

911.2:511.58 (23)

Мр Бошко Миловановић\*

## РЕЗУЛТАТИ ПРИМЕНЕ СТАТИСТИЧКИХ ПОСТУПАКА У ИСТРАЖИВАЊУ ПАДАВИНА НА СТАРОЈ ПЛАНИНИ

*Оправдано је уверење да се научно знање од других  
врста знања разликује, поред осталог, и по методу  
којим се та знања стичу и развијају  
Живан Ристић - О истраживању, методу и знању*

**Абстракт:** У раду је дат сажет приказ неколико статистичких поступака који се могу применити и у климатолошким истраживањима. За неке од њих приказана је једначина на којој се заснивају. Као конкретан бројчани пример на који су примењени статистички поступци, изабрани су подаци о средњим годишњим и средњим декадним вредностима NAO индекса и количини падавина на ширем простору Старе планине у периоду 1951-2000. година. Хомогеност података о падавинама са 11 станица које се налазе на посматраном простору је тестирана Ман-Витнијевим тестом, док су за попуњавање недостајућих података искоришћена три различита метода. Применом статистичких процедура на скупове података различитих карактеристика, покушано је да се утврди која од њих најпотпуније одсликава расподелу падавина у контексту просторне заступљености и повезаности са NAO индексом. Употребом параметарских и непараметарских поступака уочена је веза (негативна корелација) између NAO индекса и количине падавина и на годишњем и на декадном нивоу. Испитана је повезаност између климатских модификатора (експозиције нагиба) и количине падавина на Старој планини. Израчунат је тренд падавина на декадном нивоу и тестирана његова статистичка значајност.

**Кључне речи:** Параметарске и непараметарске статистичке процедуре, NAO индекс, падавине, Стара планина, Србија

### Увод

Стара планина представља део пространог Карпатско – балканског планинског лука. У Србији (њеном југоисточном делу) се налази само његов мањи, западни део који захвата површину од 1802 km<sup>2</sup>. Као морфолошка целина, Стара планина је на западу омеђена долинама Белог и Трговишког Тимока и Височице, а на истоку државном границом Србије и Бугарске. Посматрани простор се одликује изразито дисецираним рељефом

---

\* Завод за заштиту природе Србије, Др Ивана Рибара 91 - Нови Београд,  
[bohumil@verat.net](mailto:bohumil@verat.net)

и великим разликама у надморским висинама (од 42 m н.в, односно од 132 m н.в. до 2169 m н.в.) на коме је мрежа станица ретко и неравномерно распоређена (највиша станица – Дојкинци се налази на „свега“ 880 m н.в), са најчешће оскудним и непотпуним временским серијама (Миловановић, 2005).

Актуелност истраживања везаних за утицај NAO индекса на климу Европе у целини, односно њеним појединим регијама (Hurrell, 1995), усмерава на испитивање повезаности овог климатског фактора (али и неких климатских модификатора) и падавина на Старој планини. С обзиром да избор и квалитет методолошког апарата који се користи у знатној мери утиче на квалитет добијених резултата, у раду је дат и приказ основних карактеристика и функција статистичких процедура које су у истраживању коришћене.

### **Коришћени подаци**

У раду су коришћени подаци Републичког хидрометеоролошког завода Србије о средњим годишњим количинама падавина (период 1951-2000) са 11 станица (Алдинац, Димитровград, Дојкинци, Кална, Каменица Димитровградска, Неготин, Пирот, Смиловци, Височка Ржана, Витановац и Зајечар) које се налазе на ширем простору Старе планине. Хомогеност података је испитана Ман-Витнијевим непараметарским тестом. “Using a test based on the ranks of values from a time series has the benefit, that it is not greatly particularly adversely affected by outliers” (Peterson et al, 1998). Показало се да осим Алдинца и Дојкинаца, све остале станице имају релативно хомогене низове. С обзиром да су низови комплетни на свега 5 станица (Димитровград, Дојкинци, Кална, Неготин, Зајечар) извршено је попуњавање недостајућих података. У том процесу, прво се приступило одређивању коефицијента корелације између посматраних парова станица. Међутим, коефицијент корелације није установљен на нивоу годишњих вредности, већ је за сваки пар станица коефицијент корелације утврђен појединачно на нивоу месечних вредности. Добијени резултати на одређеном пару станица за одређени месец се не користе обавезно и за остале месечне вредности на истом пару станица. За попуњавање недостајућих података су коришћени: метод редукције података на исти број година, метод интерполације недостајућих података помоћу матрице (основна - са осам елемената матрице и модификована варијанта - са најмање 6 елемената матрице) и екстраполација заснована на претпоставци константног односа краћих и дужих низова.

Табела 1. Средње годишње вредности NAO индекса и количине падавина на ширем простору Старе планине (1951-2000)

Година	Падавине (mm)	NAO1 индекс	Година	Падавине (mm)	NAO индекс
1951	711,0	-0,07	1976	861,2	-0,07
1952	721,3	-0,37	1977	709,4	-0,21
1953	640,4	0,40	1978	717,5	0,21
1954	854,6	0,51	1979	689,6	0,19
1955	880,5	-0,64	1980	807,6	-0,37
1956	702,3	0,17	1981	770,1	-0,09
1957	887,5	-0,02	1982	662,9	0,67
1958	536,1	0,12	1983	570,6	0,34
1959	682,3	0,49	1984	493,0	0,26
1960	720,4	-0,30	1985	629,8	-0,47
1961	626,0	1,05	1986	618,0	0,56
1962	806,6	-0,13	1987	729,0	-0,51
1963	735,6	-0,39	1988	577,7	-0,32
1964	731,5	0,24	1989	627,6	0,57
1965	660,3	-0,23	1990	543,3	1,23
1966	819,2	-0,22	1991	648,5	0,34
1967	629,1	0,56	1992	544,9	1,11
1968	641,0	-0,62	1993	486,9	0,12
1969	736,2	-0,44	1994	518,1	0,51
1970	822,7	0,18	1995	822,2	-0,61
1971	627,7	-0,55	1996	745,0	-1,01
1972	728,7	-0,04	1997	665,2	-0,18
1973	702,7	-0,09	1998	715,3	0,26
1974	774,5	0,59	1999	741,6	0,05
1975	768,4	0,05	2000	393,8	0,04

Овакав, комбиновани приступ у попуњавању података је одабран, да би се у што већој мери подаци, односно метаподаци приближили стварним измереним вредностима. Radovanovic, Milovanovic (2003) тестирајући поменуте методе на теренима Србије изнад 1000 m н.в. наводе следеће: "Each of applied and tested methods showed its advantages and disadvantages. The method of reduction demonstrated a great applicability and satisfying accuracy. The method of interpolation showed certain limitations in the

<sup>1</sup> Подаци о NAO индексу су преузети са [www.cru.uea.ac.uk/ftpdata/nao.dat](http://www.cru.uea.ac.uk/ftpdata/nao.dat)

possibilities of its application, but also in the accuracy of the data processed. The method of extrapolation proved to have the greatest limitations in its application, but also the highest degree of accuracy. It can be concluded, that in depending of our needs and the completeness of the series of data, the combination of the methods presented in the paper seems to be an adequate approach to data processing". На посматраном простору је веома изражена варијабилност падавина. У периоду 1951-2000. година, највећа количина падавина је забележена на станици Дојкинци (880 m н.в. – 865,8 mm). Са друге стране, у Пироту који се налази на 370 m н.в. се излучи свега 583,3 mm. Овако неравномерна расподела падавина на релативно малом простору је честа у планинским теренима. „У одређеним случајевима, паралелни ниски гребени у сложеним планинским комплексима, могу представљати предиспозицију кишним сенкама. Ослабљене ваздушне масе, могу се дакле и на оваквим, релативно ниским препрекама знатно осиромашити талогом, при чему заветрене стране бивају осетно сувље” (Радовановић, 2001). Просторним осредњавањем средњих годишњих количина падавина са посматраних станица добијене су вредности за Стару планину у целини (табела 1). Овакав приступ је одабран због релативно малог простора на коме су анализиране падавине, као и због непотпуности, односно нехомогености изворне базе података.

### Методе рада и резултати

У квантификавању повезаности две појаве, односно две променљиве  $x$ ,  $y$  (којом се не потврђује узрочно-последични однос) често се користи корелациона анализа. У случају да се мери линеарна корелација (ако су вредности варијабле мерене најмање на интервалном нивоу и ако је испуњена претпоставка о нормалности расподеле) израчунава се Пирсонов коефицијент корелације  $r$ . Коефицијент корелације између годишње вредности  $NAO$  индекса и количине падавина на Старој планини (табела 1) износи  $-0,37$ . Примењујући корелациону анализу на податке о  $NAO$  индексу и падавинама са 10 станица на територији Бугарске (у периоду 1961-2000),  $Nikolova$  (2004) је израчунала сличне вредности за  $r$  (од  $-0,13$  до  $-0,47$ ).

Проверавањем нулте хипотезе по којој не постоји повезаност између посматраних варијабли ( $H_0: \rho=0$ , на супрот алтернативне хипотезе  $H_a: \rho<0$ ) и рачунањем  $t$  статистика стичемо информацију колика је вероватноћа да се вредност од  $-0,37$  појави случајно.

Провером у табелици  $t$  - дистрибуције може се закључити да је вероватноћа да је нулта хипотеза истинита веома мала (испод  $0,01$ ), те се она стога

може одбацити, а прихватити алтернативна хипотеза по којој на годишњем нивоу постоји статистички значајна повезаност између NAO индекса и количине падавина на Старој планини.

Врло често се због изоловања везе коју испитујемо, односно пречишћавања везе између варијабли од различитих „шумова“ (нпр. нехомогеност података) у климатолошким истраживањима прибегава агрегирању података (просторном и временском). У овом случају су на основу годишњих вредности за NAO индекс и падавине на ширем простору Старе планине израчунате њихове декадне вредности. Из тога проистиче да у низу постоји свега пет елемената (табела 2). Пирсонов негативни коефицијент корелације је значајно већи и износи  $-0,81$ , али за овај број елемената ( $n=5$ ,  $df=3$ ) не задовољава  $t$ -тест на нивоу поверења од 95%.

Табела 2. Агрегиране (декадне) вредности NAO индекса и количине падавина на ширем простору Старе планине

Декада	NAO индекс	Количина падавина на ширем простору Старе планине (mm)
1951-1960	0,03	733,6
1961-1970	0,00	721,6
1971-1980	-0,03	738,7
1981-1990	0,22	622,2
1991-2000	0,06	628,2

У случају да се ради о нелинеарној релацији између варијабли, могуће је трансформисати (нпр. логаритамски) једну, односно обе варијабле, или применити неки од непараметарских поступака корелације.

Овакви методи, понекад називани методи независни од расподеле, потребни су када узимамо узорке из популација које су далеко од нормалних, или када се не зна много о типу расподеле за коју се вади узорак. Спирманов поступак израчунавања корелације представља широко распрострањену, непараметарску алтернативу за Пирсонов коефицијент корелације. У суштини, информација коју стичемо употребом овог поступка је веома слична као и при употреби Пирсоновог коефицијента корелације (перфектна позитивна корелација има вредност  $+1$ ; перфектна негативна корелација вредност  $-1$ ; вредности између означавају корелацију

која није савршена). Међутим, овај поступак је заснован на ранговима и прикладан је за рачунање корелација када су варијабле мерене на ординалном нивоу (или на интервалном, али претпоставка нормалности није задовољена). Рангирање се врши рачунањем разлика између рангова додељених случајевима у оквиру варијабли (у опадајућем или растућем редоследу). Према Harnett, Murphy (1975) ако су квадриране разлике означене са  $d_i^2$  за сваки  $i$ -ти пар, онда је:

$$\sum_{i=1}^n d_i^2$$

Вредност  $r_s$  се израчунава:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}$$

Ако се образац примени на податке из табеле 2, вредност  $r_s$  износи -0,90. Провером у табели вредности за Спирманову корелацију рангова, може се закључити да постоји статистички значајна веза између декадних вредности NAO индекса и декадних вредности количине падавина на ширем простору Старе планине.

Такође, покушано је да се утврди да ли на просторну расподелу падавина, поред атмосферске циркулације, може утицати и експозиција падина на простору Старе планине. Наиме, Дуцић и др. (2003) и Миловановић (2005), анализирајући податке са укупно 28 метеоролошких станица на ширем простору Старе планине, одредили су количину падавина на групама станица са различитим експозицијама и коришћењем падавинског градијента редуковали их на исту надморску висину (табела 3).

Рангирањем сваке групе станица (табела 4) израчуната је вредност Пирсоновог коефицијента корелације од 0,97. „Може се закључити да је повезаност експонираности станица и примљене количине падавина значајна, односно високо значајна и да задовољава Студентов тест и на нивоу поверења од 95% и на нивоу поверења од 99%“ (Дуцић и др, 2003).

Табела 3: Просечне годишње количине падавина на различито експонираним станицама (према Дуцић-у и др. 2003)

	Експонираност станица						
	W	SW	S	SE	NE	N	NW
Број станица	2	11	3	3	4	2	3
Просечна надморска висина	430,0	503,2	461,7	600,3	354,3	190,0	436,7
Просечна количина падавина	769,2	649,2	645,8	691,0	668,9	627,7	758,8
Просечна количина падавина на истој надморској висини	959,6	819,2	829,4	827,0	886,5	903,1	949,2

Међутим, с обзиром да су вредности за експозицију падина мерене на номиналном нивоу (укупно седам категорија), овде је примењен хи-квадрат тест. Harnett, Murphy (1975) наводе следеће: „When the set of outcomes can be divided into two categories, then the appropriate test statistic is binominal variable. When more than two categories or classes of outcomes are involved, then the appropriate statistic is the chi-square variable”. Вредност  $\chi^2$  је израчуната према следећој формули:

Табела 4: Коефицијент корелације између различито експонираних станица и примљене количине падавина (према Дуцић-у и др, 2003)

Експонираност групе станице	Задата вредност	Количина падавина (сведена на 1000 м.н.в)
W	7	959,6
NW	6	949,2
N	5	903,1
NE	4	886,5
SE	3	827,0
S	2	829,4
SW	1	819,2
Коефицијент корелације	0,97	

$$\chi^2_{(c-1)d.f} = \sum_{i=1}^c \frac{(O_i - E_i)^2}{E_i}$$

где је:

с - број категорија,

$O_i$  - број уочених фреквенција,

$E_i$  - број очекиваних фреквенција,

$H_0$ :  $E_1=E_2...=E_7=882$ ,

$H_a$ : Фреквенције нису подједнако распоређене.

Вредност  $\chi^2$  статистика износи 23,51. Провером у табlici за хи-квадрат дистрибуцију, може се закључити да је на нивоу значајности  $\alpha < 0,005$  оправдано одбацити нулту хипотезу, односно прихватити хипотезу по којој постоји статистички значајна повезаност између експонираности падина и примљене количине падавина на ширем простору Старе планине.

Пошто је употребом Спирмановог поступка утврђена статистички значајна корелација између декадних вредности НАО индекса и декадних вредности количине падавина на ширем простору Старе планине, покушали смо да дамо одговор на питање да ли у анализираним подацима постоји тренд, колике и какве су промене посматраног елемента и да ли су те промене статистички значајне у анализираним периоду (1951-2000. година). Једна од процедура погодних за детектовање и оцену тренда је и Ман-Кендалов тест. Према Salmi et al, (2002) овај тест је погодан у случајевима када постоји монотонички тренд (без сезонских и цикличких варијација), док је Сенцов метод погодан за оцену нагиба тренда. У примени ових метода се могу користити и низови у којима постоје недостајуће вредности, а подаци су независни од типа расподеле. Осим тога, Сенцов метод је слабо осетљив на појединачне грешке и аутлејере (тј. ретке, атипично екстремне вредности). Израчунава се према формули<sup>2</sup>:

$$Q_i = \frac{x_j - x_k}{j - k}$$

где је:  $j > k$ .

Сенова оцена нагиба  $Q$ , показује промену по јединици времена и заправо представља медијану  $N$  вредности  $Q_i$ , при чему се вредност  $N$  рачуна као:

<sup>2</sup> [http://www.fmi.fi/organisation/kontakt\\_11.html](http://www.fmi.fi/organisation/kontakt_11.html).



$$N = \frac{n(n-1)}{2}$$

Ако би се образац за  $Q_i$  применио на податке из табеле 2, добила би се матрица као у табели 5.

Табела 5. Матрица добијених вредности  $Q_i$  -Сенових оцена нагиба

-12,06	2,54	-37,14	-26,37
	17,14	-49,68	-31,14
		-116,50	-55,28
			5,94

Извођењем медијане из табеле 5 добија се вредност од -28,75 mm по декади.

Израчунавањем Ман-Кендаловог тест статистика  $S$  (за низове са мање од 10 чланова) може се проверити статистичка значајност тренда. Ова величина се израчунава преко обрасца:

$$S = \sum (x_j - x_k) \begin{cases} 1 \text{ if } x_j - x_k > 0 \\ 0 \text{ if } x_j - x_k = 0 \\ -1 \text{ if } x_j - x_k < 0 \end{cases}$$

Претходно приказани, као и Вилкоксон тест су у испитивању фреквенције појављивања екстремних температурних догађаја у централној и јужној Европи такође применили и дали детаљнија објашњења Domonkos et al, (2003) и нарочито Szinell et al, (2003).

Ако би се формула за тест статистик  $S$  применила на податке из табеле 2, добила би се вредност од -4. Провером у одговарајућој табели (табела 6), може се закључити да на простору Старе планине не постоји статистички значајан тренд (може се прихватити  $H_0$  - с обзиром да је присутан негативан тренд, а добијена вредност је већа од критичне вредности  $S_{\max}$  за  $\alpha = 0,2$ ) промене декадних вредности количине падавина у периоду од 1951-2000. године.

Табела 6. Вредности Ман-Кендаловог тест статистика S

Број елемената	Опсег S	$S_{\max}$
N		$\alpha=0,2$
4	-6 до 6	-4
5	-10 до 10	-5
6	-15 до 15	-6
7	-21 до 21	-7
8	-28 до 28	-8
9	-36 до 36	-9
10	-45 до 45	-10

### Закључак

Употреба статистичких поступака представља неопходност из које се могу стећи подробне и прецизне информације о повезаности између одређених климатских елемената и појава. Приказом основних карактеристика и функција појединих процедура, као и резултата добијених њиховом применом, олакшана је употреба статистичких поступака у климатолошким проучавањима. Испитивањем хомогености података (употребом Ман-Витнијевог теста) о средњим годишњим количинама падавина са 11 станица на ширем простору Старе планине, утврђена је хомогеност низова на 9 станица (изузетак су Алдинац и Дојкинци). Попуњавање недостајућих података у циљу добијања што квалитетнијих података је извршено комбиновањем три различита метода (редукција на исти број година, интерполација помоћу матрице – основна и модификована варијанта и екстраполација заснована на претпоставци константног односа краћих и дужих низова), чиме је повећана тачност метаподатака.

Због релативно малог простора на коме су анализиране падавине, као и због нехомогености изворне базе података, извршено је просторно осредњавање, чиме је добијена средња годишња (декадна) количина падавина на Старој планини у целини. Израчунавањем Пирсоновог коефицијента корелације уочена је статистички значајна веза (негативна корелација) између NAO индекса и количине падавина на посматраном простору (на годишњем нивоу  $r = -0,37$ ). Међутим, применом истог

поступка на агрегиране (декадне) вредности, упркос порасту коефицијента корелације (али уз смањење броја елемената низа, односно степени слободе), на истом нивоу значајности та веза је изгубљена ( $r = -0,81$ ). Спирманов поступак, као непараметарска алтернатива Пирсоновом коефицијенту корелације, и на декадном нивоу је показао статистички значајну везу. Употребом хи-квадрат теста, на веома високом нивоу поверења од преко 99% је утврђен утицај експозиције падина на примљену количину падавина на ширем простору Старе планине. Израчунавањем Сенове оцене нагиба тренда, установљено је смањење количине падавина од  $-28,75$  mm по декади на посматраном простору у периоду 1951-2000 година. Применом Ман-Кендаловог теста установљено је да у посматраном периоду, а на декадном нивоу, то није статистички значајна промена.

### Литература

Domonkos P, Kysel J, Piotrowicz K, Petrovic P, Likso T. (2003): Variability of extreme temperature events in South-Central Europe during the 20<sup>th</sup> century and its relationship with large scale circulation, *International Journal of Climatology* 23, p. 987-1010.

Дуцић В, Радовановић М, Миловановић Б. (2003): Просторни распоред падавина на Старој планини у зависности од надморске висине и експозиције (Regional disposition of precipitation on Stara planina to dependence from exposition and elevation), *Зборник радова Географског факултета*, св. LI, Београд, стр. 39-55.

Миловановић Б. (2005): Климатска регионализација Старе планине (Climate regionalization of mountain Stara planina), Магистарска теза, Географски факултет, Београд.

Nikolova N. (2004): Rainfall variability in Bulgaria and its relation with North Atlantic Oscillation. BALWIOS-2004, Ohrid, FY Republic of Macedonia, 25-29 May 2004 ([http://balwois.mpl.ird.fr/balwois/administration/full\\_paper/ffp-10-047.pdf](http://balwois.mpl.ird.fr/balwois/administration/full_paper/ffp-10-047.pdf))

Peterson T, Easterling D, Karl T, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Førland E, Bauer I. H, Alexandersson H, Jones P, Parker D. (1998): Homogeneity adjustment of in situ atmospheric climate data – a review, *International Journal of Climatology* 18, p. 1493-1517.

Радовановић М. (2001): Утицај рељефа и атмосферске циркулације на диференцијацију климата у Србији (Influence of relief and atmospheric circulation on climate differentiation in Serbia), Докторска дисертација, Географски факултет, Београд.

Radovanovic M, Milovanovic B. (2003): Methods to Complete the Missing Data on Precipitation in the Mountains of Serbia – Testing and Application, *Studia Geograficzne* 75, No 2542, p. 593-598, International conference "Man and climate in the 20th Century", Wroclaw, 13 – 15 June 2002. – Institute of Geography, University of Wroclaw and Polish Geophysical Society of Marshall of the Voivod of Lower Silesia, Wroclaw.

Rimbu N, Boronean C, Buta C, Dima M. (2002): Decadal variability of the Danube river flow in the lower basin and its relation with the North Atlantic Oscillation, *International Journal of Climatology* 22, p. 1169-1179.

Ристић Ж. (2006): О истраживању, методу и знању (About scientific research, method and knowledge), Институт за педагошка истраживања, Београд.

Salmi T, Määttä A, Anttila P, Airola T, Amnell T. (2002): Detecting trends of annual values of atmospheric pollutants by the Mann-Kendall test and Sen's slope estimates-the excel template application meekness, Finish Meteorological Institute, Helsinki, Finland [http://www.fmi.fi/organisation/kontakt\\_11.html](http://www.fmi.fi/organisation/kontakt_11.html).

Szinell C, Bussay A, Szentimrey T. (1998): Drought tendencies in Hungary, *International Journal of Climatology* 18, p. 1479–1491.

Harnett D, Murphy J. (1975): Introductory statistical analysis, Addison-Wesley Publishing company, INC, Reading, Massachusetts.

Hurrell, J. (1995): Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation, *Science*, no. 269, 676-679.