

مقایسه مدل‌های رگرسیون خطی، فازی و فازی-ژنتیک در برآورد گنجایش تبادل کاتیونی خاک

حبیب پالیزوان زنده^۱ و عباس احمدی^{۲*}

(تاریخ دریافت: ۱۳۹۳/۰۸/۱۱ تاریخ پذیرش: ۱۳۹۴/۰۷/۲۳)

چکیده

گنجایش تبادل کاتیونی یکی از ویژگی‌های شیمیایی خاک است که تاثیر عمده‌ای بر سایر خواص شیمیایی، فیزیکی، حاصلخیزی و بیولوژیکی خاک دارد. در این تحقیق، کارآیی برخی روش‌های ارائه توابع انتقالی نظیر روش رگرسیونی، روش فازی و روش فازی-ژنتیک در برآورد گنجایش تبادل کاتیونی براساس خصوصیات زود یافت خاک مورد ارزیابی قرار گرفته است. برای این منظور ۷۷۰ نمونه از پایگاه اطلاعات داده‌های خاک اروپا (IES) استخراج گردید. سپس مدل‌های رگرسیون خطی چند متغیره، فازی و فازی-ژنتیک به منظور توسعه توابع انتقالی برای تخمین CEC خاک با استفاده از خصوصیات زود یافت رس و کربن آلی خاک، استفاده شد. به منظور ارزیابی مدل‌ها از معیارهای ضریب تبیین (R^2)، ریشه میانگین مربع خطا (RMSE) و میانگین مطلق خطا (MAE) استفاده شد. مقادیر R^2 ، RMSE و MAE برای مدل رگرسیون خطی به ترتیب برابر 0.72 و $7.42 \text{ cmol}_c \text{ kg}^{-1}$ و $9.13 \text{ cmol}_c \text{ kg}^{-1}$ و برای مدل فازی به ترتیب 0.78 ، $4.44 \text{ cmol}_c \text{ kg}^{-1}$ و $4.32 \text{ cmol}_c \text{ kg}^{-1}$ به دست آمد در حالی که این پارامترها برای مدل فازی-ژنتیک به ترتیب 0.84 ، $4.07 \text{ cmol}_c \text{ kg}^{-1}$ و $3.57 \text{ cmol}_c \text{ kg}^{-1}$ بود. این نتایج نشان داد که مدل فازی-ژنتیک دقت بیشتری نسبت به مدل فازی، آن هم دقت بیشتری نسبت به مدل رگرسیون خطی در برآورد گنجایش تبادل کاتیونی خاک دارد.

کلمات کلیدی: توابع انتقالی خاک، درصد رس، قواعد فازی، کربن آلی

۱- دانشجوی کارشناسی ارشد علوم خاک دانشگاه تبریز

۲- استادیار گروه علوم خاک، دانشگاه تبریز (مکاتبه کننده)

* پست الکترونیک: a_ahmadi@tabrizu.ac.ir

مقدمه

مک براتنی و همکاران (McBratney *et al.*, 2002) با استفاده از رگرسیون خطی چند متغیره با متغیرهای مستقل رس و مواد آلی میزان CEC را تخمین زدند. به دلیل صرف وقت و هزینه زیاد در اندازه‌گیری ویژگی‌هایی چون گنجایش تبادل کاتیونی، برای تخمین دقیق این نوع ویژگی‌ها باید به دنبال راه حل‌های مناسب، ساده و کم هزینه بود. به همین منظور استفاده از روش‌های تخمین غیر مستقیم مورد توجه قرار گرفته است. روش اخیر یا روش غیر مستقیم به دست آوردن داده‌های دیرپافت خاک، توابع انتقالی نام‌گذاری شده است (Bouma, 1989).

به دلیل این که توابع انتقالی خصوصیات دیرپافت را با استفاده از خصوصیات زودپافت خاک پیش‌بینی می‌کنند، دارای مزیت‌های عمده‌ای همچون صرفه‌جویی در وقت و هزینه هستند (Wosten *et al.*, 2001).

کروچ و همکاران (Krogh *et al.*, 2000) با استفاده از اطلاعات ۱۶۴۳ نمونه از خاک‌های دانمارک به کمک متغیرهای مقدار ماده آلی، رس، سیلت ریز و pH خاک در عصاره کلرور کلسیم توابعی برای برآورد گنجایش تبادل کاتیونی ارائه دادند که در این توابع مقدار رس و ماده آلی خاک ۹۰ درصد از تغییرات گنجایش تبادل کاتیونی را توجیه می‌کند. امینی و همکاران (Amini *et al.*, 2005) با استفاده از مقدار مواد آلی و رس اقدام به برآورد میزان CEC خاک در منطقه اصفهان، به وسیله شبکه عصبی مصنوعی و پنج مدل تجربی که همه بر پایه روش‌های رگرسیونی می‌باشند، نمودند. نتایج این محققان نشان داد روش شبکه عصبی مصنوعی از برتری قابل قبولی نسبت به سایر روش‌ها برخوردار است. حسینی عربلو و همکاران (Hosseini Arabloo *et al.*, 2015) با استفاده از گروه‌بندی داده‌ها بر اساس میزان رس و کربن آلی به این نتیجه رسیدند که در خاک‌هایی با مقادیر کربن آلی کمتر از ۰/۵ درصد که عمده خاک‌های بخش‌های خشک و نیمه خشک ایران را در برمی‌گیرد تنها با استفاده از درصد رس به عنوان ورودی، می‌توان CEC را با دقت قابل قبولی برآورد

گنجایش تبادل کاتیونی^۱ یکی از خصوصیات مهم خاک می‌باشد که در پایگاه داده‌های خاک مورد نیاز است (Manrique *et al.*, 1991) و به عنوان داده ورودی در اغلب مدل‌های مربوط به خاک و محیط زیست استفاده می‌شود (Keller *et al.*, 2001). به طور مثال جذب سطحی و دفع فلزات سنگین، مانند مس و روی (Arias *et al.*, 2005) و سرب (Altin & Degirmenci, 2005) به طور معنی‌دار با گنجایش تبادل کاتیونی خاک مرتبط است. این ویژگی معرف کمیت مکان‌های تبدالی موجود در سطوح ذرات خاک برای نگهداری کاتیون‌ها به وسیله نیروهای الکترواستاتیک و قابلیت رهاسازی آن‌ها برای جذب توسط گیاه می‌باشد. مقدار گنجایش تبادل کاتیونی بسته به نوع شرایط خاک متغیر است. میرخانی و همکاران (Mirkhani *et al.*, 2005) گزارش دادند که رس‌ها و مواد آلی خاک به علت سطح ویژه زیاد و باردار بودن نقش مهمی در CEC دارند و با افزایش مقدار رس و ماده آلی خاک مقدار CEC افزایش می‌یابد. کریمی‌ان (Karimian, 1996) با استفاده از ۱۵۰ نمونه خاک استان فارس، معادلات رگرسیونی چند متغیره ایجاد و سهم رس و مواد آلی در گنجایش تبادل کاتیونی را محاسبه و نیز گنجایش تبادل کاتیونی تعدادی از خاک‌ها را پس از اکسایش مواد آلی با آب اکسیژنه اندازه‌گیری کرد و سهم مواد آلی و رس را به طور مستقیم تعیین نمود.

مک دونالد (MacDonald, 1998) برای ایالت کبک^۲ کانادا رابطه $(CEC=2OC+0/5Clay)$ و برای ایالت آلبرتا^۳ در کانادا رابطه $(CEC=3/8OC+0/6Clay)$ و برای ناحیه هالیدمند نرفولک اونتاریو^۴ رابطه $(CEC=3/79+2/54OC+0/21Clay)$ را پیشنهاد کرد.

1-Cation exchange capacity

2-Quebec

3-Alberta

4-Haidimand-norfolk region of Ontario

منطق فازی به علت توانایی رقابت با هوشمندی انسانی و رهیافت سیستماتیک خود در بررسی شرایط و موقعیت‌های مبهم که ریاضیات متعارف چندان کارایی ندارد، ابزار تکنیکی طبیعی‌ای را برای ارزیابی پدیده‌ها و امور فراهم آورده است (Andriantiatsaholiniaina & Kouikoglou, 2004). منطق فازی ابزاری علمی است که امکان و اجازه شبیه‌سازی پویایی یک سیستم را بدون نیاز به توصیفات ریاضیاتی مفصل و با استفاده از داده‌های کیفی و کمی پدید آورده است (Phillis & Andriantiatsaholiniaina, 2001).

در اکثر موارد استفاده علمی از شاخص‌ها نیز به‌عنوان ابزار، به‌علت فقدان داده‌های مناسب (مرتبط، معتبر و کافی) محدود می‌شود (Kuswandari, 2004).

متاسفانه ابزارها و روش‌های سنتی کمی برای تصمیم‌گیری نیز به دلایلی از قبیل اطلاعات بیشتر کیفی تا کمی، معیارها و شاخص‌های ناکامل، تعاریف، تصورات و ادراکات متعارض، ریسک و عدم قطعیت و هزینه‌های بالای دسترسی به اطلاعات دقیق، به‌هنگام مواجهه با اطلاعات نادقیق و مفاهیم گنگ، بسیار ضعیف عمل می‌کنند که هر کدام از این موانع نیز در جای خود در خور توجه هستند (Ducy & Larson, 1999).

این پژوهش دو هدف عمده را دنبال می‌کند. هدف از این پژوهش، بررسی کارآیی مدل‌های رگرسیون خطی، فازی و فازی-ژنتیک در تخمین گنجایش تبادل کاتیونی خاک و امکان سنجی ارتقاء نتایج مدل فازی توسط سیستم الگوریتم ژنتیک بود.

مواد و روش‌ها

برای انجام پژوهش از پایگاه اطلاعات داده‌های خاک اروپا^۱ (IES) استفاده شد. به‌دلیل صرفه جویی کردن در وقت و هزینه، یعنی مدل با کمترین ورودی بیشترین دقت و صحت را داشته باشد و همچنین با توجه به تحقیقات انجام شده مک دونالد (MacDonald, 1998)، مک براتنی و همکاران

کرد. همچنین آن‌ها نشان دادند که رس (به جز در خاک‌هایی با مقدار رس ≤ 35 درصد) تأثیر گذارترین عامل در میزان CEC خاک‌های مورد مطالعه می‌باشد به طوری که با وارد نمودن متغیرهای دیگر بهبود چشم‌گیری در دقت و اعتبار توابع به‌دست نیامد.

یکی دیگر از روش‌های غیر مستقیم جهت برآورد گنجایش تبادل کاتیونی خاک، روش فازی است که می‌تواند جهت توسعه توابع انتقالی به‌کار رود. منطق فازی به‌دلیل در نظر گرفتن محدوده‌ای از امکان‌ها به جای اعداد، علاوه بر مزایای روش‌های آماری به‌دلیل قابلیت در فرموله نمودن دانش توصیفی بشری در قالب ریاضی، ابزاری سودمند در مدل‌سازی پدیده‌های طبیعی به‌شمار می‌آید.

فردریچ و همکاران (Freidrich *et al.*, 2002) از تئوری فازی جهت کمی کردن ویژگی‌های خاک بر اساس شکل ظاهرشان بهره بردند. حاج عباسی و همکاران (Hajabbasi *et al.*, 1997)، مک براتنی و همکاران (McBratney *et al.*, 2000)، محمدی و طاهری (Mohammadi & Tahery, 2005) و توربرت و همکاران (Torbert *et al.*, 2008) بکارگیری تئوری فازی را برای مشخصه‌های خاک در ارتباط با ارزیابی رویشگاه‌های مختلف استفاده کردند.

الگوریتم ژنتیک یک تکنیک برنامه‌نویسی است که از تکامل ژنتیکی به‌عنوان الگوی حل مسئله استفاده می‌کند (Ahmad & Simonovic, 2005). در این روش نخست برای تعداد ثابتی از جمعیت، مجموعه‌ای از داده‌ها و پارامترهای هدف، به‌طور تصادفی تولید می‌شود و افراد در برابر این مجموعه از داده‌ها آزمایش می‌شوند و مناسب‌ترین آن‌ها باقی می‌مانند و نسل جدید را شکل می‌دهند و این فرآیند برای نسل‌های بعدی تا بهینه‌سازی تکرار می‌شود. اینس و دروگرز (Ines & Droogers, 2002) الگوریتم ژنتیک را برای تخمین پارامترهای هیدرولیکی غیراشباع به‌کار بردند. نتایج این تحقیق نشان داد که می‌توان با کاربرد الگوریتم ژنتیک به خوبی پارامترهای هیدرولیکی خاک غیر اشباع را تخمین زد.

۱- Joint Research Center Institute for Environment and Sustainability, IES

ژنتیک استفاده گردید. سپس با استفاده از متغیرهای فوق مدل‌های رگرسیون خطی چند متغیره، فازی و فازی-ژنتیک به شرح ذیل طراحی گردید:

مدل رگرسیونی

آزمون آماری داده‌ها با استفاده از نرم افزار Statistica انجام گرفت. با استفاده از متغیرهای درصد رس و کربن آلی مناسب‌ترین معادله رگرسیون خطی چند متغیره برای برآورد گنجایش تبادل کاتیونی ارائه شد.

مدل فازی

برای طراحی مدل فازی در ابتدای امر با استفاده از روش سعی و خطا با توجه به شکل ۱ از دو نوع تابع عضویت مثلثی و دوزنقه‌ای شکل، با سه متغیر کلامی کم، متوسط و زیاد برای کربن آلی و پنج متغیر کلامی خیلی کم، کم، متوسط، زیاد و خیلی زیاد برای درصد رس و گنجایش تبادل کاتیونی تعریف شد. سپس قواعد فازی به شکل قواعد "اگر-آنگاه فازی" برای سیستم طراحی گردید. در کل در این تحقیق ۱۵ قاعده برای انجام استدلال فازی مورد استفاده قرار گرفت. در نهایت موتور استنتاج ممدانی برای انجام استدلال‌ها و معمول‌ترین روش غیر فازی کردن یعنی روش مرکز ثقل برای تعیین متغیرهای خروجی مورد استفاده قرار گرفت (Ansari & Davary, 2010).

مدل فازی - ژنتیک به منظور افزایش دقت سیستم فازی، توابع عضویت و وزن‌های قواعد استنتاج آن به وسیله الگوریتم ژنتیک بهینه شد و قواعد وزن‌دهی شده برای طراحی مدل فازی به کار رفت. روش تعیین توابع عضویت و اوزان قواعد در سیستم فازی-ژنتیک در زیر تشریح گردیده است:

تعیین جمعیت اولیه: در این پژوهش ۲۵ عدد کروموزوم به عنوان جمعیت اولیه در نظر گرفته شد.

تعیین عملگر پیوند: جهت پیوند بین کروموزوم‌های والد از عملگر محاسباتی استفاده شده است. چهار نوع عملگر پیوند وجود دارد: تک نقطه‌ای، دو نقطه‌ای، یکنواخت و محاسباتی. عملگرهای تک نقطه‌ای، دو

(McBratney *et al.*, 2002)، کشاورزی و همکاران (Keshavarzi *et al.*, 2011) و خداوردیلو و حسینی عربلو (Khodaverdiloo & Hosseini, 2014) در این پژوهش نیز اطلاعاتی نظیر میزان رس و کربن آلی و گنجایش تبادل کاتیونی برای ۷۰ نمونه خاک استخراج گردید. به منظور افزایش دقت و صحت مدلی که ارائه خواهد شد، کل مجموعه داده‌های گردآوری شده به دو دسته تقسیم شد. دسته اول که زیر مجموعه آموزشی (واسنجی مدل) نامیده می‌شود، شامل ۷۰ درصد کل مجموعه داده‌ها بود. دسته دوم زیر مجموعه صحت‌سنجی بوده و شامل ۳۰ درصد کل داده‌ها می‌باشد.

مدل‌سازی

برای تخمین گنجایش تبادل کاتیونی از مدل‌های رگرسیون خطی، فازی و فازی-ژنتیک استفاده شد و برای این منظور از نرم‌افزارهای Statistica و MATLAB استفاده گردید. به منظور افزایش دقت و کارایی مدل‌ها، متغیرهای ورودی و خروجی مدل‌ها با استفاده از رابطه (۱) استاندارد و در محدوده [۰ ۱] قرار گرفتند. برای تبدیل داده‌های تخمینی مدل از حالت استاندارد به مقادیر واقعی از رابطه (۲) استفاده شد.

$$X_n = (X_i - X_{\min}) / (X_{\max} - X_{\min}) \quad (1)$$

$$X_i = X_{\min} + (X_{\max} - X_{\min}) X_n \quad (2)$$

در این رابطه X_{\max} ، X_{\min} ، X_n ، X_i به ترتیب محدوده واقعی، محدوده نرمال، حداقل و حداکثر مقادیر داده‌ها می‌باشد. کشاورزی و همکاران (Keshavarzi *et al.*, 2011) با بررسی ۷۰ نمونه خاک از موقعیت‌های مختلف ۱۵ پروفیل خاک برای برآورد گنجایش تبادل کاتیونی خاک در منطقه زیران در استان قزوین از دو مدل منطق فازی و شبکه عصبی مصنوعی با استفاده از دو متغیر درصد رس و کربن آلی استفاده کردند و به این نتیجه رسیدند که روش فازی دارای دقت بیشتری نسبت به روش شبکه عصبی است.

در این تحقیق نیز از دو متغیر درصد رس و کربن آلی برای ایجاد مدل‌های رگرسیون خطی، فازی و فازی-

تعیین تابع هزینه: در این تحقیق هدف کمینه کردن مقدار جذر میانگین مربعات خطا (RMSE) می‌باشد، که هدف از آن این است که اختلاف بین گنجایش تبادل کاتیونی واقعی و تخمینی را به حداقل برساند یعنی میزان خطا را تقریباً صفر کند.

تعیین شرط خاتمه: در این تحقیق به خاطر این که توقف الگوریتم مستقل از نتیجه به دست آمده از الگوریتم است از شرط غیر فعال استفاده شد و حداکثر تولید نسل ۱۰۰ در نظر گرفته شد.

بعد از انجام مراحل فوق توابع عضویت بهینه شده متغیرهای کلامی برای هر کدام از متغیرهای ورودی (درصد رس و کربن آلی) و متغیر خروجی (CEC) به دست آمد. طبق جدول ۱ در سیستم فازی ژنتیک اوزان قواعد فازی دیگر برابر یک نبوده بلکه در بازه صفر الی یک قرار دارد.

نقطه‌ای و یکنواخت برای مسائل باینری و مسائل اعداد صحیح به کار می‌رود ولی عملگر محاسباتی برای مسائل اعداد حقیقی به کار می‌رود. در این پژوهش چون مسائل جزء اعداد حقیقی یا پیوسته است بنابراین از عملگر محاسباتی استفاده شد.

تعیین عملگر انتخاب: عملگر انتخاب به کار رفته، عملگر انتخاب تصادفی بدون جایگذاری (چرخ رولت) می باشد. دلیل به کارگیری عبارت «نمونه برداری تصادفی با جایگزین» برای این روش آن است که با انتخاب یک فرد به عنوان والدین، فرد مذکور، بخت انتخاب شدن دوباره برای تولید فرزند را از دست نمی‌دهد (Liu et al., 2007).

تعیین عملگر جهش: برای جلوگیری از قرارگیری نقطه بهینه سراسری در اکستریم‌های محلی از عملگر جهش با احتمال ۰/۰۲ برای هر ژن استفاده شد.

جدول ۱- اوزان قوانین در سیستم‌های فازی و فازی-ژنتیک

Table 1: Weights of rules in fuzzy and fuzzy-genetic systems

وزن قوانین سیستم فازی-ژنتیک Rules weights of fuzzy-genetic system	وزن قوانین سیستم فازی Rules weights of fuzzy system	گنجایش تبادل کاتیونی Cation exchange capacity	آنگاه Then	کربن آلی Organic carbon	و and	میزان رس Clay content	اگر If	شماره قانون Rule number
0.65	1	Very low کم خیلی	آنگاه Then	Low کم	و and	Very کم خیلی low	اگر If	1
0.17	1	Very low کم خیلی	آنگاه Then	Low کم	و and	Low کم	اگر If	2
0.58	1	Low کم	آنگاه Then	Low کم	و and	Middle متوسط	اگر If	3
...
0.33	1	Middle متوسط	آنگاه Then	high زیاد	و and	Low کم	اگر If	12
0.35	1	Middle متوسط	آنگاه Then	high زیاد	و and	Middle متوسط	اگر If	13
0.3	1	Very high خیلی زیاد	آنگاه Then	high زیاد	و and	high زیاد	اگر If	14
0.85	1	Very high خیلی زیاد	آنگاه Then	high زیاد	و and	Very خیلی زیاد high	اگر If	15

ارزیابی مدل‌ها

برای ارزیابی دقت مدل‌ها از شاخص‌های ضریب تبیین (R^2)، مجذور مربعات خطا^۱ (RMSE) و میانگین مطلق مطلق خطا^۲ (MAE) استفاده شد.

نتایج و بحث

نمای کلی از مقادیر آماری حداقل، حداکثر، میانگین و انحراف معیار برای خصوصیات درصد رس، کربن آلی و گنجایش تبادل کاتیونی خاک برای داده‌های سری واسنجی و صحت سنجی مدل‌ها به ترتیب در جدول ۲ نشان داده شده است.

معادله رگرسیون چند متغیره با استفاده از روش گام به گام^۳ به دست آمد. در نهایت معادله گنجایش تبادل کاتیونی برای خاک‌های مورد نظر طبق رابطه ۳ ارائه شد. تابع ارائه شده نشان داد که از بین متغیرهای مستقل درصد رس و کربن آلی موثرترین پارامتر برای تخمین CEC کربن آلی می‌باشد و این امر به دلیل سطح ویژه زیاد مواد آلی و داشتن گروه‌های عامل و تاثیر آن‌ها در افزایش CEC می‌باشد. گنجایش تبادل کاتیونی عموماً بر اساس بافت و مواد آلی خاک برآورد می‌شود و شکل کلی برآورد CEC به صورت (Mac Donald, 1998)

$$CEC = 0.008 + 0.616 \text{ Clay} + 1.058 \text{ OC} \quad (3)$$

Clay و OC به ترتیب نشان دهنده میزان رس و کربن آلی بوده و هر دو بر حسب درصد می‌باشد. شکل ۱ سطح برازشی داده‌های ورودی درصد رس و کربن آلی به داده‌های خروجی گنجایش تبادل کاتیونی را نشان می‌دهد. نتایج ارزیابی انواع توابع عضویت برای متغیرهای ورودی نشان داد که مناسب‌ترین تابع عضویت برای هر سه ویژگی مورد استفاده مثلثی و دوزنقه‌ای می‌باشد. کشاورزی و همکاران (Keshavarzi et al., 2011) برای تخمین گنجایش تبادل کاتیونی در منطقه زیران در استان قزوین از توابع عضویت مثلثی استفاده کردند. میترا و همکاران (Mitra et al.,

1998) نیز از توابع عضویت مثلثی و دوزنقه‌ای برای پیش‌بینی فرسایش خاک در حوضه‌های آبخیز واریگل در منطقه آراکانزاس امریکا استفاده نمودند. بازه تغییرات متغیرهای کلامی استاندارد شده پارامترهای موجود در سیستم فازی بر اساس روش آزمون و خطا، در جدول ۳ و شکل ۲ نشان داده شده است.

شکل ۳ سطح فازی ورودی‌های درصد رس و کربن آلی و خروجی CEC را نشان می‌دهد. با استفاده از این سطح فازی می‌توان با دادن مقادیر رس و کربن آلی مقادیر گنجایش تبادل کاتیونی را به دست آورد.

گرچه نوع توابع عضویت به دست آمده برای ویژگی‌های مورد استفاده در روش فازی-ژنتیک تفاوتی با روش مدل‌سازی فازی نداشت، لیکن بازه تغییرات هر متغیر کلامی در این روش متفاوت با روش فازی بود. بازه تغییرات متغیرهای کلامی استاندارد شده پارامترهای موجود در سیستم فازی-ژنتیک در جدول ۴ و شکل ۴ نشان داده شده است. شکل ۵ سطح فازی بهینه شده ورودی‌های درصد رس و کربن آلی و خروجی CEC را نشان می‌دهد. سطح فازی بهینه شده دارای تغییرات کم نسبت به خروجی برخوردار است و نسبت به سطح فازی در شکل ۳ مسطح و هموارتر است. جدول ۵ مقادیر آماره‌های مورد استفاده برای ارزیابی مدل‌ها را نشان می‌دهد. همانطور که مشاهده می‌شود، ضریب تبیین یا R^2 برای مدل فازی-ژنتیک بیشتر از مدل‌های فازی و رگرسیون خطی می‌باشد. همچنین مقادیر RMSE و MAE مدل فازی-ژنتیک کمتر از مدل‌های فازی و رگرسیون خطی است. مقدار R^2 برای مدل فازی بیشتر از مدل رگرسیون خطی و همچنین مقادیر پارامترهای RMSE و MAE برای مدل فازی کمتر از مدل رگرسیون خطی است. بنابراین تمامی آماره‌های ذکر شده نشان دهنده مزیت و کارایی بیشتر مدل فازی-ژنتیک نسبت به مدل‌های فازی و رگرسیون خطی می‌باشد. همچنین مدل فازی دارای دقت و کارایی بیشتری نسبت به مدل رگرسیون خطی برخوردار است.

1-Root mean squared error

2-Mean absolute error

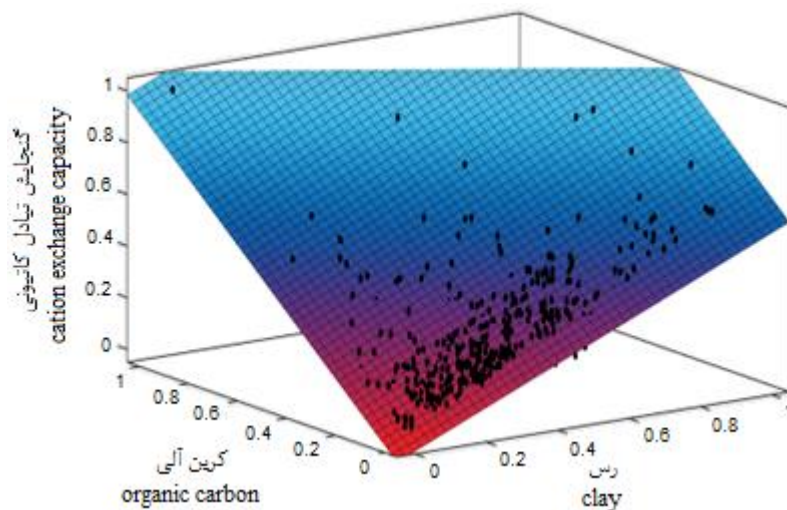
3-Stepwise

جدول ۲- مقادیر آماری داده‌های واسنجی و صحت‌سنجی مدل‌ها
Table 2: Statistical values of data to calibration and testing models

انحراف معیار Standard deviation	میانگین Average	حداکثر Maximum	حداقل Minimum	خصوصیات خاک Soil properties	داده‌های واسنجی مدل Experimental data
14.17	22.66	72	1	رس (%)	داده‌های صحت‌سنجی Testing data
1.48	1.33	14	0.03	کربن آلی (%)	
9.69	17.08	59.1	0.4	CEC (cmol _c kg ⁻¹)	
19.07	19.45	79	1	رس (%)	داده‌های صحت‌سنجی Testing data
2.33	1.41	14.5	0.03	کربن آلی (%)	
17.5	16.57	88.1	0.5	CEC (cmol _c kg ⁻¹)	

جدول ۲- مقادیر آماری داده‌های واسنجی و صحت‌سنجی مدل‌ها
Table 2: Statistical values of data to calibration and testing models

انحراف معیار Standard deviation	میانگین Average	حداکثر Maximum	حداقل Minimum	خصوصیات خاک Soil properties	داده‌های واسنجی مدل Experimental data
14.17	22.66	72	1	رس (%)	داده‌های صحت‌سنجی Testing data
1.48	1.33	14	0.03	کربن آلی (%)	
9.69	17.08	59.1	0.4	CEC (cmol _c kg ⁻¹)	
19.07	19.45	79	1	رس (%)	داده‌های صحت‌سنجی Testing data
2.33	1.41	14.5	0.03	کربن آلی (%)	
17.5	16.57	88.1	0.5	CEC (cmol _c kg ⁻¹)	

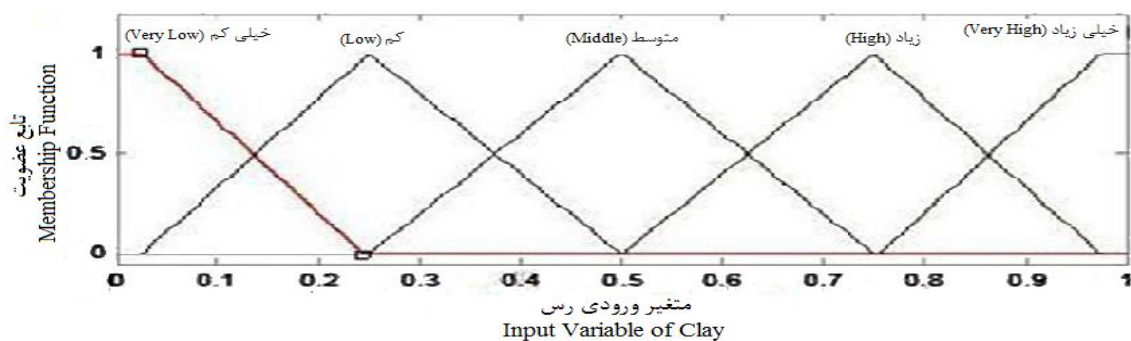


شکل ۱- سطح برازش داده‌های ورودی درصد رس و کربن آلی استاندارد شده به داده‌های خروجی گنجایش تبادل کاتیونی استاندارد شده

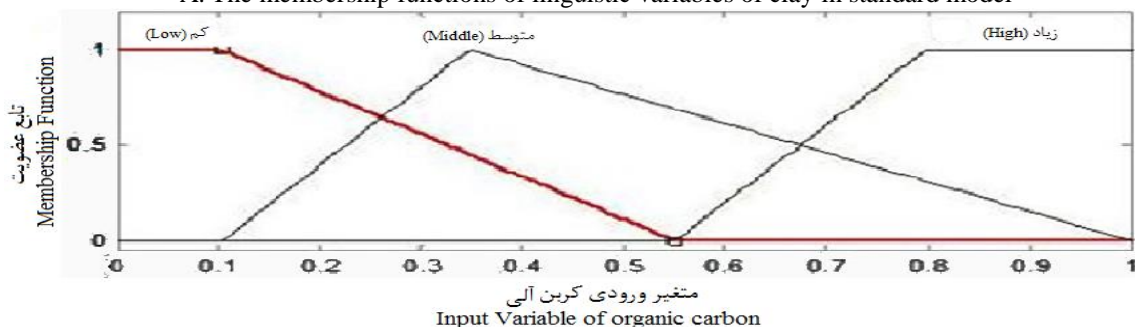
Fig.1. Fitted plan to the standardized values of organic carbon, clay content and cation exchange capacity

جدول ۳- بازه تغییرات متغیرهای کلامی استاندارد شده سیستم فازی
Table 3: Range of standardized linguistic variables in fuzzy systems

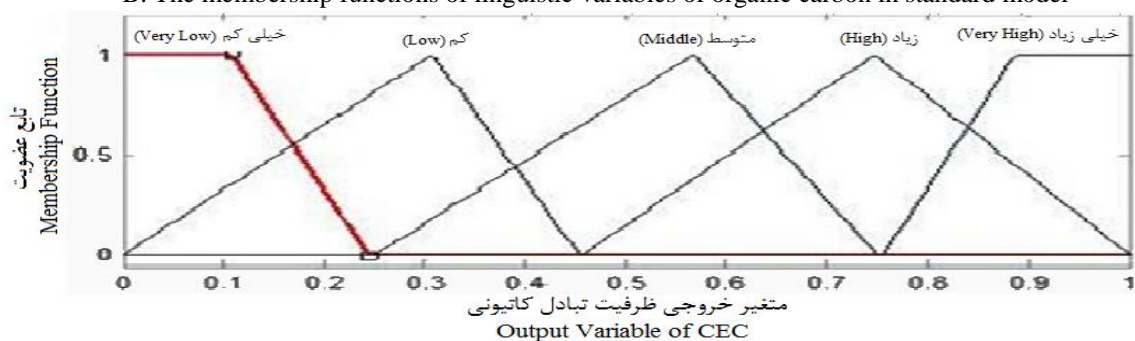
متغیر کلامی (linguistic variable)					
خیلی زیاد Very high	زیاد High	متوسط Middle	کم low	خیلی کم Very low	ویژگی Property
0.75-1	0.5-0.975	0.25-0.75	0.025-0.5	0-0.247	درصد رس (%) clay
---	0.55-1	0.104-0.348	0-0.55	---	درصد کربن آلی (%) Organic carbon
0.755-1	0.45-1	0.25-0.75	0-0.453	0-0.245	CEC (cmol _c kg ⁻¹)



الف- توابع عضویت متغیرهای کلامی ویژگی درصد رس در حالت استاندارد شده
A. The membership functions of linguistic variables of clay in standard model



ب- توابع عضویت متغیرهای کلامی ویژگی درصد کربن آلی در حالت استاندارد شده
B. The membership functions of linguistic variables of organic carbon in standard model



ج- توابع عضویت متغیرهای کلامی ویژگی گنجایش تبادل کاتیونی در حالت استاندارد شده
C. The membership functions of linguistic variables of cation exchange capacity in standard model

شکل ۲- بازه تغییرات متغیرهای کلامی در سیستم فازی
Fig.2. Range of standardized linguistic variables in fuzzy systems

جدول ۴- بازه تغییرات متغیرهای کلامی استاندارد شده سیستم فازی-ژنتیک

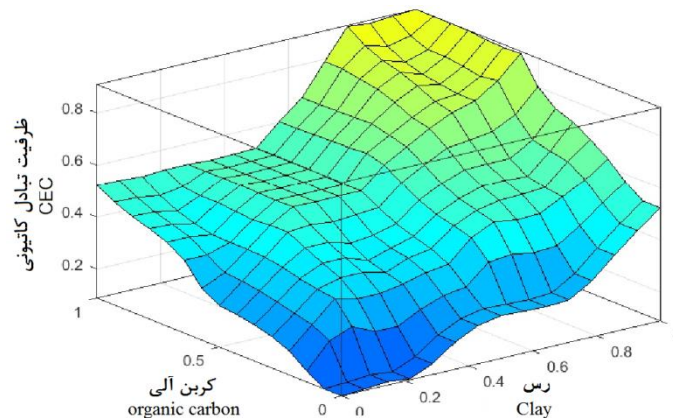
Table 4: Range of standardized linguistic variables in Fuzzy-Genetic systems

متغیر کلامی (linguistic variable)					
خیلی زیاد Very high	زیاد High	متوسط Middle	کم Low	خیلی کم Very low	ویژگی Property
0.555-1	0.405-0.625	0.255-0.555	0.11-0.405	0-0.255	درصد رس (%) clay
---	0.274-1	0.124-0.386	0-0.274	---	درصد کربن آلی (%) Organic carbon
0.505-1	0.366-0.648	0.236-0.505	0.087-0.366	0-0.236	CEC (cmol _c kg ⁻¹)

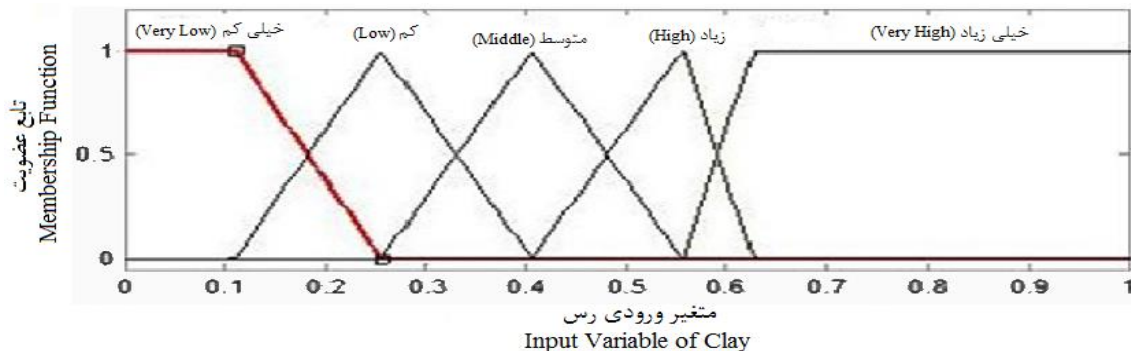
الگوریتم ژنتیک را به طور کاملاً موفق در پاسخ به مسائل غیر خطی و بزرگ مدیریت منابع آب و به دست آوردن منحنی‌های فرمان به کار بردند؛ آن‌ها در نتیجه گیری کارشان این روش را به مثابه یک ابزار قدرتمند اثبات کردند.

چانگ و همکاران (Chang *et al.*, 2005) در بررسی کارایی روش‌های IDW و فازی و نیز فازی-ژنتیک در برآورد متوسط بارندگی در یک منطقه به این نتیجه رسیدند که روش فازی ژنتیک با دقت بیشتری نسبت به سایر روش‌ها متوسط بارندگی منطقه را برآورد می‌نماید.

کای و همکاران (Cai *et al.*, 2001)، چن (Chen, 2003) و تانک و همکاران (Tung *et al.*, 2003)

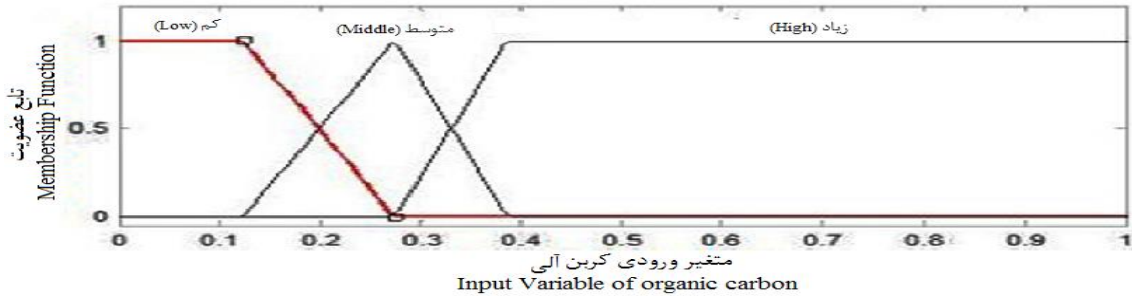


شکل ۳- سطح فازی برای ورودی درصد رس و کربن آلی استاندارد شده و خروجی گنجایش تبادل کاتیونی استاندارد شده
Fig.3. Fuzzy surface for standardized values of clay, organic carbon content and cation exchange capacity



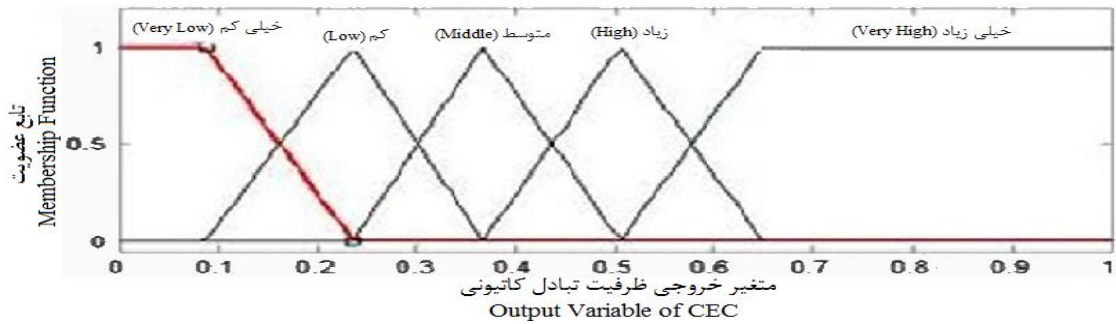
الف- توابع عضویت بهینه شده متغیرهای کلامی ویژگی درصد رس در حالت استاندارد شده

A. Optimized membership functions of linguistic variables of clay in standard model



ب- توابع عضویت بهینه شده متغیرهای کلامی ویژگی درصد کربن آلی در حالت استاندارد شده

B. Optimized membership functions of linguistic variables of organic carbon in standard model

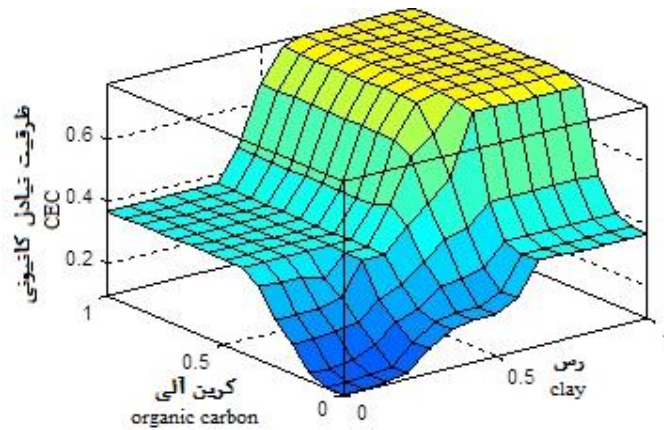


ج- توابع عضویت بهینه شده متغیرهای کلامی گنجایش تبادل کاتیونی در حالت استاندارد شده

C. Optimized membership functions of linguistic variables of cation exchange capacity in standard model

شکل ۴- بازه تغییرات توابع عضویت بهینه شده متغیرهای کلامی برای سیستم فازی-ژنتیک

Fig.4. Standardized range of optimized linguistic variables in Fuzzy-Genetic systems



شکل ۵- سطح فازی بهینه شده برای ورودی درصد رس و کربن آلی و خروجی گنجایش تبادل کاتیونی

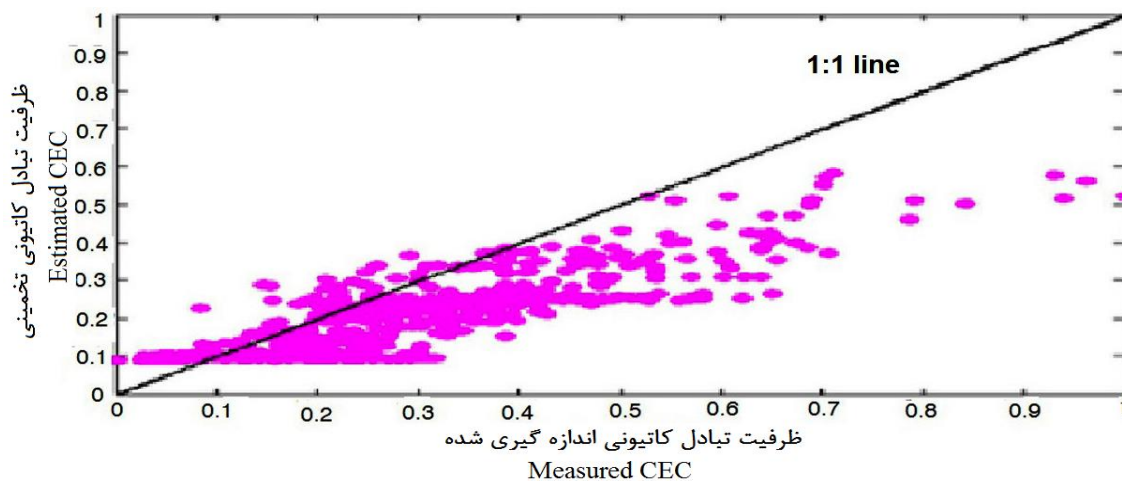
Fig.5. Optimized Fuzzy surface for standardized values of clay, organic carbon and cation exchange capacity

جدول ۵-مقادیر معیارهای ارزیابی مدل‌های ارائه شده برای تخمین گنجایش تبادل کاتیونی خاک
Table 5: Statistical criteria for proposed models in estimation of CEC

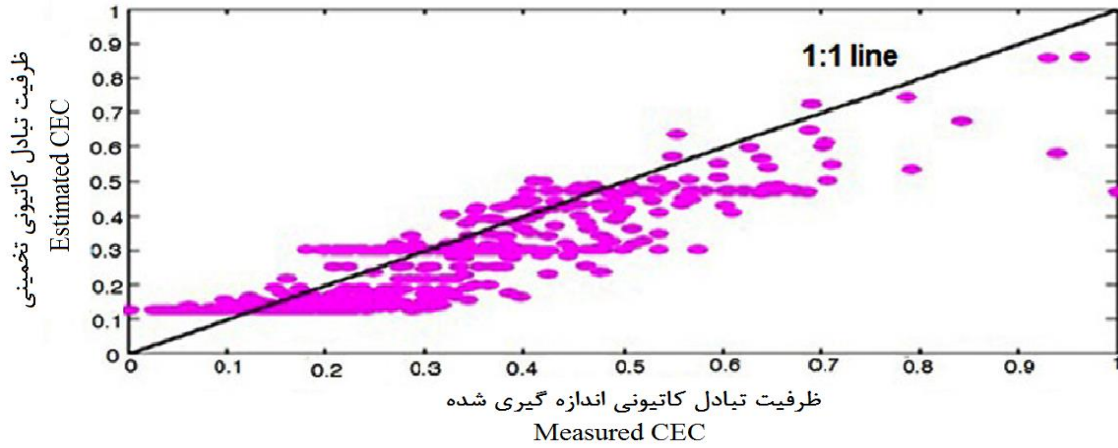
MAE	RMSE	R ²	داده‌های واسنجی مدل Experimental data	مدل (Model)
9.13	7.42	0.727		رگرسیون خطی Linear regression
4.32	5.42	0.781		فازی Fuzzy
3.57	4.7	0.844		فازی-ژنتیک Fuzzy-genetic
9.64	7.57	0.703		رگرسیون خطی Linear regression
8.2	9.5	0.758	داده‌های صحت‌سنجی Validation Data	فازی Fuzzy
7.94	9.3	0.783		فازی-ژنتیک Fuzzy-genetic

و در مدل فازی-ژنتیک داده‌ها به خط یک به یک نزدیکتر هستند که نشان دهنده دقت بیشتر مدل فازی-ژنتیک است. اسدی‌پور و همکاران (Asadipour *et al.*, 2012) با استفاده از ترکیب الگوریتم ژنتیک و منطق فازی، هیدروگراف خروجی بهینه را بر اساس هیدروگراف ورودی به مخزن سد کارون ۳ تعیین کرده‌اند. نتیجه تحقیق آن‌ها، بیانگر قابلیت مناسب روش پیشنهادی در بهینه‌سازی عملکرد مخزن است.

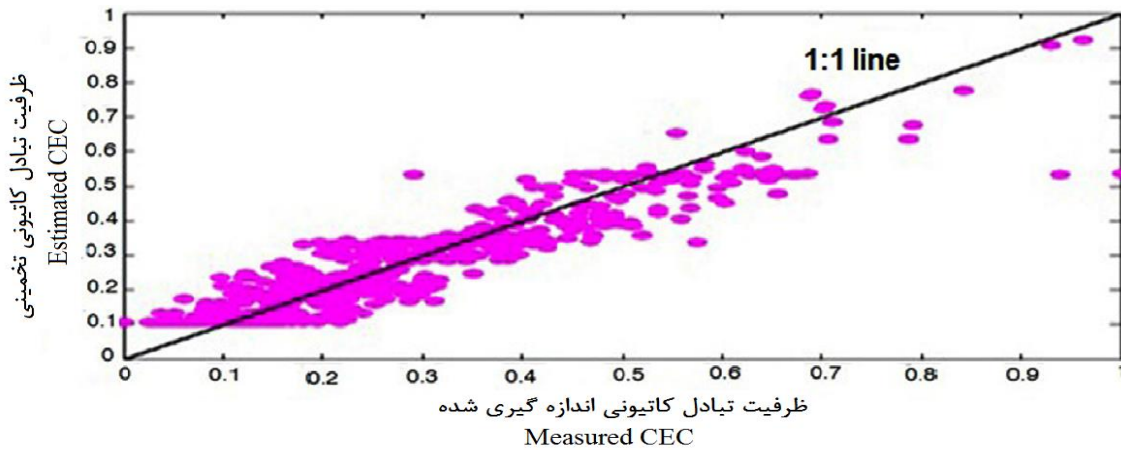
شکل ۶ به ترتیب خط یک به یک بین مقادیر اندازه‌گیری شده و برآوردی CEC توسط مدل‌های رگرسیون خطی، فازی و فازی-ژنتیک را نشان می‌دهند. هر چه پراکندگی داده‌ها حول این خط کمتر باشد مقادیر برآوردی توسط مدل به مقادیر اندازه‌گیری شده نزدیکتر خواهد شد. طبق این شکل‌ها پراکندگی داده‌ها در برآورد CEC در مدل فازی-ژنتیک کمتر از مدل‌های فازی و رگرسیون خطی است.



الف-مقادیر استاندارد اندازه‌گیری شده و برآوردی CEC بر حسب $\text{cmol}_c \text{kg}^{-1}$ توسط مدل رگرسیون خطی
A. Standard values of measured and estimated CEC ($\text{cmol}_c \text{kg}^{-1}$) by linear regression model



ب-مقادیر استاندارد اندازه‌گیری شده و برآوردی CEC بر حسب $\text{cmol}_c \text{kg}^{-1}$ توسط مدل فازی
B. Standard values of measured and estimated CEC ($\text{cmol}_c \text{kg}^{-1}$) by Fuzzy model



ج-مقادیر استاندارد اندازه‌گیری شده و برآوردی CEC بر حسب $\text{cmol}_c \text{kg}^{-1}$ توسط مدل فازی-ژنتیک
C. Standard values are measured and estimated CEC ($\text{cmol}_c \text{kg}^{-1}$) by Fuzzy-genetic model

شکل ۶- پراکنش مقادیر اندازه‌گیری شده و تخمینی گنجایش تبادل کاتیونی

Fig.6. Distribution of measured and estimated values for cation exchange capacity

نتیجه‌گیری کلی

در این پژوهش با استفاده از مدل‌های رگرسیون چند متغیره، فازی و فازی-ژنتیک مقادیر گنجایش تبادل کاتیونی برآورد گردید. مدل‌های طراحی شده با دو متغیر ورودی درصد رس و کربن آلی، توانستند میزان CEC را با دقت مناسبی برآورد نمایند. بنابراین استفاده از این مدل‌ها در پیش‌بینی CEC توصیه می‌شود.

در این مطالعه از توابع عضویت مثلثی و ذوزنقه‌ای در سیستم فازی و فازی-ژنتیک استفاده گردید. گرچه

نوع توابع عضویت به‌دست آمده برای ویژگی‌های مورد استفاده در روش فازی-ژنتیک تفاوتی با روش مدل‌سازی فازی نداشت، لیکن بازه تغییرات هر متغیر کلامی در این روش متفاوت با روش فازی بود. تعیین وزن توابع عضویت و قواعد مورد استفاده در مدل فازی توسط تکنیک الگوریتم ژنتیک سبب افزایش دقت مدل فازی طراحی شده گردید. بنابراین استفاده از این تکنیک در مطالعات مربوط به CEC برای طراحی مدل‌های فازی توصیه می‌شود.

آماره‌های ذکر شده نشان دهنده مزیت و کارایی بیشتر مدل فازی-ژنتیک نسبت به مدل‌های فازی و رگرسیون خطی می‌باشد. همچنین مدل فازی دارای دقت و کارایی بیشتری نسبت به مدل رگرسیون خطی برخوردار است.

ضریب تبیین یا R^2 برای مدل فازی-ژنتیک بیشتر از مدل‌های فازی و رگرسیون خطی می‌باشد. همچنین مقادیر RMSE و MAE مدل فازی-ژنتیک کمتر از مدل‌های فازی و رگرسیون خطی است. مقدار R^2 برای مدل فازی بیشتر از مدل رگرسیون خطی و همچنین مقادیر پارامترهای RMSE و MAE برای مدل فازی کمتر از مدل رگرسیون خطی است. بنابراین تمامی

References

- Ahmad, S., and Simonovic, S. P. (2005). An artificial neural network model for generating hydrograph from hydro-meteorological parameters. *Journal of Hydrology*, 315(1), 236-251.
- Altin, A., and Degirmenci, M. (2005). Lead (II) removal from natural soils by enhanced electrokinetic remediation. *Science of the Total Environment*, 337(1), 1-10.
- 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*, 56(4), 551-559.
- Andriantiatsaholiniaina, L. A., Kouikoglou, V. S., and Phillis, Y. A. (2004). Evaluating strategies for sustainable development: fuzzy logic reasoning and sensitivity analysis. *Ecological Economics*, 48(2), 149-172.
- Ansari, H. and Davary, K. (2010). Estimating Precipitation Data Using a Fuzzy based Technique. *Iran-Water Resources Research*, 6 (1), 39 -47.
- Arias, M., Perez-Novo, C., Osorio F., Lopez E. and Soto, B. (2005). Adsorption and desorption of copper and zinc in the surface layer of acid soils. *Journal of Colloid and Interface Science*, 288 (1), 21-29.
- Asadipour, N., Karami, M. and Shahinezhad, B. (2012). Genetic algorithm using fuzzy logic in determining the parameters for optimizing the outflow hydrograph. *Journal of Iran Water Research*, 6 (10), 37-45, (In Persian).
- Bouma, J. (1989). Using soil survey data for qualitative land evaluation. *Advances in Soil Science*, 9, 177-213.
- Cai, X., McKinney, D. C. and Lasdon, L. S. (2001). Solving nonlinear water management models using a combined genetic algorithm and linear programming approach. *Advances in Water Resources*, 24 (1), 667-676.
- Chang, C. L., Lo, S. L. and Yu, S. L. (2005). Applying fuzzy theory and genetic algorithm to interpolate precipitation. *Journal of Hydrology*, 314 (1-4), 92-104.
- Chen, L. (2003). Real coded genetic algorithm optimization of long-term reservoir operation. *Environment and Urbanization*, 39(5), 1157-1165.
- Ducey, M. J. and Larson, B.C. (1999). A fuzzy set approach to the problem of sustainability. *Forest Ecology and Management*, 115 (1), 29-40.
- Freidrich, C., Fohrer, N. and Frede, H. G. (2002). Quantification of soil properties based on external information by means of fuzzy-set theory. *Journal of Plant Nutrition and Soil Science*, 165 (4), 511-516.
- Hajbasi, M. A., Jalalian, A. and Karimzadeh, H. R. (1997). Deforestation effects on soil physical and chemical properties, Lordegan, Iran. *Plant and Soil*, 190 (2), 301-308.
- Hosseini Arabloo, N., Khodaverdiloo, H., Ghorbani Dashtaki, S. H. and Momtaz, H. R. (2015). Effect of grouping soils based on their organic carbon and clay content on performance of hierarchical pedotransfer functions of soil cation exchange capacity. *Journal of Soil Management and Sustainable Production*, 4 (4), 215-234, (In Persian).

- Ines, A. V. M. and Droogers, P. (2002). Inverse modelling in estimating soil hydraulic functions: a Genetic Algorithm approach. *Hydrology and Earth System Sciences*, 6 (1), 49-65.
- Karimian, A. 1996. The role of clay and organic matter in cation exchange capacity of calcareous soils in Fars province. *5th Soil Science Congress*, Karaj, Iran, (In Persian).
- Keller, A., von Steiger, B., van Der Zee, S. T. and Schulin, R. (2001). A stochastic empirical model for regional heavy metal balances in agro ecosystems. *Journal of Environmental Quality*, 30 (6), 1976-1989.
- Keshavarzi, A., Sarmadian, F., Labbafi, R. and Rajabi Vandechali, M. (2011). Modeling of Soil Cation Exchange Capacity Based on Fuzzy Table Look-up Scheme and Artificial Neural Network Approach. *Journal of Modern Applied Science*, 5 (1), 153-164.
- Khodaverdiloo, H. and 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 Water and Soil Science*, 18 (67), 311-320, (In Persian).
- Krogh, L., Breuning-Madsen, H. and Greve, M. H. (2000). Cation exchange capacity pedotransfer function for Danish soils. *Plant and Soil*, 50 (1), 1-12.
- Kuswandari, R. (2004). Assessment of Different Methods for Measuring the Sustainability of Forest Management. International Institute for Geo-information Science and Earth Observation. *Enschede*, Netherlands.
- Liu, S. h., Butler, D., Brazier, R., Heathwaite, L. and Khu, S. (2007). Using genetic algorithm to calibrate a water quality model. *Science of the Total Environment*, 374 (2-3), 260-272.
- MacDonald, K. B. (1998). Development of pedotransfer functions of southern Ontario soils Report from. *Greenhouse and processing crops research center. Harrow, Ontario*. No: 01686-8-0436:1-23.
- 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.
- McBratney, A., Odeh, O. A., Bishop, T. F., Dunbar, M. S. and Shatar, T. M. (2000). An overview of pedometric techniques for use in soil survey. *Geoderma*, 97 (3-4), 293-327.
- McBratney, A. B., Minasny, B., Cattle, S. R. and Vervoort, R. W. (2002). From pedotransfer functions to soil inference systems. *Geoderma*, 109 (12), 41-73.
- Mirkhani, R., Shabanpour, M. and Saadat, S. (2005). Using particle size distribution and organic carbon percentage to predict the cation exchange capacity of soils of Lorestan province. *Iranian Journal of Soil Research*, 19 (2): 235-242, (In Persian).
- Mitra, B., Scott, H. D., Dixon, J. C. and McKimmey, J. M. (1998). Applications of fuzzy logic to the prediction of soil erosion in a large watershed, *Geoderma*, 86 (3), 183-209.
- Mohammadi, J. and Taheri, M. (2005). Estimation of pedotransfer function using fuzzy regression. *Journal of Agriculture Science and Technology*, 2, 51-60, (In Persian).
- Phillis, Y. A. and Andriantiatsaholiniaina L. A. (2001). Sustainability: an ill-defined concept and its assessment using fuzzy logic. *Ecological Economics*, 37(3), 435-456.
- Torbert, H. A., Krueger, E. and Kurtener, D. (2008). Soil quality assessment using fuzzy modeling. *International Agrophysics*, 22 (4), 365-370.
- Tung, C., Hsu, S., Liu, C. M. and Li, J. (2003). Application of the genetic algorithm for optimizing operation rules of the LiYuTan reservoir in Taiwan. *JAWRA Journal of the American Water Resources Association*, 39 (3), 649-657.
- Wosten, J. H. M., Pachepsky, Y. A. and Rawls, W. J. (2001). Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251 (3), 123-150.

Comparison of linear regression, Fuzzy and Fuzzy-genetic models to predict soil cation exchange capacity

Habib Palizvanzand¹ and Abbas Ahmadi^{2*}

(Received: November 2014

Accepted: October 2015)

Abstract

Cation exchange capacity (CEC) is one of the most important soil chemical properties that affects other chemical, physical, and biological soil properties and fertility. In this study, performance of some procedures such as regression, Fuzzy and Fuzzy-genetic approaches in estimation of soil CEC has been investigated. Consequently, the required data of 770 samples from the Europe database (IES) was extracted. Then multiple linear regression, Fuzzy and Fuzzy-genetic approaches were used for development of pedotransfer functions for estimating of soil CEC. The coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE) criteria were used for evaluation of the proposed models. The values of R^2 , RMSE and MAE obtained for the linear regression model 0.72, 7.42 $\text{cmol}_c \text{kg}^{-1}$ and 9.13 $\text{cmol}_c \text{kg}^{-1}$, for Fuzzy model, 0.78, 5.44 $\text{cmol}_c \text{kg}^{-1}$ and 4.32 $\text{cmol}_c \text{kg}^{-1}$, for Fuzzy-genetic model 0.84, 4.7 $\text{cmol}_c \text{kg}^{-1}$ and 3.57 $\text{cmol}_c \text{kg}^{-1}$, respectively. These results indicated that the Fuzzy-genetic CEC model is more accurate than Fuzzy model, and Fuzzy model is more accurate than regression CEC model.

Keywords: Clay, Fuzzy rules, Organic carbon, Pedotransfer functions

1- MSc Student of Soil Science, University of Tabriz

2- Assistant Professor, Dept. of Soil Science, University of Tabriz.

* Corresponding author Email: a_ahmadi@tabrizu.ac.ir