



Creating pseudo-continuous pedotransfer functions for estimating soil water retention curve using the M5 tree method

Reza Kiani¹, Hossein Bayat^{*2}

1. Ph.D. Student of Soil Science, Dept. of Soil Science and Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran. E-mail: siavashkiani37@gmail.com
2. Corresponding Author, Associate Prof., Dept. of Soil Science and Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran. E-mail: h.bayat@basu.ac.ir, hbayat2001@gmail.com

Article Info

Article type:
Full Length Research Paper

Article history:

Received: 04.09.2024
Revised: 05.20.2024
Accepted: 05.20.2024

Keywords:

Estimation error map,
Field capacity moisture content,
Non-linear regression,
Organic matter,
Saturated hydraulic conductivity

ABSTRACT

Background and Objectives: Recently, pseudo-continuous transfer functions (PC-PTFs) have been introduced for estimating soil water retention curve (SWRC). The M5 tree method is similar to regression trees, where linear functions are located in its leaves, and it has a high capability in creating transfer functions. These functions are highly sensitive to the power of machine learning algorithms. However, so far, the powerful M5 tree method has not been used to develop PC-PTFs for a wide range of soil textures. Additionally, the effectiveness of some soil structural variables in improving PC-PTFs has not been investigated, so far. Furthermore, the dependency of the error distribution of PC-PTFs on soil textural triangles to various factors has not been deeply examined. Therefore, the objectives of this study were to develop PC-PTFs using the M5 method, investigate the effect of soil structural variables on the performance of these functions, and examine the error dependence of these functions on different factors.

Materials and Methods: A total of 120 soil samples were collected from depths of 10 to 60 centimeters, with agricultural, orchard, and pastureland uses of Tehran and Hamedan provinces, and soil texture, bulk density (BD), SWRC, saturated hydraulic conductivity (Ks), organic matter (OM), mean weight diameter (MWD) of soil aggregates, and penetration resistance at 300 hectopascals (PR300) were measured. Thirteen PC-PTFs, in three groups of inputs, were developed to estimate SWRC, using M5 tree and non-linear regression methods. The error distribution of all PC-PTFs was plotted on the soil texture triangle, according to root mean square error (RMSE).

Results: In the first function, soil suction was used as the only estimator. A non-linear regression model produced an acceptable model for the first function with a R^2 of 0.718. In PC-PTFs 3 to 6, components of soil texture, BD, and FC (at 300 hPa matric suction) and PWP (at 15000 hPa matric suction) moisture contents were used to estimate SWRC. The R^2 for these functions ranged from 0.719 to 0.990, indicating an improvement in the performance of SWRC estimation. In the M5 method, the use of FC significantly improved the model performance and created an optimal model, resulting in RMSE of 0.015 and 0.020 $\text{cm}^3\text{cm}^{-3}$, and R^2 of 0.987 and 0.973 in the training and validation stages, respectively. In the M5 method, any function using Ks

and MWD as estimators showed significant improvement compared to PC-PTF4, which used soil texture components and BD as estimators. The AIC values in both training and validation stages in the M5 method were 37% to 283% and 111% to 157% lower compared to non-linear regression, respectively. The error distribution on the soil texture triangle showed no dependence on soil texture but was related to the method of creating PC-PTFs and relevant input variables.

Conclusion: A powerful artificial intelligence methods can be employed to create a comprehensive model for SWRC. This would eliminate the need for users to rely on various SWRC models such as van Genuchten for different soils. Incorporating a set of soil texture and structure variables increases the accuracy of SWRC estimation. However, among structural variables, those indicating pore size distribution were more suitable for SWRC estimation. The greater impact of FC compared to PWP demonstrated the higher efficiency of moisture in intermediate matric suctions for SWRC estimation. The robust algorithm of the M5 tree method identified some patterns of relationships between input and output variables that were not detectable by non-linear regression. Considering the dependence of error distribution on the soil texture triangle to the method of creating PC-PTFs and input variables, categorizing error distribution maps should be done based on the mentioned factors.

Cite this article: Kiani, Reza, Bayat, Hossein. 2024. Creating pseudo-continuous pedotransfer functions for estimating soil water retention curve using the M5 tree method. *Journal of Soil Management and Sustainable Production*, 14 (3), 53-76.



© The Author(s).

DOI: 10.22069/EJSMS.2024.22346.2141

Publisher: Gorgan University of Agricultural Sciences and Natural Resources



ایجاد توابع انتقالی شبه پیوسته برای برآورد منحنی نگه‌داشت آب خاک با استفاده از روش درخت M5

رضا کیانی^۱، حسین بیات^{۲*}

۱. دانشجوی دکتری گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه بوعلی‌سینا، همدان، ایران.
رایانامه: siavashkiani37@gmail.com
۲. نویسنده مسئول، دانشیار گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه بوعلی‌سینا، همدان، ایران.
رایانامه: h.bayat@basu.ac.ir, hbayat2001@gmail.com

اطلاعات مقاله	چکیده
<p>نوع مقاله: مقاله کامل علمی- پژوهشی</p> <p>تاریخ دریافت: ۱۴۰۳/۰۱/۲۱</p> <p>تاریخ ویرایش: ۱۴۰۳/۰۲/۳۱</p> <p>تاریخ پذیرش: ۱۴۰۳/۰۲/۳۱</p> <p>واژه‌های کلیدی: رطوبت ظرفیت زراعی، رگرسیون غیرخطی، ماده آلی، نقشه خطای برآورد، هدایت هیدرولیکی اشباع</p>	<p>سابقه و هدف: اخیراً توابع انتقالی شبه پیوسته (PC-PTF) برای برآورد منحنی نگه‌داشت آب خاک (SWRC) معرفی شده‌است. این توابع به شدت به قدرت الگوریتم‌های یادگیری ماشین حساس هستند. روش درخت M5 مشابه درخت‌های رگرسیون است، که توابع خطی در برگ‌های آن قرار دارند و دارای قدرت بالایی در ایجاد توابع انتقالی است. با این وجود تاکنون از این روش برای ایجاد PC-PTF برای طیف وسیعی از بافت‌های خاک استفاده نشده است. هم‌چنین، کارایی برخی متغیرهای ساختمان خاک در بهبود PC-PTFها بررسی نشده است. علاوه بر این، وابستگی توزیع خطای PC-PTFها به عوامل مختلف مورد بررسی عمیق قرار نگرفته است. بنابراین اهداف این مطالعه ایجاد توابع انتقالی شبه پیوسته با استفاده از روش M5، بررسی تأثیر متغیرهای ساختمان خاک بر عملکرد این توابع و بررسی وابستگی خطای این توابع به عوامل مختلف بود.</p> <p>مواد و روش‌ها: تعداد ۱۲۰ نمونه خاک از استان‌های تهران و همدان از عمق ۱۵ تا ۶۰ سانتی‌متری با کاربری زراعی، باغی و مرتع برداشت شد. بافت خاک، جرم مخصوص ظاهری (BD)، هدایت هیدرولیکی اشباع (Ks)، مواد آلی (OM)، میانگین وزنی قطر خاکدانه‌ها (MWD) و مقاومت فروری در مکش ۳۰۰ هکتوپاسکال (PR300) اندازه‌گیری شد. ۱۳ تابع انتقالی شبه پیوسته، با هر کدام از روش‌های درخت M5 و رگرسیون غیر خطی، در قالب ۳ گروه متغیر ورودی، برای برآورد SWRC ایجاد شد. توزیع خطای تمام توابع انتقالی شبه پیوسته براساس آماره مجذور میانگین مربعات خطا بر روی مثلث بافت خاک ترسیم شد.</p>

یافته‌ها: در تابع اول، مکش خاک به‌عنوان تنها تخمین‌گر مورد استفاده قرار گرفت. رگرسیون غیرخطی مدل قابل‌قبولی برای تابع اول با ضریب تعیین $0/718$ ایجاد کرد. در توابع شبه‌پیوسته ۳ تا ۶ اجزای بافت، BD و رطوبت FC (مکش 300 هکتوپاسکال) و PWP (مکش 15000 هکتوپاسکال) برای برآورد SWRC مورد استفاده قرار گرفتند. ضریب تعیین این توابع $0/719$ تا $0/990$ به‌دست آمد که بهبود عملکرد برآورد SWRC را نشان داد. در روش M5، استفاده از رطوبت FC موجب بهبود قابل‌توجه عملکرد مدل گردید و با مجذور میانگین مربعات $0/015$ و $0/020 \text{ cm}^3 \text{ cm}^{-3}$ ، و ضریب تعیین $0/987$ و $0/973$ به‌ترتیب در مراحل آموزش و اعتبارسنجی، یک مدل بهینه را ایجاد کرد. در روش M5، هر تابعی که از Ks و MWD به‌عنوان تخمین‌گر استفاده کرد، بهبود معنی‌داری نسبت به تابع ۴ که از اجزای بافت خاک و BD به‌عنوان تخمین‌گر استفاده کرده بود، نشان داد. مقادیر آماره آکائیک در هر دو مرحله آموزش و اعتبارسنجی در روش M5 نسبت به رگرسیون غیرخطی، به‌ترتیب به‌میزان 37 تا 283 درصد و 111 تا 157 درصد کم‌تر به دست آمد. توزیع خطا بر روی مثلث بافت خاک، هیچ وابستگی به بافت خاک نشان نداد، ولی به روش ایجاد توابع شبه‌پیوسته و متغیرهای ورودی مرتبط بود.

نتیجه‌گیری: می‌توان با روش‌های هوش مصنوعی قدرتمند یک مدل جامع برای SWRC ایجاد کرد، که باعث عدم نیاز کاربران به مدل‌های مختلف SWRC مانند ونگونختن برای خاک‌های مختلف خواهد بود. استفاده از مجموعه‌ای از متغیرهای بافت و ساختمان خاک باعث افزایش دقت برآورد SWRC می‌گردد. ولی آن‌دسته از متغیرهای ساختمانی، که شاخصی از توزیع اندازه منافذ هستند، برای برآورد SWRC مناسب‌تر بودند. تأثیر بیش‌تر FC، در بهبود برآورد SWRC نسبت به PWP، کارایی بیش‌تر رطوبت در مکش‌های میانی در برآورد SWRC را نشان داد. الگوریتم قوی روش درخت M5 برخی الگوهای روابط میان متغیرهای ورودی و خروجی که توسط روش رگرسیون غیرخطی قابل تشخیص نبود، را تشخیص داد. با توجه به وابستگی توزیع خطا بر روی مثلث بافت خاک، به روش ایجاد توابع شبه‌پیوسته و متغیرهای ورودی، باید نقشه‌های توزیع خطا را بر اساس فاکتورهای مذکور دسته‌بندی نمود.

استناد: کیانی، رضا، بیات، حسین (۱۴۰۳). ایجاد توابع انتقالی شبه‌پیوسته برای برآورد منحنی نگه‌داشت آب خاک با استفاده از روش

درخت M5. نشریه مدیریت خاک و تولید پایدار، ۱۴ (۳)، ۵۳-۷۶.

DOI: 10.22069/EJSMS.2024.22346.2141



© نویسندگان.

ناشر: دانشگاه علوم کشاورزی و منابع طبیعی گرگان

مقدمه

ویژگی‌های هیدرولیکی خاک برای اندازه‌گیری آب موجود برای گیاهان، رفتار تغییر شکل خاک (۱) و مدل‌سازی حرکت آب و مواد محلول از طریق خاک مورد نیاز هستند (۲). این ویژگی‌ها، عملکرد خاک و مدیریت اراضی را تحت‌تأثیر قرار می‌دهند و برای استفاده در بسیاری از رشته‌ها از جمله هیدرولوژی، زراعت، هواشناسی، اکولوژی، حفاظت محیط زیست و دیگر حوزه‌های مربوط به خاک الزامی هستند (۳، ۴). روش‌های مزرعه‌ای و آزمایشگاهی زیادی برای اندازه‌گیری مستقیم ویژگی‌های هیدرولیکی وجود دارد، اما به دست آوردن ارتباط بین پتانسیل ماتریک و مقدار آب خاک هزینه‌بر و زمان‌بر است (۵، ۶، ۷، ۸). این شرایط باعث شده است تعداد زیادی از دانشمندان برای پیش‌بینی ویژگی‌های هیدرولیکی خاک مانند مقدار آب خاک در پتانسیل‌های خاص یا هدایت هیدرولیکی خاک، توابع انتقالی^۱ (PTFs) را پیشنهاد کنند. این توابع از ویژگی‌ها سهل‌الوصول برای به‌دست آوردن منحنی نگه‌داشت آب خاک^۲ (SWRC) استفاده می‌کنند. توابع انتقالی به دو گروه نقطه‌ای و پارامتریک تقسیم می‌شوند (۹). PTF‌های نقطه‌ای فقط نقاط محدودی از منحنی نگه‌داشت آب خاک را پیش‌بینی می‌کنند. PTF‌های پارامتریک پارامترهای یک معادله منحنی نگه‌داشت آب خاک مانند، مدل ون‌گونوختن (۱۹۸۰) (۱۰) را پیش‌بینی می‌کنند. هر کدام از این توابع دارای معایبی هستند. تعداد زیادی از توابع نقطه‌ای برای به دست آوردن کل منحنی نگه‌داشت آب خاک لازم است (۱۱). عیب اصلی PTF‌های پارامتریک نیز این است که، وقتی این توابع توسعه یافتند، کاربر نمی‌تواند مدل منحنی نگه‌داشت آب خاک از پیش انتخاب شده را تغییر دهد، که ممکن است این مدل به اندازه کافی شکل واقعی SWRC را نشان ندهد (۱۲). علاوه بر این،

گاهی اوقات مرتبط کردن پارامترهای معادلات هیدرولیکی خاک با خصوصیات پایه خاک چالش‌برانگیز است (۱۲). حق‌وردی و همکاران (۲۰۱۲) برای اولین بار مفهوم توابع انتقالی شبه پیوسته^۳ (PC-PTF) را به عنوان یک رویکرد جایگزین برای توابع نقطه‌ای و پارامتریک معرفی کردند (۱۳). یک PC-PTF از لگاریتم پتانسیل ماتریک خاک به عنوان پارامتر ورودی علاوه بر سایر متغیرهای ورودی مانند بافت خاک و جرم مخصوص ظاهری^۴ (BD) برای پیش‌بینی مقدار آب خاک استفاده می‌کند. با ورود یک دامنه وسیع از پتانسیل‌های ماتریک خاک به عنوان ورودی، می‌توان دامنه مناسبی از مقدار آب خاک را پیش‌بینی کرد که منجر به ایجاد یک SWRC شبه پیوسته می‌شود (۱۴). یک PC-PTF از مکش ماتریک خاک در تعداد محدود جفت نقاط اندازه‌گیری شده منحنی نگه‌داشت آب خاک به همراه برخی متغیرهای ورودی، برای برآورد SWRC استفاده می‌کند. در نتیجه برای ایجاد توابع، به شدت وابسته به قدرت الگوریتم‌های یادگیری ماشین است (۱۴) و استفاد از روش‌های قدرتمند هوش مصنوعی برای ایجاد این نوع از توابع انتقالی، باعث موفقیت آن‌ها خواهد شد. علی‌رغم این‌که روش‌های متعددی برای ایجاد توابع نقطه‌ای و پارامتریک استفاده شده است (۱۵، ۱۶) ولی با این وجود روش‌های استفاده شده برای ایجاد توابع انتقالی شبه پیوسته جهت برآورد منحنی نگه‌داشت آب خاک فقط محدود به روش‌های شبکه‌های عصبی مصنوعی (۱۳، ۱۷)، ماشین‌های بردار پشتیبان (۱۴)، k-نزدیک‌ترین همسایه و رگرسیون خطی چندگانه (۱۸) می‌باشد.

روش M5 توسط پژوهش‌گران مختلفی برای مقاصد غیر از برآورد منحنی نگه‌داشت آب خاک مورد استفاده قرار گرفته است. در زمینه خصوصیات هیدرولیکی خاک برخی پژوهش‌گران نیز برای برآورد نقطه‌ای منحنی نگه‌داشت آب خاک از روش M5

3- Pseudo-continuous pedotransfer functions

4- Bulk density

1- Pedotransfer functions

2- Soil water retention curve

هیدرولیکی اشباع و مقاومت فروروی خاک در بهبود دقت و صحت توابع انتقالی شبه‌پیوسته و ۳) بررسی احتمال وابستگی خطا به فاکتورهای مختلف بود.

مواد و روش‌ها

نمونه‌برداری و اندازه‌گیری ویژگی‌های خاک: برای انجام این پژوهش تعداد ۱۲۰ نمونه دست‌خورده و دست‌نخورده از لایه سطحی (۱۰ تا ۴۰ سانتی‌متر) و در برخی نقاط از لایه زیرسطحی (۲۰ تا ۶۰ سانتی‌متر) خاک‌های استان تهران (۹۰ نمونه) و همدان (۳۰ نمونه) برداشت شد. کاربری اراضی محل‌های نمونه‌برداری در استان تهران زراعی، باغی و مرتع بود. کاربری اراضی محل‌های نمونه‌برداری در استان همدان فقط زراعی یا باغی بود. نمونه‌های دست‌نخورده با استفاده از سیلندرهای نمونه‌برداری به ارتفاع ۴/۵ و قطر ۵/۱ سانتی‌متر تهیه گردید. نمونه‌های دست‌خورده هواخشک شده و از الک ۲ میلی‌متری عبور داده شدند.

بافت خاک با روش هیدرومتر (۲۳)، تعیین کلاس بافت خاک بر اساس سیستم وزارت کشاورزی آمریکا (USDA) و جرم مخصوص ظاهری با روش نمونه‌های دست‌نخورده اندازه‌گیری شد (۲۴). برای اندازه‌گیری منحنی نگه‌داشت آب خاک، مقدار رطوبت در مکش‌های ۰، ۱۰، ۲۰، ۴۰ و ۶۰ هکتوپاسکال با استفاده از دستگاه جعبه شن و در مکش‌های ۳۰۰، ۱۰۰۰، ۳۰۰۰، ۵۰۰۰ و ۱۵۰۰۰ هکتوپاسکال با استفاده از دستگاه صفحات تحت فشار تعیین شد (۲۵). اندازه‌گیری‌ها تا مکش ۳۰۰۰ هکتوپاسکال بر روی نمونه‌های دست‌نخورده و در مکش‌های بالاتر از آن بر روی نمونه‌های دست‌خورده انجام شد. هدایت هیدرولیکی اشباع به روش بار ثابت (۲۶) و مواد آلی به روش اکسیداسیون تر (۲۷) اندازه‌گیری شد. برای به دست آوردن میانگین وزنی قطر خاکدانه‌ها (MWD) از روش کمپر و روزنا (۱۹۸۶) استفاده

استفاده کردند (۱۹). راستگو و همکاران (۲۰۲۲) نیز از روش M5 برای برآورد نقطه‌ای و پارامتریک منحنی نگه‌داشت آب خاک استفاده کردند (۱۶). فقط ذاکری‌نیا و قربانی (۲۰۱۳) از روش M5 برای تخمین مقادیر رطوبت غیراشباع خاک به صورت توابعی وابسته به مکش که می‌توان آن را توابع شبه‌پیوسته نامید، استفاده کردند (۲۰). هرچند مدل‌های ایجاد شد توسط این پژوهش‌گران محدود به سه کلاس بافتی بوده و دقت روش M5 را با هیچ روش دیگری مقایسه نکردند. بررسی منابع نشان داد که روش M5 تاکنون در هیچ مطالعه‌ای برای ایجاد توابع انتقالی شبه‌پیوسته برای طیف وسیعی از بافت‌های خاک استفاده نشده است.

فاکتورهای متفاوتی می‌توانند بر خطای توابع انتقالی ایجاد شده تأثیر داشته باشند. برای بررسی اثر برخی فاکتورها از جمله بافت خاک بر خطای توابع انتقالی، می‌توان توزیع خطا بر روی مثلث بافت خاک را بررسی نمود. محققان مختلفی (۲۱، ۲۲) توزیع خطای ایجاد توابع انتقالی بر روی مثلث بافت خاک را گزارش کرده‌اند. پاپسکی و راولز (۲۰۰۴) گزارش کردند که، به‌علت ناهمگن بودن توزیع خطا بر روی مثلث بافت خاک نمی‌توان وابستگی خطا به بافت خاک در برآورد خصوصیات هیدرولیکی را تأیید نمود (۲۱). راستگو (۲۰۲۰) نیز نتیجه مشابهی را گزارش کرد (۲۲) ولی وابستگی توزیع خطا به عوامل مختلف کمتر مورد بررسی قرار گرفته است. تا جایی که نویسندگان اطلاع دارند تاکنون از هدایت هیدرولیکی اشباع و مقاومت فروروی خاک به‌عنوان تخمین‌گر در توابع انتقالی شبه‌پیوسته برای برآورد SWRC استفاده نشده است. هم‌چنین کارایی روش M5 در ایجاد این نوع از توابع انتقالی مورد بررسی قرار نگرفته است. بنابراین اهداف این مطالعه شامل ۱) ایجاد توابع انتقالی شبه‌پیوسته با استفاده از روش M5 و مقایسه دقت و صحت برآورد توسط آن‌ها با روش رگرسیون غیرخطی، ۲) بررسی تأثیر متغیرهای ساختمان خاک شامل میانگین وزنی قطر خاک‌دانه‌ها، هدایت

معکوس، مربع، جذر، معکوس توان دوم، کاکس باکس، جانسون، و نرمال اسکور بر روی داده‌ها انجام شد. برای هر متغیر تبدیلی که به توزیع نرمال نزدیک‌ترین نتیجه را می‌داد، انتخاب شد. سپس همه متغیرهای ورودی (به جز مکش خاک) استاندارد شدند تا میانگین صفر و انحراف استاندارد یک داشته باشند. از مکش خاک لگاریتم گرفته شد. برای بررسی چندمخطی از فاکتور عامل تورم واریانس^۳ (VIF) استفاده شد، که در همه توابع، تمام متغیرها VIF کم‌تر از ۱۰ داشتند. در مجموع ۱۳ تابع انتقالی شبه پیوسته در این پژوهش ایجاد شد. این ۱۳ تابع به ۳ گروه، بر اساس متغیرهای ورودی، تقسیم شدند. این تقسیم‌بندی در جدول ۱ مشخص گردیده است. برای جلوگیری از ایجاد چندمخطی متغیر سیلت به‌عنوان ورودی استفاده نشد.

گردید (۲۸). سری الک‌ها شامل ۴، ۲، ۱، ۰/۵ و ۰/۲۵ میلی‌متر بودند. اصلاح رطوبتی و اصلاح شن نیز در این روش انجام شد. مقاومت فروری^۱ (PR) با استفاده از دستگاه فروسنج ریز با قطر قاعده مخروط ۲ میلی‌متر و زاویه ۳۰ درجه بر روی نمونه‌های دست‌نخورده تعیین شد. ابتدا نمونه‌ها در مکش ۳۰۰ هکتوپاسکال به تعادل رطوبتی رسیدند و سپس اندازه‌گیری مقاومت فروری بر روی آن‌ها انجام شد (۲۹).

ایجاد توابع انتقالی: قبل از ایجاد توابع انتقالی فرایند پیش‌پردازش داده‌ها از جمله جایگزینی داده‌های پرت و آزمون توزیع نرمال داده‌ها انجام شد. بررسی نرمال بودن داده‌ها با استفاده از آزمون کلموگروف-اسمیرنوف^۲ در نرم‌افزار MiniTab 19 انجام شد. در صورت نرمال نبودن تبدیل‌های متعدد مانند لگاریتم،

جدول ۱- گروه‌بندی و متغیرهای ورودی برای سیزده تابع انتقالی شبه پیوسته در برآورد منحنی نگه‌داشت آب خاک.

Table 1. Grouping and inputs for thirteen pseudo-continuous pedotransfer functions in estimating soil water retention curve.

ورودی‌ها Inputs	توابع PTF	گروه Group
h^{\dagger}	1	اول (First)
hTexture	2	دوم (Second)
hSC	3	
hSCBD	4	
hSCBDFC	5	
hSCBDFCPWP	6	
hSCBDOM	7	
hSCBDKs	8	سوم (Third)
hSCBDMKs	9	
hSCBDMWD	10	
hSCBDPR300	11	
hSCBDMWDPR300	12	
hSCBDMKsMWDPR300	13	

†. در جدول فوق h : مکش ماتریک آب خاک، C: مقدار رس، S: مقدارشن، BD: جرم مخصوص ظاهری، OM: ماده آلی، Ks: هدایت هیدرولیکی اشباع، MWD: میانگین وزنی قطر خاکدانه‌ها، PR300: مقاومت فروری در مکش ۳۰۰ هکتوپاسکال و T0 - T15000 (PWP): به ترتیب رطوبت حجمی در مکش ۰ تا ۱۵۰۰۰ هکتوپاسکال، FC: رطوبت ظرفیت زراعی (رطوبت در مکش ۳۰۰ هکتوپاسکال)، PWP: رطوبت نقطه پژمردگی دائم (رطوبت در مکش ۱۵۰۰۰ هکتوپاسکال)

†. In the above Table: h: soil water matric suction, C: clay, S: sand, BD: bulk density, OM: organic matter, Ks: saturated hydraulic conductivity, MWD: mean weight diameter of soil aggregates, PR300: penetration resistance at 300 hectopascals, T0 - T15000 (PWP): volumetric moisture content at 0 to 15000 hectopascals., FC: field capacity (moisture content at 300 hPa matric suction) and PWP: permanent wilting point (moisture content at 15000 hPa matric suction)

- 1- Penetration resistance
- 2- Kolmogorov-Smirnov
- 3- Variance inflation factor

$$R_{adj}^2 = 1 - \left[\frac{(1-R^2)(N-1)}{(N-m-1)} \right] \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (2)$$

که در آن، R_{adj}^2 ضریب تبیین تعدیل شده، m تعداد متغیرهای ورودی مدل، $RMSE$ مجذور میانگین مربعات خطا، y_i مقدار اندازه‌گیری شده رطوبت، \hat{y}_i مقدار رطوبت پیش‌بینی شده توسط توابع و N تعداد کل داده‌های مرحله آموزش یا اعتبارسنجی می‌باشد (۳۲).

معیار اطلاعات آکائیک^۴، AIC ، (۳۳)

$$AIC = N \ln \left[\sum_{i=1}^N (y_i - \hat{y}_i)^2 \right] + 2n_p \quad (3)$$

که در آن، n_p تعداد پارامترهای مدل که باید برآورد شوند. هرچه AIC کوچک‌تر باشد کیفیت برازش بالاتر است. پس از بسط توابع به دو روش مختلف، برای مقایسه یک تابع با سایر توابع براساس نوع متغیرهای ورودی از مقایسه آماره آکائیک استفاده شد. روش مقایسه با بهره‌گیری از آماره AIC به این گونه می‌باشد که نخست یکی از مدل‌ها به‌عنوان مدل پایه گزینش می‌شود و مدل‌های دیگر نسبت به آن سنجیده می‌شود. اگر مقدار AIC مدل مقایسه‌شونده، از $۰/۹۵$ اندازه AIC به دست آمده برای مدل پایه کم‌تر باشد، به این معنی است که مدل مقایسه‌شونده با اطمینان ۹۵ درصد بهتر از مدل پایه است و در غیر این صورت مدل پایه ترجیح داده می‌شود (۳۴).

نتایج و بحث

ویژگی‌های آماری: ویژگی‌های آماری متغیرهای مورد مطالعه در جدول ۲ و توزیع نمونه‌ها بر روی مثلث

در این پژوهش برای تعیین نمونه‌های آموزش و اعتبارسنجی از روش k -Fold با $k=10$ استفاده شد (۳۰). با این روش تمام نمونه‌ها هم برای آموزش و هم برای اعتبارسنجی استفاده شدند. روش درخت رگرسیونی $M5$ با استفاده از نرم‌افزار Weka 3.8 و رگرسیون غیر خطی با استفاده از نرم‌افزار DataFit 8 برای ایجاد توابع انتقالی شبه‌پیوسته مورد استفاده قرار گرفتند.

درخت M5: روش درخت $M5$ یکی از روش‌های داده‌کاوی است که به تازگی توجه زیادی را به‌خود جلب کرده و در مدل‌سازی مسائل مختلف استفاده می‌شود. نظریه عمومی این روش مشابه درخت‌های رگرسیون است، با این تفاوت که به جای مقادیر ثابت و برجسب‌های کلاس، توابع خطی در برگ‌های آن قرار دارند. مانند روش $CART^1$ ، روش $M5$ نیز شاخه‌های خود را به صورت دودویی ایجاد می‌کند. در درخت $M5$ ، مدل به اجزای کوچک تقسیم شده و برای هر جزء یک مدل رگرسیون خطی چندمتغیره ایجاد می‌شود. تفاوت اصلی این روش با روش $CART$ در این است که برگ‌های مدل در درخت $M5$ ، برخلاف درخت‌های رگرسیونی $CART$ که حاوی مقادیر عددی هستند، شامل مدل‌های خطی چندمتغیره هستند. به همین علت، مدل درخت $M5$ می‌تواند معادل توابع خطی چندمتغیره تلقی شود (۳۱).

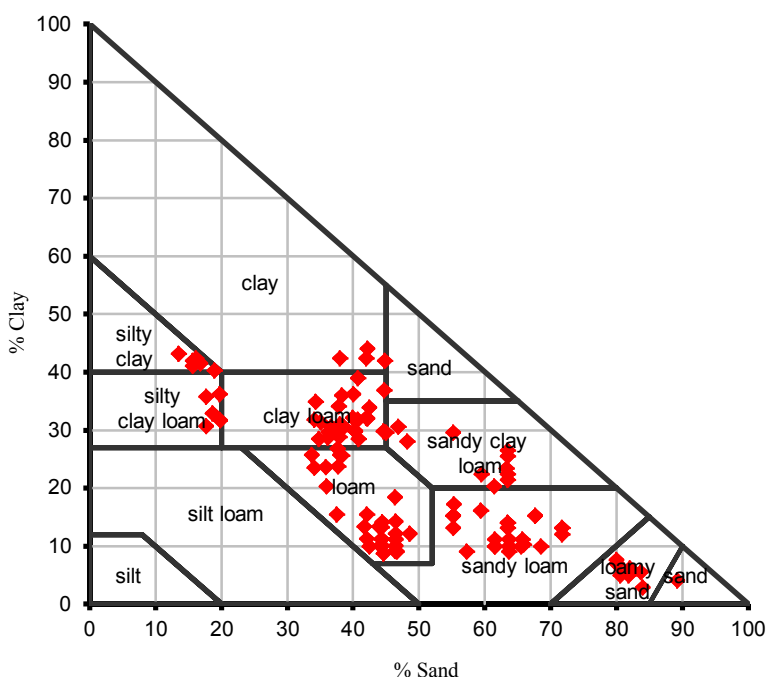
برای ارزیابی دقت^۲ و قابلیت اعتماد^۳ توابع از آماره‌های ضریب تبیین، مجذور میانگین مربعات خطا و معیار اطلاعات آکائیک استفاده شد (روابط ۱ تا ۳). برای بررسی وابستگی خطای توابع انتقالی ایجاد شده به فاکتورهای مختلف توزیع مجذور میانگین مربعات خطای مرحله اعتبارسنجی بر روی مثلث بافت خاک رسم شد.

- 1- Classification and regression trees
- 2- Accuracy
- 3- Reliability

4- Akaike's information criterion

نیز در دامنه متوسط و بالا بوده و با افزایش مکش، ضریب تغییرات نیز افزایش یافت. کم‌ترین ضریب تغییرات برابر با ۱۹ درصد برای رطوبت اشباع و بیش‌ترین ضریب تغییرات برابر با ۳۹ درصد برای رطوبت در مکش ۱۵۰۰۰ هکتوپاسکال به‌دست آمد. مثلث بافت خاک نشان می‌دهد که نمونه‌های مورد مطالعه در این پژوهش شامل ۹ کلاس بافتی بر اساس سیستم وزارت کشاورزی آمریکا (USDA) بودند (شکل ۱)، تنها کلاس‌های بافتی رس شنی، لوم سیلتی و سیلت در میان نمونه‌ها موجود نبود. بنابراین نتایج مثلث بافت خاک نیز تأییدی بر وسیع بودن دامنه متغیرهای بافتی بود.

بافت خاک نیز در شکل ۱ آورده شده است. ویژگی‌های فیزیکی و مکانیکی خاک که دارای ضریب تغییرات بالاتر از ۳۵ درصد باشند، تغییرپذیری بالایی دارند (۳۵). بر این اساس، ضریب تغییرات رس، ماده آلی و میانگین وزنی قطر خاکدانه‌ها به ترتیب ۵۳، ۴۱ و ۴۷ درصد بود که نشان‌دهنده تغییرپذیری زیاد این متغیرها است. ضریب تغییرات شن و هدایت هیدرولیکی اشباع به‌ترتیب ۳۵ و ۲۷ درصد بود که در کلاس متوسط (بین ۱۵ و ۳۵ درصد) قرار گرفتند (۳۶). مقدار ضریب تغییرات جرم مخصوص ظاهری و مقاومت فروری کم بود (ضریب تغییرات کم‌تر از ۱۵ درصد) و تأییدی بر تغییرات پایین این متغیرها است (۳۶). ضریب تغییرات رطوبت در مکش‌های مختلف



شکل ۱- توزیع بافت خاک نمونه‌های مطالعه شده بر روی مثلث بافت خاک USDA.

Figure 1. Distribution of soil texture of the studied samples on the USDA soil texture triangle.

جدول ۲- ویژگی‌های آماری متغیرهای مورد مطالعه.

Table 2. Statistical characteristics of the studied variables.

T10 (cm ³ cm ⁻³)	T0 (cm ³ cm ⁻³)	PR300 (kPa)	MWD (mm)	Ks (cm h ⁻¹)	OM (%)	BD (gcm ⁻³)	S (%)	C [†] (%)	
0.34	0.36	742	0.76	1.14	1.59	1.44	48	21	میانگین Mean
0.34	0.36	745	0.74	1.09	1.61	1.43	44	20	میانه Median
0.07	0.07	45	0.36	0.31	0.65	0.07	17	11	انحراف استاندارد Standard Deviation
0.16	0.19	626	0.15	0.66	0.27	1.28	13	3	حداقل Minimum
0.50	0.52	853	1.63	1.87	2.69	1.58	89	44	حداکثر Maximum
>0.15	>0.15	0.718	0.023	0.022	<0.01	0.092	<0.01	<0.01	p-value
			MWD ^{0.5}	J	J		NS	NS	تبدیل Transformation
	T15000, PWP (cm ³ cm ⁻³)	T5000 (cm ³ cm ⁻³)	T3000 (cm ³ cm ⁻³)	T1000 (cm ³ cm ⁻³)	T300, FC (cm ³ cm ⁻³)	T60 (cm ³ cm ⁻³)	T40 (cm ³ cm ⁻³)	T20 (cm ³ cm ⁻³)	
	0.08	0.10	0.17	0.19	0.22	0.27	0.29	0.32	میانگین Mean
	0.08	0.11	0.16	0.19	0.21	0.27	0.29	0.32	میانه Median
	0.03	0.04	0.06	0.07	0.07	0.07	0.07	0.07	انحراف استاندارد Standard Deviation
	0.01	0.02	0.04	0.05	0.08	0.10	0.12	0.14	حداقل Minimum
	0.16	0.19	0.32	0.35	0.37	0.42	0.44	0.48	حداکثر Maximum
	0.068	0.039	>0.15	>0.15	>0.15	>0.15	>0.15	>0.15	p-value
									تبدیل Transformation

† در جدول فوق C: مقدار رس، S: مقدار شن، BD: جرم مخصوص ظاهری، OM: ماده آلی، Ks: هدایت هیدرولیکی اشباع، MWD: میانگین وزنی قطر خاکدانه‌ها، PR300: مقاومت فروری در مکش ۳۰۰ هکتوپاسکال و T0 - T15000 (PWP): به ترتیب رطوبت حجمی در مکش ۰ تا ۱۵۰۰۰ هکتوپاسکال، J: تبدیل جانسون و NS: تبدیل از طریق نرمال اسکور در نرم افزار مینی تب انجام شده. p-value: p: آزمون کلموگروف-اسمیرنوف. ماده آلی با فرمول زیر نرمال شد:

$$J(OM) = -0.0633132 + 0.636356 \times \ln((OM - 0.231063) / (2.78399 - OM))$$

هدایت هیدرولیکی اشباع با فرمول زیر نرمال شد:

$$J(Ks) = 0.360800 + 0.707555 \times \ln((Ks - 0.641147) / (1.90489 - Ks))$$

† In the above Table: C: clay, S: sand, BD: bulk density, OM: organic matter, Ks: saturated hydraulic conductivity, MWD: mean weight diameter of soil aggregates, PR300: penetration resistance at 300 hectopascals, T0 - T15000 (PWP): volumetric moisture content at 0 to 15000 hectopascals, J: Johnson transformation, NS: transformation via Normal Score in MiniTab software, p: p-value of Kolmogorov-Smirnov test. Organic matter was normalized with the following formula:

$$J(OM) = -0.0633132 + 0.636356 \times \ln((OM - 0.231063) / (2.78399 - OM))$$

Saturated hydraulic conductivity was normalized using the following formula:

$$J(Ks) = 0.360800 + 0.707555 \times \ln((Ks - 0.641147) / (1.90489 - Ks))$$

کنترل کننده نگه‌داشت آب در خاک می‌باشد. به همین علت KS دارای همبستگی معنی‌دار با رطوبت خاک بود. MWD و PR300 دارای همبستگی معنی‌دار ($P < 0/05$)، به ترتیب مثبت و منفی، با رطوبت در تعداد محدودی از مکش‌ها بودند. مقاومت فروری دارای همبستگی منفی با نگه‌داشت آب در خاک است. وابستگی بین مقاومت فروری و مقدار رطوبت خاک توسط پژوهش‌گران دیگری نیز گزارش شده است (۴۰).

ضریب همبستگی تمام مکش‌ها در مقابل تمام رطوبت‌ها $0/822^{**}$ - به دست آمد، که در سطح ۱ درصد معنی‌دار بود و استفاده از آن به عنوان تخمین‌گر موجب بهبود برآورد SWRC گردید. لازم به ذکر است چون برای رطوبت در هر مکش برای تمام ۱۲۰ نمونه فقط یک عدد به عنوان مکش وجود دارد و از طرف دیگر در هر مدل همه رطوبت‌ها به طور همزمان برآورد می‌شوند، بنابراین محاسبه ضریب همبستگی برای رطوبت در مقابل مکش به صورت انفرادی منطقی نیست.

همبستگی بین متغیرهای ورودی و خروجی: همبستگی بین متغیرهای ورودی و خروجی در جدول ۳ نشان داده شده است. درصد رس و شن دارای همبستگی معنی‌دار با رطوبت در غالب مکش‌ها بودند. جرم مخصوص ظاهری دارای همبستگی منفی و معنی‌دار ($P < 0/01$) با رطوبت در تمام مکش‌ها بود. چراکه با افزایش جرم مخصوص ظاهری، تخلخل کاهش یافته و مقدار رطوبت کاهش می‌یابد (۳۷، ۳۸، ۳۹). ماده آلی دارای همبستگی مثبت و معنی‌دار ($P < 0/01$) با رطوبت در تمام مکش‌ها، به جز ۵۰۰۰ و ۱۵۰۰۰ هکتوپاسکال، بود. غالباً ماده آلی به عنوان ماده اصلاحی مناسب برای بهبود ویژگی‌های فیزیکی خاک و افزایش ظرفیت نگه‌داشت آب در خاک شناخته شده است.

هدایت هیدرولیکی اشباع نیز دارای همبستگی منفی و معنی‌دار ($P < 0/01$) با رطوبت در تمام مکش‌ها، به جز ۱۵۰۰۰ هکتوپاسکال، بود. متغیر KS به شدت وابسته به بافت و ساختمان خاک بوده و شاخصی از توزیع اندازه منافذ است (۳۸) که

جدول ۳- همبستگی بین متغیرهای ورودی و خروجی (رطوبت در مکش‌های ماتریک مختلف).

Table 3. Correlation between input and output variables (moisture content at different matric suctions).

T15000 (PWP)	T5000	T3000	T1000	T300 (FC)	T60	T40	T20	T10	T0 [†]	
0.265**	0.450**	0.409**	0.477**	0.467**	0.471**	0.456**	0.443**	0.429**	0.430**	C
-0.07	-0.185*	-0.340**	-0.400**	-0.374**	-0.378**	-0.364**	-0.359**	-0.352**	-0.369**	S
-0.11	-0.13	0.201*	0.17	0.15	0.17	0.16	0.16	0.18	0.201*	Si
-0.249**	-0.373**	-0.596**	-0.612**	-0.607**	-0.566**	-0.547**	-0.525**	-0.497**	-0.489**	BD
-0.09	0.00	0.253**	0.328**	0.326**	0.354**	0.355**	0.376**	0.394**	0.407**	OM
-0.04	-0.295**	-0.443**	-0.551**	-0.532**	-0.534**	-0.520**	-0.520**	-0.507**	-0.502**	Ks
0.17	0.221*	0.230*	0.16	0.15	0.11	0.10	0.07	0.04	0.04	MWD
-0.06	-0.05	-0.193*	-0.06	-0.07	-0.07	-0.08	-0.08	-0.09	-0.10	PR300
0.493**	0.696**	0.906**	0.992**	1.000**	0.978**	0.961**	0.930**	0.904**	0.877**	T300 (FC)
1.000**	0.833**	0.548**	0.463**	0.493**	0.521**	0.526**	0.502**	0.491**	0.481**	T15000 (PWP)

* و ** به ترتیب همبستگی معنی‌دار در سطح ۰/۰۵ و ۰/۰۱ را نشان می‌دهند

† در جدول فوق C: مقدار رس، S: مقدار شن، Si: مقدار سیلت، BD: جرم مخصوص ظاهری، OM: ماده آلی، Ks: هدایت هیدرولیکی اشباع، MWD: میانگین وزنی قطر خاکدانه‌ها، PR300: مقاومت فروری در مکش ۳۰۰ هکتوپاسکال و T0 - T15000 (PWP): به ترتیب رطوبت حجمی در مکش ۰ تا ۱۵۰۰۰ هکتوپاسکال

* and ** indicate significant correlation at the 0.05 and 0.01 levels, respectively

†. In the above Table: C: clay, S: sand, Si: silt, BD: bulk density, OM: organic matter, Ks: saturated hydraulic conductivity, MWD: mean weight diameter of soil aggregates, PR300: penetration resistance at 300 hectopascals, T0 - T15000 (PWP): volumetric moisture content at 0 to 15000 hectopascals

برای مدل‌های SWRC مانند ون‌گنوختن (۱۹۸۰) و بروکس و کوری (۱۹۶۴) است (۱۰ و ۴). در صورتی که بتوان با روش‌های هوش مصنوعی قدرتمند، تابع دقیقی تنها بر اساس استفاده از مکش خاک به‌عنوان تخمین‌گر ایجاد نمود، مشکل کهن ایجاد مدل‌های مختلف SWRC در خاک‌های مختلف، مرتفع خواهد شد. البته این موضوع نیاز به پژوهش‌های بیشتر دارد. حق‌وردی و همکاران (۲۰۱۲) و جین و همکاران (۲۰۰۴) گزارش کردند که هم مدل‌های توابع انتقالی و هم توابع برازشی SWRC را می‌توان توسط یک مدل شبکه عصبی جایگزین کرد (۱۳، ۴۲). جین و همکاران (۲۰۰۴) ثابت کردند که عملکرد شبکه‌های عصبی برای برازش SWRC، مشابه و در برخی موارد حتی بهتر از عملکرد معادله ون‌گنوختن است (۴۲). حق‌وردی و همکاران (۲۰۱۲) گزارش کردند که توابع شبه‌پیوسته کارایی بیشتری نسبت به توابع پارامتریک دارند. بنابراین وقتی که بخواهیم از مجموعه داده‌هایی با جفت داده‌های محدود برای SWRC استفاده کنیم، نیازی به استفاده از یک معادله از پیش تعریف شده مانند معادله ون‌گنوختن نیست (۱۳).

نتایج نشان داد که مقادیر RMSE توابع بسط داده شده به روش رگرسیون خطی در دو مرحله آموزش و اعتبارسنجی تقریباً مشابه به دست آمد اما مقادیر آماره آکائیک در دسته داده آموزش بسیار بالاتر از دسته داده اعتبارسنجی بود. دلیل این امر استفاده از روش k-fold با $k=10$ برای تقسیم‌بندی داده‌های آموزش و اعتبارسنجی بود. بدین معنی که داده‌ها به ۱۰ قسمت تقسیم شد و ۱۰ بار مدل اجرا شد. در هر اجرا ۹ قسمت داده‌ها برای آموزش و ۱ قسمت برای اعتبارسنجی استفاده شد، به‌گونه‌ای که تمام نمونه‌ها، هر کدام یک‌بار برای اعتبارسنجی استفاده شدند. پس از هر اجرا هر سه آماره آکائیک، R^2 و RMSE برای

نتایج برآورد رطوبت توسط توابع انتقالی شبه‌پیوسته ایجاد شده با روش رگرسیون غیر خطی و روش درخت M5: نتایج برآورد منحنی نگاه‌داشت آب خاک توسط توابع انتقالی شبه‌پیوسته ایجاد شده با روش رگرسیون غیرخطی و روش درخت M5 به‌ترتیب در جدول‌های ۴ و ۵ آورده شده است. همان‌گونه که قبلاً ذکر شد، در این مطالعه ۱۳ تابع انتقالی شبه‌پیوسته با هر کدام از روش‌های رگرسیون غیرخطی و M5 ایجاد شد، که این ۱۳ تابع به سه گروه، بر اساس متغیرهای ورودی، تقسیم شدند. دقت و قابلیت اعتماد برآورد SWRC در هر گروه مورد بررسی قرار گرفت.

گروه اول (تابع شبه‌پیوسته ۱): مکش خاک به‌عنوان تنها تخمین‌گر در تنها تابع این گروه استفاده شد. روش درخت M5 برای این تابع مدلی ایجاد نکرد (جدول ۵). چون در این روش استفاده از تنها یک متغیر تخمین‌گر موجب بروز خطا در ماتریس محاسبات گردید. ولی رگرسیون غیرخطی مدلی را ایجاد کرد (جدول ۴)، که فرمول آن در زیر آورده شده است:

$$T=0.00246 \times (\log h)^3 - 0.02219 \times (\log h)^2 - 0.01630 \times (\log h) + 0.35629 \quad (4)$$

در معادله ۴، h ، مکش خاک و T رطوبت خاک در هر مکشی می‌باشد. آماره‌های خطای معادله ۴، دقت و قابلیت اعتماد قابل‌قبول این مدل را نشان می‌دهد، چرا که ضریب تعیین 0.729 و 0.718 به‌ترتیب برای مراحل آموزش و اعتبارسنجی به‌دست آمد. هرچند با دقت در شکل پراکنش مقادیر پیش‌بینی‌شده در مقابل اندازه‌گیری‌شده (شکل ۲)، مشخص می‌شود که مدل ایجاد شده در واقع یک میانگینی از رطوبت در هر مکشی را برای آن مکش لحاظ نموده است، که چندان معتبر به نظر نمی‌رسد. در واقع تابع ۱، نوعی جایگزین

این آماره، پس از آنکه از مجموع مربعات خطا لگاریتم طبیعی گرفته می‌شود، عدد حاصل در تعداد مشاهدات ضرب می‌گردد. پس هرچه تعداد مشاهدات بیشتر باشد آکائیک بیش‌تر است. چون در این پژوهش تعداد نمونه‌های آموزش در هر اجرا ۹ برابر تعداد نمونه‌های اعتبارسنجی است، بنابراین تفاوت زیادی بین آن‌ها وجود دارد.

هر دو مرحله به صورت جداگانه محاسبه شد. در نهایت میانگین آماره‌ها در ۱۰ اجرا به عنوان آماره نهایی گزارش شد. بنابراین احتمالاً این نحوه اجرای مدل باعث نزدیک شدن نتایج آماره‌های R^2 و RMSE در دو مرحله گردید، که اعتبار بالای مدل‌های ایجاد شده را نشان می‌دهد. اما علت تفاوت بالای آماره آکائیک در دو مرحله آموزش و اعتبارسنجی، فرمول مربوطه می‌باشد. چون در فرمول

جدول ۴- نتایج روش رگرسیون غیر خطی (NLR) برای سیزده تابع انتقالی شبه پیوسته در برآورد منحنی نگه‌داشت آب خاک.

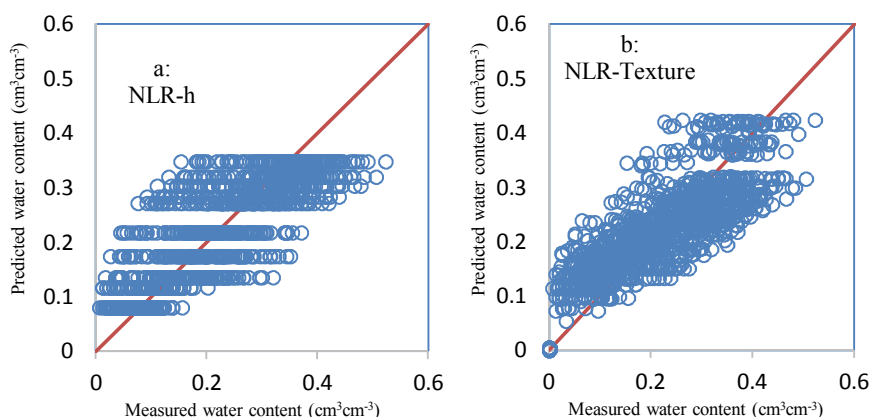
Table 4. Results of non-linear regression (NLR) method for thirteen semi-continuous pedotransfer functions in estimating soil water retention curve.

مدل مقایسه‌شونده Comparing model	اعتبارسنجی Validation			آموزش Training			توابع PTF
	AIC	R^2_{adj}	RMSE ($\text{cm}^3\text{cm}^{-3}$)	AIC	R^2_{adj}	RMSE ($\text{cm}^3\text{cm}^{-3}$)	
	-61	0.716	0.066	1950	0.727	0.065	1
1	-62	0.715	0.066	1973	0.714	0.066	2
2	-74	0.736	0.063	1851	0.739	0.063	3
3	-88	0.757	0.060	1733	0.762	0.060	4
4	-155	0.859	0.047	1108	0.858	0.046	5
5	-156	0.859	0.046	1095	0.859	0.046	6
4	-89	0.759	0.060	1720	0.763	0.060	7
4	-89	0.760	0.060	1717	0.763	0.060	8
4	-90	0.759	0.060	1705	0.764	0.059	9
4	-87	0.757	0.060	1731	0.761	0.060	10
4	-87	0.757	0.060	1732	0.760	0.060	11
4	-87	0.755	0.060	1730	0.759	0.060	12
4	-90	0.755	0.060	1694	0.761	0.059	13

† در جدول فوق RMSE: مجذور میانگین مربعات خطا، R^2 : ضریب تبیین، AIC: معیار اطلاعات آکائیک. هایلیت‌های رنگی در ستون‌های

آکائیک، تفاوت معنی‌دار تابع مربوطه با تابع مقایسه‌شونده را نشان می‌دهد

†. In the above Table: RMSE: root mean square error, R^2 : coefficient of determination, AIC: Akaike's information criterion. Colored highlights in the AIC column show the significant differences of that PTF with comparing PTF



شکل ۲- توزیع مقادیر رطوبت تخمینی در مقابل رطوبت اندازه‌گیری شده برای مرحله اعتبارسنجی توابع شبه‌پیوسته ۱ (a) و ۲ (b) ایجاد شده با روش رگرسیون غیر خطی (NLR). h: مکش ماتریک آب خاک.

Figure 2. Distribution of estimated versus measured moisture contents for the validation stage of semi-continuous pedotransfer functions 1 (a) and 2 (b) created using non-linear regression (NLR) method. †. h: soil water matric suction.

جدول ۵- نتایج روش درخت M5 برای سیزده تابع انتقالی شبه‌پیوسته در برآورد منحنی نگه‌داشت آب خاک.

Table 5. Results of M5 tree method for thirteen semi-continuous pedotransfer functions in estimating soil water retention curve.

مدل مقایسه‌شونده Comparing model	اعتبارسنجی Validation			آموزش Training			توابع PTF
	AIC	R ² _{adj}	RMSE (cm ³ cm ⁻³)	AIC	R ² _{adj}	RMSE (cm ³ cm ⁻³)	
							1
1	-131	0.830	0.051	1237	0.846	0.049	2
2	-160	0.862	0.046	584	0.911	0.037	3
3	-200	0.894	0.040	31	0.944	0.029	4
4	-376	0.972	0.020	-1635	0.986	0.015	5
5	-400	0.977	0.019	-1940	0.989	0.013	6
4	-203	0.897	0.039	15	0.944	0.029	7
4	-213	0.904	0.038	-140	0.951	0.027	8
4	-215	0.904	0.037	-195	0.953	0.027	9
4	-204	0.899	0.039	-19	0.946	0.029	10
4	-198	0.894	0.040	33	0.944	0.029	11
4	-201	0.897	0.039	-68	0.947	0.028	12
4	-219	0.907	0.037	-306	0.956	0.025	13

† در جدول فوق RMSE: مجذور میانگین مربعات خطا، R²: ضریب تبیین، AIC: معیار اطلاعات آکائیک. هایلایت‌های رنگی در ستون‌های آکائیک، تفاوت معنی‌دار تابع مربوطه با تابع مقایسه‌شونده را نشان می‌دهد.

†. In the above Table: RMSE: root mean square error, R²: coefficient of determination, AIC: Akaike's information criterion. Colored highlights in the AIC column show the significant differences of that PTF with comparing PTF

SWRC استفاده کردند در مقایسه با تابع ۲ که از کلاس بافت خاک به‌عنوان تخمین‌گر استفاده کرد، را نشان داد. اگرچه روش‌های مختلفی برای توسعه توابع پیش‌بینی SWRC براساس ورودی‌های مختلف وجود دارد، اما پیش‌بینی SWRC هم‌چنان چالشی است و موضوع پژوهش‌های بسیاری در فیزیک خاک است (۴۴). راولز و همکاران (۲۰۰۳) بهبود جزئی توابع با استفاده از اجزای بافت به‌عنوان تخمین‌گر را نسبت به استفاده از کلاس بافت خاک در برآورد SWRC را گزارش کردند (۴۵). راستگو و همکاران (۲۰۲۲) نیز نتیجه مشابهی را گزارش کردند (۱۶). اثر مفید استفاده از شاخص‌های ساختمان خاک به‌عنوان متغیرهای ورودی در افزایش دقت برآورد SWRC نیز به اثبات رسیده است (۴۶). به‌همین علت استفاده از BD موجب بهبود معنی‌دار دقت و قابلیت اعتماد برآورد SWRC گردید (جدول‌های ۴ و ۵).

چون رطوبت FC (رطوبت در مکش ۳۰۰ هکتوپاسکال) مستقیماً به توزیع اندازه منافذ مربوط بوده و یک نقطه از SWRC است، که به‌عنوان تخمین‌گر مورد استفاده قرار گرفت، به‌همین علت موجب بهبود دقت و قابلیت اعتماد مدل گردید. پژوهش‌گران بسیاری (۴۰، ۴۷، ۴۸) نشان داده‌اند که استفاده از یک یا دو نقطه SWRC در تخمین‌ها باعث افزایش دقت و قابلیت اعتماد تخمین SWRC خواهد شد.

در این پژوهش نقش PWP در بهبود برآورد SWRC از طریق توابع انتقالی شبه‌پیوسته کم‌رنگ‌تر شد. چراکه استفاده از آن در روش رگرسیون غیر خطی، مجذور میانگین مربعات خطا در مرحله آموزش و اعتبارسنجی را فقط به‌ترتیب صفر و ۲ درصد و در روش M5 به‌ترتیب ۱۳ و ۵ درصد

گروه دوم (توابع شبه‌پیوسته ۲ تا ۶): در جدول‌های ۴ و ۵ هر تابع با تابع دیگری مقایسه شده است، که تابع مقایسه‌شونده در ستون مربوطه نشان داده شده است. توابع ۲ تا ۶، همگی با تابع قبل از خود مقایسه شده‌اند (جدول‌های ۴ و ۵) و توابع ۷ تا ۱۳ نیز با تابع ۴ مقایسه شده‌اند. مقایسه معنی‌داری بر اساس آماره AIC انجام شده است. در صورت وجود تفاوت معنی‌دار سلول مربوط به آکائیک‌هایلیت شده است.

در روش رگرسیون غیر خطی استفاده از بافت خاک به‌عنوان تخمین‌گر بهبود معنی‌داری نسبت به تابع ۱ ایجاد نکرد (جدول ۴). هر چند شکل پراکنش مقادیر تخمینی رطوبت خاک در مقابل مقادیر اندازه‌گیری شده آن در تابع ۲ (شکل ۲b) توزیع بهتری را نسبت به تابع ۱ (شکل ۲a) نشان داد. بعد از تابع ۱، تابع ۲ ایجاد شده توسط هر دو روش رگرسیون غیرخطی و M5 ضعیف‌ترین تابع بود. چراکه کم‌ترین اطلاعات را وارد مدل می‌کند. غالباً دقت توابع کلاسی کم‌تر از توابع پیوسته است (۴۳). تابع ۲ در این پژوهش نیز معادل تابع کلاسی بوده ولی توابع ۳ تا ۱۳ معادل توابع پیوسته هستند.

در تمام توابع شبه‌پیوسته ۳ تا ۵ ایجاد شده توسط هر دو روش به‌ترتیب استفاده از رس و شن، جرم مخصوص ظاهری و FC بهبود معنی‌داری در دقت و قابلیت اعتماد این توابع نسبت به تابع قبل از خود نشان داد (جدول‌های ۴ و ۵). استفاده از PWP در تابع ۶، موجب بهبود معنی‌دار برآورد SWRC در روش M5 گردید، ولی در روش رگرسیون غیرخطی بهبودی ایجاد نکرد.

نتایج این پژوهش (جدول‌های ۴ و ۵) دقت و قابلیت اعتماد بالاتر توابع شبه‌پیوسته ۳ تا ۶ را که از اجزای بافت، BD و نقاطی از SWRC برای برآورد

برای برآورد نقطه‌ای و پارامتریک SWRC استفاده کرد و برتری M5 نسبت به رگرسیون غیرخطی را مشاهده نمود (۲۲). Ks نشان‌دهنده هندسه و پیوستگی منافذ بوده و وابسته به توزیع اندازه منافذ موجود در خاک است (۵۰)، به همین علت استفاده از آن موجب بهبود معنی‌دار دقت و قابلیت اعتماد برآورد SWRC گردید. میناسنی و همکاران (۲۰۰۴) نیز گزارش کردند که افزودن Ks علاوه بر سیلت، شن و رس موجب بهبود تخمین ویژگی‌های هیدرولیکی می‌شود (۵۱).

۲. از میان متغیرهای ساختمانی، Ks مؤثرترین متغیر بوده و MWD نیز تأثیر کمی در برآورد SWRC داشت، چراکه فقط در مرحله آموزش روش M5 موجب بهبود معنی‌دار برآورد SWRC گردید. این نتایج تقریباً منطبق با نتایج همبستگی میان رطوبت در مکش‌های مختلف با Ks و MWD است (جدول ۲).

۳. OM و PR300 اثر معنی‌داری در برآورد SWRC نشان ندادند. اثر PR300 بر دقت برآورد SWRC منطبق با نتایج همبستگی آن با متغیرهای خروجی است. وابستگی مقاومت فروری به مکش یا رطوبت خاک توسط پژوهش‌گران بسیاری، از جمله استوک و داوونز (۲۰۰۸)، گزارش گردیده است (۵۲). بنابراین، بر اساس روابط گزارش شده، مقاومت فروری در هر مکش، متأثر از رطوبت خاک است و می‌توان از آن برای تخمین رطوبت خاک استفاده نمود. اما علت این‌که در این مطالعه، استفاده از مقاومت فروری در برآورد منحنی نگه‌داشت آب خاک تأثیر قابل‌توجهی در بهبود برآوردهای مدل نداشت، اندازه‌گیری مقاومت فروری در مکش ۳۰۰ هکتوپاسکال بود. دلیل این امر تغییرپذیری کم و ناچیز مقاومت فروری خاک در مکش‌های کم خاک است (۵۲)، بنابراین

کاهش داد. این نتیجه از یک‌سو کاهش کارایی PWP در برآورد SWRC و از سوی دیگر توانمندی روش M5 در استخراج الگوی میان داده‌ها را نشان می‌دهد. تیت جی و تاپکن هنریش (۱۹۹۳) نیز عدم بهبود دقت برآوردهای نقاطی از SWRC با استفاده از PWP به‌عنوان تخمین‌گر را گزارش نمودند (۳۰).

گروه سوم (توابع شبه‌پیوسته ۷ تا ۱۳): در این گروه ترکیب‌های مختلفی از OM و متغیرهای ساختمان خاک به‌عنوان تخمین‌گرهای اضافی علاوه بر مکش خاک، رس، شن و جرم مخصوص ظاهری وارد مدل شدند و نتایج آن‌ها با نتایج تابع ۴ مقایسه شد (جدول‌های ۴ و ۵). در روش رگرسیون غیر خطی، استفاده از هیچ‌کدام از متغیرهای اضافی این گروه بهبود معنی‌داری در توابع انتقالی شبه‌پیوسته برآورد SWRC ایجاد نکرد. اما در روش M5، هر تابعی که در آن از Ks به‌عنوان تخمین‌گر استفاده شد، بهبود معنی‌داری نسبت به تابع ۴ در هر دو مرحله آموزش و اعتبارسنجی نشان داد. چراکه مدل‌های پیش‌بینی وابسته به روش ایجاد تابع می‌باشند (۴۹). هم‌چنین هر تابعی که از MWD به‌عنوان تخمین‌گر استفاده نمود، فقط در مرحله آموزش بهبود معنی‌داری نسبت به تابع ۴ نشان داد. اما استفاده از OM و PR300 به‌تنهایی بهبودی در برآورد SWRC نسبت به تابع ۴ نداشتند. بنابراین بر اساس نتایج این پژوهش می‌توان به چند نکته اشاره نمود: ۱. متغیرهای ساختمانی استفاده شده در این گروه، اثری به‌مراتب کم‌تر از متغیرهای بافت خاک و BD در برآورد SWRC داشتند. به‌همین علت این اثر توسط روش رگرسیون غیرخطی قابل تشخیص نبود، اما روش M5 با الگوریتم قوی‌تر خود قادر به تشخیص اثر کم‌تر هم بود. راستگو (۲۰۲۰) از روش‌های مختلفی از جمله M5 و رگرسیون غیرخطی

قابلیت اعتماد تخمین‌ها به‌ترتیب در هر دو مرحله آموزش و اعتبارسنجی در روش M5 به‌طور معنی‌دار و قابل توجهی بیشتر از رگرسیون غیر خطی بود. صمدیان‌فرد و همکاران (۲۰۱۴) برتری مدل درخت M5 در مقایسه با برنامه‌ریزی بیان ژن در مدل‌سازی حرکت آب در خاک را گزارش کردند (۵۵). ذاکری‌نیا و قربانی (۱۳۹۲) نیز از روش M5 برای برآورد رطوبت در مکش‌های مختلف، مشابه این پژوهش استفاده کردند. آن‌ها برای توابعی با متغیرهای ورودی دقیقاً معادل با توابع ۳ و ۴ به‌ترتیب ضریب تعیین و RMSE برابر با ۰/۸۶۶، ۰/۰۳۸ و ۰/۸۹۳، ۰/۰۳۴ به‌دست آوردند، که ضعیف‌تر از نتایج مرحله آموزش و معادل نتایج مرحله اعتبارسنجی این پژوهش بود (۲۰). راستگو و همکاران (۲۰۲۲) از چهار روش تابع پایه شعاعی، ماشین‌های یادگیری بینهایت، مدیریت گروهی داده‌ها و درخت M5 برای برآورد پارامتریک SWRC استفاده کردند و عملکرد مشابه و یا بهتر درخت M5 نسبت به دیگر روش‌ها را گزارش کردند (۱۶).

علت برتری مدل M5 این است که این الگوریتم ابتدا یک درخت ثابت و یکنواخت را می‌سازد و یک مدل رگرسیون خطی را به داده‌های هر گره برگ برازش داده و مقدار خطا را محاسبه می‌کند. معیار انتخاب بر اساس میزان خطا در هر گره می‌باشد. به عبارت دیگر، استفاده از واریانس یا انحراف استاندارد از مقادیر پاسخ یک رویکرد معقول برای تقسیم گره‌ها است (۵۶). الگوریتم درخت M5 بر اساس کاهش خطای حداکثری از میان تمام انشعابات^۱ ممکن، انشعاب با کم‌ترین خطا را انتخاب می‌کند (۵۷). با توجه به این‌که با این روند یک درخت بزرگ ایجاد می‌شود که احتمال بیش‌برازشی وجود

پارامتر مناسبی برای برآورد تغییرات رطوبت خاک در مکش‌های مختلف نبود. عدم وجود همبستگی معنی‌دار آن با مقدار رطوبت خاک در بسیاری از مکش‌ها (جدول ۳) نیز این نتیجه را تأیید می‌کند.

ولی در خصوص OM نتیجه حاصله عکس همبستگی بالا و معنی‌دار آن با رطوبت در مکش‌های مختلف است (جدول ۳). با توجه به این‌که OM دارای همبستگی مثبت با رطوبت در مکش‌های مختلف بوده (جدول ۳) و موجب بهبود ساختمان خاک و افزایش خاکدانه‌سازی می‌گردد (۵۳)، بنابراین انتظار بر این است که برآورد SWRC را بهبود دهد. شاید همبستگی آن با متغیرهای دیگری که قبلاً در مدل وجود دارند، عامل این نتیجه باشد. OM دارای همبستگی منفی و معنی‌دار با شن (**۰/۳۳۹-) و BD (**۰/۴۱۲-) بود. چون شن و BD قبلاً به‌عنوان تخمین‌گر وارد مدل شده‌اند، به‌همین علت استفاده از OM اطلاعات اضافی وارد مدل نکرده و بهبود معنی‌داری در دقت برآورد SWRC ایجاد نکرد. در مورد تأثیر OM بر تخمین SWRC نتیجه مشابهی توسط توماسلا و همکاران (۲۰۰۳) نیز گزارش شده است (۵۴). حق‌وردی و همکاران (۲۰۱۸) نیز گزارش کردند که در توابع انتقالی شبه‌پیوسته ایجاد شده توسط آن‌ها، OM کم‌ترین تأثیر را در بهبود برآورد SWRC داشت (۱۷).

مقایسه دقت و قابلیت اعتماد برآورد SWRC با دو روش رگرسیون غیر خطی (NLR) و درخت M5: برای مقایسه دقت و قابلیت اعتماد برآورد SWRC با دو روش، توابع با متغیرهای ورودی یکسان به‌صورت دویبه‌دو بر اساس آماره آکائیک مقایسه شدند (جدول‌های ۴ و ۵). روش M5 برتری خود را در تمام ۱۲ تابع شبه‌پیوسته نشان داد. چراکه دقت و

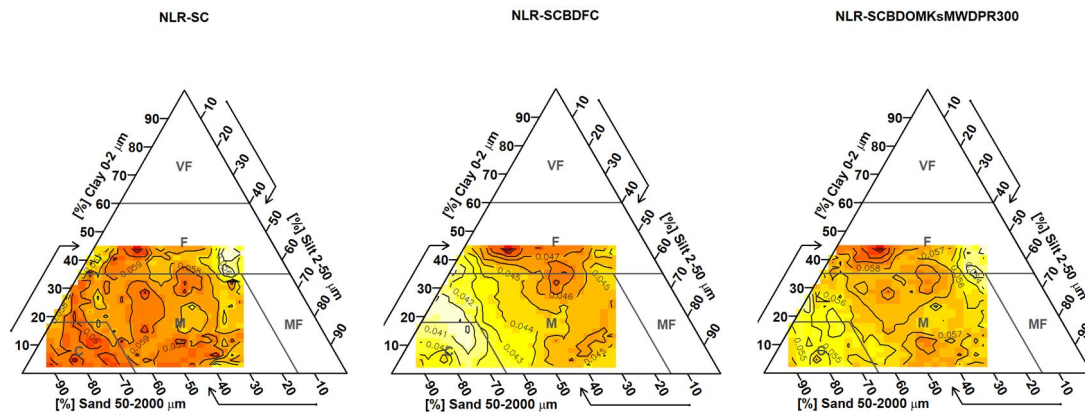
دارد، در نتیجه، در مرحله دوم درخت هرس می‌شود. اصلی‌ترین مزیت روش درخت M5 این است که تعداد بسیار بیش‌تر درخت (مدل) کوچک‌تر از درختان رگرسیون CART می‌سازد که به صورت طبیعی شامل بسیاری از متغیرها نمی‌شوند، ولی این نوع مدل‌ها دقیق‌تر هستند. روش M5 وابستگی بین متغیرهای ورودی را در نظر گرفته و در واقع ترکیبی از یادگیری درخت تصمیم‌گیری با رگرسیون خطی را ایجاد می‌کند (۵۸).

توزیع خطای توابع انتقالی شبه‌پیوسته بر روی مثلث بافت خاک: توزیع مجذور میانگین مربعات خطای مرحله اعتبارسنجی روش‌های رگرسیون غیر خطی و درخت M5 بر روی مثلث بافت خاک به‌ترتیب در شکل‌های ۳ و ۴ برای توابع ۳، ۵ و ۱۳ نشان داده شده است. انتخاب توابع ۳، ۵ و ۱۳ بر اساس گروه‌بندی توابع بر مبنای نوع متغیرهای ورودی انجام شد (جدول ۱). در واقع از هر گروه از توابع، یک تابع برای نمایش روند تغییرات خطا انتخاب شد. نتایج نشان داد که توزیع مجذور میانگین مربعات خطا بر روی مثلث بافت خاک بر اساس نکات زیر شکل گرفته است:

۱. توزیع خطای توابع انتقالی شبه‌پیوسته برآورد SWRC بر روی مثلث بافت خاک، وابستگی خاصی به بافت خاک نشان نداد. در تمام توابع هر دو روش هیچ بافت خاکی یافت نشد که روند ثابتی از خطا را در تمام توابع نشان دهد. این نکته نشان‌دهنده عدم تأثیر بافت خاک بر میزان خطای توابع شبه‌پیوسته ایجاد شده در این پژوهش برای برآورد SWRC بود. پژوهش‌گران دیگری (۲۱، ۲۲) نیز گزارش کردند که،

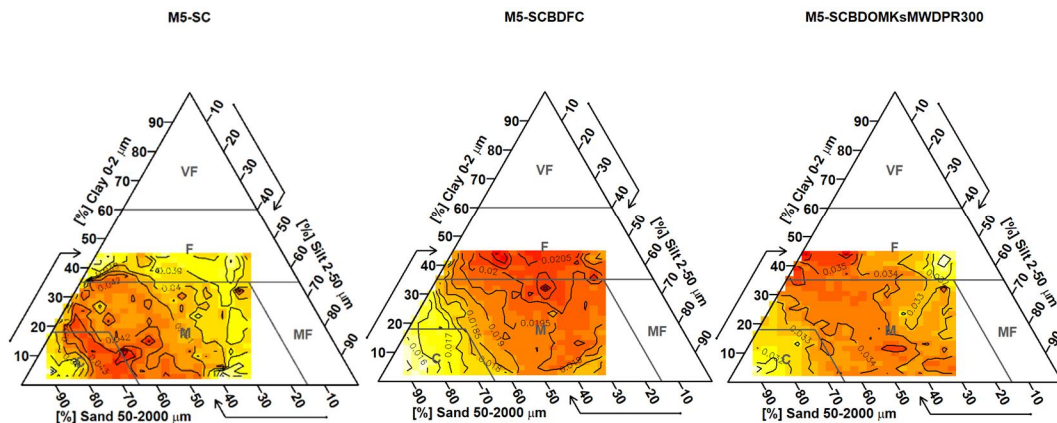
توزیع خطای برآورد خصوصیات هیدرولیکی بر روی مثلث بافت خاک وابسته به بافت خاک نیست. ۲. توزیع خطا در توابع مشابه دو روش کاملاً متفاوت بود. به‌عنوان مثال تابع ۳ که از مکش خاک و اجزای بافت خاک برای برآورد SWRC استفاده کرده است، در روش رگرسیون غیر خطی در بافت شنی بیش‌ترین خطا را نشان داد و به‌تدریج با حرکت به سمت ضلع سیلت مقدار خطا کاهش یافت. ولی در روش درخت M5، در بافت شنی خطا کم بود، در بافت‌های درشت تا متوسط بیش‌ترین مقدار خطا و سپس با حرکت به سمت ضلع سیلت مجدداً خطا کاهش یافت. این نتیجه نشان‌دهنده تأثیر روش ایجاد توابع بر توزیع خطا در کلاس‌های بافتی مختلف می‌باشد.

در هر روش متغیرهای ورودی مشابه توزیع خطای مشابهی ایجاد کردند. بر این اساس توزیع خطای توابع ایجاد شده بر روی مثلث بافت خاک به سه گروه قابل تقسیم بود. گروه اول شامل توابع ۱ تا ۴ بود، که در این توابع ترکیب‌های متفاوت اطلاعات بافت، مکش آب خاک و BD به‌عنوان تخمین‌گر استفاده شدند. گروه دوم، توابع ۵ و ۶ بودند که از FC و PWP نیز به‌عنوان تخمین‌گر استفاده کردند. گروه سوم توابع ۷ تا ۱۳ بودند که از اطلاعات ساختمانی به‌عنوان تخمین‌گرهای مضاعف استفاده کردند. لازم به ذکر است که این گروه‌بندی در هر دو روش قابل تشخیص بود، ولی در توابع مشابه الگوی تغییر در دو روش ثابت نبود. بنابراین توزیع خطا بر روی مثلث بافت خاک وابسته به نوع متغیرهای ورودی و روش ایجاد توابع است.



شکل ۳- توزیع مجذور میانگین مربعات خطای ($\text{cm}^3 \text{cm}^{-3}$) مرحله اعتبارسنجی روش رگرسیون غیرخطی (NLR) بر روی مثلث بافت خاک برای توابع شبه پیوسته (به ترتیب از چپ به راست) ۳، ۵ و ۱۳. در داخل مثلث‌ها VF: خیلی ریز، F: ریز، M: متوسط، MF: متوسط ریز و C: درشت. علائم اختصاری مذکور در بالای مثلث‌ها: C: مقدار رس، S: مقدار شن، BD: جرم مخصوص ظاهری، OM: ماده آلی، Ks: هدایت هیدرولیکی اشباع، MWD: میانگین وزنی قطر خاکدانه‌ها، PR300: مقاومت فروری در مکش ۳۰۰ هکتوپاسکال، FC: رطوبت ظرفیت زراعی (مکش ۳۰۰ هکتوپاسکال).

Figure 3. Distribution of mean squared error ($\text{cm}^3 \text{cm}^{-3}$) of non-linear regression (NLR) method at the validation stage on soil texture triangle for semi-continuous pedotransfer functions of 3, 5 and 13, from left to right, respectively. Inside the triangles, VF: Very Fine, F: Fine, M: Medium, MF: Medium Fine, and C: Coarse. The abbreviations mentioned at the top of the triangles; C: clay, S: sand, BD: bulk density, OM: organic matter, Ks: saturated hydraulic conductivity, MWD: mean weight diameter of soil aggregates, PR300: penetration resistance at 300 hectopascals, FC: volumetric moisture content at field capacity (300 hectopascals).



شکل ۴- توزیع مجذور میانگین مربعات خطای ($\text{cm}^3 \text{cm}^{-3}$) مرحله اعتبارسنجی روش درخت M5 بر روی مثلث بافت خاک برای توابع شبه پیوسته (به ترتیب از چپ به راست) ۳، ۵ و ۱۳. در داخل مثلث‌ها VF: خیلی ریز، F: ریز، M: متوسط، MF: متوسط ریز و C: درشت. علائم اختصاری مذکور در بالای مثلث‌ها: C: مقدار رس، S: مقدار شن، BD: جرم مخصوص ظاهری، OM: ماده آلی، Ks: هدایت هیدرولیکی اشباع، MWD: میانگین وزنی قطر خاکدانه‌ها، PR300: مقاومت فروری در مکش ۳۰۰ هکتوپاسکال، FC: رطوبت ظرفیت زراعی (مکش ۳۰۰ هکتوپاسکال).

Figure 4. Distribution of mean squared error ($\text{cm}^3 \text{cm}^{-3}$) of M5 tree method at the validation stage on soil texture triangle for semi-continuous pedotransfer functions of 3, 5 and 13, from left to right, respectively. Inside the triangles, VF: Very Fine, F: Fine, M: Medium, MF: Medium Fine, and C: Coarse. The abbreviations mentioned at the top of the triangles; C: clay, S: sand, BD: bulk density, OM: organic matter, Ks: saturated hydraulic conductivity, MWD: mean weight diameter of soil aggregates, PR300: penetration resistance at 300 hectopascals, FC: volumetric moisture content at field capacity (300 hectopascals).

نتیجه‌گیری کلی

باعث ایجاد شفافیت در روابط خواهد شد. عملکرد روش درخت M5 در ایجاد توابع انتقالی شبه‌پیوسته، به‌طور معنی‌داری بهتر از رگرسیون غیر خطی بود. روش M5 با الگوریتم قوی‌تر خود قادر به تشخیص برخی الگوهای روابط میان متغیرهای ورودی و خروجی گردید که رگرسیون غیرخطی در تشخیص آن‌ها ناتوان بود. بنابراین استفاده از روش M5 برای ایجاد توابع انتقالی شبه‌پیوسته توصیه می‌گردد. توزیع مجذور میانگین مربعات خطای مرحله اعتبارسنجی روش‌های رگرسیون غیرخطی و درخت M5 بر روی مثلث بافت خاک، نشان‌دهنده عدم وابستگی خطا به بافت خاک بود. در مقابل توزیع خطا به‌شدت وابسته به روش ایجاد توابع شبه‌پیوسته و متغیرهای ورودی بود. هرچند این نکته نیاز به بررسی‌های بیشتر دارد. روش درخت M5 با استفاده از FC، اجزای بافت خاک و متغیرهای ساختمانی مانند BD، Ks و MWD برای برآورد SWRC مناسب بود.

سیاسگذاری

این پژوهش با حمایت مادی و معنوی دانشگاه بوعلی‌سینا، همدان انجام شده است.

استفاده تنها از مکش خاک به‌عنوان تخمین‌گر، باعث برآورد قابل‌قبول، اما نه‌چندان قوی، برای رطوبت در هر مکشی گردید. در نتیجه، شاید بتوان با روش‌های هوش مصنوعی قدرتمند مدل مناسبی برای SWRC به‌دست آورد، که در این‌صورت دیگر نیازی به استفاده از مدل‌های مرسوم SWRC نخواهد بود. استفاده از هر دو مجموعه متغیرهای بافتی مانند توزیع اندازه ذرات خاک و ساختمانی مانند BD، Ks و MWD موجب بهبود برآورد SWRC گردید. ولی همه متغیرهای ساختمانی تأثیر معنی‌داری در بهبود برآورد SWRC نداشتند. نکته دارای اهمیت در خصوص متغیرهای ساختمانی این است که، باید متغیرهایی انتخاب شوند که تا حد امکان شاخصی از توزیع اندازه منافذ باشند. بنابراین متغیرهایی که ارتباط قوی با توزیع اندازه منافذ نداشته باشند، تأثیر معنی‌داری در عملکرد توابع انتقالی برآورد SWRC نخواهند داشت. تأثیر FC در مقایسه با PWP، در برآورد SWRC بیشتر بود. بنابراین باید از نقاط نزدیک به نقطه عطف SWRC برای برآورد آن استفاده نمود. البته پژوهش بیشتر در این خصوص

منابع

1. Yuan, S., Liu, X., & Buzzi, O. (2021). A microstructural perspective on soil collapse. *Geotechnique*, 71(2), 132-140. doi: 10.1680/jgeot.18.P.256.
2. Rawls, W. J., Brakensiek, C. L., & Saxton, K. E. (1982). Estimation of soil water properties. *Transactions - American Society of Agricultural Engineers*, 25(5), 1316-1320. doi: 10.13031/2013.33720.
3. Rawls, W. J., Nemes, A., & Pachepsky, Y. (2004). Effect of soil organic carbon on soil hydraulic properties. p. 95-114, In: Y. Pachepsky & W.J. Rawls, Editors. *Developments in Soil Science*, Elsevier. doi: [https://doi.org/10.1016/S0166-2481\(04\)30006-1](https://doi.org/10.1016/S0166-2481(04)30006-1).
4. Kuzmanovski, V., Trajanov, A., Leprince, F., Džeroski, S., & Debeljak, M. (2015). Modeling water outflow from tile-drained agricultural fields. *Science of The Total Environment*, 505, 390-401. doi: <https://doi.org/10.1016/j.scitotenv.2014.10.009>.
5. Vereecken, H., Amelung, W., Bauke, S. L., Bogaen, H., Brüggemann, N., Montzka, C., Vanderborght, J., Bechtold, M., Blöschl, G., Carminati, A., Javaux, M., Konings, A. G., Kusche, J., Neuweiler, I., Or, D., Steele-Dunne, S., Verhoef, A., Young, M., & Zhang, Y. (2022). Soil hydrology in the Earth system. *Nature Reviews Earth & Environment*, 3(9), 573-587. doi: 10.1038/s43017-022-00324-6.

6. Gupta, S., & Larson, W. (1979). Estimating soil water retention characteristics from particle size distribution, organic matter percent, and bulk density. *Water Resources Research*, 15(6), 1633-1635. doi: <https://doi.org/10.1029/WR015i006p01633>.
7. Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M., & Genuchten, M. T. (2010). Using pedotransfer functions to estimate the van Genuchten-Mualem soil hydraulic properties: A review. *Vadose Zone Journal*, 9(4), 795-820. doi: <https://doi.org/10.2136/vzj2010.0045>.
8. Zhang, Y., & Schaap, M. G. (2017). Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *Journal of Hydrology*, 547, 39-53. doi: <https://doi.org/10.1016/j.jhydrol.2017.01.004>.
9. Weihermüller, L., Lehmann, P., Herbst, M., Rahmati, M., Verhoef, A., Or, D., Jacques, D., & Vereecken, H. (2021). Choice of Pedotransfer Functions Matters when Simulating Soil Water Balance Fluxes. *Journal of Advances in Modeling Earth Systems*, 13(3), e2020MS002404. doi: <https://doi.org/10.1029/2020MS002404>.
10. 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. doi: <https://doi.org/10.2136/sssaj1980.03615995004400050002x>.
11. Wösten, J., Pachepsky, Y. A., & Rawls, W. (2001). Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251(3), 123-150. doi: [https://doi.org/10.1016/S0022-1694\(01\)00464-4](https://doi.org/10.1016/S0022-1694(01)00464-4).
12. Minasny, B., & McBratney, A. B. (2002). The neuro-m method for fitting neural network parametric pedotransfer functions. *Soil Science Society of America Journal*, 66(2), 352-361. doi: <https://doi.org/10.2136/sssaj2002.3520>.
13. Haghverdi, A., Cornelis, W. M., & Ghahraman, B. (2012). A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. *Journal of Hydrology*, 442-443, 46-54. doi: <https://doi.org/10.1016/j.jhydrol.2012.03.036>.
14. Haghverdi, A., Öztürk, H. S., & Cornelis, W. M. (2014). Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. *Geoderma*, 226-227, 31-38. doi: <https://doi.org/10.1016/j.geoderma.2014.02.026>.
15. Li, Y., & Vanapalli, S. K. (2022). Prediction of soil-water characteristic curves using two artificial intelligence (AI) models and AI aid design method for sands. *Canadian Geotechnical Journal*, 59(1), 129-143. doi: [10.1139/cgj-2020-0562](https://doi.org/10.1139/cgj-2020-0562).
16. Rastgou, M., Bayat, H., Mansoorizadeh, M., & Gregory, A. S. (2022). Estimating Soil Water Retention Curve by Extreme Learning Machine, Radial Basis Function, M5 Tree and Modified Group Method of Data Handling Approaches. *Water Resources Research*, 58(4), e2021WR031059. doi: <https://doi.org/10.1029/2021WR031059>.
17. Haghverdi, A., Öztürk, H. S., & Durner, W. (2018). Measurement and estimation of the soil water retention curve using the evaporation method and the pseudo continuous pedotransfer function. *Journal of Hydrology*, 563, 251-259. doi: <https://doi.org/10.1016/j.jhydrol.2018.06.007>.
18. Nguyen, P. M., Haghverdi, A., de Pue, J., Botula, Y. D., Le, K. V., Waegeman, W., & Cornelis, W. M. (2017). Comparison of statistical regression and data-mining techniques in estimating soil water retention of tropical delta soils. *Biosystems Engineering*, 153, 12-27. doi: <https://doi.org/10.1016/j.biosystemseng.2016.10.013>.
19. Pachepsky, Y. A., & Rawls, W. (2003). Soil structure and pedotransfer functions. *European Journal of Soil Science*, 54(3), 443-452. doi: <https://doi.org/10.1046/j.1365-2389.2003.00485.x>.

20. Zakerinia, M., & Ghorbani, K. (2013). Feasibility of decision tree application (M5 model) for determining soil moisture characteristic curve from easily available soil parameters. *Journal of Water and Soil Conservation*, 20(5), 221-230. doi: [20.1001.1.23222069.1392.20.5.14.0](https://doi.org/10.1001.1.23222069.1392.20.5.14.0). [In Persian]
21. Pachepsky, Y., & Rawls, W. J. 2004. Development of pedotransfer functions in soil hydrology. Vol. 30. p, 512, Elsevier.
22. Rastgou, M. (2020). Comprehensive comparison of the methods of developing pedotransfer functions (PTFs) and development of new algorithms to predict soil water retention curve (SWRC) and soil hydraulic conductivity curve (SHCC). in *Soil Science and Enguneering*. 2020, Bu-Ali Sina University: Hamedan, Iran. p. 363. [In Persian]
23. Gee, G. W., & Or, D. (2002). Particle-Size Analysis. p, 225-295. In: Warren, A.D. (ed) *Methods of Soil Analysis. Part 4. Physical Methods*. Soil Science Society of America Inc, Madison.
24. Grossman, R., & Reinsch, T. (2002). Bulk Density and Linear Extensibility. p. 201-228, In: J.H. Dane, Topp, G.C., Editor. *Methods of Soil Analysis: Part 4 Physical Methods*, Soil Science Society of America Inc, Madison.
25. Grossman, R. B., & Reinsch, T. G. (2002). Water Retention and Storage. p. 201–228., In: J.H. Dane, Topp, G.C., Editor. *Methods of Soil Analysis. Part 4. Physical Methods*, Soil Science Society of America, Madison.
26. Reynolds, W. D., & Elrick, D. E. (2002). Falling Head Soil Core (tank) Method. p. 809-812., In: A.D. Warren, Editor. *Methods of Soil Analysis. Part 4. Physical Methods*, Soil Science Society of America Inc, Madison.
27. Walkley, A., & Black, I. A. (1934). An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Science*, 37(1), 29-38. doi: <http://dx.doi.org/10.1097/00010694-193401000-00003>.
28. Kemper, W., & Rosenau, R. (1986). Aggregate stability and size distribution. p. 425-442, In: A. Klute, Editor. *Methods of Soil Analysis: Part 1 Physical and Mineralogical Methods*, Soil Science Society of America, Inc. Madison. doi: [10.2136/sssabookser5.1.2ed](https://doi.org/10.2136/sssabookser5.1.2ed).
29. Jones, C. A. (1983). Effect of soil texture on critical bulk densities for root growth. *Soil Science Society of America Journal*, 47(6), 1208-1211. doi: <https://doi.org/10.2136/sssaj1983.03615995004700060029x>.
30. Tietje, O., & Tapkenhinrichs, M. (1993). Evaluation of pedo-transfer functions. *Soil Science Society of America Journal*, 57(4), 1088-1095. doi: <https://doi.org/10.2136/sssaj1993.03615995005700040035x>.
31. Pal, M., Singh, N. K., & Tiwari, N. K. (2012). M5 Model Tree for Pier Scour Prediction Using Field Dataset, Kimberley Structural Consulting Engineers. *Journal of Civil Engineering*, 16(6), 1079-1084. doi: [10.1007/s12205-012-1472-1](https://doi.org/10.1007/s12205-012-1472-1).
32. Wösten, J. H. M., Pachepsky, Y. A., & Rawls, W. J. (2001). Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251(3), 123-150. doi: [https://doi.org/10.1016/S0022-1694\(01\)00464-4](https://doi.org/10.1016/S0022-1694(01)00464-4).
33. Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6), 716-723. doi: [10.1109/tac.1974.1100705](https://doi.org/10.1109/tac.1974.1100705).
34. Hwang, S. I., Lee, K. P., Lee, D. S., & Powers, S. E. (2002). Models for estimating soil particle-size distributions. *Soil Science Society of America Journal*, 66(4), 1143–1150. doi: <https://doi.org/10.2136/sssaj2002.1143>.
35. Ghavami, M. S., Ayoubi, S., Mosaddeghi, M. R., & Naimi, S. (2023). Digital mapping of soil physical and mechanical properties using machine learning at the watershed scale. *Journal of Mountain Science*, 20(10), 2975-2992. doi: [10.1007/s11629-023-8056-z](https://doi.org/10.1007/s11629-023-8056-z).

36. Wilding, L. P., & Drees, L. R. (1983). Spatial variability and pedology. p. 83–116, In: L.P. Wilding, N.E. Smeck & G.F. Hall, Editors. Pedogenesis and soil taxonomy: Concepts and interactions, Elsevier, New York.
37. Brady, N. C., & Weil, R. R. 2010. Elements of the Nature and Properties of Soils. Pearson Educational International Upper Saddle River, NJ. doi: <http://lccn.loc.gov/2016008568>.
38. Hillel, D. 1998. Environmental soil physics: Fundamentals, applications, and environmental considerations. Academic press. Waltham.
39. Nemes, A., Rawls, W. J., & Pachepsky, Y. A. (2005). Influence of organic matter on the estimation of saturated hydraulic conductivity. *Soil Science Society of America Journal*, 69(4), 1330-1337. doi: <https://doi.org/10.2136/sssaj2004.0055>.
40. Pachepsky, Y., Rawls, W., Giménez, D., & Watt, J. P. C. (1998). Use of soil penetration resistance and group method of data handling to improve soil water retention estimates. *Soil and Tillage Research*, 49(1-2), 117-126. doi: [https://doi.org/10.1016/S0167-1987\(98\)00168-8](https://doi.org/10.1016/S0167-1987(98)00168-8).
41. Brooks, R. H., & Corey, A. J. (1964). Hydraulic properties of porous media. *Hydrol. Pap. 3. Colorado State Univ., Fort Collins*.
42. Jain, S. K., Singh, V. P., & Van Genuchten, M. T. (2004). Analysis of soil water retention data using artificial neural networks. *Journal of Hydrologic Engineering*, 9(5), 415-420. doi: [10.1061/ASCE1084-0699\(2004\)9:5\(415\)](https://doi.org/10.1061/ASCE1084-0699(2004)9:5(415)).
43. Saxton, K., & Rawls, W. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Science Society of America Journal*, 70(5), 1569-1578. doi: <https://doi.org/10.2136/sssaj2005.0117>.
44. 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. doi: <https://doi.org/10.1016/j.compag.2020.105502>.
45. Rawls, W., Pachepsky, Y. A., Ritchie, J., Sobecki, T., & Bloodworth, H. (2003). Effect of soil organic carbon on soil water retention. *Geoderma*, 116(1), 61-76. doi: [https://doi.org/10.1016/S0016-7061\(03\)00094-6](https://doi.org/10.1016/S0016-7061(03)00094-6).
46. Pachepsky, Y. A., & Rawls, W. J. (2003). Soil structure and pedotransfer functions. *European Journal of Soil Science*, 54(3), 443-451. doi: [10.1016/j.1365-2389.2003.00485.x](https://doi.org/10.1016/j.1365-2389.2003.00485.x).
47. Schaap, M. G., Leij, F. J., & Van Genuchten, M. T. (2001). Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*, 251(3-4), 163-176. doi: [https://doi.org/10.1016/S0022-1694\(01\)00466-8](https://doi.org/10.1016/S0022-1694(01)00466-8).
48. Wagner, B., Tarnawski, V., Wessolek, G., & Plagge, R. (1998). Suitability of models for the estimation of soil hydraulic parameters. *Geoderma*, 86(3), 229-239. doi: [https://doi.org/10.1016/S0016-7061\(98\)00040-8](https://doi.org/10.1016/S0016-7061(98)00040-8).
49. Tajik, S., Ayoubi, S., & Zeraatpisheh, M. (2020). Digital mapping of soil organic carbon using ensemble learning model in Mollisols of Hyrcanian forests, northern Iran. *Geoderma Regional*, 20, e00256. doi: <https://doi.org/10.1016/j.geodrs.2020.e00256>.
50. Kay, B. D., & VandenBygaart, A. J. (2002). Conservation tillage and depth stratification of porosity and soil organic matter. *Soil and Tillage Research*, 66(2), 107-118. doi: [https://doi.org/10.1016/S0167-1987\(02\)00019-3](https://doi.org/10.1016/S0167-1987(02)00019-3).
51. Minasny, B., Hopmans, J. W., Harter, T., Eching, S., Tuli, A., & Denton, M. (2004). Neural networks prediction of soil hydraulic functions for alluvial soils using multistep outflow data. *Soil Science Society of America Journal*, 68(2), 417-429. doi: <https://doi.org/10.2136/sssaj2004.4170>.
52. Stock, O., & Downes, N. K. (2008). Effects of additions of organic matter on the penetration resistance of glacial till for the entire water tension range. *Soil*

- and Tillage Research*, 99(2), 191-201. doi:<https://doi.org/10.1016/j.still.2008.02.002>.
53. Beare, M., Hendrix, P., & Coleman, D. (1994). Water-stable aggregates and organic matter fractions in conventional- and no-tillage soils. *Soil Science Society of America Journal*, 58(3), 777-786. doi: <https://doi.org/10.2136/sssaj1994.03615995005800030020x>.
54. Tomasella, J., Pachepsky, Y., Crestana, S., & Rawls, W. (2003). Comparison of two techniques to develop pedotransfer functions for water retention. *Soil Science Society of America Journal*, 67(4), 1085-1092. doi: <https://doi.org/10.2136/sssaj2003.1085>
55. Samadianfard, S., Nazemi, A. H., & Sadraddini, A. A. (2014). M5 model tree and gene expression programming based modeling of sandy soil water movement under surface drip irrigation. *Agriculture Science Developments*, 3, 178-190.
56. Huynh, H. (2015). Improving M5 Model Tree by Evolutionary Algorithm. *Østfold University College, Halden, Norway. Master Thesis. uri: http://hdl.handle.net/11250/293858*.
57. Kisi, O., Shiri, J., & Demir, V. (2017). Chapter 3 - Hydrological Time Series Forecasting Using Three Different Heuristic Regression Techniques. p. 45-65, In: P. Samui, S. Sekhar & V. E. Balas, Editors. *Handbook of Neural Computation*, Academic Press. doi: <https://doi.org/10.1016/B978-0-12-811318-9.00003-X>.
58. Melucci, M., & Pretto, L. (2007). PageRank: When order changes. In *European Conference on Information Retrieval*. Springer.