Change Points Detection and Variability Analysis of some Precipitation Series

CARMEN MAFTEI[1], ALINA BARBULESCU[2], CONSTANTIN BUTA[1],
CRISTINA SERBAN (GHERGHINA)[1]
[1] Faculty of Civil Engineering, [2]Faculty of Mathematics and Computers Science
Ovidius University of Constanta
124 Mamaia Blvd.
ROMANIA
cmaftei@univ-ovidius.ro, alinadumitriu@yahoo.com
costi_buta@yahoo.com, serban.cristina@univ-ovidius.ro

Abstract: - The temporal characteristics of the evolution of precipitation are investigated herein, in order to provide a framework for sustainable resource management in the region of Dobrudja, Romania. The evolution of precipitation data from 10 meteorological stations is examined using statistical methods. First, we study the temporal characteristics of precipitation evolution, and subsequently we perform several break tests, in order to identify discontinuities in the time series. The analysis indicates an increasing trend of the mean annual precipitation. After 1995, its increment is of 82.8mm, a value in agreement with existing estimations for the whole of Europe.

Key-Words: - Precipitation, time series, break test, statistical analysis

1 Introduction

Since climatic change has an enormous impact on economic and social human activities, there is an increasing interest for research in the area. Greenhouse gas-induced climate warming could affect precipitation severely, in different ways, including changes in frequency, intensity, and timing of occurrence [12], [16], [21], [35]. A 2001 assessment of the IPCC-Intergovernmental Panel on Climate Change [14] concludes that as effect of increasing of greenhouse gas concentrations "…over many Northern Hemisphere mid- to high-latitude land areas, more intense precipitation events are very likely". Annual zonally average precipitation increased by 7% to 12% for the zones 30^0N to 85^0N in emerged landmasses in the Northern hemisphere, except in the Far East [14]. At the planetary scale, the models consistently simulate an intensification of the hydrological cycle in a future climate, warmer than the present-day one [28]. This simulation highlights the annual impact of climatic change on precipitation in Europe [7]. In winter, the precipitation increasing is more important in Northern Europe [28].

A downward pluvial trend was observed in West Africa, starting at the end of 1960s or the beginning of the 1970s, until the beginning of the 1990s [3], [10], [19], [28]. Gong and Wang [9] and Qian and Zhu [33] indicated a significant negative precipitation trend in eastern China for 1954 – 1976 and a subsequent positive trend for 1977 – 1998.

Fig. 1. is based on the scenario of the Special Report on Emission Scenarios (SRES) of the IPCC [13], [14], [15]. The effects of the climate change in Europe were estimated for the period 2071 – 2100 in comparison with the period 1961 – 1990.

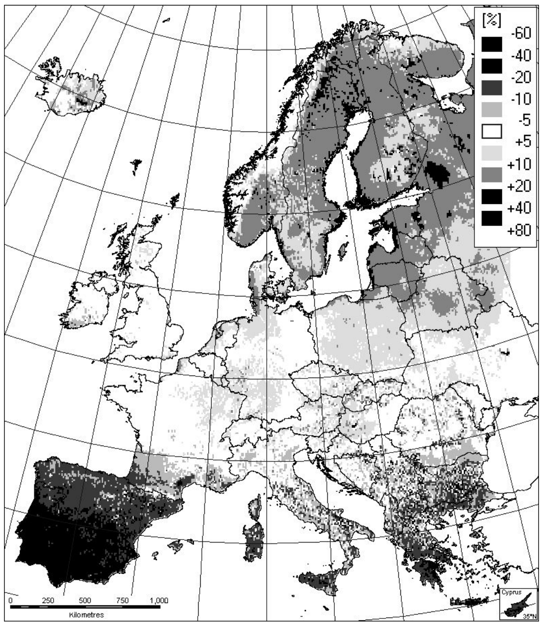


Fig. 1 Estimated evolution of annual mean precipitation for the period 2071-2100
(PRUDENCE Project; reference interval: 1961-1990)

The map is based on the data of DMI/PRUDENCE [27], processed by the Joint Research Centre (JRC) during the PESETA study [24]. The figure depicts a considerable increase of the quantities of precipitation in northern Europe (+20 - +80 % of the total annual quantity of precipitation), while in the south of the continent, the periods of drought are increasingly frequent (-20-60% of the total quantity of precipitation).

From this figure, we observe a decrease between 5 and 20% of the total annual quantity of precipitation, in Romania, excepting its northern the north-eastern part, where an increase between 5 and 10% of the total annual quantity of precipitation is predicted. In the Dobrudja region, we observe a variation between -5 and +5% of the total precipitation.

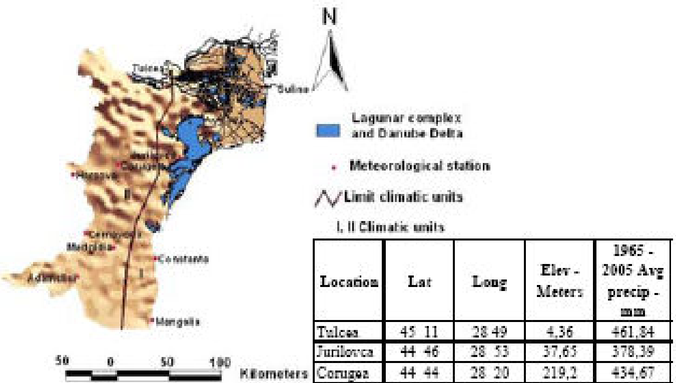
Such a situation cannot be ignored, because its effects could be tragic. There is a risk to jeopardize the projects of the agricultural development program, and to perturb the efficiency of the structures built in previous years (such as irrigation works that sometimes stopped working in summer, as a result of the dramatic decrease of the Danube water level). In this context, this paper attempts to identify discontinuities in precipitation time series, using statistical analysis, in order to determine whether or not there has been a significant change in the precipitation regime in Dobrudja.

This paper describes the methodology used for the detection of break points in the precipitation series and presents the results found.

2 Problem Formulation

2.1 Study Area

Dobrudja (in Romanian: Dobrogea) is a region situated in the South – East of Romania, between the Black Sea and the lower Danube River (Fig. 2).



Location	Lat	Long	Elev - Meters	1965 - 2005 Avg precip - mm
Tulcea	45 11	28 49	4,36	461,84
Jurilovca	44 46	28 53	37,65	378,39
Corugea	44 44	28 20	219,2	434,67
Harsova	44 41	27 57	37,51	408,82
Cernavoda	44 21	28 03	87,17	487,60
Medgidia	44 15	28 16	69,54	449,92
Constanta	44 13	28 38	12,8	423,04
Adamclisi	44 08	28 00	158	484,54
Mangalia	43 49	28 35	6	427,74
Sulina	45 09	29 39	2,08	261,63

Fig. 2 Dobrudja region and its ten meteorological stations

Dobrudja (without the Danube Delta) is a plateau with hilly aspect. In the North, remnants of Hercinic and Caledonian mountains are present, with an altitude up to some 400m (Greci Peak 467m). The altitude decreases towards the South, where the average altitude is between 100 and 180m [26]. Generally, Dobrudja's climate is temperate-continental and comprising two units (Fig. 3): (I), which contains the Danube Delta, the lagoons (Razim Lake and Sinoe Lake) and the Eastern part of the region and (II), which contains the rest of territory, where the climate is influenced by the moderate continental belt. The air temperature average is slightly over 11^0C towards the littoral area and the Danube floodplain, and no more than 10 - 11^0C in its Northern and central part [20].

2.2. Methods and Methodology

The ten time series used in this study, obtained from the archives of the National Agency of Meteorology, contain the annual average precipitation spanning the period 1965-2005 for Dobrudja region. The stations names, their locations, elevations and multi-annual means of precipitation are presented in Fig. 2. We observe that 5 of the 10 meteorological stations are situated in the first climatic unit and the other 5 in the second climatic unit. We have to point out that the Sulina station is situated 13 km offshore, on the Sulina horn dam.

The methodology presented in this paper, and applied to Dobrudja region, consists of an analysis of the temporal characteristics, and the identification of the discontinuities in precipitation time series. The steps that were followed are: estimation of the annual mean precipitation for each station, estimation of the multi-annual mean precipitation for each station, determination of the anomaly chart, and detection of the discontinuities in data series.

In order to determine the discontinuities in the precipitation regime over the period 1965–2005, some homogeneity and break tests are performed.

We define a break as a change of the probability distribution function of a process variable at a certain moment [18]. The break tests allow the detection of a change in a time series mean.

The methods used to detect a break are: Pettitt test [23], Buishand test [4], [5], Lee and Heghinian test [17] and the segmentation procedure of Hubert [10], [11]. The null hypothesis that is tested is that there is no break in the precipitation series $(X_i)_{i \in \overline{1,N}}$.

The Pettitt test is a non - parametric one. To perform it, the steps are [2]: (i) the studied series is divided into two sub-samples of sizes *m* and *n* respectively; (ii) the values of the two samples are grouped and arranged by increasing order; (iii) the sum of the ranks of the components of each sub-sample in the total sample is then calculated; (iv) a statistic, U_t, is defined using the two sums thus obtained in order to assess whether the two samples belong to the same population.

Let us define:

$$\operatorname{sgn}(x) = \begin{cases} 1\,, & x > 0 \\ 0\,, & x = 0 \\ -1, & x < 0 \end{cases} \tag{1}$$

$$D_{i,j} = \text{sgn}(x_i - x_j) \tag{2}$$

$$U_{t,N} = \sum_{i=1}^{t} \sum_{j=t+1}^{N} D_{i,j} \, , \tag{3}$$

$$K_N = \max_{t=1,N-1} \left| U_{t,N} \right| \tag{4}$$

If k is the value of K_N taken on the studied series, under the null hypothesis, then:

$$Prob(K_N > k) \approx 2\exp(-6 \cdot k^2 / (N^3 + N^2)) \tag{5}$$

If $Prob(K_N > k) < \alpha$, for a significance level α, then the null hypothesis is rejected.

The Buishand and Lee & Heghinian tests are Bayesian procedures applied under the assumption that the studied series is normally distributed. The tests are based on the following model, which supposes a change in the series mean:

$$X_i = \begin{cases} \mu + \varepsilon_i & , \quad i = 1,.....,m \\ \mu + \varepsilon_i + \delta, & i = m+1,.....N \end{cases}, \tag{6}$$

where ε_i are random variables, independent and normally distributed, with null expectance and a constant variance. The break point m and the parameters μ and δ are nknown.

The Lee & Heghinian method works in the hypothesis that $(X_i)_{i=\overline{1,N}}$ is a series of independent variables, with a constant variance. The method determines the a posteriori probability distribution function of the parameters μ and δ, considering their a priori distributions and supposing that the break time follows a uniform distribution. The range of the break time corresponds to the values of the modes of the a posteriori distributions of m and δ respectively.

Hubert's segmentation procedure detects the multiple breaks in time series. The principle is to cut the series into m segments (m>1) such that the calculated means of the neighboring sub-series significantly differ. To limit the segmentation, the means of two contiguous segments must be different to the point of satisfying Scheffe's test. The procedure gives the timing of the breaks.

CUSUM charts are constructed by calculating and plotting a cumulative sum based on the data. CUSUM charts show the cumulative sum of differences between the values and the average. Because the average is subtracted from each value, the cumulative sum also ends at zero [32].

If x_i, $i = \overline{1,n}$ represent the registered data, the cumulative sums S_i are calculated by:

$$S_0 = 0, \, S_i = S_{i-1} + (x_i - \overline{x}), \, i = \overline{1,n}, \tag{7}$$

where $\overline{x} = \sum_{i=1}^{n} x_i$ is the average.

The anomaly chart procedure calculates the overall average precipitation for the base period specified. The procedure then calculates the annual differences between each yearly mean precipitation and the base period mean precipitation. To facilitate the understanding of anomaly charting, negative anomalies are represented below the abscissa, positive anomalies above.

3 Results

3.1 Precipitation variation

Fig. 3 represents the spatial evolution of the multi-annual mean precipitation in Dobrudja. The isohyets are automatically generated in GIS ArcView®, by spline interpolation on the base of the annual mean precipitation calculated at each station.

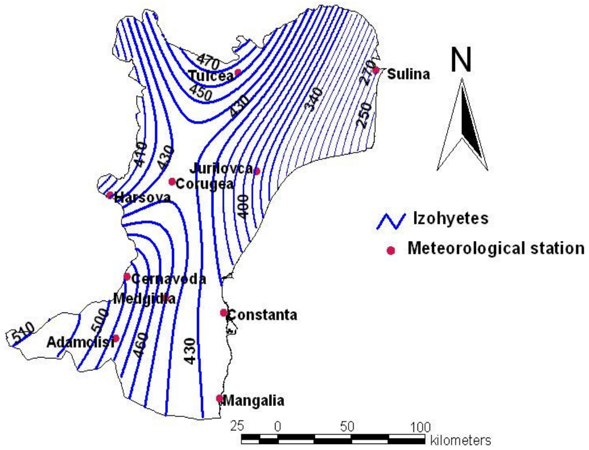


Fig. 3 Multi-annual mean precipitation

The multi-annual mean precipitation vary in large limits (260–500 mm approximately), the highest values being registered in the North and center of the region. The precipitation values increase with altitude. The lowest precipitation was registered at Sulina (262 mm at 2 m), Constanța (423 mm at 12 m) and Mangalia (427 mm at 6 m) – on the coast, respectively at Harşova (408 mm at 37.31m), on the Danube, and the largest at Tulcea (462 mm at the North).

The variation of the annual mean precipitation for each station (Fig. 4) reveals the succession of the humid and dry years during the study period. It can be observed that the same evolution pattern is present at all the stations, i.e. starting in 1995 the mean annual precipitation at each station is higher than the multi-annual mean precipitation (but 2000 and 2001, which are the driest years). This remark suggests that the mean precipitation of the sample 1965 - 1995 and that of the sample 1996 - 2005 are different. In order to demonstrate, the values of the annual mean precipitation and the multi-annual mean precipitation are represented for the Mangalia station (Fig. 5).

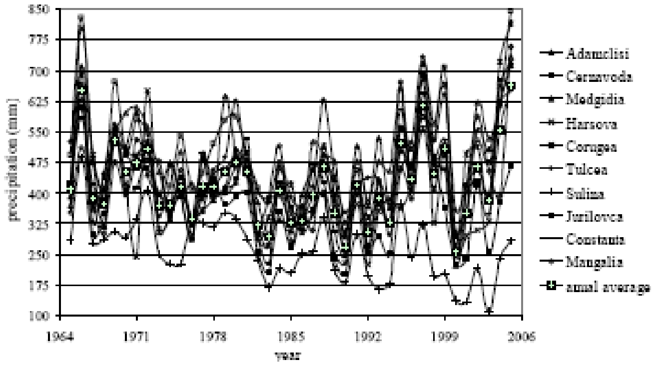


Fig. 4 The mean multi-annual precipitation at the study stations during the period 1965-2005

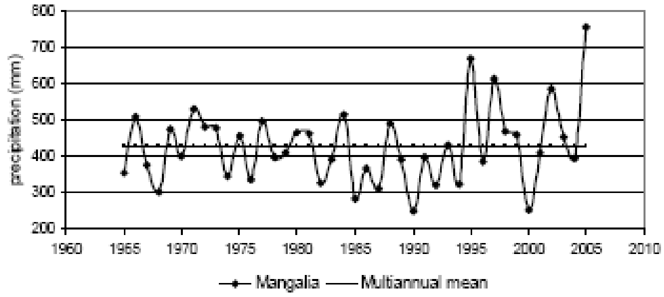


Fig. 5 The evolution of the mean precipitation at Mangalia station

The precipitation anomaly chart (Fig. 6) shows the difference between the annual precipitation and the average precipitation during a base period. The positive anomaly shows those years when the annual mean precipitation exceeded the 1965-2005 baseline average, the negative anomaly shows those years when the mean precipitation was less than the baseline average.

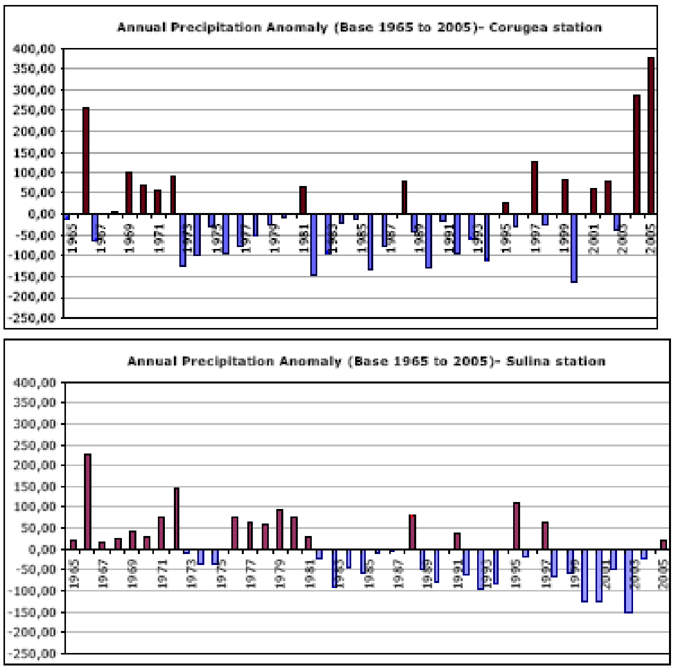


Fig. 6 Example of annual precipitation anomaly chart

The anomaly chart (Fig. 6) analysis reveals: (i) a negative anomaly, for the period 1981 – 1995 (approx.), excepting Sulina series, where the period with negative anomaly started in 1982; (ii) a positive anomaly for the period 1965 – 1973 and 1995 (1997) – 2005.

3.2 Break analysis

Firstly, the normality, correlation and homoscedasticity hypotheses were tested [1]. Since some series are not normally distributed, they were normalized by Box – Cox transformations, in order to allow the application of certain break tests. The results of these tests are: two tests (Pettitt and Buishand) give the same result, i.e. there is a break in the mean annual precipitation series only for Sulina station. The tests Lee & Heghinian test and Hubert procedure detect a break point, which appears in 2003 or 2004.

Since the results of the procedures are inconclusive, a test of homogeneity (Wilcoxon procedure) was also performed (Table 1).

Table 1 The result of the Wilcoxon test

Hypothesis: H_0 The average for the two samples are equal H_1 The average for the two samples are unequal				
Adamclisi	Cernavoda	Medgidia	Harsova	Corugea
Result -1995 $\|W\|$ =2.40 p = 0.0165 H_0 rejected at 5%, but accepted at 1%	Result -1995 $\|W\|$ = 2. p = 0.0329 H_0 *rejected* at 5%, but accepted at 1%.	Result -2000 $\|W\|$ = 1.97 p = 0.0486 H_0 *rejected* at 5 %, but accepted at 1 %	Result -1973; 2003 $\|W\|$ =1.99; 2.21 p = 0.046; 0.027 H_0 *rejected* at 5 %, but accepted at 1 %	Result -1973; 2000 $\|W\|$ =2.29; 2.25 p =0.022; 0.024 H_0 *rejected* at 5 %, but accepted at 1 %
Tulcea	Sulina	Jurilovca	Constantza	Mangalia
Result -2003 $\|W\|$ = 1.97 p = 0.0492 H_0 *rejected* at 5 %, but accepted at 1 %	Result -1981 $\|W\|$ = 3.72 p = 0.000202 H_0 *rejected* at 5 %, but accepted at 1 %	Result -1982; $\|W\|$ = 2.26 p = 0.0237 H_0 *rejected* at 5 %, but accepted at 1 %	Result -1995 $\|W\|$ = 2.43 p = 0.0152 H_0 *rejected* at 5 %, but accepted at 1 %	Result $\|W\|$ = 1.73 p = 0.0832 H_0 accepted at 5 %,

The choice of the year 2003 (or 2004) as moment of separation into two sub-samples is not conclusive, since there is not enough data. Therefore, other points have been chosen to apply this test.

Interpreting a CUSUM chart requires some practice. The CUSUM chart trend shows the way in which the individual values compare to the overall average. Each time the measurements are below the overall average, the CUSUM decreases. 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. Periods where the CUSUM chart follows a relatively straight path indicate a period where the average did not change.

Two CUSUM charts are plotted in Figs. 7 and 8, where a sudden change in direction can be seen around 1995. The results of the CUSUM charts of the series are presented in Table 2, where it can be seen that around 1995 the average shifted. The conclusion of the analysis is that the precipitation regime is changing in Dobrudja area.

ISBN: 978-1-61804-011-4

Table 2 The result for CUSUM chart for climate stations

	Adamclisi	Cernavoda	Medgidia	Corugea	Harsova	Tulcea	Sulina	Jurilovca	Constantza	Mangalia
Mean	484.54	487.6	449.92	408.82	434.67	461.84	261.63	378.39	423.04	427.74
Conf. level	93%	91%	93%	96%	95%		100%	98%	94%;	
Break pts	1995	1995	2002	1973	1973		1982	1982	1982	
Avg 1	454.5	465.3	433.8	497.5	497.2		314.6	418.8	434.8	
Avg 2	0.0	0.0	0.0	392.9	367.8		224.1	300.3	341.8	
Change pt				2002	2004			1995	1995	
Avg 1				392.9	367.8			300.3	341.8	
Avg 2				611.9	691.8			491.5	500.8	
Change pt								1999		
Avg 1								491.5		
Avg 2								353.1		

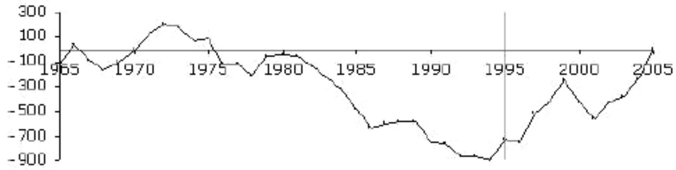


Fig. 7 CUSUM chart for Adamclisi station

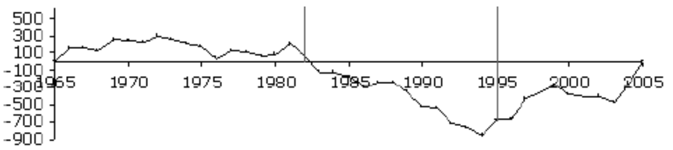


Fig. 8 CUSUM chart for Constanta station

The isohyet's maps (Figs. 9 and 10) depict a general increase of the mean multi-annual precipitation after 1995, as compared to the period 1965-1995.

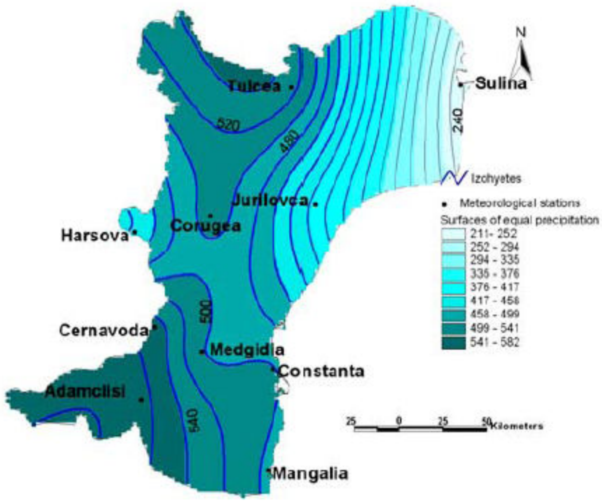


Fig. 9 The isohyets 1965 – 1995

We can conclude that, generally, starting in 1995, an important movement of the isohyets to the North-West, in the Danube floodplain and the lagoon region (where the precipitation rates decrease) and the increase of the precipitation rate over the littoral area have been remarked.

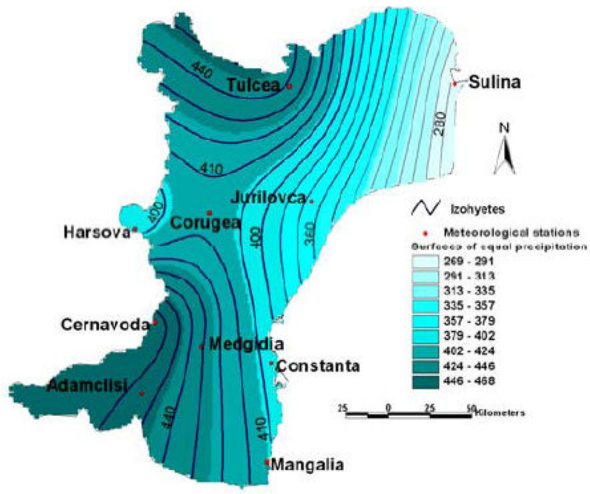


Fig. 10 The isohyets 1995 – 2005

4 Conclusion

Analyzing the results of all tests for the ten stations we observe:

(i) 1995 is the break point year,

(ii) in the climatic unit II, after 1995, the mean annual precipitation increased with 82.8 mm, a fact which is in concordance with the estimations made for Romania and Europe [6], [14], [15];

(iii) in the climatic unit I (without Jurilovca and Sulina station) the mean annual precipitation increased with 98 mm, but, unfortunately, the precipitation distribution in time is not uniform. For example, at the Constanta station, in 2004 a precipitation value of 259 mm was recorded (in August 2004), representing one third of the annual mean precipitation at this station.

Acknowledgement: This article was supported by CNCSIS – UEFISCSU Romania, under research grant PNII IDEI 262/2007

References:

[1] A. Bărbulescu, *Time series with applications*, Junimea, Iasi, 1993

[2] J.F. Boyer, Logiciel Khronostat d'analyse statistique de séries chronologique, IRD UR2, Programme 21 FRIEND AOC, Equipe Hydrologie UMRGBE, Université de Montpellier II, Ecole des Mines de Paris, 2002, http://www.hydrosciences.org/mytech/khronostat.html

[3] J.P. Bricquet, F. Bamba, G. Mahe, M. Toure, J.C. Olivry, Variabilité des ressources en eau de l'Afrique Atlantique, PHI-V, 6, 1997, pp. 83-95

[4] T. A. Buishard, Some methods for testing the homogeneity of rainfall records, *Journal of Hydrology*, 58, 1982, pp.11-27

ISBN: 978-1-61804-011-4

[5] T. A Buishard, Tests for detecting a shift in the mean of hydrological time series, *Journal of Hydrology*, 58, 1984, pp.51-69

[6] A. Busuioc, C. Boroneant, M. Baciu, A. Dumitrescu, *Observed temperature and precipitation variability in Romania,* 1995

[7] H. Douville, F. Chauvin, S. Planton, J. - F. Royer, D. Salas-Mélia, S. Tyteca, Sensitivity of the hydrological cycle to increasing amounts of greenhouse gases and aerosols, *Clim. Dyn.,* 20, 2002, pp. 45–68

[8] European Climate Support Network (ECSN), *Climate of Europe: Recent Variation, Present State and Future Prospects*, KNMI, De Bilt, 1995

[9] D.Y. Gong, S.W. Wang, Severe summer rainfall in China associated with aggravated global warming, *Climate Research*, 16, 2000, pp. 51–59

[10] P. Hubert, J. P. Carbonnel, A. Chaouche, Segmentation des séries hydrométéorologiques. Application à des séries de précipitations et de débits de l'Afrique de l'Ouest, *Journal of Hydrology*, 110, 1989, pp. 349-367

[11] P. Hubert, J. P. Carbonnel, Segmentation des séries annuelles de débits de grands fleuves africains, *Bulletin de liaison du CIEH*, 92, 1993, pp. 3-10

[12] C. Huntingford, R.G. Jones, C. Prudhomme, R. Lamb, J..H.C. Gash, D.A. Jones, Regional climate—model predictions of extreme rainfall for a changing climate, *Q. J. R. Meteorol. Soc.*, 129, 2003, pp. 1607–1621

[13] IPCC,. Climate changes 1995: Contribution of Working Group I to the Second Report of the Intergovernmental Panel on Climate Change, 1995, pp. 141–193

[14] IPCC, *Climate Change 2001: The Scientific Basis.* Cambridge University Press, Cambridge, 2001

[15] IPCC, *Climate Change 2007: The Scientific Basis,* Cambridge University Press, Cambridge, 2007

[16] V. Kharin, F. Zwiers, Changes in the extremes in the ensemble of transient climate simulations with a coupled atmosphere–ocean, *GCM J. Climate*, 13, 2000, pp. 3760– 3788

[17] A. F. S. Lee, M.S. Heghinian, A Shift of the Mean Level in a Sequence of Independent Normal Random Variables - A Bayesian Approach, *Technometrics*, 19(4), 1977, pp. 503-506

[18] H. Lubes, M. Masson, E. Servat, J.-E. Paturel, B. Kouame, J.F. Boyer, Caractérisation de fluctuations dans une série chronologique par applications de tests statistique, Rapport no 3, Programme ICCARE, 1994, 21 pp.

[19] G. Mahe, J.C. Olivry, Variations des précipitations et des écoulements en Afrique de l'ouest et centrale de 1951 à 1989, *Sécheresse*, 6 (1), 1995, pp. 109 - 117

[20] C. Maftei, A. Bărbulescu, Statistical analysis of climate evolution in Dobrudja region, *Lecture Notes in Engineering and Computer Sciences*, WCE 2008, vol. II, IAENG, 2008, pp. 1082 -1087

[21] W. May, R. Voss, E. Roeckner, Changes in the mean and extremes of the hydrological cycle in Europe under enhanced greenhouse gas conditions in a global time-slice experiment, *Advances in Global Change Research*, 10, 2002, pp. 1 – 29

[22] C.R. Păltineanu, I.F.L. Mihăilescu, I. Seceleanu, *Dobogea. Conditii climatice, consumul si necesarul apei de irigatie pentru principalele culturi agricole*, ExPonto, Constantza, 2000

[23] A. N. Pettitt, A non - parametric approach to the change-point problem, *Applied Statistics*, 28(2), 1979, pp. 126 – 135

[24] PESETA project, http://peseta.jrc.es

[25] S. Planton, M. Déqué, H. Douville, B. Spagnoli, Impact du réchauffement climatique sur le cycle hydrologique, *Geoscience*, 337, 2005, pp.193–202

[26] G. Posea, O. Bogdan, I. Zăvoianu, Geografia Romaniei, vol. 5, Ed. Academiei Române, 2005, pp. 47 - 65

[27] PRUDENCE project, http://prudence.dmi.dk

[28] E. Servat, J. - E. Paturel, H. Lubes–Niel, B. Kouame, J.M. Masson, M.Travaglio, B. Marieu, De différents aspects de la variabilité de la pluviométrie en Afrique de l'ouest et centrale non sahélienne, *Revue des sciences de l'eau*, 12(2), 1999, pp. 363 - 387

[29] A. Steinemann, M. Hayes, L. Cavalcanti, Drought Indicators and Triggers, 2005; available: www.water.washintogn.edu/research/

[30] E.H. Steinberger, N. Gazit-Yaari, Recent changes in the spatial distribution of annual precipitation in Israel, *J. Climate*, 9, 1996, pp.3328–3336

[31] N. Topor, *Ani ploiosi si secetosi în R.P. Romană*, C.S.A., I.N.M.H., 1964

[32] W. Taylor, Change-Point Analyzer, ASA, 2002; available: http://www.variation.com/library.html

[33] W. H. Qian, Y. Zhu, Climate change in China from 1880 to 1998 and its impact on the environmental condition, *Climate Change*, 50, 2001, pp. 419 – 444

[34] G. Willeke, J. R. M. Hosking, J. R. Wallis, N. B. Guttman, *The National Drought Atlas*, Institute for Water Resources Report 94-NDS-4, U.S. Army Corps of Engineers, 1994

[35] F. Zwiers, V. Kharin, V., Changes in the extremes of climate simulated by CCC GCM2 under CO_2 doubling, *J. Climate*, 11, 1998, pp. 2200– 2222