

مقایسه روش‌های شبکه عصبی مصنوعی و درخت تصمیم در تهیه نقشه رقومی خاک در منطقه اردکان

روح‌الله تقی‌زاده مهرجردی^{۱*}، فریدون سرمدیان^۲، محمود امید^۳، غلامرضا ثواقبی^۴، نورایر تومانیان^۵

محمدجواد روستا^۶، محمدحسن رحیمیان^۷

۱. استادیار، دانشکده کشاورزی و منابع طبیعی دانشگاه اردکان

۲، ۳، ۴. استادان، دانشکده مهندسی و فناوری کشاورزی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران

۵. استادیار، مرکز تحقیقات کشاورزی و منابع طبیعی اصفهان

۶، ۷. استادیار و مربی، مرکز ملی تحقیقات شوری

(تاریخ دریافت: ۱۳۹۱/۸/۲ - تاریخ تصویب: ۱۳۹۲/۳/۸)

چکیده

در پاسخ به تقاضای اطلاعات مکانی خاک، به‌کاربردن داده‌های کمکی رقومی و ارتباط آن‌ها با داده‌های مشاهداتی صحرایی در حال افزایش است. استفاده از اطلاعات رقومی از طریق روش‌های کامپیوتری، که اصطلاحاً نقشه‌برداری رقومی خاک خوانده می‌شود، قابل اعتمادتر و کم‌هزینه‌تر از روش‌های سنتی نقشه‌برداری خاک است. بنابراین، در پژوهش حاضر از مدل‌های درخت تصمیم و شبکه عصبی مصنوعی برای پیش‌بینی مکانی کلاس‌های تاکسونومیک خاک در منطقه‌ای خشک، به وسعت ۷۲۰ کیلومتر مربع در اردکان، استفاده شد. عملیات نقشه‌برداری در آنجا بسیار سخت است. در این منطقه بر اساس روش مربع لاتین مکان جغرافیایی ۱۸۷ پروفیل خاک مشخص و سپس تشریح و نمونه‌برداری شدند و طبقه‌بندی بر اساس سیستم امریکایی انجام گرفت. متغیرهای محیطی یا نمایندگان فاکتورهای خاک‌سازی، که در این پژوهش استفاده شد، شامل اجزای سرزمین، داده‌های تصویر ETM⁺ ماهواره لندست، و نقشه سطوح ژئومورفولوژی است. نتایج این تحقیق نشان داد، برای پیش‌بینی کلاس خاک، مدل درخت تصمیم نسبت به شبکه عصبی مصنوعی حدود هفت درصد دقیق‌تر است. مدل درختی و شبکه عصبی مصنوعی به ترتیب دقت کلی و ضریب تبیین و ضریب کاپای ۴۸ درصد، ۵۲ درصد، ۰/۳۴ و ۰/۴۶ و ۰/۱۳ و ۰/۲۵ دارند. نتایج نشان داد برای پیش‌بینی کلاس خاک شاخص خیسی، سطوح ژئومورفولوژی، و شاخص همواری دره با درجه تفکیک بالا مهم‌ترین پارامترها هستند. مدل‌های درختی نسبت به روش شبکه عصبی مصنوعی دقت بالاتری دارند و همچنین تفسیر نتایج مدل درختی بسیار راحت‌تر است. لذا پیشنهاد می‌شود در مطالعات آینده برای تهیه نقشه رقومی خاک از مدل‌های درختی استفاده شود.

کلیدواژگان: پیش‌بینی مکانی، گروه‌های بزرگ خاک، مربع لاتین، نقشه‌برداری رقومی.

مقدمه

در گذشته برای تهیه نقشه خاک و بررسی تغییرات خصوصیات آن اغلب روش‌های آماری معمول به‌کار می‌رفت، که در بیشتر مطالعات تغییرات ویژگی‌های خاک داخل واحدهای نقشه نادیده گرفته می‌شد (Rossiter et al., 2005). این روش‌ها نقشه‌های پلی‌گونی خاک را ایجاد می‌کنند که در حال حاضر منبع اصلی اطلاعات توزیع مکانی خصوصیات خاک است. اما استفاده از این نقشه‌ها ایرادهایی دارد که بیشتر به دلیل منطبق ناپیوسته‌ای است که در نقشه‌برداری خاک استفاده می‌شود؛ در صورتی که تغییر خصوصیات و حتی واحدهای تاکسونومیک خاک‌ها در یک منظره اراضی^۱ پیوسته است و شکستن این محیط پیوسته و

تبدیل آن به گروه‌های مکانی و طبقه‌ای مجزا باعث از دست رفتن مقدار قابل ملاحظه‌ای از اطلاعات می‌شود (Zhu, 1994; Webster and Oliver, 2007; McKenzie et al., 2008). از دیگر مشکلات روش معمول نقشه‌برداری خاک وابستگی تعیین حدود واحدهای خاک به قدرت درک کارشناس از روابط خاک و محیط است؛ به نحوی که می‌توان گفت احتمال اینکه چند کارشناس مرز واحدی از خاک‌های یک ناحیه ارائه دهند کم است. از سوی دیگر روش‌های معمول شناسایی خاک به نمونه‌برداری زیاد و تجزیه‌های آزمایشگاهی نیاز دارد و این‌گونه عملیات وقت و هزینه فراوان می‌طلبد (Zhu et al., 2001; McBratney et al., 2003).

با توجه به مشکلات موجود در نقشه‌های سنتی خاک و نیاز مبرهن به ارائه اطلاعات قابل اعتمادتر و به‌هنگام درباره خاک‌ها با هزینه منطقی و ارتقای تفسیر نتایج به نحوی که افراد

* نویسنده مسئول: rh_taghizade@yahoo.com

در حال حاضر، شناسایی و نقشه‌برداری خاک به روش‌های مرسوم در کشورهای توسعه‌یافته تقریباً به پایان رسیده و اطلاعات پایه به‌صورت رقومی برای پردازش با فنون نوین قابل دسترس‌تر است (McBratney *et al.*, 2003). ولی توقف و رکود فعالیت‌های پدولوژی در کشورهای در حال توسعه، مانند ایران، باعث شده اطلاعات خام و دقیق برای پردازش فراهم نباشد. نقشه خاک کشور با مقیاس ۱:۱۰۰۰۰۰۰ نیز، که اخیراً تهیه شده، با توجه به مقیاس، دقت و جزئیات بالایی ندارد. در حقیقت، مطالعات نقشه‌برداری رقومی خاک در ایران در ابتدای راه است. بنابراین، لزوم تهیه نقشه‌های خاک با مقیاس بزرگ در مناطقی که اهمیت و اولویت بیشتری دارند با استفاده از اطلاعات و روش‌های موجود و به‌کارگیری روش‌های نوین احساس می‌شود. پژوهش حاضر با این رویکرد طراحی شده تا برای تهیه نقشه رقومی واحدهای خاک در منطقه خشک ایران مرکزی اقدام کند.

مواد و روش‌ها

تشریح محدوده مطالعاتی

منطقه مورد مطالعه در ایران مرکزی واقع است و ۷۲۰۰۰ هکتار از اراضی دشت یزد- اردکان را شامل می‌شود. این منطقه بین طول‌های جغرافیایی $53^{\circ}50'$ تا $54^{\circ}10'$ شرقی و عرض‌های جغرافیایی $32^{\circ}16'$ تا $32^{\circ}27'$ شمالی قرار دارد. منابع آب و خاک این منطقه شور است. میانگین تبخیر و تعرق مرجع، دما، و بارندگی سالانه این منطقه به‌ترتیب برابر ۳۴۸۳ میلی‌متر، ۱۸/۵ درجه سانتی‌گراد، و ۷۵ میلی‌متر است. بارندگی بسیار کم و محدود به بارش‌های زمستانه است. بنابراین، رژیم رطوبتی و حرارتی خاک به‌ترتیب اریدیک و ترمیک طبقه‌بندی می‌شود. منطقه مورد مطالعه از نظر توپوگرافی بیشتر مسطح است و فقط رشته‌کوه‌هایی در شمال شرق منطقه مشاهده می‌شود. بلندترین ارتفاعات منطقه ۱۹۴۴ متر و کم‌ارتفاع‌ترین قسمت منطقه، که در نواحی میانی قرار گرفته، ۹۴۴ متر است. لندفرم‌های اصلی این منطقه، با روند کلی شرق به غرب، شامل کوه و مخروط‌افکنه‌های جدید و سطوح رسی و تپه‌های گچی است.

نمونه‌برداری

هایپرکیوب^۱

این روش طرحی تصادفی و طبقه‌بندی‌شده است که باعث نمونه‌برداری مؤثر به کمک توزیع چند متغیر می‌شود. یک شبکه مربع، حاوی موقعیت‌های نمونه، شبکه‌ای لاتین است اگر فقط

غیر خاک‌شناس و غیر متخصص نیز بتوانند از این اطلاعات استفاده کنند، استفاده از روش‌های نوین در تهیه نقشه خاک ضروری می‌نماید. امروزه مشخص شده دقت روش‌های پدومتری^۱ برآورد توزیع مکانی داده‌های اندازه‌گیری‌شده خاک، به‌علت در نظر گرفتن پیوستگی مکانی داده‌ها، نسبت به روش‌های معمول آماری بیشتر است (Grunwal, 2009; Rossiter and Loza, 2010; McBratney *et al.*, 2000) و نقشه‌برداری رقومی خاک^۲ (McBratney *et al.*, 2003) به‌منزله زیرمجموعه‌ای از روش‌های پدومتری، به دلیل افزایش منابع داده‌های کمکی^۳ یا محیطی فراوان، به‌صورت کاملاً عملی توانسته است به انتقادات وارده بر خاک‌شناسی مرسوم فایق آید.

تا به حال روش‌های متعددی برای پیش‌بینی مکانی کلاس‌های خاک ارائه شده است؛ مثلاً، در تحقیقی که در کشور آلمان صورت گرفته توانایی برخی تکنیک‌ها و الگوریتم‌های داده‌کاوی در پیش‌بینی کلاس‌های خاک مقایسه شدند. در جمع‌بندی کلی از این تحقیق محققان گزارش کردند بهترین رویکرد و الگوریتم در برآورد کلاس خاک شبکه‌های عصبی مصنوعی است. از سوی دیگر، آن‌ها اعلام کردند روش کوانتیزه کردن آرایه تعلیم را به‌هیچ‌وجه برای برآورد کلاس‌های خاک در این منطقه توصیه نمی‌کنند (Behrens *et al.*, 2005). Moonjun *et al.* (2010) تکنیک شبکه عصبی مصنوعی و درخت تصمیم را برای پیش‌بینی واحدهای خاک مقایسه کردند. ایشان نشان دادند بین دو تکنیک برای پیش‌بینی تفاوتی معنادار وجود ندارد. علاوه بر این، ایشان نقشه ژئومورفولوژی را متغیر کمکی مناسبی معرفی کردند. Luoto and Hjort (2005) چندین روش نقشه‌برداری، از جمله رگرسیون و تصمیم‌گیری درختی و شبکه عصبی، را برای نقشه‌برداری سطوح ژئومورفیک در فنلاند به کار بردند. نتایج کارآیی بالاتر درخت تصمیم را نشان داد.

Kheir *et al.* (2010) درخت تصمیم و کریجینگ شاخص را، برای پیش‌بینی ماده آلی خاک، در کشور دانمارک مقایسه کردند و نتیجه گرفتند که درخت تصمیم مناطق با ماده آلی بالا را بهتر طبقه‌بندی می‌کند. در ایران نیز در سال‌های اخیر محققان در زمینه نقشه‌برداری رقومی خاک تلاش‌هایی کرده‌اند؛ از جمله Hengl *et al.* (2007)، Bagheri *et al.* (2011)، Jafari *et al.* (2012)، Hamzhepour *et al.* (2013).

1. Pedometric techniques

2. Digital soil mapping

3. Ancillary data

4. Latin hypercube method

تشخیص تفاوت‌ها، استفاده شد. سپس عکس‌های هوایی و روکش مربوطه اسکن شد و در محیط سامانه جغرافیایی (پی‌سی‌آی ژئوماتیکا^۱) قرار گرفت. در مراحل بعد، عکس‌ها و روکش‌های داخل سامانه جغرافیایی، با استفاده از تصاویر ماهواره‌ای، به روش ارتوفوتوژئورفرنسینگ^۲ (Rossiter and Hengl, 2001) زمین مرجع^۳ شدند. با رقومی کردن مرزهای ترسیم‌شده بر روکش عکس‌ها، نقشه ژئومورفولوژی تهیه شد.

استخراج پارامترهای زمین‌نما

در تحقیق حاضر مدل رقومی سازمان نقشه‌برداری کشور، با قدرت تفکیک مکانی ۱۰ متر، به کار رفت. بعد از آماده کردن مدل ارتفاعی رقومی، از آن برای استخراج اطلاعات کمکی یا تصاویر تخمین‌زنده متغیرهای ژنتیکی خاک استفاده شد. پارامترهای زمین‌نما- از قبیل شیب^۴، ارتفاع^۵، ارتفاع بالای شبکه زهکشی^۶، مساحت حوزه‌های اصلاح‌شده^۷، موقعیت میانی شیب^۸، عمق دره^۹، شاخص خیسی^{۱۰}، شاخص همواری دره با درجه تفکیک بالا^{۱۱}، شاخص بالای پشته با درجه تفکیک بالا^{۱۲}، و شیب حوضه^{۱۳} - در محیط سامانه جغرافیایی ساگا^{۱۴} محاسبه و استخراج شد. روش استخراج همه پارامترهای مزبور در روش ارائه‌شده را Hengl *et al.* (2004) تشریح کرده‌اند.

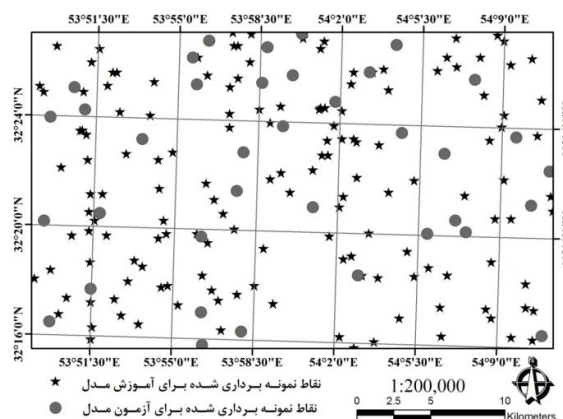
تصاویر ماهواره‌ای

خاک‌های گوناگون خصوصیات طیفی متفاوت دارند (Andronikov and Dorbrolv'skiy, 1991). با توجه به اینکه بیشتر منطقه مطالعاتی فاقد پوشش گیاهی است، تغییر خصوصیات و نوع خاک‌ها به راحتی با اطلاعات ماهواره‌ای قابل تشخیص است (Metternicht and Zinck, 2003). در این مطالعه از تصاویر ماهواره‌ای سنجهده ETM⁺ استفاده شد. برای کاهش تعداد داده‌های خام (تعداد باندها) تصویر ماهواره‌ای از یک تجزیه به مؤلفه‌های اصلی، بر اساس ماتریس همبستگی، استفاده شد. بر اساس فاکتورهای خاک‌سازی منطقه، شاخص‌هایی از قبیل شاخص گیاهی نرمال‌شده^{۱۵} (نشان‌دهنده

یک واحد نمونه‌برداری در هر ردیف و هر ستون وجود داشته باشد. یک هایپرکیوب لاتین^{۱۶} تعمیم این مفهوم به تعداد اختیاری ابعاد است؛ به طوری که هر واحد نمونه‌برداری فقط در یک صفحه آن قرار گرفته باشد. LHS شامل نمونه‌برداری n ولیو از توزیع تشریح‌شده هر متغیر است. توزیع تجمعی هر متغیر به n فاصله با احتمال مساوی تقسیم می‌شود و یک ولیو از هر فاصله به طور تصادفی انتخاب می‌شود. سپس n ولیوی به دست آمده برای هر متغیر با متغیرهای دیگر جفت می‌شود. این روش پوششی کامل از هر متغیر ایجاد می‌کند (Minasny and McBratney, 2006). بر اساس تکنیک هایپرکیوب محل ۱۸۷ پروفیل در منطقه اردکان به وسعت ۷۲۰۰۰ هکتار انتخاب شد (شکل ۱).

تجزیه آزمایشگاهی

از لایه‌های ژنتیکی کلیه مقاطع نمونه‌برداری و نمونه‌ها برای تجزیه به آزمایشگاه منتقل شد. تعدادی از ویژگی‌های فیزیکی و شیمیایی نمونه‌ها چنین اندازه‌گیری شدند: بافت خاک به روش هیدرومتری (Gee and Bauder, 1986)، واکنش نمونه‌های خاک در گل اشباع به وسیله pH متر، قابلیت هدایت الکتریکی عصاره اشباع به وسیله هدایت‌سنج، کربنات کلسیم معادل خاک به روش تیتراسیون برگشتی با اسیدکلریدریک ۲ نرمال (Allison and Moodie, 1965)، کاتیون و آنیون‌های محلول بر اساس روش‌های استاندارد (Sparks *et al.*, 1996)، درصد کربن آلی خاک به روش اکسایش تر (Walkey and Black, 1934)، و درصد گچ (Nelson, 1982).



شکل ۱. پراکنش نقاط نمونه‌برداری شده بر حسب مربع لاتین. در شکل مکان هندسی نقاط آموزشی و آزمون مدل مشخص شده است. این داده‌ها به صورت تصادفی انتخاب شدند.

تهیه داده‌های محیطی یا کمکی

نقشه ژئومورفولوژی

بر عکس‌های هوایی ۱/۵۵۰۰۰ منطقه روکش طلق‌مانندی کشیده شد تا خطوط سطوح ژئومورفیک روی آن‌ها ترسیم شود. برای تفکیک واحدهای اراضی از سیستم استریوسکوپ، به منظور

1. PCI Geomatica
2. Ortho-geo-referencing
3. Georeference
4. Slop
5. Elevation
6. Altitude above channel network
7. Modified catchments area
8. Mid-slop position
9. Valley depth
10. Wetness Index
11. Multiresolution index of Valley Bottom Flatness
12. Multiresolutionridgetop flatness index
13. Catchment slope
14. SAGA
15. Normalized difference vegetation index

وضعیت پوشش گیاهی)، شاخص رس^۱، شاخص کربنات^۲، شاخص گچ^۳، شاخص شوری^۴، و شاخص روشنایی^۵ (نشان‌دهنده ترکیبات زمین‌شناسی مواد سطحی) محاسبه شد.

آماده‌سازی داده‌ها برای مدل‌سازی

برای پهنه‌بندی کلاس‌های خاک در سطح گروه بزرگ، داده‌های کلاس خاک کدگذاری و به‌منزله متغیر وابسته در نظر گرفته شد. همچنین همه لایه‌های اطلاعاتی یا کمکی به فرمت رستری با اندازه سلولی ۳۰ متری تهیه شد. این اطلاعات متغیرهای وابسته در مدل اسکوپین شناخته می‌شوند. همه اطلاعات، یعنی داده‌های خاک و متغیرهای محیطی، به فرمت نهایی txt تبدیل و برای انجام‌گرفتن فرایند مدل‌سازی در نرم‌افزارهای مربوطه آماده شدند.

مدل‌سازی خاک

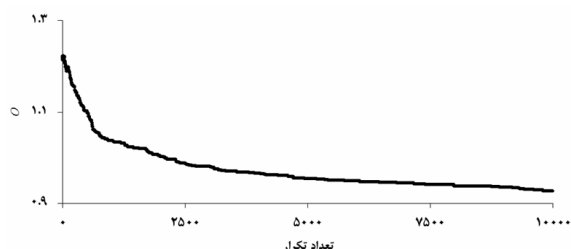
درخت تصمیم

طراحی الگوی نمونه‌برداری در تحقیق حاضر از روش طراحی الگوی نمونه‌برداری مربع لاتین برای انتخاب ۱۸۷ محل نمونه‌برداری از بین ۸۴۲۷۶۶ فضای ممکن استفاده شد. ورودی‌های این مدل‌ها بر اساس تکنیک آنتروپی شانون انتخاب شد (جدول ۱). با توجه به جدول ۱، ده پارامتری که دارای وزن بیشترند- از قبیل نقشه ژئومورفولوژی، ارتفاع، شاخص خیس، شاخص همواری دره با درجه تفکیک بالا، جهت حوضه، تجزیه به مؤلفه‌های اول و دوم تصویر ETM⁺، شیب، موقعیت شیب میانی، و انحنای شیب- وارد مدل مربع لاتین شد.

مدل درختی طبقه‌بندی و رگرسیون یک روش ناپارامتری الگوریتمی است که قادر است متغیرهای کمی یا متغیرهای طبقه‌بندی‌شده را بر اساس مجموعه‌ای از متغیرهای پیش‌بینی‌کننده کمی و کیفی پیش‌بینی کند. در این روش مجموعه‌ای از شرط‌های منطقی^۶ به‌صورت یک الگوریتم با ساختار درختی برای طبقه‌بندی یا پیش‌بینی کمی یک متغیر به‌کار می‌رود. ایجاد درخت تصمیم دو مرحله دارد. مرحله اول مرحله ایجاد و رشد درخت است. این مرحله شامل پیوند^۷ و انشعاب^۸ است. مرحله دوم مرحله توقف و هرس^۹ است. هدف این مرحله به‌حداقل‌رساندن خطای پیش‌بینی است. در تحقیق حاضر از نرم‌افزار سی‌۵^{۱۰} (Quinlan, 2001) برای ساختن درخت تصمیم و در نهایت پیش‌بینی کلاس‌های خاک استفاده شد.

شبکه عصبی مصنوعی

در این تحقیق شبکه عصبی پرسپترون چندلایه با یک لایه مخفی بررسی شد. این شبکه عصبی دارای تابع فعال‌سازی سیگموئید در لایه مخفی و تابع فعال‌سازی خطی در لایه خروجی بود و تعداد نورون‌های آن از ۱ تا ۱۰ نورون متغیر بود. بهترین تعداد نورون به‌صورت سعی و خطا تعیین شد. همچنین الگوریتم آموزشی لونبرگ مارکوارت، به‌علت کارایی و سادگی و سرعت بالا، به‌کار رفت (Amini et al., 2005). در تحقیق حاضر



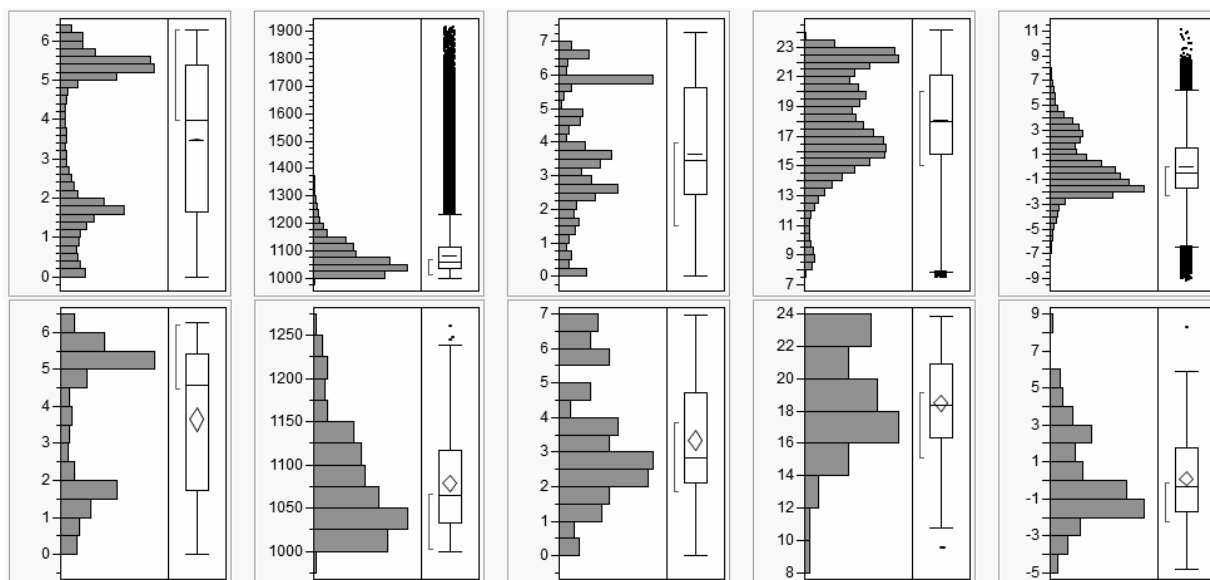
شکل ۲. تکامل معادله هدف با تعداد تکرار ۱۰۰۰۰

1. Clay Index
2. Carbonate Index
3. Gypsum Index
4. Salinity Index
5. Brightness Index
6. if-then
7. Merging
8. Splitting
9. Pruning
10. See5

11. Neurosolutions-5
12. Objective function

جدول ۱. وزن های پارامترهای ورودی انتخاب شده برای مدل های مربع لاتین و فازی کی مینز

پارامتر	ژئومورفولوژی	ارتفاع	شاخص خیسی	ام آربی اف	جهت	مؤلفه اول	مؤلفه دوم	شیب	میداسلوپ	انحنای شیب
وزن	۰٫۲۸	۰٫۱۱	۰٫۱۷	۰٫۲۵	۰٫۰۸	۰٫۰۹	۰٫۰۶	۰٫۰۰۱	۰٫۰۰۱	۰٫۰۰۱



شکل ۳. هیستوگرام روش مربع لاتین از بالا به پایین: داده های اصلی و نمونه برداری مربع لاتین؛ متغیرهای مختلف از چپ به راست: جهت شیب حوزه، ارتفاع، شاخص خیسی، شاخص همواری دره با درجه تفکیک بالا و مؤلفه درجه اول تصویر ماهواره ای

نمونه برداری گزارش کردند. *Thomas et al.* (2012) نیز برای مشخص کردن محل های نمونه برداری در منطقه ای به وسعت ۲۳۰۰۰ هکتار در کشور فیلیپین روش های تصادفی و شبکه ای و مربع لاتین را مقایسه کردند. نتایج برتری مربع لاتین را نشان داد. *Roudier et al.* (2012) نیز به نتایجی مشابه درباره مربع لاتین دست یافتند.

متغیرهای محیطی

با بازدیدهای مکرر صحرایی از منطقه و تفسیر عکس های هوایی و تطابق آن ها با واقعیت و استفاده از تصاویر ماهواره ای به هنگام، ۴ نوع زمین نما و ۲۵ سطح ژئومورفیک در منطقه مورد مطالعه تشخیص داده شد. تفکیک سطوح ژئومورفیک متفاوت در مرحله تفسیر عکس های هوایی بر اساس دانش ژئومورفولوژی و اطلاعات قابل درک از روی عکس های هوایی منطقه بود و طی بازدیدهای صحرایی بر اساس شیب، میزان تجمع مواد دانه ریز و دانه درشت، تراکم آبراهه ها، میزان رطوبت خاک، و سطح آب زیرزمینی صحت تقسیمات مشخص شد.

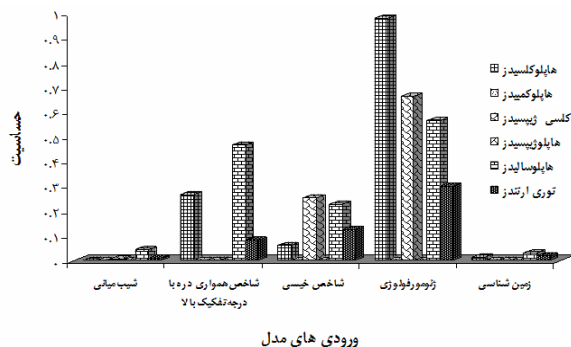
دومین فاکتور محیطی استفاده شده در تحقیق حاضر تصویر ماهواره ای ETM⁺ بود. در این پژوهش از مؤلفه های اصلی به جای باندهای ۱ تا ۶ تصویر استفاده شد. نتایج مؤلفه های

برای بررسی دقت روش نمونه برداری، توزیع هیستوگرامی متغیرهای استفاده شده مدل در محل های منتخب رسم شد. نتایج در شکل ۳ می آید. به نظر می رسد توزیع متغیرهای انتخاب شده با استفاده از مربع لاتین بیشترین تبعیت را از توزیع جامعه دارند. در حقیقت، الگوی نمونه برداری مربع لاتین طوری نقاط نمونه را از جامعه انتخاب کرده که توزیع نمونه گیری در هیستوگرام یکنواخت باشد. *Minasny and McBratney* (2006) نیز در منطقه هانتر ولی^۱، به وسعت ۱۴ کیلومتر مربع در کشور استرالیا، چهار روش نمونه برداری مربع لاتین، میانگین کا، مؤلفه های اصلی، و تصادفی را مقایسه کردند. ایشان لایه رقمی ارتفاع تصویر ماهواره ای ETM⁺ و واحدهای یکنواخت خاک/لنداسکپ را به منزله ورودی های مدل مربع لاتین به کار بردند. نتایج نشان داد روش مربع لاتین هیستوگرام متغیرهای کمی منطقه را به خوبی در برمی گیرد و نمونه برداری میانگین کا فقط توزیع جغرافیایی مناسبی دارد. *Brungard and Boettinger* (2010) نیز شباهت توزیع فراوانی روش مربع لاتین به متغیرهای جامعه را دلیل برتری آن بر سایر روش های

1. Hunter Valley

دارد. لایه ورودی شبکه ۱۵ ورودی و لایه خروجی ۶ نورون دارد. تعداد نورون‌ها و لایه‌های مخفی به پیچیدگی مسئله مورد نظر وابسته است. در این تحقیق یک لایه مخفی برای ایجاد شبکه پرسپترون استفاده شد. تعداد نورون‌های لایه مخفی و تعداد اپوک در فرایند آموزش به صورت سعی و خطا به دست آمد. در این پژوهش، ۲ تا ۱۰ نورون و ۱۰۰ تا ۱۰۰۰ تکرار آزمایش شد. برای تعیین بهترین ترکیب از شاخص صحت کلی استفاده شد. نتایج نشان داد تعداد نورون ۹ و تعداد تکرار ۲۰۰ دارای بهترین ترکیب برای پیش‌بینی گروه بزرگ خاک است. در نتیجه، ساختار ۱۵-۹-۶ برای پیش‌بینی کلاس گروه خاک پیشنهاد می‌شود. همچنین، نتایج ماتریس خطا نشان داد کلاس هاپلوسالید (۹۶/۱۵ درصد) بالاترین صحت طبقه‌بندی را دارد؛ درحالی‌که مدل در مرحله آموزش نتوانست هیچ‌یک از هاپلوکمییداها (صفر درصد) را به درستی طبقه‌بندی کند. دقت کلی این شبکه در مرحله آموزش ۷۰ درصد است.

بعد از به دست آمدن ساختار بهینه در مرحله آموزش شبکه عصبی مصنوعی، که توضیح داده شد، شبکه آموزش دیده بر داده‌های آزمون آزمایش کردند تا دقت این مدل برای پیش‌بینی کلاس خاک تا سطح گروه بزرگ مشخص شود. نتایج ماتریس خطا برای مرحله آزمون با ۳۷ داده نشان داد بالاترین دقت متعلق به گروه هاپلوسالیدز (۹۴/۴ درصد) و بدترین طبقه‌بندی مربوط به چهار کلاس خاک است که دقت پیش‌بینی صفر درصد دارد. دقت کلی مدل ۴۸ درصد برای پیش‌بینی کلاس خاک تا سطح گروه بزرگ است. همچنین، نتایج آنالیز حساسیت شبکه عصبی مصنوعی نشان داد پارامتر ژئومورفولوژی برای پیش‌بینی کلاس‌های خاک در سطح گروه بزرگ بالاترین اثر را دارد. بعد از سطوح ژئومورفیک پارامترهای سرزمین، از جمله شاخص خیزی و همواری دره با درجه تفکیک بالا، حائز اهمیت است (شکل ۵).

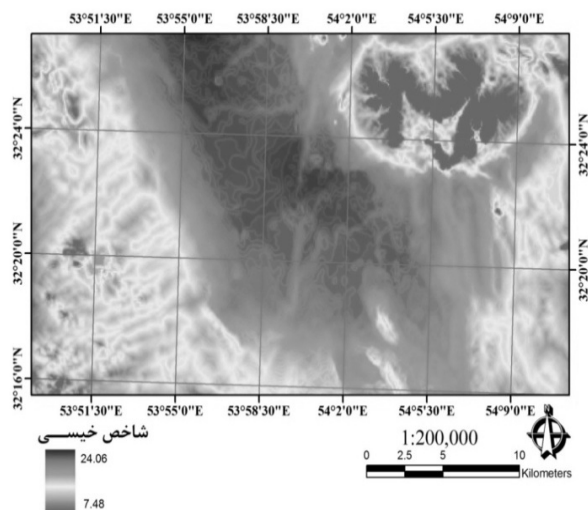


شکل ۵. آنالیز حساسیت ورودی‌های مدل با استفاده از شبکه عصبی

برخی محققان از شبکه‌های عصبی مصنوعی برای پیش‌بینی کلاس خاک استفاده کردند؛ از جمله Lehmann *et al* (1999), Zhu (2000)، و Behrens *et al* (2005). آن‌ها

اصلی نشان داد سه مؤلفه اول ۹۹ درصد تغییرات را شامل می‌شود. بنابراین، سه مؤلفه اول به منزله ورودی مدل انتخاب شدند. اولین مؤلفه ۸۹/۸۱ درصد تغییرات باندهای تصویر را در برمی‌گیرد. بالاترین مقدار ویژه متعلق به باند ۲ و کمترین آن مربوط به باند ۴ است.

سومین فاکتور محیطی استفاده شده در این پژوهش از آنالیز مدل رقومی ارتفاع^۱ به دست آمد؛ مثلاً شکل ۴ تغییرات مکانی شاخص خیزی را نشان می‌دهد. مقادیر کم شاخص خیزی (>۷) عموماً در مناطق کوهستانی واقع شده‌اند. مقادیر حد واسط، که در حدود ۱۴ است، قسمت‌هایی از لندفرم باهادا را پوشش می‌دهد و بالاترین مقادیر این شاخص با لندفرم پلایا هم‌خوانی دارد. در حقیقت، این ناحیه پتانسیل بالای تجمع مواد شور را نشان می‌دهد. پلایا به خصوص واحد P113 در انتهای حوزه دشت یزد اردکان قرار گرفته است و در نتیجه مواد با قابلیت حلالیت بالا را دریافت می‌کند. این پدیده منجر به شوری خاک‌های این واحد می‌شود. مطابق یافته‌های بسیاری از محققان، مانند Jafari *et al* (2012) و Moore *et al* (1991)، بین شاخص خیزی و شوری خاک ارتباط زیادی وجود دارد.



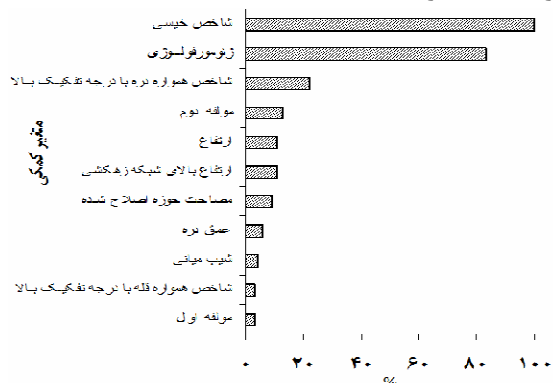
شکل ۴. نقشه شاخص خیزی به دست آمده از مدل رقومی ارتفاع

مدل سازی مکانی

شبکه عصبی مصنوعی: توپولوژی شبکه عاملی مهم در طراحی شبکه عصبی مصنوعی است؛ زیرا این ساختار بر سرعت یادگیری و صحت طبقه‌بندی نهایی اثر می‌گذارد. تعداد لایه‌های مخفی و تعداد نورون‌ها اجزای اصلی شبکه پرسپترون‌اند. تعداد اپوک نیز در مرحله یادگیری بسیار مهم است. تعداد نورون‌ها در لایه‌های ورودی و خروجی ثابت است و به تعداد ورودی (متغیرهای کمکی) و خروجی (کلاس‌های خاک) مدل استفاده شده بستگی

1. Digital Elevation Model

بلندی از عوامل بسیار مهم خاک سازی است. اجزای سرزمین در توزیع مکانی خاک و خصوصیاتش بسیار مؤثر است؛ زیرا توزیع مکانی تابش خورشیدی، درجه حرارت، رطوبت، و جریان مواد را تحت تأثیر قرار می دهد. یکی از اجزای سرزمین شاخص خیزی یا رطوبت است که برای تعیین توزیع مکانی نفوذ و جریان آب سطحی به کار می رود. در واقع تأثیر توپوگرافی بر مکان و میزان تجمع رطوبت در خاک یا آب در سطح زمین نما را نشان می دهد. این پارامتر با متغیرهای وابسته ای که با رطوبت ارتباط و تحت تأثیر آن قرار دارند همبستگی بیشتری نشان می دهد؛ به طوری که در مطالعه حاضر شاخص خیزی به منزله یک متغیر پیش بینی کننده معنادار در مدل های افق مشخصه سالیک وارد شد. شاخص خیزی و همواری دره با درجه تفکیک بالا مناطق مستعد برای دریافت رسوبات را (نواحی میانی) نشان می دهند و این قسمت از مناطق بیشتر از خاک های شور تشکیل می شوند. در حقیقت این شاخص ها به شناسایی گروه بزرگ هاپلوسالیدز کمک می کنند. Moore et al. (1991) نیز ارتباط معنادار بین خاک های شور و شاخص خیزی را تأیید کردند. Jafari et al. (2012) نیز در منطقه خشک اقدام به پیش بینی کلاس های خاک کردند و شاخص خیزی و همواری دره با درجه تفکیک بالا را پارامتری مهم برای مدل خود معرفی کردند. سطوح ژئومورفولوژی در رده دوم قرار دارد و ۸۴ درصد در مدل استفاده شد. سطح ژئومورفولوژی نیز پارامتری بسیار مهم است؛ شاید به این دلیل که سطوح ژئومورفولوژی اخیراً تشکیل شده است و بنابراین با فرایند خاک سازی در مناطق خشک ارتباط مناسبی را نشان می دهد. این نتایج با نتایج محققان دیگر هم خوانی دارد. Scull et al. (2005) و Jafari et al. (2012) تأکید کردند سطوح ژئومورفولوژی مهم ترین متغیر کمکی برای شناسایی خاک ها در مناطق بیابانی اند. نتایج همچنین نشان داد سایر متغیرهای کمکی تأثیر چندانی بر مدل برای پیش بینی گروه های بزرگ خاک ندارند.



شکل ۶. نتایج اهمیت متغیرهای محیطی در پیش بینی گروه های بزرگ خاک بر حسب درصد

شبکه های عصبی مصنوعی را نیز برای پیش بینی مکانی واحدهای خاک به کار بردند. این محققان اجزای سرزمین- نظیر شیب، انحنای شیب، ارتفاع، جهت شیب، تابش خورشیدی- را فاکتورهای کمکی مؤثر معرفی کردند. در مطالعه Behrens et al. (2005) نقشه رومی ژئومورفولوژی و زمین شناسی هم به منزله متغیرهای کمکی در پیش بینی واحدهای خاک استفاده شدند. Zhao et al. (2009) در کشور چین نیز کارایی بالای شبکه عصبی مصنوعی را در پیش بینی کلاس بافت تأیید کردند.

در روش درخت تصمیم، مجموعه ای از شرط های منطقی به صورت الگوریتمی با ساختار درختی برای طبقه بندی کلاس های خاک به کار می رود. نتایج پیش بینی در مرحله آموزش نشان داد دقت کلی مدل درخت تصمیم ۸۸ درصد است. نتایج ماتریس خطا نشان داد بهترین طبقه بندی در این روش متعلق به گروه بزرگ کلسی ژپسیدز (۹۲ درصد) و بدترین طبقه بندی مربوط به کلاس هاپلوکمبیدز (۶۶ درصد) است. بعد از به دست آمدن قوانین مرتبط، که توضیح داده شد، این مدل درختی بر داده های آزمون (۲۰ درصد از داده ها) آزمایش شد تا دقت این مدل به منظور پیش بینی کلاس خاک برای سطح گروه بزرگ مشخص شود. نتایج ماتریس خطا برای مرحله آزمون با ۳۷ داده در جدول ۲ می آید. در این گروه داده نیز بالاترین دقت متعلق به گروه هاپلوسالیدز (۸۸ درصد) و بدترین طبقه بندی مربوط به کلاس هاپلوکمبیدز خاک است؛ که دقت پیش بینی صفر درصد دارد. دقت کلی مدل ۵۲ درصد به منظور پیش بینی کلاس خاک برای سطح گروه بزرگ است. برخی محققان کارایی درختان تصمیم را در پیش بینی کلاس های خاک تأیید کرده اند (Bui et al., 1999; Moran and Bui, 2002; Bui and Moran, 2003; Henderson et al., 2005; Minasny and McBratney, 2006; Mendoca-Santos et al., 2007).

آنالیز درخت تصمیم نشان داد بعضی متغیرهای کمکی- مانند شاخص خیزی، سطوح ژئومورفولوژی، شاخص همواری دره با درجه تفکیک بالا، مؤلفه درجه دوم، ارتفاع، و ارتفاع بالای شبکه زهکشی- بالاترین تأثیر را بر پیش بینی گروه های بزرگ خاک دارند (شکل ۶). شکل ۶ تأثیر هر متغیر کمکی را به صورت درصد بر مدل درختی بیشتر نشان می دهد. با توجه به شکل ۶، چنین استنباط می شود که شاخص خیزی مؤثرترین متغیر کمکی در مدل است. این شاخص ۱۰۰ درصد در مدل استفاده شد. همچنین، متغیر کمکی مهم دیگر شاخص همواری دره با درجه تفکیک بالاست که ۲۷ درصد در مدل وارد شد. این موضوع می تواند مؤید این مطلب باشد که در این منطقه پستی و

جدول ۲. نتایج ماتریس خطا مربوط به درخت تصمیم در مرحلهٔ آزمون

دقت کاربر	توری ارتندز	هاپلوسالیدز	هاپلوژیپسیدز	کلسی ژیپسیدز	هاپلوکمیدز	هاپلوکلسیدز	دقت تولیدکننده
۰ درصد	۰	۲	۱	۰	۱	۰	هاپلوکلسیدز
۰ درصد	۰	۰	۲	۱	۰	۰	هاپلوکمیدز
۰ درصد	۱	۰	۱	۰	۰	۲	کلسی ژیپسیدز
۰ درصد	۰	۰	۰	۳	۰	۱	هاپلوژیپسیدز
۸۸ درصد	۰	۱۶	۰	۱	۱	۰	هاپلوسالیدز
۷۵ درصد	۳	۰	۰	۰	۰	۱	توری ارتندز
	۷۵ درصد	۸۸ درصد	۰ درصد	۰ درصد	۰ درصد	۰ درصد	

نتیجه گیری

نتایج مقایسهٔ مدل‌های شبکهٔ عصبی مصنوعی و درخت تصمیم برای پیش‌بینی کلاس خاک نشان داد که مدل درخت تصمیم در جداسازی خاک‌ها کارایی بالاتری دارد. بنابراین، از این مدل برای طبقه‌بندی استفاده شد. نتایج همچنین نشان داد که شاخص خیسی و سطوح ژئومورفولوژی مهم‌ترین متغیرهای کمکی‌اند و صحت طبقه‌بندی به تعداد دادهٔ مشاهده‌شده وابسته است. ارتباط قوی بین دادهٔ خاک و پارامترهای محیطی نیز یکی دیگر از عوامل تأثیرگذار بر دقت است. به‌طور کلی نتایج نشان داد تکنیک‌های خاک‌آماری، به‌خصوص درخت تصمیم، می‌توانند فرایند نقشه‌برداری خاک‌ها را در گستره‌ای وسیع و متشکل از هر گونه عوارض طبیعی به‌یک‌باره انجام دهند، روش سنتی را ارتقا بخشند، سرعت عمل و کارآمدی نقشه‌ها را در انتقال داده‌ها و اطلاعات افزایش دهند، و قابلیت استفاده از آن‌ها را برای قشر وسیعی از شاخه‌های علمی امکان‌پذیر کنند. پیشنهاد می‌شود برای تهیهٔ نقشهٔ رقومی خاک از مدل‌های درختی در مطالعات آینده استفاده شود.

نتایج مقایسهٔ مدل‌ها برای پیش‌بینی گروه بزرگ خاک در جدول ۳ می‌آید. با دقت در جدول ۳ روشن می‌شود مدل درخت تصمیم از لحاظ هر سه معیار ارزیابی برای پیش‌بینی گروه بزرگ خاک بهترین عملکرد را دارد. بعد از درخت تصمیم، شبکهٔ عصبی، با دقت کلی ۴۸ درصد، در مرتبهٔ دوم قرار می‌گیرد. Luoto and Hjort (2005) نیز کارایی بالاتر درخت تصمیم را در مقایسه با شبکهٔ عصبی نشان دادند. همچنین Kheir *et al* (2010) بدین نتیجه رسیدند که درخت تصمیم مناطق با مادهٔ آلی بالا را بهتر طبقه‌بندی می‌کند. با وجود این Moonjun *et al* (2010) تکنیک شبکهٔ عصبی مصنوعی و درخت تصمیم را برای پیش‌بینی واحدهای خاک مقایسه کردند و نشان دادند که برای پیش‌بینی تفاوتی معنادار بین دو تکنیک وجود ندارد.

جدول ۳. نتایج مقایسهٔ انواع مدل‌ها برای پیش‌بینی گروه بزرگ خاک بر

اساس داده‌های آزمون		
مدل	دقت کلی	ضریب کاپا
شبکهٔ عصبی مصنوعی	۴۸ درصد	۰٫۱۳
درخت تصمیم	۵۲ درصد	۰٫۲۵
		ضریب تبیین
		۰٫۳۴
		۰٫۴۶

REFERENCES

- Amini, M., Abbaspour, K. C., Khademi, H., Fathianpour, N., Afyuni, M., and Schulin, R. (2005). Neural network models to predict cation exchange capacity in arid regions of Iran, *European Journal of Soil Science*. 53, 748–757.
- Bagheri-Bodaghabadi, M., Salehi, M. H., Martínez-Casasnovas, J. A. Mohammadi, J., Toomanian, N., and Esfandiarpour-Borujeni, I. (2011). Using Canonical Correspondence Analysis (CCA) to identify the most important DEM attributes for digital soil mapping applications, *Catena*. 86, 66–74.
- Behrens, T., Foster, H., Scholten, T., Steinrucken, U., Spies, E., and Golshmitt, M. (2005). Digital soil mapping using artificial neural networks, *Journal of Plant Nutrition Soil Science*, 169, 434–443.
- Brungard, C. W. and Boettinger, J. L. (2010). Conditioned Latin Hypercube Sampling: Optimal Sample Size for Digital Soil Mapping of Arid Rangelands in Utah, USA.
- Bui, E. N. and Moran, C. J. (2003). A strategy to fill gaps in soil survey over large spatial extents: an example from the Murray-Darling basin of Australia, *Geoderma*, 111, 21–44.
- Bui, E. N., Loughhead, A., and Corner, R. (1999). Extracting soil-landscape rules from previous soil surveys, *Australian Journal of Soil Research*, 37, 495–508.
- Gee, G. W. and Bauder, J. W. (1986). Particle size analysis, 383–411, In: A. Klute. (ed), *Methods of Soil Analysis, Part 1, American Society of Agronomy* Madison, WI.
- Hamzhepour, N., Eghbal, M. K., Bogaert, P., Toomanian, N., and Sokouti, R. S. (2013). Spatial Prediction of Soil Salinity Using Kriging with Measurement Errors and Probabilistic Soft Data,

- Arid Land Research and Management*, 27, 128-139.
- Henderson, B. L., Bui, E. N., Moran, C. J., and Simon, D. A. P. (2005). Australia-wide predictions of soil properties using decision trees, *Geoderma*, 124, 383-398.
- Hengl, T., Huvelink, G. B. M., and Stein, A. (2004). A generic framework for spatial prediction of soil variables based on regression-kriging, *Geoderma*, 120, 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.
- Jafari, A., Finke, P. A., de Wauw, J. V., 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, 284-298.
- Kheir, B., Greve, M. H., Bøcher, P. K., Greve, M. B., Larsen, R., and McCloy, K. (2010). Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: The case study of Denmark, *Journal of Environmental Management*, 91, 1150-1160.
- Lehmann D., Billen N., and Lenz, R. (1999). Anwendung von Neuronalen Netzen in der Landschaftsökologie-Synthetische Bodenkartierung im GIS, In: Strobl, J. and T. Blaschke (Eds), *Anwendung von Bodenkartierung im GIS*, Wichmann, Heidelberg, 330-336.
- Luoto, M. and Hjort, J. (2005). Evaluation of current statistical approaches for predictive geomorphological mapping, *Geomorphology*, 67, 299-315.
- McBratney, A. B., Mendonça-Santos, M. L., and Minasny, B. (2003). On digital soil mapping, *Geoderma*, 117, 3-52.
- McBratney, A. B., Odeh, I. O. A., Bishop, T. F. A., Dunbar, M. S., and Shatar, T. M. (2000). An overview of pedometric techniques for use in soil survey, *Geoderma*, 97, 293-327.
- McKenzie, N., Grundy, M., Webster, R., and Ringrose-Voase, A. (2008). *Guideline for surveying soil and land resources*, CSIRO Publishing, 557p.
- Mendonça-Santos, M. L., McBratney, A. B., and Minasny, B. (2006). Soil prediction with spatially decomposed environmental factors, *Digital Soil Mapping — An Introductory Perspective*, 31: 269-278.
- Metternicht, G. and Zinck, J. A. (2003). Remote Sensing of Soil Salinity: Potentials and Constraints, *Remote Sensing of Environment*, 85, 1-20.
- Minasny, B. and McBratney, A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information, *Computer & Geosciences*, 32, 1378-1388.
- Minasny, B. and McBratney, A. B. (2007). Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes, *Geoderma*, 142, 285-293.
- Moonjun, R., Farshad, A., Shrestha, D. P., and Vaiphasa C. (2010). Artificial Neural Network and Decision Tree in Predictive Soil Mapping of Hoi NumRin Sub-Watershed, Thailand, *Digital Soil Mapping*, Progress in Soil Science 2, 151-164.
- Moore, I. D., Grayson, R. B., and Ladson, A. R. (1991). Digital terrain modeling: review of hydrological, geomorphological and biological applications, *Hydrology Processing*, 5, 3-30.
- Moran, C. J. and Bui, E. N. (2002). Spatial data mining for enhanced soil map modeling, *International Journal of Geographical Information Science*, 16, 533-549.
- Nelson, R. E. (1982). Carbonate and gypsum, In: A. L. Page, R. H. Miller, and R. Keeny, *Methods of soil analysis*, Part II-chemical and microbiological properties, Madison, WI, 181-196.
- Quinlan, J. R. (2001). *Cubist: An Informal Tutorial*, <http://www.rulequest.com>.
- Rossiter, D. G. and Hengl, T. (2001). Technical note: Creating geometrically-correct photo-interpretation, photo-mosaics, and base maps for a projects GIS, <http://www.itc.nl/rossiter>.
- Rossiter, D. G. and Loza, A. V. (2010). Technical note: Analyzing land cover change with logistic regression in R (Version 2.2, First version April 2004), ITC, Enschede, The Netherland.
- Rossiter, D. G., de la Geomatica, S., and Fe de Bogota, S. (2005). Digital soil mapping: towards a multiple-use soil information system, *International Journal of Geo-information, Science*, Earth Obser.
- Roudier, P., Hewitt, A. E., and Beaudette, D. E. (2012). A conditioned Latin hypercube sampling algorithm incorporating operational constraints, *Digital Soil Assessments and Beyond*, Taylor and Francis Group, 227-231
- Scull P., Franklin, J., and Chadwick, O. A. (2005). The application of classification of tree analysis to soil type prediction in a desert landscape, *Ecological Modelling*, 181: 1-15.
- Sparks, D. L., Page, A. L., Helmke, P. A., Leoppert, R. H., Soltanpour, P. N., Tabatabai, M. A., Johnston, G. T., and Summer, M. E. (1996). *Methods of Soil Analysis*, *Soil Science Society of American Journal*, Madison, Wisconsin.
- Thomas, M., Odgers, N. P., Ringrose-Voase, A., Grealish, G., Glover M., and Dowling T. (2012). Soil survey design for management-scale digital soil mapping in a mountainous southern Philippine catchment, *Digital Soil Assessments and Beyond*, Taylor and Francis Group, 233-238.
- Walkey, A. and Black, I. A. (1934). An examination of Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid in soil analysis, 1. Experimental, *Soil Science Society of American Journal*, 79, 459-465.

Webster, R., Oliver, M. (2007). Geostatistics for Environmental Scientists, 2nd Edition, John Wiley & Sons Ltd, Chichester UK.

Zhao, Z., Chow, T. L., Rees, H. W., Yang, Q., Xing Z., and Meng, F. (2009). Predict soil texture distributions using an artificial neural network

model, *Computers and electronics in agriculture*, 65, 36–48.

Zhu A. X. (2000). Mapping soil landscape as spatial continua: the neural network approach, *Water Resources Research*, 36, 663-677.