ACTA HYDROLOGICA
SLOVACA

Ročník 18, č. 1, 2017, 3 – 14

MODELOVANIE MESAČNÝCH PRIETOKOVÝCH RADOV S APLIKÁCOIU SETAR A LSTAR MODELOV

Danuše Szőkeová, Silvia Kohnová

Modelovanie prietokových radov a predpovedanie prietokov patrí medzi dôležité úlohy inžinierskej hydrológie. V práci sme analyzovali mesačné prietokové rady a vytvorili matematické modely niekoľkých radov nameraných na staniciach Banská Bystrica (Hron), Moravský Ján (Morava) a Lekárovce (Uh) v referenčnom období 1961 až 2000. Najskôr sme odhadli systematickú funkciu, ktorá popisuje trendy a cyklické zložky v časových radoch. Následne sme aplikovali viacrežimové modely typu SETAR a LSTAR na reziduálne rady systematickej funkcie a vytvorili aditívne modely priemerných mesačných prietokov. Cieľom bolo otestovať vhodnosť použitia viacrežimových modelov pri modelovaní a simulácii prietokových radov.

KĽÚČOVÉ SLOVÁ: priemerné mesačné prietoky, systematická funkcia, viacrežimové modely typu SETAR a LSTAR, aditívne modely radov

MODELLING OF THE MEAN MONTHLY DISCHARGES USING THE SETAR AND LSTAR REGIME SWITCHING MODELS. The modelling and forecasting of flows is very important problem for reservoir management. Time series analysis based on past records from Slovak stream-gaugin stations is used for building mathematical models and simulation in hydrology. In this paper we made analysis of discharge time series obtained at the stations Banská Bystrica (Hron), Moravský Ján (Morava) and Lekárovce (Uh) in the period 1961 to 2000. At first we have created systematical function to describe trends and cycles in the data. Then we have applied the SETAR and LSTAR regime switching model class for modelling systematic function residual time series and we have created additive discharge models. The aim was to test profitability of regime switching models for modelling and simulation water discharge time series.

KEY WORDS: mean monthly discharges, systematic function, SETAR and LSTAR regime switching models, monthly discharge additive models

Úvod

Poznávanie javov v prírode, porozumenie zákonitostiam, ktorými sa vyznačujú a aplikovanie získaných znalostí v praxi je dlhodobý proces, ktorý umožňuje ľuďom lepšie využívať prírodné zdroje, predvídať prírodné katastrofy a zmierniť ich následky.
Jedným z hlavných problémov hydrológie a vodného hospodárstva v súčasnosti je skúmanie klimatických zmien a ich vplyvov na hydrologické charakteristiky tokov. Zmeny klimatických pomerov môžu v nepredvídanej miere ovplyvniť prirodzené prietoky a spôsobovať ich neplánované výkyvy. Hľadajú sa odpovede na otázky, čo je príčinou povodní a dlhotrvajúcich období sucha, či sa podobné javy vyskytovali aj v minulých obdobiach a či ich intenzita bude rásť aj v budúcich rokoch. Pri hľadaní odpovedí na tieto otázky je nevyhnutné vychádzať z dôkladnej analýzy historických údajov o klimatických a hydrologických javoch (prietoky, teploty, zrážky), ktoré sú základom pre zostavovanie vhodných modelov, ktoré matematickými prostriedkami adekvátne popisujú pozorované javy a ich závislosti. Predstavujú významný prostriedok pre odhaľovanie zákonitostí hydrologických procesov prebiehajúcich v prírode a v konečnom dôsledku nástroj na predpovedanie budúcich udalostí (Komorníková a kol., 2008; Svetlíková a kol., 2008).
V oblasti analýzy meteorologických a hydrologických

časových radov sa v druhej polovici 20. storočia začali uplatňovať metódy klasickej analýzy časových radov, matematickej štatistiky a matematického modelovania. Analyzovala sa najčastejšie časová variabilita hydrologických procesov, cyklické zmeny a extrémne hydrologické javy. Vzniklo veľa odborných prác využívajúcich korelogramy, periodogramy, harmonickú a spektrálnu analýzu, regresné metódy, atď., ako aj veľa dátami riadených matematických modelov, ktoré zohľadňujú vzťahy v jednorozmerných a viacrozmerných prírodných procesoch.
Dátami riadené modely sú založené na extrahovaní a využívaní informácií, ktoré sú implicitne obsiahnuté v hydrologických dátach bez toho, aby sme do modelu priamo zahrnuli fyzikálne zákonitosti, ktoré tvoria základ odtokových procesov. Popularita dátami riadených modelov prietokov, t.j. modelov, ktoré využívajú iba historické údaje, vzrástla prudko v súvislosti s poklesom časových nárokov na odhad parametrov modelov a na spracovania veľkého množstva dát. Tieto modely poskytujú aj možnosti pre fyzikálnu interpretáciu procesov a pre pochopenie javov prebiehajúcich v povodiach, ale predovšetkým umožňujú aj relatívne presné a dostatočne rýchle predpovedanie hodnôt prietokov.
V minulosti sa pri tvorbe modelov súvisiacich s vodnými zdrojmi najčastejšie využívali lineárne modely, ich možnosti boli však vzhľadom na to, že vyžadujú stacionaritu časových radov, dosť obmedzené. Výsledky aplikácie tejto triedy modelov pri modelovaní hydrologických časových radov boli publikované vo viacerých prácach domácich aj zahraničných autorov (Salas, 1993; Koutsoyiannis, 2005; Pekárová, 2003; Szolgay, 2005; Komorníková a Szőkeová, 2008; Svetlíková a kol., 2010).
Modifikáciami lineárnych modelov, ktoré našli uplatnenie v hydrológii, sú aj autoregresné integrované ARIMA modely (Pekárová, 2003) a frakcionálne integrované ARFIMA modely (Szőkeová, 2007). Aplikovanie ARIMA a ARFIMA modelov súvisí so štatistickými vlastnosťami prietokových radov, ako aj s popularitou Box Jenkinsovej metodológie. Komorníková a Szőkeová (2004) testovali priemerné mesačné prieto-ky a ročné prietoky na niekoľkých tokoch (Kysuca, Uh, Topľa, Ľubochnianka, Torysa) vzhľadom na existenciu dlhej pamäte (Long Memory) a v závislosti od výsledku testu použili modely ARFIMA.
Mnoho autorov sa v súčasnosti venuje modelovaniu prietokových a zrážkových radov s aplikáciou nelineárnych viacrežimových modelov (regime switching models). Prvé výsledky modelovania s prepínaním režimov v závislosti od predchádzajúcich (posunutých) hodnôt sa objavili v prácach (Tong, 1978, 1990) a (Tong, Lim, 1980). Amendola (2003) predpovedala s veľmi dobrými výsledkami hodnoty hydrologických radov použitím viacrežimových modelov s prepínaním režimov vzhľadom na prahovú hodnotu indexu API (Antecedent Precipitation Index). Najjednoduchšie viacrežimové, t.j. TAR modely nenašli však v hydrológii väčšie uplatnenie.
Viacrežimové modely typu SETAR a LSTAR boli použité na modelovanie prietokov v oblasti mokrade Kláštorské lúky (Svetlíková, 2006). V práci Szőkeová, Kohnová (2012) sú uvedené pomerne dobré výsledky aplikácie dvojrežimových SETAR modelov pri popise a predpovedi priemerných mesačných prietokov na niekoľkých tokoch v povodí Hrona. Viacrežimové modely môžu byť vhodné aj pre analýzu extrémnych udalostí v časových radoch.

Výber a vstupná analýza dát prietokových radov

Hydrologické časové rady, ktorých analýzu predkladáme, majú tvar diskrétneho časového radu a jedná sa o postupnosti hodnôt priemerných mesačných prietokov. Tieto jednorozmerné časové rady získané na základe meraní vo vodomerných staniciach počas dlhšieho časového obdobia predstavujú základné údaje, pomocou ktorých popíšeme matematické zákonitosti priebehu prietokov a ich zmien spôsobených fyzikálnymi, geografickými a klimatickými činiteľmi, ktoré sa označujú ako hydrologický režim tokov.
Časové rady prietokov, ale aj hodnôt zrážok a teplôt, ktoré sú sledované na území Slovenska v niektorých staniciach už od tridsiatych rokov minulého storočia, predstavujú cenné údaje pre analýzu hydrologických javov a pre ich matematickú interpretáciu v tvare matematického modelu. Pre účely predpovedania budúcich výskytov suchých a vodných období je potrebné vystihnúť trendy i dlhodobú periodickosť v týchto radoch, vytvárať stochastické modely a vychádzať pritom z čo najdlhších historicky pozorovaných prietokových radov. V týchto radoch je potrebné v prvom rade popísať trendy, sezónnosť, ako aj krátkodobé (2 – 5 ročné) a strednodobé (5 – 15 ročné) a dlhodobé (15 – 100 ročné) cykly, ak predpokladáme vylúčenie zmien spôsobených antropogénnou činnosťou.
V tomto príspevku sme použili časové rady priemerných mesačných prietokov získané meraním v troch vodomerných staniciach situovaných na tokoch v západnej, strednej a východnej časti územia Slovenska. Prvým tokom je Hron, vodomerná stanica Banská Bystrica. Hron je druhá najväčšia slovenská rieka, ktorá pramení na svahoch Kráľovej hole a ústi do Dunaja. Dĺžka toku je 280 km a celé povodie má rozlohu takmer 5500 km^2. Vodomerná stanica Banská Bystrica leží na 175,2 km toku Hrona, merania prietokov sa vykonávajú od roku 1931. Ďalším skúmaným tokom je Morava, vodomerná stanica Moravský Ján. Rieka Morava, povodie Dunaj, pramení v Česku na severnej Morave. Na dolnom toku tvorí prirodzenú hranicu medzi Českom a Slovenskom a medzi Slovenskom a Rakúskom. Dĺžka toku je 329 km, z toho na Slovensku 114 km, celé povodie má rozlohu 26580 km^2. Vodomerná stanica Moravský Ján leží na 67,15 km. Na tejto stanici sa vykonávajú merania prietokov od roku 1922. Posledným vybraným tokom je východoslovenská rieka Uh, povodie Bodrog a Hornád. Má celkovú dĺžku 127 km, pramení na Ukrajine, iba 21,3 km preteká územím Slovenska. Stanica Lekárovce leží na 16,6 km, a mera-

nia prietokov sa vykonávajú od roku 1931. Na analýzu prietokov sme v tejto práci vybrali časť údajov, ktoré boli získané v referenčnom období 1961 až 2000, to zn. n = 480 mesiacov.

Pre vzájomné porovnanie rôznych dátových súborov sa využívajú viaceré štatistické charakteristiky, t.j. čísla, ktoré nahradzujú celý dlhý rad hodnôt sledovanej veličiny.

V tabuľke 1 sú uvedené základné štatistické charakteristiky jednotlivých vzoriek priemerných mesačných prietokov. Hodnota μ predstavuje dlhodobý priemerný mesačný prietok. Hodnoty štandardných odchýliek σ informujú o rozkolísanosti prietokov počas daného obdobia. Priemerné prietoky Hronu a Uhu sú porovnateľné, avšak tok Uh má oveľa vyššiu štandardnú odchýlku ako Hron. Tok Morava má najväčší prietok aj štandardnú odchýlku z vybraných radov.

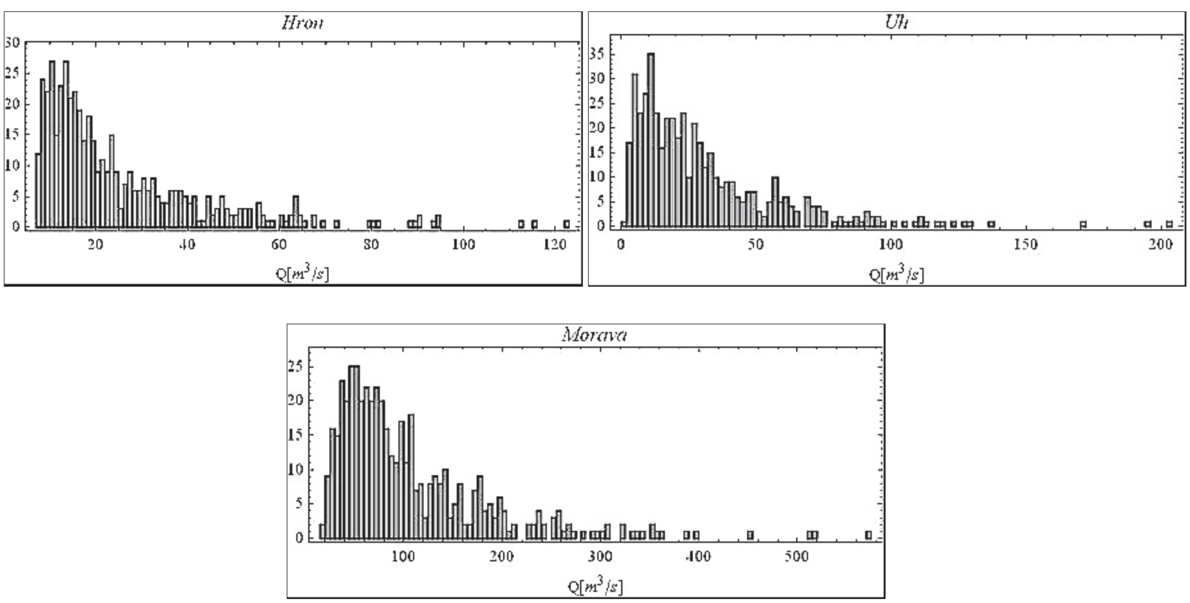
Všetky analýzy a výpočty v tejto práci sme realizovali pomocou výpočtového systému Wolfram Mathematica 9.0, v ktorom je k dispozícii veľa procedúr uľahčujúcich aplikáciu rôznych matematických metód a algoritmov.

Histogram je grafická reprezentácia, pomocou ktorej možno zistiť pravdepodobnostné rozdelenie dát. Z obr. 1 je zrejmé, že histogramy mesačných prietokov analyzovaných staníc majú niekoľko vrcholov a dáta nemajú štandardné pravdepodobnostné rozdelenie. Pravdepodobnosť výskytu extrémne vysokých prietokov je na skúmaných tokoch najväčšia v prípade toku Uh.

Tabuľka 1. Základné štatistické charakteristiky priemerných mesačných prietokov na vybraných tokoch, referenčné obdobie 1961 – 2000

Table 1. Basic statistical characteristics of the mean monthly discharges on selected rivers from the period 1961 – 2000

Č.	Tok	Vodomerná stanica	Plocha povodia [km^2]	Priemerné mesačné prietoky (1961-2000) [m^3/s]				
				Min	Max	Medián	μ	σ
1.	***Hron***	Banská Bystrica	1766,48	7,3	122,9	18,6	25,44	18,63
2.	***Morava***	Moravský Ján	24129,30	15,1	574,8	80,45	106,33	80,93
3.	***Uh***	Lekárovce	1989,41	1,8	203,0	22,3	30,64	28,25



Obr. 1. Histogramy hodnôt priemerných mesačných prietokov, obdobie 1961 – 2000.

Fig. 1. Histograms of the mean daily discharges, period1961 – 2000.

Trendy a cyklické procesy v mesačných prietokoch

Časové rady považujeme za stacionárne, ak neobsahujú trendy alebo cyklické zložky. V takom prípade sú štatistické charakteristiky ako stredná hodnota, rozptyl, štandardná odchýlka konštantné v čase. V opačnom prípade sú rady nestacionárne. Vo všeobecnosti možno povedať, že ročné hydrologické časové rady sú spravidla stacionárne. Tento predpoklad nemusí byť pravdivý v prípade dlhodobej klimatickej variability, prírodných zlomových udalostí, ako sú vulkanické erupcie alebo zmeny zapríčinených ľudskou činnosťou, napríklad stavbou priehrad a zmenou tokov. Hydrologické rady definované v časových intervaloch kratších ako rok, napríklad v mesačných intervaloch, sa vyznačujú spravidla nestacionaritou, predovšetkým ako dôsledok ročnej sezónnosti.
Prírodné procesy a ľudská činnosť spôsobujú dlhodobé zmeny, ktoré sa označujú ako trendy. Zdrojom trendov a ich zlomov v prietokoch môžu byť klimatické zmeny, ako aj zásahy človeka v povodiach, preto je vždy potrebná analýza tejto zložky časových radov (Komorníková a Szőkeová, 2005).
Časové rady, ktoré reprezentujú pozorovania prírodných procesov v intervaloch kratších ako rok (denné, týždenné, mesačné) sa často vyznačujú sezónnou periodickou zložkou. Sezónnosť alebo iné cykly v časových radoch sa prejavujú v štatistických charakteristikách, ktoré sa menia počas roka, resp. týždňa, dňa a pod.
Hodnoty časového radu y_t možno potom vyjadriť ako súčet hodnôt štyroch systematických zložiek (dekompozícia časového radu):

$$y_t = T_t + S_t + C_t + z_t \tag{1}$$

kde
T_t – je dlhodobý trend,
S_t – je sezónna zložka,
C_t – je cyklická (nesezónna) zložka,
z_t – je chybová (nesystematická, reziduálna) zložka (zahrňuje nepravidelnosti).

Časové rady priemerných mesačných prietokov vykazujú existenciu všetkých týchto zložiek, ktoré popisujeme matematickými funkciami. Pri odhade systematickej funkcie sa postupuje nasledovne:
– pomocou lineárnej alebo nelineárnej regresnej analýzy získame matematickú funkciu trendu (v prietokových radoch sme identifikovali ako najvhodnejší lineárny trend);
– z časového radu získaného odpočítaním hodnôt trendu od pozorovaných dát určíme pomocou regresnej analýzy sezónnu zložku, ktorá reprezentuje periodické zmeny opakujúce sa v priebehu roka (má tvar súčtu funkcií sínus a kosínus s periódou odpovedajúcou jednému roku);
– z časového radu získaného odpočítaním hodnôt trendu a sezónnej funkcie stanovíme pomocou spektrálnej analýzy a Fischerovho testu významné frekvencie, ktoré popisujú periodické zmeny, v ktorých dĺžka periódy nezodpovedá kalendárnym jednotkám (cyklická funkcia má tvar súčtu sínusov a kosínusov pre každú významnú frekvenciu s koeficientmi odhadnutými pomocou regresnej analýzy) (Cipra, 1986);
– reziduálnu zložku dostaneme odpočítaním hodnôt systematickej funkcie od nameraných údajov.
Odhad systematickej funkcie, sme urobili pre všetky vybrané rady priemerných mesačných prietokov. V tabuľke č. 2 sú uvedené parametre lineárneho trendu, významné cykly periodickej zložky a štatistické charakteristiky syntetických radov vyčíslených ako hodnoty systematických funkcií.
Na všetkých vybraných tokoch je smernica lineárneho trendu prietokov záporná, čo znamená veľmi mierny pokles prietokov v priebehu referenčného obdobia 1961 – 2000. Vo všetkých prietokových radoch boli zistené okrem 12-mesačného cyklu (sezónnosť radu), aj cykly s dĺžkou 4, 6, 14 a 120 mesiacov a cyklus 60 mesiacov (okrem toku Uh). Stredné hodnoty simulovaných radov na základe systematickej funkcie sú porovnateľné so strednými hodnotami pozorovaných radov (tab. 1). Štandardné odchýlky syntetických radov sú o niečo menšie.

Tabuľka 2. Parametre lineárneho trendu, významné cykly, štatistické charakteristiky hodnôt generovaných systematickou funkciou, obdobie 1961 – 2000
Table 2. Parameters of the trend function, significant cycles and statistics of the valus generated by systematical function, period 1961 – 2000

Systematické funkcie priemerných mesačných prietokov (1961 – 2000)								
Č.	**Tok**	**Lineárny trend a+b.t**		**Významné cykly**	**Štatistické charakteristiky syntetických radov**			
		a	**b**	*[mes]*	**μ**	**σ**	R^2	σ_{rez}
1.	***Hron***	29,62	–0,017	12,4,6,14,24,30,44,60,120	25,44	13,19	0,50	13,18
2.	***Morava***	116,70	–0,043	12,4,6,14,20,30, 60,120	106,40	54,58	0,46	59,62
3.	***Uh***	31,06	–0,002	12,3,4,6,8,14,24,48,120	30,64	16,10	0,32	23,21

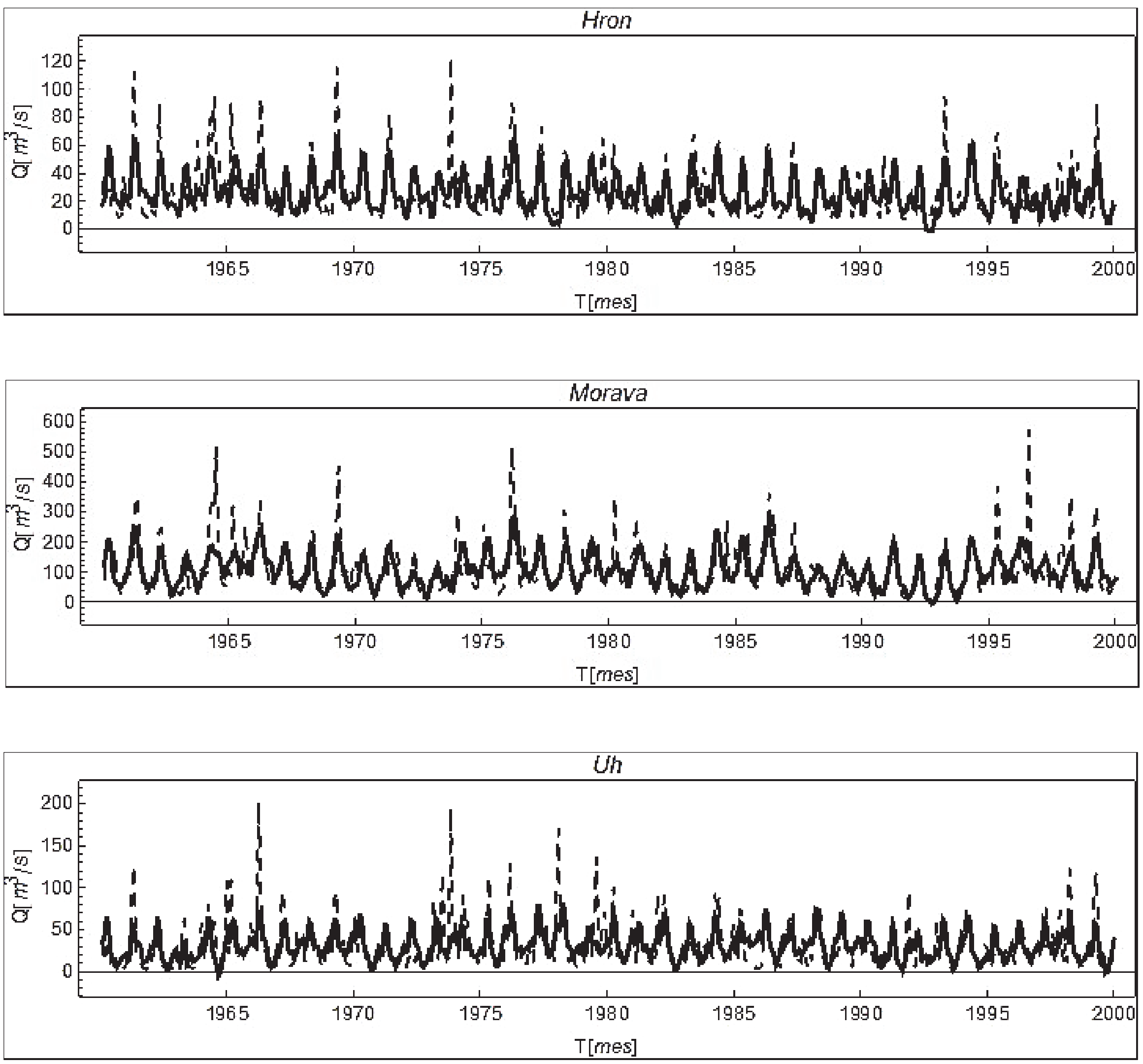
V tabuľke 2. je uvedená aj hodnota koeficientu determinácie R^2. Stručne možno povedať, že tento koeficient určuje, aký podiel variability pozorovanej premennej je vysvetlený systematickou funkciou a ilustruje praktickú cennosť takéhoto modelu. Čím je väčšia hodnota koeficientu R^2, tým adekvátnejší je matematický popis údajov. Z výsledkov je zrejmé, že systematická funkcia predstavuje z hľadiska tohto koeficientu lepší matematický model v prípade tokov Hron (50 %) a Morava (46 %), najhorší pre tok Uh (32 %).

V poslednom stĺpci tabuľky sú uvedené hodnoty smerodajnej odchýlky σ_{rez} reziduálneho časového radu, ktorú treba v procese tvorby matematických modelov minimalizovať. Stredné hodnoty rezíduí sú vo všetkých prípadoch rovné nule.

Grafická reprezentácia pozorovaných mesačných prietokov (čiarkovaná čiara) a graf systematickej funkcie mesačných prietokov (plná čiara) je na obrázku 2.

Nelineárne modely časových radov

Pred viac ako dvadsiatimi rokmi uverejnil Hamilton článok o nelineárnom modelovaní hodnôt hrubého domáceho produktu, čím sa začala etapa záujmu o problematiku nelineárnych modelov (Hamilton, 1989; Hamilton, 1994). Tieto modely sa uplatňujú v rôznych oblastiach, napríklad v ekonomike, ale aj hydrológii, geodézii, atď. Dve významné triedy nelineárnych modelov predstavujú modely s podmienenou heteroskedasticitou a viacrežimové modely.



Obr. 2. Priemerné mesačné prietoky a grafy systematických funkcií priemerných mesačných prietokov, obdobie 1961 – 2000.
Fig. 2. Mean monthly discharges and time series of systematical functions of mean monthly discharges, period 1961 – 2000.

V*iacrežimové modely* alebo ***modely s premenlivými režimami*** (regime switching) jednoduchým spôsobom zachycujú zmeny autokorelácie, resp. rozptylu časového radu, ktoré možno chápať ako zmeny správania sa procesu. Formalizujú myšlienku, že existuje určitý počet rôznych režimov generovaných stochastickým procesom. Jednotlivé režimy môžu byť reprezentované napríklad množinou jednoduchých lineárnych modelov (Hamilton, 1989; Teräsvirta, 1994). Nelineárne viacrežimové modely nemusia byť jednoznačne prostriedkom na úspešnú simuláciu procesov, majú svoje úskalia, ktoré sa týkajú metód pre odhad parametrov základných modelov a spôsobov prepínania medzi jednotlivými režimami (Franses a van Dijk, 2000).

Doteraz bolo navrhnutých niekoľko typov viacrežimových modelov, ktoré sa odlišujú napríklad spôsobom prepínania režimov. Prepínanie môže byť určené hodnotou ***pozorovateľnej veličiny*** (observable variable) alebo ***nepozorovateľnej veličiny*** (unobservable variable). Pri simulácii procesu sa v každom časovom okamihu vyčísľuje hodnota jedného režimu (prípadne ich lineárna kombinácia).

Okrem systematickej zložky obsahujú pozorované časové rady náhodnú reziduálnu zložku tvorenú fluktuáciami v priebehu časového radu, ktoré nemajú spravidla rozpoznateľný systematický charakter. Táto zložka by mala predstavovať Gussovský biely šum. Ak to tak nie je, pristúpi sa k špecifikácii a odhadu nejakého stochastického matematického modelu reziduálneho časového radu. Najčastejšie sa na ako prvý aplikuje lineárny model, tento postup sme použili a popísali vo viacerých prácach (Komorníková a Szőkeová, 2004, 2008) a model prietoku je potom aditívny model vytvorený ako súčet systematickej funkcie a stochastického lineárneho modelu rezíduí. V tomto príspevku sme reziduálny rad systematickej funkcie modelovali použitím dvoch typov modelov z triedy viacrežimových modelov s prepínaním na základe pozorovateľnej premennej, a síce trojrežimového SETAR modelu a dvojrežimového LSTAR modelu.

Modely priemerných mesačných prietokov s aplikáciou trojrežimových SETAR modelov rezíduí systematickej funkcie

Za ***jednorozmerný diskrétny časový rad*** považujeme množinu pozorovaní nejakého javu $\{y_1,y_2,\ldots,y_t,...\}$, pričom predpokladáme, že pozorovania sa vykonávajú v rovnakých časových intervaloch. Index t možno považovať za čas, to zn. y_t je hodnota pozorovania v čase t. Celkový počet prvkov v časovom rade nazývame ***dĺžka časového radu*** a budeme ju označovať n.

SETAR model **(**SelfExciting Threshold AutoRegressive) je model, v ktorom je režim v čase t určený hodnotou nejakej ***prahovej premennej*** (threshold variable), ktorá nadobúda ***posunuté (oneskorené) hodnoty*** analyzovaného časového radu y_{t-d} (lagged endogenous variable), kde ***parameter oneskorenia*** d (delay parameter), $d > 0$, je celé číslo udávajúce posunutie. V prípade dvojrežimového modelu sa hodnota prahovej premennej q_t porovnáva s ***prahovou hodnotou*** c (threshold value); prvý režim nastane, ak platí $q_t \le c$, druhý režim, ak $q_t > c$.

Vektory autoregresných parametrov lineárnych modelov jednotlivých režimov, ktoré sú závislé od hodnoty prahovej premennej q_t, označíme $\boldsymbol{\Phi}_i(q_t)=\{\phi_{i,0},\ldots,\phi_{i,p}\}$, $i = 1,2$, $p=\max(p_1,p_2)$. Hodnoty dvojrežimového SETAR(c; p_1,p_2,d) modelu s prahovou hodnotou c, so základnými autoregresnými lineárnymi modelmi AR(p), v každom režime a s oneskorením d možno popísať:

$$y_t = (\phi_{1,0}+\phi_{1,1}y_{t-1}+\ldots+\phi_{1,p}y_{t-p}+{}^1\varepsilon_t)\, I[y_{t-d}\le c] + (\phi_{2,0}+\phi_{2,1}y_{t-1}+\ldots+\phi_{2,p}y_{t-p}+{}^2\varepsilon_t)\, I[y_{t-d}>c], \qquad (2)$$

kde

$\phi_{i,0}+\phi_{i,1}\, y_{t-1}+\ldots+\phi_{i,p}\, y_{t-p}$, i=1,2, – sú autoregresné polynómy, popisujú režimy určené hodnotou prahovej premennej vzhľadom na prahovú hodnotu c,

I[A] –je indikačná funkcia s hodnotami I[A] = 1 (ak udalosť A nastane), I[A] = 0 (v opačnom prípade),

$\{{}^i\varepsilon_t\}$ –je proces i.i.d.N(0,${}^i\sigma^2$) (independent and identically distributed), i=1,2.

Tvorba dvojrežimového SETAR sa začína špecifikáciou parametrov d a p_i, i=1,2. Prahová hodnota c musí byť zvolená tak, aby obidva režimy obsahovali dostatočný počet hodnôt. Nasleduje odhad autoregresných koeficientov $\boldsymbol{\Phi}_i$ základných modelov AR(p_i), i=1,2, použitím metódy najmenších štvorcov. Nakoniec sa vyberie výsledný model z modelov odhadnutých pre všetky hodnoty c daný množinou parametrov $\boldsymbol{\theta}=\{c,p_1,p_2,\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_2,\sigma_1,\sigma_2\}$.

Pri experimentoch s aplikáciou SETAR modelov v aditívnych modeloch prietokových radov sa ukázalo, že je potrebné použiť najmenej tri režimy. Odhad trojrežimového modelu SETAR($c_1,c_2;p_1,p_2,p_3$,d) obsahuje tie isté kroky, treba však odhadovať autoregresné parametre troch režimov popísaných základnými modelmi AR(p_i), i=1,2,3, v závislosti na dvoch prahových hodnotách c_1, c_2.

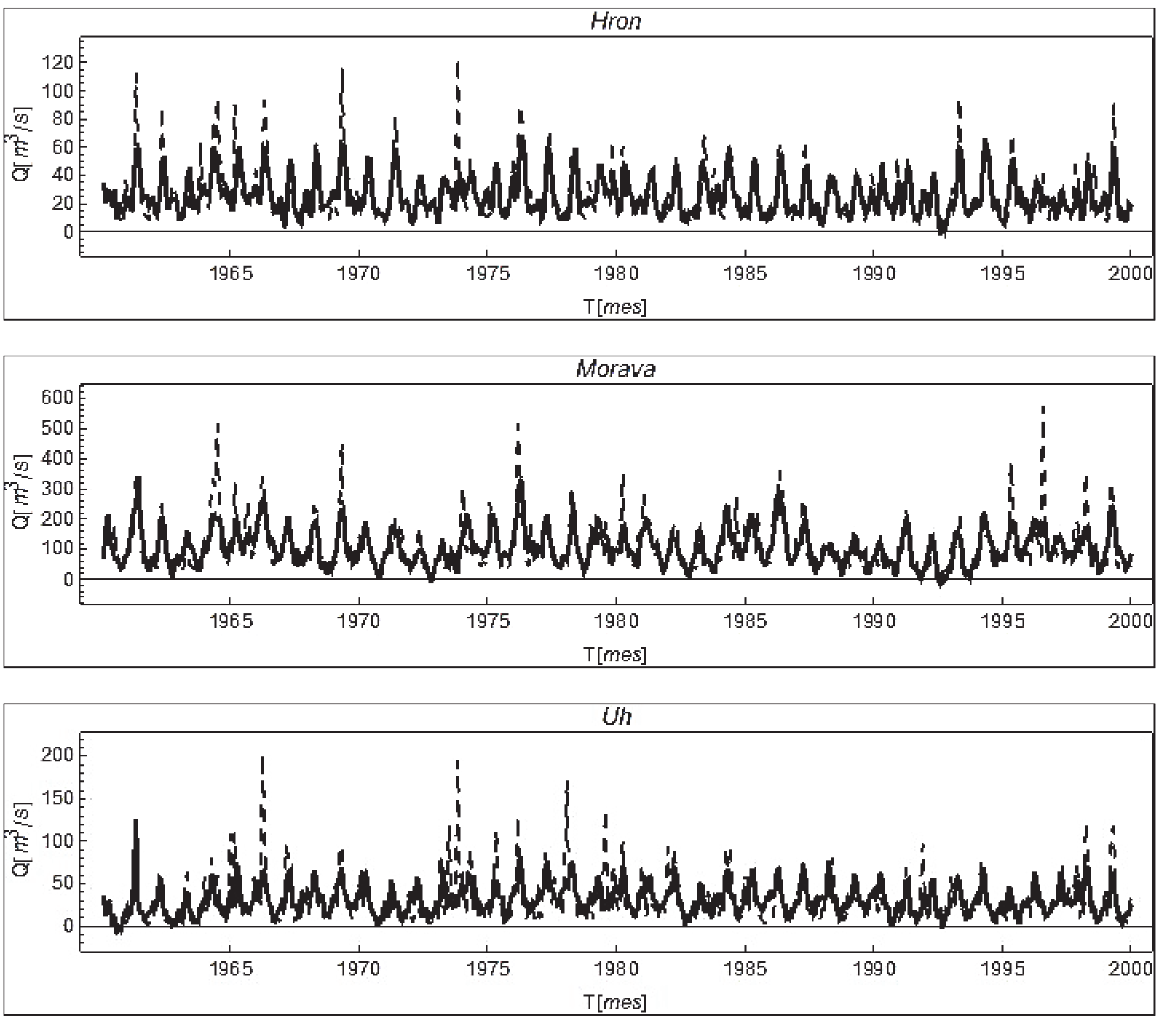
Výsledky odhadu trojrežimových SETAR modelov reziduálnych radov mesačných prietokov pre vybrané toky, ako aj štatistické charakteristiky syntetických radov generovaných pomocou aditívnych modelov, t.j. súčtu hodnoty systematickej funkcie a simulovanej hodnoty na základe trojrežimového SETAR modelu, sú v tabuľke 3. Vo všetkých SETAR modeloch sme použili základné autoregresné AR(p) modely maximálne rádu p=3, posunutie bolo identifikované vo všetkých prípadoch hodnotou d=1.

Na obrázku 3 sú zobrazené pozorované hodnoty priemerných mesačných prietokov (čiarkovaná čiara) a simulované hodnoty mesačných prietokov (plná čiara) na základe aditívnych modelov s trojrežimovými SETAR modelmi rezíduí systematickej funkcie.

Tabuľka 3. Parametre trojrežimového SETAR($c_1,c_2;p_1,p_2,p_3,d$) modelu rezíduí systematickej funkcie a štatistické charakteristiky syntetického radu priemerných mesačných prietokov generovaných aditívnym modelom (systematická funkcia + trojrežimový SETAR model rezíduí), obdobie 1961 – 2000

Table 3. Parameters of the SETAR (c1, c2; p1, p2, p3, d) residual model of the systematic function and the statistical characteristics of the synthetic timeseries of the average monthly discharges generated by the additive model (systematic function + SETAR residual model), period 1961 to 2000

Tok (1961–2000)	Parametre trojrežimového modelu rezíduí SETAR(c_1,c_2,p_1,p_2,p_3,d)									Štatistické charakteristiky simulovaných hodnôt aditívneho modelu prietokov	
	c_1	c_2	p_1	p_2	p_3	d	σ_1	σ_2	σ_3	μ	σ
Hron	–5,26	1,50	3	1	1	1	3,34	2,59	3,14	25,35	13,46
Morava	16,35	46,71	2	1	2	1	7,63	7,71	7,38	106,90	60,51
Uh	–8,53	8,94	1	3	1	1	2,67	2,97	2,94	30,45	17,73



Obr. 3. Priemerné mesačné prietoky a simulované hodnoty na základe aditívneho modelu (systematická funkcia + trojrežimový SETAR model rezíduí), obdobie 1961 – 2000.
Fig. 3. Mean monthly discharges and simulated time series based on the additive model (systematic function + SETAR residual model), period 1961-2000.

Modely priemerných mesačných prietokov s aplikáciou dvojrežimových LSTAR modelov rezíduí systematickej funkcie

Vo viacrežimových modeloch typu SETAR sa predpokladá, že hranica medzi dvoma režimami je daná prahovou hodnotou c. Nakoľko je indikačná funkcia $I[y_{t-d} > c]$, nespojitá, prechod z jedného režimu do druhého je okamžitý, skokový. Nahradením indikačnej funkcie spojitou funkciou $G(q_t;c)$, ktorej hodnota s rastúcou premennou q_t plynule prechádza od hodnoty 0 po hodnotu 1, sa dosiahne to, že prechod z jedného režimu do druhého je plynulý, spojitý. Režim v čase t je určený hodnotou ***prechodovej funkcie*** (transition function) $G(q_t;c)$, kde q_t je ***prechodová premenná*** (transition variable) a c je ***prahová hodnota***. Takýto model označujeme ako STAR (**S**mooth**T**ransition **A**uto**R**egressive).

STAR model považovať za viacrežimový model, ktorý pripúšťa dva režimy odpovedajúce extrémnym hodnotám prechodovej funkcie $G(q_t;c) = 0$ a $G(q_t;c) = 1$, pričom prechod od jedného režimu k druhému je hladký, alebo možno STAR model považovať aj za "kontinuum" režimov odpovedajúcich jednotlivým hodnotám funkcie $G(q_t;c)$ z intervalu od 0 po 1. V tejto práci budeme uvažovať prvú interpretáciu, v ktorej uvedený STAR model považujeme za dvojrežimový so spojitým prechodom jedného režimu do druhého. Rôzne voľby prechodovej funkcie $G(q_t;c)$ umožňujú modelovanie rôznych typov viacrežimového správania.

Ak je prechodovou funkciou logistická funkcia s premennou q_t (nadobúda posunuté hodnoty časového radu y_{t-d}), t.j.:

$$L(q_t;\gamma,c) = \frac{1}{1+\frac{1}{e^{\gamma(q_t-c)}}} = 1/(1+\exp(-\gamma(q_t-c)))^{-1},\ \gamma > 0 \quad (3)$$

Výsledný model sa označuje ako logistický LSTAR(γ,c; p_1,p_2,d). Parameter c v (3) možno interpretovať ako ***prahovú hodnotu*** medzi dvoma režimami odpovedajúcim hodnotám logistickej funkcie $L(q_t;\gamma,c) = 0$ a $L(q_t;\gamma,c) = 1$, pričom platí $L(c;\gamma,c) = 0{,}5$. Parameter γ je ***parameter vyhladzovania*** (smoothing parameter), určuje hladkosť prechodu od jedného režimu k druhému (Dijk, Teräsvirta, Franses, 2000).

Hodnoty dvojrežimového LSTAR(γ,c;p_1,p_2,d) modelu s prahovou hodnotou c, so základnými autoregresnými lineárnymi modelmi AR(p), $p = \max(p_1,p_2)$, v každom režime možno opísať:

$$y_t=(\phi_{0,1}+\phi_{1,1}y_{t-1}+\ldots+\phi_{p,1}y_{t-p})(1-L(y_{t-d};\gamma,c)+(\phi_{0,2}+ +\phi_{1,2}y_{t-1}+\ldots+\phi_{p,2}y_{t-p})\ L(y_{t-d};\gamma,c)\ )+\varepsilon_t \quad (4)$$

kde:

$(\phi_{0,i}+\phi_{1,i}y_{t-1}+\ldots+\phi_{p,i}y_{t-p})$, i=1,2, – sú autoregresné polynómy v jednotlivých režimoch,

$L(y_{t-d};\gamma,c)$ je prechodová logistická funkcia (neklesajúce zobrazenie prechodovej premennej q_t na interval [0,1]),

C – je prahová hodnota,

γ – je parameter vyhladzovania,

$\{\varepsilon_t\}$ –je proces i.i.d.$N(0,\sigma^2)$.

Odhad dvojrežimového LSTAR modelu začneme špecifikáciou maximálnej hodnoty parametra p, $p=\max(p_1, p_2)$. Hodnoty ďalších parametrov odhadujeme v procese iterácie, v ktorom meníme postupne hodnoty parametrov c,d a odhadujeme autoregresné parametre v jednotlivých režimoch, v závislosti od nich sa nakoniec určí parameter vyhladzovania γ. Výsledkom je viacero LSTAR modelov, z ktorých vyberieme konkrétny na základe metódy najmenších štvorcov.

Výsledky odhadu dvojrežimových LSTAR modelov rezíduí systematickej funkcie mesačných prietokoch na troch vybraných tokoch a štatistické parametre simulovaných časových radov sú uvedené v nasledujúcej tabuľke 4.

Tabuľka 4. Parametre dvojrežimového LSTAR (g,c;p_1,p_2,d) modelu rezíduí systematickej funkcie a charakteristiky simulovaných hodnôt priemerných mesačných prietokov generovaných aditívnym modelom, obdobie 1961 – 2000

Table 4. Parameters of the LSTAR (g,c;p_1,p_2,d) residual model of the systematic function and the statistical characteristics of the synthetic timeseries of the average monthly discharges generated by the additive model, period 1961 to 2000

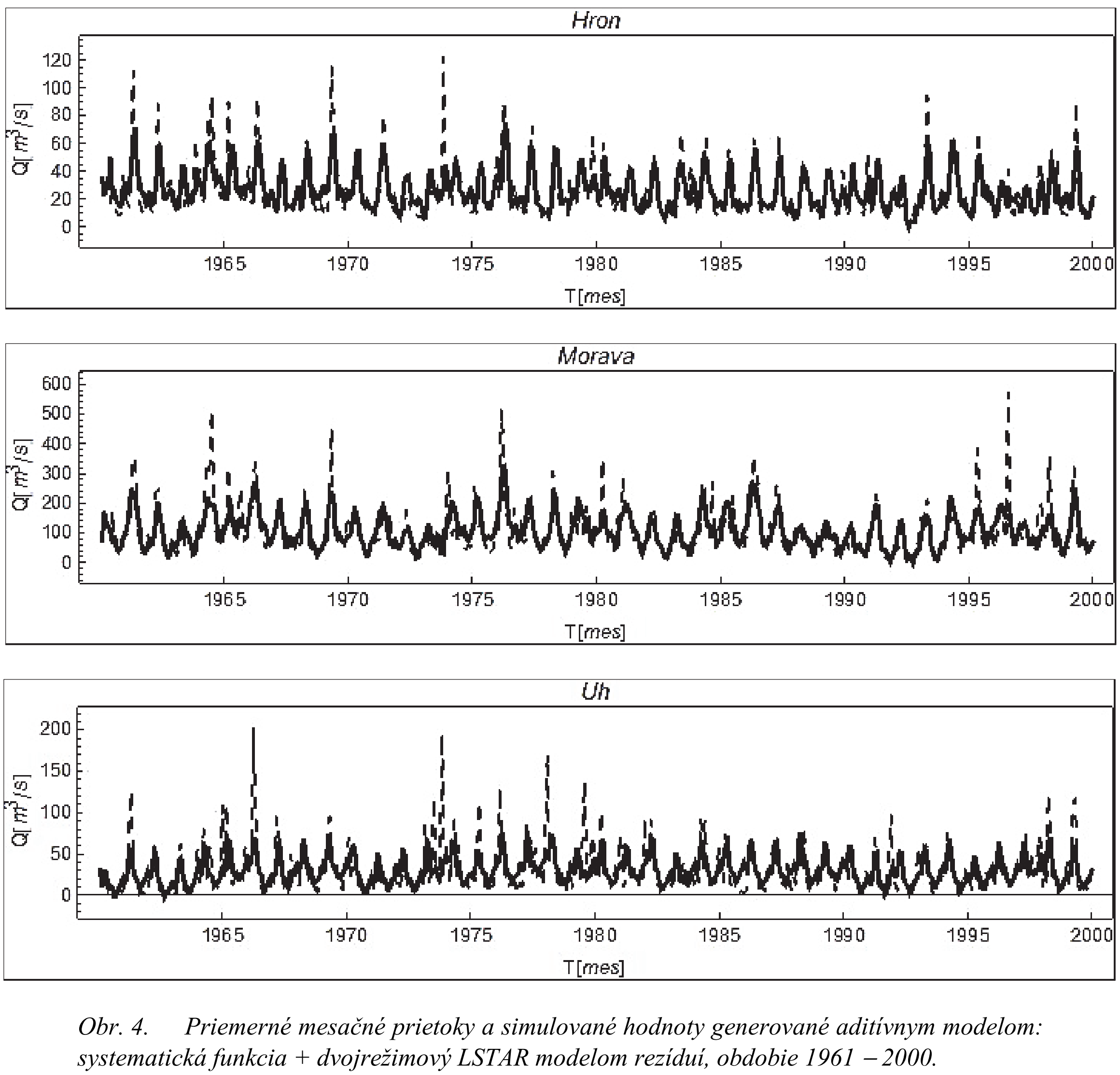
Tok (1961–2000)	Parametre dvojrežimového LSTAR(g,c;p_1,p_2,d) modelu						Štatistické charakteristiky simulovaných hodnôt aditívneho modelu	
	g	c	p_1	p_2	d	σ	μ	σ
Hron	0,1	–3,01	5	4	1	0,15	25,27	13,80
Morava	0,1	18,23	4	2	1	0,56	103,63	57,14
Uh	0,1	1,62	5	2	1	0,22	30,56	16,38

Grafy pozorovaných prietokov (čiarkovaná čiara) a syntetických časových radov (plná čiara) na základe aditívnych modelov s dvojrežimovými LSTAR modelmi rezíduí systematickej funkcie sú na nasledujúcom obr. 4.

Diagnostika a porovnanie modelov priemerných mesačných prietokov

Každý model odhadnutý pre danú množinu dát je potrebné na záver verifikovať, či je naozaj vhodným modelom. Ak model nie je adekvátny, robia sa modifikácie jeho parametrov, prípadne sa vyberie iný typ modelu a celý proces špecifikácie, odhadu parametrov a diagnostickej kontroly sa musí zopakovať. Adekvátnosť modelu možno posúdiť na základe vizualizácie modelovaných údajov spolu so syntetickými hodnotami generovanými modelom. Tieto vizualizácie pre jednotlivé modely sme prezentovali na obr. 2, 3, 4.

Okrem toho je všeobecne používaným spôsobom diagnostickej kontroly modelu testovanie jeho rezíduí. Ako prvé sa zobrazí reziduálny časový rad modelu, aby sme vizuálne zistili, či je tento rad stacionárny, t.j. či má konštantný rozptyl (homoskedasticita) alebo vykazuje v priebehu času väčšie výkyvy (heteroskedasticita). Často sa na testovanie rezíduí používajú rôzne štatistické testy, ktorými sa testuje normalita a korelovanosť rezíduí. V prípade aditívnych modelov, ktoré sme popísali, sú všetky reziduálne rady nekorelované. Hodnoty štandardných odchýliek rezíduí σ_{rez} sú uvedené v tab. 5, pričom stredné hodnoty reziduálnych radov sú vo všetkých modeloch nulové. Aditívne modely majú menšiu štandardnú odchýlku rezíduí ako systematické funkcie.



Obr. 4. Priemerné mesačné prietoky a simulované hodnoty generované aditívnym modelom: systematická funkcia + dvojrežimový LSTAR modelom rezíduí, obdobie 1961 – 2000.
Fig.4. The mean monthly discharges and the simulated values based on the additive model (systematic function + LSTAR residual model), period 1961 – 2000.

Tabuľka 5. Hodnoty koeficientu determinácie R^2, miery presnosti MAPE a štandardnej odchýlky reziduálneho radu systematickej funkcie a aditívnych modelov priemerných mesačných prietokov

Table 5. Values of the coefficient of determination, accuracy rates of MAPE and standard deviation of the residual series of systematic function and additive models

Tok (1961-2000)	Systematická funkcia			Syst.f. + SETAR3			Syst.f. + LSTAR2		
	R^2	MAPE	σ_{rez}	R^2	MAPE	σ_{rez}	R^2	MAPE	σ_{rez}
Hron	0,50	44,33	13,18	0,57	37,94	12,38	0,56	36,72	12,40
Morava	0,46	47,64	59,62	0,52	42,57	56,00	0,52	40,36	56,17
Uh	0,32	85,15	23,21	0,37	79,29	22,46	0,36	81,18	22,55

Ak odhadujeme viac typov modelov pre pozorované údaje, existuje niekoľko možností, ako ich navzájom porovnať a posúdiť ich kvalitu. Pre každý model sa vyčísľujú miery presnosti. V tomto príspevku na porovnanie použijeme ***priemernú absolútnu percentuálnu odchýlku*** (Mean Absolute Percentage Error)

$$\mathrm{MAPE}=\frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \tag{5}$$

Na základe údajov uvedených v tab. 5 možno konštatovať, že vzhľadom na mieru presnosti MAPE (čím menšia, tým lepší model), hodnoty koeficientu determinácie R^2 (čím väčší, tým lepší model) a štandardnej odchýlky reziduálneho radu σ_{rez} (čím menšia, tým lepší model), sú pre popisné účely všetkých prietokových radov medzi aditívnymi modelmi len veľmi malé rozdiely.

Aditívne modely mesačných prietokov s trojrežimovým SETAR modelom rezíduí mali pre všetky prietokové rady vyššie hodnoty parametra determinácie R^2 v porovnaní s hodnotami koeficientu determinácie systematickej funkcie, najviac v prípade toku Hron, v percentuálnom vyjadrení o 7 %, najmenej pre tok Uh, o 5 %, čo možno považovať len za mierne skvalitnenie modelu pri popise dát v porovnaní so systematickou funkciou. Štandardné odchýlky rezíduí tohto modelu σ_{rez} sú v porovnaní s rezíduami systematickej funkcie taktiež len o málo menšie.

Aditívne modely mesačných prietokov, v ktorých sú syntetické hodnoty generované ako súčet hodnôt systematickej funkcie a hodnôt stochastického dvojrežimového LSTAR modelu rezíduí, majú koeficient determinácie v porovnaní s koeficientmi determinácie hodnôt systematickej funkcie vo všetkých prípadoch mierne vyšší, najväčšie zlepšenie je v prípade toku Hron a Morava, v percentuálnom vyjadrení o 6 %, najmenšie pre tok Uh, o 4 %. Výsledky sú porovnateľné s výsledkami aditívnych modelov s trojrežimovými SETAR modelmi rezíduí. Naše praktické skúsenosti potvrdzujú, že hlavne pri modelovaní dlhých prietokových radov na veľkých tokoch je dôležité experimentovať s rôznymi typmi modelov a venovať veľkú pozornosť výberu vhodného typu modelu.

Záver

V príspevku sme sa zamerali na aplikáciu známych a doporučených metód založených na súčasných poznatkoch z hydrológie a analýzy časových radov. Analyzovali sme údaje priemerných mesačných prietokov namerané na troch tokoch situovaných v rôznych oblastiach Slovenska v období 1961 až 2000. Na základe týchto údajov sme vykonali základnú štatistickú analýzu uvedenej množiny hodnôt prietokov, odhadli sme matematickú funkciu trendu a cyklickej zložky. V ďalšom kroku sme zostavili aditívne modely a simulovali prietokové rady ako súčet hodnôt systematickej funkcie a stochastických viacrežimových modelov reziduálnych časových radov typu SETAR a LSTAR.

Existuje veľké množstvo dlhodobých hydrologických údajov pozorovaných v rôznych časových intervaloch: spojito, hodinovo, denne, mesačne, ročne alebo nepravidelne, ktoré pokrývajú rôzne časti priestoru, od izolovaných miest až po rozsiahle súvislé územia (prípadne celú Zem). Hlavná úloha, ktorá stojí pred výskumníkmi je analýza týchto dát, určenie trendov, cyklov, skokových zmien, ktoré môžu súvisieť s určitými prahovými hodnotami. Pri riešení týchto problémov sa nezaobídeme bez vývoja a aplikácie moderných metód, medzi ktoré patrí aj stochastické modelovanie.

Cieľom práce bolo poskytnúť informácie o vývoji prietokového režimu z hľadiska trendov a cyklických javov na troch vybraných tokoch a vytvoriť čo stochastické modely prietokov. Výsledky tejto analýzy pomôžu lepšie pochopiť zmeny v režime prietokov za obdobie 1961 až 2000.

Aplikácia rôznych typov modelov, interpretácia a porovnanie výsledkov modelovania môžu prispieť z hľadiska vodohospodárskeho k lepšiemu využitiu tokov. Odhadnuté modely umožňujú aj predpovedanie prietokov v mesačnom kroku v danej oblasti. Analýzy pozorovaných dát a ich matematické modely môžu byť podkladom pre rozhodovacie procesy pri ekologicko-technických zásahoch na zlepšenie hydrologických pomerov, ako aj pre manažment povodí.

Poďakovanie

Tento článok vznikol s finančnou podporou projektu VEGA č. 1/0891/17 a 1/0710/15. Autori týmto ďakujú za podporu výskumu.

Literatúra

Amendola, A. (2003): Forecasting performance of regime-switching models in hydrological time series, Metodi tatistici e Matematici per le analisi Idrologiche-Roma, 770 – 779.

Arlt, J., Arltová, M. (2007): Ekonomické časové rady. Grada Publishing.

Cipra,T. (1986): Analýza časových řad s aplikacemi v ekonomii. SNTL/Alfa, Praha.

Franses, P.H., van Dijk, D. (2000): Non-linear time series models in empirical finance. Cambridge University Press, Cambridge.

Hamilton, J.D. (1990): Analysis of time series subject to changes in regime. J. of Econometrics, 45, 39 – 70.

Hamilton, J.D. (1994): Time Series Analysis. Princeton University Press.

Hipel, K.W., Mcleod, A.I. (1994): Time Series Modelling of Water Recources and Environmental Systems, Elsevier, Amsterdam, 1013 pp.

Komorníková, M., Szőkeová, D. (2004): ARMA Models in Hydrologic Time Series. Prastan 2004, Slovak Statistical and Demograph. Society, 61 – 69, ISBN 80-88946-36-0

Komorníková, M., Szőkeová, D. (2005): Testovanie spoločných trendov v prietokoch slovenských riek. Forum Statisticum Slovacum Roč. 1, č. 3, 119 – 128, ISSN 1336-7420.

Komorníková, M., Szőkeová, D. (2008): Time series analysis of a regional system of river flows, Zborník príspevkov z medzinárodnej vedeckej konferencie. In 70 rokov SvF STU, Bratislava, ISBN 978-80-227-2979-6.

Komorníková, M., Szolgay, J., Svetlíková, D., Szőkeová, D., Jurčák, S. (2008): A hybrid modeling framework for forecasting monthly reservoir inflows. Journal of Hydrology and Hydromechanics, 56, Vol. 3.

Svetlíková, D., Komorníková, M., Szolgay, J., Kohnová, S. (2008): Hybridný model priemerných mesačných prietokov pre profil Banská Bystrica. Acta Hydrologica Slovaca, Ústav hydrológie SAV, Bratislava.

Koutsoyiannis, D. (2005): Stochastic simulation of hydrosystems. Water Encyclopedia, Wiley, New York.

Mathematica 8.0 Guide to standard extra packages, Wolfram Research (2010).

Pekárová, P. (2003): Dynamika kolísania odtoku svetových a slovenských tokov, Vydavateľstvo Slovenskej akadémie vied.

Salas, J.D. (1993): Analysis and Modeling of Hydrologic Time Series. The McGraw Hill Handbook of Hydrology.

Svetlíková, D., Kohnová, S., Szolgay, J., Komorníková, M., Hlavčová, K. (2010): Využitie hybridných metód v hydrologických predpovediach. Edice monografie, KEY Publishing s.r.o., ISBN 978–80–7418–084–2.

Svetlíková, D. (2006): Analýza časových radov prietokov v oblasti mokrade Kláštorské lúky, 18. Konferencia mladých hydrológov, Bratislava, ISBN 80-88-907-56-X.

Szőkeová, D, Kohnová, S. (2012): SETAR models in the streamflow modelling. Forum statisticum Slovacum 1/2012, 56 – 61 ISSN 1336–7420.

Szőkeová, D. (2007): Long memory process and fractional integration in hydrologic time series. Forum Statisticum Slovacum, Roč. 3, č. 2. 184 – 192. ISSN 1336-7420 (2007)

Szolgay, J. (): Multilinear flood routing using variable travel-time discharge relationships on the Hron river. Journal of Hydrology and Hydromechanics, 52,4, 303 – 317.

Teräsvirta, T. (1994): Specification, estimation, and evaluation of smooth transition models. Journal of American Statistical Association 89, 208–218.

Tong, H., Lim, K.S. (1980): Threshold autoregression, limit cycles and cyclical data. Journal of the Royal Statistical Society, Series B, Methodological, 42, 245 – 292.

Tong, H. (1990): Non-linear time series: A dynamical systems approach. Oxford University Press, Oxford.

Tong, H. (1978): On a threshold model. V: C. H. Chen (ed.), Pattern recognition and Signal Processing. Amsterdam, 101 – 141.

MODELLING OF THE MEAN MONTHLY DISCHARGES USING THE SETAR AND LSTAR REGIME SWITCHING MODELS

The modelling and forecasting of discharge time series is an important issue in the engineering hydrology. The aim of the study was to provide information on the changes of the mean monthly discharges in terms of trends and cyclical phenomena and to create various stochastic models. The analysis was based on the mean monthly discharge time series obtained at the stations Banská Bystrica (Hron), Moravský Ján (Morava) and Lekárovce (Uh) from the period 1961 to 2000. First, the systematical function to describe trends and cycles in the data have been created. Then we have applied the SETAR and LSTAR regime switching model class for modelling systematic function of residual time series and have created additive models for mean monthly discharges. Finally, we tested the profitability of various regime switching models for modelling and simulation the mentioned discharge time series. The results of this analysis will help to better understand changes in the flow regime for the period 1961-2000 as also can be used for building mathematical models for the simulation of future changes and forecasting of mean monthy discharges.

Prof. Ing. Silvia Kohnová, PhD.
Katedra vodného hospodárstva krajiny
Stavebná fakulta Slovenskej technickej univerzity v Bratislave,
Radlinského 11
813 68 Bratislava
Tel.: 0915 135 067
E-mail: silvia.kohnova@stuba.sk

RNDr. Danuše Szőkeová, PhD.
Katedra ekonómie a financií
Fakulta managementu Univerzity Komenského v Bratislave,
Odbojárov 10
820 05 Bratislava 25
Tel.: 0907 387 412
E-mail: dana@szoke.sk