Presentation of Four Centennial-long Global Gridded Datasets of the Standardized Precipitation Index

Hristo Chervenkov[1], Ivan Tsonevsky[2], Kiril Slavov[3]

[1]National Institute of Meteorology and Hydrology – Bulgarian Academy of Sciences
[2]European Centre for Medium-Range Weather Forecast
[3]National Institute of Meteorology and Hydrology – Bulgarian Academy of Sciences

Abstract— *In this article four global gridded datasets of the Standardized Precipitation Index (SPI) are presented. They are computed from four different data sources: UDEL/GEOG/CCR v3.02, GPCC/ v7.0, NOAA-CIRES 20CR v2c and ECMWF ERA-20C each covering more than a century-long period. The SPI is calculated for the most frequently used time windows of 1, 3, 6, and 12 months. UDEL/GEOG/CCR v3.02 and GPCC/ v7.0 are used in the highest native resolution of 0.5×0.5° whilst NOAA-CIRES 20CR v2c and ECMWF ERA-20C are interpolated at 1.5×1.5° and 0.5×0.5° correspondingly. In contrast to some other indices, for example the popular Palmer Drought Severity Index (PDSI), SPI has significant advantages such as simplicity, suitability on variable time scales and robustness rooted in a solid theoretical development. SPI has been selected by the World Meteorological Organization (WMO) as a key indicator for monitoring drought ('Lincoln declaration'). As a result, drought monitoring centres worldwide are effectively exploiting this index and the National Meteorological and Hydrological Services (NMHSs) are encouraged to use it for monitoring meteorological droughts. These facts and the strong conviction of the authors that the free exchange of data and software services are a basis of effective scientific collaboration, are the main motivators to provide these datasets free of charge at ftp://xeo.cfd.meteo.bg/SPI/. The paper briefly presents some possible applications of the SPI data, revealing its suitability for various objective long-term drought studies at any geographical location.*

***Keywords*— Global Gridded Data-sets of SPI, Objective Drought Assessment, Free SPI-data Download.**

I. INTRODUCTION

Drought is a natural phenomenon that poses significant problems around the world, and places huge demands on rural and urban water resources as well as enormous burdens on agricultural and energy production. It is also common in terms of geography, climate and political boundaries and may be considered as a normal, recurrent feature of the climate, although a common misconception is that it constitutes an extraordinary event. In general, drought is defined as the water scarceness due to insufficient precipitation, high evapotranspiration and over-exploitation of water resources or a combination of these factors. Despite its complex nature, there is overall agreement (Barua et al 2009) that precipitation is the primary factor controlling the formation and persistence of drought conditions. Drought Indices (DIs) have been commonly used to define drought conditions. In general, a DI is a function of several hydro-meteorological variables (e.g. rainfall, temperature, streamflow, snowmelt). They can be integrated in decision support systems as a drought management tool to trigger drought relief programs. Moreover, it has been used to quantify deficits in water resources and as a drought monitoring tool. Researchers, however, are confronted with the ambiguity of drought definitions and DIs, which has never been resolved to the satisfaction of all professionals. In attempt to overcome this issue, an Inter-Regional Workshop on Indices and Early Warning Systems for Drought was held at the University of Nebraska-Lincoln from 8 to 11 December, 2009. It was jointly sponsored by the School of Natural Resources of the University of Nebraska, the U.S. National Drought Mitigation Center, the World Meteorological Organization (WMO), the U.S. National Oceanic and Atmospheric Administration (NOAA), the U.S. Department of Agriculture (USDA), and the United Nations Convention to Combat Desertification (UNCCD). The workshop reviewed the drought indices currently in use in various regions of the world to explain meteorological, agricultural and hydrological droughts, assessed the capacity for collecting information on the impacts of drought, reviewed the current and emerging technologies for drought monitoring and discussed the need for consensus standard indices for describing different types of droughts. The outcome of the workshop was the Lincoln declaration (Hayes et al 2011), in which the Standardized Precipitation Index (SPI) was proposed to be used for characterizing meteorological droughts. Moreover, NMHSs around the world were encouraged to use the SPI to characterize meteorological droughts in addition to the indices currently in use. The free availability of digital maps for the monthly precipitation sums in the recent decades, either from objective analysis or from reanalysis, has encouraged the authors to compute the SPI for the frequently used time windows of 1, 3, 6, and 12 months (noted traditionally as SPI-1, SPI-3, SPI-6 and SPI-12) from four sources of data and for the full time length of each dataset. Consequently, following our strong conviction that the free exchange of data and software services are a basis of effective scientific collaboration, we will provide these results free of charge. Main aim of this study is to present these datasets rather than perform comprehensive drought climatology for a selected region and time spans. Thus, the examples presented in this paper should be considered as a small illustration of the wide variety of potential applications at any possible geographical location, depending on the particular interest of each end-user.

The paper is organized as follows: Section 2 provides a description of some theoretical aspects of the SPI, its strengths, limitations and application for objective drought assessment. Section 3 contains a concise description of the used precipitation datasets. The calculations performed, validation of the output and description of one problem, which arose during the computational process, are presented in Section 4. In Section 5 some illustrative examples and qualitative comparisons are provided. The short summary and conclusion are placed in Section 6.

II. THEORETICAL ASPECTS, STRENGTS, WEAKNESSES AND LIMITATIONS OF THE SPI

The definition of the SPI is part of many publications and a short summary is presented in this section.

The SPI was developed by McKee et al (1993) for monitoring drought conditions based on precipitation sums. It is computed by fitting a cumulative distribution function (CDF) to the distribution of precipitation accumulated over the time window of interest. This is performed separately for each month (or whatever the temporal basis (time window) is of the raw precipitation time series) and for each location. The precipitation time series can be modelled using different statistical distributions. The gamma distribution is used most frequently with a probability density function (PDF) defined as

$$g(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad \text{for} \quad x > 0, \tag{1}$$

where $\alpha>0$ is a shape parameter, $\beta>0$ is a scale parameter, and x>0 is the amount of precipitation. $\Gamma(\alpha)$ is the gamma function, which is defined as

$$\Gamma(\alpha) = \lim_{n\to\infty} \prod_{v=0}^{n-1} \frac{n! n^{y-1}}{y+v} \equiv \int_0^{\infty} y^{\alpha-1} e^{-y} dy \tag{2}$$

Fitting the distribution to the data requires α and β to be estimated. Edwards and McKee (1997) suggest estimating these parameters using the approximation of Thom (1958) for maximum likelihood as follows:

$$\hat{\alpha} = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right) \tag{3}$$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \tag{4}$$

where, for n observations (or, more generally, time series length)

$$A = \ln(\bar{x}) - \frac{\sum_{i=1}^{n} \ln(x_i)}{n}, \tag{5}$$

Where the over bar means temporal averaging: $\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i$

This approach can be refined using an iterative procedure (Wilks, 1995). After estimating coefficients α and β the PDF $g(x)$ is integrated with respect to x and we obtain an expression for the CDF for the amount of for a given month and for a specific time scale:

$$G(x) = \int_0^x g(x)dx = \frac{1}{\hat{\beta}^{\hat{\alpha}}\Gamma(\hat{\alpha})}\int_0^x x^{\hat{\alpha}-1}e^{-x/\hat{\beta}}dx \tag{6}$$

Substituting *t* for $x/\hat{\beta}$ reduces Equation (6) to

$$G(x) = \frac{1}{\Gamma(\hat{\alpha})}\int_0^x t^{\hat{\alpha}-1}e^{-1}dt, \tag{7}$$

which is the incomplete gamma function. Since the gamma distribution is not defined for *x=0*, and q=P(x=0)>0, where *P(x=0)* is the probability of zero precipitation, the cumulative probability becomes

$$H(x) = q + (1-q)G(x) \tag{8}$$

The cumulative probability is then transformed into a normal standardized distribution (i.e. with null average and unit variance), employing the approximate conversion provided by Abramowitz and Stegun (1965) as proposed in Edwards and McKee (1997), which will not be listed here.

The strength of the anomaly is classified into groups as shown in Table 1. Hence the SPI is uniquely related to probability and the corresponding probabilities of occurrence of each group are also provided.

TABLE 1
VALUES OF SPI AND CORRESPONDING CATEGORIES AND PROBABILITIES (AFTER LLOYD-HUGHES AND SAUNDERS, 2002)

SPI value	Category	Probability %
2.00 or more	Extremely wet	2.3
1.50 to 1.99	Severely wet	4.4
1.00 to 1.49	Moderately wet	9.2
0 to 0.99	Mildly wet	34.1
0 to −0.99	Mild drought	34.1
−1.00 to −1.49	Moderate drought	9.2
−1.50 to −1.99	Severe drought	4.4
−2 or less	Extreme drought	2.3

It is possible that precipitation in some regions or at particular time scales may be modelled better by a distribution other than the gamma. If, however, different probability distributions and models are used to describe an observed series of precipitation, then different SPI values can be obtained (Guttmann, 1999). Gutmann presented the impact of various probability models on the CDF calculation, respectively on the SPI values and the dry event characteristics. It is concluded that the three-parameter Pearson Type III distribution is the “best” universal model, and that the reliability of the SPI depends on the sample size. Lana et al (2001) found that the Poisson-gamma distribution was suitable for modelling precipitation in Catalonia.

To assess how well a given distribution describes the data, Lloyd-Hughes and Saunders (2002) compared the empirical cumulative probability distribution with the corresponding theoretical one over Europe applying the Kolmogorov–Smirnov statistical test using the dataset of the Climatic Research Unit (CRU) at the University of East Anglia for the period 1901–98. They have shown, that the gamma distribution provides a good fit for all months at the time scales of 3 and 12 months.

In the recent decade the SPI is widely used worldwide in both research and operational modes. For example, Bussay et al (1998, 1999) and Szalai and Szinell (2000) assessed the utility of the SPI for describing drought in Hungary. They concluded that the SPI was suitable for quantifying most types of drought events. Stream flow was described best by SPIs with time scales of 2–6 months. A strong relationship to ground water levels was found at time scales of 5–24 months. Agricultural drought (proxied by soil moisture content) was replicated best by the SPI on a scale of 2–3 months. Belayneh and Adamowski (2012) have used SPI values over multiple lead times in the Awash River Basin in Ethiopia as a primary drought indicator. Karavitis et al (2011) have investigated the applicability of the SPI in Greece, and in the work of Barua et al (2009) comparative study of the SPI-performance and other DI for selected drought episodes of Yarra River Catchment in Victoria, Australia, has been carried out. Lana et al (2001) used the SPI to investigate patterns of rainfall over Catalonia, Spain. Lloyd-Hughes and Saunders (2002) have carried out comprehensive European drought climatology, based partly on the SPI. A significant strength of this study is the selection of of the input gridded dataset, instead of a limited number of station measurements.

In many articles (e.g. Hayes et al, 1999) the advantages and disadvantages of using the SPI for drought severity assessment have been discussed. The SPI has three main advantages. The first and primary advantage is its simplicity. The SPI is based solely on the rainfall and requires only the computation of two parameters, compared with the 68 computational terms needed to describe the PDSI. The SPI is not affected adversely by topography as well. The second advantage of SPI is its variable time scale, which allows it to describe drought conditions important for a range of meteorological, agricultural, and hydrological applications. This temporal versatility is also helpful for the analysis of drought dynamics, especially the determination of their onset and cessation, which have always been difficult to track with other indices. The third advantage comes from its standardization, which ensures that the frequency of extreme events at any location and on any time scale is consistent. Another strength, usually dismissed in the SPI description, is that this index is defined without any empiricism, in particular 'local parametrizations', which, more or less, depend on the geographical location and/or dataset. Similarly, the consideration of different empirical relationships in the attempts to generalize the SPI, including the evapotranspiration (Vicente-Serrano, 2012) is also problematic. The SPI has three potential disadvantages, the first being the assumption that a suitable theoretical probability distribution can be found to model the raw precipitation data prior to standardization. An associated problem is the quantity and and reliability of the data used to fit the distribution. This issue will be addressed in Section 4. A second limitation of the SPI arises from the standardized nature of the index itself; extreme droughts (or any other drought threshold) measured by the SPI, when considered over a long time period, will occur with the same frequency at every location. Thus, the SPI is not capable of identifying regions that may be more ‘prone’ to drought than others. A third problem may arise when applying the SPI at short time scales (1, 2, or 3 months) to regions of low seasonal precipitation. Strictly speaking, as it will be shown later, the definition of the SPI has to be extended in order to account zero precipitation. In these cases, misleadingly large positive or negative SPI values may result. The work of Wu et al (2012) is dedicated partly on this problem. Based on the results presented in this study, the authors recommend that users should be cautious when applying short-time-scale SPIs in arid climatic regimes.

III. DATA

In the recent decades, objective analyses and reanalyses have been extensively developed by various institutions and used in many applications. Depending on the leading physical and mathematical concept, involved data streams and, correspondingly, the incorporated processing methods, they can vary greatly, but generally can be divided into two main groups. The first one accommodates methods, in which the meteorological parameter of interest is analyzed separately from the others and this technique is known as objective analysis. In contrast, in the reanalysis, the assimilation procedure is combined with a dynamical processor minimizing both model simulation and observational errors achieving 'best estimate' of the atmosphere for a given historical moment. The dynamical link between all involved quantities in the model ensures the physical consistency of all output data. The final product of all the analysis and reanalyses is a timely continuous digital map of gridded datasets. Their relatively long (in climatological sense, i. e. in order of decades) temporary extent, acceptable horizontal resolution, presence of subsets for various variables, and, last but not least, their free availability make them suitable for calculating the SPI. Moreover, the used analysis system ensures largely the spatial and temporal homogeneity, which is an inherent problem by the observational point measurements. We have used in our study two objective analysis datasets, namely the Version 3.02 of the “Terrestrial Precipitation: 1900-2010 Gridded Monthly Time Series” of the Department of Geography of the University of Delaware (Peterson et al, 1998), noted further as UDEL and the Global Precipitation Climatology Centre (GPCC) Full Data Reanalysis version 7.0 (Schneider et al, 2015), noted as GPCC. They both are global and data with a resolution of 0.5×0.5° are used for the SPI calculation. UDEL datasets cover the period from 1900 to 2010 and GPCC from 1901 to 2013. The reanalysis datasets used in this study are the version v2c of the NOAA-CIRES 20CR (Compo et al, 2011) noted as CIRES and ERA-20C produced by the European Centre for Medium-Range

Weather Forecasts (ECMWF) (Stickler et al, 2014). These datasets have also global coverage for the period 1851-2011 and 1900-2010 respectively. The raw data of CIRES are in a Gaussian irregular grid and for the purposes of this study they are interpolated with the cdo tool (http://www.mpimet.mpg.de/cdo) at the 1.5×1.5°. This gridcell size is roughly equal to the original one and a single CIRES gridcell accommodates nine UDEL or GPCC ones. Following our general concept to ensure intercomparability between all products, ERA-20C dataset is downloaded at a resolution of 0.5×0.5° which is higher than the native one of approximately 125 km.

IV. CALCULATIONS AND THEIR VALIDATION

In this paper a classical approach for the SPI computation is applied proposed by McKee et al (1993). The values of SPI-1, SPI-3, SPI-6 and SPI-12 are computed with our own code, written in FORTRAN 90/95 using subroutines from the library "Numerical recipes in Fortran 90" (Press et al, 1996) and some ideas from the original Kleist's FORTRAN77 source-code. The calculations are performed over the whole spatial and temporal coverage of each input dataset, described in the previous section.

Data for monthly precipitation sums and the SPI from CARPATCLIM project (JRC report, 2010 and the references therein) have been used to validate our calculations. The two computations were found to be in excellent agreement. The European Drought Observatory of the Joint Research Centre (EDO-JRC, http://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1000) publishes also SPI maps of recent drought episodes. Although computed with SYNOP data and valid for a different reference period they are very close to our results.

The selection of the reference period for the calculation of the PDF parameters is usually not discussed in detail, though it is very relevant (Guttman, 1999). McKee et al (1993) recommend using time series with at least 30-year length and thus the current WMO standard reference period 1961-1990 seems an appropriate choice. More generally, WMO recommends that precipitation totals for at least 30 years should be used as a reference time-line for calculating rainfall statistics. Several studies, however, suggest calculating the statistics for the SPI from even longer time periods (e.g., 50 or more years) for an accurate representation of extreme events. In order to ensure comparability of the results across the world and across scales it is highly recommended the use of a common reference period for the the SPI calculation. Considering the results of an inventory of the reference periods used in various WMO Member States, specific needs for accurately representing extreme events, and possible changes in the rainfall regimes due to climate change, the Water Scarcity and Drought Expert Group strongly recommends using the period from January 1971 to December 2010 as a reference period for the calculation of the SPI. This period is used in our study and, to ensure equal number of time windows for all SPIs, the data for the precipitation sums from February 1970 to December 1971 are supplied if necessary for the computation of SPI-3, SPI-6 and SPI-12. The SPI is set to undefined value in three cases:

• if the input precipitation data is undefined. This is the case over the sea in UDEL and GPCC.

• traditionally, if the current precipitation sum can not be calculated which is the case in the first two months for SPI-3, the first five months for SPI-6 and first eleven months for SPI-12 respectively for each database.

The third case is the most interesting one and it will be described thoroughly. Let's assume, that all precipitation sums in a given time window are very close to a certain value $\tilde{x}$. Then each member x can be represented as a sum of this value and a small deviation x_i', which can be positive or negative and $|x_i'| << \tilde{x}$. Thus equation (5) can be reformulated as follows:

$$A = \ln\left(\frac{\sum_{i=1}^{n}(\tilde{x}+x_i')}{n}\right) - \frac{\sum_{i=1}^{n}\ln(\tilde{x}+x_i')}{n} = \ln\left(\tilde{x} + \frac{\sum_{i=1}^{n}(x_i')}{n}\right) - \frac{\sum_{i=1}^{n}\ln\left[\tilde{x}\left(1+\frac{x_i'}{\tilde{x}}\right)\right]}{n} \tag{9}$$

According the basic principles of the perturbation theory $\sum_{i=1}^{n}(x_i')\Big/n = \bar{x}' = 0$. Keeping in mind that the logarithm from the product is equal to the sum of logarithms and $|x_i'|\big/\tilde{x} << 1$ the last term can be replaced with its Taylor series representation:

$$\ln(1+p)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n},\quad |p|<1 \tag{10}$$

or

$$\ln(1+p)=p-\frac{p^2}{2}+\frac{p^3}{3}-\frac{p^4}{4}=p+O(p^2), \tag{11}$$

which yields in our case

$$\ln(1+\frac{\overline{x_i'}}{\tilde{x}})\approx\frac{\overline{x_i'}}{\tilde{x}} \tag{12}$$

Thus, we obtain finally:

$$A=\ln(\tilde{x})-\ln(\tilde{x})-\frac{1}{n}\sum_{i=1}^{n}\frac{x_i'}{\tilde{x}}=-\frac{\overline{x_i'}}{\tilde{x}}=0 \tag{13}$$

or in a linear truncation approximation the parameter A is zero. In Eq. (3) this parameter is in the denominator which can lead to unlimited growth of $\hat{\alpha}$. Being iterative, the computational algorithm of the incomplete gamma function cannot converge, at least for the number of iterations prescribed in advance in the corresponding subroutine. In this case the CDF degrades to the Heaviside step function $H(\tilde{x})$ and, consequently, the PDF to the Dirac delta function $\delta(\tilde{x})$.

Although this case might seem exotic, when calculating SPI-1 it occurred in some gridpoints over Africa where the monthly sums for all 40 years was very close to 1 mm. Although, as shown above, this problem can arise for every $\tilde{x}$, such cases have not been detected during the computational process of the other three SPIs (SPI-3, SPI-6 and SPI-12). A probable explanation is that the longer accumulation periods ensure spreading of the rain sums, which prevent any possible computational instability caused by the reasons explained in this paragraph.

V. SOME EXAMPLES AND QUALITATIVE COMPARISONS

As a result of this work spatial and temporal arrays of the four SPIs are obtained in the spatial resolution described above and retaining the temporal extent of the corresponding input datasets. This structure facilitates the combined analysis in a parallel mode. Hence each map, drawn after interpolation of unstructured data, for instance the point measurements in the SYNOP stations as in the work of Karavitis et al (2011), depends on the subjective choice of the interpolation procedure, whilst presented datasets in a 'native' regular grid are much more consistent.

The output data are written in standard meteorological file formats and will be open to the scientific community.

As far as any drought climatology is beyond the scope of the presented work, only some illustrative examples and qualitative comparisons will be shown, which reveals possible applications of the SPIs datasets.

A typical issue of the long-term drought climatology is the analysis of historical drought events. The EDO-JRC, for example, maintains a website for drought episodes, which is updated regularly. A remarkable case in their database is the drought episode in Europe in 1989-1991, which is among the 21 biggest droughts since 1950. Southern Europe and the Mediterranean were particularly affected. Figures 1-4 show the spatial distribution of the SPI-6 in the middle year, 1990. The 'mildly wet' and 'mild drought' categories are shown blank.

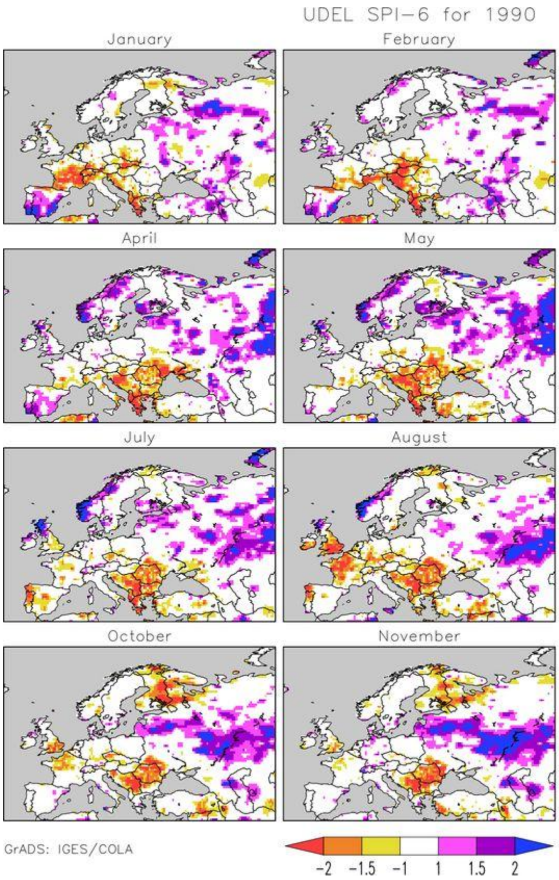

FIG. 1 DISTRIBUTION OF THE SPI-6 OVER EUROPE, OBTAINED FROM UDEL

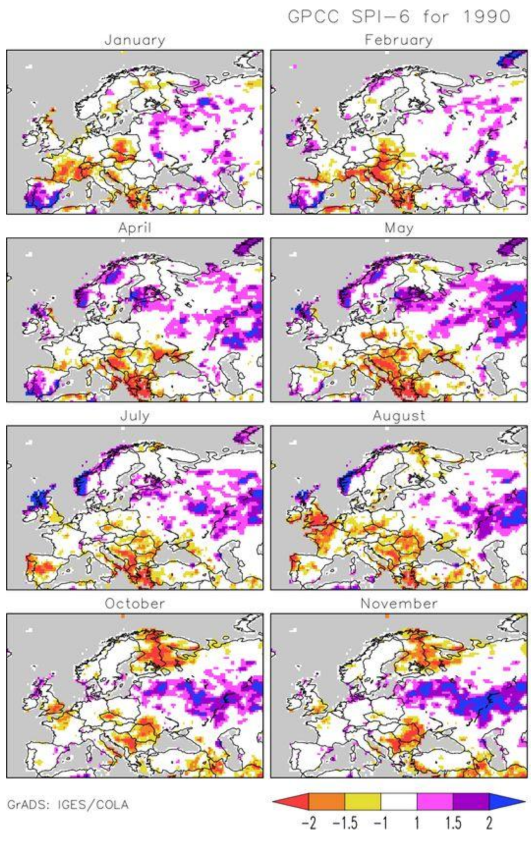

FIG. 2 AS FIGURE 1, BUT FOR GPCC

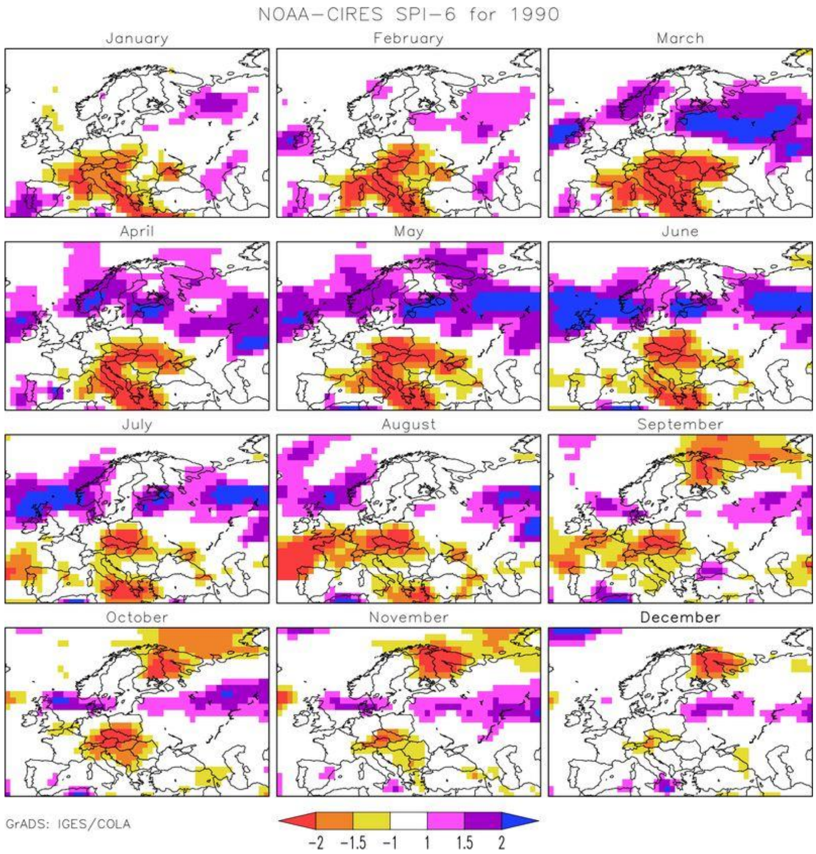

FIG. 3 AS FIGURE 1, BUT FOR NOAA-CIRES

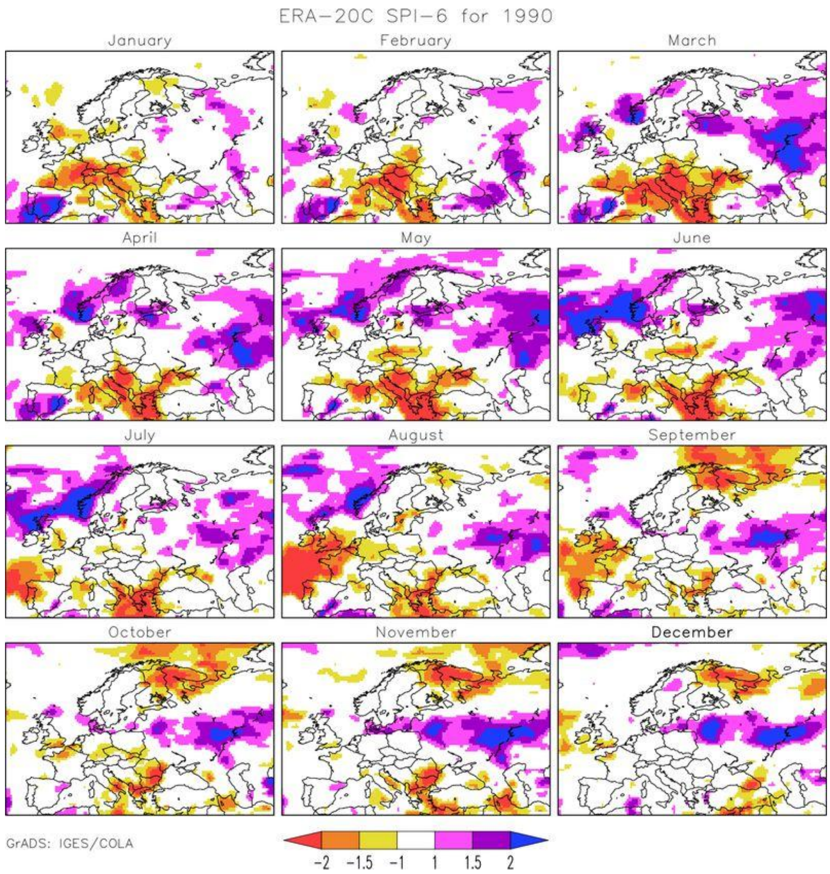

FIG. 4 AS FIGURE 1, BUT FOR ERA-20C

The situation worsened by the fact, that the most severe manifestation of the drought was during the active vegetation period from March to September 1990. The main extremes can also be seen in the other three SPIs (not shown here).

Figures 1-4 reveal that the most significant anomaly patterns (according to their magnitude and extent) occurred over the same regions in the same months, although with some discrepancies in the details. The greatest similarities can be seen between the two arrays, derived from the two measurement-only-based datasets (UDEL and GPCC) and on the other hand between the datasets with the 0.5×0.5° resolution (UDEL, GPCC and ERA-20C) and that with the coarser 1.5×1.5° resolution (NOAA-CIRES). The well-known effect of the spatial smoothing is clearly pronounced here: the structures are not so sharp, but they cover a larger area. It is also worth emphasizing that the NOAA-CIRES and ERA-20C based products, in contrast to the other two, are not limited just over the land, which can be important for some applications.

The presented datasets can be post-processed statistically in order to derive some complimentary measures, which are suitable for climatological analysis. For example, Lloyd-Hughes and Saunders (2002) show maps of the number of drought events. Similarly, on Figures 5-8 we present maps of the number of months in each category except 'near normal'. Calculated in moving windows and/or for fixed intervals (for example decades), this measure would be suitable for trend estimations. For the statistical calculations, the longest period of 110 years (1901-2010) common for all the datasets is considered, which ensures a fair comparison between quantities derived from all the datasets.

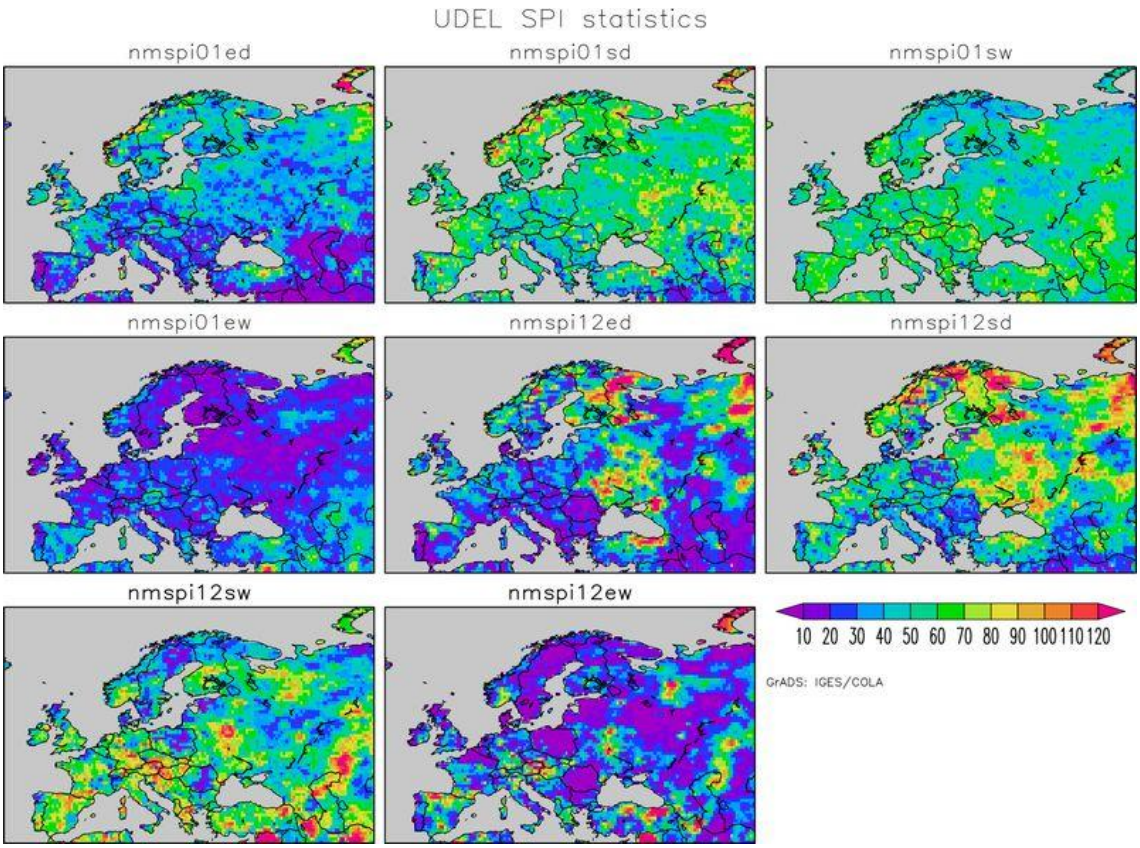

FIG. 5 SPI-STATISTICS, OBTAINED FROM UDEL. NUMBER OF EXTREME DROUGHT (NMSPI01ED), SEVERE DROUGHT (NMSPI01SD), SEVERELY WET (NMSPI01SW) AND EXTREMELY WET (NMSPI01EW) MONTHS ACCORDING SPI-1 AND THE SAME QUANTITIES, BUT ACCORDING SPI-12 NOTED CORRESPONDINGLY AS NMSPI12ED, NMSPI12SD, NMSPI12SW AND NMSPI12EW.

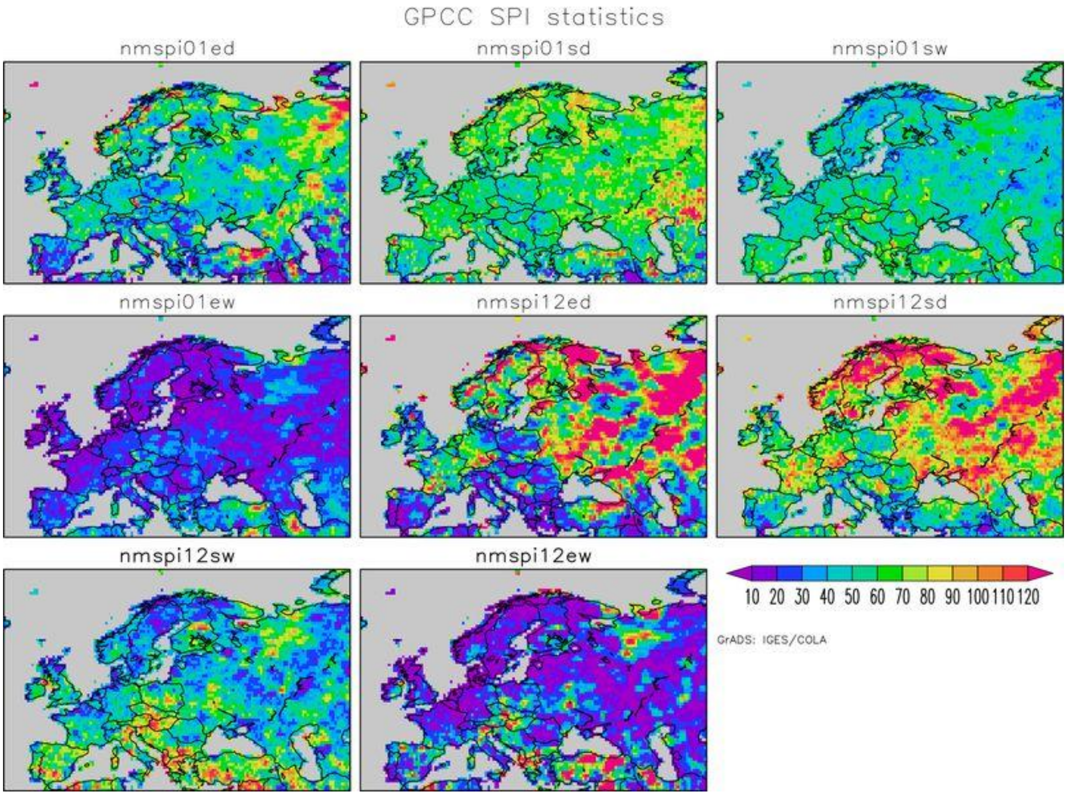

FIG. 6 AS FIGURE 5, BUT FOR GPCC.

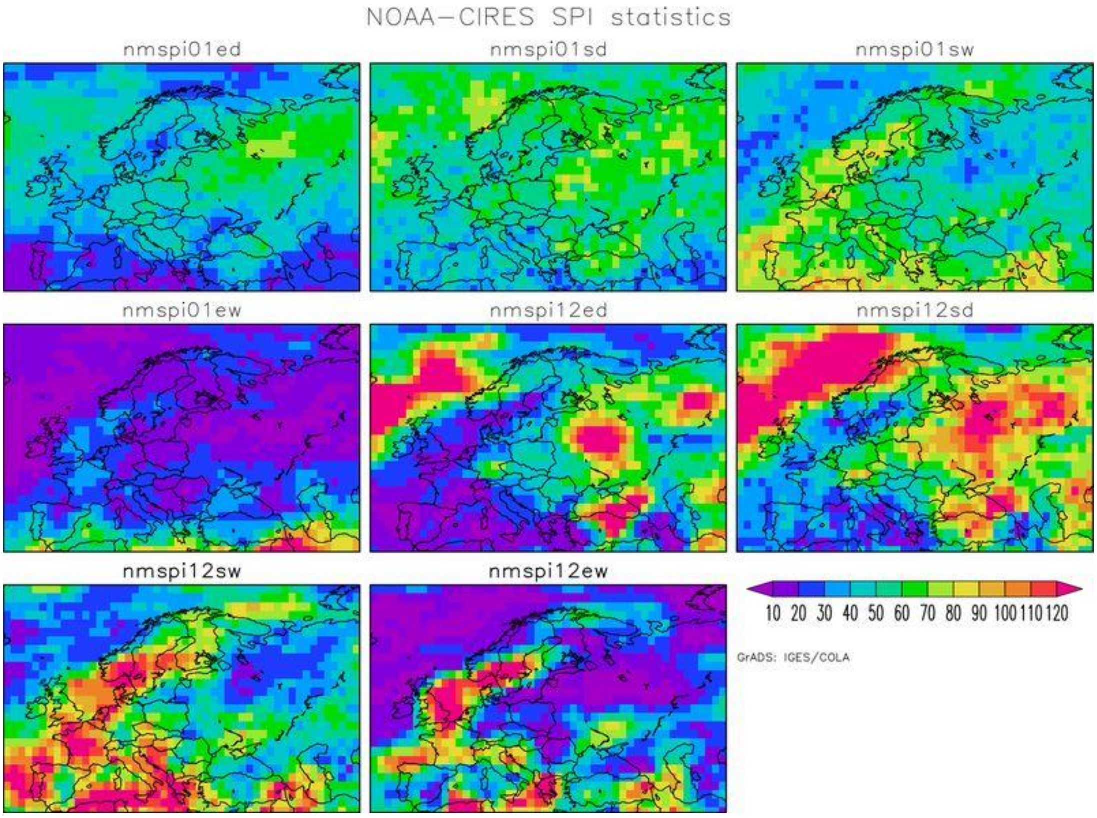

FIG. 7 AS FIGURE 5, BUT FOR NOAA-CIRES.

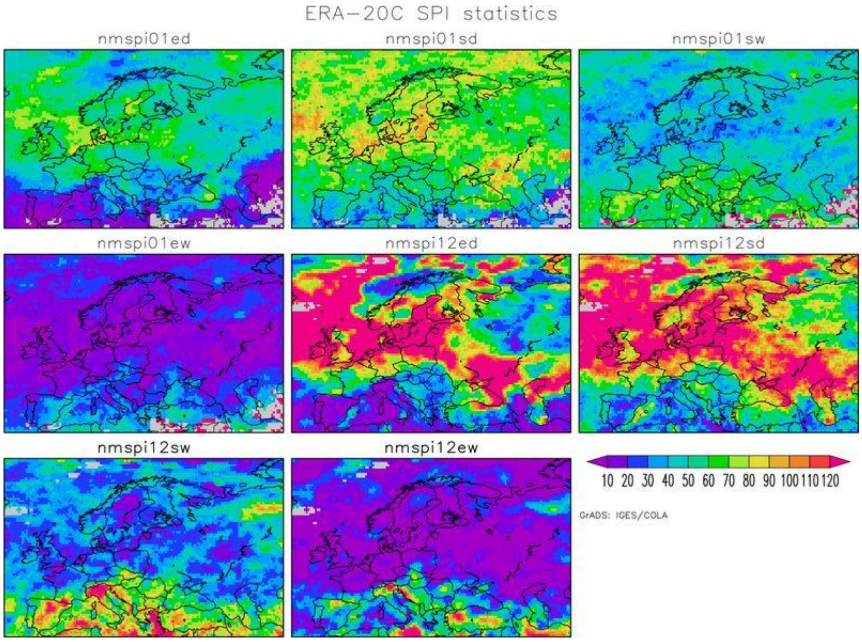

FIG. 8 AS FIGURE 5, BUT FOR ERA-20C.

Figures 5-8 show that the distribution of the number of months in different categories is quite similar for all the datasets.

Finally, the presented datasets offer the possibility to extract certain information for a particular (grid) point of interest as chronograms, which is another traditional way for data analysis and visualization, as shown on Figure 9.

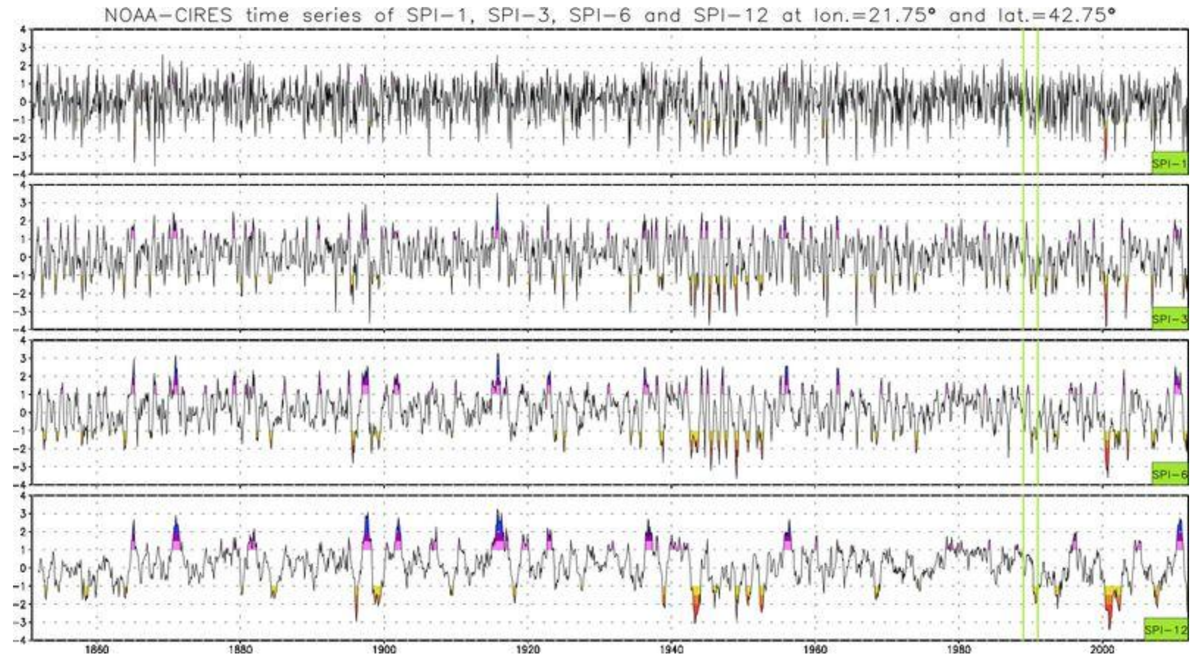

FIG. 9 TIME SERIES OF SPI-1, SPI-3, SPI-6 AND SPI-12 FOR A GRID POINT IN SOUTH-EAST EUROPE OBTAINED FROM NOAA-CIRES. THE AREA BETWEEN THE TWO GREEN VERTICAL LINES DENOTE THE DROUGHT EVENT 1989-1991.

The time-series of the SPIs of the dataset with the longest time-span, NOAA-CIRES, are plotted only. Such subsets can also be post-processed, especially applying techniques for trend analysis or for searching periods of a cyclic repetition of predefined anomalies.

VI. SUMMARY AND CONCLUSION

Although many researchers argue (e.g., Barua et al, 2009) that rainfall-based DIs are not meaningful enough to define the wider drought conditions, their suitability was proven in numerous studies for most parts in the world. Despite their limitations, these indicators offer pragmatic approach for quantitative estimation of complex phenomena. This has been especially true since the rise of the digital era, when reliable datasets comprising plenty of meteorological and hydrological parameters have become available. This allows calculations of such 'secondary' quantities as the SPI routinely for climatologically significant intervals (in the order of decades) over the whole world.

The few examples, presented in this paper, show general agreement between the four output datasets. They are not equivalent, however, and more or less each of them inherits the strengths and weaknesses of the corresponding input data. This has to be taken into account for certain applications.

In certain cases, however, as shown in Section 4 and in some other studies (e.g. Wu et al., 2012) the user should be cautious and analyze the SPI values carefully.

The mapped statistical measures, which result from very simple preprocessing, demonstrate the great variety of possibilities for quantitative analysis, based on various statistical methods.

The datasets, presented in this study, and their availability *a priori* (i.e. before the start of any drought study) save computational time and effort.

They can be used by a wide community of researchers and decision-makers either separately or, as we recommend, in conjunction with and/or in addition to other methods achieving comprehensive drought assessment in every region of interest.

ACKNOWLEDGEMENTS

This study is based entirely on public data and open source software and the authors would like to express their deep gratitude to all the providers (University of Delaware, GPCC/DWD, ECMWF, NOAA, NCEP-NCAR, Unidata, MPI-M and all others). This study would not be possible without their innovative data services and tools. Many thanks also to all third-party organizations for the funding: Support for the Twentieth Century Reanalysis Project version 2c dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. The authors are grateful to Dr. E. Dutra at ECMWF for the valuable suggestions and to C. Kodama and B. Guan for the public and useful GrADS scripts.

REFERENCES

[1] Abramowitz, M., Stegun, A. (eds). (1965) Handbook of Mathematical Formulas, Graphs, and Mathematical Tables. Dover Publications, Inc.: New York.

[2] Barua, Sh., Perera, B. J. C., Ng, A. W. M. (2009) A comparative drought assessment of Yarra River Catchment in Victoria, Australia. In: Interfacing modelling and simulation with mathematical and computational sciences: 18th IMACS World Congress, MODSIM09, Cairns, Australia 13-17 July 2009: proceedings. Anderssen, R. S and Braddock, R. D and Newham, L. T. H, eds. Modelling and Simulation Society of Australia and New Zealand, Canberra, ACT, pp. 3245-3251. ISBN 9780975840078

[3] Belayneh, A. and Adamowski, J. (2012) Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Support Vector Regression, Applied Computational Intelligence and Soft Computing, vol. 2012, Article ID 794061, 13 pages, 2012. doi:10.1155/2012/794061

[4] Bussay, A., Szinell C, Hayes, M, Svoboda, M. (1998) Monitoring drought in Hungary using the standardized precipitation index. Annales Geophysicae, Supplement 11 to Vol 16, the Abstract Book of 23rd EGS General Assembly, C450. April 1998, Nice; France.

[5] Bussay, A, Szinell, C., Szentimery, T. (1999) Investigation and Measurements of Droughts in Hungary. Hungarian Meteorological Service: Budapest.

[6] Compo, G.P., J.S. Whitaker, P.D. Sardeshmukh, N. Matsui, R.J. Allan, X. Yin, B.E. Gleason, R.S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R.I. Crouthamel, A.N. Grant, P.Y. Groisman, P.D. Jones, M. Kruk, A.C. Kruger, G.J. Marshall, M. Maugeri, H.Y. Mok, Ø. Nordli, T.F. Ross, R.M. Trigo, X.L. Wang, S.D. Woodruff, and S.J. Worley, (2011) The Twentieth Century Reanalysis Project. Quarterly J. Roy. Meteorol. Soc., 137, 1-28. DOI: 10.1002/qj.776

[7] Guttman, N. B. (1999) Accepting the standardized precipitation index: a calculation algorithm. JAWRA Journal of the American Water Resources Association, 35: 311–322. doi: 10.1111/j.1752-1688.1999.tb03592.x

[8] Hayes, M., Svoboda, MD., Wilhite, DA., Vanyarkho, OV. (1999) Monitoring the 1996 drought using the standardized precipitation index. Bulletin of the American Meteorological Society 80: 429–438.

[9] Hayes, M., Svoboda, M., Wall, N., Widhalm, M., (2011) The Lincoln Declaration on Drought Indices: Universal Meteorological Drought Index Recommended. Bull. Amer. Meteor. Soc., 92, 485–488, doi: http://dx.doi.org/10.1175/2010BAMS3103.1

[10] JRC, 2010: Climate of the Carpathian Region. Technical Specifications (Contract Notice OJEU 2010 /S 110-166082 dated 9 June 2010)

[11] Edwards DC, McKee TB. (1997) Characteristics of 20th century drought in the United States at multiple timescales. Colorado State University: Fort Collins. Climatology Report No. 97-2. Karavitis, C., Alexandris, S., Tsesmelis D., Athanasopoulos, G., Application of the Standardized Precipitation Index (SPI) in Greece, Water 2011, 3(3), 787-805; doi:10.3390/w3030787

[12] Lana, X, Serra, C, Burgueno, A. (2001) Patterns of monthly rainfall shortage and excess in terms of the standardized precipitation index. International Journal of Climatology 21: 1669–1691.

[13] Lloyd-Hughes, B; Saunders, MA; (2002) A drought climatology for Europe. INT J CLIMATOL, 22 (13) 1571 - 1592. 10.1002/joc.846.

[14] McKee, TB., Doesken, NJ., Kleist, J. (1993) The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference on Applied Climatology, 17–22 January, Anaheim, CA. American Meteorological Society: Boston, MA; 179–184.

[15] Peterson, T. C., R. S. Vose R. Schmoyer and V. Razuvaëv (1998). Global Historical Climatology Network (GHCN) Quality Control of Monthly Temperature Data. International Journal of Climatology, 18, 1169-1179.

[16] Press, W., Teukolsky, S., Vetterling, W., Flanner, B. (1996) Numerical recipes in Fortran 90 (2nd ed.): the art of parallel scientific computing, Cambridge University Press New York, NY, USA ISBN:0-521-57439-0

[17] Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B. Ziese, M. (2015) GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI: 10.5676/DWD_GPCC/FD_M_V7_050

[18] Stickler, A., S. Brönnimann, M. A. Valente, J. Bethke, A. Sterin, S. Jourdain, E. Roucaute, M. V. Vasquez, D. A. Reyes, R. Allan, and D. Dee, (2014) ERA-CLIM: Historical Surface and Upper-Air Data for Future Reanalyses. Bull. Amer. Meteor. Soc., 95(9), 1419-1430 (DOI: 10.1175/BAMS-D-13-00147.1).

[19] Szalai, S., Szinell, C. (2000) 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.

[20] Thom, HCS. (1958) A note on the gamma distribution. Monthly Weather Review 86: 117–122.

[21] Vicente-Serrano, SM., Beguería, S., Lorenzo-Lacruz, J., Camarero, JJ., López-Moreno, J., Azorin-Molina, C., Revuelto, J., Morán-Tejeda, E., Sanchez-Lorenzo, A., (2012) Performance of drought indices for ecological, agricultural, and hydrological applications. Earth Interact., 16, 1–27. doi: http://dx.doi.org/10.1175/2012EI000434.1

[22] Wilks, DS. (1995) Statistical Methods in the Atmospheric Sciences. Academic Press: London.

[23] Wu, H., Svoboda, M. D., Hayes, M. J., Wilhite, D. A. and Wen, F. (2007), Appropriate application of the standardized precipitation index in arid locations and dry seasons. Int. J. Climatol., 27: 65–79. doi:10.1002/joc.1371