



تفکیک ویژگی‌های مؤثر خاک بر منحنی مشخصه رطوبتی با استفاده از درخت تصمیم‌گیری

سمیرا مصری^۱، شجاع قربانی دشتکی^۲، حسین شیرانی^۳، ابوالقاسم کامکار روحانی^۴
و حمیدرضا متقیان^۵

^۱دانشجوی دکتری گروه علوم و مهندسی خاک، دانشگاه شهرکرد، استاد گروه علوم و مهندسی خاک، دانشگاه شهرکرد،
^۲استاد گروه علوم و مهندسی خاک، دانشگاه ولی عصر رفسنجان، ^۳دانشیار گروه مهندسی نفت، معدن و ژئوفیزیک دانشگاه صنعتی شاهرود،
^۴استادیار گروه علوم و مهندسی خاک، دانشگاه شهرکرد
تاریخ دریافت: ۱۳۹۸/۱۱/۱۶؛ تاریخ پذیرش: ۱۳۹۹/۰۴/۲۲

چکیده

سابقه و هدف: تکنیک‌های محاسباتی نرم در دهه‌های اخیر به‌طور وسیعی در تحقیقات علمی و مسائل مهندسی مطالعه و به‌کار برده شده‌اند. از آنجایی‌که اندازه‌گیری ویژگی‌های هیدرولیکی با روش‌های مستقیم آزمایشگاهی دشوار، زمانبر و هزینه‌بر است و روش‌های جایگزینی را می‌طلبد که بتوان با صرف هزینه و زمان کم‌تری آن را از روی داده‌های زودیافت خاک برآورد نمود. روش‌های ناپارامتریک از روش‌های نوین برآورد ویژگی‌های هیدرولیکی خاک همانند منحنی مشخصه رطوبتی خاک می‌باشند. این پژوهش به‌منظور مطالعه کارایی روش ناپارامتریک درخت تصمیم‌گیری برای تفکیک ویژگی‌های مؤثر در تخمین پارامترهای منحنی مشخصه رطوبتی خاک انجام شد.

مواد و روش‌ها: بدین‌منظور، ۷۲ نمونه خاک از مناطق مختلف زیرحوزه مرغملک واقع در شهرستان شهرکرد از عمق ۰-۲۰ سانتی‌متری جمع‌آوری و برخی ویژگی‌ها مانند پ‌هاش، شوری، رطوبت اشباع، کربنات کلسیم معادل، ماده آلی، فراوانی نسبی ذرات، چگالی، میانگین وزنی قطر خاکدانه خشک و مرطوب، میانگین هندسی و انحراف معیار هندسی قطر ذرات خاک اندازه‌گیری شدند. هم‌چنین، منحنی مشخصه رطوبتی در مکش‌های صفر، ۱، ۳، ۵، ۱۰، ۳۰، ۵۰، ۱۰۰، ۵۰۰، ۱۰۰۰، ۱۵۰۰ کیلوپاسکال تعیین گردید و بر مدل ون‌گونوختن برازش داده شد. متغیرهای ورودی در دو سناریو (سناریو اول: پ‌هاش، EC، درصد شن و رس، ماده آلی، کربنات کلسیم، میانگین وزنی قطر خاکدانه خشک و مرطوب، چگالی و رطوبت اشباع و سناریو دوم: پ‌هاش، EC، میانگین هندسی قطر ذرات، انحراف معیار هندسی قطر ذرات، ماده آلی، کربنات کلسیم، میانگین وزنی قطر خاکدانه خشک و مرطوب، چگالی و رطوبت اشباع) به نرم‌افزار MATLAB معرفی و به‌وسیله درخت تصمیم‌گیری و تخمین‌گرهای خطای اعتبارسنجی متقاطع و بازجایگزینی مدل‌سازی شدند. معیارهای ارزیابی در مدل‌سازی شامل ضریب تبیین، جذر میانگین مربعات خطا و درصد میانگین مربعات خطا بود.

یافته‌ها: نتایج به‌دست آمده از مدل‌سازی با درخت تصمیم‌گیری نشان داد که مهم‌ترین عامل مؤثر بر رطوبت در مکش نقطه پژمردگی (PWP) رطوبت اشباع و رس می‌باشند. در هر دو سناریو متغیر هدف PWP بیش‌ترین میزان

* مسئول مکاتبه: shoja2002@yahoo.com

همبستگی (به ترتیب (۰/۸۸) و (۰/۹۱)) و کم‌ترین میزان خطا را بین سایر متغیرها داشت، و پس از آن ظرفیت زراعی دارای بیش‌ترین میزان همبستگی (۰/۸۶) در سناریو دوم بود. متغیر هدف n نیز بیش‌ترین میزان خطا و α کم‌ترین همبستگی را در دو سناریو داشت. به‌طورکلی سناریو دوم با جایگزینی میانگین هندسی و انحراف معیار قطر ذرات با درصد رس و شن عملکرد بهتری نسبت به سناریو اول داشت. آنالیز حساسیت نشان داد که رطوبت PWP به پهاش و چگالی، کربنات کلسیم و ماده آلی و رطوبت ظرفیت زراعی (FC) به انحراف معیار هندسی و میانگین وزنی قطر خاکدانه مرطوب بیش‌ترین حساسیت را داشتند.

نتیجه‌گیری: به‌طورکلی مدل‌سازی در هر دو سناریو موفق بود، اما با جایگزینی میانگین هندسی و انحراف معیار قطر ذرات به‌جای درصد رس و شن عملکرد بهتری در برآورد متغیرهای منحنی مشخصه رطوبتی در سناریو دوم به وجود آمد.

واژه‌های کلیدی: گنجایش زراعی، معادله ون گنوختن، منحنی رطوبتی

مقدمه

منحنی مشخصه رطوبتی خاک از ویژگی‌های فیزیکی خاک می‌باشد که در برآورد کیفیت فیزیکی خاک اهمیت دارد. با استفاده از این منحنی، می‌توان به چگونگی نگهداشت آب در خاک پی برد و مقدار آب قابل‌ذخیره در خاک را در پتانسیل ماتریک به‌دست آورد. منحنی نگهداشت رطوبتی خاک از مهم‌ترین ویژگی‌های محیط غیراشباع خاک است که به‌طور وسیعی در تعیین کل آب قابل‌دسترس، دور آبیاری، تخلخل مؤثر و همچنین در مطالعات زیست‌محیطی و حفاظت خاک به‌کار برده شده است (۱). منحنی نگهداشت رطوبتی به‌دلیل تنوع بسیار زیاد پارامترهای مؤثر بر آن، حتی برای یک خاک خاص بسیار پیچیده است بنابراین کاربرد هر مدل منوط به پارامترهای مؤثر بر آن و آزمایش مدل مربوطه در آن شرایط خاص می‌باشد. ژبو و یو (۲۰۰۵) طی پژوهش‌های خود تعدادی از مهم‌ترین مدل‌های ریاضیاتی برآوردکننده منحنی مشخصه رطوبتی از جمله مدل ون‌گنوختن را ارزیابی کردند (۲۴).

به‌طورکلی روش‌های متفاوتی به‌منظور تخمین پارامترهای منحنی نگهداشت رطوبتی خاک و برازش

مدل ون‌گنوختن وجود دارد که شامل روش‌های مستقیم و غیرمستقیم می‌باشد. به‌دلیل دشواری نمونه‌برداری، پرهزینه و زمان‌بر بودن روش‌های مستقیم برآورد منحنی مشخصه رطوبتی، استفاده از روش‌های غیرمستقیم و ویژگی‌های پایه‌ای و زودیافت که عمدتاً در عملیات شناسایی خاک به‌دست می‌آیند، به‌عنوان یک راه جایگزین در برآورد ویژگی‌های هیدرولیکی مورد توجه می‌باشد (۱۳). ازجمله روش‌های غیرمستقیمی که برای برآورد ویژگی‌های هیدرولیکی استفاده می‌شود شامل روش‌های کلاسیک رگرسیونی، شبکه‌های عصبی مصنوعی (ANNs) و سیستم عصبی - فازی (NF) می‌باشد که در مدل‌سازی روابط غیرخطی مقبولیت جهانی یافته‌اند. تکنیک‌های ANNs و NF مزایایی ازجمله قابلیت استفاده در داده‌های نویزی حاصل از سیستم‌های دینامیک و غیرخطی به‌ویژه هنگامی که روابط بین متغیرها به‌طور کامل قابل‌فهم نباشد - نسبت به روش‌های سنتی دارند. بهبود عملکرد مدل، توسعه سریع‌تر مدل و زمان محاسباتی کم‌تر و ارائه تکنیک خودراه‌انداز^۱ جهت اطمینان از صحت برآوردها، از مزایای این دو تکنیک محاسباتی نرم می‌باشد (۱۷).

1- Bootstrapping technique

مدل درخت تصمیم بر خلاف مدل شبکه عصبی مصنوعی به تولید قانون می‌پردازد. در ساختار درخت تصمیم، پیش‌بینی به‌دست آمده در قالب یک سری قواعد توضیح داده خواهد شد. همچنین در درخت تصمیم برخلاف شبکه عصبی مصنوعی، ضرورتی وجود ندارد که داده‌ها لزوماً به‌صورت عددی باشند (۱۲). مونکادا و همکاران (۲۰۱۴) مدل‌های درخت تصمیم را برای تخمین ویژگی‌های هیدرولیکی در دو خاک گرمسیری و معتدل استفاده کردند. آن‌ها از ویژگی‌های مورفولوژیکی خاک علاوه بر ویژگی‌های شیمیایی و فیزیکی استفاده و مشاهده کردند که کاربرد ویژگی‌های مورفولوژیکی در کنار سایر ویژگی‌های خاک تخمین ویژگی‌های هیدرولیکی را بهبود داده و درختی با یک ساختار ساده‌تر ایجاد می‌شود (۱۴). همچنین درخت تصمیم‌گیری نسبت به شبکه عصبی مصنوعی توانایی بالاتری در تخمین ویژگی‌های هیدرولیکی با پارامترهای ورودی کم‌تر دارد (۹). پژوهش‌های متعددی بر روی کارایی درخت تصمیم در زمینه علوم طبیعی صورت گرفته است و این مدل به لحاظ سادگی و ایجاد قوانین پیش‌بینی ارجحیت دارد. از طرفی مدل درخت تصمیم قادر است بدون دخالت کاربر، ورودی‌های مهم‌تر را برای ایجاد قوانین استفاده و ورودی‌های ضعیف‌تر حذف نماید.

از جمله پژوهش‌های صورت گرفته با استفاده از درخت تصمیم می‌توان به پژوهش نمس و همکاران (۲۰۰۶) اشاره کرد که با استفاده از روش درخت تصمیم و k نزدیک‌ترین همسایه نگه داشت آب در دو پتانسیل ماتریک ۳۳- و ۱۵۰۰- کیلوپاسکال را با استفاده از داده‌های زودیافت خاک تخمین زدند و برتری این روش را در امکان استفاده از داده‌های محلی بیان نمودند (۱۶). حق‌وردی و همکاران (۲۰۱۰) برای برآورد میزان رطوبت در دو نقطه ظرفیت زراعی و پژمردگی دائم از روش درخت تصمیم استفاده کردند (۵).

روش الگوریتم طبقه‌بندی درخت تصمیم از جمله روش‌های تخمین پارامترهای مؤثر منحنی رطوبتی می‌باشد. درخت‌های تصمیم ابزار قدرتمند و در عین حال رایجی هم برای دسته‌بندی و هم برای تخمین یا رگرسیون هستند. درخت تصمیم ساختاری است که برای تقسیم مجموعه بزرگی از داده‌های جمع‌آوری شده به مجموعه‌های کوچک‌تر زنجیره‌وار داده‌ها، بر اساس یک سری قوانین ساده تصمیم‌گیری به‌کار می‌روند. در هر تقسیم‌بندی متوالی، اعضای مجموعه‌های حاصل بیش از پیش به یکدیگر شباهت پیدا می‌کنند (۲۰). درخت تصمیم برای بررسی داده‌ها، به‌منظور کسب شناخت بهتر روابط موجود بین تعداد زیادی از متغیرهای ورودی کاندید شده برای یک متغیر هدف نیز به‌کار می‌رود که مسأله بسیار مهم در این مدل، انجام درست دسته‌بندی یا پیش‌بینی است (۲۰). با این‌که گونه‌های زیادی از الگوریتم‌های درخت تصمیم وجود دارد ولی همه آن‌ها از روند مشابهی پیروی می‌کنند که از تقسیم مکرر داده‌ها به گروه‌های کوچک و کوچک‌تر به‌نحوی که با توجه به متغیر هدف، هر نسل جدید گره‌ها، خالص‌تر از پیشینیان خود می‌باشند. سرانجام وقتی امکان یافتن تقسیم‌بندی‌های بیش‌تری در هیچ‌جای درخت وجود نداشته باشد، درخت تصمیم کامل رشد کرده است (۲۰). الگوریتم‌های ساخت درخت تصمیم با تلاش در یافتن آن متغیر ورودی شروع می‌شود که بهترین تقسیم‌بندی داده‌ها را در میان گروه‌های دلخواه انجام می‌دهد. روش‌های درخت تصمیم کاربرد وسیعی در اکتشاف داده‌ها، دسته‌بندی و امتیازدهی دارد (۲۰). ساختار ایده‌آل درخت در کاهش فضای طبقه‌بندی، حداقل کردن تعداد گره‌ها، کنترل در هم‌پوشانی بین کلاس‌ها و کنترل تعداد شاخه‌ها نقش مهمی ایفا می‌کند. به همین جهت طراحی یک طبقه‌بندی‌کننده در درخت تصمیم‌گیری پیچیده است (۱۱).

سالانه ۴۰۰ میلی‌متر است که بیشتر در زمستان و بهار اتفاق می‌افتد. دمای متوسط سالانه ۲۳ درجه سانتی‌گراد با میانگین حداقل ۳/۸ و میانگین حداکثر ۳۰/۷ درجه سانتی‌گراد است. شهرستان شهرکرد نیز دارای اقلیم نیمه‌مرطوب معتدل با تابستان‌های معتدل و زمستان‌های بسیار سرد است. میانگین دمای هوا در شهرکرد ۱۱/۵ درجه سانتی‌گراد است (۱۵).

روش نمونه‌برداری: ابتدا نقشه توپوگرافی با مقیاس ۱:۲۵۰۰۰ تهیه شد. نقاط نمونه‌برداری به‌طور تصادفی بر روی محدوده مطالعاتی تعیین شدند به‌طوری‌که نقاط در تمام منطقه پخش بودند. پس از تعیین مختصات جغرافیایی نقاط، موقعیت نقاط توسط دستگاه GPS در صحرا مشخص شد. سپس ۷۲ نمونه خاک از عمق صفر تا ۲۰ سانتی‌متر با استفاده از بیلچه برداشته شده و به آزمایشگاه منتقل شد. هم‌چنین نمونه‌های دست‌نخورده نیز برای تعیین چگالی به‌وسیله سیلندر با ابعاد معین از سطح خاک برداشت شدند.

اندازه‌گیری برخی ویژگی‌های فیزیکی و شیمیایی: بافت خاک به روش هیدرومتری (۳)، پ‌هاش خاک در حالت گل اشباع با استفاده از دستگاه pH متر، قابلیت هدایت الکتریکی با استفاده از دستگاه هدایت‌سنج الکتریکی در عصاره اشباع (۱۸)، کربن آلی با اکسیداسیون توسط دی‌کرومات پتاسیم (۲۳)، چگالی نمونه‌ها با روش سیلندر با ابعاد معین (۱۰)، آهک، به روش خنثی کردن مواد خنثی‌شونده با اسید کلریدریک و تیتراسیون اسید اضافی با سود، میانگین وزنی قطر خاکدانه^۱ با استفاده از الک خشک و مرطوب (۸) و رطوبت اشباع خاک به روش تهیه گل اشباع در آزمایشگاه و سپس خشک کردن در ۱۰۵ درجه به‌مدت ۲۴ ساعت تعیین شدند.

منطقه مرغملک جزو زیرحوزه‌های آبخیز زاینده‌رود با مساحت ۹۷ کیلومترمربع واقع در شمال‌غرب شهرکرد می‌باشد و دارای کاربری‌های متنوعی است و کشت و کار به‌صورت وسیع در آن انجام می‌شود، بنابراین حفظ کیفیت خاک این منطقه و شهرستان شهرکرد برای انجام کشاورزی پایدار اهمیت زیادی دارد. ویژگی‌های اثرگذار بر کیفیت خاک در مناطق مختلف با توجه به نوع کاربری و نوع مدیریت متفاوت بوده و به‌علت مدیریت نادرست، زمین‌های مناسب کشت و کار و منابع طبیعی رو به تخریب هستند. بنابراین بررسی کیفیت خاک‌ها از طریق شاخص‌های مؤثر بر کیفیت خاک همانند ویژگی‌های مؤثر بر منحنی مشخصه رطوبتی، امری ضروری می‌باشد. هم‌چنین تاکنون پژوهش‌های کمی در خصوص ویژگی‌های مؤثر بر منحنی مشخصه رطوبتی با استفاده از سناریوهای مختلف مدل‌سازی با درخت تصمیم‌گیری انجام شده است، بنابراین این مطالعه با اهداف: شناسایی عوامل مؤثر بر منحنی مشخصه رطوبتی خاک با استفاده از مدل‌سازی با درخت تصمیم و بررسی دقت مدل‌سازی و اثر ویژگی‌های خاک بر منحنی مشخصه رطوبتی در سناریوهای مختلف و بیان همبستگی بین ویژگی‌های زود یافت خاک با ویژگی‌های هیدرولیکی خاک انجام شد.

مواد و روش‌ها

این پژوهش در زیرحوزه آبخیز مرغملک (شهرستان شهرکرد، استان چهارمحال و بختیاری) انجام شد. این زیرحوزه از زیر حوضه‌های زاینده‌رود با مساحت ۹۷ کیلومترمربع (با حذف کوه‌ها) در ۵۵ کیلومتری شمال‌غرب شهرستان شهرکرد (مرکز استان چهارمحال و بختیاری) قرار دارد. زیرحوضه مرغملک در محدوده جغرافیایی، عرض ۳۰° ۲۲' ۳۰" و طول ۵۰° ۲۲' ۳۰" تا ۳۰° ۳۴' ۵۰" شهرکرد قرار دارد. بارش متوسط

1- Mean Weight Diameter

سناریو اول: پهاش، EC، درصد شن و رس، ماده آلی، کربنات کلسیم، میانگین وزنی قطر خاکدانه خشک و مرطوب، چگالی و رطوبت اشباع
 سناریو دوم: پهاش، EC، میانگین هندسی قطر ذرات، انحراف معیار هندسی قطر ذرات، ماده آلی، کربنات کلسیم، میانگین وزنی قطر خاکدانه خشک و مرطوب، چگالی و رطوبت اشباع
 لازم به ذکر است که شاخص‌های ساختمانی در مدل‌سازی برای PWP در نظر گرفته نشد چون ارتباط کم‌تری با PWP دارند و PWP بیش‌تر به بافت خاک بستگی دارد.

برای مدل‌سازی با استفاده از درخت تصمیم از تخمین‌گرهای خطای اعتبارسنجی متقاطع و بازجایگزینی استفاده شد. در این روش مجموعه داده‌های اندازه‌گیری شده به k گروه تقسیم می‌شوند. ابتدا یک گروه کنار گذاشته شده و از سایر گروه‌ها یا داده‌ها برای طراحی و تنظیم مدل استفاده می‌شود. سپس گروه کنار گذاشته شده به‌عنوان داده تست به مدل اعمال شده و خطای آن ثبت می‌شود. در مرحله بعد گروه دوم به‌عنوان تست کنار گذاشته شده و مدل‌سازی با سایر گروه‌ها انجام می‌گردد و این روند ادامه می‌یابد تا وقتی که تمام گروه‌ها یک بار به‌عنوان داده‌های تست به مدل اعمال گردد در نهایت میانگین خطای گروه‌های تست به عنوان خطای مدل‌سازی در نظر گرفته می‌شود.

شاخص‌های ارزیابی مدل: برای سنجش دقت و اعتبار مدل از شاخص‌های ضریب تبیین (R^2) و جذر میانگین مربعات خطا (RMSE) و درصد جذر میانگین مربعات خطا (RMSE%) استفاده شد.

اندازه‌گیری منحنی رطوبتی: منحنی مشخصه رطوبتی نمونه‌های خاک دست‌نخورده (با استفاده از سیلندرهایی با ابعاد 5×5) در مکش‌های (صفر، ۱، ۳، ۵، ۱۰) کیلوپاسکال توسط Sand box و مکش‌های (۳۰، ۵۰، ۱۰۰، ۵۰۰، ۱۰۰۰، ۱۵۰۰) کیلوپاسکال به‌وسیله دستگاه صفحات فشاری اندازه‌گیری شد. رطوبت وزنی نمونه‌ها در مکش‌های گفته شده تعیین گردید و درصد رطوبت حجمی از حاصلضرب چگالی ظاهری در رطوبت وزنی خاک محاسبه شد.

مدل ون گنوختن: مدل ون گنوختن یکی از مدل‌های مهم برای توصیف منحنی مشخصه رطوبتی خاک است. در این مطالعه از این معادله برای برازش داده‌های اندازه‌گیری منحنی مشخصه رطوبتی استفاده شد. این معادله به شکل زیر است:

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

که در آن، θ_s و θ_r به‌ترتیب مقدار حجمی رطوبت اشباع و باقی‌مانده خاک، h مکش خاک (cm)، α تقریباً معادل عکس پتانسیل در نقطه ورود هوا (cm^{-1}) و n و m ضرایب بدون بعد معادله هستند که به ویژگی‌های خاک بستگی دارند. برای برازش مدل ون گنوختن از نرم‌افزار RETC استفاده شد.

مدل‌سازی نقاط مهم رطوبتی خاک و پارامترهای مدل وان گنوختن: به‌منظور مدل‌سازی از درخت تصمیم باینری و از الگوریتم C&R که برای رگرسیون به‌کار می‌رود، در نرم‌افزار MATLAB استفاده شد. متغیرهای هدف شامل رطوبت نقطه پژمردگی دائم (PWP)، ظرفیت زراعی (FC) و پارامترهای α ، n و m در مدل وان گنوختن بوده‌اند. برای مدل‌سازی از دو سناریو مختلف استفاده شد. متغیرهای ورودی در دو سناریو به‌صورت زیر بودند:

به وسیله نرم افزار MATLAB ۲۰۱۶ و رسم نمودارها به وسیله نرم افزار Excel انجام شد.

نتایج و بحث

جدول ۱ توصیف آماری داده‌ها را نشان می‌دهد. در بین کمیت‌های ورودی میانگین هندسی قطر ذرات و درصد شن بالاترین ضریب تغییرات را داشتند. کمیت‌های اندازه‌گیری شده از شش کلاس بافتی مختلف خاک بود که بافت غالب خاک لومرسی سیلتی (SCL) بود. سطح معنی‌داری (توزیع پراکندگی) بین کمیت‌های ورودی نیز نشان داد که همه کمیت‌های ورودی به جز میانگین هندسی قطر ذرات (Dg) دارای پراکندگی نرمال هستند.

شیرانی و رفیع‌نژاد (۲۰۱۲) از شبکه عصبی برای تخمین رطوبت زراعی و نقطه پژمردگی استفاده و گزارش کردند بهترین نتیجه زمانی به دست می‌آید که از همه پارامترهای ورودی شامل شن، رس، ماده آلی، کربنات کلسیم و چگالی به همراه یکدیگر در یک مدل استفاده شود (۲۱). حق‌وردی و همکاران (۲۰۱۰) از پارامترهای شن، رس، چگالی، رطوبت اشباع و هدایت هیدرولیکی اشباع برای تخمین رطوبت در نقاط ظرفیت زراعی و پژمردگی دائم استفاده کردند و گزارش نمودند زمانی که پارامتر رطوبت اشباع به مدل اضافه گردید دقت شبکه در تخمین نقاط افزایش پیدا کرد (۵).

در رابطه‌های زیر N تعداد نمونه‌ها، m میانگین داده‌های واقعی، Pi مقادیر اندازه‌گیری شده و Oi مقادیر برآورد شده می‌باشد.

$$R^2 = \frac{\sum_{i=1}^n (Pi - Oi)^2}{\sum_{i=1}^n (Pi - Oi)^2} \quad (۲)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (Pi - Oi)^2} \quad (۳)$$

$$\%RMSE = (RMSE/m) * 100 \quad (۴)$$

تحلیل حساسیت: برای بررسی اهمیت متغیرهای ورودی در مدل‌سازی از تحلیل حساسیت به روش Stat Soft استفاده شد. در این روش ابتدا مدل با تمامی متغیرهای ورودی ایجاد و پس از رسیدن به بهترین عملکرد یا کم‌ترین خطا، مقدار شاخص خطا محاسبه شده و در نظر گرفته می‌شود. سپس یک متغیر ورودی معین حذف شده و مدل دوباره با سایر ویژگی‌های ورودی ایجاد می‌شود. پس از رسیدن به مناسب‌ترین ساختار و عملکرد در مدل، مقدار شاخص خطا در این حالت نیز تعیین می‌شود. از نسبت شاخص خطا در حالت دوم (حذف یک ویژگی ورودی) به حالت اول (حضور تمام ورودی‌ها)، مقدار حساسیت خروجی نسبت به متغیر ورودی مورد نظر محاسبه می‌شود (۲۲).

نرم‌افزارهای مورد استفاده: برای محاسبه شاخص‌های آماری و توزیع نرمال داده‌ها از نرم‌افزار Minitab و آزمون Kolomogorov-Smirnov استفاده شد. مدل‌سازی

جدول ۱- پارامترهای مورد استفاده در مدل‌سازی به عنوان متغیر ورودی و هدف.

Table 1. Parameters used in modeling as input and goal variables.

سطح معنی‌داری Significance level	ضریب تغییرات Coefficient of Variation	انحراف معیار Standard deviation	میانگین Mean	حداکثر Maximum	حداقل Minimum	کمیت‌ها parameters
P>0.5	2.16	0.16	7.55	7.9	7	پ‌هانش pH
P>0.9	42.72	0.30	0.7	1.69	0.27	شوری (dS/m) Electrical conductivity
P>0.44	15.53	0.063	0.40	0.51	0.25	رطوبت اشباع (kg/kg) Saturation moisture
P>0.06	37.19	10.61	28.52	76.5	14.5	آهک (%) Calcium Carbonate
P>0.4	35.93	0.54	1.5	2.41	0.1	ماده آلی (%) Organic matter
P>0.55	10.84	0.13	1.19	1.59	0.96	چگالی Bulk density
P>0.37	19.40	5.87	30.25	44	18	رس (%) Clay percentage
P>0.12	46.29	8.25	17.81	47	4	شن (%) Sand percentage
P>0.99	22.85	0.227	0.99	1.51	0.56	میانگین وزنی قطر خاکدانه خشک Mean weight diameter of dry aggregate (mm)
P>0.78	38.97	0.150	0.38	0.8	0.12	میانگین وزنی قطر خاکدانه مرطوب Mean weight diameter of wet aggregate (mm)
P>0.05	48.7	0.01	0.02	0.064	0.01	میانگین هندسی قطر ذرات Geometric mean particle diameter (mm)
P>0.19	26	2.66	10.21	18.95	5.53	انحراف معیار هندسی قطر ذرات Geometric standard deviation of particle
P>0.2	16.64	0.03	0.23	0.33	0.15	FC
P>0.21	11.46	0.01	0.12	0.16	0.10	PWP

شاخص ضریب تبیین R^2 هرچه به یک نزدیک‌تر باشد بهتر است. مقایسه شاخص‌های ارزیابی خطا RMSE و %RMSE و ضریب تبیین در هر دو سناریو در متغیرهای مختلف نشان داد که متغیر هدف PWP دارای کم‌ترین خطا و بیش‌ترین ضریب تبیین می‌باشد. متغیر هدف η نیز بیش‌ترین میزان خطا و α کم‌ترین ضریب تبیین در بین تمامی متغیرها در هر دو

مدل‌سازی با درخت تصمیم: جدول ۲ شاخص‌های ارزیابی مدل را نشان می‌دهد. نتایج نشان داد که ضریب تبیین برای اکثر متغیرهای هدف ۸۰ درصد و یا بیش‌تر بود که اعتبار مناسب درخت تصمیم را در مدل‌سازی نشان می‌دهد. ضریب تبیین مهم‌ترین معیاری است که به کمک آن می‌توان رابطه بین دو متغیر را توضیح داد و مقداری بین صفر و یک دارد.

کردند که اگر $\%RMSE$ بین ۰ تا ۴۰ باشد تخمین خوب و مدل‌سازی قوی می‌باشد، اگر بین ۴۰ تا ۷۰ باشد متوسط و در صورتی که بالای ۷۰ باشد مدل‌سازی ضعیف است (۶). همان‌طور که در جدول ۲ مشاهده می‌شود مقدار $\%RMSE$ در این پژوهش در هر دو سناریو مقدار مطلوبی است.

سناریو داشت. $\%RMSE$ برآورد شده در سناریو دوم کم‌تر از سناریو اول است که نشان می‌دهد مدل‌سازی سناریو دوم با جایگزینی میانگین هندسی و انحراف معیار قطر ذرات به جای درصد رس و شن موفق‌تر عمل کرده است. اما به‌طور کلی اختلاف چندانی بین $\%RMSE$ در هر دو سناریو در متغیرهای هدف وجود نداشت. هنگل و هیوسانجاک (۲۰۰۶) بیان

جدول ۲- شاخص‌های ارزیابی مدل‌سازی با استفاده از درخت تصمیم.

Table 2. Comparison of training data error and testing data error.

ضریب تبیین Correlation Coefficient		خطای داده‌های آزمون Error testing data				خطای داده‌های آموزشی Error training data				هدف Goal
سناریو دوم Second scenario	سناریو اول First scenario	سناریو دوم Second scenario RMSE %RMSE	سناریو اول First scenario RMSE %RMSE	سناریو دوم Second scenario RMSE %RMSE	سناریو اول First scenario RMSE %RMSE	سناریو دوم Second scenario RMSE %RMSE	سناریو اول First scenario RMSE %RMSE			
0.91	0.88	9.14	0.011	9.64	0.0116	3.43	0.004	3.85	4.7×10^{-3}	PWP
0.86	0.83	13.99	0.032	15.34	0.036	6.19	0.014	6.61	0.015	FC
0.81	0.74	21.53	0.112	22.38	0.116	8.44	0.04	9.77	0.05	α
0.84	0.80	21.82	0.54	8.32	0.20	6.72	0.17	7.33	0.19	n
0.82	0.81	13.95	0.081	14.89	0.087	4.16	0.025	4.32	0.026	m

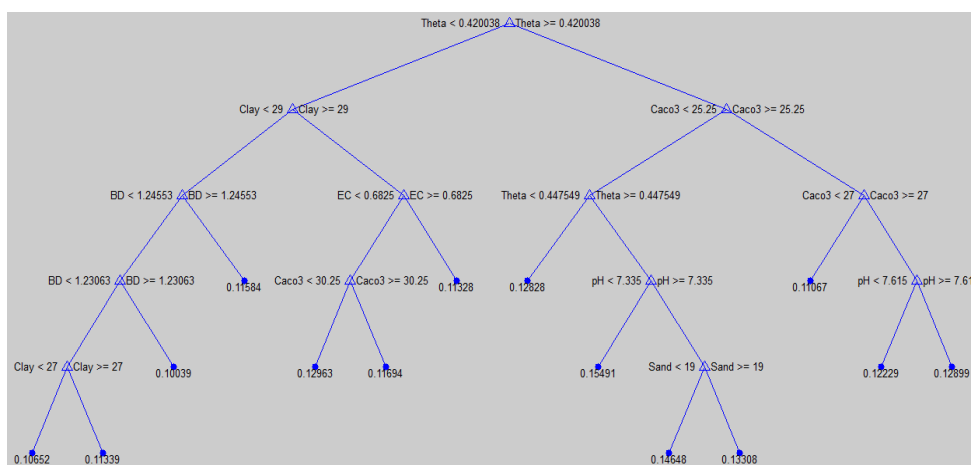
مقایسه مقدار خطا در دو مرحله آموزش و آزمون در جدول ۲ می‌توان دریافت که همانند بسیاری از مدل‌های دیگر، خطای مرحله آزمون بیش‌تر از مرحله آموزش است البته نکته قابل‌توجه این است که تفاوت خطا و دقت مدل برای دو مرحله آموزش و آزمون در هر دو سناریو مختصر است، که این به دلیل همگن بودن داده‌ها و کیفیت بالای آن‌ها می‌باشد. لازم به ذکر است که ناهمگنی داده‌ها، عملکرد الگوریتم‌های یادگیری ماشین را مختل و استفاده از آن‌ها را با مشکل مواجه می‌سازد. آگاهی از بالاترین حد توانایی مدل‌سازی خصوصیات خاک توسط مدل‌های شبکه عصبی، وابسته به آزمون ساختارهای مختلف آن‌ها و مقایسه نتایج می‌باشد. فرزاد مهر و همکاران (۲۰۱۸) با مدل‌سازی هدایت هیدرولیکی به وسیله درخت تصمیم

هاتسون و کاس (۱۹۸۷) نشان دادند که توابع انتقالی دقت بیش‌تری در تخمین رطوبت پژمردگی دائم نسبت به رطوبت ظرفیت زراعی دارد و علت این امر را ساختمان متغیر خاک (توزیع اندازه منافذ) در پژوهش خود گزارش کردند که این پارامتر تأثیر زیادی بر رطوبت ظرفیت زراعی دارد (۷). طبق نتایج به‌دست آمده در این پژوهش نیز مهم‌ترین کمیت تأثیرگذار بر ظرفیت زراعی نیز رطوبت اشباع، میانگین وزنی قطر خاکدانه مرطوب و خشک و انحراف معیار هندسی قطر ذرات بوده است. هم‌چنین مهم‌ترین پارامترهای مؤثر بر α اسیدیت، شوری، ماده آلی و میانگین وزنی قطر خاکدانه خشک بوده است که پس از n دارای بیش‌ترین درصد RMSE در هر دو سناریو و کم‌ترین ضریب تبیین در بین متغیرهای هدف بوده است. با

بیشترین تأثیر را بر رطوبت نقطه پژمردگی داشته است (۲). رضانی و همکاران (۲۰۱۱) مهم‌ترین متغیرهای تأثیرگذار در بیان میزان رطوبت در نقطه پژمردگی دائم، میزان رس و چگالی دانستند (۱۹). انشعاب پارامتر مورد بررسی باید تا جایی ادامه پیدا کند که به میزان ناخالصی مطلوب (کم‌ترین ناخالصی) برسد. پس برای رسیدن زودتر به این درجه ناخالصی، فرزندی که درجه ناخالصی کم‌تری دارد انتخاب می‌شود تا زودتر انشعاب پارامتر موردنظر متوقف و انشعاب پارامتر بعدی آغاز شود. بدین‌گونه تمامی پارامترها به ترتیب درجه اهمیتشان و بر اساس میانگین‌گیری منشعب شده تا در نهایت به گره برگ (گره‌های بدون فرزند-گره‌های پایانی) و با کم‌ترین درجه ناخالصی برسند.

بیان نمودند که در مرحله آزمون، دقت مدل برای پیش‌بینی داده‌های جدید به خوبی مرحله آموزش نیست و عملکردش کاهش یافته است (۴).

نمودار درخت تصمیم نهایی برای مدل‌سازی در شکل ۱ نشان داده شده است. به‌طورکلی در ساختار درخت انشعابات بالاتر بیانگر اهمیت بیشتر متغیر ورودی در مدل‌سازی است. طبق شکل ۱ برای متغیر هدف PWP پارامتر رطوبت اشباع مهم‌ترین پارامتر اثرگذار می‌باشد که در طبقه‌بندی در بالاترین قسمت الگوریتم، یعنی گره ریشه قرار دارد. پس از رطوبت اشباع درصد رس و آهک پارامترهای اثرگذار بر نقطه پژمردگی می‌باشند. بررسی نتایج امیرعابدی و همکاران (۲۰۱۳) نیز نشان می‌دهد که در بین پارامترهای ورودی، رس در بین سایر پارامترها



شکل ۱- طبقه‌بندی درخت تصمیم برای متغیر هدف نقطه پژمردگی در سناریو اول.

Figure 1. classification of decision tree for target variable of PWP in the first scenario.

می‌شود. طبق جدول ۳ بیشترین گره‌های فرزند و برگ مربوط به متغیرهای هدف n و m می‌باشد.

مهم‌ترین عوامل اثرگذار و آغازکننده تقسیم و تعداد گره‌های فرزند و برگ ایجاد شده در درخت تصمیم، برای متغیرهای هدف در جدول ۳ مشاهده

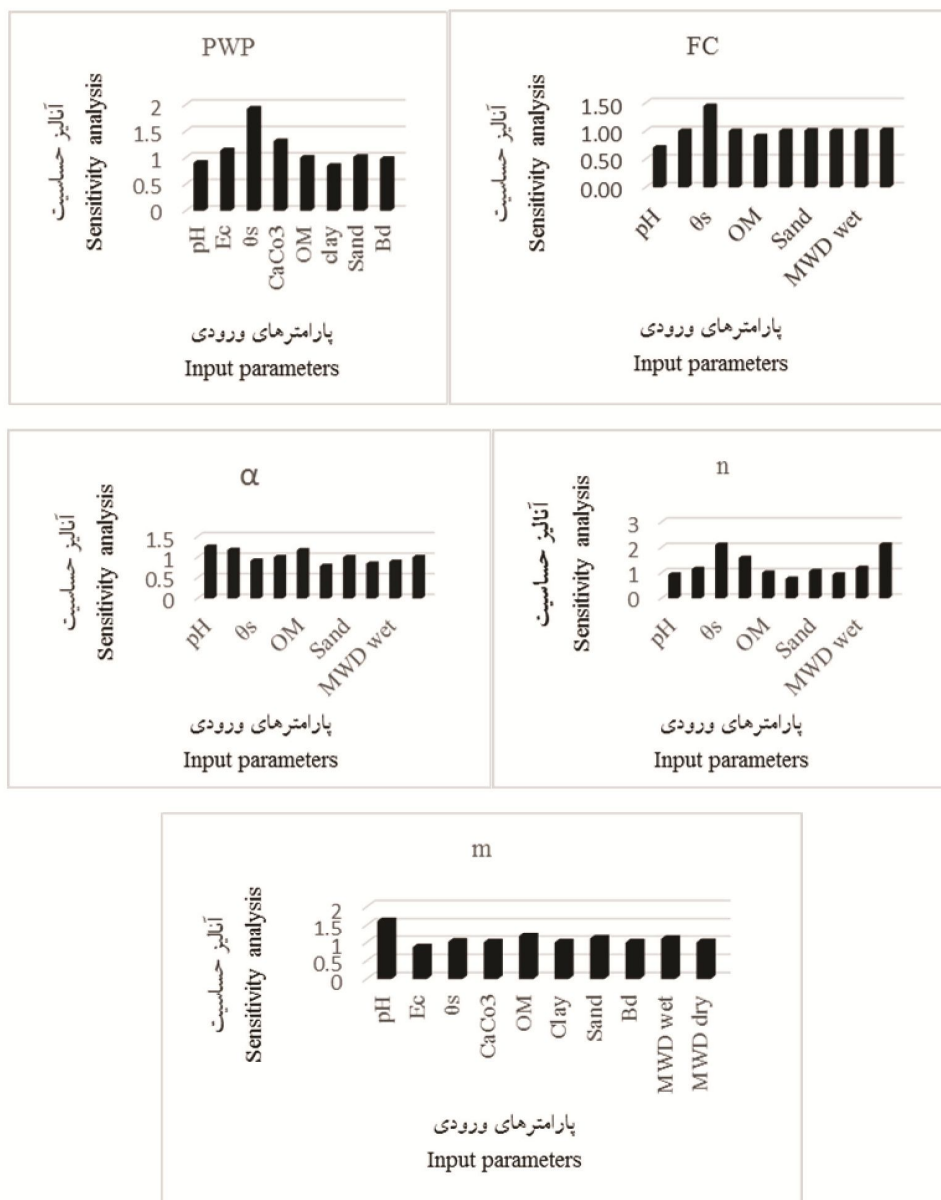
جدول ۳- نتایج مدل‌سازی درخت تصمیم برای متغیرهای هدف.

Table 3. Results of decision tree modeling for target variables.

گره برگ Leaf node	گره فرزندان Children node	مهم‌ترین عامل اثرگذار گره ریشه و آغازکننده تقسیم The most important factor affecting the root node and initiator of the division	سناریوها Scenarios	متغیرهای هدف Goal variables
14	12	رس و رطوبت اشباع Clay and saturated moisture	سناریو اول First scenario	PWP
16	14	میانگین هندسی قطر ذرات و رطوبت اشباع Geometric mean particle diameter and saturated moisture	سناریو دوم Second scenario	
13	11	رطوبت اشباع Saturated moisture	سناریو اول First scenario	FC
13	11	رطوبت اشباع Saturated moisture	سناریو دوم Second scenario	
14	12	رس و رطوبت اشباع Clay and saturated moisture	سناریو اول First scenario	α
13	11	رطوبت اشباع و شوری Saturated moisture and salinity	سناریو دوم Second scenario	
16	14	رطوبت اشباع و میانگین وزنی قطر خاکدانه مرطوب Saturated moisture and mean weight diameter of wet aggregate	سناریو اول First scenario	n
16	14	میانگین وزنی قطر خاکدانه خشک و شوری Mean weight diameter of dry aggregate and salinity	سناریو دوم Second scenario	
18	16	رطوبت اشباع و میانگین وزنی قطر خاکدانه مرطوب Saturated moisture and mean weight diameter of wet aggregate	سناریو اول First scenario	m
18	16	رطوبت اشباع و میانگین وزنی قطر خاکدانه مرطوب Saturated moisture and mean weight diameter of wet aggregate	سناریو دوم Second scenario	

په‌هاش، ماده آلی، شوری، میانگین وزنی قطر خاکدانه خشک کمیت‌های مؤثر بر α و په‌هاش، ماده آلی و درصد شن برای کمیت m اهمیت بیشتری داشته‌اند. شیرانی (۲۰۱۷) بیان نمود اگر مقدار ضریب حساسیت برای یک متغیر ورودی، بیش‌تر از یک باشد، آن متغیر سهم قابل‌توجهی در عملکرد مدل و خروجی آن دارد. اگر این نسبت از یک کم‌تر باشد، بدین معنی است که خطا در غیاب ویژگی‌های ورودی، کم‌تر از خطا در حضور آن است. بنابراین، این متغیر نه تنها اثر مثبت بر دقت مدل ندارد، بلکه تأثیر منفی نیز دارد (۲۲).

شکل‌های ۲ و ۳ اهمیت متغیرهای ورودی در مدل‌سازی با درخت تصمیم را نشان می‌دهند. همانطور که در شکل ۲ مشاهده می‌شود در سناریوی اول مهم‌ترین متغیرهای تأثیرگذار بر نقطه پژمردگی دائم، رطوبت اشباع، آهک و هدایت الکتریکی است. برای رطوبت ظرفیت زراعی، متغیرهای رطوبت اشباع و میانگین وزنی قطر خاکدانه در حالت مرطوب و خشک و درصد شن اهمیت بیش‌تری دارند. میانگین وزنی قطر خاکدانه در حالت خشک و رطوبت اشباع نیز مهم‌ترین متغیرهای ورودی برای n می‌باشند.

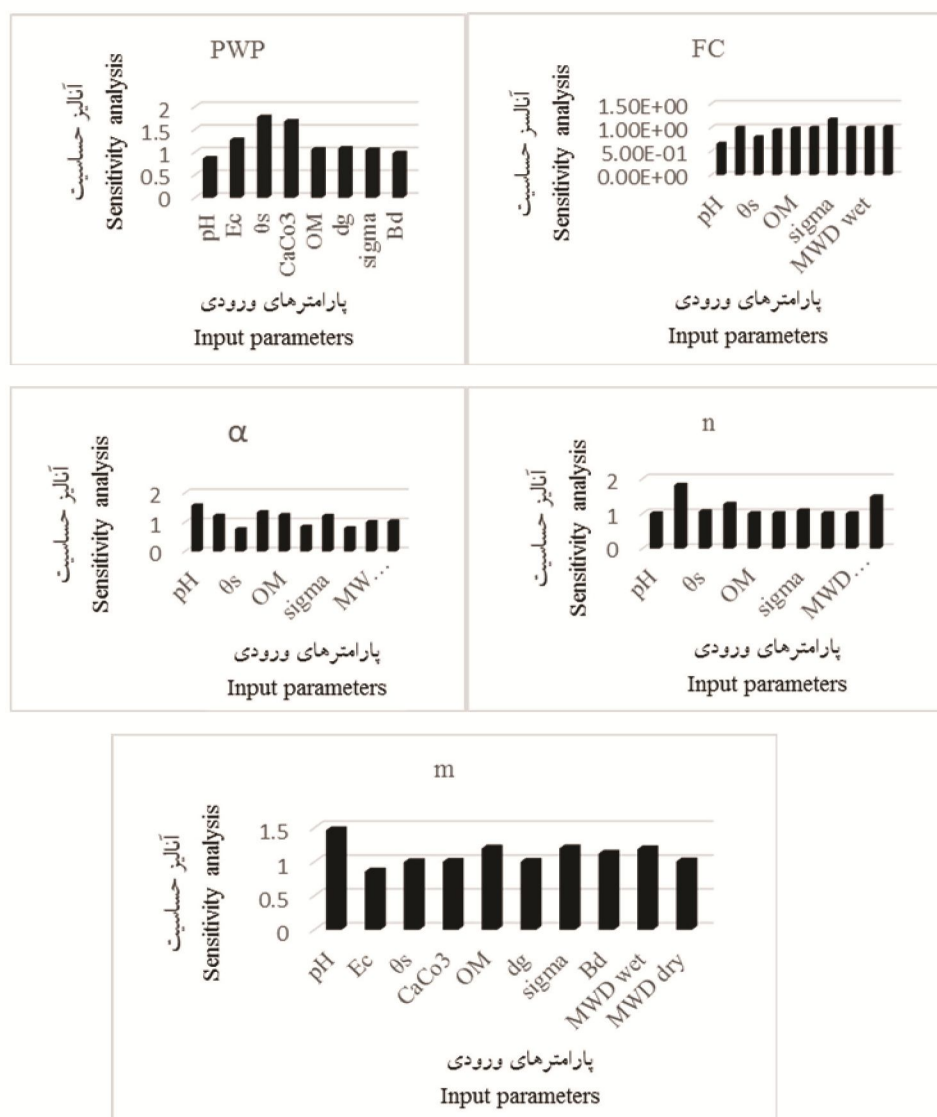


شکل ۲- آنالیز حساسیت مدل‌سازی برای متغیرهای هدف (نقطه پژمردگی PWP، ظرفیت زراعی FC، α ، n، m (فاکتورهای شکل منحنی مشخصه رطوبتی) در سناریو اول: pH اسیدیته، EC شوری، CaCO_3 کربنات کلسیم، OM ماده آلی، Clay رس، Sand شن، MWD_{wet} پایداری خاکدانه مرطوب و خشک، BD چگالی ظاهری.

Figure 2. Modelling sensitivity analysis for target variables (Permanent wilting point (PWP), Field capacity (FC), α , n, m (Moisture characteristic curve shape factors), Acidity (pH), Electrical conductivity (EC), Saturation moisture, Calcium Carbonate (CaCO_3), Organic matter (OM), Clay, Sand, mean weight diameter of wet and dry aggregate (MWD_{wet} and MWD_{dry}), Bulk density (BD)) in second scenario.

هندسی قطر ذرات می‌باشند، برای α شامل پ‌هاش، کربنات کلسیم و انحراف معیار هندسی قطر ذرات و ماده آلی، برای n شامل شوری، کربنات کلسیم و میانگین وزنی قطر ذرات و برای m شامل پ‌هاش، ماده آلی، میانگین وزنی قطر خاکدانه و انحراف معیار هندسی قطر ذرات بوده‌اند.

در شکل ۳ اهمیت متغیرهای ورودی بر PWP و FC و پارامترهای مدل و ن‌گنختن در سناریو دوم نشان داده شده است. مهم‌ترین پارامترهای اثرگذار بر PWP شامل رطوبت اشباع و کربنات کلسیم، بر FC شامل انحراف معیار هندسی قطر ذرات، میانگین وزنی قطر خاکدانه خشک و مرطوب، شوری و میانگین



شکل ۳- آنالیز حساسیت مدل‌سازی برای متغیرهای هدف (نقطه پژمردگی PWP، ظرفیت زراعی FC، α ، n ، m فاکتور شکل منحنی مشخصه رطوبتی) در سناریو دوم: pH اسیدیته، EC شوری، $CaCO_3$ کربنات کلسیم، OM ماده آلی، dg میانگین هندسی قطر ذرات، انحراف معیار هندسی قطر ذرات σ MWD_{wet} MWD_{dry} پایداری خاکدانه مرطوب و خشک، BD چگالی ظاهری.

Figure 3. Modelling sensitivity analysis for target variables (Permanent wilting point (PWP), Field capacity (FC), α , n , m (Moisture characteristic curve shape factors), Acidity (pH), Electrical conductivity (EC), Saturation moisture, Calcium Carbonate ($CaCO_3$), Organic matter (OM), Clay, Sand, mean weight diameter of wet and dry aggregate (MWD_{wet} and MWD_{dry}), Bulk density (BD)) in second scenario.

نتیجه‌گیری کلی

به‌طورکلی روابط بین ویژگی‌های یک پدیده در طبیعت غیرخطی و پیچیده است که با آنالیزهای کلاسیک و خطی نمی‌توان این روابط را استخراج نمود. در چنین مواردی استفاده از مدل‌های هوشمند مانند شبکه‌های عصبی مصنوعی و درخت تصمیم که می‌توانند روابط غیرخطی و پیچیده را شناسایی کنند، بسیار سودمند است. نتایج این پژوهش نشان داد که منحنی مشخصه رطوبتی خاک بیش‌تر تحت‌تأثیر کمیت‌های رطوبت اشباع، ماده آلی، میانگین وزنی قطر خاکدانه در حالت مرطوب است. هم‌چنین رطوبت در نقطه پژمردگی چندان متأثر از میانگین وزنی قطر خاکدانه خشک و مرطوب نبود و به همین دلیل در مدل‌سازی در نظر گرفته نشد. در مجموع با جایگزینی میانگین هندسی و انحراف معیار قطر ذرات به جای درصد رس و شن عملکرد بهتری در برآورد متغیرهای منحنی مشخصه رطوبتی به وجود آمد که نشان می‌دهد میانگین هندسی و انحراف معیار قطر ذرات مفاهیم گسترده‌تری نسبت به بافت خاک دارد، به‌عبارت دیگر بافت خاک نشان‌دهنده تغییرات (مقدار نسبی) ذرات است اما میانگین هندسی و انحراف معیار قطر ذرات هم تغییرات و هم اندازه ذرات را نشان می‌دهد. با توجه به مقادیر درصد RMSE و ضریب تبیین بهترین نتیجه در مدل‌سازی برای نقطه پژمردگی و ضعیف‌ترین برای n و α بود.

در سال‌های اخیر روش‌های مبتنی بر هوش مصنوعی، هم‌چون الگوریتم ژنتیک، درخت تصمیم‌گیری، شبکه‌های عصبی و تئوری مجموعه‌های فازی به خوبی به‌عنوان یک راه‌کار جایگزین برای غلبه بر مشکلات روش‌های سنتی ارتقا یافته‌اند. تجربه نشان داده است که پیاده‌سازی ساختارهای محاسباتی سیستم بیولوژیکی می‌تواند منجر به ایجاد الگوهای محاسباتی بهتری برای مسایل مبهم و پیچیده شود. توابع انتقالی توسعه یافته با استفاده از شبکه‌های عصبی و هوش مصنوعی توسط پژوهشگران بسیاری مورد استفاده قرار گرفته است و تقریباً نتیجه تمامی پژوهش‌های پیشین نشان می‌دهد که این مدل‌ها حداقل به اندازه بقیه روش‌های مورد استفاده خوب عمل می‌کنند و به خوبی بر مفروضات آماری درگیر با توابع انتقالی فائق می‌آیند (۵). پژوهشگران مختلفی بیان نمودند که از مزیت‌های شبکه‌های هوش مصنوعی نسبت به روش‌های رگرسیونی در تحلیل نتایج، علاوه بر دقت و همبستگی بیش‌تر با پارامترهای ورودی، جزو روش‌های تحلیل ناپارامتری هستند و بر پیش‌فرض نرمال بودن توزیع داده‌ها استوار نیستند بنابراین اگر توزیع داده‌ها نرمال نباشند، تحلیل داده‌ها به روش شبکه‌های هوش مصنوعی همانند شبکه عصبی مشکلی را به‌وجود نمی‌آورد.

منابع

1. Abbasi, Y.B., Ghanbarian-Alavijeh, A.M., Liaghat, A.M., and Shorafa. 2011. Evaluation of pedotransfer functions for estimating soil water retention curve of saline and saline-alkali soils of Iran. Soil Science Society of China. 21: 2. 230-237.
2. Amir-Abedi, H., Asghari, Sh.A., Mesri-Gandshamin, T., and Keivanbehjo, F. 2013. Estimating of field capacity, permanent wilting and available water content in Ardabil plain soils using regression and artificial neural network models. Urmia Applied Soil Research. 1: 1. 60-72. (In Persian)
3. Bouyoucos, G.J. 1962. Hydrometer method improved for making particle size analysis of soils. Agronomy Journal. 54: 464-465.
4. Farzadmehr, M., Chahmani, M., and Khayaki Siouki, A.S. 2018. Comparing decision tree and instance-based learning models to estimate soil saturated hydraulic conductivity. Journal of Soil and Water Conservation Research. 25: 5. 167-184. (In Persian)

5. Haghverdi, A., Ghahraman, B., Khoshnood Yazdi, A.A., and Arabi, Z. 2010. Estimating of water content in FC and PWP in north and north east of Iran's soil samples using k-nearest neighbor and artificial neural networks. *Journal of Water and Soil*, 24: 4. 804-814. (In Persian)
6. Hengle, T., and Husnjak, S. 2006. Evaluation adequacy and usability of soil maps in Croatia. *Soil Science Society of America Journal*, 70: 920-929.
7. Hutson, J.L., and Cass, A. 1987. A retentivity function for use in soil-water simulation models. *Journal of Soil Science*. 38: 105-113.
8. Kemper, W.D., and RoseNau, R.C. 1986. Aggregate stability and size distribution. P 425-442. In: D.L. Sparks (ed.) *Methods of soil analysis*. American society of agronomy, Madison.
9. Khashei Siuki, A., Jalali Moakhar, V.R., Noferești, A.M., and Ramazani, Y. 2015. Comparing nonparametric k-nearest neighbor technique with ANN model for predicting soil saturated hydraulic conductivity. *Journal of Soil Management in Sustainable Production*. 5: 3. 81-95. (In Persian)
10. Klute, A., and Dirksen, C. 1986. Hydraulic conductivity and diffusivity: Laboratory methods. P 687-734. In: A. Klute (ed.) *Method of soil analysis*, Part1: Agronomy Soil Science Society of America Madison, WI.
11. Matsuyama, T. 1987. Knowledge- Based Aerial Image Understanding system and Expert System for Image Processing. *IEEE Transaction on Geoscience and Remote Sensing*, 25: 305-316.
12. Meshkani, A.S., and Nazemi, A.S. 2009. *Introduction to Data Mining*. Ferdowsi University Press, Mashhad. 456p. (In Persian)
13. Minasny, B. 2007. Prediction soil properties. *Journal Ilmu Tanah dan Lingkungan*. 7: 54-67.
14. Moncada, M.P., Gabriels, D., and Cornelis, W.M. 2014. Data-driven analysis of soil quality indicators using limited data. *Geoderma*. 235: 271-278.
15. Motaghian, H.R., and Mohamadi, J. 2010. Comparison of some physical indicators of soil quality in different land uses in Marghmalak basin, Shahrekord (Chaharmahal and Bakhtiari province). *Journal of water and soil (Agricultural Sciences and Industries)*. 25: 1. 115-124. (In Persian)
16. Nemes, A., Rawls, W.J., and Pachepsky, Y.A. 2006. Use of the nonparametric nearest neighbor approach to estimate soil hydraulic properties. *Journal of Soil Science Society of America*. 70: 2. 327-336.
17. Openshaw, S., and Openshaw, C. 1997. *Artificial Intelligence in Geography*. John Wiley & Sons Ltd, Chichester. 348p.
18. Page, M.C., Sparks, D.L., Noll, M.R., and Hendricks, G.J. 1987. Kinetics and mechanisms of potassium release from sandy Middle Atlantic Coastal Plain soils. *Journal of Soil Science Society of America*. 51: 1460-1465.
19. Ramezani, M., Ganbarian, B., Liaghat, A.M., and Salehi Khoshkroudi, Sh. 2011. Developing pedotransfer functions for saline and saline- alkali soils. *Journal of Water and Irrigation Management*. 1: 1. 99-110. (In Persian)
20. Shahrabi, C. 2011. *Data Mining 2*. First edition, Amir Kabir University Industrial Jihad Press, Tehran. 300p.
21. Shirani, H., and Rafienejad, N. 2012. Estimating of some missing soil properties with regression pedotransfer functions and neural network in the Kerman. *Journal of Soil Research*. 25. 4: 349-359. (In Persian)
22. Shirani, H. 2017. *Artificial neural networks with an application in agricultural and natural resource sciences*. Rafsanjan Univ. Press. 320p. (In Persian)
23. Walkley, A., and 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. *Journal of Soil Science*. 37: 29-38.
24. Zhou, J., and Yu, J.L. 2005. Influences affecting the soil-water characteristic curve. *Journal of Zhejiang University Science*. 6: 797-804.



Separating Effective Soil Properties on Moisture Characteristic Curve Using Decision tree

S. Mesri¹, *Sh. Ghorbani Dashtaki², H. Shirani³, A. Kamkar Rohani⁴
and H.R. Motaghian⁵

¹Ph.D. Student, Dept. of Soil Science and Engineering, University of Shahrekord,

²Professor, Dept. of Soil Science and Engineering, University of Shahrekord,

³Professor, Dept. of Soil Science and Engineering, University of Vali-e-Asr Rafsanjan,

⁴Associate Prof., Dept. Oil and Mining Engineering - Geophysics, University of Shahrood,

⁵Assistant Prof., Dept. of Soil Science and Engineering, University of Shahrekord

Received: 02.05.2020; Accepted: 07.12.2020

Abstract

Background and Objectives: Soft computational techniques have been widely used in scientific research and engineering in recent decades. Since the measurement of hydraulic properties by direct laboratory methods is hard, time consuming and expensive, Thus, there is need to use alternative methods based on conveniently available soil properties to estimate it with less effort, time and cost. One of the new methods for estimating soil hydraulic properties, such as soil moisture characteristic curve, is non-parametric methods. This study was performed to determine the efficiency of the decision tree method in separation of effective properties in estimating soil moisture characteristic curve parameters.

Materials and Methods: To perform this study, number of 72 points were selected in the village of Marghmalek and Sharekord city. Samples were collected from depth of 0-20 cm and then were transferred to the laboratory for required measurements. Some properties such as pH, EC, saturated moisture, calcium carbonate, organic matter, clay and sand, bulk density, mean weight diameter of dry aggregate, mean weight diameter of wet aggregate, geometric mean and standard deviation of particle diameter were measured in the laboratory. Also, the moisture characteristic curves were determined at 0, 1, 3, 5, 10, 30, 50, 100, 150, 1000, 1500 kPa suctions and were fitted to the van Genuchten model. The input variables were introduced into the MATLAB software in two scenarios (first scenario: pH, EC, %clay and sand, organic matter, calcium carbonate, mean weight diameter of wet and dry aggregate, bulk density, saturated moisture and the second scenario: pH, EC, geometric mean and standard deviation of particle diameter, organic matter, calcium carbonate, mean weight diameter of wet and dry aggregate, bulk density, saturated moisture) and modeled by decision tree and error estimators of cross validation and resubstitution. Evaluation statistics of each model including R^2 , RMSE and %RMSE were calculated.

Results: The results obtained from decision tree modeling showed that the most important factors affecting moisture content in PWP suction, were saturated moisture and clay. The PWP target variable has the highest correlation in the first scenario (0.88) and in the second scenario (0.91) and the least error rate among the other variables and after that, FC has the highest correlation (0.86) in the second scenario. Target variables n had the highest error rate and α the lowest correlation in both scenarios. Generally, the second scenario performed better than the

* Corresponding Author; Email: shoja2002@yahoo.com

first scenario by replacing the geometric mean and standard deviation of particle diameter with the percentage of clay and sand. The sensitivity analysis showed that PWP was the most sensitive among the input parameters to pH, BD, calcium carbonate and organic matter and FC was the most sensitive to geometric standard deviation of particle and MWD_{wet} .

Conclusion: In general, modeling has been successful in both scenarios. But by substituting geometric mean and standard deviation of particle diameter instead of clay and sand percentage, a better performance was obtained in estimating moisture characteristic curve variables in the second scenario.

Keywords: Field capacity, Moisture curve, Van Genuchten equation