



Česká zemědělská univerzita v Praze

**Fakulta životního
prostředí**

Program Invers_shluk

Dokumentace
Uživatelská příručka

Martin Heřmanovský

Praha, 2011

Tato dokumentace včetně programového vybavení byla vypracována v rámci grantového projektu „Možnosti zmírnění současných důsledků klimatické změny zlepšením akumulční schopnosti v povodí Rakovnického potoka (pilotní projekt)“ po číslem NAZV QH 91247

Koordinátor Ing. Ladislav Kašpárek, CSc.,
Řešitel prof Ing. Pavel Pech, CSc.

Obsah

1	Teoretický základ	3
2	Instalace programu	6
3	Struktura programu	6
4	Požadavky na podpůrný software	6
5	Použití programu	7
5.1	Vstupní data	7
5.2	Nastavení programu	7
5.3	Výstupy z programu	8
5.3.1	Chybové hlášení	8
5.3.2	Výsledky analýzy	9
6	Příkladová regionalizační studie	9

1 Teoretický základ

Program `Invers_shluk` slouží k odhadu zájmových hydrologických charakteristik na nepozorovaných povodích (např. parametry hydrologického modelu, specifické průtoky a jiné). Program vychází z regionalizačního přístupu nazývaného v literatuře přístup fyzikální podobnosti povodí (viz např. Acreman and Sinclayr (1986), Oudin et al. (2008), Parajka et al. (2005)). V tomto přístupu se odhadují zájmové charakteristiky na nepozorovaném povodí na základě zájmových charakteristik pozorovaných povodí, která jsou v jistém smyslu nepozorovanému povodí nejpodobnější. Tradičně se využívá podobnost povodí podle zvolených deskriptorů povodí (tyto mohou být klimatické, s půdními vlastnostmi související, s vegetačním krytem související, atd.) (např. v Nathan and McMahon (1990), Young (2006)). Odůvodnění, proč se podobnost povodí vyjadřuje na základě podobnosti ve zvolených deskriptorech, souvisí s velmi úzkou spojitostí mezi nimi a hydrologickým chováním povodí.

Navržený program je založen na konceptu formování deskriptorově homogenních regionů, o kterých se předpokládá, že jsou též hydrologicky homogenní, kolem nepozorovaných povodí. Nepozorovaná povodí jsou tedy situována v geometrických střezech formovaných regionů. Tento koncept vychází z práce, kterou navrhl a prezentoval Burn (1989) jako ROI (region of influence).

Jednotlivá pozorovaná povodí jsou do regionů přiřazována na základě jejich vzdálenosti od geometrického středu (nepozorovaného povodí) vypočtené ve vícerozměrném deskriptorovém prostoru. Alternativně jsou pozorovaná povodí do shluku přiřazována na základě jejich podobnosti s nepozorovaným povodím. K tomuto účelu program využívá následující měřítka vzdálenosti (resp. podobnosti):

1. Minkowského vzdálenost,
2. Canberrská vzdálenost,
3. Bray-Curtisova vzdálenost,
4. kosínová podobnost.

Pro všechna níže popsaná měřítka vzdálenosti (resp. podobnosti) platí, že d_{UG} (resp. s_{UG}) je vzdálenost (resp. podobnost) mezi nepozorovaným povodím U a pozorovaným povodím G , X_i^U je hodnota i -tého deskriptoru X nepozorovaného povodí U a X_i^G je hodnota i -tého deskriptoru X pozorovaného povodí G . Minkowského vzdálenost mezi nepozorovaným povodím U a pozorovaným povodím G je definována předpisem:

$$d_{UG} = \left[\sum_{i=1}^N |X_i^U - X_i^G|^n \right]^{\frac{1}{n}} \quad n \geq 1 \quad (1)$$

Při $n = 1$ hovoříme o Manhattanské vzdálenosti:

$$d_{UG} = \sum_{i=1}^N |X_i^U - X_i^G| \quad (2)$$

Při $n = 2$ hovoříme o Euklidově vzdálenosti:

$$d_{UG} = \left[\sum_{i=1}^N (X_i^U - X_i^G)^2 \right]^{\frac{1}{2}} \quad (3)$$

Zvláštním případem Minkowského vzdálenosti je Chebyshevova vzdálenost, pro kterou platí, že $n = \infty$. Chebyshevova vzdálenost mezi nepozorovaným povodím U a pozorovaným povodím G je pak definována předpisem:

$$d_{UG} = \max |X_i^U - X_i^G| \quad (4)$$

Canberrská vzdálenost mezi nepozorovaným povodím U a pozorovaným povodím G je definována předpisem:

$$d_{UG} = \sum_{i=1}^N \frac{|X_i^U - X_i^G|}{|X_i^U| + |X_i^G|} \quad (5)$$

Lze si všimnout, že pokud jsou použité deskriptory nulové, pak je nutné definovat vztah:

$$0 = \frac{0}{0} \quad (6)$$

Dále, tato vzdálenost je citlivá na velmi malé změny v použitých deskriptorech, pokud jsou jejich hodnoty velmi blízko nule. Bray-Curtisova vzdálenost mezi nepozorovaným povodím U a pozorovaným povodím G je definována předpisem:

$$d_{UG} = \frac{\sum_{i=1}^N |X_i^U - X_i^G|}{\sum_{i=1}^N |X_i^U + X_i^G|} \quad (7)$$

Pro Bray-Curtisovu vzdálenost platí, že pokud jsou použité deskriptory kladné, pak je výsledná hodnota vzdálenosti v intervalu $\langle 0; 1 \rangle$, pokud mají použité deskriptory hodnotu 0, pak není tato vzdálenost definována. Kosínová podobnost mezi nepozorovaným povodím U a pozorovaným povodím G je definována předpisem:

$$s_{UG} = \frac{\sum_{i=1}^N (X_i^U \cdot X_i^G)}{\sqrt{\sum_{i=1}^N (X_i^U)^2 \cdot \sum_{i=1}^N (X_i^G)^2}} \quad (8)$$

Po přiřazení pozorovaných povodí do regionů je nutné provést kontrolu vnitroregionové homogenity a následně odstranit pozorovaná povodí, jejichž podobnost s nepozorovaným povodím je již diskutabilní. Kontrola vnitroregionové homogenity se provádí porovnáním Andrewsových křivek do shluků přiřazených pozorovaných povodí s Andrewsovou křivkou nepozorovaného povodí. Andrewsovy křivky reprezentují body (zde povodí) ve vícerozměrném prostoru o souřadnicích (hodnotách vybraných deskriptorů povodí) $x = [X_1, X_2, X_3, \dots, X_n]$ matematickou funkcí ve formě (Andrews, 1972):

$$f(t) = \frac{X_1}{\sqrt{2}} + X_2 \sin(t) + X_3 \cos(t) + X_4 \sin(2t) + X_5 \cos(2t) + \dots \quad (9)$$

v rozsahu $-\pi \leq t \leq \pi$, kde X_1, X_2, \dots, X_n jsou použité deskriptory povodí. Andrews ukázal, že podobnost dvou křivek je úměrná Euklidovské vzdálenosti. Tedy dva vícerozměrné body, které jsou prostoru situovány blízko sebe budou dávat velmi podobné křivky.

Pozorovaná povodí, jejichž Andrewsovy křivky jsou méně podobná Andrewsově křivce nepozorovaného povodí jsou z regionů odstraněna. Podobnost Andrewsových křivek je posuzována pomocí dvou kritérií – r_t^2 test a d_t test. Test r_t^2 využívá při porovnání Andrewsových křivek koeficient determinace:

$$r^2 = 1 - \frac{\sum_{b=1}^B (f_U(b) - f_G(b))^2}{\sum_{b=1}^B (f_U(b) - \bar{f}_U)^2} \quad (10)$$

kde $f_U(b)$ je funkční hodnota Andrewsovy křivky nepozorovaného povodí U v bodě b , $f_G(b)$ je funkční hodnota Andrewsovy křivky pozorovaného povodí G v bodě b a \bar{f}_U je průměr hodnot vypočtených na

základě Andrewsovy křivky nepozorovaného povodí. Pokud je vypočtená hodnota r_t^2 nižší než předem definovaná prahová hodnota r_t^2 , pak je pozorované povodí odstraněno z regionu. Při d_t testu jsou porovnávány absolutní hodnoty odchylek mezi Andrewsovou křivkou pozorovaného povodí a Andrewsovou křivkou nepozorovaného povodí pro všechna b :

$$d_t(b) = |f_U(b) - f_G(b)| \quad (11)$$

kde $d_t(b)$ je absolutní hodnota odchylky mezi Andrewsovou křivkou pozorovaného povodí G a Andrewsovou křivkou nepozorovaného povodí U v bodě b , $f_U(b)$ je funkční hodnota Andrewsovy křivky nepozorovaného povodí U v bodě b a $f_G(b)$ je funkční hodnota Andrewsovy křivky pozorovaného povodí G v bodě b . Pokud je alespoň v jednom bodě b odchylka $d_t(b)$ větší než předem definovaná prahová hodnota d_t , pak je pozorované povodí vyřazeno z regionu.

Zájmové hydrologické charakteristiky pro každé nepozorované povodí jsou v konečné fázi vypočteny jako vážené průměry známých hydrologických charakteristik přiřazených pozorovaných povodí v rámci každého regionu:

$$\theta_i^u = \frac{\sum_{p=1}^N (w_p \theta_{ip}^G)}{\sum_{p=1}^N w_p} \quad (12)$$

kde θ_i^U je zájmová hydrologická charakteristika na nepozorovaném povodí, θ_{ip}^G je známá hydrologická charakteristika pozorovaného povodí a w_p je váha aplikovaná na každou známou hydrologickou charakteristiku θ_{ip}^G . Váhou zde byla zde volena převrácená hodnota vzdálenosti (resp. podobnosti) mezi geometrickým středem regionu (nepozorovaným povodím) a přiřazovaným pozorovaným povodím.

Kvůli skutečnosti, že použité deskriptory povodí mohou mít různé velikosti a rozsahy, je nutné tyto standardizovat kvůli jejich souměřitelnosti. Program využívá různé způsoby standardizace použitých deskriptorů povodí. Tyto jsou označeny S1 až S6.

Standardizace S1:

$$S_{iA} = X_{iA} \quad (13)$$

Standardizace S2:

$$S_{iA} = \frac{X_{iA}}{\max A} \quad (14)$$

Standardizace S3:

$$S_{iA} = \frac{X_{iA}}{\max A - \min A} \quad (15)$$

Standardizace S4:

$$S_{iA} = \frac{X_{iA} - \min A}{\max A - \min A} \quad (16)$$

Standardizace S5:

$$S_A = \frac{X_{iA}}{s_A} \quad (17)$$

Standardizace S6:

$$S_A = \frac{X_{iA} - \bar{X}_A}{s_A} \quad (18)$$

kde S_{iA} je i -tá hodnota standardizovaného deskriptoru povodí A , X_{iA} je i -tá hodnota nestandardizovaného deskriptoru povodí A , $\max A$ je zjištěné maximum mezi hodnotami deskriptoru povodí A , $\min A$ je zjištěné minimum mezi hodnotami deskriptoru povodí A , s_A je směrodatná odchylka vypočtená na základě všech nestandardizovaných hodnot deskriptoru povodí A a \bar{X}_A je průměrná hodnota vypočtená na základě všech nestandardizovaných hodnot deskriptoru povodí A .

Můžeme si všimnout, že v případě standardizací S2 až S4 převádíme deskriptory povodí na čísla v intervalu $\langle 0; 1 \rangle$. V případě standardizace S6 převádíme deskriptory povodí na hodnoty, jejichž průměr je 0 a směrodatná odchylka je 1.

2 Instalace programu

Instalace programu spočívá v extrakci adresáře `invers_shluk` ze souboru `invers_shluk.zip` do oddílu C, tak aby konečná cesta k programu byla: `C:\invers_shluk`. Adresář `invers_shluk` je sice možné extrahovat na libovolný dostupný oddíl a do libovolného adresáře, avšak kvůli defaultně nastaveným cestám při tvorbě dávkových souborů pro `gnuplot` (`xxx.GNU`) je doporučena pouze výše uvedená extrakce.

3 Struktura programu

Program `Invers_shluk` je napsán v programovacím jazyku C. Primární adresář `invers_shluk` obsahuje:

1. podadresář `errors`,
2. podadresář `input_data`,
3. podadresář `output_data`,
4. podadresář `settings`,
5. spouštěcí soubor `Invers_shluk.exe`,
6. zdrojový soubor `Invers_shluk.cpp`,
7. zdrojový soubor `Invers_shluk.dev`.

Podadresář `errors` je určen pouze pro chybová hlášení a jejich výpisy a soubory v něm obsažené se vytvářejí v průběhu výpočtu. Podadresář `input_data` musí obsahovat dva soubory s příponou `txt`. Jedná se o soubor označený `parametry_modelu.txt` se známými hydrologickými charakteristikami pozorovaných povodí a soubor s deskriptory povodí označený `nenorm_deskripty_povodi.txt`. Podadresář `output_data` je určen pro soubory s výsledky a podle nastavení může obsahovat jednu nebo více složek označených `output_data_1` až `output_data.n`. Podadresář `settings` obsahuje nastavení programu uložené v ve třech souborech: `kombinace_deskripty.txt`, `settings_clus.txt` a `settings_deskripty.txt`.

Soubory obsahující zdrojový kód programu může uživatel podle potřeby dále upravovat tak, jak potřebuje. K tomuto účelu lze využít např. program Dev-C++, který je volně dostupný.

4 Požadavky na podpůrný software

Program `Invers_shluk` ke své plné funkčnosti potřebuje instalaci následujících tří programů: `gnuplot`, `Ghostgum` a `Ghostscript`. Všechny tyto programy jsou volně dostupné. Program `gnuplot` je využit při tvorbě grafů ve formátu `xxx.PS`, zbývající dva programy usnadňují práci s grafy ve formátu `xxx.PS`.

5 Použití programu

Po úspěšné extrakci a kontrole, zda podadresáře `input_data` a `settings` obsahují předepsané soubory (pokud jsou podadresáře prázdné, pak soubory musíme vytvořit, ale jejich jména se musí shodovat se jmény uvedenými v manuálu), můžeme přistoupit k požadovaným analýzám.

5.1 Vstupní data

Nejprve musíme do podadresáře `input_data` zkopírovat vstupní data. Zájmové hydrologické charakteristiky (např. parametry modelu) nakopírujeme do souboru `parametry_modelu.txt`. V případě nepozorovaného povodí nahradíme zájmové charakteristiky číslem 0. Do souboru `nenorm_deskripty_povodi.txt` nakopírujeme všechny deskriptory povodí, které máme k dispozici (se kterými budeme pracovat). Oba soubory vstupních dat musí obsahovat také číselné identifikátory každého povodí, které musí být umístěny v prvním sloupci souboru. Identifikátory musí být pouze číselné hodnoty.

5.2 Nastavení programu

V podadresáři `settings` provedeme nastavení programu. Soubor `kombinace_deskriptoru.txt` obsahuje zvolené kombinace deskriptorů, které jsou použity při formování homogenních regionů. Je důležité, aby byly testovány kombinace deskriptorů, které obsahují vždy stejný počet deskriptorů (tedy jenom dvojice, trojice, atd.).

Dále, nesmí dojít k situaci, kdy jedna nebo více kombinací deskriptorů odkazuje na neexistující sloupec v souboru `nenorm_deskripty_povodi.txt`. Např. testuje se kombinace 1-10 (dvojice obsahující první a desátý deskriptor povodí), avšak v souboru `nenorm_deskripty_povodi.txt` je pouze devět druhů deskriptorů povodí (je tedy odkazováno na neexistující desátý sloupec). V tomto případě výpočet neproběhne.

Soubor `settings_clus.txt` obsahuje celkem sedm nastavení. Nastavení jsou zde očíslována vzestupně tak, jak po sobě v souboru následují:

1. výběr měřítka vzdálenosti (resp. podobnosti),
2. Minkovského vzdálenost - nastavení 1,
3. Minkovského vzdálenost - nastavení 2,
4. prahová vzdálenost při formování primárních regionů,
5. výběr vyhodnocení testu homogenity,
6. prahová hodnota pro r_t^2 test,
7. prahová hodnota pro d_t testu,
8. výběr počtu analyzovaných povodí,
9. počet analyzovaných povodí.

Při výběru měřítka vzdálenosti se uživatel rozhoduje jaké měřítko vzdálenosti (resp. podobnosti) při analýze využije. V tabulce 1 jsou uvedena označení pro jednotlivá měřítko vzdálenosti (resp. podobnosti).

Pokud je pro analýzu vybrána Minkovského vzdálenost, je nutné tento výběr dále specifikovat. K tomu slouží nastavení č.2 a nastavení č.3. Nastavení č.2 specifikuje „typ“ Minkovského vzdálenosti (definuje n v rovnici 1). V tabulce 2 jsou uvedeny nejčastější používané možnosti.

Nastavení č.3 již pouze specifikuje, zda při výpočtu vzdálenosti dojde k odmocnění či nikoliv (zda např. bude uživatel pracovat se čtvercem Euklidovy vzdálenosti či nikoliv). Pokud zvolíme hodnotu 1

Tab. 1: Označení jednotlivých měřítek vzdálenosti (resp. podobnosti) v programu `Invers_shluk`.

označení	měřítka vzdálenosti (resp. podobnosti)
1	Minkovského vzdálenost
2	kosinová podobnost
3	Chebyshevova vzdálenost
4	Canberrská vzdálenost
5	Bray-Curtisova vzdálenost

v nastavení č.3, pak dojde o odmocnění při výpočtu vzdálenosti. Při volbě hodnoty 2 v nastavení č.3 k odmocnění při výpočtu nedojde.

Nastavení č.4 vymezuje velikost primárních regionů na základě vypočtených vzdáleností pozorovaných povodí od nepozorovaného povodí v deskriptorovém prostoru. V primárním regionu jsou v počáteční fázi přiřazena všechna pozorovaná povodí, mezi kterými je identifikováno takové pozorované povodí, jehož vzdálenost ke geometrickému středu regionu je nejmenší (d_{UG}^{min}). Nastavení č.4 pak určuje k -násobek nalezené minimální vzdálenosti a tedy prahovou hodnotu vzdálenosti pozorovaného povodí od geometrického středu (d_{UG}^t):

$$d_{UG}^t = d_{UG}^{min} \cdot \text{hodnota č.4} \quad (19)$$

Proto je možné uvažovat o nastavení č.4 jako u hrubém testu homogenity. Toto nastavení může nabývat teoreticky hodnot v intervalu $\langle 0; +\infty \rangle$.

Nastavení č.5 určuje výběr vyhodnocení testu homogenity. Při volbě 1 je použit d_t test, při volbě 2 je použit r_t^2 test. Nastavení č. 6 a č.7 pak pouze specifikují prahovou hodnotu použitého testu. Nastavení č.6 souvisí s r_t^2 testem a prahová hodnota se vybírá z intervalu $\langle 0; 1 \rangle$. Nastavení č.7 souvisí s d_t testem a prahová hodnota se vybírá z intervalu $\langle 0; 1 \rangle$. Před vlastní analýzou je vhodné stanovit optimální prahové hodnoty pro d_t test i r_t^2 test s ohledem na očekávaný počet použitých deskriptorů povodí.

Nastavení č.8 a č.9 definují počet analyzovaných nepozorovaných povodí. Nastavení č.8 rozhoduje o tom, zda se bude každé povodí v souboru řešit postupně jako nepozorované (program otestuje každé povodí). Pro tento výběr zadáváme hodnotu 1. V případě, že zadáme hodnotu 2 jako nastavení č.8, budeme analyzovat pouze zvolený počet povodí. Tato povodí je nutné v souborech se vstupními daty umístit na začátek. Nastavení č.9 pak pouze kvantifikuje počet analyzovaných povodí. Je důležité, aby počet analyzovaných povodí, zadaný v souboru nastavení, nebyl větší než reálný počet povodí v souborech se vstupními daty. V případě, že tomu tak bude, výpočet neproběhne.

Soubor `settings_deskripty.txt` obsahuje dvě nastavení. První číslo udává výběr standardizace deskriptorů povodí (přiřazení označení viz tabulka 3) a druhé číslo, které udává počet testovaných deskriptorů povodí (kdy např. číslo 2 odpovídá testování dvojic, atd.).

5.3 Výstupy z programu

5.3.1 Chybové hlášení

Pokud počet řádků v souboru `parametry_modelu.txt` (resp. počet parametrických sad analyzovaných povodí) neodpovídá počtu řádků v souboru `nenorm_deskripty_povodi.txt` (resp. počtu n -tic deskriptorů

Tab. 2: Výběr „typu“ Minkovského vzdálenosti v programu `Invers_shluk`.

označení typu	typ Minkovského vzdálenosti
1	Manhattanská vzdálenost
2	Euklidova vzdálenost

Tab. 3: Výběr nastavení standardizace deskriptorů povodí v programu `Invers_shluk`.

označení typu	typ standardizace deskriptorů povodí
1	S1 viz rovnice 13
2	S2 viz rovnice 14
3	S3 viz rovnice 15
4	S4 viz rovnice 16
5	S5 viz rovnice 17
6	S6 viz rovnice 18

povodí), spuštěný program vypíše do podadresáře `errors` chybové hlášení o tom, která povodí v jakém souboru chybí a dojde k ukončení výpočtu.

5.3.2 Výsledky analýzy

Pokud odpovídají počty řádků (a tedy povodí) v souborech vstupních dat, začnou se v podadresáři `output_data` vytvářet složky označené `output_data_X`. Počet vytvořených složek odpovídá počtu n -tic v souboru `kombinace_deskriptoru.txt`.

Každá složka `output_data_X` obsahuje čtyři různé výstupy pro každé nepozorované povodí a dva souhrnné výstupy. Pro každé nepozorované povodí je vytvořen:

- dávkový soubor pro `gnuplot(XXX_AND_krivky_graf_XX.GNU)`,
- soubor obsahující funkční hodnoty Andrewsovy křivky nepozorovaného povodí a Andrewsových křivek v regionu ponechaných pozorovaných povodí (`data_pro_andy_XX.XXX.txt`),
- soubor s hodnotami vypočtených zájmových charakteristik na nepozorovaném povodí v porovnání s optimálními hodnotami (`XXX_vazene_parametry_XX.txt`),
- soubor obsahující v regionu ponechaná pozorovaná povodí s jejich vzdáleností od geometrického středu regionu a příslušnou hodnotu d_t testu nebo r_t^2 testu (`distance_XX.XXX.txt`).

Ve všech výstupech je `XXX` identifikátor daného nepozorovaného povodí a `XX` charakterizuje test, který byl použit k vyhodnocení vnitroregionové homogenity (MD pro d_t test a CE pro r_t^2 test).

Souhrnné výstupy obsahují soubor s počty pozorovaných povodí (`pocet_povodi_XX.txt`), která byla ponechána v regionech, pro každé nepozorované povodí. Druhý soubor obsahuje vypočtené absolutní procentické odchylky odhadovaných zájmových charakteristik od jejich optimálních hodnot. Tento soubor je označen `odchylky_pro_vsechna_ungage_povodi_XX.txt`. Pro oba souhrnné výstupy platí, že `XX` charakterizuje test, který byl použit k vyhodnocení vnitroregionové homogenity (MD pro d_t test a CE pro r_t^2 test).

6 Příkladová regionalizační studie

Pro ukázkou funkce programu `Invers_shluk` obsahuje podadresář `input_data` hydrologická data projektu MOPEX (Duan et al., 2006). Jedná se o velkou skupinu deskriptorů povodí a a priori parametrů modelu Sacramento. Ve vstupních datech si můžeme všimnout, že každému povodí je přiřazen číselný identifikátor (umístěný v prvním sloupci), který se pak objeví v označení výstupních souborů.

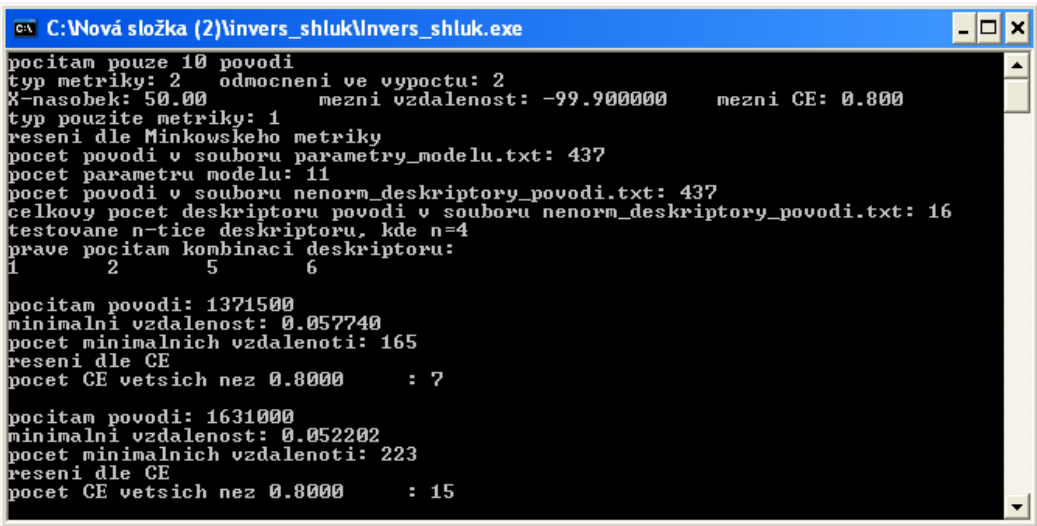
V podadresáři `settings` je uvedeno vzorové nastavení programu. Soubor `kombinace_deskriptoru.txt` obsahuje matici o rozměrech 4×4 . Z toho je zřejmé, že jsou testovány celkem čtyři čtveřice deskriptorů.

V souboru `settings_clus.txt` jsou uvedena následující čísla 1 2 2 50 2 0,8 -99,9 2 10. Tato čísla znamenají:

- použité měřítko je Minkowského vzdálenost (první číslo je 1),
- v rámci Minkowského vzdálenosti se pracuje s Euklidovskou vzdáleností (druhé číslo je 2),
- při výpočtu se použije čtverec Euklidovy vzdálenosti (třetí číslo je 2),
- v primárních shlucích jsou přiřazena pouze pozorovaná povodí, jejichž vzdálenost je menší než padesátinásobek nejmenší identifikované vzdálenosti (čtvrté číslo je 50),
- pro vyhodnocení testu homogenity je použit r_t^2 test (páté číslo je 2),
- prahová hodnota r_t^2 testu je 0,80 (šesté číslo je 0,8),
- prahová hodnota d_t testu (sedmé číslo je -99,9) není využita při výpočtu, ale v nastavení musí být na této pozici nějaké číslo,
- při analýze budou testovány pouze vybraný počet povodí (osmé číslo je 2),
- bylo vybráno 10 nepozorovaných povodí (deváté číslo je 10).

V souboru `settings_deskripty.txt` jsou uvedena dvě čísla: 4 a 6. První číslo udává, že jsou při analýze použity právě čtveřice deskriptorů povodí a druhé číslo udává použití standardizace S6.

Při spuštění programu (soubor `Invers_shluk.exe`) se objeví okno pro příkazový řádek (viz obr. 1).



```

C:\Nová složka (2)\invers_shluk\invers_shluk.exe
pocitam pouze 10 povodi
typ metriky: 2   odmocneni ve vypoctu: 2
k-nasobek: 50.00   mezni vzdalenost: -99.900000   mezni CE: 0.800
typ pouzite metriky: 1
reseni dle Minkowskeho metriky
pocet povodi v souboru parametry_modelu.txt: 437
pocet parametru modelu: 11
pocet povodi v souboru nenorm_deskripty_povodi.txt: 437
celkovy pocet deskriptoru povodi v souboru nenorm_deskripty_povodi.txt: 16
testovane n-tice deskriptoru, kde n=4
prave pocitam kombinaci deskriptoru:
1   2   5   6

pocitam povodi: 1371500
minimalni vzdalenost: 0.057740
pocet minimalnich vzdalenosti: 165
reseni dle CE
pocet CE vetsich nez 0.8000   : 7

pocitam povodi: 1631000
minimalni vzdalenost: 0.052202
pocet minimalnich vzdalenosti: 223
reseni dle CE
pocet CE vetsich nez 0.8000   : 15

```

Obr. 1: Okno příkazového řádku se spuštěným programem `Invers_shluk`.

Na obrázku je vidět jednak použité nastavení, ale také informace o tom, která kombinace deskriptorů je programem právě počítána a jaké nepozorované povodí je řešeno. Dále jsou v okně uvedeny některé výsledky analýzy (identifikovaná minimální vzdálenost pozorovaného povodí od geometrického středu v primárním regionu, počet pozorovaných povodí v primárním shluku s maximálně k -násobnou vzdáleností a počet pozorovaných povodí v regionu, jejichž Andrewsovy křivky splňují prahovou hodnotu zvoleného testu).

V podadresáři `output_data` jsou po skončení výpočtu přítomny čtyři složky: `output_data_1` až `output_data_4`. První složka pak odpovídá výsledkům testování první čtveřice deskriptorů, atd. Po analýze výsledků je vhodné odstranit v podadresáři `output_data` všechny složky. Teprve pak je možné změnit nastavení a program opět použít.

Literatura

- Acreman, A. C., Sinclayr, C. D., 1986: Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland. *Journal of Hydrology*, 84, 365-380.
- Andrews, D. F., 1972: Plots of high-dimensional data. *Biometrics*, 28, 125-136.
- Burn, D. H., 1989: Cluster analysis as applied to regional flood frequency. *Journal of Water Resources Planning and Management*, 115, 567-582.
- Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H. V., Gusev, Y. M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O. N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., Wood, E. F., 2006: Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *Journal of Hydrology*, 320, 3-17.
- Nathan, R. J., McMahon, T. A., 1990: Identification of homogeneous regions for the purposes of regionalisation. *Journal of Hydrology*, 121, 217-238.
- Oudin, L., Andréassian, V., Perrin, C., Michel, C., Le Moine, M., 2008: Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resources Research*, 44, W03413, doi:10.1029/2007WR006240.
- Parajka, J., Merz, R., Blöschl, G., 2005: A comparison of regionalisation methods for catchment model parameters. *Hydrol. Earth Syst. Sci.*, 9, 157-171.
- Young, A. R., 2006: Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model. *Journal of Hydrology*, 320, 155s-172s.