



محور مقاله: پدومتری و ارزیابی خاکها

پیش‌بینی پراکنش مکانی گروه‌های مرجع خاک در یک منطقه نیمه‌خشک به کمک مدل رگرسیون لاجستیک

حسنا پایدار^۱، پرویز شکاری^{۲*}^۱ دانش‌آموخته کارشناسی ارشد گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه رازی^۲ استادیار گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه رازی

چکیده

در دو دهه گذشته توسعه و کاربرد تکنیک‌های کمی به‌ویژه الگوریتم‌های هوش مصنوعی برای نقشه‌های خاک بطور چشمگیری افزایش یافته است. در این پژوهش از روش رایان‌آموختی نسبتاً پرکاربرد رگرسیون لاجستیک استفاده شد. بخشی از منطقه سنجابی استان کرمانشاه با مساحت بیش از ۴۸۰۰۰ هکتار به عنوان منطقه‌ی مطالعاتی برگزیده شد. تعداد ۱۰۰ خاکرخ در منطقه مطالعاتی با روش نمونه‌برداری GRTS مشخص شد. تمامی خاکرخ‌های حفارشده، براساس راهنمای تشریح و نمونه‌برداری خاک‌ها در صحرا و راهنمای مطالعات خاک‌شناسی تشریح و از تمامی افق‌های ژنتیک خاکرخ‌ها نمونه تهیه گردید. براساس ویژگی‌های ریخت‌شناختی و داده‌های آزمایشگاهی خاکرخ‌ها مطابق با سیستم مرجع جهانی (WRB) در سطح گروه مرجع خاک رده‌بندی شدند. متغیرهای محیطی مشتق شده از مدل رقومی ارتفاع و یک تصویر ماهواره لندست ۸ به عنوان متغیرهای محیطی برای پیش‌بینی کلاس‌های خاک به کار رفت. مدل رگرسیون لاجستیک توانایی قابل توجهی برای پیش‌بینی گروه‌های مرجع خاک در شرایط منطقه مطالعاتی داشت به گونه‌ای که مقدار صحت عمومی و شاخص کاپای مدل به ترتیب ۰/۵۸ و ۰/۴۶ بدست آمد. در منطقه‌ی مطالعاتی، Calcisols در مقایسه با دیگر گروه‌های مرجع گسترش بیشتری داشت. نتایج نشان داد که پیش‌بینی‌های این روش نیز مانند دیگر روش‌های رایان‌آموختی به فراوانی مشاهده‌ها و نوع و تعداد متغیرهای محیطی حساس است.

کلمات کلیدی: نقشه‌برداری رقومی، متغیرهای محیطی، رایان‌آموختی، منطقه سنجابی

مقدمه

یکی از نیازهای اساسی برای برنامه‌ریزی منابع خاک تهیه نقشه‌های دقیق خاک است. نقشه‌های توزیع مکانی کلاس‌های خاک در تصمیم‌گیری‌های مدیریتی برای برنامه‌ریزی استفاده از زمین و دیگر فعالیت‌های مرتبط با جنگلداری، کشاورزی و حفاظت محیط‌زیست ضروری است و اطلاعات زیادی در اختیار کاربران قرار می‌دهند. در گذشته تهیه نقشه‌های خاک و بررسی ویژگی آن بر روش‌های کلاسیک استوار بود که در آن تغییرات ویژگی‌های خاک درون واحدهای نقشه نادیده گرفته می‌شد (Rossiter, 2005). با استفاده از روش‌های کلاسیک، نقشه‌های چند وجهی خاک به دست می‌آید. پاسخگویی این نقشه‌ها به کاربردهای نوین مورد انتقاد قرار گرفته است و این به دلیل منطق ناپیوسته‌ای است که اساس اینگونه مطالعات را می‌سازد. این در حالی است که معمولاً تغییر ویژگی‌ها و کلاس‌های خاک در یک زمین‌چهر پیوسته بوده و تبدیل آن به گروه‌ها و کلاس‌های طبقه‌ای باعث از دست رفتن مقدار قابل توجهی از اطلاعات می‌شود. از طرفی استفاده از روش‌های مرسوم نقشه‌برداری خاک برای تهیه نقشه در مناطقی با وسعت زیاد و با مقیاس مناسب، به دلیل هزینه‌بر و وقت‌گیر بودن بسیار مشکل است. این در حالی است که امروزه نیاز به اطلاعات مکانی تفصیلی در مورد خاک و نقشه‌های دقیق بیشتر از گذشته احساس می‌شود. بنابراین با ایجاد پایگاه داده‌های مکانی، یافتن روش‌های نوین نقشه برداری که با وجود صرفه جویی در زمان و هزینه از صحت بالایی برخوردار باشند، ضروری به نظر می‌رسد. نقشه برداری رقومی خاک از طریق استفاده از مشاهدات صحرایی یا استخراج داده از نقشه‌های سنتی و ایجاد روابط کمی با فرآیندهای محیطی مؤثر بر تشکیل خاک‌ها مدیریت پایدار منابع را پشتیبانی می‌کند. از طرفی با نیاز به نقاط مشاهداتی کمتر، هزینه‌های کمتری نیز به دنبال دارد (McBratney و همکاران ۲۰۰۳). اطلاعات مکانی خاک به معنی نقشه‌ها و پایگاه‌های داده‌ای است که اطلاعات روشن و کمی از تغییرات ویژگی‌های خاک برای یک منطقه مشخص را ارائه می‌دهد. از جمله روش‌های رایان‌آموختی که در علوم خاک به‌صورت گسترده استفاده شده‌اند می‌توان روش‌های؛ درخت تصمیم‌گیری^۱ (DT)، شبکه عصبی مصنوعی^۲ (ANN)، جنگل تصادفی^۳ (RF)، درخت

* ایمیل نویسنده مسئول: pshekaari@gmail.com

^۱ Decision trees^۲ Artificial neural networks^۳ Random forests



تصمیم‌گیری و رگرسیون^۱ (CART)، ماشین بردار پشتیبان^۲ (SVM)، درختان رده‌بندی^۳ (CT) و کای نزدیک‌ترین همسایه^۴ (KNN) را نام برد. خلاصه‌ای از رویکرد نقشه برداری رقومی خاک توسط McBratney و همکاران (۲۰۰۳) نشان می‌دهد که مدل‌های خطی (مانند رگرسیون خطی چندگانه و مدل‌های خطی تعمیم یافته) به طور متداول برای نقشه‌برداری رقومی ویژگی‌های خاک و ترکیب با روش کریجینگ در رگرسیون کریجینگ به کار رفته است (Hengl و همکاران ۲۰۰۷). همچنین، برای اهداف طبقه بندی، رویکرد خطی که اغلب مورد استفاده قرار می‌گیرد، رگرسیون لاجستیک چند جمله‌ای است (Collard و همکاران ۲۰۱۴؛ Jafari و همکاران ۲۰۱۲). مدل‌های رگرسیون لاجستیک نوعی از مدل‌های خطی تعمیم یافته هستند که برای متغیرهای هدف رتبه‌ای (طبقه‌ای) مناسب می‌باشند. این مدل‌ها قادر به توصیف رابطه بین متغیرهای مستقل و یک متغیر هدف دو بخشی (دارای دو مقدار ۰ و ۱) هستند. در مورد رگرسیون لاجستیک دو جمله‌ای، خروجی رگرسیون لاجستیک به صورت احتمالی بیان می‌شود، به گونه‌ای که مقادیر نزدیک به صفر نشان دهنده کم بودن احتمال رخداد و مقادیر نزدیک به یک به معنی بالا بودن آن است (Kleinbaum و همکاران ۲۰۰۸). مطالعات نقشه برداری رقومی در ایران تاریخچه درازی ندارد و مطالعات در این زمینه مربوط به سال‌های اخیر است. به‌عنوان مثال می‌توان به نقشه‌برداری رقومی کلاس‌های خاک با استفاده از سیستم رده‌بندی (FAO, 2014) با کاربست مدل‌های جنگل تصادفی (بهمنی، ۱۳۹۵)، ماشین بردار پشتیبان (ویسی‌سرطایره، ۱۳۹۵) در دشت میان‌در بند استان کرمانشاه و روش K-نزدیکترین همسایه (کیانیان، ۱۳۹۷) و شبکه عصبی مصنوعی (ویسی، ۱۳۹۷) در منطقه سنجابی استان کرمانشاه اشاره نمود.

وجود نقشه‌های خاک از نیازهای اساسی در مطالعات علوم زمین می‌باشد. با وجودی که بیش از پنجاه سال از شروع مطالعات خاکشناسی در ایران می‌گذرد، نقشه‌های خاک کشور به طور کامل تهیه نشده‌اند. در استان کرمانشاه اطلاعات کافی در مورد خاک‌های مناطق مختلف و در نتیجه پتانسیل خاک‌ها در دست نیست. به منظور بهره‌برداری بایسته از سرزمین در استان کرمانشاه و به خصوص دشت سنجابی که جایگاه مهمی در تولید محصولات کشاورزی دارد، به نقشه‌های کارآمد خاک نیاز داریم. بنابراین هدف از این پژوهش تهیه نقشه پیش‌بینانه الگوی پراکنش کلاس‌های تاکسونومیک خاک بر اساس سیستم WRB در سطح گروه مرجع خاک در پهنه دشت سنجابی با کاربست مدل رگرسیون لاجستیک بود.

مواد و روش‌ها

منطقه مورد مطالعه در حد فاصل طول جغرافیایی ۴۶ درجه و ۴۴ دقیقه تا ۴۶ درجه و ۵۹ دقیقه شرقی و عرض جغرافیایی ۳۴ درجه و ۲۰ دقیقه تا ۳۴ درجه و ۳۱ دقیقه شمالی، با مساحت تقریبی ۴۸ هزار هکتار و ارتفاع متوسط ۱۳۳۰ متر از سطح دریا واقع در منطقه سنجابی در استان کرمانشاه است (شکل ۱). رژیم رطوبتی و دمایی خاک منطقه مطالعاتی به ترتیب زیریک و ترمیک بود. تپه، دشت دامنه‌ای و دشت آبرفتی رودخانه‌ای واحدهای فیزیوگرافی غالب در منطقه مطالعاتی بود. همچنین در بعضی مناطق، کوه و آبرفت‌های بادبزی سنگریزه‌دار دیده می‌شد. منطقه مطالعاتی قسمتی از زون خرد شده و گسله زاگرس است که از نظرساخت کلی زمین‌شناسی جزئی از مجموعه زاگرس است. به منظور تعیین موقعیت و تعداد خاکرخ‌ها، با توجه به ویژگی‌های ریخت‌شناسی منطقه و استفاده از نقشه توپوگرافی در محیط نرم افزار R (R Core Team, 2014) و براساس روش GRTS^۵ موقعیت ۱۰۰ خاکرخ تعیین گردید (**Error! Reference source not found.** ۱) و سپس بر اساس مختصات نقاط، محل آنها با استفاده از سیستم موقعیت‌یاب جهانی^۶ (GPS) در صحرا مشخص گردید. با توجه به محدودیت‌های موجود در منطقه ۵ خاکرخ در نقاطی بود که دسترسی به آنها ممکن نشد. خاکرخ‌های حفر شده، براساس راهنمای تشریح و نمونه‌برداری خاک در صحرا و با استفاده از کتاب صحرایی (Schoenberger و همکاران ۲۰۱۲) تشریح و از تمامی افق‌های ژنتیک نمونه تهیه گردید.

¹ Classification and regression trees

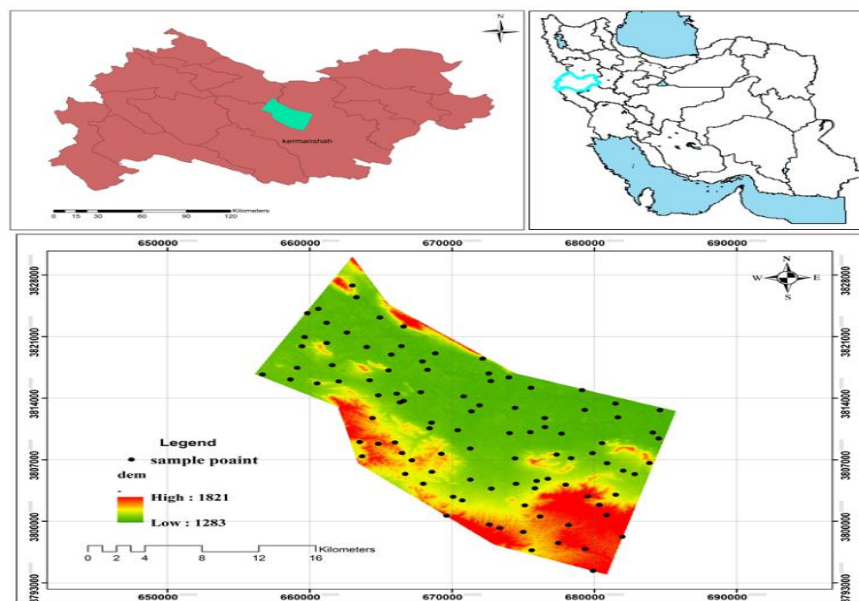
² Support vector machines

³ Classification trees

⁴ K-nearest neighbor

⁵ Generalized Random Tessellation Sampling

⁶ Global Positioning System



شکل ۱. موقعیت منطقه مطالعاتی و پراکنش نقاط مطالعاتی

پیش از تجزیه، نمونه‌های خاک از الک ۲ میلی‌متری گذرانده شده و ویژگی‌های مهم فیزیکی و شیمیایی مانند واکنش گل اشباع، رسانش الکتریکی عصاره ۲/۵:۱، بافت، کربن آلی، و کربنات کلسیم معادل به روش‌های استاندارد (Page, 1996; Klute, 1996) اندازه‌گیری شد. بر اساس نتایج ریخت‌شناختی در صحرا و تجزیه‌های آزمایشگاهی هر پدان مطابق با سیستم رده‌بندی WRB در سطح اول (RSG) رده‌بندی گردید. برای پیش‌بینی کلاس‌های خاک متغیرهای محیطی (مانند $NDVI^3$, TWI^2 , $MRVBF^1$) از مدل رقومی ارتفاع (DEM) و تصویر ماهواره Landsat 8 OLI/TIRS برگرفته شد. مدل‌سازی و پیش‌بینی کلاس‌های خاک در سطح گروه‌های مرجع خاک با کاربست مدل رگرسیون لاجستیک انجام شد. برای ارزیابی عملکرد مدل از معمول‌ترین سنجه‌های برآورد صحت شامل؛ صحت عمومی^۴، صحت کاربر^۵، قابلیت اطمینان تولید کننده^۶ و شاخص کاپا^۷ (Lu و همکاران ۲۰۰۴) استفاده شد.

نتایج و بحث

جدول ۱ نتایج رده‌بندی خاک‌ها را بر اساس سطح اول سیستم رده‌بندی WRB همراه با فراوانی مشاهده شده در هر RSG نشان می‌دهد. نتایج راستی‌آزمایی پیش‌بینی گروه‌های مرجع خاک در شکل ۲ آمده است. بررسی همزمان جدول ۱ و شکل ۲ رابطه فراوانی مشاهدات با صحت یا درستی پیش‌بینی بخوبی دریافته می‌شود در فراوانی بیشتر مشاهدات، صحت پیش‌بینی بالاتر خواهد بود. در بسیاری از پژوهش‌ها (از جمله Barthold و همکاران ۲۰۱۳) به نتیجه‌گیری‌های مشابه مبنی بر حساسیت بالای پیش‌بینی‌های رایان‌آموختی به تعداد مشاهدات اشاره شده است برای نمونه، Calcisols با بیشترین فراوانی، بالاترین صحت تولیدکننده را داشته اما صحت صفر در پیش‌بینی مکانی Leptosols (با تنها چهار مشاهده) را به احتمال قوی می‌توان با وابستگی درستی پیش‌بینی‌ها به فراوانی مشاهدات توجیه نمود. دیگر گروه‌های مرجع نیز کم و بیش از همین روند پیروی می‌کردند. اما شاید تضاد با این مطلب در Vertisols و Kastanozems باشد. به نظر می‌رسد Vertisols به‌رغم فراوانی کم در سایه ویژگی‌های منحصربه‌فرد خود و البته گزینش درست و تأثیر بالای کوواریت‌های محیطی با صحت صددرصد پیش‌بینی شده است. این نتیجه توسط Jafari و همکاران، (۲۰۱۲) نیز در پیش‌بینی کلاس‌های خاک گزارش گردیده است.

¹ Multiresolution index of valley bottom flatness

² Topographic wetness index

³ Normalized difference vegetation index

⁴ Overall map accuracy

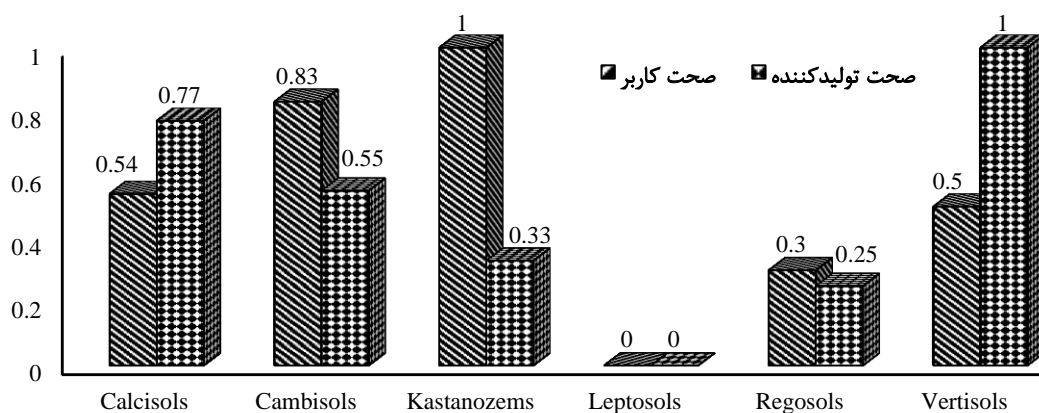
⁵ User's accuracy

⁶ Producer's reliability

⁷ Kappa Index of agreement

جدول ۱. نام و فراوانی گروه‌های مرجع خاک مشاهده شده در منطقه مورد مطالعه.

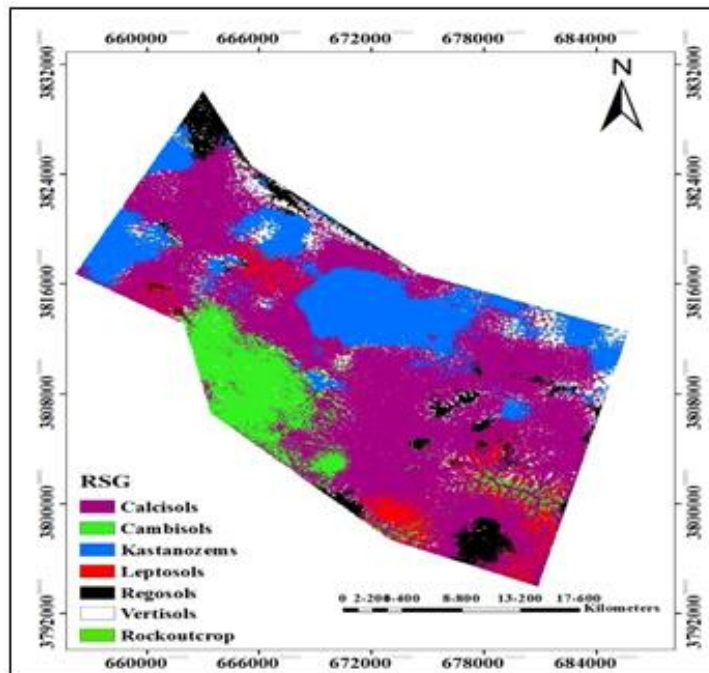
Leptosols	Vertisols	Regosols	Cambisols	Kastanozems	Calcisols
۴	۶	۱۵	۱۴	۱۷	۳۷



شکل ۲. نتایج راستی‌آزمایی پیش‌بینی گروه‌های مرجع خاک در منطقه.

در مورد Kastanozems به رغم شمار نسبی خوب مشاهدات (جدول ۱) صحت تولیدکننده ضعیف‌تری به دست آمد. به اتکاء آنچه در صحرا دیده شد و اینکه الگوریتم برخی خاک‌های این گروه مرجع را در Cambisols قرار داد، می‌توان گفت پس‌رفت ژنتیک Kastanozems زیر فشار زیاد کشت، ویژگی‌های آنها را تا اندازه‌ای به Calcisols نزدیک نموده از سوی دیگر ویژگی‌های ورتیک (نه‌چندان تیپیک) که به دلیل بافت و کانی‌شناسی ویژه منطقه میان خاک‌ها تا اندازه‌ای مشترک است از دلایل اصلی صحت کم تولیدکننده دانست. کلاس Regosols نیز به دلیل مشابه اما با شدت کمتری از درستی پیش‌بینی پایین برخوردار بود. این استدلال با آنچه توسط Hengl و همکاران (۲۰۰۷) و Jafari و همکاران (۲۰۱۲) ارائه شده هماهنگ است. راستی‌آزمایی آماری نقشه پیش‌بینی گروه‌های مرجع خاک با رگرسیون لاجستیک صحت عمومی ۵۸٪ و شاخص کاپای ۴۶٪ به دست داد. این ارقام نسبت به کاپای ۸-۱۲٪ که Brungard و همکاران (۲۰۱۵) برای همین روش و با سه مجموعه کواریت محیطی به آن دست یافتند بسیار بهتر به نظر می‌رسد، اگرچه آنان با جنگل تصادفی نتایج بهتری بدست آوردند.

در سطح دوم WRB عملکرد مدل کاملاً ناخشنودکننده بود. دلیل عمده و چشم‌گیر این وضعیت نیز کاهش شدید فراوانی مشاهدات در هر کلاس خاک در سطح دوم بود که امکان تکمیل مرحله آموزش را فراهم نمی‌کرد و نتایج مرحله‌ی راستی‌آزمایی نیز بسیار نامطلوب بود. شکل ۳ نقشه پیش‌بینی پراکنش مکانی گروه‌های مرجع خاک را در منطقه مورد مطالعه نشان می‌دهد. طبق تجربیات صحرایی Calcisols بیشترین مساحت را پوشش داده بود و Leptosols، Vertisols، Regosols، Cambisols، Kastanozems به ترتیب در مراتب بعدی قرار داشتند. چنانکه در نقشه دیده می‌شود گستره Calcisols در نقاط مسطح دشت و به عنوان خاک چیره منطقه به خوبی پیش‌بینی شده است. گروه مرجع Kastanozems پس از Calcisols بیشترین گسترش را داشت که با خاک‌های معمولاً عمیق، بیشتر در بخش‌های مرکزی و شمالی دشت سنجایی دیده می‌شد. در خاک‌ها عمدتاً افق مشخصه Cambic و در برخی نقاط Calcic نیز مشاهده شد. گروه مرجع دیگر Cambisols بود که از نظر گسترش مکانی در بخش‌های شمال‌غربی و به سمت جنوب منطقه تمرکز داشتند. گروه مرجع Regosols در این دشت عمدتاً کم‌عمق و دارای سنگریزه بالا در سطح و عمق بوده، معمولاً در حاشیه منطقه روی آبرفت‌های بادبزی شکل توسعه یافته‌اند. در Vertisols بافت سنگین و ساییده‌رخ‌های نه‌چندان تیپیک دیده شد که بیشتر آمیخته با Kastanozems و در نقاط کم ارتفاع قرار داشتند. مطابق انتظار Leptosols عمدتاً در بخش بالایی مخروط‌افکنه‌ها و دامنه ارتفاعات با پوشش خاکی نازک در سطح و لایه‌ی سنگی در زیر مشاهده گردید که معمولاً همراه با رخنمون‌های سنگی (Rockoutcrop) مشاهده گردید. این RSG کمترین گستره را در منطقه داشت و به همین دلیل فراوانی مشاهدات آن نیز کم بود. با توجه به صحت صفر مدل در پیش‌بینی این گروه مرجع خاک، موقعیت قرارگیری این گروه مرجع بر روی نقشه درست نیست و با مشاهده صحرایی همخوانی ندارد.



شکل ۳. نقشه پیش‌بینی پراکنش مکانی گروه‌های مرجع خاک WRB در منطقه‌ی مطالعاتی با استفاده از مدل رگرسیون لاجستیک.

نتیجه‌گیری

رگرسیون لاجستیک در پیش‌بینی RSG های WRB در منطقه مورد مطالعه عملکرد قابل‌قبولی داشت و تا اندازه قابل‌قبولی با الگوی پیچیده پراکنش مکانی مشاهده شده در منطقه هماهنگی داشت. مانند هر روش رایان‌آموختی دیگر، درستی پیش‌بینی‌های رگرسیون لاجستیک نیز به شدت از فراوانی مشاهدات در هر کلاس تأثیر می‌پذیرفت. به همین دلیل، کاربست الگوریتم در سطح دوم WRB به دلیل کاهش شدید مشاهدات در بسیاری از کلاس‌های خاک، نتایج مناسبی به دست نداد. نکته مهم دیگر گزینش موثرترین کواریت‌ها در منطقه بود که به نوبه خود کاملاً بر صحت پیش‌بینی RSG ها توسط مدل در منطقه‌ی مطالعاتی اثرگذار بود. همچنین عملکرد مدل نشان داد که در جداسازی خاک‌هایی با ویژگی‌های ژنتیک مشابه و نزدیک به هم و حتی به رغم شمار مناسب مشاهدات، دچار اشتباه می‌شود.

منابع

بهمنی، آ. ۱۳۹۵. توان‌آزمایی جنگل تصادفی در پیش‌بینی پراکنش مکانی خاک در بخش نیمه‌خشک استان کرمانشاه. پایان نامه کارشناسی ارشد، دانشگاه رازی. گروه علوم و مهندسی خاک. ۶۵ صفحه.

کیانیان، ن. ۱۳۹۷. بازشناخت الگوی خاک یک زمین چهار نیمه خشک در استان کرمانشاه با استفاده از الگوریتم کای نزدیکترین همسایه. پایان نامه کارشناسی ارشد، دانشگاه رازی. گروه علوم و مهندسی خاک.

ویسی، م. ۱۳۹۷. پیش‌بینی الگوی خاک در یک زمین چهار برگرزیده در استان کرمانشاه با کاربست شبکه‌های عصبی مصنوعی. پایان نامه کارشناسی ارشد، دانشگاه رازی. گروه علوم و مهندسی خاک.

ویسی‌سرطایره، غ. ۱۳۹۵. بازشناخت الگوی پراکنش خاک در دشت میان‌در بند کرمانشاه با کاربست مدل ماشین‌های بردار پشتیبان. پایان نامه کارشناسی ارشد، دانشگاه رازی. گروه علوم و مهندسی خاک. ۸۳ صفحه.

Barthold, F.K., Wiesmeier, M., Breuer, L., Frede, H.G., Wu, J. and Blank, F. B. 2013. Land use and climate control the spatial distribution of soil types in the grasslands of Inner Mongolia. *Journal of arid environments*, 88, 194-205.



- Collard, F., Kempen, B., Heuvelink, G.B.M, Saby, N.P.A., Richer de Forges, A.C., Lehmann, S., Nehlig, P. and Arrouays, D. 2014. Refining a reconnaissance soil map by calibrating regression models with data from the same map (Normandy, France). *Geoderma Regional* 1, 21-30.
- Hengl, T., Heuvelink, G.B.M. and Stein, A. 2007. About regression-kriging: From equations to case studies. *Computers & Geosciences* 33, 1301-1315.
- Jafari, A., Finke, P.A., Vande Wauw, J., Ayoubi, S. and Khademi, H. 2012. Spatial prediction of USDA-great soil groups in the arid Zarand region, Iran: comparing logistic regression approaches to predict diagnostic horizons and soil types. *European Journal of Soil Science* 63(2): 284-298.
- Kleinbaum, D.G., Kupper, L.L., Nizam, A. and Muller, K.E. 2008. Logistic regression analysis. In *Applied Regression Analysis and Other Multivariate Methods* 4th ed., Thomson Brooks/Cole, Belmont, CA: 604-634.
- Klute, A., 1996. *Methods of soil analysis. Part 1. Chemical and mineralogical properties.* Monograph No. 9. 2nd. ed. ASA monograph No. 9. SSSA, Madison, WI, USA.
- McBratney, A.B., Santos, M.M. and Minasny, B. 2003. On digital soil mapping. *Geoderma* 117(1): 3-52.
- Page, A.L. 1996. *Methods of soil analysis. Part 2. Chemical and mineralogical properties.* Monograph No. 9. 2nd. ed. ASA monograph No. 9. SSSA, Madison, WI, USA.
- Rossiter, D. 2005. Digital soil mapping: Towards a multiple-use Soil Information System. *Análisis Geográficos (Revista del Instituto Geográfico "Augustín Codazzi")*, 32(1), 7-15.
- Schoeneberger, P.J., Wysocki, D.A., Benham, E.C. and Soil Survey Staff. 2012. 2012. *Field book for describing and sampling soils, Version 3.0.* Lincoln: Natural Resources Conservation Service, National Soil Survey Center.
- IUSS Working Group WRB. G. 2015. *World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps.* World Soil Resources Reports No. 106. FAO, Rome.



16th Iranian Soil Science Congress

University of Zanjan, Iran, August 27-29, 2019



Topic for submission: Pedometry and Soil Evaluation
Spatial prediction of RSGs in a semi-arid landscape using logistic regression

Paidar¹, H., Shekaari^{2*}, P.

¹ Graduate student, Soil Science Department, Faculty of Agriculture, Razi University, Kermanshah, Iran.

² Assistant Prof., Soil Science Department, Faculty of Agriculture, Razi University, Kermanshah, Iran.

Abstract

During the last two decades application of quantitative techniques in soil mapping is considerably increased. In this work, the relatively widespread machine learning (ML) method, logistic regression (LR) applied to spatial prediction of WRB RSGs over a 48,000 ha. extent of Sandjabi area of Kermanshah province. Using GRTS sampling algorithm, 100 observation points distributed over the area. All the pedons were described, sampled, analyzed, and based on morphological and laboratory results classified at the RSG level of WRB soil classification system. Environmental covariates calculated based on a DEM file and a Landsat 8 scene, which applied as predictors in spatial prediction of RSGs. Validation measures Overall map accuracy (OA) and Kappa index of agreement (K) were calculated 0.58 and 0.46, respectively. Calcisols -the most frequent RSG in the dataset- was predicted accurately. Generally, considering sample density and study area extent, LR showed a considerable capability in predicting pattern of soil taxonomic classes over the study area. However, results confirmed that LR like other ML methods is sensitive to frequency of observations and environmental covariates.

Keywords: Digital soil mapping, environmental covariates, Sandjabi area.

* Corresponding author, Email: pshekaari@gmail.com