

UTILIZAREA SOFT-ULUI SPSS 22 PENTRU PRELUCRAREA DATELOR HIDROLOGO-HIDROGEOLOGICE

Liubov VĂTĂMANU, cercetător științific stagier
Institutul de Geologie și Seismologie, liubitaturcan@gmail.com

Abstract. Geostatistical analysis of hydrological and hydrogeological data has been realized through the international software SPSS (Statistical Package for Social Sciences) 22. Based on hydrological data, statistical descriptive analysis and statistical explore analysis were obtained. The obtained results are used for estimate verbal probability and hydro-hydrogeological data structure.

Cuvinte cheie: SPSS (Statistical Package for The Social Sciences), Q-Q Plot, Boxplot, testul Kolmogorov-Smirnov și testul Shapiro-Wilk.

INTRODUCERE

Caracteristica geostatistică a materialelor hidrologo-hidrogeologice este un element important al cercetărilor în acest domeniu. Geostatistica datelor hidrologo-hidrogeologice poate fi efectuată prin intermediul al mai multor soft-uri. Unul din cele mai folosite în plan internațional este SPSS (Statistical Package for The Social Sciences). În cazul nostru a fost utilizat SPSS 22 .

Pentru Republica Moldova, în calitate de exemplu metodic reprezentativ a estimării resurselor apelor subterane, a fost selectat bazinul hidrografic al r. Răut. În premieră, s-au folosit modele geostatistice pentru acest bazin, defalcarea hidrografului sa efectuat cu metoda Voskresenski.

MATERIALE ȘI METODE

Datele inițiale pentru acest studiu au fost colectate din lucrările anual publicate: „Гидрологический ежегодник”[1] și „Государственный водный кадастр”[2]. Pentru studiul dat în cauză sa utilizat metoda lui Voskresenski ”[3]. Prelucrării datelor geostatistice a curgerii apelor subterane pentru bazinul r. Răut se bazează pe datele înregistrate în punctele hidrologice de observație din localitatea Florești, Bălți, Răuțel (Fălești), Căzănești (Telenești), Jeloboc (Orhei), Orhei, care conțin informații cu privire la debitul de apă (Q) pe o anumită perioada de timp (tab.1).

Tabel 1: Durata observațiilor a r. Răut în punctele hidrologice

Râu	Punct hidrologic	Perioada	Anii de monitorizarea
Răut	Florești	1949-1962	15
	Bălți	1948-1990	43
	Răuțel	1970-1986	17
	Căzănești	1955-1986	32
	Jeloboc	1958-1990	34
	Orhei	1945-1956	12

Ca instrument de lucru au fost folosite softurile Microsoft Office Excel 2013, SPSS 22 (Statistical Package for The Social Sciences) [3]. Din tot spectru de obțiuni a acestui soft s-au folosit analiza statistică descriptivă și analiza statistică de cercetare. Metodologia cercetării cu soft-ul SPSS include pregătirea matricelor de date în Excel 2013, importul lor în fișierele SPSS, corecția fișierelor SPSS, calculul statisticilor și analiza rezultatelor. Pentru valorile debitele râurilor mici și curgerii apelor subterane au fost obținută statistica parametrică descriptivă, estimativă, percentile, testul de normalitate (Kolmogorov-Smirnov, Shapiro-Wilk), Q-Q Plot, BoxPlot și alți parametrii geostatistici.

DISCUȚII ȘI REZULTATE

Teritoriul de studiu este bazinul r. Răut, situat în partea centrală a Republicii Moldova, suprafața căruia este de 7760 km². Este cel mai mare afluent al Nistrului și, totodată, cel mai mare râu care izvorăște și curge integral pe teritoriul țării, având lungimea de 286 km. Răul izvorăște din satul Rediul Mare, la altitudinea de 180 m și debușează în râul Nistru de pe malul drept, la 342 km de la gura fluviului, lângă orașul Dubăsari, la altitudinea de 12 m [5]. Totodată, acest bazin este amplasat în condiții meteorologice specifice, care determină formarea excesului de umiditate, ca sursă de alimentare atât a apelor subterane, cât și a celor de suprafață. În rezultatul modelării geostatistice au fost obținute datele expuse în tab. 2.

Tabel 2: Statistica descriptivă curgerii apelor subterane după metoda Voskresenski

Râu	Nr.	M-m	m	M	Σ	\bar{x}	$\sigma\bar{x}$
Răut- Florești	15	37,31	3,66	40,98	153,39	10,23	2,39
Răut- Bălți	43	47,93	1,56	49,50	573,99	13,35	1,76
Răut- Răuțel	17	25,87	8,92	34,79	322,76	18,99	1,75
Răut- Căzănești	32	57,37	3,43	60,80	794,62	24,83	2,95
Răut- Jeloboc	34	51,59	1,41	53,00	775,32	22,80	2,50
Răut- Orhei	12	33,59	1,41	35,01	83,38	6,95	2,70

Remarcă: Valorile minimale, maximale, medii și deviația standard – sunt în mm. Nr. – numărul date; m – minimum; M-maximum; ; M-m – maximum minus minimum; \bar{x} - valoare medie; $\sigma\bar{x}$ - eroare standard a valorii medii; σ - deviația standard; σ^2 - dispersie; A- asimetrie; $A\sigma$ - eroare standard de asimetrie; E- aplatizare; $E\sigma$ - eroare standard de aplatizare; C- coeficientul de variație.

Tabel 2 (continuare): Statistica descriptivă curgerii apelor subterane după metoda Voskresenski

Râu	σ	σ^2	A	$A\sigma$	E	$E\sigma$	C
Răut- Florești	9,24	85,42	2,97	0,58	9,80	1,12	0,90
Răut- Bălți	11,55	133,44	1,42	0,36	1,67	0,71	0,87
Răut- Răuțel	7,22	52,17	0,74	0,55	-0,21	1,06	0,38
Răut- Căzănești	16,68	278,30	0,48	0,41	-0,60	0,81	0,67
Răut- Jeloboc	14,57	212,38	0,43	0,40	-0,47	0,79	0,64
Răut- Orhei	9,36	87,56	2,88	0,64	8,79	1,23	1,35

Analiza datelor tab. 2 și a fig. 1 demonstrează faptul, că datele sunt distribuite conform legii normale. Valoarea mediei \bar{x} este veridică statistic. Graficul Q-Q Plot normal este utilizat la testarea normalității datelor, cu cât reprezentarea grafică a șirului de date se desparte de diagonala din cadrul graficului, cu atât distribuția fiind asimetrică sau departe de legea normală. Cele mai importante caracteristici statistice ale unor distribuții de frecvență înglobează diagrama Boxplot. Această diagramă sunt reprezentate 5 valori de distribuție: valoarea minimă, cuantila inferioară, median, cuantila superioară și valoarea maximă (fig. 2).

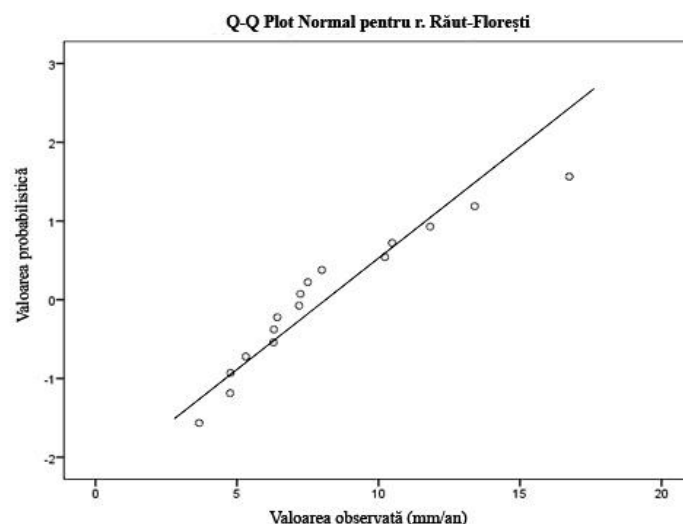


Fig. 1: Q-Q Plot normal(exemplu Răut-Florești)

În Tab. 3 sunt prezentate rezultatele testului Kolmogorov-Smirnov și testul Shapiro-Wilk. Testul Kolmogorov–Smirnov utilizează statistica D, pentru a testa apropierea unei distribuții probabilistice de o distribuție probabilistic de referință (variantea cu un eșantion), sau două distribuții diferite (variantea cu două eșantioane). Testul normalității Shapiro-Wilk utilizează statistica W, pentru a testa normalitatea unei distribuții. Ipoteza nulă este că distribuția aparține unui șir de date distribuită normal. La valori mici ipoteza nulă nu este acceptată, și normalitatea nu este acceptată, pe când valori mari confirmă normalitatea. Testul Kolmogorov-Smirnov este utilizat pentru a testa seturi mari de date în timp ce testul Shapiro-Wilk este mai potrivit pentru un eșantion mai mic, cum ar fi de 50 de numere sau mai mic. În cazul în care coloana „Sig” pentru oricare dintre teste este peste 0.05, datele sunt normal distribuite [4].

Tabel 3: Testul de normalitate

Râuri	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Răut- Florești	0,203	16,000	0,078	0,906	16,000	0,100
Răut- Bălți	0,178	5,000	0,200	0,985	5,000	0,960
Răut- Răuțel	0,174	5,000	0,200	0,944	5,000	0,694
Răut- Căzânești	0,166	10,000	0,200	0,934	10,000	0,488
Răut- Jeloboc	0,167	11,000	0,200	0,897	11,000	0,171
Răut- Orhei	0,374	4,000		0,768	4,000	0,056

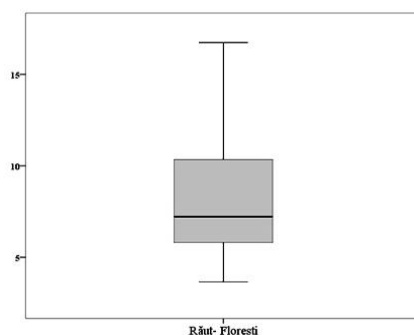


Fig. 2: Boxplot r.Răut (Florești)

CONCLUZII ȘI RECOMANDĂRI

1. Cu scopul evaluării resurselor apelor subterane în limetele teritoriului Republicii Moldova, pentru prima dată au fost testat și utilizată modelarea geostatistică a apelor subterane prin intermediu soft-ului SPSS (Statistical Package for The Social Sciences). Metoda geostatistica a permis evaluarea probalistică a parametrilor investigați pentru acest studiu.

2. Geostatistica hidrologo-hidrogeologică a fost efectuată în două compartimente: statistică descriptivă și statistică de cercetare.

3. Conform testului de normalitate datele sunt distribuite normal, deoarece datele nu depășesc valoarea 0,05 în coloana „Sig” pentru r. Răut (Florești), metoda Voskresenski. Cât acest fapt atestă graficul Q-Q Plot în care șirului de date sunt apropiate de diagonala din cadrului graficului, distribuția fiind simetrică.

REFERINȚE BIBLIOGRAFICE

1. *Гидрологический Ежегодник 1958 - 1978 г. Том 2 Бассейн Черного моря (без Кавказа) выпуск 0, 1 Бассейн Черного моря без бассейнов р. Днепр и р. Дон.* - Ленинград: Гидрометеорологическое, 1958, 480 с.

2. *Государственный водный кадастр Ежегодные данные о режиме и ресурсах поверхностных вод суши 1979 – 1986 г. Части 1 и 2 Том IX.* - Кишинев, 1979, 147 с.

3. *Труды Государственного Гидрологического Института, выпуск 29(83)* - Ленинград: Гидрометеоздат, 1951, 1148 с.

4. Statistical Package for the Social Sciences (SPSS 22). www.ibm.com/us-en/marketplace/spss-statistics (accesat 22.11.2017)

5. <http://analize-statistice.eu/testarea-normalitatii-distributiei-datelor/> (accesat 20.09.2018)

6. https://ro.wikipedia.org/wiki/R%C3%A2ul_R%C4%83ut (accesat 25.09.2018)