

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



Analýza prietoku slovenských riek pomocou neparametrických
štatistických metód

DIPLOMOVÁ PRÁCA

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

**ANALÝZA PRIETOKU SLOVENSKÝCH RIEK POMOCOU
NEPARAMETRICKÝCH ŠTATISTICKÝCH METÓD**

DIPLOMOVÁ PRÁCA

Študijný program: Ekonomicko-finančná matematika a modelovanie
Študijný odbor: 1114 Aplikovaná matematika
Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky
Vedúci práce: RNDr. Ján Pekár, PhD.



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Lucia Macháčková
Študijný program: ekonomicko-finančná matematika a modelovanie
(Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Analýza prietoku slovenských riek pomocou neparametrických štatistických metód
Non-parametric statistical analysis of Slovak rivers overflows

Anotácia: Prietok riek sa neustále mení vplyvom rôznych faktorov. Môže sa zmeniť náhle alebo dlhodobo. Tieto zmeny je nutné sledovať, aby bolo možné pripraviť sa alebo predísť katastrofám akými sú napríklad povodne. Zmeny v prietoku rieky predstavujú zmeny rozdelenia, ktorým sa rad riadi. Keďže je rozdelenie radu málokedy známe, na jeho štatistickú analýzu sa využívajú neparametrické metódy.

Cieľ: Cieľom práce je sprostredkovať prehľad metód využívaných na detekciu zmien v hydrologických časových radoch a následná aplikácia spomínaných metód na analýzu prietokov vybraných slovenských riek.

Vedúci: doc. RNDr. Ján Pekár, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.
Dátum zadania: 08.01.2019

Dátum schválenia: 08.01.2019

prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
vedúci práce

Podakovanie Rada by som podakovala vedúcemu mojej diplomovej práce RNDr. Jánovi Pekárovi, PhD. za pomoc, rady a hlavne ochotu usmerňovať moju prácu. A samozrejme ďakujem mojim rodičom a priateľom za nekonečnú podporu a trpezlivosť.

Abstrakt v štátnom jazyku

MACHÁČKOVÁ, Lucia: Analýza prietoku slovenských riek pomocou neparametrických štatistických metód [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: RNDr. Ján Pekár, PhD., Bratislava, 2020, 69 s.

Jedným z cieľov tejto diplomovej práce je oboznámenie sa s neparametrickými štatistickými metódami, ktoré sa využívajú pri analýze hydrologických časových radov. Nakoľko sa v časovom rade môžu meniť rôzne parametre, táto práca sa zaoberá zmenou ich strednej hodnoty. Pod touto zmenou chápeme nárazové zmeny, no zároveň aj zmeny v dlhodobom horizonte, ktoré predstavuje trend. Taktiež sa zaoberá aj viacnásobnými zmenami. Na analýzu takýchto zmien autorka využíva metódy aglomeratívneho a divízneho zhľukovania. Po vysvetlení teoretických poznatkov autorka aplikuje dané metódy na analýzu reálnych dát prietokov slovenských riek.

Kľúčové slová: Neparametrické štatistické metódy, Stredná hodnota, Bod zmeny, Trend, Hierarchické metódy

Abstract

MACHÁČKOVÁ, Lucia: Non-Parametric Statistical Analysis of Slovak rivers overflows [Diploma Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: RNDr. Ján Pekár, PhD., Bratislava, 2020, 69 p.

One of the main aims of this thesis is the understanding of the non-parametric statistical methods that are used to analyze hydrological time series. Although several parameters could be changed, this thesis is monitoring changes in mean of the serie. There could be included the single changes but also the long-term changes such as trend. This thesis is analyzing the multiple changes, too. For this type of changes is the author using methods of agglomerative and divisional clustering. After the explanation of the theoretical knowledge author used the given methods on the real data of the Slovak rivers' overflows.

Keywords: Non-parametric statistical methods, Mean, Change Point, Trend, Hierarchical methods

Obsah

Úvod	10
1 Homogenita v časových radoch	12
1.1 Von Neumannov test	12
1.2 Wald-Wolfowitzov test	13
2 Bod zmeny režimu	15
2.1 Štandardný test homogenity	15
2.2 Pettittov test	15
2.3 Test kumulovaných súm	17
2.4 Mann-Whitney-Wilcoxonov dvojvýberový test	17
2.5 Dvojvýberový Kolmogorov-Smirnovov test	18
3 Viacnásobné body zmeny režimu	19
3.1 E-Divisive algoritmus	19
3.2 Aglomeratívny algoritmus	21
4 Dlhodobé zmeny režimu	24
4.1 Sezónny Mann-Kendallov test	24
4.2 Cox-Stuartov test	25
5 Analýza rieky Morava	26
5.1 Homogenita	28
5.2 Bod zmeny	28
5.3 Viacnásobné body zmeny	29
5.4 Trend	33
6 Analýza rieky Hron	35
6.1 Homogenita	35
6.2 Bod zmeny	37
6.3 Viacnásobné body zmeny	38
6.4 Trend	40

7	Analýza rieky Váh	42
7.1	Homogenita	42
7.2	Bod zmeny	44
7.3	Viacnásobné body zmeny	45
7.4	Trend	46
8	Analýza rieky Belá	48
8.1	Homogenita	48
8.2	Bod zmeny	50
8.3	Viacnásobné body zmeny	51
8.4	Trend	51
9	Analýza rieky Kysuca	52
9.1	Homogenita	53
9.2	Bod zmeny	53
9.3	Viacnásobné body zmeny	55
9.4	Trend	55
10	Analýza rieky Krupinica	56
10.1	Homogenita	57
10.2	Bod zmeny	57
10.3	Viacnásobné body zmeny	59
10.4	Trend	62
11	Analýza rieky Ipeľ	64
11.1	Homogenita	64
11.2	Bod zmeny	65
11.3	Viacnásobné body zmeny	67
11.4	Trend	70
12	Analýza rieky Nitra	71
12.1	Homogenita	72
12.2	Bod zmeny	72
12.3	Viacnásobné body zmeny	74

12.4 Trend	75
13 Vzťah medzi analyzovanými radmi mesačných prietokov	76
Záver	81
Zoznam použitej literatúry	84
Príloha A	87

Úvod

Táto práca sa zaoberá analýzou hydrometeorologických časových radov pomocou neparametrických štatistických metód, ako napovedá už jej názov.

Konkrétnejšie sa budeme venovať analýze prietokov slovenských riek, ako napríklad Morava, Hron, Váh a iné.

Predpokladáme, že čitateľ disponuje základnými znalosťami v oblasti analýzy časových radov a základmi štatistiky.

Prečo sme zvolili neparametrické metódy? Dôvodov je hneď niekoľko. Asi najdôležitejším z nich je, že tieto metódy si nevyžadujú žiadne predpoklady na rozdelenie dát. Medzi ďalšie dôvody patrí napríklad aj fakt, že dokážu vypočítať relatívne presnú p-hodnotu a taktiež veľmi dobre pokrývajú intervaly spoľahlivosti. Medzi najdôležitejšie výhody však patrí aj skutočnosť, že sú častokrát jednoduché na porozumenie a na interpretáciu. Vedomosti v oblasti neparametrických štatistických metód sú značne rozsiahle. Nasvedčuje tomu aj skutočnosť, že o týchto metódach bolo napísaných viacero publikácií, ako napríklad [8]. Nakoľko sú niektoré z metód založené na usporiadaní hodnôt časového radu, budeme v práci predpokladať, že všetky hodnoty radu sú rôzne. V prípade rovnakých hodnôt sa testové štatistiky značne komplikujú.

Práve analýzou hydrologických časových radov sa zaoberá niekoľko článkov, ako napríklad [1] alebo kniha [7].

Účelom analýzy prietokov riek je poskytnutie komplexného prehľadu správania sa toku daných riek. Vďaka tomu vieme predikovať, ako sa rieka bude správať v budúcnosti, čím môžeme predísť katastrofám, ako sú povodne. Taktiež vieme odhadnúť, ako sa zmení jej prietok, ak na nej postavíme vodnú elektrárňu a aký to bude mať dopad na okolie.

Táto práca sa skladá z dvoch častí: teoretickej a praktickej. Prvé štyri kapitoly poskytujú prehľad metód použitých na analýzu. V praktickej časti vykonávame samotnú analýzu.

V prvej kapitole uvádzame metódy na zisťovanie prítomnosti zmeny v rade. Druhá kapitola obsahuje metódy zaoberajúce sa detekciou a signifikanciou jedného bodu zmeny. V ďalšej kapitole sa pozrieme na metódy používané na detekciu viacnásobných bodov zmeny. Posledná kapitola teoretickej časti sa zaoberá metódami týkajúcimi sa dlhodobých zmien v rade.

Praktická časť zahŕňa zvyšných deväť kapitol. V prvých ôsmich z nich analyzujeme rieky Morava, Hron, Váh, Belá, Kysuca, Krupinica, Ipel a Nitra. V poslednej kapitole sa zameriavame na vzťah medzi jednotlivými prietokmi riek a taktiež na porovnanie výsledkov získaných v predošlých kapitolách. Dáta týkajúce sa prietokov týchto riek pochádzajú zo Slovenského hydrometeorologického ústavu a z databázy projektu UNESCO "Flood regime of rivers in the Danube River basin" (Pekárová a Miklánek, (eds.) 2019).

1 Homogenita v časových radoch

V akejkoľvek oblasti je častokrát dôležité sledovať zmeny v pozorovaniach. Na základe týchto poznatkov je jednoduchšie predísť rôznym katastrofám. Pod spomínanými zmenami si môžeme predstaviť jednorázové zmeny, prípadne viac jednorázových zmien, ktoré sa prejavujú zmenou parametrov v rozdelení, alebo dlhodobé zmeny - trend.

Pod pojmom zmena režimu si môžeme predstaviť zmenu v rozdelení, ktorým sa daný časový rad riadi.

Inak povedané, majme časový rad $x_1, \dots, x_n$. Predpokladajme, že $x_1, x_2, \dots, x_\tau, \tau < n$ sa riadi rozdelením $F1$ a $x_{\tau+1}, x_{\tau+2}, \dots, x_n$ sa riadi rozdelením $F2$. Hovoríme, že v danom časovom rade sa nachádza nejaký bod zmeny v čase τ , ak $F1 \neq F2$. Môžeme teda povedať, že ak sa v časovom rade nachádza bod zmeny, tak je tento časový rad nehomogénny. Preto nám pri zisťovaní, či sa v danom časovom rade nachádza bod zmeny režimu, stačí testovať homogenitu časového radu, pričom vieme, resp. predpokladáme, čas zmeny. V našej práci budeme sledovať zmenu v strednej hodnote procesov.

Budeme teda testovať nulovú hypotézu H_0 oproti alternatívnej hypotéze H_1 , kde

H_0 : Časový rad je homogénny.

vs.

H_1 : V časovom rade je aspoň jeden bod zmeny režimu.

V tejto podkapitole ukážeme testy, ktoré nám povedia, či sa v danom rade nachádza nejaká zmena alebo nie. Poskytnú nám iba informáciu o prítomnosti zmeny, nie o čase jej výskytu a ani významnosti.

1.1 Von Neumannov test

Najvyužívanejším testom na testovanie (ne)homogenity v dátach je Von Neumannov test. Dôležitým predpokladom na použitie tohto testu je normálne rozdelenie dát. Práca [3] definuje testovú štatistiku tohto testu ako pomer

$$N = \frac{\sum_{i=1}^{n-1} (x_i - x_{i+1})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (1)$$

kde n je počet pozorovaní, x_i je i -té pozorovanie a $\bar{x}$ je výberový priemer. Článok [2] hovorí, že pre homogénny časový rad je očakávaná hodnota testovej štatistiky $N \approx 2$. Pre rad obsahujúci zmenu je očakávaná hodnota $N < 2$ a ak $N > 2$, časový rad má veľké výkyvy strednej hodnoty.

Bartels v článku [4] upravil Von Neumannov test na jeho rankovú verziu, aby ho bolo možné použiť aj na dáta, ktoré nie sú normálne rozdelené. Testová štatistika sa zmenila na

$$RVN = \frac{\sum_{i=1}^{m-1} (R_i - R_{i+1})^2}{\sum_{i=1}^m (R_i - \bar{R})^2} \quad (2)$$

kde R_i predstavuje rank i -tého pozorovania, $m = \max R_i$, $m \leq n$. Ranky pozorovaní určujeme z neklesajúcej postupnosti. Vyvodenie záveru testu je rovnaké ako pre nerankovú verziu.

1.2 Wald-Wolfowitzov test

Ďalší test homogenity, ktorý budeme používať, odvodili Wald a Wolfowitz v [24]. Tento test predpokladá, že bod zmeny v čase τ rozdelí časový rad na dve vzorky: $x_1, x_2, \dots, x_\tau$ a $x_{\tau+1}, x_{\tau+2}, \dots, x_n$. Podstatou tohto testu je zoradenie hodnôt oboch vzoriek do jednej neklesajúcej postupnosti. Následne sa hodnoty postupnosti rozdelia na sekvencie tak, aby v každej z nich boli hodnoty iba jednej z pôvodných vzoriek. Počet týchto sekvencií je testovou štatistikou U . Hodnotu tejto testovej štatistiky porovnávame s kritickou hodnotou Wald-Wolfowitzovho testu. Nulovú hypotézu H_0 zamietame, ak je hodnota testovej štatistiky menšia ako kritická hodnota.

Pre veľké τ a n , podľa [3] $\tau > 20, n - \tau > 20$, má testová štatistika

$$R = \frac{U - E(U)}{\sqrt{Var(U)}} \quad (3)$$

štandardizované normálne rozdelenie $N(0, 1)$. V tomto prípade porovnáme hodnotu testovej štatistiky s kritickou hodnotou normálneho rozdelenia. H_0 zamietame, ak $|R| > u_{1-\frac{\alpha}{2}}$.

2 Bod zmeny režimu

V tejto kapitole sa zameriame na tzv. náhle zmeny režimu - body zmeny režimu. V prvej časti predstavíme testy, ktoré sú popri vyhodnocovaní významnosti bodu zmeny schopné určiť presný čas, kedy zmena nastala. V ďalšej časti uvedieme testy, ktoré sa používajú v prípade, že vieme čas zmeny, ale chceme určiť jeho významnosť.

Detekcia bodu zmeny

2.1 Štandardný test homogenity

Podľa článkov [15] a [16] princíp štandardného testu homogenity (angl. Standard Normal Homogeneity Test) - SNHT spočíva v porovnaní strednej hodnoty prvých k pozorovaní so strednou hodnotou zvyšných $(n - k)$ pozorovaní. Testová štatistika tohto testu je

$$T_0 = \max T_k, \quad (4)$$

$$T_k = k \left[\frac{1}{k} \sum_{i=1}^k \frac{(X_i - \bar{X})}{S_X} \right]^2 + (n - k) \left[\frac{1}{(n - k)} \sum_{i=k+1}^n \frac{(X_i - \bar{X})}{S_X} \right]^2, 1 \leq k \leq n, \quad (5)$$

kde X_i predstavuje hodnoty radu do času i , $\bar{X}$ je výberový priemer tohto radu, S_X je výberová odchýlka X . Nulovú hypotézu H_0 zamietame, ak je hodnota testovej štatistiky T_0 väčšia ako kritická hodnota testu prínáležiaca k veľkosti dát. Pri aplikácii je p-hodnota odhadovaná pomocou Monte Carlo simulácií. Hodnota k , pre ktoré testová štatistika T_k dosiahne globálne maximum, predstavuje čas, v ktorom nastala zmena režimu.

Tento test je najvhodnejšie aplikovať na rad, v ktorom očakávame bod zmeny pri začiatku alebo konci radu.

2.2 Pettittov test

Pettittov test je jedným z najvyužívanejších testov na detekciu zmeny v procese. Je to neparametrický test využívajúci ranky, ktorý je v podstate vylepšenou verziou Mann-Whitney-Wilcoxonovho testu uvedeného v druhej časti tejto kapitoly. Ako aj ostatné

testy v tejto časti, Pettittov test dokáže určiť čas zmeny procesu. Tento test je založený na testovej štatistike

$$U_{t,n} = \sum_{i=1}^t \sum_{j=t+1}^n \text{sgn}(x_j - x_i), 1 \leq t \leq n \quad (6)$$

kde

$$\text{sgn}(x) = \begin{cases} 1, & \text{ak } x > 0 \\ 0, & \text{ak } x = 0 \\ -1, & \text{ak } x < 0 \end{cases} \quad (7)$$

V prípade dát so spojitou distribučnou funkciou je testová štatistika

$$U_{t,n} = 2 \sum_{i=1}^t R_i - t(n+1), \quad (8)$$

kde R_i je rank i-tého pozorovania. Podľa Pettittovho článku [19] môžeme vzťah (8) nahradiť

$$U_{t,n} = U_{t-1,n} + V_{t,n}, \quad t = 2, \dots, n, \quad (9)$$

kde

$$V_{t,n} = \sum_{i=1}^n \text{sgn}(x_i - x_t), \quad U_{1,n} = V_{1,n}. \quad (10)$$

Najpravdepodobnejší čas zmeny τ je taký, pre ktorý platí

$$K_\tau = |U_{\tau,n}| = \max |U_{t,n}|. \quad (11)$$

Signifikancia určená p-hodnotou vzťahujúca sa ku K_τ je daná vzťahom

$$p = 2 \exp \left(\frac{-6K_\tau^2}{n^2 + n^3} \right). \quad (12)$$

Hypotézu H_0 zamietame, ak $p < \alpha$, kde α je hranica významnosti. V tom prípade môžeme povedať, že x_τ je bodom zmeny a τ je čas, v ktorom zmena nastala. Všimnime si však, že za určitých podmienok môže p-hodnota Pettittovho testu nadobúdať hodnoty väčšie ako 1. V takom prípade považujeme výsledky tohto testu za nedôveryhodné.

2.3 Test kumulovaných súm

Výhodou testov kumulovaných súm (angl. Cumulative Sum Test) - CUSUM je, že dokážu odhaliť náhle zmeny v strednej hodnote procesu. Ich podstatou sú upravené čiastočné súčty

$$S_k^{**} = \frac{1}{\sqrt{n}} \sum_{i=1}^k \frac{(X_i - \bar{X})}{S_X}, 1 \leq k \leq n \quad (13)$$

kde $\bar{X}$ je výberový priemer a S_X je výberová štandardná odchýlka. Hodnota S_k^{**} homogénneho radu sa pohybuje v okolí 0. Lokálny extrém S_k^{**} predstavuje zmenu v časovom rade, t.j. k , pre ktoré je hodnota S_k^{**} extrémálna, predstavuje čas zmeny. Článok [1] uvádza, že z S_k^{**} boli odvodené ďalšie dve testové štatistiky

$$Q = \max |S_k^{**}| \quad (14)$$

$$R = \max(S_k^{**}) - \min(S_k^{**}) \quad (15)$$

Štatistika Q sa používa na detekciu jedného bodu zmeny, pričom použitie štatistiky R je efektívne pri zisťovaní dvoch opačných bodov zmeny. Napríklad pri posune strednej hodnoty smerom nadol dosiahne S_k^{**} maximum, ktoré následne vedie k posunu strednej hodnoty nadol, a teda S_k^{**} dosiahne minimum.

Signifikancia bodu zmeny

V praxi sa nám však môže stať, že možný čas zmeny poznáme. V tejto časti kapitoly uvedieme testy, ktoré je možné použiť na testovanie signifikancie zmeny, ktorá nastala v danom čase.

2.4 Mann-Whitney-Wilcoxonov dvojvýberový test

Neparametrickou analógiou Studentovho dvojvýberového t-testu je Mann-Whitney-Wilcoxonov test. Tento test okrem rozdelenia časového radu na dve časti nevyžaduje žiadne ďalšie predpoklady.

Princíp testu spočíva v zoradení hodnôt do neklesajúcej postupnosti a následnom priradení rankov R_i každej hodnote. Testová štatistika je daná

$$U = \min(U_1, U_2), \quad (16)$$

kde

$$U_1 = \tau(n - \tau) + \frac{\tau(\tau + 1)}{2} - \sum_{i=1}^{\tau} R_i, \quad (17)$$

$$U_2 = \tau(n - \tau) + \frac{(n - \tau)(n - \tau + 1)}{2} - \sum_{i=\tau+1}^n R_i. \quad (18)$$

Stredná hodnota testovej štatistiky U je

$$\mu_U = \frac{\tau(n - \tau)}{2} \quad (19)$$

a jej smerodajná odchýlka

$$\sigma_U = \sqrt{\frac{[\tau(n - \tau)](n + 1)}{12}}. \quad (20)$$

Hypotézu H_0 zamietame, ak je hodnota testovej štatistiky U menšia ako kritická hodnota Mann-Whitney-Wilcoxonovho testu.

Podľa [18] ak oba výbery obsahujú viac ako 8 pozorovaní, štandardizáciou štatistiky U získame štatistiku Z , ktorá sa riadi normálnym rozdelením $N(0, 1)$

$$Z = \frac{U - \frac{\tau(n - \tau)}{2}}{\sqrt{\frac{[\tau(n - \tau)](n + 1)}{12}}}. \quad (21)$$

V tomto prípade teda porovnáваме štatistiku Z s kritickou hodnotou normálneho rozdelenia.

2.5 Dvojvýberový Kolmogorov-Smirnovov test

V tomto teste opäť predpokladáme rozdelenie časového radu bodom zmeny x_τ na dva samostatné rady. Ďalej predpokladajme, že rad $x_1, \dots, x_\tau$ sa riadi rozdelením s distribučnou funkciou F_1 a rad $x_{\tau+1}, \dots, x_n$ sa riadi rozdelením s distribučnou funkciou F_2 . Podľa [5] je testová štatistika tohto testu

$$D = \max_{-\infty < x < \infty} |F_1(x) - F_2(x)|. \quad (22)$$

Hodnota testovej štatistiky sa porovná s kritickou hodnotou Kolmogorov-Smirnovovho testu s parametrami τ a $(n - \tau)$. Ak je hodnota testovej štatistiky D väčšia alebo rovná kritickej hodnote $D_{\tau, (n - \tau)}$, nulovú hypotézu H_0 zamietame.

3 Viacnásobné body zmeny režimu

Väčšina algoritmov zaoberajúcich sa detekciou viacnásobných bodov zmeny je založená na princípe bisekcie. Pomocou metód uvedených v predchádzajúcej kapitole je nájdený jeden bod zmeny, ktorý rozdelí rad na dva samostatné rady. V nich sú opäť použité metódy na nájdenie jedného bodu zmeny, ktorý znovu rozdelí testovaný rad. Tento princíp sa iteračne opakuje dovtedy, kým v rade nie sú nájdené žiadne zmeny alebo nie je homogénny.

V tejto kapitole však uvádzame aj iné metódy na detekciu viacnásobných bodov zmeny v časových radoch.

3.1 E-Divisive algoritmus

James v práci [14] predstavuje neparametrickú hierarchickú metódu na detekciu viacnásobných bodov zmeny, založenú na spojení bisekcie a viacrozmernej divergenčnej miery, nazývanú E-Divisive.

Uvažujme $Y_n = \{Y_i : i = 1, \dots, n\}$ a $Z_m = \{Z_j : j = 1, \dots, m\}$ nezávislé rovnako rozdelené výbery, pričom $Y_i, Z_j \in \mathbb{R}$. Predpokladajme, že platí $E|X|^\gamma, E|Y|^\gamma < \infty$ pre nejaké $\gamma \in (0, 2)$. Potom empirická divergenčná miera je podľa [14] daná

$$\hat{\mathcal{E}}(Y_n, Z_m; \gamma) = \frac{2}{mn} \sum_{i=1}^n \sum_{j=1}^m |Y_i - Z_j|^\gamma - \binom{n}{2}^{-1} \sum_{1 \leq i < k \leq n} |Y_i - Y_k|^\gamma - \binom{m}{2}^{-1} \sum_{1 \leq j < k \leq m} |Z_j - Z_k|^\gamma. \quad (23)$$

Ak by sa oba výbery riadili rovnakým rozdelením, potom by platilo $\hat{\mathcal{E}}(Y_n, Z_m; \gamma) \approx 0$. V opačnom prípade $\hat{\mathcal{E}}(Y_n, Z_m; \gamma) > 0$.

Pri detekcii bodu zlomu sa využíva škálovaná miera divergencie daná

$$\hat{\mathcal{Q}}(Y_n, Z_m; \gamma) = \frac{mn}{m+n} \hat{\mathcal{E}}(Y_n, Z_m; \gamma). \quad (24)$$

Predpokladajme, že $x_1, \dots, x_T$ sú nezávislé pozorovania a $1 \leq \tau \leq \kappa \leq T$ sú kladné konštanty. Rozdelíme rad $x_1, \dots, x_T$ na dva rady $Y_\tau = \{x_1, \dots, x_\tau\}$ a $Z_\tau(\kappa) = \{x_{\tau+1}, \dots, x_\kappa\}$. Bod zmeny sa potom nachádza v čase $\hat{\tau}$, ktorý vyhovuje

$$(\hat{\tau}, \hat{\kappa}) = \arg \max_{(\tau, \kappa)} \hat{\mathcal{Q}}(Y_n, Z_m(\kappa); \gamma) \quad (25)$$

Ak je známe, že v rade je iba jeden bod zmeny, potom $\kappa = T$. Ak však očakávame viac bodov, potom je vhodné považovať κ za neznámu. Dôvodom je fakt, že ak by sme

považovali $\kappa = T$, je možné, že v prípade viacerých zmien by bolo výsledné rozdelenie $Z_\tau(T)$ zmesou rôznych rozdelení. Toto $T_\tau(T)$ by tak bolo neodlíšiteľné od rozdelenia Y_τ , pričom by τ bol skutočným bodom zmeny.

Na nájdenie viacerých bodov využijeme vyššie opísanú metódu iteračne. Predpokladajme, že sme našli $k-1$ bodov zmeny v časoch $\hat{\tau}_1, \dots, \hat{\tau}_{k-1}$ takých, že platí $0 < \hat{\tau}_1 < \dots < \hat{\tau}_{k-1} < T$. Tieto body nám rozdelia náš pôvodný rad $x_1, \dots, x_T$ do k zhlukov $\hat{C}_1, \dots, \hat{C}_k$ tak, že $\hat{C}_i = \{x_{\hat{\tau}_{i-1}+1}, \dots, x_{\hat{\tau}_i}\}$, pričom $\hat{\tau}_0 = 0$ a $\hat{\tau}_k = T$. Následne v každom zhluke nájďme bod zmeny pomocou (25). O tom, ktorý z potenciálnych bodov je skutočným bodom zmeny rozhodne

$$i^* = \arg \max_{i \in \{1, \dots, k\}} \hat{Q}(Y_{\hat{\tau}_i}, Z_{\hat{\tau}_i}(\hat{\kappa}(i)); \gamma), \quad (26)$$

kde $\hat{\tau}_i$ je potenciálny čas zmeny nájdený v i -tom zhluke, $\hat{\kappa}(i)$ konštanta prináležiaca k $\hat{\tau}_i$, $Y_{\hat{\tau}_i}$ a $Z_{\hat{\tau}_i}$ časti radu v zhluke i rozdelené pomocou $\hat{\tau}_i$. Nájdením optimálneho i^* teda určíme, ktorý z potenciálnych bodov je k -tým bodom zmeny celého radu.

O významnosti nájdeného bodu rozhodne testová štatistika daná

$$\hat{q}_k = \hat{Q}(Y_{\hat{\tau}_k}, Z_{\hat{\tau}_k}(\hat{\kappa}(k)); \gamma), \quad (27)$$

kde $\hat{\tau}_k = \hat{\tau}_{i^*}$ a $\hat{\kappa}(k) = \hat{\kappa}_{i^*}$.

Veľké hodnoty $\hat{q}_k$ vedú k významnej zmene rozdelenia. Problémom však je, že na určenie skutočnej kritickej hodnoty potrebujeme vedieť, podľa akého rozdelenia sa daný rad riadi. James v [14] navrhuje využitie permutačného testu na určenie významnosti $\hat{q}_k$.

Nulovou hypotézou permutačného testu bude prípad, že v rade sa nachádza iba pôvodne nájdených $k-1$ bodov zmeny, teda žiadny ďalší tam už nie je. Podstatou permutačného testu je premiešanie hodnôt v každom zhluke a vytvorenie novej postupnosti s dĺžkou T . V novom rade sa opäť aplikuje metóda hľadania nového bodu opísaná vyššie. Tento proces sa iteruje R -krát, pričom si ukladáme hodnotu $\hat{q}_k^{(r)}$ každej iterácie. Po R permutáciách môžeme p -hodnotu aproximovať

$$p - \text{hodnota} \approx \frac{|\{r : \hat{q}_k^{(r)} \geq \hat{q}_k\}|}{R + 1}. \quad (28)$$

Voľba počtu permutácií R závisí od rozdelenia radu a taktiež od počtu a veľkosti zhlukov. Zmena v čase $\hat{\tau}_k$ je signifikantná na hladine významnosti 5%, ak aproximovaná p -hodnota

je menšia ako 0.05. V opačnom prípade nemôžeme zamietnuť nulovú hypotézu, že v rade je iba $k - 1$ bodov zmeny.

Pre lepšie pochopenie tohto algoritmu uvádzame príklad zobrazený na Obr. 1. Na začiatku máme jeden veľký zhuk - rad pozorovaní $C(123456)$. Pomocou (25) sme rozdelili pôvodný zhuk na dva zhuky: $C(12345)$ a $C(6)$.

V druhom kroku iterácie hľadáme zmeny v každom zo vstupných zhukov $C(12345)$ a $C(6)$. Pomocou (26) sme určili, že rozdelíme zhuk $C(12345)$ na zhuky $C(123)$ a $C(45)$.

Do tretieho kroku nám tak vstupujú zhuky $C(123)$, $C(45)$ a $C(6)$. Opäť v každom z nich hľadáme body zmeny. Opäť podľa (26) určíme, že sa rozdelí $C(123)$ na $C(1)$ a $C(23)$.

Štvrtý krok začíname so zhukmi $C(1)$, $C(23)$, $C(45)$ a $C(6)$. Opakujeme predošlý postup a zistíme, že sa tentokrát rozdelí zhuk $C(45)$ na zhuky $C(4)$ a $C(5)$.

Do posledného kroku nám vstupujú zhuky $C(1)$, $C(23)$, $C(4)$, $C(5)$, $C(6)$. Tu sa nám môže rozdeliť iba zhuk $C(45)$ na samostatné zhuky $C(4)$ a $C(5)$. Tento krok uzatvára celú iteračnú metódu, keďže sme už rozbili náš pôvodný rad na samostatné pozorovania.

Tento postup aplikujeme niekoľkokrát na permutovaný rad. Pomocou (28) vypočítame p-hodnotu každého nájdeného bodu. Naším záverom tak môže v tomto prípade byť, že za signifikantné body zmeny budeme považovať pozorovania 3 a 5.

3.2 Aglomeratívny algoritmus

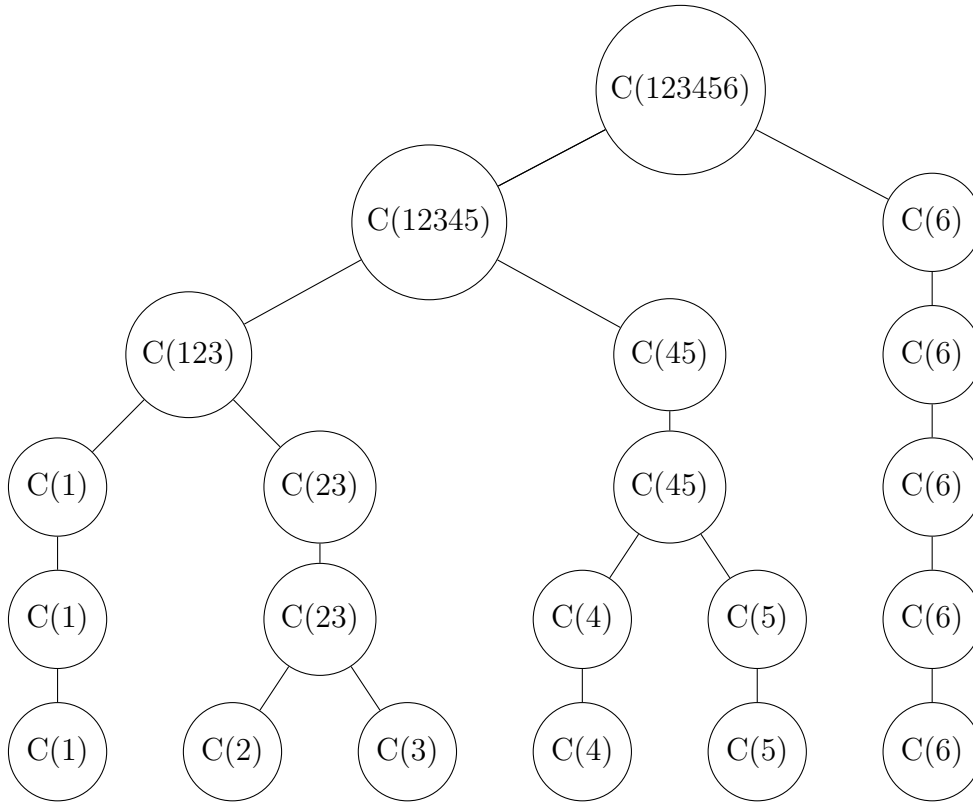
James v [14] vysvetľuje ďalšiu hierarchickú metódu - aglomeratívne zhukovanie. Táto metóda je opakom divízneho algoritmu opísaného v časti 3.1. Ďalšou odlišnosťou je aj fakt, že zachováva časové poradie hodnôt. Jej výhodou oproti E-Disisive je výpočtový čas - aglomeratívny algoritmus beží rýchlejšie. Táto metóda využíva štatistiku dobrej zhody (angl. goodness-of-fit statistics). Najskôr teda ukážeme spôsob výpočtu tejto štatistiky.

Predpokladajme množinu zhukov $\mathcal{C} = \{C_1, \dots, C_n\}$. Štatistika dobrej zhody je daná ako

$$\hat{\mathcal{S}}_n(C; \gamma) = \sum_{i=1}^{n-1} \hat{\mathcal{Q}}(C_i, C_{i+1}; \gamma), \quad (29)$$

kde C_i a C_{i+1} sú susedné zhuky a $\hat{\mathcal{Q}}$ funkcia definovaná pomocou (23).

Aglomeratívna metóda má rovnaké predpoklady na pozorovania ako E-Disisive, a to nezávislosť $x_1, \dots, x_T$ a konečnosť γ absolútneho momentu pre nejaké $\gamma \in (0, 2)$.



Obr. 1: Princíp divízneho zhukovania

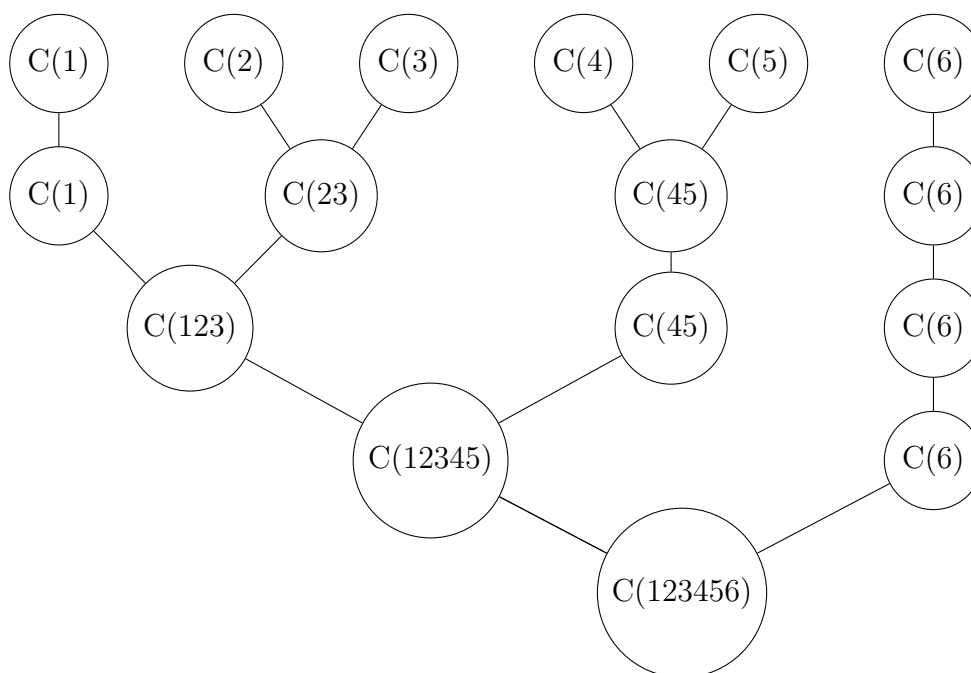
Predpokladajme, že na začiatku máme dáta rozdelené do n zhukov $\mathcal{C} = \{C_1, \dots, C_n\}$, t.j. v každom zhuku je jedno pozorovanie. Podstatou tohto algoritmu je spájanie susedných zhukov, pričom požiadavka na susednosť zhukov je nevyhnutná kvôli dodržaniu poradia pozorovaní. Matematicky túto susednosť môžeme vyjadriť: *Nech $C_i = \{x_k, x_{k+1}, \dots, x_{k+t}\}$ a $C_j = \{x_l, x_{l+1}, \dots, x_{l+s}\}$, potom zhuky C_i a C_j môžeme spojiť iba vtedy, ak $k + t + 1 = l$ alebo $l + s + 1 = k$.*

To, ktoré zhuky môžeme spojiť, nám povie štatistika dobrej zhody. Budeme spájať tie susedné zhuky, ktorých spojenie bude viesť k najväčšiemu nárastu, resp. najmenšiemu poklesu hodnoty štatistiky. Tento postup opakujeme dovtedy, kým sa nevytvorí jeden veľký zhuk.

Počet bodov zmeny potom odhadneme tým zhukovaním, pri ktorom je štatistika dobrej zhody maximálna.

Pre lepšie pochopenie tejto metódy uvádzame príklad znázornený na Obr. 2.

Na začiatku máme 6 pozorovaní, pričom každé z nich pokladáme za jeden zhuk, máme teda zhuky $C(1), \dots, C(6)$. V prvom kroku iterácie zistíme, že najväčší nárast štatistiky



Obr. 2: Princíp aglomeratívneho zhukovania

(29) sme zaznamenali pri spojení $C(2)$ s $C(3)$ a $C(4)$ s $C(5)$. Tieto hluky pospájame.

Do druhého kroku iterácie nám vstupujú 4 zhluky: $C(1)$, $C(23)$, $C(45)$, $C(6)$. Tu sa nám spoja iba prvé dva zhluky $C(1)$ a $C(23)$ do zhluku $C(123)$.

V treťom kroku začíname s tromi zhlukmi: $C(123)$, $C(45)$, $C(6)$. Zistíme, že v tomto kroku je optimálne opäť spojiť prvé dva zhluky do zhluku $C(12345)$.

Do poslednej iterácie nám vstupujú už iba 2 zhluky: $C(12345)$ a $C(6)$. Nezostáva nám už žiadna iná možnosť spojenia, ako zlúčiť tieto dva zhluky do zhluku, ktorý obsahuje všetky pozorovania - $C(123456)$.

V každej iterácii si zapamätáme optimálnu hodnotu štatistiky (29). V našom príklade predpokladajme, že spomínaná štatistika bola maximálna v druhej iterácii. To znamená, že optimálne rozdelenie pozorovaní bolo $C(123)$, $C(45)$, $C(6)$. Záverom v tomto príklade by bolo, že pozorovania 1, 2, 3 sa riadia jedným rozdelením, pozorovania 4, 5 druhým a pozorovanie 6 tretím rozdelením.

4 Dlhodobé zmeny režimu

Posledná kapitola teoretickej časti sa zaoberá testami, ktorými budeme testovať dlhodobé zmeny v časovom rade. Pod pojmom dlhodobé zmeny si môžeme predstaviť trend.

4.1 Sezónny Mann-Kendallov test

Mann-Kendallov test je najčastejšie používaným na detekciu trendu v hydrologických dátach. Tento test využíva Kendallov korelačný koeficient τ . Najskôr teda vysvetlíme, čo tento koeficient predstavuje a ako ho vypočítame.

[7] popisuje τ (korelačný koeficient) ako spôsob merania sily monotónneho vzťahu medzi dvomi premennými. Je založený na sezónnej štatistike

$$S_g = \sum_{i=1}^{n-1} \sum_{j=k+1}^n \text{sgn}(X_{jg} - X_{ig}), \quad g = 1, \dots, m, \quad (30)$$

kde g predstavuje poradové číslo "sezóny", t.j. v prípade mesačných dát $m = 12$. Potom Kendallov $\hat{S}$ -štatistika celého radu je

$$\hat{S} = \sum_{g=1}^m S_g. \quad (31)$$

Máme $\frac{n(n-1)}{2}$ možných porovnávaní. Kendallov korelačný koeficient τ je potom daný

$$\tau = \frac{\hat{S}}{\frac{n(n-1)}{2}}. \quad (32)$$

Pre τ platia rovnaké vyhodnotenia vzťahu medzi premennými ako pri jednoduchom korelačnom koeficiente, t.j. ak $\tau = 0$, premenné sú nezávislé, ak $\tau \in (0, 1]$, premenné sú priamoúmerne závislé a pre $\tau \in [-1, 0)$ premenné sú nepriamoúmerne závislé.

Testová štatistika Mann-Kendallovho testu je v podstate štatistika $\hat{S}$ Kendallovho korelačného koeficientu. Podľa [6], keď je $n \geq 10$, môžeme štatistiku $\hat{S}$ aproximovať normálnym rozdelením. Stredná hodnota štatistiky S je $\mu_{\hat{S}} = 0$ a jej odchýlka

$$\sigma_{\hat{S}} = \sqrt{\sum_{g=1}^m \frac{n_g(n_g - 1)(2n_g + 5) - \sum_{p=1}^q t_{gp}(t_{gp} - 1)(2t_{gp} + 5)}{18}}, \quad (33)$$

kde q je počet skupín s rovnakou hodnotou a t_p je počet hodnôt v p -tej skupine. Štandar-

dizovaná testová štatistika je potom

$$Z = \begin{cases} \frac{\hat{S}-1}{\sigma_{\hat{S}}}, & \text{ak } \hat{S} > 0 \\ 0, & \text{ak } \hat{S} = 0 \\ \frac{\hat{S}+1}{\sigma_{\hat{S}}}, & \text{ak } \hat{S} < 0 \end{cases} \quad (34)$$

Hodnotu testovej štatistiky Z porovnávame s kritickou hodnotou štandardizovaného normálneho rozdelenia. Nulovú hypotézu H_0 zamietame, ak $|Z| > u_{1-\frac{\alpha}{2}}$.

4.2 Cox-Stuartov test

Podstatou Cox-Stuartovho testu je porovnanie prvej a druhej polovice radu. Logicky, v prípade procesu bez trendu by mali byť hodnoty v celom rade približne rovnaké. Ak sú však hodnoty prvej polovice väčšie ako hodnoty druhej, v rade je zjavne klesajúci trend. Naopak, ak sú hodnoty prvej polovice menšie ako druhej, v rade je rastúci trend.

Vykonanie testu si vyžaduje výpočet diferencií

$$d_1 = x_{c+1} - x_1$$

$$d_2 = x_{c+2} - x_2$$

$$\vdots$$

$$d_c = x_n - x_{n-c},$$

kde $c = \frac{n}{2}$, ak n je párne, a $c = \frac{n+1}{2}$, ak n je nepárne.

Hodnotou testovej štatistiky Cox-Stuartovho testu je počet kladných diferencií.

Matematicky to môžeme formulovať nasledovne: Vyberme z diferencií $d_1, d_2, \dots, d_c$ iba kladné z nich. Označme ich $dk_1, dk_2, \dots, dk_m$. Testovú štatistiku dostaneme ako

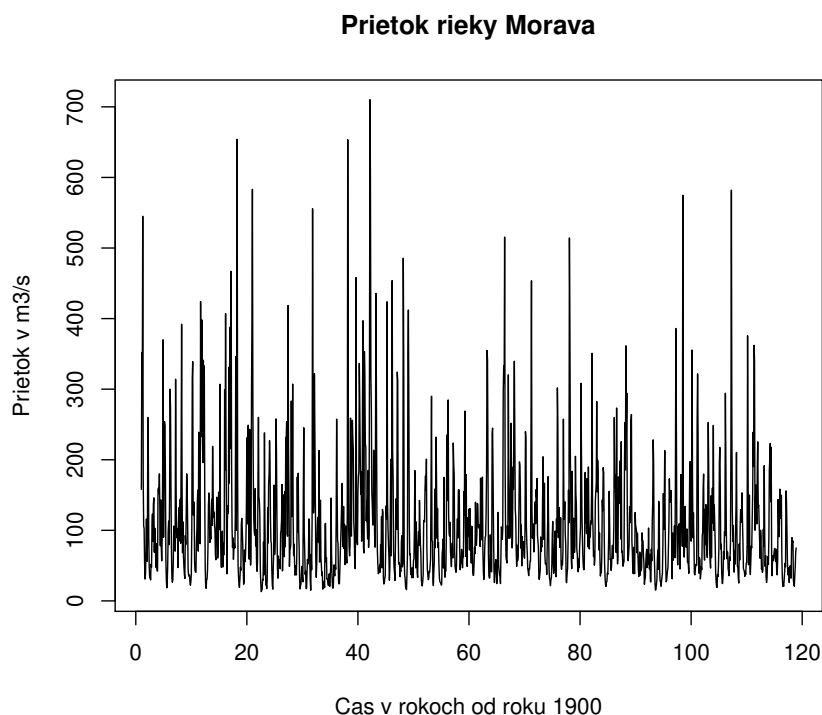
$$T = \sum_{i=1}^m \text{sgn}(dk_i). \quad (35)$$

Nulovú hypotézu H_0 zamietame, ak $|T| > t_{1-\frac{\alpha}{2}}$, kde $t_{1-\frac{\alpha}{2}}$ je 97.5%-ný kvantil binomického rozdelenia pri uvažovaní hladiny významnosti $\alpha = 5\%$. Podľa [17], ak platí $m \geq 20$, $t_{1-\frac{\alpha}{2}}$ môžeme aproximovať $t_{1-\frac{\alpha}{2}} = \frac{(m+u_{1-\frac{\alpha}{2}})\sqrt{m}}{2}$, kde $u_{1-\frac{\alpha}{2}}$ je 97.5%-ný kvantil normálneho rozdelenia.

5 Analýza rieky Morava

Podľa [21] rieka Morava pramení v Karlickom Snežníku v Českej republike. Táto rieka má dĺžku $328,9\text{km}$, z čoho 114km sa nachádza na území Slovenska, a jej povodie má rozlohu približne 26580km^2 . Vytvára prirodzenú hranicu medzi Českom, Slovenskom a Rakúskom. Pod hradom Devín sa vlieva do Dunaja.

Nasledujúci graf ukazuje hodnoty mesačných prietokov rieky Morava v m^3/s . Tieto hodnoty boli namerané na stanici v Moravskom Sv. Jáne (okres Senica) za obdobie od januára 1900 do decembra 2017. Vývoj strednej hodnoty tohto radu je zobrazený na Obr. 4.

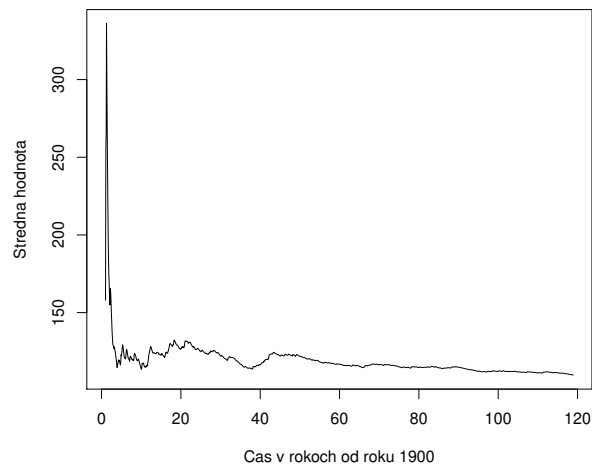


Obr. 3: Prietok Moravy od januára 1900 do decembra 2017

Obr. 4 zobrazuje strednú hodnotu prietoku Moravy za spomínané obdobie.

Z pohľadu na vývoj dát usudzujeme, že rad má aditívny charakter. Každé pozorovanie x_i tohto časového radu tak vieme zapísať ako súčet

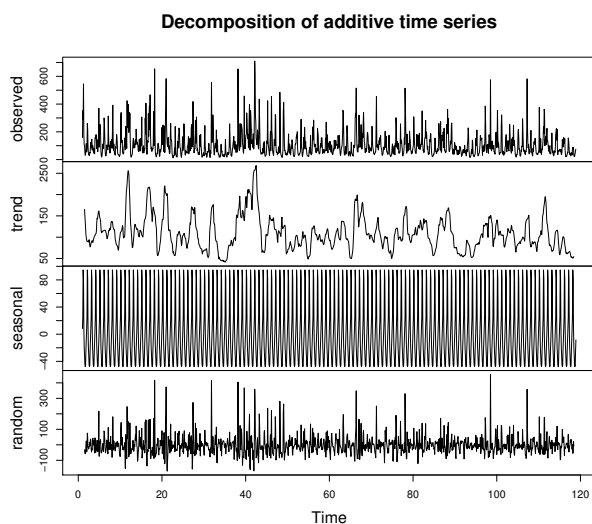
$$x_i = T_i + S_i + \epsilon_i, \quad (36)$$



Obr. 4: Stredná hodnota prietoku Moravy.

kde T_i je trend v čase i , S_i je sezónna zložka v čase i a ϵ_i je náhodná zložka v čase i , teda biely šum.

Graf na Obr. 5 ukazuje dekompozíciu nášho radu na vyššie spomínané zložky. Tento graf zachytáva mesačné hodnoty a na osi x je zobrazený čas v rokoch.



Obr. 5: Dekompozícia prietoku Moravy.

Keď vidíme, z akých zložiek sa skladá náš rad, môžeme začať s analýzou.

5.1 Homogenita

Najskôr sme vykonali testy homogenity, či sa vôbec nejaké zmeny v rade nachádzajú. Použili sme rankovú verziu Von Neumannovho testu a Wald-Wolfowitzov test. Výsledky zobrazuje Tabuľka 1.

<i>Test</i>	<i>p – hodnota</i>
Rank Von Neumann (Bartels)	$< 2.2 \cdot 10^{-16}$
Wald-Wolfowitz	$< 2.2 \cdot 10^{-16}$

Tabuľka 1: Testy homogenity prietoku Moravy.

Nakoľko pri oboch testoch je p-hodnota menšia ako 5%, hypotézu H_0 hovoriacu, že rad je homogénny, teda nie je v ňom bod zmeny, zamietame.

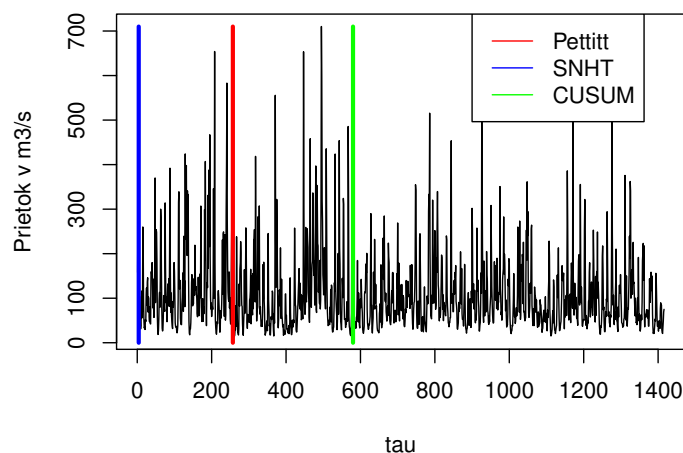
5.2 Bod zmeny

Čas, pri ktorom sa zmenila stredná hodnota procesu, budeme hľadať pomocou SNHT, Pettittovho testu a CUSUM testu. Výsledky týchto testov sú uvedené v Tabuľke 2. Keďže CUSUM neurčuje signifikantnosť zmeny, pri tejto metóde neuvádzame p-hodnotu. Okrem p-hodnoty Tabuľka 2 zobrazuje tiež poradové číslo pozorovania, pri ktorom zmena nastala a zároveň aj jeho čas.

<i>Test</i>	<i>p-hodnota</i>	τ	<i>Mesiac a rok</i>
SNHT	10^{-4}	4	Apríl 1900
Pettitt	0.01247	257	Máj 1921
CUSUM	-	580	Apríl 1948

Tabuľka 2: Detekcie jedného bodu zmeny prietoku Moravy.

SNHT označil za zmenu apríl 1900, pričom túto zmenu pokladá za signifikantnú, keďže má p-hodnotu nižšiu ako 5%. Pettittov test zasa za bod zmeny označil máj 1921 a taktiež je tento bod preňho signifikantný. Podľa CUSUM sa zmena stala v apríli 1948. Všetky tri body sú zobrazené na grafe na Obr. 6.



Obr. 6: Body zmeny nájdené pomocou SNHT, Pettittovho testu a CUSUM.

Významnosť všetkých troch bodov zmeny sme otestovali Mann-Whitney-Wilcoxonovým a Kolmogorov-Smirnovovým testom. Ich p-hodnoty sú uvedené v Tabuľke 3.

<i>Test</i>	Apríl 1900	Máj 1921	Apríl 1948
Mann-Whitney-Wilcoxon	0.002838	$3.566 \cdot 10^{-5}$	0.009959
Kolmogorov-Smirnov	0.01064	0.0001136	0.003857

Tabuľka 3: Testy signifikantnosti bodov pre prietok Moravy.

Oba testy zamietli homogenitu radu pri použití všetkých troch bodov zmeny. To znamená, že všetky tri body, t.j. apríl 1900, máj 1921 aj apríl 1948 môžeme považovať za body zmeny.

5.3 Viacnásobné body zmeny

Ďalšou časťou našej analýzy bolo hľadanie viacnásobných bodov zmeny. Toto hľadanie sme vykonávali najskôr pomocou oboch hierarchických metód, t.j. metódou E-Disisive aj aglomeratívnou metódou.

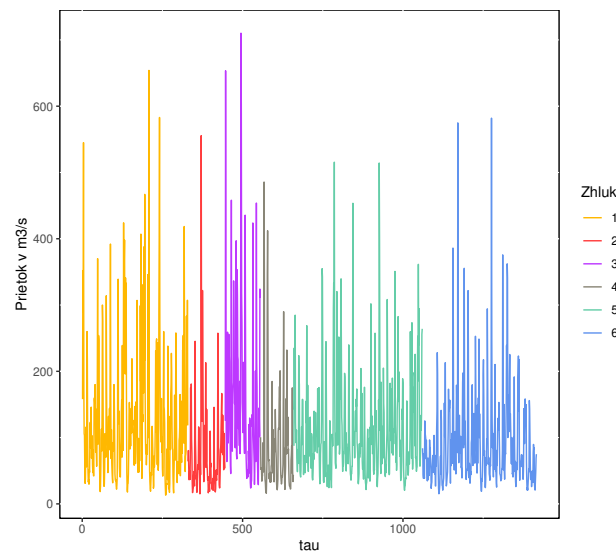
Divíznou metódou bolo nájdených päť bodov zmeny, čiže rad bol rozdelený do šiestich zhlukov, pričom každý zhluk sa riadi vlastným rozdelením.

τ	p -hodnota	Mesiac a rok
329	0.001996008	Máj 1927
445	0.001996008	Január 1937
555	0.001996008	Marec 1946
659	0.037924152	November 1954
1060	0.019960080	Apríl 1988

Tabuľka 4: Viacnásobné zmeny nájdené pomocou E-Divisive.

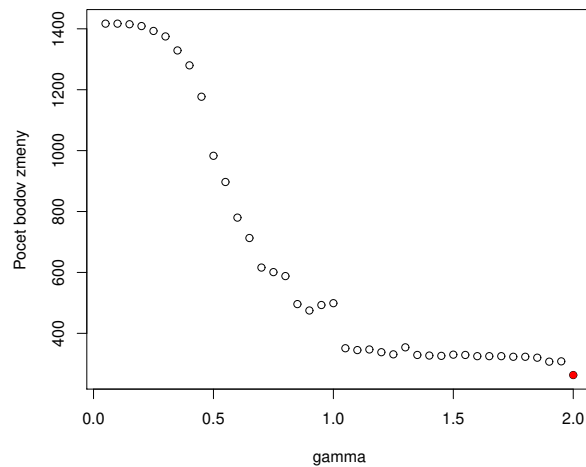
Musíme poznamenať, že algoritmus našiel ešte jeden bod, a to v $\tau = 1165$, teda v januári 1997. Avšak jeho p – hodnota je 0.129740519, čo je na hladine významnosti 5% štatisticky nevýznamný bod.

Graf na Obr. 7 zachytáva, ako nájdené body rozdeľujú náš pôvodný rad. Pozorovania zvýraznené rovnakou farbou sa riadia rovnakým rozdelením.



Obr. 7: Rozdelenie radu nájdenými bodmi pomocou E-Divisive metódy.

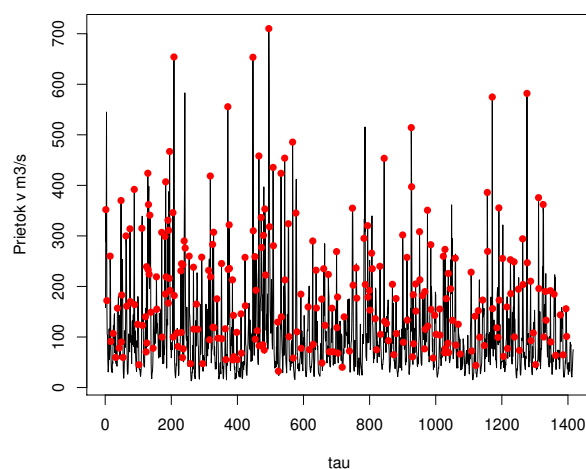
Vstupným parametrom aglomeratívneho algoritmu je parameter γ definovaný v časti 3.2. Tento parameter môže nadobúdať hodnoty z intervalu $(0; 2]$ (použitá funkcia povoľuje aj použitie krajnej hodnoty 2), pričom predstavuje momentový index, ktorý rozhoduje o vzdialenosti medzi zhlukmi. Obr.8 zachytáva závislosť počtu nájdených bodov od hodnoty parametra γ .



Obr. 8: Počet nájdených bodov zmeny pomocou aglomeratívneho algoritmu v závislosti od hodnoty γ .

Vidíme, že počet bodov zmeny pre tento rad klesá s narastajúcou hodnotou parametra γ . Nakoľko táto procedúra zjavne našla až priveľké množstvo bodov zmeny pre akýkoľvek parameter z vyššie spomínaného intervalu, zameriame sa na ten, ktorý priniesol najmenší počet bodov. V tomto prípade je to $\gamma = 2$.

Pre takúto hodnotu parametra aglomeratívna metóda našla 263 bodov zmeny. Keďže sme našli veľa bodov, nebudeme tieto body uvádzať menovite, zachytili sme ich na Obr.9. Kvôli prehľadnosti sme ich označili červenými bodmi namiesto zvislých čiar.



Obr. 9: Rozdelenie radu nájdenými bodmi pomocou aglomeratívnej metódy.

Vidíme, že nájdených bodov je príliš veľa, pričom sú relatívne husto pri sebe. To znamená, že v jednom zhluku sú v niektorých prípadoch iba dve pozorovania.

Viacnásobné body zmeny sme hľadali aj pomocou iteračného aplikovania Pettittovho testu a SNHT.

Pettittov test sme aplikovali celkovo 16-krát, pričom sme našli 8 bodov zmeny. Tie sú zaznamenané v Tabuľke 5. Môžeme si všimnúť, že bod v čase $\tau = 1061$ bol nájdený aj pomocou divíznej metódy.

τ	<i>Mesiac a rok</i>
257	Máj 1921
388	Apríl 1932
431	November 1935
468	December 1938
510	Jún 1942
778	Október 1964
1061	Máj 1988
1139	November 1994

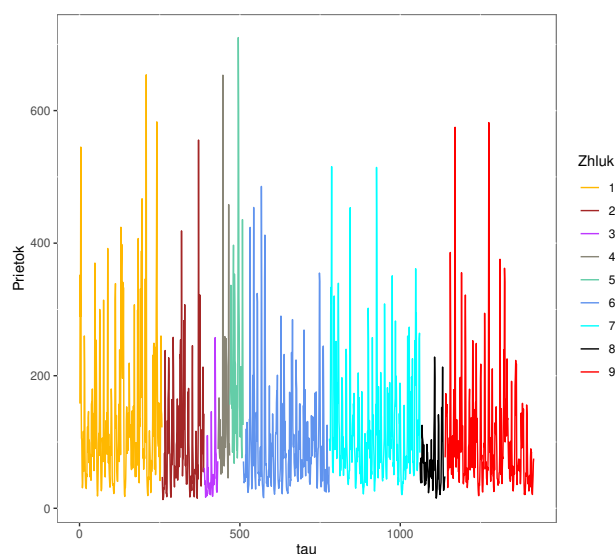
Tabuľka 5: Nájdené body zmeny prietoku Moravy pri iteračnom aplikovaní Pettittovho testu.

Aj v tomto prípade sme rozdelenie pôvodného radu zobrazili graficky - Obr. 10.

Hľadanie viacnásobných bodov zmeny pomocou iteračného SNHT si vyžadovalo vykonať 14 iterácii, pričom tento algoritmus detekoval 7 bodov zmien. Tieto body sú vypísané v Tabuľke 6. Vidíme, že zmena v čase $\tau = 446$ sa nachádza blízko jednej zo zmien detekovaných pomocou divíznej metódy. Divízna metóda našla aj zmenu v čase $\tau = 329$, teda rovnaké pozorovanie, ktoré bolo detekované iteračnou SNHT. Taká istá situácia je aj s bodom v čase $\tau = 579$, pričom nasledujúce pozorovanie označila CUSUM metóda za bod zmeny.

Opäť sme rozdelili náš pôvodný rad podľa vyššie spomínaných bodov - Obr. 11

V tomto prípade sme hodnoty vykreslili pomocou bodov pospájaných čiarami. Dôvodom je zvýraznenie pozorovania v čase $\tau = 447$. Toto pozorovanie algoritmus označil za



Obr. 10: Rozdelenie radu nájdenými bodmi pomocou iteračného aplikovania Pettittovho testu.

τ	<i>Mesiac a rok</i>
4	Apríl 1900
329	Máj 1927
446	Február 1937
447	Marec 1937
509	Máj 1942
579	Marec 1948
1384	Apríl 2015

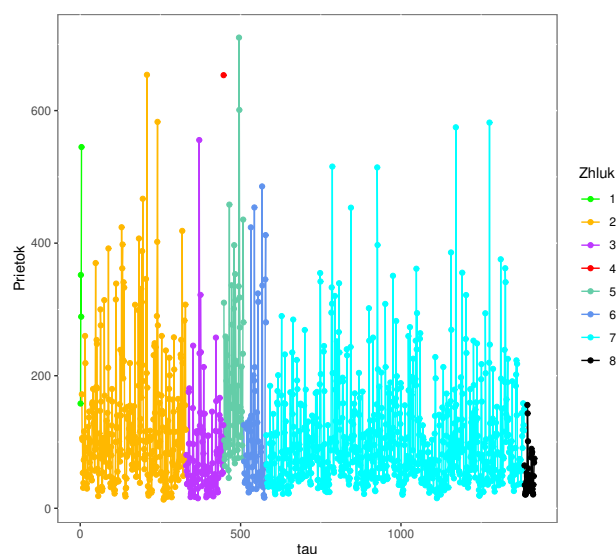
Tabuľka 6: Nájdené body zmeny prietoku Moravy pri iteračnom aplikovaní SNHT.

bod zmeny, no zjavne sa podľa tohto algoritmu riadi iným rozdelením ako jeho susedné pozorovania. Z toho by sme mohli vyvodiť, že algoritmus toto pozorovanie považuje za outliera.

5.4 Trend

Nakoniec sme testovali prítomnosť trendu. Použili sme obojstranné verzie sezónneho Mann-Kendallovho testu a Cox-Stuartovho testu.

Oba testy neprítomnosť trendu zamietli. Podľa nich sa teda v našom rade, resp. jeho



Obr. 11: Rozdelenie radu nájdenými bodmi pomocou iteračného aplikovania Pettittovho testu.

<i>Test</i>	<i>p-hodnota</i>
Sezónny Mann-Kendall	0.00057362
Cox-Stuart	0.03211

Tabuľka 7: Obojstranné testy prítomnosti trendu pre prietok Moravy.

strednej hodnote, trend nachádza.

Aby sme zistili, aký trend - rastúci alebo klesajúci, sa v rade nachádza, urobili sme jednostranný Cox-Stuartov test.

H_0	<i>p-hodnota</i>
Rastúci trend	0.01605
Klesajúci trend	0.9867

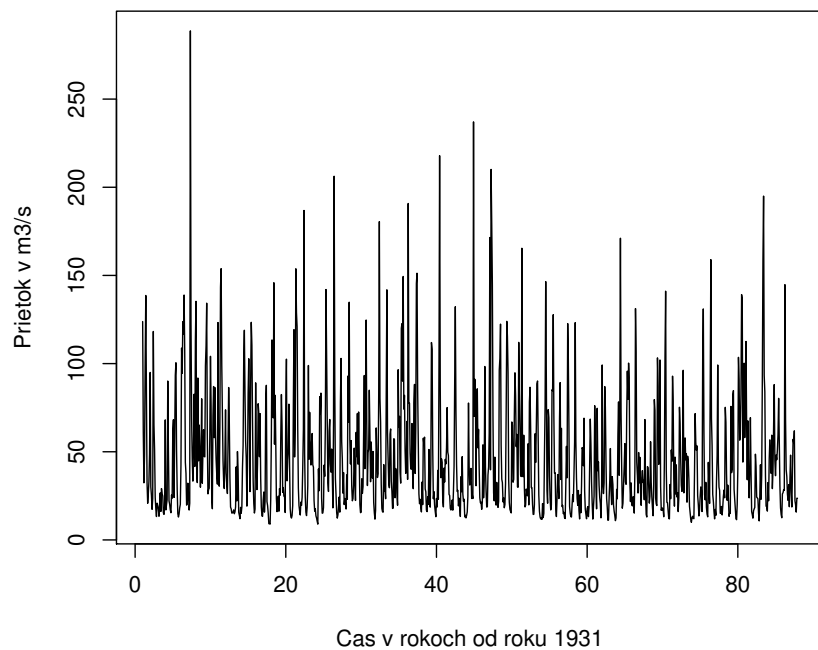
Tabuľka 8: Jednostranné Cox-Stuartove testy prítomnosti trendu pre prietok Moravy.

Tabuľka 8 najskôr vylúčila prítomnosť rastúceho trendu, no zároveň nevylúčila prítomnosť klesajúceho trendu. Tieto fakty môžeme zhrnúť tak, že v rade je klesajúci trend.

6 Analýza rieky Hron

Rieka Hron je druhou najdlhšou riekou na Slovensku s dĺžkou 298km a rozlohou povodia 5453km^2 . Pramení v Nízkych Tatrách na juhovýchodnom úpätí Kráľovej hole. Pri Štúrove sa vlieva do Dunaja [10].

V tejto práci budeme analyzovať prietok Hronu na základe mesačných dát zachytávaných obdobie od januára 1931 do decembra 2017 nameraných na stanici Brehy v Banskobystrickom regióne. Obr. 12 zachytáva priebeh týchto dát.



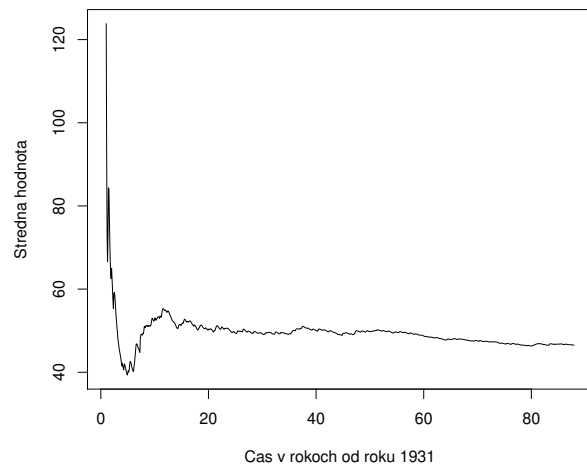
Obr. 12: Prietok Hrona od januára 1931 do decembra 2017.

Vývoj strednej hodnoty tohto radu môžeme vidieť na Obr. 13.

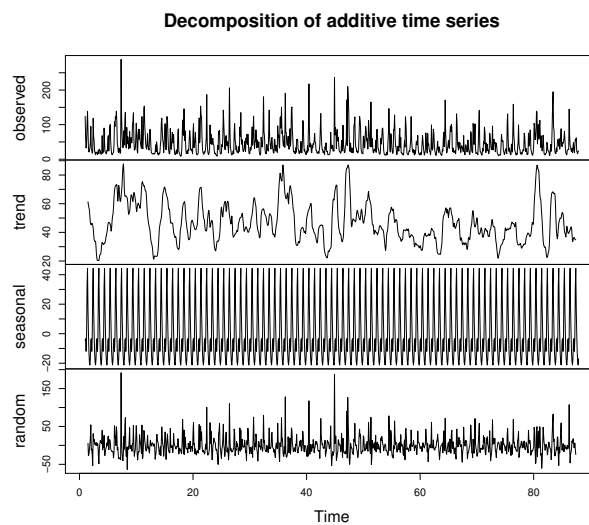
Dekompozíciu nášho radu na zložky popísané pomocou (36) ukazuje Obr. 14.

6.1 Homogenita

Opäť sme sa pri našej analýze najskôr pozreli na homogenitu tohto radu. O výsledkoch jej testov hovorí Tabuľka 9. Nakoľko majú oba testy p-hodnotu menšiu ako 5%, vylúčili homogenitu tohto radu. To znamená, že nejaká zmena sa v tomto rade nachádza.



Obr. 13: Stredná hodnota prietoku Hrona.



Obr. 14: Dekompozícia prietoku Hrona.

<i>Test</i>	<i>p – hodnota</i>
Rank Von Neumann (Bartels)	$< 2.2 \cdot 10^{-16}$
Wald-Wolfowitz	$< 2.2 \cdot 10^{-16}$

Tabuľka 9: Testy homogenity prietoku Hrona.

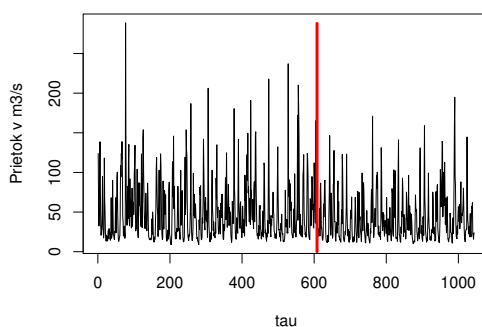
6.2 Bod zmeny

Začali sme s analýzou jedného bodu zmeny. Urobili SNH test, Pettittov test a aj CUSUM. Ich výsledky sú v Tabuľke 10.

<i>Test</i>	<i>p-hodnota</i>	τ	<i>Mesiac a rok</i>
SNHT	0.00475	608	August 1981
Pettitt	0.001261	608	August 1981
CUSUM	-	608	August 1981

Tabuľka 10: Detekcie jedného bodu zmeny prietoku Hrona.

Vidíme, že všetky tri testy označili ako potenciálny bod zmeny rovnaký bod, t.j. august 1981. Keďže p-hodnota SNHT ako aj Pettittovho testu je menšia ako 5%, tieto dva testy označili zmenu v auguste 1981 za signifikantnú. Obr. 15 zachytáva čas, v ktorom zmena strednej hodnoty nastala.



Obr. 15: Zmena v prietoku Hrona.

Signifikanciu tohto bodu sme overili aj pomocou Mann-Whitney-Wilcoxonovho testu a Kolmogorov-Smirnovovho testu. Ich závery môžeme vidieť v Tabuľke 11.

Vidíme, že oba testy zamietli H_0 , teda homogenitu. Takže môžeme konštatovať, že august 1981 je skutočne bodom zmeny.

<i>Test</i>	August 1981
Mann-Whitney-Wilcoxon	$9.94 \cdot 10^{-5}$
Kolmogorov-Smirnov	0.0009431

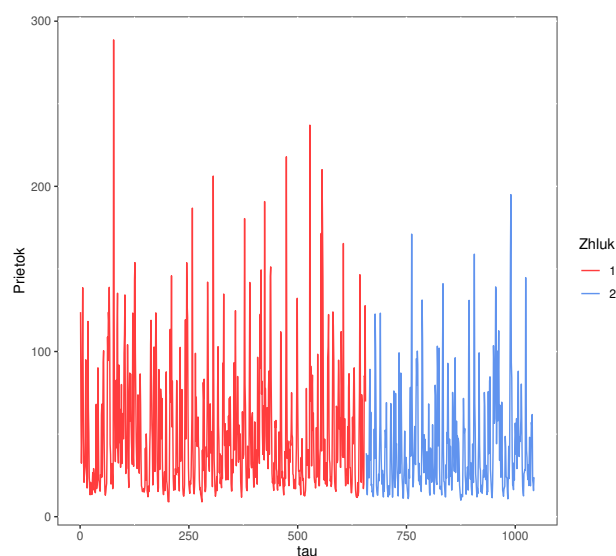
Tabuľka 11: Testy signifikantnosti pre august 1981.

6.3 Viacnásobné body zmeny

Ako pri analýze prietoku Moravy, tak aj pri analyzovaní prietoku Hrona sme hľadali viacnásobné body zmeny pomocou divízneho a aglomeratívneho algoritmu a aj pomocou iteračného Pettittovho testu a SNHT.

Divízna metóda našla iba jeden bod zmeny. Zmena podľa neho nastala v čase $\tau = 656$, teda v auguste 1985. O signifikantnosti tejto zmeny rozhodla p-hodnota, ktorá je pre tento bod 0.001996008. Keďže je p-hodnota menšia ako 5%, táto zmena je signifikantná. Táto metóda našla aj jeden nesignifikantný bod - v čase $\tau = 440$, teda august 1967. Nakoľko je jeho p-hodnota 0.153692615, čo je viac ako 5%, tento bod bol vyhodnotený za nesignifikantný.

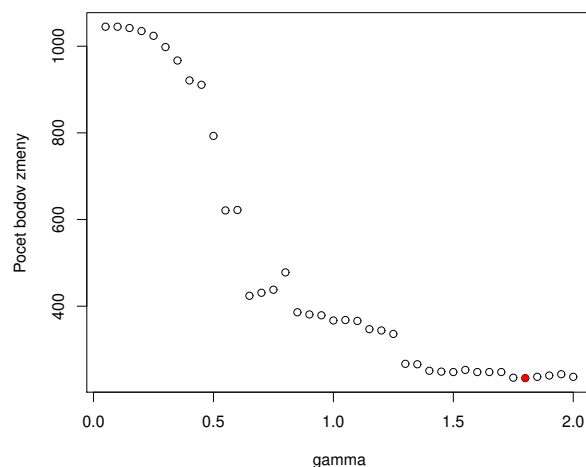
August 1985 rozdelil náš pôvodný rad na dve časti. Toto rozdelenie je zobrazené na Obr. 16.



Obr. 16: Rozdelenie prietoku Hrona na základe nájdeného bodu zmeny pomocou E-Divisive.

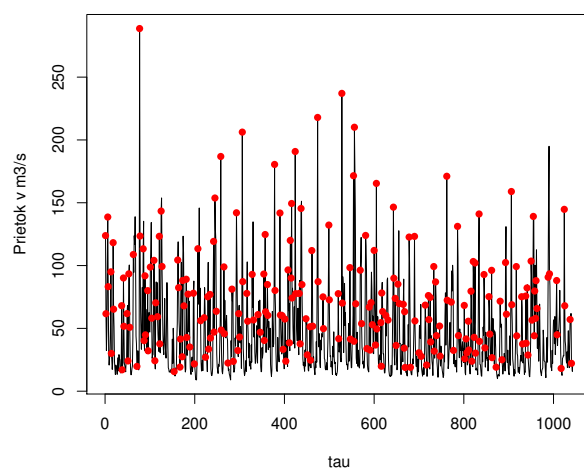
V prípade aglomeratívneho algoritmu sme opäť najskôr hľadali hodnotu parametra γ ,

pri ktorej bude počet nájdených bodov najmenší. Tento proces je zachytený na Obr. 17.



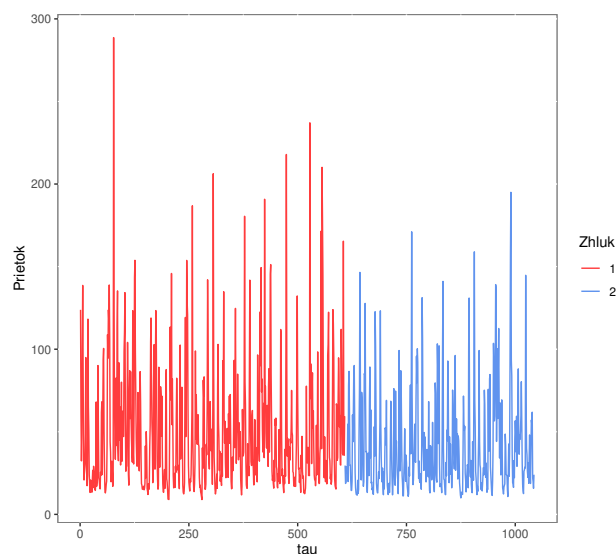
Obr. 17: Počet nájdených bodov zmeny pomocou aglomeratívneho algoritmu v závislosti od hodnoty γ .

V tomto prípade najnižší počet bodov prináša $\gamma = 1.8$, pričom nájdeme 234 bodov. Pre predstavu ako husto nájdené body ležia, prinášame Obr. 18.



Obr. 18: Body zmeny nájdené pomocou aglomeratívneho algoritmu.

Viacnásobné body zmeny sme sa pokúsili nájsť aj pomocou iteračného Pettittovho testu a SNHT. Oba testy pri hľadaní jedného bodu označili za bod zmeny pozorovanie s poradovým číslom 608, teda august 1981. Toto pozorovanie nám rozdelilo pôvodný rad na dva samostatné rady. Toto rozdelenie je zobrazené na Obr. 19.



Obr. 19: Rozdelenie prietoku Hrona na základe nájdeného bodu zmeny pomocou Pettittovho testu a SNHT.

Na oba tieto rady, ktoré vznikli rozdelením pôvodného radu, sme použili ako Pettittov test, tak aj SNHT. Testy síce body našli, no ani jeden z nich nebol signifikantný. Preto môžeme konštatovať, že oba iteračné prístupy našli iba jeden bod zmeny, ktorý sme našli v časti 6.2.

6.4 Trend

Ďalej sme sa pozreli na dlhodobé zmeny v tomto rade. Analyzovali sme teda výskyt trendu pomocou sezónneho Mann-Kendallovho testu a Cox-Stuartovho testu. Nasledujúca tabuľka zobrazuje ich výsledky.

<i>Test</i>	<i>p-hodnota</i>
Sezónny Mann-Kendall	$3.9128e - 05$
Cox-Stuart	0.07263

Tabuľka 12: Obojstranné testy prítomnosti trendu pre prietok Hronu.

Mann-Kendallov test zamietol hypotézu, ktorá hovorí o neprítomnosti trendu. Teda podľa neho sa v rade trend nachádza. Naopak, Cox-Stuartov test nevylúčil, že sa v rade trend nenachádza. Musíme však konštatovať, že hypotézu o neprítomnosti trendu na hla-

dine významnosti 5% nezamietol veľmi tesne. Z tohto dôvodu, sme vykonali Cox-Stuartove jednostranné testy, pričom prítomnosť trendu bola zisťovaná pomocou obojstranného testu. Výsledky jednostranných testov môžeme vidieť v Tabuľke 13.

H_0	p -hodnota
Rastúci trend	0.03631
Klesajúci trend	0.9701

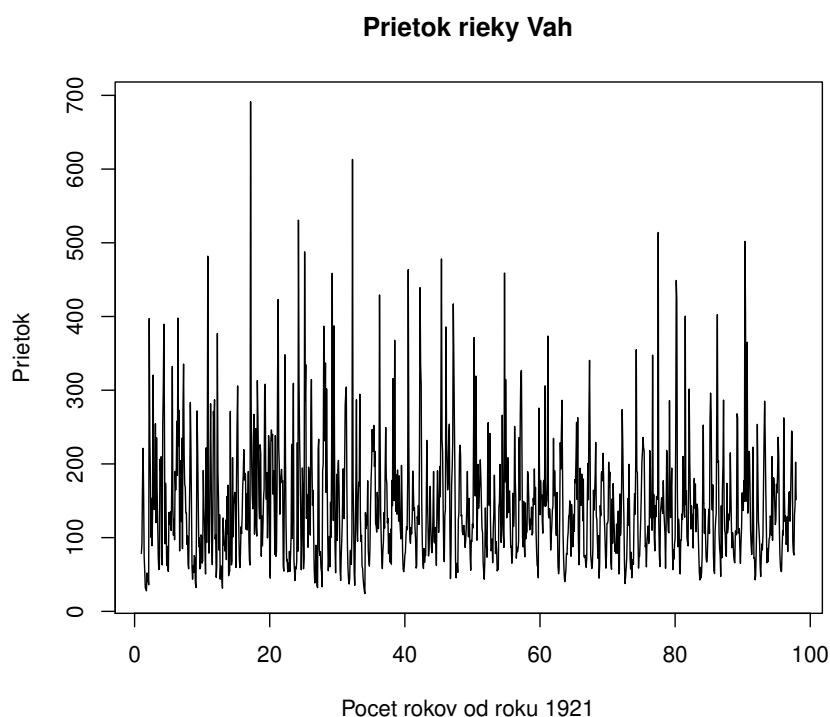
Tabuľka 13: Tabuľka p -hodnôt jednostranných Cox-Stuartových testov prítomnosti trendu pre prietok Hronu.

Vidíme, že hypotéza H_0 hovoriaca o rastúcom trende bola zamietnutá. Naopak hypotéza o klesajúcom trende zamietnutá nebola. Môžeme teda konštatovať, že v rade je klesajúci trend.

7 Analýza rieky Váh

Rieka Váh je najdlhšou riekou na Slovensku. Táto rieka vzniká sútokom riek Biely Váh, ktorý pramení na svahoch Kriváňa, a Čierny Váh prameniáci pod Kráľovou hoľou. Celková dĺžka jeho toku je 403km a jeho povodie zaberá rozlohu približne 10640km^2 . V Komárne sa Váh vlieva do Dunaja [23].

Graf na Obr. 20 zachytáva mesačný prietok Váhu v m^3/s od januára 1921 do decembra 2017. Tieto údaje boli namerané na stanici v Šali.



Obr. 20: Prietok Váhu od januára 1921 do decembra 2017.

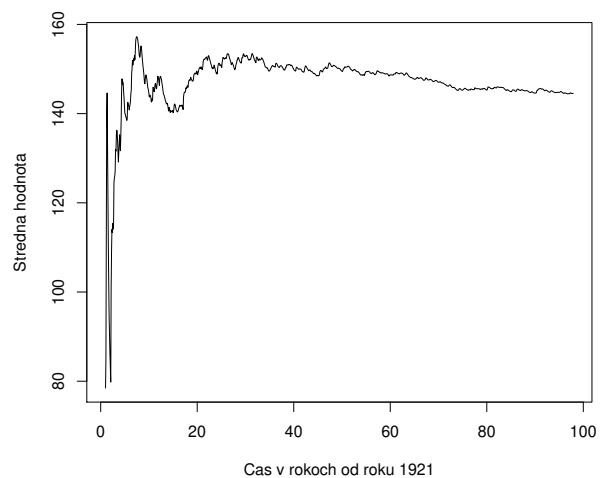
Priebeh strednej hodnoty prietoku Váhu je zobrazený na Obr. 21.

Dekompozícia tohto radu nám ukáže priebeh jeho jednotlivých zložiek - Obr. 22.

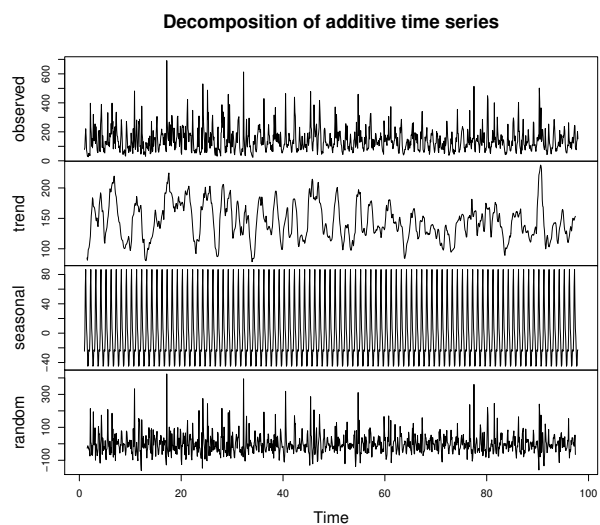
7.1 Homogenita

Na otestovanie homogenity sme opäť použili rankovú verziu Von Neumannovho testu (Bartelsov test) a Wald-Wolfowitzov test. Ich výsledky sú zachytené v Tabuľke 14.

Tak ako pri predošlých dvoch riekach, tak aj homogenita prietoku Váhu bola zamietnutá oboma testami.



Obr. 21: Pribeh strednej hodnoty Váhu.



Obr. 22: Dekompozícia prietoku Váhu.

<i>Test</i>	<i>p – hodnota</i>
Rank Von Neumann (Bartels)	$< 2.2 \cdot 10^{-16}$
Wald-Wolfowitz	$< 2.2 \cdot 10^{-16}$

Tabuľka 14: Testy homogenity prietoku Váhu.

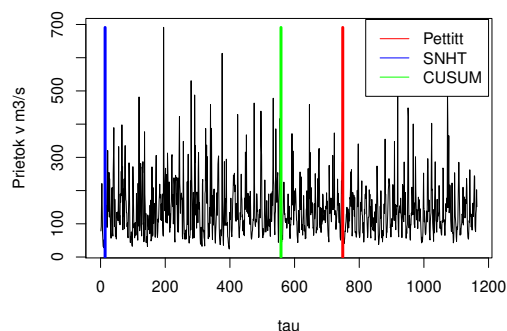
7.2 Bod zmeny

Na hľadanie bodu zmeny sme znovu urobili SNHT, Pettittov test a CUSUM na týchto dátach. Tabuľka 15 zobrazuje ich závery.

<i>Test</i>	<i>p-hodnota</i>	τ	<i>Mesiac a rok</i>
SNHT	0.1255	14	Február 1922
Pettitt	0.3768	749	Máj 1983
CUSUM	-	558	Jún 1967

Tabuľka 15: Detekcie jedného bodu zmeny prietoku Váhu.

Každý z testov označil za bod zmeny iný čas. Avšak SNHT aj Pettittov test označili nájdený bod za nesignifikantný. Všetky tri body sú zobrazené na Obr. 23



Obr. 23: Body zmeny prietoku Váhu.

Signifikanciu nájdených troch bodov sme otestovali aj Mann-Whitney-Wilcoxonovým a Kolmogorov-Smirnovovým testom.

<i>Test</i>	Február 1922	Máj 1983	Jún 1967
Mann-Whitney-Wilcoxon	0.0002548	0.05649	0.1859
Kolmogorov-Smirnov	0.001212	0.00531	0.0002395

Tabuľka 16: Testy signifikantnosti pre nájdené body zmeny prietoku Váhu.

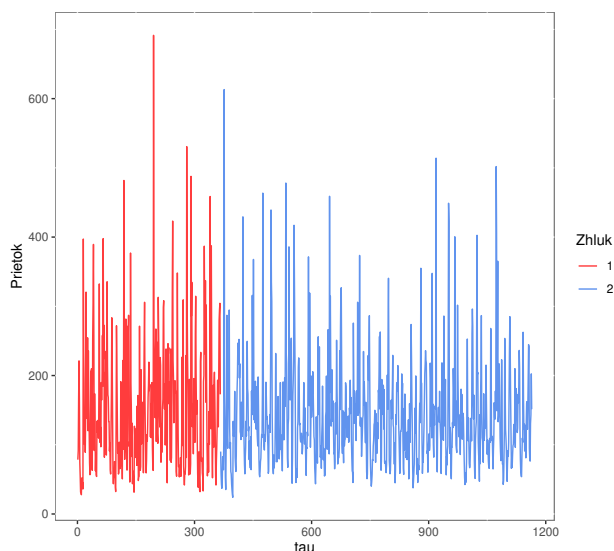
V prípade februára 1922 oba testy homogenitu zamietli, a teda určili zmenu v tomto čase za signifikantnú. Bod nájdený Pettittovým testom Mann-Whitney-Wilcoxonov test

veľmi tesne vyhodnotil ako nesignifikantný, avšak Kolmogorov-Smirnovov test ho označil ako signifikantný. Jún 1967 bol Mann-Whitney-Wilcoxonovým testom označený ako nesignifikantný, no Kolmogorov-Smirnovov test si myslí presný opak.

7.3 Viacnásobné body zmeny

Podobne ako aj v prípade predošlých dvoch riek sme viacnásobné body hľadali pomocou oboch hierarchických metód a taktiež iteračných metód.

Pomocou divíznej metódy sme našli iba jeden bod zmeny. Podľa nej zmena strednej hodnoty, teda aj zmena rozdelenia, nastala v čase $\tau = 366$. Tento bod predstavuje hodnotu nameranú v júni 1951. Rozdelenie radu môžeme vidieť na Obr. 24.



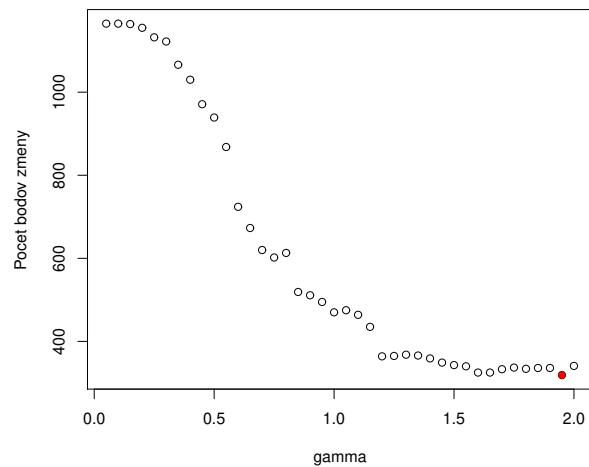
Obr. 24: Rozdelenie radu prietoku Váhu podľa bodu zmeny nájdeného pomocou divíznej metódy.

Priebeh počtu bodov nájdených pomocou aglomeratívnej procedúry je zobrazený na grafe na Obr. 25.

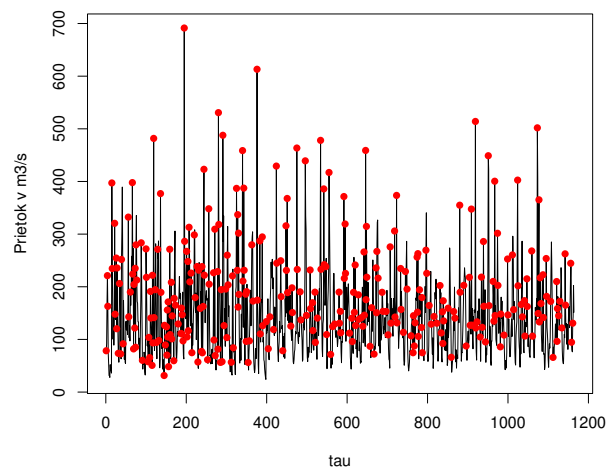
Vidíme, že najnižší počet bodov nájde, ak $\gamma = 1.95$. Pre túto hodnotu γ nájde 319 bodov zmeny. Pre predstavu uvádzame Obr. 26.

Iteračné aplikovanie Pettittovho testu skončilo s rovnakým výsledkom ako divízna metóda - nájdený bol iba jeden bod. Tento bod sme našli už v časti 7.2 pomocou Pettittovho testu. Obr. 27 zobrazuje rozdelenie pôvodného radu pomocou tohto bodu.

Napriek tomu, že SNHT označil bod nájdený v časti 7.2 za nesignifikantný, testy



Obr. 25: Počet nájdených bodov zmeny pomocou aglomeratívneho algoritmu v závislosti od hodnoty γ .



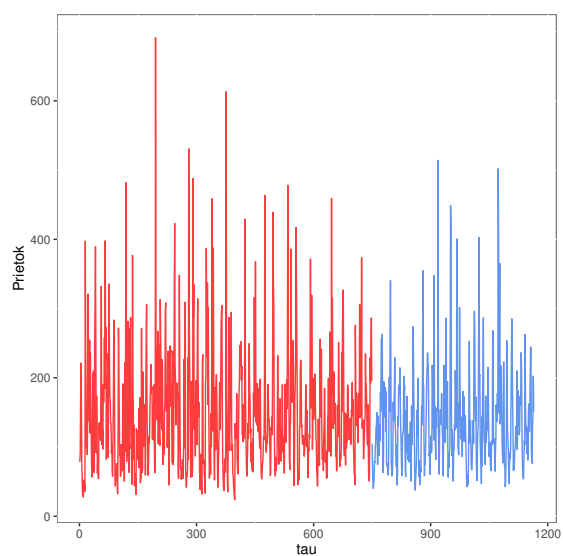
Obr. 26: Body zmeny nájdené pomocou aglomeratívneho algoritmu.

určujúce jeho signifikanciu si mysleli opak. Z tohto dôvodu sme ho aplikovali iteračne. Okrem bodu v čase $\tau = 14$, ktorý bol nájdený pomocou SNHT v časti 7.2, sme našli aj bod v čase $\tau = 5$, čo zodpovedá máju 1921.

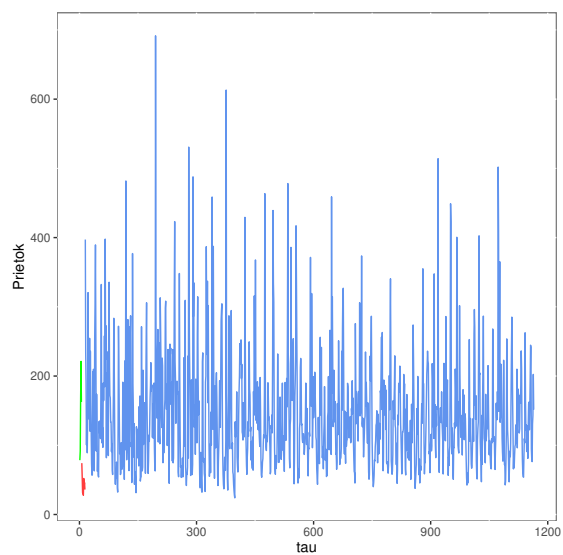
7.4 Trend

Aj na rieke Váh sme testovali prítomnosť trendu.

Oba testy prítomnosť trendu v rade radikálne zamietli. Z tohto dôvodu nevidíme dôvod



Obr. 27: Rozdelenie radu prietoku Váhu podľa bodu nájdeného pomocou Pettittovho testu.



Obr. 28: Rozdelenie radu prietoku Váhu podľa bodov nájdených pomocou iterovania SNHT.

<i>Test</i>	<i>p-hodnota</i>
Sezónny Mann-Kendall	0.42694
Cox-Stuart	0.7717

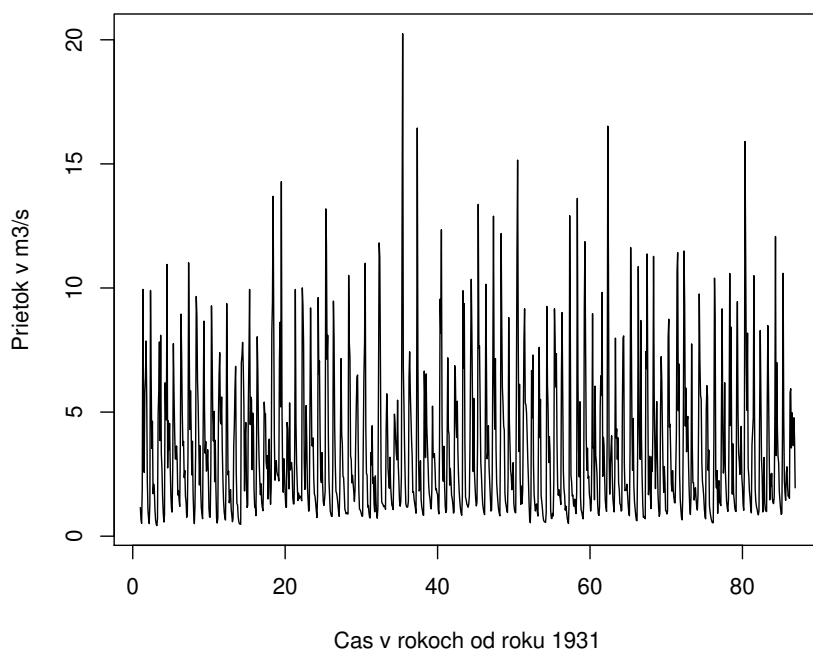
Tabuľka 17: Obojstranné testy prítomnosti trendu pre prietok Váhu.

testovať ani jednostranné hypotézy.

8 Analýza rieky Belá

Rieka Belá vzniká sútokom Tichého a Kôprovského potoka na juh od Kriváňa. Dĺžka jej toku je 24km a jej povodie zaberá plochu 245km^2 . V Liptovskom Hrádku ústí do rieky Váh. Viac o toku tejto rieky môžeme nájsť na [9].

V tejto práci analyzujeme mesačné prietoky tejto rieky za obdobie od januára 1931 do decembra 2016 namerané na stanici v Podbanskom (okres Liptovský Mikuláš). Dáta sú zachytené na Obr. 29.



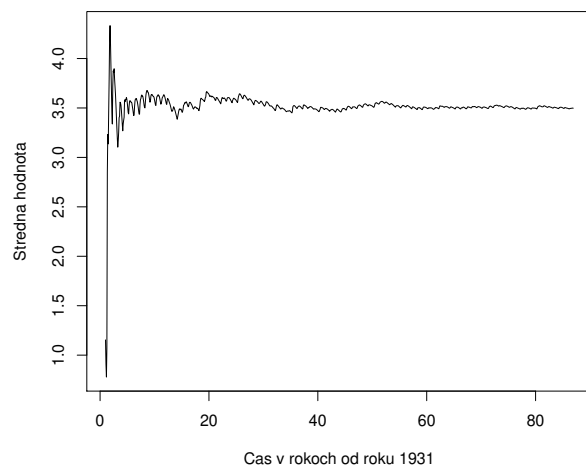
Obr. 29: Prietok Belej od januára 1931 do decembra 2016.

Pozrime sa na priebeh strednej hodnoty nášho radu na Obr. 30 a na dekompozíciu radu na trendovú, sezónnu a náhodnú zložku - Obr. 31.

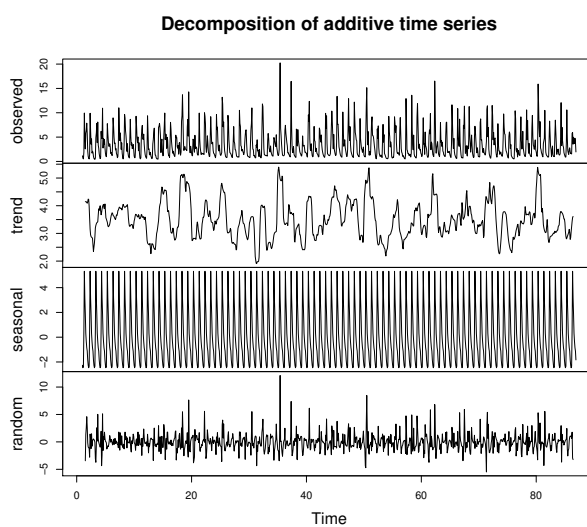
8.1 Homogenita

Testy homogenity nám povedia, či má vôbec význam hľadať zmeny v danom rade. Ich výsledky sú uvedené v Tabuľke 18.

Ako aj v predošlých prípadoch, p -hodnota je oveľa menšia ako 5%. Hypotézu, že rad je homogénny, teda v ňom nie je žiadna zmena rozdelenia, zamietame. Z toho vyplýva, že



Obr. 30: Stredná hodnota prietoku Belej.



Obr. 31: Dekompozícia prietoku Belej.

<i>Test</i>	<i>p – hodnota</i>
Rank Von Neumann (Bartels)	$< 2.2 \cdot 10^{-16}$
Wald-Wolfowitz	$< 2.2 \cdot 10^{-16}$

Tabuľka 18: Testy homogenity prietoku Belej.

by sme v rade mohli nájsť nejakú zmenu.

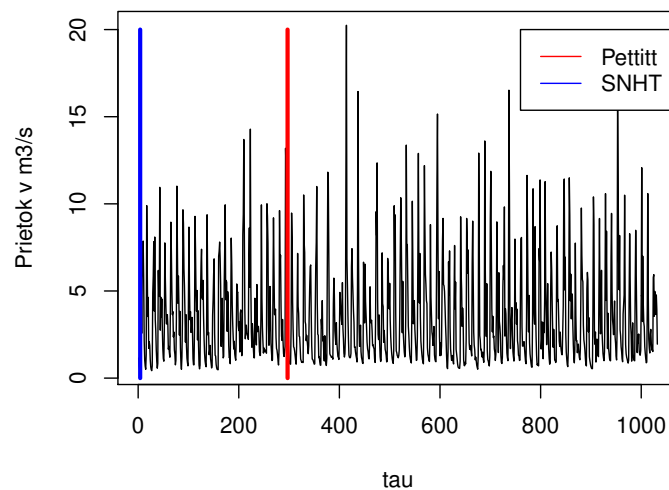
8.2 Bod zmeny

Na nájdenie bodov zmeny sme opäť použili metódy opísané v Kapitole 2. Ich výsledky sú zachytené v Tabuľke 19.

<i>Test</i>	<i>p-hodnota</i>	τ	<i>Mesiac a rok</i>
SNHT	0.9168	4	Apríl 1931
Pettitt	0.4822	297	September 1955
CUSUM	-	-	-

Tabuľka 19: Detekcie jedného bodu zmeny prietoku Belej.

SNHT našiel bod zmeny v apríli 1931 a Pettittov test v septembri 1955. Oba body však tieto testy označili za nesignifikantné. CUSUM metóda dokonca v tomto rade nenašla žiadnu zmenu v strednej hodnote. Časy zmien nájdených pomocou SNHT a Pettittovým testom sú zobrazené na Obr. 32.



Obr. 32: Body zmeny v prietoku Belej nájdené pomocou SNHT a Pettittovho testu.

Pozrime sa, čo si o nájdených bodoch myslia Mann-Whitney-Wilcoxonov test a Kolmogorov-Smirnovov test. Ich výsledky môžeme vidieť v Tabuľke 20.

Mann-Whitney-Wilcoxonov test označil september 1955 za signifikantný. Naopak zmena v apríli 1931 podľa neho nie je významná. Kolmogorov-Smirnovov test priniesol presne

<i>Test</i>	Apríl 1931	September 1955
Mann-Whitney-Wilcoxon	0.01759	0.06249
Kolmogorov-Smirnov	0.1254	0.02824

Tabuľka 20: Testy signifikantnosti zlomov v prietoku Belej.

opačné závery.

8.3 Viacnásobné body zmeny

Keďže na základe predchádzajúcej kapitoly môžeme usúdiť, že v rade nenastala žiadna zmena v strednej hodnote, nevidíme dôvod hľadať viacnásobné zmeny.

8.4 Trend

Pozrime sa však na dlhodobé zmeny režimu. Tabuľka 21 nám hovorí, že oba testy prítomnosť trendu zamietli.

<i>Test</i>	<i>p-hodnota</i>
Sezónny Mann-Kendall	0.72881
Cox-Stuart	0.8949

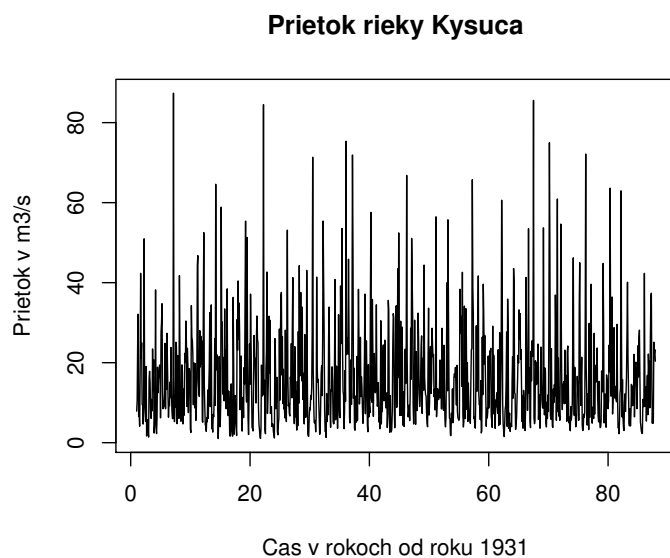
Tabuľka 21: Obojstranné testy prítomnosti trendu pre prietok Belej.

Nakoľko je p -hodnota Cox-Stuartovho obojstranného testu relatívne dosť vysoká - zamietla prítomnosť trendu, oba jednostranné testy taktiež zamietnu prítomnosť akéhokoľvek monotónneho trendu.

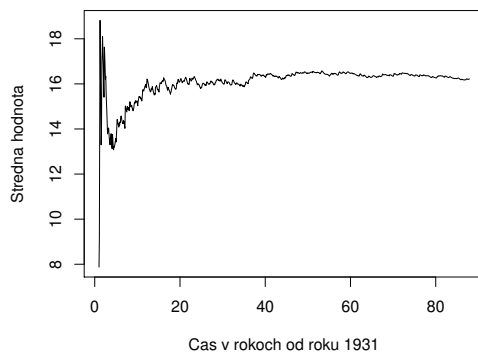
9 Analýza rieky Kysuca

Ďalšou riekou, ktorú budeme analyzovať, je rieka Kysuca. Podľa [12] táto rieka pramení v oblasti Makova a preteká celým regiónom Kysúc. Pri Žiline sa vlieva do Váhu. Jej celková dĺžka je $66,3\text{ km}$ a jej povodie má rozlohu približne 1040 km^2 .

Vstupnými dátami našej analýzy budú hodnoty mesačných prietokov spomínanej rieky za obdobie od januára 1931 do decembra 2017, ktoré boli namerané na stanici v Kysuckom Novom Meste. Pribeh týchto dát môžeme vidieť na Obr. 33 a pribeh strednej hodnoty na Obr. 34.

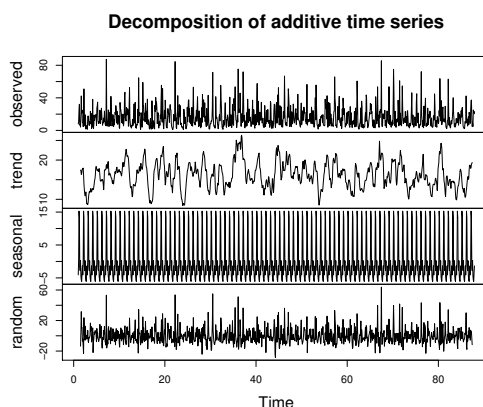


Obr. 33: Prietok rieky Kysuca od januára 1931 do decembra 2017.



Obr. 34: Stredná hodnota prietoku rieky Kysuca.

Obr. 35 zobrazuje dekompozíciu analyzovaného radu.



Obr. 35: Dekompozícia prietoku rieky Kysuca.

9.1 Homogenita

Pomocou testov homogenity zistíme, či sa v rade vôbec nejaká zmena nachádza.

<i>Test</i>	<i>p – hodnota</i>
Rank Von Neumann (Bartels)	$< 2.2 \cdot 10^{-16}$
Wald-Wolfowitz	$< 2.351 \cdot 10^{-16}$

Tabuľka 22: Testy homogenity prietoku rieky Kysuca.

Oba testy homogenity nášho radu zamietli. Môžeme povedať, že nejaká zmena sa v rade nachádza.

9.2 Bod zmeny

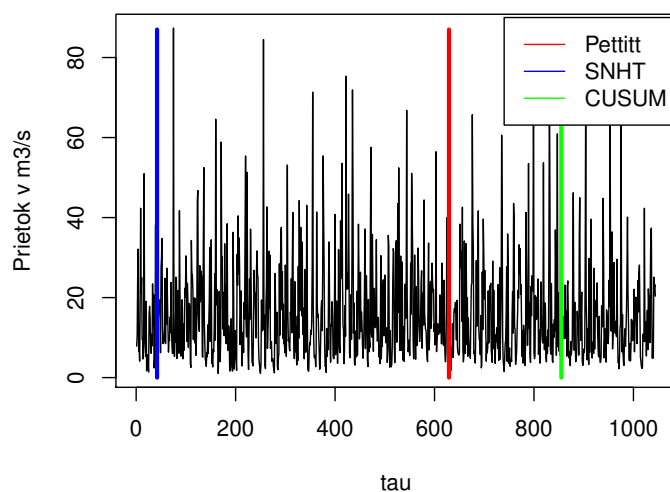
Body zmeny budeme hľadať podobne ako v predošlých kapitolách, t.j. pomocou SNHT, Pettittovho testu a CUSUM metódy. Ich závery zachytáva Tabuľka 23.

SNHT považuje za bod zmeny jún 1934 a Pettittov test máj 1983. Oba testy však nájdené body označili za nevýznamné. CUSUM metóda našla marec 2002, no zatiaľ nevieme jeho signifikantnosť. Tieto tri časy sme zaznačili aj na grafe na Obr. 36.

Signifikanciu vyššie nájdených bodov sme overili aj pomocou Mann-Whitney-Wilcoxonovho testu a Kolmogorov-Smirnovovho testu, viď Tabuľka 24.

<i>Test</i>	<i>p-hodnota</i>	τ	<i>Mesiac a rok</i>
SNHT	0.9486	42	Jún 1934
Pettitt	0.8718	629	Máj 1983
CUSUM	-	855	Marec 2002

Tabuľka 23: Body zmeny prietoku rieky Kysuca.



Obr. 36: Body zmeny v prietoku Kysuce nájdené pomocou SNHT, Pettittovho testu a CUSUM.

<i>Test</i>	Jún 1934	Máj 1983	Marec 2002
Mann-Whitney-Wilcoxon	0.04113	0.188	0.3198
Kolmogorov-Smirnov	0.08783	0.287	0.1643

Tabuľka 24: Testy signifikantnosti zlomov prietoku Kysuce.

Môžeme si všimnúť, že Mann-Whitney-Wilcoxonov test považuje za signifikantný iba jún 1934, aj to relatívne tesne. Kolmogorov-Smirnovov test nepovažuje za významný ani jeden z nájdených bodov. Môžeme teda povedať, že v prietoku Kysuce sme nenašli žiadnu signifikantnú jednorázovú zmenu strednej hodnoty. To znamená, že z pohľadu strednej hodnoty je prietok Kysuce relatívne stabilný vzhľadom na jednorázové zmeny.

9.3 Viacnásobné body zmeny

Ako v prípade rieky Belá, ktorú sme analyzovali v predošlej kapitole, tak aj v prípade rieky Kysuca nevidíme dôvod hľadať viacnásobné zmeny, keďže sme v predchádzajúcej časti tejto kapitoly nenašli ani jednu zmenu.

9.4 Trend

Hoci sa v rade nenachádzajú "nárazové" zmeny, je možné, že sa v ňom prejavujú dlhodobé zmeny - trend. Na analýzu trendu sme opäť použili už známe testy - sezónny Mann-Kendallov a Cox-Stuartov test. Výsledky tejto analýzy sú v Tabuľke 25.

<i>Test</i>	<i>p-hodnota</i>
Sezónny Mann-Kendall	0.95949
Cox-Stuart	0.6937

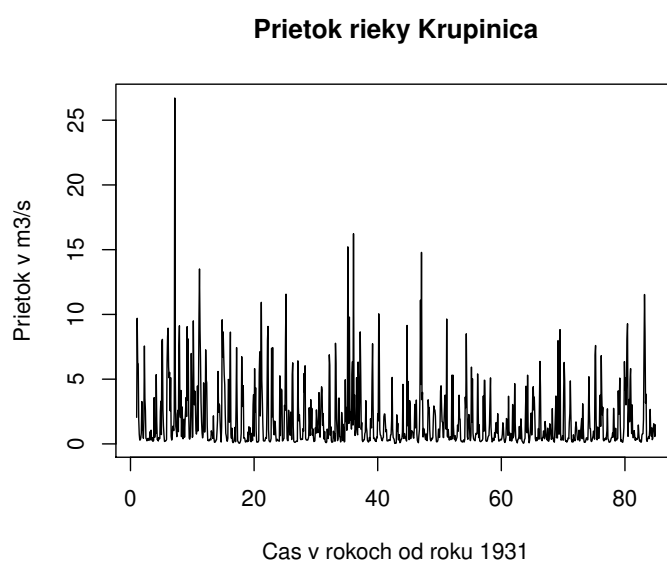
Tabuľka 25: Obojstranné testy prítomnosti trendu pre prietok Kysuce.

Oba testy sa zhodli na závere, že v našom rade sa monotónny trend nenachádza. Keďže vidíme, že p-hodnota Cox-Stuartovho testu je pomerne vysoká, predpokladáme, že aj pri aplikácii jednostranných testov bude záver podobný, t.j. zamietneme prítomnosť monotónneho trendu.

10 Analýza rieky Krupinica

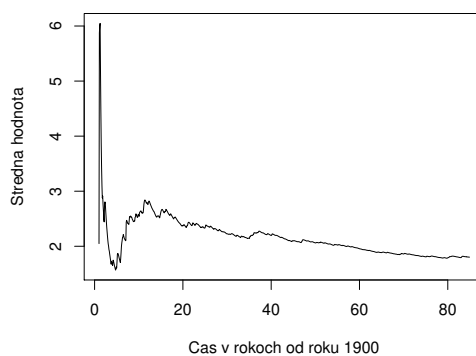
Prameň rieky Krupinica sa nachádza pri osade Podlysec v pohorí Javorina. Má dĺžku niečo viac ako 65km a jej povodie zaberá plochu 551km^2 . Západne od mesta Šahy sa vlieva do rieky Ipeľ. [22]

Prietok Krupinice budeme analyzovať na základe mesačných hodnôt nameraných na stanici v obci Plášťovce (okres Levice). Dáta sú z obdobia od januára 1931 do decembra 2014, ich priebeh je zachytený na Obr. 37.

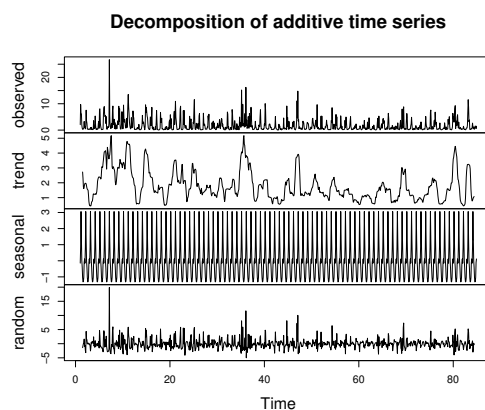


Obr. 37: Prietok Krupinice.

Na Obr. 38 vidíme priebeh strednej hodnoty tejto rieky a na Obr. 39 jej dekompozíciu.



Obr. 38: Stredná hodnota prietoku Krupinice.



Obr. 39: Dekompozícia prietoku Krupinice.

10.1 Homogenita

Z výsledkov uvedených v Tabuľke 26 vidíme, že hypotéza o homogenite nášho radu je zamietnutá. Z toho vyplýva, že sa v rade nejaká zmena nachádza.

<i>Test</i>	<i>p – hodnota</i>
Rank Von Neumann (Bartels)	$< 2.2 \cdot 10^{-16}$
Wald-Wolfowitz	$< 2.2 \cdot 10^{-16}$

Tabuľka 26: Testy homogenity prietoku Krupinice.

10.2 Bod zmeny

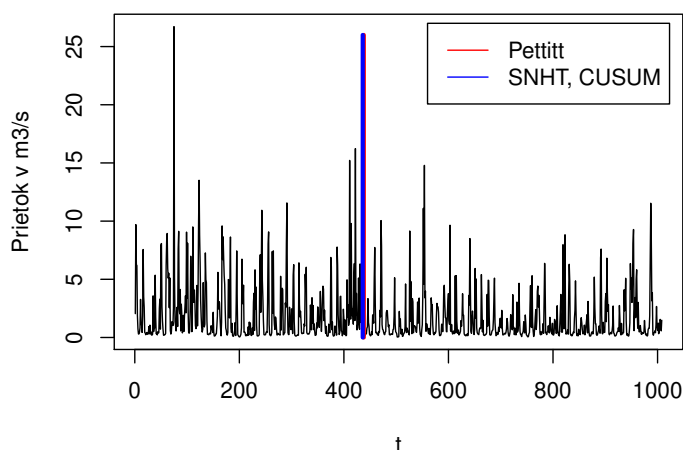
Najskôr sa pozrieme na hľadanie jedného bodu zmeny. Výsledky SNHT, Pettittovho testu a CUSUM metódy môžeme nájsť v Tabuľke 27.

<i>Test</i>	<i>p-hodnota</i>	τ	<i>Mesiac a rok</i>
SNHT	$< 2.2 \cdot 10^{-16}$	436	Apríl 1967
Pettitt	0.0004466	438	Jún 1967
CUSUM	-	436	Apríl 1967

Tabuľka 27: Detekcie jedného bodu zmeny Krupinice.

Podľa SNHT nastal zlom v apríli 1967. Pettittov test si myslí, že zmena nastala o dva

mesiace neskôr. Oba body vyhodnotil príslušný test za signifikantné. CUSUM algoritmus sa pridal na stranu SNHT a tiež označil apríl 1967 za prelomový. Čas, kedy zmeny nastali, môžeme vidieť na Obr. 40.



Obr. 40: Body zmeny v prietoku Krupinice nájdené pomocou SNHT, Pettittovho testu a CUSUM.

Významnosť apríla a júna 1967 sme overili aj pomocou testov z častí 2.4 a 2.5.

<i>Test</i>	Apríl 1967	Jún 1967
Mann-Whitney-Wilcoxon	$4.531 \cdot 10^{-5}$	$3.534 \cdot 10^{-5}$
Kolmogorov-Smirnov	$8.876 \cdot 10^{-6}$	$1.325 \cdot 10^{-5}$

Tabuľka 28: Testy signifikantnosti zlomov v prietoku Krupinice.

Z Tabuľky 28 vidíme, že homogenita pri danom τ bola zamietnutá oboma použitými testami pri oboch nájdených bodoch. Môžeme teda povedať, že zmena strednej hodnoty v nájdených časoch bola skutočne signifikantná. Keďže nájdené body sú relatívne blízko seba, môžeme povedať, že niekedy v období okolo apríla 1967 až júna 1967 skutočne nastala zmena, ktorá bola veľmi významná.

10.3 Viacnásobné body zmeny

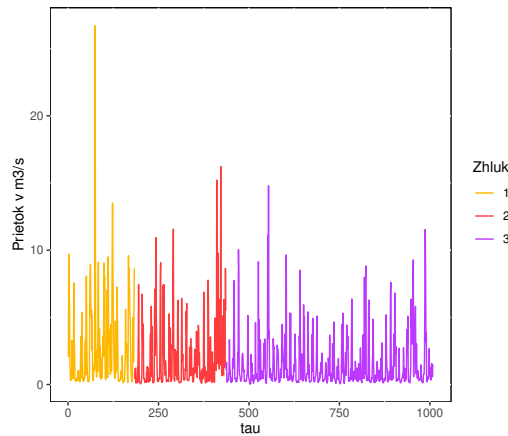
Nakolko sme v predchádzajúcej časti našli body zmeny, môžeme predpokladať, že sa ich v našom rade nachádza viac.

Najskôr sa pozrime na body nájdené pomocou divíznej metódy. Tie sú zachytené v Tabuľke 29 a rozdelenie radu podľa nich na Obr. 41.

τ	p -hodnota	Mesiac a rok
183	0.001996008	Marec 1946
436	0.039920160	Apríl 1967

Tabuľka 29: Viacnásobné zmeny v prietoku Krupinice nájdené pomocou E-Divisive.

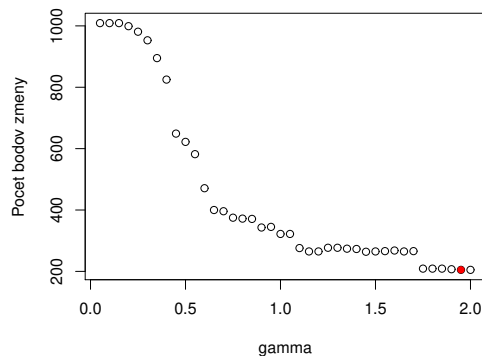
Môžeme si všimnúť, že aj divízna metóda našla bod, ktorý bol detekovaný aj pomocou SNHT a CUSUM. Tento fakt podporuje naše tušenie, že na jar 1967 nastala významná zmena prietoku Krupinice.



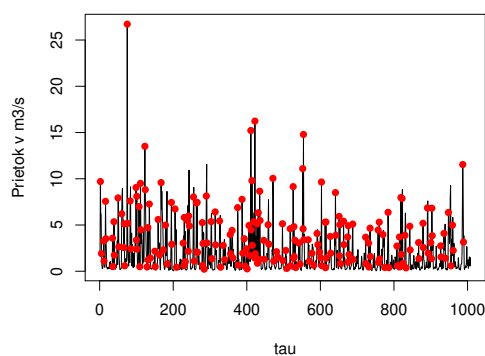
Obr. 41: Rozdelenie prietoku Krupinice nájdenými bodmi pomocou E-Divisive metódy.

Pred začiatkom aplikácie aglomeratívneho algoritmu sme opäť najskôr hľadali hodnotu parametra γ , pri ktorej bol počet nájdených bodov zmeny čo najmenší. Obr. 42 ukazuje, že takou hodnotou je $\gamma = 1.95$. Pre túto hodnotu aglomeratívna metóda našla 205 bodov zmeny. Tieto body sú zobrazené na Obr. 43.

Nakoniec sme na hľadanie viacnásobných bodov iteračne aplikovali Pettittov test a SNHT.



Obr. 42: Počet nájdených zlomov v prietoku Krupinice pomocou aglomeratívneho algoritmu v závislosti od hodnoty γ .



Obr. 43: Body zmeny Krupinice nájdené pomocou aglomeratívnej metódy.

Pomocou Pettittovho testu sme okrem júna 1967 nájdeného v časti 10.2 našli ďalšie štyri body. Výsledky iteračného Pettittovho testu vidíme v Tabuľke 30.

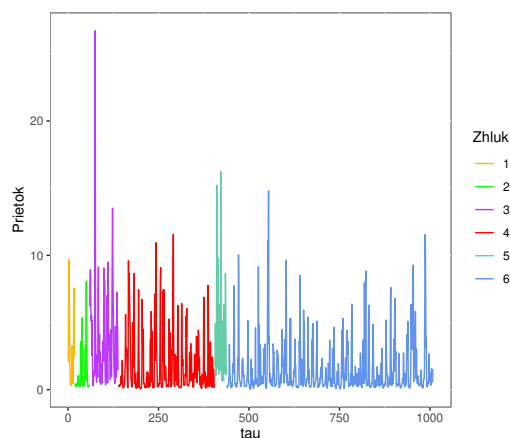
Päť nájdených bodov zlomu nám rozdelí náš rad na šesť častí, pričom každá z nich sa riadi vlastnou distribúciou. Grafické rozdelenie nášho radu zobrazuje Obr. 44.

Iteračná SNHT metóda našla okrem apríla 1967 taktiež ďalšie štyri body zachytené v Tabuľke 31. Rozdelenie radu na základe týchto bodov ukazuje Obr. 45.

Pri porovnaní Tabuľky 30 a Tabuľky 31 si môžeme všimnúť, že SNHT a Pettittov test našli 3 body, ktoré sú blízko seba. Oba testy teda našli približne rovnaké obdobia, kedy sa zmenil prietok Krupinice.

τ	<i>Mesiac a rok</i>
18	Jún 1932
59	November 1935
138	Jún 1942
405	September 1964
438	Jún 1967

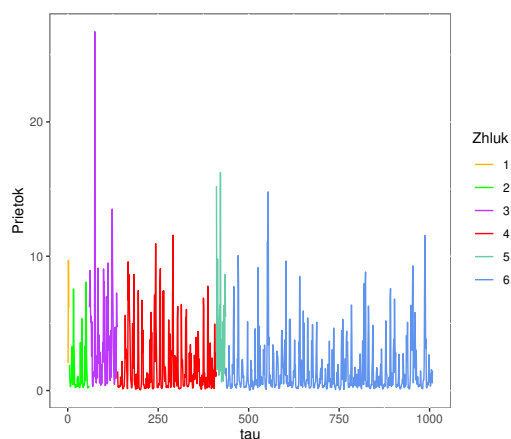
Tabuľka 30: Nájsené body zmeny prietoku Krupinice pri iteračnom aplikovaní Pettittovho testu.



Obr. 44: Rozdelenie radu prietoku Krupinice nájsenými bodmi pomocou iteračného aplikovania Pettittovho testu.

τ	<i>Mesiac a rok</i>
4	Apríl 1931
59	November 1935
137	Máj 1942
410	Február 1965
436	Apríl 1967

Tabuľka 31: Nájsené body zmeny prietoku Krupinice pri iteračnom aplikovaní SNHT.



Obr. 45: Aplikácia iteráčného SNHT na prítok Krupinice.

10.4 Trend

Okrem jednorázových zmien sa v našom rade môžu nachádzať aj dlhodobé zmeny - trend. Tabuľka 32 obsahuje výsledky sezónneho Mann-Kendallovho testu a Cox-Stuartovho testu po aplikácii na prítok Krupinice.

<i>Test</i>	<i>p-hodnota</i>
Sezónny Mann-Kendall	$3.0806 \cdot 10^{-5}$
Cox-Stuart	0.01421

Tabuľka 32: Obojstranné testy prítomnosti trendu pre prítok Krupinice.

Keďže je p-hodnota pri oboch testoch výrazne menšia ako 5%, v rade sa zrejme nachádza nejaký monotónny trend.

Na určenie toho, aký trend sa v rade nachádza, použijeme jednostranné Cox-Stuartove testy.

<i>H₀</i>	<i>p-hodnota</i>
Rastúci trend	0.007105
Klesajúci trend	0.9945

Tabuľka 33: Jednostranné Cox-Stuartove testy prítomnosti trendu pre prítok Krupinice.

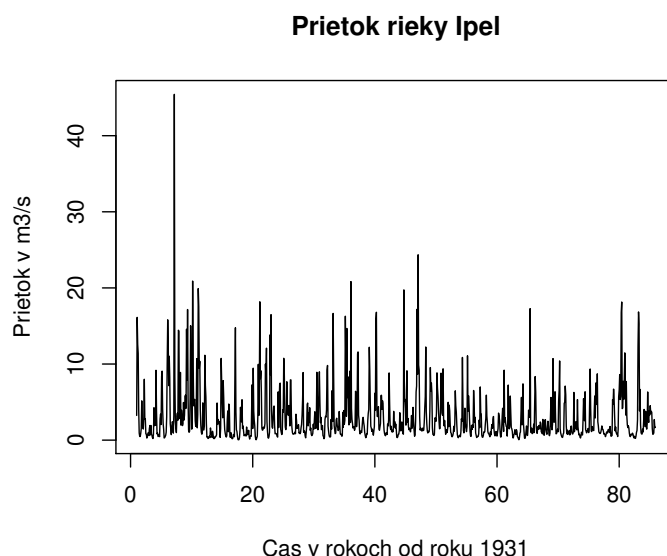
Z Tabuľky 33 vidíme, že hypotéza o rastúcom trende bola zamietnutá a hypotéza

o klesajúcom trende zamietnutá nebola. Z toho môžeme vyvodiť, že stredná hodnota prietoku rieky Krupinica s postupom času monotónne klesá.

11 Analýza rieky Ipeľ

Prameň rieky Ipeľ sa nachádza pod horou Vepor v Slovenskom rudohorí. Tok tejto rieky má približne 232.5km z čoho takmer 140km tvorí hranicu s Maďarskom. Povodie Ipľa má rozlohu asi 5151km^2 . Pri obci Chľaba v okrese Nové Zámky sa vlieva do Dunaja. Viac informácií možno nájsť na [11].

Zmeny tejto rieky budeme analyzovať na základe dát zozbieraných na stanici v obci Holiša (okres Lučenec) v období od januára 1931 do decembra 2015. Priebeh týchto dát je zobrazený na Obr. 46.

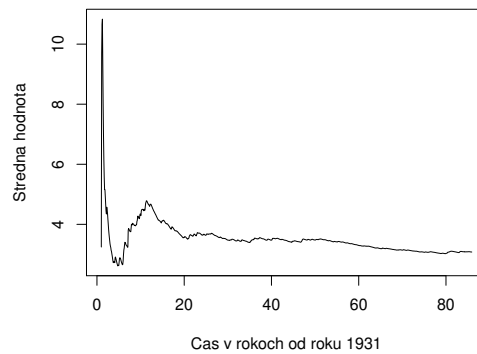


Obr. 46: Prietok Ipľa od januára 1931 do decembra 2015.

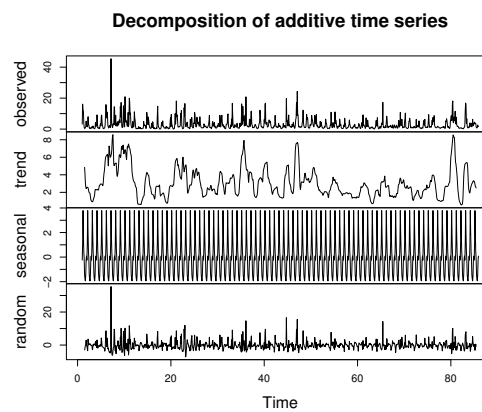
Vývoj strednej hodnoty nášho radu zachytáva Obr. 47 a jeho dekompozíciu Obr. 48.

11.1 Homogenita

Ako ukazujú výsledky testov homogenity zobrazené v Tabuľke 34, oba zamietajú hypotézu, že by sa celý rad riadil rovnakým rozdelením. To znamená, že má zmysel hľadať čas, kedy nastala nejaká zmena.



Obr. 47: Stredná hodnota prietoku Ipľa.



Obr. 48: Dekompozícia prietoku Ipľa.

<i>Test</i>	<i>p – hodnota</i>
Rank Von Neumann (Bartels)	$< 2.2 \cdot 10^{-16}$
Wald-Wolfowitz	$< 2.2 \cdot 10^{-16}$

Tabuľka 34: Testy homogenity prietoku Ipľa.

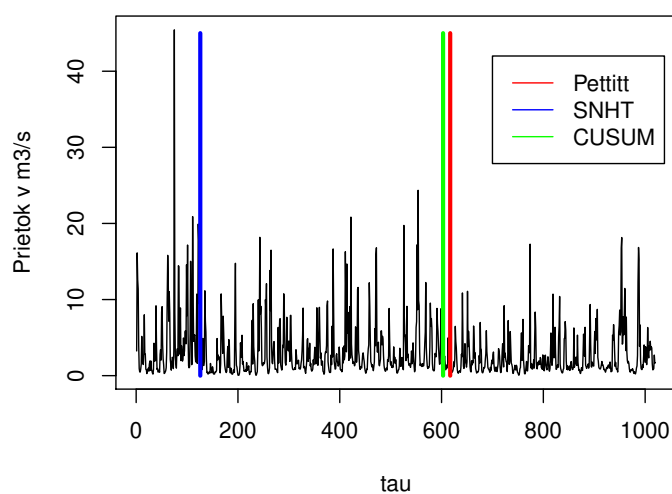
11.2 Bod zmeny

Pozrime sa na Tabuľku 35, aké body zlomu SNHT, Pettittov test a CUSUM metóda našli.

Môžeme vidieť, že jún 1941 nájdený pomocou SNHT a máj 1982 určený Pettittovým testom dané testy považujú za signifikantné. Podľa CUSUM metódy významná zmena strednej hodnoty nastala v marci 1981. Tieto tri body sú zaznačené na Obr. 49.

<i>Test</i>	<i>p-hodnota</i>	τ	<i>Mesiac a rok</i>
SNHT	$< 2.2 \cdot 10^{-16}$	126	Jún 1941
Pettitt	0.0009014	617	Máj 1982
CUSUM	-	603	Marec 1981

Tabuľka 35: Detekcie jedného bodu zmeny prietoku Ipla.



Obr. 49: Body zmeny v prietoku Ipla nájdené pomocou SNHT, Pettittovho testu a CUSUM.

Ich signifikanciu sme overili aj pomocou Mann-Whitney-Wilcoxonovho testu a Kolmogorov-Smirnovovho testu, viď Tabuľka 36.

<i>Test</i>	Jún 1941	Máj 1982	Marec 1981
Mann-Whitney-Wilcoxon	0.0001229	$5.952 \cdot 10^{-5}$	$9.62 \cdot 10^{-5}$
Kolmogorov-Smirnov	0.0001623	0.0002785	0.0004284

Tabuľka 36: Testy signifikantnosti bodov zmeny v prietoku Ipla.

Môžeme si všimnúť, že v prípade všetkých troch nájdených bodov oba testy homogenitu radu zamietli. To znamená, že významnosť všetkých troch bodov potvrdili.

11.3 Viacnásobné body zmeny

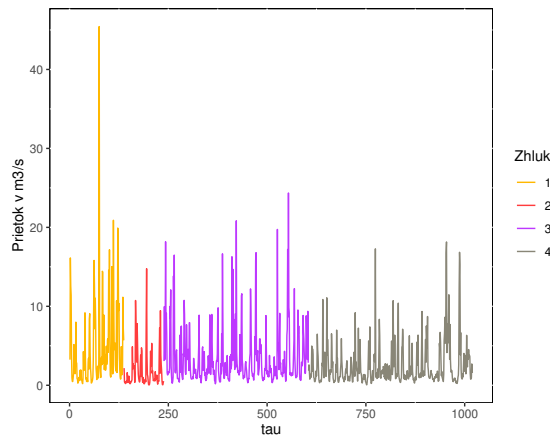
Ďalším krokom analýzy rieky Ipeľ bude hľadanie viacnásobných bodov zmeny.

Začali sme opäť divíznou metódou. Podľa tejto metódy sa rozdelenie tohto radu zmenilo celkovo trikrát. Časy, kedy tieto zmeny nastali, môžeme vidieť v Tabuľke 37. Obr. 50 zachytáva zmeny rozdelenia na pôvodných dátach.

τ	p -hodnota	Mesiac a rok
137	0.001996008	Máj 1942
238	0.001996008	Október 1950
604	0.001996008	Apríl 1981

Tabuľka 37: Viacnásobné zmeny v prietoku Ipeľa nájdené pomocou E-Divisive.

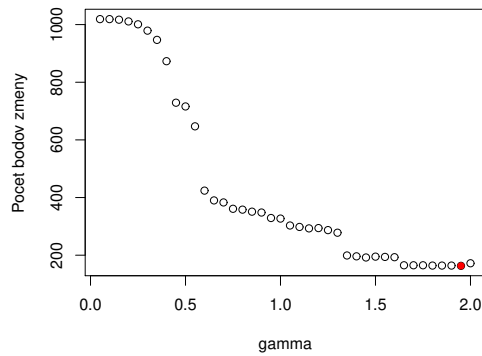
Všimnime si, že jedným z nájdených bodov pomocou E-Divisive je apríl 1981. Podľa neho teda zmena nastala iba o mesiac neskôr ako tá, ktorú našla CUSUM metóda.



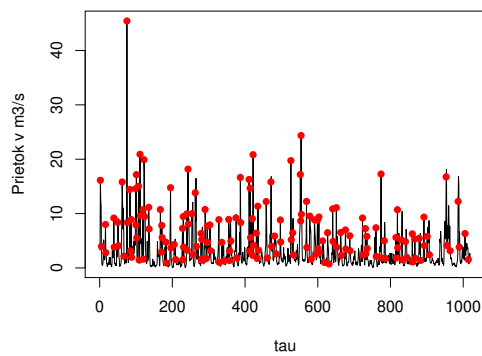
Obr. 50: Rozdelenie radu prietoku Ipeľa nájdenými bodmi pomocou E-Divisive metódy.

Pre aplikáciu aglomeratívneho algoritmu sme opäť najskôr hľadali takú hodnotu parametra γ , pri ktorej bude počet nájdených bodov minimálny. Podľa Obr. 51 sme zvolili $\gamma = 1.95$. Pre túto hodnotu algoritmus našiel 163 bodov, v ktorých sa podľa neho zmenilo rozdelenie radu. Tieto body môžeme vidieť na Obr. 52.

Okrem hierarchických metód sme viacnásobné zlomy hľadali aj pomocou iteračného Pettittovho testu a iteračného SNHT.



Obr. 51: Počet nájdených bodov zmeny pomocou aglomeratívneho algoritmu v závislosti od hodnoty γ .



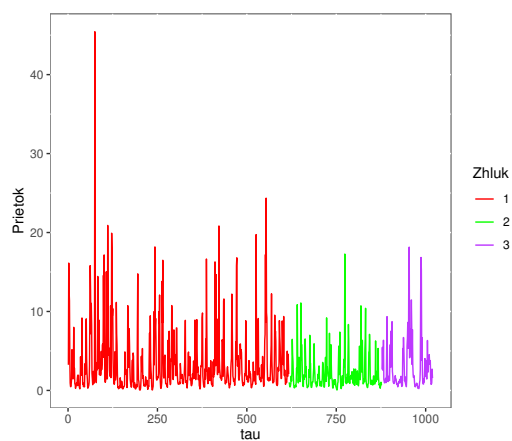
Obr. 52: Rozdelenie radu prietoku Iplá nájdenými bodmi pomocou aglomeratívnej metódy.

Pettittov test našiel okrem mája 1982 nájdeného v predchádzajúcej časti tejto kapitoly ešte jeden bod, vid' Tabuľka 38. Graficky tieto body rozdelia náš rad ako na Obr. 53.

τ	<i>Mesiac a rok</i>
617	Máj 1982
877	Január 2004

Tabuľka 38: Nájdené body zmeny prietoku Iplá pri iteračnom aplikovaní Pettittovho testu.

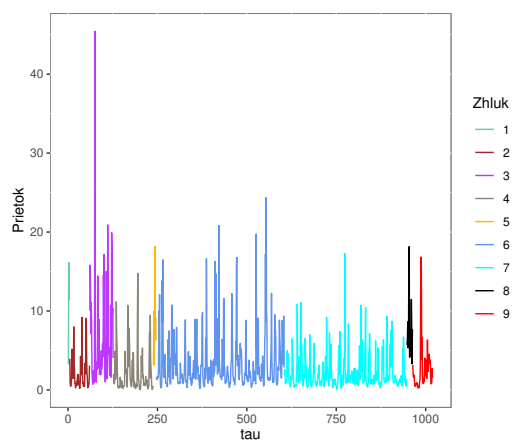
SNHT bol pri hľadaní viacnásobných bodov úspešnejší. K júnu 1941 sme ďalšími iteráciami našli ešte ďalších sedem zmien zapísaných v Tabuľke 39. Tieto body nám rad rozdelia na deväť zhlukov ako ukazuje Obr. 54.



Obr. 53: Rozdelenie radu nájdenými bodmi pomocou iteračného aplikovania Pettittovho testu.

τ	<i>Mesiac a rok</i>
4	Apríl 1931
60	December 1935
126	Jún 1941
238	Október 1950
246	Jún 1951
604	Apríl 1981
947	November 2009
963	Marec 2011

Tabuľka 39: Nájdené body zmeny prietoku Iplu pri iteračnom aplikovaní SNHT.



Obr. 54: Rozdelenie radu nájdenými bodmi pomocou iteračného aplikovania SNHT.

Všimnime si, že z bodov nájdených pomocou E-Disisive sú dva, ktoré sme našli aj pomocou iteračného SNHT. SNHT označil za prelomový aj apríl 1981, čiže nasledujúci mesiac za bodom, ktorý označila CUSUM metóda.

11.4 Trend

Poslednou časťou analýzy prietoku rieky Ipeľ je zisťovanie prítomnosti trendu. Ako aj v predošlých kapitolách, použili sme sezónny Mann-Kendallov test a Cox-Stuartov test. Ich zistenia nájdeme v Tabuľke 40.

<i>Test</i>	<i>p-hodnota</i>
Sezónny Mann-Kendall	0.049216
Cox-Stuart	0.2318

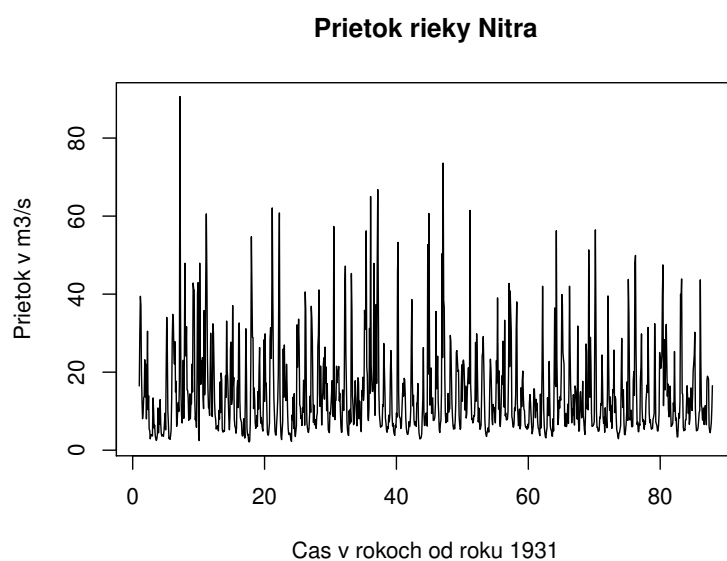
Tabuľka 40: Obojstranné testy prítomnosti trendu pre prietok Ipeľa.

Mann-Kendallov test neprítomnosť trendu veľmi tesne zamietá, teda podľa neho trend v rade je. Avšak Cox-Stuartov test neprítomnosť trendu nezamietá.

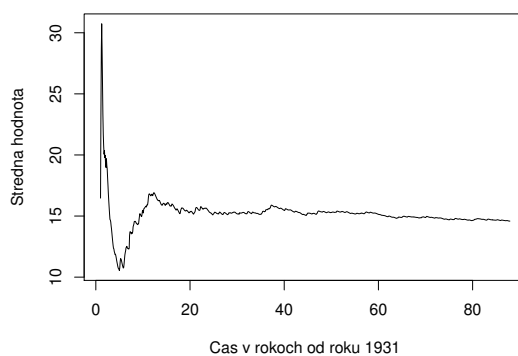
12 Analýza rieky Nitra

Poslednou riekou analyzovanou v tejto práci je rieka Nitra. Jej prameň sa nachádza pod vrchom Reváň v pohorí Malá Fatra. Jej tok je dlhý 191km a jej povodie zaberá rozlohu 5140km^2 . Pri obci Komoča (okres Nové Zámky) sa vlieva do Váhu. [13]

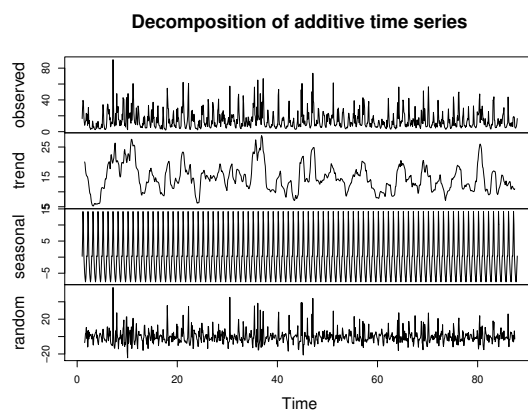
Dáta zozbierané na stanici v obci Nitrianska Streda (okres Topoľčany) za obdobie od januára 1931 do decembra 2017 sú zobrazené na Obr. 55. Strednú hodnotu tohto radu môžeme vidieť na Obr. 56 a dekompozíciu na Obr. 57.



Obr. 55: Prietok Nitry od januára 1931 do decembra 2017.



Obr. 56: Stredná hodnota prietoku Nitry.



Obr. 57: Dekompozícia prietoku Nitry.

12.1 Homogenita

Prvou časťou našej analýzy je ako aj v predošlých kapitolách testovanie homogenity.

<i>Test</i>	<i>p – hodnota</i>
Rank Von Neumann (Bartels)	$< 2.2 \cdot 10^{-16}$
Wald-Wolfowitz	$< 2.2 \cdot 10^{-16}$

Tabuľka 41: Testy homogenity prietoku Nitry.

Podľa Tabuľky 41 vidíme, že homogenita radu je zamietnutá pomocou oboch testov. Má teda význam hľadať nejakú zmenu v rade.

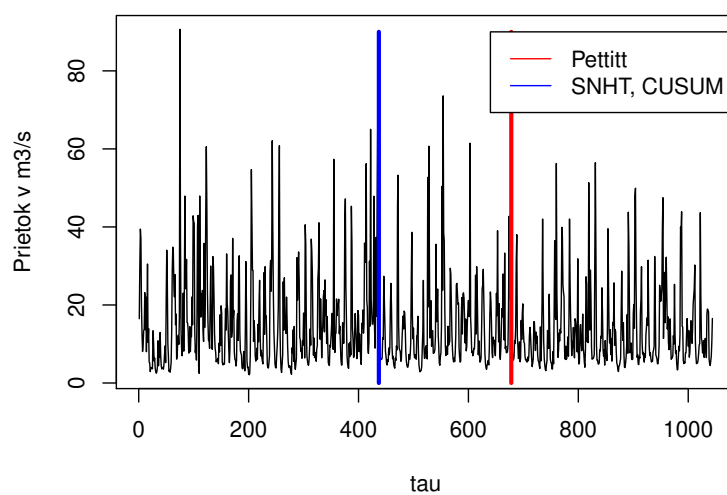
12.2 Bod zmeny

Na nájdenie zmien v strednej hodnote budeme opäť využívať SNHT, Pettittov test a CUSUM metódu.

<i>Test</i>	<i>p-hodnota</i>	τ	<i>Mesiac a rok</i>
SNHT	0.08705	437	Máj 1967
Pettitt	0.1451	678	Jún 1987
CUSUM	-	437	Máj 1967

Tabuľka 42: Detekcie jedného bodu zmeny prietoku Nitry.

Z Tabuľky 42 vidíme, že SNHT a CUSUM našli rovnaký bod - máj 1967. SNHT však tento bod vyhodnotil tesne ako nevýznamný na 5%-nej hladine významnosti. Podľa Pettittovho testu je potenciálnym bodom zmeny jún 1987. Tento test si však myslí, že nájdený bod je taktiež nesignifikantný. Nájdené časy zmeny môžeme vidieť na Obr. 58.



Obr. 58: Body zmeny v prietoku Nitry nájdené pomocou SNHT, Pettittovho testu a CUSUM.

Zisťovali sme, čo si o potenciálnych bodoch myslia testy zameriavajúce sa na signifikanciu bodov zmeny. Ich výsledky môžeme nájsť v Tabuľke 43.

<i>Test</i>	Máj 1967	Jún 1987
Mann-Whitney-Wilcoxon	0.02246	0.01639
Kolmogorov-Smirnov	0.0007312	0.006168

Tabuľka 43: Testy signifikantnosti bodov zmeny v prietoku Nitry.

Napriek tomu, že testy, ktoré dané body našli, ich signifikanciu zamietli, Mann-Whitney-Wilcoxonov test aj Kolmogorov-Smirnovov test si v prípade oboch bodov myslia presný opak. Keďže sú ich p-hodnoty menšie ako 5%, tieto testy považujú zmeny, ktoré sa udiali v máji 1967 a júni 1987, za signifikantné.

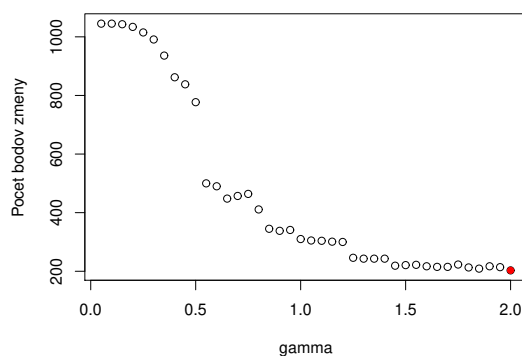
12.3 Viacnásobné body zmeny

Keďže body, ktoré našli SNHT a Pettittov test, sú podľa daných metód nesignifikantné, pri hľadaní viacnásobných bodov považujeme za zbytočné aplikovať tieto testy iteratívne. Nakoľko si testy určujúce signifikanciu nájdených bodov myslia presný opak, viacnásobné body sa pokúsime nájsť aspoň pomocou hierarchických metód.

Začneme metódou, ktorá sa nám na základe výsledkov v predošlých kapitolách javí ako spoľahlivejšia - divízna metóda E-Divisive.

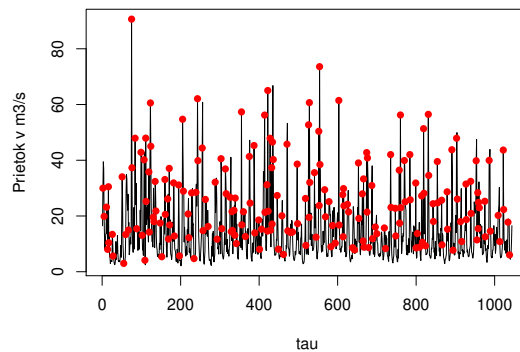
Divízna metóda nenašla ani jeden bod zmeny, ktorý by bol signifikantný. Našla však bod v čase $\tau = 437$. Jeho p-hodnota je 0.06786427, teda je nevýznamný. Môžeme tak konštatovať, že náš rad je z pohľadu divíznej metódy homogénny, inými slovami - celý rad sa riadi rovnakým rozdelením.

Pozrime sa, čo si o viacnásobných bodoch myslí aglomeratívny algoritmus. Opäť najskôr nájdeme hodnotu γ , ktorá bude minimalizovať počet nájdených bodov. Na Obr. 59 vidíme, že takou hodnotou je $\gamma = 2$. V tomto prípade algoritmus našiel 203 bodov. Ich grafické zobrazenie môžeme vidieť na Obr. 60.



Obr. 59: Počet nájdených bodov zmeny pomocou aglomeratívneho algoritmu v závislosti od hodnoty γ .

Ako sme spomínali na začiatku tejto podkapitoly, aplikácia iteratívneho SNHT a Pettittovho testu je zbytočná, keďže signifikancia nájdených bodov je zamietnutá hneď v prvej iterácii.



Obr. 60: Rozdelenie radu prietoku Nitry nájdenými bodmi pomocou aglomeratívnej metódy.

12.4 Trend

Analýzu prietoku rieky Nitra ukončíme hľadaním monotónneho trendu. Závěry testov zaoberajúcich sa trendom vidíme v Tabuľke 44.

<i>Test</i>	<i>p-hodnota</i>
Sezónny Mann-Kendall	0.067279
Cox-Stuart	0.9651

Tabuľka 44: Obojstranné testy prítomnosti trendu pre prietok Nitry.

Cox-Stuartov test si myslí, že v rade sa trend nenachádza. Mann-Kendallov test prišiel k rovnakému záveru, avšak iba veľmi tesne. Preto nevidíme dôvod robiť jednostranné Cox-Stuartove testy.

Môžeme z toho vyvodiť, že v rade sa dlhodobé zmeny nenachádzajú.

13 Vzťah medzi analyzovanými radmi mesačných prietokov

Po analýze všetkých riek je na mieste otázka, či a ako veľmi sa analyzované rieky navzájom ovplyvňujú.

Na posúdenie vzájomného vplyvu jednej rieky na druhú nám slúži korelácia.

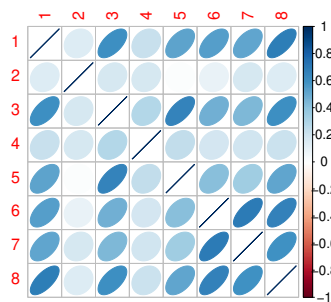
Keďže v našej práci využívame neparametrické metódy, na výpočet korelácie sme použili Kendallov korelačný koeficient (32) definovaný v časti 4.1.

Keďže máme ku každej rieke k dispozícii dáta za rôzne obdobia, pri výpočte korelácie budeme používať dáta za obdobie, ktoré majú spoločné. Avšak rieka Belá je horskou riekou, teda si môžeme všimnúť, že jej priemerný mesačný prietok je najvyšší v máji - júni, pričom pri ostatných riekach je maximálny v marci - apríli. Z týchto dôvodov budeme výpočet korelácie aplikovať na dátach za obdobie od januára 1931 do októbra 2014 pri všetkých riekach až na Belú, pri tejto rieke budeme používať dáta pochádzajúce z obdobia od marca 1931 do decembra 2014. Korelácie medzi jednotlivými riekami sú zobrazené v Tabuľke 46.

Rieka	Morava	Hron	Váh	Belá	Kysuca	Krupinica	Ipeľ	Nitra
Morava	1.00	0.15	0.60	0.22	0.52	0.54	0.51	0.68
Hron	0.15	1.00	0.17	0.17	0.01	0.09	0.17	0.14
Váh	0.60	0.17	1.00	0.28	0.65	0.47	0.44	0.60
Belá	0.22	0.17	0.28	1.00	0.24	0.18	0.19	0.21
Kysuca	0.52	0.01	0.65	0.24	1.00	0.41	0.34	0.51
Krupinica	0.54	0.09	0.47	0.18	0.41	1.00	0.69	0.66
Ipeľ	0.51	0.17	0.44	0.19	0.34	0.69	1.00	0.59
Nitra	0.68	0.14	0.60	0.21	0.51	0.66	0.59	1.00

Tabuľka 45: Korelácie riek pomocou Kendallovho korelačného koeficientu.

Na Obr. 61 vidíme grafické zobrazenie vypočítaných korelácií. Čísla 1 – 8 predstavujú rieky, pričom sú zoradené v poradí, v akom sme ich analyzovali v tejto práci. To znamená, že číslo 1 zastupuje Moravu, číslo 2 patrí Hronu, atď. až číslo 8 predstavuje Nitru.



Obr. 61: Korelácie mesačných radov prietokov riek.

Kladná hodnota korelačného koeficientu znamená priamoúmerný vzťah, záporná hodnota nepriamoúmerný vzťah. Ak je hodnota korelačného koeficientu 0, premenné (rieky) sú nezávislé. Čím vyššia, resp. v prípade zápornej korelácie nižšia hodnota, tým viac sa premenné ovplyvňujú. Logicky, korelácia rieky samej so sebou bude mať hodnotu korelačného koeficientu 1.

V našej analýze by sme očakávali, že korelácia medzi riekami, ktoré do seba ústia, bude relatívne vysoká.

Môžeme si všimnúť, že záporná hodnota korelačného koeficientu sa v Tabuľke 46 nenachádza. Taktiež sa v nej nenachádza ani jedna dvojica, ktorá by bola nezávislá.

Najvyššiu hodnotu korelačného koeficientu môžeme vidieť pri dvojici Krupinica - Ipeľ. Tento fakt je logicky zdôvodniteľný tým, že Krupinica sa vlieva do Ipľa. To znamená, že so zmenou prietoku Krupinice sa zmení prietok Ipľa.

Na počudovanie, druhú najvyššiu koreláciu má dvojica Morava - Nitra. Medzi týmito riekami nie je žiadny vzťah rovnako ako pri porovnaní Nitry s Krupinicou a Moravy s Váhom.

Podme sa pozrieť, ako dopadli korelácie medzi riekami, pri ktorých jedna je prítokom druhej. Rieka Belá je prítokom rieky Váh, avšak hodnota ich korelácie je relatívne nízka. To znamená, že prietok Belej ovplyvňuje prietok Váhu na nízkej úrovni. Môžeme si však všimnúť, že Belá má najvyššiu koreláciu práve s Váhom. Do Váhu sa vlieva aj rieka Kysuca. V tomto prípade tomu zodpovedá aj ich korelačný koeficient - patrí medzi najvyššie. Ďalším prítokom Váhu je rieka Nitra. Vzájomná korelácia týchto dvoch riek taktiež patrí medzi vyššie hodnoty v našej tabuľke. Môžeme teda povedať, že Kysuca aj

Nitra ovplyvňujú prietok Váhu vo vyššej miere ako ho ovplyvňuje Belá.

Skúsme sa teraz pozrieť, či sa rieky, ktoré by sa mali navzájom ovplyvňovať, skutočne ovplyvňujú aj z pohľadu zmien. Pozrime sa, či v období, keď sa zmenilo rozdelenie jednej rieky, sa nejakým spôsobom zmenila aj druhá rieka.

Najskôr sa pozrime na logický vplyv - jedna rieka je prítokom druhej.

Začnime s riekou Belá. Tá je prítokom Váhu. Body zmien týchto dvoch riek sú diametrálne odlišné, aj keď prihliadneme na posun spôsobený horským charakterom rieky. Spoločnou črtou je, že hoci sme nejaké body našli, väčšina z nich bola vyhodnotená ako nesignifikantná. Ďalšou spoločnou črtou je, že ani u jednej z nich sme nenašli monotónny trend.

Ďalším prítokom Váhu je rieka Kysuca. Pri porovnaní výsledkov analýzy oboch riek vidíme, že Pettittov test identifikoval rovnaký čas, kedy nastala zmena strednej hodnoty - máj 1983. Avšak pri oboch riekach bol tento bod vyhodnotený za nevýznamný. Žiadnej zo zvyšných metód sa nepodarilo nájsť rovnaký bod. Závery ohľadom trendu sú pre tieto dve rieky taktiež zhodné - podľa nich sa trend v radoch nenachádza. Vidíme, že zrejme Kysuca čiastočne ovplyvňuje Váh, čo potvrdil aj korelačný koeficient.

Do Váhu sa vlieva ešte rieka Nitra. Pri analýze Nitry sme našli iba dva body, a to máj a jún 1967. Môžeme si všimnúť, že jedine CUSUM metóda našla pri analýze Váhu bod, ktorý je blízky bodu zlomu Nitry. Taktiež sa zhoduje záver v prípade trendu.

Krupinica vteká do rieky Ipeľ. Tieto dve rieky majú medzi sebou najvyššiu koreláciu spomedzi testovaných riek. Očakávame teda, že sa budú najviac ovplyvňovať, resp. budú mať najviac rovnakých alebo blízky bodov. Naše očakávanie splňajú iba body nájdené pomocou SNHT. Táto metóda našla jeden rovnaký bod - apríl 1931, a dvojicu blízky bodov - november 1935 (Krupinica) a december 1935 (Ipeľ). V otázke trendu vplyv prietoku Krupinice na prietok Ipeľa nevieme porovnať, pretože závery použitých testov sa v prípade Ipeľa odlišujú.

Nakoniec porovnajme dvojice, pri ktorých je vysoká hodnota korelácie, ale nemajú žiaden logický vzťah. Začneme s dvojicou Morava - Nitra. Všimnime si, že napriek tomu, že korelácia medzi týmito riekami je druhou najvyššou v rámci tejto práce, výsledky analýz ich prietokov nemajú spoločný žiaden bod ani záver.

Medzi dvojice s vyššou hodnotou korelácie patrí Nitra - Krupinica. V prípade týchto

Rieka	SNHT	Pettitt	CUSUM	E-Divisive	Trend
Morava	Apríl 1900, Máj 1927, Február 1937, Marec 1937, Máj 1942, Marec 1948, Apríl 2015	Máj 1921, Apríl 1932, November 1935, December 1938, Jún 1942, Október 1964, Máj 1988, November 1994	Apríl 1948	Máj 1927, Január 1937, Marec 1946, November 1954, Apríl 1988	↘
Hron	August 1981	August 1981	August 1981	Apríl 1985	↘
Váh	Máj 1921, Február 1922	Máj 1983	Jún 1967	Jún 1951	-
Belá	Apríl 1931	September 1955	-	-	-
Kysuca	Jún 1934	Máj 1983	Marec 2002	-	-
Krupinica	Apríl 1931, November 1935, Máj 1942, Február 1965 , Apríl 1967	Jún 1967, Jún 1932, November 1935, Jún 1942, September 1964	Apríl 1967	Marec 1946, Apríl 1967,	↘
Ipeľ	Apríl 1931, December 1935, Jún 1941, Október 1950, Jún 1951, Apríl 1981, November 2009 Marec 2011	Máj 1983, Január 2004	Marec 1981	Máj 1942, Október 1950, Apríl 1981	-
Nitra	Máj 1967	Jún 1967	Máj 1967	-	-

Tabuľka 46: Korelácie riek pomocou Kendallovho korelačného koeficientu.

dvoch riek vidíme akúsi príbuznosť iba v bodoch nájdených pomocou CUSUM metódy. Iná zhoda alebo podobnosť v záveroch týchto riek nie je.

Poslednou dvojicou, ktorú budeme porovnávať je Morava a Váh. Ako aj v prípade predošlej dvojice, tak aj pri tejto sme našli iba jeden bod, ktorý je rovnaký pri oboch riekach. Tým bodom je Máj 1921. Inú spoločnú charakteristiku tieto dve rieky vzhľadom na našu analýzu nemajú.

Záver

V tejto práci sme sa zaoberali analýzou časových radov priemerných mesačných prietokov vybraných riek na Slovensku. Rieky boli vybrané na základe dĺžky pozorovaní vodných stavov. Na túto analýzu sme využívali neparametrické štatistické metódy.

Skôr ako sme sa dostali k samotnej analýze, uviedli sme využité metódy. Boli to najmä metódy zaoberajúce sa zmenou strednej hodnoty v radoch, či už sa jednalo o jednu zmenu, alebo o dlhodobú monotónnu zmenu, teda trend.

V praktickej časti sme analyzovali prietoky riek Morava, Hron, Váh, Belá, Kysuca, Krupinica, Ipel a Nitra. Na túto analýzu sme využili funkcie implementované v nadstavbe štatistického programu R - RStudio vo verzii 1.1.463.

Začali sme analýzou rieky Morava. Pomocou SNHT sme získali sedem bodov, kedy sa zmenilo rozdelenie - apríl 1900, máj 1927, február 1937, marec 1937, máj 1942, marec 1948, apríl 2015. Pettittov test našiel o jeden bod viac ako SNHT - máj 1921, apríl 1932, november 1935, december 1938, jún 1942, október 1964, máj 1988, november 1994. CUSUM metódu sme neaplikovali iteračne, našla iba jeden zlom, a to v apríli 1948. Divízna metóda, určená na hľadanie viacnásobných bodov zmeny rozdelenia, našla celkovo päť bodov - máj 1927, január 1937, marec 1946, november 1954, apríl 1988. Pri určení niektorých bodov sa použité metódy v niektorých prípadoch priblížili k rovnakému záveru, prípadne sa zhodli. Príkladmi takýchto záverov sú body v máji 1927 alebo marec-apríl 1948. Z pohľadu dlhodobých zmien, inak povedané trendu, metódy na jeho detekciu usudzujú, že sa v tomto rade nachádza klesajúci trend. Analýzu rieky Morava tak môžeme uzavrieť konštatovaním, že mesačný prietok tejto rieky podlieha relatívne často zmenám, no v posledných rokoch nie tak výrazne, ako tomu bolo v minulosti. Z dlhodobého hľadiska však jej prietok klesá.

Ďalšou analyzovanou riekou bol Hron. V prípade tejto rieky sa použité metódy zhodli na rovnakom závere - rozdelenie, ktorým sa riadi prietok Hrona sa zmenilo iba jedenkrát, a to v auguste 1981. Jediný rozdielny bod našla iba E-Divisive, ktorá s nimi súhlasí v počte bodov, no podľa nej táto zmena nastala v apríli 1985. Napriek nízkemu počtu jednorázových zmien, sme však zistili, že aj v tomto rade sa nachádza klesajúci trend. Z pohľadu jednorázových zmien môžeme povedať, že až na jeden moment je prietok tejto rieky stabilný, avšak z dlhodobého pohľadu má klesajúcu tendenciu.

Pri analýze Váhu sme pomocou SNHT našli dva body - máj 1921 a február 1922. Máj 1921 sme však našli až pri iteračnej aplikácii, pričom sme február 1922 pokladali za signifikantný, aj napriek tomu, že SNHT ho pokladá za nesignifikantný. Nevýznamnú zmenu našiel aj Pettittov test v máji 1983. V prípade júna 1967, nájdeného pomocou CUSUM, sa testy signifikancie v záveroch odlišujú. Divízna metóda našla taktiež iba jeden bod - jún 1951. Narozdiel od Moravy alebo Hrona, prietok Váhu neklesá. Môžeme teda tvrdiť, že rieka Váh je relatívne stabilná.

Prietok rieky Belá sa podľa SNHT zmenil v apríli 1931 a podľa Pettittovho testu v septembri 1955. Oba testy tieto zmeny pokladajú za nevýznamné. Zvyšné metódy nenašli žiaden zlom a testy na detekciu trendu jeho prítomnosť taktiež zamietli. Z nášho pohľadu je rieka Belá, aj napriek svojimi horskému charakteru, stabilná.

Rovnaký záver priniesla aj analýza rieky Kysuca. Pomocou SNHT sme síce označili za zlomový jún 1934, Pettittovým testom to bol máj 1983 a CUSUM našla marec 2002, no všetky tieto body boli vyhodnotené ako nesignifikantné.

Ďalšou riekou, ktorej prietok sme analyzovali, je rieka Krupinica. SNHT našiel celkovo päť jednorázových bodov zmeny - apríl 1931, november 1935, máj 1942, február 1965 a apríl 1967. Pettittovým testom sme našli taktiež päť bodov - jún 1967, jún 1932, november 1935, jún 1942 a september 1964. CUSUM metódou sme získali rovnaký bod ako pomocou SNHT - apríl 1967. Tento bod sme našli aj pomocou divíznej metódy. Tá k nemu označila za prelomový aj marec 1946. Podľa testov zameriavajúcich sa na detekciu prítomnosti trendu sa okrem jednorázových zmien v rade nachádza aj klesajúci trend. Analýzu tejto rieky môžeme uzavrieť konštatovaním, že jej rozdelenie sa relatívne častokrát zmenilo a okrem toho jej prietok klesá.

Pri analyzovaní prietoku rieky Ipeľ sme pomocou SNHT našli až osem bodov, kedy sa jej prietok zmenil. Tými bodmi sú apríl 1931, december 1935, jún 1941, október 1950, jún 1951, apríl 1981, november 2009 a marec 2011. Pettittov test našiel iba dva body - máj 1983 a január 2004. Jeden bod našla CUSUM metóda - marec 1981. Týmto bodom sa priblížila k bodu nájdenému pomocou CUSUM. Divízna metóda označila za zlomové tri body, pričom novým z nich je iba máj 1942. Okrem neho našla ešte október 1950 a apríl 1981, ktoré boli nájdené aj pomocou SNHT. Podľa Mann-Kendallovho testu prietok tejto rieky klesá, avšak Cox-Stuartov test nevylúčil možnosť neprítomnosti trendu.

Poslednou analyzovanou riekou je rieka Nitra. SNHT spolu s CUSUM našli rovnaký bod, a to máj 1967. Pettittov test označil za prelomový nasledujúci mesiac, teda jún 1967. Napriek tomu, že metódy, ktoré dané body našli, ich významnosť zamietli, testy zameriavajúce sa na určenie signifikancie si myslia presný opak. Divízna metóda v prietoku tejto rieky nenašla žiadnu zmenu rozdelenia. Taktiež v jej prietoku nebol detekovaný žiadny trend. Môžeme teda povedať, že rieka Nitra je až na obdobie máj-jún 1967 stabilná.

Pri hľadaní viacnásobných zlomov sme v prípade všetkých analyzovaných riek využívali aj aglomeratívnu metódu. Táto metóda však pri všetkých analýzach našla až priveľké množstvo bodov. Z tohto dôvodu sme jej výsledky do záveru nezahrnuli.

Praktickú časť tejto práce sme uzavreli analýzou vzájomných vzťahov prietokov spomínaných riek a porovnaním výsledkov ich analýz. Dospeli sme k záveru, že zmena prietoku rieky, ktorá sa vlieva do inej rieky, môže prietok tejto rieky čiastočne ovplyvniť. Vplyv zmeny však nemusí byť veľmi výrazný. Taktiež sme zistili, že aj napriek vyššej hodnote korelácie, sa zmeny v prietokoch daných riek nemusia navzájom ovplyvňovať.

Z hydrologického hľadiska sú nájdené zmeny veľmi užitočné. Znamenáť môžu zmeny spôsobené výmenou prístrojovej techniky v priebehu rokov, ale aj skutočné zmeny režimu odtoku v dôsledku zalesnenia/odlesnenia (napr. v prípade rieky Belá). Rovnako sa môže jednať o zmenu mesačného režimu v dôsledku výstavby vodných nádrží (napr. v povodí Moravy alebo Váhu v Šali). Tieto výsledky budú slúžiť hydrológom na ďalšiu podrobnú hydrologickú analýzu.

Zoznam použitej literatúry

- [1] Aalders, P. Warmerdam, P.M.M. Torfs, P.J.J.F.: Rainfall Generator for the Meuse Basin (2004), dostupné na internete (14.01.2020): <https://edepot.wur.nl/212001>
- [2] Ahmed, K., Shahid, S., Ismail, T., Nawaz, N., Wang, X.: Absolute homogeneity assessment of precipitation time series in an arid region of Pakistan, *Atmosfera* 31 (2018), 301-316, dostupné na internete (14.01.2020): <https://pdfs.semanticscholar.org/4443/e6007b9fd144f067929803942351a8ebc290.pdf>
- [3] Ballová, D.: Modelling of Nonlinear Time Series With Applications In Hydrometeorology, Written part of disertatation examination, Faculty of Civil Engineering STU, Bratislava, 2019
- [4] Bartels, R.: The Rank Version of von Neumann's Ratio Test for Randomness, *Journal of the American Statistical Association* 77 (1982), 40-46, dostupné na internete (14.01.2020): <https://www.semanticscholar.org/paper/The-Rank-Version-of-von-Neumann's-Ratio-Test-for-Bartels/5f2170e8261aeeab45a82eab2d515106bc391511>
- [5] Conover, W. J.: Practical Nonparametric Statistics, New York: John Wiley and Sons, 1971, p. 309-314.
- [6] Drápela, K., Drápelová, I. : Application of Mann-Kendall test and the Sen's slope estimates for trend detection in deposition data from Bílý Kříž (working paper), Mendelova univerzita v Brně, 2011, dostupné na internete (14.01.2019): <https://pdfs.semanticscholar.org/df9d/08a2ec398f722048abb88e148e6468607b2f.pdf>
- [7] Helsel, D.R., Hirsch, R.M.: Statistical Methods in Water Resources, Elsevier Sci Pub. Co., 1992, 213-214, 238-240, dostupné na internete (14.01.2020): <https://www.semanticscholar.org/paper/Statistical-methods-in-water-resources-Helsel-Hirsch/a89c397230aa22dbe7be49f87f66e968266bd0b1>
- [8] Hollander, M., Wolfe, D.A., Chicken, E.: Nonparametric Statistical Methods, John Wiley and Sons, Hoboken, New Jersey, 2014
- [9] Infoglobe, dostupné na internete (12.04.2020): <https://www.infoglobe.sk/rieky/europa/slovenska-republika/bela-liptovska/>

- [10] Infoglobe, dostupné na internete (14.01.2020):
<https://www.infoglobe.sk/rieky/europa/slovenska-republika/hron/>
- [11] Infoglobe, dostupné na internete (18.04.2020):
<https://www.infoglobe.sk/rieky/europa/slovenska-republika/ipel/>
- [12] Infoglobe, dostupné na internete (14.04.2020):
<https://www.infoglobe.sk/rieky/europa/slovenska-republika/kysuca/>
- [13] Infoglobe, dostupné na internete (20.04.2020):
<https://www.infoglobe.sk/rieky/slovenska-republika/europa/nitra/>
- [14] James, N.: Multiple Change Point Analysis od Multivariate Data Via Energy Statistics, Dissertation Thesis, Faculty of the Graduate School of Cornell University, New York, 2015, dostupné na internete (05.05.2020):
<https://davidmattesontstatistics.files.wordpress.com/2016/03/nicholas-james-thesis.pdf>
- [15] Kahya, E., Bickici Arikan, B., Akdeniz, E.: Homogeneity Analysis of Precipitation Series in Turkey (Conference Paper), Turkey, 2016
- [16] Leander, R., Buishand, T.A.: Rainfall Generator for the Meuse Basin : Inventory and homogeneity analysis of long daily precipitation records, De Bilt, Wilhelminalaan 10, The Netherlands, 2004, dostupné na internete (14.01.2020):
<https://www.knmi.nl/kennis-en-datacentrum/publicatie/rainfall-generator-for-the-meuse-basin-inventory-and-homogeneity-analysis-of-long-daily-precipitation-records>
- [17] Lehtinen, E., Pulkkinen, U., Pörn K.: Statistical Trend Analysis Methods for Temporal Phenomena, Skivlingvagen 24, SE-611 63 Nykoping, Sweden, 1997
- [18] Nachar, N.: The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution, Tutorials in Quantitative Methods for Psychology vol. 4 (2008), 13-20, dostupné na internete (14.01.2020):
<http://tqmp.org/RegularArticles/vol04-1/p013/p013.pdf>

- [19] Pettitt, A. N.: A non-parametric approach to the change-point problem, Appl. Statist. 28 (1979), 126–135
- [20] Pohlert, T.: trend: Non-Parametric Trend Tests and Change-Point Detection, R package version 0.0.1 (Technical Report), 2018, dostupné na internete (14.01.2020): <https://cran.r-project.org/web/packages/trend/vignettes/trend.pdf>
- [21] Sládek, M.: Návrh plánu povodňové ochrany obce Kopčany (diplomová práce), Univerzita Tomáše Bati ve Zlíně, Fakulta aplikované informatiky, Zlín, 2018, dostupné na internete (14.01.2020): <http://digilib.k.utb.cz/handle/10563/43364>
- [22] Slovenský Trip, dostupné na internete (16.04.2020): <http://www.slovakiatrip.sk/destinacie/Rieka-Krupinica?oid=132>
- [23] SPOZNAJ.EU, dostupné na internete (14.01.2020): <http://www.spoznaj.eu/riekavah>
- [24] Wald, A., Wolfowitz, J. : On a test whether two samples are from the same population, The Annals of Mathematical Statistics, 11(1940), 147-162, dostupné na internete (14.01.2020): <https://projecteuclid.org/euclid.aoms/1177731909>

Príloha A

Listing 1: Morava

```
1 ##### Instalacia balickov a kniznic #####
2   install.packages("openxlsx")
3   install.packages("moments")
4   install.packages("trend")
5   install.packages("Kendall")
6   install.packages("randtests")
7   install.packages("changepoint")
8   install.packages("ecp")
9   install.packages("zoo")
10  install.packages("ggplot2")
11  install.packages("paletteer")
12  library("openxlsx")
13  library("moments")
14  library("trend")
15  library("Kendall")
16  library("randtests")
17  library("changepoint")
18  library("ecp")
19  library("zoo")
20  library("ggplot2")
21  library("paletteer")
22  #####
23  morava_pov <- read.xlsx("Data.xlsx", 1, colNames=TRUE)
24  morava <- ts(morava_pov[,3], frequency = 12)
25  plot.ts(morava, xlab='Cas v rokoch od roku 1900', ylab='Prietok v m3/s') #
      Vykreslenie dat
26  title(main = 'Prietok rieky Morava')
27  # Vypocet vyvoja strednej hodnoty
28  str_hod_morava <- rep(0, length(morava))
29  for (i in 1:length(morava)) {
30    str_hod_morava[i] <- mean(morava[1:i])
31  }
32  # Vykreslenie strednej hodnoty
33  plot(ts(str_hod_morava, frequency = 12), xlab="Cas v rokoch od roku 1900",
```

```

      ylab="Stredna hodnota")
34
35 # Dekompozicia
36 morava_dec <- decompose(ts(morava_pov[,3], frequency = 12), type = "additive
    ")
37 plot(morava_dec)
38
39 ## Homogenita ##
40 brt_morava <- bartels.rank.test(morava)
41 ww_morava <- ww.test(morava)
42 print(brt_morava) # p-value < 2.2e-16
43 print(ww_morava) # p-value < 2.2e-16
44
45 ## Single Change Point ##
46 pettitt_morava <- pettitt.test(morava)
47 print(pettitt_morava) # K=257, p-value = 0.01247
48
49 snht_morava <- snh.test(morava)
50 print(snht_morava) # K=4, p-value = 1e-04
51
52 cusum_morava <- cpt.mean(morava, penalty = "Manual", pen.value = 0.05,
    method = "AMOC", test.stat = "CUSUM", class = TRUE, param.estimates =
    TRUE)
53 morava_cusum <- cpts(cusum_morava) # 581, vystupom je najblizsi bod po
    zmene => zmena je v 580
54
55 plot.ts(morava_pov[,3], xlab= 'tau', ylab='Prietok v m3/s')
56 lines(x=rep(pettitt_morava$estimate, length(0:max(morava_pov[,3]))), y=0:
    max(morava_pov[,3]), col='red', lwd=3)
57 lines(x=rep(snht_morava$estimate, length(0:max(morava_pov[,3]))), y=0:max(
    morava_pov[,3]), col='blue', lwd=3)
58 lines(x=rep(morava_cusum-1, length(0:max(morava_pov[,3]))), y=0:max(morava_
    pov[,3]), col='green', lwd=3)
59 legend(900, 738, legend=c("Pettitt", "SNHT", "CUSUM"), col = c("red", "blue"
    , "green"), lty = 1)
60
61 ## Testovanie najdenych bodov zmeny ##
62 morava_pred_pettitt <- morava[1:pettitt_morava$estimate] # data do bodu

```



```

      zmeny
63 morava_po_pettitt <- morava[(pettitt_morava$estimate+1):length(morava)] #
      data po zmene
64
65 morava_pred_snht <- morava[1:snht_morava$estimate] # data do bodu zmeny
66 morava_po_snht <- morava[(snht_morava$estimate+1):length(morava)] # data po
      zmene
67
68 morava_pred_cusum <- morava[1:morava_cusum-1] # data do bodu zmeny
69 morava_po_cusum <- morava[morava_cusum :length(morava)] # data po zmene
70
71 mww_morava_pettitt <- wilcox.test(morava_pred_pettitt, morava_po_pettitt,
      alternative = "two.sided")
72 mww_morava_snht <- wilcox.test(morava_pred_snht, morava_po_snht,
      alternative = "two.sided")
73 mww_morava_cusum <- wilcox.test(morava_pred_cusum, morava_po_cusum,
      alternative = "two.sided")
74 print(mww_morava_pettitt) # p-value = 3.566e-05
75 print(mww_morava_snht) # p-value = 0.002838
76 print(mww_morava_cusum) # p-value = 0.009959
77
78 ks_morava_pettitt <- ks.test(morava_pred_pettitt, morava_po_pettitt)
79 ks_morava_snht <- ks.test(morava_pred_snht, morava_po_snht)
80 ks_morava_cusum <- ks.test(morava_pred_cusum, morava_po_cusum)
81 print(ks_morava_pettitt) # p-value = 0.0001136
82 print(ks_morava_snht) # p-value = 0.01064
83 print(ks_morava_cusum) # p-value = 0.003857
84
85 ## Trend ##
86 smk_morava <- SeasonalMannKendall(morava)
87 print(smk_morava) # pvalue =0.00057362
88
89 cs_morava <- cox.stuart.test(morava) # p-value = 0.03211
90 print(cs_morava)
91
92 cs_left_morava <- cox.stuart.test(morava, alternative = "left.sided")
93 cs_right_morava <- cox.stuart.test(morava, alternative = "right.sided")
94 print(cs_left_morava) # p-value = 0.01605, alternative hypothesis:

```

```

    decreasing trend
95 print(cs_right_morava) # p-value = 0.9867, alternative hypothesis:
    increasing trend

```

Listing 2: Morava - viacnásobné

```

1 morava_pov <- read.xlsx("Data.xlsx", 1, colNames=TRUE)
2 morava <- ts(morava_pov[,3], frequency = 12)
3
4 ## Hierarchicke metody ##
5
6 # Aglomerativna
7 # hladanie gammy minimalizujucej pocet bodov
8 krok <- 0.05
9 m_pocet <- rep(0, (2-0)/krok)
10 m_k <- rep(0, (2-0)/krok)
11 k <- 0
12 for (i in 1:(length(m_pocet)-1)){
13   k <- k + krok
14   m_k[i] <- k
15   m_pom <- e.agglo(X = as.matrix(morava), member = 1:length(morava), alpha
    = k, penalty = function(cps,Xts) 0)
16   m_pocet[i] <- length(m_pom$estimates)
17
18 }
19 m_pocet[length(m_pocet)] <- length((e.agglo(X = as.matrix(morava), member =
    1:length(morava), alpha = 2, penalty = function(cps,Xts) 0))$estimates)
20 m_k[length(m_k)] <- 2
21 plot(m_k,m_pocet, xlab="gamma", ylab="Pocet bodov zmeny", xlim = c(0.05,2))
22 points(m_k[which.min(m_pocet)],m_pocet[which.min(m_pocet)], pch = 16, col="
    red")
23 which.min(m_pocet)
24 # e.agglo pre oprimalnu hodnotu gammy
25 morava_agglo <- e.agglo(X = as.matrix(morava), member = 1:length(morava),
    alpha = m_k[which.min(m_pocet)], penalty = function(cps,Xts) 0)
26 # Vykreslenie najdenych bodov
27 morava_agglo_est <- morava_agglo$estimates
28 morava_agglo_val <- rep(0, length(morava_agglo_est))
29 for (i in 1:length(morava_agglo_est)) {

```

```

30   morava_agglo_val[i] <- morava_pov[morava_agglo_est[i],3]
31 }
32 plot(morava_pov[,3], type='l', xlab="tau", ylab="Prietok v m3/s")
33 points(morava_agglo_est, morava_agglo_val, pch = 16, col="red")
34
35 # Divizna metoda
36 morava_div <- e.divisive(X = as.matrix(morava), sig.lvl = .05, R = 500, k =
    NULL, min.size = 100)
37 # Vykreslenie najdenych bodov
38 morava_div_est <- morava_div$estimates # 330 446 556 660 1061 =>
    nasledujuce body po zmene
39 morava_div_val <- rep(0, length(morava_div_est))
40 for (i in 1:length(morava_div_est)) {
41   morava_div_val[i] <- morava_pov[morava_div_est[i],3]
42 }
43 M <- cbind(1:length(morava), morava_pov[,3], morava_div$cluster)
44 length(morava_div$cluster)
45 ggplot(data=NULL, aes(x=M[,1], y=M[,2], color=as.factor(M[,3]))) + geom_
    line() + scale_color_manual(values=c("darkgoldenrod1", "brown1", "
    darkorchid1", "cornsilk4", "aquamarine3", "cornflowerblue")) + labs(x = "
    tau", y = "Prietok v m3/s", color = "Zhuk") + theme(panel.background =
    element_rect(fill = "white", colour = "black"), legend.key = element_
    rect(colour = "transparent", fill = "white"))
46
47 ## Iteracne metody ##
48
49 # Pettitt #
50 morava.pettitt0 <- pettitt.test(morava)
51 morava_pov[morava.pettitt0$estimate,4]
52 morava.pettitt0.pred <- morava_pov[1:morava.pettitt0$estimate, 3:4]
53 morava.pettitt0.po <- morava_pov[(morava.pettitt0$estimate+1):length(morava
    ), 3:4]
54
55 ww.test(morava.pettitt0.pred[,1]) #p-value = 3.042e-16
56 ww.test(morava.pettitt0.po[,1]) #p-value < 2.2e-16
57 morava.pettitt1a <- pettitt.test(morava.pettitt0.pred[,1]) # K=120, p-value
    = 0.1145
58 morava.pettitt0.pred[morava.pettitt1a$estimate,2]

```

```

59 morava.pettitt1b <- pettitt.test(morava.pettitt0.po[,1]) # K=174, p-value =
    0.03057
60 morava.pettitt0.po[morava.pettitt1b$estimate,2] #431
61 morava.pettitt1b.pred <- morava.pettitt0.po[1:morava.pettitt1b$estimate,
    1:2]
62 morava.pettitt1b.po <- morava.pettitt0.po[(morava.pettitt1b$estimate+1):
    length(morava.pettitt0.po[,1]), 1:2]
63
64 ww.test(morava.pettitt1b.pred[,1]) #p-value = 2.205e-11
65 ww.test(morava.pettitt1b.po[,1]) #p-value < 2.2e-16
66 morava.pettitt2a <- pettitt.test(morava.pettitt1b.pred[,1]) # K=131, p-
    value = 0.004209
67 morava.pettitt1b.pred[morava.pettitt2a$estimate,2] #388
68 morava.pettitt2a.pred <- morava.pettitt1b.pred[1:morava.pettitt2a$estimate
    ,1:2]
69 morava.pettitt2a.po <- morava.pettitt1b.pred[(morava.pettitt2a$estimate+1):
    length(morava.pettitt1b.pred[,1]),1:2]
70 morava.pettitt2b <- pettitt.test(morava.pettitt1b.po[,1]) # K=630, p-value
    = 0.001252
71 morava.pettitt1b.po[morava.pettitt2b$estimate,2] #1061
72 morava.pettitt2b.pred <- morava.pettitt1b.po[1:morava.pettitt2b$estimate
    ,1:2]
73 morava.pettitt2b.po <- morava.pettitt1b.po[(morava.pettitt2b$estimate+1):
    length(morava.pettitt1b.po[,1]),1:2]
74
75 ww.test(morava.pettitt2a.pred[,1]) # p-value = 1.251e-07
76 ww.test(morava.pettitt2a.po[,1]) # p-value = 6.013e-06
77 ww.test(morava.pettitt2b.pred[,1]) # p-value < 2.2e-16
78 ww.test(morava.pettitt2b.po[,1]) # p-value < 2.2e-16
79 morava.pettitt3aa <- pettitt.test(morava.pettitt2a.pred[,1]) # K=16, p-
    value = 0.08597
80 morava.pettitt3ab <- pettitt.test(morava.pettitt2a.po[,1]) # K = 28, p-
    value = 0.1738
81 morava.pettitt3ba <- pettitt.test(morava.pettitt2b.pred[,1]) # K = 79, p-
    value = 0.004994
82 morava.pettitt2b.pred[morava.pettitt3ba$estimate,2] #510
83 morava.pettitt3ba.pred <- morava.pettitt2b.pred[1:morava.pettitt3ba$
    estimate,1:2]

```

```

84 morava.pettitt3ba.po <- morava.pettitt2b.pred[(morava.pettitt3ba$estimate
+1):length(morava.pettitt2b.pred[,1]),1:2]
85 morava.pettitt3bb <- pettitt.test(morava.pettitt2b.po[,1]) # K = 78, p-
value = 0.1738
86 morava.pettitt2b.po[morava.pettitt3bb$estimate,2] #1139
87 morava.pettitt3bb.pred <- morava.pettitt2b.po[1:morava.pettitt3bb$estimate
,1:2]
88 morava.pettitt3bb.po <- morava.pettitt2b.po[(morava.pettitt3bb$estimate+1):
length(morava.pettitt2b.po),1:2]
89
90 ww.test(morava.pettitt3ba.pred[,1]) # p-value = 2.214e-05
91 ww.test(morava.pettitt3ba.po[,1]) # p-value < 2.2e-16
92 ww.test(morava.pettitt3bb.pred[,1]) # p-value = 9.078e-09
93 ww.test(morava.pettitt3bb.po[,1]) # p-value = 1.045e-08
94 morava.pettitt4aa <- pettitt.test(morava.pettitt3ba.pred[,1]) # K=37, p-
value = 0.004841
95 morava.pettitt3ba.pred[morava.pettitt4aa$estimate,2] # 468
96 morava.pettitt4a.pred <- morava.pettitt3ba.pred[1:morava.pettitt4aa$
estimate,1:2]
97 morava.pettitt4a.po <- morava.pettitt3ba.pred[(morava.pettitt4aa$estimate
+1):length(morava.pettitt3ba.pred[,1]),1:2]
98 morava.pettitt4ab <- pettitt.test(morava.pettitt3ba.po[,1]) # K=268, p-
value = 0.0003007
99 morava.pettitt3ba.po[morava.pettitt4ab$estimate,2] # 469
100 morava.pettitt4b.pred <- morava.pettitt3ba.po[1:morava.pettitt4ab$estimate
,1:2]
101 morava.pettitt4b.po <- morava.pettitt3ba.po[(morava.pettitt4ab$estimate+1):
length(morava.pettitt3ba.po[,1]),1:2]
102 morava.pettitt4ba <- pettitt.test(morava.pettitt3bb.pred[,1]) # K=48, p-
value = 0.2607
103 morava.pettitt4bb <- pettitt.test(morava.pettitt3bb.po[,1]) # K=31, p-value
= 0.3588
104
105 ww.test(morava.pettitt4a.pred[,1]) # p-value = 0.07275
106 ww.test(morava.pettitt4a.po[,1]) # p-value = 0.0001393
107 ww.test(morava.pettitt4b.pred[,1]) # p-value < 2.2e-16
108 ww.test(morava.pettitt4b.po[,1]) # p-value < 2.2e-16
109

```

```

110 morava.pettitt5ab <- pettitt.test(morava.pettitt4a.po[,1]) # K = 25, p-
    value = 1.099
111 morava.pettitt5ba <- pettitt.test(morava.pettitt4b.pred[,1]) # K = 149, p-
    value = 0.08948
112 morava.pettitt5bb <- pettitt.test(morava.pettitt4b.po[,1]) # K = 42, p-
    value = 0.09175
113
114 MV <- cbind(morava_pov[,4], morava_pov[,3], rep(0,length(morava_pov[,3])))
115 MV[1:257,3] <- rep(1,257)
116 MV[258:388,3] <- rep(2,131)
117 MV[389:431,3] <- rep(3,43)
118 MV[432:468,3] <- rep(4,37)
119 MV[469:510,3] <- rep(5,42)
120 MV[511:778,3] <- rep(6,268)
121 MV[779:1061,3] <- rep(7,283)
122 MV[1062:1139,3] <- rep(8,78)
123 MV[1140:1416,3] <- rep(9,277)
124
125 ggplot(data=NULL, aes(x=MV[,1], y=MV[,2], color=as.factor(MV[,3]))) + geom_
    line() + scale_color_manual(values=c("darkgoldenrod1", "brown", "
    darkorchid1", "cornsilk4", "aquamarine3", "cornflowerblue", "cyan", "black
    ", "red")) + labs(x = "tau", y = "Prietok", color = "Zhluk") + theme(
    panel.background = element_rect(fill = "white", colour = "grey50"),
    legend.key = element_rect(colour = "transparent", fill = "white"))
126
127 # SNHT #
128 morava.snht0 <- snh.test(morava)
129 morava_pov[morava.snht0$estimate,4] # 4
130 morava.snht0.pred <- morava_pov[1:morava.snht0$estimate, 3:4]
131 morava.snht0.po <- morava_pov[(morava.snht0$estimate+1):length(morava),
    3:4]
132
133 ww.test(morava.snht0.pred[,1]) #p-value = 0.2235
134 ww.test(morava.snht0.po[,1]) #p-value < 2.2e-16
135 morava.snht1a <- snh.test(morava.snht0.pred[,1]) # K=3, p-value = 0.6165
136 morava.snht0.pred[morava.snht1a$estimate,2] # 3
137 morava.snht1b <- snh.test(morava.snht0.po[,1]) # K=575, p-value = 0.0015
138 morava.snht0.po[morava.snht1b$estimate,2] #579

```

```

139 morava.snht1b.pred <- morava.snht0.po[1:morava.snht1b$estimate, 1:2]
140 morava.snht1b.po <- morava.snht0.po[(morava.snht1b$estimate+1):length(
      morava.snht0.po[,1]), 1:2]
141
142 ww.test(morava.snht1b.pred[,1]) #p-value < 2.2e-16
143 ww.test(morava.snht1b.po[,1]) #p-value < 2.2e-16
144 morava.snht2a <- snh.test(morava.snht1b.pred[,1]) # K=442, p-value = 0.003
145 morava.snht1b.pred[morava.snht2a$estimate,2] # 446
146 morava.snht2a.pred <- morava.snht1b.pred[1:morava.snht2a$estimate, 1:2]
147 morava.snht2a.po <- morava.snht1b.pred[(morava.snht2a$estimate+1):length(
      morava.snht1b.pred[,1]), 1:2]
148 morava.snht2b <- snh.test(morava.snht1b.po[,1]) # K=805, p-value = 0.02155
149 morava.snht1b.po[morava.snht2b$estimate,2] # 1384
150 morava.snht2b.pred <- morava.snht1b.po[1:morava.snht2b$estimate, 1:2]
151 morava.snht2b.po <- morava.snht1b.po[(morava.snht2b$estimate+1):length(
      morava.snht1b.po[,1]), 1:2]
152
153 ww.test(morava.snht2a.pred[,1]) #p-value < 2.2e-16
154 ww.test(morava.snht2a.po[,1]) #p-value = 4.763e-10
155 ww.test(morava.snht2b.pred[,1]) #p-value < 2.2e-16
156 ww.test(morava.snht2b.po[,1]) #p-value = 4.01e-05
157 morava.snht3aa <- snh.test(morava.snht2a.pred[,1]) # K=325, p-value = 7e-04
158 morava.snht2a.pred[morava.snht3aa$estimate,2] # 329
159 morava.snht3aa.pred <- morava.snht2a.pred[1:morava.snht3aa$estimate, 1:2]
160 morava.snht3aa.po <- morava.snht2a.pred[(morava.snht3aa$estimate+1):length(
      morava.snht2a.pred[,1]), 1:2]
161 morava.snht3ab <- snh.test(morava.snht2a.po[,1]) # K=63, p-value = 0.00445
162 morava.snht2a.po[morava.snht3ab$estimate,2] # 509
163 morava.snht3ab.pred <- morava.snht2a.po[1:morava.snht3ab$estimate, 1:2]
164 morava.snht3ab.po <- morava.snht2a.po[(morava.snht3ab$estimate+1):length(
      morava.snht2a.po[,1]), 1:2]
165 morava.snht3ba <- snh.test(morava.snht2b.pred[,1]) # K=81, p-value = 0.0806
166 morava.snht2b.pred[morava.snht3ba$estimate,2] # 660
167 morava.snht3bb <- snh.test(morava.snht2b.po[,1]) # K=7, p-value = 0.4958
168 morava.snht2b.po[morava.snht3bb$estimate,2] # 1391
169
170 ww.test(morava.snht3aa.pred[,1]) #p-value < 2.2e-16
171 ww.test(morava.snht3aa.po[,1]) #p-value = 6.068e-09

```

```

172 ww.test(morava.snht3ab.pred[,1]) #p-value = 9.486e-05
173 ww.test(morava.snht3ab.po[,1]) #p-value = 1.852e-05
174 morava.snht4aa <- snh.test(morava.snht3aa.pred[,1]) # K = 106, p-value =
    0.3943
175 morava.snht3aa.pred[morava.snht4aa$estimate,2] # 110
176 morava.snht4ab <- snh.test(morava.snht3aa.po[,1]) # K = 59, p-value =
    0.4621
177 morava.snht3aa.po[morava.snht4ab$estimate,2] # 388
178 morava.snht4ba <- snh.test(morava.snht3ab.pred[,1]) # K = 1, p-value =
    0.02595
179 morava.snht3ab.pred[morava.snht4ba$estimate,2] # 110
180 morava.snht4ba.po <- morava.snht3ab.pred[(morava.snht4ba$estimate+1):length
    (morava.snht3ab.pred[,1]), 1:2]
181 morava.snht4bb <- snh.test(morava.snht3ab.po[,1]) # K = 67, p-value =
    0.00465
182 morava.snht3ab.po[morava.snht4bb$estimate,2] # 576
183 morava.snht4bb.pred <- morava.snht3ab.po[1:morava.snht4bb$estimate, 1:2]
184 morava.snht4bb.po <- morava.snht3ab.po[(morava.snht4bb$estimate+1):length(
    morava.snht3ab.po[,1]), 1:2]
185
186 ww.test(morava.snht4ba.po[,1]) #p-value = 6.224e-05
187 ww.test(morava.snht4bb.pred[,1]) #p-value = 0.0001079
188 ww.test(morava.snht4bb.po[,1]) #p-value = 1
189 morava.snht5a <- snh.test(morava.snht4ba.po[,1]) # K=46, p-value = 0.1738
190 morava.snht5b <- snh.test(morava.snht4bb.pred[,1]) # K=20, p-value = 0.46
191
192 MVS <- cbind(morava_pov[,4], morava_pov[,3], rep(0,length(morava_pov[,3])))
193 MVS[1:4,3] <- rep(1,4)
194 MVS[5:329,3] <- rep(2,325)
195 MVS[330:446,3] <- rep(3,117)
196 MVS[447,3] <- rep(4,1)
197 MVS[448:509,3] <- rep(5,62)
198 MVS[510:579,3] <- rep(6,70)
199 MVS[580:1384,3] <- rep(7,805)
200 MVS[1385:1416,3] <- rep(8,32)
201
202 ggplot(data=NULL, aes(x=MVS[,1], y=MVS[,2], color=as.factor(MVS[,3]))) +
    geom_line() + scale_color_manual(values=c("green", "darkgoldenrod1", "

```



```

darkorchid1", "red", "aquamarine3", "cornflowerblue", "cyan", "black")) +
labs(x = "tau", y = "Prietok", color = "Zhluk") + theme(panel.background
= element_rect(fill = "white", colour = "grey50"), legend.key = element
_rect(colour = "transparent", fill = "white")) + geom_point()

```

Listing 3: Hron

```

1 hron_pov <- read.xlsx("Data.xlsx", 2, colNames=TRUE)
2 hron <- ts(hron_pov[,3], frequency = 12)
3 plot(hron, xlab='Mesiac', ylab='Prietok') # Vykreslenie dat
4 title(main = 'Prietok rieky Hron')
5 # Vypocet strednej hodnoty
6 str_hod_hron <- rep(0, length(hron))
7 for (i in 1:length(hron)) {
8   str_hod_hron[i] <- mean(hron[1:i])
9 }
10 # Vykreslenie strednej hodnoty
11 plot(ts(str_hod_hron, frequency = 12), xlab="Cas v rokoch od roku 1931",
      ylab="Stredna hodnota")
12
13 # Dekompozicia
14 hron_dec <- decompose(ts(hron_pov[,3], frequency = 12), type = "additive")
15 plot(hron_dec)
16
17 ## Homogenita ##
18 brt_hron <- bartels.rank.test(hron) # p-value < 2.2e-16
19 ww_hron <- ww.test(hron) # p-value < 2.2e-16
20 print(brt_hron)
21 print(ww_hron)
22
23 ## Single Change Point ##
24 pettitt_hron <- pettitt.test(hron)
25 print(pettitt_hron) # K=608, p-value = 0.001261
26
27 snht_hron <- snh.test(hron)
28 print(snht_hron) # K=608, p-value = 0.00475
29
30 cusum_hron <- cpt.mean(hron, penalty = "Manual", pen.value = 0.05, method =
    "AMOC", test.stat = "CUSUM", class = TRUE, param.estimates = TRUE)

```

```

31 hron_cusum <- cpts(cusum_hron) # 609 => 608
32
33 plot.ts(hron_pov[,3], xlab= 'tau', ylab='Prietok v m3/s')
34 lines(x=rep(pettitt_hron$estimate, length(0:max(hron_pov[,3]))) , y=0:max(
    hron_pov[,3]), col='red', lwd=3)
35
36 ## Testovanie najdenych bodov zmeny ##
37 hron_pred_pettitt <- hron[1:pettitt_hron$estimate] # data do bodu zmeny
38 hron_po_pettitt <- hron[pettitt_hron$estimate+1:length(hron)] # data po
    zmene
39
40 hron_pred_snht <- hron[1:snht_hron$estimate] # data do bodu zmeny
41 hron_po_snht <- hron[snht_hron$estimate+1:length(hron)] # data po zmene
42
43 hron_pred_cusum <- hron[1:hron_cusum-1] # data do bodu zmeny
44 hron_po_cusum <- hron[hron_cusum :length(hron)] # data po zmene
45
46 mww_hron_pettitt <- wilcox.test(hron_pred_pettitt, hron_po_pettitt,
    alternative = "two.sided")
47 mww_hron_snht <- wilcox.test(hron_pred_snht, hron_po_snht, alternative = "
    two.sided")
48 mww_hron_cusum <- wilcox.test(hron_pred_cusum, hron_po_cusum, alternative =
    "two.sided")
49 print(mww_hron_pettitt) # p-value = 9.94e-05
50 print(mww_hron_snht) # p-value = 9.94e-05
51 print(mww_hron_cusum) # p-value = 9.94e-05
52
53 ks_hron_pettitt <- ks.test(hron_pred_pettitt, hron_po_pettitt)
54 ks_hron_snht <- ks.test(hron_pred_snht, hron_po_snht)
55 ks_hron_cusum <- ks.test(hron_pred_cusum, hron_po_cusum)
56 print(ks_hron_pettitt) # p-value = 0.0009431
57 print(ks_hron_snht) # p-value = 0.0009431
58 print(ks_hron_cusum) # p-value = 0.0009431
59
60 ## Trend ##
61 smk_hron <- SeasonalMannKendall(hron) # pvalue =3.9128e-05
62
63 cs_hron <- cox.stuart.test(hron)

```

```

64 print(cs_hron) # p-value = 0.07263
65
66 # Jednostranne testy na zistenie charakteru trendu
67 cs_left_hron <- cox.stuart.test(hron, alternative = "left.sided")
68 cs_right_hron <- cox.stuart.test(hron, alternative = "right.sided")
69 print(cs_left_hron) # p-value = 0.03631
70 print(cs_right_hron) # p-value = 0.9701

```

Listing 4: Hron - viacnásobné

```

1 hron_pov <- read.xlsx("Data.xlsx", 2, colNames=TRUE)
2 hron <- ts(hron_pov[,3], frequency = 12)
3
4 ## Hierarchicke metody ##
5
6 # Aglomerativna
7 krok <- 0.05
8 h_pocet <- rep(0, (2-0)/krok)
9 h_k <- rep(0, (2-0)/krok)
10 k <- 0
11 for (i in 1:(length(h_pocet)-1)){
12   k <- k + krok
13   h_k[i] <- k
14   h_pom <- e.agglo(X = as.matrix(hron), member = 1:length(hron), alpha = k,
15     penalty = function(cps,Xts) 0)
16   h_pocet[i] <- length(h_pom$estimates)
17 }
18 h_pocet[length(h_pocet)] <- length((e.agglo(X = as.matrix(hron), member =
19   1:length(hron), alpha = 2, penalty = function(cps,Xts) 0))$estimates)
20 h_k[length(h_k)] <- 2
21 plot(h_k,h_pocet, xlab="gamma", ylab="Pocet bodov zmeny", xlim = c(0.05,2))
22 points(h_k[which.min(h_pocet)],h_pocet[which.min(h_pocet)], pch = 16, col="
23   red")
24 which.min(h_pocet)
25 h_k[which.min(h_pocet)]
26
27 hron_agglo <- e.agglo(X = as.matrix(hron), member = 1:length(hron), alpha =
28   h_k[which.min(h_pocet)], penalty = function(cps){0})
29 hron_agglo_est <- hron_agglo$estimates

```

```

26 hron_agglo_val <- rep(0, length(hron_agglo_est))
27
28 for (i in 1:length(hron_agglo_est)) {
29   hron_agglo_val[i] <- hron_pov[hron_agglo_est[i],3]
30 }
31 plot(hron_pov[,3], type='l', xlab="tau", ylab = "Prietok v m3/s")
32 points(hron_agglo_est, hron_agglo_val, pch = 16, col="red")
33
34 # Divizny
35 hron_div <- e.divisive(X = as.matrix(hron), sig.lvl = .05, R = 500, k =
  NULL, min.size = 100)
36 hron_div_est <- hron_div$estimates # 657 => 656
37 hron_div_val <- rep(0, length(hron_div_est))
38
39 for (i in 1:length(hron_div_est)) {
40   hron_div_val[i] <- hron_pov[hron_div_est[i],3]
41 }
42 H <- cbind(1:length(hron), hron_pov[,3], hron_div$cluster)
43 ggplot(data=NULL, aes(x=H[,1], y=H[,2], color=as.factor(H[,3]))) + geom_
  line() + scale_color_manual(values=c("brown1", "cornflowerblue")) + labs
  (x = "tau", y = "Prietok", color = "Zhluk") + theme(panel.background =
  element_rect(fill = "white", colour = "grey50"), legend.key = element_
  rect(colour = "transparent", fill = "white"))
44
45 ## Iteracne metody ##
46
47 # Pettitt #
48 hron.pettitt0 <- pettitt.test(hron)
49 hron_pov[hron.pettitt0$estimate,4] # 608
50 hron.pettitt0.pred <- hron_pov[1:hron.pettitt0$estimate, 3:4]
51 hron.pettitt0.po <- hron_pov[(hron.pettitt0$estimate+1):length(hron), 3:4]
52
53 ww.test(hron.pettitt0.pred[,1]) #p-value < 2.2e-16
54 ww.test(hron.pettitt0.po[,1]) #p-value < 2.2e-16
55 hron.pettitt1a <- pettitt.test(hron.pettitt0.pred[,1]) # K=61, p-value =
  0.5068
56 hron.pettitt0.pred[hron.pettitt1a$estimate,2] #61
57 hron.pettitt1b <- pettitt.test(hron.pettitt0.po[,1]) # K=340, p-value =

```

```

0.1245
58 hron.pettitt0.po[hron.pettitt1b$estimate,2] #948
59
60 HV <- cbind(hron_pov[,4], hron_pov[,3], rep(0,length(hron_pov[,3])))
61 HV[1:608,3] <- rep(1,608)
62 HV[609:1044,3] <- rep(2,436)
63
64 ggplot(data=NULL, aes(x=HV[,1], y=HV[,2], color=as.factor(HV[,3]))) + geom_
  line() + scale_color_manual(values=c("brown1", "cornflowerblue")) + labs
  (x = "tau", y = "Prietok", color = "Zhluk") + theme(panel.background =
  element_rect(fill = "white", colour = "grey50"), legend.key = element_
  rect(colour = "transparent", fill = "white"))
65
66 ## SNHT
67 hron.snht0 <- snh.test(hron)
68 hron_pov[hron.snht0$estimate,4] # 608
69 hron.snht0.pred <- hron_pov[1:hron.snht0$estimate, 3:4]
70 hron.snht0.po <- hron_pov[(hron.snht0$estimate+1):length(hron), 3:4]
71
72 ww.test(hron.snht0.pred[,1]) #p-value = 0.000487
73 ww.test(hron.snht0.po[,1]) #p-value < 2.2e-16
74 hron.snht1a <- snh.test(hron.snht0.pred[,1]) # K=7, p-value = 0.4231
75 hron.snht0.pred[hron.snht1a$estimate,2] # 7
76 hron.snht1a.pred <- hron_pov[1:hron.snht1a$estimate, 3:4]
77 hron.snht1a.po <- hron_pov[(hron.snht1a$estimate+1):length(hron.snht0.pred
  [,1]), 3:4]
78 hron.snht1b <- snh.test(hron.snht0.po[,1]) # K=341, p-value = 0.2352
79 hron.snht0.po[hron.snht1b$estimate,2] #949

```

Listing 5: Vah

```

1 vah_pov <- read.xlsx("Data.xlsx", 3, colNames=TRUE)
2 vah <- ts(vah_pov[,3], frequency = 12)
3 plot(vah, xlab='Pocet rokov od roku 1921', ylab='Prietok')
4 title(main = 'Prietok rieky Vah')
5
6 str_hod_vah <- rep(0, length(vah))
7 for (i in 1:length(vah)) {
8   str_hod_vah[i] <- mean(vah[1:i])

```

```

9 }
10
11 plot.ts(vah, xlab='Tau', ylab='Prietok')
12 plot(ts(str_hod_vah, frequency = 12), xlab="Cas v rokoch od roku 1921",
      ylab="Stredna hodnota")
13
14 vah_dec <- decompose(ts(vah_pov[,3], frequency = 12), type = "additive")
15 plot(vah_dec)
16
17 ## Homogenita ##
18 brt_vah <- bartels.rank.test(vah) # p-value < 2.2e-16
19 ww_vah <- ww.test(vah) # p-value < 2.2e-16
20 print(brt_vah)
21 print(ww_vah)
22
23 ## Single Change Point ##
24 pettitt_vah <- pettitt.test(vah)
25 print(pettitt_vah) # K=749, p-value = 0.3768
26
27 snht_vah <- snh.test(vah)
28 print(snht_vah) # K=14, p-value = 0.1255
29
30 cusum_vah <- cpt.mean(vah, penalty = "Manual", pen.value = 0.05, method = "
      AMOC", test.stat = "CUSUM", class = TRUE, param.estimates = TRUE)
31 vah_cusum <- cpts(cusum_vah) # 559 => 558
32
33 plot.ts(vah_pov[,3], xlab='tau', ylab='Prietok v m3/s')
34 lines(x=rep(pettitt_vah$estimate, length(0:max(vah_pov[,3]))), y=0:max(vah_
      pov[,3]), col='red', lwd=3)
35 lines(x=rep(snht_vah$estimate, length(0:max(vah_pov[,3]))), y=0:max(vah_pov
      [,3]), col='blue', lwd=3)
36 lines(x=rep(vah_cusum-1, length(0:max(vah_pov[,3]))), y=0:max(vah_pov[,3]),
      col='green', lwd=3)
37 legend(820, 715, legend=c("Pettitt", "SNHT", "CUSUM"), col = c("red", "blue"
      , "green"), lty = 1)
38
39 ## Testovanie najdenych bodov zmeny ##
40 vah_pred_pettitt <- vah[1:pettitt_vah$estimate] # data do bodu zmeny

```

```

41 vah_po_pettitt <- vah[pettitt_vah$estimate+1:length(vah)] # data po zmene
42
43 vah_pred_snht <- vah[1:snht_vah$estimate] # data do bodu zmeny
44 vah_po_snht <- vah[snht_vah$estimate+1:length(vah)] # data po zmene
45
46 vah_pred_cusum <- vah[1:vah_cusum-1] # data do bodu zmeny
47 vah_po_cusum <- vah[vah_cusum :length(vah)] # data po zmene
48
49 mww_vah_pettitt <- wilcox.test(vah_pred_pettitt, vah_po_pettitt,
    alternative = "two.sided")
50 mww_vah_snht <- wilcox.test(vah_pred_snht, vah_po_snht, alternative = "two.
    sided")
51 mww_vah_cusum <- wilcox.test(vah_pred_cusum, vah_po_cusum, alternative = "
    two.sided")
52 print(mww_vah_pettitt) # p-value = 0.05649
53 print(mww_vah_snht) # p-value = 0.0002548
54 print(mww_vah_cusum) # p-value = 0.1859
55
56 ks_vah_pettitt <- ks.test(vah_pred_pettitt, vah_po_pettitt)
57 ks_vah_snht <- ks.test(vah_pred_snht, vah_po_snht)
58 ks_vah_cusum <- ks.test(vah_pred_cusum, vah_po_cusum)
59 print(ks_vah_pettitt) # p-value = 0.00531
60 print(ks_vah_snht) # p-value = 0.001212
61 print(ks_vah_cusum) # p-value = 0.0002395
62
63 ## Trend ##
64 smk_vah <- SeasonalMannKendall(vah) # pvalue =0.42694
65
66 cs_vah <- cox.stuart.test(vah)
67 print(cs_vah) # p-value = 0.7717
68
69 ## Viacnasobne hierarchicke ##
70
71 # Aglomerativny
72 krok <- 0.05
73 v_pocet <- rep(0, (2-0)/krok)
74 v_k <- rep(0, (2-0)/krok)
75 k <- 0

```

```

76 for (i in 1:(length(v_pocet)-1)){
77   k <- k + krok
78   v_k[i] <- k
79   v_pom <- e.agglo(X = as.matrix(vah), member = 1:length(vah), alpha = k,
      penalty = function(cps,Xts) 0)
80   v_pocet[i] <- length(v_pom$estimates)
81 }
82 v_pocet[length(v_pocet)] <- length((e.agglo(X = as.matrix(vah), member = 1:
      length(vah), alpha = 2, penalty = function(cps,Xts) 0))$estimates)
83 v_k[length(v_k)] <- 2
84 plot(v_k,v_pocet, xlab="gamma", ylab="Pocet bodov zmeny", xlim = c(0.05,2))
85 points(v_k[which.min(v_pocet)],v_pocet[which.min(v_pocet)], pch = 16, col="
      red")
86 which.min(v_pocet)
87 v_k[which.min(v_pocet)]
88
89 vah_agglo <- e.agglo(X = as.matrix(vah), member = 1:length(vah), alpha = v_
      k[which.min(v_pocet)], penalty = function(cps,Xts) 0)
90 vah_agglo_est <- vah_agglo$estimates
91 vah_agglo_val <- rep(0, length(vah_agglo_est))
92
93 for (i in 1:length(vah_agglo_est)) {
94   vah_agglo_val[i] <- vah_pov[vah_agglo_est[i],3]
95 }
96 plot(vah_pov[,3],type='l', xlab="tau", ylab = "Prietok v m3/s")
97 points(vah_agglo_est,vah_agglo_val, pch = 16, col="red")
98
99 # Divizny
100 vah_div <- e.divisive(X = as.matrix(vah), sig.lvl = .05, R = 500, k = NULL,
      min.size = 100)
101 vah_div_est <- vah_div$estimates # 367 => 366
102 vah_div_val <- rep(0, length(vah_div_est))
103
104 for (i in 1:length(vah_div_est)) {
105   vah_div_val[i] <- vah_pov[vah_div_est[i],3]
106 }
107 V <- cbind(1:length(vah), vah_pov[,3], vah_div$cluster)
108 length(vah_div$cluster)

```



```

109 ggplot(data=NULL, aes(x=V[,1], y=V[,2], color=as.factor(V[,3]))) + geom_
    line() + scale_color_manual(values=c("brown1", "cornflowerblue")) + labs
    (x = "tau", y = "Prietok", color = "Zhluk") + theme(panel.background =
    element_rect(fill = "white", colour = "grey50"), legend.key = element_
    rect(colour = "transparent", fill = "white"))

```

Listing 6: Belá

```

1  bela_pov <- read.xlsx("Data.xlsx", 4, colNames=TRUE)
2  bela <- ts(bela_pov[,2], frequency = 12)
3  plot.ts(bela, xlab='Cas v rokoch od roku 1931', ylab='Prietok v m3/s')
4  title(main = 'Prietok rieky Bela')
5  str_hod_bela <- rep(0, length(bela))
6
7  for (i in 1:length(bela)) {
8    str_hod_bela[i] <- mean(bela[1:i])
9  }
10
11 plot(ts(str_hod_bela, frequency = 12), xlab="Cas v rokoch od roku 1931",
    ylab="Stredna hodnota")
12
13 bela_dec <- decompose(ts(bela_pov[,2], frequency = 12), type = "additive")
14 plot(bela_dec)
15
16 ## Homogenita ##
17 brt_bela <- bartels.rank.test(bela)
18 ww_bela <- ww.test(bela)
19 print(brt_bela) # p-value < 2.2e-16
20 print(ww_bela) # p-value < 2.2e-16
21
22 ## Single Change Point ##
23 pettitt_bela <- pettitt.test(bela)
24 print(pettitt_bela) # K=297, p-value = 0.4822
25
26 snht_bela <- snh.test(bela)
27 print(snht_bela) # K=4, p-value = 0.9168
28
29 cusum_bela <- cpt.mean(bela, penalty = "Manual", pen.value = 0.05, method =
    "AMOC", test.stat = "CUSUM", class = TRUE, param.estimates = TRUE)

```

```

30  bela_cusum <- cpts(cusum_bela) # NENASIEL
31
32  plot.ts(bela_pov[,2], xlab= 'tau', ylab='Prietok v m3/s')
33  lines(x=rep(pettitt_bela$estimate, length(0:max(bela_pov[,2]))) , y=0:max(
      bela_pov[,2]), col='red', lwd=3)
34  lines(x=rep(snht_bela$estimate, length(0:max(bela_pov[,2]))) , y=0:max(bela_
      pov[,2]), col='blue', lwd=3)
35  legend(760, 20, legend=c("Pettitt", "SNHT"), col = c("red", "blue"), lty =
      1)
36
37  ## Testovanie najdenych bodov zmeny ##
38  bela_pred_pettitt <- bela[1:pettitt_bela$estimate] # data do bodu zmeny
39  bela_po_pettitt <- bela[(pettitt_bela$estimate+1):length(bela)] # data po
      zmene
40
41  bela_pred_snht <- bela[1:snht_bela$estimate] # data do bodu zmeny
42  bela_po_snht <- bela[(snht_bela$estimate+1):length(bela)] # data po zmene
43
44  mww_bela_pettitt <- wilcox.test(bela_pred_pettitt, bela_po_pettitt,
      alternative = "two.sided")
45  mww_bela_snht <- wilcox.test(bela_pred_snht, bela_po_snht, alternative = "
      two.sided")
46  print(mww_bela_pettitt) # p-value = 0.06249
47  print(mww_bela_snht) # p-value = 0.01759
48
49  ks_bela_pettitt <- ks.test(bela_pred_pettitt, bela_po_pettitt)
50  ks_bela_snht <- ks.test(bela_pred_snht, bela_po_snht)
51  print(ks_bela_pettitt) # p-value = 0.02824
52  print(ks_bela_snht) # p-value = 0.1254
53
54  ## Trend ##
55  smk_bela <- SeasonalMannKendall(bela)
56  print(smk_bela) # pvalue =0.72881
57
58  cs_bela <- cox.stuart.test(bela)
59  print(cs_bela) # p-value = 0.8949

```

Listing 7: Kysuca

```

1 kysuca_pov <- read.xlsx("Data.xlsx", 5, colNames=TRUE)
2 kysuca <- ts(kysuca_pov[,2], frequency = 12)
3 plot.ts(kysuca, xlab='Cas v rokoch od roku 1931', ylab='Prietok v m3/s')
4 title(main = 'Prietok rieky Kysuca')
5
6 str_hod_kysuca <- rep(0, length(kysuca))
7 for (i in 1:length(kysuca)) {
8   str_hod_kysuca[i] <- mean(kysuca[1:i])
9 }
10
11 plot(ts(str_hod_kysuca, frequency = 12), xlab="Cas v rokoch od roku 1931",
      ylab="Stredna hodnota")
12
13 kysuca_dec <- decompose(ts(kysuca_pov[,2], frequency = 12), type = "additive
      ")
14 plot(kysuca_dec)
15
16 ## Homogenita ##
17 brt_kysuca <- bartels.rank.test(kysuca)
18 ww_kysuca <- ww.test(kysuca)
19 print(brt_kysuca) # p-value < 2.2e-16
20 print(ww_kysuca) # p-value < 2.351e-16
21
22 ## Single Change Point ##
23 pettitt_kysuca <- pettitt.test(kysuca)
24 print(pettitt_kysuca) # K=629, p-value = 0.8718
25
26 snht_kysuca <- snh.test(kysuca)
27 print(snht_kysuca) # K=42, p-value = 0.9486
28
29 cusum_kysuca <- cpt.mean(kysuca, penalty = "Manual", pen.value = 0.05,
      method = "AMOC", test.stat = "CUSUM", class = TRUE, param.estimates =
      TRUE)
30 kysuca_cusum <- cpts(cusum_kysuca) # 856 => 855
31
32 plot.ts(kysuca_pov[,2], xlab= 'tau', ylab='Prietok v m3/s')
33 lines(x=rep(pettitt_kysuca$estimate, length(0:max(kysuca_pov[,2]))), y=0:
      max(kysuca_pov[,2]), col='red', lwd=3)

```

```

34 lines(x=rep(snht_kysuca$estimate,length(0:max(kysuca_pov[,2]))) , y=0:max(
    kysuca_pov[,2]), col='blue', lwd=3)
35 lines(x=rep(kysuca_cusum-1,length(0:max(kysuca_pov[,2]))) , y=0:max(kysuca_
    pov[,2]), col='green', lwd=3)
36 legend(740, 90, legend=c("Pettitt", "SNHT", "CUSUM"), col = c("red", "blue",
    "green"), lty = 1)
37
38 ## Testovanie najdenych bodov zmeny ##
39 kysuca_pred_pettitt <- kysuca[1:pettitt_kysuca$estimate] # data do bodu
    zmeny
40 kysuca_po_pettitt <- kysuca[(pettitt_kysuca$estimate+1):length(kysuca)] #
    data po zmene
41
42 kysuca_pred_snht <- kysuca[1:snht_kysuca$estimate] # data do bodu zmeny
43 kysuca_po_snht <- kysuca[(snht_kysuca$estimate+1):length(kysuca)] # data po
    zmene
44
45 kysuca_pred_cusum <- kysuca[1:kysuca_cusum-1] # data do bodu zmeny
46 kysuca_po_cusum <- kysuca[kysuca_cusum :length(kysuca)] # data po zmene
47
48 mww_kysuca_pettitt <- wilcox.test(kysuca_pred_pettitt, kysuca_po_pettitt,
    alternative = "two.sided")
49 mww_kysuca_snht <- wilcox.test(kysuca_pred_snht, kysuca_po_snht,
    alternative = "two.sided")
50 mww_kysuca_cusum <- wilcox.test(kysuca_pred_cusum, kysuca_po_cusum,
    alternative = "two.sided")
51 print(mww_kysuca_pettitt) # p-value = 0.188
52 print(mww_kysuca_snht) # p-value = 0.04113
53 print(mww_kysuca_cusum) # p-value = 0.3198
54
55 ks_kysuca_pettitt <- ks.test(kysuca_pred_pettitt, kysuca_po_pettitt)
56 ks_kysuca_snht <- ks.test(kysuca_pred_snht, kysuca_po_snht)
57 ks_kysuca_cusum <- ks.test(kysuca_pred_cusum, kysuca_po_cusum)
58 print(ks_kysuca_pettitt) # p-value = 0.287
59 print(ks_kysuca_snht) # p-value = 0.08783
60 print(ks_kysuca_cusum) # p-value = 0.1643
61
62 ## Trend ##

```

```

63 smk_kysuca <- SeasonalMannKendall(kysuca)
64 print(smk_kysuca) # pvalue =0.95949
65
66 cs_kysuca <- cox.stuart.test(kysuca)
67 print(cs_kysuca) # p-value = 0.6937

```

Listing 8: Krupinica

```

1  krupinica_pov <- read.xlsx("Data.xlsx", 6, colNames=TRUE)
2  krupinica <- ts(krupinica_pov[,2], frequency = 12)
3  plot.ts(krupinica, xlab='Cas v rokoch od roku 1931', ylab='Prietok v m3/s')
4  title(main = 'Prietok rieky Krupinica')
5  str_hod_krupinica <- rep(0, length(krupinica))
6
7  for (i in 1:length(krupinica)) {
8    str_hod_krupinica[i] <- mean(krupinica[1:i])
9  }
10
11 plot(ts(str_hod_krupinica, frequency = 12), xlab="Cas v rokoch od roku 1900
    ", ylab="Stredna hodnota")
12
13 krupinica_dec <- decompose(ts(krupinica_pov[,2], frequency = 12), type = "
    additive")
14 plot(krupinica_dec)
15
16 ## Homogenita ##
17 brt_krupinica <- bartels.rank.test(krupinica)
18 ww_krupinica <- ww.test(krupinica)
19 print(brt_krupinica) # p-value < 2.2e-16
20 print(ww_krupinica) # p-value < 2.2e-16
21
22 ## Single Change Point ##
23 pettitt_krupinica <- pettitt.test(krupinica)
24 print(pettitt_krupinica) # K=438, p-value = 0.0004466
25
26 snht_krupinica <- snh.test(krupinica)
27 print(snht_krupinica) # K=436, p-value < 2.2e-16
28
29 cusum_krupinica <- cpt.mean(krupinica, penalty = "Manual", pen.value =

```

```

0.05, method = "AMOC", test.stat = "CUSUM", class = TRUE, param.
  estimates = TRUE)
30 krupinica_cusum <- cpts(cusum_krupinica) # 437 => 436
31
32 plot.ts(krupinica_pov[,2], xlab= 't', ylab='Prietok v m3/s')
33 lines(x=rep(pettitt_krupinica$estimate, length(0:max(krupinica_pov[,2]))),
      y=0:max(krupinica_pov[,2]), col='red', lwd=3)
34 lines(x=rep(snht_krupinica$estimate, length(0:max(krupinica_pov[,2]))), y
      =0:max(krupinica_pov[,2]), col='blue', lwd=3)
35 legend(560, 27, legend=c("Pettitt", "SNHT, CUSUM"), col = c("red", "blue", "
      green"), lty = 1)
36
37 ## Testovanie najdenych bodov zmeny ##
38 krupinica_pred_pettitt <- krupinica[1:pettitt_krupinica$estimate] # data do
      bodu zmeny
39 krupinica_po_pettitt <- krupinica[(pettitt_krupinica$estimate+1):length(
      krupinica)] # data po zmene
40
41 krupinica_pred_snht <- krupinica[1:snht_krupinica$estimate] # data do bodu
      zmeny
42 krupinica_po_snht <- krupinica[(snht_krupinica$estimate+1):length(krupinica
      )] # data po zmene
43
44 krupinica_pred_cusum <- krupinica[1:krupinica_cusum-1] # data do bodu zmeny
45 krupinica_po_cusum <- krupinica[krupinica_cusum :length(krupinica)] # data
      po zmene
46
47 mww_krupinica_pettitt <- wilcox.test(krupinica_pred_pettitt, krupinica_po_
      pettitt, alternative = "two.sided")
48 mww_krupinica_snht <- wilcox.test(krupinica_pred_snht, krupinica_po_snht,
      alternative = "two.sided")
49 mww_krupinica_cusum <- wilcox.test(krupinica_pred_cusum, krupinica_po_cusum
      , alternative = "two.sided")
50 print(mww_krupinica_pettitt) # p-value = 3.534e-05
51 print(mww_krupinica_snht) # p-value = 4.531e-05
52 print(mww_krupinica_cusum) # p-value = 4.531e-05
53
54 ks_krupinica_pettitt <- ks.test(krupinica_pred_pettitt, krupinica_po_

```

```

    pettitt)
55 ks_krupinica_snht <- ks.test(krupinica_pred_snht, krupinica_po_snht)
56 ks_krupinica_cusum <- ks.test(krupinica_pred_cusum, krupinica_po_cusum)
57 print(ks_krupinica_pettitt) # p-value = 1.325e-05
58 print(ks_krupinica_snht) # p-value = 8.876e-06
59 print(ks_krupinica_cusum) # p-value = 8.876e-06
60
61 ## Trend ##
62 smk_krupinica <- SeasonalMannKendall(krupinica)
63 print(smk_krupinica) # pvalue =3.0806e-05
64
65 cs_krupinica <- cox.stuart.test(krupinica)
66 print(cs_krupinica) # p-value = 0.01421
67
68 cs_left_krupinica <- cox.stuart.test(krupinica, alternative = "left.sided"
    )
69 cs_right_krupinica <- cox.stuart.test(krupinica, alternative = "right.
    sided")
70 print(cs_left_krupinica) # p-value = 0.007105
71 print(cs_right_krupinica) # p-value = 0.9945

```

Listing 9: Krupinica - viacnásobné

```

1 krupinica_pov <- read.xlsx("Data.xlsx", 6, colNames=TRUE)
2 krupinica <- ts(krupinica_pov[,2], frequency = 12)
3
4 ## Hierarchicke metody ##
5
6 # Agglomerativny
7 krok <- 0.05
8 kr_pocet <- rep(0, (2-0)/krok)
9 kr_k <- rep(0, (2-0)/krok)
10 k <- 0
11 for (i in 1:(length(kr_pocet)-1)){
12   k <- k + krok
13   kr_k[i] <- k
14   kr_pom <- e.agglo(X = as.matrix(krupinica), member = 1:length(krupinica),
       alpha = k, penalty = function(cps,Xts) 0)
15   kr_pocet[i] <- length(kr_pom$estimates)

```

```

16 }
17 kr_pocet[length(kr_pocet)] <- length((e.agglo(X = as.matrix(krupinica),
      member = 1:length(krupinica), alpha = 2, penalty = function(cps,Xts) 0))
      $estimates)
18 kr_k[length(kr_k)] <- 2
19 plot(kr_k,kr_pocet, xlab="gamma", ylab="Pocet bodov zmeny", xlim = c
      (0.05,2))
20 points(kr_k[which.min(kr_pocet)],kr_pocet[which.min(kr_pocet)], pch = 16,
      col="red")
21 which.min(kr_pocet)
22
23 krupinica_agglo <- e.agglo(X = as.matrix(krupinica), member = 1:length(
      krupinica), alpha = kr_k[which.min(kr_pocet)], penalty = function(cps,
      Xts) 0)
24 krupinica_agglo_est <- krupinica_agglo$estimates
25 krupinica_agglo_val <- rep(0, length(krupinica_agglo_est))
26
27 for (i in 1:length(krupinica_agglo_est)) {
28   krupinica_agglo_val[i] <- krupinica_pov[krupinica_agglo_est[i],2]
29 }
30 plot(krupinica_pov[,2],type='l', xlab="tau", ylab="Prietok v m3/s")
31 points(krupinica_agglo_est,krupinica_agglo_val, pch = 16, col="red")
32
33 # Divizny
34 krupinica_div <- e.divisive(X = as.matrix(krupinica), sig.lvl = .05, R =
      500, k = NULL, min.size = 100)
35 krupinica_div_est <- krupinica_div$estimates # 184 437 => 183 436
36 krupinica_div_val <- rep(0, length(krupinica_div_est))
37
38 for (i in 1:length(krupinica_div_est)) {
39   krupinica_div_val[i] <- krupinica_pov[krupinica_div_est[i],3]
40 }
41
42 Kr <- cbind(1:length(krupinica), krupinica_pov[,2], krupinica_div$cluster)
43 length(krupinica_div$cluster)
44 ggplot(data=NULL, aes(x=Kr[,1], y=Kr[,2], color=as.factor(Kr[,3]))) + geom_
      line() + scale_color_manual(values=c("darkgoldenrod1", "brown1", "
      darkorchid1", "cornsilk4", "aquamarine3", "cornflowerblue")) + labs(x = "

```



```

    tau", y = "Prietok v m3/s", color = "Zhluk") + theme(panel.background =
    element_rect(fill = "white", colour = "black"), legend.key = element_
    rect(colour = "transparent", fill = "white"))
45
46 ## Iteracne metody ##
47
48 # Pettitt
49 krupinica.pettitt0 <- pettitt.test(krupinica)
50 krupinica_pov[krupinica.pettitt0$estimate,3] # 438
51 krupinica.pettitt0.pred <- krupinica_pov[1:krupinica.pettitt0$estimate,
    2:3]
52 krupinica.pettitt0.po <- krupinica_pov[(krupinica.pettitt0$estimate+1):
    length(krupinica), 2:3]
53
54 ww.test(krupinica.pettitt0.pred[,1]) #p-value = 8.929e-16
55 ww.test(krupinica.pettitt0.po[,1]) #p-value < 2.2e-16
56 krupinica.pettitt1a <- pettitt.test(krupinica.pettitt0.pred[,1]) # K=138, p
    -value = 0.01007
57 krupinica.pettitt0.pred[krupinica.pettitt1a$estimate,2] # 138
58 krupinica.pettitt1b <- pettitt.test(krupinica.pettitt0.po[,1]) # K=508, p-
    value = 0.4814
59 krupinica.pettitt0.po[krupinica.pettitt1b$estimate,2] #946
60 krupinica.pettitt1a.pred <- krupinica.pettitt0.pred[1:krupinica.pettitt1a$
    estimate, 1:2]
61 krupinica.pettitt1a.po <- krupinica.pettitt0.pred[(krupinica.pettitt1a$
    estimate+1):length(krupinica.pettitt0.pred[,1]), 1:2]
62
63 ww.test(krupinica.pettitt1a.pred[,1]) #p-value == 3.363e-06
64 ww.test(krupinica.pettitt1a.po[,1]) #p-value = 7.681e-11
65 krupinica.pettitt2a <- pettitt.test(krupinica.pettitt1a.pred[,1]) # K=59, p
    -value = 1.48e-05
66 krupinica.pettitt1a.pred[krupinica.pettitt2a$estimate,2] #59
67 krupinica.pettitt2a.pred <- krupinica.pettitt1a.pred[1:krupinica.pettitt2a$
    estimate,1:2]
68 krupinica.pettitt2a.po <- krupinica.pettitt1a.pred[(krupinica.pettitt2a$
    estimate+1):length(krupinica.pettitt1a.pred[,1]), 1:2]
69 krupinica.pettitt2b <- pettitt.test(krupinica.pettitt1a.po[,1]) # K=267, p-
    value = 0.0237

```

```

70 krupinica.pettitt1a.po[krupinica.pettitt2b$estimate,2] #405
71 krupinica.pettitt2b.pred <- krupinica.pettitt1a.po[1:krupinica.pettitt2b$
    estimate,1:2]
72 krupinica.pettitt2b.po <- krupinica.pettitt1a.po[(krupinica.pettitt2b$
    estimate+1):length(krupinica.pettitt1a.po[,1]),1:2]
73
74 ww.test(krupinica.pettitt2a.pred[,1]) #p-value = 4.362e-05
75 ww.test(krupinica.pettitt2a.po[,1]) #p-value = 0.005509
76 ww.test(krupinica.pettitt2b.pred[,1]) #p-value = 6.56e-16
77 ww.test(krupinica.pettitt2b.po[,1]) #p-value = 0.4864
78 krupinica.pettitt3aa <- pettitt.test(krupinica.pettitt2a.pred[,1]) # K =
    18, p-value = 0.04264
79 krupinica.pettitt2a.pred[krupinica.pettitt3aa$estimate,2]
80 krupinica.pettitt3aa.pred <- krupinica.pettitt2a.pred[1:krupinica.
    pettitt3aa$estimate,1:2]
81 krupinica.pettitt3aa.po <- krupinica.pettitt2a.pred[(krupinica.pettitt3aa$
    estimate+1):length(krupinica.pettitt2a.pred[,1]),1:2]
82 krupinica.pettitt3ab <- pettitt.test(krupinica.pettitt2a.po[,1]) # K = 8, p
    -value = 0.4287
83 krupinica.pettitt3ba <- pettitt.test(krupinica.pettitt2b.pred[,1]) # K =
    100, p-value = 0.7214
84
85 ww.test(krupinica.pettitt3aa.pred[,1]) #p-value = 0.0255
86 ww.test(krupinica.pettitt3aa.po[,1]) #p-value = 0.0003782
87 krupinica.pettitt4a <- pettitt.test(krupinica.pettitt3aa.pred[,1]) # K = 5,
    p-value = 0.2779
88 krupinica.pettitt4b <- pettitt.test(krupinica.pettitt3aa.po[,1]) # K = 16,
    p-value = 0.2034
89
90 KrV <- cbind(krupinica_pov[,3], krupinica_pov[,2], rep(0,length(krupinica_
    pov[,2])))
91 KrV[1:18,3] <- rep(1,18)
92 KrV[19:59,3] <- rep(2,41)
93 KrV[60:138,3] <- rep(3,79)
94 KrV[139:405,3] <- rep(4,267)
95 KrV[406:438,3] <- rep(5,33)
96 KrV[439:1008,3] <- rep(6,570)
97

```

```

98  ggplot(data=NULL, aes(x=KrV[,1], y=KrV[,2], color=as.factor(KrV[,3]))) +
    geom_line() + scale_color_manual(values=c("darkgoldenrod1", "green", "
    darkorchid1", "red", "aquamarine3", "cornflowerblue")) + labs(x = "tau", y
    = "Prietok", color = "Zhluk") + theme(panel.background = element_rect(
    fill = "white", colour = "grey50"), legend.key = element_rect(colour = "
    transparent", fill = "white"))
99
100 # SNHT
101 krupinica.snht0 <- snh.test(krupinica)
102 krupinica.pov[krupinica.snht0$estimate,3] # 436
103 krupinica.snht0.pred <- krupinica.pov[1:krupinica.snht0$estimate, 2:3]
104 krupinica.snht0.po <- krupinica.pov[(krupinica.snht0$estimate+1):length(
    krupinica), 2:3]
105
106 ww.test(krupinica.snht0.pred[,1]) #p-value = 9.093e-16
107 ww.test(krupinica.snht0.po[,1]) #p-value < 2.2e-16
108 krupinica.snht1a <- snh.test(krupinica.snht0.pred[,1]) # K= 410, p-value =
    0.00685
109 krupinica.snht0.pred[krupinica.snht1a$estimate,2]
110 krupinica.snht1b <- snh.test(krupinica.snht0.po[,1]) # K=511, p-value =
    0.1497
111 krupinica.snht0.po[krupinica.snht1b$estimate,2] #947
112 krupinica.snht1a.pred <- krupinica.snht0.pred[1:krupinica.snht1a$estimate,
    1:2]
113 krupinica.snht1a.po <- krupinica.snht0.pred[(krupinica.snht1a$estimate+1):
    length(krupinica.snht0.pred[,1]), 1:2]
114
115 ww.test(krupinica.snht1a.pred[,1]) #p-value < 2.2e-16
116 ww.test(krupinica.snht1a.po[,1]) #p-value = 0.8099
117 krupinica.snht2a <- snh.test(krupinica.snht1a.pred[,1]) # K=137, p-value =
    0.0184
118 krupinica.snht1a.pred[krupinica.snht2a$estimate,2] #137
119 krupinica.snht2a.pred <- krupinica.snht1a.pred[1:krupinica.snht2a$estimate
    ,1:2]
120 krupinica.snht2a.po <- krupinica.snht1a.pred[(krupinica.snht2a$estimate+1):
    length(krupinica.snht1a.pred[,1]), 1:2]
121
122 ww.test(krupinica.snht2a.pred[,1]) #p-value = 3.531e-06

```

```

123 ww.test(krupinica.snht2a.po[,1]) #p-value = 1.169e-15
124 krupinica.snht3a <- snh.test(krupinica.snht2a.pred[,1]) # K=59, p-value =
    0.02665
125 krupinica.snht2a.pred[krupinica.snht3a$estimate,2] #59
126 krupinica.snht3a.pred <- krupinica.snht2a.pred[1:krupinica.snht3a$estimate
    ,1:2]
127 krupinica.snht3a.po <- krupinica.snht2a.pred[(krupinica.snht3a$estimate+1):
    length(krupinica.snht2a.pred[,1]),1:2]
128 krupinica.snht3b <- snh.test(krupinica.snht2a.po[,1]) # K=20, p-value =
    0.2516
129 krupinica.snht2a.po[krupinica.snht3b$estimate,2] # 157
130
131 ww.test(krupinica.snht3a.pred[,1]) #p-value = 4.362e-05
132 ww.test(krupinica.snht3a.po[,1]) #p-value = 0.004242
133 krupinica.snht4a <- snh.test(krupinica.snht3a.pred[,1]) # K=4, p-value =
    0.0014
134 krupinica.snht3a.pred[krupinica.snht4a$estimate,2] #4
135 krupinica.snht4a.pred <- krupinica.snht3a.pred[1:krupinica.snht4a$estimate
    ,1:2]
136 krupinica.snht4a.po <- krupinica.snht3a.pred[(krupinica.snht4a$estimate+1):
    length(krupinica.snht3a.pred[,1]),1:2]
137 krupinica.snht4b <- snh.test(krupinica.snht3a.po[,1]) # K=6, p-value =
    0.5645
138 krupinica.snht3a.po[krupinica.snht4b$estimate,2] # 65
139
140 ww.test(krupinica.snht4a.pred[,1]) #p-value = 0.5071
141 ww.test(krupinica.snht4a.po[,1]) #p-value = 0.0003309
142 krupinica.snht5b <- snh.test(krupinica.snht4a.po[,1]) # K=42, p-value =
    0.6931
143 krupinica.snht4a.po[krupinica.snht5b$estimate,2] # 46
144
145 KrS <- cbind(krupinica_pov[,3], krupinica_pov[,2], rep(0,length(krupinica_
    pov[,2])))
146 KrS[1:4,3] <- rep(1,4)
147 KrS[5:59,3] <- rep(2,55)
148 KrS[60:137,3] <- rep(3,78)
149 KrS[138:410,3] <- rep(4,273)
150 KrS[411:436,3] <- rep(5,26)

```

```

151 KrS[437:1008,3] <- rep(6,572)
152
153 ggplot(data=NULL, aes(x=KrS[,1], y=KrS[,2], color=as.factor(KrS[,3]))) +
  geom_line() + scale_color_manual(values=c("darkgoldenrod1", "green", "
  darkorchid1", "red", "aquamarine3", "cornflowerblue")) + labs(x = "tau", y
  = "Prietok", color = "Zhluk") + theme(panel.background = element_rect(
  fill = "white", colour = "grey50"), legend.key = element_rect(colour = "
  transparent", fill = "white"))

```

Listing 10: Ipeľ

```

1  ipel_pov <- read.xlsx("Data.xlsx", 7, colNames=TRUE)
2  ipel <- ts(ipel_pov[,2], frequency = 12)
3  plot.ts(ipel, xlab='Cas v rokoch od roku 1931', ylab='Prietok v m3/s')
4  title(main = 'Prietok rieky Ipeľ')
5  str_hod_ipel <- rep(0, length(ipel))
6
7  for (i in 1:length(ipel)) {
8    str_hod_ipel[i] <- mean(ipel[1:i])
9  }
10
11 plot(ts(str_hod_ipel, frequency = 12), xlab="Cas v rokoch od roku 1931",
    ylab="Stredna hodnota")
12
13 ipel_dec <- decompose(ts(ipel_pov[,2], frequency = 12), type = "additive")
14 plot(ipel_dec)
15
16 ## Homogenita ##
17 brt_ipel <- bartels.rank.test(ipel)
18 ww_ipel <- ww.test(ipel)
19 print(brt_ipel) # p-value < 2.2e-16
20 print(ww_ipel) # p-value < 2.2e-16
21
22 ## Single Change Point ##
23 pettitt_ipel <- pettitt.test(ipel)
24 print(pettitt_ipel) # K=617, p-value = 0.0009014
25
26 snht_ipel <- snh.test(ipel)
27 print(snht_ipel) # K=126 , p-value < 2.2e-16

```

```

28
29 cusum_ipel <- cpt.mean(ipel, penalty = "Manual", pen.value = 0.05, method =
    "AMOC", test.stat = "CUSUM", class = TRUE, param.estimates = TRUE)
30 ipel_cusum <- cpts(cusum_ipel) # 604 => 603
31
32 plot.ts(ipel_pov[,2], xlab= 'tau', ylab='Prietok v m3/s')
33 lines(x=rep(pettitt_ipel$estimate, length(0:max(ipel_pov[,2]))) , y=0:max(
    ipel_pov[,2]), col='red', lwd=3)
34 lines(x=rep(snht_ipel$estimate, length(0:max(ipel_pov[,2]))) , y=0:max(ipel_
    pov[,2]), col='blue', lwd=3)
35 lines(x=rep(ipel_cusum-1, length(0:max(ipel_pov[,2]))) , y=0:max(ipel_pov
    [,2]), col='green', lwd=3)
36 legend(700, 42, legend=c("Pettitt", "SNHT", "CUSUM"), col = c("red", "blue",
    "green"), lty = 1)
37
38 ## Testovanie najdenych bodov zmeny ##
39 ipel_pred_pettitt <- ipel[1:pettitt_ipel$estimate] # data do bodu zmeny
40 ipel_po_pettitt <- ipel[(pettitt_ipel$estimate+1):length(ipel)] # data po
    zmene
41
42 ipel_pred_snht <- ipel[1:snht_ipel$estimate] # data do bodu zmeny
43 ipel_po_snht <- ipel[(snht_ipel$estimate+1):length(ipel)] # data po zmene
44
45 ipel_pred_cusum <- ipel[1:ipel_cusum-1] # data do bodu zmeny
46 ipel_po_cusum <- ipel[ipel_cusum :length(ipel)] # data po zmene
47
48 mww_ipel_pettitt <- wilcox.test(ipel_pred_pettitt, ipel_po_pettitt,
    alternative = "two.sided")
49 mww_ipel_snht <- wilcox.test(ipel_pred_snht, ipel_po_snht, alternative = "
    two.sided")
50 mww_ipel_cusum <- wilcox.test(ipel_pred_cusum, ipel_po_cusum, alternative =
    "two.sided")
51 print(mww_ipel_pettitt) # p-value = 5.952e-05
52 print(mww_ipel_snht) # p-value = 0.0001229
53 print(mww_ipel_cusum) # p-value = 9.62e-05
54
55 ks_ipel_pettitt <- ks.test(ipel_pred_pettitt, ipel_po_pettitt)
56 ks_ipel_snht <- ks.test(ipel_pred_snht, ipel_po_snht)

```

```

57 ks_ipel_cusum <- ks.test(ipel_pred_cusum, ipel_po_cusum)
58 print(ks_ipel_pettitt) # p-value = 0.0002785
59 print(ks_ipel_snht) # p-value = 0.0001623
60 print(ks_ipel_cusum) # p-value = 0.0004284
61
62 ## Trend ##
63 smk_ipel <- SeasonalMannKendall(ipel)
64 print(smk_ipel) # pvalue =0.049216
65
66 cs_ipel <- cox.stuart.test(ipel)
67 print(cs_ipel) # p-value = 0.2318

```

Listing 11: Ipeľ - viacnásobné

```

1 ipel_pov <- read.xlsx("Data.xlsx", 7, colNames=TRUE)
2 ipel <- ts(ipel_pov[,2], frequency = 12)
3
4 ## Hierarchicke ##
5
6 # Aglomerativny
7
8 krok <- 0.05
9 ip_pocet <- rep(0, (2-0)/krok)
10 ip_k <- rep(0, (2-0)/krok)
11 k <- 0
12 for (i in 1:(length(ip_pocet)-1)){
13   k <- k + krok
14   ip_k[i] <- k
15   ip_pom <- e.agglo(X = as.matrix(ipel), member = 1:length(ipel), alpha = k
16     , penalty = function(cps,Xts) 0)
17   ip_pocet[i] <- length(ip_pom$estimates)
18 }
19 ip_pocet[length(ip_pocet)] <- length((e.agglo(X = as.matrix(ipel), member =
20   1:length(ipel), alpha = 2, penalty = function(cps,Xts) 0))$estimates)
21 ip_k[length(ip_k)] <- 2
22 plot(ip_k, ip_pocet, xlab="gamma", ylab="Pocet bodov zmeny", xlim = c
23   (0.05,2))
24 points(ip_k[which.min(ip_pocet)], ip_pocet[which.min(ip_pocet)], pch = 16,
25   col="red")

```

```

22 which.min(ip_pocet)
23
24 ipel_agglo <- e.agglo(X = as.matrix(ipel), member = 1:length(ipel), alpha =
    ip_k[which.min(ip_pocet)], penalty = function(cps,Xts) 0)
25 ipel_agglo_est <- ipel_agglo$estimates
26 ipel_agglo_val <- rep(0, length(ipel_agglo_est))
27
28 for (i in 1:length(ipel_agglo_est)) {
29   ipel_agglo_val[i] <- ipel_pov[ipel_agglo_est[i],2]
30 }
31 plot(ipel_pov[,2], type='l', xlab="tau", ylab="Prietok v m3/s")
32 points(ipel_agglo_est, ipel_agglo_val, pch = 16, col="red")
33
34 # Divizny
35 ipel_div <- e.divisive(X = as.matrix(ipel), sig.lvl = .05, R = 500, k =
    NULL, min.size = 100)
36 ipel_div_est <- ipel_div$estimates # 138 239 605 => 137 238 604
37 ipel_div_val <- rep(0, length(ipel_div_est))
38
39 for (i in 1:length(ipel_div_est)) {
40   ipel_div_val[i] <- ipel_pov[ipel_div_est[i],3]
41 }
42 ip <- cbind(1:length(ipel), ipel_pov[,2], ipel_div$cluster)
43 length(ipel_div$cluster)
44 ggplot(data=NULL, aes(x=ip[,1], y=ip[,2], color=as.factor(ip[,3]))) + geom_
    line() + scale_color_manual(values=c("darkgoldenrod1", "brown1", "
    darkorchid1", "cornsilk4", "aquamarine3", "cornflowerblue")) + labs(x = "
    tau", y = "Prietok v m3/s", color = "Zhuk") + theme(panel.background =
    element_rect(fill = "white", colour = "black"), legend.key = element_
    rect(colour = "transparent", fill = "white"))
45
46 ## Iteracne ##
47
48 # Pettitt
49 ipel.pettitt0 <- pettitt.test(ipel)
50 ipel_pov[ipel.pettitt0$estimate,3] # 617
51 ipel.pettitt0.pred <- ipel_pov[1:ipel.pettitt0$estimate, 2:3]
52 ipel.pettitt0.po <- ipel_pov[(ipel.pettitt0$estimate+1):length(ipel), 2:3]

```



```

53
54 ww.test(ipel.pettitt0.pred[,1]) #p-value < 2.2e-16
55 ww.test(ipel.pettitt0.po[,1]) #p-value < 2.2e-16
56 ipel.pettitt1a <- pettitt.test(ipel.pettitt0.pred[,1]) # K=238, p-value =
    0.05422
57 ipel.pettitt0.pred[ipel.pettitt1a$estimate,2] # 238
58 ipel.pettitt1b <- pettitt.test(ipel.pettitt0.po[,1]) # K=260, p-value =
    0.03788
59 ipel.pettitt0.po[ipel.pettitt1b$estimate,2] #877
60 ipel.pettitt1a.pred <- ipel.pettitt0.po[1:ipel.pettitt1b$estimate, 1:2]
61 ipel.pettitt1a.po <- ipel.pettitt0.po[(ipel.pettitt1b$estimate+1):length(
    ipel.pettitt0.po[,1]), 1:2]
62
63 ww.test(ipel.pettitt1a.pred[,1]) #p-value = 2.254e-09
64 ww.test(ipel.pettitt1a.po[,1]) #p-value < 2.2e-16
65 ipel.pettitt2a <- pettitt.test(ipel.pettitt1a.pred[,1]) # K=216, p-value =
    0.7245
66 ipel.pettitt1a.pred[ipel.pettitt2a$estimate,2] #833
67 ipel.pettitt2b <- pettitt.test(ipel.pettitt1a.po[,1]) # K=58, p-value =
    0.3254
68 ipel.pettitt1a.po[ipel.pettitt2b$estimate,2] #935
69
70 IpV <- cbind(ipel_pov[,3], ipel_pov[,2], rep(0,length(ipel_pov[,2])))
71 IpV[1:617,3] <- rep(1,617)
72 IpV[618:877,3] <- rep(2,260)
73 IpV[878:1020,3] <- rep(3,143)
74
75 ggplot(data=NULL, aes(x=IpV[,1], y=IpV[,2], color=as.factor(IpV[,3]))) +
    geom_line() + scale_color_manual(values=c("red", "green", "darkorchid1"))
    ) + labs(x = "tau", y = "Prietok", color = "Zhluk") + theme(panel.
    background = element_rect(fill = "white", colour = "grey50"), legend.key
    = element_rect(colour = "transparent", fill = "white"))
76
77 # SNHT
78 ipel.snht0 <- snh.test(ipel)
79 ipel_pov[ipel.snht0$estimate,3] # 126
80 ipel.snht0.pred <- ipel_pov[1:ipel.snht0$estimate, 2:3]
81 ipel.snht0.po <- ipel_pov[(ipel.snht0$estimate+1):length(ipel), 2:3]

```

```

82
83 ww.test(ipel.snht0.pred[,1]) #p-value = 4.273e-05
84 ww.test(ipel.snht0.po[,1]) #p-value < 2.2e-16
85 ipel.snht1a <- snh.test(ipel.snht0.pred[,1]) # K= 60, p-value = 0.0035
86 ipel.snht0.pred[ipel.snht1a$estimate,2] # 60
87 ipel.snht1b <- snh.test(ipel.snht0.po[,1]) # K=478, p-value = 0.03245
88 ipel.snht0.po[ipel.snht1b$estimate,2] #604
89 ipel.snht1a.pred <- ipel.snht0.pred[1:ipel.snht1a$estimate, 1:2]
90 ipel.snht1a.po <- ipel.snht0.pred[(ipel.snht1a$estimate+1):length(ipel.
    snht0.pred[,1]), 1:2]
91 ipel.snht1b.pred <- ipel.snht0.po[1:ipel.snht1b$estimate, 1:2]
92 ipel.snht1b.po <- ipel.snht0.po[(ipel.snht1b$estimate+1):length(ipel.snht0.
    po[,1]), 1:2]
93
94 ww.test(ipel.snht1a.pred[,1]) #p-value = 5.148e-07
95 ww.test(ipel.snht1a.po[,1]) #p-value = 0.08168
96 ww.test(ipel.snht1b.pred[,1]) #p-value < 2.2e-16
97 ww.test(ipel.snht1b.po[,1]) #p-value < 2.2e-16
98 ipel.snht2a <- snh.test(ipel.snht1a.pred[,1]) # K= 4, p-value < 2.2e-16
99 ipel.snht1a.pred[ipel.snht2a$estimate,2] # 4
100 ipel.snht2a.pred <- ipel.snht1a.pred[1:ipel.snht2a$estimate, 1:2]
101 ipel.snht2a.po <- ipel.snht1a.pred[(ipel.snht2a$estimate+1):length(ipel.
    snht1a.pred[,1]), 1:2]
102 ipel.snht2ba <- snh.test(ipel.snht1b.pred[,1]) # K=112, p-value = 0.0093
103 ipel.snht1b.pred[ipel.snht2ba$estimate,2] #238
104 ipel.snht2ba.pred <- ipel.snht1b.pred[1:ipel.snht2ba$estimate, 1:2]
105 ipel.snht2ba.po <- ipel.snht1b.pred[(ipel.snht2ba$estimate+1):length(ipel.
    snht1b.pred[,1]), 1:2]
106 ipel.snht2bb <- snh.test(ipel.snht1b.po[,1]) # K=343, p-value = 1e-04
107 ipel.snht1b.po[ipel.snht2bb$estimate,2] #947
108 ipel.snht2bb.pred <- ipel.snht1b.po[1:ipel.snht2bb$estimate, 1:2]
109 ipel.snht2bb.po <- ipel.snht1b.po[(ipel.snht2bb$estimate+1):length(ipel.
    snht1b.po[,1]), 1:2]
110
111 ww.test(ipel.snht2a.pred[,1]) #p-value = 0.8051
112 ww.test(ipel.snht2a.po[,1]) #p-value = 0.0001962
113 ww.test(ipel.snht2ba.pred[,1]) #p-value = 3.723e-08
114 ww.test(ipel.snht2ba.po[,1]) #p-value < 2.2e-16

```

```

115 ww.test(ipel.snht2bb.pred[,1]) #p-value = 5.826e-14
116 ww.test(ipel.snht2bb.po[,1]) #p-value = 1.271e-09
117 ipel.snht3ab <- snh.test(ipel.snht2a.po[,1]) # K = 30, p-value = 0.7878
118 ipel.snht3ba <- snh.test(ipel.snht2ba.pred[,1]) # K = 11, p-value = 0.6771
119 ipel.snht3bb <- snh.test(ipel.snht2ba.po[,1]) # K = 8, p-value = 0.016
120 ipel.snht2ba.po[ipel.snht3bb$estimate,2] # 246
121 ipel.snht3bb.pred <- ipel.snht2ba.po[1:ipel.snht3bb$estimate, 1:2]
122 ipel.snht3bb.po <- ipel.snht2ba.po[(ipel.snht3bb$estimate+1):length(ipel.
    snht2ba.po[,1]), 1:2]
123 ipel.snht3bba <- snh.test(ipel.snht2bb.pred[,1]) # K = 303, p-value =
    0.7885
124 ipel.snht3bbb <- snh.test(ipel.snht2bb.po[,1]) # K = 16, p-value = 5e-05
125 ipel.snht2bb.po[ipel.snht3bbb$estimate,2] # 963
126 ipel.snht3bbb.pred <- ipel.snht2bb.po[1:ipel.snht3bbb$estimate, 1:2]
127 ipel.snht3bbb.po <- ipel.snht2bb.po[(ipel.snht3bbb$estimate+1):length(ipel.
    snht2bb.po[,1]), 1:2]
128
129 ww.test(ipel.snht3bb.pred[,1]) #p-value = 0.8785
130 ww.test(ipel.snht3bb.po[,1]) # p-value < 2.2e-16
131 ww.test(ipel.snht3bbb.pred[,1]) #p-value = 0.2191
132 ww.test(ipel.snht3bbb.po[,1]) #p-value = 6.513e-10
133 ipel.snht4a <- snh.test(ipel.snht3bb.po[,1]) # K = 303, p-value = 0.2316
134 ipel.snht4b <- snh.test(ipel.snht3bbb.po[,1]) # K = 21, p-value = 0.06055
135
136 IpS <- cbind(ipel_pov[,3], ipel_pov[,2], rep(0,length(ipel_pov[,2])))
137 IpS[1:4,3] <- rep(1,4)
138 IpS[5:60,3] <- rep(2,56)
139 IpS[61:126,3] <- rep(3,66)
140 IpS[127:238,3] <- rep(4,112)
141 IpS[239:246,3] <- rep(5,8)
142 IpS[247:604,3] <- rep(6,358)
143 IpS[605:947,3] <- rep(7,343)
144 IpS[948:963,3] <- rep(8,16)
145 IpS[964:1020,3] <- rep(9,57)
146
147 ggplot(data=NULL, aes(x=IpS[,1], y=IpS[,2], color=as.factor(IpS[,3]))) +
    geom_line() + scale_color_manual(values=c("aquamarine3", "brown", "
    darkorchid1", "cornsilk4", "darkgoldenrod1", "cornflowerblue", "cyan", "

```

```

    black", "red")) + labs(x = "tau", y = "Prietok", color = "Zhuk") + theme
    (panel.background = element_rect(fill = "white", colour = "grey50"),
    legend.key = element_rect(colour = "transparent", fill = "white"))

```

Listing 12: Nitra

```

1  nitra_pov <- read.xlsx("Data.xlsx", 8, colNames=TRUE)
2  nitra <- ts(nitra_pov[,2], frequency = 12)
3  plot.ts(nitra, xlab='Cas v rokoch od roku 1931', ylab='Prietok v m3/s')
4  title(main = 'Prietok rieky Nitra')
5  str_hod_nitra <- rep(0, length(nitra))
6
7  for (i in 1:length(nitra)) {
8    str_hod_nitra[i] <- mean(nitra[1:i])
9  }
10
11 plot(ts(str_hod_nitra, frequency = 12), xlab="Cas v rokoch od roku 1931",
      ylab="Stredna hodnota")
12
13 nitra_dec <- decompose(ts(nitra_pov[,2], frequency = 12), type = "additive")
14 plot(nitra_dec)
15
16 ## Homogenita ##
17 brt_nitra <- bartels.rank.test(nitra)
18 ww_nitra <- ww.test(nitra)
19 print(brt_nitra) # p-value < 2.2e-16
20 print(ww_nitra) # p-value < 2.2e-16
21
22 ## Single Change Point ##
23 pettitt_nitra <- pettitt.test(nitra)
24 print(pettitt_nitra) # K=678, p-value = 0.1451
25
26 snht_nitra <- snh.test(nitra)
27 print(snht_nitra) # K=437, p-value = 0.08705
28
29 cusum_nitra <- cpt.mean(nitra, penalty = "Manual", pen.value = 0.05, method
    = "AMOC", test.stat = "CUSUM", class = TRUE, param.estimates = TRUE)
30 nitra_cusum <- cpts(cusum_nitra) # 438 => 437
31

```

```

32 plot.ts(nitra_pov[,2], xlab= 'tau', ylab='Prietok v m3/s')
33 lines(x=rep(pettitt_nitra$estimate, length(0:max(nitra_pov[,2]))) , y=0:max(
    nitra_pov[,2]), col='red', lwd=3)
34 lines(x=rep(snht_nitra$estimate, length(0:max(nitra_pov[,2]))) , y=0:max(
    nitra_pov[,2]), col='blue', lwd=3)
35 legend(640, 90, legend=c("Pettitt", "SNHT, CUSUM"), col = c("red", "blue", "
    green"), lty = 1)
36
37 ## Testovanie najdenych bodov zmeny ##
38 nitra_pred_pettitt <- nitra[1:pettitt_nitra$estimate] # data do bodu zmeny
39 nitra_po_pettitt <- nitra[(pettitt_nitra$estimate+1):length(nitra)] # data
    po zmene
40
41 nitra_pred_snht <- nitra[1:snht_nitra$estimate] # data do bodu zmeny
42 nitra_po_snht <- nitra[(snht_nitra$estimate+1):length(nitra)] # data po
    zmene
43
44 nitra_pred_cusum <- nitra[1:nitra_cusum-1] # data do bodu zmeny
45 nitra_po_cusum <- nitra[nitra_cusum :length(nitra)] # data po zmene
46
47 mww_nitra_pettitt <- wilcox.test(nitra_pred_pettitt, nitra_po_pettitt,
    alternative = "two.sided")
48 mww_nitra_snht <- wilcox.test(nitra_pred_snht, nitra_po_snht, alternative =
    "two.sided")
49 mww_nitra_cusum <- wilcox.test(nitra_pred_cusum, nitra_po_cusum,
    alternative = "two.sided")
50 print(mww_nitra_pettitt) # p-value = 0.01639
51 print(mww_nitra_snht) # p-value = 0.02246
52 print(mww_nitra_cusum) # p-value = 0.02246
53
54 ks_nitra_pettitt <- ks.test(nitra_pred_pettitt, nitra_po_pettitt)
55 ks_nitra_snht <- ks.test(nitra_pred_snht, nitra_po_snht)
56 ks_nitra_cusum <- ks.test(nitra_pred_cusum, nitra_po_cusum)
57 print(ks_nitra_pettitt) # p-value = 0.006168
58 print(ks_nitra_snht) # p-value = 0.0007312
59 print(ks_nitra_cusum) # p-value = 0.0007312
60
61 ## Trend ##

```

```

62 smk_nitra <- SeasonalMannKendall(nitra)
63 print(smk_nitra) # pvalue =0.067279
64
65 cs_nitra <- cox.stuart.test(nitra)
66 print(cs_nitra) # p-value = 0.9651

```

Listing 13: Nitra - viacnásobné

```

1  nitra_pov <- read.xlsx("Data.xlsx", 8, colNames=TRUE)
2  nitra <- ts(nitra_pov[,2], frequency = 12)
3
4  ## Hierarchicke ##
5
6  # Aglomerativny
7
8  krok <- 0.05
9  n_pocet <- rep(0, (2-0)/krok)
10 n_k <- rep(0, (2-0)/krok)
11 k <- 0
12 for (i in 1:(length(n_pocet)-1)){
13   k <- k + krok
14   n_k[i] <- k
15   n_pom <- e.agglo(X = as.matrix(nitra), member = 1:length(nitra), alpha =
      k, penalty = function(cps,Xts) 0)
16   n_pocet[i] <- length(n_pom$estimates)
17 }
18 n_pocet[length(n_pocet)] <- length((e.agglo(X = as.matrix(nitra), member =
      1:length(nitra), alpha = 2, penalty = function(cps,Xts) 0))$estimates)
19 n_k[length(n_k)] <- 2
20 plot(n_k, n_pocet, xlab="gamma", ylab="Pocet bodov zmeny", xlim = c(0.05,2))
21 points(n_k[which.min(n_pocet)], n_pocet[which.min(n_pocet)], pch = 16, col="
      red")
22 which.min(ip_pocet)
23
24 nitra_agglo <- e.agglo(X = as.matrix(nitra), member = 1:length(nitra),
      alpha = n_k[which.min(n_pocet)], penalty = function(cps,Xts) 0)
25 nitra_agglo_est <- nitra_agglo$estimates
26 nitra_agglo_val <- rep(0, length(nitra_agglo_est))
27

```

```

28 for (i in 1:length(nitra_agglo_est)) {
29   nitra_agglo_val[i] <- nitra_pov[nitra_agglo_est[i],2]
30 }
31 plot(nitra_pov[,2],type='l', xlab="tau", ylab="Prietok v m3/s")
32 points(nitra_agglo_est,nitra_agglo_val, pch = 16, col="red")
33
34 # Divizny
35 nitra_div <- e.divisive(X = as.matrix(nitra), sig.lvl = .05, R = 500, k =
  NULL, min.size = 100)
36 nitra_div_est <- nitra_div$estimates # ziadny
37 nitra_div_val <- rep(0, length(nitra_div_est))
38
39 for (i in 1:length(nitra_div_est)) {
40   nitra_div_val[i] <- nitra_pov[nitra_div_est[i],3]
41 }
42 N <- cbind(1:length(nitra), nitra_pov[,2], nitra_div$cluster)
43 length(nitra_div$cluster)
44 ggplot(data=NULL, aes(x=N[,1], y=N[,2], color=as.factor(N[,3]))) + geom_
  line() + scale_color_manual(values=c("darkgoldenrod1", "brown1", "
  darkorchid1", "cornsilk4", "aquamarine3", "cornflowerblue")) + labs(x = "
  tau", y = "Prietok v m3/s", color = "Zhuk") + theme(panel.background =
  element_rect(fill = "white", colour = "black"), legend.key = element_
  rect(colour = "transparent", fill = "white"))
45
46 ## Iteracne ##
47
48 # Pettitt
49 nitra.pettitt0 <- pettitt.test(nitra)
50 nitra_pov[nitra.pettitt0$estimate,3] # 678
51 nitra.pettitt0.pred <- nitra_pov[1:nitra.pettitt0$estimate, 2:3]
52 nitra.pettitt0.po <- nitra_pov[(nitra.pettitt0$estimate+1):length(nitra),
  2:3]
53
54 ww.test(nitra.pettitt0.pred[,1]) # p-value < 2.2e-16
55 ww.test(nitra.pettitt0.po[,1]) # p-value < 2.2e-16
56 nitra.pettitt1a <- pettitt.test(nitra.pettitt0.pred[,1]) # K = 59, p-value
  = 0.1162
57 nitra.pettitt1b <- pettitt.test(nitra.pettitt0.po[,1]) # K = 77, p-value =

```

0.06217

```
58
59 NV <- cbind(nitra_pov[,3], nitra_pov[,2], rep(0,length(nitra_pov[,2])))
60 NV[1:678,3] <- rep(1,678)
61 NV[679:1044,3] <- rep(2,366)
62
63 ggplot(data=NULL, aes(x=NV[,1], y=NV[,2], color=as.factor(NV[,3]))) + geom_
  line() + scale_color_manual(values=c("red", "green")) + labs(x = "tau",
  y = "Prietok", color = "Zhluk") + theme(panel.background = element_rect(
  fill = "white", colour = "grey50"), legend.key = element_rect(colour = "
  transparent", fill = "white"))
64
65 # SNHT
66 nitra.snht0 <- snh.test(nitra)
67 nitra_pov[nitra.snht0$estimate,3] # 437
68 nitra.snht0.pred <- nitra_pov[1:nitra.snht0$estimate, 2:3]
69 nitra.snht0.po <- nitra_pov[(nitra.snht0$estimate+1):length(nitra), 2:3]
70
71 ww.test(nitra.snht0.pred[,1]) # p-value < 2.2e-16
72 ww.test(nitra.snht0.po[,1]) # p-value < 2.2e-16
73 nitra.snht1a <- snh.test(nitra.snht0.pred[,1]) # K = 410, p-value < 2.2e-16
74 nitra.snht1a.pred <- nitra.snht0.pred[1:nitra.snht1a$estimate, 1:2]
75 nitra.snht1a.po <- nitra.snht0.pred[(nitra.snht1a$estimate + 1):length(
  nitra.snht0.pred[,1]), 1:2]
76 nitra.snht1b <- snh.test(nitra.snht0.po[,1]) # K = 88, p-value = 0.1975
77
78 ww.test(nitra.snht1a.pred[,1]) # p-value < 2.2e-16
79 ww.test(nitra.snht1a.po[,1]) # p-value = 0.3063
80 nitra.snht2a <- snh.test(nitra.snht1a.pred[,1]) # K = 59, p-value = 0.08085
81
82 NS <- cbind(nitra_pov[,3], nitra_pov[,2], rep(0,length(nitra_pov[,2])))
83 NS[1:410,3] <- rep(1,410)
84 NS[411:437,3] <- rep(2,27)
85 NS[438:1044,3] <- rep(3,607)
86
87 ggplot(data=NULL, aes(x=NS[,1], y=NS[,2], color=as.factor(NS[,3]))) + geom_
  line() + scale_color_manual(values=c("red", "green", "darkorchid1")) +
  labs(x = "tau", y = "Prietok", color = "Zhluk") + theme(panel.background
```



```

    = element_rect(fill = "white", colour = "grey50"), legend.key = element
    _rect(colour = "transparent", fill = "white"))

```

Listing 14: Korelácie

```

1  install.packages("corrplot") # potrebný balíček
2  library(corrplot) # nainštalovanie knižnice
3  # Nainštalovanie dát
4  morava_pov <- read.xlsx("Data.xlsx", 1, colNames=TRUE)
5  morava <- ts(morava_pov[,3], frequency = 12)
6
7  hron_pov <- read.xlsx("Data.xlsx", 2, colNames=TRUE)
8  hron <- ts(hron_pov[,3], frequency = 12)
9
10 vah_pov <- read.xlsx("Data.xlsx", 3, colNames=TRUE)
11 vah <- ts(vah_pov[,3], frequency = 12)
12
13 bela_pov <- read.xlsx("Data.xlsx", 4, colNames=TRUE)
14 bela <- ts(bela_pov[,2], frequency = 12)
15
16 kysuca_pov <- read.xlsx("Data.xlsx", 5, colNames=TRUE)
17 kysuca <- ts(kysuca_pov[,2], frequency = 12)
18
19 krupinica_pov <- read.xlsx("Data.xlsx", 6, colNames=TRUE)
20 krupinica <- ts(krupinica_pov[,2], frequency = 12)
21
22 ipel_pov <- read.xlsx("Data.xlsx", 7, colNames=TRUE)
23 ipel <- ts(ipel_pov[,2], frequency = 12)
24
25 nitra_pov <- read.xlsx("Data.xlsx", 8, colNames=TRUE)
26 nitra <- ts(nitra_pov[,2], frequency = 12)
27
28 # Vytvorenie matice, ktorej stĺpce sú dáta za obdobie 01/1931 (03/1931 –
    Bela) až 10/2014 (12/2014 – Bela)
29 data <- cbind(morava[373:1378], hron[1:1006], vah[121:1126], bela[3:1008],
    kysuca[1:1006], krupinica[1:1006], ipel[1:1006], nitra[1:1006])
30 kor <- round(cor(data, method = "kendall"), 2) # vytvorenie korelačnej
    matice
31 corrplot(kor, method = "ellipse") # vykreslenie

```