ارزیابی تاثیر گروه‌بندی بر پایه ویژگی‌های مختلف بر عملکرد توابع در تخمین ظرفیت تبادل کاتیونی خاک

چکیده

ظرفیت تبادل کاتیونی خاک یکی از مهمترین عوامل موثر در حاصلخیزی خاک است که اندازه‌گیری آن دشوار، زمان‌بر و هزینه‌بر است. استفاده از مدل‌ها و معادلات مختلف یکی از ساده‌ترین، ارزان‌ترین و سریع‌ترین روش‌های ارزیابی ظرفیت تبادل کاتیونی خاک است. لذا هدف از مطالعه حاضر ارزیابی تاثیر گروه‌بندی بر پایه ویژگی‌های مختلف بر عملکرد توابع در تخمین ظرفیت تبادل کاتیونی خاک و معرفی نوعی از گروه‌بندی که بهترین نتایج تخمین را دربرداشته باشد و همچنین مقایسه قابلیت تخمین ظرفیت تبادل کاتیونی با استفاده از روش شبکه‌های عصبی مصنوعی است. این مطالعه در سال ۱۴۰۰ در دانشگاه بوعلی سینا همدان انجام شد. در این پژوهش از ۴۵۹۴۸ نمونه خاک مربوط به پایگاه اطلاعاتی یکنواخت شده خاک‌های جهان استفاده گردید. ابتدا نمونه خاک‌های پایگاه اطلاعاتی در حالت‌های مختلف گروه‌بندی شدند. سپس برای کل داده و کلاس‌های مختلف هر گروه با استفاده از ۹ متغیر تخمینگر شامل اجزای بافت خاک، کربن آلی، سولفات کلسیم، کربنات کلسیم، جرم مخصوص ظاهری، درصد اشباع بازی، مجموع کاتیون‌های بازی قابل تبادل واکنش خاک در ۱۱ مدل ارزیابی شد. نتایج نشان داد در کلاس‌های بافتی ضریب بهبود نسبی در بخش آزمون شبکه عصبی مصنوعی برابر ۸۷ درصد بود. همچنین نتایج نشان داد که RMSE در بخش آزمون در کلاس درشت بافت برابر ۰/۲۵۷ و برای کلاس ریز بافت برابر با ۰/۳۶۴ بود. به طورکلی نتایج نشان داد که استفاده از توابع به دست آمده که گروه‌بندی در آن‌ها موجب بهبود تخمین ظرفیت تبادل کاتیونی شده روشی آسان و کم هزینه در تخمین ظرفیت تبادل کاتیونی به شمار می‌رود.

کلیدواژه‌ها: پایگاه اطلاعات خاک، شبکه عصبی مصنوعی، قابلیت اعتماد مدل، مدل‌سازی

Evaluation of the Effect of Grouping Based on Different Characteristics on the Performance of Functions in Estimating Soil Cation Exchange Capacity

Abstract

This study addresses the challenge of measuring soil cation exchange capacity (CEC), a vital factor influencing soil fertility, by exploring the impact of grouping soil samples based on different characteristics on the performance of estimation models. Recognizing the difficulties associated with traditional CEC measurement methods, the study employs a cost-effective and rapid approach using various models and equations. The research, conducted at Bu Ali Sina University in Hamedan, utilizes a substantial dataset of 45,948 soil samples from the standardized database of world soils. Soil samples are initially categorized into different groups, and nine estimator variables are examined across 11 models for the entire dataset and specific classes within each group. These variables include soil texture components, organic carbon, calcium sulfate, calcium carbonate, bulk density, base saturation percentage, total exchangeable base cations, and soil reaction. The results demonstrate that grouping soil samples, especially based on texture classes, significantly improves the performance of artificial neural network models, with a remarkable 87% relative improvement coefficient in the test section. The study reveals that data grouping enhances the model's estimation capabilities, as evidenced by reduced root mean square error (RMSE) values in the test sections for different texture classes. In conclusion, the findings suggest that utilizing functions derived from grouped data offers an effective and cost-efficient method for estimating soil cation exchange capacity. This approach provides valuable insights for soil fertility management, offering a simplified yet accurate means of assessing this critical soil parameter.

Keywords: Soil database, Linear regression, Artificial neural network, Model reliability

مقدمه

حاصلخیزی خاک عبارت است از قابلیت خاک برای فراهمی عناصر غذایی مورد نیاز گیاه به مقدار کافی و نسبت مناسب در طی دوره رشد گیاه که در تغذیه گیاهان و سلامتی و بقاء موجودات اهمیت بسیاری دارد (Marschner and Rengel, 2023; Tessler et l. 2023). ظرفیت تبادل کاتیونی[1] (CEC) از جمله خصوصیات خاک است که در تعیین حاصلخیزی خاک نقش بسزایی دارد (Emamgholizadeh et al. 2023). ویژگی جذب سطحی خاک که از مهم‌ترین عوامل تعیین کننده حاصلخیزی خاک محسوب می‌شود با CEC خاک مرتبط است. به‌طوری‌که هر چه مقدار CEC زیاد باشد ظرفیت جذب سطحی خاک نیز بیش‌تر بوده و توان

[1]. Cation Exchange Capacity

نگهداری آب و عناصر غذایی افزایش خواهد یافت (Yuksele and Kaya, 2006; Firat Pulat et al. 2014). CEC نقش مهمی در کیفیت خاک دارد (Mishra et al. 2019). CEC می‌تواند تحت تاثیر ویژگی‌های فیزیکی (مانند بافت خاک)، شیمیایی (مانند pH، کانی‌شناسی) و بیولوژیکی (مانند ماده آلی خاک) خاک باشد. بنابراین، CEC جنبه‌های هر سه گروه نشانگر را که معمولاً برای تعیین کیفیت خاک استفاده می‌شوند ادغام می‌کند (Mukherjee and Zimmerman, 2013).

CEC خاک دربردارنده اطلاعات بسیار مهمی مانند مینرالوژی رس‌های خاک نیز است. CEC به صورت بافر عمل کرده و با تعدیل واکنش خاک مانع آبشویی عناصر غذایی می‌گردد. در نتیجه ذخیره عناصر غذایی برای ریشه گیاهان را افزایش می‌دهد (Krogh et al. 2000; Obalum et al. 2012). همچنین CEC یکی از مهمترین شاخص‌های خاک است که کیفیت محیط زیست را کنترل می‌کند. بنابراین اطلاع از CEC خاک‌ها به انتخاب عملیات مدیریت کشاورزی مناسب، به ویژه در خاک‌های با حاصلخیزی پایین کمک خواهد کرد (Ibrahim et al. 2022). تعیین CEC به دلایل مختلفی مانند تغییرپذیری مکانی–زمانی این ویژگی (Sayegh et al. 1978; Cai et al. 2023) و عدم وجود روش‌های آزمایشگاهی استاندارد مستقیم و آسان برای اندازه‌گیری آن (غالب روش‌ها شامل مراحل زیاد و نیازمند ترکیبات شیمیایی مختلف هستند) (Soares et al. 2005) دشوار بوده و نیازمند صرف وقت و هزینه زیادی است. از طرف دیگر توسعه روش‌های آسان، دقیق و کم هزینه برای تعیین خواص خاک در ارزیابی کیفیت خاک به منظور بهبود مدیریت سیستم خاک در مناطقی که داده‌ها کم و یا غیر ممکن هستند مورد نیاز است (Costa et al. 2015; Pulido et al. 2018). از این رو در سال‌های اخیر پژوهشگران سعی در برآورد CEC خاک با بهره‌گیری از روش‌های غیرمستقیم نموده‌اند. توابع انتقالی با استفاده از داده‌هایی از مشاهدات خاک یا منابع دیگر برای پیش‌بینی خواصی از خاک که تعیین آن‌ها سخت، هزینه‌بردار و زمان‌بر است، توسعه یافته‌اند و به همین دلیل از ویژگی‌های زودیافت خاک مانند بافت، واکنش خاک و کربن آلی تخمین استفاده می‌کنند (MacDonald, 1998).

پیشینهٔ پژوهش

برای توسعه توابع انتقالی روش‌های متعددی وجود دارد. بر مبنای مطالعات انجام شده رایج‌ترین این روش‌ها شبکه‌های عصبی مصنوعی و معادلات رگرسیونی می‌باشند، که هر یک ویژگی‌های خاص خود را دارند. نتایج Tamari et al. (1996) نشان داد که در صورت بالا بودن ناپایداری داده‌ها، شبکه‌های عصبی مصنوعی نسبت به مدل‌های رگرسیونی خطی عملکرد بهتری نخواهند داشت. اما اگر از داده‌های با دقت بالا استفاده شود شبکه‌های عصبی کارایی بالاتری را نشان می‌دهند. بررسی منابع در مورد نحوه عملکرد دو روش شبکه‌های عصبی مصنوعی و رگرسیونی در پیش‌بینی CEC نشان داده است که در بیش‌تر تحقیقات روش شبکه‌های عصبی مصنوعی در مقایسه با روش رگرسیونی کارایی بالاتر و قابلیت تخمین بهتری را داشته‌اند (Amini et al. 2005; Tang et al. 2008; Rahimi Lake et al. 2009). ولی چون استفاده از شبکه‌های عصبی مصنوعی به علت نامعلوم بودن روابط بین ورودی‌ها و خروجی‌ها و مشکل بودن تحلیل توابع حاصله دشوار می‌باشد، و از سوی دیگر مدل‌های رگرسیونی برای استفاده و تحلیل روابط ساده‌تر می‌باشند، لذا همچنان از این روش برای ایجاد توابع استفاده می‌شود. Olorunfemi et al. (2016) نیز بیان کردند که توابع انتقالی ایجاد شده با روش‌های رگرسیون خطی چند گانه برای پیش‌بینی CEC با استفاده از مشخصات پایه‌ای خاک به عنوان تخمین‌گر موفقیت‌آمیز بود.

تفاوت خاک‌های بررسی شده بدون در نظر گرفتن شباهت‌های میان آن‌ها، مانند همگونی در ترکیب کانی‌شناسی، قدرت پیش‌بینی خواص آن‌ها را کاهش می‌دهد. بنابراین در صورت مناسب بودن گروه‌بندی انجام شده، ایجاد تابع مجزا برای هر گروه قدرت پیش‌بینی ویژگی‌های خاک را افزایش خواهد داد (Syers et al. 1970). در این راستا برخی از پژوهشگران قبل از ایجاد توابع پیش‌بینی‌کننده CEC، ابتدا نمونه خاک‌های مورد مطالعه را بر اساس برخی از خصوصیات خاک مانند رده، افق و کلاس بافت گروه‌بندی کردند و در بیش‌تر موارد نتایج آن‌ها نشان داد که با قرار دادن خاک‌های مشابه در یک گروه، دقت و صحت توابع به مقدار قابل توجهی افزایش یافت (بیات و همکاران، ۱۳۹۰؛ Seybold et al. 2005; Kar et al. 2008). با این حال در غالب موارد تعداد خصوصیاتی که خاک‌ها بر پایه آن‌ها طبقه‌بندی شده‌اند معدود است. در واقع گروه‌بندی بر اساس کربنات کلسیم، رژیم

رطوبتی خاک، ظرفیت آب قابل استفاده و برخی از ویژگی‌هایی که برای تعریف واحد خاک و کشاورزی مناسب هستند انجام نشده و گروه‌بندی برمبنای یک ویژگی خاص به عنوان بهترین ویژگی معرفی نشده است. در حالی که اگر گروه‌بندی با بالاترین کارایی ممکن انجام شود شاید بتوان مقادیر CEC را با کم‌ترین خطای ممکن نسبت به مقادیر اندازه‌گیری شده این ویژگی پیش‌بینی نمود. همچنین با معرفی نوعی از گروه‌بندی که بهترین نتایج تخمین را در بر دارد می‌توان در صرف زمان لازم برای برآورد CEC خاک با بالاترین دقت ممکن صرفه‌جویی نمود. ایجاد توابع انتقالی با قابلیت تخمین قابل قبول برای خاک‌های گروه‌بندی شده بر اساس ویژگی‌های مختلف، شاید بتوانند این امکان را برای دانشمندان علوم خاک فراهم نمایند تا به راحتی از این توابع برای تخمین CEC تحت شرایطی که اندازه‌گیری مستقیم آن ممکن نیست استفاده کنند. بنابراین اهداف این تحقیق عبارت‌اند از: (۱) ارزیابی تاثیر گروه‌بندی بر پایه گروه بافت، کلاس بافت، رده، افق، رژیم رطوبتی، ظرفیت آب قابل استفاده، کربنات کلسیم و برخی از ویژگی‌ها که برای تعریف واحد خاک و کشاورزی مناسب هستند بر عملکرد توابع در تخمین CEC خاک و معرفی نوعی از گروه‌بندی که بهترین نتایج تخمین را دربردارد، (۲) مقایسه قابلیت تخمین CEC توسط روش شبکه‌های عصبی مصنوعی است.

مواد و روش‌ها

در مطالعه حاضر داده‌های حدود ۴۵۹۴۸ نمونه خاک جمع‌آوری شده از خاک سطحی (۰-۳۰ سانتی‌متر) و عمقی (۱۰۰-۳۰ سانتی‌متر) مربوط به پایگاه اطلاعاتی یکنواخت شده خاک‌های جهان (HWSD[1]) استفاده شد (پایگاه اطلاعاتی یکنواخت شده خاک‌های جهان، ۲۰۱۲). مجموعه اطلاعات HWSD با استفاده از چهار پایگاه داده‌ای شامل پایگاه نقشه خاک‌های جهان[2]، پایگاه اطلاعاتی خاک و نوع زمین منطقه‌ای مختلف[3]، پایگاه اطلاعاتی خاک‌های اروپا[4] (ESDB) و انجمن علوم خاک چین[5] (نقشه خاک چین با مقیاس ۱:۱ میلیون) جمع‌آوری و تکمیل شده است (پایگاه اطلاعاتی یکنواخت شده خاک‌های جهان، ۲۰۱۲). در مطالعه حاضر به منظور برآورد CEC خاک متغیرهای اجزاء بافت خاک، کربن آلی، واکنش خاک، جرم مخصوص ظاهری، کربنات کلسیم، سولفات کلسیم، نسبت سیلت به شن، درصد رس، ظرفیت آب قابل استفاده، مجموع کاتیون‌های بازی تبادلی، برخی از ویژگی‌های مناسب برای تعریف واحد خاک کشاورزی و CEC از پایگاه اطلاعاتی مذکور در فوق مورد استفاده قرار گرفتند.

آماده‌سازی داده‌ها

قبل از ایجاد توابع، نمونه خاک‌های پایگاه داده‌ای HWSD بر اساس ویژگی‌های گروه بافت خاک، کلاس بافت خاک، رده، افق، مقدار کربنات کلسیم، مقدار سولفات کلسیم، ظرفیت آب قابل استفاده، رژیم رطوبتی و برخی از ویژگی‌ها که برای تعریف واحد خاک و کشاورزی مناسب هستند گروه‌بندی شدند. هر یک از این گروه‌ها نیز دربردارنده کلاس‌های متفاوتی هستند (جدول ۱). علاوه بر کلاس‌های آورده شده در جدول ۱، برای هر گروه یک کلاس تحت عنوان کل داده نیز وجود دارد. در جدول ۲ تغییرات رژیم رطوبتی در گروه‌های مختلف نمایش داده شده است.

جدول ۱. کلاس‌های مختلف هر گروه برای داده‌های مورد استفاده از HWSD

کلاس	کلاس بافت	رده	رژیم رطوبتی خاک	برخی از ویژگی-های مناسب برای تعریف واحد خاک	گروه بافت[b]	کربنات کلسیم (درصد)	سولفات کلسیم (درصد)
۱	رسی [a]	ورتی سول	۱۵۰	Pertic	درشت	۲-۵	۲۵-۰
۲	رسی سیلتی	انتی سول	۱۲۵	Gelic	متوسط	۵-۱۵	

1. Harmonized World Soil Database
2. Soil map of the world
3. Various regional soil and terrain databases (SOTWIS databases)
4. European soil database
5. Chinese Academy of Sciences – The soil map of China 1:1 million scale

۳	رسی	اسپودوسول	۱۰۰	Vertic	ریز
۴	لوم رسی سیلتی	آلفی سول	۷۵		
۵	لوم رسی	اکسی سول	۵۰		
۶*	–	اندی سول	۱۵		
۷	لوم سیلت	اولتی سول			
۸	رسی شنی	اریدی سول			
۹	لوم	مالی سول			
۱۰	لوم رسی شنی				
۱۱	لوم شنی				
۱۲	شن لومی				
۱۳	شنی				

[a]. خاک‌های که محتوی بیش از ۶۰ درصد رس و کم‌تر از ۴۰ درصد سیلت و شن می‌باشند در این کلاس قرار می‌گیرند.

[b]. گروه بافت درشت: این گروه محتوی خاک‌های شنی، شن لومی و لوم شنی با مقدار رس کم‌تر از ۱۸ درصد و

مقدار شن بیش‌تر از ۶۵ درصد است؛ گروه بافت متوسط: این گروه شامل خاک‌های لوم شنی، لوم، لوم رسی شنی، لوم سیلت، سیلت، لوم رسی سیلتی و لوم رسی با مقدار رس کم‌تر از ۳۵ درصد و مقدار شن کم‌تر از ۶۵ درصد است، گروه بافت ریز: این گروه شامل خاک‌های رسی، رسی شنی، لوم رسی و لوم رسی سیلتی با مقدار رس بیش‌تر از ۳۵ درصد است.

*برای کلاس ۶ گروه کلاس بافت تعداد داده کافی برای ایجاد تابع وجود نداشت.

جدول ۲. تغییرات گروه‌های مختلف رژیم رطوبتی خاک‌های مورد مطالعه

گروه	رژیم رطوبتی خاک	افق
۱	در عمق ۸۰ سانتی‌متر برای بیش‌تر از ۳ ماه مرطوب نیست، یا مرطوب در عمق ۴۰ سانتی‌متر برای ۱ ماه	Stony[c]
۲	مرطوب در عمق ۸۰ سانتی‌متر برای ۳ تا ۶ ماه، اما غیر مرطوب در عمق ۴۰ سانتی‌متر برای بیش‌تر از ۱ ماه	Lithic
۳	مرطوب در عمق ۸۰ سانتی‌متر برای بیش‌تر از ۶ ماه، اما غیر مرطوب در عمق ۴۰ سانتی‌متر برای بیش‌تر از ۱۱ ماه	Pertic
۴	مرطوب در عمق ۴۰ سانتی‌متر برای بیش‌تر از ۱۱ ماه	Petrocalcic

[c]. Stony: در مناطقی که دارای این افق می‌باشند به دلیل وجود تکه سنگ‌های بزرگ و یا رخنمون سنگی در لایه‌های سطحی استفاده از ادوات مکانیزه کشاورزی دشوار است. Lithic: هنگامی‌که در ۵۰ سانتی‌متری سطح خاک، چسبندگی مداوم و یا سخت کفه وجود داشته باشد نام افق Lithic خواهد بود. Pertic: در صورتی که در عمق ۱۰۰ سانتی‌متری سطح خاک، ۴۰ درصد یا بیش‌تر اکسیدهای سخت یا پلینتایت[1] یا سنگ آهن و یا تکه سنگ‌های بزرگ با ضخامت کم‌تر از ۲۰ سانتی‌متر وجود داشته باشد نام افق Pertic خواهد بود. Petrocalcic: خاک‌های محتوی بیش از ۴۰ درصد آهک در عمق ۱۰۰ سانتی‌متری سطح خاک را گویند.

به منظور ارزیابی تاثیر گروه‌بندی داده‌ها بر تخمین CEC خاک و اینکه گروه‌بندی بر مبنای کدام خصوصیت نتیجه تخمین بهتری را خواهد داشت، یک بار برای کل داده‌ها در هر گروه و بار دیگر برای هر کلاس توابع انتقالی توسعه یافت. در هر کلاس پس از حذف داده‌های تکراری، تعداد معینی داده جهت ایجاد توابع مورد استفاده قرار گرفت که در جدول ۳ نشان داده شده است. کلیه متغیرها با استفاده از نرم‌افزار مینی تب[2] (آزمون کولموگروف – اسمیرنوف) از نظر نرمال بودن بررسی شدند. متغیرهایی که توزیع نرمال نداشتند با تبدیلات مختلف نرمال گردیدند. به این صورت که کربنات کلسیم، درصد اشباع بازی و سولفات کلسیم با استفاده از روش لگاریتم‌گیری و شن با استفاده از تابع x^2 نرمال شدند.

جدول ۳. تعداد داده‌های مورد استفاده برای ایجاد توابع برای نمونه‌های مربوط به پایگاه داده‌ای HWSD

	کل داده	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳
کلاس بافت	۲۳۵۰	۲۰۰	۱۴۶	۲۰۰	۲۰۰	۲۰۰	–	۲۰۰	۲۰۴	۲۰۰	۲۰۰	۲۰۰	۲۰۰	۲۰۰
افق	۱۳۱۴	۳۲۸	۲۲۸	۶۶	۵۸	۸۰	۱۴۶	۹۸	۶۸	۵۸	۱۸۴			
رده	۱۵۰۶	۲۰۰	۲۰۰	۶۰	۲۰۰	۲۰۰	۱۲۸	۱۱۸	۲۰۰	۲۰۰				
ظرفیت آب قابل	۱۱۵۲	۲۰۰	۲۰۰	۲۰۰	۱۷۴	۲۰۰	۱۷۸							
رژیم رطوبتی خاک	۸۳۰	۱۱۲	۲۷۰	۲۰۲	۱۵۸	۸۸								

1. Plinthite
2. MINITAB

گروه بافت	۶۰۰	۲۰۰	۲۰۰	۲۰۰
برخی از ویژگی‌های مناسب برای تعریف واحد خاک و	۳۰۸	۷۴	۳۴	۲۰۰
کربنات کلسیم	۴۳۶	۲۴۷	۱۸۹	
سولفات کلسیم	۱۶۷			

توسعه توابع انتقالی

برای ایجاد توابع انتقالی برای هر کلاس و کل داده‌های هر گروه دو سوم داده‌ها برای آموزش و یک سوم برای آزمون توابع ایجاد شده به صورت تصادفی انتخاب گردیدند. جهت اطمینان از تصادفی بودن داده‌ها میانگین داده‌های آموزش و آزمون با آزمون t مقایسه شدند. CEC خاک به عنوان متغیر وابسته و سایر متغیرها (درصد رس، نسبت سیلت به شن، کربن آلی، جرم مخصوص ظاهری، واکنش خاک، درصد اشباع بازی، مجموع کاتیون‌های بازی تبادلی، کربنات کلسیم و سولفات کلسیم) به عنوان متغیر مستقل استفاده شدند. استفاده همزمان از متغیرهای مستقلی که همبستگی بالایی بین آن‌ها وجود دارد منجر به همراستایی[1] یا چند هم‌خطی می‌گردد که در ایجاد روابط رگرسیونی باید از آن دوری شود. چرا که باعث ایجاد مدل‌های رگرسیونی نامعتبر می‌گردد (Merdun et al. 2006). به دلیل همبستگی بالایی که بین فراوانی ذرات رس، سیلت و شن وجود داشت برای جلوگیری از همراستایی در توابع به دست آمده، در ایجاد توابع انتقالی به جای مقادیر شن و سیلت از نسبت سیلت به شن استفاده شد. ایجاد توابع برای گروه‌های ویژگی‌های بافت خاک، کلاس بافت خاک، رده، افق، ظرفیت آب قابل استفاده، رژیم رطوبتی و برخی از ویژگی‌ها که برای تعریف واحد خاک و کشاورزی مناسب هستند و کلاس‌های مربوطه در ۱۱ مرحله و با دو روش رگرسیون خطی و شبکه‌های عصبی مصنوعی صورت گرفت. ورودی‌های مدل‌ها در جدول ۴ نشان داده شده است. انتخاب متغیرهای ورودی بر اساس قابلیت دسترسی آن‌ها برای کاربران و بررسی میزان بهبود تخمین‌ها با اضافه کردن هر کدام از آن‌ها انجام شده است. در انتخاب مدل مناسب بایستی مسئله هزینه انجام آزمایش‌های مورد نیاز را در نظر گرفت، به عبارتی مدلی مناسب است که علاوه بر دقت مناسب از نظر هزینه نیز مقرون به صرفه باشد. همانگونه که در جدول ۴ مشاهده می شود با افزایش تعداد پارامترها به مدل طبیعتا هزینه انجام آزمایش‌ها افزایش می‌یابد، بنابراین بایستی بین دقت و هزینه یک تعادل برقرار کرد.

جدول ۴. متغیرهای ورودی استفاده شده در هر مدل برای نمونه‌های مربوط به پایگاه داده‌ای HWSD

مدل	ورودی‌های مدل
۱	رس و نسبت سیلت به شن
۲	رس، نسبت سیلت به شن و کربنات کلسیم
۳	رس، نسبت سیلت به شن و سولفات کلسیم
۴	رس، نسبت سیلت به شن، کربنات کلسیم و سولفات کلسیم
۵	رس، نسبت سیلت به شن، کربنات کلسیم، سولفات کلسیم و کربن آلی
۶	رس، نسبت سیلت به شن، درصد کربن آلی و واکنش خاک
۷	رس، نسبت سیلت به شن، درصد کربن آلی، واکنش خاک و جرم مخصوص ظاهری
۸	رس، نسبت سیلت به شن، درصد کربن آلی، واکنش خاک و درصد اشباع بازی
۹	رس، نسبت سیلت به شن، درصد کربن آلی، واکنش خاک و مجموع کاتیون های بازی تبادلی
۱۰	رس، نسبت سیلت به شن، درصد کربن آلی، واکنش خاک، درصد اشباع بازی و مجموع کاتیون‌های بازی تبادلی
۱۱	رس، نسبت سیلت به شن، درصد کربن آلی، واکنش خاک، جرم مخصوص ظاهری، کربنات کلسیم و سولفات کلسیم

در مدل اول فقط از ویژگی‌های توزیع اندازه ذرات (رس و نسبت سیلت به شن) به عنوان متغیر ورودی استفاده گردید و CEC تخمین زده شد. در مراحل بعدی متغیرهای جدیدی به جمع ورودی‌های مدل اضافه شدند و صحت و قابلیت اعتماد آن‌ها با مدل اول مقایسه شد. در تمام توابع ایجاد شده به منظور بررسی وجود یا عدم وجود مشکل هم‌راستایی چندگانه از آماره فاکتور

1. Multicolinearity

تورم واریانس[1] استفاده شد. به‌گونه‌ای که اگر مقدار آماره فاکتور تورم واریانس بزرگ‌تر از ۱۰ باشد، ضعف برآورد ضرایب رگرسیون بر اثر هم‌راستایی چندگانه مشهود خواهد بود (Ho, 2006). در مطالعه حاضر برای تمامی مدل‌ها مقدار آماره مذکور کم‌تر از ۱۰ بود (داده‌ها نشان داده نشده‌اند). در این پژوهش از شبکه‌های عصبی پرسپترون چند لایه با یک لایه مخفی و تابع فعال‌سازی تانژانت هایپربولیک استفاده شد. تعداد سلول‌های لایه پنهان از ۳ تا ۸ در نظر گرفته شد. سپس از بین نتایج، شبکه با تعداد سلول پنهان که بهترین نتیجه (بر اساس RMSE و R^2) را داشت به عنوان شبکه نهایی انتخاب شد.

معیارهای ارزیابی عملکرد توابع ایجاد شده

برای بررسی دقت و قابلیت اعتماد توابع ساخته شده و مقایسه عملکرد آن‌ها از معیارهای جذر میانگین مربعات خطا[2](RMSE) (مقدم، ۱۳۸۰) و بهبود نسبی[3](RI) (Bayat et al. 2008) بهره‌گیری شد.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - y_m)^2} \quad (۱)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \quad (۲)$$

$$AIC = N\ln\left[\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}\right] + 2n_p \quad (۳)$$

$$RI = \frac{RMSE_1 - RMSE_n}{RMSE_1} \times 100 \quad (۴)$$

در روابط فوق N تعداد نمونه‌های خاک، y_i، $\hat{y}_i$ و y_m به ترتیب مقادیر اندازه‌گیری شده، تخمین زده شده و میانگین مقادیر اندازه‌گیری شده ظرفیت تبادل کاتیونی می‌باشند. در رابطه ۲ $RMSE_1$ میانگین مجذور میانگین مربعات خطای مدل‌های ایجاد شده با کل داده و $RMSE_i$ میانگین مجذور میانگین مربعات خطای مدل‌های ایجاد شده در هر کلاس می‌باشد. از آماره بهبود نسبی برای مشخص نمودن اندازه برتری توابع انتقالی در هر گروه نسبت به کل داده‌ها و همچنین مقایسه عملکرد از روش شبکه‌های عصبی مصنوعی استفاده شد. شایان ذکر است آماره RMSE از طریق تقسیم بر مقادیر میانگین CEC اندازه‌گیری شده نرمال گردید و فاقد واحد است. برای مقایسه میانگین بین گروه‌ها از آزمون دانکن در سطح پنج درصد با استفاده از نرم‌افزار SPSS 22 انجام شد.

بحث و نتایج

همبستگی بین متغیرها

در جدول ۵ ضرایب همبستگی خطی (پیرسون) بین متغیرهای ورودی و CEC به عنوان متغیر خروجی مربوط به نمونه‌های پایگاه داده‌ای HWSD نشان داده شده است. همبستگی CEC با رس مثبت و در سطح یک درصد معنی‌دار بود (جدول ۵). در مطالعات بسیاری از پژوهشگران بین CEC و رس همبستگی مثبت مشاهده شده است (Horn et al. 2005; Rahimi Lake et al.

1. Variance inflation factor
2. Root mean square error
3. Relative improvement

2009; Sarmadian et al. 2013). اغلب این پژوهشگران دلیل وجود همبستگی بین این دو ویژگی را بار سطحی زیاد و سطح ویژه بسیار بالای ذرات رس بیان نموده‌اند. همبستگی مثبت CEC با رس و کربن آلی به دلیل مشارکت آنها در ایجاد بارهای منفـی و پدیـده تبادل کاتیونی است که توسط بسیاری از پژوهشگران به آن اشاره شده است (Amini et al. 2005).

همانطور که در جدول ۵ مشاهده می‌شود همبستگی بین درصد شن و CEC منفی و در سطح یک درصد معنی‌دار بود. وجود همبستگی منفی و معنی‌دار بین درصد شن و CEC نشان می‌دهد که وجود مقادیر بالای شن در خاک باعث کاهش CEC خواهد شد. نوربخش و همکاران (۱۳۸۲) در پژوهش خود همبستگی منفی و معنی‌داری بین CEC و درصد شن مشاهده کردند. این پژوهشگران بیان نمودند که همبستگی منفی CEC با درصد شن از آن‌رو است که با افزایش شن از مقدار ذراتی مانند رس که دارای توانایی ایجاد بارهای منفی هستند و سهم بسزایی را در CEC خاک‌ها دارند کاسته می‌شود. Obalum et al. (2012) نیز گزارش کردند که در گروه خاک با پایداری خیلی کم، همبستگی شن با CEC منفی و معنی‌دار بود. در حالی که در سایر گروه‌ها (خاک با پایداری کم، پایداری متوسط، پایداری بالا و خاک سطحی و عمقی) ارتباط معنی‌داری بین آن‌ها مشاهده نشد. همبستگی بین سیلت و CEC در کل داده مثبت و در سطح یک درصد معنی‌دار بود (جدول ۵). خداوردی‌لو و حسینی عربلو (۱۳۹۲) نیز گزارش کردند که ارتباط سیلت با CEC در کلاس بافت رسی، در سطح یک درصد معنی‌دار بود. برخی از شواهد در مقالات نشان داد که جزء سیلت سهم معنی‌دار و بسزایی را در CEC خاک دارند (Thompsonet al. 1989; Asadu et al. 1997).

ارتباط منفی و معنی‌داری در سطح یک درصد بین CEC و جرم مخصوص ظاهری مشاهده شد (جدول ۵). در واقع جرم مخصوص ظاهری تابعی از بافت و ماده آلی خاک است که از مهم‌ترین عوامل موثر بر CEC به شمار می‌روند. از این رو ارتباط بین BD و CEC با توجه به نوع بافت خاک و مقدار ماده آلی آن متفاوت خواهد بود. به طوری که با افزایش مقدار ماده آلی و رس خاک BD کاهش پیدا می‌کند در حالی که این روند برای CEC معکوس خواهد بود افزایش می‌یابد. در برخی مواقع حضور پوشش‌های رسی و یا ماده آلی بیش‌تر در فرکشن‌های سیلت و شن سبب می‌شود با افزایش این ذرات هر دو خصوصیت BD و CEC خاک نیز افزایش یابند (Morras, 1995; Turpault et al. 1996). در پژوهش فولادمند (۱۳۸۶) نیز جرم مخصوص ظاهری با CEC همبستگی منفی داشت. ارتباط کربن آلی با CEC مثبت و در سطح یک درصد معنی‌دار بود (جدول ۵). شمار بسیاری از پژوهشگران پس از انجام آنالیزهای آماری دریافتند که همبستگی معنی‌داری بین کربن آلی و CEC خاک وجود دارد که اغلب آن‌ها این ارتباط را با نقش فوق‌العاده مهم مواد آلی در فراهمی بارهای موثر در CEC توجیه نمودند (Rahimi Lake et al. 2009; Keshavarzi et al. 2010; Bayat et al. 2013). Sarmadian et al. (2013) مشاهده کردند که استفاده از کربن آلی به عنوان تخمینگر در مدل‌های شبکه‌های عصبی مصنوعی باعث بهبود قابلیت تخمین CEC گردید. همبستگی بین واکنش خاک و CEC در سطح یک درصد مثبت و معنی‌دار بود (جدول ۵). خداوردی‌لو و حسینی عربلو (۱۳۹۲) گزارش کردند که همبستگی بین واکنش خاک و CEC فقط در کلاس بافت لوم مثبت و معنی‌دار بود و در سایر کلاس‌های بافتی این همبستگی منفی و غیرمعنی‌دار شده است. آن‌ها بیان کردند که وجود همبستگی مثبت و معنی‌دار در سطح یک درصد بین واکنش خاک و CEC در کلاس بافت لومی نشان‌دهنده وجود بارهای وابسته به واکنش خاک در خاک است، به طوری که با افزایش واکنش خاک، مقدار بارهای منفی و لذا CEC افزایش می‌یابد.

در این مطالعه همبستگی مثبت و معنی‌داری (در سطح یک درصد) بین CEC و درصد اشباع بازی مشاهده شد (جدول ۵). وجود ارتباط معنی‌دار بین دو متغیر CEC و درصد اشباع بازی را می‌توان این گونه توجیه کرد که خاک‌های با درصد اشباع بازی بالا، میزان pH بالایی نیز دارند، بنابراین ظرفیت بافری بیش‌تری در برابر کاتیون‌های اسیدی (مانند Al^{3+}) و فرآیندهایی (مانند باران اسیدی و نیتریفیکاسیون) که منجر به اسیدی شدن خاک و به تبع آن کاهش CEC می‌شوند، خواهد داشت (Kissel and Sonon, 2008).

با توجه به نتایج آورده شده در جدول ۵، همبستگی CEC با مجموع کاتیون‌های بازی تبادلی مثبت و در سطح یک درصد

معنی‌دار بود. می‌توان وجود همبستگی بالا بین مجموع کاتیون‌های بازی تبادلی و CEC را به این موضوع ارتباط داد که از مجموع ظرفیت تبادلی هر خاک مقدار قابل توجهی در اختیار بازهای تبادلی است. هرچند نقش کاتیون‌های تبادلی در CEC بسته به نوع خاک و واکنش خاک آن متفاوت خواهد بود (محمدی، ۱۳۸۵). نتایج Manrique et al. (1991) نشان داد که در همه خاک‌ها (در افق‌های A، B و C و همه رده‌ها) مجموع کاتیون‌های بازی تبادلی همبستگی بسیار بالا و معنی‌داری با CEC دارد.

جدول ۵. ضریب همبستگی بین متغیرهای ورودی و خروجی (CEC) در خاک‌های پایگاه داده‌ای HWSD

Sa	سیلت	نسبت سیلت به شن	رس	جرم مخصوص ظاهری	کربن آلی
کل داده	-۰/۵۱۱**	۰/۳۱۵**	۰/۱۹۸**	-۰/۴۲۸**	-۰/۵۴۶**

واکنش خاک	درصد اشباع بازی	مجموع کاتیون های بازی تبادلی	کربنات کلسیم	سولفات کلسیم	واکنش خاک
۰/۶۰۶**	۰/۱۹۶**	۰/۲۴۵**	۰/۸۱۸**	۰/۰۲۶	۰/۶۰۶**

* و ** به ترتیب معنی‌داری در سطح ۵ و ۱ درصد را نشان می‌دهند.

ارزیابی تاثیر گروه‌بندی بر عملکرد توابع ایجاد شده

به منظور ارزیابی نحوه تاثیر گروه‌بندی بر اساس ویژگی‌های مختلف (گروه بافت، کلاس بافت، رده، رژیم رطوبتی، افق، ظرفیت آب قابل استفاده، کربنات کلسیم و برخی از ویژگی‌های مناسب برای تعریف واحد خاک و کشاورزی) بر عملکرد توابع ایجاد شده با دو روش شبکه‌های عصبی مصنوعی و رگرسیون خطی در پیش‌بینی CEC از معیار بهبود نسبی و آزمون میانگین جفتی بین آماره RMSE استفاده گردید.

ارزیابی تاثیر گروه‌بندی بر پایه گروه بافت

مقایسه آماره‌های RMSE و RI در دو مجموعه آموزش و آزمون برای کلاس کل داده و کلاس‌های تفکیک شده بر اساس گروه بافت خاک نشان داد که (جدول ۶) در همه گروه‌های بافتی (درشت، متوسط و ریز) تفکیک داده‌ها سبب بهبود معنی‌دار در برآورد CEC توسط دو روش رگرسیون خطی و شبکه‌های عصبی مصنوعی شد. احتمالا از دلایل اصلی بهبود دقت مدل‌ها کاهش نسبی هم‌راستایی چندگانه و کاسته شدن از میزان اختلاط اثرات متغیرهای تخمینگر پس از طبقه‌بندی داده‌ها بر پایه گروه بافت است. همچنین جداسازی داده‌ها برحسب شباهت در ویژگی‌ها سبب افزایش یکنواختی ویژگی‌های خاک‌ها می‌گردد (Salchow et al. 1996 ؛ مهاجر و همکاران، ۱۳۸۷).

جدول ۶. نتایج مقایسه میانگین دقت توابع ایجاد شده توسط شبکه‌های عصبی مصنوعی برای کل داده‌ها با کلاس‌های آن‌ها در گروه‌های بافت

کلاس	آموزش		آزمون	
	میانگین RMSE[a]	میانگین RI	میانگین RMSE	میانگین RI
کل داده	۰/۴۲۹		۰/۵۵۹	
درشت	۰/۱۹۰*	۵۵/۷	۰/۲۵۷*	۵۴/۱
متوسط	۰/۲۸۲*	۳۴/۳	۰/۴۰۴*	۲۷/۸
ریز	۰/۲۵۹*	۳۹/۶	۰/۳۶۴*	۳۵

[a]. آماره‌های RMSE و RI به ترتیب نشان دهنده مقادیر میانگین مجذور میانگین مربعات خطا و بهبود نسبی یازده مدل شبکه‌های عصبی مصنوعی هستند.

* معنی‌دار بودن اختلاف بین مقادیر RMSE کل داده و کلاس‌های مختلف هر گروه در مرحله آموزش و آزمون را نشان می‌دهد.

ارزیابی تاثیر گروه‌بندی بر پایه کلاس بافت

با توجه به نتایج آورده شده در جدول ۷ گروه‌بندی خاک‌ها به تفکیک کلاس‌های بافتی و به منظور افزایش کارایی مدل‌های شبکه‌های عصبی مصنوعی بر اثر همگون‌سازی داده‌ها، فقط در برخی از کلاس‌ها موجب بهبود تخمین CEC گردید. در کلاس رسی در مجموعه آموزش RMSE به طور معنی‌داری نسبت به کلاس کل داده کاهش یافت و قابلیت تخمین CEC را به مقدار اندک افزایش داد. در کلاس‌های لوم رسی و لوم شنی نیز مانند کلاس رسی فقط در مجموعه آموزش و در کلاس لوم رسی سیلتی در هر دو مجموعه آموزش و آزمون مقدار آماره RMSE که معیاری از صحت برآورد است، کم‌تر از کلاس کل داده بود و گروه‌بندی داده‌ها در این کلاس‌ها به طور موفقیت‌آمیزی موجب بهبود معنی‌دار قابلیت تخمین CEC نسبت به کلاس کل داده گردید. مقادیر RI نیز به خوبی صحت این موضوع را نشان می‌دهند (جدول ۷).

Wagner et al. (2001) نیز بر پایه کلاس‌های بافتی مختلف توابعی را برای پیش‌بینی CEC خاک‌های مناطق شمالی آلمان ایجاد کردند. این پژوهشگران مشاهده نمودند تفکیک داده‌ها باعث افزایش قابلیت تخمین CEC توسط توابع در کلاس سیلتی نسبت به سایر کلاس‌های بافتی گردید. احتمالا کاهش درصد کربن آلی خاک با افزایش درصد شن را می‌توان به سرعت بالای فرآیندهای تجزیه مواد آلی در خاک‌های شنی که تهویه بهتری دارند نسبت داد، در حالی که در خاک‌های ریزبافت، رس‌ها با نگهداری مواد آلی از تجزیه مواد آلی جلوگیری می‌کنند (نوربخش و همکاران، ۱۳۸۲) این امر و افزون بر آن تفاوت در فراوانی رس موجب شده است که بیشترین پراکندگی مقادیر CEC در کلاس بافتی رسی و کمترین آن در کلاس بافتی لوم شنی باشد (خداوردی لو و حسینی عربلو، ۱۳۹۲). معلمی و دواتگر (۱۳۸۹) نیز تاثیر گروه‌بندی خاک‌ها بر پایه کلاس‌های بافتی و کربن آلی بر بهبود قابلیت تخمین CEC توسط دو روش شبکه‌های عصبی مصنوعی و رگرسیون خطی چندگانه را مورد بررسی قرار دادند. نتایج آن‌ها نشان داد که گروه‌بندی داده‌ها در کلاس‌های بافتی شن و شن رسی لومی در روش شبکه‌های عصبی مصنوعی، موجب بهبود محسوس پیش‌بینی‌ها نسبت به کل داده‌ها شد. در سایر کلاس‌های بافتی (لوم شنی، لوم، لوم سیلتی، لوم رسی، لوم رسی سیلتی، رسی سیلتی و رسی) پس از گروه‌بندی نه تنها بهبودی در قابلیت تخمین توابع نسبت به کل داده مشاهده نشد بلکه مقادیر و RMSE به ترتیب کاهش و افزایش یافتند. زیرا به نظر می‌رسد CEC خاک علاوه بر بافت خاک تحت تاثیر نوع کانی‌های رسی نیز قرار دارد. این پژوهشگران دلیل بهبود مشاهده شده در قابلیت پیش‌بینی توابع بسط یافته در کلاس‌های بافتی شن و شن رسی لومی بر اثر گروه‌بندی داده‌ها را همگونی بیش‌تر نمونه خاک‌های این دو کلاس از نظر منشاء و نوع کانی‌های رسی بیان کردند. در پژوهشی دیگر Salchow et al. (1996) به منظور برآورد CEC با استفاده از ویژگی‌های زودیافت خاک (درصد شن، سیلت، رس و ماده آلی و مقدار جرم مخصوص ظاهری) در اوهایوی جنوبی ابتدا داده‌ها را بر پایه کلاس بافت گروه‌بندی نمودند، سپس توابع را بسط دادند. نتایج آن‌ها نشان داد که بر اثر تفکیک داده‌ها هم‌راستایی چندگانه بین متغیرهای مستقل به طور معنی‌داری کاهش یافت و به تبع آن عملکرد توابع رگرسیونی بهبود قابل توجهی پیدا کرد. هزارجریبی و همکاران (۱۳۹۲) نشان دادند که اضافه شدند تعداد پارامترها به مدل سبب افزایش دقت تخمین CEC می‌شود.

جدول ۷. نتایج مقایسه میانگین دقت توابع ایجاد شده توسط شبکه‌های عصبی مصنوعی برای کل داده‌ها با کلاس‌های آن‌ها در گروه کلاس بافت

کلاس	آموزش		آزمون	
	میانگین RMSE[a]	میانگین RI	میانگین RMSE	میانگین RI
کل داده	۰/۲۰۴		۰/۲۰۶	
رسی سنگین	۰/۲۳۰*	-۱۲/۹	۰/۲۵۸*	−۲۵/۳
رسی سیلتی	۰/۳۴۴*	−۶۸/۸	۰/۴۵۶*	−۱۲۱/۸
رسی	۰/۱۹۲*	۵/۹	۰/۲۰۹	−۱/۷
لوم رسی سیلتی	۰/۱۵۴*	۲۴/۶	۰/۱۷۶*	۱۴/۳
لوم رسی	۰/۱۶۹*	۱۷/۲	۰/۲۱۵	−۴/۴
لوم سیلت	۰/۲۷۴*	−۳۴/۳	۰/۲۹۷*	−۴۴/۶
رسی شنی	۰/۲۱۹	−۷/۲	۰/۲۶۴*	−۲۸/۵
لوم	۰/۳۱۷*	−۵۵/۴	۰/۳۳۲*	−۶۱/۷

لوم رسی شنی	۰/۱۹۵	۴/۲	۰/۲۲۴*	−۸/۸
لوم شنی	۰/۱۷۲*	۱۵/۹	۰/۲۱۰	−۲/۲
شن لومی	۰/۲۶۴*	−۲۹/۳	۰/۲۵۱*	−۲۲/۱
شنی	۰/۳۴۳*	−۶۷/۹	۰/۴۲۲*	−۱۰۵/۴

[a]. آماره‌های RMSE و RI به ترتیب نشان دهنده مقادیر میانگین مجذور میانگین مربعات خطا و بهبود نسبی یازده مدل شبکه‌های عصبی مصنوعی هستند.

* معنی‌دار بودن اختلاف بین مقادیر RMSE کل داده و کلاس‌های مختلف هر گروه در مرحله آموزش و آزمون را نشان می‌دهد.

ارزیابی تاثیر گروه‌بندی بر پایه رژیم رطوبتی

نمونه‌های پایگاه داده‌ای HWSD بر پایه ویژگی رژیم رطوبتی به ۵ کلاس (۰، ۱، ۲، ۳ و ۴) گروه‌بندی شدند. مقایسه معیارهای RMSE و RI به دست آمده برای کلاس کل داده و سایر کلاس‌ها نشان داد که (جدول ۸) بر اثر تفکیک داده‌ها بر پایه ویژگی رژیم رطوبتی در کلاس‌های ۲، ۳ و ۴ (در هر دو مجموعه آموزش و آزمون) و در کلاس ۱ در مجموعه آموزش قابلیت تخمین CEC توسط توابع ایجاد شده توسط دو روش شبکه‌های عصبی مصنوعی و رگرسیون خطی به طور معنی‌داری افزایش یافت (جدول ۸). اگرچه در کلاس صفر (در هر دو مجموعه آموزش و آزمون) و ۱ (مجموعه آزمون) RMSE نسبت به کلاس کل داده افزایش یافت و گروه‌بندی برآورد CEC این کلاس‌ها را بهبود نداد. اما همانطور که ملاحظه می‌شود پس از جداسازی داده‌ها برمبنای ویژگی رژیم رطوبتی خاک، به طور عمده نتایج تخمین بهبود یافت. از این رو می‌توان چنین بیان کرد که این نوع گروه‌بندی سبب افزایش قابلیت تخمین CEC شده است. همانطور که قبلا اشاره شد ویژگی CEC علاوه بر مقدار ماده آلی و رس تابعی از نوع رس خاک نیز است (Edmeades, 1982). به طوری که نسبت CEC مربوط به بارهای دایم و وابسته به واکنش خاک در بین کانی‌های رسی متفاوت است (Seybold et al. 2005). احتمالا از دلایل تاثیر رژیم رطوبتی بر میزان CEC خاک‌ها آن باشد که این ویژگی خاک علاوه بر اینکه میزان مواد آلی خاک را تحت تأثیر قرار می‌دهد بر نوع کانی‌های رسی خاک نیز موثر است. به طوری که در رژیم‌هایی که شرایط برای تغییر و تبدیل کانی‌ها مناسب نیست بیش‌تر در خاک شاهد کانی‌های اولیه مانند کائولینایت، که نقش آن‌ها در CEC خاک نسبت به اسمکتایت‌ها کم‌تر است (Sparks, 1995)، خواهیم بود (زراعت پیشه و همکاران، ۱۳۹۱). در حالی که در رژیم‌های رطوبتی مرطوب‌تر شرایط برای ایجاد کانی اسمکتایت، که دربردارنده سهم مهمی از CEC خاک می‌باشند، مناسب‌تر است (Khormali et al. 2003). Pachepsky and Rawls (1999) گزارش کردند هنگامی‌که نمونه خاک‌های مورد ارزیابی بر اساس رژیم رطوبتی طبقه‌بندی شدند دقت مدل‌ها در پیش‌بینی منحنی رطوبتی خاک افزایش یافت.

جدول ۸. نتایج مقایسه میانگین دقت توابع ایجاد شده توسط شبکه‌های عصبی مصنوعی برای کل داده‌ها با کلاس‌های آن‌ها در گروه رژیم رطوبتی

کلاس	آموزش		آزمون	
	میانگین RMSE[a]	میانگین RI	میانگین RMSE	میانگین RI
کل داده	۰/۲۶۷		۰/۲۷۰	
۰	۰/۷۵۶*	−۱۸۳/۶	۱/۰۱۲*	−۲۷۴/۳
۱	۰/۲۴۴	۸/۴	۰/۲۷۹	−۳/۳
۲	۰/۱۴۱*	۴۷/۰	۰/۱۵۸*	۴۱/۶
۳	۰/۱۴۰*	۴۷/۶	۰/۰۵۰*	۸۱/۴
۴	۰/۱۲۹*	۵۱/۶	۰/۱۱۷*	۵۶/۷

[a]. آماره‌های RMSE و RI به ترتیب نشان دهنده مقادیر میانگین مجذور میانگین مربعات خطا و بهبود نسبی یازده مدل شبکه‌های عصبی مصنوعی هستند.

* معنی‌دار بودن اختلاف بین مقادیر RMSE کل داده و کلاس‌های مختلف هر گروه در مرحله آموزش و آزمون را نشان می‌دهد.

ارزیابی تاثیر گروه‌بندی بر پایه رده

همانطور که در جدول ۹ مشاهده می‌شود گروه‌بندی داده‌ها برمبنای رده‌های مختلف تاثیر متفاوتی بر کارایی توابع ایجاد شده توسط دو روش شبکه‌های عصبی مصنوعی و رگرسیون خطی در این رده‌ها داشت. در همه رده‌ها (به جز رده انتی‌سول) در مجموعه آموزش، پس از تقسیم‌بندی داده‌ها معیار RMSE به طور معنی‌داری کاهش یافت. مقادیر بهبود نسبی نیز بیش از ۵ درصد بود. در مجموعه آزمون نیز از ۹ رده آورده شده در این مطالعه در ۵ رده آلفی‌سول، اندی‌سول، اولتی‌سول، اریدی‌سول و مالی‌سول، تفکیک داده‌ها موجب کاهش معنی‌دار خطای تخمین شد (جدول ۹). هرچند در مجموعه آزمون در ۴ رده (ورتی‌سول، انتی‌سول، اسپودوسول و اکسی‌سول) و در مجموعه آموزش در رده‌های انتی‌سول و اسپودوسول، مقادیر RMSE نسبت به کل داده افزایش یافت (جدول ۸)، و در کارایی توابع بهبودی مشاهده نشد. اما در کل نتایج به دست آمده بیانگر بهبود قابلیت تخمین توابع پس از گروه‌بندی داده‌ها برمبنای رده، به ویژه در مجموعه آموزش است. رده اولتی‌سول به دلیل RMSE برابر با ۰/۱۶۱ و ۰/۱۷۲ و RI برابر با ۴۷/۳ و ۴۵/۱ درصد به ترتیب در مجموعه آموزش و آزمون، بیش‌ترین دقت و قابلیت تخمین را در مقایسه با سایر رده‌ها داشت (جدول ۹).

تفاوت مشاهده شده در قابلیت تخمین توابع بسط یافته برای رده‌های مختلف می‌تواند حاصل تفاوت در مقدار ماده آلی، میزان و نوع رس خاک‌های هر رده باشد. با افزایش میزان رس‌های سیلیکاتی و مواد آلی، که از مهم‌ترین منابع بارهای منفی موثر بر CEC هستند، در خاک‌های هر رده CEC خاک نیز افزایش خواهد یافت (Khormali et al. 2005). Seybold et al. (2005) به منظور افزایش دقت توابع در پیش‌بینی CEC، نمونه خاک‌های مورد مطالعه را بر پایه ۵ ویژگی‌های (مقدار کربن آلی، واکنش خاک، کلاس مینرالوژی فامیل[1]، کلاس فعالیت CEC[2] و رده) به گروه‌های همگن‌تر طبقه‌بندی نمودند. این پژوهشگران مشاهده نمودند در بین رده‌های مورد ارزیابی رده‌های آلفی‌سول (با کربن آلی بیش از ۳ درصد) و اینسپتی‌سول با کم‌ترین مقدار میانگین RMSE (۰/۱۲۵) بهترین نتایج تخمین را داشتند. به طور کلی، تفکیک خاک‌ها بر اساس رده و تفکیک هر خاک و فامیل به لایه‌های زیرین و سطحی موجب افزایش دقت مدل‌ها در تخمین CEC شده است. دلیل اصلی افزایش دقت مدل‌ها را می‌توان به افزایش همگنی و یکنواختی ویژگی‌ها و همچنین کاهش نسبی اختلاط اثرات این ویژگی‌ها پس از انجام تفکیک در بین گروه‌ها دانست (مهاجر و همکاران، ۱۳۸۷). نتایج Manrique et al. (1991) نیز نشان دادند که گروه‌بندی خاک‌ها بر اساس رده منجر به افزایش قابلیت تخمین CEC توسط رس و کربن آلی شد. به عنوان مثال کربن آلی و رس بیش از ۶۷ درصد از تغییرات CEC در آلفی‌سول‌ها، اینسپتی‌سول‌ها، مالی‌سول‌ها و ورتی‌سول‌ها، و بیش از ۷۸ درصد از تغییرات CEC در انتی‌سول‌ها و اسپودوسول‌ها را توجیه می‌کنند. در پژوهشی دیگر (Asadu and Akamigbo, 1990) برای برآورد CEC چهار رده خاک، یکسری توابع انتقالی بسط داده شد. در این پژوهش گزارش گردید هنگامی‌که داده‌های خاک برمبنای رده طبقه‌بندی شدند در پیش‌بینی CEC بهبود قابل توجهی مشاهده شد.

جدول ۹. نتایج مقایسه میانگین دقت توابع ایجاد شده توسط شبکه‌های عصبی مصنوعی برای کل داده‌ها با کلاس‌های آن‌ها در گروه رده

کلاس	آموزش		آزمون	
	میانگین RMSE[a]	میانگین RI	میانگین RMSE	میانگین RI
کل داده	۰/۲۱۶		۰/۲۲۰	
ورتی‌سول	۰/۱۸۹*	۱۲/۴	۰/۲۸۶*	−۲۹/۸
انتی‌سول	۰/۲۱۷	−۰/۳	۰/۲۷۳*	−۲۴/۰
اسپودو-آلفی‌سول	۰/۲۱۷	−۰/۵۰	۰/۲۹۴*	−۳۳/۵
	۰/۱۸۵*	۱۴/۵	۰/۱۷۹*	۱۸/۷
اکسی-اندی‌سول	۰/۱۹۷*	۹/۱	۰/۲۴۰	−۸/۹
	۰/۱۷۱*	۲۱/۰	۰/۲۱۰	۴/۸

1. Family mineralogy class
2. CEC activity class

اولتی‌سول	۰/۱۱۴*	۴۷/۳	۰/۱۲۱*	۴۵/۱
اریدی-	۰/۱۶۱*	۲۵/۶	۰/۱۷۲*	۲۱/۹
مالی‌سول	۰/۱۴۹*	۳۱/۱	۰/۱۴۷*	۳۳/۱

[a]. آماره‌های RMSE و RI به ترتیب نشان دهنده مقادیر میانگین مجذور میانگین مربعات خطا و بهبود نسبی یازده مدل شبکه‌های عصبی مصنوعی هستند.

* معنی‌دار بودن اختلاف بین مقادیر RMSE کل داده و کلاس‌های مختلف هر گروه در مرحله آموزش و آزمون را نشان می‌دهد.

ارزیابی تاثیر گروه‌بندی بر پایه افق‌های مختلف

هنگامی‌که داده‌ها بر پایه افق‌های مختلف گروه‌بندی شدند عملکرد توابع بسط یافته به‌وسیله شبکه‌های عصبی مصنوعی به گونه قابل توجهی در مقایسه با حالتی که از کل داده استفاده شد افزایش یافت (جدول ۱۰). زیرا به استثناء یک کلاس (Inundic) در همه کلاس‌ها (در ۹ افق) در هر دو مرحله آموزش و آزمون آماره‌های RMSE به طور معنی‌داری کاهش پیدا کرند. مقادیر بهبود نسبی نیز تاثیر مثبت گروه‌بندی داده‌ها بر اساس افق بر قابلیت تخمین CEC خاک را نشان می‌دهد (جدول ۱۰). احتمالا نتایج حاصله به دلیل افزایش یکنواختی و همگنی ویژگی‌های خاک‌ها و کاهش تغییرات خاک تحت تأثیر جداسازی داده‌ها برحسب شباهت در ویژگی‌ها (ویژگی افق) است (Drake and Motto, 1982; Salchow et al. 1996).

جدول ۱۰. نتایج مقایسه میانگین دقت توابع ایجاد شده توسط شبکه‌های عصبی مصنوعی برای کل داده‌ها با کلاس‌های آن‌ها در گروه افق

کلاس	آموزش		آزمون	
	میانگین RMSE[a]	میانگین RI	میانگین RMSE	میانگین RI
کل داده	۰/۳۸۴		۰/۳۹۱	
Stony	۰/۲۷۹*	۳۰/۱	۰/۱۹۷*	۴۹/۶
Lithic	۰/۱۴۲*	۶۴/۶	۰/۱۴۵*	۶۲/۹
Pertic	۰/۲۷۶*	۳۱/۰	۰/۳۶۰*	۷/۸
Petrocalcic	۰/۱۲۸*	۶۷/۹	۰/۱۷۹*	۵۴/۱
Phreatic	۰/۱۶۹*	۵۷/۷	۰/۲۱۵*	۴۴/۹
Saline	۰/۱۱۰*	۷۲/۵	۰/۱۴۱*	۶۴/۰
Sodic	۰/۰۷۶*	۸۰/۹	۰/۱۰۱*	۷۴/۱
Inundic	۰/۶۲۱*	−۵۵/۵	۰/۷۷۵*	−۹۸/۳
Erosion	۰/۳۰۴*	۲۳/۹	۰/۳۰۲*	۲۲/۶
Gravelly	۰/۱۶۲*	۵۹/۵	۰/۱۷۸*	۵۴/۴

[a]. آماره‌های RMSE و RI به ترتیب نشان دهنده مقادیر میانگین مجذور میانگین مربعات خطا و بهبود نسبی یازده مدل شبکه‌های عصبی مصنوعی هستند.

* معنی‌دار بودن اختلاف بین مقادیر RMSE کل داده و کلاس های مختلف هر گروه در مرحله آموزش و آزمون را نشان می‌دهد.

نوع افق در تعیین CEC خاک دارای اهمیت بسزایی است. شاید به این دلیل که کاتیون‌های موثر در CEC خاک به طور عمده توسط بارهای منفی ایجاد شده به وسیله ذرات رس و مواد آلی خاک تحت تأثیر قرار می‌گیرند، از سوی دیگر ویژگی بار مواد آلی نیز در افق‌های مختلف به جهت متفاوت بودن مراحل تجزیه متغیر است. همچنین بسیاری از ویژگی‌های خاک (مقدار ماده آلی، pH، ساختمان خاک و مقدار و نوع رس) که از عوامل موثر بر CEC خاک محسوب می‌شوند، در افق‌های مختلف متفاوت هستند (Tang et al. 2008). Matos et al. (2001) نیز دریافتند که افق‌های خاک فاکتورهایی که ارتباط معنی‌داری با CEC دارند را تحت تأثیر قرار می‌دهند. Madeira et al. (2003)، CEC ۳۱ افق را اندازه‌گیری کردند. این پژوهشگران مشاهده کردند در افق‌های سطحی به دلیل وجود کربن آلی بالا در مقایسه با افق‌های زیرسطحی، ظرفیت نگه‌داری کاتیون‌های بازی قابل تبادل

بیش‌تر بوده است. احتمال می‌رود وجود اختلاف معنی‌دار میان ماده آلی دو عمق به خاطر روند معمول توزیع مواد آلی در پروفیل خاک باشد زیرا مقدار مواد آلی در سطح خاک بیشتر از عمق است. احتمالا علت بالاتر بودن رس در لایه زیرین هر گروه نسبت به لایه سطحی ناشی از پدیده انتقال است که باعث جابه‌جایی رس از افق‌های سطحی (اپی پدون‌ها) به عمق‌های پایین‌تر و تجمع آن در افق‌های تحت‌الارضی می‌گردد.

تحقیقات نوربخش و همکاران (۱۳۸۲) نیز همانند نتایج مطالعه حاضر نشان داد که تفکیک خاک‌ها بر اساس افق‌های مختلف سبب افزایش دقت توابع گردید. همچنین در مطالعات این پژوهشگران معلوم شد CEC مواد آلی در افق‌های زیرین بیش‌تر از CEC افق A است. هرچند سهم مواد آلی در CEC افق A بیش‌تر از افق‌های زیرین است. در افق‌های زیرین به دلیل کاهش مواد آلی و افزایش فراوانی رس با افزایش عمق، معمولا نقش رس در CEC بیش‌تر خواهد بود. مطالعات نشان داده است که رده‌بندی خاک‌ها بر اساس تاکسونومی و یا گروه‌بندی بر اساس افق‌های مشخصه، از آنجایی که باعث همگون شدن ویژگی‌های خاک‌ها می‌شود، منجر به کاهش تغییرات در خصوصیات خاک می‌گردد و توابع انتقالی حاصل جهت تخمین CEC می‌تواند به مقدار قابل توجهی دقیق‌تر و صحیح‌تر باشد (Asadu and Akamigbo, 1990; Drake and Motto, 1982).

ارزیابی تاثیر گروه‌بندی بر پایه ظرفیت آب قابل استفاده

توابع انتقالی توسعه یافته در کلاس کل داده و در کلاس‌های تفکیک شده (۶ کلاس) بر اساس ویژگی ظرفیت آب قابل استفاده در مجموعه داده‌های آموزش و آزمون با استفاده از آماره‌های RMSE و RI مقایسه شدند (جدول ۱۱). بر پایه آماره RI، در مجموعه آموزش کلاس‌های ۴، ۵ و ۶ نسبت به کل داده به ترتیب از ۳۴/۹، ۲۰/۹ و ۲۹/۹ درصد بهبود نسبی برخوردار هستند (جدول ۱۱). در مجموعه آزمون نیز کلاس‌های ۵ و ۶ نسبت به کل داده به ترتیب ۱۵/۸ و ۱۹/۰ درصد بهبود داشتند. همچنین مقادیر RMSE در این کلاس‌ها (مجموعه آموزش کلاس ۴ و هر دو مجموعه آموزش و آزمون کلاس ۵ و ۶) به گونه معنی‌داری کاهش پیدا کردند (جدول ۱۱). در سایر کلاس‌ها (۱، ۲، ۳ و مجموعه آزمون کلاس ۴) نه تنها گروه‌بندی باعث بهبود عملکرد توابع نگردید بلکه ناکارآمدی توابع را افزایش داد (جدول ۱۱). شاید بتوان چگونگی تاثیر مقدار آب قابل استفاده خاک بر میزان CEC خاک را به تغییرات مقدار ماده آلی خاک بر اثر تغیر در رطوبت خاک ارتباط داد. در خاک‌هایی که میزان رطوبت قابل استفاده خاک بیش‌تر است اکسیداسیون مواد آلی به دلیل خیس بودن خاک و تهویه کم‌تر کاهش می‌یابد و به تبع آن مقدار ماده آلی خاک (که به دلیل دارا بودن گروه‌های عاملی از منابع مهم ایجاد بارهای منفی موثر بر CEC محسوب می‌شوند) بیش‌تر خواهد بود (Wilding et al. 1983).

جدول ۱۱. نتایج مقایسه میانگین دقت توابع ایجاد شده توسط شبکه‌های عصبی مصنوعی برای کل داده‌ها با کلاس‌های آن‌ها در گروه AWC

کلاس	آموزش		آزمون	
	میانگین RMSE[a]	میانگین RI	میانگین RMSE	میانگین RI
کل داده	۰/۲۲۶		۰/۲۰۱	
۱	۰/۳۵۶*	−۵۷/۸	۰/۴۱۶*	−۱۰۷/۳
۲	۰/۲۵۸	−۱۴/۲	۰/۲۶۸*	−۳۳/۷
۳	۰/۲۳۳*	−۳/۴	۰/۲۳۶*	−۱۷/۸
۴	۰/۱۴۷*	۳۴/۹	۰/۲۰۵	−۲/۲
۵	۰/۱۷۸*	۲۰/۹	۰/۱۶۹*	۱۵/۸
۶	۰/۱۵۸*	۲۹/۹	۰/۱۶۲*	۱۹/۰

[a]. آماره‌های RMSE و RI به ترتیب نشان دهنده مقادیر میانگین مجذور میانگین مربعات خطا و بهبود نسبی یازده مدل شبکه‌های عصبی مصنوعی هستند.

* معنی‌دار بودن اختلاف بین مقادیر RMSE کل داده و کلاس‌های مختلف هر گروه در مرحله آموزش و آزمون را نشان می‌دهد.

ارزیابی تاثیر گروه‌بندی بر پایه برخی از ویژگی‌های مناسب برای تعریف واحد خاک و کشاورزی

جدول ۱۲ نتایج پیش‌بینی CEC توسط توابع انتقالی را قبل و پس از جداسازی داده‌های منبع HWSD برمبنای برخی ویژگی‌های مناسب برای تعریف واحد خاک و کشاورزی نشان می‌دهد. ارزیابی آماره‌های RMSE و RI در کل داده و کلاس‌های Pertic، Gelic و Vertic نشان داد که پس از جداسازی داده‌ها قابلیت پیش‌بینی توابع در هر ۳ کلاس به ویژه در کلاس Vertic بهبود چشمگیری داشت. کاهش مقادیر RMSE به جز مجموعه آزمون کلاس Pertic و Gelic در سایر موارد معنی‌دار بود.

جدول ۱۲. مقایسه میانگین دقت توابع ایجاد شده توسط شبکه‌های عصبی مصنوعی برای کل داده‌ها با کلاس‌های آن‌ها در گروه برخی از ویژگی‌های مناسب برای تعریف واحد خاک و کشاورزی

کلاس	آموزش		آزمون	
	میانگین RMSE[a]	میانگین RI	میانگین RMSE	میانگین RI
کل داده	۰/۲۱۶		۰/۲۱۸	
Pertic	۰/۱۷۵*	۱۹/۲	۰/۱۹۰	۱۳/۰
Gelic	۰/۱۵۲*	۲۹/۸	۰/۱۹۸	۹/۵
Vertic	۰/۰۷۰*	۶۷/۸	۰/۰۶۱*	۷۲/۲

[a]. آماره‌های RMSE و RI به ترتیب نشان دهنده مقادیر میانگین مجذور میانگین مربعات خطا و بهبود نسبی یازده مدل شبکه‌های عصبی مصنوعی هستند.

* معنی‌دار بودن اختلاف بین مقادیر RMSE کل داده و کلاس‌های مختلف هر گروه در مرحله آموزش و آزمون را نشان می‌دهد.

ارزیابی تاثیر گروه‌بندی بر پایه مقادیر مختلف کربنات کلسیم

همانطور که در جدول ۱۳ مشاهده می شود گروه‌بندی داده‌ها بر پایه مقادیر مختلف کربنات کلسیم در هر دو مجموعه آموزش و آزمون کلاس ۲ سبب بهبود معنی‌داری در قابلیت تخمین CEC توسط توابع گردید. ارزیابی معیار RI (جدول ۱۳) نشان داد که عملکرد توابع ایجاد شده در مجموعه آموزش و آزمون کلاس ۲ نسبت به عملکرد توابع ایجاد شده درکل داده به ترتیب از ۴۴/۰ و ۳۹/۹ در صد بهبود نسبی برخوردار هستند. برای کلاس ۱ پس از تفکیک داده‌ها بر اساس مقدار کربنات کلسیم بهبودی در قابلیت تخمین توابع مشاهده نگردید.

جدول ۱۳. نتایج مقایسه میانگین دقت توابع ایجاد شده توسط شبکه‌های عصبی مصنوعی برای کل داده‌ها با کلاس‌های آن‌ها در گروه کربنات کلسیم

کلاس	آموزش		آزمون	
	میانگین RMSE[a]	میانگین RI	میانگین RMSE	میانگین RI
کل	۰/۴۴۴		۰/۵۵۰	
۱	۰/۴۷۳*	-۶/۵	۰/۵۷۴	-۴/۲
۲	۰/۲۴۹*	۴۴/۰	۰/۳۳۱*	۳۹/۹

[a]. آماره‌های RMSE و RI به ترتیب نشان دهنده مقادیر میانگین مجذور میانگین مربعات خطا و بهبود نسبی یازده مدل شبکه‌های عصبی مصنوعی هستند.

* معنی‌دار بودن اختلاف بین مقادیر RMSE کل داده و کلاس‌های مختلف هر گروه در مرحله آموزش و آزمون را نشان می‌دهد.

به طور کلی از مقایسه نتایج جداول (۵ تا ۱۲) چنین برداشت می‌شود که از ۴۹ کلاس آورده شده در ۲۹ کلاس در هر دو مجموعه آموزش و آزمون و در ۵ کلاس در یک مجموعه (آموزش یا آزمون) گروه‌بندی داده سبب افزایش عملکرد توابع در پیش‌بینی CEC خاک گردید (جدول ۱۴). همچنین ارزیابی نتایج به دست آمده از مقایسه میانگین غیرجفتی بین آماره RI حاکی از آن است که در مجموع، گروه بافت نسبت به سایر گروه‌ها با دقت بالاتری CEC را پیش‌بینی کرد (جدول ۱۴). علت این امر

شاید همگونی بیش‌تر ویژگی‌های خاک در گروه بافت در مقایسه با سایر گروه‌ها باشد. بهبود نسبی گروه‌های رده، گروه بافت و برخی از ویژگی‌های مناسب برای تعریف واحد خاک و کشاورزی در برابر گروه کلاس بافت به طور معنی‌داری بیش‌تر بود (جدول ۱۴). مهاجر و همکارانش در پژوهشی (۱۳۸۷) با تفکیک خاک‌ها در دو گروه عنوان کردند که در هر دو گروه CEC بیشترین ضریب تغییرات را نشان می‌دهد و ضریب تغییرات کربنات کلسیم و کاتیون‌های تبادلی بین دو فامیل مشابه است.

جدول ۱۴. نتایج ارزیابی بهبود نسبی توابع ایجاد شده در گروه‌های مختلف

گروه	گروه بافت	رده	برخی از ویژگی‌های مناسب برای تعریف واحد خاک و کشاورزی	کربنات کلسیم	افق	ظرفیت آب قابل استفاده	رژیم رطوبتی خاک	کلاس بافت
میانگین RI[a]	۴۱/۰۸	۳۸/۴۲	۳۵/۲۶	۲۴/۹۶	۱۲/۲۹	-۹/۶۶	-۱۲/۶۹	-۲۵/۸۴[b]

[a]. آماره RI نشان دهنده مقادیر میانگین بهبود نسبی کلاس‌های هر گروه نسبت به کل داده است.

[b]. نشان دهنده وجود تفاوت معنی‌دار بین کلاس بافت و گروه‌های رده، گروه بافت و برخی از ویژگی‌های مناسب برای تعریف واحد خاک و کشاورزی است.

نتیجه‌گیری

با توجه به مشکل بودن اندازه‌گیری CEC خاک، در مطالعه حاضر سعی شد از طریق گروه‌بندی داده‌ها بر پایه تعداد خصوصیات متفاوت بیش‌تر قبل از ایجاد مدل، علاوه بر دست یافتن به مدل‌های با دقت تخمین بالاتر، بهترین نوع از گروه‌بندی نیز معرفی شود. نیل به این هدف شاید این امکان را برای کاربران فراهم آورد که بتوانند با انتخاب نوع بهتر گروه‌بندی راحت‌تر و با صرف وقت و هزینه کم‌تری CEC خاک را با بالاترین دقت برآورد کنند. بر اساس نتایج به دست آمده توابع ایجاد شده در گروه بافت نسبت به توابع ایجاد شده در سایر گروه‌ها با میانگین خطای کم‌تری CEC را پیش‌بینی کردند. احتمالا به این دلیل که همگونی به وجود آمده در خصوصیات خاک بر اثر این نوع از گروه‌بندی بیش‌تر از انواع دیگر گروه‌بندی بود. به این دلیل که در طراحی شبکه‌های عصبی مصنوعی نیازی به نوع خاصی از معادلات نیست و با ایجاد رابطه مناسب بین داده‌های خروجی و ورودی می‌توان به نتایج صحیح و مناسب دست یافت. پیشنهاد می‌شود با توجه به وجود مدل‌های مختلف هوش مصنوعی موجود و امکان استفاده از آنها به جای شبکه عصبی مصنوعی همانند Deep learning و الگوریتم ژنتیک تخمین CEC مورد ارزیابی قرار گیرد.

منابع

بیات، حسین؛ دوات‌گر، ناصر؛ و معلمی، سمیه. (۱۳۹۰). استفاده از سطح ویژه برای بهبود تخمین ظرفیت تبادل کاتیونی خاک از طریق شبکه‌های عصبی مصنوعی. *نشریه دانش آب و خاک*، ۴۰، ۱۱۹-۱۰۶.

خداوردی‌لو، حبیب؛ و حسینی عربلو، نوش آفرین. (۱۳۹۲). ایجاد، ارزیابی و مقایسه توابع انتقالی کلاسی و پیوسته برای برآورد ظرفیت تبادل کاتیونی خاک در چند کلاس بافتی. *مجله علوم و فنون کشاورزی و منابع طبیعی، علوم آب و خاک*، ۶۷، ۳۲۰-۳۱۱.

زراعت پیشه، مجتبی؛ خرمالی، فرهاد؛ کیانی، فرشاد؛ و پهلوانی، محمد هادی. (۱۳۹۱). مطالعه کانی‌های رسی در خاک‌های تشکیل شده بر روی مواد مادری لسی در یک توالی اقلیمی در استان گلستان. *مجله پژوهش‌های خاک (علوم خاک و آب)*، ۳، ۳۱۶-۳۰۳.

فولادمند، حمید رضا. (۱۳۸۶). تخمین ظرفیت تبادل کاتیونی خاک با استفاده از برخی خصوصیات فیزیکوشیمیایی خاک. *مجله علوم کشاورزی و منابع طبیعی*، ۱، ۸-۱.

محمدی، مهدی (۱۳۸۵). خاک شناسی کشاورزی. تهران. *مرکز نشر سپهر*. ۲۴۵ صفحه.

معلمی، سمیه؛ و دوات‌گر، ناصر. (۱۳۸۹). مقایسه توابع انتقالی رگرسیونی و شبکه عصبی مصنوعی در برآورد گنجایش تبادل کاتیونی خاک-های گیلان. *مجله علوم و فنون کشاورزی و منابع طبیعی*، ۵۵، ۱۸۱-۱۶۹

مقدم، محمدرضا. (۱۳۸۰). آمار و توصیف اکولوژی پوشش گیاهی. انتشارات دانشگاه تهران. ۲۸۵ صفحه.

مهاجر، رضا؛ صالحی محمد حسن؛ و بیگی هرچگانی حبیب‌اله. (۱۳۸۸). تخمین ظرفیت تبادل کاتیونی خاک با استفاده از رگرسیون و شبکه عصبی و اثر تفکیک داده‌ها بر دقت و صحت توابع. *علوم و فنون کشاورزی و منابع طبیعی علوم آب و خاک*، ۴۹، ۹۷-۸۳.

نوربخش، فرشید؛ جلالیان، احمد؛ و شریعتمداری، حسین. (۱۳۸۲). تخمین گنجایش تبادلی خاک با استفاده از برخی ویژگی‌های فیزیکوشیمیایی خاک. *علوم و فنون کشاورزی و منابع طبیعی*، ۳، ۱۱۷-۱۰۷

هزارجریبی، ابوطالب؛ نصرتی کاریزک، فهمیه؛ عبداله نژاد، کامل؛ قربانی، خلیل. (۱۳۹۲). بررسی امکان پیش‌بینی ظرفیت تبادل کاتیونی خاک با استفاده از پارامترهای زودیافت خاک. آب و خاک، ۲۷(۴)، ۷۱۹-۷۱۲.

Amini, M., Abbaspour, K. C., Khademi, H., Fathianpour, N., Afyuni, M., & Schulin, R. (2005). Neural network models to predict cation exchange capacity in arid regions of Iran. *European Journal of Soil Science*, 56, 551-559.

Asadu, C. L. A., & Akamigbo, F. O. R. (1990). Relative contribution of organic matter and clay fractions to cation exchange capacity of soils in southern Nigeria. Sa*maru: Journal of Agriculture Research*, 7, 17–23.

Asadu, C. L. A., Diels, J., & Vanlauwe, B. (1997). A comparison of the contributions of clay, silt and organic matter to the effective CEC of soils of sub-Saharan Africa. *Soil Science*, 162, 785-794.

Bayat, B. M., Neyshabouri, R., Hajabbasi, M. A., Mahboubi, A. A., & Mosaddeghi, M. R. (2008). Comparing neural networks, linear and nonlinear regression techniques to model penetration resistance. *Turkish Journal of Agriculture and Forestry*, 32, 425–433.

Bayat, H., Davatgar, N., & Jalali, M. (2013). Prediction of CEC using fractal parameters by artificial neural networks. *International Agrophysics*, 28, 143-152.

Bayat, H., Davatgar, N., & Moallemi, S. (2012). Using of specific surface to improve the prediction of Ssoil CEC by artificial neural networks. *Journal of Water and Soil Science*, 21(4), 105-119. (In Persian).

Cai, Z., Yang, C., Du, X., Zhang, L., Wen, S., & Yang, Y. (2023). Parent material and altitude influence red soil acidification after converted rice paddy to upland in a hilly region of southern China. *Journal of Soils and Sediments*, 1-13.

Costa, J. L., Aparicio, V., & Cerdà, A. (2015). Soil physical quality changes under different management systems after 10 years in the Argentine humid pampa. *Solid Earth*, *6*(1), 361-371.

Drake, E. H., & Motto, H. L. (1982). An analysis of the effect of clayand organic matter content on the cation exchange capacity of New Jersey soils. *Soil Science*, 133, 281–288.

Edmeades, D. C. (1982). Effects of lime on effective cation exchange capacity and exchangeable cations on a range of New Zealand soils. *New Zealand Journal of Agricultural Research*, 25, 27-33.

Emamgholizadeh, S., Bazoobandi, A., Mohammadi, B., Ghorbani, H., & Sadeghi, M. A. (2023). Prediction of soil cation exchange capacity using enhanced machine learning approaches in the southern region of the Caspian Sea. *Ain Shams Engineering Journal*, 14(2), 101876.

FAO/IIASA/ISRIC/ISS-CAS/JRC. (2012). *Harmonized World Soil Database* (version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Firat Pulat, H., Tayfur, G., & Yukselen-Aksoy, Y. (2014). Developing cation exchange capacity and soil index properties relationships using a neuro-fuzzy approach. *Bulletin of Engineering Geology and the Environment*, 73, 1141–1149.

Fouladmand, H. R. (2007). Estimation of soil cation exchange capacity using some soil physicochemical properties. *Journal of Agricultural Sciences and Natural Resources*, 1, 1-8. (In Persian).

Hezarjaribi, A., Nosrati Karizak, F., Abdollahnezhad, K., & Ghorbani, Kh. (2013). The prediction possibility of soil cation exchange capacity by using of easily accessible soil parameters. *Journal of Water and Soil*, 27(4), 712-719. (In Persian).

Ho, R. (2006). *Handbook of univariate and multivariate data analysis and interpretation with SPSS*. Chapman and Hall/CRC, 403p.

Horn, A. L., Düring, R. A., & Gäth, S. (2005). Comparison of the prediction efficiency of two pedotransfer functions for soil cation-exchange capacity. *Journal of Plant Nutrition and Soil Science*, 168, 372-374.

Ibrahim, O. M., El-Gamal, E. H., Darwish, K. M., & Kianfar, N. (2022). Modeling main and interactional effects of some physiochemical properties of Egyptian soils on cation exchange capacity via artificial neural networks. *Eurasian Soil Science*, 55(8), 1052-1063.

Kar, R., Bose, P. C., and Bajpai, A. K. (2008). Prediction of cation exchange capacity of soils of mulberry garden based on their clay and organic carbon content in Eastern India. *Journal of Crop and Weed*, 4(2), 47-49.

Keshavarzi, A., Sarmadian, F., Sadeghnejad, M., & Pezeshki, P. (2010). Developing pedotransfer functions for estimating some soil properties using artificial neural network and multivariate regression approaches. *Pro Environment*, 3, 322 – 330.

Khodaverdiloo, H., & Hosseini Arablu, N. (2014). Derivation, validation and comparison of class and continuous pedotransfer functions for predicting soil cation exchange capacity in several textural classes. *Journal of Science and Technology of Agriculture and Natural Resources*, 18(67), 311-320. (In Persian).

Khormali, F., Abtahi, A., Mahmoodi, S., & Stoops, G. (2003). Argillic horizondevelopment in calcareous soils of arid and semiarid regions of southern Iran. *Catena*, 53, 273-301.

Khormali, F., Ghorbani, R., and Amoozadeh Omrani, R. (2005). Variations in soil properties as affected by deforestation on loess-derived hillslopes of Golestan Province, northern Iran. Sociedade and Natureza, Uberlândia, Brazil, Pp: 440-445.

Kissel, D. E., & Sonon, L., (eds). (2008) . Soil Test Handbook for Georgia. http://aesl.ces.uga.edu/publications/soil/STHandbook.pdf

Krogh, L., Breuning-madsen, H., & Greve, M. H. (2000). Cation exchange capacity pedotransfer function for Danish soils. *Plant and Soil*, 50, 1-12.

MacDonald, K.B. (1998). Development of Pedotransfer Functions of Southern Ontario Soils, pp: 1–23. Report from greenhouse and processing crops research center, Harrow, Ontario, No: 01686-8- 0436

Madeira, M., Auxtero, E., & Sousa, E. (2003). Cation and anion exchange properties of Andisols from the Azores, Portugal, as determined by the compulsive exchange and the ammonium acetate methods. *Geoderma*, 117-225.

Manrique, L. A., Jones, C. A., & Dyke, P. T. (1991). Predicting cation exchange capacity from soil physical and chemical properties. Soil Science Society of America Journal, 55:787–794.

Marschner, P., & Rengel, Z. (2023). *Nutrient availability in soils*. In Marschner's Mineral Nutrition of Plants (pp. 499-522). Academic press.

Matos, A. T., Fontes, M. P. F., Costa, L. M., & Martinez, M. A. (2001). Mobility of heavy metals as related to soil chemical and mineralogical characteristics of Brazilian soils. *Environmental Pollution*, 111, 429-435.

Merdun, H., Meral, O. C., & Apan, R. M. (2006). Comparison of artificial neural network and regression pedo transfer funection for predict of water retention and standard hydraulic conductivity. *Soil and Tillage Research*, 90, 108-116.

Mishra, G., Das, J., & Sulieman, M. (2019). Modelling soil cation exchange capacity in different land-use systems using artificial neural networks and multiple regression analysis. *Current Scienc*e, 116(12), 2020-2027.

Moallemi, S., & Davatgar, N. (2011). Comparison of artificial neural network and regression pedotransfer functions for prediction of cation exchange capacity in Guilan Province soils. *Journal of Science and Technology of Agriculture and Natural Resources*, 15(55), 169-182. (In Persian).

Moghadam, M. R. (2001). Statistics and description of vegetation ecology. Tehran University Publications. 285 pages. (In Persian).

Mohajer, R., Salehi, M., & Beigi Herchegani, H. (2009). Estimating soil cation exchange capacity (in view of pedotransfer functions) using regression and artificial neural networks and the effect of data partitioning on accuracy and precision of functions. *Journal of Science and Technology of Agriculture and Natural Resources*, 13(49), 83-97. (In Persian).

Mohammadi, M. (2006). Agricultural soil science. Tehran. *Sepehr Publishing Center*, 245 pages. (In Persian).

Morras, H. J. M. (1995). Mineralogy and cation exchange capacity of the fine silt fraction in teo soils from the Chaco region (Argentina). *Geoderma*, 64, 281-295.

Mukherjee, A., & Zimmerman. A.R. (2013). Organic carbon and nutrient release from a range of laboratory-produced biochars and biochar–soil mixtures. Geoderma.193–194(0):122- 30.

Nourbakhsh, F., Jalalian, A., & Shariatmadari, H. (2003). Estimation of cation exchange capacity from some soil physical and chemical properties. *Journal of Science and Technology of Agriculture and Natural Resources*, 7(3), 107-118. (In Persian).

Obalum, S. E., Watanabe, Y., Igwe, C. A., Obi, M. E., & Wakatsuki, T. (2012). Improving on the prediction of cation exchange capacity for highly weathered and structurally contrasting tropical soils from their fine-earth fractions. *Taylor and Francis Group*, 44,1831–1848.

Olorunfemi, I., Fasinmirin, J., & Ojo, A. (2016). Modeling cation exchange capacity and soil water holding capacity from basic soil properties. *Eurasian Journal of Soil Science*, *5*(4), 266-274.

Pachepsky, Y. A., & Rawls, W. J. (1999). Accuracy and reliability of pedotransfer functions as affected by grouping soils. *Soil Science Society of America Journal*, 63,1748–1757.

Pulido, M., Schnabel, S., Lavado Contador, J. F., Lozano- Parra, J., & Gonzalez, F. (2018). The impact of heavy grazing on soil quality and pasture production in rangelands of SW Spain. *Land Degradation & Development*, *29*(2), 219-230.

Rahimi Lake, H., Akbarzadeh, A., & Taghizadeh Mehrjardi, R. (2009). Development of pedo transfer functions (PTFs) to predict soil physico-chemical and hydrological characteristics in southern coastal zones of the Caspian Sea. *Journal of Ecology and The Natural Environment*, 1(7), 160-172.

Salchow, E., Lal, R., Fausey, N. R., & Ward, A. (1996). Pedotransfer functions for variable alluvial soils in southern Ohio. *Geoderma*, 73, 165-181.

Sarmadian, F., Azimi, S., Keshavarzi Ahmadi, A. (2013). Neural computing model for prediction of Soil Cation Exchange Capacity: A Data Mining Approach. *International Journal of Plant Production*, 4, 1706-1712.

Sayegh, A. H., Khan, N. A., Khan, P., & Ryan, J. (1978). Factors affecting gypsum and CEC determinations in gypsiferous soils. *Soil Science*, 125(5), 294-300.

Seybold, C. A., Grossman, R. B., & Reinsch, T. G. (2005). Preicting Cation Exchange Capacity for Soil Survey Using Linear Models. *Soil Science Society of America Journal*, 69, 856-86.

Soares, M. R., Alleoni, L. R. F., Vidal-Torrado, P., & Cooper, M. (2005). Mineralogy and ion exchange properties of the particle-size fractions of some Brazilian soils in tropical humid areas. *Geoderma*, 125, 355–367.

Sparks, D. L. (1995). *Enviromental Soil Chemistry*. In Academic Press Inc. University of Delaware London.

Syers, J. K., Campbell, A. S., & Walker, T. W. (1970). Contribution of organic carbon and clay to cation exchange capacity in a chronosequence of sandy soils. *Plant and Soil*, 33, 104–112.

Tamari, S., Wosten, J. H. M., & Ruz-suarez, J. C. (1996). Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Science Society of America Journal*, 60, 1732-1741.

Tang, L., Zeng, G. M., Nourbakhsh, F., & Shen, G. L., (2008). Artificial neural network approach for predicting cation exchange capacity in soil based on physico-chemical properties. *Environmental Engineering Science*, 26(2), 1-10.

Tessler, M., David, F. J., Cunningham, S. W., & Herstoff, E. M. (2023). Rewilding in miniature: suburban meadows can improve soil microbial biodiversity and soil health. *Microbial Ecology*, 1-10.

Thompson, M. L., Zhang, H., Kazemi, M., & Sandor. J. A. (1989). Contribution of organic matter to cation exchange capacity and specific surface area of fractionated soil materials. *Soil Science*, 148: 250-257.

Turpault, M. P., Bonnaud, P., Fichter, J., Ranger, J., & Dambrine, E. (1996). Distribution of cation exchange capacity between organic matter and mineral fractions in acid forest soils (Vosges mountains, France). *European Journal of Soil Science*, 47, 545-556.

Wagner, B., Hennings, V., Muller, U., Wessolek, G., & Plagge, R. (2001). Evaluation of pedotransfer functions for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma*, 102, 275-279.

Wilding, L. P., Smeck, N. E., & Hall, G. F. (1983). Pedogenesis and soil taxonomy. I. Concepts and interactions. *Elsevier Publishing Company*, 303p.

Yukselen, Y., & Kaya, A. (2006). Prediction of cation exchange capacity from soil index properties. *Clay Minerals*, 41, 827–837.

Zeraat Pishe, M., Khormali, F., Kiani, F., & Pahlavani, M. H. (2013). Studying clay minerals in soils formed on loess parent materials in a climatic gradient in Golestan Province. *Soil Research*, 26(3), 303-316. (In Persian).

Evaluation of the Effect of Grouping Based on Different Characteristics on the Performance of Functions in Estimating Soil Cation Exchange Capacity

Abstract

Introduction

The cation exchange capacity of the soil is one of the most important characteristics, as it greatly influences soil fertility. This capacity is influenced by various physical and chemical characteristics of the soil, resulting in significant variability. However, measuring soil cation exchange capacity can be challenging and expensive. The objective of this study is to assess the impact of grouping based on different characteristics on the accuracy of estimating soil cation exchange capacity. Additionally, we aim to introduce a grouping method that yields the best estimation results and compare the effectiveness of linear regression and artificial neural networks in estimating cation exchange capacity.

Material and Methods

For this research, we utilized 45,948 soil samples from a standardized global soils database. Initially, these samples were grouped into different categories. Subsequently, we evaluated nine predictor variables including soil texture components, organic carbon content, calcium sulfate and carbonate levels, apparent specific gravity, base saturation percentage, total exchangeable base cations, and soil reaction in 11 models for both the entire dataset and each group separately. To determine which grouping method yields more accurate estimations for soil cation exchange capacity and which characteristic contributes to better results, transfer functions were developed for each group using all data and then again for each class within the groups. In this study, we employed multilayer perceptron neural networks with one hidden layer and a hyperbolic tangent activation function. The number of cells in the hidden layer ranged from 3 to 8. The network with the optimal number of hidden cells was selected as the final network based on its performance. Additionally, linear regression modeling was conducted using Datafit 9 software to compare its effectiveness with artificial neural networks in estimating cation exchange capacity.

Results and Discussion

The correlation between the cation exchange capacity of soil and clay, apparent specific gravity, the ratio of silt to positive sand, calcium carbonate, total exchangeable base cations, and soil reaction was found to be significant at the 1% level. When comparing the RMSE and RI statistics in two sets of training and test data for the entire dataset and for classes separated based on soil texture groups, it was observed that data separation improved the mean in estimating the cation exchange capacity of the soil. Two methods, linear regression and artificial neural networks, were used to estimate the cation exchange capacity. The results showed that for all coarse, medium, and fine texture groups, the artificial neural network test section had a relative improvement coefficient of 87%, while linear regression had a value of 17%. The grouping of data generally increased the estimation ability of the models. Among the groups studied in this research, the texture group showed higher accuracy in predicting cation exchange capacity compared to other groups. Overall, using functions obtained through grouping proved to be an easy and cost-effective method for estimating cation exchange capacity.

Conclusion

In general, artificial neural networks outperformed linear regression in estimating cation exchange capacity in most classes within each group. This may be due to not requiring a specific type of equation in designing artificial neural networks. By establishing a suitable relationship between output and input data, accurate results can be achieved.

Keywords: Soil database, Linear regression, Artificial neural network, Model reliability