

ارزیابی دقت نقشه برداری رقومی خاک با محدودیت داده در بخشی از فلات لسی استان گلستان

صدیقه ملکی^{۱*}، فرهاد خرمالی^۲، محسن باقری بذاغ آبادی^۳، جهانگیر محمدی^۴

(تاریخ دریافت: ۱۳۹۷/۰۵/۱۳ تاریخ پذیرش: ۱۳۹۷/۰۷/۲۴)

چکیده

روش‌های نقشه برداری رقومی خاک نیازمند داده‌های کمی برای تخمین هستند که به کارگیری آن‌ها در مناطق با تعداد داده خاک کم، و عدم دسترسی مناسب بسیار سخت می‌باشد. انجام عملیات میدانی به‌ویژه در سطح وسیع کاری زمان‌بر و همراه با هزینه زیاد است که عملیات نقشه برداری را مشکل‌ساز می‌نماید. از دیگر سو، باور در نقشه برداری رقومی بر این است که وجود یک خاک منحصر به فرد در منطقه وابستگی زیادی به متغیرهای محیطی آن منطقه و شناسایی دقیق آن‌ها دارد. در این مطالعه از الگوریتم جنگل تصادفی (RF) به همراه اطلاعات طبقه‌بندی ۶۴ خاکرخ و ۱۹ متغیر محیطی، شامل خصوصیات توپوگرافی، واحدهای ژئومورفولوژی، کاربری اراضی و شاخص پوشش گیاهی، برای تهیه نقشه کلاس خاک در بخشی از فلات لسی استان گلستان استفاده گردید. ژئومورفولوژی، ارتفاع، جهت شیب و کاربری اراضی تقریباً در همه سطوح رده‌بندی دارای بیش‌ترین اهمیت در تخمین کلاس‌های خاک بودند. نتایج ارزیابی دقت الگوریتم RF با متغیرهای ورودی مختلف نشان داد شاخص‌های صحت مدل از جمله صحت کلی و کاپا به ترتیب از ۰/۹۱ و ۰/۸۳ برای گروه بزرگ، به ۰/۷۸ و ۰/۵۶ برای زیرگروه، و ۰/۵۰ و ۰/۳۲ برای فامیل کاهش می‌یابد. کم‌ترین و بیش‌ترین مقدار خطای تخمین نمونه‌های اعتبارسنجی در مدل‌سازی ۳۲/۶۹ و ۶۵/۳۸ درصد به ترتیب برای سطح گروه بزرگ و فامیل خاک دست آمد و کلاس‌های دارای تعداد نمونه بیش‌تر، خطای کم‌تری داشتند. مطالعه حاضر نشان داد که در مناطق با محدودیت داده از ایران، نقشه برداری رقومی خاک، استفاده از متغیرهای محیطی با دقت بالا همراه با تعداد نمونه خاک کم می‌تواند نتایج مطلوبی را در سطوح بالای رده‌بندی ارائه دهد.

واژه‌های کلیدی: جنگل تصادفی، دقت نقشه، متغیرهای محیطی، محدودیت نمونه خاک

ملکی ص.، خرمالی ف.، باقری بذاغ آبادی م.، محمدی ج. ۱۳۹۸. ارزیابی دقت نقشه برداری رقومی خاک با محدودیت داده در بخشی از فلات لسی استان گلستان. تحقیقات کاربردی خاک. جلد ۷، شماره ۴. صفحه: ۱۶۳-۱۴۸.

۱- دانش آموخته دکتری گروه علوم خاک، دانشکده مهندسی آب و خاک، دانشگاه علوم کشاورزی و منابع طبیعی گرگان (مکاتبه کننده)

۲- استاد گروه علوم خاک، دانشکده مهندسی آب و خاک، دانشگاه علوم کشاورزی و منابع طبیعی گرگان

۳- موسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج

۴- استادیار گروه علوم جنگل، دانشکده جنگل، دانشگاه علوم کشاورزی و منابع طبیعی گرگان

* پست الکترونیک: elymaleki@yahoo.com

مقدمه

دنیاستفاده شده است. این موارد شامل توابع انتقالی^۲، زمین آمار، مدل‌های آماری تجربی، طبقه‌بندی فازی، تصمیم‌گیری درختی^۳ و شبکه‌های عصبی مصنوعی هستند (McBratney *et al.*, 2003). یکی دیگر از این مدل‌ها، مدل جنگل (RF) است که شامل رشد دادن مجموعه‌ای از درختان در مدل درخت تصمیم (Breiman & Cutler, 2004) می‌باشد. این مدل دقیق بوده و از لحاظ محاسباتی نیز سریع می‌باشد (Breiman, 2001) و به جای رشد دادن یک درخت، تعداد زیادی درخت مثلا صدها یا هزارها درخت، مستقل رشد داده می‌شود. تقریبا یک سوم داده‌های اصلی در ایجاد هر درخت استفاده نمی‌شوند که نمونه‌های بیرون از سبد (OOB) نامیده می‌شوند. چون این نمونه‌ها در آموزش درخت استفاده نشده‌اند، در تست مورد استفاده قرار می‌گیرند و بنابراین نیازی به اعتبارسنجی مستقل در این روش نمی‌باشد. وانگ و همکاران (Wang *et al.*, 2018) در شمال شرق چین با استفاده از این روش اقدام به تهیه نقشه پیش‌بینی نیتروژن با استفاده از ۱۲ متغیر کمکی (شامل داده‌های توپوگرافی، اقلیم و سنجش از دور) نموده که نتایج ارزیابی صحت آن‌ها، شاهدهی بر کارایی بالای الگوریتم RF با استفاده از متغیرهای محیطی بوده است. همچنین پهلوان‌راد و همکاران (Pahlevan Rad *et al.*, 2014) برای روز کردن نقشه‌های قدیمی از RF استفاده نمودند. نتایج آن‌ها نشان داد که RF یک روش قابل اعتماد و با دقت مناسب است که با تعداد نمونه کم تخمین قابل قبول ارائه می‌دهد.

تنگ و همکاران (Teng *et al.*, 2018) نیز توانمندی الگوریتم RF را در به روزرسانی نقشه طبقه‌بندی خاک استرالیا با دقت ۵۵/۶۰ درصد گزارش نموده‌اند. سری‌نیواس و همکاران (Sreenivas *et al.*, 2016) نیز کربن آلی و غیرآلی خاک را تا عمق یک متری در سراسر هندوستان با استفاده از ۱۱۹۸ نمونه خاک و الگوریتم RF مدل‌سازی کردند. نتایج آن‌ها نشان داد که ضریب تبیین برای کربن آلی و غیرآلی به ترتیب ۰/۸۲ و ۰/۸۶ بود.

وجود نقشه‌های خاک از نیازهای اساسی در مطالعات علوم زمین می‌باشد (Pahlevan Rad *et al.*, 2014). با وجودی که بیش از ۵۰ سال از شروع مطالعات خاک‌شناسی در ایران می‌گذرد، نقشه‌های خاک کشور به طور کامل تهیه نشده‌اند. علاوه بر این، نقشه خاک با مقیاس ۱:۱۰۰۰۰۰۰ منبع اصلی اطلاعات خاک در ایران است و در اوایل سال ۲۰۰۰ میلادی تهیه و کامل گردیده است (Banaei, 2000). علاوه بر این، حدود ۷۵ درصد از خاک‌های کشور نیز فاقد اطلاعات نقشه‌برداری می‌باشد (Pahlevan Rad *et al.*, 2016). با توجه به وسعت زیاد کشور ایران، استفاده از روش‌های مرسوم نقشه‌برداری خاک برای تهیه نقشه‌های خاک کل مناطق کشور با مقیاس مناسب، به دلیل هزینه‌بر و وقت‌گیر بودن، بسیار مشکل است. تکنیک‌های نقشه‌برداری سنتی (به عنوان مثال، Soil Survey Staff, 1996) نیازمند تعداد مشاهدات صحرایی زیادی می‌باشند (Stoorvogel *et al.*, 2009). بنابراین برای به تصویر کشیدن تغییرات مکانی خاک‌ها با هزینه کمتر و در مدت زمان کوتاه‌تر به گونه‌ای که نقشه‌های قابل قبول از خاک ارائه شود، روش‌های تخمین مکانی خاک با تاکید بر منابع داده‌های قابل دسترس ارائه شده است، که در اصطلاح نقشه‌برداری رقومی خاک نامیده می‌شود (McBratney *et al.*, 2003).

پایه و اساس نقشه‌برداری رقومی خاک مبتنی بر معادله اسکورپن^۱ می‌باشد. در این معادله ویژگی‌ها یا کلاس‌های خاک با یکسری متغیرهای محیطی (متغیرهای کمکی) ارتباط داده می‌شوند. به این صورت که در هر موقعیت مشاهده خاک، یک سری متغیرهای محیطی وجود دارد که همبستگی بالایی با ویژگی‌ها یا کلاس‌های خاک دارند. یافتن این ارتباط و پیش‌بینی خصوصیات خاک در نقاط دیگر، منجر به تهیه نقشه رقومی خاک در تمام منطقه مورد مطالعه می‌گردد (Abbaszadeh Afshar *et al.*, 2018). با توجه به توضیحات داده شده، یک جز اصلی معادله اسکورپن فرآیند یادگیری یا به عبارتی مدل پیش‌بینی‌کننده می‌باشد. برای پیش‌بینی ویژگی‌های خاک یا کلاس خاک، مدل‌های مختلفی در نقاط مختلف

2. Pedotransfer
3. Decision tree
4. Out of bag

1. Scorpan

بسیار بحث برانگیز بوده و در گذشته براساس دوره‌های وجود یا عدم وجود بارندگی و یا یخچالی و بین‌یخچالی مورد بحث قرار گرفته است (Wang *et al.*, 2016). اگرچه مطالعات انجام شده در فلات لسی استان گلستان نیز نشان می‌دهند که این لس‌ها حاوی اطلاعاتی مفیدی برای تغییر اقلیم و تکامل زمین‌نما در دوره کواترنری می‌باشند (Khormali & Kehl, 2011). اما با وجود پتانسیل زیاد این منطقه به عنوان مرتع بهاره و تولید گندم به‌عنوان یک محصول استراتژیک، مطالعات خاک‌شناسی به‌ویژه در خاک‌های این مناطق بسیار کم صورت گرفته است و با توجه به پیچیدگی‌های توپوگرافی و عدم مسیر دسترسی مناسب، اطلاعات بسیار کم و ناچیزی در مورد تکامل خاک، ژئومورفولوژی و پتانسیل‌ها و مسائل مربوط به کاربری اراضی منطقه در دسترس می‌باشد. از این رو در این مطالعه، سعی بر آن است که با استفاده از الگوریتم RF، دانش فرد خیره در جداسازی خاک‌های منطقه و استفاده از تعداد مشاهدات خاکی کمتر، نقشه کلاس خاک در بخشی از منطقه فلات لسی استان گلستان در سه سطح گروه بزرگ، زیر گروه و فامیل خاک تهیه شود.

مواد و روش‌ها

مشخصات عمومی منطقه مورد مطالعه

منطقه‌ی مطالعاتی به مساحت تقریبی ۵۳۹۰/۳۷ هکتار در بین طول‌های جغرافیایی "۳۶' ۰۹" ۵۵ تا "۲۶' ۱۳" ۵۵ شرقی و عرض‌های جغرافیایی "۳۷' ۳۶" ۳۷ تا "۴۱' ۳۷" شمالی در فلات لسی استان گلستان قرار گرفته است (شکل ۱). مشاهدات سن‌یابی و چینه‌شناسی این منطقه نشان می‌دهد که این رسوبات لسی مربوط به دوران کواترنری می‌باشند. از طرفی بین این رسوبات و رسوبات لسی در جنوب شرقی اروپا و بخش مرکزی آسیا می‌توان ارتباط برقرار کرد (Kehl *et al.*, 2005; Frechen *et al.*, 2009). میانگین بارش و متوسط دمای سالانه منطقه، به ترتیب ۳۵۰ میلی‌متر و ۱۷ درجه‌ی سلسیوس می‌باشد. رژیم‌های رطوبتی و حرارتی خاک منطقه مطالعاتی، به ترتیب، زیریک خشک و ترمیک می‌باشند.

با تمام مباحث ذکر شده، هنوز در برخی مناطق دنیا به علت عدم وجود داده‌های خاک در دسترس، تکنیک‌های نقشه‌برداری رقومی به خوبی قابل استفاده نیستند (Stoorvogel *et al.*, 2009; Zhu *et al.*, 2008). اکثر مطالعات در زمینه نقشه‌برداری رقومی کلاس‌های خاک در ایران و جهان نیز با استفاده از تعداد داده بیشتر از ۱۰۰ نمونه خاکرخ انجام شده است (به عنوان مثال، جعفری و همکاران (Jafari *et al.*, 2013) و عباس‌زاده افشار و همکاران (Abbaszadeh Afshar *et al.*, 2018) با ۱۲۶ نمونه به ترتیب در منطقه زرنند و بم استان کرمان، میرکزه‌ی و همکاران (Mirakzahi *et al.*, 2018) ۱۰۸ نمونه در منطقه هیرمند استان سیستان و بلوچستان، تقی‌زاده مهرجردی و همکاران (Taghizadeh-Mehrjardi *et al.*, 2014) با ۱۷۳ نمونه در منطقه اردکان استان یزد، استام و همکاران (Stum *et al.*, 2010) و برونگارد (Brungard, 2009) به ترتیب با ۶۷۲ و ۳۰۰ مشاهده خاکرخ در مناطق خشک از امریکا و هرینگ و همکاران (Häring *et al.*, 2012) با ۹۹۲۴ مشاهده در خاک‌های جنگلی باوریا و در مقابل مطالعات اندکی (همچون Zhu *et al.*, 2008; Stoorvogel *et al.*, 2009; Zeraatpisheh *et al.*, 2017)، اقدام به بررسی اثر تعداد نمونه محدود در تهیه نقشه خاک پرداخته‌اند. این در حالی است که برای کشوری مانند ایران که در حال حاضر نگرانی درباره منابع آب و خاک برای تامین مواد غذایی راه‌های مدیریت خاک را تغییر داده است، نیاز به اطلاعات و نقشه‌های خاک بیش از پیش احساس می‌گردد. بنابراین، استفاده بهینه از علم نقشه‌برداری رقومی، به کارگیری متغیرهای کمکی دقیق‌تر و استفاده از دانش افراد خبره در زمینه شناسایی خاک بایستی افزون بر تسهیل در مطالعات خاک و انتخاب واحدهای نقشه یکنواخت، منجر به برداشت تعداد نمونه کمتر و کاهش هزینه‌های نمونه‌برداری و آزمایشگاهی در جهت پیش‌برد نقشه‌برداری مناطق با عدم داده در دسترس گردد.

رسوبات لسی شمال ایران، منعکس‌کننده چندین چرخه تغییر اقلیم و تکامل سیمای سرزمین برای دوره میانی تا انتهای کواترنری هستند. تغییر اقلیم در ارتفاعات ایران و مناطق اطراف آن در اواسط و اواخر دوره کواترنری

آمریکا (Soil Survey Staff, 2014) تشریح، طبقه‌بندی و مورد نمونه‌برداری قرار گرفتند. برای هر نقطه مطالعاتی اطلاعاتی از محل خاکرخ در زمین‌نما، واحد ژئومورفیک، نوع مواد مادری، جهت شیب، میزان سنگریزه‌های سطحی، نوع و میزان فرسایش نیز یادداشت‌برداری شد.

نمونه‌های خاک تهیه شده از هر افق، پس از هوا خشک شدن و عبور از الک ۲ میلی‌متری، براساس دستورالعمل نقشه‌برداری خاک (Soil Survey Staff, 1996) جهت انجام آزمایش‌های فیزیکی و شیمیایی شامل کربن آلی خاک، کربنات کلسیم معادل، هدایت الکتریکی، pH، بافت خاک، گچ و ظرفیت تبادل کاتیونی مورد آنالیز قرار گرفتند.

است که براساس تعداد نمونه مورد نظر کاربر، داده‌های متغیرهای محیطی به همان تعداد خوشه‌های خواسته شده تقسیم می‌گردند و یک نمونه تصادفی از داده‌های ورودی متغیرهای محیطی هر خوشه انتخاب می‌شود. در این مطالعه، مقادیر همه متغیرهای کمکی (جدول ۱) در پیکسل‌های ۵×۵ متری تهیه و به نرم‌افزار متلب وارد گردید. لازم به ذکر است به دلیل عدم وجود اطلاعات پایه و خاکی در ارتباط با منطقه مطالعاتی، کلیه نقاط مشخص شده با روش cLHS در صحرا توسط فرد خیره (Bagheri Bodaghabadi *et al.*, 2015) مورد بازبینی قرار گرفت و با توجه به ویژگی‌های خاک‌ساز مهم منطقه تعداد ۶۴ خاکرخ حفر گردید (شکل ۱ج). با استفاده از GPS مکان دقیق نقاط نمونه‌برداری در صحرا مشخص و تمامی مقاطع خاک براساس روش طبقه‌بندی خاک

جدول ۱- داده‌های محیطی مورد استفاده به عنوان متغیر پیش‌بینی کننده در مدل‌سازی

Table 1. Environmental data that used as predictor variables in modeling

Environmental data	Soil forming factor	Parameter	Symbol	Variable type	Reference
Topographic attributes	Topography	Elevation	ELV	Quantitative	Wilson & Gallant, 2000
		Slope	Slope		
		Aspect	ASP		
		Curvature	Curv		
		Profile curvature	Profc urv		
		Plan curvature	Planc urv		
		Convergence	Conv erg		
		Classcurvature	Class curv		
		Relative slope	Reltsl p		
		Flow accumulation	Flowa cc		
		Stream power index	LS		
		Topographic wetness index	TWI		
		Valley depth	valley		
Multi-resolution valley bottom flatness index	mrv				
Multi-resolution of ridge top flatness index	mrr				
Topographic position index	TPI	Weiss, 2001			
Remote sensing attributes	Vegetation	Soil adjusted vegetation index	SAVI		Huete, 1988
Landuse map		Landuse	landu se	Categorical	-
Geomorphology map	Parent material, topography, soil	Geomorphology units	landf orm	Categorical	Toomanian <i>et al.</i> , 2006

متغیرهای کمکی

مدل رقومی ارتفاع (DEM) منطقه مطالعاتی با پیکسل‌هایی به ابعاد ۵×۵ متر از خطوط توپوگرافی با مقیاس ۱:۲۵۰۰۰ سازمان نقشه برداری کشور تهیه شد. لازم به ذکر است که این DEM توسط ۶۰۰ نقطه ارتفاعی برداشت شده در منطقه توسط دوربین تئودولیت تصحیح گردید. ویژگی‌های اولیه و ثانویه پستی و بلندی با استفاده از محاسبات روی DEM در محیط سامانه جغرافیایی SAGA محاسبه و استخراج شد (جدول ۱). روش استخراج همه پارامترهای مزبور را، هنگل و همکاران (Hengl et al., 2004) تشریح کردند. از SAVI (Huete, 1988) به عنوان شاخص پوشش گیاهی استفاده گردید، که از تصاویر ماهواره لندست ۸ مربوط به اواخر اسفند ماه ۱۳۹۳ (۲۵ فوریه ۲۰۱۴) با استفاده از نرم‌افزار پردازشی ENVI 4.4 محاسبه شد. واحدهای ژئومورفیک با استفاده از روش تفسیر عکس‌های هوایی (Toomanian et al., 2006) تعیین شدند (جدول ۲). این روش دارای چهار سطح زمین‌نما، لندفرم، لیتولوژی و سطح ژئومورفیک می‌باشد. در تفکیک واحدهای اراضی، از ارتباط خاک- اراضی استفاده و براساس دانش ژئومورفولوژی و شناخت فرآیندهای تکامل واحدهای اراضی، مرز واحدها ترسیم و با ارتباط دادن آن‌ها با تکامل خاک‌ها همراه با تاثیر دادن فاکتورهای زمین‌شناسی، توپوگرافی و پوشش گیاهی یا کاربری اراضی، سطوح ژئومورفیک یکنواختی ترسیم گردید. سپس در محیط نرم‌افزار گوگل ارث با استفاده از تصاویر ماهواره Quickbird که برای منطقه مورد نظر وجود داشت، در حالت سه‌بعدی اقدام به جداسازی و رقومی کردن مرز واحدهای مختلف گردید. در نهایت لایه تهیه شده با فرمت KML وارد محیط نرم‌افزار ArcGIS 10.2 شده و نقشه ژئومورفولوژی استخراج گردید (شکل ۱ ج). به منظور صحت‌مربندی، در طی عملیات صحرایی، مرز واحدهای موجود در نقشه تفسیری اولیه (ژئوform) کنترل و تصحیح شد. یکنواختی واحدها و سطوح زمین‌ریختی با نمونه‌برداری صحرایی و تشریح خاک‌رخ مورد بررسی و صحت حدود ترسیم شده برای واحدها مورد آزمون قرار گرفت. از این رو با توجه به اثر جهت شیب در تکامل

خاک در منطقه مورد نظر سطح فاز اراضی به سطوح قبلی جدول ژئوform اضافه گردید که پیش از این توسط ملکی و همکاران (Maleki et al., 2017) به طور کامل بحث گردیده است.

جنگل تصادفی (RF)

الگوریتم RF به دو روش میانگین کاهش دقت و میانگین کاهش ضریب جینی اهمیت متغیرها را نشان می‌دهد (Myles et al., 2004). در روش میانگین کاهش دقت که بیش‌ترین کاربرد را دارد، مقادیر درست متغیرها با مقادیر که به طور تصادفی برای هر درخت تولید شده است، جایگزین می‌شود و اثر این تغییر را روی طبقه‌بندی اندازه‌گیری می‌کند. اگر این جایگزینی اثری روی خطای اندازه‌گیری نداشته باشد، اهمیت آن متغیر کم است و اگر مقدار خطای اندازه‌گیری افزایش یابد، آن متغیر مهم می‌باشد (Breiman & Cutler, 2004). میانگین کاهش ضریب جینی همچنین اهمیت متغیرها را به وسیله جابجا کردن مقادیر متغیرهای محیطی در نمونه‌های OOB انجام می‌دهد و متغیرهای محیطی همراه با یک افزایش نسبی در OOB با اهمیت هستند. متعاقباً متغیر دارای همگنی بالا، یک میانگین کاهش ضریب جینی بالایی خواهد داشت (Breiman, 2001). در بحث مزایای این الگوریتم، بایستی بیان نمود که مسئله خودآزمایی^۲ الگوریتم در بحث داده‌های OOB، شبیه مسئله واریانس اعتبار می‌باشد که هر بار روی قسمت‌بندی‌های متفاوتی از داده اولیه محاسبه می‌شوند. بنابراین با استفاده از این قابلیت RF، حتی اگر کل داده آموزشی به‌عنوان ورودی الگوریتم استفاده شود، امکان خودآزمایی در طی یادگیری فراهم است. علاوه بر آن، این قابلیت در مورد داده‌های کوچک دارای اهمیت زیادی است. چرا که نیازی به کنار گذاشتن قسمتی از داده برای تست نمی‌باشد. بنابراین از جمله الگوریتم‌های پرکاربرد در خصوص اعمال بر روی داده‌های با حجم نمونه کم محسوب می‌شود. از دیگر نتایج وجود داده‌های OOB در هنگام ساخت درخت، مقاومت در برابر مسئله بیش‌برازش و عمومیت بخشی در برابر داده‌های جدید می‌باشد. همچنین با انتخاب زیر مجموعه‌ای از خصیصه‌ها، نیازی به اعمال الگوریتم‌های انتخاب ویژگی

1. Soil adjusted vegetation index

2. Self- testing

۲۰ نمونه به طور مستقل اعتبارسنجی گردید. پارامترهای دقت کلی (Brus *et al.*, 2011) و شاخص کاپا (Marchetti *et al.*, 2011; Pahlevan Rad *et al.*, 2016) برای نقشه‌های سطوح مختلف رده‌بندی تعیین شد.

نیز نخواهد بود (Breiman & Cutler, 2004; Zeraatpisheh *et al.*, 2017) در این پژوهش الگوریتم RF با استفاده از نرم‌افزار R Development Core (Team, 2013) اجرا خواهد گردید. اگرچه الگوریتم خود اعتبارسنجی انجام می‌دهد ولی در این تحقیق برای بررسی دقت روش، همه نقشه‌های خاک تولید شده با

جدول ۲- راهنمای سطوح ژئومورفیک تفکیک شده در منطقه مطالعاتی (شکل ۱ ج)

Table 2. The legend of distinguished geomorphic units in the study area (Figure 1c)

Landscape	Landform	Lithology	Geomorphic surface	PGS	Symbol	Area of PGS (ha)		
Flood plain	River valley	Reworked loess	Meanders belt, erosional surface, cultivated		FP121	116.05		
	Terrace		Erosional or depositional terraces, cultivated	-	FP221	71.69		
	Loess bluff		Gully, uncultivated		FP321	11.96		
Valley	Plain	Loess	Moderately flat, cultivated		VA111	497.64		
			Moderately flat with sinkhole and gully erosion in some part, cultivated	-	VA112	246.13		
			Narrow valley with sinkhole and gully erosion in some part, cultivated		VA113	352.40		
Tableland	Loess tableland	Loess	A relatively flat- topped area with steep side slopes, cultivated	-	TB111	127.72		
Hills	High hill	Loess	Very steep complex slopes	South aspect, low density rangeland	Hi111_1	46.29		
				North aspect, dense rangeland	Hi111_2	144.08		
		Clay deposits and reddish brown LPL	Steep complex slopes	Complex aspect, no rangeland	Hi121_3	43.81		
		Limestone	Rock outcrops	Complex aspect, no rangeland	Hi131_3	52.20		
		Loess, marl mixed with siltstone	Steep complex slopes with dendritic drainage pattern and limestone outcrop	Complex aspect, low density rangeland	Hi141_4	118.92		
		Marl	Steep complex slopes with dendritic drainage pattern and gully	Complex aspect, no rangeland	Hi151_3	23.68		
		Marl dominant with partly shale and loess	Colluvium-mantled slopes with gully erosion, dendritic drainage network	Complex aspect, low density rangeland	Hi152_4	144.98		
		Loess dominant with partly clay deposits	Steep complex slopes associated with erosional characteristics and limestone outcrop	Complex aspect, moderate density rangeland	Hi161_6	43.44		
		Middle hill	Loess		Moderately complex slopes	North aspect, dense rangeland	Hi211_2	1202.58
						South aspect, low density rangeland	Hi211_1	1725.82
Very steep complex slopes with sinkhole in some part of footslope	North aspect, moderate dense rangeland				Hi212_5	131.55		
Very steep complex slopes associated with limestone outcrop	Complex aspect, low density rangeland				Hi213_4	8.77		
Reworked loess	Steep complex slopes with erosional characteristics and lithological discontinuity in some pedons				Complex aspect, low density rangeland	Hi221_4	174.59	

نتایج و بحث

تکوین و تکامل خاک‌ها

نتایج تشریح خاکرخ‌ها نشان داد که در منطقه سه راسته خاک اینسپتی‌سولز، انتی‌سولز و آلفی‌سولز وجود دارند. افزایش تنوع از راسته به فامیل خاک با افزایش زیادی همراه است که در بخش‌های بعدی به تفصیل شرح داده خواهد شد. از دلایل پایین بودن وضعیت تنوع زیر راسته و گروه بزرگ در منطقه مورد مطالعه، می‌توان به اقلیم خشک و نیمه‌خشک منطقه اشاره کرد. بیش‌ترین تنوع در سطح زیر گروه و فامیل خاک مربوط به راسته اینسپتی‌سولز می‌باشد که سطح وسیعی از منطقه مطالعاتی را به خود اختصاص داده است. افزایش تنوع در سطح فامیل خاک عمدتاً مربوط به تفاوت در مواد مادری، فرسایش و رسوب‌گذاری‌های ریز و درشت می‌باشد که در این سطح ظاهر شده است.

بایستی ادعان داشت که خاک مربوط به آلفی‌سولز که در زیر گروه مربوط Calcic Haploxeralfs قرار گرفته است، جزو خاک‌های قدیمی لسی با رنگ قرمز رنگ به حساب می‌آید، که در محیطی با متوسط بارندگی سالانه ۴۵۰-۶۵۰ میلی‌متر به وجود آمده‌اند و نشان‌دهنده اقلیم مرطوب‌تر زمان تشکیل آن‌ها نسبت به لس‌های اواخر پلیستوسن و خاک‌های مدرن می‌باشد. از جمله پدیده‌های شکل گرفته در این خاک‌ها تجمع رس، وجود افق آرجلیک، پرشدگی خلل و فرج خاک به صورت متراکم و گاهی غیرمتراکم از گچ ثانویه و تراکم اکسیدهای آهن و منگنز می‌باشد.

به طور کلی نتایج خصوصیات فیزیکی و شیمیایی خاک‌های منطقه نشان داد که با توجه به غالب بودن مواد مادری لسی در منطقه، کلاس بافتی خاک تغییرات چندانی ندارد و در بیشتر موارد در کلاس‌های لومی و لومی سیلتی قرار می‌گیرد. همچنین خاک‌های منطقه دارای کربنات کلسیم معادل زیاد، در بعضی نقاط دارای گچ زیاد و در بیشتر موارد غیر از شیب‌های شمالی دارای کربن آلی کم می‌باشد که همگی از خصوصیات خاک‌های مناطق خشک و نیمه‌خشک می‌باشد. ضمناً بایستی ذکر شود که خاک‌های منطقه دارای هدایت الکتریکی پایین می‌باشند.

مدل‌سازی مکانی خاک

کم‌ترین مقدار خطای تخمین نمونه‌های OOB که توسط الگوریتم RF در مدل‌سازی استفاده نشده‌اند و جهت اعتبارسنجی مستقل به کار می‌روند، ۳۲/۶۹ درصد بود که با تعداد ۵۰۰ درخت و ۴ متغیر در هر درخت برای سطح گروه بزرگ به دست آمد. بالعکس بیش‌ترین میزان خطای تخمین نمونه‌های OOB و کم‌ترین میزان شاخص کاپا و صحت کلی برای سطح فامیل خاک به دست آمد (جدول ۳). استام و همکاران (Stum *et al.*, 2010) با ۶۷۲ مشاهده در سطحی معادل ۴۷۰۰۰ هکتار، مقدار خطای تخمین OOB به دست آمده در سطح فامیل خاک را برای الگوریتم RF ۵۵/۲ درصد گزارش کردند. برونگارد (Brungard, 2009) با ۳۰۰ نمونه خاکرخ در سطح ۳۰۰۰۰ هکتار در سطح زیر گروه بزرگ خاک مقدار خطای تخمین ۶۲ درصد را به دست آورد.

همچنین مطابق با جدول ۳، بهترین نتایج شاخص کاپا و صحت کلی برای سطح گروه بزرگ می‌باشد. زراعت پیشه و همکاران (Zeraatpisheh *et al.*, 2017) نیز میزان شاخص کاپا ۰/۶۰، ۰/۴۸، ۰/۴۴ و ۰/۳۶ را به ترتیب برای سطوح راسته، زیر راسته، گروه بزرگ و زیرگروه در منطقه بروجن از استان چهارمحال و بختیاری با تعداد ۱۰۰ نمونه خاکرخ گزارش نموده‌اند. پهلوان‌راد و همکاران (Pahlevan Rad *et al.*, 2014) مقدار OOB ۴۸/۵ درصد برای سطح گروه بزرگ، ۵۱/۵ درصد برای سطح زیرگروه، و ۵۶/۵ درصد برای سطح سری خاک در محدوده مطالعاتی به وسعت ۸۵۰۰۰ هکتار از استان گلستان را مشاهده نموده‌اند. ضمناً نتایج صحت کلی آن‌ها ۴۸/۵، ۵۱/۵ و ۵۶/۶ درصد برای سطوح گروه بزرگ، زیر گروه و سری خاک بوده است. نتایج صحت کلی میرکزی و همکاران (Mirakzehi *et al.*, 2018) نیز در پیش‌بینی نقشه کلاس خاک، بخشی از منطقه سیستان با استفاده از الگوریتم RF به ترتیب برای سطوح گروه بزرگ، زیر گروه و فامیل خاک عبارتند از: ۴۶/۰۰، ۴۴/۰۰ و ۴۶/۴۰ درصد و شاخص کاپا ۰/۰۶، ۰/۰۴ و ۰/۱۳ می‌باشد.

مقایسه نتایج به دست آمده از این مطالعه با نتایج سایر پژوهشگران که در بالا ذکر گردید، نشان می‌دهد که خطای به دست آمده کم‌تر از بیش‌تر آن‌ها بوده است و این در شرایطی است، که از نظر تعداد مشاهدات خاکرخ

جدول‌های ۴ تا ۶ نشان می‌دهد که عموماً کلاس‌هایی که تعداد نقاط مشاهداتی بیش‌تری داشتند، دارای خطای بیش‌بینی کم‌تری هستند. از این رو نتایج جدول ۳ و ۴ گویای این مطلب است که سطح گروه بزرگ با فراوانی تعداد نمونه بیشتر در هر کلاس مربوطه خاک، دارای دقت بیش‌تری می‌باشد. همان‌طور که از نتایج جدول ۵ و شکل ۲ واضح است، بیش‌ترین تعداد فراوانی خاکرخ‌ها مربوط به کلاس‌های Typic Haploxerepts و Typic Xerorthents بود که نیمی از محدوده مطالعاتی را شامل می‌شوند و با مقدار خطای ۳۵/۲۰ و ۱۷/۶۰ درصد دارای کم‌ترین خطای بیش‌بینی در سطح زیر گروه بوده‌اند. جعفری و همکاران (Jafari et al., 2013) نتایج مشابهی را گزارش دادند که در آن کلاس‌های خاک با نمونه کم‌تر دارای خطای بیش‌بینی بیش‌تری بودند. محققان دیگری (همچون Pahlevan Rad et al., Hengl et al., 2007; Brungard, 2009; Kempen et al., 2009; 2014) نیز بیان می‌دارند که کلاس‌های خاک با تعداد فراوانی بیش‌تر دارای صحت تخمین بالاتری هستند.

کم‌تر از سایر پژوهشگران و سطح رده‌بندی خاک هم در سطوح پایین‌تری بوده است. دلیل اصلی این امر را تهیه شدن متغیرهای کمکی با دقت بالا به‌خصوص نقشه ژئومورفولوژی و برخی خصوصیات توپوگرافی دانست که در بخش اهمیت متغیرها به طور کامل بحث خواهد گردید. ژو و همکاران (Zhu et al., 2008) نیز اقدام به تهیه نقشه رقومی خاک در سطح زیر گروه برای منطقه‌ای با وسعت ۶۰ کیلومتر مربع در کشور چین با تعداد داده خاک محدود نمودند. آن‌ها فقط ۴۵ نقطه را در صحرا مورد بازدید زمینی قرار دارند که ۳۴ نقطه از این تعداد در نقشه نهایی با نقاط مشاهده شده در صحرا همخوانی داشت و نقشه خروجی دارای صحت ۷۶ درصد می‌باشد. آنان همچنین بیان می‌دارند که صحت اغلب نقشه‌های با مقیاس ۱:۲۴۰۰۰ در آمریکا حدود ۶۰ تا ۷۶ درصد می‌باشد. بنابراین، آن‌ها معتقدند که استفاده از نظر فرد خیره، بازدید خاک‌ها و برداشت نمونه کم در مناطقی که دارای مشابهت خاکرخ زیادی هستند، یک روش نمونه‌برداری سودمند به شمار می‌رود. لازم به ذکر است نتایج خطای تخمین کلاس‌های بیش‌بینی شده در

جدول ۳- نتایج خطای OOB، شاخص کاپا و صحت کلی نقشه برای سطوح مختلف رده‌بندی براساس سیستم رده‌بندی آمریکایی

Table 3. The result of OOB error, kappa index and overall accuracy of map for different soil Taxonomic levels

Taxonomic level	Frequency	OOB error (%)	Overall accuracy	Kappa index
Great Group	4	32.69	0.91	0.83
Subgroup	6	50.00	0.78	0.56
Family	10	65.38	0.50	0.32

جدول ۴- نتایج خطای OOB برای هر کلاس در سطح گروه بزرگ

Table 4. The result of OOB error for class of great group level

Great Group	Frequency	Area (ha)	OOB error (%)
Haploxerepts	27	2648.82	26.08
Xerorthents	29	2249.68	18.18
Calcixerepts	4	65.34	100.00
Haploxeralfs	4	60.40	100.00

جدول ۵- نتایج خطای OOB برای هر کلاس در سطح زیر گروه

Table 5. The result of OOB error for class of subgroup level

Subgroup	Frequency	Area (ha)	OOB error (%)
Gypsic Haploxerepts	7	223.41	100.00
Typic Haploxerepts	20	2224.37	35.20
Typic Xerorthents	26	2420.98	17.60
Typic Calcixerepts	4	42.61	100.00
Calcic Haploxeralfs	4	109.12	100.00
Lithic Xerorthents	3	3.75	100.00

جدول ۶- نتایج خطای OOB برای هر کلاس در سطح فامیل

Table 6. The result of OOB error for class of family level

Family	Frequency	Area (ha)	OOB error (%)
Coarse-loamy, mixed, active, thermic, Gypsic Haploxerepts	4	402.19	75.00
Fine-loamy, mixed, active, thermic, Typic Haploxerepts	16	1957.53	38.46
Fine-loamy, mixed, active, thermic, Gypsic Haploxerepts	3	8.19	100.00
Fine-loamy, mixed, active, thermic, Typic Xerortent	13	1404.13	66.66
Coarse-loamy, mixed, active, thermic, Typic Xerortents	10	582.79	57.14
Coarse-Loamy, mixed, active thermic, Typic Haploxerepts	4	155.04	75.00
Fine-loamy, mixed, active, thermic, Typic Calcixerepts	4	188.54	100.00
Fine, mixed, active, thermic, Calcic Haploxeralfs	4	104.07	100.00
Fine-loamy, mixed, active, thermic, Lithic Xerortents	3	11.37	100.00
Fine, mixed, active, thermic, Typic Xerortents	3	207.35	33.33

نتایج اهمیت متغیرها در نقشه‌های تولید شده توسط الگوریتم RF در شکل ۳ ارائه شده است که براساس آن، متغیرهای دارای بیشترین اهمیت در تخمین نقشه به دست آمده، نمایش داده شده است. ژئومورفولوژی، جهت شیب و ارتفاع در هر دو روش میانگین کاهش دقت و میانگین کاهش ضریب جینی در همه سطوح رده‌بندی جزو با اهمیت‌ترین متغیرها در تخمین نقشه تولید شده می‌باشند. بنابراین براساس نتایج به دست آمده از مدل‌سازی، نقش انکارناپذیر فرآیندهای ژئومورفولوژی در این مطالعه تایید می‌شود. اثر مثبت استفاده از نقشه ژئومورفولوژی در مطالعات قبلی نیز گزارش شده است (Jafari *et al.*, Behrens *et al.*, 2005; Zeraatpisheh *et al.*, 2017; Taghizadeh- 2013; Mehrjardi *et al.*, 2014). معمولا در مناطق خشک و نیمه‌خشک، تغییرپذیری خصوصیات و انواع خاک اغلب متأثر از مواد مادری و موقعیت توپوگرافی هستند (Tajik *et al.*, 2012) که بسیار خوب توسط نقشه ژئومورفولوژی نمایش داده می‌شود. این موضوع را می‌توان این‌گونه توضیح داد که سطوح ژئومورفیک در طول یک دوره زمین‌شناسی، همراه با شرایط تشکیل خاک نزدیک به فرآیندهای خاک‌سازی کنونی در مناطق خشک و نیمه‌خشک تشکیل شده‌اند.

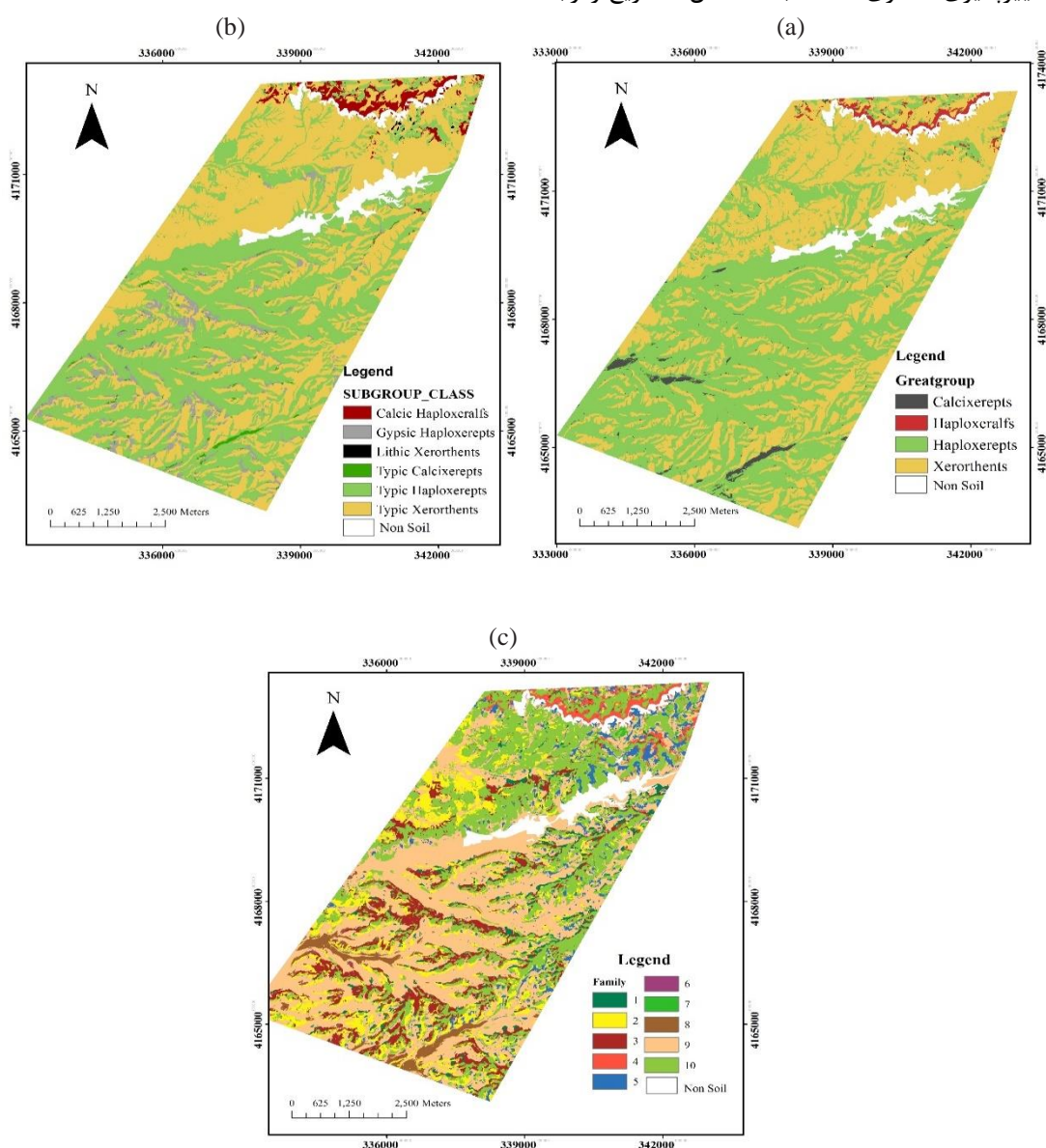
بنابراین حضور خصوصیات توپوگرافی به همراه فرآیندهای ژئومورفیک در پیش‌بینی کلاس‌های خاک نشان داد که این فاکتورها برآیند و نتیجه نیروهای موثر خاک‌سازی در منطقه مورد مطالعه هستند که با توجه به وضعیت توپوگرافی منطقه و نظر کارشناس خبره در جداسازی خاک‌ها، بسیار همخوانی دارد. نتایج برون‌گارد و همکاران (Brungard *et al.*, 2015) نشان داد که پیش‌بینی‌های مکانی دقیق‌ترین مدل‌ها معمولا با روابط

ضمن عملکرد پیش‌بینی برای کلاس‌های فامیل خاک که دارای تعداد فراوانی ۱۶ و ۱۳ خاکرخ و بخش وسیع‌تری از منطقه مطالعاتی را پوشش می‌دهند (شکل ۲ ج)، تفاوت قابل ملاحظه‌ای با مواردی داشت که دارای تعداد نمونه کمتر حدود ۳ یا ۴ مورد بودند (جدول ۶) و احتمالا به تعداد نمونه بیشتر این فامیل‌های خاک مربوط می‌شود. بنابراین، به نظر می‌رسد مدل به تعداد نمونه حساس هستند. به طوری که هرچه داده‌های بیشتری برای برازش مدل استفاده شود، تغییرپذیری کلاس‌های خاک و پیش‌بینی آن‌ها بهتر نشان داده می‌شود. بنابراین، اندازه نمونه نقش اساسی را در فرآیند ارزیابی دقت در این سطح رده‌بندی ایفا می‌کند. لازم به ذکر است نتایج جدول ۶ نشان می‌دهد که دقت پیش‌بینی برای همه انواع خاک‌ها یکسان نیست؛ برخلاف تعداد کم نمونه در کلاس فامیل خاک Fine, mixed, active, thermic, Typic Xerortents نتیجه مطلوبی حاصل شده است. این خلوص بالا در نتیجه ارتباط قوی بین پیش‌بینی‌کننده‌ها و این کلاس می‌باشد. این ارتباط قوی در اثر تمایز شدید خصوصیات متغیرهای کمکی برای این کلاس در مقایسه با بقیه کلاس‌ها با تعداد نمونه کم است. بنابراین مطالعه حاضر نشان داد که استفاده از متغیرهای محیطی با دقت بالا همراه با تعداد نمونه خاک کم می‌تواند نتایج مطلوبی را برای مناطق این چنینی ارائه دهد. اگرچه نتایج خطا حاکی از آن است که مدل RF برای پیش‌بینی نقشه سطح خاک در حد فامیل نیاز به تعداد داده بیشتری برای دستیابی به نتایج با دقت بالاتر خواهد داشت.

اهمیت متغیرها در تخمین کلاس‌های خاک با RF

پیش‌بینی‌کننده و متغیر وابسته را تشخیص می‌دهند و ارتباط برقرار می‌کنند. این موضوع در مورد هماهنگی توزیع مکانی زیر گروه Typic و Typic Haploxerepts و Xerorthents و شاخص سنجش از دور همچون SAVI و خصوصیات سطحی مؤثر بر بازتاب نوری و همچنین خصوصیات سرزمین مانند شیب، جهت شیب و درجه خیزی کاملاً صادق است.

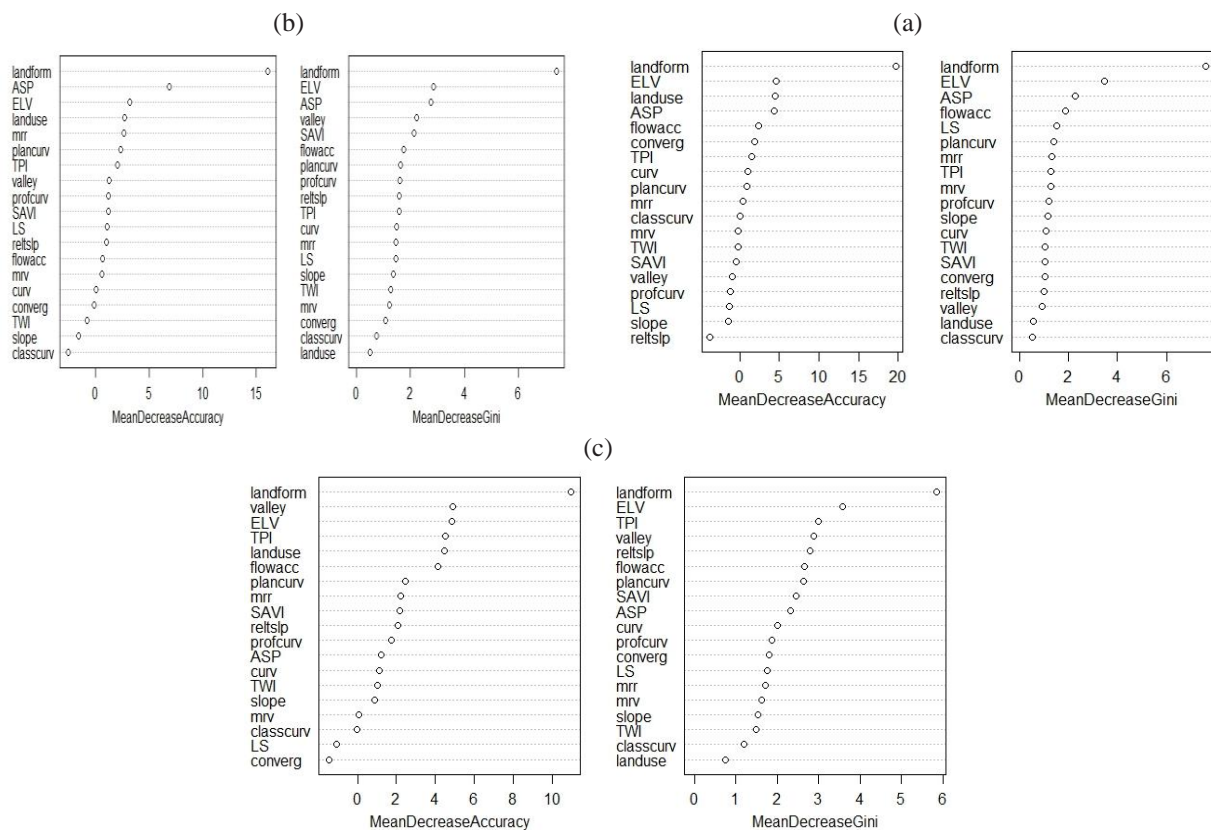
خاک- زمین‌نما و خصوصیات توپوگرافی مورد انتظار در توافق بودند. به طوری که کمترین عدم قطعیت پیش‌بینی مکانی در مناطق نسبتاً پست در منطقه مورد مطالعه آن‌ها برآورد شد. از این رو زمانی که پارامترهای حاصل از توپوگرافی، سنجش از دور و فرآیندهای ژئومورفولوژی همه با هم در پیش‌بینی شرکت کنند، بهترین نتایج پیش‌بینی حاصل می‌شود. به نظر می‌رسد جایی که توزیع پیش‌بینی‌کننده‌ها یکنواخت‌تر باشد و تغییرپذیری کمتری داشته باشد، مدل‌ها سریع‌تر رابطه



شکل ۲- نقشه پیش‌بینی خاک (الف) گروه بزرگ، (ب) زیر گروه، (ج) فامیل

Figure 2. Predicted soil map (a) Great group, (b) Subgroup, (c) Family

Code of Family: 1. Coarse- Loamy, mixed, active thermic, Typic Haploxerepts, 2. Coarse-loamy, mixed, active, thermic, Calcic Haploxeralfs, 3. Coarse-loamy, mixed, active, thermic, Gypsic Haploxerepts, 4. Fine, mixed, active, thermic, Typic Xerorthents, 5. Fine, mixed, active, thermic, Lithic Xerorthents, 6. Fine-loamy, mixed, active, thermic, Typic Calcixerepts, 7. Fine-loamy, mixed, active, thermic, Typic Haploxerepts, 8. Fine-loamy, mixed, active, thermic, Typic Xerorthents, 9. Fine-loamy, mixed, active, thermic, Typic Calcixerepts, 10. Fine-loamy, mixed, active, thermic, Typic Xerorthent



شکل ۳- اهمیت متغیرها به دو روش میانگین کاهش دقت و میانگین کاهش ضریب جینی در (الف) سطح گروه بزرگ، (ب) سطح زیر گروه، (ج) سطح فامیل خاک. نماد کلمات در جدول ۱ توصیف شده است.

Figure 3. Covariate importance in two method mean decrease of accuracy and mean decrease of Gini at the (a) Great group level (b) Subgroup level (c) Family level. Symbol of covariates are explained in Table 1

یکی از مهمترین متغیرهای ورودی به مدل نقش بسیار بسزایی در افزایش صحت نقشه خاک تولید شده داشته است. ضمن این که بهترین انتخاب در تفکیک قدرت مکانی DEM در هر منطقه، وابسته به علم پدولوژی و تشخیص خصوصیات پدوژن و عوامل تاثیرگذار در تشکیل و تکامل خاک خواهد بود. در این صورت با داشتن یک DEM مناسب و دانش کیفی و کمی خاک نتیجه بسیار مناسبی در بهبود تخمین خاک حاصل خواهد شد. پیشنهاد می گردد در مطالعات آتی، تصویر بزرگتری از منطقه مطالعاتی توسط پرنده های بدون سرنشین (تکنیک UAV) جهت تهیه DEM منطقه انجام گردد و نتایج حاصل از متغیرهای توپوگرافی و نقشه سطح ژئومورفیک آن برای پیش بینی نقشه خاک استفاده گردد. در نهایت نقشه حاصل از این رویکرد با نقشه خاک تهیه شده از DEM بدست آمده از روش های دیگر (DEM های استخراج شده از تصاویر ماهواره همانند SRTM ASTER, Pléiades stereo imagery, و

نتیجه گیری کلی

نتایج این مطالعه نشان می دهد، ویژگی های توپوگرافی و سطوح ژئومورفیک هم به لحاظ منطقی و ریاضی، و هم از نظر تجربی دارای رابطه نزدیکی با ویژگی های خاک در منطقه هستند. بنابراین استفاده از این ویژگی ها با دقت و قدرت تفکیک مکانی بالا در شناسایی خاک برخی مناطق، می تواند تا حد زیادی ضمن افزایش دقت کار و صرفه جویی در هزینه و زمان، اطلاعات ارزشمندی را در اختیار مدیران، برنامه ریزان و کارشناسان به منظور اعمال مدیریت صحیح در حفاظت از منابع خاک و آب قرار دهد. نتایج این تحقیق نشان داد که استفاده از DEM سازمان نقشه برداری کشور و تصحیح آن با نقاط ارتفاعی زمینی با توجه به استخراج مولفه های توپوگرافی، نتایج قابل قبولی بخصوص در سطوح بالای رده بندی ارائه داده است. از دیگر سو، نباید فراموش کنیم که نقش تصاویر ماهواره گوگل ارث و تهیه نقشه سطوح ژئومورفیک از آن ها در سطح فاز اراضی، به عنوان

توجه به مطالعه حاضر، با توسعه دانش ارتباط دقیق خاک و متغیرهای محیطی می‌توان با استفاده از تعداد نمونه خاک کمتر، کاهش تراکم نمونه‌برداری و داده‌های کمکی با دقت بالا در برخی از مناطق ایران، که اطلاعات خاکی قابل دسترس اندکی موجود می‌باشد، نقشه‌برداری رقمی خاک را روش مناسبی برای مطالعات آتی نقشه‌برداری در کشور دانست.

حاصل از نقشه‌های توپوگرافی با مقیاس ۱:۵۰۰۰۰ و ۱:۲۵۰۰۰ مقایسه شود تا کارایی روش استفاده از تعداد نمونه کمتر در تهیه نقشه پیش‌بینی خاک منطقه بخصوص در سطوح پایین رده‌بندی بهتر مشخص گردد. علاوه بر این، استفاده از نظرات افراد کارشناس، به ویژه در مواقعی که نمونه‌های موجود برای مدل کافی نباشند، و یا نماینده خوبی برای معرفی منطقه مورد نظر نیستند، می‌تواند در مناطق این چنینی کارگشا باشد. بنابراین با

References

- Abbaszadeh Afshar F., Ayoubi Sh., and Jafari A. 2018. The extrapolation of soil great groups using multinomial logistic regression at regional scale in arid regions of Iran. *Geoderma*, 315: 36–48.
- Bagheri Bodaghabadi M., Esfandiarpour Borujeni I., Salehi M.H., Mohammadi J., and Toomanian N. 2015. Assessment of the expert knowledge's effect in digital soil mapping and soil sampling. *14th Congress of Soil Science, Rafsanjani University, Iran*. (In Persian)
- Banaei M.H. 2000. The map of resources and land capability of Iran soils. Soil and Water Research Institute, Karaj, Iran. (In Persian)
- Behrens T., Förster H., Scholten T., Steinrücken U., Spies E.D., and Goldschmitt M. 2005. Digital soil mapping using artificial neural networks. *Journal of Plant Nutrition and Soil Science*, 168(1): 21-33.
- Breiman L. 2001. Random forests. *Machine Learning*, 45(1): 5-32.
- Breiman L., and Cutler A. 2004. Random Forests homepage. Retrieved April 23rd.
- Brungard C.W. 2009. Alternative sampling and analysis methods for digital soil mapping in southwestern Utah. Thesis for Master of Science, Utah State University, USA, 284p.
- Brungard C.W., Boettinger J.L., Duniway M.C., Wills S.A., and Edwards Jr T.C. 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*, 239-240: 68-83.
- Brus D.J., Kempen B., and Heuvelink G.B.M. 2011. Sampling for validation of digital soil maps. *European Journal of Soil Science*, 62: 394–407.
- Frechen M., Kehl M., Rolf C., Sarvati R., and Skowronek, A. 2009: Loess chronology of the Caspian Lowland in Northern Iran. *Quaternary International*, 128(1-2): 220-233.
- Gallant J.C., and Dowling T.I. 2003. A multi resolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research*, 39: 1347-1360.
- Häring T., Dietz E., Osenstetter S., Koschitzki T., and Schröder B. 2012. Spatial disaggregation of complex soil map units: a decision-tree based approach in Bavarian forest soils. *Geoderma*, 185–186: 37–47.
- Hengl T., Heuvelink G.B.M., and Stein A. 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*, 120(1-2): 75-93.
- 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.
- Huete A.R. 1988. A soil adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25: 295–309.
- Jafari A., Ayoubi S., Khademi H., Finke P.A., and Toomanian N. 2013. Selection of a taxonomic level for soil mapping using diversity and map purity indices: A case study from an Iranian arid region. *Geomorphology*, No of Pages 12.
- Kehl M., Sarvati R., Ahmadi H., Frechen M., and Skowronek A. 2005. Loess paleosol-sequences along a climatic gradient in Northern Iran. *Eiszeitalter u. Gegenwart*, 55: 149-173.
- Kempen B., Brus D.J., Heuvelink G.B.M., and Stoorvogel J.J. 2009. Updating the 1:50000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma*, 151: 311-326.
- Khormali F., and Kehl M. 2011. Micromorphology and development of loess-derived surface and buried soils along a precipitation gradient in Northern Iran. *Quaternary International*, 234: 109–123.

- MathWorks. 2009. Matlab. The Math Works., Inc., Natick, MA.
- Maleki S., Khormali F., Bagheri Bodaghabadi M., Mohammadi J., Kehl M., Hoffmeister D., Ghaffary M. 2017. Using Unmanned Aerial Vehicle in future studies of digital soil mapping? Accuracy, coverage and the effects on preparing of geomorphology map. *15th Soil Congress*, University of Isfahan, Iran. (In Persian)
- Marchetti A., Piccini C., Santucci S., Chiuchiarelli I., and Francaviglia R. 2011. Simulation of soil types in Teramo province (central Italy) with terrain parameters and remote sensing data. *Catena*, 85: 267-273.
- McBratney A.B., Mendonc Santos M.L., and Minasny B. 2003. On digital soil mapping. *Geoderma*, 117: 3-52.
- Mirakzahi K.h., Pahlavan-Rad M.R., Shahriari, A., and Bameri, A. 2018. Digital soil mapping of deltaic soils: A case of study from Hirmand (Helmand) river delta. *Geoderma*, 313: 233–240.
- Minasny B. and McBratney A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*, 32: 1378–1388.
- Myles A.J., Feudale R.N., Liu Y., Woody N.A., and Brown S.D. 2004. An introduction to decision tree modeling. *Journal of Chemometrics*, 18(6): 275–285.
- 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.
- Pahlavan-Rad M.R., Khormali F., Toomanian N., Brungard C.W., Kiani F., and Komaki C.B., Bogaert P. 2016. Legacy soil maps as a covariate in digital soil mapping: A case study from Northern Iran. *Geoderma*, 279: 141–148.
- R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org>.
- Sreenivas K., Dadhwal V.K., Kumar S., Harsha G.S., Mitran T., Sujatha G., Janaki Rama Suresh G., Fyze M.A., and Ravisankar T. 2016. Digital mapping of soil organic and inorganic carbon status in India. *Geoderma*, 269: 160-173.
- Stoorvogel J.J., Kempen B., Heuvelink G.B.M., and Bruin S. 2009. Implementation and evaluation of existing knowledge for digital soil mapping in Senegal. *Geoderma*, 149: 161–170.
- Stum A.K., Boettinger J.L., White M.A., and Ramsey R.D. 2010. Random Forests applied as a soil spatial predictive model in arid Utah, P 179-189. In: Boettinger J.L., Howell D., Moore W, A.C., Hartemink A., Kienast-Brown E.S. (Ed.), *Digital Soil Mapping: Bridging Research, Environmental Application, and Operation*. Progress in Soil Science. Springer, Logan, USA.
- Soil Survey Staff. 1996. Soil survey laboratory methods manual. Report No. 42, USDA, NRCS, NCSS.
- Soil Survey Staff. 2014. Keys to Soil Taxonomy (12th Ed.), U.S. Department of Agriculture, Natural Resources Conservation Service. 372p.
- Taghizadeh-Mehrjardi R., Minasny B., Sarmadianc F., and Malone B.P. 2014. Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma*, 213: 15–28.
- Tajik S., Ayoubi S., and Nourbakhsh F. 2012. Prediction of soil enzymes activity by digital terrain analysis: comparing artificial neural network and multiple linear regression models. *Environmental Engineering Science*, 29(8): 798-806.
- Teng H., Viscarra Rossel R.A., Shi Zh., and Behrens Th. 2018. Updating a national soil classification with spectroscopic predictions and digital soil mapping. *Catena*, 164: 125-134.
- Toomanian N., Jalalian A., Khademi H., Karimian Eghbal M., and Papritz A. 2006. Pedodiversity and pedogenesis in Zayandeh-rud Valley, Central Iran. *Geomorphology*, 81: 376–393.
- Wang X., Wei H., Khormali F., Taheri M., Kehl M., Frechen M., Lauer M., and Chen M. 2016. Grain-size distribution of Pleistocene loess deposits in northern Iran and its palaeoclimatic implications. *Quaternary International*, 1-11.
- Wang Sh., Jin X., Adhikari K., Li W., Yu M., Bian Zh., and Wang Q. 2018. Mapping total soil nitrogen from a site in northeastern China. *Catena*, 166: 134-146.
- Weiss A.D. 2001. Topographic position and landforms analysis. *Proceedings of the ESRI User Conference*, 9–13 July, San Diego, CA, USA.
- Wilson J.P., and Gallant J.C. 2000. Terrain analysis. Wiley & Sons, New York.

- Zeraatpisheh M., Ayoubi Sh., Jafari A., and Finke P. 2017. Comparing the efficiency of digital and conventional soil mapping to predict soil types in a semi-arid region in Iran. *Geomorphology*, 285: 186–204.
- Zhu A.X., Yang L., Li B., Qin Ch., English E., Burt J.E., and Zhou Ch. 2008. Purposive Sampling for Digital Soil Mapping for Areas with Limited Data. *In: Digital Soil Mapping with Limited Data, Part 12*. Springer Science, pp. 223- 245.

Evaluation of Accuracy of Digital Soil Mapping with Limited Data in a Part of Loess Plateau, Golestan Province

Sedigheh Maleki^{1*}, Farhad Khormali², Mohsen Bagheri Bodaghabadi³, Jahangir Mohammadi⁴

(Received: August 2018

Accepted: October 2018)

Abstract

Digital soil mapping approaches that require quantitative data for prediction are difficult to implement in countries with limited data on soil and auxiliary variables. Extensive field sampling is very labor intensive and costly that is problematic for mapping missions. On the other side, it is believed in digital soil mapping approaches that unique soil conditions (soil types or soil properties) can be associated with unique combination and configuration of environmental variables. In this study we used a Random Forest (RF) algorithm combined with classification information of 64 soil profiles and 19 environmental variables (including terrain attribute, geomorphology units, land use and vegetation cover index) to map soil classes in the part of loess plateau, Golestan Province Iran. Geomorphology, elevation, slope aspect and land use were the most important parameters in prediction of soil map in different taxonomic level. The results of accuracy assessment of RF with different entrance variables revealed that accuracy of model including overall accuracy and kappa index respectively decreased of 0.91 and 0.83 for great group, 0.78 and 0.56 for subgroup, 0.50 and 0.32 for family. The minimum and maximum Out of bag (OOB) estimate error rate in modeling were 32.69% and 65.38% for great group and soil family, respectively and the soil classes with higher frequency had the lower OOB error. The present study showed that in regions of Iran with limited data, digital soil mapping and high resolution ancillary data with smaller sample size can be led to an effective result in higher taxonomic levels.

Keywords: Environmental variables, Limited data, Map Accuracy, Random forest

Maleki S., Khormali F., Bagheri Bodaghabadi M, Mohammadi J. 2019. Evaluation of accuracy of digital soil mapping with limited data in a part of loess plateau, Golestan province. *Applied Soil Research*, 7(4): 148-163.

1. PhD graduated, Department of Soil Science, Faculty of Water and Soil Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

2. Professor, Department of Soil Science, Faculty of Water and Soil Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

3. Soil and Water Research Institute, Agricultural Research, Education and Extension Organization, Karaj, Iran

4. Assistant Professor, Department of Forestry Science, Faculty of Forestry, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

* Corresponding Author Email: elymaleki@yahoo.com