تهیه نقشه پیش‌بینانه پراکنش گروه بزرگ‌های خاک در یک منطقه نیمه‌خشک با کاربست الگوریتم درخت تصمیم‌گیری C5.0

مرضیه زندی باغچه‌مریم و پرویز شکاری
به‌ترتیب دانشجوی کارشناسی ارشد و استادیار گروه علوم ومهندسی خاک، پردیس کشاورزی و منابع طبیعی، دانشگاه رازی، کرمانشاه

چکیده

بررسی تغییرات مکانی خاک از نخستین اولویت‌های مدیریت منابع طبیعی است. هدف از این مطالعه بررسی توانایی الگوریتم بهینه شده درخت تصمیم‌گیری[1] (BDTA) در آشکار ساختن الگوی پراکنش خاک‌ها در سطح گروه بزرگ Soil Taxonomy (ST) در دشت میان‌دربند استان کرمانشاه بود. پس از طراحی یک الگوی سیستماتیک ساده، ۷۸ نقطه مشاهداتی در سطح منطقه بررسی و ۶ گروه بزرگ خاک شناسایی شد. ۲۸ متغیر محیطی از مدل رقومی ارتفاع منطقه و نیز تصویر ماهواره‌ای OLI با تفکیک ۳۰متر برگرفته شد. پس از آموزش الگوریتم درخت تصمیم‌گیری، مراحل پیش‌بینی و تهیه نقشه انجام گرفت. راستی آزمایی عملکرد الگوریتم بر اساس دو سنجه صحت عمومی نقشه (OA) و نمایه سازگاری کاپا (K) به ترتیب برابر با ۰/۸۰ و ۰/۷۴ محاسبه گردید. موثرترین متغیرهای محیطی در پیدایش گروه بزرگ‌های خاک در منطقه عمدتا شامل ویژگی‌های توپوگرافی و به مقدار کمتر مواد مادری بود. بر پایه نتایج BDTA از عملکرد چشم‌گیری برخوردار بود. این نتیجه‌گیری تنها بر پایه مقادیر OA و K نبوده و دلیل دیگر آن توپوگرافی هموار در بخش بزرگی از منطقه مطالعاتی بود که هر گونه استنتاج را بسیار دشوار می‌ساخت. همچنین BDTA زیر اثر بیش برازش قرار نگرفت. به طور کلی نتایج نشان از توان بالای پیش بینی BDTA داشت که آن را به یک جایگزین نیرومند در کارهای آتی تبدیل می‌کند.

واژه‌های کلیدی: نقشه برداری رقومی خاک، درخت تصمیم‌گیری، دشت میان‌دربند، متغیرهای محیطی.

مقدمه

روی آوردن به روش‌های جدید در راستای کاربرد بهینه خاک به گونه‌ای که هم نیازهای آدمی برآورده شود و هم این منبع تجدید ناپذیر برای آیندگان حفظ شود ضروری است. نقشه‌های خاک از منابع پایه اطلاعات برای بسیاری مطالعات محیطی و از جمله مهم‌ترین آنها در این رابطه ارزیابی و آمایش سرزمین است (Daigle et al., 2005). تغییرپذیری خاک پدیده‌ای تصادفی نیست و نتیجه‌ی تغییر عوامل خاک‌سازی از نقطه‌ای به نقطه دیگر است. به همین دلیل، این عوامل را می‌توان در پیش‌بینی کلاس خاک و تهیه نقشه الگوی پراکنش خاک‌ها به کار برد. نقشه‌ی خاک وسعت و توزیع جغرافیایی خاک‌ها و ویژگی‌های آن را تا حدی که مقیاس نقشه اجازه دهد، ارائه می‌کند (علیخانی و همکاران، ۱۳۹۳). به دیگر سخن، نقشه خاک یک نمایش تصویری برای انتقال اطلاعات در مورد پراکنش مکانی ویژگی‌های خاک است. با توجه به محدودیت‌ها و کمبودهای موجود در نقشه‌های سنتی، ایجاد نقشه‌های جدید که کاستی‌های گذشته را نداشته و توانایی آشکار کردن پراکنش مکانی و تغییرات خاک را دارا باشد، ضروری است. بطور کلی نقشه‌برداری خاک‌ها به صورت نقشه‌های سنتی (چندگوشه‌ای[2]) به علت ناتوانی نمایش تغییرپذیری خاک به صورت پیوسته (Burrough et al.,1989; McBratney et al., 2003; Ryan et al., 2000; Scull et al., 2003)، مستند نشدن بیشتر دانش درگیر در تولید این نقشه‌ها و در نتیجه منتقل نشدن دانش به نسل آینده (Hudson, 1992)، عدم وجود اطلاعات مکانی در مورد دقت و صحت نقشه (Burrough et al., 1971)، وابستگی به مقیاس و حذف اطلاعات هنگام تلفیق واحدهای کوچک‌تر در واحدهای بزرگتر نقشه (بداغ آبادی، ۱۳۹۱)، و ماهیت ذهنی و کیفی مدلی که نقشه بر اساس آن تولید می‌شود، مورد انتقاد واقع شده است.

1. Boosted Decision Tree Algorithm
2. Polygonal

چالش درک، دسته‌بندی و استفاده بهینه از این اطلاعات به ایجاد و گسترش ابزارهای نوینی مانند روش‌های داده‌کاوی[1] و رایان‌آموختی[2] انجامیده است. همزمان با آن، توسعه‌ی سیستم‌های اطلاعات جغرافیایی، سنجش از دور و مدل‌های رقومی ارتفاع با درجه تفکیک مکانی بالا، بستر پیشرفت دانش نقشه‌برداری رقومی خاک (DSM) را فراهم کرده است (McBratney et al., 2003). هدف اصلی از DSM استفاده از داده‌های موجود برای مطالعه پراکنش مکانی خاک‌ها به صورت واقعی‌تر، مؤثرتر، ارزان‌تر و سریع‌تر است (Elnaggar, 2007). اساس DSM برگرفته از مدل ۵ عاملی ینی است که در قالب مدل اسکورپن روزآمد شده و امکان توصیف کمی روابط خاک و متغیرهای محیطی را فراهم می‌کند.

DSM با در نظر گرفتن خطای کمی مدل، پیش‌بینی مکانی واقعی ایجاد می‌کند (Lagacherie, 2008). در سال‌های اخیر خاک‌شناسان برای تولید پایگاه داده‌های مکانی خاک و تکمیل نقشه رقومی خاک‌ها در سطح جهانی تحقیقات جامعی را آغاز کرده‌اند. از آن میان می‌توان به مقایسه روش‌های رایان آموختی برای هدف‌های طبقه‌بندی در نقشه‌برداری رقومی خاک (Heung et al, 2016; Brungard et al., 2014)، پیش‌بینی نقشه خاک با استفاده از شبکه عصبی مصنوعی و درخت تصمیم‌گیری (Moonjun et al., 2010)، نقشه‌برداری رقومی خاک با استفاده از پدان‌های مشاهده شده در یک مطالعه خاک‌شناسی سنتی با کاربرد الگوریتم ژنتیک [3](GARP) و درخت تصمیم‌گیری (Nelson and odeh, 2009). اشاره نمود.

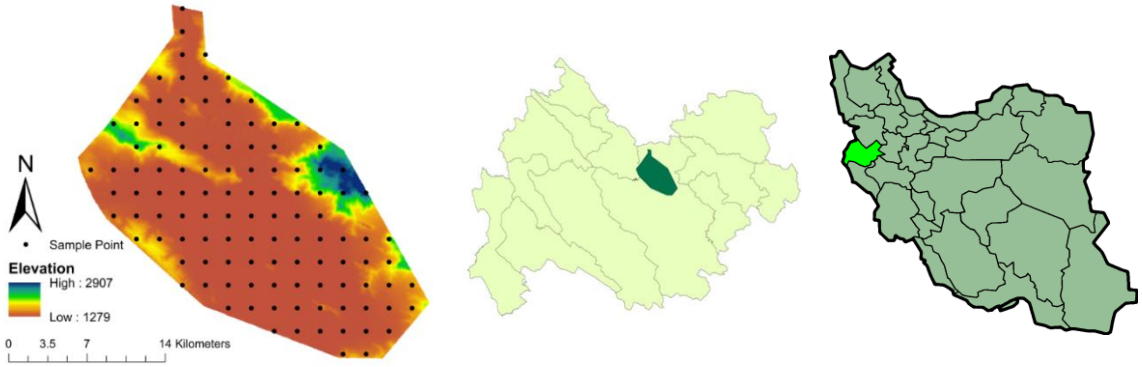
در ایران، شمار پژوهش‌های مرتبط با DSM بسیار کم بوده که تنها در سالهای اخیر انجام شده‌اند. از آن میان می‌توان کاربرد مدل‌های رگرسیون لاجیستیک دودویی۴ و رگرسیون درختی۵ (جعفری و همکاران، ۱۳۹۱)، پیش‌بینی کلاس‌های خاک با کاربرد درخت تصمیم‌گیری تصادفی (پهلوان راد و همکاران، ۱۳۹۴)، تهیه نقشه رقومی بافت خاک با درخت تصمیم‌گیری و شبکه عصبی مصنوعی (مهرجردی و همکاران ، ۱۳۹۳) و همچنین تهیه نقشه رقومی کلاس‌های خاک در استان یزد با استفاده از الگوریتم درخت تصمیم‌گیری (Taghizadeh-Mehrjardi et al., 2014) را برشمرد. برای مدیریت بهینه سرزمین به منظور دست یافتن به سطح قابل قبولی از تولید محصولات و در کنار آن حفظ منابع خاک در ایران به ویژه در مناطق خشک و نیمه خشک، به مطالعات بیشتر و به کارگیری روش‌های جدید نقشه‌برداری نیاز است. هدف از پژوهش حاضر بررسی توان BDTA در پیش‌بینی پراکنش خاک در سطح گروه بزرگ ST در دشت میان‌دربند استان کرمانشاه بود.

مواد و روش‌ها

منطقه مطالعاتی منطقه مورد مطالعه دشت میان‌دربند واقع در استان کرمانشاه با مساحت حدود ۵۰/۰۰۰ هکتار است (شکل ۱). این منطقه در دامنه طول جغرافیایی ´۴۵ °۴۶ تا ´۰۵ °۴۷ و عرض جغرافیایی ´۲۳ °۳۴ تا ´۳۵ °۳۴ واقع شده است. این دشت با ارتفاع متوسط ۱۳۵۷/۷ متر از سطح دریا دارای آب و هوای معتدل با زمستان‌های سرد است. میانگین سالیانه‌ی دما ۱۴°C، میزان تبخیر ۱۳۷۷/۳ میلی‌متر در سال و میانگین سالیانه بارندگی ۴۶۲/۶ میلی‌متر است. رژیم رطوبتی و دمایی خاک منطقه به ترتیب زریک و ترمیک است (Banaii, 1977).

طرح نمونه‌برداری به منظور تعیین مکان نقاط نمونه برداری خاک و استخراج متغیرهای محیطی با استفاده از روش نمونه برداری سیستماتیک در محیط نرم افزار R (R Core Team, 2014) یک شبکه منظم به ابعاد ۲×۲ کیلومتر (شکل۱) طراحی شده و بر اساس آن، مکان هر نقطه توسط GPS دستی با دقت ±۳ متر پیمایش شد. با توجه به محدودیت‌های موجود در منطقه ۵ نقطه از طرح نمونه برداری حذف گردید. تعداد ۷۸ نقطه مشاهداتی در پهنه خاکی و ۳۷ نقطه در کوهستان قرار گرفت. به جز نقاطی که سنگ بستر یا عوامل محدود کننده مانع شد، در محل هر نقطه یک گودال بررسی تا عمق حداقل ۱۲۰ سانتی‌متر برای مطالعه پدان حفر شد.

1. Data mining
2. Machine learning
3. Genetic Algorithm
4. Binary logistic regression
5. Tree regression



شکل ۱- موقعیت منطقه مطالعاتی و طرح نمونه برداری.

مطالعات صحرائی و تجزیه‌های آزمایشگاهی تشریح ریخت‌شناختی پدان‌ها بر اساس راهنمای شناسایی خاک در صحرا (شونبرگر و همکاران، ۲۰۱۲) انجام شد. نمونه‌های تمام افق‌ها برای انجام تجزیه‌های فیزیکی و شیمیایی از الک ۲ میلی‌متری عبور داده شد. ویژگی‌های خاک شامل واکنش، بافت، کربن آلی و کربنات کلسیم معادل به روش‌های استاندارد (Klut et al., 1996; Page et al., 1996) اندازه‌گیری شد. سپس بر اساس نتایج ریخت‌شناختی و تجزیه‌های آزمایشگاهی هر یک از پدان‌ها مطابق با ST (Soil survey staff, 2014) در سطح گروه‌بزرگ رده‌بندی شد.

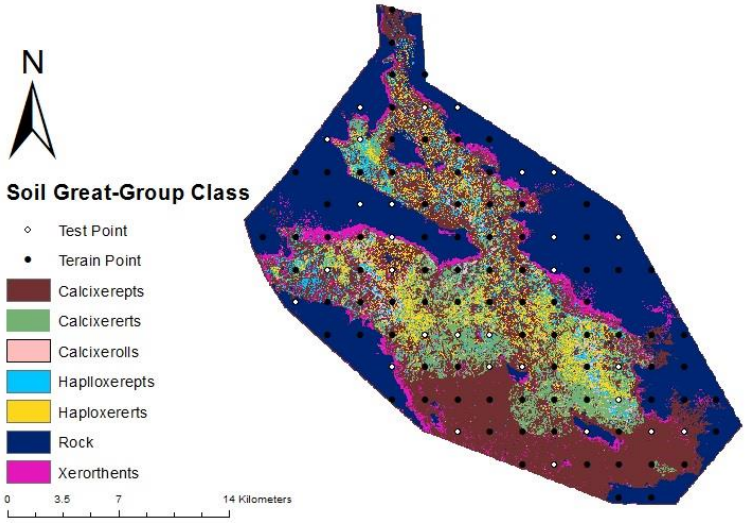
استخراج متغیرهای محیطی به منظور تعیین متغیرهای پیش‌بین، ۱۹ ویژگی زمین (توپوگرافیک و آب‌شناختی) به عنوان متداول‌ترین متغیرهای محیطی و ۱۱ نمایه پوشش گیاهی و مواد مادری به ترتیب از فایل DEM و تصویر ماهواره Landsat 8 OLI/TIRS (هر دو با وضوح ۳۰ متر) در محیط نرم افزار SAGA 3.0.0 محاسبه شد.

تولید نقشه‌های رقومی و ارزیابی مدل در این مطالعه به منظور پیش‌بینی گروه بزرگ خاک از الگوریتم نظارتی درخت تصمیم‌گیری به نام C5.0 که به کمک روش بوستینگ تقویت شده در استنتاج اطلاعات به کاررفت. برای ارزیابی صحت پیش‌بینی‌ها از نمایه سازگاری کاپا[1] (Cohen, 1960) و صحت عمومی[2] (Jensen, 1996) بهره برده شد.

نتایج و بحث

مطالعه صحرایی ساختاری پیچیده و درهم از گروه‌بزرگ‌های شناسایی شده را آشکار ساخت. الگوی پراکنش خاک چنان بود که پیش‌بینی فراوانی و گسترش هر خاک به روش سنتی و با صحتی پذیرفتنی، تقریبا غیرممکن می‌نمود. گروه بزرگ‌های شناسایی شده شامل Calcixererts با ۱۰، Calcixerolls با ۴، Calcixerepts با ۳۷، Haploxererts با ۸، Calcixererts با ۱۰ و Haploxerepts با ۱۰ مشاهده بود. نقشه پیش‌بینانه پراکنش گروه بزرگ خاک‌ها در منطقه مطالعاتی با کاربرد BDTA در شکل ۳ دیده می‌شود. در دید کلی، نقشه با تجارب و مشاهدات صحرایی، هماهنگی بالایی دارد. به ویژه توان استنتاج الگوریتم در بخش تقریبا هموار و بزرگی از دشت که متغیرهای توپوگرافیک کمک چندانی نمی‌کردند (منطقه زیر پوشش Calcixerepts به رنگ قهواه‌ای)، جالب توجه بود. جدول ۲ نتایج راستی آزمایی پیش‌بینی مدل را پس از بارها اجرای الگوریتم روی مجموعه داده‌ها با صحت عمومی ۰/۸۰ نشان می‌دهد. نمایه سازگاری کاپا نیز ۰/۷۴ محاسبه شد. پیش‌بینی پراکنش ۵ گروه بزرگ خشنودکننده بود اما مدل در پیش‌بینی گروه بزرگ Haploxerepts موفق نبود. کمینه دقت تولید کننده در گروه بزرگ Calcixererts برابر با ۶۰ درصد و بیشینه آن در گروه بزرگ Calcixerolls و Calcixerepts برابر با ۱۰۰ درصد بود. همچنین کمینه دقت کاربر در گروه بزرگ Haploxererts برابر با ۵۰ درصد و بیشینه آن در گروه بزرگ Calcixererts و Calcixerolls بود. میزان اثر برخی از متغیرهای محیطی در شکل ۴ خلاصه شده است. متغیرهای تابشی، آب‌شناختی، توپوگرافیک و پس از آن نمایه‌های آهن، گچ و رس در گوناگونی خاک‌ها در منطقه نقش چشم‌گیری داشتند. به طور کلی می‌توان گفت که عامل

1. Kappa index of agreement
2. Overall Accuracy



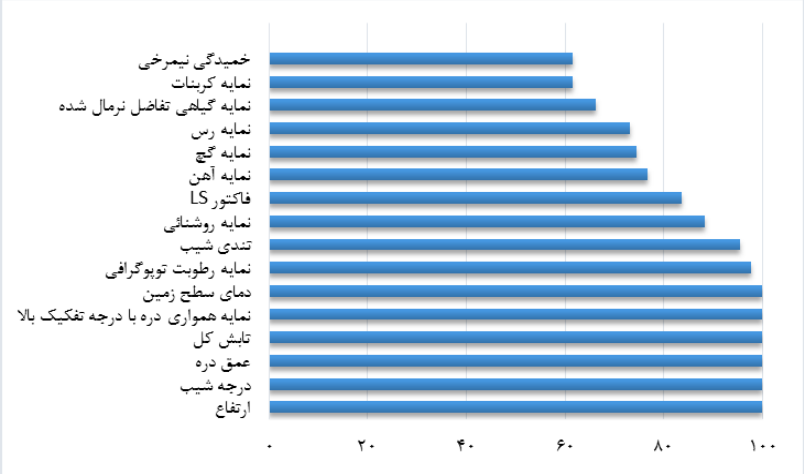
شکل۳- نقشه پیش‌بینی پراکنش گروه بزرگ خاک در سطح منطقه مطالعاتی با BDTA.

جدول۲- ماتریس درهمی پیش‌بینی خاک منطقه در سطح گروه بزرگ.

گروه بزرگ	a	b	c	d	e	f	g	دقت کاربر
a: Calcixerepts	۶	۱	۰	۰	۰	۰	۱	۰/۷۵
b: Calcixererts	۰	۳	۰	۰	۰	۰	۰	۱۰۰
c: Calcixerolls	۰	۰	۱	۰	۰	۰	۰	۱۰۰
d: Haploxerepts	۱	۰	۰	۰	۰	۱	۰	۰
e: Haploxererts	۱	۱	۰	۰	۲	۰	۰	۰/۵۰
f: Rock	۰	۰	۰	۰	۰	۱۲	۰	۱۰۰
g: Xerorthents	۰	۰	۰	۱	۰	۰	۴	۰/۸۰
دقت تولید کننده	۰/۷۵	۰/۶۰	۱۰۰	۰	۱۰۰	۰/۹۲	۸۰	۰/۸۰

توپوگرافی و پس از آن مواد مادری بیشترین تأثیر را در تشکیل گروه‌بزرگ‌های خاک دشت میان‌دربند دارد.

نتایج به دست آمده با دیگر مطالعات انجام شده در مناطق خشک و نیمه خشک ایران شباهت دارد. مثلا در مطالعه انجام شده توسط (Taghizadeh-Mehrjardi et al., 2014) انواع متغیرهای محیطی و کاربرد الگوریتم‌های مختلف الگوریتم C5.0 با صحت عمومی ۶۷/۵ درصد موفق‌ترین پیش‌بینی پراکنش گروه بزرگ‌های خاک را ارائه نمود. همچنین مهرجردی و همکاران (۱۳۹۲) در مطالعه‌ای دیگر از ترکیب فاصله تاکسونومیک و درخت تصمیم‌گیری برای پیش‌بینی پراکنش خاک در سطح گروه بزرگ بهره گرفته و صحت عمومی ۷۰ درصد به دست آوردند. براین اساس، پیشنهاد کردند که الگوریتم درخت تصمیم‌گیری در دیگر مناطق ایران به کار گرفته شود. در پژوهش دیگر (Jafari et al., 2012) صحت عمومی پیش‌بینی پراکنش گروه بزرگ خاک ۶۷/۵ درصد به دست آمد. همچنین (Scull et al., 2003) در پیش‌بینی گروه بزرگ خاک با کاربرد الگوریتم درخت تصمیم‌گیری به صحت عمومی ۷۰ درصد دست یافتند. (Moran and Bui, 2002) نیز با کاربرد بوستینگ درخت تصمیم‌گیری نقشه پیش‌بینانه گروه بزرگ خاک را با دقت ۷۰ درصد به دست آوردند. نتایج پژوهش‌های گوناگون درخت تصمیم‌گیری توان بالای این الگوریتم را در آشکارسازی الگوها نشان داده است. بر این پایه، به نظر می‌رسد آمیختن آن با دیگر روش‌هایی که توان مناسبی دارند می‌تواند گام‌هایی برای پژوهش‌های آتی باشد.



شکل ۴- اهمیت متغیرهای محیطی در پیش‌بینی BDTA در سطح گروه بزرگ

منابع

باقری بداغ آبادی، م. ۱۳۹۱. نقشه‌برداری رقومی خاک با استفاده از تکنیک محاسبات نرم. پایان‌نامه‌ی دکترای خاک‌شناسی، دانشکده‌ی کشاورزی، دانشگاه شهرکرد.

پهلوان راد، م.، خرمالی، ف.، تومانیان، ن.، کیانی، ف.، و کمکی، ب. ۱۳۹۴. پیش‌بینی کلاس‌های خاک با روش‌های درختان تصمیم‌گیری تصادفی و رگرسیون منطقی چند متغیره در استان گلستان. مجموعه مقالات چهاردهمین کنگره علوم خاک ایران. ص. ۱۱۵-۱۰۱. دانشگاه ولی‌عصر (عج)، رفسنجان، ایران.

جعفری، ا.، خادمی، ح.، و ایوبی، ش. ۱۳۹۱. نقشه‌برداری رقومی افق‌های مشخصه در گزوه‌های بزرگ خاک در منطقه زرند کرمان. مجله علوم و فنون کشاورزی و منابع طبیعی، علوم آب و خاک. شماره ۶۲. ص. ۱۷۷-۱۹۱.

مهرجردی، ر.، سرمدیان، ف.، امید، م.، تومانیان، ن.، روستا، م.، و رحیمیان، م. ۱۳۹۳. نقشه‌برداری رقومی کلاس‌های خاک با استفاده از انواع روش‌های داده کاوی در منطقه اردکان استان یزد. مهندسی زراعی (مجله علمی کشاورزی). جلد ۳۷، شماره۲، ص. ۳۹-۵۳.

مهرجردی، ر.، سرمدیان، ف.، امید، م.، تومانیان، ن.، روستا، م.، و رحیمیان، م. ۱۳۹۲. ترکیب فاصله تاکسونومیکی خاک‌ها و درخت تصمیم‌گیری برای پیش‌بینی مکانی کلاس‌های خاک در اردکان. دو فصلنامه علمی-پژوهشی خشک بوم. دوره ۳، شماره ۲، ص ۲۷-۳۹.

علیجانی، ز.، سرمدیان، ف.، و موسوی، ر. ۱۳۹۳.مقایسه دقت نقشه خاک تهیه شده به روش ژئوپدولوژی و روش معمول ایران (مطالعه موردی: کوهین). نشریه مرتع وآبخیزداری، مجله منابع طبیعی ایران دوره۶۷، شماره ۱، ص. ۹۳-۱۰۲.

Banaii, M. H. 1977. Soil moisture and temperature regimes map of Iran. Soil and Water research Institute of Iran. Ministry of Agriculture, Tehran, Iran.

Brungard, C. W., Boettinger, J. L., Duniway, M. C., Wills, S. A., and Edwards, T. C. 2015. Machine learning for predicting soil classes in three semi-arid landscapes. Geoderma 239: 68-83.

Burrough, P. A. 1989. Fuzzy mathematical methods for soil survey and land evaluation. European Journal of Soil Science, 40(3): 477-492.

Burrough, P. A., Beckett, P. H. T., and Jarvis, M. G. 1971. The relation between cost andutility in soil survey. European Journal of Soil Science 22(2): 368-381.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20: 37-46.

Daigle, J. J., Hudnall, W. H., Gabriel, W. J., Mersiovsky, E., and Nielson, R. D. 2005. The National Soil Information System (NASIS): Designing soil interpretation classes for military land-use predictions. Journal of terramechanics, 42(3): 305-320.

Elnaggar, A. A. 2007. Development of predictive mapping techniques for soil survey and salinity mapping Oregon State University (Doctoral dissertation). pp. 7-11.

Heung, B., Ho, H. C., Zhang, J., Knudby, A., Bulmer, C. E., and Schmidt, M. G. 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma 265: 62-77.

Hudson, B. D. 1992. The soil survey as paradigm-based science. Soil Science Society of America Journal 56: 836-841.

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.

Jensen, J. R. 1996. Introductory digital image processing: A remote sensing perspective. 2nd Ed., Prentice Hall, Inc., Upper Saddle River, New Jersey, USA.

Klute, A., and Ch, Dirksen. (ed.). 1996. Methods of soil analysis. Part 1. Physical and mineralogical properties. Monograph No. 9. 2nd. Ed. ASA monograph No. 9. SSSA, Madison, WI, USA.

Lagacherie, P. 2008. Digital soil mapping: a state of the art. In Digital soil mapping with limited data. Springer Netherlands. pp. 3-14.

McBratney, A. B., Santos, M. M., and Minasny, B. 2003. On digital soil mapping. Geoderma 117(1): 3-52.

Moonjun, R., Farshad, A., Shrestha, D. P., & Vaiphasa, C. 2010. Artificial neural network and decision tree in predictive soil mapping of Hoi Num Rin sub-watershed, Thailand. In Digital Soil Mapping. Springer Netherlands. pp. 151-164.

Moran, C. J., and Bui, E. N. 2002. Spatial data mining for enhanced soil map modelling. International Journal of Geographical Information Science 16(6): 533-549.

Nelson, M. A., and Odeh, I. O. A. 2009. Digital soil class mapping using legacy soil profile data: a comparison of a genetic algorithm and classification tree approach. Soil Research 47(6): 632-649.

Page, A. L., Miller, R. H., and Jeeney, D. R. (Ed.). 1996. Methods of soil analysis. Part 2. Chemical and mineralogical properties. Monograph No. 9. 2nd. ed. ASA monograph No. 9. SSSA, Madison, WI, USA.

Ryan, P. J., McKenzie, N. J., O'Connell, D., Loughhead, A. N., Leppert, P. M., Jacquier, D., and Ashton, L. 2000. Integrating forest soils information across scales: spatial prediction of soil properties under Australian forests. Forest Ecology and Management 138(1): 139-157.

Schoeneberger, P. J., D. A. Wysocki, E.C. Benham, and Soil Survey Staff. 2012. Field book for describing and sampling soils, Version 3.0. Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE.

Scull, P., Franklin, J., Chadwick, O. A. and McArthur, D. 2003. Predictive soil mapping: a review. Progress in Physical Geography 27(2): 171-197.

Shirali, R. 2016. Classification Trees and Rule-Based Modeling Using the C5. 0 Algorithm for Self-Image across Sex and Race in St. Washington University St. Louis.

Soil Survey Staff. 2014. Keys to Soil Taxonomy (12th. ed). USDA-Natural Resources Conservation Service. National Soil Survey Center, Lincoln, NE.

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 the Ardakan region, Iran. Arid Land Research and Management, 28(2): 147-168.

Predictive mapping of soil Great groups distribution in a semi-arid area using boosted decision tree algorithm

M. Zandi Baghcheh-Maryam and P. Shekaari,

MSc student and assistant professor of department of soil science, college of agriculture, Razi University, Kermanshah.

Abstract

Soil spatial variability is of prime concern in environmental management. This study conducted to verify how boosted decision tree algorithm (BDTA) is capable in predicting distribution pattern of Soil Taxonomy's Great groups in Miandarband area of Kermanshah province. According to a simple systematic sampling design, 115 pedons studied and 6 Great groups identified. 28 environmental covariates calculated from 30-meter resolution DEM and OLI image of Landsat 8. Training step of the BDTA followed by prediction and mapping stages to construct a predictive map of Great groups. Calculation of Overall accuracy (OA) and Kappa index of agreement (K) led to 0.80 and 0.74 prediction accuracies, respectively. Generally, topographic covariates were recognized most effective ones, while parent material showed a minor effect on differentiation of Great groups in the study area. Based on the results, BDTA showed a considerable performance. This conclusion is not just based on OA and K values but because of nearly flat topography in a large extent of the area which made any inference too difficult. Furthermore, it did not affected by overfitting. Considering all the points, BDTA showed good predictability potential that make it a powerful alternative for future works.

Key Words: Digital Soil Mapping, Decision Tree, Miandarband Plain, Environmental covariates.