

## پیش‌بینی شوری خاک با استفاده از رگرسیون درختی و شبکه عصبی مصنوعی در منطقه قروه استان کردستان

شیرین مرادیان<sup>۱</sup>، \* کمال نبی‌اللهی<sup>۲</sup> و روح‌الله تقی‌زاده مهرجردی<sup>۳</sup>

<sup>۱</sup> دانش‌آموخته کارشناسی ارشد گروه علوم و مهندسی خاک، دانشگاه کردستان، آستادیار گروه علوم و مهندسی خاک، دانشگاه کردستان،

<sup>۲</sup> آستادیار گروه مدیریت مناطق خشک و بیابانی، دانشگاه اردکان

تاریخ دریافت: ۹۶/۲/۱۰؛ تاریخ پذیرش: ۹۶/۶/۱۵

### چکیده

**سابقه و هدف:** شوری خاک یکی از مشکلات عمده در مناطق خشک و نیمه‌خشک است. در این شرایط، نمک‌های محلول در سطح خاک تجمع یافته و باعث کاهش عملکرد و حاصلخیزی خاک می‌شوند. شناسایی و نقشه‌برداری خاک‌های مبتلا به نمک می‌تواند به بهبود مدیریت این خاک‌ها کمک کند. بررسی تغییرات شوری خاک به شیوه‌های مرسوم گران و زمان‌بر است. بنابراین یکی از راه‌های چاره جهت حل این چالش استفاده از نقشه‌برداری رقومی خاک است که خصوصیات خاک با استفاده از داده‌های کمکی نقشه‌برداری می‌شوند. هدف از این پژوهش استفاده از مدل‌های رگرسیون درختی و شبکه عصبی مصنوعی و داده‌های کمکی برای تهیه نقشه شوری خاک می‌باشد.

**مواد و روش‌ها:** با استفاده از روش نمونه‌برداری هایپرکیوب تعداد ۱۵۰ نمونه خاک از عمق ۰-۳۰ سانتی‌متری خاک‌های منطقه قروه استان کردستان (با وسعت ۳۰۰۰۰ هکتار) برداشت شده و هدایت الکتریکی خاک اندازه‌گیری شد. متغیرهای محیطی در این پژوهش اجزاء سرزمین و داده‌های تصویر ETM<sup>+</sup> ماهواره لندست ۸ بودند. پارامترهای سرزمین (شامل ۱۵ پارامتر) و شاخص شوری (SI) و شاخص گیاهی نرمال شده (NDVI) به ترتیب با استفاده از نرم‌افزار SAGA و ArcGIS محاسبه و استخراج گردید. جهت ایجاد ارتباط بین شوری خاک و متغیرهای کمکی از مدل رگرسیون درختی و شبکه عصبی بهره گرفته شد و با استفاده از روش اعتبارسنجی مورد ارزیابی قرار گرفت. در نهایت نقشه شوری خاک با استفاده از مدل بهتر تهیه شد.

**یافته‌ها:** برای پیش‌بینی شوری خاک، متغیرهای کمکی شامل شاخص شوری، شاخص خیسی، شاخص همواری دره، شاخص NDVI، باند ۳ و باند ۷ مهم‌ترین بودند. نتایج این پژوهش نشان داد که مدل شبکه عصبی مصنوعی (با ۰/۸۳، ۰/۳۶ و ۰/۱۹۰ به ترتیب ۰ برای ضریب تبیین، میانگین خطا و میانگین ریشه مربعات خطا) دارای دقت بیشتری نسبت به رگرسیون درختی برای پیش‌بینی شوری خاک می‌باشد. شوری خاک در محدوده بین ۶/۹۳-۰/۲۳ دسی‌زیمنس بر متر قرار داشت و بیش‌ترین مقادیر شوری خاک در مناطق مرکزی (اراضی پست و بایر) قرار داشت. در این مناطق مرکزی، داده‌های کمکی شامل شاخص شوری، شاخص همواری دره، شاخص خیسی، باند ۷ و باند ۳ بیش‌ترین مقدار و شاخص NDVI کم‌ترین مقدار را داشتند.

\* مسئول مکاتبه: [k.nabiollahi@uok.ac.ir](mailto:k.nabiollahi@uok.ac.ir)

**نتیجه‌گیری:** مهم‌ترین متغیر کمکی در پیش‌بینی شوری خاک در منطقه شاخص شوری می‌باشد و وجود ارتباط قوی بین داده خاک و داده‌های کمکی می‌تواند بر روی دقت مدل اثرگذار باشد. به‌طورکلی نتایج نشان داد که تکنیک‌های پدومتری می‌تواند در گستره‌ای وسیع جهت نقشه‌برداری رقومی خصوصیات خاک‌ها به‌کار گرفته شود. پیشنهاد می‌گردد که جهت تهیه نقشه خصوصیات خاک از مدل شبکه عصبی مصنوعی، و داده‌های کمکی هم‌چون اجزاء سرزمین و تصاویر ماهواره‌ای در مطالعات آینده استفاده شود.

**واژه‌های کلیدی:** تصویر ماهواره‌ای، مدل رقومی ارتفاع، پدومتری

### مقدمه

شوری خاک یکی از مشکلات عمده در مناطق خشک و نیمه‌خشک می‌باشد (۷). در این شرایط، نمک‌های محلول در سطح خاک تجمع یافته و باعث کاهش عملکرد و حاصلخیزی خاک می‌شوند. شناسایی و نقشه‌برداری خاک‌های مبتلا به نمک می‌تواند به اصلاح و بهبود مدیریت آن‌ها کمک کند. این در حالی است که بررسی تغییرات شوری خاک به شیوه‌های مرسوم به‌ویژه در مناطق با وسعت زیاد، گران و زمان‌بر است. در روش‌های سنتی صحت نقشه متأثر از کیفیت داده‌های جمع‌آوری شده می‌باشد، مثلاً تجربه صحرایی فردی که شناسایی خاک را انجام می‌دهد می‌تواند مؤثر باشد (۱۳) و همچنین به‌روز کردن نقشه‌ها و اطلاعات خاک به هزینه و زمان زیادی نیاز دارد. از این‌رو برای پیش و بررسی روند تغییرات شوری خاک در سطوح گسترده نیاز به روش‌هایی می‌باشد که بتواند به‌راحتی سطح گسترده‌ای از مناطق را پیش کرده و روند شوری خاک را مورد بررسی قرار دهد. یکی از راه‌های چاره جهت غلبه بر این چالش استفاده از روش‌های پدومتری است. امروزه مشخص شده دقت روش‌های پدومتری جهت برآورد توزیع مکانی داده‌های اندازه‌گیری شده خاک، به‌علت در نظر گرفتن پیوستگی مکانی داده‌ها، نسبت به روش‌های معمول آماری بیشتر است (۱۶). نقشه‌برداری رقومی خاک به‌عنوان زیرمجموعه‌ای از روش‌های پدومتری، به‌دلیل

افزایش منابع داده‌های کمکی یا محیطی فراوان، به‌صورت کاملاً عملی توانسته است به انتقادات وارده به خاکشناسی مرسوم فائق آید. در نقشه‌برداری رقومی خاک، خصوصیات خاک بر اساس ارتباطشان با متغیرهای محیطی به‌صورت دیجیتالی نقشه‌برداری می‌شوند (۱۷). نقشه‌برداری رقومی در واقع توسعه یک مدل عددی یا آماری از رابطه بین متغیرهای محیطی و خصوصیات خاک است که برای داده‌های جغرافیایی به‌منظور تولید نقشه رقومی به‌کار می‌رود. در نقشه‌برداری رقومی رابطه تغییرات یک ویژگی خاک با عوامل خاک‌سازی، مانند توپوگرافی، آب‌وهوا، کاربری اراضی، پوشش گیاهی و نوع خاک بیان می‌شود به این صورت که در هر موقعیت مشاهده خاک، یک سری متغیرهای محیطی وجود دارد که همبستگی بالایی با خصوصیات خاک دارند که یافتن این ارتباط و پیش‌بینی خصوصیات یا کلاس‌های خاک در نقاط دیگر، منجر به تهیه نقشه رقومی خصوصیات خاک در تمام منطقه مورد مطالعه می‌گردد (۱۷). بر اساس مطالعات پیشین نقشه مدل ارتفاع رقومی و داده‌های سنجش از دور از جمله مهم‌ترین داده‌های کمی هستند که جهت پیش‌بینی خصوصیات خاک می‌توانند مفید باشند (۲ و ۲۹). شاخص شوری، معیاری از میزان شوری است، مقادیر زیاد آن نشان‌دهنده میزان شوری بیش‌تر می‌باشد. اژیرایی و همکاران (۲۰۱۵)

نشان دادند که بین مقادیر هدایت الکتریکی عصاره اشباع خاک با باند ۴، شاخص روشنایی، شاخص سبزیگی و شاخص به دست از ادغام باند پانکروماتیک با باند ۴ ارتباط معنی دار مثبت در سطح ۱ درصد وجود دارد (۲۵).

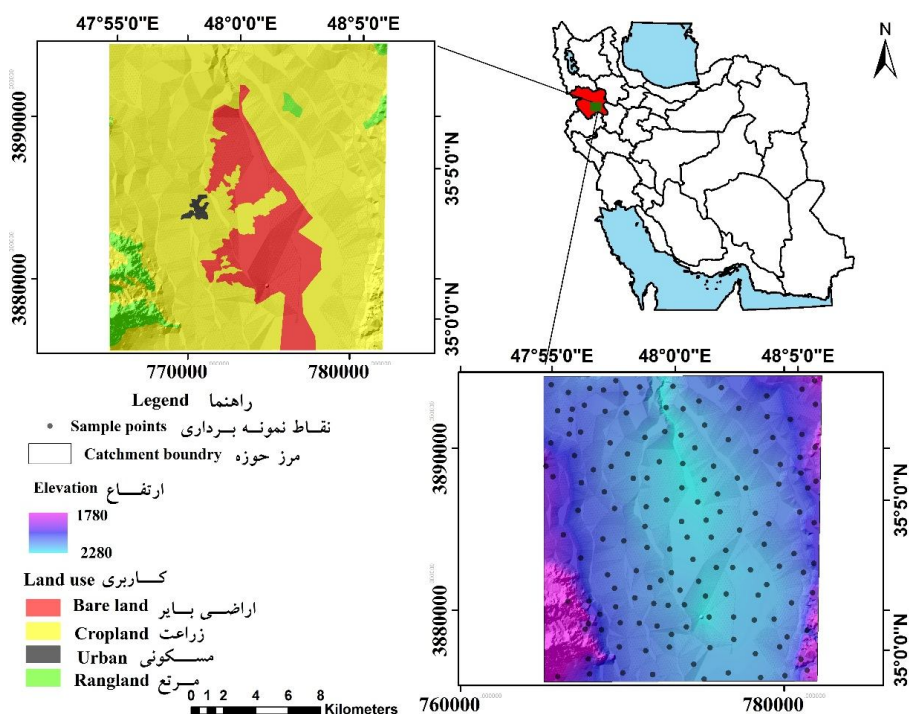
در سال‌های اخیر، تکنیک‌های مختلفی هم‌چون جنگل تصادفی (۲۲)، فازی (۱۱)، شبکه عصبی مصنوعی (۲۹ و ۶)، رگرسیون درختی (۲۸ و ۱۲)، رگرسیون لوجستیک (۳۱) و زمین‌آمار (۲۳ و ۳۲) توسط پژوهشگران جهت نقشه‌برداری رقومی به کار برده شده است. مدل درختی طبقه‌بندی و رگرسیون، یک روش ناپارامتری الگوریتمی است (۲۷) که قادر به پیش‌بینی متغیرهای کمی (شوری خاک) یا متغیرهای طبقه‌بندی شده (کلاس‌های خاک) بر اساس مجموعه‌ای از متغیرهای پیش‌بینی‌کننده کمی و کیفی است که برای اولین بار این روش توسط بریمن و همکاران (۱۹۸۴) گسترش داده شد (۴). شبکه عصبی مصنوعی هم به‌عنوان یکی از معمول‌ترین تکنیک‌های نقشه‌برداری رقومی با شبیه‌سازی و الهام از مطالعه سیستم مغز و شبکه عصبی موجودات زنده اشتقاق یافته است (۱۹). در این روش سعی می‌شود، بر اساس روابط ذاتی میان داده‌ها، مدلی خطی یا غیرخطی بین متغیرهای مستقل و وابسته برقرار شود (۱۵). خیر و همکاران (۲۰۱۰) از مدل رگرسیون درختی برای پهنه‌بندی نوزیع مقدار روی با استفاده از داده‌های کمکی پارامترهای سرزمین در لبنان استفاده کردند (۱۴). تقی‌زاده و همکاران (۲۰۱۴) در منطقه اردکان اقدام به تهیه نقشه سه‌بعدی شوری خاک با استفاده از مدل رگرسیون درختی کردند (۲۸). همچنین تقی‌زاده و همکاران (۲۰۱۳) از مدل‌های نروفازی، الگوریتم ژنتیک، شبکه عصبی، و رگرسیون چندمتغیره در پیش‌بینی شوری خاک استفاده کردند (۳۰).

در پهنه‌بندی شوری خاک در مزرعه نمونه ارتش گرگان با استفاده از شاخص‌های مختلف استخراج شده از تصاویر ماهواره لندست نشان دادند که شاخص‌های شوری بیش‌ترین همبستگی را با داده‌های زمینی دارند (۳). مترنیچ و زینک (۲۰۰۳) نشان دادند که به دلیل تأثیر گیاهان در بازتاب شوری، گیاهان می‌توانند به‌عنوان شاخص غیرمستقیم در سنجش از دور خاک مورد استفاده قرار گیرد، و بیش‌ترین همبستگی بین داده‌های NDVI و هدایت الکتریکی خاک وجود دارد که تفکیک خاک‌های شور و قلیا را از مناطق غیرشور ممکن می‌سازد (۱۸) و از این فاکتور به‌طور وسیع در مطالعات نقشه‌برداری رقومی استفاده شده است (۶، ۸ و ۱۰). شاخص خیسی، تأثیر توپوگرافی بر مکان و میزان تجمع رطوبت در خاک، یا در سطح زمین‌نما را نشان می‌دهد که با افزایش درصد شیب میزان آن کاهش می‌یابد. با کاهش شاخص خیسی و کم شدن میزان تجمع رطوبت در خاک، نفوذپذیری خاک نیز افزایش پیدا می‌کند و به‌تبع آن شوری خاک هم می‌تواند تحت تأثیر قرار گرفته و کاهش یابد. همچنین شاخص همواری دره، قسمت‌های هموار پایین دره را نشان می‌دهد که با افزایش درصد شیب میزان آن کاهش می‌یابد. با کاهش شاخص همواری دره مناطق مسطح و هموار در سطح زمین‌نما و نفوذپذیری خاک، کاهش یافته که به دنبال آن شوری خاک می‌تواند افزایش پیدا کند. عبدی‌نام (۲۰۰۴) در تهیه نقشه شوری خاک با استفاده از ایجاد همبستگی بین داده‌های ماهواره‌ای با مقادیر شوری خاک در دشت قزوین به دلیل وجود همبستگی بالای ارقام رقومی باند ۷ تصاویر ETM<sup>+</sup> با مقادیر شوری، از داده‌های این باند استفاده کرد (۱). تاج‌گردان و همکاران (۲۰۰۹) در تهیه نقشه شوری در منطقه آق‌قلای استان گلستان با استفاده از داده‌های ETM<sup>+</sup>

### مواد و روش‌ها

**محدوده مطالعاتی:** منطقه مورد مطالعه حوزه‌ای با وسعت ۳۰/۰۰۰ هکتار بوده که در شمال شرقی شهرستان قروه در استان کردستان (شامل روستاهای دزج، صندوق‌آباد، آب‌باریک، زرینه، زنگ‌آباد، شهاب‌الدین، گرمخانی، نارنجک و شکوه‌آباد) واقع شده است (شکل ۱). ارتفاع بلندترین نقطه ۲۲۸۰ متر و ارتفاع پایین‌ترین نقطه ۱۷۸۰ متر از سطح دریا است. متوسط بارندگی سالیانه ۳۶۹/۸ میلی‌متر و متوسط درجه حرارت سالیانه ۱۰/۲ درجه سانتی‌گراد می‌باشد. رژیم رطوبتی و حرارتی خاک منطقه به ترتیب زیریک و مزیک می‌باشد. واحدهای فیزیوگرافی منطقه شامل اراضی پست، دشت‌های دامنه‌ای، فلات و تپه و کاربری‌های اصلی منطقه شامل زراعت و مرتع می‌باشد. از لحاظ زمین‌شناسی منطقه عمدتاً حاوی آبرفت، مرمر، شیست، سنگ آهک، مارن و دولومیت می‌باشد.

منطقه قروه یکی از قطب‌های مهم کشاورزی در استان کردستان و ایران است که بخش‌های از آن از مشکل شوری رنج می‌برد. بنابراین تهیه نقشه شوری خاک جهت افزایش تولید محصولات و مدیریت بهتر اراضی مناطق منابع طبیعی و جلوگیری از تخریب اراضی ضروری می‌باشد. با توجه به عدم وجود نقشه‌ای مناسب در منطقه، نیاز به تهیه نقشه شوری خاک با مقیاس‌های بزرگ احساس می‌شود، که در این راستا با بهره‌گیری از نقشه‌برداری رقومی که سبب کاهش چشمگیر در مطالعات میدانی، زمان و هزینه شده، امکان مطالعه و تهیه نقشه‌های خصوصیات خاک هم‌چون شوری فراهم شده است. بنابراین هدف از این مطالعه پیش‌بینی شوری خاک با استفاده از روش‌های رگرسیون درختی و شبکه عصبی مصنوعی و داده‌های کمکی در منطقه قروه استان کردستان است.



شکل ۱- موقعیت منطقه مورد مطالعاتی و پراکنش مکانی نقاط نمونه‌برداری.

Figure 1. Location of study area and spatial variability of sample points.

نمونه برداری و آنالیزهای آزمایشگاهی: بر اساس روش هایپیرکیوب (طرحی تصادفی و طبقه بندی)، محل ۱۵۰ نمونه در منطقه قروه استان کردستان به وسعت ۳۰۰۰۰ هکتار انتخاب شده و سپس نمونه برداری از عمق ۳۰-۰ سانتی متری صورت گرفت (از متغیرهای کمکی هم چون تصویر ماهواره ای (شامل باندهای اصلی ۱ تا ۷ و شاخص های شوری (SI) پوشش گیاهی نرمال شده (NDVI))، مدل رقومی ارتفاع در این روش نمونه برداری استفاده شد). (شکل ۱). نمونه های خاک پس از هوا خشک شدن و عبور از الک ۲ میلی متری هدایت الکتریکی (۲۴) آن ها اندازه گیری شد.

داده های کمکی توپوگرافی و سنجش از دور: در نقشه برداری رقومی خاک، علاوه بر داده های خاک نیاز به داده های کمکی می باشد. در مطالعه حاضر از مدل رقومی ارتفاعی (SRTM) سازمان زمین شناسی آمریکا (قدرت تفکیک مکانی ۹۰ متر) استفاده گردید. بعد از آماده کردن مدل رقومی ارتفاع، از آن برای استخراج کردن اطلاعات کمکی استفاده گردید. پارامترهای زمین نما (۱۵ پارامتر) شامل ارتفاع، فاکتور LS<sup>۱</sup>، شاخص خیزی، شاخص همواری دره، شیب، ارتفاع، انحنا، پروفیلی، انحنا، عمق دره، موقعیت نسبی شیب، مساحت حوضه آبخیز، الگوی خمیدگی، مشخصات خمیدگی، خمیدگی کلی، بیشترین و کمترین خمیدگی در محیط سامانه جغرافیایی SAGA محاسبه و استخراج گردید (۹). همچنین در این مطالعه از تصاویر ماهواره ای لندست ۸ سنجنده ETM<sup>+</sup> مربوط به ۲۰۱۴ با ردیف ۳۶ و مسیر ۱۶۷ استفاده گردید. این تصویر از باندهای (باند ۱: ۰/۴۵-۰/۵۲ میکرومتر، باند ۲: ۰/۵۲-۰/۶۰ میکرومتر، باند ۳: ۰/۶۳-۰/۶۹ میکرومتر، باند ۴:

۰/۷۶-۰/۹۰ میکرومتر، باند ۵: ۱/۵۵-۱/۷۵ میکرومتر، باند ۷: ۲/۰۸-۲/۳۵ میکرومتر) تشکیل شده است. همچنین شاخص های همچون شاخص شوری (SI) شاخص پوشش گیاهی نرمال شده (NDVI) نیز با استفاده از نرم افزار ArcGIS از ترکیب باندها محاسبه گردید. جهت از بین بردن اطلاعات زائد از تصاویر ماهواره ای از موجک دوبعدی استفاده شد (۲۰). سپس تمام لایه های اطلاعاتی یا کمکی به فرمت رستری با اندازه سلولی ۳۰ متری تهیه شد. در نهایت تمامی اطلاعات (هدایت الکتریکی خاک و متغیرهای محیطی) به فرمت نهایی txt تبدیل شده و برای انجام فرآیند مدل سازی در نرم افزار مربوطه Matlab آماده شدند.

آنالیز حساسیت: در اغلب موارد بسیاری از ویژگی ها (لایه های اطلاعاتی)، نامربوط و زائد هستند و کارایی الگوریتم یادگیری را کاهش داده و در بعضی موارد دقت و سرعت مدل سازی را پایین می آورند. بنابراین انتخاب لایه های اطلاعاتی مؤثر در فرآیند پیش بینی دارای اهمیت می باشد. بسیاری از روش ها برای انتخاب ویژگی ها در دهه اخیر معرفی شده اند. الگوریتم های انتخاب ویژگی بسته به روند ارزیابی آن ها به دو دسته عمده تقسیم می شوند. اگر انتخاب ویژگی مستقل از هر گونه الگوریتم یادگیری انجام شود (یعنی به صورت پیش پردازنده کاملاً مجزا)، آن را روش فیلتر یا حلقه باز می گویند. در این مورد ویژگی های نامطلوب قبل از استنتاج دور ریخته می شوند. اما، اگر روند ارزیابی با یک الگوریتم طبقه بندی در ارتباط باشد، روش انتخاب ویژگی را روش رپر<sup>۲</sup> یا حلقه بسته می نامند. این روش، جستجو در فضای زیرمجموعه ها را بر اساس تخمین دقت ناشی از انتخاب یک زیرمجموعه خاص، تحت شرایط الگوریتم طبقه بندی مورد استفاده انجام می دهد.

1- LS factor

2- Wrapper

مارکوات به ترتیب برای طراحی رگرسیون درختی و شبکه عصبی مصنوعی و در نهایت پیش‌بینی شوری خاک استفاده گردید. سپس برای پهنه‌بندی شوری خاک تمام لایه‌های اطلاعاتی یا کمکی به فرمت رستری با اندازه پیکسل ۳۰ متری تهیه شدند. تمام اطلاعات (داده‌های شوری خاک و متغیرهای محیطی) به فرمت نهایی txt تبدیل شد و برای انجام فرآیند مدل‌سازی در نرم‌افزارهای مربوطه آماده شدند. در نهایت نقشه رقومی شوری خاک با استفاده از مدل بهتر تهیه شد.

**ارزیابی مدل و نقشه:** برای بررسی ارزیابی نقشه شوری خاک، اعتبارسنجی به روش حذف یک داده و پیش‌بینی آن انجام گرفت. در شرایطی که زمان و بودجه کافی برای برداشت نمونه احتمالی وجود ندارد، بهترین گزینه اعتبارسنجی به روش حذف داده است (۵). بر خلاف روش تقسیم داده‌ها به تست و آموزش، در این روش تقسیم کردن به‌طور تکرارپذیری انجام می‌شود و این روش را مؤثرتر می‌کند. پایگاه داده (n) به n-۱ موقعیت برای واسنجی و یک موقعیت برای اعتبارسنجی تقسیم می‌شود. در هر تکرار، مدل برای موقعیت حذف شده اجرا می‌گردد و متغیر حذف شده پیش‌بینی می‌گردد. سپس با مقدار واقعی مقایسه می‌گردد و خطای پیش‌بینی محاسبه می‌شود. این فرآیند برای همه موقعیت‌های نمونه‌برداری انجام می‌شود. پس از تعیین مدل مناسب، با استفاده از ضریب تبیین بین مقادیر واقعی و برآورد شده شوری خاک و همچنین با استفاده از محاسبه شاخص‌های میانگین ریشه مربعات خطا (رابطه ۱) و میانگین خطای مدل (رابطه ۲) که واحد اندازه‌گیری‌شان درصد است اعتبارسنجی شد (۵).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z_o - Z_p)^2} \quad (1)$$

الگوریتم‌های دسته دوم معمولاً نتایج بهتری به دست می‌دهند. مهم‌ترین بخش در روش انتخاب ویژگی حلقه بسته، الگوریتم جستجویی است که در آن به‌کار رفته است. در این مقاله از روش رپر با الگوریتم جستجوی تکاملی از نوع الگوریتم ژنتیک استفاده گردید. این روش به جعبه سیاه معروف است. دلیل استفاده از الگوریتم ژنتیک این است که این الگوریتم می‌تواند یک جستجوی تصادفی را انجام دهد و مستعد گیر افتادن در کمینه محلی نمی‌باشد (۲۱).

**مدل‌سازی مکانی:** جهت برقراری ارتباط بین متغیرهای کمکی و شوری از مدل رگرسیون درختی و شبکه عصبی مصنوعی استفاده شد. مدل درختی طبقه‌بندی و رگرسیون، یک روش ناپارامتری الگوریتمی است که قادر به پیش‌بینی متغیرهای کمی یا متغیرهای طبقه‌بندی‌شده بر اساس مجموعه‌ای از متغیرهای پیش‌بینی‌کننده کمی و کیفی است. در این روش، مجموعه‌ای از شرط‌های منطقی<sup>۱</sup> به‌صورت یک الگوریتم با ساختار درختی برای طبقه‌بندی یا پیش‌بینی کمی یک متغیر به‌کار می‌رود. ایجاد درخت تصمیم شامل دو مرحله است. مرحله اول ایجاد و رشد درخت است. این مرحله شامل پیوند<sup>۲</sup> و انشعاب<sup>۳</sup> می‌باشد. مرحله دوم، مرحله توقف و هرس<sup>۴</sup> است. هدف از این مرحله به حداقل رساندن خطای پیش‌بینی است (۲۷). همچنین شبکه عصبی مصنوعی پرسپترون چندلایه با یک لایه مخفی که دارای تابع فعال‌سازی سیگموئید در لایه مخفی و تابع فعال‌سازی خطی در لایه خروجی بوده و تعداد نرون‌های آن از دو تا ۱۰ نرون متغیر بوده استفاده شد و بهترین تعداد نرون به‌صورت سعی و خطا تعیین شد. در پژوهش حاضر از نرم‌افزار Matlab و الگوریتم CARDA و لونبرگ

- 1- If-then
- 2- Merging
- 3- Splitting
- 4- Pruning

## نتایج و بحث

در جدول ۱ خلاصه خصوصیات آماری شوری خاک در عمق ۰-۳۰ سانتی متری آمده است. میانگین شوری خاک در منطقه نشان‌دهنده شوری کم تا زیاد منطقه مورد بررسی است. بیشترین مقدار آن نیز در مرکز، جنوب و جنوب غربی منطقه مطالعاتی (عمدتاً شوره‌زارهای بدون پوشش) و کمترین مقدار در شمال غربی، جنوب شرقی، شمال شرقی و شمال منطقه مطالعاتی (عمدتاً بلندی‌های با پوشش مرتعی) مشاهده شد.

$$ME = \frac{1}{n} \sum_{i=1}^n (Z_0 - Z_p)^2 \quad (2)$$

$$R^2 = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \quad (3)$$

که در آن‌ها،  $Z_0$  مقادیر پیش‌بینی شده،  $Z_p$  مقادیر مشاهداتی،  $Z_{ave}$  متوسط مقادیر مشاهداتی،  $n$  تعداد داده‌ها،  $x$  و  $y$  به ترتیب مقادیر مطلق مشاهداتی و پیش‌بینی شده،  $\bar{x}$  و  $\bar{y}$  به ترتیب مقدار متوسط مقادیر مشاهداتی و پیش‌بینی شده است.

جدول ۱- پارامترهای آماری مربوط به هدایت الکتریکی خاک و داده‌های کمکی.

Table 1. Statistic parameters of soil electrical conductivity and auxiliary data.

کشیدگی Kurtosis	چولگی Skewness	میانگین Average	حداکثر Maximum	حداقل Minimum	خصوصیت Characteristic
18.48	3.98	0.81	6.93	0.23	هدایت الکتریکی (دسی‌زیمنس بر متر) (Electrical conductivity) ( $\text{dSm}^{-1}$ )
37.146	-5.549	-0.0861	-0.03	-0.50	شاخص شوری (Salinity index)
-0.887	0.163	3.701	7.29	-0.03	شاخص همواری دره (MrVBF index)
-0.971	-0.219	9.022	12.40	5.08	شاخص خیسبی (Wetness index)
30.943	5.049	0.0830	0.39	0.04	شاخص NDVI (NDVI)
0.151	-0.06	10845.46	13783.00	7997.00	باند ۳ (Band 3)
1.789	-0.795	13816.23	18294.00	8107.00	باند ۷ (Band 7)

شود که نتایج ریشه مربعات خطا، میانگین خطا و ضریب تبیین برای شوری خاک در جدول ۲ آمده است. نتایج نشان داد که در مدل شبکه عصبی مصنوعی تعداد نرون ۶ و تکرار ۱۰۰۰ دارای بهترین ترکیب برای پیش‌بینی شوری خاک می‌باشد. نتایج سه

مدل‌سازی مکانی: بعد از به‌دست آوردن قوانین مرتبط در مرحله آموزش شبکه عصبی مصنوعی و رگرسیون درختی که در بالا توضیح داده شد، شبکه آموزش‌دیده با روش اعتبارسنجی آزمایش کرده تا دقت این مدل‌ها برای پیش‌بینی شوری خاک مشخص

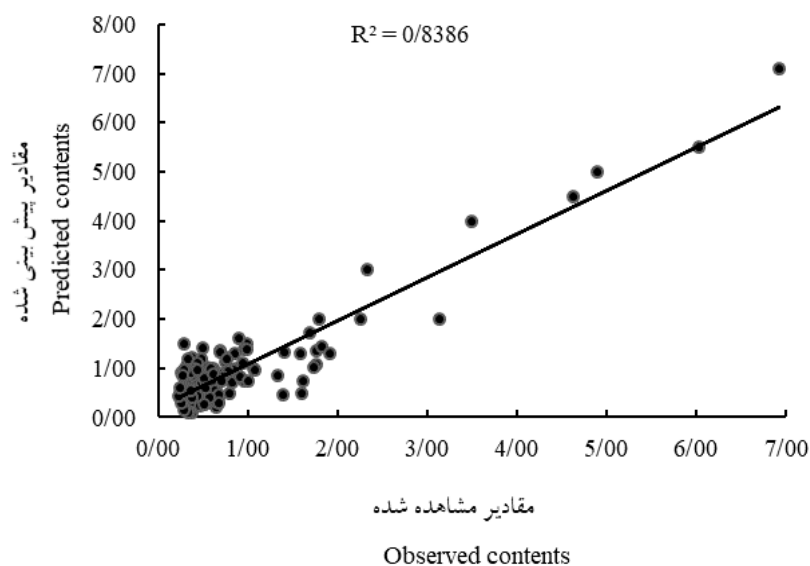
پیش‌بینی شده برای مدل شبکه عصبی مصنوعی ارائه شده است. با توجه به این شکل، ملاحظه می‌شود بهترین خط برازش شده دارای زاویه‌ای نزدیک به ۴۵ درجه بوده که نشان‌دهنده دقت بالای برآورد توسط مدل شبکه عصبی مصنوعی است. در نهایت نقشه شوری خاک با استفاده مدل شبکه عصبی مصنوعی در محیط ArcGIS تهیه شد که دامنه تغییرات آن بین ۶/۵-۰ دسی‌زیمنس بر متر می‌باشد (شکل ۳).

عدد ریشه مربعات خطا، میانگین خطا و ضریب تبیین برای مدل شبکه عصبی مصنوعی جهت پیش‌بینی شوری خاک به ترتیب ۰/۱۹۰، ۰/۰۳۶ و ۰/۸۳ و برای مدل رگرسیون درختی به ترتیب ۰/۲۰۵، ۰/۰۴۲ و ۰/۶۳ به دست آمد (جدول ۲). مطابق با این نتایج مدل شبکه عصبی مصنوعی دارای دقت بیشتری نسبت به رگرسیون درختی به منظور پیش‌بینی شوری خاک می‌باشد. در شکل ۲ نمودار پراکنش داده‌های

جدول ۲- نتایج مدل‌ها برای پیش‌بینی شوری خاک.

Table 2. Results of models to estimate soil salinity.

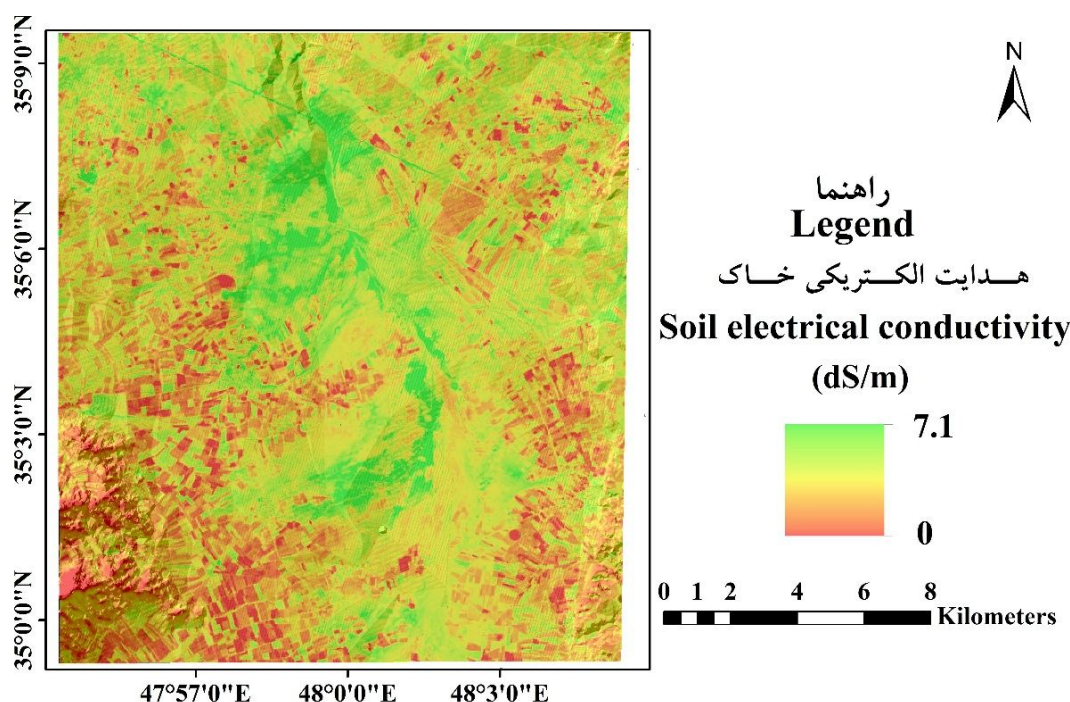
میانگین خطا Mean error	ضریب همبستگی Correlation coefficient	ریشه مربعات خطا Root mean square error	مدل Model
0.042	0.63	0.205	رگرسیون درختی (Regression tree)
0.036	0.83	0.190	شبکه عصبی مصنوعی (Artificial neural network)



شکل ۲- نتایج مدل شبکه عصبی مصنوعی برای تخمین شوری خاک.

Figure 2. Results of artificial neural network model to estimate soil salinity.



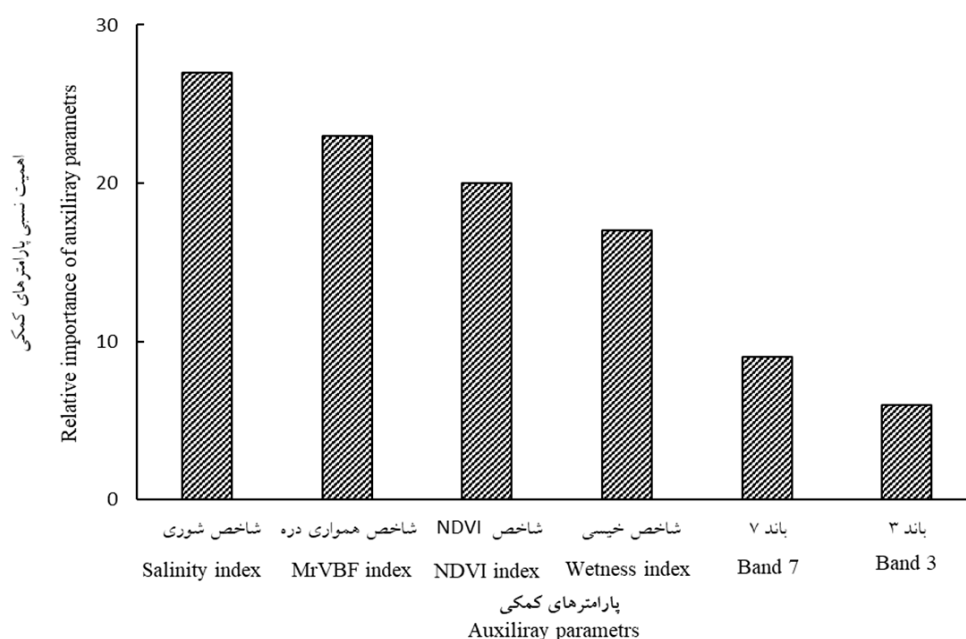


شکل ۳- نقشه شوری خاک منطقه مورد مطالعه با استفاده از مدل شبکه عصبی مصنوعی.

Figure 3. Soil salinity map of the study area using artificial neural network.

ویژگی‌های مهم و تأثیرگذار بر ویژگی‌های خاک آن منطقه از جمله شوری خاک می‌باشد. نتایج پژوهش‌های زیادی بیانگر رابطه پارامترهای سرزمین و تصاویر ماهواره‌ای با ویژگی‌های خاک بوده است. تقی‌زاده و همکاران (۲۰۱۴) جهت نقشه‌برداری رقومی شوری خاک در منطقه اردکان یزد از متغیرهای محیطی شامل اجزای سرزمین، داده‌های تصویر  $ETM^+$  و نقشه ژئومورفولوژی استفاده کردند و نتایج آن‌ها نشان داد مهم‌ترین متغیرهای کمکی، نقشه ژئومورفولوژی، شاخص خیسی و شاخص همواری دره با درجه تفکیک بالا بودند (۲۸).

آنالیز حساسیت: آنالیز حساسیت مدل رپر نشان داد که متغیرهای کمکی شاخص شوری (۲۷)، شاخص خیسی (۱۷ درصد)، شاخص همواری دره (۲۳ درصد)، باند ۷ (۹ درصد)، شاخص NDVI (۲۰ درصد) و باند ۲ (۶ درصد) بالاترین تأثیر را بر روی پیش‌بینی شوری خاک در منطقه مورد مطالعه دارند (شکل ۵). این موضوع نشان‌دهنده آن است که در منطقه مورد مطالعه، پستی و بلندی و پوشش گیاهی از مهم‌ترین فاکتورهای خاک‌سازی بوده و در توزیع مکانی خاک و ویژگی‌های آن مؤثر می‌باشد، چرا که توپوگرافی و پوشش گیاهی هر منطقه یکی از

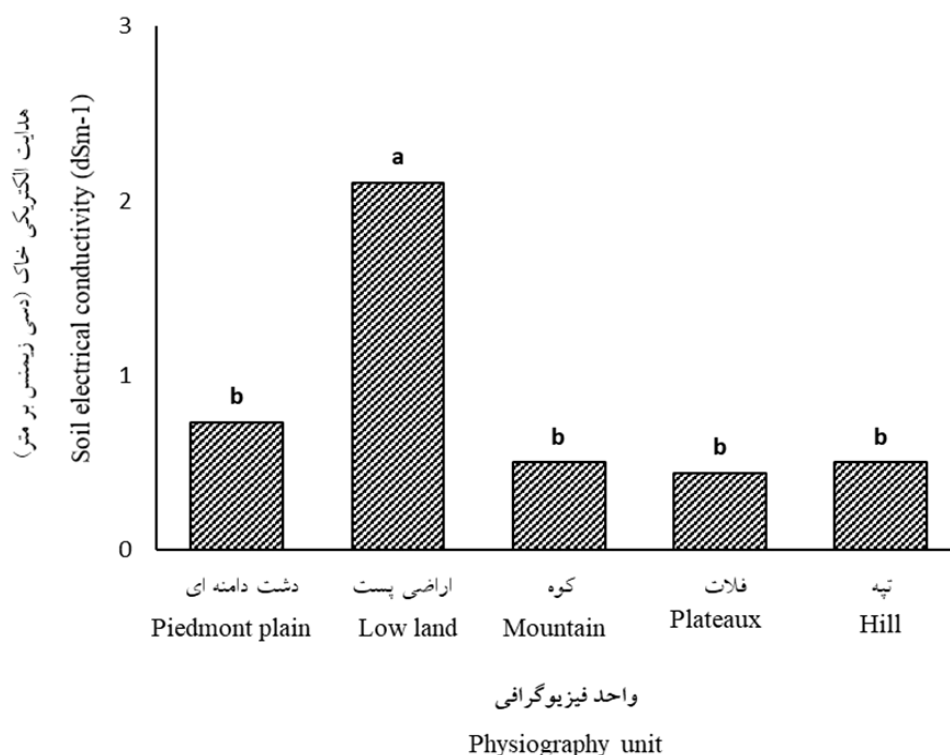


شکل ۴- نتایج اهمیت متغیرها در پیش‌بینی شوری خاک بر اساس آنالیز حساسیت (بر اساس درصد).

Figure 4. Results of variables importance to estimate soil salinity based on analysis test (based on percentage).

بود. به دلیل آن‌که اراضی پست در منطقه عمدتاً در قسمت‌های مرکزی با شیب کم و توپوگرافی پست مقعر واقع شده‌اند بنابراین املاح مربوط به قسمت‌های مرتفع منطقه شستشو پیدا کرده و در این قسمت‌ها تجمع پیدا کرده و منجر به شوری بیش‌تر این واحد فیزیوگرافی نسبت به سایر واحدهای فیزیوگرافی شده است. واحدهای فیزیوگرافی کوه، تپه و فلات‌ها عمدتاً در اراضی مرتفع با شیب زیاد واقع شده‌اند، بنابراین شوری کمی دارند و واحد فیزیوگرافی دشت دامنه‌ای دارای شوری کم می‌باشند که در قسمت‌های کم شیب با زهکشی مناسب‌تر نسبت به اراضی پست قرار گرفته‌اند.

آنالیز مکانی شوری خاک و مهم‌ترین داده‌های کمکی: مطابق با شکل ۳ بیش‌ترین مقدار شوری در قسمت‌های مرکزی منطقه مشاهده می‌شود که به دلیل داشتن شیب کم و توپوگرافی مقعر، تجمع املاح منتقل شده از قسمت‌های مرتفع اطراف منطقه در این قسمت صورت گرفته است. این نواحی با شوری زیاد دارای واحد فیزیوگرافی اراضی پست می‌باشند که از مشخصه‌های این واحد فیزیوگرافی وجود خاک‌های شور می‌باشد. نتایج تست آنالیز واریانس یک‌طرفه نشان داد (شکل ۵) که پنج واحد فیزیوگرافی دارای اختلاف آماری معنی‌داری بودند و نتایج آزمون دانکن نشان داد که میزان شوری خاک در واحد فیزیوگرافی اراضی پست بیش‌تر از سایر واحدهای فیزیوگرافی



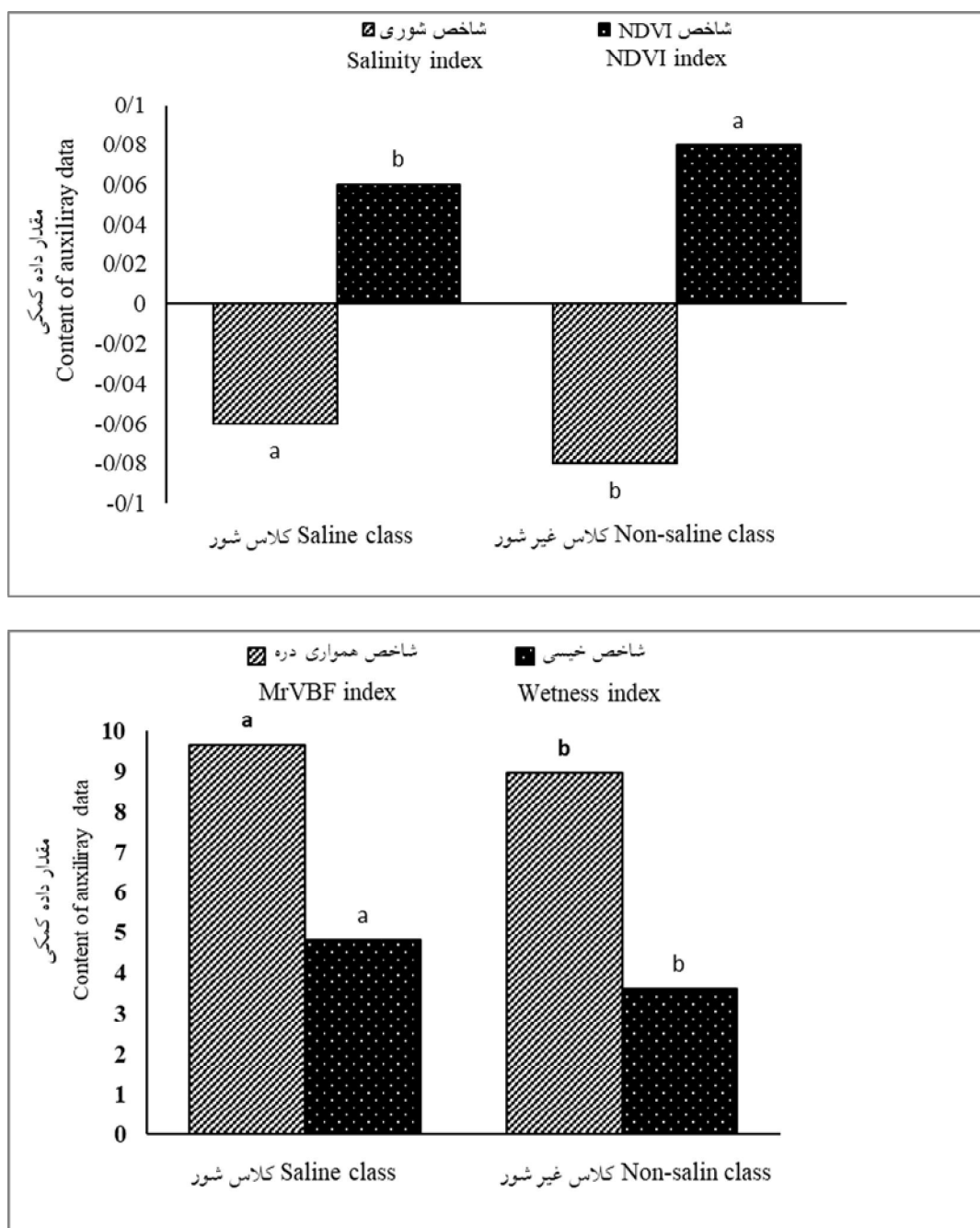
شکل ۵- میانگین شوری خاک در واحدهای فیزیوگرافی منطقه مورد مطالعه (میانگین‌های که حروف مشترک دارند مطابق با آزمون دانکن در سطح ۱ درصد اختلاف معنی‌داری با هم ندارند (مقدار  $P > 0.05$ )).

**Figure 5.** Mean value of soil salinity in physiography units of the study area using. Means ( $n = 5$ ) that share a letter are not significantly different at the level 0.01 ( $P > 0.05$ ) according to Duncan's test.

کلاس شور می‌باشد و از لحاظ آماری اختلاف معنی‌داری دارند.

در منطقه مورد مطالعه مقدار داده‌های کمکی باند ۳ و باند ۷ دارای اختلاف معنی‌داری از لحاظ کلاس‌های شور و غیرشور نمی‌باشند. اگرچه مقادیر باندهای ۳ و ۷ از لحاظ کلاس‌های شور و غیرشور اختلاف معنی‌داری نداشتند، اما میانگین مقادیر این باندها در کلاس شور بیش‌تر از کلاس غیرشور می‌باشد که نشان‌دهنده بازتاب بیش‌تر در کلاس شور می‌باشد، که می‌تواند به دلیل پوشش گیاهی کم‌تر و متعاقباً بازتاب بیش‌تر در باند ۳ و بیش‌تر بودن میزان کانی‌ها و املاح شور و متعاقباً بازتاب بیش‌تر در باند ۷ در مناطق شور باشد.

در مناطق با شوری بیش‌تر شاخص NDVI دارای کم‌ترین مقدار و شاخص شوری، شاخص خیسی و شاخص MrVBF، باند ۳ و باند ۷ دارای بیش‌ترین مقادیر می‌باشند که نشان‌دهنده تطابق این پارامترها با شاخص شوری خاک می‌باشد. شکل ۶ میانگین مهم‌ترین داده‌های کمکی در پیش‌بینی شوری خاک را در دو کلاس شور (هدایت الکتریکی بیش از ۲ دسی‌زیمنس بر متر) و غیرشور (هدایت الکتریکی کم‌تر از ۲ دسی‌زیمنس بر متر) در منطقه مورد مطالعه نشان می‌دهد. میانگین مقادیر شاخص شوری، شاخص MrVBF و شاخص خیسی در کلاس غیرشور کم‌تر از کلاس شور و در مقابل میانگین مقادیر شاخص NDVI در کلاس غیرشور بیش‌تر از



شکل ۶- میانگین مهم‌ترین داده‌های کمکی در کلاس شور و غیرشور منطقه مورد مطالعه (میانگین‌های که حروف مشترک ندارند مطابق با آزمون تی در سطح ۱ درصد اختلاف معنی‌داری باهم دارند (مقدار  $P < 0.05$ )).

Figure 6. Mean value of the most auxiliary data in saline and non-saline classes of the study area using. Means ( $n = 2$ ) that share a letter are not significantly different at the level 0.01 ( $P > 0.05$ ) according to T's test.

## نتیجه گیری کلی

بیشترین میزان شوری خاک در قسمت‌های مرکزی منطقه با واحد فیزیوگرافی اراضی پست و پوشش گیاهی ضعیف می‌باشد. مهم‌ترین متغیر کمکی در پیش‌بینی شوری خاک در منطقه شاخص شوری می‌باشد و وجود ارتباط قوی بین داده خاک و داده‌های کمکی می‌تواند بر روی دقت مدل اثرگذار باشد. به‌طورکلی نتایج نشان داد که تکنیک‌های

پدومتری می‌تواند در گستره‌ای وسیع جهت نقشه‌برداری رقومی خصوصیات خاک‌ها به‌کار گرفته شود. پیشنهاد می‌گردد که جهت تهیه نقشه خصوصیات خاک از مدل شبکه عصبی مصنوعی، و داده‌های کمکی هم‌چون اجزاء سرزمین و تصاویر ماهواره‌ای در مطالعات آینده استفاده شود.

## منابع

1. Abdinam, A. 2004. An investigation on preparing of the soil salinity map using correlation method between imagery and soil salinity data in the Qazvin plain. *J. Pashouhesh and Sazandegi*. 64: 33-38. (In Persian)
2. Adhikari, K., Minasny, B., Greve, B.G., and Greve, M.H. 2014. Constructing a soil class map of Denmark based on the FAO legend using digital techniques. *Geoderma*. 214-215: 101-113.
3. Azhirabi, R., Kamkar, B., and Abdi, O. 2015. Comparison of different indices adopted from Landsat images to map soil salinity in the army field of Gorgan. *J. Soil Manage. Sust. Prod*. 5: 1. 173-176. (In Persian)
4. Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. 1984. *Classification and Regression Trees*. Chapman & Hall, New York, 355p.
5. Brus, D.J., Kempen, B., and Heuvelink, G.B.M. 2011. Sampling for validation of digital soil maps. *Eur. J. Soil Sci*. 62: 394-407.
6. Dai, P.F., Qigang, Z., Zhiqiang, L.V., Xuemei, W., and Gangcai, W.L. 2014. Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. *Ecol. Ind*. 45: 184-194.
7. Farifte, J., Farshad, A., and George, R.J. 2005. Assessing salt - affected soils using remote sensing, solute modeling, and geophysics. *Geoderma*. 130: 191-206.
8. Grinand, C., Arrouays, D., Laroche, B., and Martin, M.P. 2008. Extrapolating regional soil landscapes from an existing soil map: sampling intensity, validation procedures, and integration of spatial context. *Geoderma*. 143: 180-190.
9. Hengel, T., Rossiter, D.G., and Stein, A. 2003. Soil sampling strategies for spatial prediction by correlation with auxiliary maps. *Geoderma*. 120: 75-93.
10. Hengl, T., Toomanian, N., Reuter, H., and Malakouti, M.J. 2007. Methods to interpolate soil categorical variables from profile observations: Lessons from Iran. *Geoderma*. 140: 417-427.
11. Heung, B., Bulmer, C.E., and Schmidt, M.G. 2014. Predictive soil parent material mapping at a regional-scale: a random forest approach. *Geoderma*. 214-215: 141-154.
12. Jafari, A., Khademi, H., Finke, P., Wauw, J.V.D., and Ayoubi, S. 2014. Spatial prediction of soil great groups by boosted regression trees using a limited point dataset in an arid region, southeastern Iran. *Geoderma*. 232-234: 148-163.
13. Kempen, B., Brus, D.J., Heuvelink, G.B.M., and Stoorvogel, J.J. 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma*. 151: 311-326.
14. Kheir, R.B., Greve M.H., Abdallah, C., and Dalgaard, T. 2010. Spatial soil zinc content distribution from terrain parameters: A GIS-based decision-tree model in Lebanon. *Environ. Pollut*. 158: 520-528.

15. Marcel, G.S., Feike, J.L., Martinus, T., and Van Genuchten, H. 1998. Neural Network Analysis for Hierarchical Prediction of Soil Hydraulic Properties. *Soil Sci Soc. Am. J.* 62: 847-855.
16. McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., and Shatar, T.M. 2000. An overview of pedometric techniques for use in soil survey. *Geoderma.* 97: 293-327.
17. McBratney, A.B., Santos, M.L.M., and Minasny, B. 2003. On digital soil mapping. *Geoderma.* 117: 3-52.
18. Metternicht, G., and Zinck, J.A. 2004. Remote sensing of soil salinity: Potentials and constraints. *Remote Sens Environ.* 64: 33-38.
19. Minasny, B., and McBratney, A. 2002. The method for fitting neural network parametric pedotransfer functions. *Soil Sci. Soc. Am. J.* 66: 2. 352-361.
20. Minasny, B., and McBratney, A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput. Geosci.* 32: 1378-1388.
21. Nosrati, H., and Eftekhari, M. 2014. A new approach for variable selection using fuzzy logic. *Computational Intelligence in Electrical Engineering.* 4: 71-83. (In Persian)
22. Pahlavan-Rad, M.R., Toomanian, N., Khormali, F., Brungard, C.W., Komaki, C.B., and Bogaert, P. 2014. Updating soil survey maps using random forest and conditioned Latin hypercube sampling in the loess derived soils of northern Iran. *Geoderma.* 232-234: 97-106.
23. Piccini, C., Marchetti, A., and Francaviglia, R. 2014. Estimation of soil organic matter by geostatistical methods: use of auxiliary information in agricultural and environmental assessment. *Ecol. Ind.* 36: 301-314.
24. Sparks, D.L., Page, A.L., Helmke, P.A., Leppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, G.T., and Summer, M.E. 1996. *Methods of Soil Analysis.* Soil. Sci. Soc. Am. J. Madison, Wisconsin.
25. Tajgardan, T., Aubi, Sh., Shatai, Sh., and Khormali, F. 2009. Mapping soil surface salinity using remote sensing data of ETM<sup>+</sup>. (Case study: North of Agh Ghala, Golestan province). *J. Soil Water Cons.* 88: 9-15. (In Persian)
26. Taghizadeh-Mehrjardi, R., Nabiollahi, K., Minasny, B., and Triantafyllis, J. 2015. Comparing data mining classifiers to predict spatial distribution of USDA-family soil groups in Baneh region, Iran. *Geoderma.* 253-254: 67-77.
27. Taghizadeh-Mehrjardi, R. 2016. *Modern concepts in Soil Science (PEDOMETRICS).* Ardakan Univ. Press, 311p.
28. Taghizadeh-Mehrjardi, R., Minasny, B., Sarmadian, F., and Malone, B.P. 2014. Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma.* 213: 15-28.
29. Taghizadeh-Mehrjardi, R., Nabiollahi, K., and Kerry, R. 2016. Digital mapping of soil organic carbon at multiple depths using different data mining techniques in Baneh region, Iran. *Geoderma.* 253-254: 67-77.
30. Taghizadeh-Mehrjardi, R., Sarmadian, F., Savaghebi, G.H., Omid, M., Tomanian, N., Rosta, M.J., and Rahimian, M.H. 2013. The comparison of efficiency of neuro-fuzzy, genetic algorithm, neural network and multivariate regression models to prediction of soil salinity (Case study: Ardakan). *J. Natur. Resour.* 66: 2. 207-222. (In Persian)
31. Vasques, G.M., Dematte, J.A.M., Viscarra Rossel, R.A., Ramirez-Lopez, L., and Terra, F.S. 2014. Soil classification using visible/near-infrared diffuse reflectance spectra from multiple depths. *Geoderma.* 223-225: 73-78.
32. Veronesi, F., Corstanje, R., and Mayr, T. 2014. Landscape scale estimation of soil carbon stock using 3D modeling. *Sci. Total Environ.* 487: 578-586.



---

## **Prediction of soil salinity using tree regression and artificial neural network in Ghorveh soils, Kurdistan province**

**Sh. Moradian<sup>1</sup>, \*K. Nabiollahi<sup>2</sup> and R. Taghizadeh-Mehrjardi<sup>3</sup>**

<sup>1</sup>M.Sc. Graduate, Dept. of Soil Science and Engineering, University of Kurdistan,

<sup>2</sup>Assistant Prof., Dept. of Soil Science and Engineering, University of Kurdistan,

<sup>3</sup>Assistant Prof., Dept. of Arid and Desert Regions Management, University of Ardakan

Received: 04/30/2017; Accepted: 09/06/2017

---

### **Abstract**

**Background and Objectives:** Soil salinity is one of the major problems in arid and semi-arid area. In this condition, soluble salts accumulate in the soil surface and reduce yield and soil fertility. Soils survey and mapping can help to improve these soils. The investigation of variability of soil salinity using traditional methods is expensive and time consuming. Therefore, one of the ways to solve this challenge is using digital soil mapping that soil characteristics were mapped using auxiliary data. The aim of this research is using tree regression (TR) and artificial neural network (ANN) models and auxiliary data to prepare soil salinity map.

**Materials and Methods:** Using Hypercube soil sampling method, 150 soil samples in depths 0-30 cm of Ghorveh soils, Kurdistan Province (covers 30000 ha) were taken and soil electrical conductivity was measured. Auxiliary data in this study were terrain attributes and Landsat 8 ETM<sup>+</sup> data. Terrain parameters (include 15 parameters) and salinity index (SI) and normalized difference vegetative index (NDVI) were computed and extracted using SAGA and ArcGIS software, respectively. To make a relationship between soil salinity and auxiliary data, TR and ANN models were applied and were validated using cross validation method. Finally, soil salinity map were made using better model.

**Results:** To predict soil salinity, auxiliary variables include salinity index, wetness index, index of valley bottom flatness, NDVI index, Band 3 and Band 7 were the most important. The results of the study showed that ANN model (0.83, 0.036 and 0.190, respectively for determination of coefficient, mean error and root mean square root) has more accuracy compared to TR model to predict soil salinity. Soil salinity content ranged between 0.23 to 6.93 dSm<sup>-1</sup> and the highest content of soil salinity located in central regions (lowland and bare land). In these central regions, auxiliary data include salinity index, index of valley bottom flatness, wetness index, band 7 and band 3 had the highest values and NDVI index had the lowest values.

**Conclusion:** Salinity index is the most important auxiliary data to predict soil salinity of the study area. Strong link between soil data and auxiliary data can impact on the accuracy of the model. In general, the results showed that pedometrics techniques in a wide range can be used for digital mapping of soil properties. It is suggested ANN model and auxiliary data such as terrain attributes and satellite images were applied to prepare map of soil properties in future studies.

**Keywords:** Satellite image, Digital elevation model, Pedometrics

---

\* Corresponding Author; Email: [k.nabiollahi@uok.ac.ir](mailto:k.nabiollahi@uok.ac.ir)

