina (1961–2011), in W. LEAL FILHO *et al.* (eds.), *International perspectives on climate change, Climate Change Management*, doi 10.1007/978-3-319-04489-7, Springer International Publishing Switzerland, 2014.
- [33] Hosking, J.R.M., Fortran code written for inclusion in IBM research report RC20525, Fortran routines for use with method of L-moments, IBM Research Division, T.J. Watson Research Center, New York, USA, 1996.