

INFORMAČNÍ BULLETIN



České statistické společnosti

Ročník 28, číslo 3, září 2017

VYUŽITÍ ROZDĚLENÍ LN5 PRO ANALÝZU CHOVÁNÍ PRŮTOKŮ

USING LN5 DISTRIBUTION FOR ANALYSIS OF DISCHARGE CURVES

Ladislav Budík, Marie Budíková

Adresa: Český hydrometeorologický ústav, pobočka Brno, Kroftova 43, 667 16 Brno a Ústav matematiky a statistiky, Přírodovědecká fakulta Masarykovy univerzity, Kotlářská 2, 611 37 Brno

E-mail: budik@chmi.cz, budikova@math.muni.cz

Abstrakt: Pro analýzu chování průtoků lze využít tzv. rozdělení LN5, které je mocninným zobecněním rozdělení LN3. Porovnávání parametrů ročních křivek překročení denních průměrných průtoků umožňuje odhalit určité změny v chování průtoků a jejich reakce na změny přírodních podmínek povodí (např. zánik lesa, změna charakteru zim apod.). Běžným porovnáváním průtoků za různá období se tyto změny odhalit nedaří.

Klíčová slova: rozdělení LN5, roční křivka překročení, přírodní podmínky povodí.

Abstract: We use LN5 distribution for analysis of discharge curves. It is power generalization of LN3 distribution. Comparing curves parameters of average daily discharges allows to see some time changes in behaving of discharges and their reactions on changes of natural conditions (forest destruction, change of winter pass and so on). Using comparing only discharges it is not usually possible to detect this changes.

Keywords: LN5 distribution, cumulative frequency curve, natural conditions in a catchment.

1. Úvod

K popisu průtokových charakteristik vodních toků se téměř 100 let používá tzv. M-denních průtoků a křivky překročení. Křivka překročení je speciální instrument vyvinutý pro potřeby hydrologie. Vzniká tak, že sestupně seřadíme průměrné denní průtoky za pozorované období n let (mohou to být i průměrné průtoky počítané za jinou časovou jednotku, např. hodinu či měsíc). Závislost těchto průtoků na jejich pořadovém čísle (výhodné je přepočítat rozsah pořadových čísel na interval $(0, 1)$) se nazývá (empirická) křivka

překročení. Tato empirická křivka překročení je odhadem neznámé teoretické křivky překročení, tj. odhadem pravděpodobnosti překročení. Používání křivky překročení v hydrologii vzniklo historicky na počátku 20. století a jedná se vlastně o funkci přežití, kde je zaměněna role závisle a nezávisle proměnné. U funkce přežití je na ose x hodnota a na ose y pravděpodobnost přežití. V hydrologii je zobrazení obrácené, na ose x je pravděpodobnost překročení a na ose y je hodnota – průtok (jakoby zpětně zobrazené přežití).

Křivku překročení, tj. přepočítané denní průměrné průtoky seřazené podle velikosti, musíme nahrazovat teoretickými křivkami proto, abychom mohli odhadovat co nejlépe průtoky v nepozorovaných profilech (skládáním teoretických časových řad viz [1], [2] a [3]), což je jeden z důležitých úkolů hydrologů, a nebo abychom mohli odhadnout velikost možných povodní či such. Ale to již je jiná záležitost, než zde chceme probírat. Ze známých rozdělení bylo vybráno LN3 rozdělení.

Teoretické křivky (rozdělení LN3) se historicky prokládaly právě výběrem po $30 \cdot n$ hodnotách průtoků, kde n je počet let pozorování. Křivka překročení je pak dána výběrem 12 hodnot. Bylo to dáno výpočetními možnostmi. Změna mohla přijít až s výrazným rozvojem stolní výpočetní techniky. Prokládáním teoretické křivky tak řídkým výběrem samozřejmě dochází ke zkreslení jejího průběhu, zvláště v místech, kde jsou prudší změny průběhu křivky. V prokládání teoretických křivek tak byly dvě nepřesnosti, jednak z výběru dat po $30 \cdot n$ hodnotách a jednak z rozdílu možností teoretického rozdělení LN3, které se používalo, oproti vlastnostem naměřených dat. Pro aktualizovaná data 1981–2010 se prokládání teoretických křivek provedlo stejně, tj. z 12 hodnot, jenom se použilo rozdělení LN5, vzniklé z normálního rozdělení trochu složitější transformací.

Nechť X je náhodná veličina s normálním rozdělením se střední hodnotou μ a rozptylem $\sigma^2 > 0$. Buď $y_0 \in \mathbf{R}^+$. Rozdělení náhodné veličiny Y , která vzniká transformací

$$Y = e^X + y_0 \quad (1)$$

se nazývá LN3 rozdělení s parametry μ , σ^2 a y_0 .

Rozdělení náhodné veličiny Y , která vzniká transformací

$$Y = a \cdot e^{\text{sign}(X) \cdot X^b} + y_0 \quad (2)$$

nazýváme LN5 rozdělením s parametry μ , σ^2 , y_0 , a a b . V LN3 rozdělení je parametr a i b roven 1, tedy LN3 je zvláštním případem LN5.

Při odhadu parametrů LN3 rozdělení se postupovalo tak, že z naměřených hodnot průtoků v dané vodoměrné stanici seřazených do křivky překročení

se vybralo (v pravidelných intervalech po $30 \cdot n$) jen 12 hodnot a těmi se prokládala teoretická křivka, s kterou se pak dále pracovalo místo empirických dat. Stejný postup byl původně použit i pro odhad parametrů LN5 rozdělení. Nebyl tedy k dispozici plný počet hodnot a bylo nutno při numerické optimalizaci teoretické křivky řešit problém více lokálních minim použité metriky (vážená metoda nejmenších čtverců). Takto se postupovalo vlastně ze setrvačnosti (dříve při ručním zpracování nebylo možné použít všechna data). Tento postup se volil i tehdy, kdy již počítačová technika umožňovala použití všech naměřených hodnot a tím i mnohem jednoznačnější výsledek. (Jsou určité věrohodnější odhady parametrů teoretické křivky o třech či dokonce pěti parametrech, prokládáme-li ji např. 3650 hodnotami deseti let pozorování, než prokládáme-li pouze vybranými dvanácti hodnotami, přičemž jako nejvyšší použitá pro určení parametrů je brána až 36. nejvyšší hodnota.) Použití všech dat k pokusu o prokládání seřazených hodnot průtoků v jednotlivých hydrologických rocích (od 1. 11. do 31. 10.) teoretickými křivkami překročení LN5 je postup, který dosud nikdo ve větším měřítku nepoužil. Výsledkem jsou odhady pěti parametrů pro každý rok pozorování. Jejich průběh či změny budou dále diskutovány.

2. Popis metodiky

Aby bylo možné získané parametry pro různé profily a roky mezi sebou porovnávat, byly průtoky pro každý jednotlivý rok normovány na průměr 1. Hodnota byla odvozena z normování normálního rozdělení, které má střední hodnotu rovnou 0. Rozdělení průtoků jsou zhruba exponenciální transformací normálního rozdělení a e^0 je rovno 1. Bez tohoto normování se nedá výsledků umožňujících následující srovnání dosáhnout, protože průměr ovlivňuje velikost jednotlivých parametrů LN rozdělení. Vynásobíme-li LN3 rozdělení ve vztahu (1) konstantou a budeme-li je chtít znovu proložit LN3 rozdělením, dostaneme odlišné parametry. Podobně, i když o něco lépe, dopadneme u rozdělení LN5, kde tomuto postupu bude překážet jenom parametr y_0 v rovnici (2). Bude-li roven 0, pak tento problém s normováním nenastane, protože normování zastane parametr a . Obecně však je parametr y_0 různý od 0. Tímto normováním docílíme toho, že povodí se stejnými hydrologickými vlastnostmi, a tedy i podobnými křivkami překročení (jako podobné trojúhelníky), budou mít všech 5 parametrů stejných. Stejně by to dopadlo i při normování a použití LN3 rozdělení, jen s tím rozdílem, že by byla podstatně horší rozlišovací schopnost parametrů vůči vlastnostem povodí. Tato úprava je to, co hydrolog potřebuje, aby mohl s vyšší pravděpodobností říct, že dvě různá povodí s různým průměrným dlouhodobým odtokem (výběro-

vým průměrem), mají stejné parametry odtoku a tedy i podobné vlastnosti a mohl z toho vyvozovat, jaké jsou průtoky tam, kde nemá pozorování.

Po výše zmíněném normování budeme označovat ve vzorci (2) pro transformaci normálního rozdělení na LN5 rozdělení hodnotu posunu y_0 jako $K100\%$. Písmenem K označujeme v hydrologii normované průtoky a hodnota 100% označuje, že hodnota je překročena ve 100 % případů, tedy je to teoreticky nejnižší dosažitelná hodnota vyplývající z teoretického rozdělení s danými parametry.

2.1. Význam parametrů a , b , μ , σ a $K100\%$ po normování průtoků

Význam a vzájemné závislosti jednotlivých parametrů jsou složitější než u rozdělení LN3, ale do jisté míry podobné a lze z nich vycházet. Zcela logicky jsou parametry v určitém rozsahu každý rok jiné. Pokud jsou podmínky stacionární, měly by se parametry jednotlivých let pohybovat v určitém rozsahu a z něj nevybočovat. Pokud se rozsah parametrů začne měnit, mění se též podmínky, za nichž průtok vzniká. Pokusíme se popsat význam parametrů pro interpretaci křivek.

Parametr $K100\%$ pouze posouvá o svoji hodnotu křivku překročení podél osy y (průtoky) a vlastně určuje minimální teoreticky dosažitelnou hodnotu (průtok). Jeho význam je podobný jako parametr y_0 pro LN3 rozdělení. Největší relativní vliv má u nízkých hodnot křivky překročení. Jeho záporná hodnota u roční křivky překročení však nemusí jednoznačně znamenat skutečné vyschnutí toku (máme pouze výběrové hodnoty).

Parametry σ a b jsou vzájemně volně závislé. Na jejich vztah má však určitý vliv i parametr μ . Proto mezi σ a b nemůžeme mluvit o funkční závislosti. V zásadě však přibližně platí, že s růstem parametru b nad hodnotu 1 parametr σ pod hodnotu 1 klesá a naopak. To je dáno tím, že hydrologické poměry jsou poměrně stabilní (geologie, pedologie, rostlinstvo, srážky) a příroda má tedy k dispozici jen určité množství vody, které rozděluje podobným způsobem – proto se za stacionárního prostředí hodnoty všech parametrů pohybují v určitém pásu hodnot a změna tohoto pásu znamená změnu podmínek prostředí. Pokud je parametr b roven 1, máme rozdělení LN3, neboť funkce parametru μ v exponentu rovnice (2) splyne s funkcí parametru a a nové $\mu' = \mu + \ln(a)$, což vyplývá z rovnice (2) po dosazení $b = 1$. Pokud $b > 1$, narůstají vysoké hodnoty rozdělení na úkor nízkých a křivka překročení se více prohýbá (oproti průběhu LN3 se blíží „stoličce“), pro $b < 1$ je tomu naopak. Do těchto změn oproti LN3 rozdělení ještě promlouvá parametr μ . Pokud je roven 0, je transformované normální rozdělení umístěno symetricky

na střed intervalu $(0, 1)$ a transformované rozdělení má viditelné „ohnuté chvosty“ jak do hodnot vysokých, tak do hodnot blízkých $K100\%$. Posun μ ke kladným hodnotám potlačuje „chvost“ do nízkých hodnot, posun μ do hodnot záporných potlačuje „chvost“ do vysokých hodnot. Přes tyto změny ohybů funkce překročení promlouvá parametr μ do vzájemného vztahu parametrů a a b (i při jeho změnách se mění průběh „chvostů“, i když trochu jiným způsobem), a logicky je to i obráceně. Tyto tři parametry b , μ a σ svou vzájemnou vhodnou kombinací docílí žádaný tvar křivky překročení, je však požadované křivce jen podobná (jako dva trojúhelníky). Žádanou „velikost“ křivky překročení potom dosáhneme vhodnou velikostí parametru a , s jehož pomocí a s pomocí parametru $K100\%$ tak dosáhneme žádaného průměru normovaného výběru rovného 1. Toto je také důvod, proč velikost parametru a reaguje na změny jakéhokoli parametru. Ze změny hodnoty parametru a ovšem nelze obecně usoudit, o jakou změnu se jedná.

Parametr b může při hodnotě odlišné od 1 způsobit dvouvrcholovou funkci hustoty rozdělení LN5. Uplatnění této vlastnosti souvisí s velikostí parametru μ . Pokud je absolutní hodnota μ dostatečně velká, pak ke dvouvrcholovosti prakticky dojít nemůže. Střed normálního rozdělení (kde jsou hodnoty v absolutní hodnotě blízké nule a tedy při umocnění na exponent různý od jedné způsobí průhyb křivky překročení), je odsunut až na okraj (pravý či levý) křivky překročení. Významné zvýšení hustoty hodnot určité velikosti potom viditelně nezpůsobí. Faktem je, že mnohé roční křivky překročení průtoků dvouvrcholovou hustotu mají. Rozdělením LN5 tyto případy popsat lze (např. případy, kdy po vlhkém období následuje období suché současně se začátkem vegetace rostlin a tím zvýšeným výparem a následně zrychleným poklesem průtoků).

Jako metrika pro optimalizaci výsledku byla použita vážená metoda nejmenších čtverců. Zde je nutno dodat, že metoda nejmenších čtverců (nikoliv vážená) byla odvozena pro normálně rozložená data a pro ně by měla fungovat bez systematických ovlivnění výsledků. Při použití na data rozložená podle LN3 či LN5 tato metoda silně preferuje přesnost u vysokých hodnot. Z toho vyplývá, že velmi silně narůstá relativní chyba proložení směrem k malým hodnotám. Tato skutečnost neodpovídá rozložení chyb při měření průtoků, kde je velmi přibližně stejná relativní chyba dat podél celého pásma měřených hodnot (odhlédneme-li od toho, že chyba není u konkrétní hodnoty rozložena směrem nahoru a dolů symetricky). Proto jako metrika lépe vyhovuje minimalizace sumy druhých mocnin rozdílů mezi hodnotou naměřenou a teoretickou dělených hodnotou naměřenou.

2.2. Výběr řad pro analýzu

Aby výše uvedený způsob prokládání křivky LN5 popsal změnami parametrů změny v hydrologickém chování průtoků, je nutné mít k dispozici řady pokud možno bez vlivu manipulace průtoků prováděné člověkem – tzv. ovlivněné průtoky. U způsobu jejich stanovení se musíme na chvíli zastavit. I takové průtoky mají totiž do průtoků přirozených stále velmi daleko (ukáže to i rozbor dosažených výsledků). Ovlivněné řady denních průměrných průtoků jsou opraveny o všechny známé lidské manipulace s průtoky (nebo je neobsahují). Jsou to např. manipulace na přehradách, odběry povrchových i podzemních vod, vypouštění odpadních vod nebo převody vod mezi toky (povodími). Do těchto oprav nejsou zahrnuty vlivy obtížně kvantifikovatelné. Jsou to např. změny ve využívání krajiny (místo lesa pole – zde si musíme uvědomit, že na našem území by byl bez lidského vlivu prakticky souvislý lesní porost), odvodnění pozemků, do roku 1992 nejsou započítávány do ovlivnění odtoky dešťových srážek ze zpevněných ploch, zhoršení kvality lesních porostů způsobené exhalacemi atd.

Při prvních analýzách provedených výše popsáním způsobem bylo často potřeba více pokusů optimalizace parametrů s různými startovacími hodnotami, aby se podařilo najít parametry, které by poskytly absolutní minimum váženého součtu čtverců odchylek. Optimalizace LN5 rozdělení může mít až tři různá lokální minima a záleží na startovacích hodnotách, jestli se trefíme do minima absolutního. Také bylo nutno se naučit výsledky interpretovat – s tím doposud nebyla žádná zkušenost. Došlo se k překvapivým závěrům, jak bude uvedeno dále.

Protože výpočty jsou dosti pracné, bylo prozatím spočteno jen několik stanic, které jsou bez antropogenního ovlivnění ve smyslu výše popsaném nebo jen s malým podílem tohoto ovlivnění. Jedná se zatím o 12 povodí.

Prakticky bez nebo s nepatrným antropogenním ovlivněním ve výše uvedeném smyslu (většinou les, ale s nepříliš původní skladbou nebo pastviny, málo pole) jsou povodí: Kychová, Kychovka; Zdechová, Zdechovka; Blanický Mlýn, Blanice; Modrava, Modravský potok; Bílý potok, Smědá.

S malým ovlivněním: Borovnice, Svatka; Dolní Bory, Oslava; Rejštejn, Otava; Jeseník, Bělá; Koryčany nad přehradní nádrží, Kyjovka.

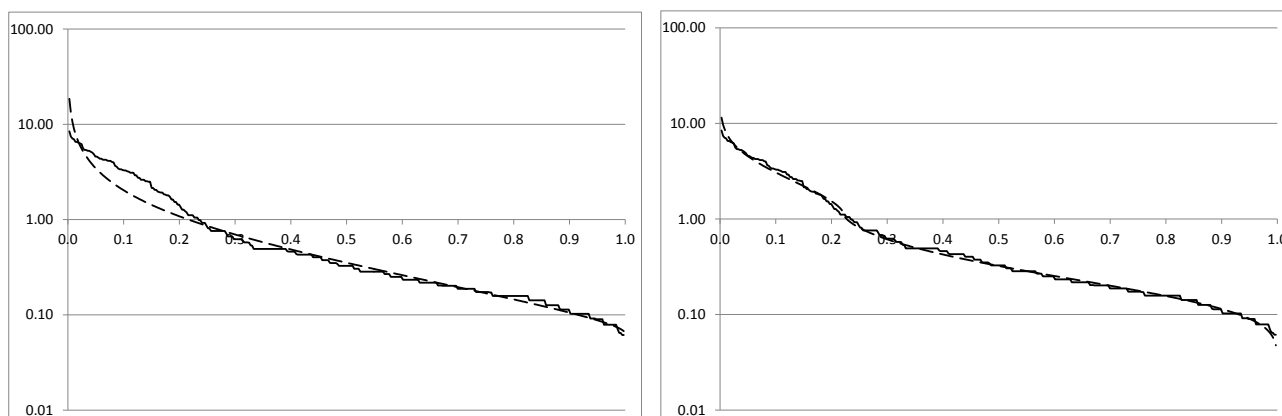
S větším podílem ovlivnění: Punkva, Skální Mlýn; Rychmanov, Litava.

Doby pozorování nejsou stejné. S tím je nutno se smířit. Pro každý hydrologický rok byly stanoveny parametry Q_a (roční průměr), a , b , μ , σ a $K100\%$ (ve vztahu (2) je to y_0). Výsledky byly porovnány mezi sebou a též s historickým vývojem povodí v období pozorování.

3. Popis výsledků a jejich analýzy

Velkým překvapením bylo, že chceme-li co nejlépe popsat data proloženou křivkou, má tato křivka velmi často parametr $b < 1$, což znamená, že hustota rozdělení je často dvouvrcholová. Zároveň s tím se ale vyskytují roky, kde parametr b může být výrazně větší než 1. Pokud máme křivku překročení za více let (např. 10), pak výsledný parametr $b \geq 1$ a dojde k vyhlazení detailů.

Vysvětlení pro dvouvrcholovou hustotu křivky překročení je relativně jednoduché: v průběhu roku dojde k víceméně rychlé proměně období se srážkami vyššími než je výpar za období se srážkami nižšími než výpar. Podzemní vody zásobující tok v období sucha pak přestávají být dotovány srážkovou vodou a naopak z nich vegetace, pokud do nich dosáhne, začne vodu odčerpávat. K tomu může dojít ve vrcholném jaru s olistěním vegetace nebo při změně vydatnějšího srážkového období za sušší s vyššími teplotami prakticky kdykoli v průběhu vegetace. Proto se zhoupnutí do dvou vrcholů hustoty (nebo zlom v křivce překročení) může vyskytnout na různých místech roční křivky překročení. Ve výrazně vlhkých letech se toto zhoupnutí obvykle nevyskytuje. Na obrázcích jsou příklady takových křivek s proložením LN3 (vlevo) a LN5 (vpravo).



3.1. Vztah parametrů LN5 a ovlivnění

Tzv. ovlivnění je přítok do toku, který vzniká lidskou činností. Mohou to být odběry vody (vodovody, závlahy apod.), pak je hodnota přítoku záporná. Nebo to může být vypouštění (čistírny odpadních vod, průmyslové vody apod.), potom je hodnota přítoku kladná. Mohou to být i manipulace na vodních nádržích, potom je hodnota buď kladná (odpouštění většího množství vody než je celkový přítok) nebo záporná (nádrž vodu nadržuje – hromadí).

Ze zkušenosti vyplývá, že většina ovlivnění má křivky překročení blízké normálnímu rozdělení a koeficienty korelace ovlivnění s přirozenými průtoky

se nachází v intervalu $(-0,3; 0,3)$. Z normality ovlivnění a slabé korelace tedy usuzujeme, že ovlivnění působí především na dlouhodobý průměr Q_a , $K100\%$. Parametr Q_a je ovlivněn plně, protože průměrné hodnoty skládaných řad se sčítají bez ohledu na koeficient korelace. Parametr $K100\%$ může být změněn ještě více než průměr, protože jeho výsledná hodnota je závislá na koeficientu korelace a kolísání hodnot v řadě ovlivnění. Zde nutno vzít v úvahu, že nejmenší hodnota ovlivnění může být i záporná (to je způsobeno odběrem). Potom výsledný parametr může být i 0. Změna parametru μ je dána tím, že rozdělení ovlivnění je blízké normálnímu, a tudíž téměř symetrické. Po složení normálního rozdělení s LN5 rozdělením (což je exponenciálně mocninná transformace normálního rozdělení) dochází ve výsledném LN5 rozdělení podle poměru velikostí dat v jednotlivých skládaných rozděleních k posunu parametru μ směrem k nule, a to zdola i shora. Parametr σ je složením řad ovlivněn méně, protože koeficient korelace je obvykle blízký nule. Parametr b je složením obou řad zasažen nejméně, protože jeho velikost je závislá především na průběhu nejvyšších průtoků (často za extrémních průtoků), kde vliv ovlivnění bývá zanedbatelný (ne vždy – nádrže).

Výběrem stanice pro prováděné analýzy se tento stav dá ovlivnit jen málo, protože většina stanic s dlouhou dobou pozorování je ovlivněna provozem nádrží a velkých uživatelů vody a zpětně polohu stanice pochopitelně změnit nelze. Těch několik výjimek, kde ovlivnění popsaného typu není třeba uvažovat, zde rozebereme.

3.2. Hydrologický význam změn jednotlivých parametrů

Dlouhodobý průměr Q_a není třeba příliš komentovat. Jeho změna obvykle souvisí s nárůstem či poklesem srážek nebo/a poklesem či nárůstem výparu.

Parametr a převádí tvar funkce překročení získaný pomocí parametrů b , μ , σ a $K100\%$ tak, abychom získali potřebné Q_a , tedy aritmetický průměr seřazených průtoků. Je to podobný proces jako u podobných trojúhelníků. Charakterizuje tedy, jak daleko je křivka překročení získaná ze čtyř zmíněných parametrů od výsledné potřebné velikosti průměru (jako u podobných trojúhelníků). Jeho velikost je ovlivněna změnou jakéhokoli ze čtyř zbývajících parametrů. Z toho plyne, že je citlivý na jakékoli změny v průběhu křivky překročení, a to jak jejího tvaru, tak i změny průměru beze změny tvaru. Z jeho změny tedy nelze určit, jaká konkrétně se stala změna v průběhu křivky překročení (změna fyzicko-geografických podmínek či schopnosti vegetace ovlivnit odtok, velké sucho či intenzivní srážky), ale víme, že nějaká změna se děje. Jeho změna je ovšem závislá i na rozdělení srážek během roku (např. výskyt či neexistence prudkého lijáku, který může změnit para-

metr b). Určité kolísání velikosti parametru v nějakém stabilním rozmezí mezi jednotlivými roky v průběhu pozorování je normální a je v pořádku. Problém nastává při systematické změně jeho velikosti nebo rozsahu jeho kolísání. To potom znamená, že došlo ke změně podmínek pro vytváření odtoku. Prudkou změnu parametru a může vyvolat povodňová epizoda či extrémní sucho, ale pak se po jejím odeznění parametry vrátí k původnímu rozsahu kolísání.

Parametr b je parametrem, který obecně vzato, nejvíce charakterizuje podíl rychlého odtoku (většinou povrchového – ne podzemního odtoku, který mívá zpoždění za povrchovým). Čím vyšší parametr b , tím větší je podíl tohoto rychlého odtoku. Souvisí tedy s intenzitou srážek v daném místě a se schopností celého ekosystému povodí transformovat srážkovou epizodu na dlouhodobější mírný odtok. Souvisí samozřejmě i s kvalitou půd a vegetace na nich a jejich prostřednictvím se způsobem obhospodařování krajiny. Povodňová epizoda v daném roce může parametr b prudce zvýšit, ale po jejím odeznění zase klesne. Pro vyhodnocení jeho změn tedy platí totéž, co v předchozím případě. Kolísá v určitém rozmezí, může prudce vyskočit při povodňové situaci, ale až systematická změna hodnot nebo rozsahu kolísání znamená trvalou změnu podmínek v povodí.

U dalších parametrů je vysvětlení komplikovanější. Parametr μ zvyšuje o konstantu hodnotu normálního rozdělení, které transformujeme na rozdělení LN5, a tím mění tvar rozdělení. Při stejné střední hodnotě snižuje méně minimální průtoky než zvětšení parametru σ . Zvětšuje poměr mezi nejvyšší a nejnižší hodnotou rozdělení. Ale méně než u parametru σ , kterým data normálního standardizovaného rozdělení násobíme. Tím se rozptyl ovlivní mnohem více – zvětší se více poměr mezi nejvyšší a nejnižší hodnotou transformovaného rozdělení. Obecně pro parametry rozdělení LN5 křivek překročení průtoků platí, že existuje volná vazba mezi dvojicemi parametrů a a μ a parametrů b a σ . Pokud parametr μ vzroste, parametr a to kompenzuje poklesem, neboť aritmetický průměr (tj. průměrný průtok) se příliš nemění. Podobná vazba je mezi parametry b a σ . Žádná z uvedených vazeb není matematickou funkcí druhého parametru, je volnější. Hodnoty parametru $K100\%$ jsou relativně nezávislé na ostatních parametrech a protože jeho hodnota je obecně malá, zřetelně ovlivňuje pouze velikost nejnižších hodnot průtoků křivky překročení.

3.3. Výsledky analýz

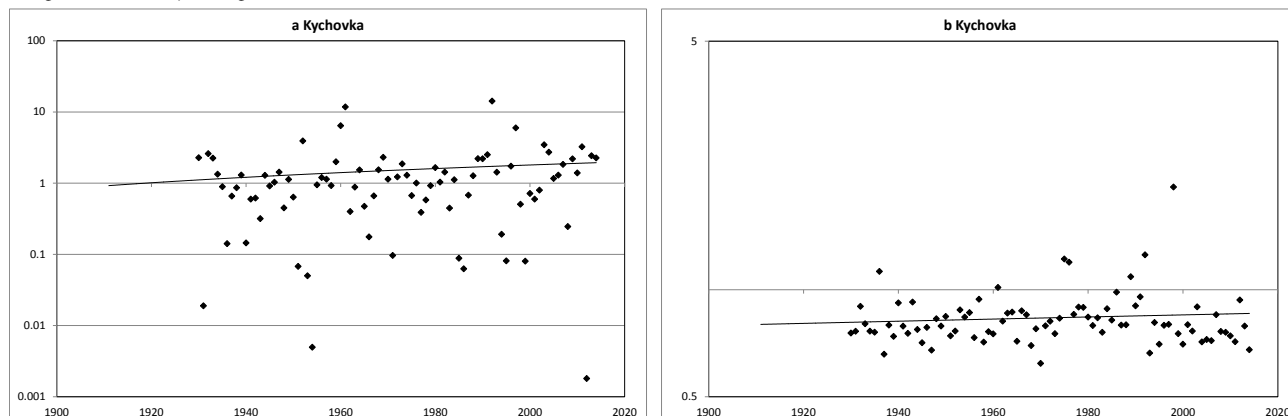
Byly poněkud překvapivé. A to proto, že výzkumníci, kteří se zabývají modelováním důsledků změn v krajině, prohlašují, co by se při změnách v ní mělo dít, ale hydrologové to nemohou potvrdit. Dlouhodobý průměr odtoku

se v podstatě nemění (kolísání od 0 % do 4 %), ve srážkách je to obdobné. Koeficient variace denních průtoků se také nijak významně nemění a ostatní parametry jako koeficient asymetrie a vyšší momenty jsou už silně závislé na momentálním výskytu významné průtokové události, které jsou obecně pozorovány po delších obdobích (významná sucha, povodně).

Zkoumané stanice byly vybrány především tak, aby měly co nejdelší řadu pozorování, zanedbatelné nebo žádné antropogenní ovlivnění (ve smyslu popsaném výše) a pokud možno různý a známý typ vývoje krajiny. Jejich názvy jsou uvedeny výše. Na obrázcích jsou tečkové diagramy vývoje parametrů a (vlevo) a b (vpravo) v čase po hydrologických rocích (od 1. 11. do 31. 10.). U profilů s jejich významnější změnou bude uvedena poznámka v textu. Všeobecně nelze říci, že by se parametry v čase vyvíjely shodně, naopak je zde značná proměnlivost. Nicméně se po delším zkoumání zdá, že lze najít systém ve vysvětlení rozdílností v chování jednotlivých povodí.

Pro vysvětlení, jakým způsobem byla data vyhodnocována, uvedeme podrobný popis chování dvou zcela rozdílných povodí – Kychovku v Kychové a Smědou v Bílém potoce.

Kychová, Kychovka



Jde o experimentální povodí v blízkosti Vsetína s vysokým zastoupením lesa, jež se nemění. Téměř zde není orná půda.

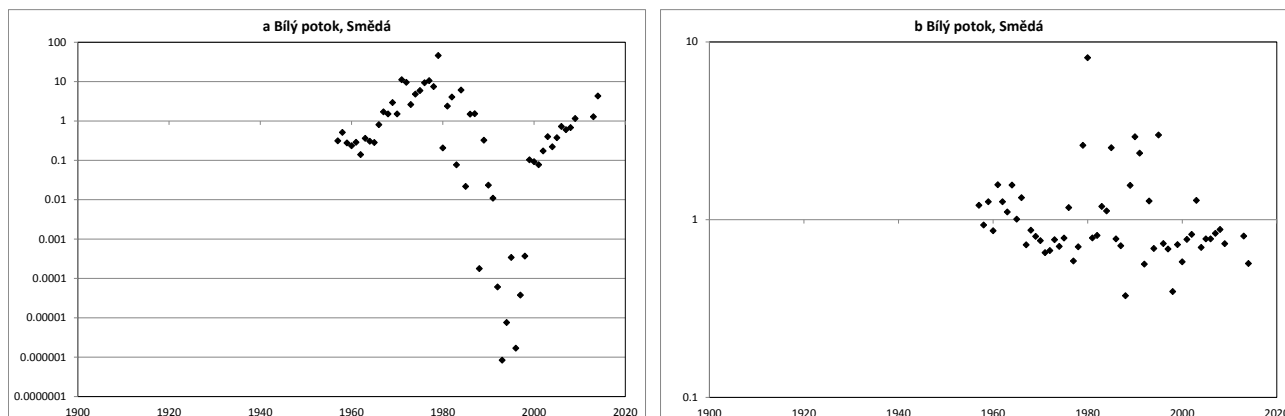
Dlouhodobý průměrný průtok lehce klesá a snižuje rozptyl – to může být způsobeno zkracováním zimního období s hromaděním sněhu a následně zvyšováním výparu v delší vegetační době.

Parametr $K100\%$ se zdá lehce narůstat a zvyšovat rozptyl, což může souviset buď se zvyšováním zásob podzemní vody, nebo s možností uvolnit i zásoby, které dříve uvolnit nešly (např. odvodnění pozemků).

Parametry μ a σ normálního rozdělení lehce klesly, b se téměř nemění. Logicky by tedy mělo vzrůst $K100\%$, což se děje. Zároveň s tím roste parametr a a medián. To tedy znamená, že obecně klesají nejvyšší roční průtoky, máme-li dodržet malinko klesající Q_a . Parametr b vykazuje slabou změnu

v 70. a 80. letech k vyšším hodnotám (zřetelnější po logaritmování hodnot). Zřejmě se jedná o narušení kvality lesních porostů oxidem siřičitým. Ale tato změna je zde jen malá a v 90. letech se, zřejmě díky snížení exhalací oxidu siřičitého, hodnoty vrací na původní úroveň. Vyskytuje se i několik hodnot s výrazně zvýšeným parametrem b v letech s extrémními srážkami.

Bílý potok, Smědá



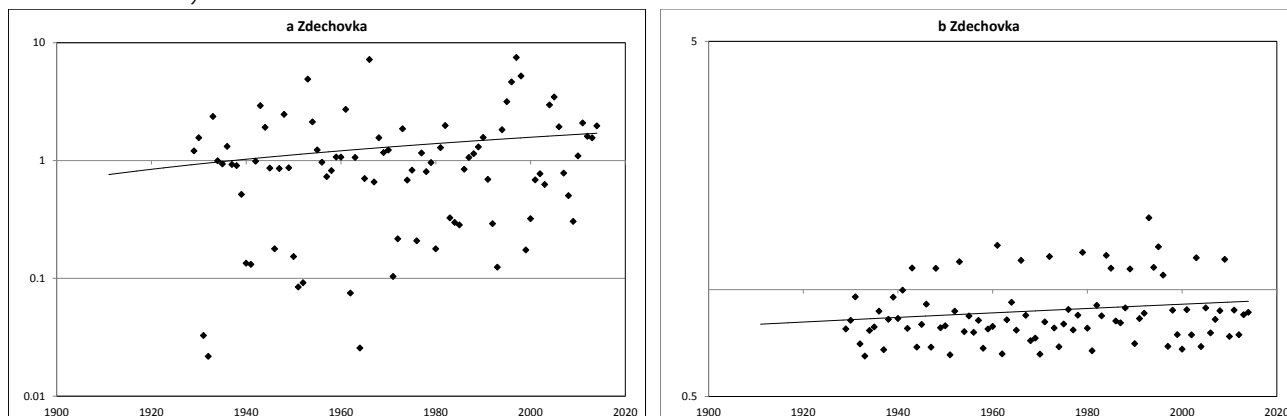
Jedná se o povodí v Jizerských horách se stanicí na jejich úpatí a s minimálním ovlivněním. Mělo by být téměř celé zalesněno kromě úplného závěru sledovaného období. Problémem zde je, že povodí v 70. a 80. letech 20. století bylo postiženo kalamitou způsobenou emisemi oxidu siřičitého. Bylo téměř zcela odlesněno kromě severní části, kde byl bukový les. Ale i ten silně utrpěl na životaschopnosti. Po roce 1995, kdy bylo dokončeno odsíření elektráren, se sem začal les pomalu vracet.

Parametry rozdělení LN5 a , b , σ a μ tuto situaci velmi věrně kopírují. Lze říci, že i medián. Z dat vyplývá, že charakteristiky odtoku zatím nejsou úplně stejné jako v 50. a 60. letech minulého století. To může být způsobeno tím, že zatím chybí dostatečná vrstva organického odpadu, která se v průběhu neexistence lesa rozložila a snížila kapacity půd. Důvod může být i v delší vegetační době a kratších zimách s menším množstvím sněhu. Nejpravděpodobnější však je, že množství kyseliny sírové dopadající na povodí je stále ještě vyšší než ekosystémy dlouhodobě unesou.

Průtoková data jsou, bohužel, k dispozici až od poloviny 50. let 20. století. Změnu odezvy na systematicky měnící se podmínky je vidět velmi pěkně na průběhu parametru a po logaritmování. Od 1. poloviny 60. let začíná parametr a pomalu narůstat. Nárůst vrcholí na konci 70. let a potom prudce padá až do roku 1995. Od té doby roste. Průběh parametru b vykazuje nejvyšší hodnoty v letech 1979–1995, což odpovídá období největšího poškození lesních ekosystémů, kdy les zde prakticky přestal existovat. (Pozn.: Nelze si nevšimnout, že u parametru b se v tomto období objevuje stejná „dvojkolejnost“ výskytu hodnoty parametru jako v jiných povodích, kde je málo lesa,

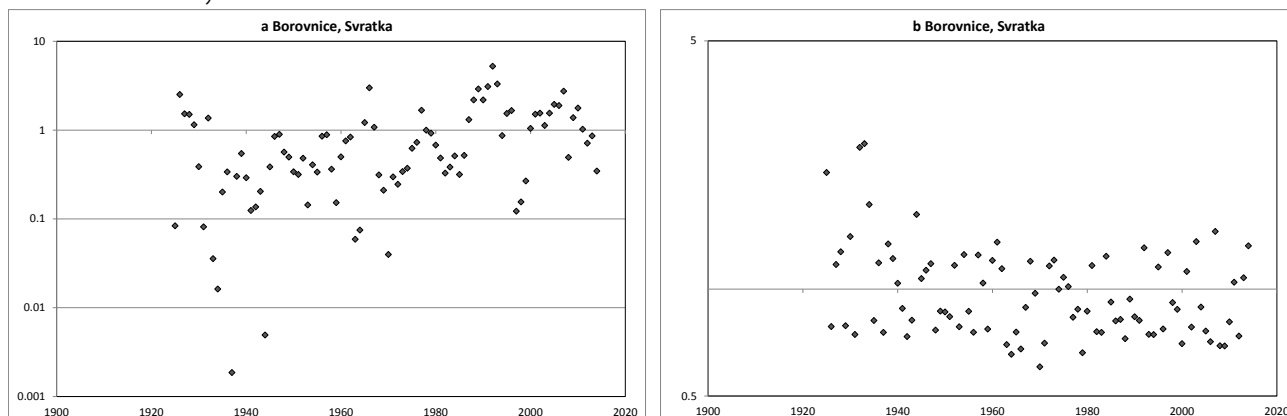
viz např. níže povodí Zdechovky. Pouze hodnoty ve vyšším pásu jsou vyšší než u jiných, níže položených povodí s nižšími srážkami. Z porovnání charakteristik ročních křivek překročení před kalamitou v 80. letech a nyní by ale vyplývalo, že les zde není dosud stabilní, a jak stromy mohutní, blíží se znovu rozpad lesa.)

Zdechová, Zdechovka



Jedná se o experimentální povodí sousedící s Kychovkou, ale s nízkým zastoupením lesa. V průběhu parametru b je vidět jasná dvojkolejnost a vyšší zastoupení hodnot větších než 1. To zřejmě souvisí s častějšími výskyty průtokových vln se strmějším vzestupem i sestupem průtoků asi způsobeném povrchovým odtokem. Navíc v průběhu času se zvětšuje vzdálenost obou pruhů v tečkovém diagramu. Nejvyšší hodnoty b jsou po roce 1960. Medián i $K100\%$ v průběhu doby rostou. Roční Q_a lehce klesá. Je vidět, že povodí má nižší schopnost transformovat vysoké srážky na nižší průtoky než povodí Kychovky a tato schopnost se v průběhu času zhoršuje. Zvláště na změnách parametru b je vidět vyšší podíl rychlého odtoku při silnějších srážkách (oproti Kychovce je zřetelně „dvojkoľejný“). Rostoucí medián může mít stejnou souvislost jako u Kychovky.

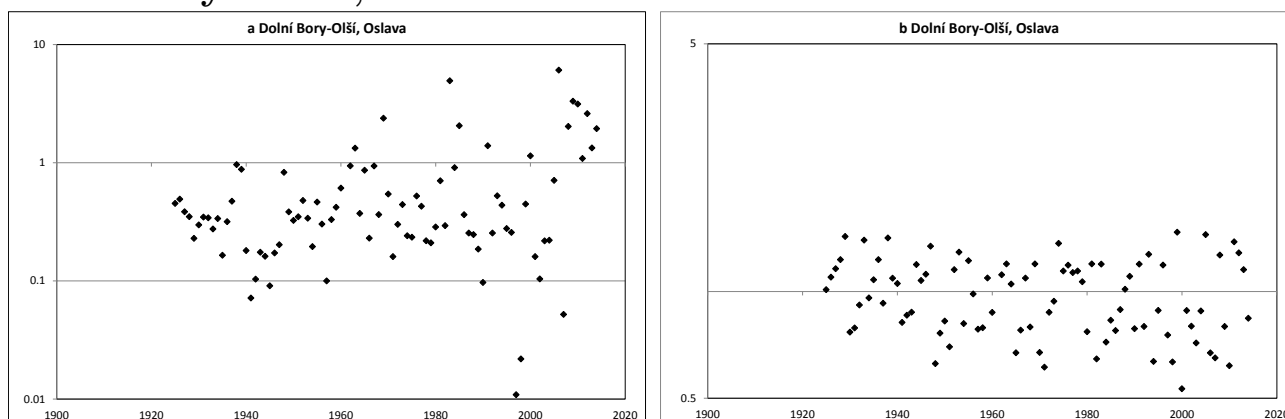
Borovnice, Svratka



Toto povodí je převážně zalesněno ve vyšší nadmořské výšce s postupně rostoucím podílem odvodňované půdy zhruba od konce 2. světové války. Nelesní

půda je zde odvodněna ve velmi vysokém podílu. Orné půdy je zde málo. Parametr $K100\%$ lehce roste, b a μ klesají, u parametru σ se výrazně snižuje rozptyl, i když v průměru neklesá. Takovéto chování odpovídá snižování zastoupení přímého odtoku. Půda s hladinou podzemní vody blízko povrchu se chová, pokud není velké sucho, jako beton a větší část srážkové vody odvádí do přímého odtoku. Po odvodnění, i když by byla půda utužena, převede větší podíl srážky odvodněním uprázdněným půdním horizontem se zpožděním do podzemního odtoku přes odvodňovací trubky. Mělo by tudíž lehce narůstat $K100\%$, což se i děje. Lehce se zvyšuje i medián. Výrazný několikaletý vzestup parametru a v letech 1985 až 1995 může souviset s narušováním funkcí vegetace a lesů zvláště v době nejvyšších exhalací oxidu siřičitého.

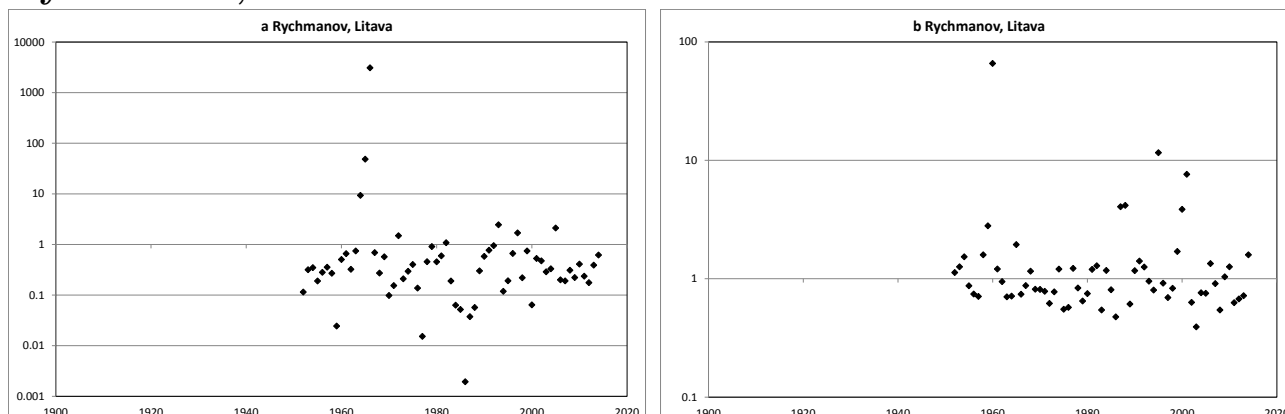
Dolní Bory – Olší, Oslava



V povodí se nachází značný počet velkých rybníků, a to po celou dobu pozorování. Zemědělská a orná půda se hodně odvodňovala (patrně na poklesu obou pruhů parametru b od začátku pozorování do 60. let podobně jako v povodí Svratky v Borovnici). Lesní půdy je méně než polovina. Ovlivnění lidskou činností ve smyslu výše popsaném je málo.

Parametr $K100\%$ pomalu klesá, medián stoupá, Q_a nepatrně klesá. Parametry b a μ lehce klesají a zvyšují rozptyl. Parametr σ lehce roste a zvyšuje rozptyl. Z těchto změn charakteristik je vidět, že voda z povodí se po dešti snaží odtéci dříve, leckdy i více jako povrchový odtok bezprostředně po dešti (zvyšující se parametr b zvláště po roce 1990). To může souviset s utužováním zemědělských půd i nárůstem zpevněných ploch či jejich přímým odkanalizováním do vodoteče, nikoli do okolních volných ploch jako tomu bylo dříve (i když v tomto povodí to asi nebude příliš dramatické). Rostoucí medián může souviset s ovlivněním, ale i s výtoky z odvodňovacích soustav.

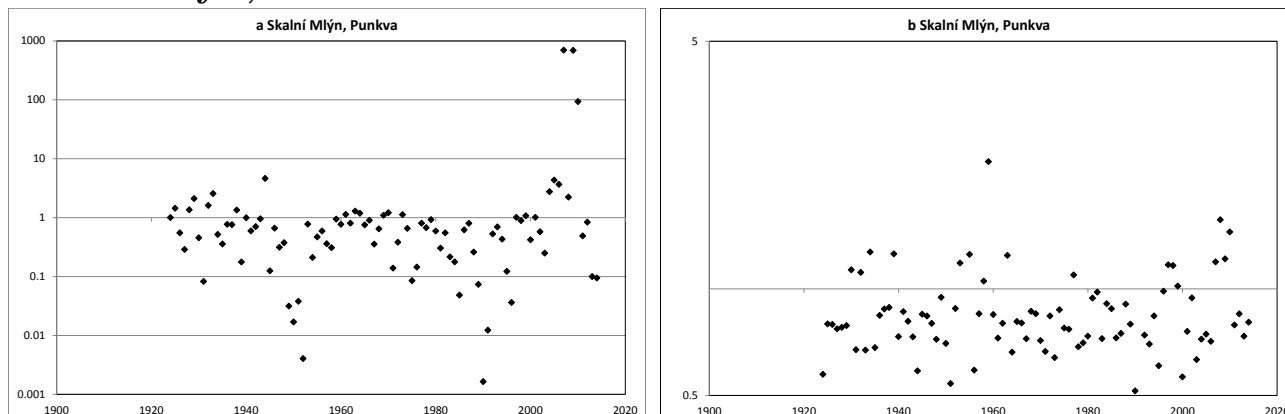
Rychmanov, Litava



Povodí má již vyšší ovlivnění vypouštěním vod do toků i odběry podzemních vod. Má většinou těžší zemědělské půdy, a to i na svazích. Roviny se vyskytují jen v nivách toků. Vrcholové partie jsou zalesněny. Lze tedy očekávat, že sklony k silnějšímu povrchovému odtoku při vydatnějších srážkách byly po zornění povodí vždy.

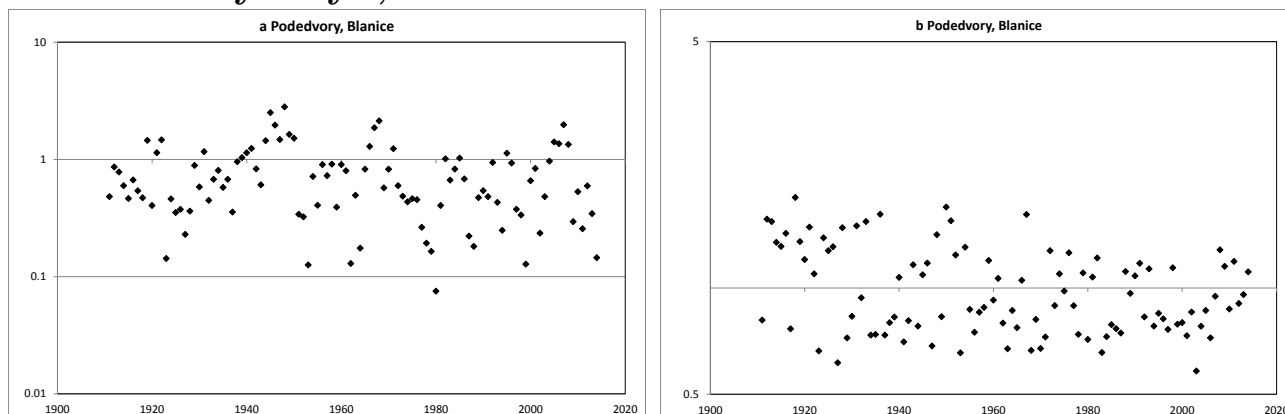
Zde je prokazatelné pouze zvyšování mediánu a pokles $K100\%$. Znamená to, že opět dochází k přesunu části odtoku k vyšším průtokům. Ostatní parametry vykazují velký rozptyl, což zřejmě plyne z výše uvedené charakteristiky povodí.

Skalní Mlýn, Punkva



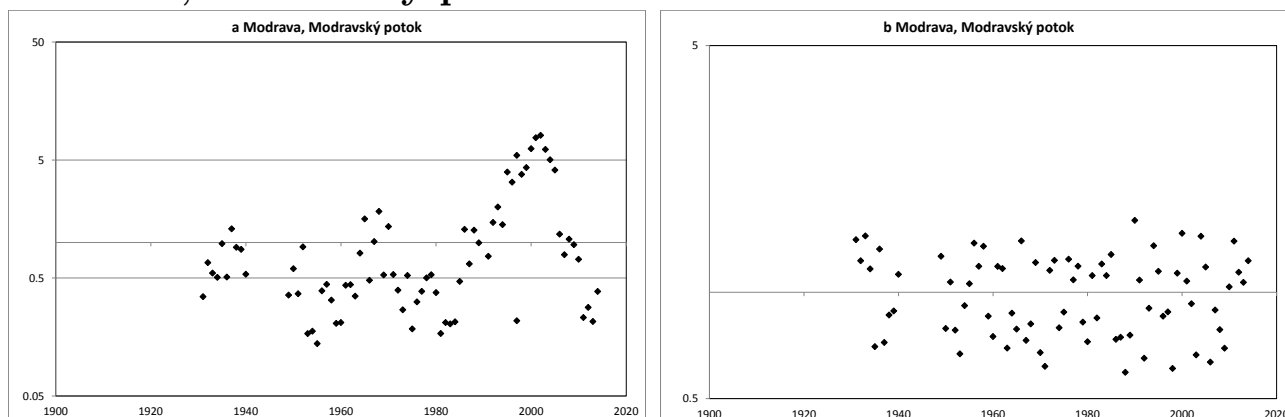
Stanice je pod jeskyněmi Moravského krasu. Zde se pravděpodobně žádné významné změny neudály. Povodí zřejmě reaguje na suchá období změnou parametrů kvůli vsakům vody do jeskynních systémů puklinami. Za sucha je proto zřejmě relativně silnější reakce na vyšší srážky než je obvyklé u jiných povodí. Vysoké hodnoty parametru a ke konci období odpovídají výskytu katastrofických dešťů v letech 2007 a 2008.

Podedvorský Mlýn, Blanice



Jedná se o povodí s lesy a loukami zhruba půl na půl. Je to povodí podobného charakteru jako povodí Zdechovky s malým ovlivněním, jen s obráceným vývojem charakteristik odtoku. Změny parametrů mají také podobný průběh. Historie povodí je taková, že před rokem 1945 bylo osídleno německou menšinou, a proto s velkým podílem polí. Po roce 1945 vesnice zanikly a rozšířil se les do dnešního stavu. Došlo k poklesu $K100\%$ i mediánu. Na průměr tyto změny neměly žádný vliv. Rozdíl je vidět hlavně na parametru b – výrazný pokles po roce 1945, což svědčí o klesající intenzitě reakce odtoku na srážky.

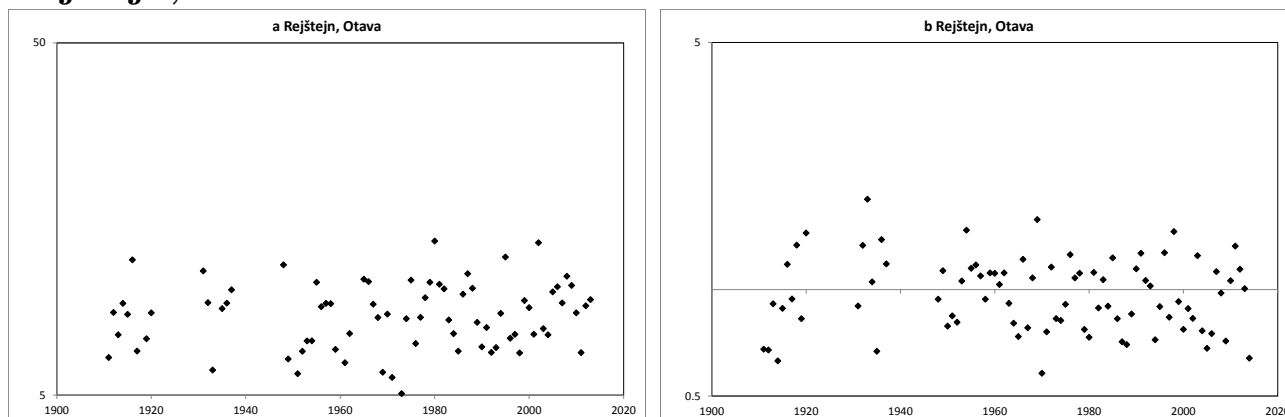
Modrava, Modravský potok



Nad touto stanicí od počátku 90. let byl likvidován smrkový les kůrovcem. Tato situace je na průběhu parametru a velmi dobře vidět, nastává prudká změna velikosti od 80. let.

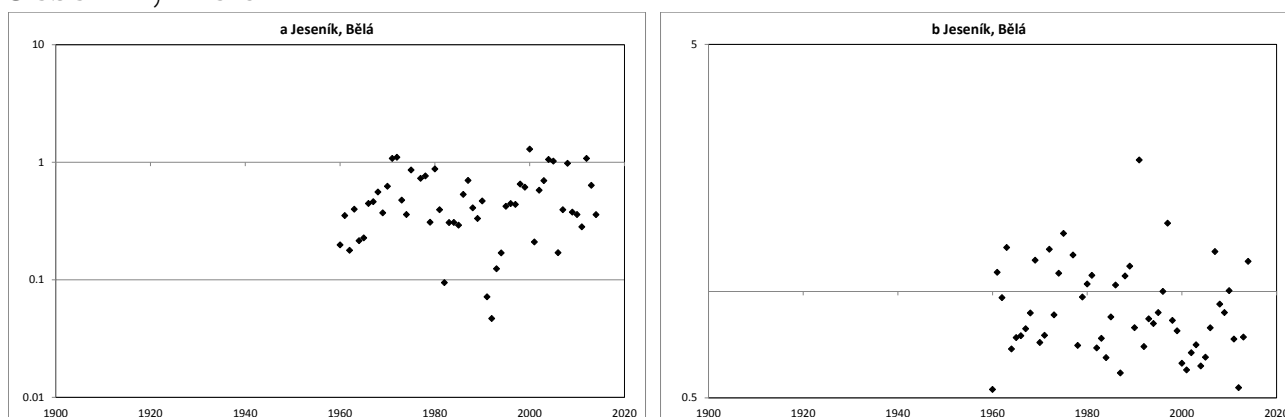
Na vývoji parametru b se rozpad lesa způsobený kůrovcem nijak významně neprojevil zřejmě proto, že důvodem rozpadu lesa nebylo zhoršení podmínek tak, že v půdě může růst jen máloco, ale že les byl zničen kůrovcem a vzápětí byl nahrazen novými dřevinami, i když malými, ale schopnými růst. To v povodí Smědé (viz výše) vzhledem ke kyselým srážkám možné nebylo.

Rejštejn, Otava



Je profilem, jímž Otava opouští Šumavu. Většina lesa v povodí je vizuálně v pořádku (kromě nejvyšších partií – např. nad Modravou viz předchozí povodí). Z průběhu parametru a ale je zřejmé, že i na tomto povodí dochází ke změnám charakteristik odtoku a tudíž vegetace není úplně v pořádku. Parametr b i σ od 50. let narůstají, takže se zesiluje průtoková odezva na stejnou srážku.

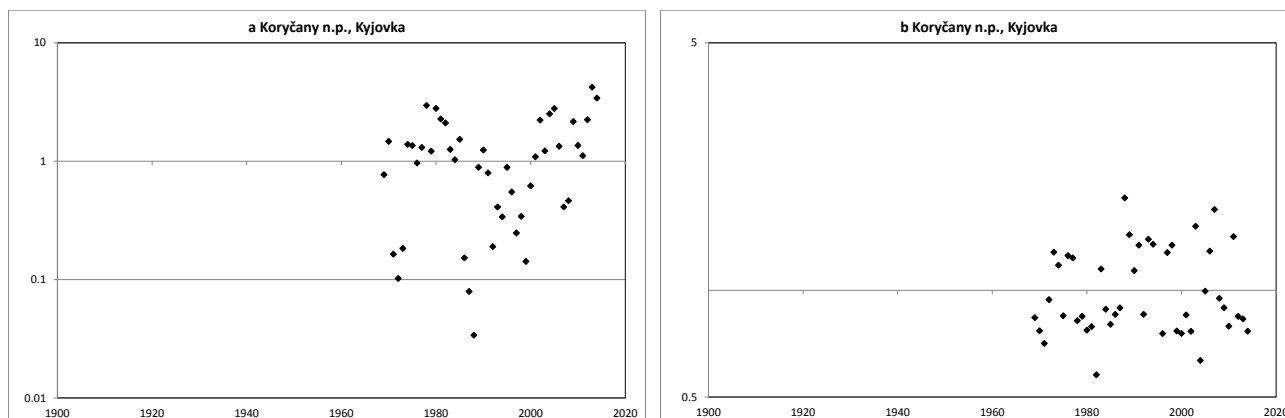
Jeseník, Bělá



Parametr a zde roste. Ke změně tedy dochází, ale je obtížně interpretovatelná. Ve vrcholových partiích je mladý les vyrostlý po poškození lesa, který byl oslaben imisemi oxidu siřičitého v 80. letech minulého století a následně vyvrácen vichřicemi. Poškození lesa nemělo zřejmě za následek významně rychlejší reakci odtoku na srážky, protože parametr b se nijak významně nemění.

Koryčany n.p., Kyjovka

Jedná se o povodí s již větší plochou zemědělské půdy a to i orné a do 25 % lesa. Parametr b se lehce zvyšuje zvláště po roce 1980. Hodnota z roku 1995 je mimo graf. Medián se téměř nemění, $K100\%$ výrazně klesá. Znamená to, že povodí snižuje uvolňování vody v době sucha. To je vidět i na kombinaci změn parametrů: σ klesá, μ také spíše klesá a parametr a má obdobné tendence. Je zde tedy tendence k narůstání přímého odtoku.



4. Závěr

Z uvedených analýz parametrů několika křivek překročení je vidět, že použijeme-li vhodnou transformaci normálního rozdělení, potom se změny ve využívání ploch zřetelně projevují i v charakteristikách odtoků. Teoretické rozdělení LN5 popíše tvary hydrologických křivek překročení daleko lépe než rozdělení LN3. Z výsledných analýz parametrů ročních teoretických křivek lze dovodit, že narušování vegetačního krytu nebo snižování propustnosti půd (ať už zhutněním či jejich zaplavením) vede ke zvýšení podílu povrchového odtoku a ke změně parametrů ročních křivek překročení. Parametry teoretických křivek LN5 na tyto změny reagují velmi citlivě a určitě by si tyto změny zasloužily detailnější a rozsáhlejší analýzy u více povodí než tato víceméně pouze kvalitativní a stručná analýza bez důsledných rozborů srážek a dějů v povodí.

Literatura

- [1] Budík, L.: *Řešení katastru M-denních průtoků pomocí teoretického rozdělení LN5*. 9th Int. Conference Aplimat, Bratislava, 2010, pp. 857–866.
- [2] Budík, L., Kotrnec, J.: *Stručný metodický popis a manuál k programu pro výpočet hydrologických návrhových veličin program KATABR – M-denní vody*. ČHMÚ, pobočka Brno, 1994, nepublikováno.
- [3] Novický, O., Kašpárek, L., Kolářová, S.: *Hydrologická data pro návrhové účely*. ČHMÚ, Praha, 1992.

MANIFEST CENTRA PRO BAYESOVSKOU INFERENCI

MANIFESTO OF THE BAYESIAN INFERENCE CENTER

**Halina Šimková, Tomáš Fürst, Jiří Drábek, Jan Zimmer,
Vlastimil Stenzl, Jan Strojil a Ondřej Vencálek**

Těžko najít kontext, v němž je správné usuzování z dat důležitější než v trestně právním procesu a v medicíně. V medicíně jde o život, v trestním právu o rozhodnutí, která často velmi radikálně mění lidské osudy. Lidský mozek však nebyl evolucí třiben k porozumění světu, ale k tomu, aby v něm zajistil svému nositeli co nejvyšší rozmnožovací úspěšnost, což je funkce diametrálně odlišná. Čím složitější svět je, tím hůře mozek odhaluje skutečnou váhu a vzájemné vazby všech informací, které jsou mu k dispozici. Důsledkem jsou omyly, nechtěná nespravedlnost a smrt.

Rádi si namlouváme, že lidé vysoce inteligentní a vzdělaní jsou před takovým selháváním úsudku automaticky chráněni a že obecně celý „inferenční systém západní civilizace“ je nastaven správně. Stále rostoucí replikační krize v mnoha výzkumných odvětvích [1], kdy v některých oblastech výzkumu to vypadá, že většina publikovaných výsledků není replikovatelná, stejně jako inferenční a interpretační krize řady forenzních oborů [2] však odhalují hrozivou skutečnost: Ani ve vědeckém světě není správná inference vůbec samozřejmostí.

Znalec je běžně u soudu tázán: „Tak co, pane znalče, je to jeho podpis nebo ne?“, přičemž jiná než zcela deterministická odpověď „určitě ano“ nebo „určitě ne“ není vítána. Lidské myšlení extrémně nerado akceptuje nejistotu jako integrální součást tohoto světa a nejeden soudce při zmínce o pravděpodobnosti zbledne, zamračí se a prohlásí: „Jaká pravděpodobnost? My musíme nade vší pochybnost!“. Centrum pro bayesovskou inferenci si klade za cíl tuto situaci změnit a nejen komunitu znalců, soudců, žalobců, obhájců a vyšetřovatelů v principech správné inference vzdělávat.

Proto si dovoluujeme předložit čtenářům Informačního bulletinu Manifest Centra pro bayesovskou inferenci. Jsme si vědomi, že text je výrazně populárnější, než je na těchto stránkách zvykem. Je to proto, že není cílen na profesionální statistiky, ale na široké „laické“ publikum. O pozornost profesionálních statistiků ovšem samozřejmě stojíme a velmi si vážíme možnosti Manifest na těchto stránkách prezentovat.

Reference

- [1] Baker, M.: 1,500 scientists lift the lid on reproducibility, *Nature* 533, 452–454, 2016 May 26. doi:10.1038/533452a
- [2] Roux, C., et al.: The end of the (forensic science) world as we know it? The example of trace evidence, *Philos Trans R Soc Lond B Biol Sci.* 370(1674): 20140260, 2015 Aug 5. doi: 10.1098/rstb.2014.0260

Manifest 4BIN

Držíme v prstech minci logiky a obdivně hledíme na její lícovou stranu, která je nám už dávno důvěrně známa: Když prší, budou mokré chodníky – A implikuje B. Málokoho však napadne peníz otočit: Když jsou mokré chodníky, pršelo? Ne nutně, mohl třeba projet kropicí vůz nebo prasknout vodovod, nicméně mokré chodníky *zvyšují pravděpodobnost*, že pršelo. První, kdo si existenci rubové strany mince zřetelně uvědomil, byl reverend Thomas Bayes. Tak se zrodila **bayesovská inference**. Od onoho okamžiku uplynulo čtvrt tisíciletí, a přesto velká část z nás stále odmítá uvěřit, že **každá mince má dvě strany**.

Logika je základem matematiky a matematika je jazykem přírodních věd. Poněkud nešťastně jsme pravděpodobnost dlouho považovali a často stále považujeme za jeden příklad z teorie míry¹, což nám brání pochopit, že **pravděpodobnost má stejně fundamentální význam jako logika**. Bayesovská revoluce nám teď dává šanci odvrácenou stranu logiky správně uchopit.

Naše civilizace je poháněna čtyřtákním motorem²: *pochopit jak a proč funguje příroda – na základě toho vymyslet technologie a inovace – tyto vyrobit – a prodat*. Poslední dobou motor škytá: **Velká část výsledků výzkumu není replikovatelná³; velká část výsledků výzkumu přináší predikce, které se nenaplní**. Tato replikační a predikční krize je do značné míry důsledkem užívání jiné inference, než je ta správná, tedy bayesovská. Neriskujme, že nám motor chcípne, a přepneme na Bayese hned, neboť

$P(a|f)$ **Bayesovská inference je princip jednotící a všeobjímající.**

Je univerzálně aplikovatelná na rozmanité typy vědeckých problémů, ale též na nejrůznější situace v běžném životě⁴.

¹<http://www.4bin.org/teorie-miry/>

²<http://www.4bin.org/motor-sveta/>

³<http://www.4bin.org/replikacni-krize/>

⁴<http://www.4bin.org/bezny-zivot/>

$P(b|f)$ **Bayesovská inference je princip plně racionální a logický.**

Všechny parametry, s nimiž pracuje, lze snadno definovat, přisoudit jim běžnou logikou uchopitelný význam a interpretovat je.

$P(c|f)$ **Bayesovská inference je princip pevný i pružný jako Lycra.**

Pevný, neboť užívá stále též formát řešení, a přitom pružný, neboť umožňuje – pokud je to třeba – vytvářet komplexní modely pro tu kterou specifickou situaci.

$P(d|f)$ **Bayesovská inference přidává další rozměry selskému rozumu, aniž by ho popírala.**

Selský rozum neboli *intuitivní inference* je nástrojem lidské mysli, který umí zacházet s odhadem pravděpodobností v jednoduchých situacích s přehledným uspořádáním; ve vysoce komplexním světě však selhává⁵. Bayesovská inference se umí i s komplexním informačním prostředím vypořádat, přitom ale stále staví na základech, které vybudovala inference intuitivní.

$P(e|f)$ **Bayesovská inference propojuje objektivní a subjektivní vnímání pravděpodobnosti.**

Je zcela jedinečná tím, že umožňuje skloubit objektivní fakta se subjektivní mírou přesvědčení o platnosti určité hypotézy a z tohoto spojení vyvodit závěry.

Jistota je v reálném světě nedosažitelná; bayesovská inference však umí dát nejistotě rozměr. Jistota je ochranným konstruktem, který si vytváří lidská mysl, aby eliminovala nepříjemné emoce a navodila pocit bezpečí. V komplexním světě se však lze mnohem bezpečněji opřít o správně kvantifikovanou nejistotu než o iluzi jistoty. Zacházet s podmíněnými pravděpodobnostmi není úplně jednoduché, ale je to nezbytně nutné ve všech oblastech lidské činnosti.

Nejlepší mírou úspěchu v inferenci je správnost predikcí na ní založených. Obrovské úspěchy metod strojového učení (*machine learning*⁶), které z bayesovské inference vyrostly, svědčí o tom, že **jdeme po správné cestě.**

Tento vesmír je **logický** – a **bayesovský**.

Bayes se i v češtině čte ['beiz]⁷ a *inference* znamená usuzování.

⁵<http://www.4bin.org/komplexni-svet/>

⁶<http://www.4bin.org/machine-learning/>

⁷<http://www.4bin.org/lingvisticky-koutek/>

Obsah

Vědecké a odborné články

Ladislav Budík, Marie Budíková

Využití rozdělení LN5 pro analýzu chování průtoků 1

Zprávy a informace

*Halina Šimková, Tomáš Fürst, Jiří Drábek, Jan Zimmer, Vlastimil Stenzl,
Jan Strojil a Ondřej Vencálek*

Manifest Centra pro bayesovskou inferenci 18

Informační bulletin České statistické společnosti vychází čtyřikrát do roka v českém vydání. Příležitostně i mimořádné české a anglické číslo. Vydavatelem je Česká statistická společnost, IČ 00550795, adresa společnosti je Na padesátém 81, 100 82 Praha 10. Evidenční číslo registrace vedené Ministerstvem kultury ČR dle zákona č. 46/2000 Sb. je E 21214. Časopis je na Seznamu recenzovaných neimpaktovaných periodik vydávaných v ČR, více viz server <http://www.vyzkum.cz/>.

The Information Bulletin of the Czech Statistical Society is published quarterly.
The contributions in the journal are published in English, Czech and Slovak languages.

Předseda společnosti: RNDr. Marek MALÝ, Státní zdravotní ústav, Šrobárova 48, Praha 10, 100 42, e-mail: mmaly@szu.cz.

Redakce: prof. RNDr. Gejza DOHNAL, CSc. (šéfredaktor), prof. RNDr. Jaromír ANTOCH, CSc., prof. Ing. Václav ČERMÁK, DrSc., doc. Ing. Jozef CHAJDIK, CSc., doc. RNDr. Zdeněk KARPÍŠEK, CSc., RNDr. Marek MALÝ, CSc., doc. RNDr. Jiří MICHÁLEK, CSc., prof. Ing. Jiří MILITKÝ, CSc., doc. Ing. Iveta STANKOVIČOVÁ, PhD., doc. Ing. Josef TVRDÍK, CSc., Mgr. Ondřej VENCÁLEK, Ph.D.

Redaktor časopisu: Mgr. Ondřej VENCÁLEK, Ph.D., ondrej.vencalek@upol.cz.
Informace pro autory jsou na stránkách společnosti, <http://www.statspol.cz/>.

DOI: 10.5300/IB, <http://dx.doi.org/10.5300/IB>
ISSN 1210–8022 (Print), ISSN 1804–8617 (Online)

Toto číslo bylo vytištěno s laskavou podporou Českého statistického úřadu.