

SINTEZA CERCETĂRILOR - 2007 - 2010

Cod: ID_262

Contract de cercetare nr. 429/2007

Titlul:

MODELAREA MATEMATICĂ A PRECIPITAȚIILOR DIN DOBROGEA

Perioada acoperită: 2007 - 2010

Data prezentării: 1 decembrie 2010

Director proiect: Conf. Dr. ALINA BĂRBULESCU

ECHIPA DE CERCETARE

1. Conf. dr. ing. Maftai Carmen Elena
2. Lect. dr. Pelican Elena
3. Dr. ing. Teodorescu Dacian Constantin
4. As. drd. Gherghina (căsătorită Șerban) Cristina

CUPRINS

1. Obiective
2. Date generale
3. Rezultate
4. Concluzii
5. Bibliografie

1. OBIECTIVE

- Descrierea statistica a seriilor de date constituite.
- Elaborarea de modele și predicția evoluției precipitațiilor.
- Analiza frecvențială.
- Determinarea dependenței lungi/scurte a seriilor precipitațiilor.
- Caracterizarea seriilor precipitațiilor cu ajutorul dimensiunilor fractionare.
- Analiza multifractală a seriilor de precipitații.
- Modele pentru precipitațiilor extreme.

2. DATE GENERALE

Dobrogea este o regiune situată în sud-estul României, între Marea Neagră și cursul inferior al Dunării. Structura sa este de platou cu aspect deluros. În general, climatul este continental temperat și este divizat în două unități: una, care include Delta Dunării, sudul său, lacurile Razim și Sinoe și o fâșie de 10 - 20 km lățime de-a lungul mării și o alta, care include restul teritoriului, unde climatul este influențat de centura continentală moderată.

Temperatura medie anuală este peste 11⁰C/an de-a lungul Litoralului și în Câmpia Dunării și sub 10⁰C - 11⁰C în nord și centru.

S-a dispus de serii de date anuale și lunare privind precipitațiile pe o perioadă de 41 ani (1965 - 2005), de la 10 stații hidro-meteorologice principale, aparținând ANM.

3. REZULTATE

3.1. Analiza statistică

Variabilele caracteristice definite au fost precipitațiile medii, maxime sau minime anuale, trimestriale și lunare.

După trasarea graficelor a fost verificată *ipoteza normalității* distribuției seriilor, utilizând testele Kolmogorov – Smirnov, Shapiro – Wilk, Jarque – Bera sau Q-Q plot. A rezultat că seriile anuale sunt normal distribuite, iar cele lunare nu au această proprietate.

Detectarea omogenității a fost realizată prin aplicarea testului Wilcoxon, la nivelul de semnificație de 1% și 5%. Pentru seriile anuale, exceptând Mangalia, ipoteza omogenității a fost respinsă la nivelul de 5%, dar acceptată la nivelul de 1%. Pentru seria anuală Mangalia s-a acceptat ipoteza omogenității la nivelul de 5%, omogenă; această ipoteză a fost respinsă pentru seriile lunare.

În vederea determinării punctelor de discontinuitate în evoluția seriilor au fost aplicate teste de ruptură. Rezultatele acestora au fost diferite. Pentru seria anuală Sulina, momentul de ruptură a fost 1981, iar pentru celelalte serii, testele Buishand și Pettitt nu au evidențiat schimbări de evoluție, în timp ce testele Lee & Heginian și Hubert au dat ca moment de ruptură 2003. Având în vedere că datele de după 2003 sunt insuficiente, nu putem avea în vedere acest moment ca punct de ruptură. Seriile lunare au prezentat cel puțin un punct de ruptură.

Aplicând testul Mann – Kendall, a fost respinsă ipoteza existenței unui trend pentru toate seriile anuale, exceptând Sulina.

În majoritatea seriilor lunare există valori aberante, care au fost eliminate înaintea determinării modelelor matematice.

Analiza corelației datelor a fost făcută utilizând funcția de autocorelație, ducând la concluzia că patru dintre seriile anuale nu sunt independente și toate seriile lunare sunt corelate.

Din analiza izohietelor s-a observat că, în general, valorile precipitațiilor medii anuale au crescut în perioada 1995 - 2005, exceptând Câmpia Dunării și stația Sulina.

3.2. Modele matematice

Tipurile de modele determinate pentru seriile de precipitații au fost următoarele:

I. *Modele multiplicative*, de forma:

$$X_t = Y_t \cdot S_t \cdot \varepsilon_t, t \geq 0,$$

unde: X_t este seria, Y_t - trendul, S_t - componenta sezonieră și ε_t - reziduul.

II. Modele de tip ARIMA/ FARIMA.

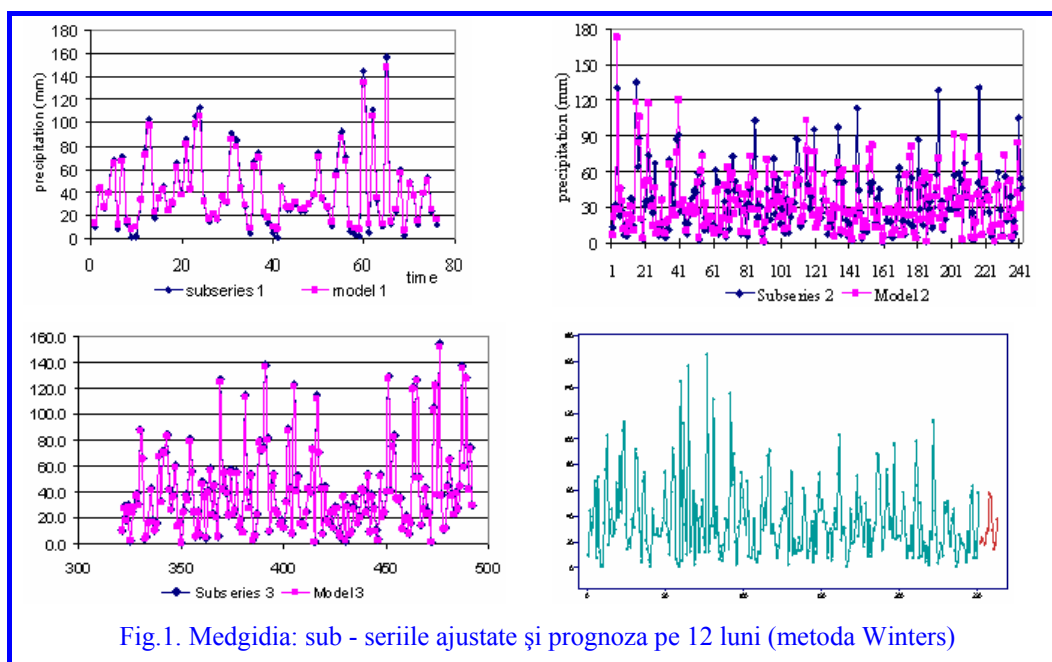
III. Modele obținute prin utilizarea rețelelor neurale.

IV. Modele obținute prin utilizarea algoritmilor genetici (GEP, AdaGEP) și hibridi (ARIMA – AdaGEP).

V. Modele frecvențiale (trasarea curbelor IDF).

În cele ce urmează exemplificăm.

1. Pentru seria lunară *Medgidia*, care prezintă punctele de ruptură mai 1971 și iulie 1991, s-au determinat următoarele modele (Fig.1):



- Pentru perioada ianuarie 1965 - mai 1971, după aplicarea unei transformări Box-Cox cu $\lambda = 0.42$, seriei inițiale, pentru seria rezultată, (X_t) , s-a determinat un model:

$$X_t = Z_t - 0.2242Z_{t-4}, t \in \overline{5, 76},$$

unde $(Z_t)_{t \in \overline{1, 76}}$ este un zgomot alb.

- Pentru perioada iulie 1971-iulie 1991 s-a determinat un model multiplicativ, trendul fiind calculat metoda mediilor mobile, factorii de sezonaliitate fiind dați în Tabelul 1.

Tabel 1. Factori sezonaliitate (%)

ian.	febr.	martie	aprilie	mai	iunie	iulie	august	sept.	oct.	nov.	dec.
182.2	144.9	114.1	75.6	83.3	102.2	74.2	64.0	68.0	62.4	104.9	0.991

- Pentru perioada septembrie 1991 - decembrie 2005, după o diferențiere de ordin 1, modelul propus, după înlăturarea coeficienților ne semnificativi și aplicarea criteriului verosimilității maxime, a fost de tip medie mobilă, de ordin 11:

$$X_t = Z_t - 0.6399Z_{t-11}, t \geq 12,$$

unde $(Z_t)_{t \geq 1}$ este un zgomot alb.

2. Pentru seria lunară *Sulina*, după o transformare Box - Cox cu parametrul $\lambda = 0.34$, se acceptă ipoteza că seria (Z_t) este normal distribuită. După analiza corelogramei se respinge ipoteza că seria este necorelată, la nivelul de încredere 95%. Ipoteza homoscedasticității a fost respinsă după aplicarea testului Bartlett.

- Sub - seria (Z_t) pe perioada ianuarie 1965 – august 1982 formează un zgomot alb.

- Pentru sub-seria (Z_t) pe perioada septembrie 1982 – decembrie 2005, s-a determinat după un model multiplicativ, in care trendul are ecuația:

$$\bar{Z}_t = 4.2768 + 0.6475 * \cos(0.05348t - 2.9331) \text{ (Fig.2),}$$

varianța reziduală a fost 0.326515, iar testele făcute conduc la acceptarea ipotezei că erorile formează un zgomot gaussian.

Tabel 2. Indici de sezonaliitate

ian.	febr.	martie	aprilie	mai	iunie	iulie	august	sept.	oct.	nov.	dec.
0.834	0.875	0.888	0.911	0.978	1.430	0.858	0.956	1.082	1.048	1.148	0.991

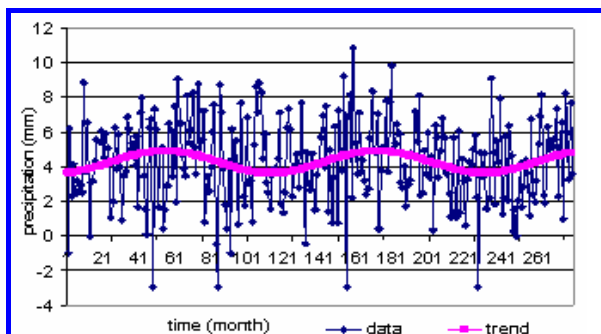


Fig. 2. Datele și trendul pentru sub - seria septembrie 1982 – decembrie 2005

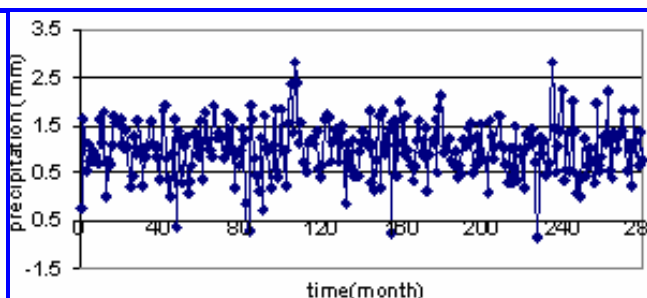


Fig. 3. Reziuul în modelul multiplicativ pentru sub - seria septembrie 1982 – decembrie 2005

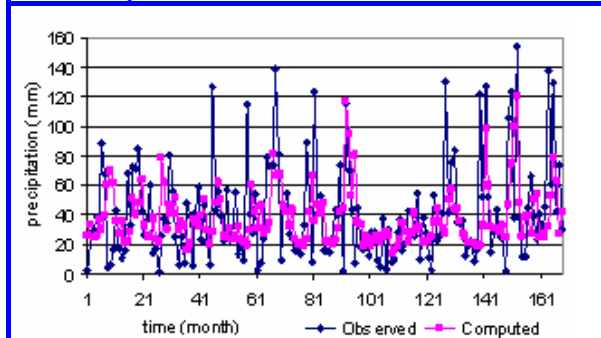


Fig. 4. Predicție pe 170 luni folosind rețele neurale pentru seria lunară Sulina

Predicția evoluției precipitațiilor s-a făcut prin antrenarea unei rețele neurale (Fig.4). Datele din fiecare sub-serie au fost împărțite în 3 mulțimi: de training, de validare și de testare. S-a folosit o rețea „two-layer feed forward” cu un layer de output și layere ascunse având între 5 - 10 neuroni *tansig*sau *purelin* și output - ul un neuron *tansig*.

3. Pentru seria anuală Sulina, cel mai bun model determinat (FARIMA) după extragerea mediei are ecuația

$$(1 - B)^{0.28} X_t = Z_t,$$

cu (Z_t) zgomot alb. (Fig.5)

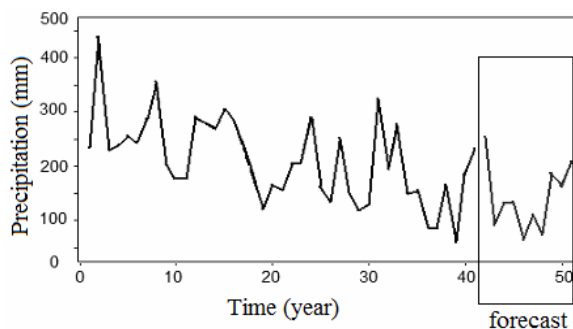
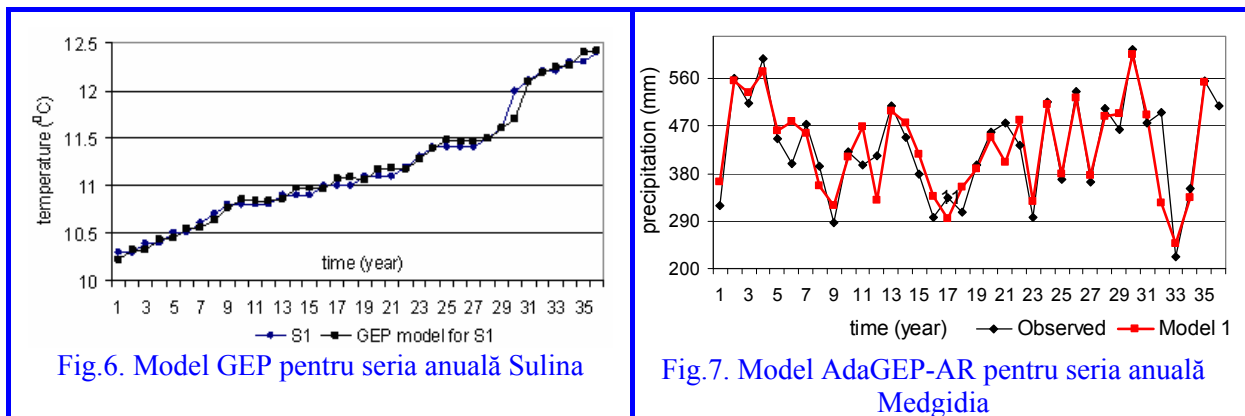


Fig. 5. Seria anuală Sulina

4. Deoarece în general seriile meteorologice prezintă variabilitate mare și evoluția neliniară nu poate fi captată totdeauna de modelele ARIMA/FARIMA, am utilizat programarea genetică (GEP) pentru determinarea de modele alternative (Fig.6).

5. S-a combinat metodologia Box-Jenkins (Fig.7), cu GEP, obținând modele hibride, AdaGEP - ARIMA care le îmbunătățesc pe cele obținute printr-un algoritm genetic adaptiv propus (AdaGEP).



6. Ajustarea modelelor frecvențiale și construirea curbelor intensitate - durată – frecvență pentru seriile de date a fost făcută utilizând legea lui Gumbel:

$$F(x) = \exp[-\exp(-(x - a)/b)].$$

Cu variabila redusă $u = (x - a)/b$, funcția de repartiție se scrie: $F(u) = \exp[-\exp(-u)]$ și

$$u(F) = -\ln(-\ln(-F(x))).$$

Avantajul folosirii variabilei reduse este acela că expresia quantilei este liniară:

$$x_q = a + bu_q.$$

Etapele ajustării sunt următoarele:

i. Pregătirea datelor pentru trasarea poligonului frecvențelor cumulate:

- trierea valorilor în ordine crescătoare și atașarea unui rang, r ,
- calculul frecvențelor empirice, cu formula Hazen:

$$F(x_r) = (r - 0.5) / n,$$

unde r este rangul, iar n – numărul de valori.

ii. Calculul variabilei reduse Gumbel, u ;

iii. Reprezentarea grafică a perechilor (u_i, x_i) și determinarea parametrilor a și b ai dreptei

de regresie, prin metoda momentelor sau a celor mai mici pătrate;

iv. Pentru timpi de întoarcere diferiți, T , calculul frecvenței $F = 1 - 1/T$;

v. Calculul variabilei reduse $u(F) = -\ln(-\ln(-F(x)))$;

vi. Calculul valorii estimate a quantilei x_q .

Curbele IDF pentru seriile Medgidia și Sulina sunt prezentate în Fig.8 și 9.

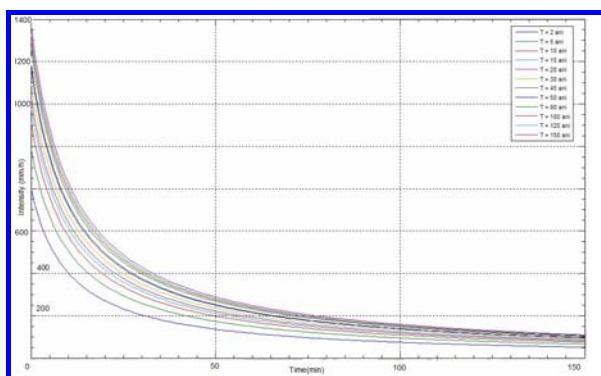


Fig.8. Curbe IDF pentru seria Medgidia

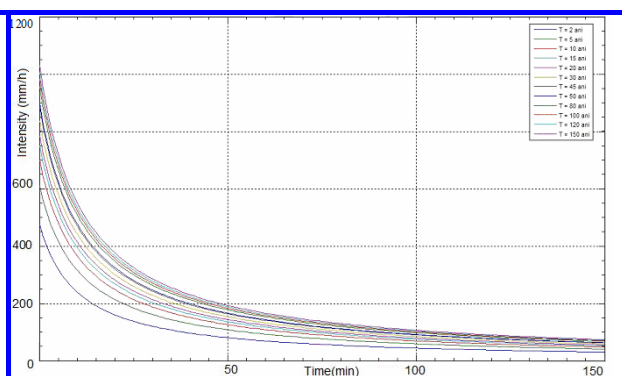


Fig.9. Curbe IDF pentru seria Sulina

3.3. Analiza dependenței lungi

Aceasta a fost realizată prin calculul coeficientului Hurst, H , prin metode diferite, ca:

- metoda R/S și statistica R/S modificată a lui Lo;
- metoda varianței agregate;
- metoda momentelor absolute;
- DFA;
- GPH;
- metoda raportului varianțelor reziduurilor;
- Periodograma;
- Wavelets.

Dacă procesul este aleator, atunci $H = 0,5$. Dacă procesul are proprietatea dependenței lungi, $H > 0,5$, iar dacă nu este persistent în timp, $H < 0,5$.

Rezultatele obținute pentru seriile anuale de date au fost diferite, funcție de metoda utilizată. În Tabelele 3-5 se găsesc rezultatele obținute pentru seriile anuale prin aplicarea unora dintre metodele menționate.

Tabel 3. Coeficientul Hurst pentru seriile anuale, calculat prin metoda R/S

Station	Adamclisi	Cernavoda	Constanta	Corugea	Harsova
H	0.7207	0.9775	0.9738	0.9224	0.8028
Station	Jurilovca	Medgidia	Mangalia	Sulina	Tulcea
H	0.9301	0.8727	0.7049	0.9567	0.7057

Tabel 4. Coeficientul Hurst pentru seriile anuale, calculat prin metoda Lo

Station	Adamclisi	Cernavoda	Constanta	Corugea	Harsova
H	0.0988	0.1177	0.11	0.1262	0.1167
Station	Jurilovca	Medgidia	Mangalia	Sulina	Tulcea
H	0.1104	0.1304	0.1165	0.119	0.1121

Tabel 5. Coeficientul Hurst pentru seriile anuale calculat prin metoda periodogramei

Station	Adamclisi	Cernavoda	Constanta	Corugea	Harsova
H	0.44505	0.4491	0.36745	0.2826	0.55745
Station	Jurilovca	Medgidia	Mangalia	Sulina	Tulcea
H	0.46905	0.3967	0.4559	0.32415	0.58925

Tabel 6. Coeficientul Hurst pentru seriile anuale calculat prin metoda GPH

Station	Adamclisi	Cernavoda	Constanta	Corugea	Harsova
H	0.44505	0.4491	0.36745	0.2826	0.55745
Station	Jurilovca	Medgidia	Mangalia	Sulina	Tulcea
H	0.46905	0.3967	0.4559	0.32415	0.58925

Un program care calculează coeficienții prin metodele precizate a fost construit (Fig.10).

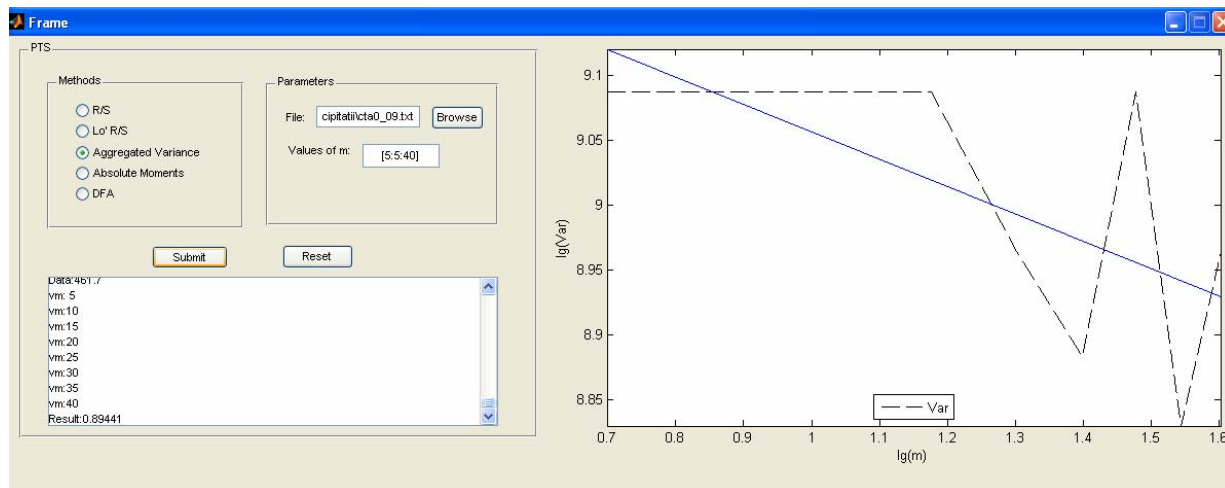


Fig.10. Print - screen al interfaței grafice a programului

În figurile de mai jos prezentăm graficul R/S și al statisticii V asociate, pentru seria Adamclisi. Grafice asemănătoare au fost trasate pentru toate seriile în studiu.

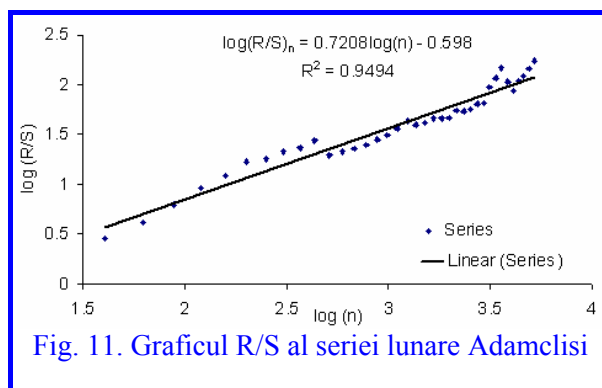


Fig. 11. Graficul R/S al seriei lunare Adamclisi

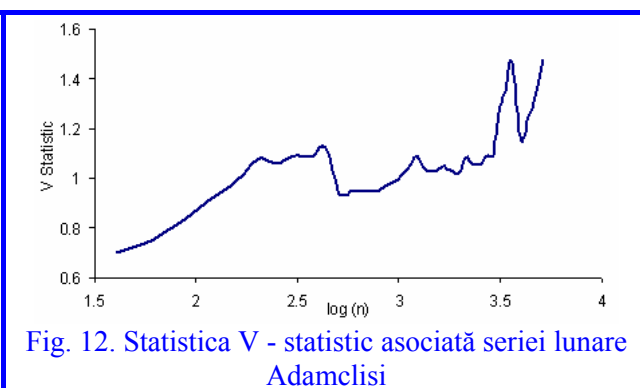


Fig. 12. Statistica V - statistic asociată seriei lunare Adamclisi

3.4. Caracterizarea seriilor precipitațiilor cu ajutorul dimensiunilor fracționare

Studiul caracterului fractal al seriilor precipitațiilor se poate face cu ajutorul dimensiunilor fracționare, ca *box*, *ruler*, *informational*, *fragmentation dimension* etc.

Tabel 6. Box dimension pentru seriile lunare

	Box dimension	Std. deviation
Adamclisi	1.89359	0.0058726
Cernavoda	1.89878	0.0078325
Constanta	1.90286	0.0055319
Corugea	1.89716	0.0055399
Harsova	1.89729	0.0054955
Jurilovca	1.90063	0.0057935
Mangalia	1.90238	0.0056315
Medgidia	1.88917	0.0062289
Sulina	1.90197	0.0057572
Tulcea	1.89563	0.0055636

Din Tabelul 6 rezultă omogenitatea dimensiunilor *box*, în jurul lui 1.9, pentru seriile de precipitații lunare.

3.5. Analiza multifractală a seriilor de precipitații

Studiul caracterului multifractal al seriilor de precipitații lunare presupune determinarea exponentului de scală, $\tau(q)$ și a spectrului singularităților $f(\alpha)$. El a fost realizat prin două tehnici diferite (box-counting și wavelets transform maxima modulus - WTMM).

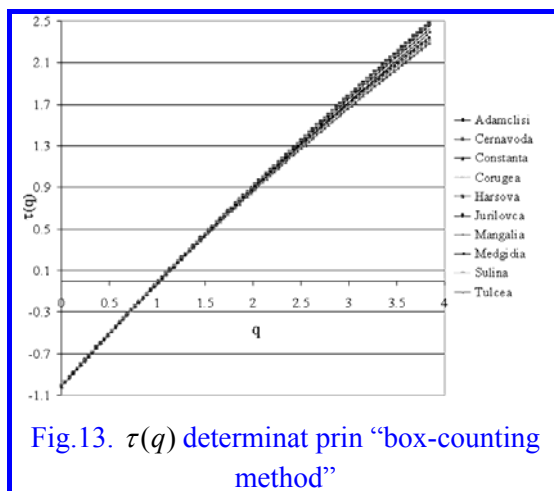


Fig. 13. $\tau(q)$ determinat prin "box-counting method"

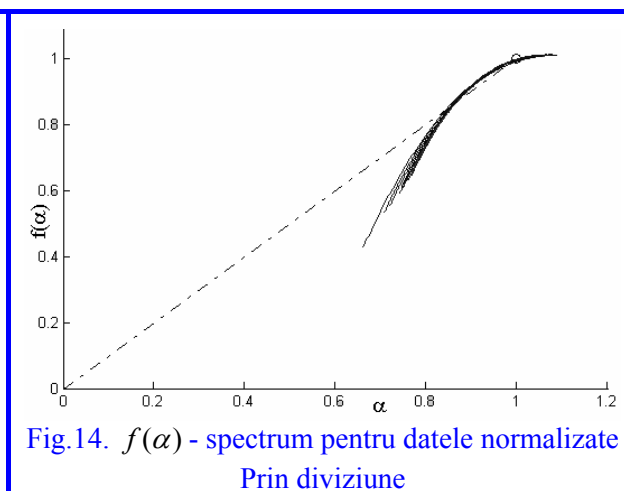


Fig. 14. $f(\alpha)$ - spectrum pentru date normalizate Prin diviziune

Rezultatele în extenso sunt prezentate într-un articol trimis spre publicare [A. Bărbulescu s.a., 2010].

3.6. Modele pentru precipitațiilor extreme

Studiul precipitațiilor maxime zilnice și lunare a fost realizat prin determinarea unor modele de tip Poisson generalizat (GPD), rezultatele fiind prezentate în două articole, unul în curs de apariție în IJMC, și unul submitis la Int. Journal of Climatology. A se vedea și cartea [Bărbulescu s.a., NOVA Publishing, 2010]. Graficele de diagnostic pentru modelul propus pentru seria zilnică Tulcea se găsește în Fig. 15.

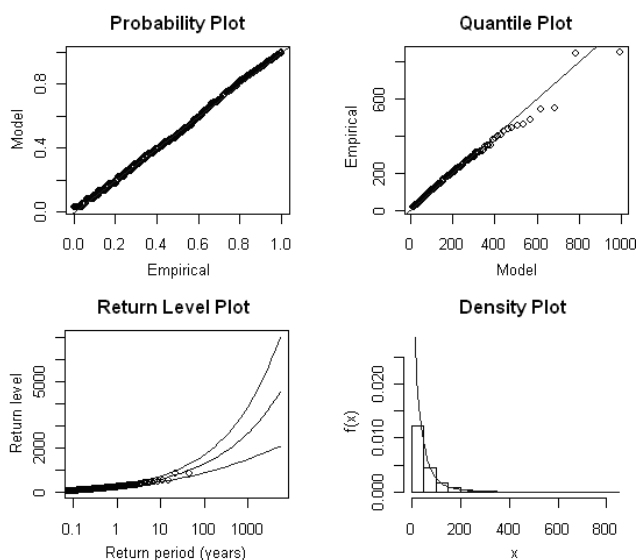


Fig. 15. Model GPD pentru seria zilnică Tulcea [IJMC, 2010]

De asemenea a fost realizat studiul precipitației maxime probabile – PMP (fig.16), rezultatele fiind curs de apariție în lucrarea [Bărbulescu s.a., LAP Publishing, 2010].

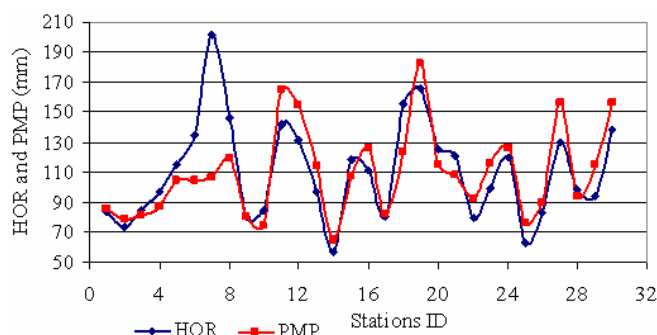


Fig.16. Rezultatele metodei Hershfield pentru precipitațiile maxime în 24 ore

4. CONCLUZII

În cadrul contractului de cercetare s-a realizat un studiu sistematic al precipitațiilor din regiunea Dobrogea. Modelele obținute pot fi utilizate cu succes pentru predicția evoluției precipitațiilor din zonă. De asemenea, modelul regional obținut pentru Dobrogea, poate fi încadrat într-un model global, pentru descrierea evoluției climatice în sud-estul Europei.

5. BIBLIOGRAFIE

- H. Akaike, Information theory and an extension of the maximum likelihood principle, 2nd International Symposium on Information Theory, B. N. Petrov and F. Csaki (eds.), Akademia Kiaofo, Budapest, pp. 267 - 281
- A. Bărbulescu, *Time series with applications*, Junimea, Iași, 2002
- A. Bărbulescu, E. Băutu, Mathematical models of climate evolution in Dobrudja, *Theoretical and Applied Climatology*, vol.100, nos.1 – 2/March, 2010, pp. 29 - 44, DOI 10.1007/s00704 – 009 – 0160 – 7
- A. Bărbulescu, E. Băutu, Meteorological Time Series Modelling Based on Gene Expression Programming, *Recent Advances in Evolutionary Computing*, WSEAS Press, 2009, pp. 17-23
- A. Bărbulescu, E. Băutu, ARIMA Models versus Gene Expression Programming in Precipitation Modeling, *Recent Advances in Evolutionary Computing*, WSEAS Press, 2009, pp.112-117
- A. Bărbulescu, E. Băutu, ARIMA and GEP models for climate variation, *International Journal of Mathematics and Computation*, June 2009, Volume 3, No J09, pp. 1-7
- A. Bărbulescu, E. Băutu, Time Series Modeling Using an Adaptive Gene Expression Programming, *International Journal of Mathematical Models and Methods in Applied Sciences*, Issue 2, Volume 3, 2009, pp. 85 – 93
- A. Bărbulescu, J. Deguenon, D. Teodorescu, *Study of water resources in the Black Sea region. Some mathematical models*, Nova Publishing, USA, 2010 (in print)
- A. Bărbulescu, C. Maftai, E. Băutu, *Modeling the hydro-meteorological time series. Applications to Dobrudja region*, Lambert Academic Publishing GmbH & Co. KG, Germany, 2010 (ISBN 978-3-8433-6578-9)
- A. Bărbulescu, E. Pelican, ARIMA models for the analysis of the precipitation evolution, *Recent Advances in Computers*, 2009, pp. 221 – 226
- A. Bărbulescu, D - C. Toncu, Modelling Precipitation Influence on Tăbăcărie Lake Water Quality, *Mathematical Methods, Computational Techniques, Intelligent Systems*, 2010, pp.152 - 158
- A. Bărbulescu, C. Șerban (Gherghina), C. Maftai, Evaluation of Hurst exponent for precipitation time series, *Latest Trends on Computers*, Vol.II, 2010, pp. 590 – 595

A. Bărbulescu, The analysis of correlation of some ions concentration in rainwater in an urban area, *International Journal of Mathematical Models and Methods in Applied Sciences*, Issue 2, Vol. 4, 2010, pp.105 – 112

A. Bărbulescu, C. Șerban (Gherghina), C. Maftai, Statistical Analysis and Evaluation of Hurst Coefficient for Annual and Monthly Precipitation Time Series, *WSEAS Transactions on Mathematics*, Issue 10, Vol. 9, October 2010, pp. 791 – 800 (ISSN 1109 - 2769)

A. Bărbulescu, C. Șerban, A. Cârsteanu, Statistical and multifractal analysis of rainfall of Romania, 2010, submis

P. J. Brockwell, R.A. Davis, *Time series analysis, forecasting and control*, Holden - Day, San Francisco, 1976

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

T. A. Buishand, Tests for detecting a shift in the mean of hydrological time series, *Journal of Hydrology* 58, 1984, pp. 51-69.10.

A. Chhabra, R. Jensen, Direct determination of the $f(\alpha)$ singularity spectrum, *Physical Review Letters* 62 (12), 1989, pp.1327–1330.

J. Deguenon, A. Bărbulescu, Study of Extreme Daily Rainfall Using GPD Model, IJMC, 2010, în curs de apariție

J. Deguenon, A. Bărbulescu, GPD Models for extreme rainfall with cyclic trend detection over ten stations in Dobrudja, *International Journal of Climatology*, 2010, submis 5.08.2010

P. Hubert et al, 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

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

Khronostat, <http://www.hydrosciences.org/spip.php?article239&lang=en>

A. F. S. Lee, S. M., 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

H. Lubes-Niel et al., Caractérisation de fluctuations dans une série chronologique par applications de tests statistiques - Etude bibliographique, *Rapport interne ICCARE* 3, 1994, ORSTOM – Hydrologie

C. Maftai, A. Bărbulescu, Computing the intensity - duration - frequency curves for a Romanian catchments, *Proceedings of The third International Conference of Mathematical Sciences*, 3-6.03.2008, Al Ain, p.1178-1186.

A. N. Pettitt, A non - parametric approach to the change-point problem, *Applied Statistics* 28, n°2, 1979, pp.126 - 135.

M. Taqqu, V. Teverovsky, W. Willinger, Estimators for long range dependence: an empirical study, *Fractals*, vol.3, no.4, pp.785-788.

W. Taylor, Change point analyser 2.0 shareware program, Taylor Entreprises, Libertyville, Illinois, 2003, <http://variation.com/cpa>

V. Venugopal, S.G. Roux, E. Foufoula-Georgiou, A. Arneodo, Revisiting multifractality of high resolution temporal rainfall using a wavelet-based formalism, *Water Resources Research*, 42, 2006, W06D14, doi:10.1029/2005WR004489

R. Weron, Estimating long range dependence: finite sample properties and confidence intervals, arXiv: cond-mat/0103510v2 9 May 2001