

پیش‌بینی ظرفیت تبادل کاتیونی خاک‌های ایران با استفاده از روش‌های گوناگون

روح‌الله تقی‌زاده مهرجردی^۱، فریدون سرمدیان^{۲*}، علی اصغر ذوالفقاری^۳، اعظم جعفری^۴

۱- استادیار دانشکده کشاورزی و منابع طبیعی دانشگاه اردکان

۲- استاد گروه مهندسی علوم خاک، پردیس کشاورزی و منابع طبیعی دانشگاه تهران

۳- استادیار دانشکده کویرشناسی دانشگاه سمنان

۴- استادیار گروه مهندسی علوم خاک دانشگاه شهید باهنر کرمان

تاریخچه مقاله	چکیده
دریافت: ۱۳۹۲/۱۰/۰۱	اندازه‌گیری ظرفیت تبادل کاتیونی خاک در سطوح وسیع، بسیار پرهزینه و وقت‌گیر است. برآورد این خصوصیت به‌وسیله ویژگی‌های زود یافت خاک، از طریق روش‌های پارامتریک و غیر پارامتریک (توابع انتقالی) می‌تواند رویکرد مناسبی باشد. در این پژوهش، روش پارامتریک رگرسیون‌های خطی و غیر خطی و روش‌های غیر پارامتریک شبکه‌های عصبی مصنوعی، رگرسیون درختی و فازی عصبی در تخمین ظرفیت تبادل کاتیونی خاک مورد استفاده قرار گرفت. برای این منظور ۱۷۷۰ نمونه خاک از مناطق مختلف ایران انتخاب شدند که ۱۴۱۴ عدد برای آموزش و ۳۵۶ عدد از آن‌ها به عنوان داده‌های آزمون مدل‌ها استفاده شدند. بررسی همبستگی‌ها نشان داد که پارامترهای رس و درصد ماده آلی خاک بیش‌ترین ارتباط را با ظرفیت تبادل کاتیونی خاک دارند؛ بنابراین این ویژگی‌ها به عنوان متغیر مستقل ورودی (ویژگی زود یافت) و ظرفیت تبادل کاتیونی به عنوان متغیر وابسته خروجی انتخاب شدند. ریشه میانگین مربعات خطا و ضریب تبیین در روش رگرسیون خطی برابر ۴/۷۴ و ۰/۵۰ و در روش رگرسیون غیر خطی برابر ۴/۷۱ و ۰/۵۲ بود که نشان می‌دهد که هر دو روش با دقت نسبتاً خوب و یکسانی قادر به پیش‌بینی ظرفیت تبادل کاتیونی خاک می‌باشند؛ همچنین نتایج نشان داد که روش رگرسیون غیر خطی فقط سبب بهبود ۰/۶ درصدی دقت پیش‌بینی ظرفیت تبادل کاتیونی خاک شده است. نتایج نشان داد که استفاده از روش‌های شبکه عصبی مصنوعی سبب بهبود معنی‌داری در دقت برآورد ظرفیت تبادل کاتیونی خاک نمی‌شود. بیش‌ترین بهبود در پیش‌بینی مدل به نسبت توابع انتقالی خطی در روش شبکه عصبی پس انتشار
پذیرش: ۱۳۹۳/۰۹/۰۵	
کلمات کلیدی: ظرفیت تبادل کاتیونی، توابع انتقالی، شبکه‌های عصبی، فازی عصبی، رگرسیون درختی	
* عهده‌دار مکاتبات E-mail: fsarmad@ut.ac.ir	

مشاهده شد. این روش سبب بهبود ۳ درصدی پیش‌بینی ظرفیت تبادل کاتیونی خاک گردید. دقت برآورد روش درخت تصمیم، اندک بهتری از روش‌های شبکه عصبی مصنوعی بود. بهبود نسبی این روش نسبت به رگرسیون خطی برابر با $4/4$ درصد بود؛ اما بیشترین بهبود نسبی در روش فازی عصبی مشاهده شد. این روش سبب کاهش ۱۵ درصدی خطا به نسبت معادلات رگرسیونی خطی گردید؛ لذا این نتایج نشان می‌دهد که یکی از بهترین روش‌ها در پیش‌بینی ظرفیت تبادل کاتیونی خاک‌های ایران، روش فازی عصبی می‌باشد.

مقدمه

ظرفیت تبادل کاتیونی^۱ یکی از ویژگی‌های مهم خاک است که بسیاری از عملکردهای اساسی خاک در ارتباط با نوسانات اسیدیته، مواد غذایی و آلاینده‌ها را کنترل می‌نماید و به همین دلیل این ویژگی، یکی از شاخص‌های کلیدی در مدیریت حاصل‌خیزی خاک و بهره‌وری زمین محسوب می‌شود (۱۰). ظرفیت تبادل کاتیونی خاک‌های زراعی می‌تواند تفسیر روشن‌تر و کامل‌تری از فرایندهای تغذیه گیاه، کاربرد کودها و اصلاح شیمیایی خاک‌ها را در اختیار پژوهشگران قرار دهد. معمولاً روش‌های اندازه‌گیری این پارامتر نسبتاً پر زحمت، پر هزینه و البته بسیار وقت‌گیر می‌باشند (۳۰ و ۳۸). به همین دلیل تاکنون بسیاری از پژوهشگران تحقیقات خود را بر روش‌های غیر مستقیم برآورد ظرفیت تبادل کاتیونی با استفاده از ویژگی‌های زودیافت خاک متمرکز نموده‌اند (۳۸). برقراری معادلات چند متغیره بین ویژگی‌های پایه‌ای خاک و مقادیر ظرفیت تبادل کاتیونی به دست آمده در آزمایشگاه، یکی از روش‌هایی است که می‌تواند تخمین قابل قبولی از ظرفیت تبادل کاتیونی خاک مناطق همگن را در اختیار پژوهشگران قرار دهد. مطالعات گوناگونی در زمینه استفاده از معادلات رگرسیونی در برآورد ظرفیت تبادل کاتیونی خاک در ایران و دنیا صورت گرفته است؛ به

عنوان مثال کریمیان (۲۰) سهم رس و مواد آلی در ظرفیت تبادل کاتیونی ۱۵۰ نمونه از خاک‌های آهکی مناطقی استان فارس را در قالب توابع انتقالی بیان نمود. در مطالعه دیگری معماریان فرد و بیگی هرچگانی (۳۱) پارامترهای درصد مواد آلی خاک، رطوبت اشباع، درصد شن و رس را به عنوان متغیرهای ورودی در به-دست آوردن معادلات رگرسیونی مورد استفاده قرار دادند. این پژوهشگران ضریب تبیین $0/69$ را بین ظرفیت تبادل کاتیونی پیش‌بینی شده و اندازه‌گیری بدست آورند؛ آنان همچنین متوسط خطای پیش‌بینی توابع انتقالی به دست آمده را برابر با $4/69$ به دست آوردند. از معایب توابع انتقالی و به طور کلی توابع پارامتریک در تخمین ویژگی‌های دیریافت خاک، تعیین مقدار عددی هر یک از پارامترها است که از طریق بهترین برازش معادلات مذکور بر داده‌ها به دست می‌آید. در مقابل، روش‌های غیر پارامتریک مانند مدل‌های شبکه عصبی مصنوعی معمولاً تخمین‌های دقیق‌تری از مقدار ظرفیت تبادل کاتیونی ارائه نموده‌اند (۳، ۲۵ و ۳۱). معماریان فرد و بیگی هرچگانی (۳۱) از شبکه عصبی مصنوعی برای برآورد ظرفیت تبادل کاتیونی استفاده کردند. این پژوهشگران بیان کردند که دقت برآورد ظرفیت تبادل کاتیونی با استفاده از شبکه عصبی مصنوعی بیشتر از معادلات رگرسیونی خطی می‌باشد. در مطالعه دیگری تقی‌زاده و همکاران (۳۹) نیز بیان داشتند که دقت شبکه

1- Cation Exchange Capacity (CEC)

از چندین مدل تجربی، شبکه عصبی و مدل نروفازی تخمین زدند. نتایج نشان داد که مدل نروفازی دارای کارایی بیشتری بر اساس معیار ریشه مربعات خطا نسبت به دو مدل دیگر بود. کالکان^۴ و همکاران (۱۹) برای پیش‌بینی ویژگی مکانیکی خاک‌های فشرده شده از توزیع اندازه ذرات ۸۴ نمونه خاک استفاده کردند. برای مدل کردن ویژگی مذکور از دو روش نروفازی و شبکه عصبی استفاده نمودند. نتایج نشان داد که مدل نروفازی دارای کارایی بالاتری نسبت به مدل شبکه عصبی مصنوعی بود. کیسی^۵ (۲۴) نشان داد که مدل نروفازی در مقایسه با مدل‌های شبکه عصبی مصنوعی و رگرسیون چند متغره خطی دارای دقت بالاتری در پیش‌بینی مقدار رسوب معلق بود.

از آنجا که تاکنون مطالعات کمتری در ایران با هدف مقایسه عملکرد روش پارامتریک توابع رگرسیون خطی و روش‌های غیر خطی شبکه عصبی مصنوعی، فزای عصبی و رگرسیون درختی در برآورد ظرفیت تبادل کاتیونی خاک‌های کل کشور (۱۷۷۰ نمونه) انجام گرفته است، بنابراین در این پژوهش ضمن بررسی سودمندی کاربرد روش توابع رگرسیونی، دقت تخمین این روش با پنج روش شبکه عصبی پسانتشار، پایه شعاعی و آبشاری، مدل نروفازی و رگرسیون درختی نیز مورد ارزیابی قرار خواهد گرفت.

مواد و روش‌ها

نمونه برداری

تعداد ۱۷۷۰ نمونه خاک از استان‌های مختلف ایران از افق‌های ژنتیکی خاک برداشت شد. از این تعداد، ۳۵۶ عدد خاک (۲۰ درصد خاک‌ها) به صورت تصادفی انتخاب و به عنوان خاک‌های مورد آزمون جهت تخمین ظرفیت تبادل کاتیونی خاک در نظر گرفته شدند و باقی خاک‌ها برای مشتق سازی معادلات رگرسیونی و یا برای

عصبی مصنوعی در برآورد ظرفیت تبادل کاتیونی خاک بیشتر از توابع انتقالی خطی (معادلات رگرسیونی خطی) می‌باشد (۶ و ۱۶). به تازگی ذوالفقاری و همکاران (۴۵) شبکه عصبی مصنوعی را با یک روش غیر پارامتریک دیگری تحت عنوان k نزدیکترین همسایگی در برآورد ظرفیت تبادل کاتیونی خاک مقایسه کردند. این پژوهشگران اختلاف معنی‌داری را بین دقت برآورد دو روش مشاهده نکردند؛ هرچند که با توجه به سادگی الگوریتم روش k- نزدیکترین همسایگی آن‌ها بیان داشتند که این روش به نسبت شبکه عصبی مصنوعی برتری دارد.

رگرسیون درختی یکی دیگر از روش‌های غیر پارامتریک است که برای ایجاد توابع انتقالی مورد استفاده قرار می‌گیرد این روش برای اولین بار توسط بریمن^۱ و همکاران (۸) گسترش داده شد. پژوهشگران مختلفی از این روش جهت پیش‌بینی ویژگی‌های خاک از جمله شوری (۴۰)، ماده آلی (۳۴)، واحدهای خاک (۴۱) استفاده کرده‌اند و همگی بیان داشتند که مدل‌های درختی در مقایسه با سایر روش‌ها دارای مزایایی از جمله قابلیت تفسیر بالا، عدم وابستگی به داده‌های پرت، انعطاف‌پذیری به روابط غیر خطی و عدم وابستگی به توزیع داده‌ها می‌باشند (۳۵).

استفاده از منطق فازی یکی دیگر از روش‌هایی است که بر اساس آن می‌توان ارتباط بین ورودی‌ها و خروجی‌ها را با استفاده از سیستم استنتاج فازی برقرار نمود (۲ و ۲۲). از انواع مدل‌های نروفازی می‌توان به مدل‌های سیستم فازی تطبیق‌پذیر مبتنی بر شبکه عصبی اشاره کرد که برای اولین بار توسط ژانگ^۲ (۱۸) معرفی شد. مطالعات اخیر، کارایی بیشتر مدل‌های نروفازی را در مقایسه با دیگر توابع انتقالی نشان می‌دهد. به عنوان مثال مقدم‌نیا^۳ و همکاران (۳۲) میزان تبخیر را با استفاده

1- Breiman *et al.*

2- Jang

3- Moghaddamnia *et al.*

4- Kalkan *et al.*

5- Kisi

برآورد ظرفیت تبادل کاتیونی خاک‌ها مورد استفاده قرار گرفت. یکی از مشکلات موجود در استفاده از توابع انتقالی غیر پارامتریک این است که یک واحد اختلاف در یک متغیر (صفت) ممکن است که به اندازه یک واحد اختلاف در متغیر دیگر اهمیت نداشته باشد. به عنوان مثال مقدار رس به عنوان یک متغیر ممکن است که از ۵ تا ۷۰ درصد تغییر نماید در حالی که تغییرات کرن آلی بسیار کم‌تر از تغییرات رس در خاک‌ها می‌باشد. بنابراین در این پژوهش پیش از انجام هر گونه محاسباتی، ابتدا داده‌های مورد مطالعه با استفاده از معادله (۱) استاندارد شدند. به عبارت دیگر رابطه (۱) سبب می‌شود که تمامی داده‌ها در محدوده ۰/۱ تا ۰/۹ قرار گیرند (۲۳).

$$y = 0.8 \times \frac{x_i - x_{min}}{x_{max} - x_{min}} + 0.1 \quad (1)$$

که x_i خاک نام از متغیر x و x_{min} و x_{max} به ترتیب کوچک‌ترین و بزرگ‌ترین مقادیر متغیر x می‌باشند.

سومین روش مورد استفاده در این پژوهش، روش غیرپارامتریک نروفازی بود. این مدل یک سیستم فازی را در یک ساختار عصبی اجرا می‌کند و برای فرآیند آموزش از ترکیبی شامل روش‌های آموزش شامل شیب نزولی و حداقل مربعات خطا استفاده می‌کند (۱۲ و ۲۴). علاوه به سه روش ذکر شده، در پژوهش حاضر از روش غیر پارامتریک رگرسیون درختی نیز در ایجاد توابع انتقالی برای پیش‌بینی ظرفیت تبدالی کاتیونی خاک استفاده شد. در این روش، مجموعه‌ای از شرط‌های منطقی به صورت یک الگوریتم با ساختار درختی برای طبقه‌بندی یا پیش‌بینی کمی یک متغیر به کار می‌رود. ایجاد درخت تصمیم شامل دو مرحله است که عبارتند از ایجاد و رشد درخت و مرحله توقف و هرس می‌باشد. هدف از مرحله دوم به حداقل رساندن خطای پیش‌بینی است. جهت طراحی رگرسیون درختی از نرم‌افزار کیوبیست (۳۶) استفاده شد و در نهایت با توجه به

آموزش روش‌های غیر پارامتریک مورد استفاده قرار گرفتند. کلیه نمونه‌ها پس از هوا خشک شدن و کوبیده شدن از الک ۲ میلی‌متری عبور داده شدند و برای انجام آزمون‌های فیزیکی و شیمیایی آماده گردیدند. توزیع اندازه ذرات خاک با استفاده از روش هیدرومتري (۱۴)، ماده آلی به روش سوزاندن تر (۴۴) و اندازه گیری ظرفیت تبادل کاتیونی نمونه‌های خاک به روش باور انجام شد (۷).

مراحل تحقیق

گام نخست: بر اساس تحلیل‌های همبستگی بین متغیرهای ورودی و ظرفیت تبادل کاتیونی، پارامترهای موثر بر آن انتخاب گردیدند؛ سپس بررسی هم‌راستایی بین داده‌های مستقل صورت گرفت. به منظور بررسی هم‌راستایی از رابطه هم‌بستگی بین متغیرهای مستقل استفاده شد و متغیرهایی که دارای هم‌بستگی زیادی با یک‌دیگر بودند، حذف گردیدند (۱۵). در این مطالعه بر اساس هم‌بستگی بین متغیرهای مستقل، دو متغیر درصد مواد آلی و رس خاک به عنوان متغیرهای مستقل در برآورد ظرفیت تبادل کاتیونی به عنوان متغیر وابسته استفاده شدند.

گام دوم: در مرحله دوم با توجه به متغیرهای مستقل انتخاب شده، معادلات رگرسیونی بر اساس روش رگرسیون گام به گام در داده‌های آموزش تعیین شد. همچنین مدل‌های رگرسیونی غیر خطی که توسط دیگر پژوهشگران به دست آمده بود نیز مورد آزمون قرار گرفتند. برای این کار در ابتدا پارامترهای معادلات مورد نظر با استفاده از داده‌های آموزش تعیین شدند. برای تعیین این پارامترها از جعبه ابزار Solver نرم‌افزار اکسل استفاده شد.

در این مطالعه علاوه بر استفاده از رگرسیون چند متغیره خطی از دیگر روش‌های غیر پارامتریک نیز برای پیش‌بینی ظرفیت تبادل کاتیونی خاک‌ها استفاده شد. یکی از این روش‌های مورد استفاده در این مطالعه، روش شبکه عصبی مصنوعی بود و انواع شبکه‌های عصبی مصنوعی (پس انتشار، آبخاری و پایه شعاعی) به منظور

نتایج و بحث

خلاصه آماری

جدول ۱ خلاصه‌ای از ویژگی‌های اندازه‌گیری شده خاک‌های بانک داده آموزش و آزمون را نشان می‌دهد. همان‌گونه که مشاهده می‌شود، در خاک‌های مورد مطالعه ویژگی‌هایی مانند توزیع اندازه ذرات و کربن آلی که از مهم‌ترین عوامل موثر بر ظرفیت تبادل کاتیونی خاک‌ها هستند، دامنه وسیعی را نشان می‌دهند. به عنوان مثال کربن آلی در خاک‌های مورد مطالعه بین ۰/۰۱ و ۲/۹۴ درصد متغیر بود و یا درصد رس در خاک‌های مورد مطالعه بین ۵ الی ۷۸ درصد متغیر بود. این امر را می‌توان از دلایل اصلی تنوع مقادیر ظرفیت تبادل کاتیونی خاک‌های مورد مطالعه در این پژوهش نیز به حساب آورد. جدول ۱، میانگین و انحراف معیار تقریباً یکسانی را بین خصوصیات خاک‌های مورد آموزش و بانک داده آزمون نشان می‌دهد. این امر می‌تواند به بهبود تخمین‌های ظرفیت تبادل کاتیونی خاک با استفاده از روش‌های پارامتریک و غیرپارامتریک مورد مقایسه در این تحقیق کمک نماید.

رگرسیون چندمتغیره

کربن آلی و رس خاک بیش‌ترین تاثیر را بر تغییرات ظرفیت تبادل کاتیونی خاک داشت. به عنوان مثال در صورتی که دو متغیر کربن آلی و درصد رس به عنوان متغیرهای مستقل در معادله رگرسیون در نظر گرفته شدند (رابطه ۳)، میانگین ریشه مربعات خطا برابر با ۴/۷۳ بود؛ اما با افزودن متغیر مستقل رس در معادله رگرسیون، میانگین ریشه مربعات خطا به مقدار جزئی کاهش و برابر با ۴/۷۲ به دست آمد. هم‌چنین افزودن متغیر سیلت نیز اثر معنی‌داری بر کاهش میانگین ریشه مربعات خطا نداشت. در این مطالعه از نسبت رس به سیلت به عنوان متغیر مستقل دیگر استفاده شد.

رگرسیون درختی طراحی شده، ظرفیت تبادل کاتیونی خاک با استفاده از این روش پیش‌بینی گردید.

گام سوم: هدف از این مرحله مقایسه کارایی روش‌های گوناگون در برآورد ظرفیت تبادل کاتیونی بود. در تحقیق حاضر از روش ارزیابی واقعی استفاده گردید. بروس^۱ و همکاران (۹) معتقد هستند که ارزیابی بر روی داده‌های آزمون مستقل بهترین نوع ارزیابی مدل‌ها می‌باشد. برای این منظور از آماره‌های ضریب هم‌بستگی و مجذور میانگین مربعات خطا استفاده گردید.

علاوه بر آماره‌های فوق‌الذکر برای بررسی کارایی روش‌های مختلف از شاخص درصد کاهش ریشه مربعات خطا نیز استفاده گردید (رابطه ۲).

$$RI = \left(\frac{RMSE_a - RMSE_b}{RMSE_a} \right) \times 100 \quad (2)$$

که در این مدل $RMSE_a$ میزان ریشه مربعات خطا در روش رگرسیون و $RMSE_b$ ریشه مربعات خطا در سایر روش‌ها است. این آماره نشان‌دهنده کاهش میزان خطا در روش‌های مختلف نسبت به روش رگرسیون می‌باشد؛ در واقع مدلی مناسب‌تر است که شاخص بهبود بیش‌تری داشته باشد (۳ و ۲۹).

گام چهارم: هدف از این مرحله بررسی عدم قطعیت مدل‌ها می‌باشد. با فرض نرمال بودن توزیع آماری پارامترها و با استفاده از روش مونت کارلو تعداد ۱۰۰۰ نمونه تصادفی از پارامترهای ورودی به طور تصادفی انتخاب شد. سپس خروجی معادل این پارامترها به دست آمد. و در نهایت عدم قطعیت مدل از طریق محاسبه شاخص‌های آماری تابع توزیع احتمالات برآورد شد. در این مطالعه از توزیع آماری دهک‌های اول و نهم و میانگین داده‌های تصادفی برای بررسی عدم قطعیت استفاده شد. هرچه فاصله (پهنای باند) بین این پارامترهای آماری بیشتر باشد، نشان‌دهنده عدم قطعیت بالاتر مدل مورد استفاده می‌باشد.

تقی زاده مهرجردی و همکاران: پیش‌بینی ظرفیت تبادل کاتیونی خاک‌های...

جدول (۱) خلاصه آماری داده‌های مورد استفاده

Table (1) Statistical summary of used data

ضریب تغییرات (درصد) (Coefficient of variation)	انحراف استاندارد (Standard Deviation)	میانگین (Mean)	کمینه (Min)	بیشینه (Max)	واحد (Unit)	ویژگی (Properties)
87.97	0.54	0.62	0.04	2.94	%	کربن آلی (SOC)
61.59	20.07	23.58	4.00	90	%	شن (Sand)
40.95	13.18	32.18	4.00	77	%	سیلت (Silt)
42.07	14.80	35.18	5.00	78	%	رس (Clay)
39.91	6.70	16.78	1.85	35	Meq/100g	ظرفیت تبادل کاتیونی (Cation exchange capacity)
93.73	0.62	0.66	0.01	2.97	%	کربن آلی (SOC)
59.64	20.81	34.90	4.72	90	%	شن (Sand)
42.71	13.37	31.28	4	71.4	%	سیلت (Silt)
43.09	14.59	33.87	5	75.05	%	رس (Clay)
42.49	6.97	16.41	2.11	35.1	Meq/100g	ظرفیت تبادل کاتیونی (Cation exchange capacity)

داده‌های آموزش (1414 داده)
(Training data set (1414 data))

داده‌های آزمون (356 داده)
(Validation data set (356 data))

تبادل کاتیونی خاک می‌باشد. این پژوهشگران رابطه ۴ را ارائه کردند. آزمون معادله ارائه شده توسط بن و ون کولن (رابطه ۴) بر خاک‌های مورد مطالعه نشان داد که این معادله به طور نسبی قادر به توصیف ظرفیت تبادل کاتیونی خاک می‌باشد.

$$CEC = 0.3 \times Clay + 4.8 \quad (4)$$

ضریب تبیین این معادله در خاک‌های مورد آزمون برابر با 0.39^{**} به دست آمد. نتایج این مطالعه نشان داد که استفاده از متغیر درصد کربن آلی خاک سبب افزایش ۱۰ درصدی ضریب تبیین ($R^2=0.49^{**}$) معادله

نتایج نشان داد که استفاده از این متغیر می‌تواند سبب افزایش یک درصدی ضریب تبیین معادله رگرسیونی شود؛ بنابراین از این متغیر نیز در برآورد ظرفیت تبادل کاتیونی صرف نظر شد.

$$CEC = 3.5 \times OC + 0.2 \times Clay + 5.2 \quad (3)$$

بل و ون کولن^۱ (۵) بیان داشتند که درصد رس به عنوان مهم‌ترین عامل توصیف کننده تغییرات ظرفیت

1- Bell and Van-Kulen

پیچیده غیر خطی سبب بهبود کارایی معادلات رگرسیونی نمی‌شود. به عنوان مثال استفاده از رابطه (۶) در برآورد ظرفیت تبادل کاتیونی خاک سبب اختلاف معنی‌داری در میانگین ریشه مربعات خطا (۴/۷۱) معادله رگرسیونی نشده است.

شبکه عصبی مصنوعی

مدل شبکه عصبی با یک لایه مخفی که تعداد نرون‌های آن به بین ۲ تا ۲۰ نرون بود، طراحی و ساختار بهینه شبکه به روش سعی و خطا و بر اساس آماره میانگین ریشه مربعات خطا تعیین گردید. با توجه به این که در توسعه رگرسیون خطی در برآورد ظرفیت تبادل کاتیونی خاک، درصد رس و کربن آلی تاثیر معنی‌داری داشتند، لذا در طراحی شبکه عصبی نیز فقط از متغیرهای مذکور استفاده شد. شکل (۱) تغییرات میانگین ریشه مربعات خطا در برآورد ظرفیت تبادل کاتیونی خاک را در تعداد نرون‌های مخفی مختلف با استفاده از شبکه عصبی نشان می‌دهد. با توجه به این شکل ملاحظه می‌شود که کمینه مقدار میانگین ریشه مربعات خطا در زمانی حاصل شد که شبکه ای با سه نرون در لایه مخفی، ۸۵۰ اپوک و تابع سیگموئیدی و خطی در لایه مخفی و خروجی طراحی گردید. مقادیر میانگین ریشه مربعات خطا و ضریب تبیین برآورد ظرفیت تبادل کاتیونی خاک در بهترین ساختار شبکه عصبی مصنوعی، به ترتیب برابر با ۴/۶۱ میلی‌اکی‌والان بر ۱۰۰ گرم و ۰/۵۰** بدست آمد. کاشی و همکاران (۲۱) نیز نشان دادند که کم‌ترین میانگین ریشه مربعات خطا در برآورد ظرفیت تبادل کاتیونی خاک زمانی حاصل می‌شود که ۲ نرون در لایه مخفی وجود داشته باشد. شکل (۱) نشان می‌دهد که تغییرات میانگین ریشه مربعات خطا در تعداد نرون مختلف، دارای روند مشخص و معلومی نیست. علت نامشخص بودن روند تغییرات میانگین ریشه مربعات خطا تصادفی بودن اوزان اولیه می‌باشد. به همین علت بایستی با سعی و خطا بهترین تعداد نرون لایه مخفی را به دست آورد.

رگرسیون به دست آمده می‌شود.

مهاجر و همکاران (۳۳) از متغیرهای مستقل درصد رس و درصد مواد آلی خاک به عنوان متغیرهای مستقل در برآورد ظرفیت تبادل کاتیونی خاک استفاده کردند. این پژوهشگران ضریب تبیین برابر با ۰/۶۷* را برای معادله رگرسیون ارائه شده به دست آوردند. شاید تعداد نمونه زیاد خاک‌های مورد مطالعه در این تحقیق یکی از دلایل عمده بالاتر بودن ضریب تبیین این محققان نسبت به مطالعه حاضر می‌باشد. کاشی و همکاران (۲۱) نشان دادند که استفاده از متغیرهایی همانند چگالی ظاهری خاک، درصد آهک، درصد شن، درصد رس به سیلت و نسبت جذب سدیم امکان‌پذیر می‌باشد. این پژوهشگران بیان داشتند چگالی ظاهری خاک بیش‌ترین تاثیر را در برآورد ظرفیت تبادل کاتیونی خاک دارد.

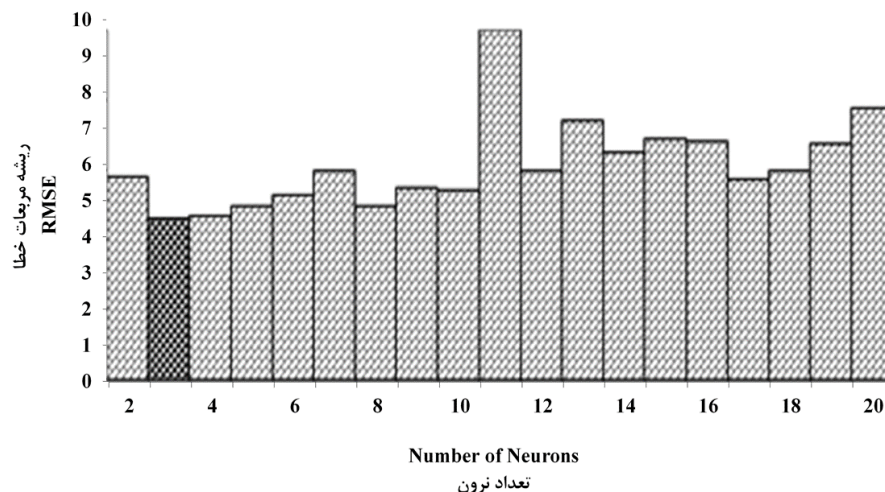
مانریکو^۱ و همکاران (۲۸) رابطه پارامتریک غیر خطی زیر را برای برآورد ظرفیت تبادل کاتیونی خاک با استفاده از داده‌های درصد رس و کربن آلی خاک ارائه کردند.

$$CEC = a_1 \times OC^b + a_2 \times Clay + a_0 \quad (5)$$

مانریکو و همکاران (۲۸) مقادیر ۴/۵۶، ۵/۲ و ۰/۲ را به ترتیب برای پارامترهای a_0 ، a_1 و a_2 به دست آوردند. در این مطالعه پارامترهای رابطه (۵) با توجه خاک‌های بانک داده آموزش تعیین شدند. رابطه (۶) معادله به دست آمده را در خاک‌های مورد مطالعه نشان می‌دهد.

$$CEC = 3.1 \times OC^{1/4} + 0.2 \times Clay + 5.4 \quad (6)$$

مقایسه روابط (۵) و (۶) نشان می‌دهد که ضرایب متغیرهای مستقل دو رابطه به یک‌دیگر نزدیک می‌باشد. هم‌چنین توان کربن آلی خاک (۱/۱۴) نیز به عدد یک نزدیک می‌باشد. به عبارت دیگر این نتایج نشان می‌دهد ارتباط بین ظرفیت تبادل کاتیونی خاک و درصد رس و کربن آلی خاک یک رابطه خطی است و استفاده از روابط



شکل (۱) مقادیر ریشه مربعات خطا برای تعداد نرون متفاوت در پیش‌بینی پارامتر ظرفیت تبادل کاتیونی در روش شبکه عصبی مصنوعی پرسپترون چند لایه

Figure (1) Root mean square error values for different neuron to predict cation exchange capacity using multi-layer perceptron artificial neural network

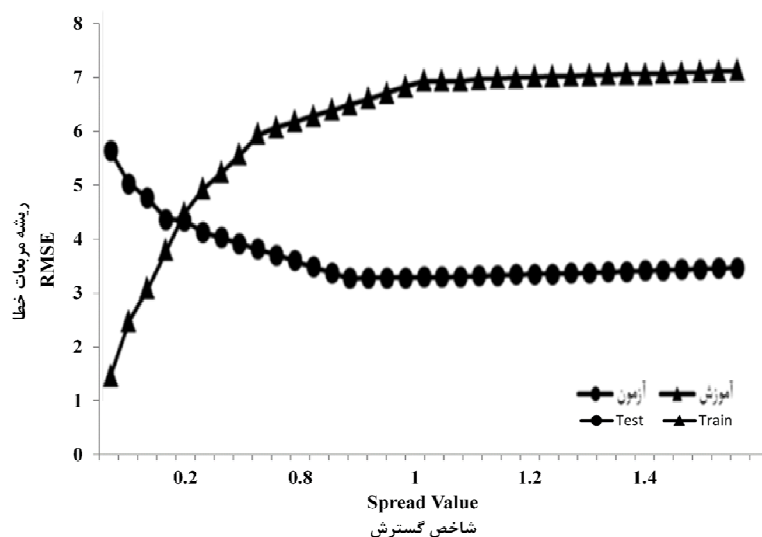
که در این شکل مشاهده می‌گردد، روند تغییرات میانگین ریشه مربعات خطا در مقادیر متفاوت این شاخص برای داده‌های آموزش و آزمون متفاوت است و بنابراین دو مقدار شاخص گسترده‌گی برای ارائه انتخاب گردید (۳). زمانی که این شاخص سبب ایجاد خطای کمینه در داده‌های آزمون می‌شود که نمودارهای آموزش و آزمون یک‌دیگر را قطع می‌کنند. با توجه به این دو رویکرد دو مدل پایه شعاعی توسعه داده شد. در مدل پایه شعاعی ۱ از شاخص گسترده‌گی استفاده شد که در آن میانگین ریشه مربعات خطا داده‌های آزمون کمینه بود و در مدل پایه شعاعی ۲ از شاخص گسترده‌گی استفاده شد که در آن نمودارهای آموزش و آزمون همدیگر را قطع می‌کنند. نتایج نشان داد که آماره‌های میانگین ریشه مربعات خطا و ضریب تبیین برای مدل پایه شعاعی ۱ به ترتیب $4/73$ و $0/49$ و برای مدل پایه شعاعی ۲ به ترتیب $4/72$ و $0/50$ به دست آمد؛ بنابراین نتایج مذکور نشان می‌دهد که اختلاف معنی‌داری بین این دو روش در برآورد ظرفیت تبادل کاتیونی خاک وجود ندارد.

شبکه عصبی مصنوعی آبخاری

برای اجرای این شبکه نیازی به بهینه‌سازی ساختار اولیه (تعداد نرون در لایه مخفی) نمی‌باشد که این نکته از محاسن این شبکه می‌باشد. نتایج خطای این روش بر اساس معیارهای میانگین ریشه مربعات خطا و ضریب تبیین به ترتیب $4/71$ و $0/50$ به دست آمد. نتایج نشان داد که اختلاف معنی‌داری بین روش شبکه عصبی مصنوعی آبخاری، شبکه عصبی مصنوعی پرسپترون و توابع انتقالی پارامتریک (معادلات رگرسیونی) وجود ندارد؛ البته یکی از برتری‌های استفاده از روش شبکه عصبی مصنوعی آبخاری این است که در این روش نیازی به بهینه‌سازی تعداد نرون ورودی نمی‌باشد.

شبکه عصبی مصنوعی پایه شعاعی

حسن این شبکه، الگوریتم یادگیری سریع آن است؛ یعنی شبکه با یک دوره آموزش پیدا می‌کند. در این شبکه، تنها پارامتری که نیاز به بهینه‌شدن دارد، شاخص گسترده‌گی می‌باشد. مقدار این شاخص حساسیت تابع به ورودی را نشان می‌دهد. در این مطالعه مقدار بهینه این شاخص با استفاده از شکل (۲) به دست آمد. همان‌طور



شکل (۲) مقادیر ریشه میانگین مربعات خطا برای شاخص گسترده‌گی متفاوت در پیش‌بینی پارامتر ظرفیت تبادل کاتیونی در روش شبکه عصبی مصنوعی پایه شعاعی

Figure (2) Root mean square error values for different spread parameters to predict cation exchange capacity using radial based neural network

فازی عصبی

مجموعه‌های فازی را می‌توان به عنوان وزن ارتباطات، و متغیرهای ورودی و خروجی را به عنوان نرون‌ها تفسیر کرد و الگوریتم یادگیری ساختارها، پارامترها و یا هر دوی آن‌ها را اصلاح می‌کند. برای ایجاد این شبکه باید چند پارامتر بهینه شوند که شامل نوع و تعداد تابع عضویت، روش یادگیری و تعداد اپیپاک هستند. برای این کار تمام حالات به وسیله سعی و خطا به دست آمد. نتایج خصوصیات مدل فازی عصبی به کار رفته برای پیش‌بینی ظرفیت تبادل کاتیونی خاک نشان داد که تعداد چهار تابع عضویت مثلثی و تعداد قوانین فازی ۹ بهترین ساختار را می‌دهد. همچنین این ساختار در تکرار ۱۴۱۷ با روش یادگیری پس‌خور و روش دفازی کردن میانگین وزنی به دست آمد. نتایج آماره‌های میانگین ریشه مربعات خطا و ضریب تبیین برای مدل فازی عصبی به ترتیب $۴/۰۲$ و $۰/۵۸^{**}$ به دست آمد؛ لذا این نتایج نشان می‌دهد که استفاده از

روش فازی عصبی سبب بهبود نسبی در پیش‌بینی برآورد ظرفیت تبادل کاتیونی خاک شده است.

رگرسیون درختی

در این روش، در ابتدا ساختار درختی با استفاده از داده‌های آموزش ایجاد شده و سپس با استفاده از قوانین به دست آمده در مرحله آموزش، دقت مدل بر روی داده‌های مستقل آزمون، کنترل شدند. نتایج معیارهای میانگین ریشه مربعات خطا و ضریب تبیین برای مدل رگرسیون درختی به ترتیب $۴/۵۳$ و $۰/۵۳^{**}$ به دست آمد. این نتایج نشان می‌دهد که دقت این روش از روش‌های شبکه عصبی بیش‌تر اما از روش فازی عصبی کمتر می‌باشد.

مقایسه مدل‌ها

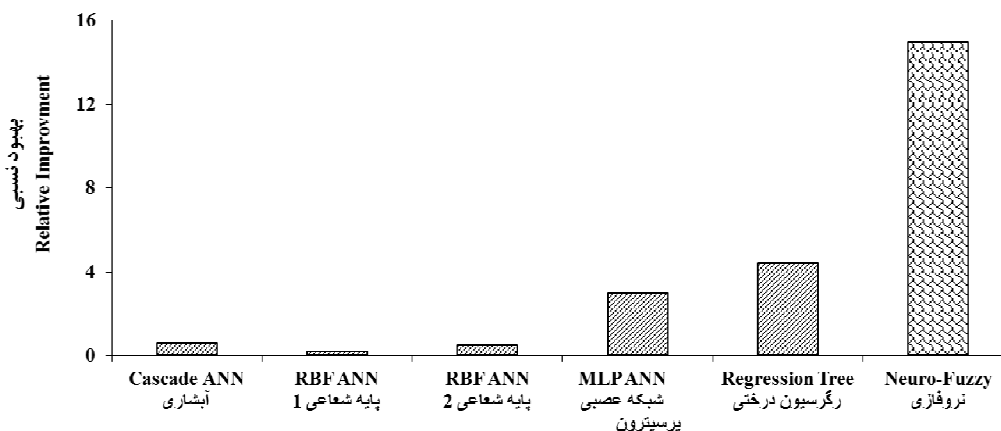
با توجه به جدول ۲ ملاحظه می‌شود در کل بهترین عملکرد در پیش‌بینی ظرفیت تبادل کاتیونی مربوط به مدل نروفازی می‌باشد. از لحاظ هر سه معیار، روش فازی عصبی بهتر از رگرسیون درختی، شبکه عصبی و

تقی زاده مهرجردی و همکاران: پیش‌بینی ظرفیت تبادل کاتیونی خاک‌های...

رگرسیون چند متغیره می‌باشد. پس از روش فازی- عصبی، روش‌های رگرسیون درختی، شبکه عصبی مصنوعی پس انتشار، آبشاری، معادله رگرسیون غیر خطی، شبکه‌های عصبی پایه شعاعی و معادلات رگرسیون خطی به ترتیب بهترین عملکرد را به خود اختصاص داده است. مقدار کوچک میانگین ریشه مربعات خطا برای بیش‌تر مدل‌ها، نشان می‌دهد که این روش‌ها تا اندازه‌ای قادر به برآورد ظرفیت تبادل کاتیونی خاک می‌باشند. در چنین شرایطی که آماره‌های میانگین ریشه مربعات خطا و ضریب تبیین به یکدیگر نزدیک می‌باشد و قضاوت درباره کارایی مدل‌ها مشکل است، می‌توان از شاخص‌های دیگری از قبیل شاخص بهبود نسبی استفاده نمود. این آماره نشان-دهنده میزان کاهش خطا در مدل‌های مختلف نسبت به روش رگرسیون خطی می‌باشد. همان‌گونه که در جدول (۲) و شکل (۳) ارائه شده است، مدل نروفازی دقت پیش‌بینی را در برآورد ظرفیت تبادل کاتیونی خاک نسبت به روش رگرسیون خطی به میزان ۱۵ درصد افزایش داده است. بعد از روش فازی عصبی روش‌های رگرسیون درختی و شبکه عصبی پس انتشار به ترتیب ۴/۴ و ۳ درصد، سبب بهبود نسبی پیش‌بینی‌ها به نسبت رگرسیون چند متغیره خطی شده‌اند.

جدول (۲) مقایسه نتایج ارزیابی خطا در روش‌های مختلف برای پیش‌بینی ظرفیت تبادل کاتیونی در بانک داده آزمون
Table (2) Comparison of different techniques based on error criteria for prediction of cation exchange capacity (Validation)

مدل Model	ریشه مربعات خطا (Root mean square error meq 100g ⁻¹)	ضریب تبیین (Coefficient of Determination)	بهبود نسبی (درصد) (Relative Improvement %)
رگرسیون چند متغیره خطی (Multi-Linear regression)	4.74	0.49	0.00
رگرسیون چند متغیره غیر خطی (Multi-nonLinear regression)	4.71	0.52	0.60
شبکه عصبی آبشاری (Cascade Neural Network)	4.71	0.50	0.60
شبکه عصبی پایه شعاعی ۱ (Radial Based Neural Network 1)	4.73	0.49	0.20
شبکه عصبی پایه شعاعی ۲ (Radial Based Neural Network 2)	4.72	0.50	0.50
شبکه عصبی پس انتشار (Multi-layer perceptron neural network)	4.61	0.50	3.00
رگرسیون درختی (Regression Tree)	4.53	0.53	4.40
نروفازی (Nerou-fuzzy)	4.02	0.58	15.00



شکل (۳) مقایسه نتایج بهبود نسبی در روش‌های مختلف در بانک داده آزمون
Figure (3) Comparison of relative improvement in different methods (Validation data set)

و دانش گذشته‌ی سیستم در طراحی شبکه، مورد استفاده قرار نمی‌گیرد؛ همچنین تعیین ساختار شبکه عصبی مصنوعی که با روش سعی و خطا صورت می‌گیرد، معمولاً با مشکل همراه می‌باشد. الگوریتم یادگیری آن همیشه ممکن است در کمینه‌های محلی قرار گیرد. توانایی شبکه عصبی مصنوعی در برون‌یابی بسیار کم می‌باشد، این مساله زمانی خود را نشان می‌دهد که داده جدید بسیار از داده‌های آموزش دیده فاصله داشته باشد.

نتایج همچنین نشان داد که مدل رگرسیون درختی نسبت به شبکه عصبی مصنوعی دارای دقت بالاتری است؛ به‌طور مشابه لوتو و جرت^۳ (۲۷) کارایی بالاتر درخت تصمیم را در مقایسه با شبکه عصبی نشان دادند. همچنین گریم^۴ و همکاران (۱۷) نتیجه گرفتند که درخت تصمیم به نسبت شبکه عصبی مصنوعی به میزان بهتری قادر به پیش‌بینی ماده آلی خاک است. با این وجود مونجون و همکاران (۳۶) اقدام به مقایسه شبکه عصبی مصنوعی و درخت تصمیم برای پیش‌بینی واحدهای خاک نموده و نشان دادند که تفاوت معنی‌داری بین دو روش برای پیش‌بینی وجود ندارد.

نتایج این مطالعه نشان داد که شبکه‌های عصبی مصنوعی (پرسپترون چندلایه) نسبت به معادله رگرسیونی کارایی بهتر به میزان ۳ درصد (جدول ۲) دارند. که احتمالاً به علت رابطه غیر خطی بین متغیرهای وابسته و مستقل می‌باشد؛ زیرا بر طبق نظر شاپ^۱ و همکاران (۳۷) در طراحی شبکه‌های عصبی نوع خاصی از معادلات لازم نیست و با ایجاد رابطه مناسب بین داده‌های ورودی و خروجی می‌توان به نتایج مناسب دست پیدا کرد و به علت وجود روابط غیر خطی میان متغیرهای وابسته و پیش‌بینی شونده شبکه عصبی عملکرد بهتری نسبت به مدل‌های رگرسیون داشته است. این نتایج با نتایج تاماری^۲ و همکاران (۴۲) همخوانی داشت، این پژوهشگران نیز میانگین ریشه مربعات خطا کم‌تری از برآورد خصوصیات خاک با شبکه عصبی مصنوعی نسبت به رگرسیون خطی چندگانه به دست آوردند؛ اما باید دقت شود که شبکه عصبی مصنوعی دارای معایبی نیز می‌باشد. شبکه عصبی به صورت جعبه سیاهی است که احتیاج به تعریف شدن دارد. یکی از معایب شبکه عصبی، قابلیت تفسیر ضعیف آن است. همچنین این شبکه برای مسائلی با پایگاه داده‌ای بزرگ مناسب است

3- Luoto and Hjort
4- Grimm *et al.*

1- Schaap *et al.*
2- Tamari *et al.*

تمامی روش‌های ذکر شده در این تحقیق بر پایه فرض دقیق بودن متغیرهای مورد مطالعه و مشاهده‌های مربوط به آن‌ها استوار است و روابط بین متغیرها نیز دقیق فرض می‌گردد؛ حال آن‌که در سیستم‌های طبیعی مانند خاک، عموماً با مشاهدات روابط ناشناخته بین متغیرها وجود دارد. در چنین شرایطی باید از مدل‌هایی بهره گرفت که قادر به ارائه الگوهای مناسب‌تر باشند، به گونه‌ای که این مدل‌ها بتوانند انطباق بیشتری با دنیای واقعی داشته باشند. یکی از این مدل‌ها که می‌تواند انطباق بیشتری با دنیای واقعی داشته باشد، روش فازی عصبی است. سیستم‌های استنتاج فازی، می‌توانند از دانش گذشته‌ی سیستم در چارچوب قوانین فازی استفاده نمایند. بعلاوه در این سیستم‌ها، امکان تعریف ورودی و خروجی غیر دقیق، میسر می‌باشد. اگر عملکردهای منطق فازی در شبکه‌های عصبی و یادگیری را در سیستم‌های فازی شریک گردانیم، آنگاه قصور و نواقص در شبکه‌های عصبی و سیستم‌های فازی را می‌توان پوشش داد. نتیجه این کار، یک مدل فازی عصبی خواهد بود. مدل‌های فازی عصبی، با استفاده از مجموعه‌ی داده‌ها به یادگیری رفتار سیستم و استخراج قوانین می‌پردازد. همچنین این مدل‌ها دارای قابلیت تعمیم‌پذیری بوده و نقص کلیدی سیستم‌های فازی را که عدم توانایی یادگیری و تعمیم‌پذیری است، پوشش می‌دهند. قسمت شبکه عصبی به طور خودکار، قواعد منطق فازی و توابع عضویت را در طول دوره تناوب یادگیری ایجاد می‌کند. نتایج نروفازی در پیش‌بینی ظرفیت تبادل کاتیونی خاک نشان داد که دقت این مدل از تمام انواع شبکه‌های عصبی و درخت تصمیم بیش‌تر است. به دلیل ساختار غیر خطی مدل‌های شبکه عصبی و ابهام در متغیرهای خاک سبب شده است که امروزه بیش‌تر پژوهشگران تمایل به استفاده از مدل‌های ترکیبی مانند سیستم فازی تطبیق‌پذیر مبتنی بر شبکه عصبی نشان دهند (۲۶)؛ به عنوان مثال

آلی^۱ و همکاران (۱) نشان دادند که مدل فازی عصبی در پیش‌بینی درصد رطوبت اشباع خاک دارای دقت بالاتری نسبت به مدل شبکه عصبی مصنوعی بود و یا آزمون‌تالا^۲ و همکاران (۴) مدل فازی عصبی را در پیش‌بینی بار بستر چندین رودخانه در مالزی مناسب‌تر از مدل‌های رگرسیونی دیگر تشخیص دادند. به طور مشابه سیویلکولگو^۳ و همکاران (۱۱) در مدل‌سازی هدر رفت نیتروژن و کربن از خاک به این نتیجه رسیدند که مدل فازی عصبی در مقایسه با مدل‌های رگرسیونی دیگر برآورد دقیق‌تری دارد. ترزی^۴ و همکاران (۴۳) نیز مدل فازی عصبی را بهترین روش در پیش‌بینی تبخیر سطحی معرفی کردند. بدین ترتیب با توجه به وجود عدم قطعیت در پدیده‌های مرتبط با خاک یا تقریبی بودن مقادیر اندازه‌گیری شده خصوصیات مختلف خاک، به نظر می‌رسد کارایی بالاتر مدل مبتنی بر مجموعه‌های فازی در پردازش توابع انتقالی خاک به همین علت باشد. بنابراین نتایج این مطالعه با نتایج پژوهشگرانی که بیان می‌کنند که روش ترکیبی (مانند فازی عصبی) نتایج بهتری را ایجاد می‌کند، مطابقت داشت. شکل (۴) رابطه بین ظرفیت تبادل کاتیونی برآورد شده و اندازه‌گیری شده را در خاک‌های بانک داده آموزش و آزمون با استفاده از روش فازی عصبی نشان می‌دهد. در صورتی که نقاط برآورد شده و اندازه‌گیری شده ظرفیت تبادل کاتیونی در اطراف خط ۱:۱ پراکنده باشند، نشان دهنده ناریب بودن خطا در روش استفاده شده می‌باشد و اگر این نقاط در یک طرف خط ۱:۱ پراکنده باشند، نشان دهنده این است که مدل بیش برآورد و یا کم برآورد دارد. نتایج این مطالعه نشان داد که خاک‌های مورد مطالعه به خوبی در اطراف خط ۱:۱ پراکنده می‌باشند.

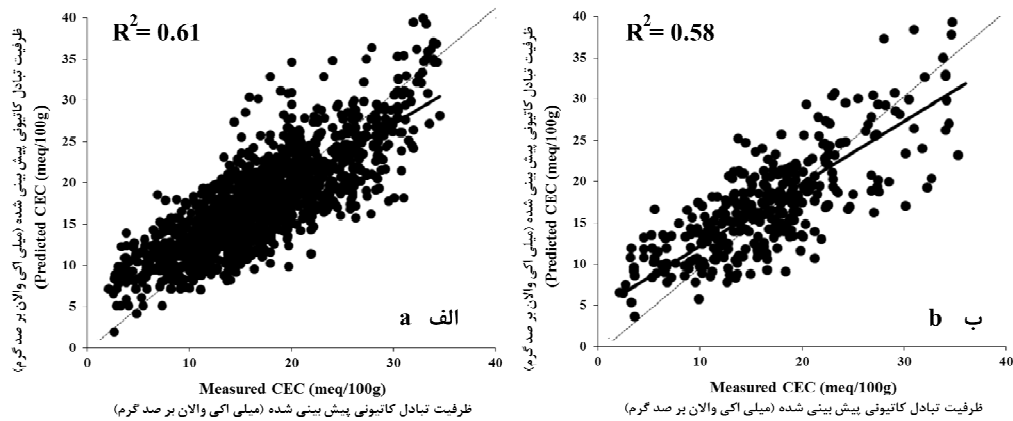
با توجه به آنچه بیان شد، مدل فازی عصبی ارائه شده با ضریب تبیین بالا (۰/۵۸) ناریب نبود و به خوبی قادر به

1- Aali *et al.*

2- Azmathulla *et al.*

3- Civelekoglu *et al.*

4- Terzi *et al.*



شکل (۴) نمودار پراکنش برای مقادیر مشاهداتی و شبیه‌سازی شده داده‌های آموزش (الف) و آزمون در فازی عصبی (ب)

Figure (4) Scatter gram predicted and measured for validation (a) and training using ANFIS (b)

خوشه‌بندی

در این مطالعه به منظور بررسی ناهمگنی داده بر نتایج پیش-بینی مدل‌ها، در ابتدا داده‌ها با استفاده از روش میانگین فازی کی همگن شدند. در این روش داده‌ها در گروه‌هایی که اعضای هر گروه از جنبه‌های خاصی بیش‌ترین شباهت را با هم داشته و با اعضای خوشه‌های دیگر هیچ شباهتی ندارند، قرار داده می‌شوند. برای انتخاب تابع فاصله برای خوشه‌بندی در مطالعه حاضر از نرُم ماهالانویس استفاده شد؛ پس از انتخاب تابع فاصله مهم-ترین پارامتر در طبقه‌بندی فازی تعیین مقدار نمای فازی بودن است. در این مطالعه خوشه‌بندی بر اساس تعداد کلاس‌های فازی در گستره ۷-۲ و نمای فازی بودن در دامنه ۱/۱-۲ با گام ۰/۱ در نظر گرفته شد و این الگوریتم بر روی داده‌های ماده آلی، درصد رس و ظرفیت تبادل کاتیونی اجرا گردید. از مهم‌ترین شرایط برای قابل پذیرش بودن یک خوشه‌بندی آن است که میانگین مقدار عضویت برای هر نمونه در کلاس‌های عادی و کلاس نامعمول در حد امکان نزدیک به هم بوده و در حالت آرمانی برابر باشد؛ در این حالت شانس قرار گرفتن یک داده در همه کلاس‌ها برابر می‌باشد و آن چه در تعلق آن تاثیر دارد، فاصله آماری آن از مراکز کلاس‌ها است؛ بنابراین در مطالعه حاضر تمامی حالت‌های خوشه‌بندی انجام گرفت و انتخاب

پیش‌بینی ظرفیت تبادل کاتیونی خاک می‌باشد. باید دقت نمود با توجه به مطالعات اخیر توابع انتقالی ضریب همبستگی ۰/۵۸ قابل قبول می‌باشد. مک‌براتنی و همکاران^۱ (۳۰) بیان داشتند که در مطالعات خاکشناسی ضریب تبیین ۰/۴ الی ۰/۷ بسیار معمول می‌باشد؛ اما پایین بودن نسبی ضریب تبیین ۰/۵۸ نسبت به بسیاری از مطالعات انجام شده در کشور به منظور برآورد ظرفیت تبادل کاتیونی خاک که ضرایب تبیین در دامنه ۰/۵۰ الی ۰/۶۸ را به دست آورده‌اند؛ می‌تواند به این علت باشد که در اکثر مطالعات گذشته تعداد خاک-های مورد استفاده بندرت بیش از ۱۰۰ عدد می‌رسد و همچنین تمامی خاک‌ها از یک منطقه مشخص به دست آمده‌اند؛ لذا تغییر خصوصیات خاک‌ها در مطالعات صورت گرفته تاکنون زیاد نیست و این در حالی است که در مطالعه حاضر از تعداد بیشتری نمونه‌های خاک با طیف گسترده‌ای از خصوصیات استفاده گردید. بنابراین همگن بودن خاک‌ها (نمونه‌ها) در بانک داده‌های کوچک موجب به بدست آوردن ضریب تبیین بالاتر می‌شود. این در حالی است که در این مطالعه با توجه به پراکندگی زیاد خاک‌ها، ضریب تبیین نیز اندکی کاهش یافته است.

1- McBratney *et al.*

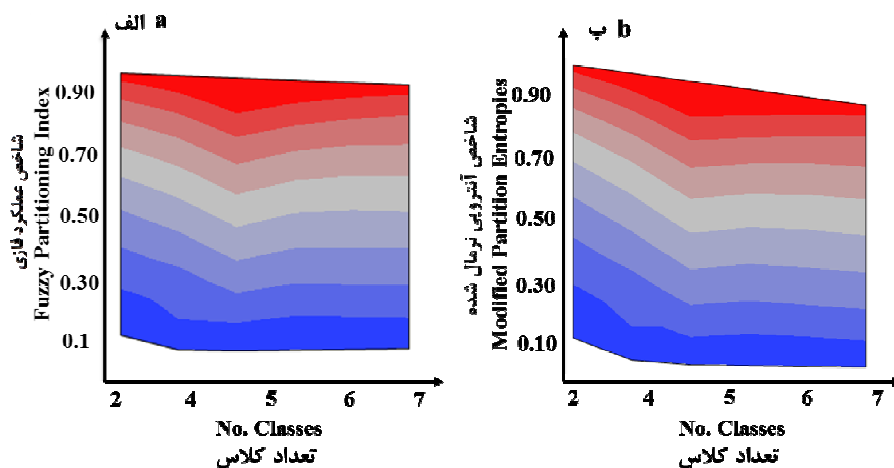
تقی زاده مهرجردی و همکاران: پیش‌بینی ظرفیت تبادل کاتیونی خاک‌های...

بروی هر یک از گروه‌های چهارگانه اجرا گردید. نتایج نشان داد که ضریب همبستگی گروه‌های چهارگانه به ترتیب ۰/۶۷، ۰/۷۳، ۰/۷۰ و ۰/۶۵ می‌باشد. این نتایج نشان می‌دهد که مدل‌سازی برای گروه‌های همگن که دارای خصوصیات خاکی تقریباً مشابه هستند، باعث افزایش ضریب همبستگی به میزان ۱۵، ۲۶، ۲۰ و ۱۲ درصد نسبت به حالت اولیه (مدل نروفازی با کل داده‌ها) شده است.

عدم قطعیت

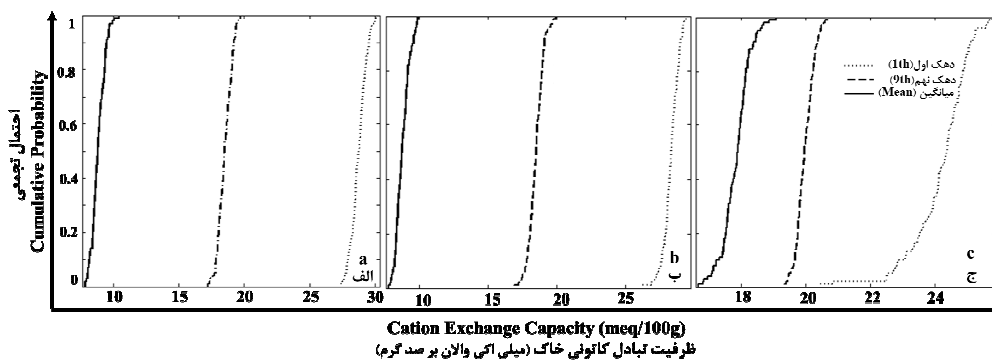
شکل (۶) عدم قطعیت سه مدل نروفازی، شبکه عصبی مصنوعی و مدل رگرسیونی را نشان می‌دهد. نتایج نشان می‌دهد که فاصله بین دهک اول و نهم در مدل نروفازی به

تعداد بهینه کلاس‌ها با استفاده از توابع شاخص عملکرد فازی و آنتروپی نرمال شده در هر نمای فازی بودن محاسبه و نسبت به تغییرات مقادیر شاخص‌ها بر حسب تعداد کلاس‌ها ترسیم شد. همان‌طور که در شکل ۵ ملاحظه می‌شود، با افزایش تعداد کلاس‌ها تابع شاخص عملکرد فازی، روند کاهشی دارد. در تعداد کلاس چهارم کمترین اندازه از این مقدار ملاحظه می‌شود. به‌طور مشابه شاخص آنتروپی نرمالیزه شده در کلاس چهارم حداقل مقدار خود را دارد. این شرایط نشان دهنده آن است که می‌توان داده‌ها را به چهار گروه همگن تقسیم کرد. گروه اول، دوم، سوم و چهارم به ترتیب دارای ۵۸۴، ۲۰۱، ۶۰۱ و ۳۸۴ داده می‌باشد. در مرحله بعد، مدل نروفازی (بهترین تابع انتقالی انتخاب شده در مطالعه حاضر) را



شکل (۵) تغییرات آنتروپی نرمالیزه شده (راست) و تغییرات عملکرد فازی (چپ) بر حسب تعداد کلاس‌ها با نمای فازی مختلف.

Figure (5) Variation of fuzzy partitioning index (a) and modified partition entropies (b) using different fuzzification values for three to seven classes.



شکل (۶) مقادیر احتمال تجمعی ظرفیت تبادل کاتیونی خاک با استفاده از روش مونت کارلو برای مدل‌های ال رگرسیون خطی، ب) شبکه عصبی مصنوعی، ج) نروفازی

Figure (6) Cumulative probability of cation exchange capacity using mont-carlo method for a) multi-linear regression method, b) artificial neural network and c) ANFIS.

پیش‌بینی ظرفیت تبادل کاتیونی کمک کنند. همچنین باید دقت نمود که اندازه‌گیری آزمایشگاهی ظرفیت تبادل کاتیونی با اشکالاتی مواجه می‌باشد. این مساله بخصوص در خاک‌های آهکی و گچی دوچندان خواهد شد؛ بنابراین اگر خوشه‌بندی بر اساس مقادیر آهک و گچ نیز انجام گیرد، شاید باعث بهبود مدل نهایی شود. در پایان باید گفت که با توجه به ماهیت ابهامی پدیده‌های مرتبط با خاک یا تقریبی بودن مقادیر اندازه‌گیری شده ویژگی‌های گوناگون خاک، به نظر می‌رسد کارایی بالاتر مدل مبتنی بر مجموعه‌های فازی در پردازش توابع انتقالی خاک به همین علت باشد.

در مجموع هر یک از روش‌های داده‌کاوی دارای مزیت و معایبی است که با توجه به ماهیت داده مورد بررسی، باید انتخاب شوند. به عنوان مثال، رگرسیون خطی ساده‌ترین شکل از بیان ارتباط بین داده‌ها می‌باشد؛ ولی در عین حال فقط ارتباط خطی را در بین متغیرهای مستقل و وابسته در نظر می‌گیرد. در مقابل، شبکه‌های عصبی مصنوعی روابط غیر خطی را در نظر گرفته و همچنین دارای قابلیت پردازش موازی می‌باشند و به داده‌های پرت نیز حساس نمی‌باشند. شبکه‌های عصبی مصنوعی بر خلاف روش‌های آماری کلاسیک، مفروضاتی در خصوص ویژگی‌های توزیع داده‌ها و متغیرهای مستقل ورودی ندارند و قابلیت حل مسائل پیچیده را دارند. از معایب شبکه‌های عصبی آن است که تفسیر روشی که این شبکه‌ها در آن به تصمیمات خود می‌رسند، بسیار دشوار است. از دیگر انتقادهای واره بر شبکه‌های عصبی می‌توان به تعدادی از عوامل نظیر توپولوژی شبکه اشاره کرد که باید به طور تجربی، تعریف شوند؛ بنابراین شبکه‌های عصبی معمولاً زمان بسیار زیادی را برای آموزش نیاز دارند. همچنین شبکه‌ها به صورت جعبه سیاه هستند و قابلیت تفسیر کاربر را کاهش و در نتیجه بسیار پیچیده می‌باشند. در مقابل پیچیدگی شبکه‌های عصبی مصنوعی، مهم‌ترین

نسبت مدل‌های شبکه عصبی مصنوعی و رگرسیونی کمتر می‌باشد و بیشترین پهنای باند مربوط به مدل رگرسیونی است. این نتایج نشان می‌دهد که اگر از میانگین نمونه‌ها (پارامترهای ورودی) برای برآورد ظرفیت تبادل کاتیونی خاک استفاده شود، بزرگی خطا در مدل رگرسیونی به نسبت مدل‌های نروفازی و شبکه عصبی مصنوعی بیشتر خواهد بود. در حقیقت بررسی عدم قطعیت مدل‌ها با استفاده از شبیه‌سازی مونت کارلو نشان داد که مدل نروفازی دارای دقت بیشتری نسبت به سایر مدل‌ها می‌باشد. فرخ‌نیا و مرید (۱۳) نیز نشان دادند که مدل نروفازی دارای عدم قطعیت کمتری نسبت به شبکه عصبی مصنوعی می‌باشد.

نتیجه‌گیری

برای پیش‌بینی ظرفیت تبادل کاتیونی، از عوامل رس و کربن آلی استفاده گردید. معیارهای ارزیابی نتایج نشان داد که مدل فازی عصبی دارای بالاترین ضریب تبیین و کم‌ترین میانگین ریشه مربعات خطا در پیش‌بینی پارامتر خاک می‌باشد، به طوری که توانسته دقت پیش‌بینی را به میزان ۱۵ درصد نسبت به روش رگرسیون خطی افزایش دهد. بعد از مدل فازی عصبی، رگرسیون درختی، شبکه عصبی مصنوعی پس انتشار، پایه شعاعی و آبخاری به ترتیب بهترین عملکرد را به خود اختصاص داده است. همچنین نتایج نشان داد که استفاده از روش خوشه‌بندی باعث افزایش دقت مدل نروفازی می‌شود؛ به طوری که مدل‌سازی برای گروه‌های همگن که دارای خصوصیات خاکی تقریباً مشابه هستند، باعث افزایش ضریب همبستگی به میزان ۱۵، ۲۶، ۲۰ و ۱۲ درصد نسبت به حالت اولیه (مدل نروفازی با کل داده‌ها) شده است. البته جهت افزایش بیشتر ضریب همبستگی می‌توان از داده‌های کمکی دیگر نیز استفاده نمود. به عنوان مثال استفاده از درصد آهک، وزن مخصوص ظاهری، واکنش خاک و کانی‌شناسی رس به نوعی می‌توانند در افزایش دقت

کاربردها دانش مربوط به توصیف رفتار سیستم در خود مجموعه داده‌های آن مستتر می‌باشد که استخراج اطلاعات آن به صورت دستی از میان آن همه اطلاعات، زمان بسیار زیاد و وقت و سعی و افری را طلب می‌کند. هنگامی که مجموعه‌ای از داده‌ها شامل اطلاعاتی درباره چگونگی عملکرد سیستم تحت طراحی باشند؛ استفاده از شبکه عصبی می‌تواند نوید بخش ارائه راه حلی مناسب برای این منظور محسوب گردد؛ زیرا شبکه‌های عصبی توانایی یادگیری از روی مجموعه داده‌ها را دارند. مدل نروفازی با استفاده کارایی پردازش موازی شبکه عصبی مصنوعی، سعی در یافتن قوانین فازی دارد؛ در نتیجه از کارایی نسبتاً بالایی برخوردار می‌باشد.

سپاس‌گزاری

بدین وسیله، از دانشگاه اردکان، در تأمین بودجه و امکانات لازم تحقیق حاضر سپاس‌گزاری می‌شود.

خصوصیت درخت‌های تصمیم، قابلیت آن‌ها در شکستن فرآیند پیچیده تصمیم‌گیری به مجموعه‌ای از تصمیمات ساده‌تر است که براحتی قابل تفسیر هستند. درخت تصمیم به دلیل سادگی و قابل فهم بودن تکنیک محبوبي در داده کاوی محسوب می‌شود و بسیار راحت‌تر با داده‌های گم‌شده یا داده‌های پرت برخورد می‌کند، همچنین به روابط غیر خطی بسیار انعطاف‌پذیر است و هیچ فرضیه‌ای را برای توزیع داده‌ها لازم ندارد و نسبت به تغییرات بر روی داده‌های پیش‌بینی کننده مقاوم می‌باشد. رمز موفقیت منطق فازی نیز در امکان توصیف رفتار سیستم مورد نظر با دستورات ساده شرطی if-then است که در بیشتر کاربردها، این امکان راه حل ساده‌تر و صرف مدت زمان کوتاهتری را برای طراحی سیستم میسر می‌کند و بعلاوه تمامی اطلاعات و دانش مهندسی مربوط به چگونگی عملکرد سیستم برای بهینه سازی کارایی آن به طور مستقیم قابل استفاده است. این امکان جالب و قدرتمند منطق فازی در عین حال یک محدودیت محسوب می‌شود؛ زیرا در بسیاری از

منابع

1. Aali, K.A., Parsinejad, M., and Rahmani, B. 2009. Estimation of saturation percentage of soil using multiple regression, ANN, and ANFIS techniques. *Computer and Information Science*, 2: 127–136.
2. Akbarzadeh, A., Taghizadeh Mehrjardi, R., Rouhipour, H., Gorji M., and Refahi, H.G. 2009. Estimating of soil erosion covered with rolled erosion control systems using rainfall simulator (neuro-fuzzy and artificial neural network approaches). *Journal of Applied Sciences Research*, 5: 505–514.
3. Amini, M., Abbaspour, K.C., Khademi, H., Fathianpour, N., Afyuni, M., and Schulin, R. 2005. Neural network models to predict cation exchange capacity in arid regions of Iran. *European Journal of Soil Science*, 53: 748–757.
4. Azmathulla, H.M., Chang, C.K., Ghani, A., Ariffin, J., Zakaria, N.A., and Abu Hasan, Z. 2008. An ANFIS-based approach for predicting the bed load for moderately sized rivers. *Journal of Hydro-environment Research*, 3: 35–44.
5. Bell, M.A., and Van Kulen, H. 1995. Soil pedotransfer function for four Mexican soils. *Soil Science Society of America Journal*, 59: 865-871.

6. Bouma, J. 1989. Using soil survey data for quantitative land evaluation. *Advanced Soil Science*, 9: 177–213.
7. Bower, C.A., Reitmeir, R.F., and Fireman, M. 1952. Exchangeable cation analysis of saline and alkali soils. *Soil Science*, 73:251-261.
8. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. 1984. *Classification and regression*. Tress. Wadsworth, Belmont, CA.
9. Brus, D.J., Kempen, B., Heuvelink, G.B.M. 2011. Sampling for validation of digital soil maps. *European Journal of Soil Science*, 62 (3): 394–407.
10. Carrow, R.N., Waddingtone, D.V., and Rieke, P.E. 2001. *Turfgrass soil fertility and chemical problems*. John Wiley and Sons, New York.
11. Civelekoglu, G., Perendeci, A., Yigit, N. O., and Kitis, M. 2007. Modeling carbon and nitrogen removal in an industrial wastewater treatment plant using an adaptive network-based fuzzy inference system. *Clean-Soil, Air, Water*, 35: 617–625.
12. Drake, J.T. 2000. *Communications phase synchronization using the adaptive network fuzzy inference system*. Ph.D. dissertation, New Mexico State University, Las Cruces, New Mexico, USA.
13. Farokhnia, A., and Morid, S. 2013. Uncertainty analysis of artificial neural network and neuro-fuzzy models in prediction of river flow. *Journal of Water Resources Research*, 3: 14–27.
14. Gee, G.W., and Bauder, J.W. 1986. Particle size analysis. p. 383-411, In: A. Klute. (ed), *Methods of Soil Analysis. Part 1*, American Society of Agronomy Madison, WI.
15. Ghorbani, S., and Homaei, M. 2003. Evaluation of soil un-saturated hydrologic parameters using pedotransfer functions. *Journal of Agricultural Engineering Research*, 3: 3–15.
16. Givi, J., Prasher, S.O., and Patel, R.M. 2004. Evaluation of pedotransfer functions in predicting the soil water contents at field capacity and wilting point. *Agricultural Water Management*, 70: 83–96.
17. Grimm, R., Behrens, T., Marker, M., and Elsenbeer, H. 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island - digital soil mapping using random forests analysis. *Geoderma*, 146: 102–113.
18. Jang, J.S.R. 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23: 665–685.
19. Kalkan, E., Akbulut, S., Tortum, A., and Celik, S. 2008. Prediction of the unconfined compressive strength of compacted granular soils by using inference systems. *Environmental Geology*, 12: 254–267.
20. Karimian, A. 1995. Influence of clay and organic matter in cation exchange capacity

of calcareuou soils of Fars. 5th Soil Science Congress, Karaj, Iran.

21. Kashi, H., Emamgholizadeh, S., and Hashemi, A. 2013. Prediction of cation exchange capacity in the bare and cultivated lands using multiple regression and artificial neural network. *Journal of Soil and Water*, 27: 472–484.
22. Kayadelen, C., Günaydın, O., Fener, M., Demir, A., and Özvan, A. 2009. Modeling of the angle of shearing resistance of soils using soft computing systems. *Expert Systems with Applications*, 36: 11814–11826.
23. Kimar, M., Raghuwanshi, N.S., Singh, R., Wallender, W.W., and Pruitt, W.O. 2002. Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, 128: 224–233.
24. Kisi, O. 2005. Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrological Sciences–Journal–des Sciences Hydrologiques*, 50: 683–696.
25. Krogh, L., Madsen, H.B., and Greve, M.H. 2000. Cation exchange capacity pedotransfer functions for Danish soils. *Acta Agriculture Scandinavica, Section B — Soil and Plant Science*, 50: 1-12.
26. Lee, E.S. 2000. Neuro-Fuzzy estimation in spatial statistics. *Journal of Mathematical Analysis and Applications*, 249: 221–231.
27. Luoto, M. and Hjort, J. 2005. Evaluation of current statistical approaches for predictive geomorphological mapping. *Geomorphology*, 67: 299-315.
28. Manrique, L.A., Jones, C.A., and Dyke, P.T. 1991. Predicting cation exchange capacity from soil physical and chemical properties. *Soil Science Society of America Journal*, 50: 787–794.
29. Manyame, C., Morgan, C.L., Heilman, J.L., Fatondji, D., Gerard, B., and Payne, W.A. 2007. Modeling hydraulic properties of sandy soils of Niger using pedotransfer functions. *Geoderma*, 141: 407–415.
30. McBratney, A.B., Minasny, B., Cattle, S.R., and Vervoort, R.W. 2002. From pedotransfer functions to soil inference systems. *Geoderma*, 109:41-73.
31. Meamarian-Fard, M., and Beigi, H. 2009. Comparison of multiple regression and artificial neural network pedotransfer functions for prediction of cation exchange capacity in soils of Chaharmahal-Bakhtiari. *Journal of Soil and Water*, 23: 90–99.
32. Moghaddamnia, A., Ghafari-Gousheh, M., Piri, J., Amin, S., and Han, D. 2009. Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Advances in Water Resources*, 32: 88–97.
33. Mohajer, R., Salehi, M., and Beigi, H. 2009. Prediction of cation exchange capacity using multiple regression and artificial neural network and effect of data partitioning on the accuracy of models. *Journal of Soil and Water*, 49: 83–97.

34. Moonjun, R., Farshad, A., Shrestha, D.P., and Vaiphasa, C. 2010. Artificial neural network and decision tree in predictive soil mapping of Hoi Num Rin sub-watershed, Thailand. *Digital Soil Mapping, Progress in Soil Science 2*, pp. 151- 164
35. Moran, C.J., and Bui, E.N. 2002. Spatial data mining for enhanced soil map modelling. *International Journal of Geographical Information Science*, 16: 533-549.
36. Quinlan, J.R. 2001. *Cubist: An Informal Tutorial*. <http://www.rulequest.com>.
37. Schaap, M.G., Leij, F.J., and van-Genuchten, M.T. 1998. Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Science Society of America Journal*, 62: 847–855.
38. Seybold, C.A., Grossman, R.B., and Reinsch, T.G. 2005. Predicting cation exchange capacity for soil survey using linear models. *Soil Science Society of America Journal*, 69: 856–86.
39. Taghizadeh-Mehrjardi, R., Mahmoodi, S., Heidari, A., and Akbarzadeh, A. 2009. Estimation of cation exchange capacity using multiple regression and artificial neural network techniques in Khezrabd region, Yazd. *Journal of Research in Agriculture*, 1: 1–11. (in Persian).
40. Taghizadeh-Mehrjardi, R., Minasny, B., Sarmadian, F., and Malone, B.P. 2014. Digital mapping of soil salinity in Ardakan, central Iran. *Geoderma*, 213: 15-28.
41. Taghizadeh-Mehrjardi, R., Sarmadian, F., Minasny, B., Triantafilis, J., and Omid, M. 2014. Digital mapping of soil classes using decision tree and auxiliary data in Ardakan. *Arid Land Research and Management*, 28:147–168.
42. Tamari, S., Wosten, J.H.M., and Ruz-suarez, J.C. 1996. Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Science Society of America Journal*, 60: 1732–1741.
43. Terzi, O., Keskin, M.E., and Taylan, E.D. 2007. Estimating evaporation using ANFIS. *Journal of Irrigation and Drainage Engineering*, 132:503–507.
44. Walkey, A., and Black, I.A. 1934. An examination of degtjareff method for determining soil organic matter and a proposed modification of the chromic acid in soil analysis. 1. Experimental. *Soil Science Society of America Journal*, 79: 459-465.
45. Zolfaghari, A., Soltani, M., Afshari, T., and Sarmadian, F. 2013. Comparison of K-nearest neighbor and artificial neural network techniques in prediction of cation exchange capacity. *Journal of Soil Management and Sustainable*, 3: 77–94.