



19th Iranian Soil Science Congress
2-4 December, 2025



نوزدهمین کنگره علوم خاک ایران
۱۱ تا ۱۳ آذرماه ۱۴۰۴



۰۴۲۵۰-۳۲۰۳۱

مدیریت جامع نگر و هوشمند خاک و آب

Holistic and Smart Soil and Water Management

دانشکده کشاورزی و منابع طبیعی دانشگاه تهران

College of Agriculture & Natural Resources, University of Tehran



برآورد پارامترهای منحنی نگهداری آب خاک در مدل ون گنوختن با استفاده از توابع انتقالی خاکی و طیفی

چیمین مهدی زاده^۱، حسین بیات^{۲*}، مسعود داوری^۳

۱- دانشجوی دکتری گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران

(Chimanmehdizadeh@gmail.com)

۲- * استاد گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران

(h.bayat@basu.ac.ir)

۳- دانشیار گروه علوم و مهندسی خاک دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران

(m.davari@uok.ac.ir)

چکیده

استفاده از توابع انتقالی خاک (PTFs) و تحلیل رفتار طیفی خاک برای برآورد ویژگی‌های فیزیکی، شیمیایی و هیدرولیکی مورد توجه قرار گرفته است. این پژوهش با هدف برآورد پارامترهای منحنی نگهداری آب خاک (SWRC) شامل رطوبت حجمی باقی مانده (θ_r) و اشباع (θ_s)، پارامتر α و n با بهره‌گیری از بازتاب طیفی و ویژگی‌های خاک با استفاده از جنگل تصادفی (RF) انجام شد. در این راستا، ۱۳۰ نمونه خاک از پنج استان ایران جمع‌آوری، ویژگی‌های آنها اندازه‌گیری و بازتاب طیفی نمونه‌ها ثبت گردید. از روش‌های پیش‌پردازش متعددی نظیر تصحیح پخشیده چندگانه (MSC)، مشتق اول و دوم همراه با فیلتر ساویتزکی-گلای (FD-SG, FD-SG2) و متغیر نرمال استاندارد (SNV) استفاده شد. نتایج نشان داد MSC بهترین عملکرد را در تخمین θ_s و پارامتر n داشت و معیار اطلاعات آکایک را به طور قابل توجهی کاهش داد. برای پارامتر α ، روش بدون پیش‌پردازش (NP) بهترین نتیجه را ارائه داد. تحلیل مؤلفه‌های اصلی، طول‌موج‌های ۸۰۴، ۱۰۰۰، ۱۳۳۷ و ۲۱۳۷ نانومتر در ناحیه Vis-NIR را به عنوان کلیدی‌ترین طول موج‌ها در توصیف SWRC شناسایی کرد. این مطالعه نشان داد که ترکیب داده‌های طیفی و خاکی همراه با الگوریتم جنگل تصادفی، رویکردی دقیق و مقرون به صرفه برای برآورد پارامترهای SWRC فراهم می‌کند.

کلید واژه: بازتاب طیفی Vis-NIR، توابع انتقالی، جنگل تصادفی، منحنی نگهداری آب خاک.

مقدمه

منحنی نگه‌داری آب خاک (SWRC)^۱ یکی از ویژگی‌های اساسی هیدرولیکی خاک است که برای توصیف حرکت آب و انتقال املاح در خاک به کار می‌رود (لی و همکاران، ۲۰۲۳). این منحنی رابطه کمی بین مکش خاک و محتوای آب خاک را بیان می‌کند (داسیلوا و همکاران، ۲۰۲۱) و یکی از پارامترهای کلیدی برای مطالعه حرکت و نگهداشت آب در خاک محسوب می‌شود (لی و همکاران، ۲۰۲۰). اندازه‌گیری مستقیم SWRC در آزمایشگاه پرهزینه و زمان‌بر است (بوتولا و همکاران، ۲۰۱۴). همچنین SWRC دارای تغییرپذیری زمانی و مکانی قابل توجهی نیز است (میناسنی و همکاران، ۱۹۹۹). بنابراین به کار بردن روش‌های غیرمستقیم مانند توابع انتقالی (PTFs)^۲ به منظور تخمین پارامترهای معادله ون‌گنوختن برای منحنی نگه‌داری آب خاک ضروری است.

پیشرفت‌های اخیر در حوزه داده کاوی و یادگیری ماشین، به ویژه جنگل تصادفی (RF)^۳، رگرسیون خطی چندگانه و شبکه عصبی مصنوعی، امکان توسعه PTFs با دقت بالاتر را فراهم کرده است (راستگو و همکاران، ۲۰۲۰، بیات و همکاران، ۲۰۱۳). مطالعات متعددی نشان داده‌اند که ترکیب داده‌های خاکی با الگوریتم‌های یادگیری ماشین می‌تواند دقت تخمین پارامترهای SWRC را به طور قابل توجهی بهبود بخشد. در طول چند دهه گذشته، طیف سنجی مرئی-مادون قرمز نزدیک (Vis-NIR)^۴ به عنوان روش غیرمخرب، سریع و اقتصادی بدون ضرر و تخریب برای محیط زیست برای برآورد ویژگی‌های فیزیکی و شیمیایی خاک، از جمله ماده آلی، رطوبت سطحی، توزیع اندازه ذرات و ظرفیت نگهداشت آب، مورد توجه قرار گرفته است (سیلوا و همکاران، ۲۰۲۰، بابائیان و همکاران، ۲۰۱۵). با وجود اثبات کارایی این روش در پیش بینی ویژگی‌های پایه خاک (داوری و همکاران، ۲۰۲۲، پینیرو و همکاران، ۲۰۱۷)، پژوهش‌های محدودی به کاربرد داده‌های طیفی در محدوده Vis-NIR در برآورد پارامترهای SWRC پرداخته‌اند. از طرف دیگر عملکرد بهتر جنگل تصادفی در تخمین برخی از ویژگی‌های هیدرولیکی خاک در مقایسه با سه روش رگرسیون خطی چندگانه، رگرسیون خطی ساده و الگوریتم کویبست توسط برخی محققین مانند بنیچ و همکاران (۲۰۲۳) نشان داده شده است. بر اساس مطالعات صورت گرفته در خصوص برآورد پارامترهای SWRC در معادله ون‌گنوختن با استفاده از داده‌های طیفی و خاکی از طریق توابع انتقالی خاک با روش جنگل تصادفی مطالعات اندکی صورت گرفته است. بنابراین هدف از انجام این تحقیق برآورد پارامترهای SWRC در معادله ون‌گنوختن (and $\theta_r, \theta_s, \alpha_n$) با استفاده از داده‌های طیفی و خاکی از طریق توابع انتقالی خاک با روش جنگل تصادفی بود.

مواد و روش‌ها

در این مطالعه ۱۳۰ نمونه خاک دست خورده و دست نخورده از لایه‌های سطحی خاک (۲۰-۰ سانتی‌متر) با استفاده از سیلندرهای استیل از اراضی ۵ استان همدان، کرمانشاه، آذربایجان غربی، آذربایجان شرقی و مازندران جمع‌آوری شد. نمونه‌ها پس از انتقال به آزمایشگاه و هوا خشک شدن، از الک ۲ میلی‌متری عبور داده شده و برای انجام آزمایش‌های فیزیکی، شیمیایی، منحنی نگه‌داری آب خاک و طیف سنجی آماده شدند. بافت خاک با روش هیدرومتر اندازه‌گیری شد (گی و ار، ۲۰۰۲). جرم مخصوص ظاهری خاک از نمونه‌های دست نخورده به روش استوانه‌های نمونه‌برداری تعیین گردید (بلاک و هارتگه، ۱۹۸۶). برای تعیین رطوبت خاک در مکش‌های ۱، ۲، ۴، ۶ کیلو پاسکال از دستگاه جعبه شن و در مکش‌های ۱۰، ۳۳، ۱۰۰، ۴۰۰ و ۱۵۰۰ کیلو پاسکال از دستگاه صفحات فشاری استفاده شد. ضرایب مدل ون‌گنوختن- معلم توسط نرم افزار RETC به دست آمد. از مدل ون‌گنوختن^۵ (۱۹۸۰) با فرض معلم برای برازش بر داده‌های SWRC استفاده شد.

بازتاب طیفی نمونه‌ها با استفاده از دستگاه اسپکترومتر زمینی (FieldSpec@3, ASD, FR, USA)، در دامنه Vis-NIR (۳۵۰ تا ۲۵۰۰ نانومتر) در گروه خاکشناسی دانشگاه تربیت مدرس اندازه‌گیری شد. نمونه‌های خاک هوا خشک شده و از الک ۲ میلی‌متری عبور داده شدند، ۱۰۰ گرم از هر نمونه در پتری دیش قرار داده شد و منحنی‌های طیفی آنها اندازه‌گیری شد (ویسکارا روسل، ۲۰۰۸). به منظور افزایش دقت، بازتاب طیفی هر نمونه ۵ بار در موقعیت‌های مختلف قرائت شد. داده‌های

¹ Soil water retention curve

² Pedotransfer functions

³ Random forest

⁴ Visible -near-infrared spectroscopy

⁵ Van Genuchten

طیفی پس از ثبت توسط نرم افزار View Spec Pro 6.0.15، میانگین‌گیری و سپس این منحنی‌ها به شکل کتابخانه‌های طیفی در قالب فایل متنی ذخیره شدند و جهت اعمال روش‌های مختلف پیش پردازش تصحیح پخشیده چندگانه (MSC)⁶، مشتق اول به همراه فیلتر ساویتزگی و گلائی (FD-SG1)⁷، مشتق دوم به همراه فیلتر ساویتزگی و گلائی (FD-SG2)⁸ و متغیر نرمال استاندارد (SNV)⁹، به نرم افزار Parles v3.1 منتقل شدند (ویسکارا روسل، ۲۰۰۸). بخش ابتدایی و انتهای طیف‌ها (محدوده بین ۴۰۰-۳۵۰ و ۲۵۰۰-۲۴۵۱) به دلیل آشفتگی‌های زیاد طیفی از تمامی طیف‌ها حذف شده و در مدل‌سازی نیز مورد استفاده قرار نگرفت (چتر نوری و همکاران، ۲۰۲۳).

قبل از توسعه PTFs پیش پردازش داده‌ها که شامل حذف داده‌های پرت، آزمون نرمال بودن (آزمون کولموگروف-اسمیرنوف)، تبدیل داده‌های غیرنرمال و استانداردسازی (میانگین صفر و واریانس ۱) بر روی داده‌ها اعمال شد. همبستگی بین متغیرهای ورودی و خروجی با نرم افزار spss16 بررسی شد. برای کاهش ابعاد داده‌های طیفی و استخراج مهم‌ترین ویژگی‌ها، از روش تجزیه به مولفه‌های اصلی (PCA)¹⁰ استفاده شد (جولیف و کادیم، ۲۰۱۶).

برای توسعه PTFs، ۶ PTF برای پیش بینی پارامترهای SWRC توسعه یافتند. داده‌ها به دو مجموعه آموزش (۹۰ نمونه) و تست (۴۰ نمونه) تقسیم شدند، و توابع براساس الگوریتم جنگل تصادفی، در نرم افزار STATISTICA ساخته و اعتبار سنجی شدند. PTFs در دو گام توسعه یافتند: در گام اول PTF1 (شن، رس و جرم مخصوص ظاهری) به عنوان متغیرهای ورودی استفاده شد و مدل پایه برای تخمین پارامترهای SWRC ایجاد شد. در گام دوم PTF2 تا PTF6 به ترتیب داده‌های طیفی خام (NP)¹¹، روش‌های پیش پردازش MSC، FD-SG1، FD-SG2 و SNV، به عنوان متغیرهای ورودی جدید به متغیرهای ورودی PTF1 اضافه شدند. از بین روش‌های پیش پردازش که از PTF2 تا PTF6 می‌باشد. هر کدام که نسبت به روش بدون پیش پردازش موجب منفی‌تر شدن معیار اطلاعات آکایک (AIC)¹² شدند با PTF1 مقایسه شدند. در نهایت برای بررسی دقت و قابلیت اطمینان توابع انتقالی توسعه یافته از آماره‌های معیار اطلاعات آکایک (AIC)، مجذور میانگین مربعات خطا (RMSE)¹³، ضریب تعیین (R²)¹⁴ و میانگین خطا (ME)¹⁵ استفاده شد.

نتایج و بحث

تخمین رطوبت حجمی باقی مانده، اشباع و پارامتر n با استفاده از داده‌های طیفی

با مقایسه بین PTFهای مرحله دوم (PTF2-PTF6)، در مرحله آموزش PTF3 (MSC) و در مرحله تست (تنها پارامتر n) PTF2 (NP) دارای AIC منفی‌تر نسبت به مدل پایه و توابع دیگر بودند (جدول ۱). ولی هیچ کدام از توابع ایجاد شده در مرحله دوم بر اساس معیار AIC در برآورد رطوبت حجمی باقی مانده بهبود معنی‌داری ایجاد نکردند (نتایج گزارش نشده است) (جدول ۱). از طرفی نتایج تحلیل PCA نشان داد که اولین مولفه اصلی داده‌های طیفی پردازش شده با روش تصحیح پخشیده چندگانه PCMSC در محدوده طول موج‌های ۸۰۴، ۱۰۰۰، ۱۳۳۷ و ۲۱۳۷ نانومتر دارای ارتباط با متغیر خروجی بود (شکل ۱). تخمین‌گرهای استفاده شده در توابع ایجاد شده در مرحله دوم در واقع خروجی روش‌های بدون پیش پردازش و پیش پردازش داده‌های طیفی هستند، که علاوه بر متغیرهای ورودی PTF1 به عنوان تخمین‌گر برای تخمین رطوبت حجمی اشباع و پارامتر n استفاده شدند.

جدول ۱- نتایج تخمین رطوبت حجمی اشباع و پارامتر n از طریق RF با استفاده از متغیرهای ورودی مختلف

⁶ Multiplicative signal correction

⁷ First derivative with savitzi gavlay filter

⁸ The second derivative with a savitzi-golay filter

⁹ Standard normal variate

¹⁰ Principal component analysis

¹¹ No-preprocessing

¹² Akaike Information Criterion

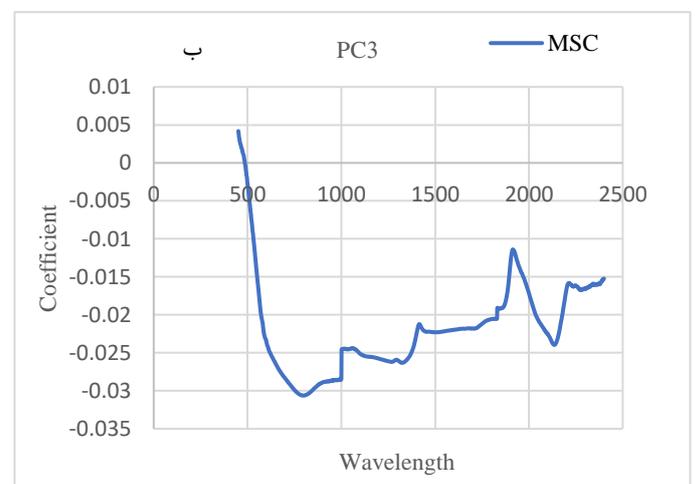
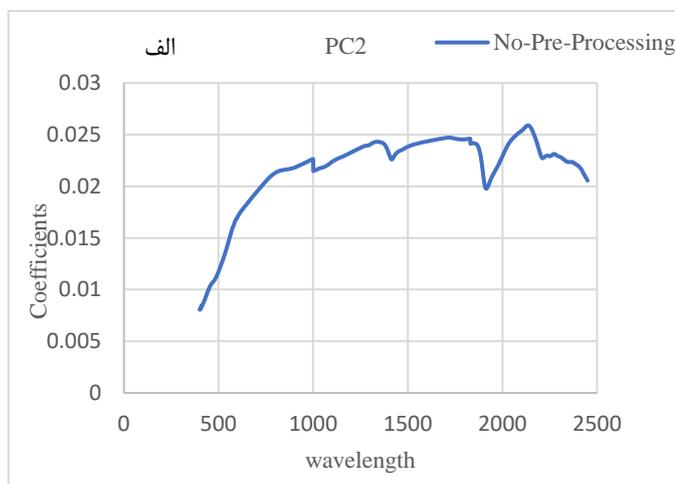
¹³ Root mean square error

¹⁴ Coefficient of determination

¹⁵ Mean error

CM	Test				CM	Train				
	AIC	R ²	RMSE	ME		AIC	R ²	RMSE	ME	
θ_s										
	-۲۲۷	۰/۲۹	۰/۰۶	-۰/۰۰۲		-۴۹۴/۶	۰/۳۶	۰/۰۶۵	۰/۰۰۰۵	PTF1
1	-۲۳۲/۸	۰/۳۱	۰/۰۵۸	-۰/۰۰۵	1	-۴۹۷/۳	۰/۴۰	۰/۰۶۴	۰/۰۰۰۴	PTF2(NP)
1	-۲۱۹/۹	۰/۲۹	۰/۰۷۱	-۰/۰۰۱	1	-۵۲۷/۸*	۰/۴۵	۰/۰۵۵	۰	PTF3 (MSC)
1	-۲۱۹/۷	۰/۲۹	۰/۰۷۱	-۰/۰۰۲	1	-۵۲۵/۵*	۰/۴۴	۰/۰۵۷	۰/۰۰۰۲	PTF4 (FD-SG1)
1	-۲۱۸/۷	۰/۲۵	۰/۰۷۲	-۰/۰۰۵	1	-۵۲۲/۳*	۰/۴۲	۰/۰۵۸	۰/۰۰۰۴	PTF5 (FD-SG2)
1	-۲۱۸/۹	۰/۲۶	۰/۰۷۱	-۰/۰۰۵	1	-۵۲۳/۸*	۰/۴۳	۰/۰۵۷	۰/۰۰۰۱	PTF6 (SNV)
n										
	-۴/۸*	۰/۳۱	۰/۹۲	۰/۲۲۴		-۵۱/۶*	۰/۴۹	۰/۷۴	-۰/۰۰۰۵	PTF1
1	-۲/۵	۰/۳۰	۰/۹۴	۰/۱۵۴	1	-۶۰/۳*	۰/۵۵	۰/۷۰	۰/۰۰۵	PTF2 (NP)
1	-۱۱*†	۰/۲۵	۰/۸۸	۰/۰۷۱	1	-۷۴/۲*†	۰/۴۹	۰/۶۶	-۰/۰۰۱	PTF3 (MSC)
1	-۱/۲	۰/۳۶	۰/۹۶	۰/۱۲۳	1	-۶۲/۴*	۰/۵۴	۰/۷۰	۰/۰۱۳	PTF4 (FD-SG1)
1	-۴/۷*	۰/۳۵	۰/۹۲	۰/۲۳۲	1	-۵۳/۱	۰/۵۱	۰/۷۳	۰/۰۰۵	PTF5 (FD-SG2)
1	-۳/۳	۰/۳۳	۰/۹۴	۰/۱۵۲	1	-۵۶/۱*	۰/۵۲	۰/۷۲	۰/۰۰۰۶	PTF6 (SNV)

* نشان دهنده تفاوت معنی دار بین هر تابع با تابع مقایسه شونده آن می باشد. † نشان دهنده تفاوت معنی دار بین هر کدام از توابع ۲ تا ۶ (روش های پیش پردازش) با PTF2 (روش بدون پیش پردازش) می باشد. θ_r : رطوبت حجمی باقی مانده، θ_s : رطوبت حجمی اشباع، n : پارامتر شکل، وابسته به توزیع اندازه منافذ، ME: میانگین خطا، RMSE: مجذور میانگین مربعات خطا، R^2 : ضریب تعیین، AIC: معیار اطلاعات آکاییک، CM: مدل مقایسه شونده، PTF2 (NP): داده های بدون پیش پردازش، PTF3 (MSC): تصحیح پراکندگی چندگانه، PTF4 (FD-SG1): مشتق اول به همراه فیلتر ساویتزکی و گلائی، PTF5 (FD-SG2): مشتق دوم به همراه فیلتر ساویتزکی گلائی، PTF6 (SNV): متغیر نرمال استاندارد و RF: جنگل تصادفی.



شکل ۱- ضرایب مولفه های اصلی در مقابل طول موج برای PC2: اولین مولفه اصلی داده های طیفی پردازش شده با روش بدون پیش پردازش (PCNP) و PC3: اولین مولفه اصلی داده های طیفی پردازش شده با روش MSC (PCMSC)

PTF3: در این تابع از اولین مولفه های اصلی داده های طیفی پردازش شده با روش MSC (PCMSC) به همراه تخمین گرهای PTF1 در برآورد رطوبت حجمی اشباع و پارامتر n بهره گیری شد. MSC یکی از تکنیک های نرمال سازی است، به این صورت که این تکنیک با تاثیرگذاری بر منحنی های طیفی در تخمین رطوبت حجمی اشباع و پارامتر n می تواند اثرات تغییرات شدت کلی یا تصادفی را حذف کند (استوری و هلمای، ۲۰۱۹). در واقع هدف اصلی MSC حذف غیرخطی های ناشی از پراکندگی ذرات در نمونه ها است (رینان و همکاران، ۲۰۰۹). در PCMSC بیشترین ضرایب با مقادیر بازتاب طیفی در محدوده NIR در

طول موج ۸۰۴، ۱۰۰۰، ۱۳۳۷ و ۲۱۳۷ نانومتر بود (شکل ۱-ب). در مرحله تست PTF2، یعنی روش بدون پیش پردازش احتمالاً به این دلیل باشد که ویژگی‌های طیفی خاک‌های مورد مطالعه به گونه‌ای است که نویز پراکندگی و اثرات سطحی کم بوده و پیش پردازش‌های استاندارد نتوانسته‌اند مزیت قابل توجهی ایجاد کنند که موجب معنی‌داری این روش نسبت به روش‌های پیش پردازش شده باشد.

رطوبت حجمی اشباع مستقیماً به تخلخل کل خاک بستگی دارد ($r=0.304^{**}$). نواحی اصلی نزدیک به ۱۳۰۰ تا ۱۴۰۰ نانومتر از نظر طیفی برای کانی‌های رسی خاک مانند اسمکتیت، ایلیت و کائولینیت فعال هستند (کندل و همکاران، ۲۰۱۳)، که در این پژوهش با طول موج ۱۳۳۷ نانومتر مطابقت دارد. محدوده طیفی ۲۱۰۰ تا ۲۳۰۰ نانومتر تحت تاثیر تغییرات محتوای رس قرار می‌گیرد که به کشش O-H در شبکه رسی مرتبط است (کاستالدی و همکاران، ۲۰۱۵). رطوبت حجمی به طور قابل توجهی تحت تاثیر توزیع اندازه منافذ خاک بافت و ساختار خاک قرار می‌گیرد (آنیدی کچی و همکاران، ۲۰۲۲) که این عوامل به شدت تحت تاثیر نوع و مقدار رس خاک هستند. علاوه بر این، ماده آلی و رس بر بازتاب طیفی در ناحیه مادون قرمز و اکسیدهای آهن بر بازتاب در ناحیه مرئی تاثیرگذار می‌باشند (بابائیان و همکاران، ۲۰۱۵)، بنابراین احتمالاً حضور مواد آلی و کانی‌های رس خاک سبب خاکدانه‌سازی و افزایش حجم منافذ خاک و در نتیجه افزایش ظرفیت نگهداری آب خاک می‌شود (پیتولو و همکاران، ۲۰۱۸). از طرف دیگر، مواد آلی و نوع و مقدار کانی‌های رس خاک جذب پرتو در ناحیه مادون قرمز نزدیک را افزایش می‌دهند. بنابراین یک همبستگی قوی بین مواد آلی و نوع و مقدار کانی‌های رس خاک با جذب پرتو و ظرفیت نگهداری آب خاک وجود دارد. به همین علت استفاده از داده‌های طیفی موجب بهبود تخمین θ_s و پارامتر n گردید. پارامتر n به توزیع اندازه منافذ و بافت خاک بستگی دارد (مرر و همکاران، ۲۰۲۱)، که از طریق توزیع اندازه ذرات به طور غیرمستقیم می‌تواند بازتاب طیفی را کاهش دهد. این پارامتر با رس همبستگی منفی (-0.44^{**}) و با شن همبستگی مثبت (0.33^{**}) دارد. لذا با افزایش منافذ ریز خاک مقدار آن کاهش می‌یابد، که این امر به معنای کاهش شیب منحنی رطوبتی خاک و کاهش از دست دادن آب خاک است. رطوبت خاک موجب کاهش کلی بازتاب طیفی می‌شود، زیرا آب و رطوبت امواج الکترومغناطیسی را جذب می‌کنند (باورس و هانگس، ۱۹۶۵). بنابراین حضور کانی‌های رسی با افزایش نگهداشت آب در خاک، بازتاب طیفی خاک را کاهش می‌دهد. بابائیان و همکاران (۲۰۱۵) یک رابطه همبستگی بین رطوبت حجمی اشباع با مقادیر بازتاب در طول موج‌های ۲۱۳۷ نانومتر و ۲۲۴۲ نانومتر را گزارش کردند. جذب نزدیک به ۱۰۶۷-۱۰۴۷ نانومتر که در این پژوهش با طول موج ۱۰۰۰ نانومتر مطابقت دارد، احتمالاً به حضور گروه‌های آمین و هیدروکسیل مرتبط است (داوری و همکاران، ۲۰۲۲). بالاترین ضریب در PCMSC برای تخمین رطوبت حجمی اشباع و پارامتر n در طول موج ۸۰۴ نانومتر مشاهده شد که احتمالاً به محتوای رس مربوط می‌شود. بازتاب طیفی خاک، در کل طیف‌ها، عمدتاً تحت تاثیر بافت، محتوای کربن آلی و pH است (بابائیان و همکاران، ۲۰۱۵). پارامتر n نیز به توزیع اندازه منافذ و بافت خاک بستگی دارد و از طریق توزیع اندازه ذرات به طور غیرمستقیم می‌تواند بازتاب طیفی را کاهش دهد (مرر و همکاران، ۲۰۲۱). چتر نوری و همکاران (۲۰۲۳) بیان کردند که محدوده‌های طیفی ۵۵۰-۵۸۰، ۷۸۰-۸۸۰ و ۱۱۰۰ سهم قابل توجهی در مدل‌سازی محتوای رس را دارند. همچنین، استنبرگ (۲۰۱۰)، حداقل بازتابش در طول موج‌های نزدیک به ۸۸۰، ۶۲۰ و ۵۱۰ نانومتر را به حضور هماتیت نسبت دادند.

تخمین پارامتر α با استفاده از داده‌های طیف

با مقایسه بین PTF‌های مرحله دوم (PTF2-PTF6)، در مرحله آموزش و تست (NP) PTF2 دارای AIC منفی‌تر نسبت به مدل پایه و توابع دیگر بودند (جدول ۲). تحلیل نتایج PCA نشان داد که روش بدون پیش پردازش (PCNP) در محدوده طول موج‌های ۱۴۱۴، ۱۹۱۲ و ۲۱۵۰ نانومتر، دارای ارتباط معنی‌دار با پارامتر α بودند (شکل ۱-ب).

جدول ۲- نتایج تخمین پارامتر α از طریق RF با استفاده از متغیرهای ورودی مختلف

CM	Test				CM	Train				α
	AIC	R ²	RMSE	ME		AIC	R ²	RMSE	ME	

	-۴۵/۶	۰/۰۷	۰/۵۵	-۰/۰۹۶		-۱۵۲/۴	۰/۴۲	۰/۴۲	۰/۰۰۴	PTF1
1	-۵۰/۱*	۰/۱۵	۰/۵۲	-۰/۰۶۸	1	-۱۶۲/۷*	۰/۵۰	۰/۴۰	-۰/۰۰۲	PTF2(NP)
1	-۴۵/۸	۰/۱۱	۰/۵۵	-۰/۰۹۴	1	-۱۶۱/۲*	۰/۴۸	۰/۴۰	۰	PTF3 (MSC)
1	-۴۶	۰/۱۲	۰/۵۵	-۰/۰۲۷	1	-۱۶۱/۶*	۰/۵۰	۰/۴۰	۰/۰۰۰۷	PTF4 (FD-SG1)
1	-۴۷/۲*	۰/۱۴	۰/۵۴	-۰/۰۸۷	1	-۱۵۸/۱	۰/۴۵	۰/۴۱	۰/۰۰۰۸	PTF5 (FD-SG2)
1	-۴۷/۹*	۰/۱۴	۰/۵۳	-۰/۱۳۶	1	-۱۵۹/۴	۰/۵۱	۰/۴۰	-۰/۰۰۰۸	PTF6 (SNV)

* نشان دهنده تفاوت معنی‌دار بین هر تابع با تابع مقایسه شونده آن می‌باشد. † نشان دهنده تفاوت معنی‌دار بین هر کدام از توابع ۸ تا ۱۱ (روش‌های پیش پردازش) با تابع ۲ (روش بدون پیش پردازش) می‌باشد. α : معکوس مکش ورود هوا، ME: میانگین خطا، RMSE: مجذور میانگین مربعات خطا، R^2 : ضریب تعیین، AIC: معیار اطلاعات آکاییک، CM: مدل مقایسه شونده، PTF2 (NP): داده‌های بدون پیش پردازش، PTF3 (MSC): تصحیح پراکندگی چندگانه، PTF4 (FD-SG1): مشتق اول به همراه فیلتر ساویتزی و گلائی، PTF5 (FD-SG2): مشتق دوم به همراه فیلتر ساویتزی گلائی، PTF6 (SNV): متغیر نرمال استاندارد و RF: جنگل تصادفی.

PTF2: در این تابع از اولین مولفه اصلی داده‌های طیفی بدون پیش پردازش (PCNP) به همراه تخمین‌گرهای PTF1 در برآورد پارامتر α بهره‌گیری شد. داده‌های طیفی خام در این تابع موجب کاهش معیار AIC نسبت به PTF1 شدند. در PTF2 با توجه به اینکه قبل از روش‌های پیش پردازش ابتدا و انتهای منحنی‌های طیفی خام به علت نویز حذف شدند و بعد وارد مدلسازی شدند و با توجه به اینکه معیار اطلاعات آکاییک در داده‌های طیفی خام با داده‌های پیش پردازش شده تفاوت چندانی نداشت احتمالاً مقدار نویز در منحنی‌های داده‌های خام به حدی نبوده است که موجب بهبود دقت تخمین پارامتر α نسبت به PTF1 شده است.

بزرگترین ضرایب PC اول در روش بدون پیش پردازش بازتاب طیفی در ناحیه مادون قرمز نزدیک در طول موج ۲۱۵۰ نانومتر و با دو مشخصه جذب در طول موج‌های ۱۴۱۴ و ۱۹۱۸ نانومتر بود (شکل ۱). پیک‌های طیفی طول موج ۲۴۵۰-۲۱۵۰ نانومتر را می‌توان به گروه‌های عاملی OH شبکه رسی نسبت داد (ویسکارا روسل و همکاران، ۲۰۰۶). یو و همکاران (۲۰۲۳) گزارش کردند که طول موج نزدیک به ۲۰۰۰ نانومتر را می‌توان به مولکول‌های آلی CH_2 ، CH_3 ، NH_3 ، پیوندهای Si-OH و پیوندهای کاتیونی OH در کانی‌هایی مانند کائولینیت و مونت موریلونیت نسبت داد. با توجه به اینکه پارامتر α نشان دهنده معکوس مکش ورود هواست و در نتیجه وابسته به ساختار منفذی دینامیک خاک است (مرر و همکاران، ۲۰۲۰) و با رس و روش بدون پیش پردازش همبستگی مثبت و معنی‌داری ($r=0.308^{**}$ ، $r=0.22^{**}$) دارد (جدول ۲). این ارتباط می‌تواند ناشی از بازتاب دهی طیفی مواد رسی باشد که به طور غیرمستقیم بر ساختار منافذ خاک اثر می‌گذارند. هم‌چنین از آنجایی که ویژگی‌های طیفی خاک در نواحی خاصی می‌توانند با مقدار رس در ارتباط باشند، بنابراین ذرات رس با تشکیل خاکدانه‌ها و بهبود ساختمان خاک (سیمانسی و همکاران، ۲۰۱۹) موجب تشکیل منافذ درشت شده که این افزایش اندازه منافذ سبب کاهش مکش ورود هوا و در نتیجه افزایش پارامتر α شده است. رس هم با ساختار منافذ و در نتیجه با پارامتر α در ارتباط است. بنابراین رابطه مشاهده شده بین داده‌های طیفی و پارامتر α از نظر فیزیکی توجیه پذیر است. هم‌چنین ذرات رس به علت سطح ویژه زیاد خاصیت نگه‌داشت آب در خاک را دارند. از طرفی ذرات رس عمدتاً از کانی‌های رسی تشکیل شده است. بنابراین می‌توان تاثیر کانی شناسی بر طیف‌های Vis-NIR را یک ویژگی ارزشمند برای پیش بینی پارامتر α فرض کرد.

نتیجه‌گیری

در این مطالعه، ۱۳۰ نمونه خاک از مناطق مختلف (همدان، کرمانشاه، آذربایجان غربی، آذربایجان شرقی و مازندران) جمع‌آوری و بازتاب طیفی آن‌ها در بازه ۲۵۰۰-۳۵۰ نانومتر ثبت گردید. یافته‌های این پژوهش نشان داد که بهره‌گیری از ترکیب داده‌های طیفی خاک در کنار الگوریتم‌های یادگیری ماشین، به ویژه RF، می‌تواند روشی موثر، دقیق و مقرون به صرفه برای برآورد پارامترهای کلیدی SWRC فراهم آورد. تحلیل مدل‌ها نشان داد که روش‌های پیش پردازش طیفی، به ویژه MSC، به طور معنی‌داری دقت پیش‌بینی را برای پارامترهای θ_s و n افزایش داده‌اند. هم‌چنین، روش بدون پیش پردازش در تخمین پارامتر α عملکرد مناسبی از خود نشان داد. علاوه بر این، تحلیل مولفه‌های اصلی نقش کلیدی برخی طول موج‌های خاص در محدوده Vis-NIR، به ویژه طول موج‌های ۸۰۴، ۱۰۰۰، ۱۳۳۷ و ۲۱۳۷ نانومتر، را در تبیین تغییرات SWRC برجسته ساخت. در مجموع، این مطالعه نه تنها قابلیت مدل‌های طیفی-یادگیری ماشین را در پیش بینی دقیق پارامترهای

SWRC تایید می‌کند، بلکه بستری علمی برای به کارگیری روش‌های نوین دیجیتال در مدیریت پایدار منابع خاک و آب فراهم می‌آورد.

فهرست منابع

- Babaeian E., Homae M., Montzka C., Vereecken H., and Norouzi A.A. 2015a. Towards retrieving soil hydraulic properties by hyperspectral remote sensing. *Vadose Zone Journal*, 14(3): 1-17.
- Babaeian E., Homae M., Vereecken H., Montzka C., Norouzi, A.A., and van Genuchten M.T. 2015b. A comparative study of multiple approaches for predicting the soil–water retention curve: hyperspectral information vs. basic soil properties. *Soil Science Society of America Journal*, 79: 1043–1058.
- Bayat H., Neyshaburi M.R., Mohammadi K., Nariman-Zadeh N., Irannejad M., and Gregory A.S. 2013. Combination of artificial neural network and fractal theory to predict soil water retention curve. *Computers and Electronics in Agriculture*, 92: 92-103.
- Beniaich, A., Otten, W., Shin, H. C., Cooper, H. V., Rickson, J., Soulaïmani, A., & El Gharous, M. (2023). Evaluation of pedotransfer functions to estimate some of soil hydraulic characteristics in North Africa: A case study from Morocco. *Frontiers in Environmental Science*, 11, 1090688.
- Blake, G. R., & Hartge, K. H. (1986). Particle Density. In *Methods of Soil Analysis* (pp. 377-382). <https://doi.org/https://doi.org/10.2136/sssabookser5.1.2ed.c14>
- Botula, Y.-D., Nemes, A., Mafuka, P., Van Ranst, E., & Cornelis, W. M. (2013). Prediction of water retention of soils from the humid tropics by the nonparametric-nearest neighbor approach. *Vadose Zone Journal*, 12. <https://doi.org/10.2136/vzj2012.0123>
- Bowers, S.A.; Hanks, R.J. (1965) Reflection of radiant energy from soils. *Soil Sci.* 100, 130–138. [CrossRef]
- Castaldi, F., Palombo, A., Pascucci, S., Pignatti, S., Santini, F., & Casa, R. (2015). Reducing the Influence of Soil Moisture on the Estimation of Clay from Hyperspectral Data: A Case Study Using Simulated PRISMA Data. *Remote Sensing*, 7(11), 15561-15582. <https://www.mdpi.com/2072-4292/7/11/15561>
- Chatrenour, M., Landi, A., Bahrami, H., & Mirzaei, S. (2023). Dust source clay content and salinity estimation using VNIR spectrometry. *Arid Land Research and Management*, 37(3), 369-388. <https://doi.org/10.1080/15324982.2023.2170837>
- da Silva, A. C., Armindo, R. A., Minasny, B., & Prevedello, C. L. (2021). Evaluating the Splintex model for estimating the soil water retention curve for a wide range of soils. *Soil and Tillage Research*, 209, 104974
- Davari, M., Fahmideh, S., & Mosaddeghi, M. R. (2022). Rapid assessment of soil water repellency indices using Vis-NIR spectroscopy and pedo-transfer functions. *Geoderma*, 406, 115486. <https://doi.org/https://doi.org/10.1016/j.geoderma.2021.115486>
- Gee, G. W., & Or, D. (2002). 2.4 Particle-size analysis. *Methods of soil analysis: Part 4 physical methods*, 5, 255-293.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.
- Knadel, M., Viscarra Rossel, R. A., Deng, F., Thomsen, A., & Greve, M. H. (2013). Visible–near infrared spectra as a proxy for top soil texture and glacial boundaries. *Soil Science Society of America Journal*, 77(2), 568-579.
- Li, P., Zha, Y., Zuo, B., & Zhang, Y. (2023). A family of soil water retention models based on sigmoid functions. *Water Resources Research*, 59(3), e2022WR033160
- Meurer, K. H. E., Chenu, C., Coucheney, E., Herrmann, A. M., Keller, T., Kätterer, T., & Jarvis, N. (2020). Modelling dynamic interactions between soil structure and the storage and turnover of soil organic matter. *Biogeosciences Discussions*, 2020, 1-34
- Minasny, B., McBratney, A. B., & Bristow, K. L. (1999). Comparison of different approaches to the development of pedotransfer functions for water-retention curves. *Geoderma*, 93, 225–253. [https://doi.org/10.1016/s0016-7061\(99\)00061-0](https://doi.org/10.1016/s0016-7061(99)00061-0)
- Onyedikachi, S. N., Gomes, C., & van Collier, J. (2022, October). Equations for predicting the volumetric water content affecting the resistivity of soil in the vicinity of a driven earth electrode in variably saturated homogeneous soil. In *2022 36th International Conference on Lightning Protection (ICLP)* (pp. 1-6). IEEE
- Pinheiro É.F., Ceddia M., Clingensmith C., Grunwald S., and Vasques G. 2017. Prediction of soil physical and chemical properties by visible and near-infrared diffuse reflectance spectroscopy in the Central Amazon. *Remote Sensing*, 9, 293.
- Pituello, C., Dal Ferro, N., Francioso, O., Simonetti, G., Berti, A., Piccoli, I., Pisi, A., & Morari, F. (2018). Effects of biochar on the dynamics of aggregate stability in clay and sandy loam soils. *European Journal of Soil Science*, 69(5), 827-842. <https://doi.org/https://doi.org/10.1111/ejss.12676>
- Rastgou, M., Bayat, H., Mansoorizadeh, M., & Gregory, A. S. (2020). Estimating the soil water retention curve: Comparison of multiple nonlinear regression approach and random forest data mining technique. *Computers and Electronics in Agriculture*, 174, 105502. <https://doi.org/10.1016/j.compag.2020.105502>
- Rinnan, Å., Berg, F. v. d., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28(10), 1201-1222. <https://doi.org/https://doi.org/10.1016/j.trac.2009.07.007>
- Silva, L. S., Marques Júnior, J., Barrón, V., Gomes, R. P., Teixeira, D. D. B., Siqueira, D. S., & Vasconcelos, V. (2020). Spatial variability of iron oxides in soils from Brazilian sandstone and basalt. *Catena*, 185, 104258. <https://doi.org/https://doi.org/10.1016/j.catena.2019.104258>

- Šimanský, V., Juriga, M., Jonczak, J., Uzarowicz, Ł., & Stepień, W. (2019). How relationships between soil organic matter parameters and soil structure characteristics are affected by the long-term fertilization of a sandy soil. *Geoderma*, 342, 75-84.
- Stenberg, B. (2010). Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma*, 158(1), 15-22. <https://doi.org/https://doi.org/10.1016/j.geoderma.2010.04.008>
- Storey, E. E., & Helmy, A. S. (2019). Optimized preprocessing and machine learning for quantitative Raman spectroscopy in biology. *Journal of Raman Spectroscopy*, 50(7), 958-968. <https://doi.org/https://doi.org/10.1002/jrs.5608>
- Van Genuchten, M.T., (1980). "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils". *Soil Science Society of America Journal* 44(5), 892-898.
- Viscarra Rossel, R. A. (2008). ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, 90(1), 72-83. <https://doi.org/https://doi.org/10.1016/j.chemolab.2007.06.006>
- Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1), 59-75. <https://doi.org/https://doi.org/10.1016/j.geoderma.2005.03.007>
- Yu, B., Yan, C., Yuan, J., Ding, N., & Chen, Z. (2023). Prediction of soil properties based on characteristic wavelengths with optimal spectral resolution by using Vis-NIR spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 293, 122452. <https://doi.org/https://doi.org/10.1016/j.saa.2023.122452>

Estimation of soil water retention curve parameters in the Van Genuchten model using soil and spectral pedotransfer functions

Ch.Madizadeh ¹, H. Bayat ^{*2}, M. Davari

1- Ph. D. Student of Soil Science and Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran.

[\(Chimanmehdzadeh@gmail.com\)](mailto:Chimanmehdzadeh@gmail.com)

2- Professor, Department of Soil Science and Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran.

[\(h.bayat@basu.ac.ir\)](mailto:h.bayat@basu.ac.ir)

3- Professor, Department of Soil Science, Faculty of Agriculture, University of Kurdistan, Sanandaj, Iran.

m.davari@uok.ac.ir

Abstract

Use of indirect methods such as pedotransfer functions (PTFs) and spectral behavior analysis of soil has been considered for estimating soil physical, chemical and hydraulic properties. This study aimed to estimate the parameters of the soil water retention curve (SWRC) including residual (θ_r) and saturated (θ_s) volumetric moisture contents, α and n parameters by utilizing spectral reflectance and soil properties using random forest (RF) method. 130 soil samples were collected from five provinces of Iran, their properties were measured and the spectral reflectance of the samples was recorded. Several preprocessing methods such as multiple scatter correction (MSC), first and second derivatives with Savitzky-Glay filter (FD-SG, FD-SG2) and standard normal variate (SNV) were used. The results showed that MSC had the best performance in estimating θ_s and n parameter and significantly reduced the AIC criterion. For the α parameter, the no preprocessing (NP) method provided the best results. Principal component analysis identified wavelengths of 804, 1000, 1337 and 2137 nm in the Vis-NIR region as the most key wavelengths in describing SWRC. This study showed that combining spectral and soil data with the random forest algorithm provides an accurate and cost-effective approach to estimate SWRC parameters.

Keywords: pedotransfer functions, random forest, soil water retention curve, Vis-NIR spectral reflectance