



دانشگاه گواران، دانش‌پژوهی آب و خاک

نشریه پژوهش‌های حفاظت آب و خاک

جلد بیست و پنجم، شماره دوم، ۱۳۹۷

<http://jwsc.gau.ac.ir>

DOI: 10.22069/jwsc.2018.13870.2860

## انبوهش‌زدایی مکانی نقشه نیمه‌تفصیلی خاک با استفاده از روش دسمارت

\*شاهرخ فاتحی<sup>۱</sup>، کامران افتخاری<sup>۲</sup> و جلال قادری<sup>۱</sup>

استادیار پژوهش، بخش تحقیقات خاک و آب، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان کرمانشاه، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرمانشاه، ایران، استادیار پژوهش، مؤسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران  
تاریخ دریافت: ۹۶/۵/۳۱؛ تاریخ پذیرش: ۹۷/۲/۲

### چکیده

**سابقه و هدف:** مدیریت چالش‌های جهان امروز مانند امنیت غذایی و مسائل زیست‌محیطی، نیازمند اطلاعات رقومی خاک با صحت و دقت کافی و با قدرت تفکیک مکانی ریز است. نقشه‌های خاک عمدتاً در مقیاس کوچک قابل دسترس هستند. با وجود این، طی دو دهه اخیر با ظهور روش‌های نقشه‌برداری و مدل‌سازی رقومی خاک، امکان انبوهش‌زدایی مکانی<sup>۱</sup> واحدهای نقشه خاک فراهم شده است. انبوهش‌زدایی مکانی واحدهای نقشه خاک، روشی است که توزیع مکانی کلاس‌های منفرد خاک را مدل‌سازی می‌کند. طی این فرایند اطلاعات نقشه خاک از مقیاس کوچک (قدرت تفکیک مکانی درشت) به مقیاس بزرگ (قدرت تفکیک مکانی ریز) تبدیل می‌شود و برای اجرای آن، روش‌های آماری و داده‌کاوی مورد استفاده قرار می‌گیرد. هدف از انجام این پروژه، استفاده از روش انبوهش‌زدایی و یکنواخت‌سازی واحدهای نقشه خاک از طریق بازچینی توسط درختان طبقه‌بندی<sup>۲</sup> (دسمارت<sup>۳</sup>) برای انبوهش‌زدایی واحدهای نقشه نیمه‌تفصیلی خاک است.

**مواد و روش‌ها:** ناحیه مورد مطالعه با وسعتی حدود ۱۴۰۸۳/۹ هکتار در استان کرمانشاه قرار دارد. نقشه خاک منطقه از ۵ واحد نقشه تشکیل شده است و در کل واحدهای نقشه مرکب از ۴ زیرگروه خاک می‌باشند. در این پژوهش، با استفاده از مدل‌های رقومی ارتفاع با قدرت تفکیک مکانی ۱۰ متر، نقشه متغیرهای کمکی شامل ارتفاع، درصد شیب، جهت شیب، تحدب، تابش مستقیم، شاخص رسوب و شاخص همواری دره با درجه تفکیک بالا، شاخص خیسی، عمق دره و فاصله عمودی تا شبکه آبراهه تهیه گردید. همچنین از تصاویر ماهواره‌ای لندست 7<sup>+</sup> ETM، برای تولید متغیر کمکی شاخص رس، شاخص اندازه ذرات و شاخص پوشش گیاهی نرمال شده استفاده شد. نقشه زمین‌شناسی با مقیاس ۱:۱۰۰/۰۰۰ نیز به‌عنوان متغیر کیفی ورودی به‌کار رفت. سپس مدل دسمارت برای انبوهش‌زدایی نقشه خاک منطقه مورد مطالعه اجرا شد. روش دسمارت از درون واحدهای نقشه به‌طور تصادفی نمونه‌برداری نموده و از درختان طبقه‌بندی الگوریتم C5.0 برای تولید نقشه‌های سطوح احتمال توزیع کلاس‌های منفرد خاک استفاده می‌کند. ارزیابی

\* مسئول مکاتبه: [shahrokh.fatehi@gmail.com](mailto:shahrokh.fatehi@gmail.com)

1- Spatial disaggregation

2- Disaggregation and Harmonization of Soil Map Units Through Resampled Classification Trees

3- DSMART

نقشه‌های سطوح احتمال با استفاده از داده‌های اعتبارسنجی شامل ۸۲ نیمرخ خاک و معیارهای مستخرج از ماتریس درهمی شامل صحت تولیدکننده، صحت کاربر، صحت کلی و ضریب کاپا انجام شد. همچنین شاخص آسیمگی بین نقشه احتمال کلاس خاک مرتبه اول و دوم محاسبه گردید. شاخص آسیمگی بیان مختصری از درجه آسیمگی کلاس‌های خاک است.

**یافته‌ها:** فاصله عمودی تا شبکه آبراهه، ارتفاع، لیتولوژی، شاخص اندازه ذرات و شاخص همواری دره با درجه تفکیک بالا از مهم‌ترین متغیرهای پیش‌بینی‌کننده مدل درختان طبقه‌بندی بودند. شاخص آسیمگی نزدیک به یک، که گسترش فراوانی در ناحیه مورد مطالعه دارد احتمال وقوع یکسان زیرگروه‌های خاک در هر موقعیت مکانی در هر دو نقشه احتمال کلاس خاک مرتبه اول و دوم را نشان می‌دهد. اعتبارسنجی نقشه‌های سطوح احتمال نشان می‌دهد که صحت کلی احتمال کلاس خاک مرتبه اول (۴۴ درصد)، احتمال کلاس خاک مرتبه دوم (۲۸ درصد) و احتمال کلاس خاک مرتبه سوم (۱۱ درصد) می‌باشد. این نتایج بیانگر کارایی نسبتاً خوب این روش برای تولید نقشه رقوم کلاس‌های منفرد خاک است. اما ضرایب کاپا برای نقشه‌های سطوح احتمال خاک مرتبه اول، دوم و سوم به ترتیب ۰/۰۴، ۰/۰۲ و ۰/۰۸- به دست آمد. پایین بودن ضرایب کاپا را می‌توان ناشی از ماهیت واقعی داده‌ها یعنی غالب بودن زیرگروه تیپیک کلسی زیرپتزر در نقشه سنتی خاک، نقشه پیش‌بینی مدل دسمارت و داده‌های اعتبارسنجی دانست.

**نتیجه‌گیری:** روش دسمارت قادر است احتمال وقوع تمام کلاس‌های خاک ناحیه مورد مطالعه که توزیع آن‌ها در واحدهای نقشه نامشخص است را پیش‌بینی نماید و فرصتی برای تولید نقشه‌های رقوم کلاس خاک هنگامی که داده‌های میراثی خاک و متغیرهای کمکی در دسترس هستند فراهم می‌کند. چنین خروجی‌هایی می‌تواند به فهم روابط سیمای اراضی و خاک کمک کند.

**واژه‌های کلیدی:** انبوهش‌زدایی مکانی، نقشه رقوم خاک، کلاس خاک، واحد نقشه

## مقدمه

نقشه‌های سنتی خاک تنها منبع داده‌های کمی و کیفی خاک در کشور هستند که بیش از نیم قرن است در مقیاس‌های مختلف اجمالی، نیمه‌تفصیلی و تفصیلی تهیه شده‌اند (۱). نقشه‌های سنتی خاک موجود در کشور را می‌توان از نظر روش تهیه، به دو دسته تفکیک نمود. دسته اول نقشه‌های هستند که تفکیک واحدهای نقشه براساس تفسیر عکس‌های هوایی (با مقیاس ۱:۵۰۰۰۰ و ۱:۲۰۰۰۰) و واحدهای فیزیوگرافی نه‌گانه انجام شده است. در این نوع نقشه‌ها، واحد نقشه به‌صورت واحد همگن و دارای تنها یک کلاس خاک (سطوح مختلف رده‌بندی خاک از رده تا سری خاک) غالب است و معمولاً حدود تغییرات خاک در

واحد نقشه مشخص نشده و تنها در متن گزارش الحاقی به آن اشاره می‌گردد. در مواردی که مقیاس نقشه اجازه جدایی واحدهای فیزیوگرافی را نداده باشد واحدهای رده‌بندی به‌صورت اجتماع یا کمپلکس بدون درج سهم نسبی هر کدام از کلاس‌های خاک در داخل واحد نقشه و راهنمای آن نشان داده شده‌اند. دسته دوم نقشه‌های سنتی خاک مبتنی بر روش ژئوپدولوژیک می‌باشد که طی دو دهه اخیر در مؤسسه تحقیقات خاک و آب تهیه شده‌اند (۱۰). در این روش، واحدهای نقشه بر اساس تجزیه و تحلیل سیمای اراضی تفکیک می‌شوند و نامگذاری واحدهای نقشه بر اساس مفاهیم ارائه شده توسط گروه شناسایی و نقشه‌برداری خاک (۱۹۹۳) مانند همگون، اجتماع،

(۲۰۱۰) به منظور پهنه‌بندی کلاس‌های منفرد خاک، از قواعد حاکم بر روابط خاک و سیمای اراضی موجود در یک گزارش خاکشناسی استفاده نمودند (۲۴). پس از آن، این رویکرد با یک روش طبقه‌بندی نظارت شده توسط نیومان و تامسون (۲۰۱۲) تلفیق شد (۱۸). به دلیل ماهیت نیمه‌خودکار روش‌های مذکور، هر دوی این مطالعات انبوه‌ش‌زدایی نقشه خاک، زمان‌بر و حساس به نظر متخصص خاکشناس بود. برخی از پژوهشگران سعی کردند از مدل‌های کمی در انبوه‌ش‌زدایی نقشه خاک استفاده کنند و برای مدل‌سازی از برخی الگوریتم‌های داده‌کاوی استفاده نمایند. در این شیوه، ابتدا مدل آموزش داده می‌شود و سپس بر اساس برخی معیارهای حداقل‌سازی خطا، مدل بهینه می‌شود. سپس به منظور پهنه‌بندی کلاس‌های خاک موجود در واحدهای نقشه خاک انبوه‌ش‌زدایی شده، میان‌یابی بر اساس مدل برازش یافته اجرا می‌گردد. چنین روش‌های داده‌کاوی برای انبوه‌ش‌زدایی نقشه خاک توسط بوی و موران (۲۰۰۱) در استرالیا، هارینگ و همکاران (۲۰۱۲) در آلمان و سابوریلو (۲۰۱۴) در ایالات متحده آمریکا استفاده شده است (۲، ۱۲ و ۲۲). در روش‌های طبقه‌بندی نظارت شده، مدل‌های مورد استفاده در نقشه‌برداری رقومی خاک معمولاً با داده‌های مشاهداتی واسنجی می‌شوند اما موقعی که منابع مالی برای جمع‌آوری نقاط مشاهداتی جدید در صحرا محدود است. تهیه پایگاه داده‌های واسنجی مدل به وسیله نمونه‌برداری تصادفی از درون خود واحدهای نقشه خاک، گزینه جایگزین و جذابی است هر چند واضح و مبرهن است ویژگی‌ها و نوع خاک‌های که بدین شیوه از واحدهای نقشه خاک استخراج می‌گردند، جایگزین واقعی داده‌های مشاهداتی خاک نیستند. این شیوه نمونه‌برداری تصادفی از درون واحدهای نقشه، پیش از این توسط لاگچری و همکاران (۱۹۹۵)، گریناند و همکاران (۲۰۰۸) استفاده

کمپلکس و گروه‌های جدانشده صورت می‌گیرد (۲۳). در این نقشه‌ها، اطلاعات راجع به سهم نسبی کلاس‌های خاک در هر واحد نقشه وجود دارد. نقشه‌های سنتی خاک یا به عبارتی بهتر داده‌های میراثی خاک<sup>۱</sup>، تاکنون جوابگوی بسیاری از نیازهای عمرانی کشور در بخش‌های مختلف کشاورزی، منابع طبیعی و برنامه‌ریزی شهری بوده است (۱) اما با توجه به نقش کارکردهای خاک در حل مسائل جهان امروز مانند امنیت غذایی، امنیت آب و حفظ تنوع زیستی؛ نیاز به داده‌های دقیق و رقومی خاک با عدم قطعیت مشخص فزونی یافته است (۱۶). در این راستا، تولید نقشه‌هایی از خاک که توزیع مکانی منفرد کلاس‌های خاک را نشان دهد یکی از نیازهای کشور می‌باشد. در این نقشه‌ها، کلاس‌های خاک در هر واحد نقشه شناخته شده است و گاهی اوقات اطلاعات راجع به سهم نسبی هر کلاس هم وجود دارد. آنچه نامشخص است، آرایش و وضوح مکانی کلاس‌های خاک یاد شده در درون واحد نقشه می‌باشد. در مطالعات انبوه‌ش‌زدایی، بزرگترین مسأله، تعیین روش کمی تعیین آرایش مکانی کلاس‌های خاک در درون هر واحد نقشه است. در واقع، انبوه‌ش‌زدایی مکانی واحدهای نقشه خاک، روشی است که توزیع مکانی کلاس‌های منفرد خاک را مدل‌سازی می‌کند. طی این فرایند اطلاعات نقشه خاک از مقیاس کوچک (قدرت تفکیک مکانی درشت) به مقیاس بزرگ (قدرت تفکیک مکانی ریز) تبدیل می‌شود و برای اجرای آن روش‌های آماری و داده‌کاوی مورد استفاده قرار می‌گیرد. این یک موضوع رایج در به‌روزرسانی و تجدید نمودن نقشه‌های میراثی خاک می‌باشد. بعضی ایده‌های اساسی راجع به انبوه‌ش‌زدایی نقشه‌های خاک، در بحث عمیق‌تر مقیاس‌سازی اطلاعات خاک به وسیله مک‌برتنی (۱۹۹۸) صورت‌بندی و ارائه شده‌اند (۱۵). تامسون و همکاران

مطابق با آمار طولانی مدت هواشناسی، متوسط درجه حرارت سالانه هوا ۱۳/۲ درجه سانتی‌گراد و میانگین بارندگی سالانه ۴۸۱/۴ میلی‌متر است. اراضی مورد مطالعه بر روی رسوبات دوره کواترنری گرفته‌اند و سنگ آهک، دولومیت، مارن، سنگ‌رس و ماسه‌سنگ، سنگ‌های اصلی محدوده مطالعاتی را تشکیل داده‌اند. اراضی دیم (۱۳۱۰۲/۷۹ هکتار)، اراضی آبی (۲۷۲/۰۷ هکتار)، مرتع (۵۷۳/۲۳ هکتار) و اراضی مخلوط جنگل و مرتع (۱۳۳/۱۱ هکتار) مهم‌ترین کاربری‌های ناحیه مورد مطالعه می‌باشند (۱۰).

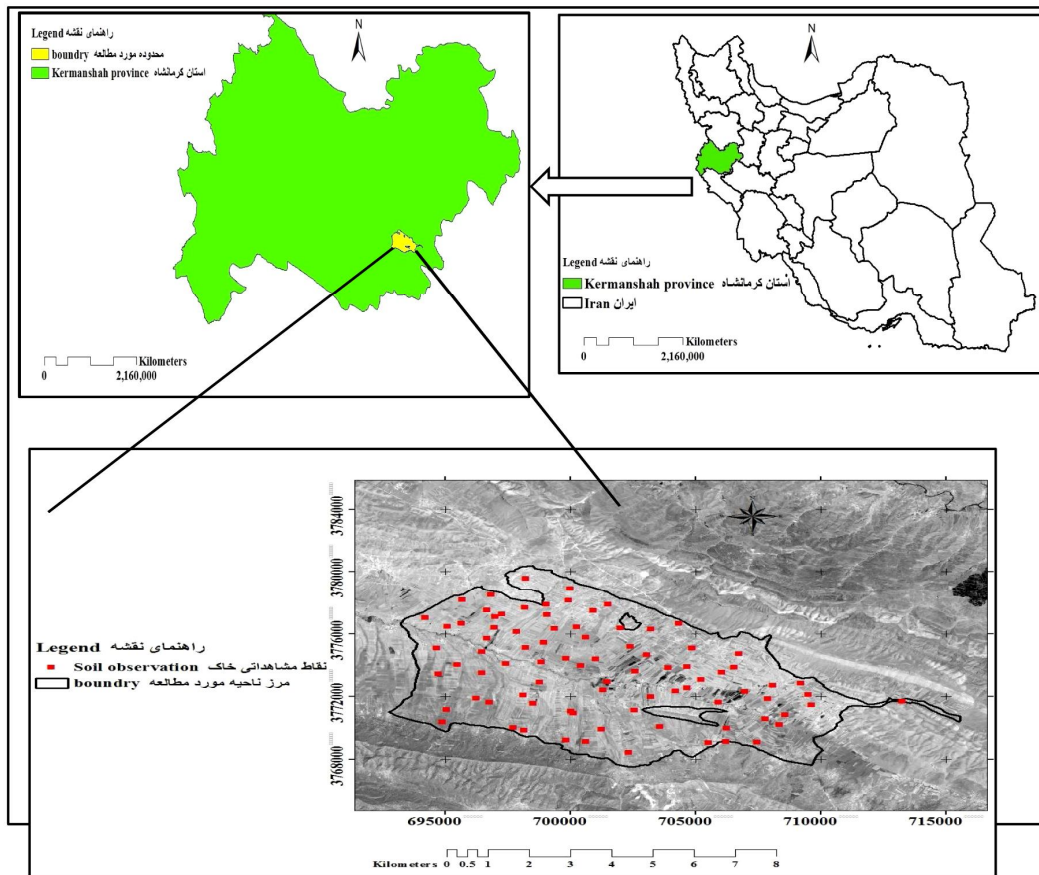
**داده‌های خاک:** نقشه خاک منطقه از ۵ واحد نقشه اجتماع تشکیل شده است (شکل ۲) و در مجموع واحدهای نقشه، مرکب از ۴ زیرگروه خاک می‌باشند (۱۰). چون نقشه خاک منطقه مورد مطالعه در مقیاس درشت‌تر تهیه شده است اگر وارد مدل شود امکان دارد بر تعمیم دادن تغییرپذیری کلاس‌های خاک در سیمای اراضی تأثیر بگذارد (۱۹). بنابراین، نقشه خاک نیمه‌تفصیلی منطقه مورد مطالعه با مقیاس کارتوگرافی ۱:۵۰۰۰۰، به‌عنوان یک متغیر کمکی در فرایند مدل‌سازی وارد نگردید و تنها از اطلاعات ترکیب نسبی کلاس‌های خاک در هر واحد نقشه به‌عنوان راهنمای نمونه‌برداری و تعیین داده‌های واسنجی در درون هر واحد نقشه استفاده شد (ستون ششم و هفتم جدول ۱). ستون اول جدول ۱ سیمای اراضی منطقه مورد مطالعه یعنی دشت دامنه‌ای و ستون دوم رلیف (پستی و بلندی محلی) منطقه را نشان می‌دهد که شامل دشت سر فوقانی، میانی و تحتانی می‌باشد. ستون سوم سنگ‌شناسی و رخساره زمین محدوده مورد مطالعه شامل سنگ‌های غالب رسی، مارنی و مواد آبرفتی را معرفی می‌کند، ستون چهارم بیانگر پنج نوع لندفرم است. ستون پنجم جدول، علائم واحد نقشه بر روی نقشه خاک را نشان می‌دهد (شکل ۱).

شده است (۹ و ۱۴). کولارد و همکاران (۲۰۱۴) و سابوریالو و همکاران (۲۰۱۴) این روش نمونه‌برداری درون‌نقشه‌ای را برای انبوهش‌زدایی نقشه‌های کلاس خاک دارای اطلاعات ترکیب و درصد اجزا واحد نقشه مورد استفاده قرار دادند (۷ و ۲۲).

وی و همکاران (۲۰۱۰) روش جدیدی برای انبوهش‌زدایی نقشه خاک معرفی نمودند در این روش به‌جای پیش‌بینی یک کلاس خاک منفرد، توزیع احتمال تمام کلاس‌های بالقوه خاک برآورد می‌شود (۲۷). اجرز و همکاران (۲۰۱۴) این الگوریتم را تصحیح و به‌منظور تهیه داده‌های واسنجی مدل، از اطلاعات سهم نسبی کلاس‌های خاک در درون هر واحد نقشه استفاده نمودند و آن را انبوهش‌زدایی یکنواسازی واحدهای نقشه خاک از طریق بازچینی درختان طبقه‌بندی (دسمارت) نام نهادند (۱۹). روش دسمارت به‌وسیله چانی و همکاران (۲۰۱۴)، هولمز و همکاران (۲۰۱۵) و وینسنت و همکاران (۲۰۱۶) برای انبوهش‌زدایی برخی پایگاه داده‌های جغرافیای مطالعه خاک به‌ترتیب در ایالات‌متحده آمریکا، استرالیا و فرانسه مورد استفاده قرار گرفت (۶، ۱۳ و ۲۶). در پژوهش حاضر، کارایی استفاده از الگوریتم دسمارت برای انبوهش‌زدایی نقشه نیمه‌تفصیلی کلاس خاک به مقیاس ۱:۵۰۰۰۰ (با ترکیب نسبی کلاس‌های خاک مشخص) در بخشی از حوضه آبخیز رودخانه کرخه واقع در استان کرمانشاه مورد بررسی قرار گرفته است.

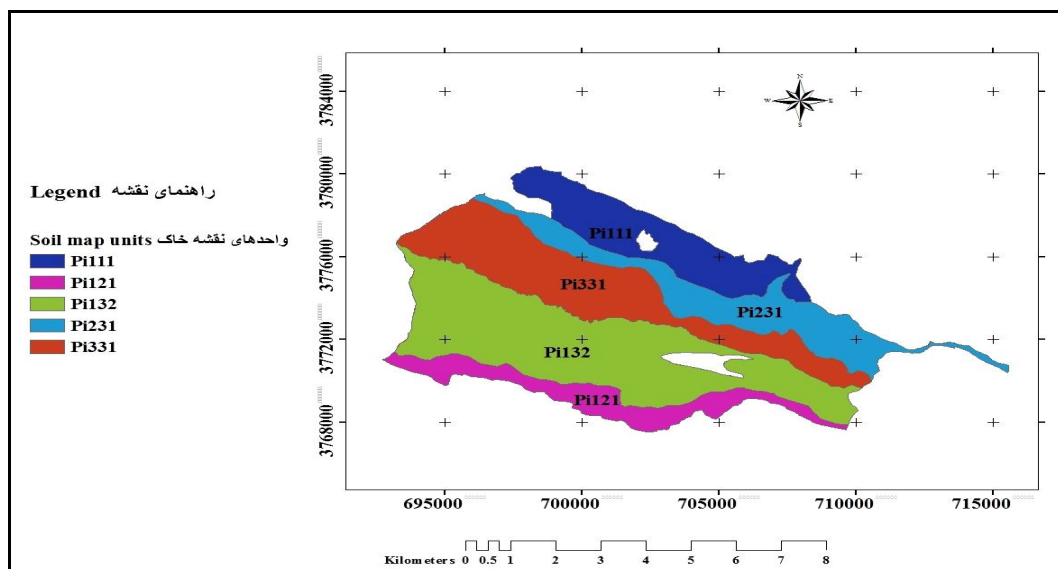
### مواد و روش‌ها

**موقعیت و وضعیت عمومی منطقه مطالعاتی:** ناحیه مورد مطالعه بخش کوچکی از زیرحوضه آبخیز کرخه بوده و با وسعتی حدود ۱۴۰۸۳/۹ هکتار در بیست کیلومتری جنوب‌شرقی شهر کرمانشاه، در بین مختصات جغرافیایی ۵' ۴۷° تا ۲۰' ۴۷° طول شرقی و ۲' ۳۴° تا ۸' ۳۴° عرض شمالی واقع شده است (شکل ۱).



شکل ۱- موقعیت منطقه مورد مطالعه در ایران و استان کرمانشاه و نقاط مشاهداتی.

Figure 1. Location of the study area in Iran and the Kermanshah province and observation points.



شکل ۲- نقشه میراثی خاک منطقه مورد مطالعه.

Figure 2. Legacy soil map of the study area.

جدول ۱- نسبت وقوع کلاس‌های خاک (زیرگروه‌های خاک) در هر واحد نقشه.

Table 1. Proportions of occurrence of the soil classes (soil subgroups) in each map unit.

سبب‌های اراضی Landscape	ریف Relief	لیتولوژی/رخساره Lithology/Facies	لاندفرم Landform	واحد نقشه Map unit	کلاس خاک Soil class	مشاهدات خاک Soil observation	سهم نسبی Proportion (%)	مساحت (هکتار) Area (ha)
Piedmonts	کلاسی میانی middle glacia	Alluvium	Slope facet complex	P1231	Typic Calcixerepts	5	60	2181
					Fluventic Haploxerepts	2	40	
	کلاسی فوقانی High glacia	Claystone, sandstone, conglomerate	Upper upper glacia	P1121	Vertic Calcixerepts	3	37.5	1401.1
					Fluventic Haploxerepts	2	25	
	کلاسی فوقانی High glacia	Marl with intercalatin of argillaceous, limestone	Slope facet complex	P1111	Typic Calcixerepts	7	70	2133.4
					Vertic Calcixerepts	3	30	
	کلاسی فوقانی High glacia	Alluvium	Middle upper glacia	P1132	Vertic Calcixerepts	18	72	5035
					Fluventic Haploxerepts	7	28	
	کلاسی تحتانی Low glacia	Alluvium	Slope facet complex	P1331	Typic Calcixerepts	15	60	3333.2
					Fluventic Haploxerepts	8	32	
					Chromic Calcixerepts	2	8	
Sum جمع								14083.9

این شاخص‌ها دارای اندازه پیکسل ۳۰ متر هستند و با استفاده از دستور بازچینی<sup>۲</sup> به نقشه‌های با اندازه پیکسل ۱۰ متر تبدیل شدند. لازم به ذکر است متغیرهای محیطی که در پژوهش حاضر انتخاب شده‌اند، پیش از این در پژوهش فاتحی و همکاران (۱۳۹۴) در فرآیند انبوه‌ش‌زدایی کلاس‌های خاک با روش‌های داده‌کاوی لاجستیک رگرسیون چند کلاسه و الگوریتم درختان تصمیم C5.0 در همین منطقه مورد مطالعه اهمیتشان به اثبات رسیده است (۱۱). در جدول ۳ خلاصه‌ای از خصوصیات متغیرهای کمکی مذکور ارائه شده است.

**الگوریتم دسمارت:** آجرز و همکاران (۲۰۱۴) به تفصیل الگوریتم دسمارت را شرح داده‌اند هدف از دسمارت، پیش‌بینی توزیع مکانی کلاس‌های خاک به‌وسیله انبوه‌ش‌زدایی واحدهای نقشه ستی خاک است. واحد نقشه خاک شامل ترکیبی از کلاس‌های خاک در یک الگوی مکانی معین با سهم نسبی مشخص می‌باشد (۱۹). روش دسمارت نمایش توزیع کلاس‌های خاک انبوه‌ش‌زدایی شده به‌صورت نقشه سطوح احتمال<sup>۳</sup> رستری است. نمایش رستری هر کلاس خاک به‌صورت احتمال وقوع آن صورت می‌گیرد. به‌منظور تهیه نقشه سطوح احتمال در مرحله اول از روش نمونه‌برداری تصادفی مکرر در درون هر واحد نقشه استفاده می‌شود سپس n موجودیت<sup>۴</sup> یا به‌عبارت بهتر، n مدل برازش‌یافته درختان طبقه‌بندی الگوریتم C5.0 ایجاد می‌گردد. هر مدل درخت طبقه‌بندی توسط داده‌های حاصل از m بار نمونه‌برداری تصادفی از درون هر پلی‌گون واسنجی می‌شود. به‌طور تصادفی به هر نقطه نمونه‌برداری، یک کلاس خاک (از میان کلاس‌های خاک تعریف شده در هر واحد نقشه) اختصاص داده می‌شود و از سهم نسبی کلاس‌های خاک درون واحد نقشه برای وزن‌دهی استفاده می‌گردد. اختصاص کلاس خاک به هر نقطه نمونه‌برداری لازم

**جمع‌آوری متغیرهای محیطی و انتخاب متغیرهای کمکی:** بر اساس معادله اسکورپن، متغیرهای کمکی که معرف عوامل خاک‌سازی غالب در منطقه یعنی توپوگرافی، مواد مادری و موجودات زنده (پوشش گیاهی) و اقلیم بودند تهیه و مورد استفاده قرار گرفت. در این پژوهش، با استفاده از مدل‌های رقومی ارتفاع با قدرت تفکیک ۱۰ متر، مشتقات مدل رقومی ارتفاع شامل ارتفاع، درصد شیب، جهت شیب، تحدب، تابش مستقیم، شاخص رسوب و شاخص همواری دره با درجه تفکیک بالا<sup>۱</sup>، شاخص خیسی، عمق دره و فاصله عمودی تا شبکه آبراهه با استفاده از نرم‌افزارهای ArcGIS 10.1، ILWIS3.8 و SAGA2.2 تهیه گردید. علاوه بر این، نقشه موضوعی سنگ‌شناسی که معرف عامل خاک‌سازی مواد مادری است پس از رقومی‌سازی در فرآیند مدل‌سازی مورد استفاده قرار گرفت. نقشه رقومی زمین‌شناسی از نقشه زمین‌شناسی ۱:۱۰۰۰۰۰ شیت کرمانشاه تهیه گردید. از باندهای ۱ تا ۷ تصاویر چندطیفی ماهواره‌ای لندست 7 ETM<sup>+</sup> نیز برای تولید شاخص پوشش گیاهی نرمال شده (روز و همکاران، ۱۹۷۳)، شاخص رس (واندونتر، ۱۹۹۷) و شاخص اندازه ذرات (زیائو و همکاران، ۲۰۰۶) به‌عنوان یک متغیر کمکی در عملیات مدل‌سازی استفاده شد (۲۰، ۲۵ و ۲۸). معادله‌های تهیه این شاخص‌ها به‌شرح زیر است:

الف) شاخص پوشش گیاهی نرمال شده

$$NDVI = \frac{\text{band4} - \text{band3}}{\text{band4} + \text{band3}} \quad (1)$$

ب) شاخص رس

$$\text{Clay index} = \frac{\text{band7}}{\text{band5}} \quad (2)$$

ج) شاخص اندازه ذرات

$$\text{Grain size index} = \frac{\text{band3} - \text{band1}}{\text{band1} + \text{band2} + \text{band3}} \quad (3)$$

2- Resampling  
3- Probability surfaces  
4- realization

1- Multi-resolution index of valley bottom flatness

آن‌ها در جدول ۳ آمده به‌عنوان ورودی‌های مدل دسمارت به‌کار رفت. تعداد نمونه تصادفی در هر پلی‌گون (m) ۱۵، تعداد مدل برازش‌یافته (n) ۱۰ و تعداد گره‌های محاسباتی برای تجزیه تحلیل ۴ در نظر گرفته شد. به بیانی دیگر برای اجرای فرایند انبوهش‌زدایی نقشه کلاس خاک از پارامترهای پیش فرض نرم‌افزار استفاده گردید.

**شاخص آسیمگی<sup>۱</sup>:** در مدل دسمارت، نقشه سطوح احتمال توزیع کلاس‌های خاک از n بار اجرای مدل درختان تصمیم تولید می‌شود. در هر پیکسل تعداد دفعاتی که یک کلاس خاک پیش‌بینی می‌شود، شمارش می‌شود. احتمال وقوع هر کلاس خاک از تقسیم فراوانی وقوع آن بر n به‌دست می‌آید. بنابراین یک سری نقشه احتمال کلاس‌های خاک مرتبه اول، دوم، سوم تا مرتبه n ام قابل تولید است. برو و همکاران (۱۹۹۷) شاخص آسیمگی بین احتمالات وقوع کلاس‌های خاک به‌صورت معادله زیر صورت‌بندی نمودند (۵):

$$CI = 1 - (P_{\max} - P_{\max-1}) \quad (4)$$

که در آن،  $P_{\max}$  نقشه احتمال کلاس خاک مرتبه اول و  $P_{\max-1}$  نقشه احتمال کلاس خاک مرتبه دوم است (۵). اگر مقدار شاخص آسیمگی نزدیک به صفر باشد یک کلاس خاک غالب است و اگر شاخص نزدیک به یک باشد نشان‌دهنده این است که کلاس خاک غالب وجود ندارد و کلاس‌های خاک در هر دو نقشه با احتمال نزدیک به هم ممکن است وجود داشته باشند (۵).

**اعتبارسنجی:** داده‌های مکانی پیش‌بینی شده از نوع کلاس‌های خاک با استفاده از ماتریس درهمی شامل صحت تولیدکننده<sup>۲</sup>، صحت کاربر<sup>۳</sup>، صحت کلی<sup>۴</sup> و ضریب کاپا<sup>۵</sup> اعتبارسنجی می‌شوند (۴).

است زیرا کلاس خاک واقعی در هر نقطه نمونه‌برداری ناشناخته است (۱۹). اجرای الگوریتم دسمارت را می‌توان در ۶ مرحله زیر خلاصه نمود (۱۷):

(۱) استخراج n نمونه تصادفی از هر پلی‌گون نقشه خاک.  
(۲) اختصاص کلاس‌های خاک و وزن تصادفی به هر نقطه نمونه‌برداری.

در اینجا سهم نسبی کلاس‌های خاک در داخل هر واحد نقشه به‌عنوان وزن در نظر گرفته می‌شود.

(۳) از نقاط نمونه‌برداری شده و مقدار متغیرهای کمکی مرتبط به آن‌ها برای ایجاد مدل درخت طبقه‌بندی به‌منظور پیش‌بینی توزیع مکانی کلاس‌های خاک استفاده می‌شود.

(۴) مدل برازش‌یافته، برای پیش‌بینی احتمال وقوع کلاس‌های خاک در گستره ناحیه مورد مطالعه با استفاده از لایه‌های متغیرهای کمکی به‌کار می‌رود.

(۵) گام ۱ تا ۴، i بار، برای تولید i موجودیت (مدل برازش‌یافته) از توزیع کلاس‌های خاک تکرار می‌گردد.  
(۶) با استفاده از i موجودیت به‌دست آمده، نقشه سطوح احتمال برای کلاس‌های خاک تولید می‌شود.

نوع مدلی که اجرز و همکاران (۲۰۱۴) استفاده نمودند الگوریتم درختان تصمیم C5.0 بود که توسط کوینلان (۲۰۰۴) معرفی شده است (۲۱). الگوریتم دسمارت همچنین در یک بسته نرم‌افزاری به همین نام در محیط R قابل‌دسترس می‌باشد (۱۷). بسته نرم‌افزاری dsmart به سه منبع اصلی از داده‌ها، نیاز دارد:

۱- پلی‌گون‌های واحد نقشه خاک که انبوهش‌زدایی می‌شوند.  
۲- اطلاعاتی درباره سهم نسبی کلاس‌های خاک در هر واحد نقشه خاک.

۳- متغیرهای کمکی رستری زمین مرجع شده به نمایندگی فاکتورهای اسکورپن که گستره ناحیه مورد مطالعه را به‌طور کامل و پیوسته پوشش دهند.

در این الگوریتم محدودیتی در استفاده از انواع داده‌های پیوسته یا رسته‌ای وجود ندارد. در پژوهش حاضر ۱۳ متغیر کمکی که توصیف و خلاصه آماری

- 1- Confusion index
- 2- Producer's Accuracy
- 3- User's Accuracy
- 4- Overall accuracy
- 5- Kappa coefficient



است (۸ و ۹). محاسبه بر مبنای اختلاف بین توافق موجود (توافق مشاهده شده) و توافق مورد انتظار (توافق پیش‌بینی شده) است. توافق مشاهده شده همان درصد صحت کلی است. علاوه بر این لازم است مشخص گردد میزان تفاوت بین توافق مشاهده شده و پیش‌بینی شده چقدر است. ضریب کاپا معیاری از این تفاوت است و مقدار آن بین -۱ تا +۱ قرار می‌گیرد. عدد یک توافق کامل، صفر، آنچه که واقعا با شانس انتظار می‌رود و مقادیر منفی، توافق کمتر از شانس (یعنی عدم توافق سیستماتیک بین مشاهدات و پیش‌بینی‌ها) را نشان می‌دهند. ضریب کاپا از معادله زیر به دست می‌آید (۹):

$$K = \frac{P_o - P_e}{1 - P_e} \quad (8)$$

که در آن،  $P_o$  صحت مشاهده شده یا صحت کلی است و  $P_e$  صحت مورد انتظار است که از رابطه زیر به دست می‌آید:

$$P_e = \sum_{i=1}^n \frac{\text{colsum}_i}{TO} \frac{\text{rowsum}_i}{TO} \quad (9)$$

که در آن،  $\text{colsum}_i$  جمع ستون  $i$  در جدول ماتریس درهمی،  $\text{rowsum}_i$  جمع ردیف  $i$  در جدول ماتریس درهمی،  $TO$  تعداد مشاهدات و  $n$  تعداد کلاس‌ها.

در این پژوهش، اطلاعات طبقه‌بندی خاک ۸۲ پروفیل به‌عنوان داده‌های اعتبارسنجی مورد استفاده قرار گرفتند (جدول ۲).

صحت تولیدکننده، احتمال این که در یک پیکسل بر روی نقشه، کلاس خاک مربوط به آن درست پیش‌بینی شده باشد را نشان می‌دهد.

$$PA_j = \frac{X_{jj}}{\sum_{i=1}^r X_{ij}} \quad (5)$$

صحت کاربر، میانگین احتمالی را بیان می‌کند که کلاس خاک پیش‌بینی شده مربوط به یک پیکسل بر روی نقشه مطابق با کلاس خاکی باشد که در واقعیت زمینی وجود دارد.

$$UA_j = \frac{X_{jj}}{\sum_{i=1}^r X_{ij}} \quad (6)$$

صحت کلی نقشه ارتباط بین همه داده‌های مورد استفاده و داده‌های صحیح طبقه‌بندی شده را نشان می‌دهد و در واقع میانگینی از پیکسل‌هایی است که به‌طور صحیح پیش‌بینی شده‌اند:

$$OA = \frac{\sum_{i=1}^C E_{ij}}{N} \quad (7)$$

که در آن‌ها،  $X_{ij}$  عنصر قطری از هر کلاس خاک در یک ردیف ماتریس درهمی،  $X_{ij}$  عنصر قطری از هر کلاس خاک در یک ستون ماتریس درهمی،  $E_{ij}$  مجموع عناصر قطری در ماتریس درهمی،  $C$  تعداد کلاس‌های خاک و  $N$  تعداد مشاهدات را نشان می‌دهند. یک روش خوب دارای مقادیر صحت کلی، صحت کاربر و تولیدکننده نقشه نزدیک به یک می‌باشد. ضریب کاپا یک معیار آماری دیگر برای صحت بین مشاهدات و پیش‌بینی‌های یک مدل طبقه‌بندی‌کننده

جدول ۲- زیرگروه‌های خاک در پایگاه داده‌های خاک.

Table 2. Soil subgroups in soil database.

زیرگروه Subgroup	تعداد Number
Chromic Calcixererts	8
Fluventic Haploxerepts	11
Typic Calcixerepts	54
Vertic Calcixerepts	9
جمع	82

جدول ۳- متغیرهای کمکی استفاده شده در مدل.  
Table 3. Used auxiliary variables in model.

تعداد کلاس No. class	خلاصه آماری متغیرهای کمکی Summary statistics of covariates		شرح مختصر Brief description	قدرت تفکیک مکانی (متر) / مقیاس Resolution (m) / scale	عامل اسکورپان Scorpan factor	متغیر کمکی Auxiliary variables
	میانگین mean	حداکثر maximum				
-	4202	4305	مدت تابش (ساعت) Duration of direct incoming solar radiation (hour)	10	اقلیم Climate	مدت تابش مستقیم (Direct Duration)
-	1562	1877	ارتفاع (متر) سطح اراضی نسبت به سطح دریا Elevation of the sea surface (m)	10		ارتفاع (Elevation)
-	0.2	0.5	تحدب سطح اراضی Convexity of the land surface	10		تحدب (Convexity)
-	2.4	5.9	نشان‌دهنده نواحی رسوبگذاری Identifies the depositional areas	10		شاخص همواری دره با درجه تفکیک بالا (MRVBF)
-	0	0.8	حداکثر میزان تغییر ارتفاع بین پیکسل‌ها و پیکسل‌ها مجاور (رادین) Maximum rate of change between the cells and neighbors (radian)	10	توپوگرافی Relief	شیب (Slope)
-	15.7	29244	محاسبات برای اثر توپوگرافی بر فرسایش Accounts for the effect of topography on erosion	10		شاخص رسوب (Sediment Index)
-	10	24.4	معیار کنترل توپوگرافی بر وضعیت رطوبتی خاک A criterion for controlling topography on soil moisture status	10		شاخص خیسگی توپوگرافی (Topographic Wetness Index)
-	5.3	103.4	موقعیت نسبی دره Relative position of the valley	10		عمق دره (Valley Depth)
-	1.9	59	بیانگر فاصله عمودی تا شبکه آبراه‌ها برای هر پیکسل Expresses Vertical distance to channel network for each cell	10		فاصله عمودی تا شبکه آبراهه (Vertical Distance To Channel Network)
-	10.3	17.1	نسبت تفاوت باند ۴ و ۳ به مجموع این دو باند The ratio of the difference between band4 and band3 to the sum of band4 and band3	30	موجودات زنده یا پوشش گیاهی Organisms or vegetation	شاخص پوشش گیاهی نرمال‌شده (NDVI)
-	9.8	14.9	نسبت تفاوت باند ۱، ۲ و ۳ به مجموع باند ۱، ۲ و ۳ The ratio of the difference between band1 and band2 to the sum of band1, band2 and band3	30		شاخص اندازه ذرات (Grain Size Index)
-	19.7	33.6	نسبت باند لندست ۵ به باند ۷ Landsat band ratio 5/7	30	مواد مادری Parent material	شاخص رس (Clay Index)
6	-	-	نقشه زمین‌شناسی اسکن و زمین مرجع شده Scanned and registered geological map	1:100000		نقشه زمین‌شناسی (Geology Map)

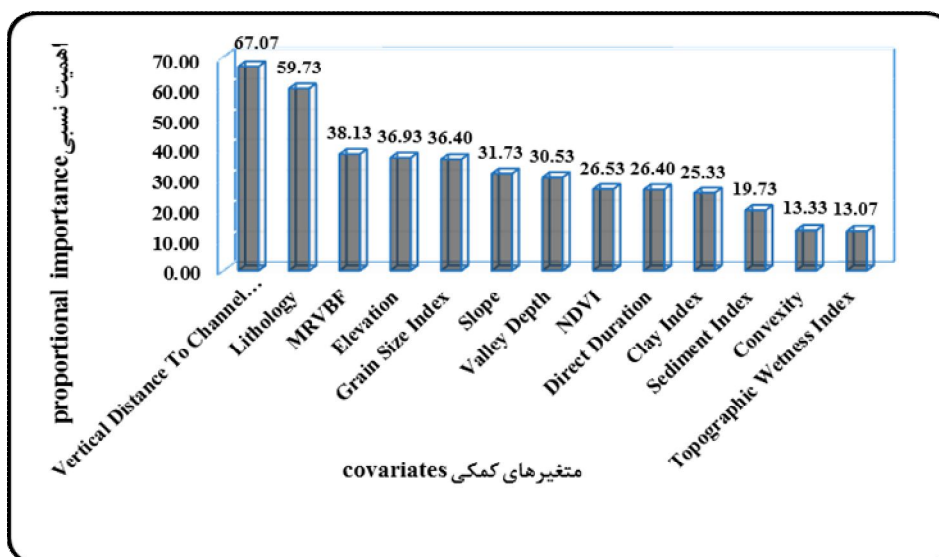
## نتایج و بحث

همان‌طور که در جدول ۳ نشان داده شده است مقدار هدایت الکتریکی عصاره گل اشباع از ۰/۳۴ تا ۱/۶۶ دسی‌زیمنس بر متر، واکنش خاک از ۷/۳ تا ۸/۱ کربن آلی خاک از ۰/۴ تا ۳/۲ درصد، مقدار رس خاک از ۳۰/۸ تا ۵۴/۸ درصد و مقدار کربنات کلسیم از ۶ تا ۴۶/۸ درصد در خاک‌های ناحیه مورد مطالعه متغیر است که نشان می‌دهد خاک‌های منطقه مورد مطالعه فاقد شوری و دارای رس و آهک زیاد هستند. شکل ۳، میانگین اهمیت نسبی متغیرهای کمکی (ورودی‌های مدل) را نشان می‌دهد. این نتایج بیان می‌کند که پارامترهای توپوگرافی شامل فاصله عمودی تا شبکه آبراهه، ارتفاع و شاخص همواری دره با درجه تفکیک بالا به‌عنوان عوامل مؤثر بر جریان آبشویی، فرسایش خاک، رسوبگذاری و لیتولوژی و شاخص اندازه ذرات معرف نوع و اندازه مواد مادری خاک، از عوامل مهم و تعیین‌کننده الگوی قرارگیری انواع خاک بر روی سیمای اراضی ناحیه مورد مطالعه هستند. نقشه احتمال کلاس خاک مرتبه اول، نقشه احتمال کلاس خاک مرتبه دوم و نقشه احتمال کلاس خاک مرتبه سوم به‌ترتیب در شکل‌های ۴، ۵ و ۶ نمایش داده شده‌اند. این اشکال نشان می‌دهند اگرچه مدل دسمارت داده‌های واسنجی را بر اساس نمونه‌برداری تصادفی از درون هر واحد نقشه به‌دست می‌آورد اما پیش‌بینی زیرگروه‌های خاک را برای کل ناحیه مورد مطالعه انجام می‌دهد. به این معنی که در واحد نقشه Pi111 مطابق جدول ۱ تنها دو زیرگروه تیپیک کلسی زریپتز و ورتیک کلاس زریپتز مشاهده می‌شود اما در هنگام انبوهش‌زدایی دو زیرگروه دیگر نیز در محدوده این واحد نقشه نیز پیش‌بینی شده است (شکل‌های ۲ و ۴). همچنین بررسی این نقشه‌ها نشان می‌دهد که در نقشه احتمال کلاس خاک مرتبه اول، زیرگروه‌های خاک غالب در منطقه یعنی تیپیک کلسی‌زریپتز بیش‌ترین مساحت خاک منطقه را تشکیل می‌دهد اما

زیرگروه فرعی کرومیک کلسی زریپتز که اساساً سهم کم‌تری در واحدهای نقشه سستی دارد (جدول ۱) وسعت کم‌تری از اراضی منطقه را در نقشه حاصل از مدل پیش‌بینی دسمارت به خود اختصاص داده است (جدول ۳ و شکل ۴). این وضعیت در نقشه احتمال کلاس خاک مرتبه سوم (شکل‌های ۵ و ۶) برعکس شده و خاک فرعی مانند کرومیک کلسی‌زریپتز وسعت بیش‌تری از نقشه پیش‌بینی را به خود اختصاص می‌دهد (جدول ۳). شاخص آسیمگی بین نقشه احتمال کلاس خاک مرتبه اول و دوم در شکل ۷ ارائه شده است. شاخص‌های آسیمگی نزدیک به صفر (کم‌تر از ۰/۲) به‌صورت پراکنده و لکه‌های کم وسعت در نواحی شمال، شمال‌شرقی و جنوب‌شرقی و برخی نواحی مرکزی نقشه دیده می‌شود. این وضعیت، نشان می‌دهد که احتمال وقوع یک کلاس خاک غالب در یک سلول شبکه (پیکسل) نقشه کم است. اما شاخص آسیمگی نزدیک به یک (بیش از ۰/۸) دارای توزیع گسترده‌ای در ناحیه مورد مطالعه می‌باشد که این وضعیت، توزیع چند زیرگروه خاک را با احتمال وقوع نزدیک به هم در هر پیکسل نقشه احتمال کلاس خاک مرتبه اول و دوم را نشان می‌دهد. ارزیابی نقشه‌های سطوح احتمال با استفاده از داده‌های اعتبارسنجی نشان می‌دهد که صحت کلی نقشه احتمال کلاس خاک مرتبه اول (۴۴ درصد)، نقشه احتمال کلاس خاک مرتبه دوم (۲۸ درصد) و نقشه احتمال کلاس خاک مرتبه سوم (۱۱ درصد) می‌باشد (شکل ۸). اجرز و همکاران (۲۰۱۴) در فرایند انبوهش‌زدایی ۳۰۸۵ پلی‌گون از واحدهای نقشه خاک مختلف با استفاده از روش دسمارت صحت کلی برای نقشه‌های احتمال کلاس خاک مرتبه‌های اول، دوم و سوم را به‌ترتیب ۲۲/۵ درصد، ۱۶/۸ درصد و ۹/۱ درصد به‌دست آوردند (۱۹). نتایج ارزیابی انبوهش‌زدایی ۲۷۳۷۶ کیلومترمربع از منطقه بریتانی در شمال‌غرب فرانسه توسط وینسنت و همکاران (۲۰۱۶) نیز نشان

کلسی‌زرپتیز) دارای صحت کاربر و تولیدکننده نقشه بیش‌تری است. همان‌طور که در جدول ۳ مشاهده می‌شود در نقشه احتمال کلاس خاک مرتبه اول وسعت خاک‌های غالب تیپیک کلسی‌زرپتیز و ورتیک کلسی‌زرپتیز بالا است. اما وسعت خاک فرعی پیش‌بینی‌شده در نقشه احتمال کلاس خاک مرتبه سوم، افزایش یافته و وسعت خاک غالب منطقه یعنی تیپیک کلسی‌زرپتیز کاهش یافته است. در نقشه احتمال کلاس خاک مرتبه دوم، خاک‌های غالب منطقه (تیپیک کلسی‌زرپتیز، ورتیک کلسی‌زرپتیز و فلونیتیک هاپلوزرپتیز) به وسعت تقریباً یکسانی پیش‌بینی شده‌اند این افزایش وسعت و پراکنش خاک‌های فرعی در نقشه احتمال دوم و سوم باعث افزایش صحت کاربر و تولیدکننده نقشه در آن‌ها شده است. به‌نظر می‌رسد سهم بیش‌تر خاک‌های فرعی در نقشه احتمال کلاس خاک مرتبه دوم و سوم ناشی از نمونه‌برداری تصادفی و وزن‌دهی به کلاس‌های خاک در روش دسمارت باشد. با این وجود، صحت کلی نقشه احتمال کلاس خاک مرتبه اول در سطح زیرگروه در طیف ۴۰ تا ۷۰ درصد قرار دارد که سایر پژوهشگران به‌دست آورده‌اند (۷، ۱۲ و ۱۸).

از صحت کلی ۲۳/۹، ۱۶/۷ و ۹/۶ درصد به‌ترتیب برای اولین، دومین و سومین سطح احتمال داشت (۲۷). هولمز و همکاران (۲۰۱۵) نیز از الگوریتم دسمارت برای انبوهش‌زدایی ۲/۵۰۰/۰۰۰ کیلومترمربع از اراضی استرالیای غربی استفاده کردند و برای نقشه احتمال کلاس خاک مرتبه اول، صحت کلی را ۴۰ درصد برآورد نمودند (۱۳). مقایسه نتایج پژوهش حاضر با پژوهش‌های پیشین نشان می‌دهد که در همه این پژوهش‌ها صحت کلی از اولین سطح احتمال تا سومین سطح احتمال کلاس‌های خاک کاهش می‌یابد. به‌نظر می‌رسد یکی از دلایل بالا بودن صحت کلی سطوح احتمال پیش‌بینی‌شده در پژوهش حاضر نسبت به پژوهش‌های سایر پژوهشگران، کم بودن تعداد واحدهای نقشه و کلاس‌های خاک است. صحت کاربر (شکل ۹) و صحت تولیدکننده نقشه (شکل ۱۰) نشان می‌دهد که در نقشه احتمال کلاس خاک مرتبه اول، زیرگروه غالب خاک (تیپیک کلسی‌زرپتیز) بالاترین صحت کاربر و تولیدکننده نقشه را دارد و در نقشه احتمال کلاس خاک مرتبه دوم و سوم، زیرگروه خاک فرعی در منطقه مورد مطالعه (ورتیک



شکل ۳- میانگین اهمیت نسبی متغیرهای کمکی در ده بار برازش مدل C5.0.

Figure 3. Mean of relative importance auxiliary variables in ten times C5.0 model fitting.

جدول ۴- پارامترهای آماری برخی ویژگی‌های خاک در نمونه‌های اعتبارسنجی.

Table 4. Statistical parameters of some soil properties in validation samples.

پارامترهای آماری Statistical Parameters								ویژگی‌های خاک (واحد) Soil properties (unit)
کشیدگی Kurtosis	چولگی Skewness	ضریب تغییرات Coefficient of variation	واریانس Variance	دامنه Range	حداکثر Maximum	حداقل Minimum	میانگین Mean	
2.18	1.7	51.96	0.13	1.32	1.66	0.34	0.68	هدایت الکتریکی خاک (دسی‌زیمنس بر متر) EC (dsm <sup>-1</sup> )
0.58	0.42	2.24	0.03	0.79	8.07	7.28	7.7	واکنش خاک PH (-log(H <sup>+</sup> ))
5.01	1.86	57.46	50.22	36.6	39.2	2.6	12.33	شن (%) Sand (%)
1.41	-7.4	10.22	20.04	21.4	51.4	30	43.78	سیلت (%) Silt (%)
-0.2	-0.45	12.00	27.75	21.6	52.4	30.8	43.89	رس (%) Clay (%)
8.19	2.12	39.81	0.23	2.78	3.16	0.38	1.2	کربن آلی (%) Organic carbon
1.22	-0.58	27.54	70.04	40.8	46.8	6	30.39	کربنات کلسیم معادل (%) Equivalent carbonate calcium (%)

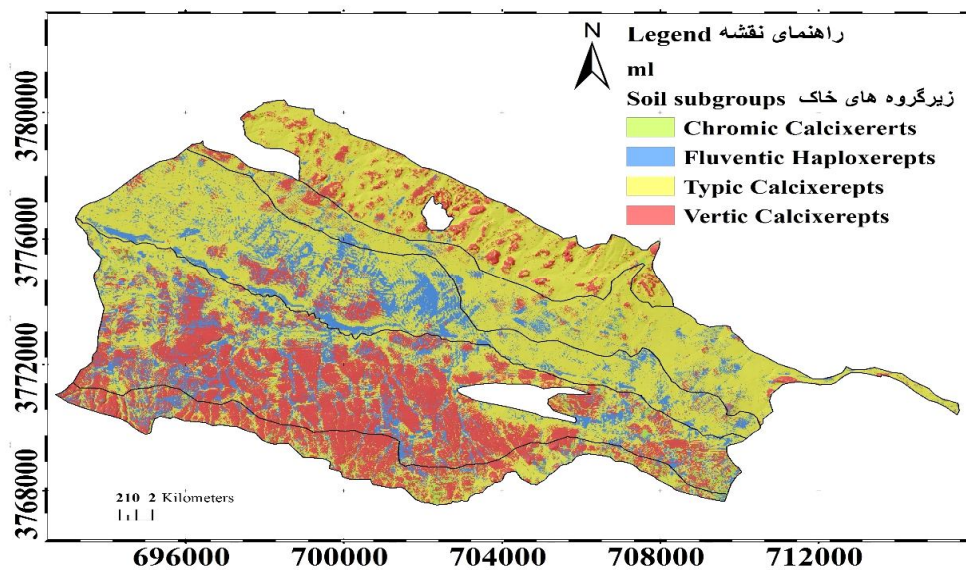
کلی است و بیانگر عدم توافق بین مشاهدات و پیش‌بینی می‌باشد. ضریب کاپای پایین در این مطالعه را می‌توان ناشی از بالا بودن وسعت زیرگروه خاک غالب در منطقه (تپیک کلسی‌زرپتز) هم در نقشه سنتی و هم نقشه پیش‌بینی حاصل از مدل دسمارت و همچنین بالان بودن تعداد مشاهدات واقعی مربوط به این زیرگروه در داده‌های اعتبارسنجی دانست (۳). بنابراین پایین بودن شاخص کاپا را نمی‌توان به‌عنوان دلیلی بر پایین بودن کارایی مدل دسمارت در نظر گرفت.

ضرایب کاپا برای نقشه‌های سطوح احتمال خاک مرتبه اول و دوم به ترتیب ۰/۰۴، ۰/۰۲ به دست آمد (شکل ۸). این نتایج نشان‌دهنده اختلاف کم بین صحت کل و صحت مورد انتظار است. از آنجا که مقادیر کاپا در نقشه‌های احتمال کلاس مرتبه اول و دوم نزدیک به صفر است می‌توان گفت که عدم توافق بین مشاهدات واقعی و پیش‌بینی ناشی از شانس می‌باشد. در نقشه سطوح احتمال مرتبه سوم مقدار ضریب کاپا کم‌تر از صفر (۰/۰۸-) برآورد شد. که به معنای بالا بودن صحت مورد انتظار نسبت به صحت

جدول ۵- مساحت زیرگروه‌های خاک در نقشه‌های سطوح احتمال.

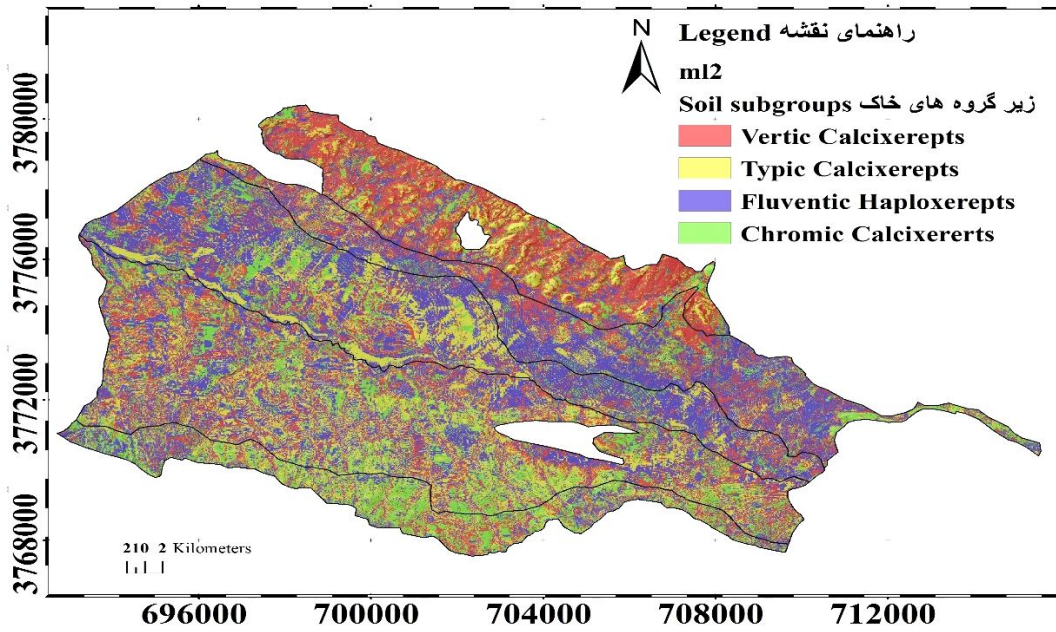
Table 5. Area of soil subgroups in probability surface maps.

زیرگروه‌های خاک Soil subgroups	مساحت (هکتار) Area (ha)		
	سطوح احتمال probability surface		
	نقشه احتمال کلاس خاک مرتبه اول the most probable soil class	نقشه احتمال کلاس خاک مرتبه دوم second probable soil class	نقشه احتمال کلاس خاک مرتبه سوم third probable soil class
Chromic Calcixererts	403.92	2299.3	5715.65
Fluventic Haploxerepts	2138.89	4679.24	3356.96
Typic Calcixerepts	7757.15	3243.92	1717.84
Vertic Calcixerepts	3783.94	3861.44	3293.45
جمع sum	14083.9	14083.9	14083.9

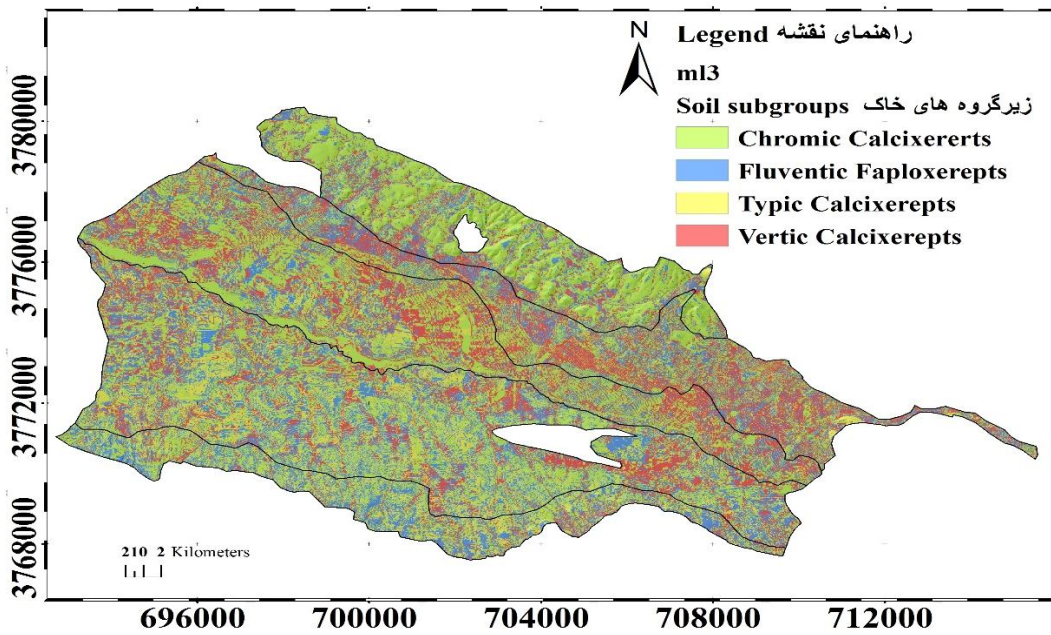


شکل ۴- نقشه احتمال کلاس خاک مرتبه اول.

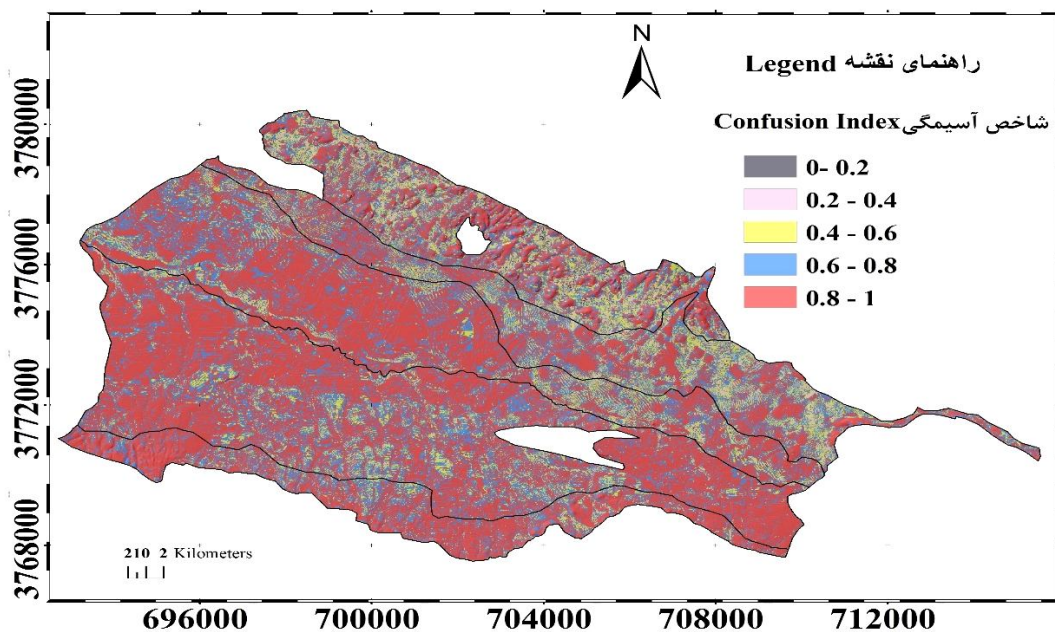
Figure 4. Most probable soil class map.



شکل ۵- نقشه احتمال کلاس خاک مرتبه دوم.  
 Figure 5. Second most probable soil class map.

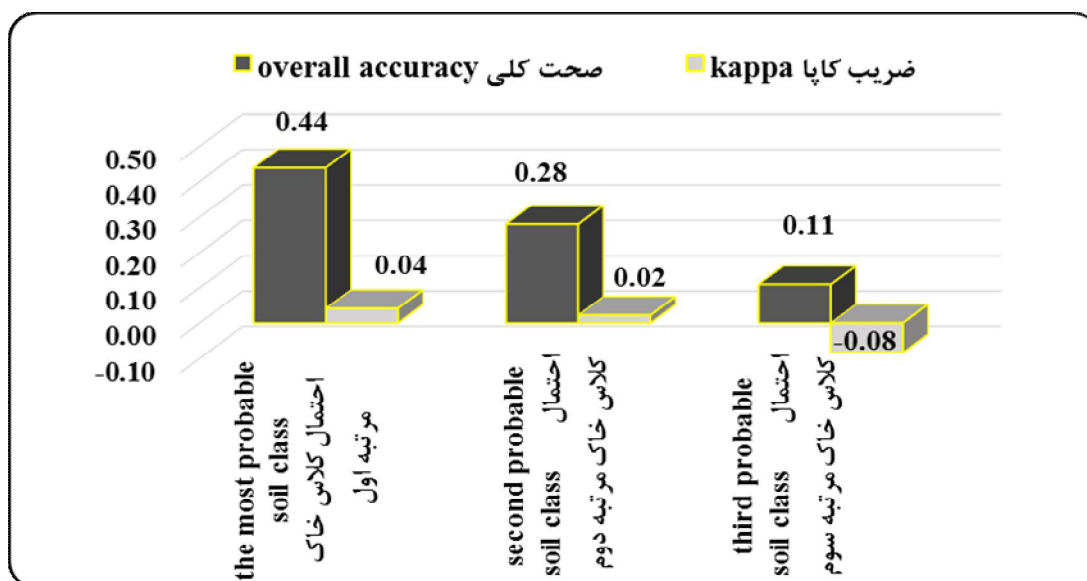


شکل ۶- نقشه احتمال کلاس خاک مرتبه سوم.  
 Figure 6. Third most probable soil class map.



شکل ۷- شاخص آسیمگی بین نقشه احتمال کلاس خاک مرتبه اول و نقشه احتمال کلاس خاک مرتبه دوم.

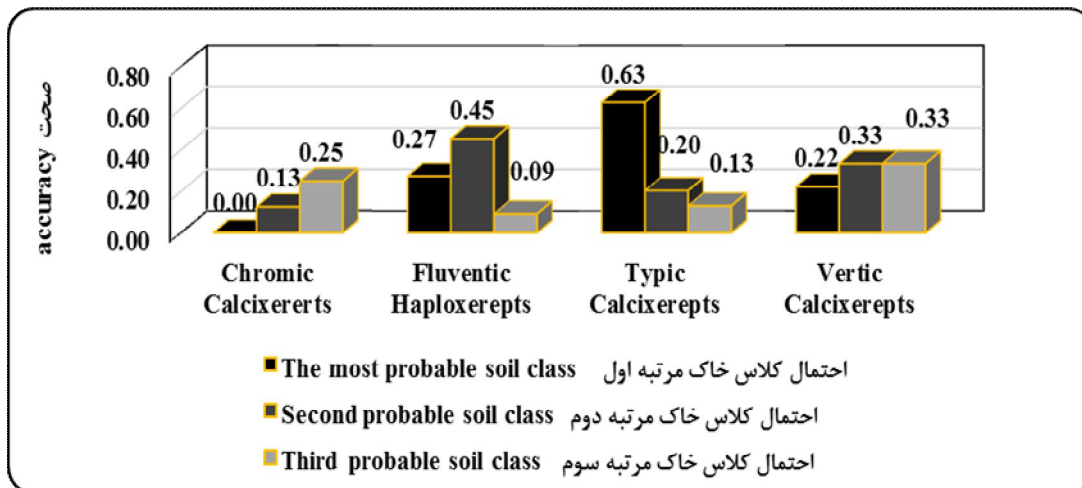
Figure 7. Confusion index between the most probable soil class and Second most probable soil class.



شکل ۸- صحت کلی و ضریب کاپا مدل دسمارت برای پیش‌بینی کلاس‌های خاک.

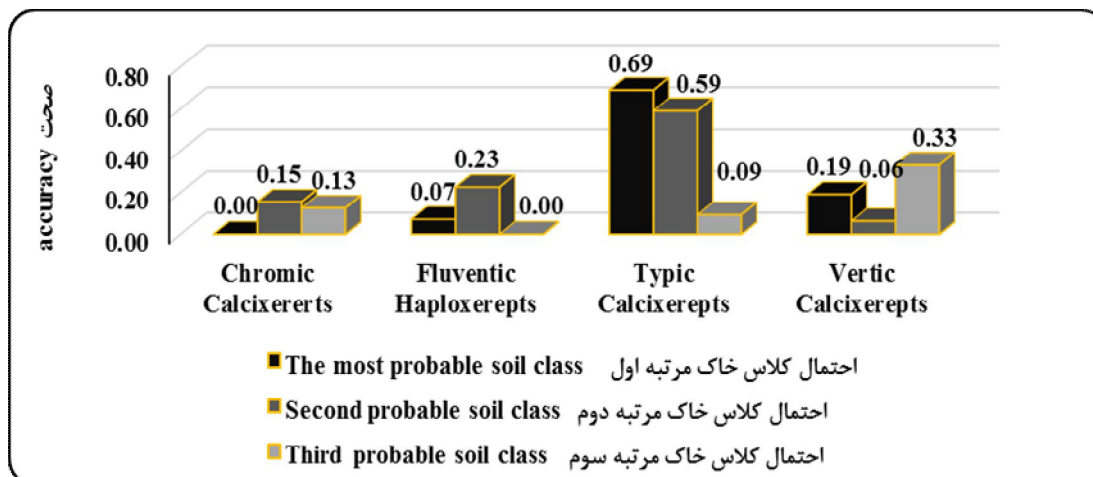
Figure 8. Overall accuracy and kappa coefficient of dsmart model for prediction soil class.





شکل ۹- صحت کاربر مدل دسمارت برای پیش‌بینی کلاس‌های خاک.

Figure 9. User's Accuracy of dsmart model for prediction soil class.



شکل ۱۰- صحت تولیدکننده نقشه مدل دسمارت برای پیش‌بینی کلاس‌های خاک.

### نتیجه‌گیری کلی

روش انبوه‌ش‌زدائی و یکنواخت‌سازی واحدهای نقشه خاک از طریق بازچینی توسط درختان طبقه‌بندی (دسمارت)، الگوریتم جدیدی برای انبوه‌ش‌زدایی نقشه‌های چندگوش (پلی‌گونی) میراثی خاک است. این روش است قادر است احتمال وقوع تمام کلاس‌های خاک ناحیه مورد مطالعه که توزیع آن‌ها در واحدهای نقشه نامشخص است را به صورت نقشه‌های رستری نمایش دهد و علاوه بر این قادر به تولید

به نظر می‌رسد استفاده از متغیرهای کمکی مانند نقشه لندفرم و داده‌های سنجنش از نزدیک<sup>۱</sup> می‌تواند به بهبود نقشه‌های پیش‌بینی تولیدشده کمک کند. به‌طورکلی می‌توان گفت این روش کارایی نسبتاً خوبی برای تولید نقشه‌های رقومی کلاس‌های منفرد خاک تنها با استفاده از نقشه چهارگوش خاک و ترکیب نسبی کلاس‌های خاک در هر واحد نقشه و متغیرهای کمکی دارد.

1- Proximal sensing

کلاس خاک با دقت نسبتاً خوب و قدرت تفکیک مکانی مناسب تبدیل شوند. از مزایای این روش عدم استفاده از نمونه‌برداری جدید برای تولید نقشه می‌باشد و انجام نمونه‌برداری تنها برای اعتبار سنجی نقشه‌های تولید شده استفاده می‌شود. در شرایط ایران، از این روش می‌توان برای انبوه‌زدایی آن دسته از نقشه‌های خاک کشور که سهم نسبی کلاس‌های خاک در آن‌ها مشخص است و همچنین انبوه‌زدایی نقشه یک میلیونم خاک‌های کشور استفاده نمود.

تعداد  $n$  نقشه احتمال کلاس خاک و نقشه درجه آسیمیگی بین آن‌ها می‌باشد. چنین خروجی‌هایی به فهم بهتر روابط سیمای اراضی - خاک کمک می‌کنند. واضح است که مدل درختان طبقه‌بندی (الگوریتم C5.0) توانسته است رابطه بین کلاس‌های خاک و الگوی قرارگیری آن‌ها بر روی سیمای اراضی را به‌خوبی استخراج نماید. این روش فرصتی را فراهم نموده است که نقشه‌های سنتی خاک را تنها بر اساس اطلاعات ترکیب واحد نقشه، به نقشه‌های رقومی

### منابع

1. Banaei, M.H., Momeni, A., Baybordi, M., and Malakouti, M.J. 2005. The soils of Iran (new achievements in perception, managements and use). Soil and Water Research Institute, AREO, Tehran, Iran, 482p. (In Persian)
2. Bui, E.N., and Moran, C.J. 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. *Geoderma*. 103: 2. 79-94.
3. 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*. 240: 68-83.
4. Brus, D.J., Kempen, B., and Heuvelink, G.B.M. 2011. Sampling for validation of digital soil maps. *Europ. J. Soil Sci.* 62: 394-407.
5. Burrough, P.A., van Gaans, P.F.M., and Hootsmans, R. 1997. Continuous classification in soil survey: spatial correlation, confusion and boundaries. *Geoderma*. 77: 115-135.
6. Chaney, N., Hempel, J.W., Odgers, N.P., McBratney, A.B., and Wood, E.F. 2014. Spatial disaggregation and harmonization of gSSURGO. In: ASA, CSSA and SSSA international annual meeting, Long Beach. ASA, CSSA and SSSA.
7. Collard, C., 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.
8. Dobos, E., Bialkó, T., Micheli, E., and Kobza, J. 2010. Legacy Soil Data Harmonization and Database Development. P 309-324, J.L. Boettinger et al. (eds.), *Digital Soil Mapping, Progress in Soil Science 2*, Springer.
9. 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.
10. Fatehi, Sh. 2008. Semi-detailed soil survey of Merek plain in Karkheh river basin. Soil and Water Research Institute, 54p. (In Persian)
11. Fatehi, Sh., Mohammadi, J., Salehi, M.H., Toomanian, N., Momeni, A., and Jafari, A. 2015. Spatial disaggregating conventional soil map using multiple logistic regression and classification tree, (Case study: Merek sub catchment in Kermanshah province). 14<sup>th</sup> Iranian soil science congress, September 28-30, Rafsanjan, Iran, Pp: 208-213. (In Persian)
12. Häring, T., Dietz, E., Osenstetter, S., Koschitzki, T., and Schroder, B. 2012. Spatial disaggregation of complex soil map units: A decision-tree based approach in Bavarian forest soils. *Geoderma*. 37: 185-186.

13. Holmes, K.W., Griffin, E.A., and Odgers, N.P. 2015. Large-area spatial disaggregation of a mosaic of conventional soil maps: evaluation over Western Australia. *Soil Research*. 53: 865-880.
14. Lagacherie, P., Legros, J.P., and Burrough, P. 1995. A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. *Geoderma*. 65: 4. 283-301.
15. McBratney, A.B. 1998. Some considerations on methods for spatially aggregating and disaggregating soil information. *Nutr. Cycl. Agroecosyst*. 50: 3. 51-62.
16. McBratney, A.B., Field, D.J., and Koch, A. 2014. The dimensions of soil security. *Geoderma*. 13: 203-213.
17. Malone, B.P., Minasny, B., and McBratney, A.B. 2017. *Using R for Digital Soil Mapping*. Springer, the Netherlands, Pp: 221-230.
18. Nauman, T.W., and Thompson, J.A. 2014. Semi-automated disaggregation of conventional soil maps using knowledge driven data mining and classification tree. *Geoderma*. 213: 385-399.
19. Odgers, N.P., Sun, W., McBratney, A.B., Minasny, B., and Clifford, D. 2014. Disaggregating and harmonizing soil map units through resampled classification trees. *Geoderma*. 215: 91-100.
20. Rouse, J.W., Hass, R.H.J., Schell, A., and Deering, D.W. 1973. Monitoring vegetation systems in the Great Plains with ERTS. Third ERTS Symposium, NASA SP-351, Vol. 1, Washington, DC. Pp: 309-317.
21. Quinlan, J.R. 1994. C4.5: Programs for machine learning. *Machine Learning*. 16: 235-240.
22. Subburayalu, S., Jenhan, I., and Slater, B.K. 2014. Disaggregation of component soil series using possibilistic decision trees from an Ohio County soil survey map. *Geoderma*. 213: 334-345.
23. Soil Survey Staff. 1993. *Soil survey manual*. U. S. Department of Agriculture Handbook. United States Department of Agriculture Soil Conservation Service.
24. Thompson, J.A., Prescott, T., Moore, A.C., Bell, J., Kautz, D.R., Hempel, J.W., Walt man, S.W., and Perry, C. 2010. Regional approach to soil property mapping using legacy data and spatial disaggregation. Techniques. In: 19<sup>th</sup> world congress of soil science. IUSS, Brisbane.
25. Van Deventer, A.P., Ward, A.D., Gowda, P.H., and Lyon, J.G. 1997. Using Thematic Mapper Data to Identify Contrasting Soil Plains and Tillage Practices. *Photogrammetric Engineering & Remote Sensing*. 63: 87-93.
26. Vincent, S., Lemercier, B., Berthier, L., and Walter, C. 2018. Spatial disaggregation of complex Soil Map Units at the regional scale based on soil-landscape relationships. *Geoderma*. 311: 130-142.
27. Wei, S., McBratney, A., Hempel, J., Minasny, B., Malone, B., D'Avello, T., Burras, L., and Thompson, J. 2010. Digital harmonisation of adjacent analogue soil survey areas-4 Iowa counties. 19<sup>th</sup> World Congress of Soil Science: Soil solutions for a changing world 2010, Wien (Vienna), Austria: International Union of Soil Sciences (IUSS).
28. Xiao, J., Shen, Y., Tateishi, R., and Bayaer, W. 2006. Development of topsoil grain size index for monitoring desertification in arid land using remote sensing. *Inter. J. Rem. Sens*. 27: 2411-2422.



Gorgan University of Agricultural  
Sciences and Natural Resources

*J. of Water and Soil Conservation, Vol. 25(2), 2018*

<http://jwsc.gau.ac.ir>

DOI: 10.22069/jwsc.2018.13870.2860

## **Spatial disaggregation of semi-detailed soil map using DSMART approach**

**\*Sh. Fatehi<sup>1</sup>, K. Eftekhari<sup>2</sup> and J. Ghaderi<sup>1</sup>**

<sup>1</sup>Research Assistant Prof., Soil and Water Research Department, Kermanshah Agricultural and Natural Resources Research and Education Center, AREEO, Kermanshah, Iran,

<sup>2</sup>Research Assistant Prof., Soil and Water Research Institute, Agriculture Research, Education and Extension Organization (AREEO), Karaj, Iran

Received: 08.22.2017; Accepted: 04.22.2018

### **Abstract**

**Background and Objectives:** Digital soil data with high spatial resolution and enough accuracy and precision are necessary for management of global challenges such as food security, environment problems. Generally, soil data are available in small scale. Nevertheless, in the last decades, with the advent of soil digital mapping and modeling approaches, it is possible to disaggregate soil map units. The spatial disaggregation of soil map units is a method for modeling the spatial distribution of individual soil classes. During this process, the soil map data from a small scale (coarse resolution) is converted to a large scale (fine resolution). The statistical and data mining methods are used for its implementation. The purpose of this research was to predict the spatial distribution of soil classes by disaggregating the soil map units of a semi detailed soil map using disaggregating and harmonizing soil map units through resampled classification trees algorithm (DSMART method).

**Materials and Methods:** The study area is located in Kermanshah province. The total area of the study was approximately 14083.9 ha. Soil polygon map include 5 map units and 4 soil subgroups. In this study, elevation, slope, aspect, convexity, direct duration, sediment index, topographic wetness index, valley depth and vertical distance to channel network as covariates produced using DEM 10 m. Grain size index, clay index and NDVI were also calculated using Landsat 7 ETM<sup>+</sup> imagery. Geological map at scale of 1:100,000 were also used as a qualitative covariate. Then, dsmart method is run as a novel approach for disaggregation soil maps. DSMART samples randomly within the soil map units and uses classification trees (C5.0 algorithm) to produce probability surface maps of soil class distribution. External validation was performed using 82 profiles. The validation dataset was intersected with the corresponding probability surface maps and validation quantified by overall accuracy, producer's accuracy, user's accuracy and kappa coefficients. Furthermore, confusion index was calculated between the most probable and second-most-probable soil class. The CI expresses concisely degree of confusion of given soil classes.

**Results:** The most important predictive variables in the tree classification model were Vertical distance to channel network, Elevation, lithology, grain size index and MRVBF. The confusion index close to 1 has a large extent in the study area. It shows that occurrence probability of soil subgroups is near equal in each location in both the most probable soil class and second probable soil class maps. Validation of probability surfaces showed that overall accuracy of the most probable soil class, second probable soil class and third probable soil class are 44%, 28% and 11%, respectively. These results indicated the relatively good performance of dsmart method for generating digital individual soil class map. However, kappa coefficients for first, second and third probable surfaces soil maps were obtained 0.04, 0.02, -0.08, respectively. Low

---

\* Corresponding Author; Email: shahrokh.fatehi@gmail.com

kappa coefficients can be attributed to the actual nature of the data, i.e the dominance of the Typic Calcixerepts subgroup as compared to other subgroups of the soil in the traditional soil map and the dsmart model prediction map and validation data.

**Conclusion:** Dsmart method can predict the probability of occurrence of all soil classes which their distribution is unclear in soil map unit. This provides the opportunity to produced digital soil class maps when legacy soil data and covariate information are available. Such outputs may help to understand the relationship between landscape and soil.

**Keywords:** Spatial disaggregation, Digital soil mapping, Soil class, Map unit

