



**SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE**  
**STAVEBNÁ FAKULTA**  
Katedra vodného hospodárstva krajiny

**Ing. František Cyprich**

Autoreferát dizertačnej práce

**VYUŽITIE KLIMATICKÝCH DATABÁZ PRI ANALÝZE A  
INTERPOLÁCII PLAVENÍN NA DUNAJI**

na získanie akademického titulu: **philosophiae doctor (PhD.)**  
v doktorandskom študijnom programe: **krajinárstvo**  
v študijnom odbore: **poľnohospodárstvo a krajinárstvo**  
Forma štúdia: **denná forma**

Miesto a dátum: **Bratislava 2021**



Dizertačná práca bola vypracovaná v dennej forme doktorandského štúdia na Katedre vodného hospodárstva krajiny Stavebnej fakulty Slovenskej technickej univerzity v Bratislave.

**Predkladateľ:** Ing. František Cyprich  
Stavebná fakulta STU  
Radlinského 11  
810 05 Bratislava

**Školiteľ:** prof. Ing. Milan Čistý, PhD.,  
Katedra vodného hospodárstva krajiny,  
Stavebná fakulta, STU Bratislava  
Radlinského 11, 810 05 Bratislava

**Oponenti:** 1. prof. Ing. Viliam Bárek, PhD.,  
Fakulta záhradníctva a krajinného inžinierstva,  
Slovenská poľnohospodárska univerzita v Nitre,  
Trieda A. Hlinku 2, 949 76 Nitra

2. prof. Ing. Andrej Šoltész, PhD.,  
Stavebná fakulta, STU Bratislava,  
Radlinského 11, 810 05 Bratislava

3. Ing. Katarína Jeneiová, PhD.,  
Špecialista pre hydrológiu, Odbor Kvantita povrchových vôd  
Slovenský hydrometeorologický ústav  
Jeséniova 17, 833 15 Bratislava 37  
katarina.jeneiova@shmu.sk

Autoreferát bol rozoslaný dňa: .....

Obhajoba dizertačnej práce sa bude konať dňa ..... o ..... hod. na Katedre vodného hospodárstva krajiny Stavebnej fakulty Slovenskej technickej univerzity v Bratislave, Radlinského 11, v miestnosti číslo .....

.....  
prof. Ing. Stanislav Unčík, PhD.  
Dekan fakulty

## Obsah

<b>Úvod</b> .....	<b>3</b>
<b>1. Súčasný stav riešenej problematiky</b> .....	<b>4</b>
<b>2. Ciele práce</b> .....	<b>5</b>
<b>3. Metodika práce a metódy skúmania</b> .....	<b>5</b>
3.1. Práca s údajmi .....	5
3.2. Modely a kalibrácia.....	6
<b>4. Výsledky práce a diskusia</b> .....	<b>7</b>
4.1. Štandardná metóda merná krivka .....	7
4.2. Výpočet pomocou údajov z predchádzajúcej stanice.....	8
4.3. Pomocou prietokov .....	9
4.4. Pomocou údajov o klíme.....	13
<b>5. Závery a diskusia</b> .....	<b>15</b>
<b>6. Summary</b> .....	<b>16</b>
<b>7. Zoznam použitej literatúry</b> .....	<b>16</b>
<b>8. Zoznam vlastnej publikačnej činnosti</b> .....	<b>22</b>

## Úvod

Plaveniny sú prirodzenou súčasťou riečnych systémov. Plaveninami nazývame rozličné minerálne a organické častice unášané tečúcou vodou (Bridge, 2003; Fryirs a Brierley, 2013). Plaveniny zohrávajú dôležitú úlohu pri štruktúrovaní krajiny, vytváraní ekologických biotopov a preprave živín (Dean a kol., 2016; Koiter a kol., 2013; Vercruyssen, Grabowski a Rickson, 2017). Napriek tomu, že sú nevyhnutnou súčasťou riečného systému, súvisia tiež s celým radom problémov týkajúcich sa znečistenia, ekologického zhoršovania, záplav a poškodzovania infraštruktúry v čoraz viac zastavanom svete (Bilotta a Brazier, 2008; Horowitz, 2009; Taylor a Owens, 2009). Vznikajú následkom erozívnej činnosti vody, vetra, alebo iných činiteľov buď v povodí, alebo eróziou vo vlastnom koryte toku a v jeho prítokoch. Množstvo a kvalita sedimentov má negatívny vplyv na ekologické spoločenstvá, zvyšujú povodňové riziko a skracujú životnosť vodnej infraštruktúry. Na riešenie je dôležité dôkladné pochopenie zdrojov sedimentu, dráhy a dynamiky transportu a faktorov, ktoré sú základom priestorovej a časovej variability suspendovaných sedimentov v riekach (Vercruyssen, Grabowski a Rickson, 2017).

Vo všeobecnosti sa pevné častice v koryte toku podľa jeho unášacej schopnosti a veľkosti častíc buď usadzujú, alebo sú transportované prúdom. Pokiaľ sa usadia hovoríme o sedimentoch. Podľa veľkosti rozdeľujeme vodou unášané častice na dve skupiny:

- splaveniny: hrubozrnné častice, ktoré voda strháva, valí a posúva po dne koryta,
- plaveniny: jemnozrnné častice, ktoré sú vo vode rozptýlené, tiež sa nazývajú suspendované splaveniny (SHMU, 2019).

Dizertačná práca sa venovala transportu sedimentov na Dunaji, a bola zameraná na suspendovaný sediment. Kvantifikácia plavenín umožňuje vyhodnotiť rozsah prebiehajúcich erózných procesov v povodiach a predikovať jeho následky na vodné diela, nádrže a na prietokové a plavebné pomery v samotnom toku. Určovanie množstva plavenín je zložitou úlohou pretože kontrolujúce procesy tohto javu sú ťažko kvantifikovateľné a predvídateľné.

Transport sedimentov v koryte toku je ovplyvňovaný morfológiou, klímou, využitím územia, pôdnymi pomermi jeho povodia, ďalej hydrologickým režimom toku, vodnými stavbami rôzneho druhu, údržbou koryta, bagrovaním dna, rýchlostnými pomermi prúdenia vody v koryte, atď. Pokles unášacej schopnosti prúdu spôsobuje, že jednotlivé častice sa začnú postupne usadzovať, pričom tento jav je najvýraznejší v nádržiach, pretože pri vstupe do nádrže tok výrazne spomaľuje svoju rýchlosť. Pri tomto procese sa znižuje užitočný objem nádrží, čo predstavuje nemalé škody vo vodnom hospodárstve. Problémy vznikajú pri zmenách rovnováhy medzi eróziou dna a usadzovaním sedimentov, čoho následkom sú zmeny morfológie dna, ktoré môžu mať negatívny vplyv na plavbu, prietokové pomery v sútokoch riek, atď. Akumulácia sedimentu v riečnych korytách je jav, ktorý je ovplyvnený nielen morfológiou koryta, ale aj fyzikálnymi a geografickými charakteristikami, endogénnymi a exogénnymi procesmi prebiehajúcimi v povodí (Barabas a kol., 2017). Vodou nesené sedimenty sú navyše zvyčajne bohaté na živiny a chemikálie, ktoré ovplyvňujú kvalitu vody, vodné prostredie, inžinierske stavby, plavbu a celkovo využiteľný potenciál vody. Je preto potrebné množstvo plavenín v tokoch kvantifikovať a poznať ako sa ich transport vyvíja v čase.

Prietok plavenín a jeho kvantifikácia je preto dôležitou informáciou o životnom prostredí, nakoľko umožňuje vyhodnotiť rozsah prebiehajúcich procesov v povodiach a v samotnom toku. O porozumenie, vyhodnotenie transportu a ukladania sedimentov sa snažia viacerí autori napr. (Nones, 2020; Hajigholizadeh, Melesse a Fuentes; 2018) vzhľadom na množstvo ovplyvňujúcich faktorov. Existuje mnoho vzorcov, ktoré možno použiť na popísanie javov erózie a transportu sedimentov v riečnom systéme, vzorce pritom pracujú na veľmi podobnom spôsobe (Whipple a Tucker, 2002; Varrani a kol., 2019). Pri aplikácii modelov v rôznych typoch riek vzniká veľa rozdielov (Attal a kol., 2008; Franzoia a Nones, 2017), čo naznačuje potrebu použitia experimentov na ich testovanie (Tucker, 2009). Pre efektívny opis prírodného prostredia a formovania riečneho systému musí model správne reprodukovat' povrch v závislosti od faktorov, ako je výška reliéfu, miestne podnebie a litológia povodí (Kirkby, 1987; Di Silvio a Nones, 2014).

Najčastejšie sa pre kvantifikáciu plavenín používa merná krivka sedimentov, ktorá určuje závislosť medzi hlavnou ovplyvňujúcou premennou – prietokom vody v danom profile a prietokom alebo koncentráciou plavenín (rating curve). Závislosť prietokov a plavenín sa touto metódou vyhodnocuje cez funkciu  $SS = a \cdot Q^b$ , kde SS je prietok plavenín (suspended sediment load) a Q je prietok. Logaritmovaním tejto závislosti dostávame  $\log(SS) = b \cdot \log(Q) + a$ , čo je lineárna rovnica, z ktorej možno pomerne jednoducho určiť koeficienty a, b. Merná krivka sedimentov je štatistický typ modelu a tieto koeficienty nemajú priamy fyzikálny význam. Avšak koeficient b do istej miery indikuje transportné charakteristiky toku a vysoké hodnoty koeficientu a zase indikujú prítomnosť intenzívne zvetraných materiálov v povodí, ktoré je možné ľahko prepravovať.

Základom aj najspoľahlivejším spôsobom bilancovania predchádzajúcich alebo odhadu budúcich údajov o kvantite plavenín je využitie záznamov z meraní (Muhammad a kol., 2019). Odber vzoriek plavenín je však pomerne náročný ak sa má naozaj reprezentatívne určovať a vyžaduje skúsených odborníkov. Dôvodmi sú napríklad jeho výrazného rozdiely po profile rieky (Morris a Fan, 1998) a podobne. Navyše je to činnosť časovo náročná a tým aj pomerne nákladná, takže nie vždy sú merania každodenné a následne nie sú z minulosti dostupné súvislé denné údaje vhodné pre ďalšie analýzy (Melesse a kol., 2011; Wang, Lu a Kumm, 2011; Walling a Fang, 2003), čo je aj prípad mostu Lafranconi.

Na simulovanie erózie v povodí a následného prietoku plavenín, splavenín, ich usadzovania a následkov na morfológiu toku boli použité aj matematicko-fyzikálne modely s distribuovanými parametrami (zväčša v podobe rastrových GIS vrstiev), ktoré však vyžadujú v porovnaní s rating curve veľmi veľa podrobných informácií. Rozsah potrebných údajov a zložitosť transportu sedimentov vedú pri praktickom použití k mnohým zjednodušujúcim predpokladom a k rôznej úrovni zjednodušovania vstupov, čoho dôsledkom sú značné neistoty výsledkov. V tejto práci je preto navrhnutý iný prístup, ktorý využíva metódy strojového učenia.

## 1. Súčasný stav riešenej problematiky

V dizertačnej práci uvádzame desiatky autorov venujúcim sa plaveninám v riekach, nádržiach alebo v prímorských oblastiach. Niektorý skúmali použitie mernej krivky plavenín

(rating curve) na odhad koncentrácie plavenín (Hapsari a kol., 2019; Jung a kol., 2020; De Girolamo, Pappagallo a Lo Porto, 2015). Ďalší autori porovnávali viacero druhov výpočtov koncentrácie plavenín (rating curve a umelá neurónová sieť) (Rezapour, Shui a Ahmad, 2010; Boukhrissa a kol., 2013) a veľa autorov používa na výpočet koncentrácie plavenín umelú neurónovú sieť (UNS) (Himanshu a kol., 2017; Yadav a kol., 2018; Kisi a Yaseen, 2019). Sú autori, ktorí skúmajú využitie diaľkových a satelitných meraní koncentrácie plavenín (Onderka a Pekárová, 2008; Onderka a Rodný, 2010; Dethier, Renshaw a Magilligan 2020). Na Slovensku sa venuje plaveninám VUVH, SHMU a autori (SHMU, 2019; Borodajkevyčová, 2015; Sočuvka, 2017; Čuban, 2017). Niektorí autori skúmajú fyzikálno-chemické zloženie sedimentov v nádržiach a vodných dielach (Jurík, Sedmáková a Fuska, 2017; Hucko a kol., 2017; Hiller, Jurkovič a Šutriepka 2010).

## 2. Ciele práce

Cieľom dizertačnej práce bolo nájsť vhodný spôsob na doplnenie alebo interpolovanie nepravidelne meraných plavenín v riekach. V práci sa navrhli rôzne štatistické metódy, metódy strojového učenia a doplnené údaje boli verifikované.

Hlavným cieľom dizertačnej práce bolo zistenie vhodnej metódy na dopočítanie nameraných koncentrácií plavenín vo zvolenom profile rieky Dunaja – most Lafranconi. Vo všeobecnosti sa pre tieto metódy používajú prevažne prietoky. Naším ďalším cieľom bolo posúdenie možnosti zaradenia údajov z klimatických databáz ako sú zrážky a teploty ako vstupné údaje do modelov, nakoľko tieto údaje bývajú celistvé a ľahšie prístupné.

Posledná fáza práce bola zameraná na posúdenie presnosti použitých modelov a rôznu kombináciu údajov. Kombinácie vstupných údajov boli zamerané na zovšeobecnenie získaných poznatkov, pretože nie všetky druhy údajov sú všade dostupné.

## 3. Metodika práce a metódy skúmania

### 3.1. Práca s údajmi

Podľa mnohých autorov spomenutých v predchádzajúcich kapitolách sa na zistenie objemu koncentrácie plavenín používa len prietok. V práci sme spolu s prietokom používali aj klimatické údaje ako sú zrážky a teploty s ktorými sme verili v lepší výsledok.

Medzi najdôležitejšie kroky pri riešení regresných problémov metódami strojového učenia patrí analýza a určitý „inžiniering“ so vstupmi, ktoré budú použité pri riešení daného problému, v našom prípade pri výpočte neznámych koncentrácií plavenín v profile Dunaja Lafranconi. Termín používaný pre túto časť hĺbkovej analýzy dát (data miningového) procesu je feature engineering (inžiniering vstupných údajov) (Kuhn a Johnson, 2019).

Pri predbežnej analýze vstupných údajov sme ukázali, že nemáme problém s chýbajúcimi údajmi (nezávislými premennými). Následne bol vykonaný preprocesing a transformácia údajov, ktorá zahrňovala premenu jednotiek, tvorbu nových premenných ako bolo kombinácia dvoch či viacerých vstupov (priemer zrážok za posledné 3 dni). Niektoré modely nedokážu spracovávať niektoré typy údajov ako napr. textové hodnoty, dni, týždne atď. Preto je bežnou praxou pri týchto modeloch tvorba fiktívnych premenných popisujúcich tieto kategorické premenné zobrazené ako množina binárnych (0 alebo 1) premenných, ktoré už spĺňajú

technické požiadavky algoritmov. Kategorické premenné sú údaje, ktoré ovplyvňujú premenné ale priamo sa s nimi nepočíta, napr. textové hodnoty, dátumy, PSČ atď.

K dispozícii sme mali koncentrácie plavenín nepravidelne merané, Počet meraných denných údajov bol len 17% z 8 rokov. Ďalej sme využili údaje o prietokoch a tiež klimatické údaje (zrážky a teploty) z databázy ECA&D nakoľko išlo o pomerne veľké územie v niekoľkých štátoch (Klein Tank a kol., 2002).

Pri rozsiahlom území povodia Dunaja po most Lafranconi sme použili údaje o zrážkach a teplotách z databázy ECA&D - 258 bodov údajov s denným krokom. Bolo potrebné rozdeliť územie na menšie celky s podobnými zrážkovými a teplotnými pomermi. Údaje o zrážkach mali byť rozdelené homogénne pre odtokové pomery a zrážky. Teploty mali byť rovnako homogénne rozdelené aby nebolo toľko vysoko korelujúcich premenných vo výpočtoch ale menej rovnako reprezentatívne zastúpených údajov. Preto bolo riešené povodie rozdelené na 4 približne rovnako veľké čiastkové medzipovodia v profiloch: Regensburg, Linz, Viedeň a Bratislava. Všetky ECA&D body (údaje o zrážkach a teplotách - časové rady) týchto medzipovodií boli následne podrobené zhlukovej analýze, ktorá bola uskutočnená pomocou k-means algoritmu, pričom počet zhlukov, bol určený pomocou silhouette grafu (Rousseeuw, 1987). V jednotlivých medzipovodiach bolo identifikovaných dva až šesť zhlukov. Silhouette graf zobrazuje počet zhlukov do ktorých sa rozdelili údaje podľa zvolenej súdržnosti a vlastnosti údajov. Vstupné údaje (zrážky a teploty) sú v následných regresných výpočtoch (prezentovaných neskôr) reprezentované stredmi týchto zhlukov, t.j., ich aritmetickým priemerom.

V dizertačnej práci sme testovali vplyv použitých údajov (nezávislých premenných) na modelovanie koncentrácie plavenín pre potreby interpolácie jej neúplného denného časového radu. Úlohou bolo počítať koncentrácie v profile Lafranconi, avšak v záujme získania poznatkov, ktoré by sa dali zovšeobecniť aj na iné profily na Dunaji alebo na iných tokoch boli pri modelovaní aplikované rôzne kombinácie údajov. Predpokladali sme, že v iných profiloch bude dostupnosť údajov variabilná a nebudú vždy k dispozícii tie isté údaje ako v prípade profilu Lafranconi na Dunaji. Celkovo bolo použitých 117 premenných: 6 premenných o koncentráciách, 16 prem. o prietokoch, 43 prem. o zrážkach a 13 premenných o teplotách. Tieto premenné sa alternatívne používali/nepoužívali na výpočet koncentrácie plavenín. V každom výpočte teda neboli použité všetky premenné. Všetky tieto údaje boli časové rady s denným krokom. Aby boli rôzne alternatívy výpočtov, ktoré sa vykonávali porovnateľné, museli sa pre tvorbu modelu (tiež kalibrovanie, trénovanie modelu) používať vždy údaje z tých istých dní. Takýto nakalibrovaný model sa testoval na údajoch, ktoré neboli použité pri jeho tvorbe, a tiež sa pri tom museli pri každej výpočtovej alternatíve použiť údaje z tých istých dní (ale iných než boli dni určené na trénovanie). Bolo preto potrebné rozdeliť 496 dní s údajmi na trénovacie a testovacie, čo bolo uskutočnené náhodným výberom. V oboch súboroch údajov (trénovací a testovací) poznáme merané koncentrácie plavenín v Lafranconi, súbory údajov museli byť relatívne podobné.

### *3.2. Modely a kalibrácia*

Na modelovanie koncentrácií plavenín boli použité rôzne typy štatistických lineárnych a machine learning (ML) modelov (modelov strojového učenia). Boli použité modely: štandardný

- porovnávací model rating curve (merná krivka sedimentov); lineárne modely: LASSO, Best subset regresia a Principal Components Regression; nelineárne modely: Support Vector Machines, Stacking Ensemble h2o Model a Deep Learning Neural Network. Pred samotnými výpočtami sme kalibrovali modely aby sme dosiahli najmenšiu chybovosť údajov. Výsledky z kalibrovania sa porovnávajú s testovacou sadou údajov.

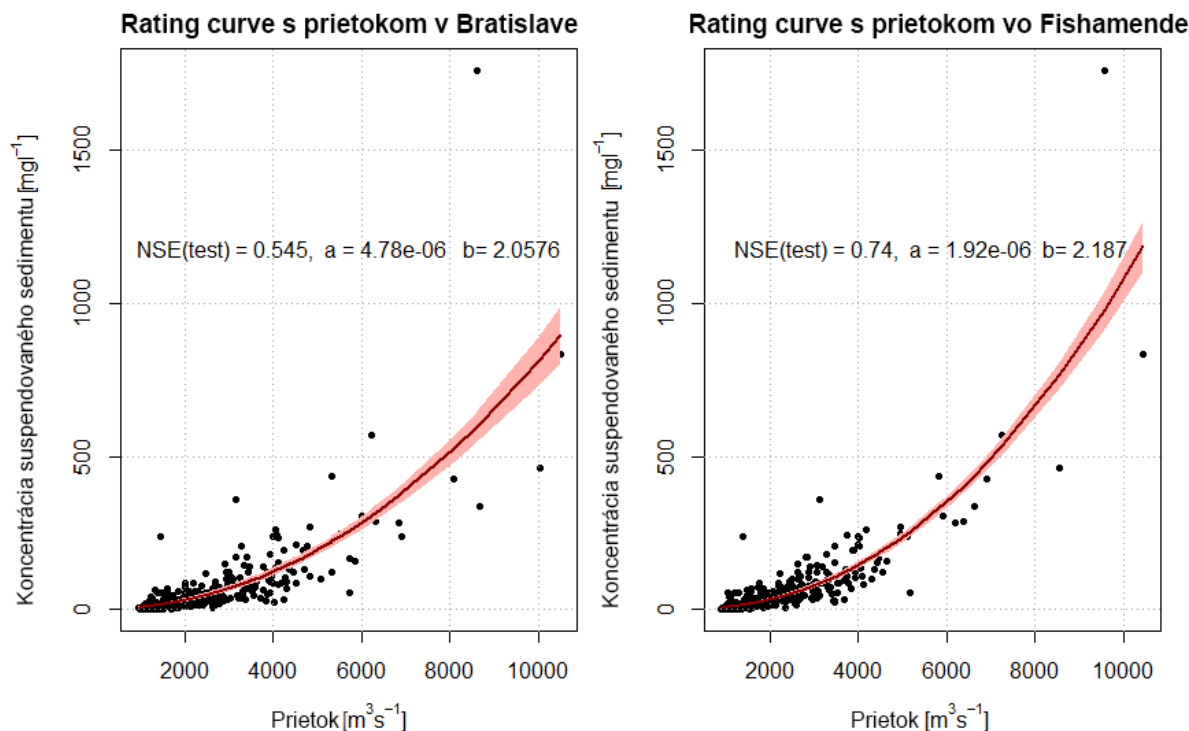
Proces kalibrácie modelovania nie je lineárny postup od začiatku do konca, ale opakovaný postup so spätnou väzbou, ktorý môže viesť pri nedostatku údajov k zmene pôvodne navrhutej metódy modelovania. Rovnako natrénovaný model môže dobre fungovať (s dobrou presnosťou) s tréningovými údajmi, ale s testovacími už nie. Model sa počas kalibrácie postupne zlepšuje a chyba s tréningovými a testovacími údajmi klesá do určitého bodu, ale po prekročení toho istého bodu kalibrovania sa začína zvyšovať. Tento jav sa nazýva „pretrénovaním“. V práci sme použili krížovú validáciu na vyhnutie sa spomenutým problémom pri kalibrácii modelov. V krížovej validácii boli tréningové údaje náhodne rozdelené na  $k$  častí. Výpočet s jednou kombináciou parametrov sa vykonal  $k$  krát a to tak, že vždy sa jedna z  $k$  častí ponechala ako testovacia, model sa natrénoval pomocou zlúčených zvyšných  $k-1$  skupín údajov a pri použití práve testovanej kombinácie parametrov. Pri  $n$  ladených hyperparametroch modelu bola navrhnutá  $n$ -rozmerná matica so sadou predpokladaných hodnôt každého z hľadaných parametrov, ktorá definuje všetky ich možné kombinácie. Najlepšia kombinácia parametrov sa našla preverením všetkých kombinácií na tréningových údajoch pomocou krížovej validácie. Pomocou zvoleného štatistického kritéria pre presnosť výpočtu sa vyhodnotila: 1) presnosť výpočtu z testovacích skupín údajov, čo je aj očakávaná presnosť modelu pri aktuálnej kombinácii parametrov, ale aj 2) sumárna presnosť výpočtu na údajoch z  $k-1$  tréningových skupín. Najlepší výsledok z hľadiska zvoleného štatistického kritéria na testovacích skupinách tréningových údajov indikoval najvhodnejšiu kombináciu parametrov. Vyhodnotenie modelov bolo vždy založené na úspešnosti modelu na testovacích údajoch, t.j. údajoch, ktoré neboli použité pri kalibrácii modelu. Bolo použitých viacero štatistických ukazovateľov.

## 4. Výsledky práce a diskusia

### 4.1. Štandardná metóda merná krivka

Merná krivka (Rating curve) je štandardná metóda pre výpočet koncentrácie alebo prietoku plavenín. Jej vyhodnotenie malo za účel stanoviť minimálnu úroveň presnosti, od ktorej má pre interpoláciu plavenín zmysel použitie zložitejších modelov. Merná krivka bola po výpočte vyhodnotená na grafe na ľavej strane obrázku 1. Ďalšie štatistické indikátory tohto modelu (na testovacích údajoch) sú v tabuľke 1. S týmito hodnotami možno porovnávať úspešnosť ďalších modelov.





Obr. 1. Merná krivka z testovacích údajov.

Tab. 1. Vyhodnotenie mernej krivky na testovacích údajoch.

model	ME	RMSE	PBIAS	NSE	r	R2
merná krivka (prietok Bratislava)	-16.2	140.8	-19.3	0.545	0.817	0.668
merná krivka (prietok Fischamend)	-17.1	106.4	-20.4	0.740	0.902	0.814
mk GAM (prietok Bratislava)	-14.3	134.4	-17.1	0.586	0.833	0.694
mk GAM (prietok Fischamend)	-13.5	101.3	-16.1	0.765	0.918	0.843

#### 4.2. Výpočet pomocou údajov z predchádzajúcej stanice

V prípade interpolácií koncentrácií bola v riečnom profile Lafranconi výhoda v tom, že koncentrácie sú merané aj vyššie v blízkom Hainburgu (15.2 km). Tieto dve merania pochopiteľne vysoko korelujú a regresné modely vyhodnotené v tabuľke 2, používajúce tento údaj poskytujú výsledky s pomerne vysokou presnosťou. Model teda dosiahol takúto presnosť iba na základe dvoch premenných, čo je jeho výhodou. Dostupnosť blízkych údajov o koncentracii ako vstupných údajov je však vo všeobecnosti skôr výnimkou.

Tab. 2. Vyhodnotenie testovania výpočtu koncentrácie z koncentrácií z Hainburgu.

model	ME	MAE	RMSE	PBIAS	NSE	r	R2
Lineárna regresia	-3.273	23.38	84.56	-3.9	0.836	0.936	0.877
Support Vector Regression	-1.128	21.85	78.13	-1.3	0.860	0.937	0.878

### 4.3. Výpočet pomocou prietokov

V predchádzajúcej časti sa na určenie koncentrácie plavenín používal iba jeden prietok, v tejto časti sme uvažovali s viacerými premennými o prietokoch (prietok na rôznych miestach, v rôznom čase). Výpočet koncentrácie plavenín bol realizovaný lineárnymi a machine learning metódami, z ktorých reprezentatívny výber je stručne charakterizovaný v tabuľkách 3 až 6.

Do vstupných údajov o prietokoch boli zahrnuté stanice Hainburg a Fischamed nad skúmaným riečnym profilom, prietoky v Bratislave, v stanici Medveďov pod skúmaným riečnym profilom a v staniciach v Moravskom Jáne a v Záhorskej Vsi na prítoku Moravy do Dunaja. Vo výpočtoch boli použité prietoky z daného dňa a z predchádzajúcich a nasledujúcich dní.

#### Lineárne modely

Bol použitý model LASSO s regularizáciou, nakoľko prietoky ako vysvetľujúce premenné medzi sebou korelujú. Tento model bol pomerne citlivý na to, aby jeho závislá premenná mala normálnu distribúciu, a preto bola normalizovaná (logaritmovaním). V prípade výpočtu bez normalizácie bol výsledok síce lepší než základný výpočet mernej krivky pomocou prietoku v Bratislave, ale horší než výpočet mernej krivky pomocou prietoku v stanici Fischamend. Dôvodom je, že metóda merná krivka je založená na nelineárnom modeli, avšak LASSO model je v podstate viacnásobná lineárna regresia, teda lineárny model. V tabuľke 3 sú výpočty s údajmi zahrnujúcimi interakčné premenné (ktoré veľmi výrazne výsledok neovplyvnili) a s logaritmom závislej premennej, čo výrazne zlepšilo úspešnosť modelu, ako ukazuje hodnota NSE. Pri použití modelu LASSO s regularizáciou bol regularizačný parameter lambda hľadaný pomocou krížovej validácie na tréningových údajoch.

Tab. 3. Vyhodnotenie LASSO regresie pri rôznych modifikáciách údajov.

model	ME	RMSE	PBIAS	NSE	r	R2
LASSO	-10.7	116.1	-12.8	0.691	0.893	0.797
LASSO s interakciami	-14.9	110.8	-17.8	0.719	0.897	0.804
LASSO s logaritmovaním	-52	108.4	-62.3	0.731	0.917	0.841
LASSO s inter. a log.	-17.8	89.7	-21.2	0.815	0.917	0.840

Best subset regresia – ďalšia použitá lineárna metóda - hľadá najvhodnejšiu podmnožinu vysvetľujúcich premenných pre výpočet závislej premennej. Výpočty potvrdili, že najlepšie výsledky v tejto úlohe dosahuje, keď sa pri nej použije interakcia premenných. Pretože v tejto úlohe ide o interpoláciu, v nájdenej rovnici sa uplatnili aj prietoky z nižšie položeného Medveďova z nasledujúceho dňa po predpovedi plavenín. Získaný bol lineárny model s dvoma kombinovanými premennými  $Q_{\text{hainburg}} * Q_{\text{fischamend}}$  a  $Q_{\text{hainburg}} * Q_{\text{medv+1}}$ . Súčin prietokov je pomerne veľké číslo, takže rezíduá tréningového súboru sú od 350 do 527 so strednou hodnotou (medián) rovnou -6.21. Obidva koeficienty boli určené so spoľahlivo nízkou p-hodnotou, prakticky nulovou. Veľkosť koeficientov je vzhľadom na veľké hodnoty súčinov prietokov malá:

$$C_{\text{lafranconi}} = 3.229768\text{E-}05 * Q_{\text{hainburg}} * Q_{\text{fischamend}} - 2.282738\text{E-}05 * Q_{\text{hainburg}} * Q_{\text{medv+1}} + 4.879137$$

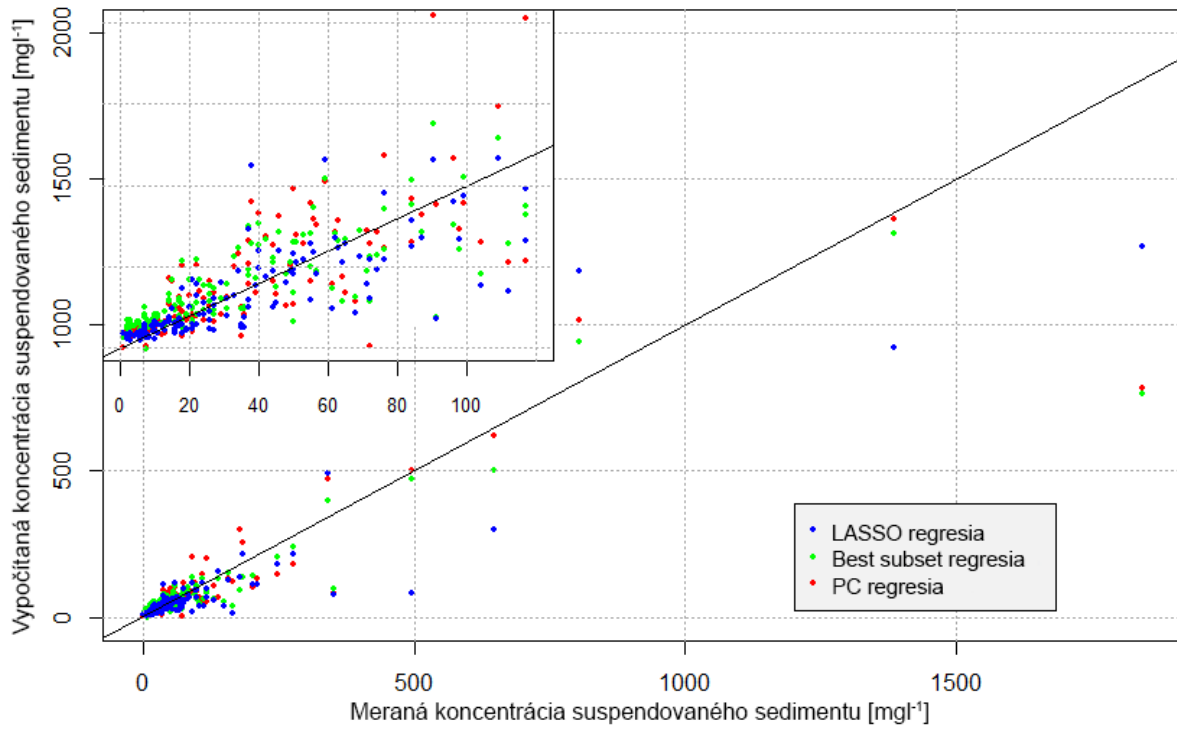
Tento model sme uviedli podrobnejšie so zreteľom na jeho dobrú interpretovateľnosť a jednoduchú aktualizáciu koeficientov v prípade zmeny údajov. Výsledný model je vyhodnotený v tabuľke 4. Je o niečo menej presný než LASSO regresia, avšak je jednoduchší, a preto možno pri jeho použití na nové údaje predpokladať dobrú generalizáciu.

Ďalšou použitou lineárnou metódou bola regresia pomocou principal components (PC). Počet použitých PCR bol optimalizovaný krížovou. Hodnota tohto parametru je 8, takže výsledná rovnica má 8 premenných - principal components namiesto 196 premenných, ktoré pri zahrnutí interakcií tvorili pôvodné vstupné údaje, čo je zrejma výhodou ak nie je na úkor presnosti, výsledky sú v tabuľke 4. Testované boli aj príbuzné metódy independent component regression a partial least square regression, avšak nedosiahli lepšie výsledky, a preto ich neuvádzame. Názvy metód nie sú zaužívané v slovenčine, preto ich uvádzame v angličtine.

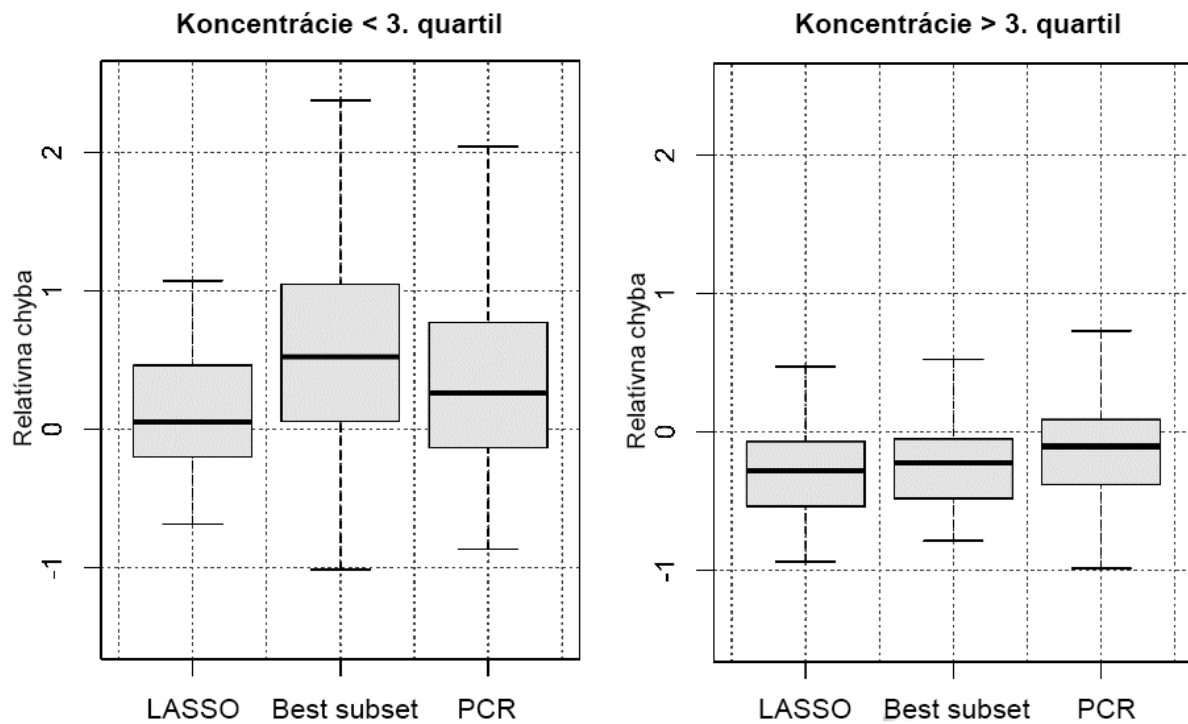
Najlepšie varianty lineárnych modelov sú vyhodnotené v tabuľke 4 a na obrázkoch 2 a 3. Na základe parametrov RMSE, NSE a R2 možno povedať, že najlepšie výsledky dosahoval model LASSO (s interakciami a logaritmovaním počítanej premennej). LASSO však oproti druhým dvom modelom o niečo viac podhodnocuje, čo odrážajú ukazovatele ME, PBIAS. Nedá sa preto povedať, že je jednoznačne najlepší, skôr platí, že modely sú podobne presné. Na obrázku 2 môžeme vidieť, že najlepšie hodnoty nedosahuje vždy model LASSO ale relatívna úspešnosť modelov je variabilná – niektoré hodnoty počíta lepšie LASSO, iné hodnoty iný model. Na obrázku 3 sú vyjadrené relatívne chyby jednotlivých modelov pri predpovedaní menších a väčších koncentrácií. Ako hranica medzi týmito dvoma kategóriami hodnôt koncentrácií bol zvolený tretí kvartil ich meraných hodnôt. Aj tento obrázok poukazuje na variabilitu relatívnej úspešnosti modelov. Pri menších koncentráciách na ľavom obrázku najčastejšia hodnota relatívnej chyby (medián) je najbližšie k nule v prípade modelu LASSO, avšak pri väčších hodnotách koncentrácie (napravo) poukazuje relatívna chyba výpočtu na najväčšiu presnosť modelu PCR. Pri podobnej presnosti týchto modelov možno predpokladať, že najstabilnejšie a najpresnejšie hodnoty získame nejakým typom ensemblevej predpovede, kde sa použijú všetky modely. V našom prípade sme tento krok uskutočnili jednoduchým priemerovaním predpovedí. Tento experiment je tiež vyhodnotený v tabuľke 4. Takto bol získaný model s najlepšimi štatistickými ukazovateľmi presnosti.

Tab. 4. Vyhodnotenie lineárnych modelov.

model	ME	RMSE	PBIAS	NSE	r	R2
LASSO	-17.8	89.7	-21.2	0.815	0.917	0.840
Best subset regresia	-10.6	96.3	-12.7	0.787	0.904	0.818
PC regresia	-6.6	97.3	-7.9	0.783	0.891	0.793
Priemer z 3 modelov	-11.6	87.6	-13.9	0.824	0.925	0.856



Obr. 2. Porovnanie úspešnosti lineárnych modelov na testovacích údajoch.



Obr. 3. Porovnanie relatívnej chyby na menších údajoch (< 3rd kvartil) a väčších údajoch

## Nelineárne modely

V tejto časti je popis výsledkov modelov Support Vector Regression (SVR), gradient boosting algoritmus CatBoost využívajúci regresné stromy a Deep Learning Neural Network (DLNN).

Pri výpočte SVR sa optimalizovala voľba kernelu, jeho parametrov a hodnota C (cost of constraints violation). Porovnaný bol výpočet SVR s radial basis kernel a s lineárnym kernelom. Tento model preukazoval pri rôznom feature engineeringu (normalizácia počítanej premennej, začlenenie vysvetľujúcich premenných zohľadňujúcich interakcie) rozdielnu presnosť. Výpočty sú vyhodnotené v tabuľke 5, a ako ukazuje hodnota NSE (primárny indikátor presnosti), uvedené manipulácie so vstupnými údajmi zvýšili presnosť výpočtu koncentrácie plavenín. Optimálne parametre modelov boli nájdené pomocou krížovej validácie pri použití trénovacích údajov.

CatBoost pri nižšom počte iterácií ovplyvňuje najviac miera učenia; pri vyššom počte iterácií sa tento vplyv redukuje. Ladenie tohto modelu je o niečo zložitejšie než predchádzajúceho, keďže má viac parametrov. Výsledný model je vyhodnotený pomocou testovacích údajov v tabuľke 6. Model CatBoost nevykazoval lepšie výsledky pri použití interakcií alebo pri normalizácii pomocou logaritmovania.

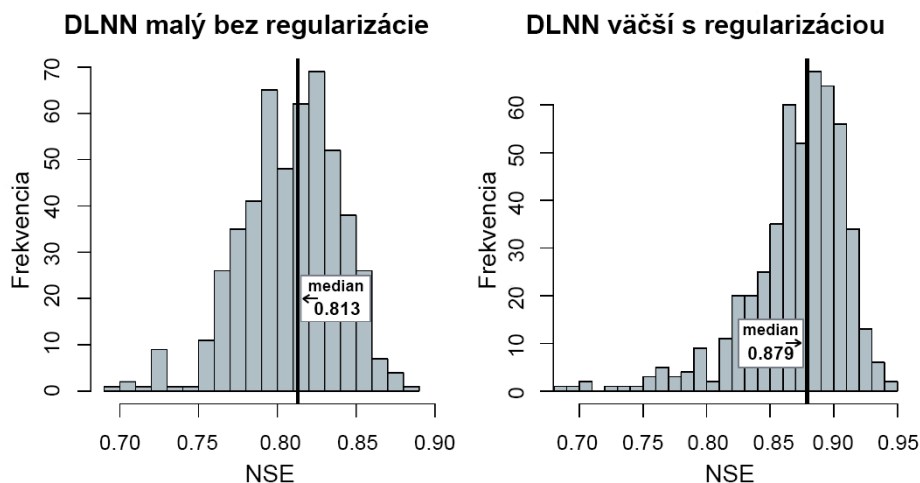
Pri výpočtoch pomocou Deep Learning Neural Net (DLNN) bola použitá platforma h2o. Chceli sme preveriť, či tento model nie je vzhľadom na množstvo použitých údajov príliš silný. Ako je známe efektívnosť deep learning sa prejavuje až pri väčších objemoch údajov, aké sa vyskytujú v oblastiach ako computer vision, rozoznávanie reči alebo pri riadení zložitých procesov. Dátový súbor používaný v tejto práci je menší, avšak zámerom bolo testovať aplikáciu DLNN, nakoľko to, že ich architektúra pozostáva z väčšieho množstva vrstiev a neurónov (ktoré sa pri menších dátových súboroch neuplatnia), nie je ich jediným vylepšením oproti klasickým viacvrstvovým perceptrómom z 90-tych rokov. To znamená, že nie je nutné používať veľké neurónové siete a predsa možno využiť výhody tejto vetvy vývoja v AI. Pri výpočtoch pomocou DLNN boli testované dve alternatívy – v prvej bola navrhnutá relatívne malá sieť (jedna skrytá vrstva s 5-15 neurónmi) a v druhej zložitejšia, s niekoľkými vrstvami a s väčším množstvom neurónov. Na rozdiel od prvej alternatívy, kde je model „regularizovaný“ iba svojou jednoduchosťou ale nevyužíva moderné regularizačné metódy ako L1 a L2 regularizáciu alebo dropout, v druhej alternatíve sme tieto inovácie deep learning použili.

Platforma h2o, ktorú sme pre výpočty používali, neumožňuje pri paralelizácii výpočtu stabilizovať generovanie náhodných čísiel vo všetkých clustroch (najmä v štádiu inicializácie výpočtu). Z toho dôvodu pri následných výpočtoch s tými istými nastaveniami možno dostať rôzne výsledky. Obidve alternatívy boli preto spustené 500 krát a ako výsledný výpočet prezentujeme ten, ktorý poskytol výsledky, ktorých Nash Sutcliffe Efficiency (hlavný ukazovateľ presnosti modelu) je medián NSE hodnôt z týchto 500 výpočtov. Na obrázku 4 sú histogramy týchto 500 NSE z oboch alternatív architektúry neurónovej siete. Model využívajúci regularizáciu ma vyššiu hodnotu mediánu a táto vyššia hodnota je aj stabilnejšia, ako možno usúdiť zo zošíkmenia histogramu doľava. Menšia neurónová sieť obsahovala jednu skrytú vrstvu s jedenástimi neurónmi, pričom aktivačnou funkciou neurónu bol rectifier. Počet

neurónov a aktivačná funkcia sa zistila pomocou krížovej validácie. Pri väčšej, regularizovanej sieti boli štyri skryté vrstvy s počtom neurónov 50, 25, 15, 10 a ako aktivačná funkcia bol použitý rectifier. Regularizácia sa realizovala pomocou input dropout a pomocou L1-regularizácie. Pri tréňovaní oboch modelov bola použitá 10-násobná krížová validácia. Tento výpočtový experiment bol pomerne náročný na výpočtový čas, na počítači AMD Ryzen Threadripper 1950X 16-Core Processor trval skoro 9 hodín.

Tab. 5. Vyhodnotenie Support Vector Regression pri rôznych parametroch a modifikáciách údajov.

model	ME	RMSE	PBIAS	NSE	r	R2
SVR linárny kernel	-20.1	130.3	-24.0	0.611	0.887	0.787
SVR radial basis kernel	-17.9	114.2	-21.3	0.701	0.897	0.805
SVR s logaritmovaním	-15.3	108.1	-18.3	0.732	0.859	0.738
SVR s log. a inter.	-15.5	90.2	-18.6	0.813	0.914	0.835



Obr. 4. Nash Sutcliffe efficiency vyjadrujúca presnosť modelu na testovacích údajoch 500 opakovaných výpočtov.

Tab. 6. Vyhodnotenie ML metód.

model	ME	RMSE	PBIAS	NSE	r	R2
SVR s log. a inter.	-15.5	90.2	-18.6	0.813	0.914	0.835
CatBoost	-12.5	90.3	-15.0	0.813	0.943	0.890
DLNN – malý	-6.9	90.3	-8.3	0.813	0.902	0.814
DLNN – väčší s regularizáciou	-8.2	72.6	-9.8	0.879	0.939	0.882

#### 4.4. Výpočet pomocou údajov o klíme

Pri predikcii plavenín sa štandardne používajú ako vstupné údaje prietoky. V práci bolo skúmané aj riešenie tejto úlohy aj pomocou klimatických údajov – zrážok a teplôt. Tieto veličiny sú základnými činiteľmi hydrologického cyklu a rozhodujú o množstve odtečenej vody z povodia a tým súčasne aj o odnose sedimentov a množstve plavenín v toku. Nie sú však takými priamymi činiteľmi ako prietok vody, ktorá plaveniny obsahuje a nesie, a preto sa pri

použití meteorologických vstupov predpokladá nižšia presnosť výpočtov. Vo svetovej literatúre týkajúcej sa modelovania koncentrácie plavenín sa klimatické údaje ako jediný typ premennej pre výpočet plavenín používajú zriedkavo (t.j., bez použitia hydrologických premenných – prietokov). V práci je preukázaná relatívne dobrá presnosť, ktorá sa môže pomocou údajov tohto typu dosiahnuť. V niektorých miestach môže byť výpočet pomocou klimatických údajov veľmi výhodný a to najmä tam, kde nie sú merania prietokov k dispozícii.

Použité klimatické údaje tvorí 74 časových radov denných hodnôt súčtu topenia snehu a tekutých zrážok. V teplom období roku, keď nedochádza ku sneženiu je hodnota tejto súčtovej premennej rovná zrážkam. Do výpočtov boli začlenené aj interaktívne premenné, nakoľko je zrejmé, že súčasný dážď (alebo topenie snehu) na rôznych miestach povodia, prípadne s rozličným časovaním pred meraním plavenín, môže mať významný synergický účinok. Interakcie neboli zadané ako vstupné údaje len pri algoritmoch CatBoost a h2o neural net, nakoľko tieto algoritmy majú pre postihnutie interakcií svoje vlastné mechanizmy.

Vyhodnotenie výsledkov jednotlivých algoritmov sa nachádza v tabuľke 7. LASSO regularizácia lineárnej regresie bola použitá vzhľadom na koreláciu klimatických premenných v geograficky blízkyh lokalitách povodia. Best subset regresia dosahovala najlepšie výsledky, pri použití interakčných premenných. Získaný lineárny model má sedem kombinovaných premenných, všetky koeficienty boli určené so spoľahlivo nízkou p-hodnotou, prakticky nulovou. Model CatBoost nevykazoval pri použití týchto údajov lepšie výsledky ako lineárne modely. Pri výpočtoch pomocou Deep Learning Neural Net (DLNN) h2o boli podobne ako v prípade výpočtov pomocou prietokov testované dve alternatívy. Výsledky prezentované štatistickými ukazovateľmi v tabuľke 7 sa týkajú výpočtu, ktorého hodnota Nash-Sutcliffe Eficiency bola mediánom z 500 výpočtov. Pri trénovaní oboch modelov bola použitá 10-skupinová krížová validácia.

Pri výpočte SVR sa optimalizovala voľba kernelu, jeho parametrov a hodnota C (cost of constraints violation). Porovnaný bol výpočet SVR s kernelom radial basis a s lineárnym kernelom. SVR preukazoval lepšie výsledky pri normalizácii počítanej premennej (logaritmovaním) a pri využití interakcie premenných. Vzhľadom na väčšie množstvo vstupných údajov ako pri výpočte koncentrácií plavenín pomocou prietokov, bol vykonaný výpočtový experiment s redukovaním a kombinovaním vstupných hodnôt pomocou principal components (ktoré sú lineárnou kombináciou pôvodných premenných, takže môžu tiež v určitom zmysle postihovať interakcie premenných).

Tab. 7. Vyhodnotenie testovania výpočtu koncentrácie na základe premennej zrážky plus odtok zo snehu.

model	ME	RMSE	PBIAS	NSE	r	R2
<b>LASSO s interakciami</b>	-5.9	103.7	-7.1	0.753	0.926	0.858
<b>Best subset s interakciami</b>	-4.7	100.1	-5.6	0.770	0.909	0.827
<b>DLNN h2o malý</b>	-7.2	127.4	-8.6	0.628	0.814	0.662
<b>DLNN – väčší s regul.</b>	1.4	100.9	1.7	0.766	0.894	0.800
<b>SVR s principal components</b>	-22.6	92.8	-27.0	0.803	0.940	0.883
<b>SVR s interakciami</b>	-16.7	96.9	-19.9	0.785	0.912	0.831



## 5. Závery a diskusia

V dizertačnej práci boli študované možnosti modelovania koncentrácie plavenín s cieľom ich interpolácie, t.j. pre ich stanovenie v dňoch, v ktorých neboli merania vykonávané. Sady meraní týchto údajov s nepravidelným časovým krokom sú časté najmä z minulosti vzhľadom na náročnosť merania tejto veličiny. Kompletné údaje umožňujú následne ďalšie analýzy potrebné v manažmente vodných tokov. Ako riešené územie pre verifikáciu výsledkov bola zvolená rieka Dunaj v Bratislave. V práci bolo potvrdené, že výpočet koncentrácie plavenín je možné v porovnaní so štandardnou metódou „mernou krivkou“ veľmi zásadne upresniť pri použití rôznych zdrojov údajov a metód strojového učenia. Najpresnejšie výsledky poskytujú výpočty, ktoré používajú merania koncentrácie plavenín z blízkej stanice Hainburg. Cieľom autorov však bola práca, ktorá poskytne všeobecnejšie poznatky a taká situácia ako je blízkosť meracích miest plavenín Hainburg a Lafranconi je zriedkavá. Z tohto dôvodu boli výpočty vykonané tiež na základe údajov o prietokoch, ale aj pomocou klimatických údajov, ktoré sú zvyčajne najľahšie dostupné. Keďže ide o regresný problém, je potrebné použiť modely kalibrovať, a to na meraných údajoch o koncentráciách plavenín. Pri voľbe dní, v ktorých sa vykonávali merania, sa prihliadalo na to, aby sa meralo hlavne vtedy, keď sú väčšie prietoky, čo prispelo k tomu, že trénovacia a testovacia sada údajov pre modelovanie je reprezentatívna, keďže distribúcia rozdelenia denných koncentrácií plavenín je výrazne zošikmená doprava. Toto znamená, že koncentrácií s malými hodnotami je nadbytok (a tieto boli práve merané menej často, avšak dostatočne reprezentatívne, čo de facto viedlo k vhodnému smplovaniu údajov). Pri teoretickom zaradení príliš veľkého množstva údajov o malých koncentráciách v obdobiach nižších prietokov do výpočtov by tieto dominovali pri regresných výpočtoch a pravdepodobne aj znižovali presnosť výpočtu väčších hodnôt (Čistý a Bezák, 2013).

Presnosť jednotlivých výpočtov je podrobnejšie uvedená v tabuľkách v štvrtej kapitole pomocou viacerých ukazovateľov. Tu zosumarizujeme výsledky použitím Nash Sutcliffe-Efficiency, aký sa určil pri výpočte testovacej sady údajov, keďže sa tento koeficient často používa v hydrologických prácach. Štandardný model rating curve  $NSE = 0.545$  (nelinárna regresia), najlepší výpočet s použitím údajov o viacerých prietokoch nad aj pod bratislavským riečnym profilom, kde sa merajú koncentrácie  $NSE = 0.879$  (DLNN), najlepší výsledok pri použití koncentrácie plavení v blízkom Hainburgu  $NSE = 0.860$  a najlepší model s výhradne klimatickými údajmi  $NSE = 0.803$  bol Support Vector Regression v ktorom sa vstupný súbor klimatických údajov transformoval pomocou principal components analysis. Tento posledný výsledok považujeme za významný príspevok k možnostiam výpočtu plavenín. V ďalšom výskume by bolo vhodné testovať, či dobrý výsledok s klimatickými údajmi možno použiť aj v menších povodiach, než bolo skúmané povodie.

Ide teda o významný nárast presnosti výpočtov voči štandardnej mernej krivke. Najúspešnejším modelom bol Deep Learning Neural Networks, pri ktorom bolo zistené, že je výhodnejšie zostaviť väčšiu sieť s viacerými vrstvami a zároveň ju regularizovať, než použiť malú sieť s jednou vrstvou a menším počtom neurónov bez regulovania, ktorá by na prvý pohľad bola vhodnejšia z pohľadu veľkosti použitého dátového súboru. To je dobrý signál týkajúci sa toho, že vývoj v oblasti Deep Learning možno zúžitkovať aj v prípade výpočtov so stredne veľkými dátovými súbormi. Dobré a konzistentné výsledky boli aj pomocou ďalších



lineárnych a nelineárnych metód, napríklad pomocou Support Vector Regression s redukovaním vstupných premenných pomocou Principal Component Analysis a tiež dobré výsledky sa dosiahli lineárnou metódou LASSO, pri ktorej sa aplikovala interakcia premenných a normalizácia cieľovej premennej. Významným dôvodom prekvapujúcej presnosti výpočtov využívajúcich iba klimatické údaje je v tejto práci navrhnutý spôsob feature engineering (zhlukovanie klimatických údajov) zakomponovanie interakcie premenných a doménných vedomostí pri transformácii teplôt a zrážok na súčtovú premennú topiaceho sa snehu a tekutých zrážok.

## 6. Summary

Measured records of vessel concentrations are important to support water management activities. The dissertation thesis was focused on the possibility of the vessel concentrations modeling for their determination in periods when measurements were not performed. The profile of the Danube in Bratislava was chosen as a case study. The thesis compared regressions using the operator of the smallest absolute selection and shrinkage, support for vector regression and a neural network with deep learning in solving this problem using different data sources. The obtained results showed a significant increase in the accuracy of modeling of vessel concentrations compared to the standard method, which is the measuring curve. Several variables were used to determine the concentration of vessels, as the same data as in this work may not be available everywhere.

In particular, the use of climatic (precipitation and temperature) and hydrological inputs (flows) was considered to support the more general benefits of the thesis. The thesis proposed an original method of modifying the input data on the climate, which significantly increased the accuracy of modeling. The author demonstrated that using the proposed methodology, the use of climate data, which are usually better available than hydrological data, would result in a comparable degree of accuracy than standard modeling based on river flow data.

## 7. Zoznam použitej literatúry

ATTAL, M. et al. 2008. Modeling fluvial incision and transient landscape evolution: Influence of dynamic channel adjustment. In *Journal of geophysical research Earth Surface* [online], vol. 113, no. F3, pp. F03013 [cit. 2021-03-27]. ISSN 2169-9011. Dostupné na: <<https://doi.org/10.1029/2007JF000893>>.

BARABAS, Dušan - BÓNA, Ján - KLEIN, Daniel – Balážovičová, Lenka. 2017. Morphometric and geological conditions for sediment accumulation in the Udava River, Outer Carpathians, Slovakia. In *Journal of Geographical Sciences* [online], vol. 27, pp. 981–998 [cit. 2021-03-29]. ISSN 1861-9568. Dostupné na: <<https://doi.org/10.1007/s11442-017-1416-2>>.

BILOTTA, G. S. – BRAZIER, R. E. 2008. Understanding the influence of suspended solids on water quality and aquatic biota. In *Water Research* [online], vol. 42, no. 12, pp. 2849-2861

[cit. 2021-04-05]. ISSN 0043-1354. Dostupné na: <<https://doi.org/10.1016/j.watres.2008.03.018>>.

BORODAJKEVYČOVÁ, M. 2015. Evaluation of the regime of suspended sediments on the Slovak section of the Danube. In *Hydrological Days 2015, Slovak Republic, 6.-8. Oct 2015*. Bratislava: SAV.

BOUKHRISA, Z. A. et al. 2013. Prediction of sediment load by sediment rating curve and neural network (ANN) in El Kebir catchment, Algeria. In *Journal of Earth System Science* [online], vol. 122, no. 5, pp. 1303-1312 [cit. 2021-02-22]. ISSN 0973-774X. Dostupné na: <<https://link.springer.com/content/pdf/10.1007/s12040-013-0347-2.pdf>>.

BRIDGE, John S. 2003. Rivers and Floodplains: Forms, Processes, and Sedimentary Record. Oxford: Wiley-Blackwell. 504 s. ISBN 978-0-632-06489-2.

ČISTÝ, Milan - BEZÁK, Juraj. 2013. The impact of sampling on flow prediction using support vector machines. In *Proceedings of the third international conference on soft computing technology in civil, structural and environmental engineering*. Stirling, United Kingdom: Civil-Comp Press, s. 22. ISBN 978-1-905088-58-4.

ČUBAN, R. 2017. Sedimenty vo vzťahu k zmene objemu VN Krpeľany. In *zborník prednášok z viii. Konferencie s medzinárodnou účasťou, sedimenty vodných tokov a nádrží*. Banská Bystrica: DALI-BB, s.r.o., s. 139-148. ISBN 978-80-89740-13-0.

DE GIROLAMO, A. M. - PAPPAGALLO, G. - LO PORTO, A. 2015. Temporal variability of suspended sediment transport and rating curves in a Mediterranean river basin: The Celone (SE Italy). In *CATENA* [online], vol. 128, pp. 135-143 [cit. 2021-03-06]. ISSN 0341-8162. Dostupné na: <<https://doi.org/10.1016/j.catena.2014.09.020> DOI>.

DEAN, David J. et al. 2016. Sediment supply versus local hydraulic controls on sediment transport and storage in a river with large sediment loads. In *Journal of geophysical research Earth Surface* [online], vol. 121, no. 1, pp. 82-110 [cit. 2021-03-25]. ISSN 2169-9011. Dostupné na: <<https://doi.org/10.1002/2015JF003436>>.

DETHIER, E. N. - RENSHAW, C. E. - MAGILLIGAN, F. J. 2020. Toward Improved Accuracy of Remote Sensing Approaches for Quantifying Suspended Sediment: Implications for Suspended-Sediment Monitoring. In *Journal of geophysical research Earth Surface* [online], vol. 125, no. 7, pp. e2019JF005033 [cit. 2021-03-06]. ISSN 2169-9011. Dostupné na: <<https://doi.org/10.1029/2019JF005033>>.

DI SILVIO, G. – NONES, M. 2014. Morphodynamic reaction of a schematic river to sediment input changes: analytical approaches. In *Geomorphology* [online], vol. 215, pp. 74-

82 [cit. 2021-03-29]. ISSN 0169-555X. Dostupné na: <<https://doi.org/10.1016/j.geomorph.2013.05.021>>.

FRANZOIA, M. – NONES, M. 2017. Morphological reactions of schematic alluvial rivers: long simulations with a 0-D model. In *International Journal of Sediment Research* [online], vol. 32, no. 3, pp. 295-304 [cit. 2021-03-10]. ISSN 1001-6279. Dostupné na: <<https://doi.org/10.1016/j.ijsrc.2017.04.002>>.

FRYIRS, Kristie A. – BRIERLEY, Gary J. 2013. *Geomorphic Analysis of River Systems: An Approach to Reading the Landscape*. Oxford: Wiley-Blackwell. 345 s. ISBN 978-1-405-19274-3.

HAJIGHOLIZADEH, Mohammad - MELESSE, Assefa M. - FUENTES, Hector R. 2018. Erosion and sediment transport modelling in shallow waters: a review on approaches, models and applications. In *International Journal of Environmental Research and Public Health* [online], vol. 15, no. 3, pp. 518 [cit. 2021-03-24]. ISSN 1660-4601. Dostupné na: <<https://doi.org/10.3390/ijerph15030518>>.

HAPSARI, Diana et al. 2019. The Use of Sediment Rating Curve under its Limitations to Estimate the Suspended Load. In *Reviews in Agricultural Science* [online], vol. 7, pp. 88-101 [cit. 2021-03-14]. ISSN 2187-090X. Dostupné na: <[http://dx.doi.org/10.7831/ras.7.0\\_88](http://dx.doi.org/10.7831/ras.7.0_88) I>.

HILLER, Edgar - JURKOVIČ, Ľubomír - ŠUTRIEPKA, Michal. 2010. Metals in the Surface Sediments of Selected Water Reservoirs, Slovakia. In *Bulletin of Environmental Contamination and Toxicology* [online], vol. 84, pp. 635-640 [cit. 2021-03-12]. ISSN 1432-0800. Dostupné na: <<https://doi.org/10.1007/s00128-010-0008-y>>.

HIMANSHU, Sushil Kumar - PANDEY, Ashish – YADAV, Basant. 2017. Assessing the applicability of TMPA-3B42V7 precipitation dataset in wavelet-support vector machine approach for suspended sediment load prediction. In *Journal of Hydrology* [online], vol. 550, pp. 103-117 [cit. 2021-03-15]. ISSN 0022-1694. Dostupné na: <<https://doi.org/10.1016/j.jhydrol.2017.04.051>>.

HOROWITZ, Arthur. J. 2009. A quarter century of declining suspended sediment fluxes in the Mississippi River and the effect of the 1993 flood. In *Hydrological Processes* [online], vol. 24, no. 1, pp. 13-34 [cit. 2021-03-22]. ISSN 1099-1085. Dostupné na: <<http://dx.doi.org/10.1002/hyp.7425>>.

HUCKO, P. et al. 2017. Výskyt stopových prvkov v sedimentoch vodných nádrží SR – výsledky monitoringu 2016. In *zborník prednášok z viii. Konferencie s medzinárodnou účasťou*,

---

*sedimenty vodných tokov a nádrží*. Banská Bystrica: DALI-BB, s.r.o., s. 207-218. ISBN 978-80-89740-13-0.

JUNG, Bárbara M. et al. 2020. Estimating Suspended Sediment Concentrations from River Discharge Data for Reconstructing Gaps of Information of Long-Term Variability Studies. In *Water* [online], vol. 12, no. 9, pp. 2382 [cit. 2021-03-10]. ISSN 2073-4441. Dostupné na: <<https://doi.org/10.3390/w12092382>>.

JURÍK, Luboš - SEDMÁKOVÁ, Miroslava – FUSKA, Jakub. 2017. Analýza sedimentov tajchu Veľká Richňava. In *zborník prednášok z viii. Konferencie s medzinárodnou účasťou, sedimenty vodných tokov a nádrží*. Banská Bystrica: DALI-BB, s.r.o., s. 139-148. ISBN 978-80-89740-13-0.

KIRKBY, M. J. 1987. Modelling some influences of soil erosion, landslides and valley gradient on drainage density and hollow development. In *Catena*, vol. 10, pp. 1–14. ISSN 0341-8162.

KISI, Özgür - YASEEN, Zaher Mundher. 2019. The potential of hybrid evolutionary fuzzy intelligence model for suspended sediment concentration prediction. In *CATENA* [online], vol. 174, pp. 11-23 [cit. 2021-03-21]. ISSN 0341-8162. Dostupné na: <<https://doi.org/10.1016/j.catena.2018.10.047>>.

KLEIN TANK, A. M. G. et al. 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. In *International Journal of Climatology* [online], vol. 22, no. 12, pp. 1441-1453 [cit. 2021-03-25]. ISSN 1097-0088. Dostupné na: <<https://doi.org/10.1002/joc.773>>.

KOITER Alexander J. et al. 2013. The behavioural characteristics of sediment properties and their implications for sediment fingerprinting as an approach for identifying sediment sources in river basins. In *Earth-Science Reviews* [online], vol. 125, pp. 24-42 [cit. 2021-03-30]. ISSN 0012-8252. Dostupné na: <<https://doi.org/10.1016/j.earscirev.2013.05.009>>.

KUHN, Max - JOHNSON, Kjell. 2019. *Feature Engineering and Selection: A Practical Approach for Predictive Models*. Chapman and Hall/CRC Press. 310 s. ISBN 978-1138079-22-9.

MELESSE, Assefa Mekonnen et al. 2011. Suspended sediment load prediction of river systems: An artificial neural network approach. In *Agricultural Water Management* [online], vol. 98, no. 5, pp. 855-866 [cit. 2021-03-13]. ISSN 0378-3774. Dostupné na: <<https://doi.org/10.1016/j.agwat.2010.12.012>>.

MORRIS, Gregory L. - FAN, Jiahua. 1998. Reservoir Sedimentation Handbook: Design and Management of Dams, Reservoirs, and Watersheds for Sustainable Use. New York: McGraw-Hill. 805 s. ISBN 0-07-043302-X.

MUHAMMAD, Nda et al. 2019. A review of field methods for suspended and bedload sediment measurement. In *World Journal of Engineering*, vol. 16, no. 1, pp. 147-165. ISSN 1708-5284. Dostupné na: <<http://dx.doi.org/10.1108/WJE-07-2018-0226>>.

NONES, Michael. 2020. On the main components of landscape evolution modelling of river systems. In *Acta Geophysica* [online], vol. 68, pp. 459–475 [cit. 2021-03-05]. ISSN 1895-7455. Dostupné na: <<https://doi.org/10.1007/s11600-020-00401-8>>.

ONDERKA, Milan - PEKÁROVÁ, Pavla. 2008. Retrieval of suspended particulate matter concentrations in the Danube River from Landsat ETM data. In *Science of The Total Environment* [online], vol. 397, no. 1-3, pp. 238-243 [cit. 2021-03-03]. ISSN 0048-9697. Dostupné na: <<https://doi.org/10.1016/j.scitotenv.2008.02.044>>.

ONDERKA, Milan - RODNÝ, Marek. 2010. Can suspended sediment concentrations be estimated from multispectral imagery using only image-derived information?. In *Journal of the Indian Society of Remote Sensing* [online], vol. 38, no. 1, pp. 85-97 [cit. 2021-02-17]. ISSN 0974-3006. Dostupné na: <<http://dx.doi.org/10.1007/s12524-010-0014-x>>.

REZAPOUR, O. M. - SHUI, Lee Teang - AHMAD, D. B. 2010. Review of artificial neural network model for suspended sediment estimation. *Australian Journal of Basic and Applied Sciences*, vol. 4, no. 8, pp. 3347-3353. ISSN 1991-8178.

ROUSSEEUW, Peter J. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. In *Journal of Computational and Applied Mathematics* [online], vol. 20, pp. 53-65 [cit. 2021-04-01]. ISSN 0377-0427. Dostupné na: <[https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)>.

SLOVENSKÝ HYDROMETEOROLOGICKÝ ÚSTAV. 2019. *Plaveniny, Hodnotenie plaveninového režimu na Slovenských tokoch*. [online]. 2019 [cit 2021-03-07]. Dostupné na: <[http://out01.shmu.sk/File/Hydrologia/Monitoring\\_PV\\_PzV/Monitoring\\_kvantity\\_PV/PVkvant2019/Hodnotenie\\_plaveninoveho\\_rezimu\\_2019.pdf](http://out01.shmu.sk/File/Hydrologia/Monitoring_PV_PzV/Monitoring_kvantity_PV/PVkvant2019/Hodnotenie_plaveninoveho_rezimu_2019.pdf)>.

SOČUVKA, V. 2017. Analýza zanášania vodného diela Veľké Kozmálovce. In *zborník prednášok z viii. Konferencie s medzinárodnou účasťou, sedimenty vodných tokov a nádrží*. Banská Bystrica: DALI-BB, s.r.o., s. 139-148. ISBN 978-80-89740-13-0.

TAYLOR, Kevin G. - OWENS, Philip N. 2009. Sediments in urban river basins: a review of sediment–contaminant dynamics in an environmental system conditioned by human

---

activities. In *Journal of Soils and Sediments* [online], vol. 9, pp. 281-303 [cit. 2021-03-08]. ISSN 1614-7480. Dostupné na: <<https://doi.org/10.1007/s11368-009-0103-z>>.

TUCKER, Gregory E. 2009. Natural experiments in landscape evolution. In *Earth Surface Processes and Landforms* [online], vol. 34, no. 10, pp. 1450–1460 [cit. 2021-03-01]. ISSN 1096-9837. Dostupné na: <<https://doi.org/10.1002/esp.1833>>.

VARRANI, Arianna – NONES, Michael – GUPANA, Remika. 2019. Long-term modelling of fluvial systems at the watershed scale: examples from three case studies. In *Journal of Hydrology* [online], vol. 574, pp. 1042–1052 [cit. 2021-03-01]. ISSN 0022-1694. Dostupné na: <<https://doi.org/10.1016/j.jhydrol.2019.05.012>>.

VERCRUYSSSE Kim. - GRABOWSKI, Robert C. - RICKSON, Richmal Jane. 2017. Suspended sediment transport dynamics in rivers: Multi-scale drivers of temporal variation. In *Earth-Science Reviews* [online], vol. 166, pp. 38-52 [cit. 2021-03-11]. ISSN 0012-8252. Dostupné na: <<https://doi.org/10.1016/j.earscirev.2016.12.016>>.

WALLING, Desmond E. - FANG, D. 2003 Recent trends in the suspended sediment loads of the world's rivers. In *Global and Planetary Change* [online], vol. 39, no. 1-2, pp. 111-126 [cit. 2021-02-20]. ISSN 0921-8181. Dostupné na: <[https://doi.org/10.1016/S0921-8181\(03\)00020-1](https://doi.org/10.1016/S0921-8181(03)00020-1)>.

WANG, Jianjun - LU, Xi Xi - KUMMU, Matti. 2011. Sediment load estimates and variations in the Lower Mekong River. In *River Research and Applications* [online], vol. 27, no. 1, pp. 33-46 [cit. 2021-03-06]. ISSN 1535-1467. Dostupné na: <<http://dx.doi.org/10.1002/rra.1337>>.

WHIPPLE, Kelin X. - TUCKER, Gregory E. 2002. Implications of sediment-flux-dependent river incision models for landscape evolution. In *Journal of Geophysical Research Atmospheres* [online], vol. 107, no. B2, pp. ETG 1-20 [cit. 2021-03-05]. ISSN 2156-2202. Dostupné na: <<http://dx.doi.org/10.1029/2000JB000044>>.

YADAV, Basant et al. 2018. Data-based modelling approach for variable density flow and solute transport simulation in a coastal aquifer. In *Hydrological Sciences Journal/Journal des Sciences Hydrologiques* [online], vol. 63, no. 2, pp. 210-226 [cit. 2021-03-21]. ISSN 0262-6667. Dostupné na: <<http://dx.doi.org/10.1080/02626667.2017.1413491>>.

## 8. Zoznam vlastnej publikačnej činnosti

### ADC - Vedecké práce v zahraničných karentovaných časopisoch

ADC01 ČISTÝ, Milan - SOLDÁNOVÁ, Veronika – CYPRICH, František – HOLUBOVÁ, Katarína – ŠIMOR, Viliam. 2021. Suspended sediment modelling with hydrological and climate input data. In *Journal of Hydroinformatics*, vol. 23, no. 1, pp. 192-210. ISSN 1465-1734. V databáze: SCOPUS: 2-s2.0-85099740645, WOS: 000613071400013

### AFC - Publikované príspevky na zahraničných vedeckých konferenciách

AFC01 ČISTÝ, Milan - CYPRICH, František - SOLDÁNOVÁ, Veronika. Prediction of soil moisture data by various regression techniques. In *SGEM 2018. 18th International Multidisciplinary Scientific GeoConference. Volume 18. Science and Technologies in Geology, Oil and Gas Exploration, Water Resources, Forest Ecosystems*. Sofia: STEF 92 Technology, 2018, s. 383--389. ISBN 978-619-7408-72-0. V databáze: SCOPUS: 2-s2.0-85063075166

AFC02 ČISTÝ, Milan - CYPRICH, František - SOLDÁNOVÁ, Veronika. A comparison of linear and machine learning models for the simulation of soil moisture. In *PERNER, P. Machine Learning and Data Mining in Pattern Recognition*. Leipzig: ibai-publishing, 2019, s. 104--116. ISBN 978-3-942952-62-0.

AFC03 ČISTÝ, Milan - CYPRICH, František - DEŽERICKÝ, Dávid. Interpolation of irregular soil moisture measurements with machine learning methods. In *5th World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium - WMCAUS 2020*. Bristol: IOP Publishing, 2020. V databáze: SCOPUS: 2-s2.0-85097977402, WOS: 000646533100268

AFC04 ČISTÝ, Milan - CYPRICH, František. Evaluation of linear and machine learning models for determining pedotransfer functions. In *6th World Multidisciplinary Earth Sciences Symposium - WMESS 2020*. Bristol: IOP Publishing, 2020. V databáze: SCOPUS: 2-s2.0-85098008168

AFC05 ČISTÝ, Milan - CYPRICH, František - HOLUBOVÁ, Katarína – ŠIMOR, Viliam. Comparison of modelling of suspended sediment concentration using linear and machine learning methods. In *SGEM 2020. 20th International Multidisciplinary Scientific GeoConference. Volume 20. Water Resources. Forest, Marine and Ocean Ecosystems*. Sofia: STEF 92 Technology, 2020, s. 51--58. ISBN 978-619-7603-08-8. V databáze: SCOPUS: 2-s2.0-85099729598

AFC06 CYPRICH, František. Doplnenie nameraných koncentrácií plavenín. In *ŠUHAJDOVÁ, E. -- NOVÁKOVÁ, J. -- VELIKOVSKÁ, K. -- MORAVČÍKOVÁ, S. -- KURUC, M. -- NOSEK, J. Juniorstav 2021*. Brno: ECON publishing, 2021, s. 416--421. ISBN 978-80-86433-75-2.

### **AFD - Publikované príspevky na domácich vedeckých konferenciách.**

- AFD01 CYPRICH, František. Hodnotenie zmien využívania krajiny v obci Podhorany v skúmanom časovom horizonte 1949-2012. In *29. konferencia mladých hydroológov, 16. konferencia mladých vodohospodárov, 18. konferencia mladých meteorológov a klimatológov*. Bratislava: Slovenský hydrometeorologický ústav, 2017, ISBN 978-80-88907-95-4.
- AFD02 CYPRICH, František. Interpolácia pôdnej vlhkosti na základe satelitných údajov. In *Advances in Architectural, Civil and Environmental Engineering*. Bratislava: Spektrum STU, 2018, s. 182--190. ISBN 978-80-227-4864-3.
- AFD03 CYPRICH, František. Využitie gridových databáz pri vodohospodárskych výpočtoch. In *Advances in Architectural, Civil and Environmental Engineering*. Bratislava: Spektrum STU, 2019, s. 153--159. ISBN 978-80-227-4972-5.
- AFD04 CYPRICH, František. Význam klimatických údajov pre doplnenie údajov plavenín. In *Advances in Architectural, Civil and Environmental Engineering*. Bratislava: Spektrum STU, 2020, s. 231--236. ISBN 978-80-227-5052-3.

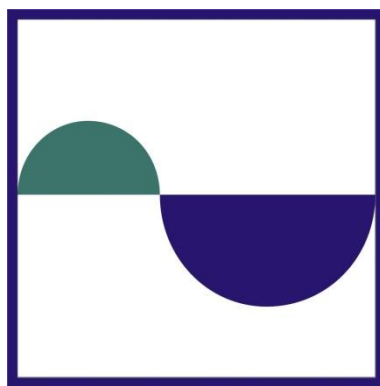
### **AFG - Abstrakty príspevkov zo zahraničných vedeckých konferencií**

- AFG01 ČISTÝ, Milan - SOLDÁNOVÁ, Veronika – CYPRICH, František. Advantages of the Simultaneous Use of the Hydrologic and Stochastic Flow Generators. In *Geophysical Research Abstracts*. Volume 20/2018. Göttingen: Copernicus Publications, 2018.
- AFG02 ČISTÝ, Milan - SOLDÁNOVÁ, Veronika – CYPRICH, František. Random forest based reclassification of soil texture for hydrological modelling. In *Geophysical Research Abstracts*. Volume 21/2019. Göttingen: Copernicus Publications, 2019.

### **AFK - Postery zo zahraničných konferencií**

- AFK01 SOLDÁNOVÁ, Veronika – ČISTÝ, Milan - CYPRICH, František. On the use of grid-based databases for r-r modelling. In *KALICZ, P. -- HLAVČOVÁ, K. -- KOHNOVÁ, S. -- RATTAYOVÁ, V. -- GRIBOVŠKI, Z. HydroCarpath 2018. Catchment Processes in Regional Hydrology: Field Experiments and Modelling in Carpathians Basins*. Sopron: University of Sopron Press, 2018, s. 36. ISBN 978-963-334-199-5.





## **Katedra vodného hospodárstva krajiny**

blok C, 12. posch., Radlinského 11

810 05 Bratislava 1

+421 2 32 888 498

[info@kvhk.sk](mailto:info@kvhk.sk)