

Možnosti využití spektroskopie pro hodnocení vlastností lesních půd

JOSEF KRATINA, VÁCLAVA MAŤAŠOVSKÁ

Klíčová slova: lesní půda – VNIR spektroskopie – predikce půdních vlastností

SOUHRN

Cílem práce bylo objektivní zhodnocení použitelnosti VNIR spektroskopie (spektroskopie ve viditelné a blízké infračervené části elektromagnetického spektra) pro predikci vlastností lesních půd. Pro každou půdní vlastnost byly hledány nejvhodnější kombinace metod statistického předzpracování (bez předzpracování, continuum removal, 1. a 2. derivace) a zpracování (PLSR – regrese částečných nejmenších čtverců, PCR – regrese hlavních komponent, SVM – support vector machines) v určitých spektrálních pásmech. Jako všeobecně nejúspěšnější se ukázala kombinace metod 1. derivace a SVM v celém VNIR spektrálním pásmu (400–2 500 nm). V některých případech (různé formy hořčíku, manganu, železa či hliníku) se však osvědčily modely jiné. Mezi nejlépe predikovatelné vlastnosti ($R^2 > 0,6$) patří pH (půdní reakce), obsah oxidovatelného uhlíku, obsahy hliníku, železa, křemíku nebo vápníku (ve vyšších koncentracích). Nepříliš vysoká úspěšnost predikce ($R^2 < 0,3$) byla zjištěna u ukazatelů, které nabývají nízkých hodnot (obsah sodíku, manganu nebo dvojmocné komplexy hliníku). Výsledky ukazují, že VNIR spektroskopie je použitelnou metodou pro predikci vlastností lesních půd. Nemůže zcela nahradit klasickou analýzu, ale může ji velice dobře doplňovat, a to především v praxi. Například při půdním mapování umí pomoci zahustit síť údajů a zpřesnit informace lépe než při použití jiných metod prostorového odhadu. Je použitelná v případech, kdy je zapotřebí získat značné množství údajů v krátkém časovém horizontu a s minimálními náklady. Je vhodná pro sledování trendů v čase nebo pro rychlý průzkum určité oblasti.

ÚVOD

Informace o půdních vlastnostech jsou vyžadovány pro různé účely, například pro precizní zemědělství či lesnictví, pro hodnocení kvality půdy, pro její mapování nebo ochranu. Při zkoumání půd je zapotřebí získat velké množství analytických dat. Odebírání a následné analyzování půdních vzorků tradičními metodami je časově náročné a finančně nákladné [1, 2]. Stále častěji se proto využívá nepřímých měření a predikcí půdních vlastností pomocí matematických modelů. Některé studie prokázaly, že spektroskopie ve viditelné (VIS) a blízké infračervené (NIR) části elektromagnetického spektra je vhodnou metodou pro hodnocení půdních vlastností. Doposud uveřejněné modely nejsou univerzální, jsou relevantní jen za specifických podmínek a pro určité skupiny půd. Cílem této studie je posouzení použitelnosti spektroskopie při hodnocení vlastností lesních půd v České republice na základě vztahů spektrálních znaků a půdních vlastností, stanovených tradičními laboratorními metodami. Pro toto hodnocení bylo použito více než 4 500 vzorků odebraných z celých půdních profilů. Pro tvorbu prediktivních modelů byly testovány vhodné kombinace

metod přípravy dat (1. a 2. derivace [3], continuum removal) a statistických technik partial least squares regression (PLSR), support vector machines (SVM) a principal component regression (PCR).

METODIKA

Při zpracování studie bylo použito 4 680 vzorků odebraných z celých půdních profilů pomocí sondýrky nebo z kopaných půdních sond. Část půdních vzorků pochází ze zdrojů Katedry pedologie a ochrany půd (ČZU v Praze), ostatní vzorky byly zapůjčeny z jiných pracovišť. Místa odběrů byla zvolena tak, aby pokryla celé území ČR a aby se do zkoumání zahrnuly různé lesní půdní typy. Odběrová stanoviště byla umístěna v různých nadmořských výškách a v lesích s různou druhovou skladbou. Studie se nezabývala měřením v terénu, byly používány pouze vysušené vzorky upravené na jemnozem I (velikosti frakce < 2 mm) [4]. Tímto byl eliminován vliv vlhkosti půdy, který je pro průběh spektrálních křivek zásadní a značně znesnadňuje terénní použití metody. Na půdních vzorcích byly provedeny vybrané analýzy konvenčními metodami (tab. 1).

Měření spekter bylo prováděno na vzorcích upravených na jemnozem I v Petriho miskách spektrometrem FieldSpec 3 (ASD Inc., USA) s High Intensity Contact Probe (obr. 1). Rozsah spektrometru je 350–2 500 nm.



Obr. 1. Spektrometr FieldSpec 3 (foto: Josef Kratina)

Fig. 1. FieldSpec 3 spectrometer

Tab. 1. Použité metody konvenční analýzy
Tab. 1. Used methods of conventional analysis

Vlastnost	Jednotky	Popis analýzy
pH _{H₂O}	–	Aktivní půdní reakce (H ₂ O) [4]
pH _{CaCl₂}	–	Výměnná půdní reakce ve výluhu CaCl ₂ [5]
pH _{KCl}	–	Výměnná půdní reakce ve výluhu KCl [6]
Cox	%	Obsah oxidovatelného uhlíku modifikovanou Tjurinovou metodou [4]
KVK	mmol.100 g ⁻¹	Kationtová výměnná kapacita podle Bowera [4]
BS	mmol.100 g ⁻¹	Stupeň nasycení sorpčního komplexu bazickými kationty [4]
N	mg.kg ⁻¹	NIR spektroskopické stanovení [7]
P _{M3}	mg.kg ⁻¹	Fosfor extrahovaný roztokem Mehlich III [4]
P _{AR}	mg.kg ⁻¹	Fosfor extrahovaný lučavkou královskou [8]
K _{M3}	mg.kg ⁻¹	Draslík extrahovaný roztokem Mehlich III [4]
K _{AR}	mg.kg ⁻¹	Draslík extrahovaný lučavkou královskou [8]
K _{BaCl₂}	mg.kg ⁻¹	Výměnné kationty (draslík), výluh BaCl ₂ [9]
Ca _{M3}	mg.kg ⁻¹	Vápník extrahovaný roztokem Mehlich III [4]
Ca _{AR}	mg.kg ⁻¹	Vápník extrahovaný lučavkou královskou [8]
Ca _{BaCl₂}	mg.kg ⁻¹	Výměnné kationty (vápník), výluh BaCl ₂ [9]
Mg _{M3}	mg.kg ⁻¹	Hořčík extrahovaný roztokem Mehlich III [4]
Mg _{AR}	mg.kg ⁻¹	Hořčík extrahovaný lučavkou královskou [8]
Mg _{BaCl₂}	mg.kg ⁻¹	Výměnné kationty (hořčík), výluh BaCl ₂ [9]
Na _{AR}	mg.kg ⁻¹	Sodík extrahovaný lučavkou královskou [8]
Na _{BaCl₂}	mg.kg ⁻¹	Výměnné kationty (sodík), výluh BaCl ₂ [9]
Mn _{BaCl₂}	mg.kg ⁻¹	Výměnné kationty (mangan), výluh BaCl ₂ [9]
Mn _{AR}	mg.kg ⁻¹	Mangan extrahovaný lučavkou královskou [8]
Mn _{KCl}	mg.kg ⁻¹	Mangan extrahovaný roztokem KCl [10]
Mn _{ox}	mg.kg ⁻¹	Mangan extrahovaný šťavelanem [11]
Mn _{dit}	mg.kg ⁻¹	Mangan extrahovaný dithioničnanem [12]
Fe _{BaCl₂}	mg.kg ⁻¹	Výměnné kationty (železo), výluh BaCl ₂ [9]
Fe _{AR}	mg.kg ⁻¹	Železo extrahované lučavkou královskou [8]
Fe _{KCl}	mg.kg ⁻¹	Železo extrahované roztokem KCl [10]
Fe _{ox}	mg.kg ⁻¹	Železo extrahované šťavelanem [11]
Fe _{dit}	mg.kg ⁻¹	Železo extrahované dithioničnanem [12]
Al _{BaCl₂}	mg.kg ⁻¹	Výměnné kationty (hliník), výluh BaCl ₂ [9]
Al _{AR}	mg.kg ⁻¹	Hliník extrahovaný lučavkou královskou [8]
Al _{KCl}	mg.kg ⁻¹	Hliník extrahovaný roztokem KCl [10]

Vlastnost	Jednotky	Popis analýzy
Al _{ox}	mg.kg ⁻¹	Hliník extrahovaný šťavelanem [11]
Al _{dit}	mg.kg ⁻¹	Hliník extrahovaný dithioničnanem [12]
Al (X) 1+	mg.kg ⁻¹	Formy hliníku ve výluhu KCl [10]
Al (Y) 2+	mg.kg ⁻¹	Formy hliníku ve výluhu KCl [10]
Al 3+	mg.kg ⁻¹	Formy hliníku ve výluhu KCl [10]
VA	mmol.kg ⁻¹	Výměnná acidita Al+H [6]
Si _{ox}	mg.kg ⁻¹	Křemík extrahovaný šťavelanem [11]
Si _{dit}	mg.kg ⁻¹	Křemík extrahovaný dithioničnanem [12]

Ke stanovení základních statistických popisných charakteristik byl použit program Statistica 12 (StatSoft). K předzpracování spektrálních dat, konkrétně k vyhlazení spektrálních křivek (splice correction), byl používán program ViewSpec Pro 6.0 (ASD Inc.). Program R (R Core Team) byl využit pro úpravu spekter pomocí continuum removal. Pro jejich kalibraci (partial least squares regression – regrese metodou částečných nejmenších čtverců, support vector machines, principal components regression) sloužily programy Unscrambler X 10.3 (CAMO Software) a R (R Core Team).

Byl statisticky hodnocen vztah spektrálních znaků a půdních vlastností, které byly získány tradičními laboratorními metodami. Byla testována vhodnost použití různých metod přípravy dat, jako je 1. a 2. derivace nebo continuum removal (sjednocení průběhu spektrálních křivek, jež jsou získány různými přístroji nebo za různých světelných podmínek). Publikované modely pro predikci půdních charakteristik ze spektrálních znaků byly testovány, upraveny jejich parametry, případně byly tvořeny nové modely pomocí statistických metod PLSR, PCR a SVM. Pro statistické hodnocení byla používána nejen všechna data dohromady, ale rovněž dělená do podsouborů podle oblastí odběrů a podle horizontů, aby bylo možno popsat nevhodnější způsob zadávání dat pro úspěšnou predikci. Dále byl testován vliv použitého spektrálního pásma na úspěšnost předpovědi. Některé vlastnosti jsou lépe predikovatelné při použití celého VNIR spektra, pro jiné je vhodnější použití pouze určitého vybraného spektrálního pásma, které je vybíráno buď pokusně, nebo na základě literatury [13–16].

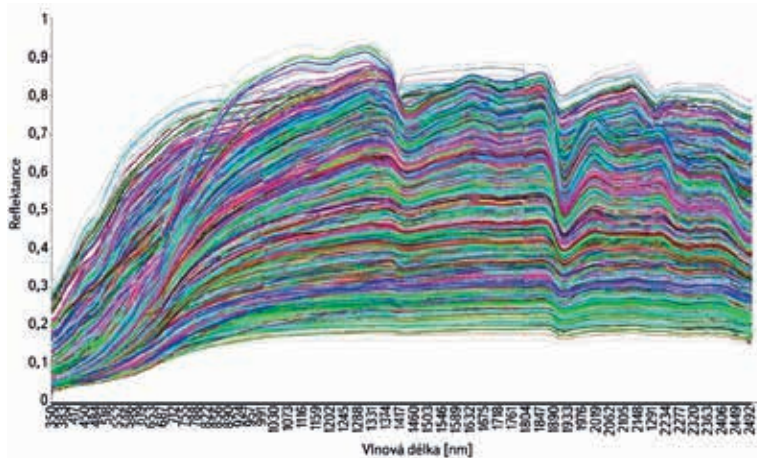
Nové modely byly ověřovány (validovány). Uváděné předpovědi vyjádřené hodnotami R² (koeficient determinace) a RMSE (střední kvadratická chyba) jsou výsledkem cross validace, při níž se soubor dat rozdělí na více podsouborů, jeden (10 % z celku) se vyjme a zbylé se použijí pro kalibraci modelu. Poté je model aplikován na soubor, který byl předtím vyjmut, jsou srovnány hodnoty predikované modelem a změřeny v laboratoři. To se opakuje pro všechny podsoubory. Následně jsou spočítány parametry R² a RMSE. Pro takto vzniklé skupiny byl kalibrován model. Posléze byly nové modely ještě podrobeny externí validaci, při níž byl model aplikován na jiný soubor dat a zjišťována úspěšnost predikce.

VÝSLEDKY A DISKUZE

Úspěšnost metody při predikci půdních vlastností je vysoce závislá na vhodném způsobu zadávání vstupních dat a jejich statistickém předzpracování a vyhodnocování. Cílem práce bylo nalezení nevhodnější kombinace způsobu zadávání, statistického předzpracování a zpracování spektrálních dat pro dosažení nejlepších výsledků předpovědi půdních vlastností.

Celý datový soubor a jeho rozdělení podle horizontů a oblastí

Jako první byl statisticky hodnocen souhrnný datový soubor bez ohledu na příslušnost k oblasti odběru či horizontu. Byla použita spektra bez předzpracování v celém rozsahu 350–2 500 nm. Statistickou metodou byla – v literatuře často doporučovaná – regrese částečných nejmenších čtverců (PLSR). Na obr. 2, který znázorňuje všechna spektra dohromady, vidíme velkou variabilitu jejich průběhu. Ta může být způsobena různými půdními vlastnostmi, např. rozdílným množstvím půdní organické hmoty v minerálních a organických horizontech.



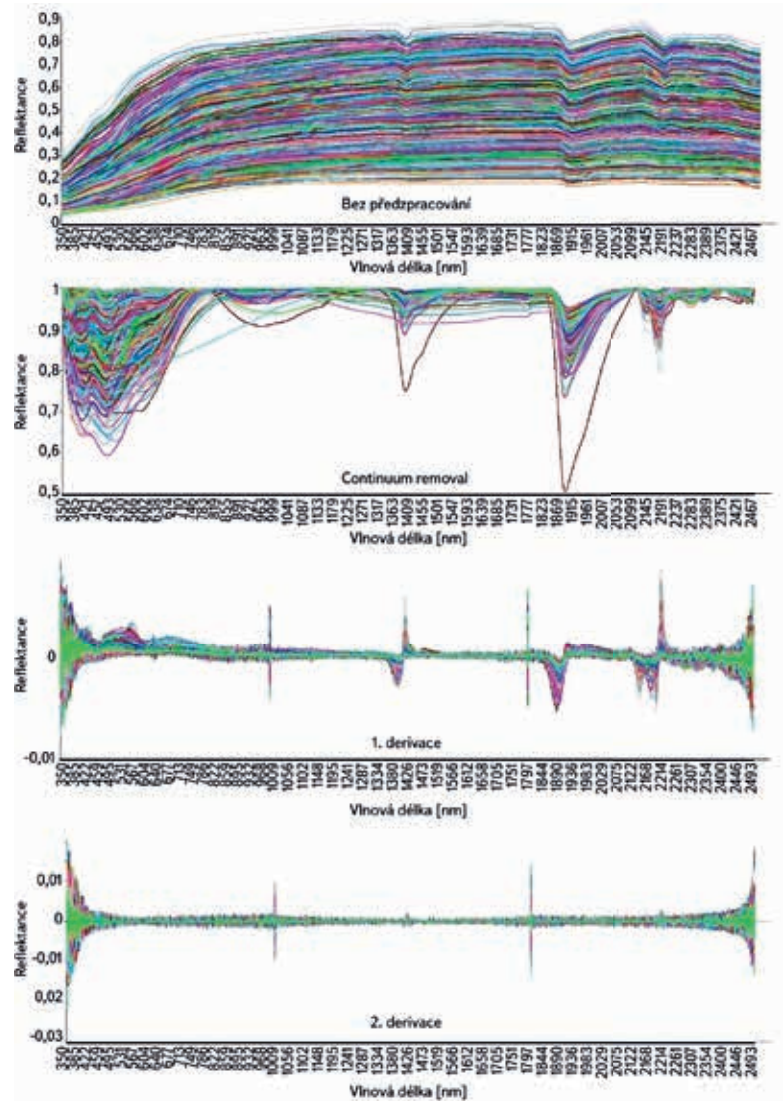
Obr. 2. Znáznornění spekter – souhrnný soubor dat
Fig. 2. Spectra representation – the summary data set

Byly hledány vlastnosti společné pro co největší množství měřených vzorků. Konkrétně šlo o množství oxidovatelného uhlíku – Cox, celkový obsah dusíku a výměnnou půdní reakci – pH_CaCl₂. Výsledky se zdají být velice dobré (Cox – R² = 0,92, obsah dusíku – R² = 0,77, pH_CaCl₂ – R² = 0,51), jejich publikace by však byla správná jen v případě výměnné reakce. Dle rozdělení četností dat jednotlivých vlastností je normální rozdělení pouze v případě výměnné půdní reakce. Výsledky obsahu oxidovatelného uhlíku a dusíku tvoří dva shluky. Ty poskytují při proložení regresní přímkou vysokou hodnotu koeficientu determinace. Výsledky jsou ovšem zkreslené, nelze je správně interpretovat. Predikce výměnné reakce je i přes nižší hodnotu koeficientu determinace úspěšnější. Jelikož pro uplatnění PLSR je normální rozdělení dat podmínkou, bylo třeba soubor podrobit jinému statistickému zpracování.

V tab. 2 jsou znázorněny výsledky predikce výměnné půdní reakce a obsahu dusíku při použití celého souboru i při jeho rozdělení na organické a minerální horizonty. Z výsledků je patrné, že na rozdíl od předchozích výsledků vykazuje celý soubor vyšší hodnotu koeficientu R², avšak spolu s ním roste i střední kvadratická chyba (RMSE), jež by měla v ideálním případě naopak klesat.

Tab. 2. Úspěšnost predikce podle horizontů
Tab. 2. Predictions according to horizons

	R ²	RMSE
pH_CaCl ₂ – celý soubor	0,51	0,46
N [mg.kg ⁻¹] – celý soubor	0,77	0,31
pH_CaCl ₂ – minerální horizonty	0,32	0,44
pH_CaCl ₂ – organické horizonty	0,64	0,44
N [mg.kg ⁻¹] – minerální horizonty	0,44	0,11
N [mg.kg ⁻¹] – organické horizonty	0,37	0,25



Obr. 3. Průběh spektrálních křivek – různé způsoby předzpracování
Fig. 3. The course of spectral curves – various pre-processing methods

Z průběžných výsledků je patrné, že rozdělování datového souboru podle oblastí odběru vzorků nezvyšuje vždy prokazatelně úspěšnost predikce půdních vlastností, a nelze tedy tento způsob přípravy dat jednoznačně doporučit. V otázce dělení dat podle příslušnosti k půdnímu horizontu je závěr komplikovanější. V některých případech se takové rozdělení jeví jako výhodnější, v jiných je úspěšnost předpovědi vyjádřená koeficientem determinace výrazně lepší ve prospěch neděleného souboru. V takové situaci je však třeba sledovat i další veličinu popisující úspěšnost predikce, střední kvadratickou chybu, která by měla na rozdíl od koeficientu determinace klesat, k čemuž však ve výše uvedených případech nedošlo.

Organické a minerální horizonty mají zásadně odlišný charakter a většinou se u nich laboratorně stanovují rozdílné vlastnosti. Pokud jsou některé vlastnosti společně stanoveny pro horizonty celého půdního profilu, pak jsou na nich jasně patrné rozdíly. Data nemají normální rozdělení (je bimodální) a výsledky nelze správně interpretovat. Tento průběžný závěr byl dále ověřen použitím metody SVM, jež není tak zásadně citlivá na rozdělení dat jako regresní (lineární) metody.

Tab. 3. Nejlepší predikce podle metody a pásma – shrnutí
 Tab. 3. The best prediction by method and band – summary

Vlastnost	Horizont (oblast)	Předzpracování	Statistická metoda	Spektrální pásmo [nm]	R ² validační	RMSE validační
pH_H ₂ O	min. hor.	1. derivace	SVM	400–2 500	0,46	0,40
pH_CaCl ₂	min. hor.	1. derivace	SVM	400–2 500	0,45	0,40
pH_CaCl ₂	org. hor.	1. derivace, CR	SVM	400–2 500	0,72	0,39
pH_KCl	A horizonty	1. derivace	SVM	400–2 500	0,55	0,12
Cox	min. hor.	1. derivace	SVM	400–2 500	0,68	1,85
Cox	org. hor.	1. derivace	SVM	400–2 500	0,84	3,71
Cox	A horizonty	1. derivace	SVM	400–2 500	0,48	1,63
KVK	A horizonty	Žádné	PLSR	400–2 500	0,37	14,37
KVK	min. hor.	1. derivace	SVM	400–2 500	0,64	22,52
BS	min. hor.	1. derivace	SVM	400–2 500	0,44	20,65
N	min. hor.	1. derivace	SVM	400–2 500	0,57	0,10
N	org. hor.	1. derivace	SVM	400–2 500	0,62	0,19
P_M3	min. hor.	1. derivace	SVM	400–2 500	0,10	26,68
P_AR	org. hor.	1. derivace	SVM	400–2 500	0,34	243,91
K_M3	min. hor.	1. derivace	SVM	400–2 500	0,31	54,00
K_AR	org. hor.	1. derivace	SVM	400–2 500	0,57	657,13
K_BaCl ₂	min. hor.	1. derivace	SVM	400–2 500	0,51	20,29
K_BaCl ₂	A horizonty	1. derivace	SVM	750–2 500	0,42	0,87
Ca_M3	min. hor.	CR	PLSR	1 100–2 500	0,36	687,14
Ca_AR	org. hor.	1. derivace	SVM	400–2 500	0,76	2 198,70
Ca_BaCl ₂	A horizonty	Žádné	PLSR	1 100–2 500	0,33	0,19
Ca_BaCl ₂	min. hor.	1. derivace	SVM	400–2 500	0,27	13,33
Mg_M3	min. hor.	CR	PLSR	1100–2 500	0,32	146,92
Mg_AR	org. hor.	1. derivace	SVM	400–2 500	0,43	1 303,50
Mg_BaCl ₂	A horizonty	1. derivace	PLSR	400–2 500	0,43	15,18
Mg_BaCl ₂	min. hor.	1. derivace	SVM	400–2 500	0,32	3,00
Na_AR	org. hor.	1. derivace	SVM	400–2 500	0,23	40,40
Na_BaCl ₂	min. hor.	1. derivace	SVM	400–2 500	0,35	0,18
Na_BaCl ₂	A horizonty	1. derivace	SVM	400–2 500	0,16	7,93
Mn_BaCl ₂	min. hor.	1. derivace	SVM	400–2 500	0,28	1,49
Mn_BaCl ₂	A horizonty	1. derivace	SVM	400–2 500	0,51	32,22
Mn_AR	org. hor.	1. derivace	SVM	400–2 500	0,51	1 141,80
Mn_KCl	A horizonty	1. derivace	SVM	400–2 500	0,48	58,28
Mn_ox	A horizonty	CR	PLSR	400–2 500	0,56	107,65

Vlastnost	Horizont (oblast)	Předzpracování	Statistická metoda	Spektrální pásmo [nm]	R ² validační	RMSE validační
Mn_dit	A horizonty	1. derivace	SVM	400–750	0,55	121,04
Fe_BaCl ₂	min. hor.	1. derivace	SVM	400–2 500	0,38	1,38
Fe_BaCl ₂	A horizonty	Žádné	PLSR	400–2 500	0,65	44,59
Fe_AR	org. hor.	1. derivace	SVM	400–2 500	0,51	5 757,50
Fe_KCl	A horizonty	Žádné	PCR	400–2 500	0,67	68,64
Fe_ox	A horizonty	Žádné	PLSR	1 100–2 500	0,68	2 033,40
Fe_dit	A horizonty	Žádné	PLSR	750–2 500	0,69	2 370,50
Al_BaCl ₂	min. hor.	1. derivace	SVM	400–2 500	0,59	14,99
Al_BaCl ₂	A horizonty	Žádné	PLSR	400–800	0,58	120,53
Al_AR	org. hor.	1. derivace	SVM	400–2 500	0,70	2 395,20
Al_KCl	A horizonty	Žádné	PCR	400–2 500	0,62	106,41
Al_ox	A horizonty	Žádné	PLSR	400–2 500	0,63	617,01
Al_dit	A horizonty	1. derivace	SVM	400–2 500	0,45	824,18
Al (X) 1+	A horizonty	Žádné	PCR	600–800	0,63	15,56
Al (Y) 2+	A horizonty	1. derivace	SVM	1 100–2 500	0,44	19,21
Al 3+	A horizonty	Žádné	PCR	400–2 500	0,58	91,29
VA	min. hor.	1. derivace	SVM	400–2 500	0,58	16,58
VA	A horizonty	Žádné	PCR	400–2 500	0,49	16,64
Si_ox	A horizonty	1. derivace	SVM	400–2 500	0,35	120,40
Si_dit	A horizonty	2. derivace	SVM	400–750	0,59	979,57

Vliv použité statistické metody na úspěšnost predikce

Dále byla data podrobena různým kombinacím předzpracování (bez předzpracování, 1. a 2. derivace, continuum removal) a statistických metod (PLSR, PCR, SVM). Příklad změny průběhu spektrálních křivek podle použitého způsobu předzpracování je znázorněn na obr. 3. Takto byly pro každou vlastnost vždy vybrány dvě nejúčinnější kombinace, které byly jednotlivě precizovány.

Viscarra Rossel a Behrens ve své práci [17] určili jako dvě nejúspěšnější metody SVM a PLSR. Výsledky této práce to potvrzují, především ve prospěch SVM, a přinášejí další informace díky kombinaci těchto metod s různými způsoby předzpracování. Ve většině případů byla zjištěna nejvyšší úspěšnost predikce při kombinaci 1. derivace spektrálních dat a metody SVM, dále pak PLSR na datech bez předzpracování a SVM po provedení continuum removal. Další úspěšnou kombinací bylo v některých případech použití 2. derivace a SVM. Naopak regresní metody uplatněné na datech předzpracovaných 2. derivací byly jednoznačně nejméně úspěšné. Regresní metody PLSR a PCR poskytují velmi podobné výsledky, většinou v mírný prospěch PLSR.

Predikce jednotlivých vlastností

Na základě literatury [13, 14, 16] a výše uvedených poznatků byly hledány nejlepší kombinace použitého typu předzpracování, statistické metody a nově i vybraného spektrálního pásma. Pro každou vlastnost byly vybrány nejúspěšnější kombinace metod podle předchozích zjištění, které byly podrobeny dalšímu testování. Pro všechny vlastnosti byla společná úprava spekter ořiznutím pásma 350–400 nm, jež je na rozhraní UV a viditelného záření výrazně rušeno šumem. Jako všeobecně nejlepší se ukazuje kombinace metody předzpracování spektrálních dat pomocí 1. derivace a statistické metody support vector machine při použití celého VNIR spektrálního pásma (400–2 500 nm). Existují ale případy, kdy se jako nejvhodnější ukázaly být jiné kombinace metod a jiná, užší spektrální pásma. Například vápník a hořčík stanovené ve výluhu lučavky královské byly nejlépe predikovatelné metodou PLSR aplikovanou na spektrální data upravená funkcí continuum removal v blízké infračervené části spektra 1 100–2 500 nm. Pro prvky stanovené v oxalátu byla místo SVM ve většině případů nejlepší metoda PLSR (Mg, Mn, Fe, Al). Rozmanitost modelů byla sledována především v případě predikce železa a hliníku, tedy dvou velmi sledovaných prvků v lesních půdách. Značně rozdílná je spektrální detekovatelnost jednotlivých forem hliníku. Nejméně úspěšně predikovatelná dvojmocná forma, která se váže na organickou hmotu, avšak je obsažena v extrémně malých množstvích, se může alternativně predikovat odečtením obsahu

Tab. 4. Nejlepší univerzální statistické modely pro predikci jednotlivých vlastností
 Tab. 4. The best universal statistical models for the prediction of individual properties

Vlastnost	Předzpracování	Statistická metoda	Spektrální pásmo [nm]
pH_H ₂ O	1. derivace	SVM	400–2 500
pH_CaCl ₂	1. derivace	SVM	400–2 500
pH_KCl	1. derivace	SVM	400–2 500
Cox	1. derivace	SVM	400–2 500
KVK	1. derivace	SVM	400–2 500
BS	1. derivace	SVM	400–2 500
N	1. derivace	SVM	400–2 500
P	1. derivace	SVM	400–2 500
K	1. derivace	SVM	400–2 500
Ca	1. derivace	SVM	400–2 500
Mg_M3	continuum removal	PLSR	1 100–2 500
Mg_AR	1. derivace	SVM	400–2 500
Mg_vym (BaCl ₂)	1. derivace	SVM	400–2 500
Na	1. derivace	SVM	400–2 500
Mn_AR, KCl	1. derivace	SVM	400–2 500
Mn_ox	continuum removal	PLSR	400–2 500
Mn_dit	1. derivace	SVM	400–750
Fe_vym (BaCl ₂)	1. derivace	SVM	400–2 500
Fe_AR	1. derivace	SVM	400–2 500
Fe_KCl	1. derivace	SVM	400–2 500
Fe_ox	Bez předzpracování	PLSR	1 100–2 500
Fe_dit	Bez předzpracování	PLSR	750–2 500
Al_vym (BaCl ₂)	1. derivace	SVM	400–2 500
Al_AR	1. derivace	SVM	400–2 500
Al_KCl	Bez předzpracování	PLSR	400–2 500
Al_ox	Bez předzpracování	PLSR	400–2 500
Al_dit	1. derivace	SVM	400–2 500
Al (X) 1+	Bez předzpracování	PCR	600–800
Al (Y) 2+	2. derivace	SVM	1 100–2 500
Al 3+	Bez předzpracování	PLSR	400–2 500
VA	1. derivace	SVM	400–2 500
Si	1. derivace	SVM	400–2 500

jednomocných a trojmocných komplexů od celkového obsahu stanoveného ve společném výluhu, v tomto případě výluhu KCl. V tab. 3 jsou znázorněny nejlepší modely pro predikci jednotlivých půdních vlastností, včetně validačního R² a RMSE. Soubor je rozdělen podle příslušnosti k půdnímu horizontu nebo skupině horizontů.

Testování a úprava nalezených modelů

Dále byly nalezené modely aplikovány na nezávislé soubory dat. Byly porovnávány úspěšnosti predikcí jednotlivých vlastností před a po aplikaci modelů. Data použitá jako ta „před aplikací modelu“ byla podrobena standardnímu statistickému zpracování, tedy metodě PLSR v celém VNIR pásmu na nepředzpracovaných spektrech. Modely, které zlepšily výsledky predikce, byly uznány za vhodné a univerzální. Pokud úspěšnost predikce vzrostla jen velmi málo, zůstala beze změny, nebo dokonce klesla, byly hledány jiné modely na základě dosavadních znalostí.

V tab. 4 jsou znázorněny vítězné modely pro predikci vlastností lesních půd. Kromě úspěšnosti predikce byl kladen důraz i na univerzálnost použitých daných modelů. V případě, že byla určitá vlastnost stanovena různými způsoby, avšak model je pro všechny tyto způsoby společný, je v tabulce označena pouze daná vlastnost. Pokud způsob stanovení měl vliv na spektrální detekovatelnost, a vyžadoval tedy použití jiného modelu, jsou vlastnosti jednotlivě popsány i se způsobem stanovení.

ZÁVĚR

Cílem studie bylo objektivní zhodnocení použitelnosti spektroskopie ve viditelné a blízké infračervené části spektra pro predikci vlastností lesních půd. Tyto půdy se od zemědělských zásadně liší svým vzhledem, vývojem, fyzikálními i chemickými procesy, přítomností organických horizontů atd. Také se u nich obvykle sledují jiné vlastnosti. Bylo zjištěno, že dělení datového souboru podle oblastí odběru vzorku není významným vstupním kritériem, podstatnější je rozložení dat. Z důvodů velké rozdílnosti mezi organickými a minerálními horizonty se na základě výsledků doporučuje zkoumání těchto horizontů zvlášť.

Jelikož bylo k dispozici velké množství dat, bylo možno tato data rozdělit na větší soubor trénovací, na němž byly postupně důkladně trénovány jednotlivé modely, a soubor testovací, na kterém byly tyto modely testovány a na základě výsledků dále případně upravovány. Takto byly pro každou půdní vlastnost nalezeny nejhodnější kombinace metod statistického předzpracování a zpracování v určitých spektrálních pásmech. Jako všeobecně nejúspěšnější se ukazuje kombinace metod 1. derivace a support vector machine v celém VNIR spektrálním pásmu (400–2 500 nm). V některých případech se však osvědčily modely jiné. Mezi nejlépe predikovatelné vlastnosti (R² > 0,6) patří pH (půdní reakce), obsah oxidovatelného uhlíku, obsahy hliníku, železa, křemíku nebo vápníku (ve vyšších koncentracích). Nepříliš vysoká úspěšnost predikce (R² < 0,3) byla zjištěna u ukazatelů, jež nabývají nízkých hodnot (obsah sodíku, manganu nebo dvojmocné komplexy hliníku).

Výsledky ukazují, že VNIR spektroskopie je použitelnou metodou pro predikci vlastností lesních půd. Není schopna zcela nahradit klasickou analýzu, ale může ji doplňovat. Například při půdním mapování pomůže zahustit síť údajů a zpřesnit informace lépe než při použití jiných metod prostorového odhadu. Je použitelná v případech, kdy je zapotřebí získat velké množství údajů v krátkém časovém horizontu a s minimálními náklady. Vhodná je pro sledování trendů v čase nebo pro rychlý průzkum vzorků z určité oblasti.

Poděkování

Autoři děkují Katedře pedologie a ochrany půd České zemědělské univerzity v Praze za poskytnutí technického vybavení a pomoc při zpracování této studie, která byla prováděna v rámci doktorského studia.

Literatura

- [1] KOOISTRA, L., WEHRENS, R., LEUVEN, R. S. E. W., BUYDENS, L. M. C. Possibilities of Visible-Near-Infrared Spectroscopy for the Assessment of Soil Contamination in River Floodplains. *Analytica Chimica Acta*. 2001, 446, s. 97–105.
- [2] JIANG, Q. et al. Estimation of Soil Organic Carbon and Total Nitrogen in Different Soil Layers Using VNIR Spectroscopy: Effects of Spiking on Model Applicability. *Geoderma*. 2017, 293, s. 54–63.
- [3] MARTENS, H., NAES, T. *Multivariate Calibration*, 2nd ed. Chichester, UK: John Wiley and Sons Ltd, 1989. 419 s.
- [4] ZBÍRAL, J. *Analýza půd I*. Brno: Ústřední kontrolní a zkušební ústav zemědělský, 2002. 197 s.
- [5] WHITE, R. E. On the Measurement of Soil pH. *Journal of the Australian Institute of Agricultural Science*. 1969, 35, s. 3–14.
- [6] GILLMAN, G. P., SUMPTER, M. E. Modification of the Compulsive Exchange Method for Measuring Exchange Characteristics of Soils. *Australian Journal of Soil Research*. 1986, 17, s. 61–66.
- [7] SHEPHERD, K. D., WALSH, M. G. Development of Reflectance Spectral Libraries for Characterization of Soil Properties. *Soil Science Society of America Journal*. 2002, 66, s. 988–998.
- [8] ISO 11466:1995. *Soil Quality — Extraction of Trace Elements Soluble in Aqua Regia. ISO/TC 190/SC 3 Chemical and Physical Characterization*. 1995-03.
- [9] BERNHARDT, C. *Particle Size Analysis – Classification and Sedimentation Methods*. London: Springer, 1994. 428 s.
- [10] DRÁBEK, O., BORŮVKA, L., MLÁDKOVÁ, L., KOČÁREK, M. Possible Method of Aluminium Speciation in Forest Soils. *Journal of Inorganic Biochemistry*. 2003, 97, s. 8–15.
- [11] COURCHESNE, F., TURMEL, M. C. Extractable Al, Fe, Mn and Si. In: CARTER, M. R., GREGORICH, E. G. (eds.), *Soil Sampling and Methods of Analysis*. 2nd ed. B. m.: Society of Soil Science. CRC Press, 2008, s. 307–315.
- [12] McKEAGUE, J. A., BRYDON, J. E., MILES, N. M. Differentiation of Forms of Extractable Ion and Aluminium in Soils. *Soil Science Society of America Journal*. 1971, 35, s. 33–38.
- [13] VISCARRA ROSSEL, R. A., WALVOORT, D. J. J., MCBRATNEY, A. B., JANIK, L. J., SKJEMSTAD, J. O. Visible, Near Infrared, Mid Infrared or Combined Diffuse Reflectance Spectroscopy for Simultaneous Assessment of Various Soil Properties. *Geoderma*. 2006, 131, s. 59–75.
- [14] BROWN, D. J., SHEPHERD, K. D., WALSH, M. G., MAYS, M. D., REINSCH, T. G. Global Soil Characterization with VNIR Diffuse Reflectance Spectroscopy. *Geoderma*. 2006, 132, s. 273–290.
- [15] BILGILI, A. V., van ES, H. M., AKBAS, F., DURAK, A., HIVELY, W. D. Visible-Near Infrared Reflectance Spectroscopy for Assessment of Soil Properties in a Semi-Arid Area of Turkey. *Journal of Arid Environments*. 2010, 74, s. 229–238.
- [16] GHOLIZADEH, A. et al. Soil Organic Carbon Estimation Using VNIR–SWIR Spectroscopy. The Effect of Multiple Sensors and Scanning Conditions. *Soil and Tillage Research*. 2021, 211. ISSN 0167-1987.
- [17] VISCARRA ROSSEL, R. A., BEHRENS, T. Using Data Mining to Model and Interpret Soil Diffuse Reflectance Spectra. *Geoderma*. 2010, 158, s. 46–54.

Autoři

Ing. Josef Kratina, Ph.D.

✉ josef.kratina@vuv.cz

ORCID: 0000-0001-6095-586X

Ing. Bc. Václava Maťašovská

✉ vaclava.matasovska@vuv.cz

ORCID: 0000-0001-9229-463X

Výzkumný ústav vodohospodářský T. G. Masaryka, Praha

Příspěvek prošel lektorským řízením.

DOI: 10.46555/VTEI.2021.11.003

POSSIBILITIES OF USING SPECTROSCOPY FOR EVALUATION OF FOREST SOIL PROPERTIES

KRATINA, J.; MAŤAŠOVSKÁ, V.

T. G. Masaryk Water Research Institute, Prague

Keywords: forest soil – VNIR spectroscopy – prediction of soil properties

The aim of this study was an objective assessment of application of VNIR spectroscopy for predicting properties of forest soils. For each soil property were found the most appropriate combination of statistical methods for pre-processing (continuum removal, 1. derivation, 2. derivation) and processing (PLSR, PCR, SVM) of certain spectral bands. As generally successful shows a combination of methods 1. derivation and support vector machine throughout the VNIR spectral range (400–2 500 nm). In some cases, however, they proved to other models. Among the best predictable features include pH, content of oxidizable carbon, aluminum, iron, silicon, or calcium (at higher concentrations). Not very high success rate prediction was found in indicators that take low values (sodium, manganese, aluminum or ferrous complexes). The results show that VNIR spectroscopy method is applicable for predicting properties of forest soils. It can not completely replace traditional analysis, but it can very well complement, especially in practice. For example, when the soil mapping can help thicken network data and refine the information better than other methods of spatial estimation. It is applicable in cases where it is required large amounts of data in a short timeframe and at minimal cost. It is suitable for monitoring trends over time, or for a quick survey of an area.