VYUŽITÍ UMĚLÝCH NEURONOVÝCH SÍTÍ A GENETICKÝCH ALGORITMŮ V MODELOVÁNÍ ČASOVÝCH ŘAD

Jan Šrejber[1], Roman Neruda[2], Martin Neruda[3]

Abstrakt

Článek prezentuje postupy a výsledky experimentů s predikcí časové řady průtoků za využití modelovacích technik založených čistě na vztahu vstupních a výstupních dat bez nutnosti znalostí fyzikálních principů modelovaného procesu jako je lineární regrese nebo při aplikaci umělé neuronové sítě (UNS). Ty umožňují modelování i nelineárních vícerozměrných vztahů.

Zájmovou oblastí bylo povodí horní Ploučnice po limnigrafickou stanici Mimoň, která sloužila jako předpovědní profil. V povodí o rozloze 268 km^2 jsou umístěny dvě srážkoměrné stanice Stráž pod Ralskem a Jablonné v Podještědí. Datové řady poskytnuté Českým hydrometeorologickým ústavem v Ústí nad Labem zahrnovaly hodinová data průtoků ze stanice Mimoň, dále srážková a teplotní data ze stanic Stráž a Jablonné za období dvou let (2006–2007). Nakonec byl přidán jako další proměnná ukazatel předchozích srážek API. Data byla rozdělena na dvě části, 90 % tvořila trénovací a 10 % testovací množina.

Byly testovány vícevrstvé perceptronové UNS různých architektur s jednou i dvěmi skrytými vrstvami, různou historií i dobou předstihu od *t+1* do *t+6*. Predikce na *t+6* hodin byla testována i na jednoduchém lineárním prediktoru, který sloužil pro nastavení měřítka úspěšnosti metod jednovrstvého perceptronu i vícevrstvé neuronové sítě. Jako kritérium úspěšnosti předpovědních modelovacích technik byl použit zejména koeficient efektivity (Nash-Sutcliffe). Pro výběr relevantních hodnot vstupů byly použity i genetické algoritmy (GA).

Klíčová slova: hydrologické modelování, neuronové sítě, genetické algoritmy

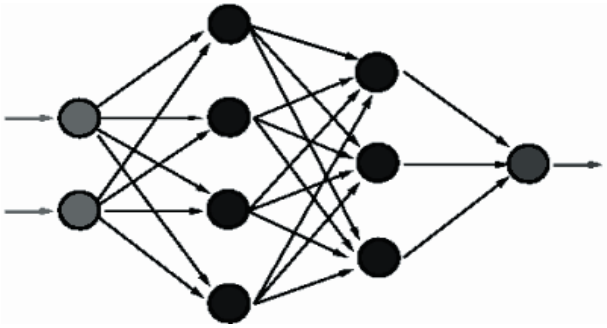
1. Úvod

Modelování přírodních tedy i hydrologických jevů se snaží o vysvětlení a matematické vyjádření fyzikálních zákonitostí. V praxi se často pro modelování procesů hydrologického cyklu využívá koncepčních deterministických modelů, které se snaží o co nejpřesnější popis těchto fyzikálních zákonitostí. Obecně se předpokládá, že čím přesněji jsou tyto vztahy v modelu popsány, tím jsou i výsledky modelu přesnější.

Na druhé straně existují modely, které jsou založeny čistě na vztahu vstupních a výstupních dat bez nutnosti znalostí fyzikálních principů modelovaného procesu. Spoléhají se na vypovídající schopnosti vstupních a výstupních dat při řešení daného problému. Příkladem jsou lineární regresní modely nebo umělé neuronové sítě umožňující modelování i nelineárních vícerozměrných vztahů např. transformace srážky na odtok. Zároveň umožňují i extrapolaci tohoto procesu v čase, tj. konstrukci předpovědních modelů.

2. Umělé neuronové sítě (UNS)

Umělá neuronová síť je výpočetní model sestavený na základě abstrakce vlastností biologických nervových systémů. Základní stavební jednotkou UNS je tedy podobně jako u sítě biologické neuron, resp. jeho matematický model – umělý neuron. Hlavní funkcí umělého neuronu, je transformace vstupních signálů do výstupního a jeho distribuce dalším neuronům.



Obr. 1*:* Schéma vícevrstvé perceptronové sítě

2.1 Učení UNS

K základním vlastnostem neuronových sítí je jejich schopnost učit se. Učení (trénování) UNS je analogický termín pro kalibraci koncepčních modelů. Jde o dynamický proces, při kterém dochází k postupné modifikaci souboru parametrů sítě podle zvoleného učícího algoritmu. Cílem tréninkového procesu je nalézt takový soubor parametrů, který nejlépe odpovídá předkládaným vzorům. V případě predikce časových řad předkládáme neuronové síti data z minulosti s nadějí, že objeví skryté závislosti a bude schopna jich využít pro předpovídání do budoucnosti.

Síť je naučena pokud je splněna předem zadaná podmínka (např. konečný počet iterací, pokles celkové chyby pod předem stanovenou mez atp.). Kriteriem může být např. střední kvadratická chyba, směrodatná odchylka, koeficient determinace apod.

Algoritmů učení je celá řada. V našem případě jsme pro učení dopředné perceptronové sítě použili algoritmus zpětného šíření chyby (BPG - Back-propagation of Gradient Algorythm). Princip tohoto algoritmu učení lze zjednodušeně popsat následovně. Do sítě postupně vstupuje vektor vstupních parametrů. Pro každý vzor je a na základě nastavení vah proveden výpočet. Výsledek je porovnán s požadovanou hodnotou a je vypočtena chyba vzoru. Následně pak probíhá zpětně (tj. ve směru negativního gradientu chybové funkce) postupný výpočet korekcí jednotlivých vah. Do opravené sítě je znovu přiveden vstupní vektor a proces se opakuje. Jde o iterativní proces s cílem nalézt minimum celkové chybové funkce.

2.2. Princip tvorby modelu UNS

Učení NS začíná vytvořením vhodné trénovací množiny vzorů tj. množinu vstupních a korespondujících požadovaných (výstupních) hodnot veličin. Na jejich kvalitě a dostatečném množství závisí úspěšnost řešení daného problému. Praxe ukazuje že je vhodné vstupní časové řady podrobit statistické analýze a použít některé optimalizační metody pro získání vhodných trénovacích dat. Část vytvořené tréninkové množiny ponecháme pro testování, tzn. nepředkládáme tuto část časové řady při učení, ale použijeme ji pro testování naučenosti sítě, schopnosti generalizovat. Kontrolujeme tak kvalitu sítě na jiných datech, než které byly síti předkládány při učení.

Někdy bývá pro přesnější predikci vhodné dodat i doplňkové informace, které by mohly mít vazbu na předpovídanou veličinu. V případě našich experimentů na Ploučnici jsme např. k časovým řadám srážek a průtoků přidali agregovanou hodnotu ukazatele předchozích srážek (API).

Další fází je volba topologie sítě, tedy určení počtu skrytých vrstev a počtu neuronu v těchto vrstvách. Úkolem skrytých vrstev obecně je zvýšení aproximačních vlastností UNS jako celku. Stanovení počtu skrytých vrstev a neuronů je věcí každého řešitele.

Pro správný chod koncepčních hydrologických modelů bývá důležité i nastavení počátečních hodnot parametrů. Analogicky u UNS může ovlivnit výsledek trénovacího procesu nastavení

synaptických vah na počátku učení. Váhy mohou být nastaveny např. explicitně, ale často se používá jejich náhodné nastavení. Podobně jako vhodné stanovení topologie, je i počáteční nastavení synaptických vah věcí experimentální.

3. Predikce průtoku na řece Ploučnici – experimentální část

Experimenty s predikcí časové řady průtoků pomocí modelu umělých neuronových sítí byly zkoušeny na povodí horní Ploučnice po profil limnigrafické stanice Mimoň, která sloužila jako předpovědní profil. V povodí o rozloze 268 km^2 jsou umístěny dvě srážkoměrné stanice Stráž pod Ralskem a Jablonné v Podještědí (viz. Obr. 2) Datové sety poskytnuté Českým hydrometeorologickým ústavem v Ústí nad Labem zahrnovaly hodinová data průtoků ze stanice Mimoň, dále srážková a teplotní data ze stanic Stráž a Jablonné za období dvou let (2006–2007). Data byla rozdělena na dvě části, 90 % tvořila trénovací a 10 % testovací množina.

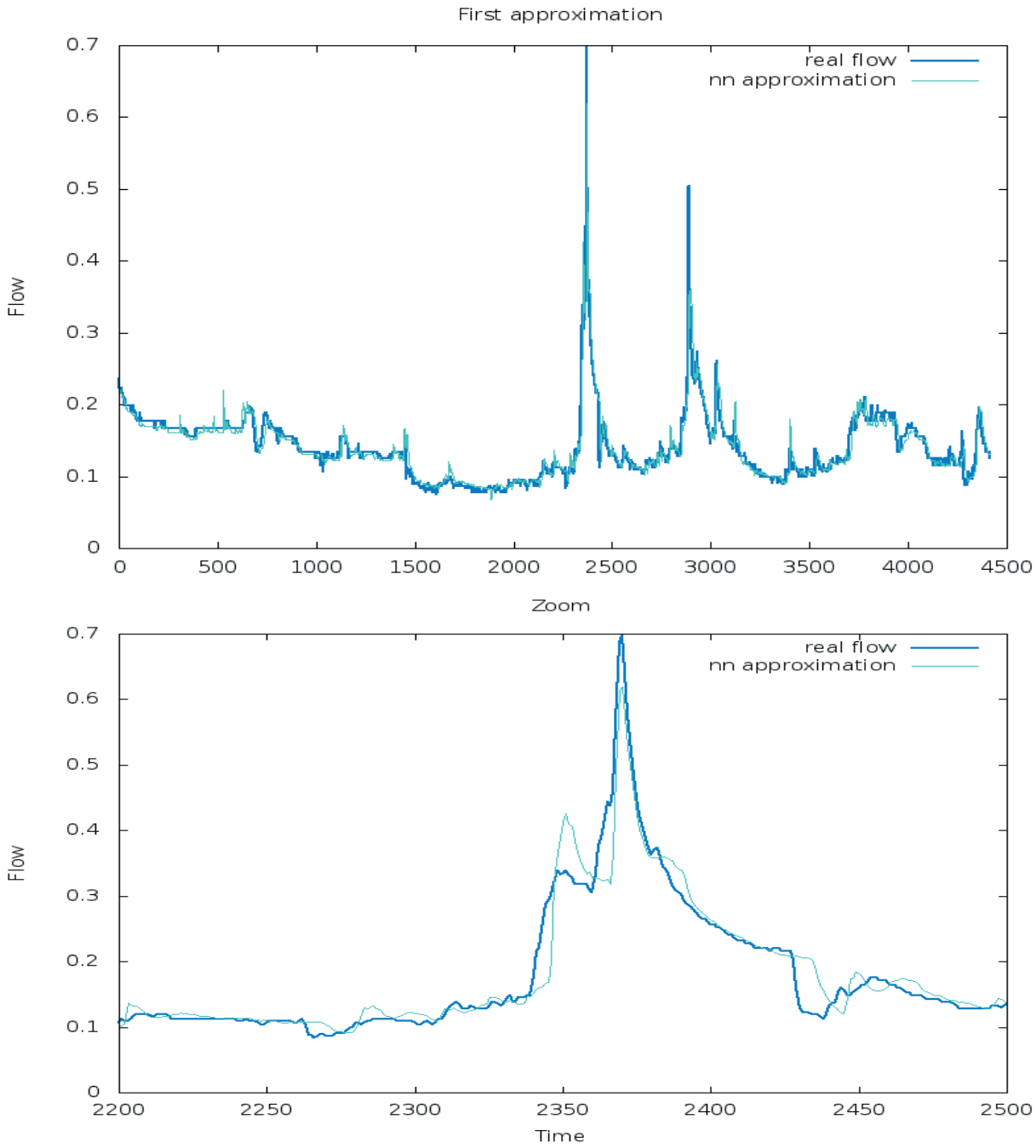


Obr. 2: Zájmová oblast povodí Ploučnice

Při těchto experimentech jsme počítali s 1, 2 a 3 hodinovou historií dat. Predikci průtoku jsme zpočátku počítali na *t+1* a *t+2* hodiny. Byla použita vícevrstvá dopředná síť typu perceptron s algoritmem učení BPG a jeho modifikace - gradientní učení s momentem. Byly zkoušeny různé topologie sítí s jednou či dvěma skrytými vrstvami a s různým počtem neuronů v těchto vrstvách (5-10 v první a 2-4 v druhé skryté vrstvě). Výsledky experimentů na trénovacích a testovacích datech shrnují následující tabulky a grafy.

Tab 1: Výsledky experimentů na trénovacích datech

Network architecture	Correlation coefficient	Mean absolute error	Root mean squared error	Relative absolute error %	Root relative squared error %
1 hour history					
3;5;1	0.9991	0.0394	0.0931	4.3732	4.5065
3;10;1	0.9992	0.0428	0.0938	4.759	4.5363
3;5;2;1	0.9991	0.0394	0.0931	4.3732	4.5065
2 hour history					
6;5;1	0.9997	0.0246	0.0575	2.9193	3.1496
6;10;1	0.9997	0.0253	0.0562	3.0022	3.0761
6;5;2;1	0.9997	0.0263	0.0597	3.1195	3.2687
6;10;4;1	0.9996	0.0267	0.0629	3.1719	3.4446
3 hour history					
9;5;1	0.9997	0.0629	0.0748	7.1259	3.7007
9;10;1	0.9997	0.0581	0.0703	6.5807	3.4809
9;5;2;1	0.9997	0.0644	0.0766	7.2986	3.7906
9;10;4;1	0.9996	0.0682	0.0817	7.7289	4.0442



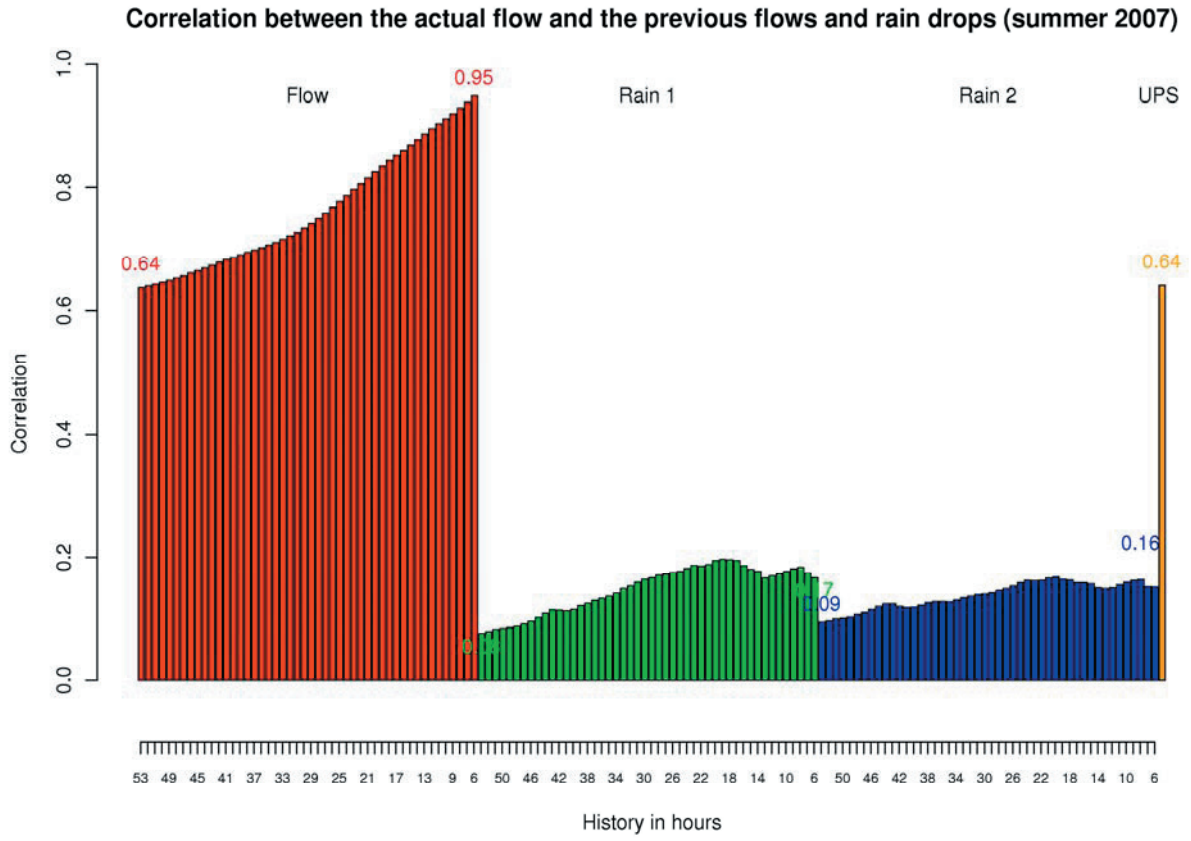
Obr. 3: Grafické porovnání předpovídaných a měřených průtoků *t+2*. Hodnoty průtoků v grafu jsou transformovány do intervalu <0,1>

Porovnáme-li absolutní chyby na trénovacích datech, nejlepší výsledky dosáhly sítě s 2hodinovou historií. Naopak velké sítě vykazují větší chybu na testovacích datech, což může ukazovat na nižší schopnost generalizace. Také se prokázalo, že sítě s 3hodinovou historií vyžadují více času na učení, jelikož obsahují více parametrů, které se musí stanovit během učení. Nicméně všechny sítě úspěšně zvládly proces učení u této úlohy.

V další fázi jsme vyzkoušeli předpovídat na delší časový úsek dopředu. Predikce na *t+6* hodin byla testována na jednoduchém lineárním prediktoru, který sloužil pro nastavení měřítka úspěšnosti metod, jednovrstvém perceptronu a vícevrstvé neuronové sítě. Kritériem byl v tomto případě koeficient efektivity (EC). Výsledné hodnoty koeficientu ukazují, že

vícevrstvý perceptron (0,94) je úspěšnější než jednoduchý perceptron (0,91), který je dokonce horší než lineární predikce (0,92). Dá se říci, že výsledky se od sebe neliší výrazně.

Poslední experiment sloužil k získání informací o tom, jaké historické údaje jsou pro predikci důležité. Míru důležitosti jsme zjišťovali korelací mezi výstupní hodnotou průtoku v čase *t+6* a historickými hodnotami průtoku a srážek a ukazatele UPS. Výsledky tohoto zkoumání ukazuje graf na Obr. 4. Je vidět, že velké míry korelace dosahují předchozí hodnoty průtoku, přičemž míra závislosti klesá s počtem hodin do historie (od 0,95 do 0,64). Podobnou míru závislosti 0,64 má i agregovaný ukazatel předchozích srážek (UPS). Oproti tomu je míra korelace průtoku na srážková historická data poměrně malá (kolem 0,16). Zajímavé je, že korelace je největší pro hodnoty srážek kolem časů *t-18*. Nakonec dodejme, že korelace postihuje především lineární závislost veličin, přičemž problém srážko-odtokových vztahů je nelineární.



Obr. 4: Grafické znázornění korelace mezi predikovanou hodnotou na *t+6* a historickými vstupy průtoků a srážek

4. Genetické algoritmy

Pro výběr relevantních hodnot vstupů byly použity i genetické algoritmy (GA). Genetické algoritmy jsou vyhledávací algoritmy založené na základních principech evoluční biologie (přirozený výběr, křížení, mutace). Principem je proces postupného vylepšování populace jedinců (v našem případě vstupních hodnot jednotlivých časových řad) opakovanou aplikací genetických operátorů, který vede k výběru takových jedinců, kteří lépe vyhovují stanoveným podmínkám než jedinci původní.

GA byl použit jako meta-učící algoritmus, který vybere relevantní vstupy s ohledem na úspěšnost evolučně optimalizovaných neuronových sítí. GA pracuje s populací binárních vektorů, které definují masku výběru vstupů (1 na daném místě znamená, že vstup je vybrán). Byly použity standardní genetické operátory: dvoubodové křížení, náhodnou bitovou mutaci a turnajovou selekci. Ohodnocovací funkce každého jedince spočívá v několika iteracích postupu, který vytvoří a otestuje odpovídající neuronovou síť. V každém jednotlivém ověření nejprve fixujeme počet skrytých neuronů (náhodně mezi 2 a 32) a pak zvolíme vstupní neurony na základě příslušné bitové masky. Váhy sítě jsou náhodně inicializovány, síť se pak po určený čas učí algoritmem zpětného šíření chyby a nakonec otestuje na testovacích datech. Po několika tisících generacích genetického algoritmu je vybrán nejlepší jedinec – vstupní filtr - který během evoluce vznikl. Celá procedura je poněkud náročná na výpočet a na běžném desktopovém PC znamená několik hodin až dnů výpočtů. Z tohoto důvodu byl pro tyto experimenty použit vícejádrový sálový počítač.

5. Závěr

Jedním z hlavních úkolů v hydrologii je vývoj a aplikace matematických modelů schopných správně simulovat a predikovat reakci povodí na spadlou srážku. Jak již bylo řečeno neuronové sítě jsou výborným nástrojem k rychlému rozpoznání závislosti mezi vstupními a výstupními daty.

Obecně můžeme konstatovat, že použité modely neuronových sítě přinášejí oproti lineárním modelům mírné zlepšení.

Tab. 2: Výsledky experimentů

TABLE 2: THE FINAL EXPERIMENT RESULTS

model	EC	MAE	MSE_{pd}	time
last value estimation	0.938	0.1036	0%	0:00
linear regression	0.937	0.1240	−1%	0:00
multilayer perceptron	0.944	0.1252	10%	0:01
evolved multilayer perceptron	0.947	0.1200	16%	0:18

Tabulka 2 ukazuje porovnání úspěšnosti jednotlivých metod. V prvním řádku je pro porovnání odhad průtoku pouze na základě předcházející hodnoty, následuje lineární regresní model. V třetím řádku je vícevrstevný perceptron s jednou skrytou vrstvou předpovídající relativní odtok, který dával doposud nejlepší výsledky. Použitím vstupních filtrů založených na GA se výsledky zlepšily o dalších přibližně 5 %. Jako kritérium byl opět použit koeficient efektivity (EC). Zlepšení je patrno na procentuálním vyjádřením střední kvadratické chyby (MSE_{pd}).

Z experimentů je patrné, že modelu UNS lze úspěšně použít jako alternativní metodu pro krátkodobou predikce průtoků. Podmínkou je dostatečné množství a rozsah kvalitních vzorů pro trénování sítě a nalezení optimálního modelu UNS. Výsledky také ukazují, že použití optimalizačních metod založených na GA pro výběr relevantních vstupů z časových řad přispívá k úspěšnosti modelu UNS.

V současné době chceme využít zkušenosti z experimentů na Ploučnici pro vyžití modelů neuronových sítí pro předpověď průtoků v „on-line“ režimu jako alternativy k již provozovanému srážko-odtokovému modelovacímu systému Aqualog na povodí Smědé pramenící v Jizerských horách. V tomto povodí dochází k častým povodňovým situacím s rychlým průběhem. Předpokládá se, že vhodný model UNS bude sloužit k předpovědi

průtoků za využití reálných datových řad již provozovaného systému. Domníváme se, že to v některých situacích může zlepšit úspěšnost hydrologických předpovědí. Nasazení dvou koncepcí modelovacích technik na jedno povodí může přinést i zajímavé porovnání z hlediska vhodnosti jejich použití.

Literatura

[1] ASCE Task Committee in Application of Artificial Neural Networks in Hydrology, 2000. Artificial Neural Networks in Hydrology II: Hydrologic Applications. In: *Journal of Hydrologic Engineering.* April 2000, pp 124-137.

[2] NACHÁZEL, K. – STARÝ, M. – ZEZULÁK, J. a kol., 2004. Využití metod umělé inteligence ve vodním hospodářství. Academia, Praha. 318 s. ISBN 80-200-0229-4

[3] PAŠČENKO, P., 2009. Computational intelligence models for hydrological predictions, [Matematicko-fyzikální fakulta UK, magisterská práce]. 63 s.

[4] ŠÍMA, J. – NERUDA, R., 1996. Teoretické otázky neuronových sítí. Matfyzpress, Praha. 390 s. ISBN 5-93108-016-3.

[5] TUČKOVÁ, J., 2003. Úvod do teorie a aplikací umělých neuronových sítí. Vydavatelství ČVÚT, Praha. ISBN 80-01-02800-3.

Kontakty na autora

1 Český hydrometeorologický ústav, pobočka Ústí nad Labem, e-mail: srejber@chmi.cz

2 Ústav informatiky AV ČR, Praha; e-mail: roman.neruda@gmail.com

3 Fakulta životního prostředí, Univerzita J. E. Purkyně, Ústí n. L.; e-mail: martin.neruda@ujep.cz