

Česká zemědělská univerzita v Praze

Fakulta agrobiologie, potravinových a přírodních zdrojů

Katedra vodních zdrojů



**Česká zemědělská
univerzita v Praze**

**Možnosti stanovení nasycené hydraulické vodivosti
nepřímými metodami**

Diplomová práce

Bc. Eva Hruzová

Ochrana a využívání přírodních zdrojů

**Vedoucí práce: Ing. Kamila Bářková, MSc., Ph.D.
Konzultant práce: Ing. Markéta Miháliková, Ph.D.**

© 2022 ČZU v Praze

Čestné prohlášení

Prohlašuji, že svou diplomovou práci „Možnosti stanovení nasycené hydraulické vodivosti nepřímými metodami“ jsem vypracovala samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autorka uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 13.4.2022

Poděkování

Ráda bych touto cestou poděkovala Ing. Kamile Bářkové, MSc., Ph.D. za vedení mé diplomové práce, za cenné rady a připomínky k práci i ochotu mi kdykoliv pomoci. Děkuji také Ing. Markétě Mihálikové, Ph.D. za poskytnutí potřebných dat, bez kterých by se má diplomová práce nemohla obejít.

Poděkování patří i mé rodině a Honzovi za podporu během celého studia.

Možnosti stanovení nasycené hydraulické vodivosti nepřímými metodami

Souhrn

Diplomová práce je zaměřena na problematiku odhadu nasycené hydraulické vodivosti (K_S), jedné z nejdůležitějších hydraulických charakteristik půdy.

V literární rešerši jsou popsány nejčastěji používané přímé a nepřímé metody stanovení K_S . Mezi nepřímé metody stanovení K_S patří metody, které její hodnotu odhadují na základě jednodušejí stanovitelných půdních charakteristik neboli prediktorů. Nazývají se pedotransferové funkce (PTF). Mezi základní prediktory patří zrnitostní složení, objemová hmotnost suché půdy, obsah organického uhlíku, případně hydrolimity – polní kapacita a bod vadnutí.

V praktické části diplomové práce jsou provedeny odhady K_S pomocí metod využívajících strojového učení – aplikace Araya a Ghezzehei (2019) a metody neuronových sítí v programu Rosetta (Schaap et al. 2001).

Aplikace Araya a Ghezzehei (2019) je novým programem pro odhad K_S využívající dva algoritmy: random forest (RF) a boosted regression trees (BRT). Ačkoliv byla tato aplikace vytvořena a otestována pro půdy USA, díky velkému množství dat (více než 18 000 údajů) použitých při jejím vývoji lze předpokládat, že by mohla být použitelná obecně i pro půdy v jiných částech světa.

Otestování tohoto předpokladu je předmětem diplomové práce, ve které jsou srovnány výsledky odhadů těchto dvou algoritmů pro tři úrovně prediktorů na celkem 126 vzorcích půd ze 13 lokalit České republiky, pro které byly k dispozici potřebné prediktory a zároveň změřena hodnota K_S . Výsledné odhadnuté K_S jsou statisticky porovnány se změřenými daty a také ve vztahu k odhadům získaným s již zavedenou metodou využívající analýzu neuronových sítí v programu Rosetta. Pro porovnání bylo použito korelačních a krabicových grafů a také koeficientu determinace (R^2), střední chyby (ME), a střední kvadratické chyby (RMSE). Vzhledem k heterogenitě půd na testovaných lokalitách, bylo testování provedeno nejen pro jednotlivé lokality, ale i pro seskupené zrnitostní třídy USDA. Všech osm testovaných modelů PTF bylo seřazeno na základě dosažené hodnoty RMSE (pořadí 1-8); hodnota 1 byla přiřazena modelu PTF s nejmenší RMSE označující nejlepší odhad.

Kvalita odhadů K_S byla srovnatelná s kvalitou odhadů autorů testovaných PTF i dříve publikovaných prací. Nebylo pozorováno významné zpřesnění při použití metod strojového učení. Jednou z možných příčin by mohlo být značně rozdílné zrnitostní složení půd použitých při testování, oproti půdám obsaženým v tréninkovém souboru, obsahující zrnitostně lehčí půdy. Pro seskupené zrnitostní třídy lehká a lehčí střední zemina (třídy 1 a 2) byly odhady velmi kvalitní.

Klíčová slova: nasycená hydraulická vodivost, odhad, pedotransferové funkce, strojové učení

Determination of saturated hydraulic conductivity by indirect methods

Summary

The diploma thesis is aimed at the estimation of saturated hydraulic conductivity (K_S), one of the most important hydraulic characteristics of soil.

The literature review provides overview of the most frequently used methods of direct and indirect determination of K_S . Indirect methods of determining K_S include methods estimating K_S values based on easily measurable and/or available soil data (predictors). The indirect methods translating predictors into the K_S are referred to as pedotransfer functions (PTFs). The basic predictors covers information about the soil texture (attribution to the soil texture class or % of sand, silt and clay particles). Additionally to that, the dry bulk density, the organic carbon content or the hydrolimits such as the field capacity, and the wilting point, can be employed as predictors.

The practical part of the diploma thesis presents K_S estimates achieved by machine learning methods: applications of Araya and Ghezzehei (2019) and by neural network methods in the Rosetta program (Schaap et al. 2001).

The application of Araya and Ghezzehei (2019) is up to date program for K_S estimates based on two algorithms: random forest (RF) and boosted regression trees (BRT). Although the PTFs were derived and tested for the US soils, the large background soil database consisting of more than 18 000 data sets, leads to an assumption, that it may be applicable for soils of different origin in other parts of the world.

The assumption has been tested based on 126 soil samples from 13 localities in the Czech Republic. Estimates of the RF and BRT algorithms for three levels of predictors were compared to the measured K_S data and compared to the K_S estimates obtained by two models of PTFs within the Rosetta computer program. The results were graphically compared by means of correlation graphs and boxplots. The statistical evaluation based on log-transformed K_S data includes calculations of coefficient of determination (R^2), mean error (ME), and root mean squared error (RMSE). With regard to the heterogeneity of the soils of the various localities, the quality of the K_S estimates was evaluated in terms of individual localities and also in terms of the grouped USDA soil texture classes. All eight tested PTFs models were ranked based on the achieved RMSE value (1-8); rank 1 was attributed to the PTF model with the smallest RMSE denoting the best estimate.

The RMSE values for the K_S estimates were comparable to the RMSE values found in other published studies. Using the machine-learning method did not lead to any significant improvement in K_S estimates. One of the possible reasons could be the different texture composition of the soils used in this study, when soils of medium to fine texture were involved to a greater extent. The high quality K_S estimates were obtained for the course textured soils (grouped USDA soil texture classes 1 and 2).

Keywords: saturated hydraulic conductivity, estimation, pedotransfer functions, machine learning

Obsah

1 Úvod.....	9
2 Vědecká hypotéza a cíle práce	10
2.1 Cíle práce	10
2.2 Vědecká hypotéza.....	10
3 Literární rešerše	11
3.1 Nasycená hydraulická vodivost.....	11
3.2 Historie stanovení nasycené hydraulické vodivosti.....	11
3.3 Přímé stanovení nasycené hydraulické vodivosti	15
3.3.1 Jednosondová metoda	15
3.3.2 Dvouválcový infiltrometr (Parr & Bertrand 1960).....	16
3.3.3 Tlakový infiltrometr (Matula & Kozáková 1997)	17
3.3.4 SATURO	17
3.3.5 Přístroj s konstantním spádem	17
3.3.6 Přístroj s proměnným spádem.....	18
3.4 Nepřímé stanovení nasycené hydraulické vodivosti.....	20
3.4.1 Pedotransferové funkce.....	20
3.4.2 Rozdělení pedotransferových funkcí	20
Texturní pedotransferové funkce	20
Kontinuální pedotransferové funkce	22
Bodové pedotransferové funkce.....	23
Parametrické pedotransferové funkce	23
3.4.3 Vývoj pedotransferových funkcí	23
3.5 Vybrané metody nepřímého stanovení nasycené hydraulické vodivosti ..	24
3.5.1 Úvod.....	24
3.5.2 Program Rosetta.....	25
Umělá neuronová síť (ANN).....	26
Hydraulické parametry	28
3.5.3 PTF App - aplikace autorů S. N. Araya a T. A. Ghezzehei	30
Modely boosted regression trees	32
Modely random forest	33
3.5.4 Vstupní data	34
Textura	34

Objemová hmotnost půdy	36
Půdní hydrolimity	37
Organický uhlík.....	37
Organický uhlík komplexní	38
4 Metodika	38
4.1 Zdroje dat	38
4.2 Dostupné informace o zájmových oblastech.....	40
4.2.1 Lokalita Ruzyně.....	40
4.2.2 Lokalita Strašov	41
4.2.3 Lokalita Uherčice.....	41
4.3 Postup práce	43
4.3.1 Prediktory	43
Zrnitostní třídy	43
Objemová hmotnost suché půdy (BD).....	45
Obsah organického uhlíku	45
4.3.2 Testované modely PTF	46
Modely BRT	46
Modely RF	46
Rosetta.....	47
4.3.3 Určení kvality odhadu a statistické vyhodnocení	47
5 Výsledky.....	49
5.1 Srovnání kvality odhadu použitých PTF	49
5.1.1 Porovnání pro jednotlivé lokality	49
5.1.2 Porovnání pro seskupené zrnitostní třídy	60
6 Diskuse	66
7 Závěr	69
8 Literatura.....	70
9 Seznam tabulek	74
10 Seznam obrázků	74

1 Úvod

Voda je jedním z nejcennějších přírodních zdrojů na naší planetě. Chemický průmysl, papírenský průmysl, metalurgie, zemědělství, stavebnictví i jiné každodenní činnosti člověka způsobují znečištění podzemních i povrchových zdrojů vody a ohrožují tím stav vodních ekosystémů i ekosystémů v jejich okolí. Znečištění vody je jedním ze závažných problémů současného světa, protože omezuje určité části lidské populace přístup k pitné vodě. V zemích hospodářsky méně rozvinutých může být zdrojem znečištění přelidnění a s tím související odpad komunální, absence čistíren odpadních vod nebo nedostatek finančních zdrojů nutných k nápravě znečištění. Naopak v zemích hospodářsky více rozvinutých se setkáváme s konzumním způsobem života, který v důsledku přináší více odpadů a tím negativně ovlivňuje životní prostředí.

Půda je dalším velmi cenným přírodním zdrojem, který poskytuje potravu a produkuje biomasu pro výživu lidí, krmení zvířat či energetické rostliny. Půda je také významným rezervoárem sladké vody. Zde se setkávají všechny hlavní sféry Země – atmosféra, hydrosféra, biosféra i litosféra. Tato práce je zaměřena především na působení hydrosféry. Hydrosféra ovlivňuje přítomnost vody v půdě, zároveň i půda ovlivňuje pohyb a rozpouštění látek ve vodě, působení sfér je vzájemné (Šimek et al. 2019).

Na ochranu povrchové, půdní a podzemní vody se v současnosti klade stále větší důraz, a to jak z hlediska kvalitativního, tak kvantitativního. Abychom dokázali vhlédnout do problematiky pohybu vody v půdě, potřebujeme znát důležité hydraulické charakteristiky půdy. Rychlost zasakování vody do půdy záleží na mnoha vlastnostech půdy, ale i na terénu, množství vsakované látky i vegetaci.

Mezi nejdůležitější hydraulické charakteristiky půdy patří retenční křivka a hydraulická vodivost (nasyčená i nenasycená). Nasyčená hydraulická vodivost charakterizuje schopnost nasyčené půdy vodu transportovat.

Stanovení nasyčené hydraulické vodivosti je důležité pro mnoho různých aplikací, neboť je klíčovým parametrem pro vědní obory jako hydrogeologie, hydroopedologie nebo ochrana životního prostředí. Stanovení hydraulických charakteristik je nezbytným krokem při numerickém modelování proudění vody a transportu chemických i biologických látek v půdním prostředí.

Přímé měření nasyčené hydraulické vodivosti, obzvláště pro velká území, je značně časově, a tedy i finančně náročné. Proto se stále častěji využívá nepřímého stanovení – odhadu, založeného na využití dostupných a snadněji měřitelných charakteristik neboli prediktorů. Termín pedotransferové funkce (PTF) poprvé použil Bouma (1989), dále pak Minasny et al. (1999) PTF zavádějí jako metodu, která přeloží data, která máme, do dat, která potřebujeme.

V diplomové práci jsou pro půdy České republiky otestovány PTF založené na principu strojového učení, které odvodili autoři Araya a Ghezzehei v roce 2019 pro půdy USA. Je to zatím nejnovější dostupná aplikace využívající strojového učení pro odhad nasyčené hydraulické vodivosti.

2 Vědecká hypotéza a cíle práce

2.1 Cíle práce

Cílem této diplomové práce je otestovat a vyhodnotit použitelnost PTF založených na principu strojového učení odvozených pro půdy v USA (Araya & Ghezzehei 2019) pro odhad nasycené hydraulické vodivosti (K_S) půd v České republice.

Cílem je zvolit co nejvhodnější varianty prezentovaných PTF na základě dostupných prediktorů a kvalitu odhadu statisticky vyhodnotit na základě porovnání se změřenými daty pro konkrétní lokality v rámci České republiky.

Porovnat kvalitu odhadu K_S s odhady již dříve publikovanými a stále používanými PTF v programu Rosetta (Schaap et al. 2001) založeném na analýze neuronových sítí.

2.2 Vědecká hypotéza

Pedotransferové funkce založené na principu strojového učení odvozené na základě velkého množství dat (více než 18 000 záznamů) pro půdy v USA je možné aplikovat pro odhad nasycené hydraulické vodivosti půd v České republice se srovnatelnou přesností.

3 Literární řešerše

3.1 Nasycená hydraulická vodivost

Nasycená hydraulická vodivost je schopnost vodou nasycené půdy vodu transportovat. Čím je hydraulická vodivost vyšší, tím je tok vody půdním prostředím rychlejší. Pokud jsou půdní póry zcela zaplněny vodou, jedná se o nasycené proudění.

Naproti tomu k nenasyčenému proudění dochází v půdní vrstvě, která není zcela nasycena vodou a část pórů je vyplněna vzduchem (Kozák et al. 2011).

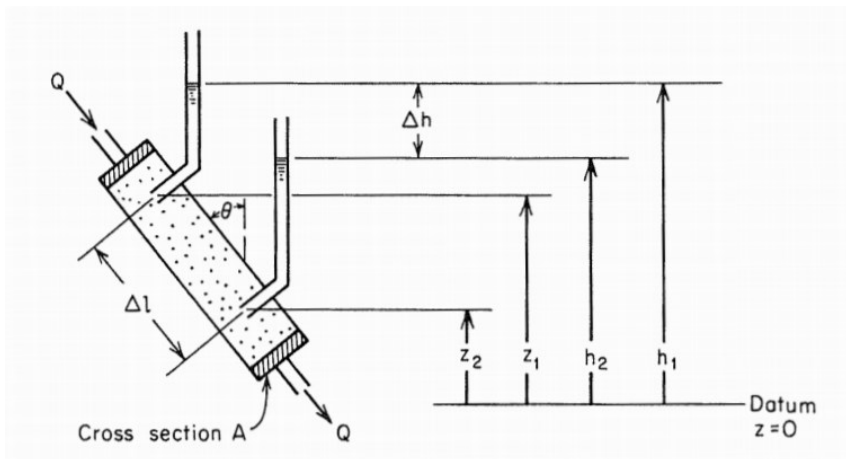
Nasycená hydraulická vodivost (K_S) je v přírodních vědách důležitým parametrem. Udává rychlost proudění vody pod hydraulickým gradientem, jak je stanoveno Darcyho rovnicí pro nasycené podmínky, ale také hraje roli v mnoha modelech pro nenasyčené proudění. Bez znalosti nasycené hydraulické vodivosti by bylo obtížné přesněji popsat transport vody, chování suspendovaných částic v půdě a sedimentech, nebo odhadnout zásoby podzemní vody a jejich doplňování (Zhang & Schaap 2019).

Měření K_S mají dlouhou tradici především v nasycených podmínkách, například při odvodňování zemědělských půd, dopravních a vodohospodářských staveb, či stavenišť (Valla et al. 2008).

Pro popis a předpověď dějů v rámci soustavy půda – hydrosféra – atmosféra, je nutné mít správně parametrizované hydraulické veličiny (retenční křivku, křivku nenasyčené hydraulické vodivosti a nasycenou hydraulickou vodivost). Numerické modely sloužící k monitorování a řízení transportu vody a roztoků v půdě vyžadují výše zmíněné vstupní parametry. Tyto klíčové údaje jsou shromažďovány do národních i mezinárodních databází. Pro Českou republiku vznikla databáze HYPRESCZ, která vychází ze struktury a metodiky databáze evropských půd HYPRES (Miháliková et al. 2013).

3.2 Historie stanovení nasycené hydraulické vodivosti

Pojem nasycená hydraulická vodivost použil poprvé v 19. století Henry Darcy, francouzský technik a inženýr, který se zabýval především hydraulikou, například distribucí vody za pomoci akvaduktů. V jeho díle „Les Fontaines Publiques de la Ville de Dijon“ (1856) popsal mnoho odlišných experimentů, například průtok vody přes pískové filtry (Obrázek 1) (Brown 2002).



Obrázek 1 – Ilustrace Darcyho pokusu od autorů R. Alan Freeze a John A. Cherry z knihy Groundwater (1979).

Darcy ve svém díle uvádí, že průtok je přímo úměrný tlakovému rozdílu a nepřímo úměrný tloušťce filtru, což mu umožnilo vytvořit původní rovnici (1):

$$Q = \frac{k \cdot s}{e} (H + e) \quad (1)$$

kde:

Q – je průtok vody [m^3/s]

s – plocha filtru [m^2]

e – tloušťka filtru [m]

H – výška vody nad filtrem [m]

Proměnnou k lze přeložit jako „koeficient závislý na povaze písku“.

Po úpravě pro zjednodušený případ vertikálního proudění je Darcyho zákon často zapsán s použitím veličin proudění a hydraulického gradientu jako rovnice (2):

$$q = -K_s \cdot \frac{\Delta H}{\Delta z} \quad (2)$$

kde:

ΔH – rozdíl mezi hladinami vody při přítoku a odtoku vody [m]

(v rovnici (1) Darcy označuje jako $H + e$)

Δz – podél délky [m]

Dále lze pozorovat, že písmeno k z rovnice (1) je přeznačeno na K_s , nasycenou hydraulickou vodivost, kterou můžeme použít pro homogenní propustné porézní prostředí, které je nasyceno vodou, za podmínky laminárního proudění (Zhang & Schaap 2019).

Přestože výsledek Darcyho pokusu je známý a uznávaný po celém světě, o procesu jeho objevení toho není k dispozici mnoho. Částečně za to může fakt, že Darcy žil před více než 150 lety, kopie jeho spisů jsou málo dostupné a ve francouzštině (Brown 2002).

První aplikaci nasycené hydraulické vodivosti v půdním modelování lze nalézt v knize „Weather Prediction by Numerical Process“ od L. Richardsona (1922). Lewis Fry Richardson byl anglický matematik, fyzik, meteorolog, průkopník numerického modelování v oblasti meteorologie. Během první světové války Richardson vymyslel vizionářskou představu pro předpověď počasí – pomocí základních rovnic hydrodynamiky a termodynamiky k určení budoucího stavu atmosféry. Zajímavé je, že jeho dílo „Weather Prediction by Numerical Process“ bylo ztraceno během bitvy o Champagne v dubnu 1917. Naštěstí (naštěstí i pro vývoj meteorologie) bylo nalezeno o měsíce později pod hromadou uhlí (Vulpiani 2014).

Richardsonovo dílo bylo publikováno roku 1922. V současnosti se uvádí Richardsonova rovnice (3) pro vertikální nenasyčené proudění podél osy z:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[K(\theta) \left(\frac{\partial h}{\partial z} + 1 \right) \right] \quad (3)$$

kde:

$K(\theta)$ – hydraulická vodivost [m/s] jako funkce obj. vlhkosti θ [m³/m³]

h – vlhkostní potenciál [m]

t – čas [s]

Derivace poté značí změnu – jak se mění objemová vlhkost půdy v čase (Zhang & Schaap 2019).

Další důležitá rovnice je Kozeny-Carman rovnice. Josef Alexandr Kožený poprvé dokázal poskytnout teoretický základ hydraulické vodivosti zapojením nových proměnných – tvaru a objemu půdních pórů. V roce 1927 sestavil rovnici, kterou později v roce 1937 doplnil a upravil Philip C. Carman na známý tvar Kozeny-Carman rovnice (4), ve které je hydraulická vodivost definována pomocí průměru částic d .

$$k = C \frac{d^2 \varphi^3}{(1 - \varphi)^2} \quad (4)$$

kde:

d – průměr půdních částic [m]

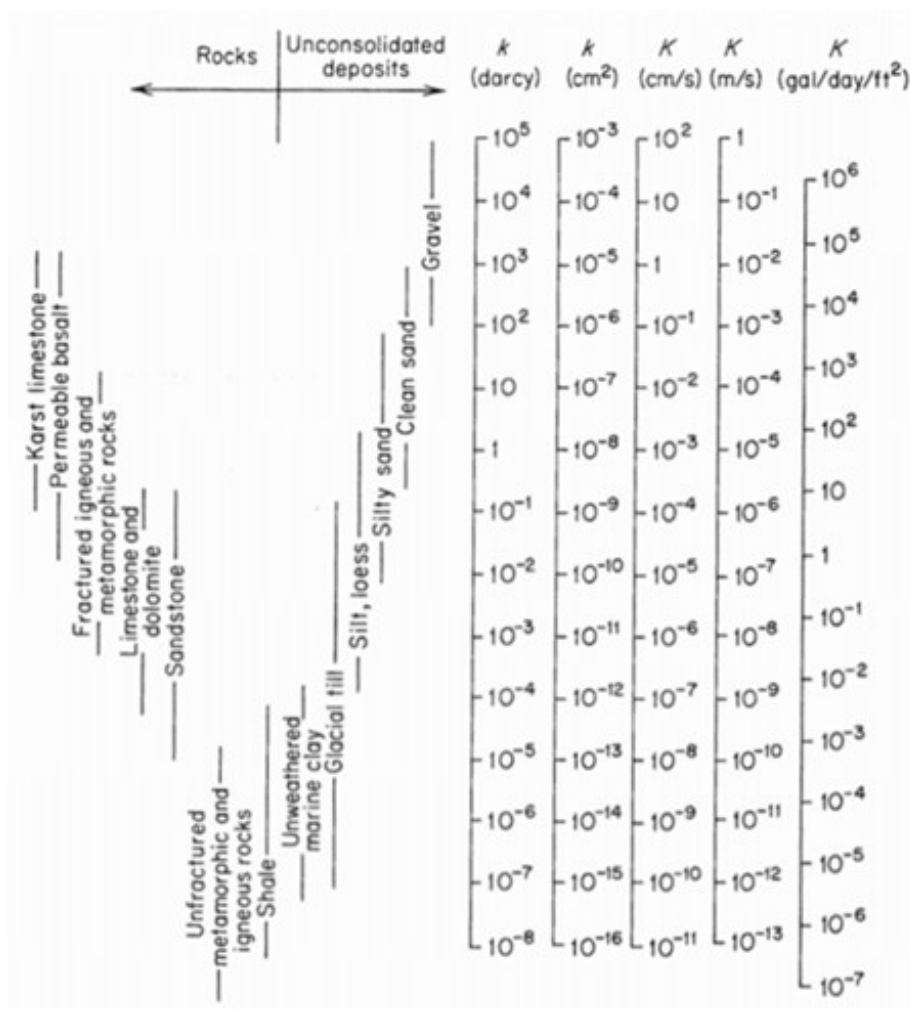
φ – pórovitost [-, %]

C – konstanta úměrnosti (Zhang & Schaap 2019).

Rovnice definuje nasycenou hydraulickou vodivost pomocí průměru půdních částic, veličiny, která je snadno měřitelná. Neposkytuje ale celkový prostorový náhled na chování půdní vody, protože velikost půdních částic se může v různých místech značně lišit.

Alan Freeze a John Cherry v roce 1979 uvedli ve své knize „Groundwater“ přehled hodnot hydraulické vodivosti a půdní propustnosti pro různé textury půdy (Obrázek 2).

Hodnoty nasycené hydraulické vodivosti se nachází v rozsahu od 10^{-11} do 10^2 cm/s v závislosti na materiálu a zrnitosti (Freeze & Cherry 1979).



Obrázek 2 - Hodnoty nasycené hydraulické vodivosti a propustnosti podle autorů R. Alan Freeze a John A. Cherry z knihy Groundwater (1979). Zleva: krasový vápenec, propustný čedič, zvětralé vyvřelé a metamorfované horniny, vápence a dolomity, pískovce, neporušené metamorfované nebo vyvřelé horniny, břidlice, nezvětralý jíl, ledovcový till, prach, spraš, prachovitý písek, písek, štěrk.

Campbell v roce 1974 vytvořil funkci pro stanovení nenasyčené hydraulické vodivosti s použitím určených pěti hodnot nasycené hydraulické vodivosti pro různé zrnitostní třídy (Zhang & Schaap 2019). Cosby et al. v roce 1984 odvodili parametry pro určení retenční křivky a nasycené hydraulické vodivosti s využitím 1448 vzorků od autorů Holtan, Rowls et al. Zajímavostí je, že ze zmíněných údajů vychází např. Community Land Model (CLM), který se zabývá koncepty ekologické klimatologie (Oleson et al. 2013).

3.3 Přímé stanovení nasycené hydraulické vodivosti

Ke stanovení nasycené hydraulické vodivosti pomocí terénního měření lze použít různé metody v závislosti na podmínkách stanoviště. V lokalitách s dostupnou hladinou podzemní vody lze s výhodou použít jednosondovou metodu, zatímco v lokalitách, kde hladina podzemní vody není dostupná se používají infiltrometry různé konstrukce a rozměrů, např.: dvouválcový infiltrometr (Parr & Bertrand 1960), tlakový infiltrometr (Matula & Kozáková 1997), SATURO (METER) nebo Hood infiltrometr (Umwelt-Geräte-Technik GmbH).

K_S je také standardně měřena v laboratoři na odebraných neporušených nebo uměle hutněných půdních vzorcích na přístrojích s konstantním nebo proměnným spádem, např. přístroj Chameleon (Soilmoisture Equipment Corp.) nebo přístroj K_{SAT} (METER), které lze použít pro oba typy měření.

Jednotlivé metody se liší jak principem měření, ale i velikostí půdního vzorku, který je proměřován. Pro měření nasycené hydraulické vodivosti není zavedena referenční metoda, ke které by bylo možné ostatní metody vztáhnout. Navíc je to charakteristika proměnlivá, jak v místě (půdní heterogenita), tak v čase. To je patrné obzvláště u zemědělských půd, kde orbou či jinými agrotechnickými zásahy dochází k dramatickým změnám vlastností svrchní vrstvy půdy. Ve zprávách a databázích údaje o použité metodě nebývají vždy dostupné.

3.3.1 Jednosondová metoda

Jednosondová metoda se využívá v lokalitách s dostupnou hladinou podzemní vody v měřené vrstvě. Ke stanovení se tak využívá přírodní voda z místa výskytu. Po vyvrtání se stanoví výška ustálené hladiny podzemní vody, voda se pomocí kalovky vyčerpá a sleduje se postupné zvyšování hladiny vody v závislosti na čase (Obrázek 3).

Nasycená hydraulická vodivost se vypočte z naměřených polních parametrů pomocí rovnice Hooghoudta a Ernsta (Hooghoudt 1936; Ernst 1950), pokud byla použita sonda o poloměru v rozmezí od 3 cm do 7 cm, hloubka vodního sloupce H v době ustáleného stavu se nacházela v rozmezí od 20 cm do 200 cm a interval pro pozorování nástupu hladiny byl větší než $0,2H$.

Může také být použita rovnice Kirkhama a van Bavela (Kirkham & van Bavel 1948), u které není hodnota ustálené hladiny podzemní vody H nijak omezena.

Pro mělké homogenní nasycené systémy se využívá Hvorslevova metoda (Hvorslev 1951), která je z těchto tří metod nejsnazší. Měření jednosondovou metodou je rychlé a nenáročné na přístrojové vybavení (Bátková et al. 2013).



Obrázek 3 - Sonda před odstraněním vody kalovkou nebo čerpadlem (Bát'ková et al. 2013).

3.3.2 Dvouválcový infiltrometr (Parr & Bertrand 1960)

Používanou klasickou metodou infiltračních testů je dvouválcový infiltrometr (Obrázek 4). Používá se v lokalitách, kde není dostupná hladina podzemní vody. Skládá se ze dvou soustředných válců a děrované vnitřní desky s měřicími hroty. V terénu se pomocí dvouválcového infiltrometru měří čas potřebný k zainfilrování konkrétního množství vody, které je indikováno objevením hrotu děrované vnitřní desky. Změřená data lze vyhodnotit pomocí Philipových (Philip 1957) infiltračních rovnic (Bát'ková et al. 2013).



Obrázek 4 - Dvouválcový infiltrometr (Bát'ková et al. 2013).

3.3.3 Tlakový infiltrometr (Matula & Kozáková 1997)

Pro stanovení nasycené hydraulické vodivosti se používá tlakový infiltrometr (Obrázek 5), který pracuje na mechanicko-hydraulickém základu bez nároku na zdroj energie, proto je výhodný pro práci v polních podmínkách. S jeho pomocí lze určit kumulativní infiltraci. Výše popsaná metoda vyžaduje znalost počátečních podmínek experimentu, především vlhkosti půdy v blízkosti tlakového infiltrometru. Pro vypočtení nasycené hydraulické vodivosti se v tomto případě musí použít upravená rovnice Elricka a Reynoldse (Bát'ková et al. 2013).



Obrázek 5 - Tlakový infiltrometr (Matula & Kozáková 1997), foto: S. Matula.

3.3.4 SATURO

Moderní infiltrometr SATURO od společnosti METER je alternativou dvouválcového infiltrometru. Je schopen poskytovat přesná měření hydraulické vodivosti a hydraulických vlastností půdy v terénních podmínkách.

Přístroj SATURO byl navržen jako automatický přístroj pro měření propustnosti a nasycené hydraulické vodivosti v půdě. Využívá vícetlakou hlavici a zjednodušuje korekce pro trojrozměrný tok z jednoválcového infiltrometru, což umožňuje rychlé měření hydraulické vodivosti bez nutnosti dalšího zpracování (METER).

3.3.5 Přístroj s konstantním spádem

Ke zjištění koeficientu nasycené hydraulické vodivosti v laboratorních podmínkách se využívá přístroj, tzv. permeametr (propustoměr), s konstantním spádem. Obvykle stanovení

probíhá na neporušených vzorcích odebraných v terénu, je však možné použít i vzorky uměle připravené. Nejprve se vzorky půdy kapilárně nasytí na filtračním papíře. Po dosažení maximální kapilární kapacity se vzorky vloží do permeamtru s konstantním spádem, kde dojde k jejich dosycení. Vzorkem nasycené půdy protéká voda ustálenou rychlostí při konstantním spádu hydraulických výšek. Hodnota nasycené hydraulické vodivosti se vypočítá z ustáleného průtoku pomocí rovnice (5):

$$K_S = \left(\frac{q}{S}\right) \cdot \left(\frac{l}{H}\right) \quad (5)$$

kde:

K_S – nasycená hydraulická vodivost [m/s]

q – ustálený průtok [m^3/s]

S – plocha Kopeckého válečku s odebraným vzorkem půdy [m^2]

l – výška válečku [m]

H – konstantní rozdíl hladin [m]

Ustálený průtok q je definován jako podíl objemu vody, která proteče za čas t (Salač 2018).

U vzorku musí dojít k dokonalému nasycení, jinak se koeficient nasycené hydraulické vodivosti nezměří přesně.

3.3.6 Přístroj s proměnným spádem

Obdobně lze ke zjištění nasycené hydraulické vodivosti v laboratořích použít permeametr s proměnným spádem. Při měření při proměnném spádu dochází ke změně (poklesu) hladiny vody v měřicí trubici přístroje, mění se tedy hydraulická tlaková výška, při které dochází k infiltraci. Po nasycení vzorků půdy ve válečcích na filtračním papíře se umístí válečky do přístroje a měří se čas potřebný k protečení konkrétního měřitelného objemu vody při současném stanovení poklesu hladiny vody v měřicí trubici. Koeficient K_S se vypočítá s využitím Darcyho rovnice integrované přes hydraulické výšky (6).

$$v = K_S \cdot \frac{\Delta h}{l} \quad (6)$$

kde:

v – rychlost poklesu hladiny vody [m/s]

Δh – rozdíl hladiny na začátku a na konci měřeného úseku [m]

l – výška Kopeckého válečku [m]

K_S – nasycená hydraulická vodivost [m/s]

Rychlost poklesu hladiny se vyjádří ve tvaru $v = \frac{\partial h}{\partial t}$, jako rozdíl hladiny vody za určitý čas t [s]. Dosazení do Darcyho rovnice (6) a následná integrace vede k získání výsledné rovnice (7), která se využívá k výpočtu koeficientu:

$$K_s = \frac{l}{t} \ln \left(\frac{h_0}{h_1} \right) \quad (7)$$

kde:

h_0 – úroveň hladiny na počátku měřeného úseku [m]

h_1 – úroveň hladiny na konci měřeného úseku [m] (Salač 2018).

Přístroj s konstantním spádem se používá u hrubozrnných, velmi propustných zemin. Naproti tomu přístroj s proměnným spádem použijeme pro méně propustné zeminy.

Mezi komerčně dostupné přístroje patří přístroj K_{SAT} od společnosti METER group, který používá metody s konstantním spádem i s proměnným spádem (Obrázek 6), nebo permeametr Chameleon společnosti Soilmoisture Equipment Corp.



Obrázek 6 - K_{SAT} . Přístroj k laboratornímu měření nasycené hydraulické vodivosti (METER)

3.4 Nepřímé stanovení nasycené hydraulické vodivosti

3.4.1 Pedotransferové funkce

V mnoha případech, obzvláště pro velká území, jsou terénní a laboratorní stanovení, i přes neustálé zlepšování měřících metod hydraulických vlastností, velmi časově a finančně náročné. V reakci na náročnost experimentálního stanovení byly zavedeny postupy odhadování, které využívají data, která jsou již známá, nebo jsou snadněji měřitelná.

Pedotransferové funkce, zkráceně označované jako PTF, jsou obecně zavedeným termínem pro metody odhadu hydrofyzikálních charakteristik půdy, mezi které náleží i K_S . Termín pedotransferové funkce pochází ze spojení dvou výrazů „transfer function“ a „pedofunction“. Jako první termín „Pedotransfer functions“ použil Bouma (1989) a Minasny et al. (1999) označil proces odhadu jako přeložení dat, která máme, na data, která potřebujeme. PTF nejprve označovaly statistické regresní rovnice, které vyjadřovaly vztahy mezi půdními vlastnostmi (Miháliková 2011).

Současně s rozvojem informačních technologií dochází i k rozvoji pedotransferových funkcí. Navzdory tomuto rozvoji obecný model pro stanovení nasycené hydraulické vodivosti pomocí PTF nelze jednoznačně stanovit.

Kvalita odhadu K_S se vyjadřuje na základě porovnání s naměřenými daty, která ovšem mohou vykazovat značný rozptyl odrážející přirozenou heterogenitu půdy v kombinaci s proměnlivostí K_S v čase a zvolenou metodou měření.

3.4.2 Rozdělení pedotransferových funkcí

Pedotransferové funkce se používají nejen pro odhad nasycené hydraulické vodivosti, ale i dalších hydrofyzikálních charakteristik, retenční čáry nebo nenasyčené hydraulické vodivosti.

Snahou pedotransferových funkcí je odhadnout obtížně získatelné hydraulické, ale i tepelné charakteristiky, nebo jiné vlastnosti půdy ze snadno měřitelných a široce dostupných parametrů půdy (např. zrnitostního složení a objemové hmotnosti suché půdy).

PTF lze klasifikovat podle více kritérií. Například klasifikace na texturní a kontinuální funkce je založená na použitelnosti a charakteru vstupních dat. Klasifikace na bodové a parametrické je podle charakteru odhadovaných dat (Tomasella et al. 2003; Miháliková 2011).

Texturní pedotransferové funkce

Texturní pedotransferové funkce odhadují průměrné hydraulické charakteristiky pro jednotlivé půdní druhy na základě znalosti zrnitosti půdy. Nejprve se data zařadí do 5 zrnitostních tříd (dle FAO – Organizace pro výživu a zemědělství), a poté se prokládají retenční křivky pro získání Mualem van Genuchtenových parametrů v každé zrnitostní třídě (Wösten et al. 1999; Kresic 2006).

Obrázek 7 znázorňuje Mualem-van Genuchtenovy parametry pro 5 tříd (dle FAO), rozdělené na ornici a podorničí a jednou třídou pro organické půdy. Mezi nimi je i nasycená hydraulická vodivost (K_s).

Obrázek 8 znázorňuje 7 hydraulických parametrů pro 12 zrnitostních tříd (dle USDA). Údaje z tabulky využívá program Rosetta (H1) (Schaap et al. 2001).

		θ_r	θ_s	α	n	m	l	K_s
Topsoils Ornice	Coarse	0.025	0.403	0.0383	1.3774	0.2740	1.2500	60.000
	Medium	0.010	0.439	0.0314	1.1804	0.1528	-2.3421	12.061
	Mediumfine	0.010	0.430	0.0083	1.2539	0.2025	-0.5884	2.272
	Fine	0.010	0.520	0.0367	1.1012	0.0919	-1.9772	24.800
	Very Fine	0.010	0.614	0.0265	1.1033	0.0936	2.5000	15.000
Subsoils Podorničí	Coarse	0.025	0.366	0.0430	1.5206	0.3424	1.2500	70.000
	Medium	0.010	0.392	0.0249	1.1689	0.1445	-0.7437	10.755
	Mediumfine	0.010	0.412	0.0082	1.2179	0.1789	0.5000	4.000
	Fine	0.010	0.481	0.0198	1.0861	0.0793	-3.7124	8.500
	Very Fine	0.010	0.538	0.0168	1.0730	0.0680	0.0001	8.235
Organic		0.010	0.766	0.0130	1.2039	0.1694	0.4000	8.000

van Genuchtenovy parametry

Hodnoty nasycené hydraulické vodivosti

Obrázek 7 - Mualem-van Genuchtenovy parametry pro 11 tříd (Wösten et al. 1999).

Texture Class	N	qr cm ³ /cm ³	qs cm ³ /cm ³	log(a) log(1/cm)	log(n) log10	Ks log(cm/day)	Ko log(cm/day)
Clay	84	0.098	0.459	-1.825	0.098	1.169	0.472
C loam	140	0.079	0.442	-1.801	0.151	0.913	0.699
Loam	242	0.061	0.399	-1.954	0.168	1.081	0.568
L Sand	201	0.049	0.390	-1.459	0.242	2.022	1.386
Sand	308	0.053	0.375	-1.453	0.502	2.808	1.389
S Clay	11	0.117	0.385	-1.476	0.082	1.055	0.637
S C L	87	0.063	0.384	-1.676	0.124	1.120	0.841
S loam	476	0.039	0.387	-1.574	0.161	1.583	1.190
Silt	6	0.050	0.489	-2.182	0.225	1.641	0.524
Si Clay	28	0.111	0.481	-1.790	0.121	0.983	0.501
Si C L	172	0.090	0.482	-2.076	0.182	1.046	0.349
Si Loam	330	0.065	0.439	-2.296	0.221	1.261	0.243

Obrázek 8 - Rosetta – van Genuchtenovy hydraulické parametry (USDA 2019).

Kontinuální pedotransferové funkce

Kontinuální pedotransferové funkce představuje sada regresních rovnic, které zkoumají závislost parametru v dané tlakové výšce na hodnotách vstupních dat (prediktorů). V práci Wösten et al. (1995) jsou prezentovány PTF na základě dat shromážděných v databázi hydraulických vlastností evropských půd HYPRES. Jsou zde odvozeny rovnice pro výpočet van Genuchtenových parametrů pro odhad retenčních křivek evropských půd.

„Kontinuální pedotransferové funkce charakterizují hydrofyzikální vlastnosti konkrétního stanoviště, protože využívají proměnné, které byly změřeny přímo na daném stanovišti.“ (Špongrová 2003, str. 22).

K_S je v práci Wösten et al. (1995) vyjádřena následující rovnicí (8). Většina kontinuálních funkcí má podobnou formu. Charakteristika K_S nevykazuje obvykle normální rozdělení, nejčastěji se používá logaritmická transformace (\ln při základě e nebo \log při základě 10)

$$\ln(K_S) = 9,5 - 1,471D^2 - 0,688 OM + 0,0369 OM^2 - 0,332\ln(CS) \quad (8)$$

kde:

D – objemová hmotnost suché půdy [g/cm^3]

OM – organická hmota [%]

CS – jíl + prach [%] (Wösten et al. 1995)

Bodové pedotransferové funkce

Bodové pedotransferové funkce odhadují vlhkost v předem určených hodnotách potenciálů – nejčastěji v bodech polní vodní kapacity a bodu vadnutí (-33 kPa, -1500 kPa), které jsou důležité pro určení množství vody, která je dostupná pro rostliny.

Parametrické pedotransferové funkce

Tyto funkce, které se hojně používají v modelování transportních procesů, jsou ve formě rovnic s určitým počtem parametrů. Nejsou závislé na měření vlhkosti v daných potenciálech, a proto lze při jejich tvoření použít data naměřená v různých potenciálech. Jedná se například o rovnice Brookse a Coreyho, či van Genuchtenovu rovnici (Miháliková et al. 2013).

3.4.3 Vývoj pedotransferových funkcí

Mezi prvními publikacemi, kde se používají jednodušeji zjistitelné veličiny (textura půdy) ke stanovení nasycené hydraulické vodivosti využitelné například k modelování pohybu půdní vody při absenci měřených hydraulických parametrů, je práce Cosbyho et al. (1984).

Dále například Puckett, který provedl studii s cílem určit vztahy mezi fyzikálními, mineralogickými a půdně hydraulickými vlastnostmi půd Ultisols (podle půdní klasifikace USDA) (Soil Survey Staff 1999) vyvinutých na nezpevněných sedimentech pobřežních nížin. Retenční křivka, nasycená hydraulická vodivost, kvantitativní mineralogie i distribuce velikosti částic byly určeny horizontem pro sedm pedonů a vztahy mezi nimi byly analyzovány regresní analýzou. Procentuální zastoupení jemného písku, písku a jílu vysoce korelovalo s nasycenou hydraulickou vodivostí. Korelační koeficienty $>0,95$ byly získány pro regresní rovnice odhadující retenční křivku z údajů o velikosti půdních částic. Mezi půdními hydraulickými

a mineralogickými vlastnostmi jílové frakce nebyly nalezeny žádné významné korelace (Puckett et al. 1985).

Dalšími v literatuře používanými funkcemi jsou PTF od Saxtona. Ve své práci se věnuje především potenciálu půdní vody (určování retenční křivky), ale v závěru neopomíná ani K_S . Zmiňuje, že první modely pro predikci vodivosti půdy byly buď empirické nebo teoretické vztahy pracující s velikostí půdních částic nebo pórovitostí půdy. Tyto modely byly obvykle založeny na vodivosti při nasycení a nebylo možné je snadno aplikovat na nenasyčené podmínky (Saxton & Rawls 2006).

V roce 2014 publikovali autoři B. Tóth, Weynants, Nemes, Makó, Bilas, G. Tóth kontinuální PTF pro střední a východní Evropu. Do té doby existovaly kontinuální PTF využívající databázi HYPRES, která obsahovala údaje o půdách především ze západní Evropy. Autoři využili databáze EU-HYDI (EUropean HYdropedological Data INventory) (Weynants et al. 2013). K vytvoření kontinuálních funkcí použili lineární regresi, i regresních stromů (regression trees). Funkce autoři rozdělili do dvou kategorií: TEST_BASIC a TEST_CHEM+. TEST_BASIC používá jako prediktory zrnitostní složení, objemovou hmotnost suché půdy a obsah organického uhlíku. TEST_CHEM+ používá jako prediktory také chemické složení půdy. Výsledky odhadu nasycené hydraulické vodivosti vycházely s chybou RMSE v rozmezích 0,90-1,36 $\log K_S$ (cm/den). PTF většinou K_S předpovídaly vyšší, než byla změřena. Model využívající regresní stromy je podle autorů využitelný v praxi, neboť v průměru poskytoval odhady s RMSE 1,05 $\log K_S$ (cm/den) (Tóth et al. 2014). Podobných výsledků bylo dosaženo i ve studii (Lilly et al. 2008), kde průměrná RMSE odhadované K_S vycházela 0,97 $\log K_S$ (cm/den).

Dále Aguilera et al. publikovali v roce 2022 studii „Building inexpensive topsoil saturated hydraulic conductivity maps for land planning based on machine learning and geostatistics“ ve které popisují použití umělé neuronové sítě (ANN) k odhadu K_S pro orní vrstvu ostrova Mallorca. Autoři otestovali různé algoritmy strojového učení, jmenovitě k-nearest neighbors (k-nejbližší sousedé), random forest (RF, náhodný les), boosted regression trees (BRT), artificial neural networks (ANN, umělé neuronové sítě). Nicméně pouze algoritmus ANN poskytl přijatelné výsledky, přibližující se široce používaným PTF (např. Puckett et al. 1985, Cosby et al. 1984) (Aguilera et al. 2022).

Pedotransferové funkce tvoří pomyslný most mezi dostupnými údaji o půdě a požadovanými informacemi. Vývoj a další zpřesňování jejich výstupů bude stěžejním pro modelování transportu půdní vody i rozpuštěných látek ve velkém měřítku.

3.5 Vybrané metody nepřímého stanovení nasycené hydraulické vodivosti

3.5.1 Úvod

Reálné aplikace PTF jsou často komplikovány potřebou velmi specifických dat. V roce 1998 byly zavedeny PTF s hierarchickým uspořádáním autory Schaap et al., kde jsou funkce seřazeny podle úrovně dostupnosti vstupních dat (zrnitost, zrnitost a objemová hmotnost apod.), což uživateli poskytuje více možností v použití dostupných dat (Miháliková 2011).

Terénní morfologie	Terénní odhad	Zrnitostní rozbor	Organická hmota	ρ_d	ϕ	θ_s	θ_{-33}	θ_{-1500}	K_s	$K(\theta)$
--------------------	---------------	-------------------	-----------------	----------	--------	------------	----------------	------------------	-------	-------------

Obrázek 9 - Úrovně dostupnosti vstupních dat. Vlevo jsou nejsnáze dostupná data. (Miháliková 2011)

Z obrázku 9 lze soudit náročnost stanovení některých veličin laboratorním, či terénním měřením. Obecně platí, že by se neměly odhadovat veličiny, pokud je jejich měření snazší než měření jejich prediktorů (Miháliková 2011). Také obecně platí, že čím více použijeme vstupních prediktorů, tím se zvyšuje přesnost odhadu hledaných hydraulických veličin.

Nicméně většina PTF poskytuje nejlepší výsledky při používání dat autorů, kteří dané funkce sami odvodili. Běžný uživatel se při zjišťování hydraulických veličin může potýkat také s problémem chybějících některých vstupních dat, nebo s výše zmíněnou nepřesností PTF od jiných autorů.

Jedním z prvních programů ke stanovení hydraulických veličin, který využívá strojového učení, konkrétně umělé neuronové sítě je program Rosetta (Schaap et al. 2001). V porovnání s tradičními pedotransferovými funkcemi má využití neuronové sítě výhodu v tom, že nevyžaduje žádný předchozí koncept.

V této práci bylo testováno použití metody strojového učení v aplikaci Araya a Ghezzehei (2019), využívajícího databáze amerických půd (které jsou v některých případech značně odlišné od středoevropských) a její výsledky porovnány s měřenými daty a také s výsledky již zavedeného programu Rosetta. Níže je uveden detailnější popis těchto dvou vybraných programů.

3.5.2 Program Rosetta

Program Rosetta – pojmenovaný po granodioritové desce, která umožnila překlad starověkých egyptských hieroglyfů do staré řečtiny, umožňuje uživatelům stanovit parametry jako retenční křivku a nasycenou hydraulickou vodivost, společně s nenasycenou hydraulickou vodivostí. Pro odhad retenční křivky a nasycené hydraulické vodivosti jsou modely s označením H1, H2, H3, H4 a H5, pro nenasycenou pak modely s označením C2-H1 až C2-H5.

Rosetta je program primárně určený pro operační systém Windows, v jiných operačních systémech (např. LINUX) lze využít verzi přes příkazový řádek. Rosetta je zdarma ke stažení přes webové rozhraní, nebo zabudovaná v softwaru Hydrus-1D, popř. 2 D/3 D.

Program má hierarchicky rozdělené modely H1 – H5, aby optimálně využíval zadaná vstupní data. Jednotlivé úrovně jsou znázorněny v tabulce 1.

Prostředí programu Rosetta zobrazuje obrázek 11.

Tabulka 1 - Modely H1-H5 programu Rosetta.

Model	Vstupní data
H1	Půdní textura (USDA zrnitostní třídy)
H2	Procentuální zastoupení písku, prachu a jílu
H3	Procentuální zastoupení písku, prachu a jílu, objemová hmotnost suché půdy
H4	Procentuální zastoupení písku, prachu a jílu, objemová hmotnost suché půdy, polní vodní kapacita při 330 cm vodního sloupce (h=33 kPa)
H5	Procentuální zastoupení písku, prachu a jílu, objemová hmotnost suché půdy, polní vodní kapacita při 330 cm vodního sloupce (h=33 kPa), bod vadnutí při 15 000 cm vodního sloupce (h=1500 kPa)

Podle Schaapa et al. (2001) znalost více prediktorů určuje přesnější výsledky hledané hydraulické veličiny. Jak uvádí Schaap et al. (2001) odhady retenční křivky a hydraulické vodivosti modelu H5 dosahují přesnějších výsledků, než je tomu v případě modelu H3. Nicméně pokud některá data chybí, modely s označením H1, H2, H3 jsou stále použitelné.

Model H1 pracuje na funkci jednoduché vyhledávací tabulky s průměrnými hodnotami hydraulických parametrů pro každou zrnitostní třídu klasifikačního systému USDA. Další modely využívají ANN v kombinaci se softwarem bootstrap (Schaap et al. 2001).

V rámci diplomové práce jsou použity modely H2 a H3, které jsou dále označovány jako Rosetta (SSC) pro model H2 a Rosetta (SSC+BD) pro model H3.

Umělá neuronová síť (ANN)

Umělá neuronová síť je jedním z výpočetních modelů, které se používají v umělé inteligenci. Počátek vzniku se datuje do roku 1943, kdy Warren McCulloch a Walter Pitts vytvořili jednoduchý matematický model neuronu – základní buňky nervové soustavy. Přestože autoři nepočítali s využitím svého matematického modelu v praxi, měla jejich práce velký ohlas. V roce 1957 Frank Rosenblatt zobecnil model neuronu i pro parametry z množiny reálných čísel (na tzv. perceptron). Pro původní model musely být číselné hodnoty z množiny čísel {0, 1}. Frank Rosenblatt je také autorem jedné z prvních knih o neurovýpočtech Principles of Neurodynamics.

První neuropočítač sestrojili v letech 1957 a 1958 v návaznosti na svůj výzkum vědci Frank Rosenblatt a Charles Wightman. Původním cílem bylo, aby počítač dokázal rozpoznávat znaky (např. klasifikovat, o jaké se jedná písmeno). Jednalo se tak o simulaci rozpoznávání vzorů na lidské sítnici. Daný znak byl promítán na světelnou tabuli, dále byl snímán polem

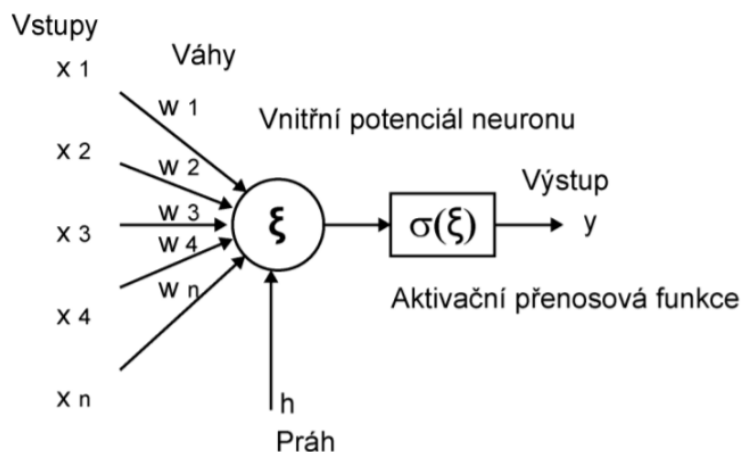
fotovodičů (20x20). Na rozdíl od programovatelných počítačů, neuropočítač nebyl závislý na zapojení. Náhodným zapojením (jednotlivé perceptrony bylo možné spojit libovolným způsobem) Rosenblatt dokazoval schopnost perceptronu učit se požadované vzory bez přesného zapojení drátů. Rosenblatt je dodnes považován za zakladatele tohoto směru.

Umělé neuronové sítě se používají nejen k rozpoznávání a kompresi obrazu a zvuku, ale také k filtrování spamů, předvídání vývoje časových řad, nebo v lékařství k objasnění a prohlubování znalostí o funkci nervových soustav organismů.

I přestože byla původně neurofyziologie modelem pro neuronové sítě, jsou ANN nyní dále rozvíjeny bez ohledu na to, zda modelují mozek, či nikoliv.

Základní buňkou nervové soustavy je neuron, který slouží k přenosu, zpracování a uchování informací, které jsou nezbytné pro fungování organismu. Biologický neuron má vstupní a výstupní přenosové kanály: dendrity a axon. Axon se větví na jednotlivé terminály zakončené blánou, která se dotýká dendritů jiných neuronů. K přenosu informace pak dochází na mezineuronovém rozhraní, tzv. synapsi. Synapse lze rozdělit na dva typy: excitační a inhibiční. Excitační synapse umožňují rozšíření vzruchu v nervové soustavě, inhibiční pak jeho útlum.

Matematický model neuronové sítě vychází ze základní jednotky – formálního neuronu. Jeho schéma je znázorněno na obrázku 10. Neuron má n reálných vstupů, které modelují dendrity. Vstupy jsou ohodnoceny synaptickými váhami (na obrázku např. $w_1, w_2, w_3, w_4, \dots, w_n$), které mohou být i záporné. Stejně jako u biologického neuronu může docházet nejen k excitaci, ale i k inhibici. Synaptické váhy určují vstupní propustnost. Vnitřní potenciál neuronu představuje vážená suma vstupních hodnot. Pokud hodnota vnitřního potenciálu neuronu dosáhne prahové hodnoty, indikuje výstup neuronu, (Schaap & Leij 1998; Volná 2002).



Obrázek 10 – Matematický model neuronu.

(Zdroj: <https://portal.matematickabiologie.cz/index.php?pg=analyza-a-hodnoceni-biologickych-dat--umela-inteligence--neuronove-site-jednotlivy-neuron--jednotlivy-neuron--matematicky-model-a-aktivni-dynamika-neuronu>)

Neuronová síť vzniká propojením neuronů, každý s malou lokální pamětí, které si mezi sebou předávají data.

Hydraulické parametry

Program Rosetta je schopen odhadnout parametry retenční křivky a nasycené hydraulické vodivosti podle upravené rovnice van Genuchtena (9): (USDA 2019)

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha h)^n]^{1 - \frac{1}{n}}} \quad (9)$$

kde:

$\theta(h)$ – objemová vlhkost [cm^3/cm^3] jako funkce tlakové výšky h [cm]

θ_r – vlhkost reziduální [cm^3/cm^3]

θ_s – vlhkost nasycení [cm^3/cm^3]

α – empirický koeficient (>0) [cm^{-1}]

n – empirický koeficient (>1) [-]

Rovnice vychází z původního vztahu pro retenční křivku (10):

$$\theta_e = \frac{1}{[1 + (\alpha h)^n]^{1 - \frac{1}{n}}} \quad (10)$$

kde:

θ_e – efektivní vlhkost [-]

α – empirický koeficient (>0) [cm^{-1}]

n – empirický koeficient (>1) [-] závislý na distribuci velikosti pórů

h – tlaková výška [m]

Pro odhad nasycené hydraulické vodivosti program používá upravenou rovnici (11) podle Mualema:

$$K(S_e) = K_0 S_e^L \left\{ 1 - \left[1 - S_e^{n/(n-1)} \right]^{1 - \frac{1}{n}} \right\}^2 \quad (11)$$

kde:

S_e – relativní nasycení [-]

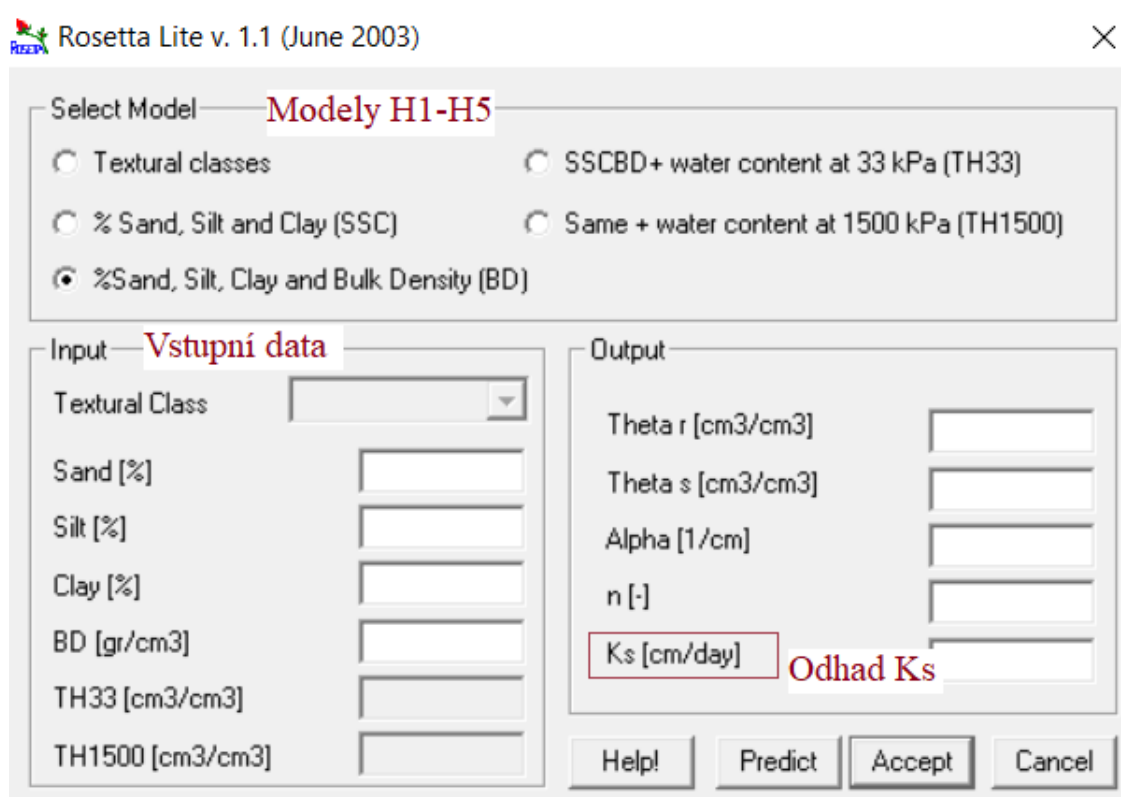
K_0 – bod nasycení [cm/den], který je podobný, ale nemusí být nezbytně stejný, koeficientu nasycené hydraulické vodivosti K_s .

L – empirický koeficient, který podle Mualema (1976) je roven 0,5

Rosetta předpovídá parametr L ve většině případů záporný. Ačkoliv záporná hodnota parametru L vede k teoretickým komplikacím, program Rosetta takto dává přesnější výsledky (Schaap et al. 2001).

Relativní nasycení S_e lze vyjádřit pomocí rovnice (12):

$$S_e = \frac{\theta(h) - \theta_r}{\theta_s - \theta_r} \quad (12)$$



Obrázek 11 - Prostředí programu Rosetta.

3.5.3 PTF App - aplikace autorů S. N. Araya a T. A. Ghezzehei

Ačkoliv v posledních letech docházelo k velkému rozvoji informačních technologií, odhad nasycené hydraulické vodivosti pomocí strojového učení zůstal (oproti jiným odvětvím) pozadu. S. N. Araya a T. A. Ghezzehei proto společně navrhli a vytvořili nový program k odhadu K_S , který využívá různých samoučících algoritmů – strojového učení.

Existují studie, u kterých se využívalo strojového učení k řešení různorodých problémů vztahujících se k hydraulickým parametrům. Některé například využívají algoritmů KNN (k-nearest neighbors algorithm, překládáno jako algoritmus k-nejbližších sousedů) k odhadu hydraulických parametrů. Jednu ze studií publikovala např. M. Miháliková (Miháliková et al. 2014), nebo A. Nemes (Nemes et al. 2007).

Program je vytvořen v programovacím jazyku a prostředí R, které jsou určeny pro statistickou analýzu dat a jejich grafické zobrazení. Lze ho najít na stránkách S. N. Araya: https://github.com/saraya209/soil_ksat. Program je k dispozici uživatelům bezplatně, avšak bez využití pro komerční účely. K jeho stažení a následnému spuštění je potřeba dobrá znalost informačních technologií. Na obrázku 12 lze vidět spuštěný program k odhadu K_S společně s prostředím R studia.

V programu se setkáme se dvěma druhy učících algoritmů; algoritmem boosted regression trees (BRT) a algoritmem random forest (RF). Základem obou algoritmů BRT i RF jsou tzv. regression trees algoritmy. Několik studií ukázalo, že pro odhad půdních i hydraulických vlastností, jsou dané algoritmy velmi přesné (Araya & Ghezzehei 2019).

Select PTF Model Výběr modelu

BRT 3-2

Load Model

Run Mode: Single File Upload

Soil Properties Vstupní data

Single Input

Soil Separates (%)

Clay Silt

Sand

Sand Subcategories (%)

Very coarse sand

Fine sand

Coarse sand

Very fine sand

Medium sand

Bulk Density (g cm⁻³)

Organic Carbon (%)

Model Name: BRT 3-2. (A gradient boosted model with gaussian loss function.)
Required Input: Clay, Silt, Sand, BD, OC
Run Mode: Manual Input.
Model Loaded?: TRUE

Model Prediction [cm/day]

Show 10 entries

Clay	Silt	Sand	BD	OC	Predicted_Ks	Model_Name
30	35	35	1.4	1.01	5.33503	BRT 3-2

Odhad Ks

Showing 1 to 1 of 1 entries

Remarks

Vstupní data pro modely 7-0, 7-1, 7-2

Objemová hmotnost suché půdy

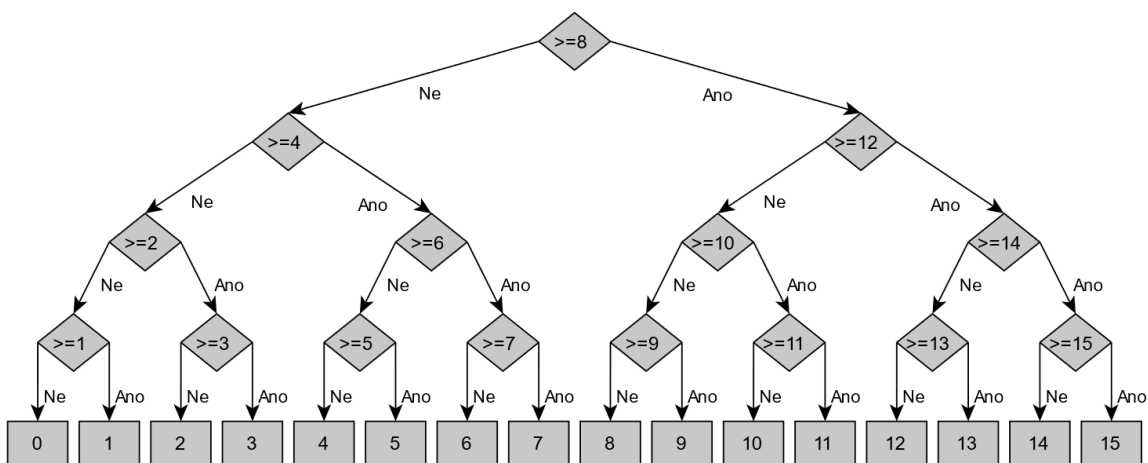
Organický uhlík

Obrázek 12 - Program ke stanovení nasycené hydraulické vodivosti v prostředí R studia.

Modely boosted regression trees

Model BRT neboli boosted regression tree, vznikl kombinací algoritmů decision tree - rozhodovacích stromů, a boosting methods.

Rozhodovací strom je jedním z často používaných nástrojů rozhodovací analýzy. Je rychlý, přehledný a výsledky jsou snadno interpretovatelné. Skládá se z jednotlivých rozhodovacích uzlů, ze kterých vychází rozhodovací, nebo situační hrany. Rozhodovací hrany představují dostupná řešení daného problému, situační hrany pak znázorňují možné důsledky. Daný název algoritmus dostal nejspíše proto, že zde probíhá rozhodovací proces, při kterém dochází k větvení řešení i jejich důsledků, a tak jeho náskres může připomínat rozvětvený strom (Straková 2016). Na obrázku 13 je znázorněno schéma rozhodovacího stromu pro nalezení celého čísla v rozsahu 0-15, který využívá metodu půlení intervalů.



Obrázek 13 - Rozhodovací strom pro nalezení čísla v rozsahu 0-15.

(Zdroj: https://popelka.ms.mff.cuni.cz/~lessner/mw/index.php/U%C4%8Debnice/Informace/Rozhodovac%C3%AD_stromy_a_chytr%C3%A9_ot%C3%A1zky)

Iterativní procedura boosting má pak za cíl zlepšit klasifikační přesnost libovolného strojového učení tím, že pružně mění rozložení učicích dat. Na začátku přiřadí všem vzorům stejnou váhu, později vyčlení vzory špatně klasifikované tak, že jim zvýší váhu a tím je opět zařadí do učicí množiny. Správně klasifikovaným datům váhu snižuje (Schapire 1999).

BRT modely se tak hodí pro „méně vyčištěná“ data, se kterými se během odhadu nasycené hydraulické vodivosti můžeme setkat. Vstupní data mohou pocházet z různých zdrojů a mohou být získány odlišnými metodami měření, proto může docházet k jejich nekonzistentnosti.

Program obsahuje 6 hierarchicky sestavených modelů podle vstupních parametrů (Tabulka 2).

Tabulka 2 - Modely BRT programu autorů Araya, Ghezzehei.

Model	Vstupní data
BRT 3-0	Procentuální zastoupení písku, prachu a jílu
BRT 3-1	Procentuální zastoupení písku, prachu a jílu, objemová hmotnost suché půdy
BRT 3-2	Procentuální zastoupení písku, prachu a jílu, objemová hmotnost suché půdy, obsah organického uhlíku
BRT 7-0	Procentuální zastoupení prachu a jílu, velmi jemného písku (VFS), jemného písku (FS), středního písku (MS), hrubého písku (COS), velmi hrubého písku (VCOS)
BRT 7-1	Procentuální zastoupení prachu a jílu, velmi jemného písku (VFS), jemného písku (FS), středního písku (MS), hrubého písku (COS), velmi hrubého písku (VCOS), objemová hmotnost suché půdy
BRT 7-2	Procentuální zastoupení prachu a jílu, velmi jemného písku (VFS), jemného písku (FS), středního písku (MS), hrubého písku (COS), velmi hrubého písku (VCOS), objemová hmotnost suché půdy, obsah organického uhlíku

Modely random forest

Model RF neboli random forest, v českém překladu „náhodný les“ je učící se algoritmus navržený L. Breimanem v roce 2001, který je velmi úspěšný při více-účelových klasifikačních a regresních metodách. Byl úspěšně použit při řešení rozličných problémů jako např. při předpovědích kvality ovzduší, v ekologii, v rozpoznávání trojrozměrných objektů, v bioinformatice, i v ekonometrii (Biau & Scornet 2016).

Pro metodu náhodného lesa se používají jednotlivé rozhodovací stromy představené výše. Předpověď každého pozorování je poté průměr ze všech stromů. Les je soubor stromů (T_1, T_2, \dots, T_n), výsledná regrese je zapsána pomocí rovnice (13):

$$f_{\text{rf}}(\mathbf{x}) = \frac{1}{n} \left(\sum_{i=1}^n T_i(\mathbf{x}) \right) \quad (13)$$

Stromy se „učí“ tak, že hledají vhodná vstupní data (prediktory), která zajistí nejlepší rozdělení, při kterém v modelu vzniká nejmenší chyba. Jednotlivé stromy využívají tzv. bootstrapové výběry (náhodná tréninková data) z tréninkové množiny. Při každém rozdělení je brána v úvahu pouze malá skupina prediktorových proměnných, což zajišťuje minimální korelaci mezi stromy (Araya & Ghezzehei 2019).

Program obsahuje 6 hierarchicky sestavených modelů podle vstupních parametrů (Tabulka 3). Vstupní data jsou shodná s modely BRT.

Tabulka 3 - Modely RF programu autorů Araya, Ghezzehei.

Model	Vstupní data
RF 3-0	Procentuální zastoupení písku, prachu a jílu
RF 3-1	Procentuální zastoupení písku, prachu a jílu, objemová hmotnost suché půdy
RF 3-2	Procentuální zastoupení písku, prachu a jílu, objemová hmotnost suché půdy, obsah organického uhlíku
RF 7-0	Procentuální zastoupení prachu a jílu, velmi jemného písku (VFS), jemného písku (FS), středního písku (MS), hrubého písku (COS), velmi hrubého písku (VCOS)
RF 7-1	Procentuální zastoupení prachu a jílu, velmi jemného písku (VFS), jemného písku (FS), středního písku (MS), hrubého písku (COS), velmi hrubého písku (VCOS), objemová hmotnost suché půdy
RF 7-2	Procentuální zastoupení prachu a jílu, velmi jemného písku (VFS), jemného písku (FS), středního písku (MS), hrubého písku (COS), velmi hrubého písku (VCOS), objemová hmotnost suché půdy, obsah organického uhlíku

3.5.4 Vstupní data

Textura

„Velikost minerálních částic v půdě, respektive zastoupení jednotlivých velikostních kategorií zásadně ovlivňuje fyzikální a chemické vlastnosti půdy.“ (Šimek et al. 2019, s. 326).

Jsou vytvořeny různé systémy třídění částic podle jejich velikosti. Nejjednodušší systém tvoří kategorie jíl, prach, písek a štěrk.

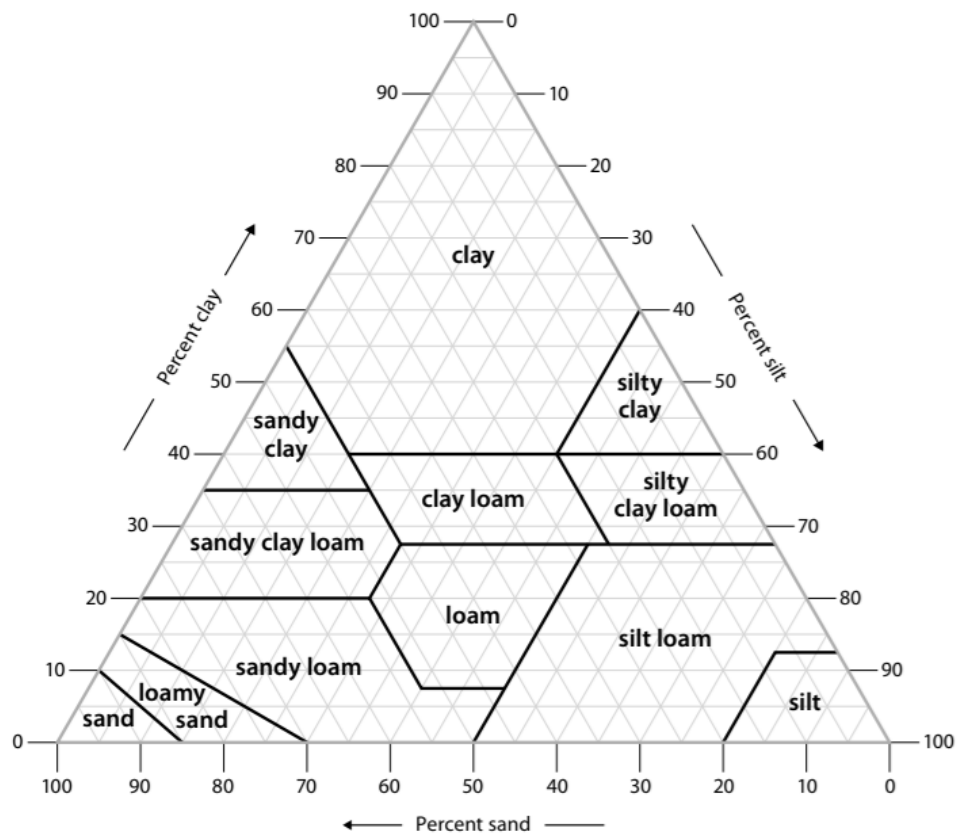
Oba programy vychází z amerického systému (Obrázek 14) a rozlišují 12 zrnitostních tříd, pro které je v ČR zavedeno označení podle Němečka (2001) (Tabulka 4).

Velikost částic se v různých systémech (mezinárodní systém, americký systém, v ČR dle Kopeckého) může lišit. Na obrázku 15 jsou proto uvedeny průměry částic a jejich zařazení (podle USDA), které je použito i v této práci.

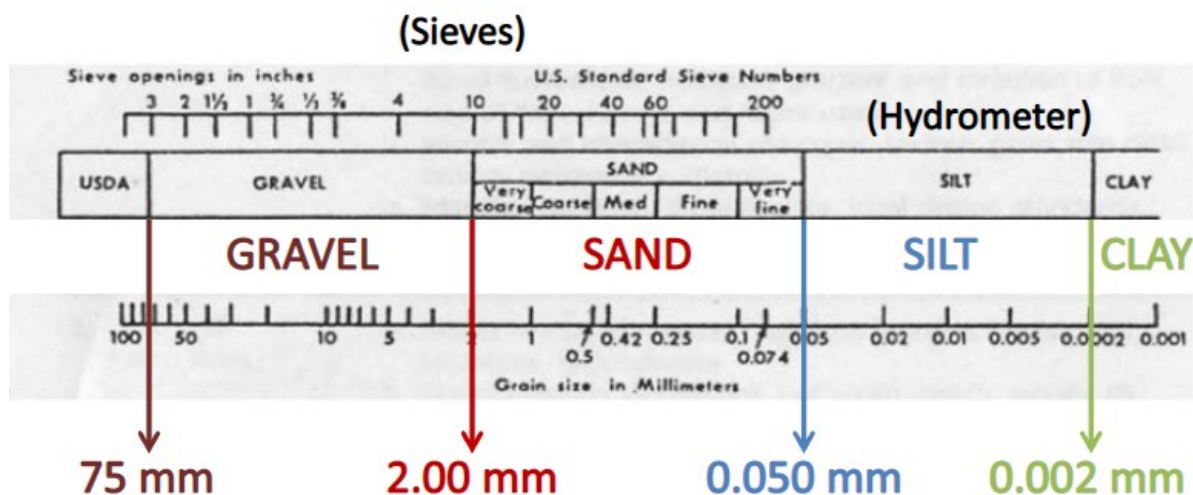
Podle textury se také rozlišují tři hlavní půdní druhy: půdy jílovité, hlinité a písčité. Pro jílovitou půdu se používá termín těžká půda, pro hlinitou střední půda a pro písčitou lehká půda.

Tabulka 4 - Zrnitostní třídy. Označení tříd podle USDA a jejich český ekvivalent.

ZRNITOSTNÍ TRÍDY	
USDA	Označení dle Němečka (2001)
Clay	Jíl
Clay Loam	Jílovitá hlína
Loam	Hlína
Loamy Sand	Hlinitý písek
Sand	Písek
Sandy Clay	Písčité jíl
Sandy Clay Loam	Písčité jílovité hlína
Sandy Loam	Písčité hlína
Silt	Prach
Silty Clay	Prachovité jíl
Silty Clay Loam	Prachovité jílovité hlína
Silty Loam	Prachovité hlína



Obrázek 14 - Trojúhelníkový diagram na stanovení druhu půdy podle obsahu jílu, prachu a písku (USDA 1951).



Obrázek 15 – Velikostní kategorie půdních minerálních částic USDA (J. Fox, Ohio EPA); Zrnitostní kategorie: clay = jííl, silt = prach, sand = písek (very fine sand = velmi jemný písek, fine sand = jemný písek, medium sand = střední písek, coarse sand = hrubý písek, very coarse sand = velmi hrubý písek, gravel = štěrk)

Objemová hmotnost půdy

Objemová hmotnost půdy ρ_b je definována jako poměr hmotnosti půdy k jejímu objemu (včetně pórů) v neporušeném stavu. Jednotkou je zpravidla g/cm^3 , nebo kg/m^3 .

Objemovou hmotnost půdy lze rozdělit na neredukovanou a redukovanou (objemová hmotnost suché půdy). Neredukovaná objemová hmotnost se určuje přímo po odběru, protože závisí na vlhkosti půdy. Vzorky půdy se odebírají do Kopeckého válečků, nejčastěji s objemem 100 cm^3 , nebo 250 cm^3 . Objemová hmotnost suché půdy se poté vypočítá z vysušeného vzorku. Vzorek minerální půdy sušíme v laboratoři při $105 \text{ }^\circ\text{C}$ až do konstantní hmotnosti (Hillel 1998; Fiala et al. 2018).

Objemovou hmotnost suché půdy ρ_b určíme pomocí vztahu (14):

$$\rho_b = \frac{m_z}{V_s} \quad (14)$$

kde:

m_z – hmotnost suché půdy [g]

V_s – objem vzorku [cm^3]

Nejnižší objemovou hmotnost mají organické půdy (ty se suší při nižší teplotách $60\text{-}80 \text{ }^\circ\text{C}$), jinak se objemová hmotnost většiny půd pohybuje v rozmezí od $1,1$ do $1,6 \text{ g/cm}^3$ (Šimek et al. 2019).

Anglické označení pro objemovou hmotnost půdy je *bulk density*, zkratka BD (zkratka používána např. v programu Rosetta).

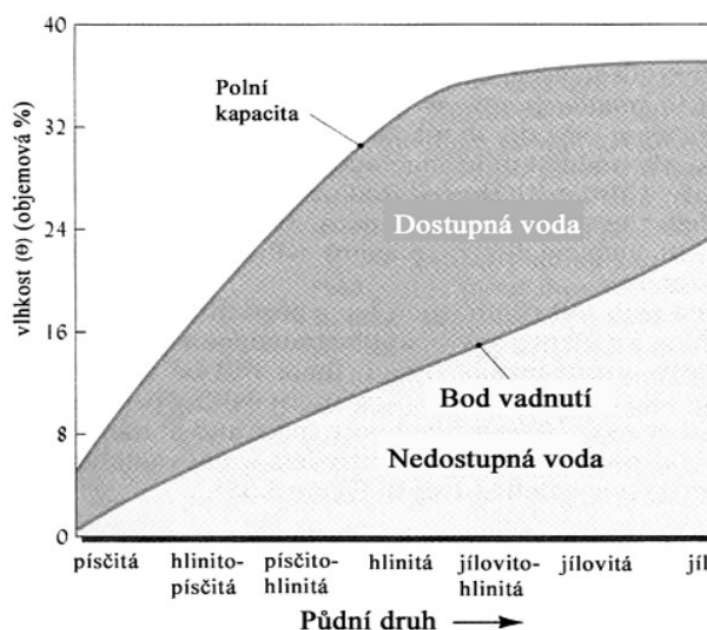
Půdní hydrolimity

„Půdní hydrolimity jsou charakteristiky vyjadřující určité konkrétní stavy půdní vlhkosti vztažené buď k prospívání vegetace, či k poutání vody v konkrétních pórech či konkrétními silami. Jsou charakterizovány vlhkostí půdy a pF hodnotami.“ (Pavlů 2018, s.37)

V programu Rosetta se můžeme setkat se dvěma aplikovanými hydrolimity – bodem vadnutí a polní vodní kapacitou.

Bod vadnutí popisuje dolní hranici využitelnosti půdní vody vegetací. Je to minimální hranice fyziologicky využitelné vody, nebo také půdní vlhkost, při které jsou rostliny trvale nedostatečně zásobeny vodou a trvale vadnou. Hodnota je stanovena domluvou – 1500 kPa. Čím více jemných částic půda obsahuje (např. půdy jílovité, jílovito-hlinité), tím více poutá vody nedostupné. Obrázek 16 ukazuje, že směrem od písčité půdy k jílovité se zvyšuje hodnota bodu trvalého vadnutí. V programu Rosetta je bod vadnutí označen TH1500.

Polní vodní kapacita popisuje horní hranici využitelnosti půdní vody. Je to ustálený stav vlhkosti půdního profilu po nadměrném zavlažení shora. Její hodnota je stanovena domluvou – 33 kPa (Pavlů 2018). Po silném zavlažení může půda obsahovat více vody, než je její polní vodní kapacita, avšak je to voda nadbytečná, která působením gravitačních sil odtéká směrem k hladině podzemní vody. V programu Rosetta nalezneme pro polní vodní kapacitu označení TH33.



Obrázek 16 - Hydrolimity – závislost na půdním druhu a objemové vlhkosti půdy (Pavlů 2018)

Organický uhlík

Obsah organických látek v půdě je další rutinně měřenou veličinou, na které závisí vlastnosti půdy. „Je to soubor všech neživých organických látek nacházejících se na povrchu půdy a v půdě. Půdní organická hmota je tvořena rostlinnými a živočišnými zbytky a odumřelými buňkami mikroorganismů, které jsou v různém stadiu rozkladu a přeměny na nové látky.“ (Šimek et al. 2019, s. 270).

Má méně přímou, ale důležitou roli na strukturu půdy. Obsah organické hmoty ovlivňuje strukturu půdy do značné míry kvůli jejímu vlivu na půdní soudržnost, stabilitu a pórovitost. Obecně zvýšení organické hmoty má za následek zvýšení půdní soudržnosti a stability, a tedy i zvýšení koeficientu nasycené hydraulické vodivosti. Studie (Rawls et al. 2003) ukazuje, že účinky organické hmoty na půdní hydrauliku jsou podobné účinkům jílu. Vysoký obsah jílu snižuje účinky zvýšeného obsahu organické hmoty. Na druhou stranu organická hmota podstatně ovlivňuje fyzikální chování půdy (např. pórovitost) pouze v případě, že obsah jílu je nad prahovou hodnotnou vzhledem k uhlíku C (Araya & Ghezzehei 2019).

Organický uhlík komplexní (C_x)

Pojem komplexního organického uhlíku zavedli Dexter et al. v roce 2008, aby lépe popsali vliv organické hmoty na fyzikální chování půdy. Komplexní organický uhlík je poměr uhlíku, se kterým tvoří komplexy frakce jílu (Cl) a vypočítá se z předpokladu, že 1 gram uhlíku je v komplexu s n gramy jílové hmoty. Pro dostatečně vysoký obsah jílu ($jíl > n \cdot C$) lze veškerý uhlík považovat za komplexní. C_x se vypočítá ze vztahu (15):

$$C_x = \min \left[C, \frac{Cl}{n} \right] \quad (15)$$

Dexter et al. zjistili, že $n=10$ nejlépe popsalo fyzické chování studovaných půd z francouzských a polských databází. V programu je proto použit stejný poměr pro výpočet C_x (Dexter et al. 2008).

Běžný uživatel do programu autorů Araya, Ghezzehei zadává pouze obsah organického uhlíku (v procentech). Ze znalosti procentuálního zastoupení jednotlivých frakcí a obsahu organického uhlíku program určí pro další potřebné výpočty komplexní organický uhlík (Araya & Ghezzehei 2019). Program Rosetta tohoto parametru nevyužívá.

4 Metodika

4.1 Zdroje dat

Data použitá v diplomové práci pocházejí ze dvou zdrojů, z již dříve shromážděných dat v rámci databáze HYPRESCZ (Miháliková et al. 2013) a z nedávných terénních průzkumů experimentálních lokalit realizovaných na Katedře vodních zdrojů ČZU v Praze. Celkem data pocházejí ze 13 lokalit v rámci celé České republiky, z toho nedávné průzkumy probíhaly na třech lokalitách: experimentální plochy VÚRV v Praze – Ruzyni, experimentální plochy VÚMOP v okrese Pardubice (lokality Strašov) a v lokalitě Uherčice v okrese Břeclav.

Databáze dat HYPRESCZ vychází z mezinárodní databáze HYPRES – HYdraulic Properties of European Soils. Pro Českou republiku bylo shromážděno více než 2000 vzorků (Miháliková et al. 2013), ovšem pouze malá část datových souborů obsahovala údaje

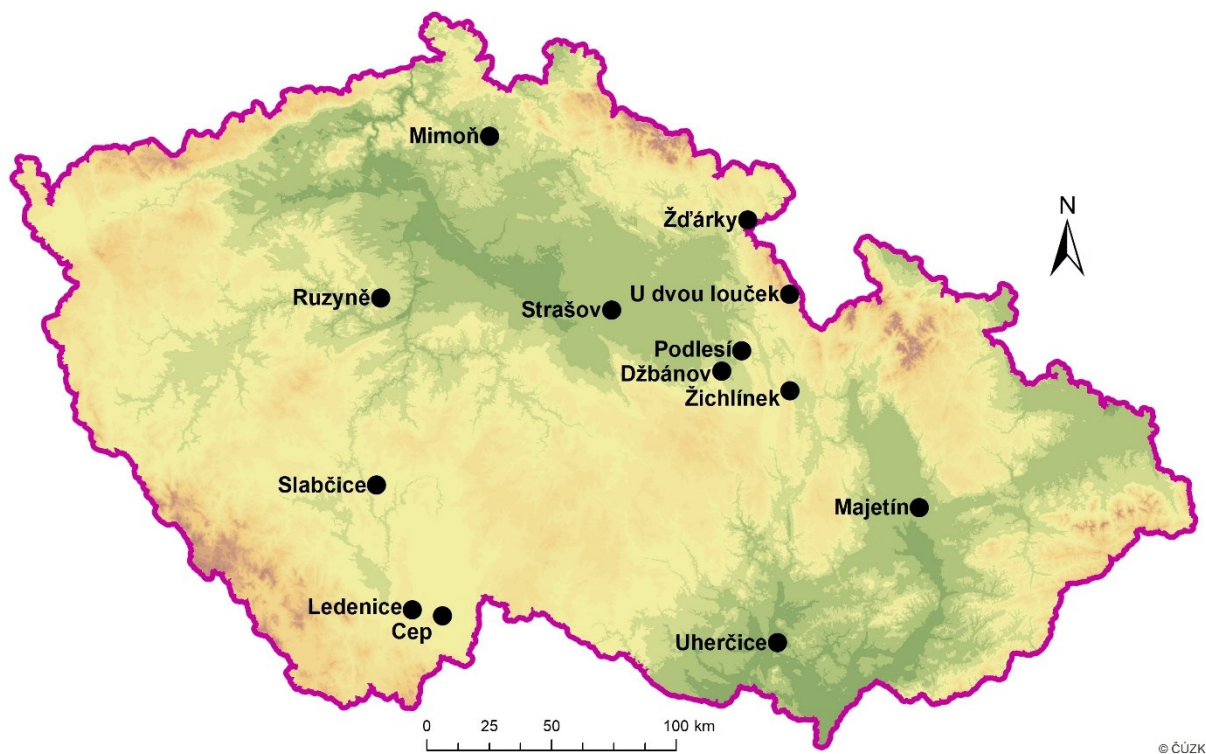
o hydraulické vodivosti. V diplomové práci bylo použito celkem 126 datových souborů z následujících lokalit: Majetín, Cep – Třeboň, Žichlínek, Žďárky, Slabčice, Ledenice, U dvou louček, Mimoň, Džbánov, Podlesí. Na jednotlivých územích se nacházejí rozdílné půdní typy, je tak zajištěna různorodost zkoumaných dat.

V tabulce 5 je uveden počet datových souborů obsahujících informaci jak o potřebných prediktorech, tak změřených hodnot K_S a na obrázku 17 je zobrazeno rozmístění zájmových lokalit na podkladové mapě České republiky.

U lokality U dvou louček byl vyřazen jeden vzorek; jednalo se o jediný případ organogenní půdy, která se od ostatních vzorků významně lišila a významně by ovlivňovala statistické vyhodnocování.

Tabulka 5 - Počet vzorků zájmových lokalit. *Pro lokalitu Ruzyně jsou dostupné 2 sady (24) naměřených hodnot K_S (tlakový infiltrometr, Hood infiltrometr).

Lokalita	Počet vzorků
Cep - Třeboň	15
Džbánov	4
Ledenice - Mysletínský potok	8
Majetín JZD	15
Mimoň	9
Podlesí	4
Ruzyně	12*
Slabčice JZD	7
Strašov	21
U dvou louček	4
Uherčice	12
Žďárky	12
Žichlínek	4
Celkem	127



Obrázek 17 - Zájmové lokality. Podkladová mapa: ČÚZK. Autor: M. Miháliková.

4.2 Dostupné informace o zájmových oblastech

V tabulce 6 jsou zaznamenány jednotlivé zájmové lokality. U každé lokality je vždy uvedeno: počet vzorků (v závorce), zrnitostní třídy, rozmezí hodnot objemové hmotnosti suché půdy (g/cm^3), organického uhlíku (%), změřených hodnot K_s a také zdroj dat.

Níže jsou podrobněji popsány lokality Ruzyně, Strašov a Uherčice, jejichž průzkum probíhal na Katedře vodních zdrojů FAPPZ ČZU. Všechny tři lokality se využívají k pěstování zemědělských plodin.

Lokality pocházející z databáze HYPRESCZ jsou popsány v práci „HYPRESCZ - Database of Soil Hydrophysical Properties in the Czech Republic“ a „Pedotransferové funkce pro odhad hydrofyzikálních vlastností půd“ (Miháliková 2011; Miháliková et al. 2013).

4.2.1 Lokalita Ruzyně

Experimentální lokalita Ruzyně se nachází na $50^\circ 08'$ severní šířky, $14^\circ 30'$ východní délky v nadmořské výšce 345 m n.m. Dlouhodobý roční úhrn srážek je 472 mm, patří tak v České republice k místům s menším úhrnem srážek. Průměrná roční teplota je pak $7,9^\circ\text{C}$ (Český hydrometeorologický ústav 2021).

Na lokalitě se nachází hnědozem modální (Taxonomický klasifikační systém půd ČR, Němeček et al. (2011)).

Na této lokalitě byla K_S měřena v terénu pomocí tlakového infiltrometru (Matula & Kozáková 1997) a Hood infiltrometrem IL-2700 (UGT, GmbH).

4.2.2 Lokalita Strašov

Lokalita Strašov se nachází v Pardubickém kraji v blízkosti města Pardubice na $50^{\circ} 05'$ severní šířky a $15^{\circ} 31'$ východní délky v nadmořské výšce 240 m. n. m. Průměrné roční úhrny srážek se pohybují v rozmezí od 550 do 650 mm, v porovnání s lokalitou Ruzyně, je zde roční úhrn srážek výrazně vyšší. Průměrná roční teplota je 8-9 °C (Český hydrometeorologický ústav 2021).

Půdy nacházející se na této lokalitě byly klasifikovány jako černice a v menší míře zde byly zastoupeny regozemě (Taxonomický klasifikační systém půd ČR, Němeček et al. (2011)).

Měření K_S probíhalo v laboratorních podmínkách na neporušených půdních vzorcích o objemu 250 cm³ odebraných z hloubky 5-15 cm na zařízení KSAT od společnosti METER Group, Inc.

4.2.3 Lokalita Uherčice

Lokalita Uherčice se nachází v okrese Břeclav v Jihomoravském kraji. Lokalita se nachází v nadmořské výšce 180-200 m. n.m. ve velmi teplém a suchém klimatickém regionu. Průměrné roční úhrny srážek se pohybují v rozmezí 500-600 mm, podobně jako u lokality Strašov. Průměrná roční teplota je 9-10 °C.

Převažujícím půdním typem jsou zde fluvizemě (Taxonomický klasifikační systém půd ČR, Němeček et al. (2011)).

V roce 2021 zde byly prováděny průzkumné práce a terénní měření Katedrou vodních zdrojů FAPPZ ČZU zaměřené na stanovení hydrofyzikálních vlastností půd, při kterém byla nasycená hydraulická vodivost měřena v terénu jednosondovou metodou, kdy k měření bylo využito regulované hladiny podzemní vody v rámci existujícího drenážního systému.

Dále byly na neporušených půdních vzorcích stanoveny hodnoty K_S pomocí přístroje KSAT (METER Group, Inc.).

Tabulka 6 - Základní informace o jednotlivých lokalitách.

LOKALITA	Zrnitostní třída (dle USDA)	Seskupená zrnitostní třída	Objemová hmotnost suché půdy (g/cm ³)	C _{ox} (%)	Změřená Ks (cm/day)	Zdroj dat
Podlesí (4)	Silty clay loam (1) Silt loam (3)	4 (1) 3 (3)	1.48-1.59	0.2-5.0	7.8 - 48.2	HYPRES CZ
U dvou louček (3)	Loam (1) Sandy loam (2)	3 (1) 2 (2)	1.30-1.59	1.05-5.1	40.5 - 370.8	HYPRES CZ
Slabčice (8)	Loam (1) Silt loam (2) Sandy loam (5)	3 (3) 2 (5)	1.01-1.89	0.46-2.79	0.3 - 38.8	HYPRES CZ
Džbánov (4)	Silt loam (4)	3 (4)	1.45-1.62	0.1-5.0	5.6 - 22.4	HYPRES CZ
Žichlínek (4)	Clay loam (1) Silty clay loam (1) Silt loam (2)	4 (2) 3 (2)	1.38-1.74	0.2-5.0	2.5 - 75.4	HYPRES CZ
Ledenice (8)	Loam (5) Silt loam (2) Sandy loam (1)	3 (7) 2 (1)	1.15-1.79	0.3-3.36	0.1 - 21.9	HYPRES CZ
Cep (15)	Clay loam (7) Sandy clay loam (3) Loam (3) Sandy loam (2)	4 (10) 3 (3) 2 (2)	1.39-1.75	0.1-2.94	0.1 - 158.2	HYPRES CZ
Mimoň (9)	Silt loam (2) Loamy sand (4) Sand (3)	3 (2) 1 (7)	1.04-1.41	0.47-1.80	121.0 - 699.8	HYPRES CZ
Žďárky (12)	Clay loam (1) Silty clay loam (7) Silt loam (4)	4 (8) 3 (4)	1.43-1.70	0.14-2.03	30.0 - 550.0	HYPRES CZ
Majetín (15)	Clay (2) Clay loam (6) Loam (4) Sandy loam (1) Silt loam (1) Silty clay (1)	5 (3) 4 (6) 3 (5) 2 (1)	1.18-1.70	0.3-4.5	0.1 - 37.7	HYPRES CZ
Strašov (21)	Clay (1) Clay loam (1) Sandy clay loam (12) Sandy loam (3) Loamy sand (2) Sand (2)	5 (1) 4 (13) 2 (3) 1 (4)	1.22-1.74	0.48-3.31	1.0 - 964	Vlastní měření (ČZU)
Ruzyně (12)	Silty clay loam (6) Silt loam (6)	4 (6) 3 (6)	1.13-1.36	1.05-1.87	10.2 - 1261.2	Vlastní měření (ČZU)
Uherčice (12)	Silty clay (4) Silty clay loam (8)	5 (4) 4 (8)	1.13-1.45	1.29-1.85	0.3 - 524.5	Vlastní měření (ČZU)

4.3 Postup práce

4.3.1 Prediktory

Zrnitostní třídy

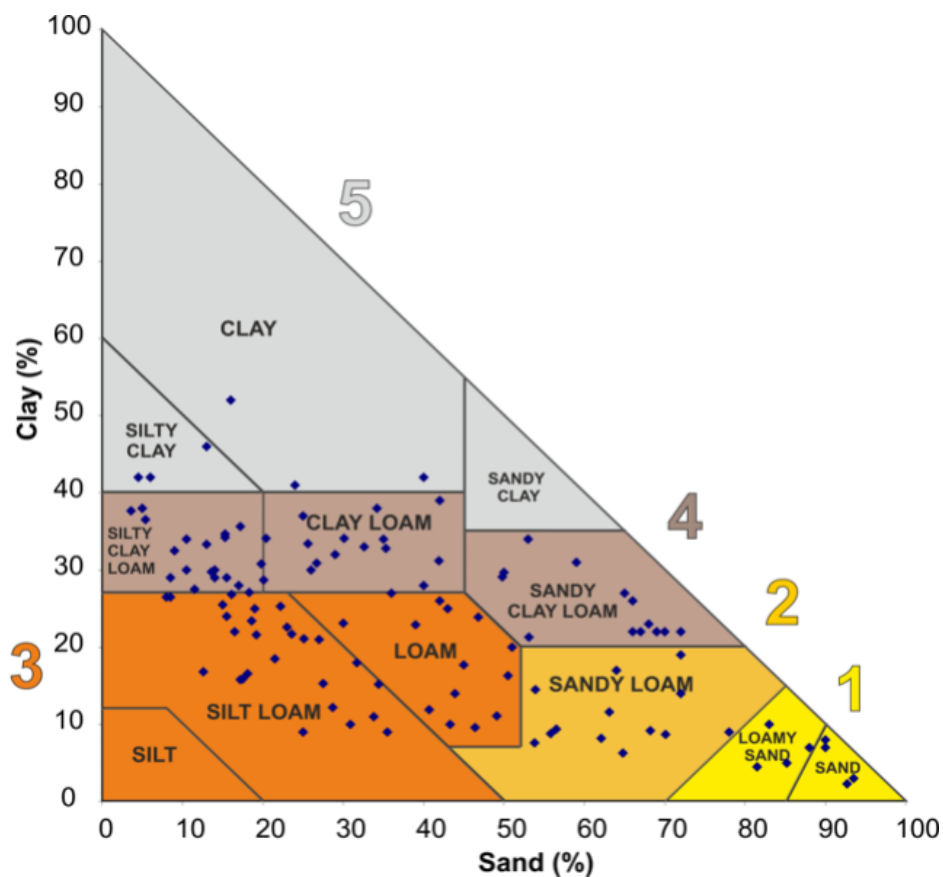
Vzhledem k tomu, že oba použité programy (Araya & Ghezzehei 2019, Rosetta) pracují s 12 texturními třídami podle USDA, pracovalo se s nimi i v rámci diplomové práce (Tabulka 6). Zastoupení jednotlivých texturních tříd testovaných dat je zobrazeno na obrázku 18. Největší zastoupení dat je ve třídách Silty clay loam, Clay loam, Silt loam, Loam, Sandy clay loam a Sandy loam. Naproti tomu pro program PTFapp byla použita tréninková data s největším zastoupením ve třídách Sand, Loamy sand, Sandy loam a Sandy clay loam (Obrázek 20).

V České republice se pro zobecnění výsledků používá seskupení tříd USDA do 5 skupin, viz. tabulka 7 (Němeček et al. 2001; Kozák et al. 2011).

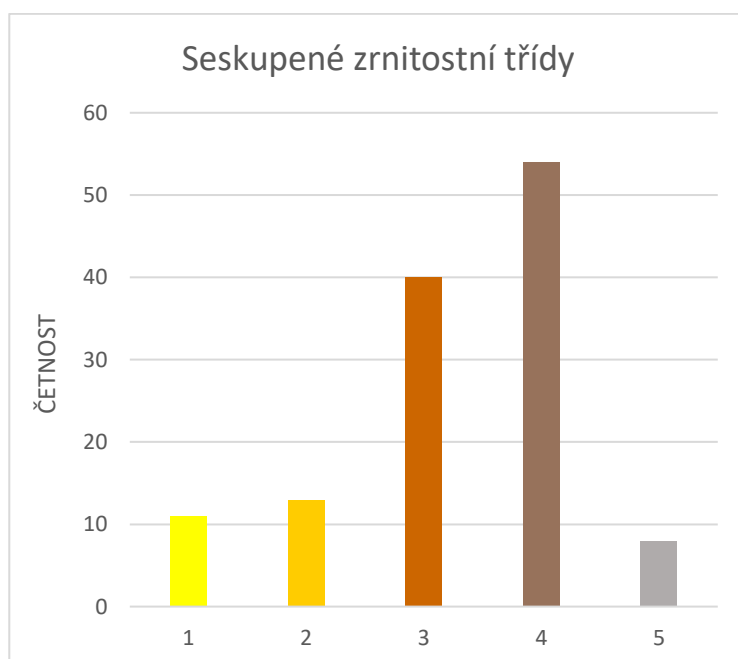
Anglické názvy jednotlivých tříd se běžně používají i v České republice, proto názvy zůstaly zachovány (Miháliková 2011).

Tabulka 7 - Seskupení zrnitostních tříd podle Němečka et al. (2001).

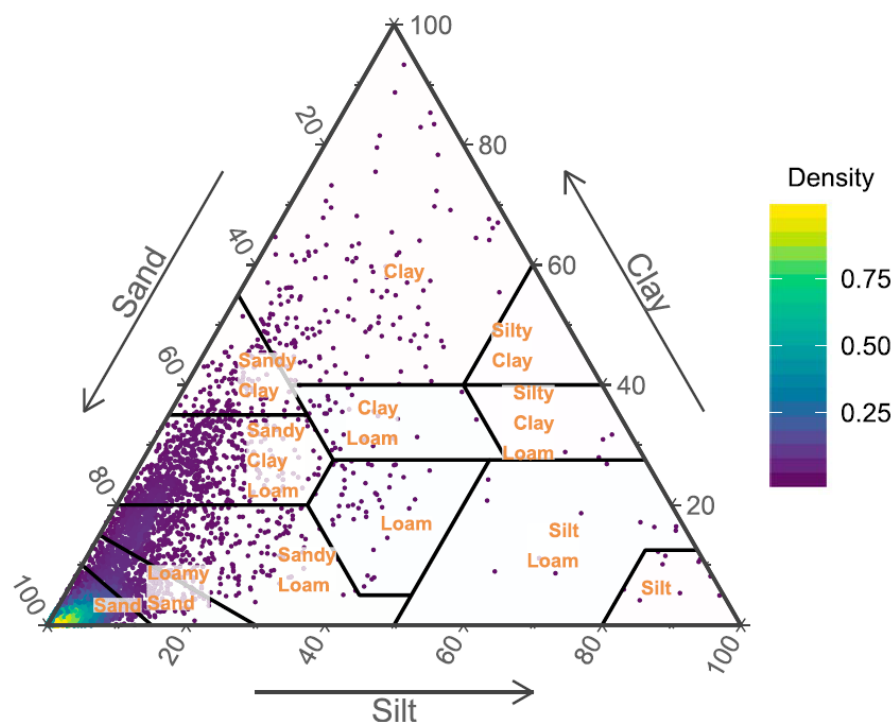
Třída	Název	Seskupené třídy
1	Lehká zemina	Sand, Loamy sand
2	Lehčí střední zemina	Sandy loam
3	Střední zemina	Loam, Silt loam, Silt
4	Těžká zemina	Silty clay loam, Clay loam, Sandy clay loam
5	Velmi těžká zemina	Silty clay, Clay, Sandy clay



Obrázek 18 – Zrnitostní složení půd na testovaných lokalitách a jejich zařazení podle texturního trojúhelníku USDA se zařazením do 5 seskupených zrnitostních tříd (1-5) podle Němečka et al. (2001).



Obrázek 19 - Počet vzorků v seskupených zrnitostních třídách. 1 – Lehká zemina, 2 – Lehčí střední zemina, 3 – Střední zemina, 4 – Těžká zemina, 5 – Velmi těžká zemina.



Obrázek 20 - Texturní trojúhelník – rozmístění vzorků použitých jako tréninková data (75 %) pro aplikaci PTFapp (Araya & Ghezzehei 2019).

Zrnitostní složení půd na testovaných lokalitách a jejich zařazení podle texturního trojúhelníku USDA se zařazením do 5 seskupených zrnitostních tříd (1-5) podle Němečka et al. (2001) je uvedeno na obrázku 18 a 19, na kterých lze pozorovat vyšší zastoupení vzorků pocházejících ze středně těžkých a těžkých půd (seskupené zrnitostní třídy 3 a 4). Oproti tomu obrázek 20 indikuje vyšší zastoupení lehkých půd ve zdrojových datech aplikace PTFapp (Araya & Ghezzehei 2019).

Objemová hmotnost suché půdy (BD)

Informace dostupná, pokud byly odebrány neporušené půdní vzorky. Vyšší hodnoty objemové hmotnosti suché půdy mohou indikovat ztuhnutí půdy a následně i nižší hodnoty nasycené hydraulické vodivosti. Hodnoty BD pro půdy na testovaných lokalitách se pohybovaly v rozmezí od 1,01 do 1,89 g/cm³ (Tabulka 6).

Obsah organického uhlíku

Testované modely PTF používají obsah organického uhlíku C_{ox} . Pro lokality Ruzyně, Strašov a Uherčice byly k dispozici informace o C_{ox} , zatímco v databázi HYPRESCZ byly uvedeny údaje o obsahu organické hmoty (%). Z těchto hodnot bylo potřeba vyvodit obsah organického uhlíku.

Obsah organického uhlíku v půdě byl přepočten z procentuálního zastoupení organické hmoty pomocí Welteho přepočtového koeficientu (Pribyl 2010) rovnicí 16, který vychází z 58% obsahu uhlíku v humusu:

$$\text{Obsah organického uhlíku [\%]} \cdot 1,724 = \text{obsah organické hmoty [\%]} \quad (16)$$

4.3.2 Testované modely PTF

Modely BRT

V tabulce 8 jsou popsány potřebné prediktory k odhadu K_S pomocí algoritmu BRT. V literární rešerši je popsáno celkem 6 hierarchických úrovní modelu BRT (Tabulka 2). Z důvodu nedostatku některých údajů pro úrovně 7-0, 7-1, 7-2 je v diplomové práci využito pouze modelů 3-0, 3-1, 3-2.

Tabulka 8 - Data potřebná k odhadu K_S algoritmem BRT.

Prediktory	
BRT 3-0	% písek, % prach, % jílu
BRT 3-1	% písek, % prach, % jílu, objemová hmotnost suché půdy (g/cm^3)
BRT 3-2	% písek, % prach, % jílu, objemová hmotnost suché půdy (g/cm^3), % organický uhlík

Modely RF

V tabulce 9 jsou popsány potřebné prediktory k odhadu K_S pomocí algoritmu RF. V literární rešerši je popsáno celkem 6 hierarchických úrovní modelu RF (Tabulka 3). Z důvodu nedostatku některých údajů pro úrovně 7-0, 7-1, 7-2 je v diplomové práci, stejně jako v případě předchozího modelu, využito pouze modelů 3-0, 3-1, 3-2.

Tabulka 9 - Data potřebná k odhadu K_S algoritmem RF.

Prediktory	
RF 3-0	% písek, % prach, % jílu
RF 3-1	% písek, % prach, % jílu, objemová hmotnost suché půdy (g/cm^3)
RF 3-2	% písek, % prach, % jílu, objemová hmotnost suché půdy (g/cm^3), % organický uhlík

Rosetta

V tabulce 10 jsou popsány potřebné prediktory k odhadu K_S . Na rozdíl od výše zmíněných modelů program Rosetta nepočítá s hodnotami organického uhlíku. Lze použít známé hodnoty hydrolimitů, ale v této diplomové práci (z důvodu srovnatelnosti s modely BRT, RF) s nimi nebylo počítáno.

Tabulka 10 - Data potřebná k odhadu K_S programem Rosetta.

Prediktory	
Rosetta (SSC)	% písek, % prach, % jílu
Rosetta (SSC+BD)	% písek, % prach, % jílu, objemová hmotnost suché půdy (g/cm^3)

4.3.3 Určení kvality odhadu a statistické vyhodnocení

Výsledky odhadů jednotlivých modelů PTF jsou graficky znázorněny pomocí tzv. boxplotů, a to jak pro jednotlivé lokality, kde je dobře patrný rozsah změřených dat a schopnost jednotlivých modelů PTF tento rozsah na základě daných prediktorů odhadnout, tak i pro jednotlivé seskupené zrnitostní třídy.

Kvalita odhadů ve vztahu ke změřeným hodnotám je graficky znázorněna pomocí korelačních grafů, kde jsou proti sobě vyneseny hodnoty změřených a odhadnutých K_S . V případě rovnosti změřené a odhadnuté K_S by příslušný bod ležel přímo na indikované křivce 1:1.

Statistické vyhodnocení je provedeno na základě střední chyby odhadu (mean error, ME), střední kvadratické chyby odhadu (root mean square error, RMSE) a koeficientu determinace (coefficient of determination, R^2). K_S je charakteristika půdy, která nevykazuje normální rozdělení. Logaritmická transformace K_S dat je považována za vhodnou, někteří autoři upřednostňují přirozený logaritmus, jiní dekadický. Tyto charakteristiky se běžně používají k vyhodnocení přesnosti odhadů PTF (Nemes et al. 2003; Lilly et al. 2008; Tóth et al. 2014; Araya & Ghezzehei 2019). Veškeré statistiky jsou v této práci počítány na základě \log_{10} K_S hodnot. Při použití log-transformovaných K_S dat pak jsou výsledné statistiky bezrozměrné (Meerschaert et al. 2013; Tóth et al. 2014).

Hodnocení a určení pořadí (umístění) jednotlivých modelů PTF je provedeno pro jednotlivé seskupené zrnitostní třídy i pro jednotlivé lokality na základě vypočtených hodnot RMSE, přičemž nejkvalitnější odhad s nejnižší hodnotou RMSE získal umístění č. 1. Ve výsledném součtu pro všech 5 seskupených zrnitostních tříd se pak nejlépe umístil model PTF s nejnižším součtem (s nejlepším umístěním).

ME (mean error) rovnice (17) – udává vychýlení hodnot. Kladné, nebo záporné hodnoty naznačují průměrnou tendenci předpovídaných hodnot mít větší, nebo menší hodnotu než naměřené hodnoty.

$$ME = \frac{1}{n} \sum_{i=1}^n (y_i - x_i) \quad (17)$$

Kde y_i jsou odhadnuté hodnoty K_S a x_i jsou naměřené hodnoty K_S . Počet dvojic $x_i y_i$ označuje písmeno n .

RMSE (root mean square error) rovnice (18) – udává průměrnou odchylku předpovídaných dat od naměřené hodnoty. Čím menší hodnota RMSE, tím lépe daný algoritmus předpovídá hodnoty K_S .

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (18)$$

Kde y_i jsou odhadnuté hodnoty K_S a x_i jsou naměřené hodnoty K_S . Počet dvojic $x_i y_i$ označuje písmeno n .

R² (koeficient determinace) rovnice (19) – označuje shodu mezi odhadnutými a naměřenými daty. Čím vyšší hodnota R^2 , tím silnější korespondence. Může nabývat hodnot od 0-1 (0 % - 100 %). 100 % znamená dokonalou predikci hodnot závisle proměnné. Naopak hodnota 0 % znamená, že model nepřináší pro poznání závisle proměnné žádnou informaci (Anděl 2011).

$$R^2 = \left\{ \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{[n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2][n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2]}} \right\}^2 \quad (19)$$

Kde y_i jsou odhadnuté hodnoty K_S a x_i jsou naměřené hodnoty K_S . Počet dvojic $x_i y_i$ označuje písmeno n (Anděl 2019).

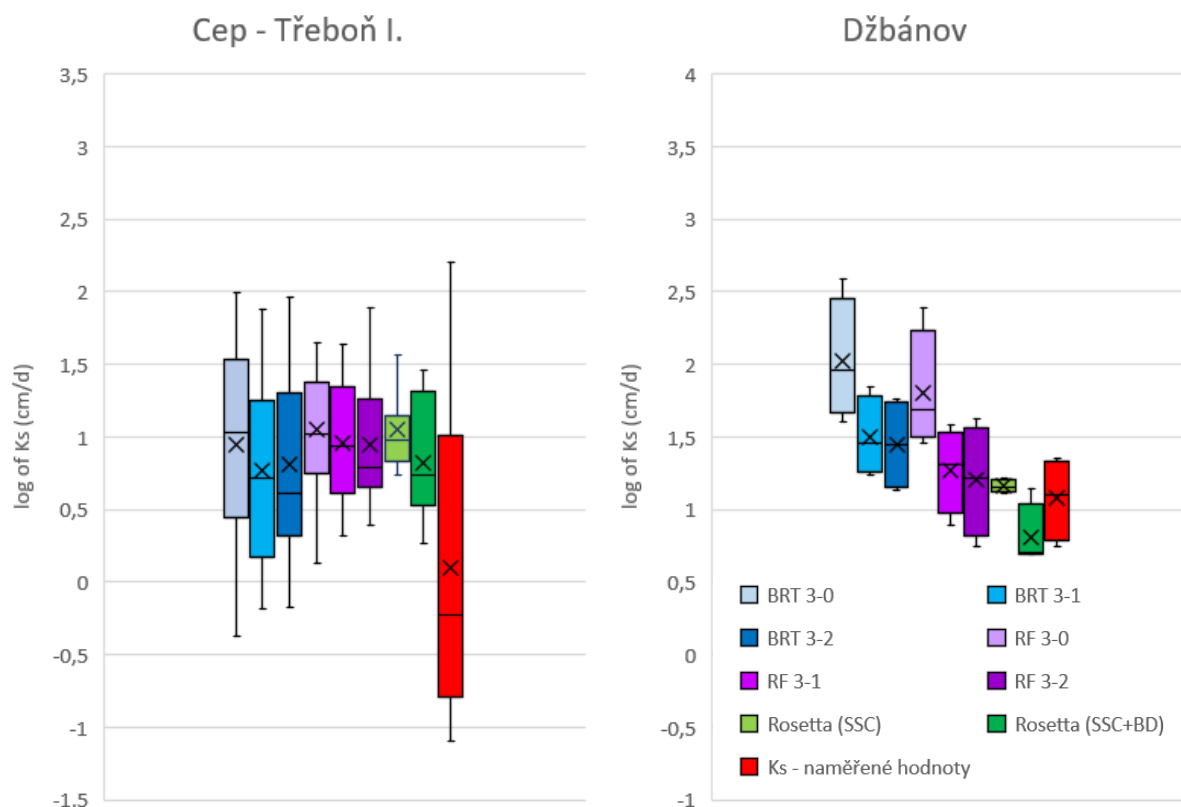
5 Výsledky

5.1 Srovnání kvality odhadu použitých PTF

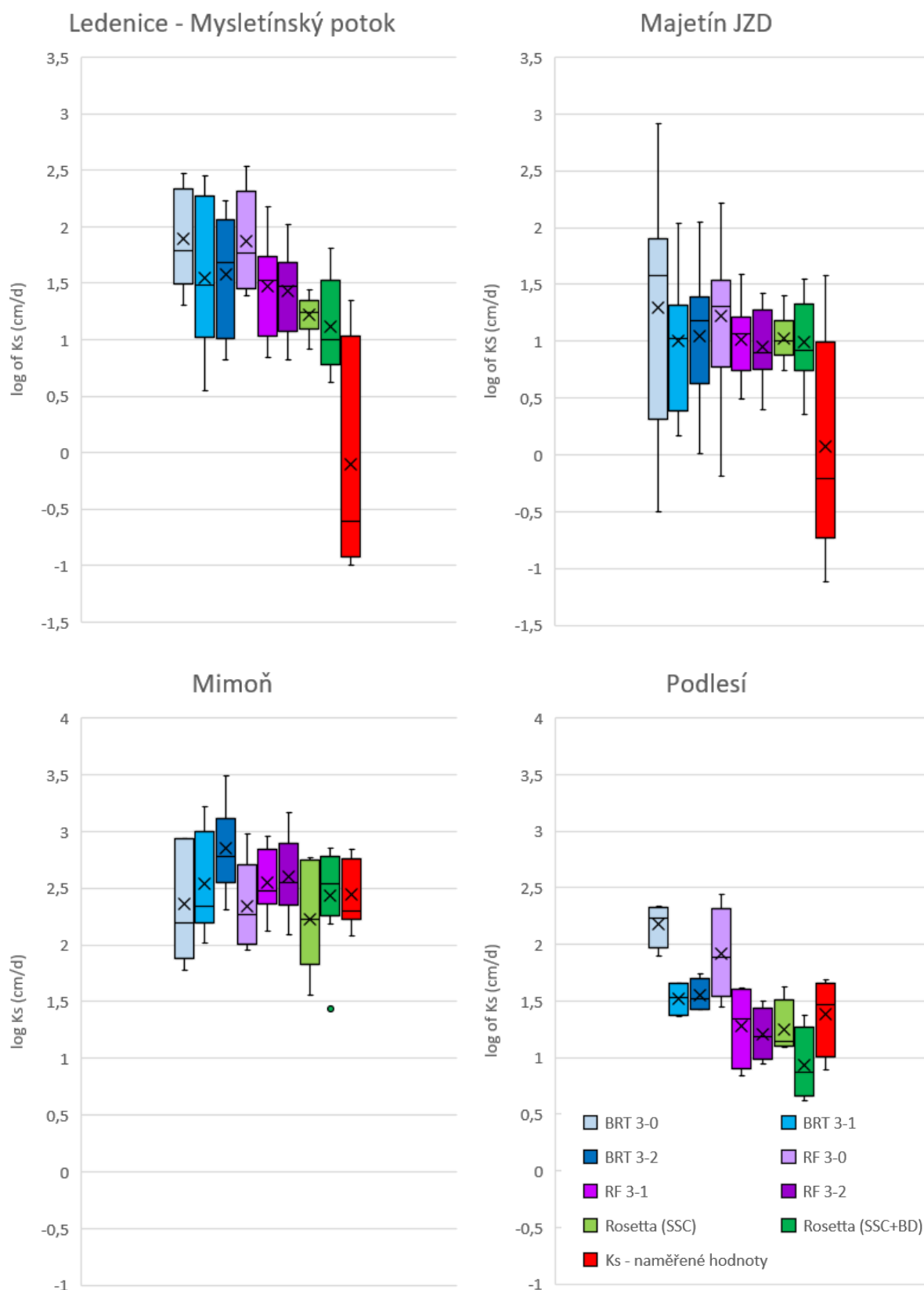
5.1.1 Porovnání pro jednotlivé lokality

Krabicové grafy pro jednotlivé lokality zobrazující poměrně velké rozsahy naměřených hodnot K_S (červená barva) jsou zobrazeny na obrázku 21. Jednotlivé grafy se liší počtem dat.

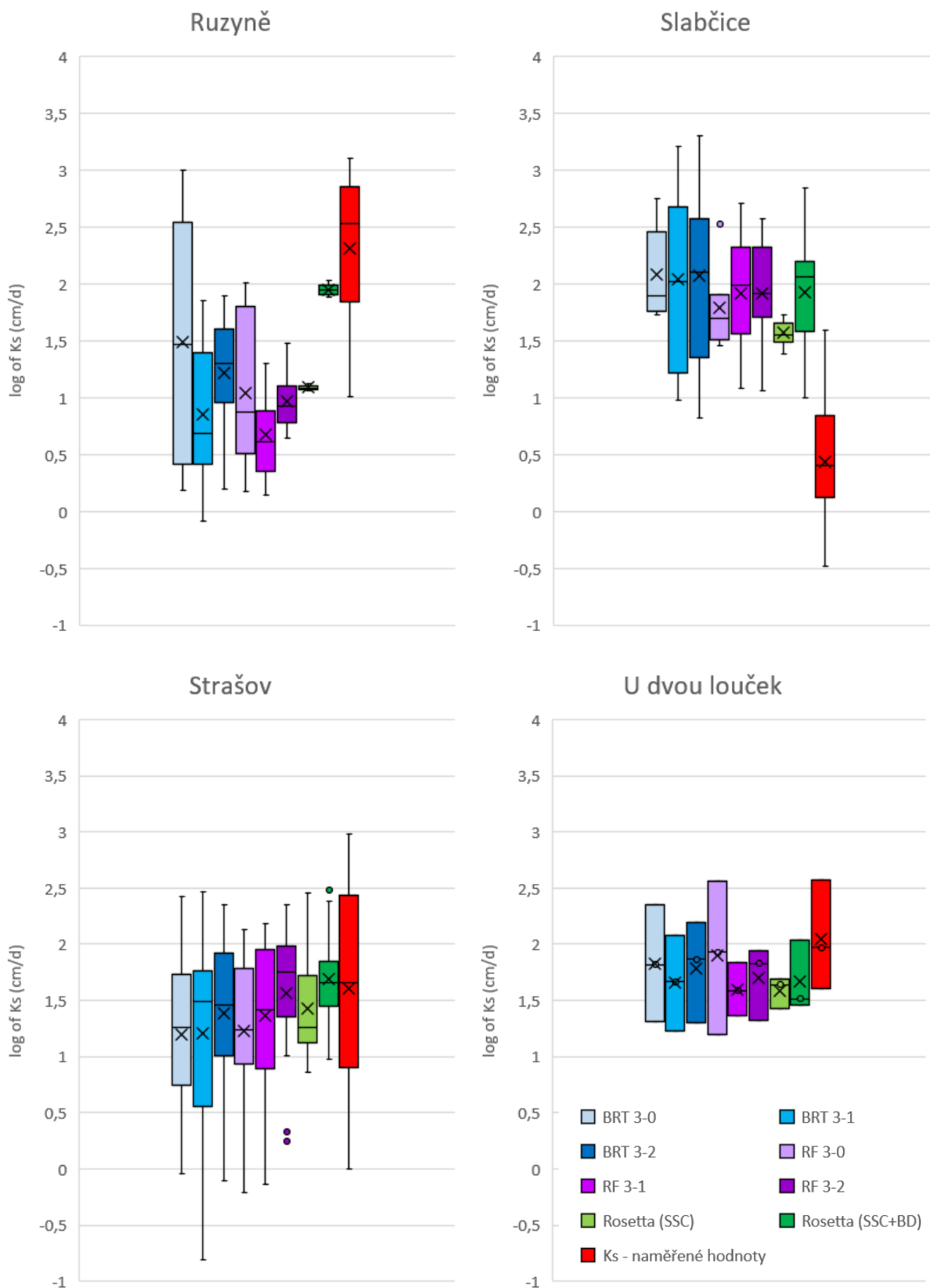
Boxploty ukazují rozdíly mezi datovými soubory. Při podrobném srovnání jednotlivých boxplotů lze zjistit, že Rosetta (SSC) vykazuje nejmenší variabilitu odhadnutých hodnot. „Box“ znázorňuje hodnoty mezi prvním a třetím kvantilem (tj. 25–75 % dat), je přepůlen horizontální úsečkou – mediánem. Křížek označuje průměrnou hodnotu. Odlehlé hodnoty jsou zobrazeny buďto jako tzv. vousy, nebo jednotlivými body (např. u lokality Žďárky).



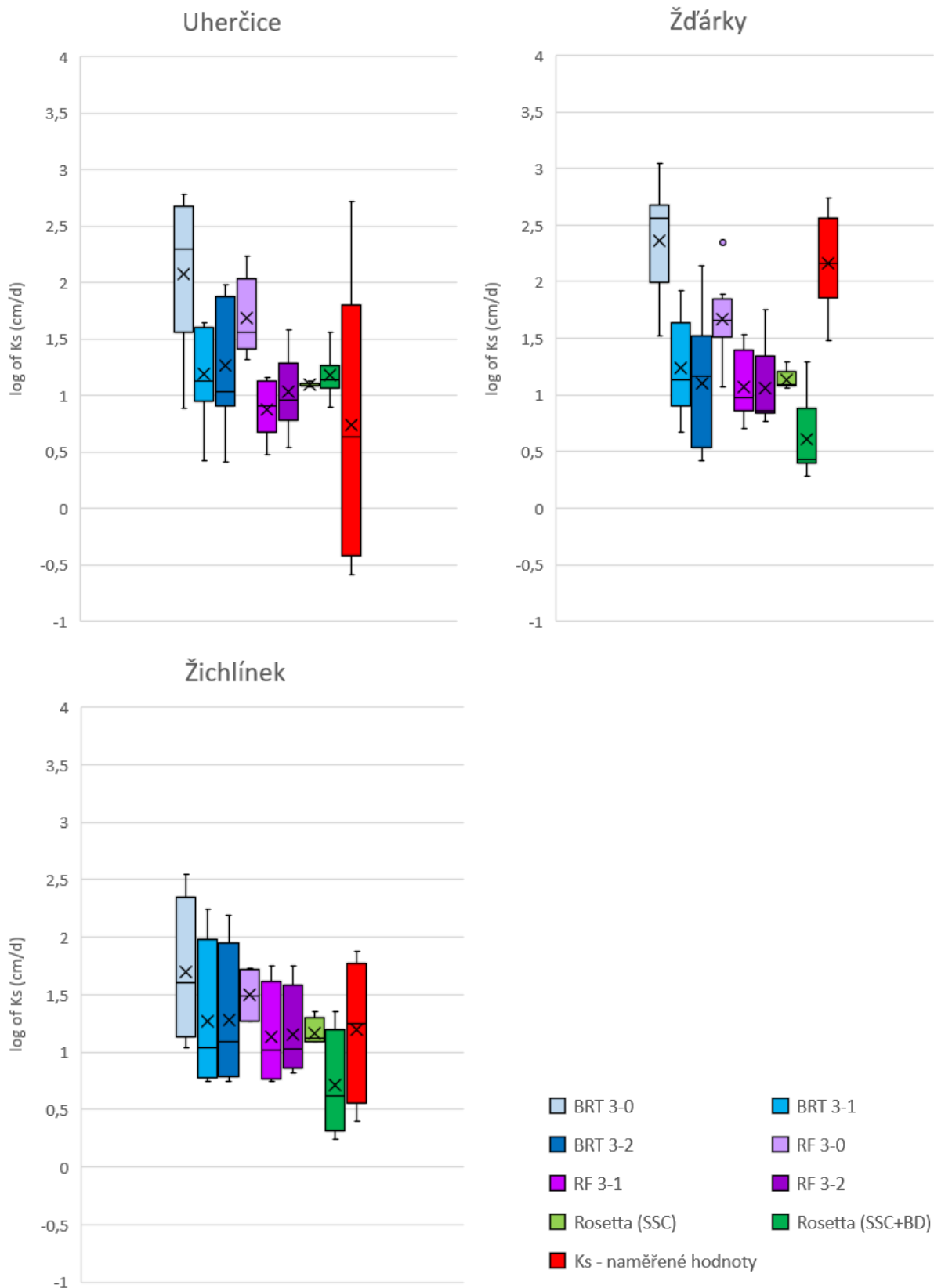
Obrázek 21a – Srovnání jednotlivých lokalit pomocí boxplotů – lokality Cep, Džbánov.



Obrázek 21b – Srovnání jednotlivých lokalit pomocí boxplotů – lokality Ledenice, Majetín, Mimoň, Podlesí.

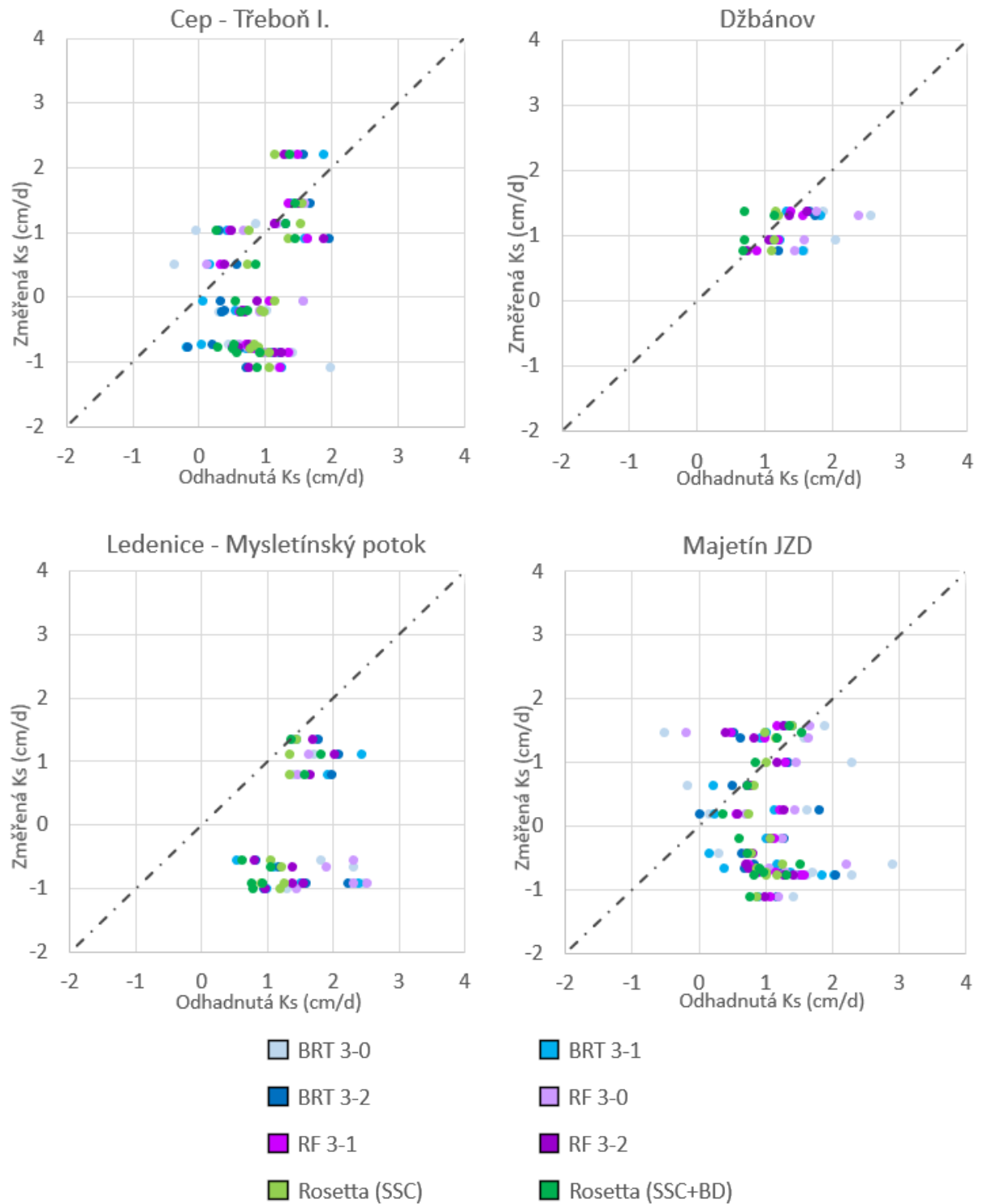


Obrázek 21c – Srovnání jednotlivých lokalit pomocí boxplotů – lokality Ruzyně, Slabčice, Strašov, U dvou louček.

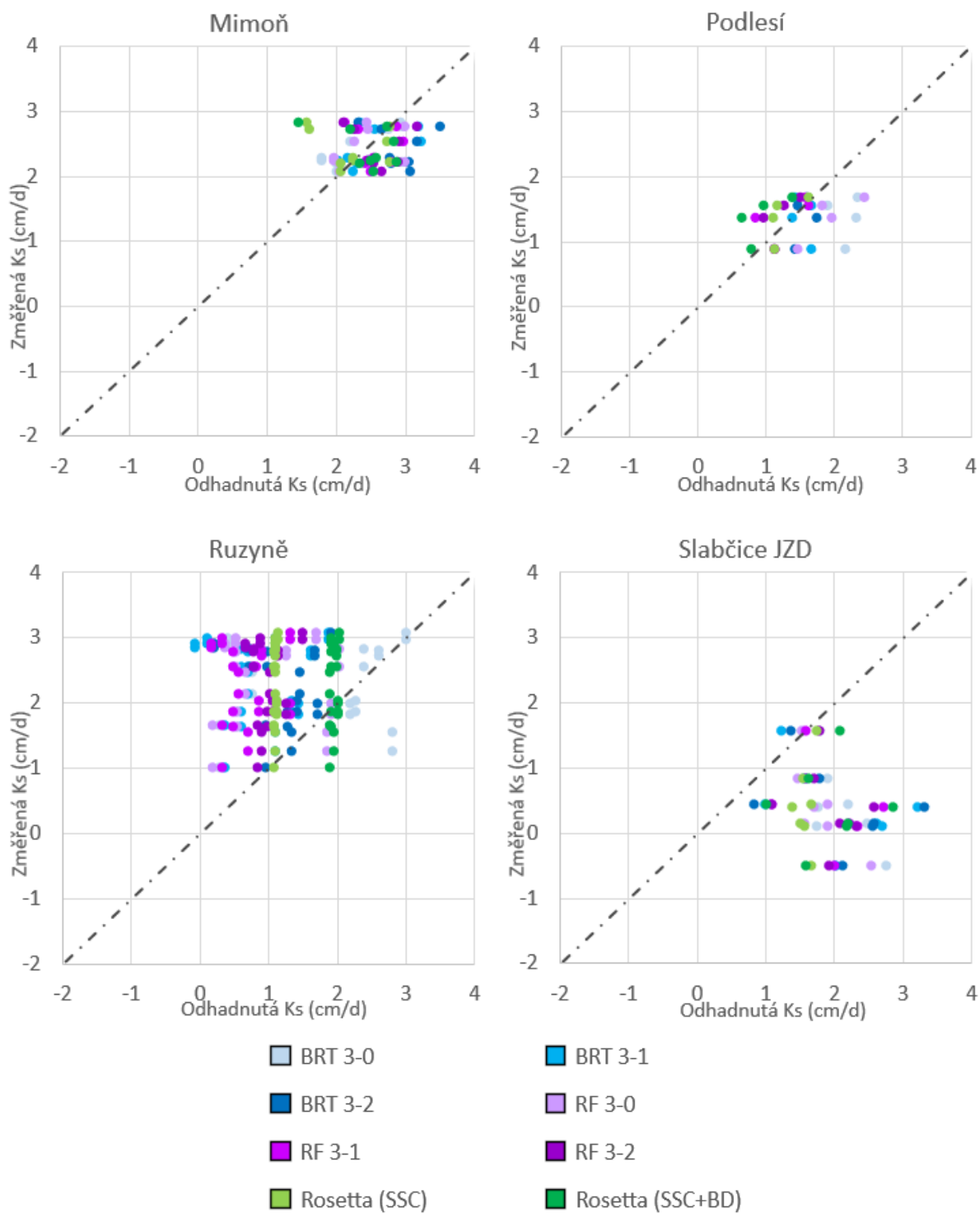


Obrázek 21d – Srovnání jednotlivých lokalit pomocí boxplotů – lokality Uherčice, Žďárky, Žichlínek.

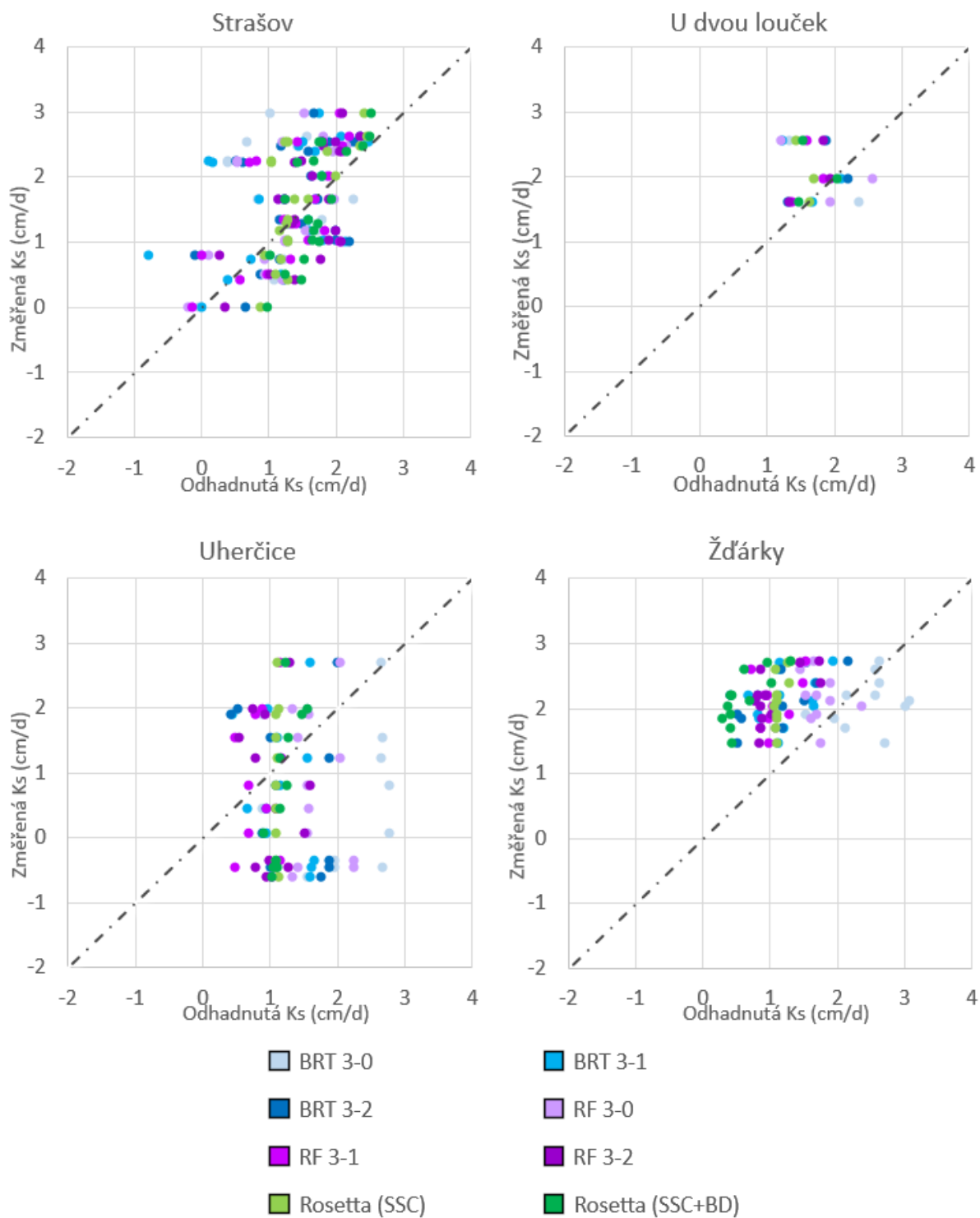
Kvalita odhadů jednotlivých modelů PTF je graficky vyhodnocena pomocí korelačních grafů, kdy je proti sobě vynesena změřená hodnota K_S a příslušná odhadnutá hodnota K_S (Obrázek 22). Při naprosté shodě hodnot by pak zobrazený bod ležel na indikované křivce 1:1.



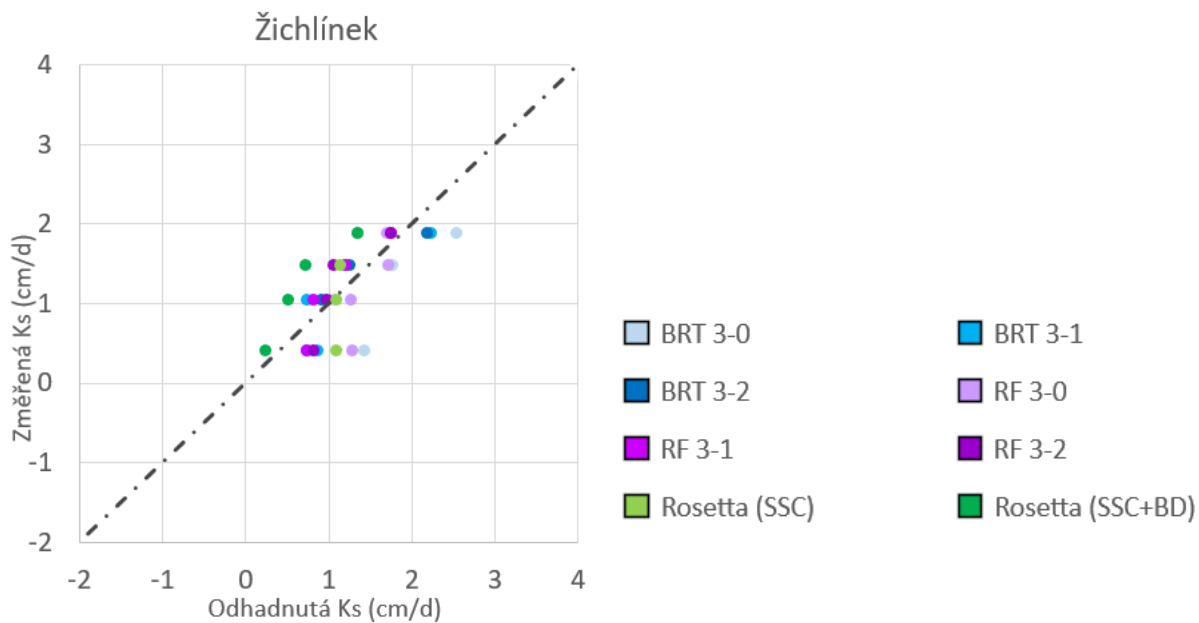
Obrázek 22a – Korelace mezi změřenou a odhadnutou K_S – lokality Cep, Džbánov, Ledenice a Majetín.



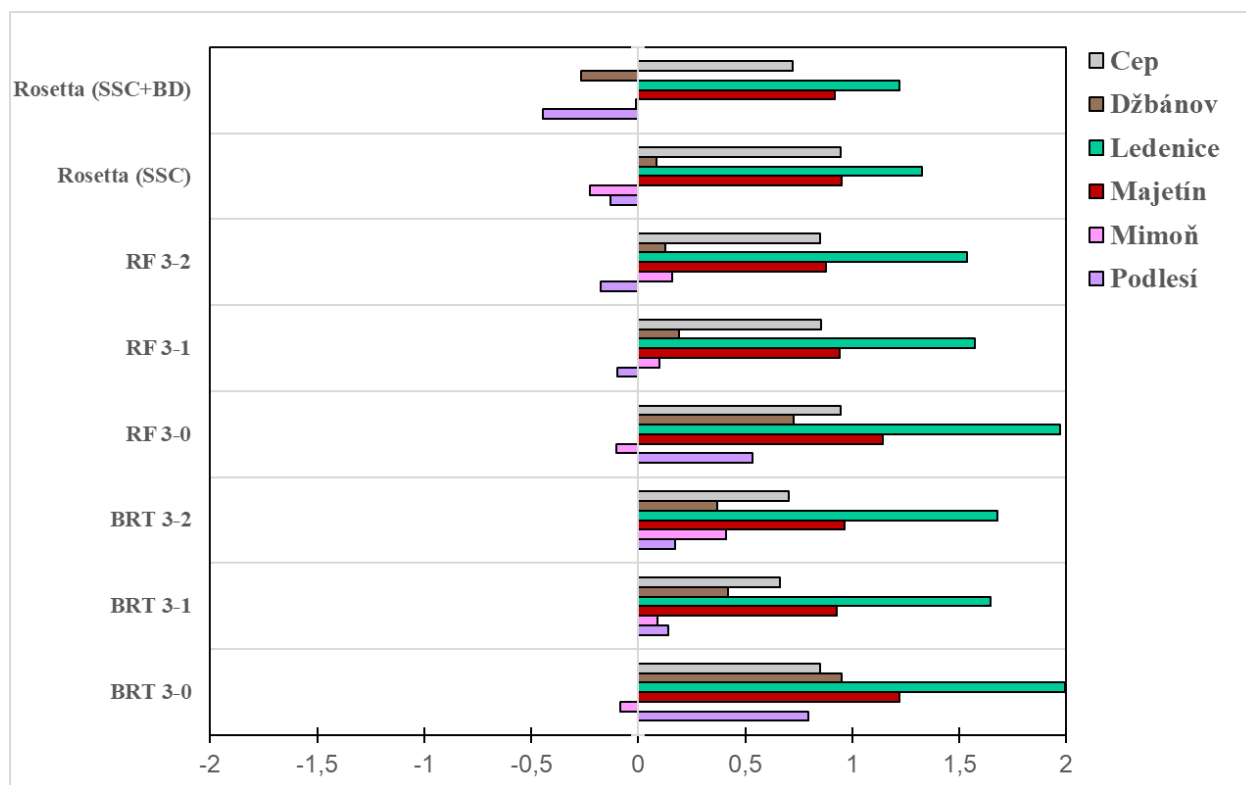
Obrázek 22b– Korelace mezi změřenou a odhadnutou K_S – lokality Mimoň, Podlesí, Ruzyně a Slabčice.



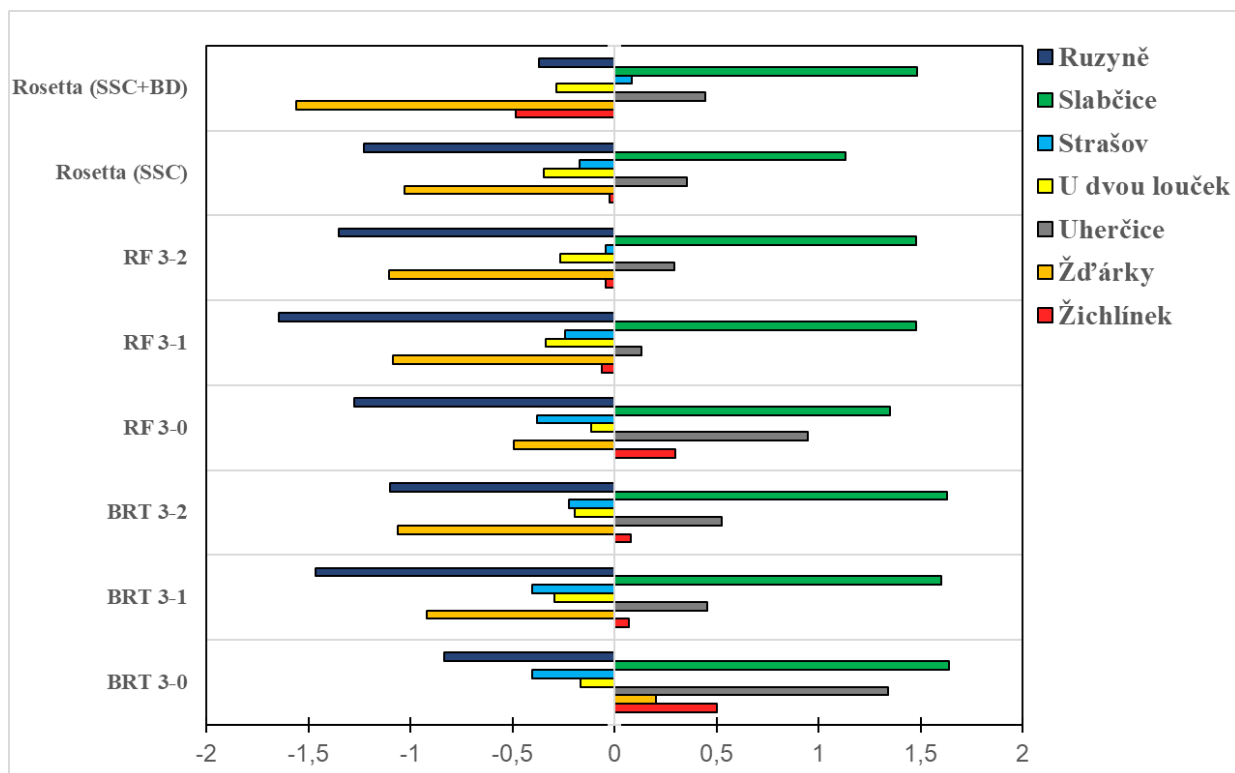
Obrázek 22c– Korelace mezi změřenou a odhadnutou K_S – lokality Strašov, U dvou louček, Uherčice a Žďárky.



Obrázek 22d– Korelace mezi změřenou a odhadnutou K_S – lokalita Žichlínek.



Obrázek 23a – Mean error testovaných PTF pro jednotlivé lokality.



Obrázek 23b – Mean error testovaných PTF pro jednotlivé lokality.

Na obrázku 23 jsou vyneseny hodnoty ME. U lokalit Slabčice, Ledenice, Majetín, Cep lze pozorovat podobné hodnoty ME. Všechny modely odhad nadhodnocují. U lokality Ruzyně lze vidět značné podhodnocení odhadu K_S . Přesný odhad byl u všech algoritmů na lokalitách Žichlínek (s výjimkou BRT 3-0, RF 3-0, Rosetta SSC+BD) a Mimoň (s výjimkou BRT 3-2). U dvou louček jsou mírně podhodnocené všechny odhady K_S . Pro Žďárky nejmenší ME vykazuje model BRT 3-0, ostatní modely hodnotu K_S podhodnocují.

Tabulka 11 - Koeficient determinace [%] pro jednotlivé modely PTF a jednotlivé lokality.

LOKALITA	Použitá PTF							Rosetta (SSC)	Rosetta (SSC+BD)
	Araya; BRT 3-0	Araya; BRT 3-1	Araya; BRT 3-2	Araya; RF 3-0	Araya; RF 3-1	Araya; RF 3-2	Araya; RF 3-2		
Podlesí	0,00	17,63	11,68	70,87	36,71	33,34	37,10	41,35	
U dvou louček	97,61	38,66	25,57	41,69	11,93	45,55	69,87	0,34	
Slabčice	43,83	22,20	16,59	59,08	15,75	5,80	6,97	0,50	
Džbánov	35,70	2,99	85,30	52,94	82,52	95,29	75,71	24,99	
Žichlínek	56,72	66,14	80,66	72,91	84,79	75,50	66,80	88,60	
Ledenice	18,96	14,56	24,39	42,56	27,24	50,54	55,44	76,40	
Cep	0,00	24,64	32,00	6,88	10,37	15,85	26,04	40,48	
Mimoň	46,88	21,52	0,29	22,81	0,96	1,95	4,50	21,31	
Žďárky	0,01	13,07	45,14	0,00	19,60	39,13	45,90	57,43	
Majetín	6,53	1,29	6,16	4,97	6,52	2,33	7,82	8,88	
Strašov	10,07	19,59	13,23	31,14	37,48	37,28	42,69	57,37	
Ruzyně	0,02	0,06	1,46	0,02	0,27	0,68	6,13	4,66	
Uherčice	0,00	6,82	6,59	0,31	0,00	2,15	14,03	55,61	

Nejvyšší koeficient determinace (97,6 %) byl stanoven pro lokalitu U dvou louček, texturní třída USDA Loam a Sandy loam. Džbánov má nejvyšší hodnotu koeficientu determinace rovnu 95,29 %, zrnitostní třída je zde Silt loam. Žichlínek má nejvyšší koeficient determinace 88,6 %, vzorky byly zařazeny do zrnitostních tříd Clay loam, Silty clay loam a Silt loam. Ledenice mají nejvyšší koeficient determinace roven 76,4 %, třídy jsou zde Loam, Silt loam, Sandy loam (Tabulka 11).

Pokud $R^2 > 80\%$, pak je těsnost velmi vysoká, hodnoty mezi 50-80 % udávají těsnost vysokou, hodnoty mezi 25-50 % těsnost význačnou.

Při použití PTF pro odhad K_S je však třeba vzít v úvahu nízký počet naměřených K_S dat, jejichž odhady byly takto přesné. Nízká korelace a s tím související koeficient determinace u lokality Majetín může být způsobena tím, že zde byla naměřena nízká hodnota K_S , dle tabulky 6 v rozmezí od 0,1-37,7 cm/den. Studie (Araya & Ghezzehei 2019) zmiňuje větší přesnost programu, pokud je $K_S > 10$ cm/den. U lokality Ruzyně je nízká korelace a hodnota koeficientu determinace způsobena pravděpodobně vlivem preferenčního proudění, jelikož půda je zemědělsky využívaná a vlivem sucha zde vznikají hluboké praskliny.

Tabulka 12 dává přehled hodnot RMSE jednotlivých algoritmů pro jednotlivé lokality ve srovnání s výsledky autorů Araya a Ghezzehei. Autoři ve své studii konstatují, že nejlepší výkon podával algoritmus BRT s RMSE $< 0,7$ (Araya & Ghezzehei 2019). Tučně jsou zvýrazněny výsledky, kde hodnota RMSE pro log K_S data (K_S v cm/den) $< 0,7$.

Tabulka 12 - Hodnoty RMSE pro vybrané lokality; tučně jsou označené hodnoty menší než 0,7 (srovnatelné s výsledky autorů Araya, Ghezzehei, 2019).

LOKALITA	Použitá PTF							
	Araya; BRT 3-0	Araya; BRT 3-1	Araya; BRT 3-2	Araya; RF 3-0	Araya; RF 3-1	Araya; RF 3-2	Rosetta (SSC)	Rosetta (SSC+BD)
Podlesí	0,869	0,403	0,333	0,568	0,296	0,306	0,276	0,508
U dvou louček	0,733	0,671	0,400	0,762	0,508	0,397	0,586	0,534
Slabčice	1,862	1,985	1,994	1,617	1,746	1,694	1,272	1,676
Džbánov	0,993	0,524	0,385	0,767	0,221	0,162	0,234	0,350
Žichlínek	0,633	0,356	0,262	0,483	0,249	0,299	0,466	0,529
Ledenice	2,316	1,880	1,868	2,340	1,767	1,698	1,560	1,368
Cep	1,461	1,092	1,079	1,352	1,271	1,243	1,297	1,077
Mimoň	0,369	0,387	0,610	0,360	0,406	0,473	0,604	0,586
Žďárky	0,655	1,016	1,133	0,688	1,144	1,148	1,082	1,578
Majetín	1,893	1,437	1,497	1,614	1,387	1,310	1,280	1,253
Strašov	0,978	0,961	0,872	0,803	0,720	0,665	0,662	0,595
Ruzyně	1,484	1,693	1,356	1,545	1,779	1,488	1,367	0,700
Uherčice	1,852	1,320	1,430	1,486	1,109	1,202	1,130	1,053

Tabulka 13 ukazuje srovnání použitých algoritmů k určení K_S u jednotlivých lokalit. Pořadí bylo určeno podle RMSE. Nejlepší pořadí (1) získal model PTF, jehož RMSE byla nejnižší.

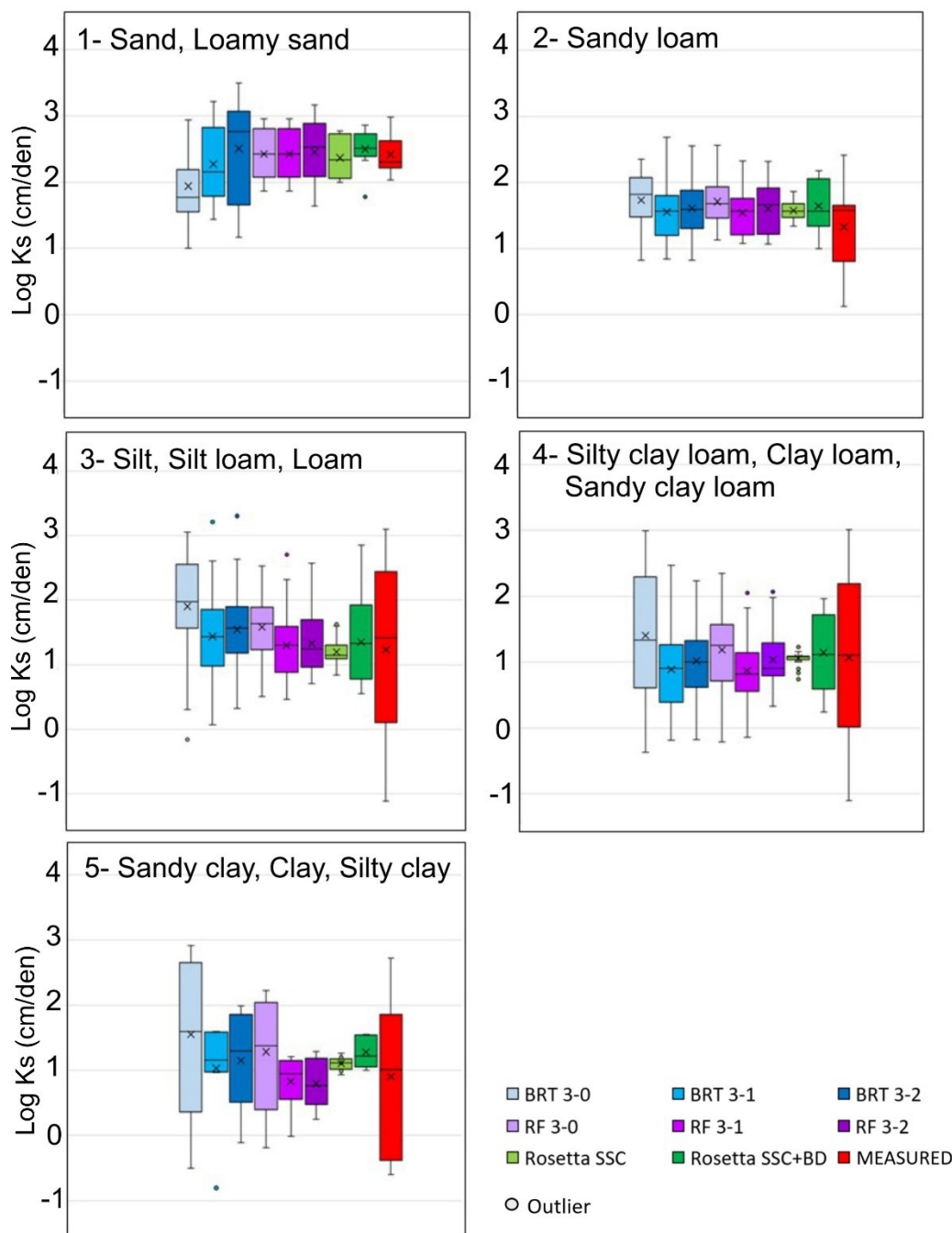
Celkově nejmenší RMSE vykazuje program Rosetta využívající pouze procentuálního zastoupení jednotlivých textur: sand, silt, clay (viz Obrázek 15 – Velikostní kategorie půdních minerálních částic USDA). Na druhém místě je také program Rosetta, kde je navíc přidána znalost objemové hmotnosti suché půdy (BD – bulk density). Algoritmus RF 3-2 je na třetím místě se ztrátou 2-3 bodů oproti algoritmům Rosetta (SSC) a Rosetta (SSC+BD). Nicméně i modely RF 3-0 a RF 3-1 měly nejlepší odhady pro jednu lokalitu (Mimoň a Žichlínek) a model BRT 3-0 pro lokalitu Žďárky.

Tabulka 13 - Srovnání použitých algoritmů k určení K_S u jednotlivých lokalit; pořadí 1-8 podle RMSE (pořadí 1 získal model s nejnižší hodnotou RMSE reprezentující nejkvalitnější odhad).

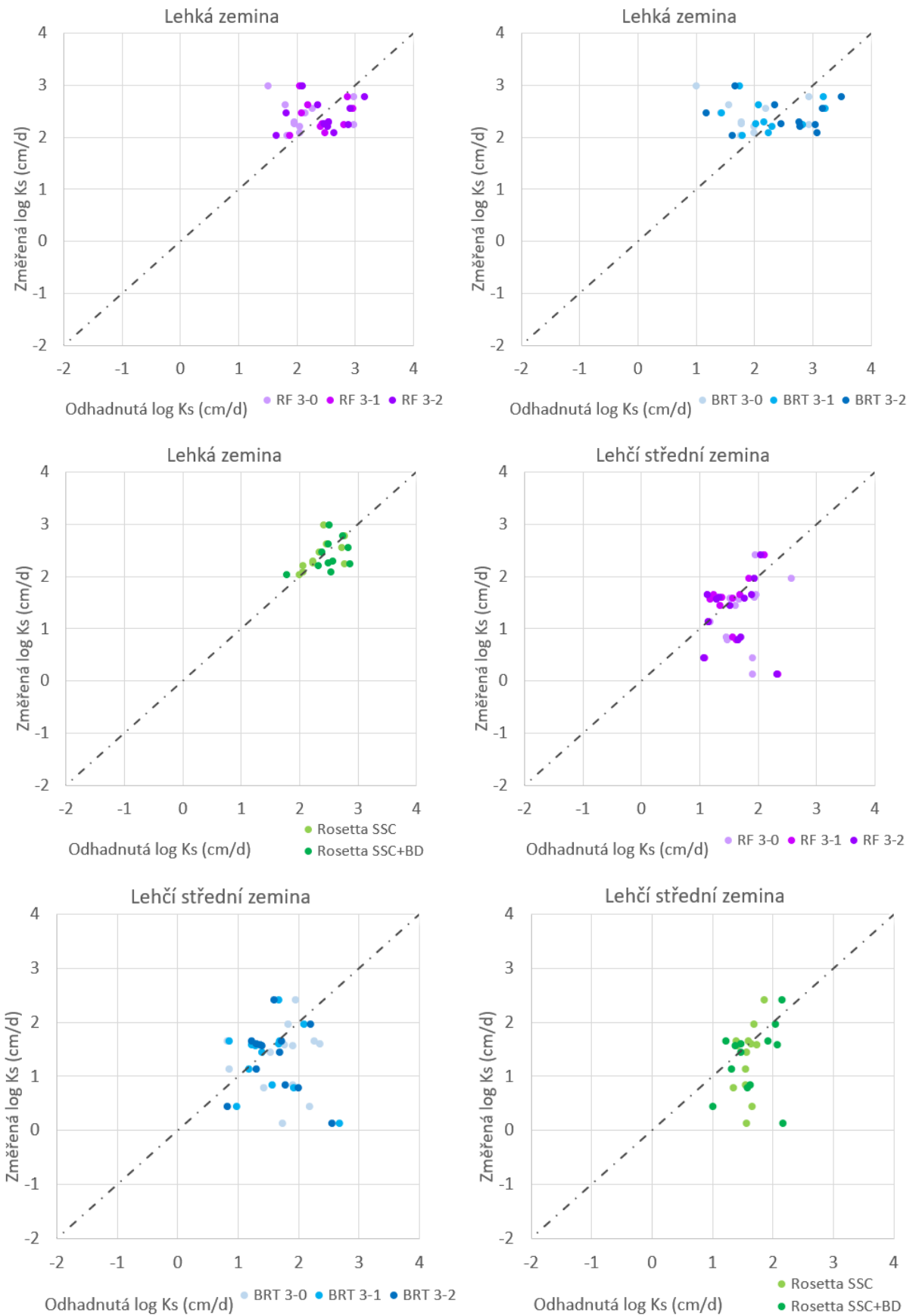
LOKALITA	Použitá PTF							
	Araya; BRT 3-0	Araya; BRT 3-1	Araya; BRT 3-2	Araya; RF 3-0	Araya; RF 3-1	Araya; RF 3-2	Rosetta (SSC)	Rosetta (SSC+BD)
Podlesí	8	5	4	7	2	3	1	6
U dvou louček	6	8	7	5	4	1	2	3
Slabčice	6	7	8	2	5	4	1	3
Džbánov	8	6	5	7	2	1	3	4
Žichlínek	8	4	2	6	1	3	5	7
Ledenice	7	6	5	8	4	3	2	1
Cep	8	3	2	7	5	4	6	1
Mimoň	2	3	8	1	4	5	7	6
Žďárky	1	3	5	2	6	7	4	8
Majetín	8	5	6	7	4	3	2	1
Strašov	8	7	6	5	4	3	2	1
Ruzyně	4	7	2	6	8	5	3	1
Uherčice	8	5	6	7	2	3	4	1
Suma	82	69	66	70	51	45	42	43
Celkové pořadí	8	6	5	7	4	3	1	2

5.1.2 Porovnání pro seskupené zrnitostní třídy

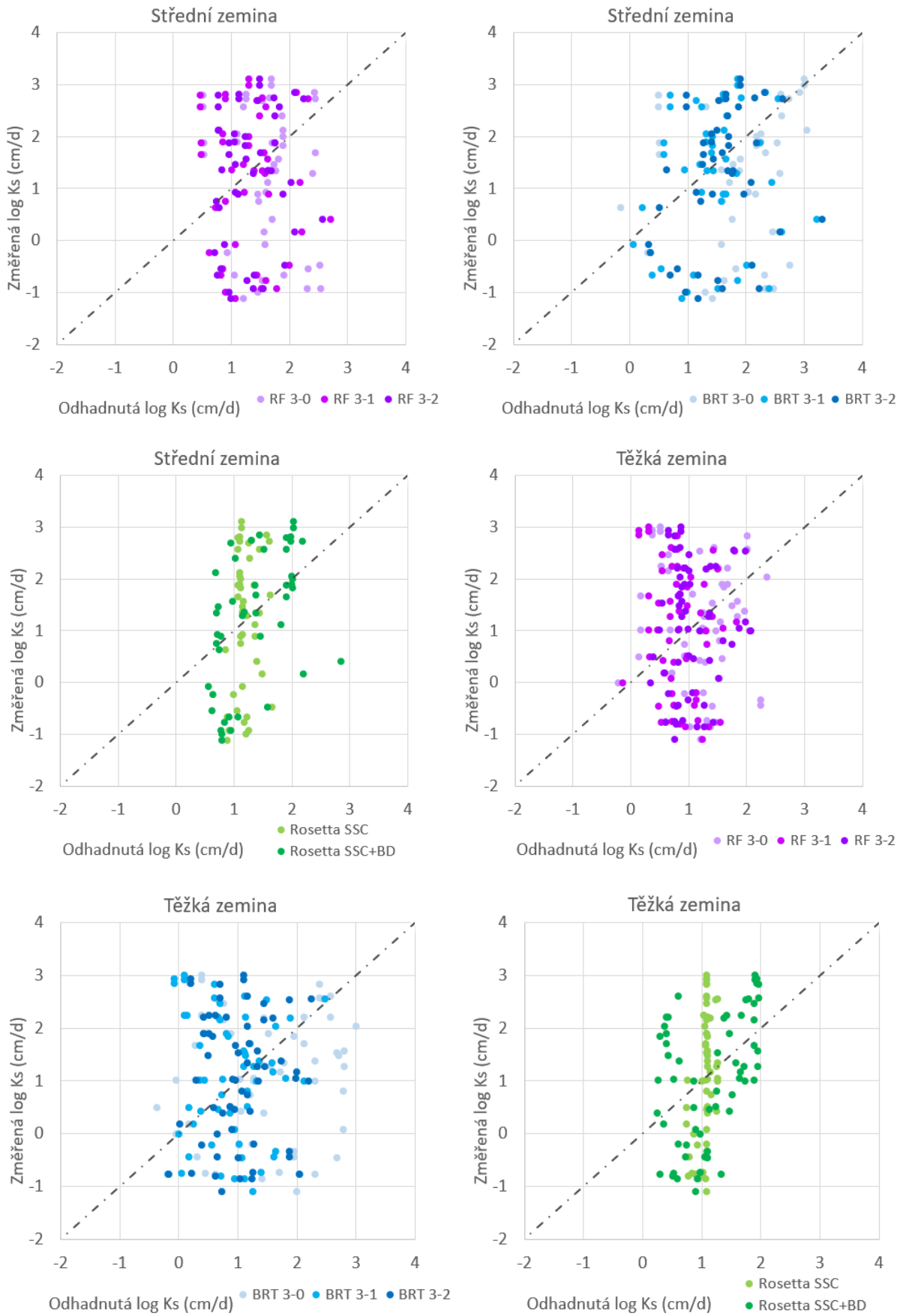
Porovnání odhadovaných a změřených K_S hodnot pro 5 seskupených zrnitostních tříd (1-5, Tabulka 7 - Seskupení zrnitostních tříd) je graficky uvedeno pomocí boxplotů na obrázku 24. Z boxplotů je patrný relativně nižší rozsah měřených hodnot K_S pro lehčí půdy (seskupené třídy 1 a 2). Korelační grafy pro každý model testovaných PTF pro jednotlivé seskupené zrnitostní třídy (1-5) jsou uvedeny na obrázku 25. Kategorie střední zemina, těžká zemina a velmi těžká zemina nevykazují příliš velkou shodu mezi naměřenými a odhadnutými hodnotami.



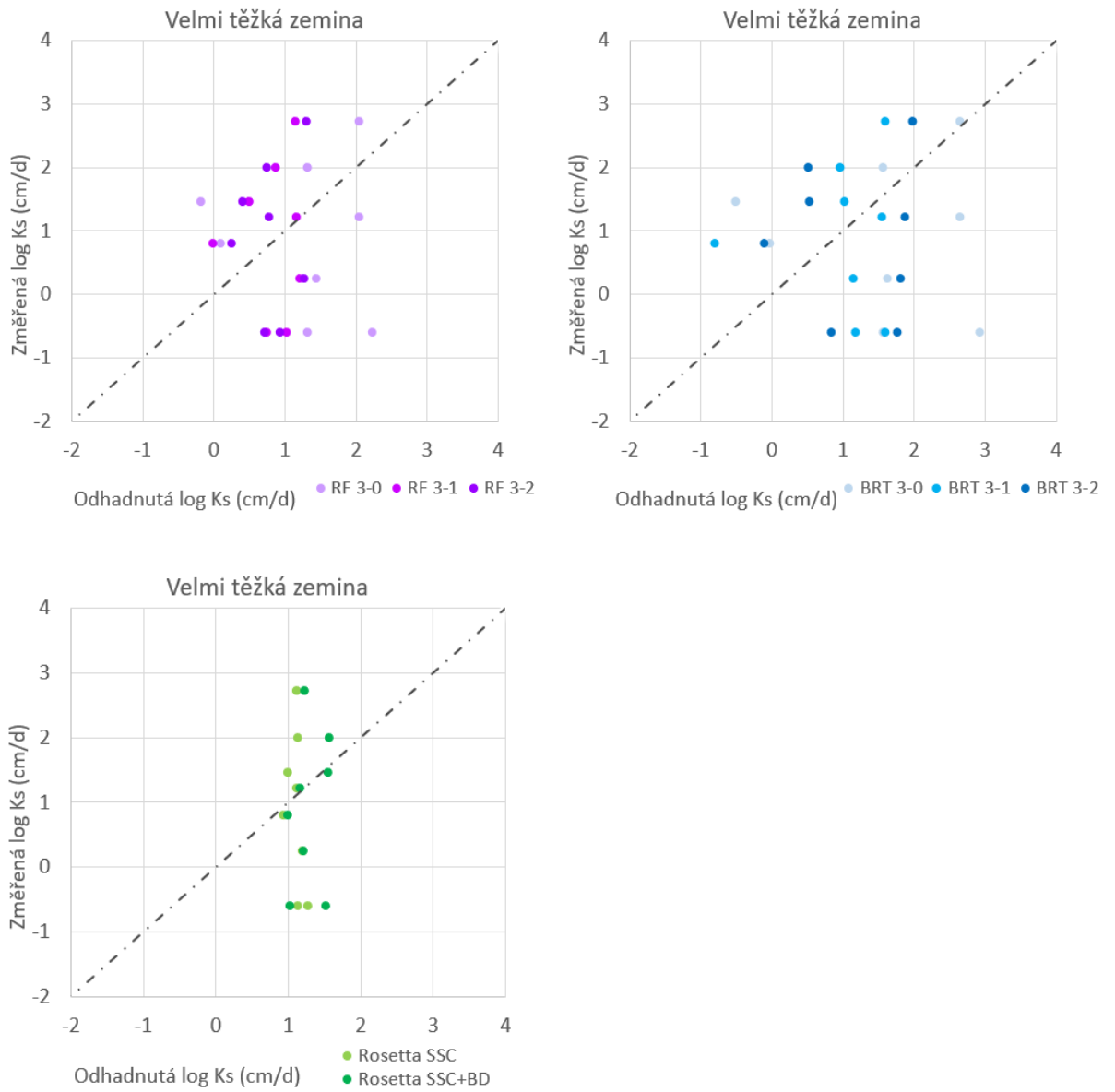
Obrázek 24 – Srovnání odhadnutých hodnot K_S pro seskupené zrnitostní třídy (1-5) pomocí boxplotů.



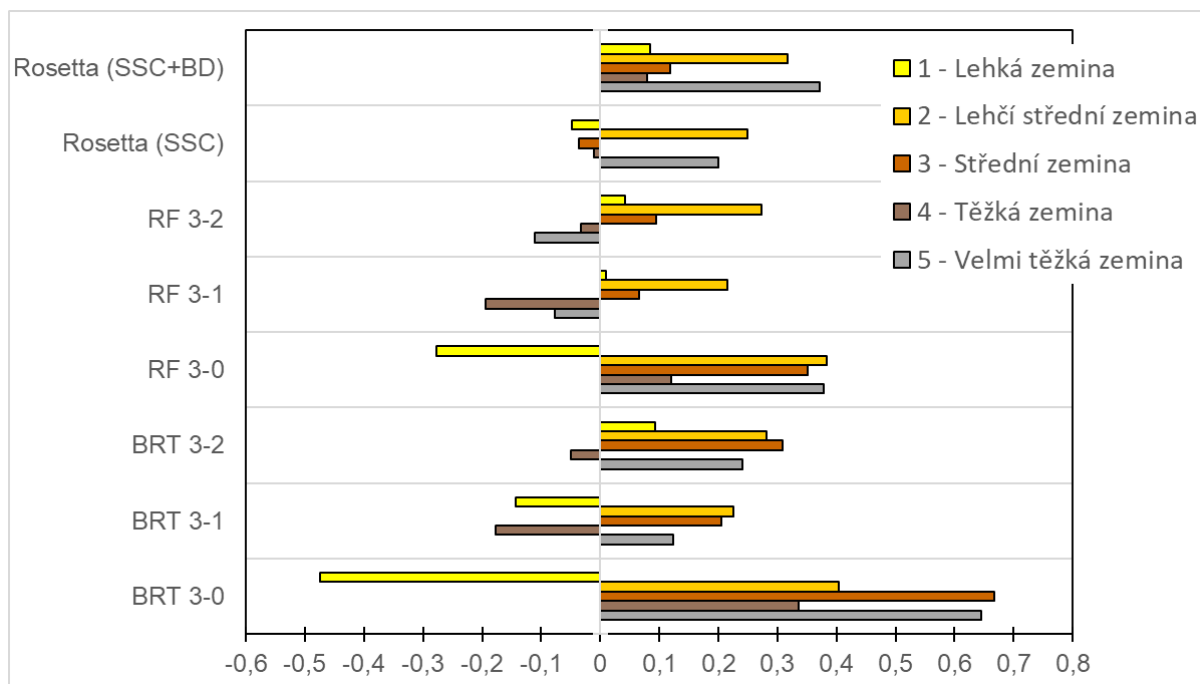
Obrázek 25a – Korelace mezi odhadnutou a změřenou K_S pro seskupené zrnitostní třídy 1 a 2.



Obrázek 25b – Korelace mezi odhadnutou a změřenou K_S pro seskupené zrnitostní třídy 3 a 4.



Obrázek 25c – Korelace mezi odhadnutou a změřenou K_S pro seskupenou zrnitostní třídu 5.



Obrázek 26 – Mean error použitých algoritmů pro seskupené zrnitostní třídy 1-5.

Pro první třídu – lehká zemina je nejnižší hodnota ME 0,01 u algoritmu RF 3-1. Pro druhou třídu – lehčí střední zemina je nejnižší hodnota 0,21 u algoritmu RF 3-1. Pro střední zeminu a těžkou zeminu jsou nejnižší hodnoty ME u algoritmu Rosetta SSC, -0,04 a -0,01. Poslední kategorie velmi těžká zemina má nejnižší hodnotu ME -0,08 u algoritmu RF 3-1 (Obrázek 26).

Tabulka 14 ukazuje hodnoty koeficientu determinace, kolik procent dat odpovídá danému regresnímu modelu. U nízkých hodnot R^2 je odhad velmi vzdálený změřené skutečnosti. Jedním z důvodů, proč byl odhad takto nekvalitní je možnost ovlivnění K_S preferenčním prouděním. Pokud $R^2 > 80\%$, pak je těsnost velmi vysoká, hodnoty mezi 50-80 % udávají těsnost vysokou, hodnoty mezi 25-50 % těsnost význačnou. Intenzita podkladu značí vyšší hodnoty R^2 .

Tabulka 14 - Koeficient determinace [%] pro jednotlivé modely PTF a seskupené zrnitostní třídy.

Seskupená třída	Použitá PTF						Rosetta (SSC)	Rosetta (SSC+BD)
	Araya; BRT 3-0	Araya; BRT 3-1	Araya; BRT 3-2	Araya; RF 3-0	Araya; RF 3-1	Araya; RF 3-2		
1 - Lehká zemina	3,58	1,63	0,49	0,07	0,39	2,19	35,40	16,55
2 - Lehčí střední zemina	0,41	4,96	2,97	4,90	0,06	0,19	16,08	6,88
3 - Střední zemina	5,19	2,55	6,62	2,63	0,38	2,01	3,31	24,96
4 - Těžká zemina	0,15	0,69	0,09	0,02	1,56	0,40	17,51	20,74
5 - Velmi těžká zemina	0,50	0,05	0,00	0,67	0,33	0,96	14,91	4,64

Tabulka 15 - Srovnání použitých PTF u jednotlivých tříd. Pořadí podle RMSE.

Seskupené zrnitostní třídy	Použitá PTF							
	Araya; BRT 3-0	Araya; BRT 3-1	Araya; BRT 3-2	Araya; RF 3-0	Araya; RF 3-1	Araya; RF 3-2	Rosetta (SSC)	Rosetta (SSC+BD)
1	8	6	7	5	3	4	1	2
2	6	8	7	4	3	5	1	2
3	7	5	4	8	6	3	2	1
4	8	7	4	5	6	3	2	1
5	8	5	6	7	3	1	4	2
Suma	37	31	28	29	21	16	10	8
Celkové pořadí	8	7	5	6	4	3	2	1

První tři místa v pořadí PTF pro seskupené zrnitostní třídy korespondují s pořadím PTF pro jednotlivé lokality. Pro seskupené zrnitostní třídy se na prvním místě umístil model (SSC+BD) programu Rosetta, využívající znalost textury a objemové hmotnosti suché půdy, na druhém místě model (SSC) programu Rosetta, využívající pouze znalosti textury. Na třetím a čtvrtém místě je model RF (3-2; 3-1) využívající stejných prediktorů jako model (SSC+BD), avšak u modelu 3-2 jsou navíc obsaženy hodnoty organického uhlíku (Tabulka 15).

6 Diskuse

Tato kapitola se věnuje shrnutí výsledků odhadů nasycené hydraulické vodivosti pomocí osmi modelů PTF a jejich vzájemným srovnáním se změřenými daty a také s již publikovanými výsledky odhadů K_S pomocí různých modelů PTF.

Pedotransferové funkce založené na strojovém učení odvozené na základě velkého množství dat pro půdy USA Araya a Ghezzehei (2019) lze, na základě výsledků pro 123 testovaných datových sad, vyhodnotit jako použitelné pro půdy České republiky se srovnatelnou přesností, a to i přes skutečnost, že RMSE odhadů K_S pomocí programu Rosetta využívající analýzu ANN byla nižší než pro modely strojového učení RF a BRT. Rozdíl mezi průměrnou hodnotou RMSE pro jednotlivé lokality mezi modely Rosetta (SSC a SSC+BD) a RF3-2 byl velmi malý, 0,02 (log K_S v cm/den).

Průměrná chyba RMSE programu Rosetta (SSC) vyhodnocená pro jednotlivé lokality byla 0,91 log K_S (cm/den) a u modelu (SSC+BD) také 0,91 log K_S (cm/den). U PTF v programu Rosetta tedy nebyl potvrzen předpoklad, že s přidáním dalšího prediktoru se zpřesní odhad K_S . Průměrná chyba RMSE u algoritmu RF3-2 byla 0,93 log K_S (cm/den), dále u RF3-1 byla 0,97 a u RF3-0 1,11 log K_S (cm/den). Tedy s přidáním dalších prediktorů (objemové hmotnosti suché půdy a obsahu organického uhlíku v půdě) se zpřesnil odhad nasycené hydraulické vodivosti. Pro testované lokality České republiky se nejméně hodil algoritmus BRT. Hodnoty RMSE byly pro všechny tři úrovně modelu BRT vyšší než 1. Hodnota RMSE u BRT3-0 byla 1,24 log K_S (cm/den), u BRT3-1 1,06 a u BRT3-2 1,02. Pro modely BRT se také potvrdil předpoklad, že se odhady zpřesňují s přidáním dalšími prediktory.

Výsledky lze srovnat se zahraničními pracemi. Přímo autoři Araya a Ghezzehei udávají RMSE <0,7, avšak jejich aplikace vznikla na podkladě dat USDA. 75 % dat autoři použili jako tréninková data pro samoučící algoritmy BRT a RF a zbylých 25 % využili k ověření správnosti PTF. Obrázek 20 popisuje rozmístění použitých dat pro tréninkovou množinu určenou samoučícím algoritmům strojového učení programu PTFapp. Velké zastoupení zde má kategorie písčitých půd – lehké půdy (1). Rozmístění použitých dat pro tréninkovou množinu koresponduje s ME.

V práci „New generation of hydraulic pedotransfer functions for Europe“ (Tóth et al. 2014) udávají autoři výsledky RMSE nasycené hydraulické vodivosti v rozmezí 0,90-1,36 log K_S (cm/den). Studie „Probabilistic Approach to the Identification of Input Variables to Estimate Hydraulic Conductivity“ (Lilly et al. 2008) udává hodnoty RMSE v rozmezí 0,83-1,14 log K_S (cm/den), průměr hodnot RMSE je 0,97. Výkon jednotlivých modelů závisí na charakteristice použitých dat. Údaje RMSE zde zmíněné poukazují na obtížnost odhadů charakteristiky K_S , která je v půdě velmi variabilní v místě i čase. Hodnoty RMSE je vždy třeba posoudit ve vztahu k rozsahu měřených hodnot K_S .

Nejnižší hodnota RMSE, tj. nejpřesnější odhady K_S , poskytoval algoritmus RF3-2 pro lokalitu Džbánov, a to 0,16. Všechny vzorky lokality Džbánov jsou zařazeny do třídy Silt loam. Nejvyšší hodnota RMSE, tj. nejméně přesně odhady K_S , poskytoval algoritmus RF3-0, pro lokalitu Ledenice, a to 2,34. Vzorky lokality Ledenice jsou zařazeny do tříd Loam (5 vzorků), Silt loam (2 vzorky) a Sandy loam (1 vzorek).

U výše zmíněných studií nebyla uvedena metoda přímého měření, se kterou se odhadnuté hodnoty K_S srovnávají. Data poskytnutá pro diplomovou práci byla změřena různými metodami (v laboratoři – přístroj s konstantním spádem, přístroj s proměnným spádem, terénní měření – tlakový infiltrometr, jednosondová metoda), zpravidla pro orníční vrstvu. Neexistuje referenční metoda pro měření nasycené hydraulické vodivosti, proto je obtížné změřená a odhadnutá data z různých lokalit mezi sebou srovnávat. V České republice je třeba při hodnocení K_S zohlednit i roční období a aktuální počasí, které mohou mít na její hodnotu značný vliv. Rozdílnou hodnotu K_S naměříme na jaře po vydatných deštích, či v srpnu po dlouhodobém suchu.

U lokality Ruzyně probíhalo přímé stanovení K_S dvěma metodami: tlakovým infiltrometrem (Matula & Kozáková 1997) a Hood infiltrometrem (UGT). Vyhodnocení a srovnání odhadů bylo stanoveno pro obě sady měření (vždy 12 měření – tedy 24 údajů). Odhady K_S u lokality Ruzyně u všech modelů PTF byly systematicky podhodnocovány. Jedním z důvodů je pravděpodobnost výskytu preferenčního proudění, jelikož měření probíhala na přelomu června a července, kdy byla půda relativně suchá a vyskytovaly se praskliny. Jedná se o zemědělskou půdu, která je sice zrnitostně poměrně homogenní (v rámci seskupených zrnitostních tříd 3 a 4), nicméně jílnaté částice reagují na změnu půdní vlhkosti bobtnáním a smršťováním. Významný vliv na změřenou hodnotu K_S bude mít rozdílné zpracování půdy, které je na experimentální lokalitě dlouhodobě aplikováno, jedná se o konvenční orbu, redukovanou orbu a bezorebný systém zpracování půdy. Koeficient determinace odhadů na této lokalitě dosahuje pouze do výše 6,13 % (model Rosetta SSC), a RMSE je u všech modelů vysoká (0,7-1,78).

Ani odhady v lokalitách Majetín, Ledenice a Slabčice nevykazují výrazně kvalitnější odhady. Nejvyšší hodnota R^2 pro Majetín není vyšší než 9 % (model Rosetta SSC+BD), RMSE je v rozmezí 1,25-1,89. Půdy jsou zde zařazeny do seskupených zrnitostních tříd 4 - těžká zemina (6 vzorků), 5 - velmi těžká zemina (3 vzorky), 3 - střední zemina (5) a 2 - lehčí střední zemina (1). Z celkových 15 vzorků převažují těžší půdy. Právě vyšší zastoupení těžších půd by mohlo vysvětlovat nízkou přesnost všech modelů PTF. Z 8 vzorků pro zájmovou oblast Ledenice je 7 vzorků ze seskupené zrnitostní třídy 3 - střední zemina. Přestože odhadnuté hodnoty K_S dobře korelují se změřenými daty, nejvyšší R^2 je pro model Rosetta (SSC+BD) a dosahuje výše 76,4 %, všechny hodnoty jsou nadhodnocené, což je ukázáno na kladných hodnotách ME (Obrázek 23) a relativně vysokými hodnotami RMSE 1,36-2,34. Podobné je tomu u lokality Slabčice, kde koeficient determinace je nejvyšší u modelu RF3-0 a to 59,08 %, nicméně RMSE je opět vysoká (1,27-1,99).

Vysoký koeficient R^2 poukazuje na vysokou korelaci mezi daty, ale ne vždy ukazuje na kvalitní odhad. Z tohoto důvodu je konečné pořadí testovaných modelů PTF určeno na základě RMSE.

Nejlepší výsledky podávaly modely v lokalitách Podlesí, U dvou louček, Džbánov, Žichlínek a Mimoň. Při nezapočítání modelu BRT3-0, který se při vyhodnocení umístil vždy na posledním místě, je u všech modelů pro lokalitu Podlesí RMSE <0,6. R^2 je maximální pro RF3-0 a to 70,87 %. Z obrázku 21 je patrné, že modely BRT3-0 i RF3-0 výrazně nadhodnocovaly hodnoty nasycené hydraulické vodivosti. Pro přesnější odhady je důležitá znalost objemové hmotnosti suché půdy. Po přidání objemové hmotnosti suché půdy jako dalšího prediktoru došlo u algoritmů k významnému zlepšení odhadů K_S . Toto tvrzení bude

patrné i v dalších popsaných lokalitách. Model Rosetta (SSC+BD) podhodnocoval odhadnuté hodnoty K_S ve srovnání se změřenými hodnotami. Avšak pro lokalitu Podlesí byly dostupné pouze 4 vzorky, pro vyslovení dalších závěrů by bylo potřeba otestovat více vzorků. Bez opětovného nezapočtení modelu BRT3-0 i odhady u lokality U dvou louček vychází srovnatelně; RMSE <0,7. Koeficient determinace i přes vyšší hodnotu RMSE je nejlepší u modelu BRT3-0 a to 97,61 %. Takto vysoká hodnota je nejspíše proto, že byly dostupné pouze 3 vzorky pro testování. Výsledky mohou být zkreslené nedostatkem dat. Pro Džbánov a Žichlínek byly otestovány vždy 4 vzorky, RMSE (bez BRT3-0) je pro Džbánov <0,6 a pro Žichlínek <0,7, u obou lokalit je těsnost velmi vysoká (>80 %). Pro Džbánov je $R^2=95,29$ % pro algoritmus RF3-2, pro Žichlínek 88,60 % pro Rosettu (SSC+BD). Dle obrázku 21 je zde patrné, že model RF3-0 a BRT3-0 u obou zmíněných lokalit opět značně nadhodnocuje. Potvrzuje tím důležitost dalších prediktorů, především objemové hmotnosti suché půdy. Program Rosetta založený na analýze neuronových sítí i přes použití objemové hmotnosti suché půdy jako dalšího prediktoru (Rosetta SSC+BD) znatelně hodnoty K_S podhodnocuje. Nej kvalitnější odhady byly pozorovány u lokality Mimoň. Pokud zohledníme vyšší počet vzorků (9 vzorků oproti 3-4 u výše zmiňovaných lokalit Džbánov, Žichlínek, Podlesí a U dvou louček), výsledné statistiky ukazují RMSE <0,7 u všech modelů a R^2 je maximální pro BRT3-0 (46,88 %, význačná těsnost). Srovnání odhadovaných hodnot pomocí boxplotů na obrázku 21 ukazuje, že si všechny modely vedly srovnatelně. Pouze s přidáním dalšího prediktoru k modelům BRT, RF, a to hodnoty organického uhlíku v půdě (%), došlo k mírnému nadhodnocování hodnot K_S . Naproti tomu Rosetta (SSC) hodnoty mírně podhodnocuje. V zájmovém území Mimoň se nachází lehčí půdy – seskupené zrnitostní třídy 1 - lehká zemina (7 vzorků) a 2 - lehčí střední zemina (2 vzorky). Takovéto zrnitostní zařazení koresponduje s tréninkovou množinou autorů Araya, Ghezzehei pro ML. Je zde patrný jistý trend a bylo by zajímavé tento trend sledovat a otestovat na dalších vzorcích z podobné texturní třídy.

V této studii byly srovnatelně kvalitní výsledky s hodnotou RMSE <1 získány u všech osmi aplikovaných modelů PTF pouze pro seskupené zrnitostní třídy 1 - lehká zemina (Sand) a 2 - lehčí střední zemina (Loamy sand, Sandy loam) a to pravděpodobně z důvodu velkého zastoupení této texturní třídy v rámci tréninkové množiny PTF Araya a Ghezzehei (2019).

Kromě vlivu již zmíněných preferenčních cest, by bylo třeba při odhadech K_S pomocí PTF zohlednit i vliv půdních organismů, především zoedafon (živočichové žijící v půdě, např. krtci a hraboši), či vliv kořenového systému rostlin. V současné době je na lokalitě Ruzyně zkoumán i vliv hnojení půdy ve vztahu k K_S . Tyto vlivy je ale velmi obtížné nějakým způsobem při odhadech zohlednit, obzvláště vliv přítomnosti zoedafonu. Strojovým učením se díky používaným mechanismům lze přiblížit průměrným měřeným hodnotám, které v praxi mohou být změřeny ve značném rozsahu.

Z literární rešerše je patrná důležitost přesných odhadů nasycené hydraulické vodivosti pro její využití v přírodních vědách i v rámci ochrany životního prostředí. Se zlepšující se technikou není obtížné pro menší plochy území stanovit nasycenou hydraulickou vodivost přímo měřením, avšak především pro velké plochy je nepřímé stanovení klíčové.

Zpřesňování odhadů nasycené hydraulické vodivosti nadále zůstává tématem pro další výzkum.

7 Závěr

- V teoretické části práce byly shromážděny informace o přímém i nepřímém stanovení K_S , PTF a jejich použití pro odhad půdních hydrofyzikálních charakteristik, se zaměřením na odhad K_S .
- V praktické části byla otestována a vyhodnocena použitelnost PTF pro odhad K_S založených na principu strojového učení (RF a BRT) odvozených pro půdy USA Araya a Ghezzehei (2019) pro použití na půdy v České republice. Bylo otestováno celkem 126 údajů ze 13 různých lokalit s různým zrnitostním složením. Data pocházela z databáze HYPRESCZ (Miháliková et al. 2013) a z měření provedených na Katedře vodních zdrojů FAPPZ ČZU v Praze (Ruzyně, Strašov, Uherčice). Pro srovnání byly odhadnuty hodnoty K_S i pomocí programu Rosetta (Schaap et al. 2001), který je pro odhad hydrofyzikálních charakteristik půdy včetně K_S používán již 20 let. Veškeré odhady z celkem 8 modelů PTF, byly porovnány a statisticky vyhodnoceny vzhledem k naměřeným hodnotám K_S ($\log K_S$) pomocí statistik ME, R^2 a RMSE. Na základě hodnot RMSE bylo určeno pořadí jednotlivých modelů PTF jak pro použití na jednotlivých lokalitách, tak i pro seskupené zrnitostní třídy.
- Autoři testovaných PTF založených na strojovém učení (RF a BRT) uváděli RMSE $<0,7 \log K_S$ (cm/den) pro všechny své modely testované pro půdy USA. Pro půdy České republiky vycházela chyba RMSE v rozmezí 0,16-2,34 $\log K_S$ (cm/den) pro jednotlivé lokality. Nejpřesnější odhady (pořadí 1-4), u kterých RMSE <1 v průměru poskytoval program Rosetta (modely SSC, SSC+BD) a to 0,91, dále RF3-2 0,93 a RF3-1 0,97 $\log K_S$ (cm/den).
- Výsledky odhadů PTF vyhodnocených pro jednotlivé seskupené zrnitostní třídy ukazují důležitost zdrojových dat a důležitost jejich podobnosti s půdami, pro které je hodnota K_S odhadována. Velmi nízké hodnoty RMSE poukazující na vysokou kvalitu odhadovaných hodnot K_S byly pozorovány pro seskupenou zrnitostní třídu 1 (Sand, Loamy sand), pro kterou byla zjištěna nejnižší hodnota ME u algoritmu RF3-1, a to 0,01, nejnižší hodnota RMSE byla pozorována pro model Rosetta (SSC) a to 0,26 $\log K_S$ (cm/den), pro model Rosetta (SSC+BD) 0,32 a pro model RF3-1 0,43. Jak ukazuje obrázek 25 korelace mezi změřenou a odhadnutou hodnotou K_S je velká především u lehkých zemin a lehčích středních zemin. Pro střední a těžké zeminy, které tvořily většinu testovaných půd v ČR nebyly odhady tak kvalitní jako pro půdy lehké, avšak vzhledem k velkému rozsahu měřených hodnot na jednotlivých lokalitách stále použitelné. Pokud je K_S odhadována pro půdy, u kterých se mohou vlivem orby, sucha či zoedafonu vyskytovat preferenční cesty, je velmi obtížné tento vliv pomocí PTF odhadnout, obzvláště pak model Rosetta (SSC), který pouze na základě údajů o zrnitosti nemůže tento vliv reflektovat.
- Výše uvedené závěry naznačují určité trendy, které by bylo vhodné ověřit na rozsáhlejším souboru dat. Pro další výzkum by bylo kromě rozšíření základny zdrojových dat také vhodné otestovat možnost použití dat o půdách v České republice v rámci tréninkových dat aplikace využívající ML techniky (RF a BRT). Tento krok by mohl zvýšit přesnost odhadu algoritmu RF programu Araya, Ghezzehei (2019), který i přes odvození pro půdy USA vykazuje velký potenciál pro půdy ČR.

8 Literatura

- (1) Aguilera H, Guardiola-Albert C, Moreno Merino L, Baquedano C, Díaz-Losada E, Robledo Ardila PA, Durán Valsero JJ. 2022. Building inexpensive topsoil saturated hydraulic conductivity maps for land planning based on machine learning and geostatistics. *CATENA* **208**: 105788.
<https://doi.org/10.1016/J.CATENA.2021.105788>
- (2) Anděl J. 2019. *Statistické metody*. Matfyzpress, Praha. ISBN: 978-80-7378-381-5.
- (3) Anděl J. 2011. *Základy matematické statistiky*. Matfyzpress, Praha. ISBN: 978-80-7378-162-0.
- (4) Araya SN, Ghezzehei TA. 2019. Using Machine Learning for Prediction of Saturated Hydraulic Conductivity and Its Sensitivity to Soil Structural Perturbations. *Water Resources Research* **55**: 5715–5737. <https://doi.org/10.1029/2018WR024357>
- (5) Bářková K, Matula S, Miháliková M. 2013. *Multimediální učebnice hydroopedologických terénních měření*. 2. doplněné vydání. Česká zemědělská univerzita v Praze. Praha. Available from <http://hydropedologie.agrobiologie.cz> (accessed 13.3.22).
- (6) Biau G, Scornet E. 2016. A random forest guided tour. *Test* **25**: 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- (7) Bouma J. 1989. Using Soil Survey Data for Quantitative Land Evaluation. *Advances in Soil Science* **9**: 177–213. https://doi.org/10.1007/978-1-4612-3532-3_4
- (8) Brown GO. 2002. Henry Darcy and the making of a law. *Water Resources Research* **38**: 11-1-11-12. <https://doi.org/10.1029/2001wr000727>
- (9) Cosby BJ, Hornberger GM, Clapp RB, Ginn TR. 1984. A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils. *Water Resources Research*. **20**: 682-690.
- (10) Český hydrometeorologický ústav, 2021. Úhrn srážek v roce 2020. Available from https://www.chmi.cz/files/portal/docs/meteo/ok/charakteristiky_klimatu/img/S_RA_2020.gif (accessed 13.3.22).
- (11) Hillel D. 1998. *Environmental Soil Physics. Fundamentals, Applications, and Environmental Considerations*, 1st edition, Academic Press. ISBN: 978-01-2348-5250.
- (12) Dexter AR, Richard G, Arrouays D, Czyz EA, Jolivet C, Duval O. 2008. Complexed organic matter controls soil physical properties. *Geoderma* **144**: 620–627. <https://doi.org/10.1016/J.GEODERMA.2008.01.022>
- (13) Ernst LF. 1950. A new formula for the calculation of the permeability factor with the auger hole method. T.N.O. Groningen. Translated from the Dutch by H. Bouwer, Cornell Univ. Ithaca, N.Y., 1955.
- (14) Fiala P, Poláková Š, Sáňka M, Vácha R. 2018. *Kritéria pro hodnocení produkčních a ekologických vlastností půd*. Ministerstvo životního prostředí, Praha. ISBN: 978-80-7212-627-9.
- (15) Freeze RA, Cherry JA. 1979. *Groundwater*. ISBN: 978-01-3365-312-0.
- (16) Hooghoudt SB. 1936. Bepaling van den doorlaatfactor van den grond met behulp van pomproeven, (z.g. boorgatenmethode). *Verslag Landbouw* **42**: 449–541.

- (17) Hvorslev M. 1951. Time lag and soil permeability in groundwater observations. Experiment St. Corps of Eng., U.S. Army 26.
- (18) Kirkham D, van Bavel CHM. 1948. Theory of seepage into auger holes. Soil Science Society of America Proceedings **13**: 75–82.
- (19) Kozák J, Němeček J, Matula S, Borůvka L, Valla M. 2011. Pedologie. Česká zemědělská univerzita. Fakulta agrobiologie, potravinových a přírodních zdrojů, Praha. ISBN: 978-80-213-0907-4.
- (20) Kresic N. 2006. Hydrogeology and Groundwater Modeling. CRC Press, Florida. ISBN: 978-04-291-2210-1.
- (21) Kulhavý Z, Jakoubek J, Matula S, Šťastná M. 2021. Snižování zátěže povrchových vod zdroji plošného zemědělského znečištění při uplatnění regulace drenážního odtoku na stávajících stavbách zemědělského odvodnění. Redakčně upravená roční zpráva výzkumného projektu. VÚMOP, ČZU, MENDELU, Agroprojekce Litomyšl.
- (22) Lilly A, Nemes A, Rawls WJ, Pachepsky YaA. 2008. Probabilistic Approach to the Identification of Input Variables to Estimate Hydraulic Conductivity. Soil Science Society of America Journal **72**: 16–24. <https://doi.org/10.2136/SSSAJ2006.0391>
- (23) Matula S, Kozáková H. 1997. A simple pressure infiltrometer for determination of soil hydraulic properties by in situ infiltration measurements. Rostlinná výroba **43**: 405-413.
- (24) Meerschaert MM, Dogan M, van Dam RL, Hyndman DW, Benson DA. 2013. Hydraulic conductivity fields: Gaussian or not? Water Resources Research **49**: 4730-4737.
- (25) METER group, SATURO. Field Saturated Hydraulic Conductivity Available from <https://www.metergroup.com/en/meter-environment/products/saturo-infiltrometer-field-saturated-hydraulic-conductivity> (accessed 4.3.22).
- (26) Miháliková M. 2011. Pedotransferové funkce pro odhad hydrofyzikálních vlastností půd [Ph.D. Thesis]. Česká zemědělská univerzita, Praha.
- (27) Miháliková M, Matula S, Doležal F. 2014. Application of k-nearest code for the improvement of class pedotransfer functions and countrywide field capacity and wilting point maps. Soil and Water Research **9**: 1–8. <https://doi.org/10.17221/44/2013-SWR>
- (28) Miháliková M, Matula S, Doležal F. 2013. HYPRESCZ – Database of soil hydrophysical properties in the Czech Republic. Soil and Water Research **8**: 34–41. <https://doi.org/10.17221/58/2012-swr>
- (29) Minasny B, Mc Bratney AB, Bristow KY. (1999): Comparison of different approaches to the development of pedotransfer functions for water retention curves. Geoderma, **93** (3-4): 225-253.
- (30) Nemes A, Roberts R T, Rawls WJ, Pachepsky YA, van Genuchten MT. 2007. Software to estimate -33 and -1500 kPa soil water retention using the non-parametric k-Nearest Neighbor technique. Environmental Modelling and Software. **23**: 254-255. <https://doi.org/10.1016/j.envsoft.2007.05.018>
- (31) Nemes A, Schaap MG, Wösten JHM. 2003. Functional Evaluation of Pedotransfer Functions Derived from Different Scales of Data Collection. Soil Science Society of America Journal **67**: 1093–1102. <https://doi.org/10.2136/SSSAJ2003.1093>

- (32) Němeček J, Muhlanselová M, Macků J, Vokoun J, Vavříček D, Novák P. 2011. Taxonomický klasifikační systém půd České republiky. 2. uprav.vyd. Česká zemědělská univerzita, Praha. ISBN: 978-80-213-2155-7.
- (33) Němeček J, Macků J, Vavříček D, Novák P. 2001. Taxonomický klasifikační systém půd České republiky. ČZU Praha spolu s VÚMOP Praha, Praha. ISBN: 80-238-8061-6.
- (34) Oleson KW, Lawrence DM, Bonan GB. 2013. Technical description of version 4.5 of the Community Land Model (CLM). Near Tech. Note NCAR/TN-503+STR. National Center for Atmospheric Research, Boulder. *Geophysical Research Letters* **37**.
- (35) Parr JF, Bertrand AR. 1960. Water Infiltration Into Soils. *Advances in Agronomy* **12**: 311–363. [https://doi.org/10.1016/S0065-2113\(08\)60086-3](https://doi.org/10.1016/S0065-2113(08)60086-3)
- (36) Pavlů L. 2018. Základy pedologie a ochrany půdy. Česká zemědělská univerzita v Praze, Praha. ISBN: 978-80-213-2876-1.
- (37) Philip JR. 1957. The theory of infiltration: 4. Sorptivity and algebraic infiltration equations. *Soil Science* **84**: 257–264.
- (38) Pribyl DW. 2010. A critical review of the conventional SOC to SOM conversion factor. *Geoderma* **156**: 75–83. <https://doi.org/10.1016/J.GEODERMA.2010.02.003>
- (39) Puckett WE, Dane JH, Hajek BF. 1985. Physical and Mineralogical Data to Determine Soil Hydraulic Properties. *Soil Science Society of America Journal* **49**: 831–836. <https://doi.org/https://doi.org/10.2136/sssaj1985.03615995004900040008x>
- (40) Rawls WJ, Pachepsky YA, Ritchie JC, Sobecki TM, Bloodworth H. 2003. Effect of soil organic carbon on soil water retention. *Geoderma* **116**: 61–76. [https://doi.org/10.1016/S0016-7061\(03\)00094-6](https://doi.org/10.1016/S0016-7061(03)00094-6)
- (41) Salač J. 2018. Stanovení hydraulických charakteristik půdy ve vybraném zájmovém území [MSc. Thesis]. Vysoké učení technické, Brno.
- (42) Saxton KE, Rawls WJ. 2006. Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions. *Soil Sci. Soc. Am. J.* **70**:1569–1578
- (43) Schaap MG, Leij FJ. 1998. Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil & Tillage Research* **47**: 37–42.
- (44) Schaap MG, Leij FJ, van Genuchten MT, Brown GE. 2001. Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology* **251**: 163-176 [https://doi.org/10.1016/S0022-1694\(01\)00466-8](https://doi.org/10.1016/S0022-1694(01)00466-8)
- (45) Schapire RE. 1999. A Brief Introduction to Boosting. Pages 1-6. Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence.
- (46) Soilmoisture Equipment Corp. Chameleon KSAT. Available from <https://www.soilmoisture.com/Chameleon/> (accessed 5.3.2022).
- (47) Soil Survey Staff 1999. Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys., 2nd ed. Natural Resources Conservation Service. U.S. Department of Agriculture Handbook 436.
- (48) Straková K. 2016. Datamining a využití rozhodovacích stromů při tvorbě Scorecards [MSc. Thesis]. Vysoká škola ekonomická v Praze, Praha.
- (49) Šimek M et al. 2019. Živá půda. Ekologie, využívání a degradace půdy. Academia, Praha. ISBN: 978-80-200-2976-8.

- (50) Špongrová K. 2003. Pedotransferové funkce jako metody odhadu hydrofyzikálních vlastností půd [MSc. Thesis]. Česká zemědělská univerzita, Praha.
- (51) Tomasella J, Pachepsky Y, Crestana S, Rawls WJ. 2003. Comparison of Two Techniques to Develop Pedotransfer Functions for Water Retention. *Soil Science Society of America Journal* **67**: 1085–1092. <https://doi.org/10.2136/SSSAJ2003.1085>
- (52) Tóth B, Weynants M, Nemes A, Makó A, Bilas G, Tóth G. 2014. New generation of hydraulic pedotransfer functions for Europe. *European Journal of Soil Science* **66**: 226–238. <https://doi.org/10.1111/EJSS.12192>
- (53) Umwelt-Geräte-Technik GmbH. Hood-infiltrimeter/IL-2700. Available from <https://www.ugt-online.de/en/products/soil-science/field-instruments/hood-infiltrimeter/> (accessed 9.4.2022).
- (54) U. S. Department of Agriculture. 2019. ROSETTA Hydraulic Functions: USDA ARS Available from <https://www.ars.usda.gov/pacific-west-area/riverside-ca/agricultural-water-efficiency-and-salinity-research-unit/docs/model/rosetta-hydraulic-functions/> (accessed 13.3.22).
- (55) U. S. Department of Agriculture. 1951. Soil Survey Manual. Soil Conservation Service. U.S. Department of Agriculture Handbook. US Government Printing Office, Washington DC.
- (56) Valla M, Kozák J, Němeček J, Matula S, Borůvka L, Drábek O. 2008. Pedologické praktikum. Česká zemědělská univerzita v Praze. Fakulta agrobiologie, potravinových a přírodních zdrojů, Praha. ISBN: 978-80-213-0914-2.
- (57) Volná E. 2002. Neuronové sítě 1. Ostravská univerzita, Ostrava.
- (58) Vulpiani A. 2014. Lewis Fry Richardson: scientist, visionary and pacifist. *Lettera Matematica* **2**: 121-128. <https://doi.org/10.1007/s40329-014-0063-z>
- (59) Weynants MA, et al. 2013. European HYdropedological Data Inventory (EU-HYDI). Pages 64-68. European Commission Joint Research Centre Institute for Environment and Sustainability. <https://doi.org/10.2788/5936>
- (60) Wösten JHM, Finke PA, Jansen MJW. 1995. Comparison of class and continuous pedotransfer functions to generate soil hydraulic characteristics. *Geoderma* **66**: 227-237. [https://doi.org/10.1016/0016-7061\(94\)00079-P](https://doi.org/10.1016/0016-7061(94)00079-P)
- (61) Wösten JHM, Lilly A, Nemes, A, le Bas C. 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* **90**: 169–185. [https://doi.org/10.1016/S0016-7061\(98\)00132-3](https://doi.org/10.1016/S0016-7061(98)00132-3)
- (62) Zhang Y, Schaap MG. 2019. Estimation of saturated hydraulic conductivity with pedotransfer functions: A review. *Journal of Hydrology* **575**: 1011–1030. <https://doi.org/10.1016/j.jhydrol.2019.05.058>

9 Seznam tabulek

Tabulka 1 - Modely H1-H5 programu Rosetta.	26
Tabulka 2 - Modely BRT programu autorů Araya, Ghezzehei.	33
Tabulka 3 - Modely RF programu autorů Araya, Ghezzehei.	34
Tabulka 4 - Zrnitostní třídy. Označení tříd podle USDA a jejich český ekvivalent.	35
Tabulka 5 - Počet vzorků zájmových lokalit. *Pro lokalitu Ruzyně jsou dostupné 2 sady (24) naměřených hodnot Ks (tlakový infiltrometr, Hood infiltrometr).	39
Tabulka 6 - Základní informace o jednotlivých lokalitách.	42
Tabulka 7 - Seskupení zrnitostních tříd podle Němečka et al. (2001).	43
Tabulka 8 - Data potřebná k odhadu Ks algoritmem BRT.	46
Tabulka 9 - Data potřebná k odhadu Ks algoritmem RF.	46
Tabulka 10 - Data potřebná k odhadu Ks programem Rosetta.	47
Tabulka 11 - Koeficient determinace [%] pro jednotlivé modely PTF a jednotlivé lokality.	57
Tabulka 12 - Hodnoty RMSE pro vybrané lokality; tučně jsou označené hodnoty menší než 0,7 (srovnatelné s výsledky autorů Araya, Ghezzehei, 2019).	58
Tabulka 13 - Srovnání použitých algoritmů k určení Ks u jednotlivých lokalit; pořadí 1-8 podle RMSE (pořadí 1 získal model s nejnižší hodnotou RMSE reprezentující nejkvalitnější odhad).	59
Tabulka 14 - Koeficient determinace [%] pro jednotlivé modely PTF a seskupené zrnitostní třídy.	64
Tabulka 15 - Srovnání použitých PTF u jednotlivých tříd. Pořadí podle RMSE.	65

10 Seznam obrázků

Obrázek 1 - Ilustrace Darcyho pokusu od autorů R. Alan Freeze a John A. Cherry z knihy Groundwater (1979).	12
Obrázek 2 - Hodnoty nasycené hydraulické vodivosti a propustnosti podle autorů R. Alan Freeze a John A. Cherry z knihy Groundwater (1979). Zleva: krasový vápenec, propustný čedič, zvětralé vyvřelé a metamorfované horniny, vápence a dolomity, pískovce, neporušené metamorfované nebo vyvřelé horniny, břidlice, nezvětralý jíl, ledovcový till, prach, spraš, prachovitý písek, písek, štěrk.	14
Obrázek 3 - Sonda před odstraněním vody kalovkou nebo čerpadlem (Bátková et al. 2013).	16
Obrázek 4 - Dvouválcový infiltrometr (Bátková et al. 2013).	16
Obrázek 5 - Tlakový infiltrometr (Matula & Kozáková 1997), foto: S. Matula.	17
Obrázek 6 - KSAT. Přístroj k laboratornímu měření nasycené hydraulické vodivosti (METER)	19
Obrázek 7 - Mualem-van Genuchtenovy parametry pro 11 tříd (Wösten et al. 1999).	21
Obrázek 8 - Rosetta – van Genuchtenovy hydraulické parametry (USDA 2019).	22
Obrázek 9 - Úrovně dostupnosti vstupních dat. Vlevo jsou nejsnáze dostupná data. (Miháliková 2011)	25

Obrázek 10 - Matematický model neuronu. (Zdroj: https://portal.matematickabiologie.cz/index.php?pg=analyza-a-hodnoceni-biologickych-dat--umela-inteligence--neuronove-site-jednotlivy-neuron--jednotlivy-neuron--matematicky-model-a-aktivni-dynamika-neuronu)	27
Obrázek 11 - Prostředí programu Rosetta.	29
Obrázek 12 - Program ke stanovení nasycené hydraulické vodivosti v prostředí R studia. ...	31
Obrázek 13 - Rozhodovací strom pro nalezení čísla v rozsahu 0-15. (Zdroj: https://popelka.ms.mff.cuni.cz/~lessner/mw/index.php/U%C4%8Debnice/Informace/Rozhodovac%C3%AD_stromy_a_chytr%C3%A9_ot%C3%A1zky)	32
Obrázek 14 - Trojúhelníkový diagram na stanovení druhu půdy podle obsahu jílu, prachu a písku (USDA 1951).....	35
Obrázek 15 - Velikostní kategorie půdních minerálních částic USDA (J. Fox, Ohio EPA); Zrnitostní kategorie: clay = jíl, silt = prach, sand = písek (very fine sand = velmi jemný písek, fine sand = jemný písek, medium sand = střední písek, coarse sand = hrubý písek, very coarse sand = velmi hrubý písek, gravel = štěrk).....	36
Obrázek 16 - Hydrolimity – závislost na půdním druhu a objemové vlhkosti půdy (Pavlů 2018)	37
Obrázek 17 - Zájmové lokality. Podkladová mapa: ČÚZK. Autor: M. Miháliková.	40
Obrázek 18 - Zrnitostní složení půd na testovaných lokalitách a jejich zařazení podle texturního trojúhelníku USDA se zařazením do 5 seskupených zrnitostních tříd (1-5) podle Němečka et al. (2001).	44
Obrázek 19 - Počet vzorků v seskupených zrnitostních třídách. 1–Lehká zemina, 2–Lehčí střední zemina, 3–Střední zemina, 4–Těžká zemina, 5–Velmi těžká zemina.....	44
Obrázek 20 - Texturní trojúhelník – rozmístění vzorků použitých jako tréninková data (75 %) pro aplikaci PTFapp (Araya & Ghezzehei 2019).....	45
Obrázek 21 - Srovnání jednotlivých lokalit pomocí boxplotů	49
Obrázek 22 - Korelace mezi změřenou a odhadnutou K_s	53
Obrázek 23 - Mean error testovaných PTF pro jednotlivé lokality.....	56
Obrázek 24 - Srovnání odhadnutých hodnot K_s pro seskupené zrnitostní třídy (1-5) pomocí boxplotů.	60
Obrázek 25 - Korelace mezi odhadnutou a změřenou K_s pro seskupené zrnitostní třídy.....	61
Obrázek 26 - Mean error použitých algoritmů pro seskupené zrnitostní třídy 1-5.	64

