Statistical analysis of precipitation time series in Dobrudja region

CARMEN MAFTEI and ALINA BARBULESCU

Ovidius University of Constanta, 124, Mamaia Bd., 900527, Constanta, Romania

(Received 09 April 2010, Modified 28 May 2012)
e-mails: cmaftei@univ-ovidius.ro; alinadumitriu@yahoo.com

ABSTRACT. Temporal characteristics of precipitation evolution in Dobrudja, a region situated in the Southeastern part of Romania, are analyzed in this article, using a data base of ten monthly series, collected in the period January 1965-December 2005. This paper describes different methods to detect the break points existence in order to detect changes in evolution of the monthly precipitation series. The study indicates a constant trend of precipitation before 2000 and an increasing one after 2000, in concordance with the predictions for this region.

Key words – Precipitation, Time series, Break tests, Statistical analysis, CUSUM.

1. Introduction

After the assessment of the IPCC-Intergovernmental Panel on Climate Change 2001 (IPCC, 2001), which asserted that the climate has been affected by the greenhouse gas emission, that provoked a climate warming, the conclusions of the IPPC 2007 Report (IPCC, 2007) came to confirm the climate change. Analyzing the period between 1816 and 2000 the IPPC 2001 Report emphasized that over the 20th century the global average surface temperature increased of about 0.6 °C and the precipitation augmented between 0.5 and 1% per decade over most mid and high latitudes of the Northern Hemisphere continents. The biggest part of Europe is affected by the augmentation of average precipitation per wet day (Frich et al., 2002), (Klein Tank et al., 2002), excepting the Mediterranean area, where a negative trend in precipitation was observed from 1950 (Piervitali et al., 1998). Significant increment of precipitation quantity has been registered in Eastern North America and South America, Northern Europe and Northern and central Asia. Based on these conclusions, and taking account of the enormous impact of the climate change on water availability and accessibility, its importance for economic and social activities, a wide interest for the research in this area is remarked. It has been proved that the climate warming could have serious impact on the precipitation occurrence, frequency and intensity (Huntingford et al., 2003), (May et al., 2002), (Zwiers and Kharin, 1998) (Planton et al., 2005). For example, decreasing precipitation rates have been registered in Western part of Africa in the period 1970-1990 (Bricquet et al., 1997), (Mahe & Olivry, 1995), (Gong and Wang, 2000); the frequency of extreme rainfall occurrence in many parts of Asia, causing severe floods, increased, while the number of rainy days and total annual amount of precipitation decreased (Zhai et al., 1999), (Khan et al., 2000) (Shrestha et al., 2000), (Mirza et al., 2002) (Lal, 2003), (Min et al., 2003), (Zhai and Pan, 2003), (Gruza and Rankova, 2004), (Zhai, 2004).

Studies concerning the climate evolution in Romania (Busuioc et al., 1995) indicate a diminution of annual precipitation, excepting some regions situated in the Northeastern part and Southeastern parts of the territory (Fig. 1). The study of Busuioc (Busuioc et al., 1995) reveals a precipitation augmentation between 0 and 25 mm in the Southern part of Dobrudja and a narrowing between 150 and 200 mm in its northeastern part.
In November 2009, Romanian Meteorological National Agency presented a study concerning the evolution of precipitation regime in Romania, showing that during the period 1901-2008 a decrease of annual precipitation trend was observed, even if an augmentation of 6.7% with respect to the multiannual average was noticed for 2001-2008. The distribution of precipitation is not uniform, a significant positive trend on large areas being noticed in autumn, while in winter and spring, negative trend were detected, in the biggest part of the country (Fig. 2).

In this context, in this paper we present the results of the analysis of temporal characteristic and spatial variability of monthly average precipitation series - official data from the archives of the National Agency of Meteorology - collected between January 1965 and December 2005 at ten main meteorological stations in Dobrudja region (Fig. 3). Break point detection and the anomaly study have been performed, in order to determine the significant change in the precipitation regime in this region.
2. Study area

Dobrudja region is located between the lower Danube River and the Black Sea including the Danube Delta and Romanian Littoral (Fig. 3), in a temperate continental climatic area, influenced by the Black Sea. Its average altitude is about 200 - 300 m. The average annual temperatures range from 11 °C inland and along the Danube to 11.8 °C on the coast and less than 10 °C in the higher parts of the plateau. Dobrudja is a windy region once known for its windmills. About 85 - 90% of all days experience some kind of wind, which usually comes from the north or northeast. Due to the limited precipitation and the proximity to the sea, rivers in Dobruja are usually short and with low discharge (Wikipedia).

3. Data and methodology

The steps followed in our study, for each station, were the estimation of the mean monthly precipitation, the estimation of the multi-monthly mean precipitation, the determination of the anomaly chart, and the discontinuities detection.

Extensive literature exists on change point problems. The break tests allow the detection of changes in a time series mean or of the probability distribution function of a process variable at a certain moment (Lubes et al., 1994). Tests for a single break point in mean were proposed by different authors (Sen and Srivastava, 1975), (Pettitt, 1979), (Buishand, 1984), (James et al., 1987). Methods for multiple change points detection (Hubert, 2000) (Olshen et al., 2004) use the Scheffé test or the Schwarz’ Bayesian information criteria for segmentation.

In this article the Pettitt, Buishand, Lee and Heghinian (Lee & Heghinian, 1977) tests, as well as the segmentation procedure of Hubert and the CUSUM method have been used for the break points detection in the study series.

The Pettitt test (Pettitt, 1979) is a non-parametric one, performed for testing the null hypothesis that there is no break in a series, against the alternative that a stepwise shift in the mean is present. It can be applied in the hypothesis that the data is independent, identically distributed. The test statistic is related to the Mann-Whitney statistic. The Pettit test is more sensitive to breaks in the middle of a time series.

The Buishand and Lee & Heghinian tests are Bayesian procedures applied under the assumption that the studied series is Gaussian, independent and identically distributed. They treat the timing of change as uncertain and the location of a change point is a parameter to be estimated. The Lee & Heghinian method determines the a posteriori probability distribution function of the parameters, considering their a priori distributions and
supposing that the break time follows a uniform distribution. This approach allows evidence for a change before a hypothesized date to count against the hypothesis (Western and Kleykamp, 2004).

The segmentation procedure of Hubert yields an optimal partition (from a least square point of view) of the original series into as many subseries as possible, all differences between two contiguous means remaining simultaneously significant. This last requirement is ensured using the Scheffé test of contrast (Hubert, 2000).

The Cumulated Sum Analysis (Taylor, 2002) aims at detecting any change in the mean of a process, assuming that the series is identically distributed and independent (or there is no strong autocorrelation). It characterizes better than other methods the changes detected by providing associated confidence levels and confidence intervals for the times of changes. If \( x_i, i = 1, n \); represent the registered data, the CUSUM charts are constructed following the steps:

(i) Determining the series average, \( \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \)

(ii) Computing the cumulative sums, \( S_i \), by

\[
S_0 = 0, \quad S_i = S_{i-1} + (x_i - \bar{x}), \quad i = 1, n;
\]

(iii) Plotting CUSUM series: the point farthest from 0 denotes a change point.

(iv) Break into two sections at change points and analyze each subseries for additional significant change points. Bootstrapping provides a measure of change point significance (Kass - Hout).

The distribution-free CUSUM test is a non-parametric rank-based method, which tests whether the means in two parts of a record are different for an unknown time of change. In particular, successive observations are compared with the median of the series in order to detect a change in the mean of a time series after a number of observations. The test statistic is the cumulative sum (CUSUM) of the \( k \) signs of the difference from the median starting from the beginning of the series.

In order to facilitate the trend computation in the precipitation patterns in Israel, Steinberger and Gazit-Yaari (Steinberger and Gazit-Yaari, 1996) introduced the
concept of “normalized variable”, calculated by dividing the actual precipitation by the average rainfall at one station and multiplying by 100%. The same index (named “Percent of Normal”) is used by Willeke (Willeke et al., 1994) in the drought analysis. A disadvantage of the use of this index is that the mean precipitation is often not the same as the median, which is the value exceeded by 50% of the precipitation occurrences in a long-term climate record (Steinemann et al., 2005). We use this concept, with one modification: the mean monthly precipitation is normalized by the regional multi-annual mean monthly precipitation (considered to be calculated by Thiessen method - Fig. 4.) (Maftei and Barbulescu, 2008). This index will be called “Normal Precipitation”. We remember that in Thiessen method the precipitation registered at a meteorological station is proportional with the surface of the polygonal area associated to the station.

4. Result and discussions

4.1. The analysis of precipitation variation

The chart of the studied monthly precipitation series is presented in Fig. 5 & Fig. 6 represents the spatial evolution of the multi-monthly mean precipitation (calculated as the average of the monthly precipitation registered in the period January 1965 - December 2005) in the region of Dobrudja. The isohyets are automatically generated in GIS ArcView®, by splines interpolation on the base of multi-monthly mean precipitation calculated at each station.

The multi-monthly mean precipitation varied between 22 and 42 mm approximately, the highest values being registered in the Northern part and the Southwestern part of the territory and augmented with the altitude. The lowest precipitation was registered at Sulina, and the largest at Tulcea, Adamclisi and Cernavoda.

Table 1 presents the descriptive statistics of the study series. Accordingly, the lowest mean precipitation occurred at five meteorological stations (0 mm: August 1986 at Adamclisi; April 1968 at Cernavoda; May 1968 at Jurilovca; October 1969, August 1989 and March 1990 at Mangalia; June 1968, August 1975, August 1986, 1989, 2001, July 1995 at Sulina) and the highest amount of precipitation has been registered in September 2005, at Mangalia. Overall, the smallest mean monthly precipitation has been registered at Sulina, in the Danube Delta, and the highest on the Littoral (Constanta and Mangalia). All the coefficients of variation are greater than 0.77, proving a high dispersion of multi-monthly average precipitation.

After performing the variance analysis, we found enough evidence to reject the hypothesis that there is no difference between the multi-monthly means of the precipitation series. To emphasize the station whose multi-monthly mean is statistically different from the others, the Tukey HSD and the Scheffé tests have been used at the level of significance $\alpha = 1\%$. It was found that the multi-monthly mean of Sulina series is not equal to the other series means, point of view of statistics (Barbulescu and Deguenon, 2012).

The precipitation anomaly charts have also been determined (Fig. 7).

Fig. 7 shows the difference between the mean annual precipitation and the average precipitation during a regional base period. The positive anomaly shows the years when the precipitation exceeded the 1965-2005 baseline average and the negative anomaly shows the years when the mean precipitation was less than the baseline average. We remark three different cases: (i) the
anomaly is mostly positive in the territory near the Danube river (Adamclisi, Cernavoda, Medgidia, Harsova, Tulcea); (ii) the anomaly is mostly negative in the Northeastern part (Sulina and Jurilovca); (iii) the positive anomaly and the negative ones are approximately equal (Corugea, Mangalia and Constanta). We also remark a trend of positive anomaly for all the series but Sulina, starting from 2000.

4.2. Break points analysis

First, the normality, the autocorrelation and homoscedasticity hypotheses were tested, respectively by the Lilliefors test, the autocorrelation function and the Levene test (Bărbulescu, 2010). Since some series were not normally distributed, they were normalized by Box-Cox transformations, in order to allow the application of certain break tests. The results of the break tests are given in Table 2, where no signifies that the null hypothesis is accepted. The Pettitt and Buishand tests (at the level of significance of 5%) gave the same result, excepting for Constanta series, for which the normality couldn’t be reached by transformations, so the last test couldn’t be performed. The Lee & Heghinian test found April 2004 as a break point for six series, August 2005 for one, January 1965, for another one. The results of segmentation procedure of Hubert are in concordance with those of Lee & Heghinian test for Cernavoda, Harsova and Mangalia (Table 2).

The CUSUM charts were also studied and the results were compared to those obtained by the break tests. From...
the definition of the sums in the CUSUM procedure, it is obvious that each time the measurements are below the overall average, the CUSUM decreases and each time the values are above the overall average, the CUSUM increases. A slope change in the CUSUM graph indicates a sudden change in the average (Taylor, 2002).

Since the result of a CUSUM procedure is influenced by the outliers’ presence, the non-parametric CUSUM analysis (Hackl and Maderbacher, 1999) has been performed in the cases of the aberrant values existence, as for Adamclisi series. In Figs. 8 & 9, we present the boxplot for Adamclisi and Tulcea series, respectively the associated CUSUM charts, for comparison. The outliers of
Fig. 8. Boxplots for Adamclisi and Tulcea monthly series

Figs. 9 (a&b). CUSUM: a. Adamclisi, b. Tulcea

TABLE 3
The results of CUSUM charts

<table>
<thead>
<tr>
<th>Location</th>
<th>Date</th>
<th>Confidence interval</th>
<th>conf level</th>
<th>from</th>
<th>to</th>
<th>level</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adamclisi</td>
<td>Jun-94</td>
<td>(218, 381)</td>
<td>95%</td>
<td>37.62</td>
<td>47.385</td>
<td>3</td>
<td>Rank</td>
</tr>
<tr>
<td>Cernavoda</td>
<td>Apr-81</td>
<td>(132, 223)</td>
<td>98%</td>
<td>43.803</td>
<td>32.511</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jun-94</td>
<td>(298, 396)</td>
<td>93%</td>
<td>32.511</td>
<td>41.609</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>May-04</td>
<td>(414, 477)</td>
<td>92%</td>
<td>41.609</td>
<td>68.085</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Medgidia</td>
<td>Jun-02</td>
<td>(191, 457)</td>
<td>95%</td>
<td>35.999</td>
<td>53.1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Harsova</td>
<td>Sep-82</td>
<td>(141, 245)</td>
<td>93%</td>
<td>37.046</td>
<td>27.06</td>
<td>9</td>
<td>Rank</td>
</tr>
<tr>
<td></td>
<td>Jun-95</td>
<td>(298, 376)</td>
<td>98%</td>
<td>27.06</td>
<td>40.056</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jan-00</td>
<td>(402, 431)</td>
<td>93%</td>
<td>40.056</td>
<td>22.441</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sep-03</td>
<td>(458, 468)</td>
<td>100%</td>
<td>32.74</td>
<td>54.407</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Corugea</td>
<td>Nov-72</td>
<td>(66, 202)</td>
<td>92%</td>
<td>41.955</td>
<td>31.332</td>
<td>2</td>
<td>Rank</td>
</tr>
<tr>
<td></td>
<td>Sep-96</td>
<td>(304, 406)</td>
<td>98%</td>
<td>31.332</td>
<td>43.898</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Tulcea</td>
<td></td>
<td>No break</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulina</td>
<td>Sep-82</td>
<td>(141, 276)</td>
<td>92%</td>
<td>26.126</td>
<td>19.572</td>
<td>1</td>
<td>Rank</td>
</tr>
<tr>
<td></td>
<td>Feb-00</td>
<td>(315, 428)</td>
<td>95%</td>
<td>19.572</td>
<td>12.296</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>May-04</td>
<td>(455, 477)</td>
<td>90%</td>
<td>12.296</td>
<td>23.53</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Jurilovea</td>
<td>Mar-95</td>
<td>(137, 369)</td>
<td>98%</td>
<td>30.447</td>
<td>47.736</td>
<td>2</td>
<td>Rank</td>
</tr>
<tr>
<td></td>
<td>Dec-98</td>
<td>(397, 440)</td>
<td>100%</td>
<td>47.736</td>
<td>27.578</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Constanta</td>
<td></td>
<td>No break</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rank</td>
</tr>
<tr>
<td>Mangalia</td>
<td></td>
<td>No break</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rank</td>
</tr>
</tbody>
</table>
Figs. 10(a-c). Evolution of the normal precipitation index: (a) Hârșova, (b) Jurilovca and (c) Sulina
Adamclisi series are depicted by small balls, the numbers near them denoting the month number, zero being associated to January 1965 and 492 to December 2005. In Figs. 8 & 9 we remark the existence of a change point for Adamclisi, respectively the change point’s absence for Tulcea.

The change points detected for all the series are presented in Table 3, together with the confidence levels, the average of the neighbour segments, the level of change point and the method used.

Since the results of the tests and of CUSUM procedure were not in concordance, a homogeneity test (Wilcoxon) was also performed (Table 4).

We remark that the Wilcoxon test and CUSUM procedure gives the same results concerning the break point year, excepting for Constanta series.

4.3. Analysis of the Normal Precipitation index

Analyzing the results of “Normal Precipitation” index we observe that its pattern is different (i) in the central and western zone, the index was slightly decreasing in the period 1965-1992, constant between 1992 and 1997 and increasing after 1997 [Fig. 10(a)]; (ii) on the Danube Delta and the Lagunar Complex, its trend was slightly increasing in the period 1965 - 1972, then constant [Fig. 10(b) - Jurilovca station]; (iii) on the
Therefore, a more complete image of the precipitation after extracting the seasonality from the series. This different seasons (winter, spring, autumn and summer), or evolution could be given performing such an analysis on approach will be done in a future article.

The data were provided by A.N.M. (National Meteorological Agency). This article was supported by CNCSIS - UEFISCSU by Grant PN II IDEI 262/2007.

References


Kass-Hout, T. https://sites.google.com/site/changepointanalysis/


Acknowledgement

The methods used here have also limitations, since the seasonality existence has not been considered. Therefore, a more complete image of the precipitation evolution could be given performing such an analysis on different seasons (winter, spring, autumn and summer), or after extracting the seasonality from the series. This approach will be done in a future article.

5. Conclusions

Statistical tools were used to study the precipitation variation in a region of Romania. The isohyets analysis emphasized the precipitation increase from the boundary of the region to its interior and from the territory with low altitude, to that with higher one. Since the break point and CUSUM results were not concordant, the Normal index, introduced by us, has been calculated and the Mann Kendall tests have been performed in order to detect the trend existence in the annual mean precipitation. The results are concordant, the single precipitation series that registered a decreasing trend being Sulina, which has a particular position in the Danube Delta, 13 km offshore.

The Littoral (without the Danube Delta and the Lagunar complex), the was slightly decreasing in the study period; (iv) for Sulina the trend was decreasing.

Since the Normal index values are very small in all cases, but Sulina, the Mann - Kendall test and Sen’s slope estimate (www.emep.int/) for the annual average precipitation have been performed. The single series for which the slope was significant is Sulina (Fig. 11).

5. Conclusions

Statistical tools were used to study the precipitation variation in a region of Romania. The isohyets analysis emphasized the precipitation increase from the boundary of the region to its interior and from the territory with low altitude, to that with higher one. Since the break point and CUSUM results were not concordant, the Normal index, introduced by us, has been calculated and the Mann Kendall tests have been performed in order to detect the trend existence in the annual mean precipitation. The results are concordant, the single precipitation series that registered a decreasing trend being Sulina, which has a particular position in the Danube Delta, 13 km offshore.

The methods used here have also limitations, since the seasonality existence has not been considered. Therefore, a more complete image of the precipitation evolution could be given performing such an analysis on different seasons (winter, spring, autumn and summer), or after extracting the seasonality from the series. This approach will be done in a future article.

Acknowledgement

The data were provided by A.N.M. (National Meteorological Agency). This article was supported by CNCSIS - UEFISCSU by Grant PN II IDEI 262/2007.


