



محور مقاله: پدومتری و ارزیابی خاکها

تأثیر تبدیل داده‌های بافت و مدلسازی رقومی اجزاء بافت بر آب قابل استفاده خاک

علیرضا امیریان چکان^{۱*}، زهرا درویش‌پسند^۱، رخسار اکبری فضلی^۲، روح‌الله تقی‌زاده مهرجردی^۴ صاحب‌خوردبین^۲^۱ استادیار گروه علوم خاک دانشکده کشاورزی، دانشگاه لرستان^۲ کارشناس ارشد، دانشگاه صنعتی خاتم‌الانبیاء بهبهان^۳ دانشجوی دکترا گروه علوم خاک، دانشگاه آزاد اسلامی واحد علوم و تحقیقات خوزستان^۴ استادیار گروه مرتع و آبخیزداری دانشکده کشاورزی و منابع طبیعی، دانشگاه اردکان

چکیده

وقتی اجزاء مختلف بافت خاک (PSFs) با استفاده از نقشه‌برداری رقومی خاک (DSM) به صورت مستقل از هم تخمین زده می‌شوند تضمینی وجود ندارد که مجموع تخمین‌ها برابر با ۱۰۰ گردد. انجام تبدیل‌های log-ratio قبل از مدل‌سازی تضمین‌کننده برابر با واحد شدن مجموع تخمین‌هاست. اطلاعات کمی درباره اثر تبدیل داده‌ها وقتی از خروجی DSM برای تخمین خواص خاک با استفاده از توابع انتقالی خاک استفاده می‌شود در دسترس است. این مطالعه در استان خوزستان برای بررسی کارایی مدل جنگل تصادفی (RF) در تخمین داده‌های PSFs تبدیل شده (TD) و تبدیل نشده (UT) انجام گرفت. همچنین، اثر تبدیل داده‌ها و نوع تبدیل روی مقدار کل آب آبیاری مورد نیاز برای گندم (TIW) بررسی شد. نتایج نشان داد هر چند که تبدیل داده‌ها منجر به تخمین‌های اریب PSFs گردید ولی مدل RF بر اساس RMSE کارایی مشابهی در تخمین داده‌های UT و TD داشت. درصد کلاس‌های بافتی درست طبقه‌بندی شده برای UT تخمینی بیشتر از TD بود. توزیع مکانی مجموع اجزاء بافت خاک نشان داد فقط در بخش‌های کوچکی از منطقه مجموع مقادیر برابر با ۱۰۰ شده بود. همچنین هر چند که تبدیل داده‌ها و نوع تبدیل تأثیر معنی‌داری روی تخمین‌های TIW نداشت ولی تخمین‌های TIW با استفاده از TD اریب بود.

کلمات کلیدی: آب قابل استفاده، داده مرکب، نقشه‌برداری رقومی، توابع انتقالی خاک

مقدمه

بافت خاک یکی از ویژگی‌های کلیدی است که خواص و فرایندهای فیزیکی و شیمیایی زیادی را تحت تأثیر قرار می‌دهد. اجزاء اندازه ذرات (PSFs) معمولاً به عنوان ورودی توابع انتقالی خاک (PTFs) برای مدل‌سازی فرایندهای هیدرولوژیکی، اکولوژیکی و زیست محیطی مورد استفاده قرار می‌گیرند (Van Looy et al., 2017; Minasny and McBratney, 2018). علیرغم اهمیت زیادی که بافت خاک دارد، داده‌های آن معمولاً در مقیاس مورد نیاز برای مدیریت کشاورزی و مدل‌سازی‌های زیست محیطی در دسترس نیست (Roman Dobarco et al., 2017). نقشه‌برداری رقومی خاک (DSM) رویکردی است که برای مدل‌سازی مکانی PSFs در مقیاس‌های مختلف به کار رفته است (Adhikari et al., 2013; Hengl et al., 2017).

بافت خاک رایج‌ترین داده مرکب در علوم خاک است. داده مرکب داده‌ای است که دارای چند جزء غیر منفی است که مجموع آنها برابر با واحد است (Aitchison, 1986). وقتی عناصر یک داده مرکب به صورت جداگانه تخمین زده می‌شوند تضمینی برای اینکه مجموع اجزاء تخمینی برابر با واحد شود وجود ندارد (Lark and Bishop, 2007). به عنوان مثال اگر درصد‌های رس، سیلت و شن در یک موقعیت بدون داده تخمین زده شوند، مجموع مقادیر تخمینی الزاماً برابر با ۱۰۰ نخواهد بود. یک روش رایج برای تضمین برابر با ۱۰۰ شدن تخمین‌ها این است که دو جزء تخمین زده شوند و جزء سوم از اختلاف دو جزء تخمینی از ۱۰۰ به دست آید. در این صورت بسته به اینکه کدام دو جزء ابتدا تخمین زده شوند ممکن است نتایج متفاوتی به دست آید و در برخی مطالعات که نیاز است این اجزا قبل از ورود به PTFs تخمین زده شوند، ممکن است خروجی‌های متفاوتی به دست آید. راهکار دیگر برای تضمین برابر با ۱۰۰ شدن تخمین‌ها استفاده از تبدیل‌های log-ratio معرفی شده توسط Aitchison (۱۹۸۶) است. هر چند که در مطالعات متعددی از این تبدیل‌ها برای مدل‌سازی PSFs استفاده شده است (Poggio and Gimona, 2017; Wang and Shi, 2017) ولی تأثیر این تبدیل‌ها و برابر با ۱۰۰ نشدن تخمین‌ها بر خروجی PTFs تا کنون مورد بررسی قرار نگرفته است. این تحقیق برای بررسی کارایی تکنیک‌های DSM در تخمین PSFs

* ایمیل نویسنده مسئول: amirian.ar@lu.ac.ir

با استفاده از داده‌های تبدیل شده و داده‌های تبدیل نشده در استان خوزستان انجام گرفت. همچنین کارآیی سه نوع از تبدیل‌های رایج در تخمین PSFs مورد بررسی قرار گرفت. بعلاوه تاثیر تبدیل داده‌ها روی تخمین ظرفیت آب قابل استفاده (AWC) و مقدار کل آب آبیاری مورد نیاز (TIW) برای گندم در طول یک فصل رشد مورد بررسی قرار گرفت.

مواد و روش‌ها

منطقه مورد مطالعه به وسعت حدود ۴۶۰۰ هکتار در شمال شرق شهر بهبهان در استان خوزستان واقع شده است. تعداد ۱۵۰ نمونه خاک از لایه سطحی (۳۰-۰ سانتی‌متر) برداشته شد و بافت خاک به روش هیرومتر اندازه‌گیری شد. چون بافت خاک یک داده مرکب است (composition) برای تضمین برابر ۱۰۰ بودن تخمین‌های رس، سیلت و شن در یک نقطه، از تبدیل‌های خاصی استفاده می‌شود. بردار $x = [x_1, x_2, \dots, x_D]$ وقتی یک داده ترکیبی D قسمتی در نظر گرفته می‌شود که همه اجزاء آن اعداد حقیقی مثبت و در برگزیده اطلاعات نسبی باشند. فضای نمونه مناسب برای داده‌های ترکیبی یک simplex (داده خام) به صورت زیر است:

$$S^D = \{x = [x_1, x_2, \dots, x_D] \mid x_i > 0, i = 1, 2, \dots, D; \sum_{i=1}^D x_i = \kappa\} \quad (1)$$

در این رابطه S^D بیانگر بردارهای یک ترکیب D قسمتی و κ مقدار ثابتی است که برابر با ۱۰۰ یا یک است. داده‌های بافت با استفاده از سه نوع تبدیل رایج شامل additive log-ratio (alr), centroid log-ratio (clr) و isometric log-ratio (ilr) تبدیل شدند. به عنوان نمونه فقط روابط مربوط به تبدیل alr ارائه می‌شود. برای تبدیل داده‌های خام (Simplex) به فضای حقیقی (مختصات alr) از رابطه زیر استفاده می‌شود:

$$y = alr(x) = \left[\ln \frac{x_1}{x_D}, \ln \frac{x_2}{x_D}, \dots, \ln \frac{x_{D-1}}{x_D} \right] \quad (2)$$

در این رابطه $y \in \mathbb{R}^{D-1}$ یک فضای حقیقی با D-1 بعد است. برای اینکه داده‌های تبدیل شده PSFs به عنوان ورودی PTFs مورد استفاده قرار گیرند باید این داده‌ها مجدداً تبدیل شوند (Back-transformed) تا داده‌های اولیه (درصدهای رس، سیلت و شن) به دست آیند. برای تبدیل داده‌ها از مختصات alr به داده‌های خاک از رابطه زیر استفاده می‌شود:

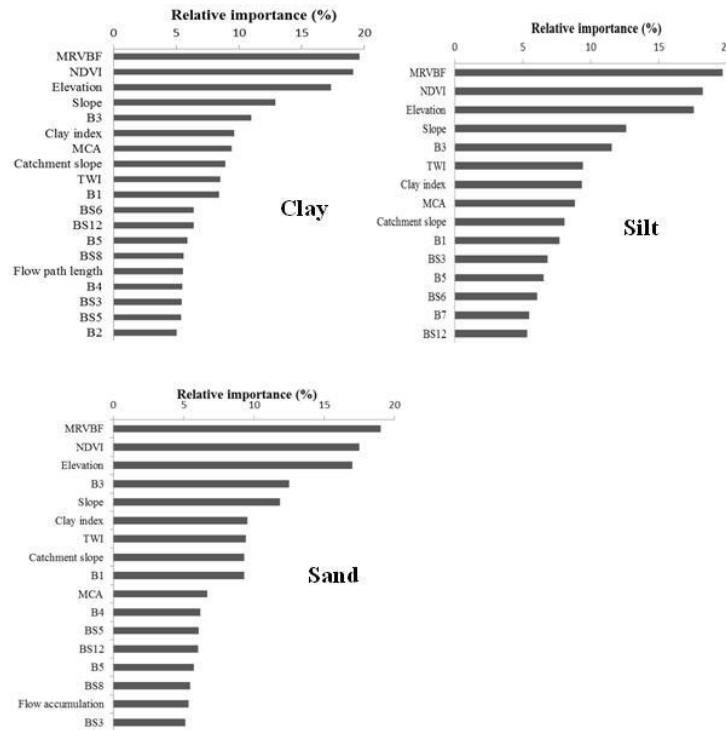
$$x = agl(y) = \left[\frac{\exp(y_1)}{1 + \sum_{i=1}^{D-1} \exp(y_i)}, \dots, \frac{\exp(y_{D-1})}{1 + \sum_{i=1}^{D-1} \exp(y_i)}, \frac{1}{1 + \sum_{i=1}^{D-1} \exp(y_i)} \right] \quad (3)$$

برای مدل‌سازی داده‌های PSFs تبدیل شده و تبدیل نشده، از متغیرهای کمکی استخراج شده از مدل رقومی ارتفاع، تصویر لندست ۸ و تصویر Sentinel-2 به عنوان ورودی و از جنگل تصادفی (RF) به عنوان مدل تخمین‌گر استفاده گردید.

برای بررسی تاثیر PSFs تخمینی تبدیل شده و تبدیل نشده روی خروجی PTFs، از توابع معرفی شده توسط خداوردیلو و همکاران (۲۰۱۱) برای تخمین AWC استفاده گردید. همچنین مقدار TIW برای گندم بر اساس مقدار آب سهل‌الوصول و تعداد دفعات آبیاری محاسبه گردید. مقادیر AWC با استفاده از داده‌های PSFs تبدیل نشده و داده‌های Back-transformed محاسبه گردید.

نتایج و بحث

اهمیت نسبی متغیرهای کمکی در تخمین PSFs در شکل (۱) نشان داده شده است. NDVI، MRVBF، ارتفاع و باند ۳ لندست ۸ مهمترین متغیرها در تخمین رس، سیلت و شن هستند. نتایج نشان می‌دهد باندهای طیفی لندست ۸ (بجز باند ۳) و Sentinel-2 اهمیت کمتری در تخمین PSFs دارند. پارامترهای اراضی استخراج شده از DEM با اثر مستقیم و غیر مستقیم روی وضعیت آب خاک و فرایندهای فرسایش و رسوب‌گذاری روی توزیع PSFs تاثیر دارند. به عنوان مثال مقادیر بالای MRVBF بیانگر مکان‌های رسوب‌گذاری است یعنی جایی که ذرات ریزتر بیشتر تجمع می‌یابند. شیب منطقه مقدار رواناب سطحی و میزان انتقال ذرات را تحت تاثیر قرار می‌دهد. معمولاً رابطه بین شیب با ذرات ریز منفی و با ذرات درشت مثبت است (Umali et al., 2012). همبستگی بین شاخص خیسی (TWI) و PSFs برعکس شیب است. باندهای طیفی که بیانگر عوارض و پوشش سطح زمین هستند می‌توانند به عنوان نماینده‌ای از عامل خاک‌سازی موجودات زنده (Malone et al., 2009) روی توزیع و تغییرات بافت خاک تاثیرگذار باشند.



شکل ۱- اهمیت نسبی متغیرهای محیطی استخراج شده از مدل رقمی ارتفاع و تصاویر ماهواره‌ای در تخمین اجزاء بافت خاک

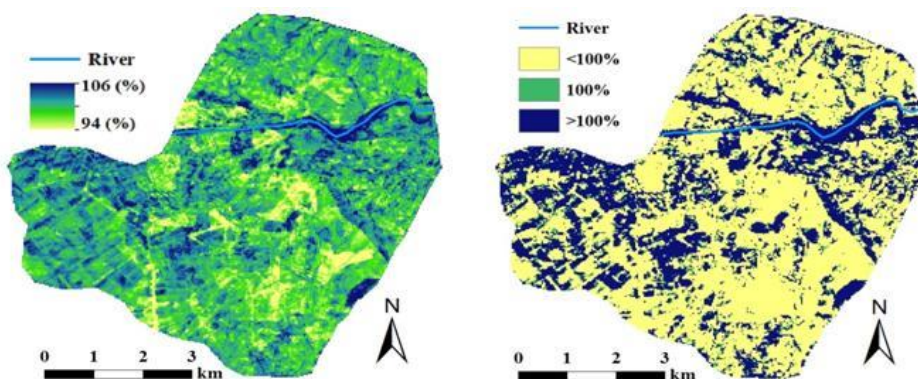
کارایی مدل جنگل تصادفی در تخمین داده‌های تبدیل شده و تبدیل نشده در جدول (۱) ارائه شده است. برای داده‌های تبدیل نشده دو رویکرد در نظر گرفته شد. در رویکرد ۱ هر سه جزء بافت خاک به طور جداگانه و با استفاده از متغیرهای کمکی توسط مدل RF تخمین زده شدند. در رویکرد ۲ هر بار دو جزء با استفاده از متغیرهای کمکی توسط مدل RF تخمین زده شد و جزء سوم از تفاوت دو جزء تخمینی از ۱۰۰ به دست آمد. نتایج نشان داد رویکرد ۲ بخصوص برای رس منجر به تخمین‌هایی با دقت کمتری شده است. بر اساس مقادیر RMSE، کارایی مدل RF در تخمین داده‌های تبدیل شده و تبدیل نشده تقریباً شبیه است ولی استفاده از داده‌های تبدیل شده منجر به تخمین‌های اریب برای هر سه جزء گردیده است. فاصله Aitchison در حالتی که داده‌ها تبدیل نشده‌اند بیشتر است. نتایج همچنین نشان می‌دهد در حالتی که از داده‌های تبدیل نشده استفاده شده است درصد بیشتری کلاس بافت خاک درست طبقه‌بندی شده است (۷۳/۶٪).

توزیع مکانی مجموع رس، سیلت و شن تبدیل نشده تخمینی با استفاده از مدل RF و داده‌های کمکی در شکل (۲) نشان داده شده است. این شکل نشان می‌دهد در اکثر قسمت‌های منطقه شرط اصلی یک داده ترکیبی برقرار نیست. یعنی مجموع درصدهای این سه جزء برابر با ۱۰۰ نشده است و از ۹۴ درصد تا ۱۰۶ درصد متغیر است. این وضعیت نیز توسط Odeh و همکاران (۲۰۰۳) برای ۶۵ درصد از منطقه مورد بررسی گزارش شده است. این نتیجه به این دلیل است که حداقل یک همبستگی منفی بین دو جزء از اجزاء یک داده ترکیبی وجود دارد و اجزاء نمی‌توانند به صورت مستقل از هم تغییر کنند (Pawlowsky-Glahn and Egozcue, 2006). برای اجزاء بافت خاک چون مجموع درصدها برابر با ۱۰۰ است، اگر مثلاً درصد رس زیاد شود درصد شن باید کم شود. در مدل‌سازی مکانی PSFs چون این همبستگی در نظر گرفته نمی‌شود، مجموع سه جزء تخمینی در یک مکان الزاماً برابر با ۱۰۰ نمی‌شود.

جدول ۱- کارایی مدل جنگل تصادفی در مدل سازی داده های تبدیل نشده (UT) و داده های تبدیل شده

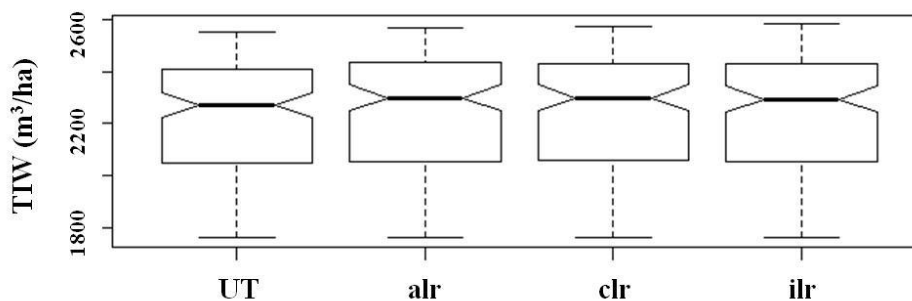
نوع داده	جزء بافت	RMSE	ME	AD	بافتهای درست طبقه بندی شده (%)
UT (رویکرد ۱)	رس	۲/۹۸	۰/۰۶		
	سیلت	۳/۲۵	-۰/۰۲	۰/۲۳۷	۷۳/۶
	شن	۴/۳۳	-۰/۰۹		
UT (رویکرد ۲)	رس	۳/۰۵	۰/۱۱		
	سیلت	۳/۲۹	۰/۰۳	۰/۲۳۴	۶۹/۶
	شن	۴/۳۰	-۰/۰۳		
alr	رس	۲/۹۲	-۰/۳۶		
	سیلت	۳/۳۵	-۰/۶۶	۰/۲۰۵	۷۰/۳
	شن	۴/۳۷	۱/۰۱		
clr	رس	۲/۹۴	-۰/۳۶		
	سیلت	۳/۲۱	-۰/۶۷	۰/۲۰۶	۷۰/۳
	شن	۴/۴۵	۱/۰۳		
ilr	رس	۲/۹۴	-۰/۲۹		
	سیلت	۳/۲۱	-۰/۷۵	۰/۲۰۵	۷۰/۹
	شن	۴/۴۱	۱/۰۴		

AD: فاصله Aitchison که معیاری است برای نشان دادن دقت تخمین ها در داده های ترکیبی؛ alr، clr و ilr به ترتیب تبدیل های centroid، additive و isometric هستند؛ رویکرد ۱: هر کدام از اجزاء بافت جداگانه با استفاده از متغیرهای کمکی تخمین زده شده اند؛ رویکرد ۲: هر جزء بافت خاک از اختلاف دو جزء تخمینی از ۱۰۰ درصد به دست آمده است.



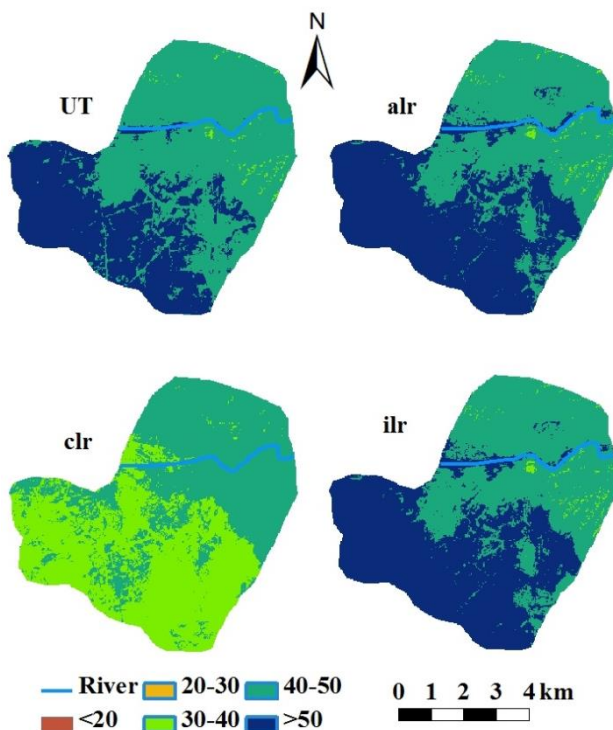
شکل ۲- توزیع مکانی مجموع اجزاء تخمینی بافت (چپ) و طبقه بندی مجموع آنها به سه گروه کمتر از ۱۰۰، ۱۰۰ و بیشتر از ۱۰۰ (راست)

اگر داده های تخمینی بافت خاک به عنوان ورودی توابع انتقالی مورد استفاده قرار گیرند تبدیل یا عدم تبدیل و همچنین نوع تبدیل ممکن است روی خروجی این توابع تاثیر بگذارد. به همین منظور با استفاده از یک PTF ابتدا مقدار آب قابل استفاده و سپس آب سهل الوصول با استفاده از داده های تبدیل نشده و داده های تبدیل شده محاسبه گردید. در نهایت مقدار کل آب آبیاری مورد نیاز (TIW) برای یک هکتار تحت کشت گندم در طول یک فصل رشد تا عمق ۳۰ سانتی متری محاسبه شد. نتایج در شکل (۳) نشان داده شده است. این نتایج بیانگر عدم وجود اختلاف معنی دار بین مقادیر TIW است. به عبارت دیگر تبدیل یا عدم تبدیل داده های PSFs و نوع تبدیل تاثیر معنی داری روی TIW نداشته است.



شکل ۳- نمودارهای جعبه‌ای مقدار کل آب آبیاری (TIW) مورد نیاز برای گندم در طول فصل رشد که با استفاده از داده‌های تبدیل نشده (UT) و تبدیل شده (ilr و clr، alr) تخمین زده شده است. همپوشانی قسمت‌های فرورفته نمودارهای جعبه‌ای بیانگر عدم وجود اختلاف معنی‌دار است

تأثیر تبدیل داده‌های PSFs روی توزیع مکانی ظرفیت آب قابل استفاده (AWC) تخمینی با استفاده از PTF به کار رفته در شکل (۴) نشان داده شده است. روند کلی توزیع مکانی AWC در دو حالت استفاده از داده‌های تبدیل شده و داده‌های تبدیل نشده به عنوان ورودی PTF تقریباً مشابه است. فقط تبدیل clr منجر به توزیع مکانی متفاوتی بخصوص در قسمت‌های جنوبی و جنوب غربی گردیده است. مقادیر بیشتر AWC مربوط به بخش‌های جنوبی و جنوب غربی منطقه است. الگوی کلی توزیع مکانی AWC منطبق با الگوی توزیع مکانی PSFs است و جاهایی که رس بیشتر است مقدار AWC هم بیشتر است. نتایج مشابهی در بررسی‌های محققین دیگر گزارش شده است (Haghverdi et al., 2015; Lai et al., 2018). در نظر نگرفتن تغییرات مکانی خواصی از خاک که روی AWC تأثیرگذارند ممکن است منجر به آبیاری بیشتر یا کمتر از حد نیاز گردد. بنابراین در مدیریت آبیاری بهتر است نوع داده (تبدیل شده یا تبدیل نشده) و نوع تبدیل را در نظر گرفت.



شکل ۴- توزیع مکانی ظرفیت آب قابل استفاده (AWC) تخمینی با استفاده از داده‌های تبدیل شده (ilr و clr، alr) و داده‌های تبدیل نشده (UT)



نتیجه‌گیری

نتایج این تحقیق نشان داد با تلفیق تکنیک‌های نقشه‌برداری رقومی خاک با الگوریتم‌های یادگیری ماشین می‌توان نقشه‌های مفیدی از PSFs تهیه نمود. همچنین عوارض سرزمین به دست آمده از DEM و داده‌های سنجش از دور، متغیرهای قابل اعتمادی برای مدل‌سازی PSFs هستند. بر اساس RMSE، مدل RF کارآیی تقریباً مشابهی در تخمین داده‌های تبدیل نشده و داده‌های تبدیل شده داشت؛ هر چند که تبدیل داده‌ها باعث به دست آمدن تخمین‌های اریب گردید. همچنین نوع تبدیل داده‌ها تأثیر معنی‌داری در کارآیی مدل RF نداشت. یافته‌ها نشان داد اگر سه جزء بافت خاک جداگانه تخمین زده شوند، دقت تخمین‌ها بیشتر از حالتی است که دو جزء تخمین زده شوند و جزء سوم از اختلاف دو جزء تخمینی از ۱۰۰ به دست آید. نتایج بیانگر این بود که تبدیل داده‌ها و نوع تبدیل وقتی داده‌های PSFs به عنوان ورودی یک PTF برای تخمین آب مورد نیاز برای گندم به کار رفتند، تأثیر معنی‌داری در مقادیر تخمینی نداشت ولی داده‌های تبدیل شده منجر به تخمین‌های اریب گردید. بنابراین برای جلوگیری از تخمین‌های اریب PSFs و آب مورد نیاز، بهتر است از داده‌های تبدیل نشده استفاده کرد.

منابع

- Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P., Minasny, B., McBratney, A.B. and Greve, M.H. 2013. High-resolution 3-d mapping of soil texture in Denmark. *Soil Sci. Soc. Am. J.* 77, 860–876.
- Aitchison, J. 1986. *The Statistical Analysis of Compositional Data*. Chapman & Hall, London.
- Haghverdi, A., Leib, B.G., Washington-Allen, R.A., Ayers, P.D. and Buschermohle, M.J. 2015. High-resolution prediction of soil available water content within the crop root zone. *J. Hydrol.* 530, 167–179.
- Hengl, T., Mendes de Jesus, J., Heuvelink G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al., 2017. SoilGrids250m. Global gridded soil information based on machine learning. *PLoS ONE* 12(2): e0169748.
- Khodaverdiloo, H., Homaei, M., van Genuchten, M.T. and Ghorbani Dashtaki, S. 2011. Deriving and validating pedotransfer functions for some calcareous soils. *J. Hydrol.* 399, 93–99.
- Lai, X., Zhou, Z., Zhu, Q. and Liao, K. 2018. Comparing the spatio-temporal variations of soil water content and soil free water content at the hillslope scale. *Catena*, 160, 366-375.
- Lark, R.M. and Bishop, T.F.A., 2007. Cokriging particle size fractions of the soil. *Eur. J. Soil Sci.* 58, 763–774.
- Malone, B., McBratney, A.B., Minasny, B. and Laslett, G. 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma*, 154, 138–152.
- Minasny, B. and McBratney, A.B. 2018. Limited effect of organic matter on soil available water capacity. *Eur. J. Soil Sci.* 69, 39-47.
- Odeh, I.O.A., Todd, A.J. and Triantafyllis, J. 2003. Spatial prediction of soil particle-size fractions as compositional data. *Soil Sci.* 168, 501–515.
- Pawłowsky-Glahn, V. and Egozcue, J.J. 2006. Compositional data and their analysis: an introduction, in: Bucciani, A., Mateu-Figueras, G., Pawłowsky-Glahn, V. (Eds.), *Compositional Data Analysis in the Geosciences: From Theory to Practice*. Geological Society, London, Special Publications, pp. 1 - 10.
- Poggio, L. and Gimona, A. 2017. 3D mapping of soil texture in Scotland. *Geoderma Regional*, 9, 5–16.
- Roman Dobarco, M., Arrouays, D., Lagacherie, P., Ciampalini, R. and Saby, N.P. 2017. Prediction of topsoil texture for Region Centre (France) applying model ensemble methods. *Geoderma*. 298, 67-77.
- Umali, B.P., Oliver, D.P., Forrester, S., Chittleborough, D.J., Hutson, J.L., Kookana, R.S. and Ostendorf, B. 2012. The effect of terrain and management on the spatial variability of soil properties in an apple orchard. *Catena*, 93, 38–48.
- Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y.A., Padarian, J. and Schaap, M.G. 2017. Pedotransfer functions in Earth system science: Challenges and perspectives. *Rev. Geophys.* 55, 1199-1256.
- Wang, Z. and Shi, W. 2017. Mapping soil particle-size fractions: A comparison of compositional kriging and log-ratio kriging. *J. Hydrol.* 546, 526–541.

Topic for submission: Pedometry and Soil Evaluation

Effects of transformation of soil texture data and digital mapping of texture components on available water

Amirian-Chakan^{*1}, A., Darvishpasand², Z., Akbarifazli³, R., Taghizadeh-Mehrjardi⁴, R., Khordehbin², S.

¹ Assistant Prof., Soil Science Department, Faculty of Agriculture, Lorestan University, Iran

² MSc. Department of Rangeland and Watershed Management, Behbahan Khatam Alanbia University of Technology, Iran

³ PhD. Student, Department of Soil Science, Khuzestan Science and Research Branch, Islamic Azad University, Iran

⁴ Assistant Prof., Faculty of Agriculture and Natural Resources, Ardakan University, Iran

Abstract

When separate components of the texture (sand, silt, and clay) are predicted independently in digital soil mapping (DSM), there is no guarantee that the separate estimates will sum to 100%. Log-ratio transformations before DSM modelling guarantee a constant sum of the estimates. Little is known about the effect of transformation of particle-size fractions (PSFs) when DSM products were used to predict soil hydraulic properties using pedotransfer functions. This study was conducted in Khuzestan province to assess the performance of random forest models (RF) to predict untransformed (UT) and transformed (TD) PSFs and to study the effects of log-ratio transformations (additive, centroid and isometric) on the estimations of total amount of irrigation water (TIW) for wheat. However, transformation resulted in biased estimates of PSFs, RF, in terms of RMSE, provided similar accuracies in predicting PSFs for both UT and TD. In addition, RF prediction based on UT data resulted in more correctly soil texture classes allocation. The spatial distribution of the sum of the predicted UT fractions indicated only small parts of the area conformed to the 100% sum. Data transformation can result in biased estimates of TIW. The findings indicated no significant difference between transformation methods in predicting TIW.

Keywords: Available water capacity; Compositional data; Digital soil mapping; Pedotransfer function

* Corresponding author, Email: amirian.ar@lu.ac.ir