

VLNAS, R., 2023. Detekce trendu v časových řadách hydrometeorologických veličin: vizuální dojmy a statistické poznatky. In: TOLASZ, R., POLCAROVÁ, E. (Eds.), 2023. Sborník příspěvků z První konference projektu PERUN (TA ČR, SS02030040). Praha: ČHMÚ, 152–161, ISBN 978-80-7653-063-8.
<https://doi.org/10.59984/978-80-7653-063-8.20>

Detekce trendu v časových řadách hydrometeorologických veličin: vizuální dojmy a statistické poznatky

(Trend detection in time series of hydrometeorological variables: visual impressions and statistical insights)

Radek Vlnas, Český hydrometeorologický ústav, radek.vlnas@chmi.cz

Abstrakt: Cílem příspěvku je na sadě malých povodí ve Frýdlantském výběžku ukázat: (i) jak různé verze Mann-Kendallova testu detekují existenci trendu v časových řadách srážek, teploty a relativní vlhkosti vzduchu, sněhové vodní hodnoty a průtoku), (ii) jak se opticky liší vizuální vnímání trendu v časové řadě od statistického testu, (iii) nakolik zavádějící je určení trendu metodou nejmenších čtverců oproti statistickému testu. Přestože se jedná o geograficky příbuzná povodí s podobným režimem tvorby odtoku, analýza trendu ukázala překvapivou rozmanitost.

Klíčová slova: trend – autokorelace

Abstract: The aim of the paper is to show on a set of small catchments in the Frýdlant area: (i) how different versions of the Mann-Kendall test detect the existence of a trend in the time series of precipitation, air temperature, relative air humidity, snow water values and discharge), (ii) how the visual perception of a trend in the time series differs from the statistical test, (iii) how misleading the least squares method is in determining a trend compared to the statistical test. Although these are geographically related catchments with similar runoff generation regimes, the trend analysis showed a surprising diversity.

Keywords: trend – autocorrelation

1. Úvod

Mann-Kendallův test je často používán pro analýzu trendu v datových sadách environmentálních systémů (Kendall 1975; Libiseller a Grimvall 2002). Jedná se o test neparametrický, který nepředpokládá normalitu dat a není citlivý vůči odlehkým hodnotám a nelineárním trendům nízkého stupně. Test lze použít v případech, kdy je v datových sadách předpokládána spojitá monotónní rostoucí nebo klesající funkce času a rezidua rozložení s nulovou střední hodnotou. Jinými slovy za předpokladu, že rozptyl rozložení je v čase konstantní.

Von Storch (1995) ukázal, že přítomnost autokorelace v časové řadě zvyšuje pravděpodobnost, že Mann-Kendallův test detekuje statisticky významný trend, tzn. že autokorelace zvyšuje pravděpodobnost chyby I. typu. Funkce omezující vliv autokorelace obsahuje např. balík zyp (Bronaugh et al. 2023) v prostředí R (R Core Team 2020), a to metodou podle Zhang et al. (2000) nebo podle Yue et al. (2002). Porovnání obou metod je jedním z cílů příspěvku.

2. Data

K hodnocení trendu byly použity časové řady měsíčních a ročních srážkových úhrnů, vodní hodnoty sněhu, teploty a relativní vlhkosti vzduchu, měřených a přirozených průtoků pro povodí a mezipovodí vodoměrných stanic Lužická Nisa a Smědé (Vlnas 2023). Průtoky jsou k dispozici za období evidence ovlivnění průtoků, tzn. 1979–2021 (43 let), ostatní veličiny od roku 1971 (51 let). Z důvodu kompatibility výsledků bylo šetření trendu provedeno na stejných řadách, tzn. na období 1980–2021 (42 let).

Trend byl šetřen zejména v mezipovodích vodoměrných stanic 320000m Lužická Nisa-Hrádek n. Nisou a 326000m Smědá-Předlánce. Za účelem ověření výsledků pak také v celých povodích 3200000 Lužická Nisa-Hrádek n. Nisou a 326000 Smědá-Předlánce a v souvisejících povodích horních částí toků 316000 Lužická Nisa-Liberec, 323000 Smědá-Frýdlant v Čechách a v povodích přítoků 319000 Jeřice-Chrastava, 323100 Řasnice-Frýdlant v Čechách a 325000 Bulovský potok-Předlánce.

Mezipovodí Lužické Nisy nad Hrádkem nad Nisou:

320000m Lužická Nisa – Hrádek n. Nisou (320000 – 316000 – 317000 – 319000)

- 316000 Lužická Nisa – Liberec
- 317000 Černá Nisa – Stráž. N. Nisou
- 319000 Jeřice – Chrastava

Mezipovodí Smědé nad Předlánci:

326000m Smědá – Předlánce (326000 – 323000 – 323100 – 325000)

- 323000 Smědá – Frýdlant v Čechách
- 323100 Řasnice – Frýdlant v Čechách
- 325000 Bulovský potok – Předlánce

Specifickou datovou sadu pak představuje soubor účelového pozorování ČGS v zájmové oblasti, sledující průtoky na drobných příhraničních vodotečích v období 1997–2022 (26 let): Oldřichovský p., Václavický p., Vítkovský p., Heřmanický p., Minkovický p., Višňovský p. a Saňský p. Vzhledem ke kvalitě a spojitosti pozorovaných řad lze však za spolehlivé považovat pouze řady průtoků z Václavického, Vítkovského a Višňovského potoka.

Statistická významnost trendu byla testována na hladině významnosti $\alpha = 0,05$ a $\alpha = 0,10$. Do hodnocení byly zahrnuty pouze statisticky významné trendy. Výsledný trend sledovaných veličin je uváděn jako průměrná změna za desetileté období. V grafech jsou okrovou barvou vyznačeny časové řady s trendem z bilančního hlediska negativním, modře s trendem pozitivním. Z bilančního hlediska negativním trendem se rozumí pokles srážek, průtoků, zásob vody ve sněhu a relativní vlhkosti vzduchu a naopak vzestup teploty vzduchu. Slabou čarou je vyznačen statisticky méně významný trend ($\alpha = 0,10$), silně pak statisticky významný trend ($\alpha = 0,05$). Daty je proložena střední hodnota (šedá) a přímka trendu (černá).

3. Metoda

Existence a velikost trendu byla šetřena pomocí neparametrického Mann-Kendallova testu a Theil-Senova odhadu směrnice trendu v prostředí R (R Core Team, 2020).

Nulová hypotéza testu H_0 je, že pozorovaná data $\{X_i, i=1,2,...,n\}$ jsou nezávislá a se shodným rozdělením. Alternativní hypotéza H_1 je, že ve výběru dat je přítomen monotónní trend. Statistika S Mann-Kendallova testu je definována jako:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(X_j - X_i) \quad (1)$$

$$\text{kde} \quad \text{sgn}(\theta) = \begin{cases} 1 & \text{když } \theta > 0 \\ 0 & \text{když } \theta = 0 \\ -1 & \text{když } \theta < 0 \end{cases} \quad (2)$$

Mann (1945) a Kendall (1975) ukázali, že pro $n \geq 8$ je statistika S přibližně normálně rozdělená se střední hodnotou:

$$E(S) = 0 \quad (3)$$

a rozptylem:

$$V(S) = \frac{n(n-1)(2n+5) - \sum_{m=1}^n t_m m(m-1)(2m+5)}{18} \quad (4)$$

kde t_m je počet skupin (dvojic, trojic, atd.) shodných párových hodnot v rozsahu m , přičemž $m = 2$ pro dvojice shodných hodnot, $m = 3$ pro trojice atd. Standardizovaná testová statistika Z je potom:

$$Z = \begin{cases} \frac{S-1}{\sqrt{V(S)}} & \text{když } S > 0 \\ 0 & \text{když } S = 0 \\ \frac{S+1}{\sqrt{V(S)}} & \text{když } S < 0 \end{cases} \quad (5)$$

Testová statistika Z má standardní normální rozdělení s nulovou střední hodnotou a rozptylem rovným jedné.

Výsledná směrnice trendu b je počítána metodou Theil-Sen (Sen, 1968), tedy jako medián směrnic všech párů dat podle rovnice:

$$b = \text{med} \left(\frac{X_j - X_i}{j - i} \right), \quad \forall i < j \quad (6)$$

3.1 Vliv autokorelace

Balík *zyp* (Bronaugh et al. 2023) obsahuje dvě funkce za účelem omezení vlivu autokorelace, a to metodou podle Zhang et al. (2000) nebo podle Yue et al. (2002). Nutno podotknout, že obě mají za cíl omezit pouze vliv autokorelace 1. řádu (AR(1)). Tyto metody se liší přístupem k předběžnému vybělení (pre-whitening) za účelem odstranění AR(1). Do verze programu *zyp* 0.11-1 byla nastavena jako defaultní metoda Yue-Pilon.

V metodě Yue-Pilon je nejprve spočtena směrnice trendu metodou Theil-Sen (1968). Je-li směrnice nenulová a statisticky významná, je z časové řady odstraněn trend daný touto směrnicí a v takto upravené řadě posouzena přítomnost AR(1). Tento postup se označuje jako TFPW (Trend Free Pre-whitening). Výsledná rezidua by měla být nezávislá. Je-li autokorelační koeficient statisticky významný, je směrnice trendu původní řady uplatněna na rezidua

autokorelace a na této řadě je teprve posouzena významnost trendu pomocí Mann-Kendallova testu (Bronaugh et al. 2023).

Při Zhangově metodě se z řady odstraní trend, pokud je významný, a vypočítá se autokorelace. Tento proces pokračuje, dokud rozdíly v odhadech sklonu a AR(1) ve dvou po sobě jdoucích iteracích nejsou menší než 1 %. Na výsledné časové řadě se poté provede Mann-Kendallův test trendu a k výpočtu sklonu trendu se použije metoda Theil-Sen (Bronaugh et al. 2023).

Dokumentace programu *zyp 0.11-1* z dubna 2022 přitom uvádí: V současné době se v hydrologické komunitě *nedoporučuje používat metodu Yue-Pilon* (Yue et al. 2002). Důvodem je skutečnost, že postup TFPW použitý v metodě Yue Pilon vykazuje s rostoucí autokorelací *velmi vysokou chybovost typu I, a proto je míra falešné detekce trendu touto metodou nepřijatelná* (Buerger 2017; Zhang a Zwiers 2004). Verze programu *zyp 0.11-1* byla proto upravena tak, aby standardně používala metody detekce trendů podle Zhang et al. (2000).

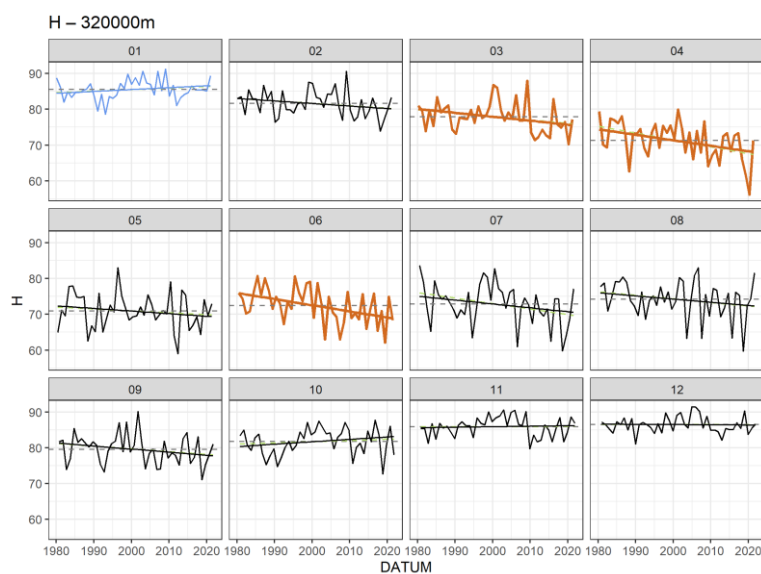
4. Výsledky

Zpracované časové řady bilančních veličin ukazují, že AR(1) detekována testem *acf* se v nich vyskytuje zřídka. Lze ji nalézt např. v časových řadách průměrných ročních průtoků, v měsíčních řadách bilančních veličin se vyskytuje zřídka a v podstatě náhodně. Z popisu metod k řešení problému AR(1) metodami Zhang a Yue-Pilon vyplývá, že není-li v časových řadách autokorelace, není třeba s nimi nijak manipulovat a pro ověření existence trendu se rovnou aplikuje Mann-Kendallův test. Čili ve většině případů (není přítomna autokorelace), by výsledky šetření trendu oběma přístupy měly být shodné. Toto tvrzení ale neplatí. A dále, přijmeme-li tvrzení Buerger (2017) a Zhang a Zwiers (2004) o vysoké míře falešné detekce trendu metodou Yue-Pilon s rostoucí autokorelací, měla by být metoda podle Zhang přísnější, tzn., že trend může být potvrzen podle Yue-Pilon a nepotvrzen podle Zhang, ale ne naopak. Ani toto tvrzení však neplatí bez výjimek.

Obr. 1 uvádí trend vlhkosti vzduchu v mezipovodí Lužické Nisy nad Hrádkem n. N. 320000m (320000 – 316000 – 317000 – 319000) v jednotlivých měsících (1 až 12) podle Yue-Pilon (a) a Zhang (b). AR(1) se podle *acf* vyskytuje pouze v řadách v prosinci. Výsledky šetření trendu se ale podle obou metod liší v lednu a v březnu. Přitom samotný Mann-Kendall test existenci rostoucího trendu v lednu podle Yue-Pilon zamítne (*pvalue* = 0,255). A dále, vizuálně oproti lednu se ale spíše trend vyskytuje v červenci, kde ale potvrzen nebyl.

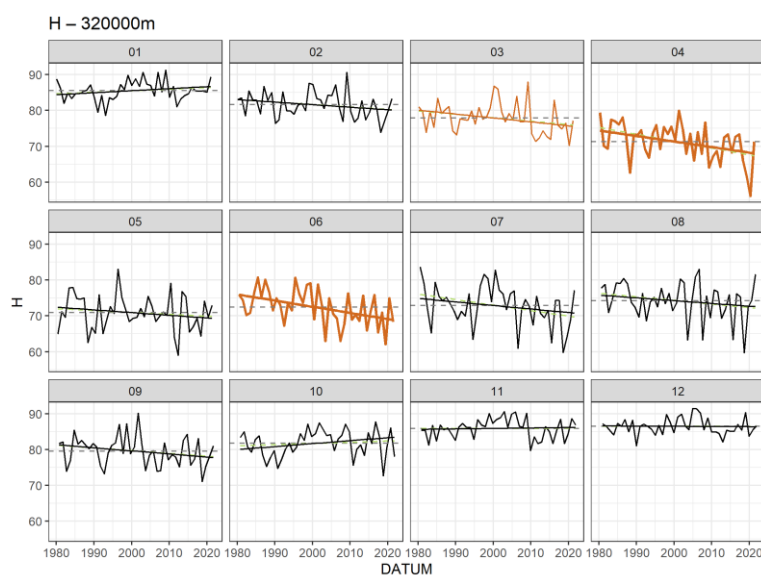
Obr. 2 uvádí trend přirozeného průtoků v mezipovodí Smědé nad Předlánci 326000m. Oproti očekávání byl naopak potvrzen trend v únoru podle Zhang.

Obr. 3 uvádí trend přirozeného průtoků v povodí Bulovského p. v Předláncích 325000 v jednotlivých měsících (1 až 12) podle Yue-Pilon (a) a Zhang (b). Výsledky šetření trendu metodou Zhang jsou podle očekávání, metoda existenci trendu častěji zamítá nebo potvrzuje s menší statistickou významností. Zvláštní je nicméně nepotvrzení trendu v prosinci, zejména např. ve srovnání s trendem v srpnu (*acf* 0,0). Při podrobnějším zkoumání vyjde najevo, že prosinec je jediný měsíc s významnou autokorelací (*acf* 0,4, limit je 0,3). Zdá se tedy, že metoda Zhang odstraňuje autokorelaci tak dlouho, až nakonec zamítne samotný trend (Zhang *pvalue* 0,113, Yue-Pilon *pvalue* 0,0072). Přitom čistý Mann-Kendallův test existenci trendu potvrdí jednoznačně (*tau* 0,266, *pvalue* 0,013). Nabízí se tedy otázka, proč odstraňovat autokorelaci právě a pouze v prosinci s tímto výsledkem. Vyskytuje-li se pouze v prosinci, jedná se pravděpodobně o náhodu.



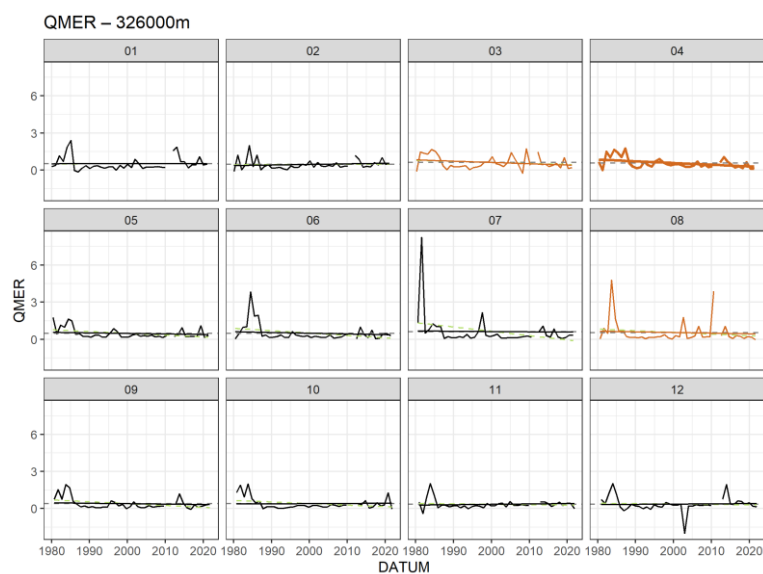
Obr. 1a Trend vlhkosti vzduchu (p.b.) v mezipovodí Lužické Nisy nad Hrádkem n. N. 320000m (320000 – 316000 – 317000 – 319000) v jednotlivých měsících (1 až 12) podle Yue-Pilon.

Fig. 1a Trend in relative air humidity (p.b.) in the subbasin Lužická Nisa above Hrádek n. N. 320000m (320000 – 316000 – 317000 – 319000) in individual months (1 to 12) according to Yue-Pilon.

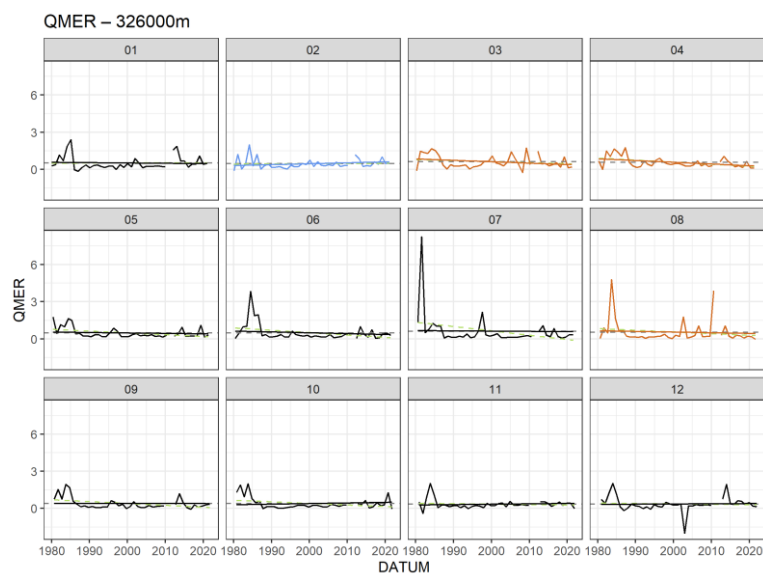


Obr. 1b Trend vlhkosti vzduchu (p.b.) v mezipovodí Lužické Nisy nad Hrádkem n. N. 320000m (320000 – 316000 – 317000 – 319000) v jednotlivých měsících (1 až 12) podle Zhang.

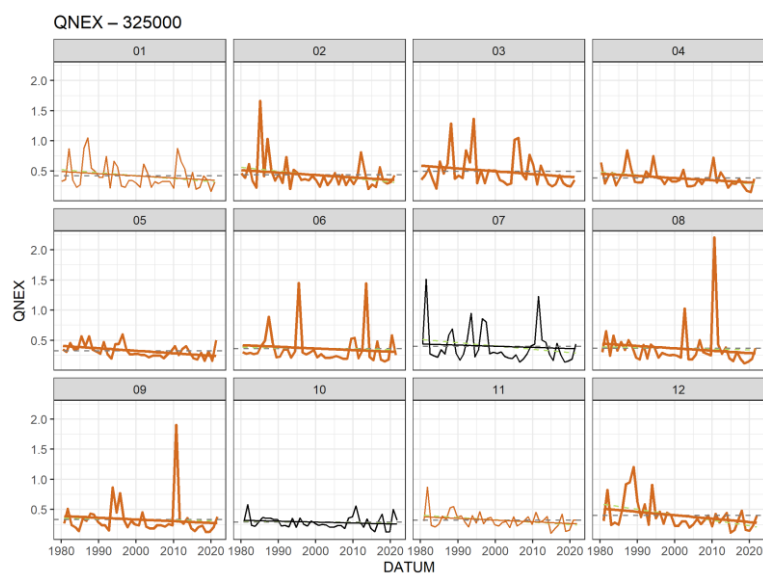
Fig. 1b Trend in relative air humidity (p.b.) in the subbasin Lužická Nisa above Hrádek n. N. 320000m (320000 – 316000 – 317000 – 319000) in individual months (1 to 12) according to Zhang.



Obr. 2a Trend měřeného průtoku ($\text{m}^3 \cdot \text{s}^{-1}$) v mezipovodí Smědé nad Předlánci 326000m (326000 – 323000 – 323100 – 325000) v jednotlivých měsících (1 až 12) podle Yue-Pilon.
 Fig. 2a Trend in natural discharge ($\text{m}^3 \cdot \text{s}^{-1}$) in subbasin Smědá nad Předlánci 326000m (326000 – 323000 – 323100 – 325000) in individual months (1 to 12) according to Yue-Pilon.

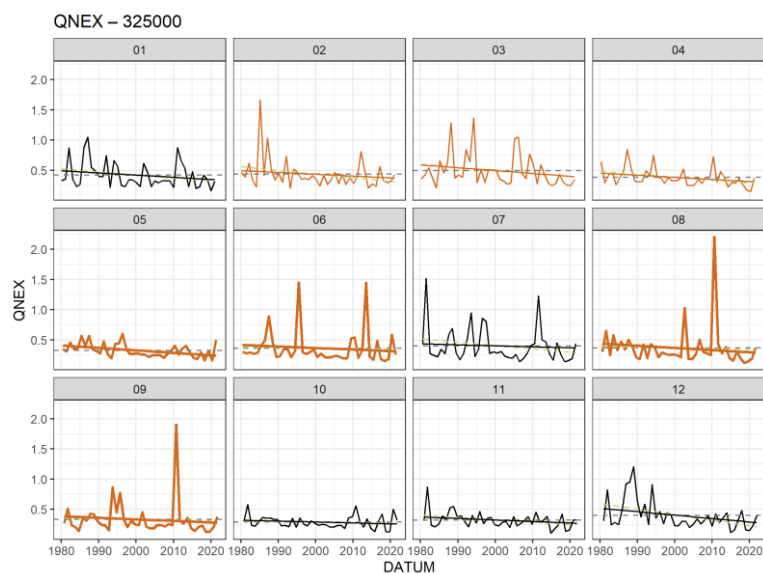


Obr. 2a Trend přirozeného průtoku ($\text{m}^3 \cdot \text{s}^{-1}$) v mezipovodí Smědé nad Předlánci 326000m (326000 – 323000 – 323100 – 325000) v jednotlivých měsících (1 až 12) podle Zhang.
 Fig. 2a Trend in observed discharge ($\text{m}^3 \cdot \text{s}^{-1}$) in subbasin Smědá nad Předlánci 326000m (326000 – 323000 – 323100 – 325000) in individual months (1 to 12) according to Zhang.



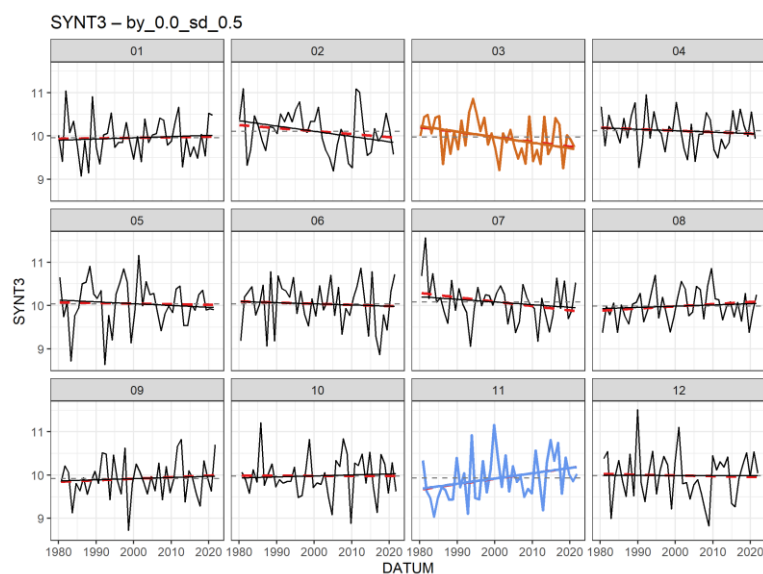
Obr. 3a Trend měřeného průtoku ($\text{m}^3 \cdot \text{s}^{-1}$) v povodí Bulovského p. v Přebělce 325000 v jednotlivých měsících (1 až 12) podle Yue-Pilon.

Fig. 3a Trend in observed discharge ($\text{m}^3 \cdot \text{s}^{-1}$) in basin Bulovský p. in Přebělce 325000 in individual months (1 to 12) according to Yue-Pilon.



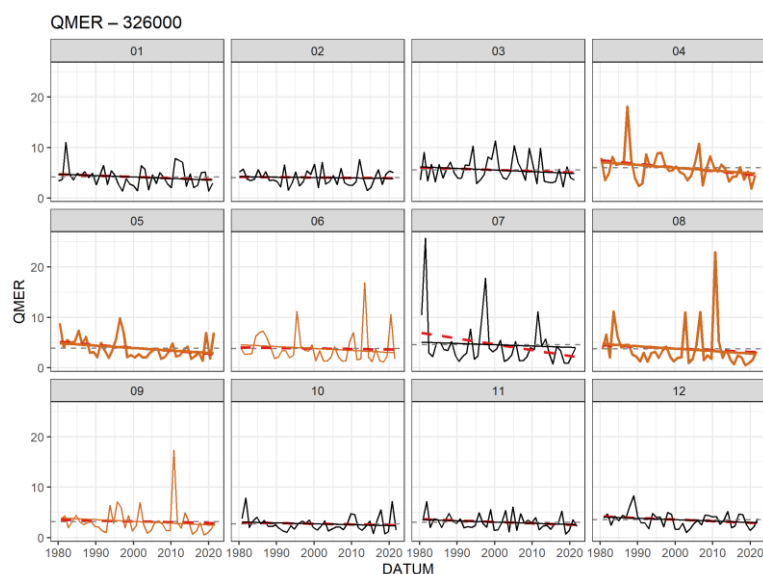
Obr. 3a Trend přirozeného průtoku ($\text{m}^3 \cdot \text{s}^{-1}$) v povodí Bulovského p. v Přebělce 325000 v jednotlivých měsících (1 až 12) podle Zhang.

Fig. 3a Trend in natural discharge ($\text{m}^3 \cdot \text{s}^{-1}$) in basin Bulovský p. in Přebělce 325000 in individual months (1 to 12) according to Zhang.



Obr. 4 Trend v konstantní syntetické řadě se šumem sd 0,5 podle Zhang.

Fig. 4 Trend in constant synthetic serie with the noise sd 0,5 podle Zhang.



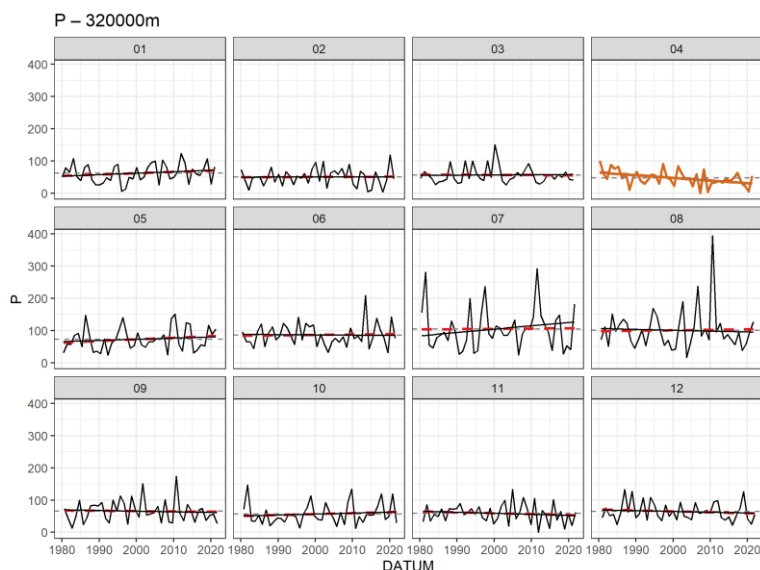
Obr. 5 Trend přirozeného průtoku ($m^3 \cdot s^{-1}$) v povodí Smědé v Předláncích 326000 v jednotlivých měsících (1 až 12) podle Yue-Pilon. Červeně proložení lineární regresí.

Fig. 5 Trend in natural discharge ($m^3 \cdot s^{-1}$) in basin Smědá. in Předlánc 326000 in individual months (1 to 12) according to Yue-Pilon. Linear regression line in red.

Zajímavý pohled na celou problematiku nabízí šetření trendu v syntetických řadách. Vytvoříme-li např. konstantní řadu s hodnotou rovnou 10 a šumem s normálním rozdělením a střední hodnotou rovnou 0 a sd rovnou 0,5, můžeme nalézt trend i v takové řadě (obr. 4). V tomto konkrétním příkladu byl nalezen statisticky významný trend poklesu v březnu a naopak růstu v listopadu.

Dosud se lze často i v odborných studiích setkat s prokládáním časových řad lineární regresí metodou nejmenších čtverců s tím, že se jedná o potvrzení trendu. Tak tomu ale v žádném případě není. Především zde chybí test statistické významnosti. A dále, metoda nejmenších čtverců je z definice silně ovlivněna extrémními hodnotami na začátku a konci řady. Obr. 5 ukazuje, jak lineární regrese v řadách přirozeného průtoku v červenci (řada s extrémním

průtokem na začátku) zvětšuje velikost trendu. Obr. 6 naopak ukazuje, jak lineární regrese v řadách srážkových úhrnů opět v červenci (vyrovnaná řada) naopak velikost trendu snižuje.



Obr. 6 Trend srážek (mm) v mezipovodí Lužické Nisy nad Hrádkem n. N. 320000m (320000 – 316000 – 317000 – 319000) v jednotlivých měsících (1 až 12) podle Zhang. Červeně proložení lineární regresí.

Fig. 6 Trend in precipitation (mm) in the subbasin Lužická Nisa above Hrádek n. N. 320000m (320000 – 316000 – 317000 – 319000) in individual months (1 to 12) according to Zhang. Linear regression line in red.

5. Závěr

Z uvedených zjištění je patrné, že existenci trendu je vždy nutné testovat na statistickou významnost, tzn., zda trend není výsledkem náhody. I tak lze ukázat, že trend může vzniknout právě jako výsledek náhodně generovaných dat. V případě reálně měřených veličin se pak může jednat o vliv náhodných proměnných. Metody odstranění vlivu autokorelace v časových řadách před testováním existence trendu pomocí Mann-Kendallového testu se ukazují jako nespolehlivé.

Poděkování

Tento příspěvek vznikl s podporou TA ČR, projektu SS02030040 „Predikce, hodnocení a výzkum citlivosti vybraných systémů, vlivu sucha a změny klimatu v Česku (PERUN)“.

Literatura

BRONAUGH, D., SCHOENEBERG, A., ZEMAN, L., 2023. zyp: Zhang + Yue-Pilon Trends Package (v. 0.11-1) [cit. 2023-03-22]. <https://CRAN.R-project.org/package=zyp>

BUERGER, G., 2017. On trend detection. *Hydrological Processes*, **31**, 4039-4042.

KENDALL, M. G., 1975. Rank correlation methods. London: Griffin.

LIBISELLER, C., GRIMVALL, A., 2002. Performance of partial Mann Kendall tests for trend detection in the presence of covariates. *Environmetrics*, **13**, 71–84.

MANN, H. B., 1945. Nonparametric tests against trend. *Econometrica*, **13**, 45–259.

R Core Team., 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Dostupné z: <http://www.R-project.org>

SEN, P. K., 1968. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, **63**, 1379–1389.

VLNAS, R., 2023. Zpracování hydrologického modelu pro příhraniční oblasti frýdlantského výběžku a Hrádku nad Nisou v Libereckém kraji. Souhrnná výzkumná zpráva za rok 2022. VÚV TGM.

VON STORCH, V. H., 1995. Misuses of statistical analysis in climate research. In *Analysis of Climate Variability: Applications of Statistical Techniques*, von Storch, H., and Navarra, A. (eds). Berlin: Springer-Verlag, 11–26.

WANG, X. L., SWAIL, V. R., 2001. Changes in extreme wave heights in northern hemisphere oceans and related atmospheric circulation regimes. *Journal of Climate*, **14**: 2204–2221.

YUE, S., PILON, P., PHINNEY, B., CAVADIAS, G., 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes*, **16**, 1807–1829.

ZHANG, X., VINCENT, L. A., HOGG, W. D. AND NIITSOO, A., 2000. Temperature and Precipitation Trends in Canada during the 20th Century. *Atmosphere-Ocean*, **38**(3), 395–429.

ZHANG, X., ZWIERS, F. W., 2004. Comment on "Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test" by Sheng Yue and Chun Yuan Wang. *Water Resources Research*, **40**.