AGRICULTURAL DROUGHT ASSESSMENT IN LATIN AMERICA BASED ON A STANDARDIZED SOIL MOISTURE INDEX

Hugo Carrão, Simone Russo, Guadalupe Sepulcre, and Paulo Barbosa

Climate Risk Management Unit, Institute for Environment and Sustainability, Joint Research Centre, European Commission, Via E. Fermi 2749 TP 28, 21027 Ispra (VA), Italy, Email: hugo.carrao@jrc.ec.europa.eu, simone.russo@jrc.ec.europa.eu, guadalupe.sepulcre@jrc.ec.europa.eu, paulo.barbosa@jrc.ec.europa.eu

ABSTRACT

We propose a relatively simple, spatially invariant and probabilistic year-round Standardized Soil Moisture Index (SSMI) that is designed to estimate drought conditions from satellite imagery data. The SSMI is based on soil moisture content alone and is defined as the number of standard deviations that the observed moisture at a given location and timescale deviates from the long-term normal conditions. Specifically, the SSMI is computed by fitting a non-parametric probability distribution function to historical soil moisture records and then transforming it into a normal distribution with a mean of zero and standard deviation of one. Negative standard normal values indicate dry conditions and positive values indicate wet conditions. To evaluate the applicability of the SSMI, we fitted empirical and normal cumulative distribution functions (ECDF and nCDF) to 32-years of averaged soil moisture amounts derived from the Essential Climate Variable (ECV) Soil Moisture (SM) dataset, and compared the root-mean-squared errors of residuals. SM climatology was calculated on a 0.25° grid over Latin America at timescales of 1, 3, 6, and 12 months for the long-term period of 1979-2010. Results show that the ECDF fits better the soil moisture data than the nCDF at all timescales and that the negative SSMI values computed with the non-parametric estimator accurately identified the temporal and geographic distribution of major drought events that occurred in the study area.

Key words: Standardized Soil Moisture Index; Satellite images; Agricultural Drought; Latin America.

1. INTRODUCTION

Drought is a damaging environmental disaster and affects more people than any other natural hazard [1]. There are numerous conceptual and operational drought definitions proposed according to different disciplinary perspectives [2]. Fundamentally, drought is a temporary water supply deficiency relative to some long-term average condition. [3] and [1] proposed a drought typology based on four distinct types, namely meteorological, agricultural, hydrological and socio-economic. The various drought types represent different stages of a continuous meteorological process and reflect the perspectives of different sectors on water shortages. Thus, they occur at different timescales, but are intimately interrelated with each other. The longer the meteorological drought is, the more likely other types of droughts (namely agricultural or hydrological) will occur as a result. The focus of this paper is on agricultural drought. Agricultural drought occurs when there is not enough soil moisture to support average crop production on farms or average grass production on range land [1]. Agricultural drought diminishes crop and forest productivity, and increases fire hazards [4].

Soil moisture volume is a good indicator of agricultural drought intensity, reflecting recent precipitation and antecedent availability, and indicating cropping potential [5]. Previous studies used soil moisture as an indicator of agricultural drought onset, duration, end and severity [e.g. 6, 7]. Unfortunately, there does not exist a comprehensive and representative global network of soil moisture monitoring instruments and those indicators are based on model estimates that perform a water balance assessment of the soil column, using variables such as precipitation, air temperature, soil temperature, soil porosity, and infiltration [5]. Simulated soil moisture fields are potential surrogates for actual observations and are frequently the only viable alternative over large scales. However, for many regions of the world, land-atmosphere feedback mechanisms are not well understood [8] and, as a result, soil moisture model estimates in these regions may be prone to large uncertainties [9].

Since the 1980s, many studies have capitalized on the synoptic, timely and spatially continuous characteristics of remote sensing data from active and passive microwave sensors to analyze and monitor drought conditions over large areas where monitoring instruments are sparse or non-existent [10]. Recently, a merged soil moisture product from different satellite sensors into a single dataset covering the period 1978-2010 was presented by [11, 12, 13]. Based on this dataset, we propose an index of agricultural drought as the normalized deficit of soil moisture relative to its monthly, seasonal, half-yearly or yearly climatology at a location. This allows us to com-

Proc. 'ESA Living Planet Symposium 2013', Edinburgh, UK
9–13 September 2013 (ESA SP-722, December 2013)

pare the occurrence of drought between different locations in a consistent and meaningful way. To compute our index, we propose to model the probability distribution of soil moisture amounts with a non-parametric Kernel Density Estimator (KDE). We chose a non-parametric estimator because it avoids distribution hypotheses and therewith the associated problems with small sample datasets. Using the Standardized Soil Moisture Index (SSMI) computed with the non-parametric Kernel, we identify major drought events in Latin America region and explore their spatiotemporal characteristics.

2. DATA AND METHODS

In this section we present the soil moisture dataset, the study area, the computation process of the Standardized Soil Moisture Index (SSMI) and the validation process of the SSMI.

2.1. The Essential Climate Variable (ECV) Soil Moisture (SM) dataset

The Standardized Soil Moisture Index (SSMI) is calculated by using the ECV-SM dataset. The theoretical and algorithmic base of the product is described in [14]. This global gridded product with a spatial resolution of $0.25°$ has been generated by blending data from active and passive microwave spaceborne instruments through the method described by [11, 12, 13]. The active dataset was generated based on observations from the C-band scatterometers on board of ERS-1, ERS-2 and METOP-A. The passive data set was generated based on passive microwave observations from Nimbus 7 SMMR, DMSP SSM/I, TRMM TMI and Aqua AMSR-E. The homogenized and merged product covers a 32-year period between November 1978 and December 2010, with a daily temporal resolution. The soil moisture data are provided in volume-metric units (m^3m^{-3}).

2.2. Study area

The data used cover the whole Latin America region (the domain of analysis is limited to land surface grid-cells between $56°$S–$35°$N, $33°$–$120°$W). Latin America spans a vast range of latitudes and has a wide variety of climates. It is characterized largely by humid and tropical conditions, but important areas have been extremely affected by meteorological droughts in the past [e.g. 15, 6, 16, 17] and the climate change scenarios foresee an increased frequency of these events for the region [e.g. 18, 19]. Given the significant reliance of Latin American economies on agricultural production, and the exposure of agriculture to a variable climate, there is a large concern in the region about present and future climate and climate-related impacts [20]. Thus, it constitutes an convenient case study region to validate the SSMI.

2.3. Standardized Soil Moisture Index (SSMI)

The SSMI can be estimated for different timescales by using soil moisture average (or accumulation, maximum, median, or others) amounts from daily up to yearly, or even longer, periods. The computation of the SSMI follows a three step process (Figure 1), as similar as for the Empirical Standardized Precipitation Index (ESPI) [21]:

1. a statistical distribution is fitted to the observed soil moisture records averaged over a reference period, e.g. n months (where n is, e.g. 1-, 3-, 6-, or 12-months).

   In the ESPI formulation [21], precipitation observations are modeled by means of a non-parametric distribution. The non-parametric Probability Density Function (PDF) and the corresponding Empirical Cumulative Density Function (ECDF) are estimated by using the Kernel Density Estimator (KDE) [22]. The KDE is a non-parametric method for estimating PDF from a set of training observations x_1, x_2,..., x_N. It is also known as Parzen Windows. Given independent and identically distributed observations x_1, x_2,..., x_N, having the common PDF $f(x)$, the general kernel estimator is defined as:

   $$\hat{f}_h(x) = \frac{1}{Nh}\sum_{i=1}^{N} K\frac{(x - x_i)}{h} \qquad (1)$$

   where K is some kernel (window) function and h is the bandwidth (smoothing parameter). Frequently, K is chosen to the Gaussian function with mean 0 and variance 1:

   $$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad (2)$$

   As shown in many studies [23, 22, 24, 25, 26], since the normal kernel estimator is symmetric, it could have a significant boundary bias for bounded data like the ECV-SM, which can only take values on the interval $(0, 1)$, i.e. $f(x) = 0$ for $x \notin (0, 1)$. In such situations, it is clearly desirable that the estimator $\hat{f}(x)$ has the same support as $f(x)$. Direct application of kernel smoothing methods does not guarantee this property, and so they need to be corrected when $f(x)$ has bounded support. To overcome this limitation, boundary bias correction methods are used in kernel density estimation [23, 27]. In the case of a non-negative variable with a two-sided bounded support, like the ECV-SM, the simplest method of solving this problem is to transform the data [24]. The idea is to estimate the PDF of a transformed random variable $Y = t(X)$ which has unbounded support. Suppose that the PDF of Y is given by $g(y)$. Then the relationship between f and g is given by

   $$f(x) = g(t(x))t'(x). \qquad (3)$$

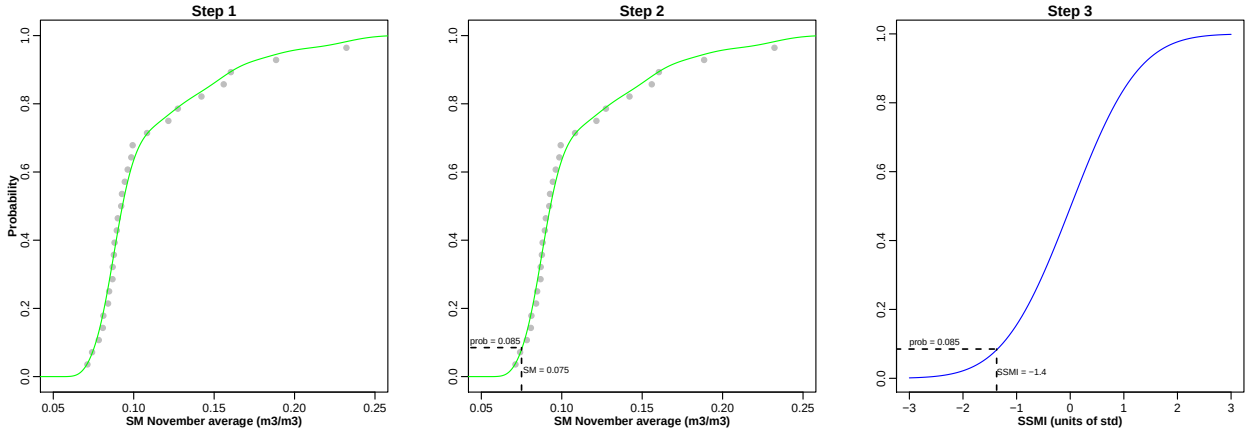

Figure 1. The three step computation process of the SSMI. The example is for the average soil moisture in November at a grid-cell in Bahia, Brazil (38.875° W, 10.875° S). The green curves in Steps 1 and 2 represent the ECDF estimated from the averaged monthly soil moisture data (gray filled circles) by means of the kernel method. The blue curve on the right represents the standard normal CDF. The exceedance probability is estimated in Step 2 for a soil moisture amount of 0.075 m^3m^{-3} and transformed to the standard normal variable in Step 3 to find the SSMI value (-1.4).

Therefore, to estimate $\hat{f}(x)$, one carries out the following steps:

(a) Transform the observations $y_i = t(x_i)$; $i = 1, 2, \ldots, n$;

(b) Apply the kernel function to estimate the PDF of $g(y)$;

(c) Estimate $f(x)$ using $\hat{f}(x) = \hat{g}(t(x))t'(x)$.

In the case of soil moisture, the support of $f(x)$ is $(0, 1)$ and a simple transformation $t : (0, 1) \to (-\infty, \infty)$ is the log-transformation $t(x) = log\left(\frac{x}{1-x}\right)$. Here $t'(x) = \frac{1}{x} + \frac{1}{1-x}$ and

$$\hat{f}(x) = \hat{g}\left(log\left(\frac{x}{1-x}\right)\right)\left(\frac{1}{x} + \frac{1}{1-x}\right). \quad (4)$$

As similar as for the ESPI, the ECDF is used in steps 2 and 3 to compute the SSMI values (Figure 1). The probabilistic characteristics of the SSMI allow us to classify several drought categories. Here, we used the drought intensity classification system suggested by [28], based on the frequency distribution of SSMI values (Table 1). In the classification system proposed by [28], every location is "near normal" 68% of the time, in "moderate" drought 9.2% of the time, in "severe" drought 4.4% of the time, and in "extreme" drought 2.3% of the time.

2. the non-exceedance probability of a soil moisture observation is computed related to the non-parametric distribution: this is simply done by evaluating the fitted ECDF at the values of the soil moisture data.

3. the non-exceedance probability is transformed to the standard normal variable (mean=0 and variance=1) and the SSMI value is found.

Table 1. Drought classification by standardized soil moisture index (SSMI) value and corresponding event probabilities according to [28]

SSMI values	**Drought class**	**Probability of Event (%)**
SSMI > 2.00	Extreme wet	2.3
1.50 < SSMI ≤2.00	Severe wet	4.4
1.00< SSMI ≤1.50	Moderate wet	9.2
-1.00< SSMI ≤1.00	Near normal	68.2
-1.50< SSMI≤-1.00	Moderate dry	9.2
-2.00 < SSMI ≤-1.50	Severe dry	4.4
SSMI≤-2.00	Extreme dry	2.3

2.4. Validation of the SSMI

To validate the SSMI, first we evaluate the goodness-of-fit of the non-parametric probability distribution fitted with the Kernel estimator to the sample frequency of soil moisture data. In detail, we compute the root-mean-square error (RMSE) of residuals for grid-cells in the study area at monthly, seasonal (December-February: DJF, March-May: MAM, June-August: JJA, September-November: SON), half-yearly (December-May: DM, June-November: JN) and yearly averaged soil moisture amounts. Then, we compare the frequency of RMSEs for the study area at different timescales with those of parametric normal distributions (nCDF) fitted to the frequency of soil moisture data for each grid-cell. The goal is to show that the non-parametric approach is more flexible and impose less restrictions than the parametric on the shapes that $f(x)$ can have. It is hypothesized that the

non-parametric fittings are more suitable for representing the frequency distribution of soil moisture values in some regions and at specific timescales where the densities are skewed or bimodal.

To compute the RMSE at all timescales and provide comparable results among regions and fitted distributions, we need the number n of soil moisture observations to be analogous for each grid-cell. Thus, daily observations are averaged to a monthly time step and these progressively averaged into seasonal, half-yearly and yearly averages. However, the daily temporal resolution of the ECV-SM product is not absolute and the number n of observations for each grid-cell is not fixed for the different timescales (Figure 2). To obtain an homogeneous dataset for the analysis, we first fixed the minimum sampling rate at one observation per month, i.e. every 30-days. Next, we selected the set of soil moisture grid-cells fulfilling this condition for at least 80% of the 32-years of data available in the time-series, i.e. $\sim$25 years. These conditions are in agreement with the recommendations of the World Meteorological Organization (WMO), regarding the length and completeness of the reference period used to compute climate normals that serve as benchmark against which anomalies are estimated [29].

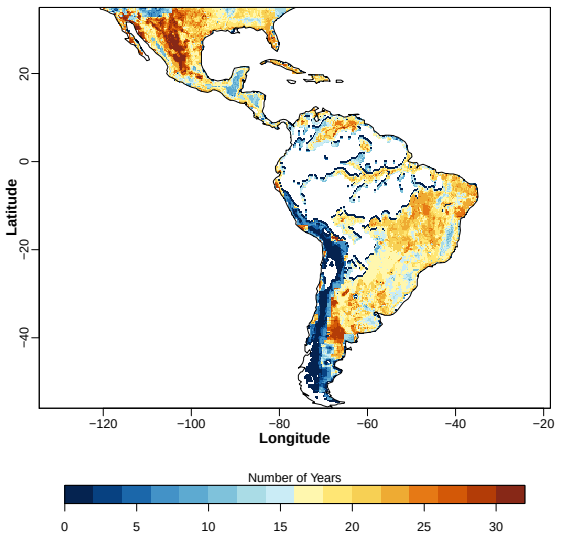

Figure 2. Number of years per grid-cell with a temporal resolution superior to 30-days (i.e. at minimum 1 soil moisture observation each month) in Latin America. No-data values are masked out in white.

Finally, to evaluate the spatiotemporal consistency of the SSMI, we analyse a time-series of the index values for a region in Bahia (northeast Brazil). We selected this region because it gathers together a set of neighboring grid-cells fulfilling the minimum sampling rate and the length of soil moisture time-series observations needed to compute the anomalies, as described above.

3. RESULTS AND DISCUSSION

The statistical distribution of the root-mean-squared errors (RMSEs) from residuals of non-parametric ECDFs and parametric nCDFs fitted to the sample frequency of soil moisture values from the selected grid-cells in the study area are presented in Figure 3. Overall, the results show that the non-parametric fits have smaller RMSEs than the parametric fits for at least 75% of grid-cells in the study area. This is valid for all timescales of analysis, thus demonstrating that the ECDF is more flexible than the nCDF in approximating the frequency shapes of soil moisture values at each grid-cell and better for fitting the frequency of soil moisture data. The differences between the flexibility of the non-parametric ECDF and parametric nCDF for fitting the frequency of soil moisture values is shown in the graphical example of Figure 4. It is noticeable that because the soil moisture data is positively skewed, the bell-shape of the nCDF is not able to accurately represent the empirical probability distribution of sample observations and the curve thus not approximate the data. On the other hand, the ECDF captures the regularities underlying the frequency of soil moisture values at this grid-cell and timescale. The accurate approximation of the ECDF curve to the data in the lower tail of the distribution allows an improved estimation of the intensity of soil dryness conditions. Consequently, more effective and efficient actions can be timely triggered to mitigate adverse impacts on vegetation growing conditions.

The results presented in Figure 3 also show that the dispersion of RMSEs, and respective upper and lower quartiles, for non-parametric fits to the frequency of soil moisture values from the selected grid-cells in the study area is analogous for all timescales of analysis. This suggests that the KDE is not unduly affected by outliers in the data and has a good performance for fitting a wide range of sample frequency shapes. The non-parametric method is robust and precise in fitting the soil moisture data averaged at different timescales and collected at geographic locations with different climate variability. On the other hand, it is noticeable that the upper and lower quartiles of RMSEs for the parametric fits to the data start decaying as long as the timsecale of soil moisture averaging periods increases. These results suggest that the shape of soil moisture frequency at each grid-cell approximates better the normal probability distribution at larger timescales. Similarly, the Interquartile Range (IQR), i.e. the difference between the upper and lower quartiles, of RMSEs for the parametric fits to the data decreases as long as the timescale increases. This finding points towards the idea that soil moisture frequency shapes become more alike at larger timescales for the grid-cells in the study area. In resume, since the dispersion and magnitude of the fitting error values in the study area is much smaller for the ECDF than for the nCDF fitted to the frequency of soil moisture values at all timescales, it can be argued that computing the SSMI with the non-parametric method provides more accurate and precise triggers of spatiotemporal drought occurrences.

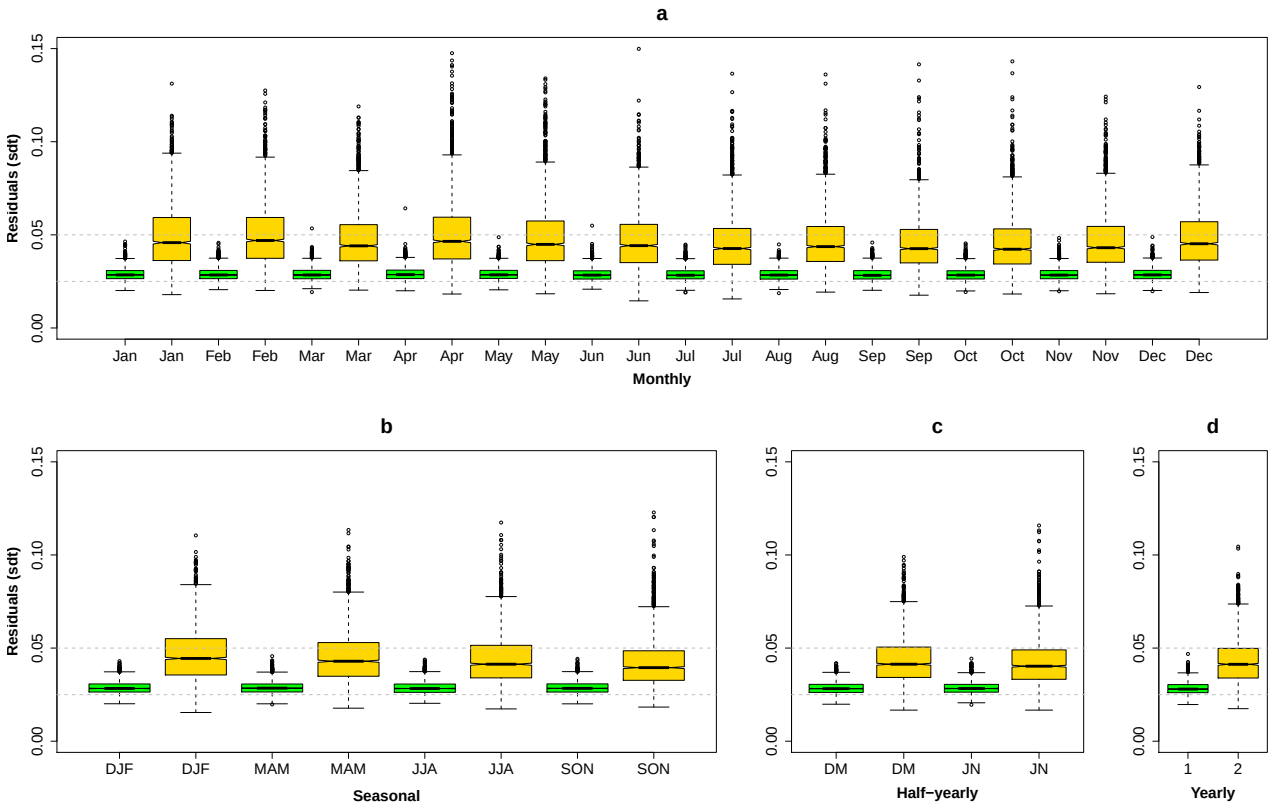

Figure 3. Statistical distribution of root-mean-squared errors (RMSEs) in Latin America for the non-parametric ECDF (green) and normal CDF (orange) fitted to the frequency of average: a) monthly soil moisture; b) seasonal soil moisture; c) half-yearly soil moisture; d) yearly soil moisture. Only grid-cells with at least 1 soil moisture observation per month for 25 consecutive years were considered for computing the results.

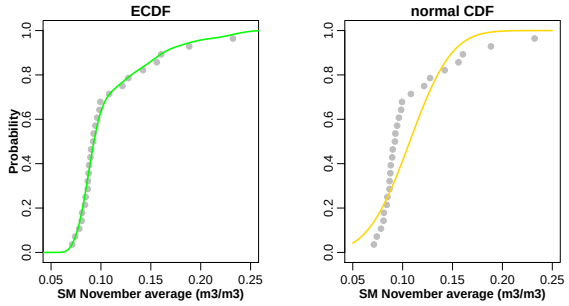

*Figure 4. ECDF (green on the left) and normal CDF (orange on the right) curves fitted to the frequency of average soil moisture (gray filled circles) in November at a grid-cell in Bahia, Brazil (*38.875° *W,* 10.875° *S).*

As a final result of this study, the non-parametric SSMI has been used to classify the dry and wet years that occurred between 1979 and 2010 in the Latin America region of Bahia, northeast Brazil. A geographic time-series of SSMI values is presented in Figure 5. The results appear to indicate that the most intense drought event taking place for the region occurred in 1993. Indeed, the whole region was under dry conditions and most of the area experienced severe and extreme drought intensity during that year, according to the classification system presented in Table 1. Our experiments are in line with previous results present by [30], which extensively described the occurrence of a severe drought in northeast Brazil in 1993. [30] suggest that the drought of 1993 is connected at least partially to unusual *El Niño*-Southern Oscillation (ENSO) conditions that occurred during that year. In Brazil, weather-related effects associated to *El Niño* and *La Niña* vary considerably across regions. During the *El Niño* phase, there is a reduction in precipitation in the north and northeast regions, while in *La Niña* years, the north-northeast regions experience higher than average rainfalls [18]. It is interesting to note that there is a connection and temporal consistency between climatic regimes during *El Niño* years and negative SSMI values for the region. For example, according to [31], the *El Niño* events were severe for the years of 1983 and 1998 and the region experienced below normal dry conditions. The negative SSMI values for 1983 and 1998 (Figure 5) are consistent with the occurrence of those events. Similarly, the wet years that are caused by *La Niña* events, such as for 1985, show a good agreement with the positive SSMI values.

Let us now look at the spatial distribution of SSMI values presented in Figure 5 for the study area. The results show

a spatial consistency of SSMI values along time for the analysed region, i.e. there are smooth transitions of SSMI categories between neighbouring grid-cells. This physical consistency is well representative of regional patterns of climate variability and spatial coherence of droughts, which tend to evolve slowly and affect large areas simultaneously [32]. The results appear to indicate that the ECV-SM product [11, 12, 13], which was derived from resampled and harmonised soil moisture datasets collected by several active and passive remote sensors, is well calibrated and provide an homogenized soil moisture climatology that matches the spatial patterns of dry and wet events for the region.

4. CONCLUSIONS

In this study, the new Standardized Soil Moisture Index (SSMI) has been defined. Its definition and computation are very similar to Empirical Standardized Precipitation Index (ESPI) [21]. The SSMI is based on soil moisture content alone and is computed by fitting a non-parametric probability distribution function to historical soil moisture records and then transforming it into a normal distribution with a mean of zero and standard deviation of one. The main difference between the SSMI and the ESPI is on the boundary bias correction method used in the non-parametric Kernel density estimation.

As demonstrated in this study, the SSMI provides a more robust correspondence between soil moisture data and the respective probability values of occurrence than a parametric model. The non-parametric model is very flexible and distribution free. It has been shown that the residuals of the fitting with the non-parametric Kernel are always inferior to the residuals of the fitting performed with a parametric model. Since the values of soil moisture are bounded at (0,1) and the shape of its frequency distribution may vary at each grid-cell, and depends on the considered averaging period, by using a non-parametric method, such as the SSMI, no problem occurs with the "best" model selection since no model has to be selected.

In order to give a practical application of the SSMI, it has been used to detect the driest and wettest years that have occurred between 1979 and 2010 in the region of Bahia, northeast Brazil. Few severe and large events were shown. The SSMI indicates 1993 as the driest year in the whole time-series. It corresponds to a severe drought event that took place that year and was linked to unusual *El Niño*-Southern Oscillation (ENSO) conditions [30]. The results also show that the new blended ECV-SM product is spatially coherent and can be used to detect the geographic patterns of dry and wet conditions for the region.

The new SSMI index tested here for Latin America over the current climate conditions, promises very good ability in monitoring dryness and wetness over other particular geographical regions and, in general, over the entire world. Moreover it is expected to be very useful in the projection of future dry and wet extreme periods under climate change, as it already fits extremely well current climate conditions.

ACKNOWLEDGMENTS

This research received support from the EUROCLIMA Initiative Programme (N DCI-ALA/2009/021-126) of the Directorate General for Development and Cooperation (DG DEVCO) of the European Commission (EC).

REFERENCES

[1] Wilhite, D. & Glantz, M. H. (1985). Understanding the drought phenomenon: The role of definitions. *Water Int.*, **10**, 111–120.

[2] Heim, R. (2002). A Review of Twentieth-Century Drought Indices Used in the United States. *Bull. Am. Meteorol. Soc.*, **83**, 1149–1165.

[3] Dracup, J., Lee, K. & Paulson Jr., E. (1980). On the Definition of Droughts. *Water Resour. Res.*, **16**, 297–302.

[4] Caccamo, G., Chisholm, L. A., Bradstock, R. A. & Puotinen, M. L. (2011). Assessing the sensitivity of MODIS to monitor drought in high biomass ecosystems. *Rem. Sens. Environ.*, **115**, 2626–2639.

[5] Keyantash, J. & Dracup, J.A. (2002). The quantification of drought: An evaluation of drought indices. *Bull. Am. Meteorol. Soc.*, **83**, 1167–1180.

[6] Sheffield, J., Andreadis, K. M., Wood, E. F. & Lettenmaier, D. P. (2009). Global and Continental Drought in the Second Half of the Twentieth Century: SeverityAreaDuration Analysis and Temporal Variability of Large-Scale Events. *J. Clim.*, **22**, 1962–1981.

[7] Dutra, E., Viterbo, P. & Miranda, P. M. A. (2008). ERA-40 reanalysis hydrological applications in the characterization of regional drought. *Geophys. Res. Lett.*, **35**, L19402.

[8] Sheffield, J., Goteti, G., Wen, F. & Wood, E. F. (2004). A simulated soil moisture based drought analysis for the United States. *J. Geophys. Res.*, **109**, D24108.

[9] Dorigo, W. A., de Jeu, R. A. M., Chung, D., Parinussa, R. M., Liu, Y. Y., Wagner, W. & Fernandez-Prieto, D. (2012). Evaluating global trends (1988-2010) in harmonized multi-satellite surface soil moisture. *Geophys. Res. Lett.*, **39**, L18405.

[10] Brown, J., Wardlow, B., Tadesse, T., Hayes, M. & Reed, B. (2008). The Vegetation Drought Response Index (VegDRI): A New Integrated Approach for Monitoring Drought Stress in Vegetation. *GIScience & Remote Sensing*, **45**, 16–46.

[11] Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., McCabe, M. F. & Evans, J. P. (2011). Developing

an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrol. Earth Syst. Sci.*, **15**, 425–436.

[12] Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., de Jeu, R.A.M. , Wagner, W., McCabe, M.F., Evans, J.P. & van Dijk, A.I.J.M. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Rem. Sens. Environ.*, **123**, 280–297.

[13] Wagner, W., Dorigo, W., de Jeu, R., Fernandez, D., Benveniste, J., Haas, E. & Ertl, M. (2012). Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume I–7, XXII ISPRS Congress, Melbourne, Australia, pp315–321.

[14] Chung, D., de Jeu, R.A.M., Dorigo, W., Hahn, S., Melzer, T., Parinussa, R.M., Paulik, C., Reimer, C., Vreugdenhil, M. & Wagner, W. (2012). *ESA CCI Soil Moisture*, Algorithm Theoretical Baseline Document (ATBD), European Space Agency (ESA), 80pp.

[15] Phillips *et al.* (2009). Drought Sensitivity of the Amazon Rainforest. *Science*, **323**, 1344–1347.

[16] Zhao, M. & Running, S. W. (2010). Drought-Induced Reduction in Global Terrestrial Net Primary Production from 2000 Through 2009. *Science*, **329**, 940–943.

[17] Vargas, W. M., Naumann, G. & Minetti, J. L. (2011). Dry spells in the River Plata Basin: an approximation of the diagnosis of droughts using daily data. *Theor. Appl. Climatol.*, **104**, 159–173.

[18] Magrin, G., Garca, C., Choque, D., Gimnez, J. C., Moreno, A. R., Nagy, G. J., Nobre, C. & Villamizar, A. (2007). Latin America. Climate Change 2007: Impacts, Adaptation and Vulnerability. In *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Eds. M. L. Parry, O. F. Canziani, J. P. Palutikof, P. J. van der Linden & C. E. Hanson), Cambridge University Press, Cambridge, UK, pp581–615.

[19] Intergovernmental Panel on Climate Change (IPCC) (2012). Summary for Policymakers. In *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (Eds. C.B. Field, V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor & P.M. Midgley), Cambridge University Press, Cambridge, UK, pp1–19.

[20] Trenberth, K.E. & Stepaniak, D.P. (2001). Indices of El Niño evolution. *J. Climate*, **14**, 1697–1701.

[21] Russo, S. Carrão, H., Singleton, A., Barbosa, P., Dosio, A. & Vogt, J. (2013). European Dryness and Wetness in the last decade with the Empirical Standardized Precipitation Index. *Water Resour. Manag.*, **manuscript in submission**.

[22] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London, UK, 176pp.

[23] Schuster, E. F. (1985). Incorporating support constraints into nonparametric estimators of densities. *Comm. Stat. Theor. Meth.*, **14**, 1123–1136.

[24] Marron, J. S. & Ruppert, D. (1994). Transformations to reduce boundary bias in kernel density estimation. *J. Roy. Stat. Soc.*, **56**, 653–671.

[25] Chen, S. X. (2000). Probability Density Function Estimation using Gamma Kernels. *Ann. Inst. Statist. Math.*, **52**, 471–480.

[26] Bouezmarni, T., El Ghouch, A. & Mesfioui, M. (2011). Gamma Kernel Estimators for Density and Hazard Rate of Right-Censored Data. *J. of Prob. and Stat.*, **2011**, Article ID 937574.

[27] Jones, M. C. (1993). Simple boundary correction for kernel density estimation. *Stat. Comput.*, **3**, 135–146.

[28] McKee, T. B., Doeskin, N. J. & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *8th Conf. on Applied Climatology*, Am. Meteorol. Soc., Anaheim, Canada, pp179–184.

[29] World Meteorological Organization (WMO) (2011). Guide to Climatological Practices, Third edition (WMO No. 100), Geneva.

[30] Rao, V. B., Hada, K. & Herdies, D. L. (1995). On the severe drought of 1993 in north-east Brazil. *Int. J. Climatol.*, **15**, 697–704.

[31] Centro de Previsão de Tempo e Estudos Climáticos, Instituto Nacional de Pesquisas Espaciais (CPTEC-INPE) (2013). Evolution of ENSO regimes during 1970-2012. Online at http://www.cptec.inpe.br/ (as of 6 September 2013).

[32] Hannaford, J., Lloyd-Hughes, B., Keef, C., Parry, S. & Prudhomme, C. (2011). Examining the large-scale spatial coherence of European drought using regional indicators of precipitation and streamflow deficit. *Hydrol. Process.*, **25**, 1146–1162.

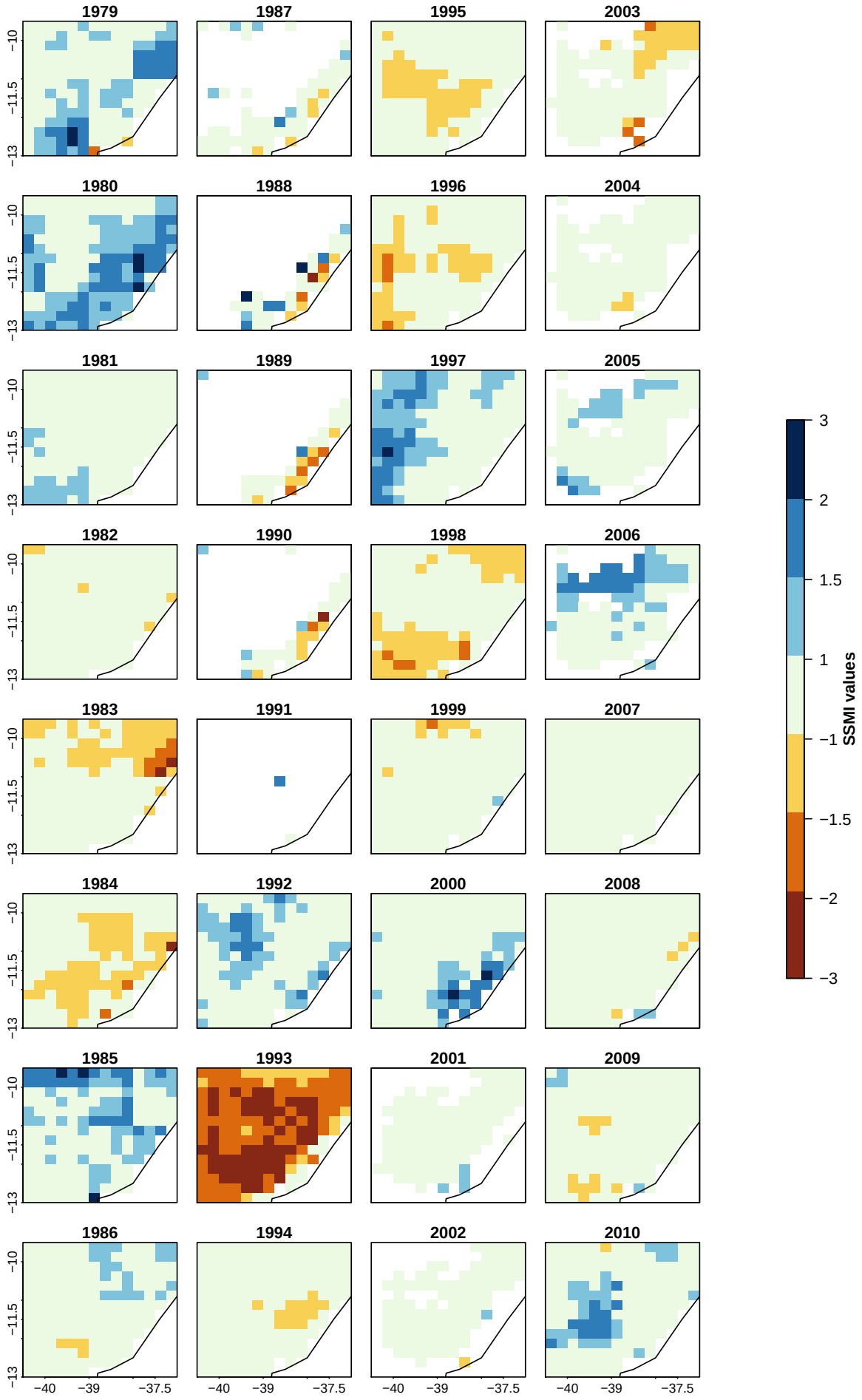

Figure 5. Time-series of yearly SSMI values computed for a region located in Bahia (northeast Brasil). No-data values are masked out in white.