

Spatial Prediction of Soil Saturated Hydraulic Conductivity by Integrating Soil Properties and Environmental Covariates

Abstract:

Soil hydraulic properties, particularly saturated hydraulic conductivity (K_s), play a crucial role in addressing problems related to soil and water management in agricultural, ecological, and environmental systems. Direct measurement of these properties is often difficult, costly, and time-consuming; hence, indirect estimation methods are commonly employed. In this study, the efficiency of multiple linear regression (MLR), decision tree (DT), and artificial neural network (ANN) methods was evaluated for estimating and mapping the spatial distribution of K_s in parts of the Cherdawel and Chamshir sub-basins (Ilam Province, Iran), using readily measurable soil properties along with environmental covariates (terrain attributes and remote sensing data). For this purpose, K_s was measured at 95 sampling points using a Guelph permeameter, and several readily measurable soil properties along with environmental covariates were also obtained at the same locations. The validity of the derived models for K_s estimation was assessed using the coefficient of determination (R^2), root mean square error (RMSE), and mean error (ME). The results demonstrated that the ANN model outperformed both MLR and DT models in estimating K_s . While the MLR and DT tended to underestimate K_s , the ANN model produced more accurate and reliable predictions. Furthermore, the spatial variability map of K_s could be successfully generated by integrating soil properties with environmental covariates, suggesting the usefulness of this approach for developing agro-hydrological models in data-limited regions. Overall, the findings suggest that incorporating environmental covariates alongside soil properties can significantly enhance the accuracy of K_s estimation, particularly when using the ANN model.

Keywords: Artificial neural network, Decision tree, Remote sensing, soil hydraulic properties, terrain attributes

برآورد مکانی هدایت هیدرولیکی اشباع خاک با استفاده از ویژگی‌های خاک و متغیرهای محیطی

چکیده

ویژگی‌های هیدرولیکی خاک، به‌ویژه هدایت هیدرولیکی اشباع (K_s)، نقشی کلیدی در حل مسائل مرتبط با مدیریت خاک و آب در سیستم‌های کشاورزی، بوم‌شناسی و زیست‌محیطی دارند. اندازه‌گیری مستقیم این ویژگی‌ها دشوار، پرهزینه و زمان‌بر است؛ از این‌رو، استفاده از روش‌هایی غیرمستقیم برای برآورد آنها رایج است. در این پژوهش با بهره‌گیری از ویژگی‌های زودپافت خاک و متغیرهای کمکی محیطی (داده‌های عوارض زمین و سنجش از دور)، کارایی روش‌های رگرسیون خطی چندگانه (MLR)، درخت تصمیم (DT) و شبکه‌های عصبی مصنوعی (ANN) در برآورد و نقشه‌سازی توزیع مکانی K_s بخش‌هایی از حوضه‌های آبخیز چرداول و چمشیر (استان ایلام) ارزیابی شد. بدین منظور، K_s در ۹۵ نقطه از محدوده مورد مطالعه با استفاده از دستگاه نفوذپذیری آب به خاک گولف اندازه‌گیری شد. برخی از ویژگی‌های زودپافت خاک و متغیرهای کمکی محیطی نیز در این نقاط تعیین شد. سپس اعتبار مدل‌های مشتقاق یافته در برآورد K_s با استفاده از ضریب تبیین (R^2)، جذر میانگین مربعات خطا (RMSE) و میانگین خطا (ME) ارزیابی شد. نتایج ارزیابی این مدل‌ها نشان داد که مدل ANN عملکرد بهتری نسبت به مدل‌های MLR و DT در تخمین K_s داشت. مدل‌های MLR و DT مقادیر K_s را کمتر از مقدار واقعی برآورد کردند، در حالی که ANN تخمین‌هایی دقیق‌تر و پایاتر ارائه کرد. نتایج همچنین نشان داد که نقشه تغییرپذیری مکانی K_s می‌تواند با تلفیق ویژگی‌های زودپافت خاک و متغیرهای کمکی محیطی تهیه شود؛ این رویکرد می‌تواند به توسعه مدل‌های اگروهیدرولوژیکی در مناطق با محدودیت داده کمک کند. در مجموع یافته‌ها بیان‌گر آن است که افزودن داده‌های کمکی محیطی به ویژگی‌های خاک می‌تواند باعث افزایش دقت برآورد K_s ، به‌ویژه در مدل ANN، شود.

کلمات محیطی: درخت تصمیم، سنجش از دور، شبکه عصبی مصنوعی، عوارض زمین، ویژگی هیدرولیکی خاک

مقدمه

خاک و آب دو منبع طبیعی حیاتی هستند که بقای انسان‌ها بیش از هر چیز به آن‌ها وابسته است. ویژگی‌های هیدرولیکی خاک، که تعیین‌کننده رفتار جریان آب در محیط متخلخل خاک هستند، در مطالعات هیدرولوژی، اکولوژی، طراحی و شبیه‌سازی سامانه‌های آبیاری و زهکشی و همچنین پایداری اکوسیستم‌های کشاورزی نقشی مهم ایفا می‌کنند (Saha et al., 2021). تعیین دقیق پارامترهای هیدرولیکی خاک و لایه‌های آبخوان از اهمیت ویژه‌ای برخوردار است، زیرا این ویژگی‌ها در ارزیابی جریان آب در خاک نقش کلیدی دارند. به‌طور معمول، فرآیندهایی همچون جریان آب در خاک، حرکت آب و مواد به سمت ریشه‌های گیاه، جریان آب به سمت زهکش‌ها و چاه‌ها و همچنین تبخیر از سطح خاک، همگی تحت تأثیر ویژگی‌های هیدرولیکی لایه‌ی سطحی خاک هستند (Motaghian and Mohammadi, 2011; Gabriel et al., 2019).

هدایت هیدرولیکی مقدار جریان آب درون خاک را کنترل کرده و بیان‌گر سرعت جریان آب در خاک است (Kirkham, 2019; Cheng et al., 2005). به‌طور کلی، مقدار هدایت هیدرولیکی در خاک‌های گوناگون می‌تواند تا دوازده مرتبه از نظر بزرگی متفاوت باشد (Fitts, 2002). ویژگی‌های هیدرولیکی خاک، به‌ویژه هدایت هیدرولیکی اشباع (K_s)، به‌عواملی همچون بافت و ساختمان خاک، مقدار ماده آلی و چگالی ظاهری خاک وابسته است. افزون بر این، حجم، شکل، نوع و پیوستگی منافذ و همچنین توزیع اندازه‌ی منافذ خاک از مهم‌ترین عوامل مؤثر بر مقدار K و بر جریان آب در لایه‌های سطحی و زیر سطحی خاک می‌باشد (Pan et al., 2026). در میان انواع خاک‌ها، خاک‌های رسی معمولاً دارای هدایت هیدرولیکی اندکی می‌باشند (در محدوده‌ی 10^{-6} تا 10^{-10} سانتی‌متر بر ثانیه) (Neuzil, 1994)، در حالی که خاک‌های شنی و سنگ‌ریزه‌ای هدایت هیدرولیکی بالاتری دارند (در محدوده‌ی 10^{-1} تا 10^0 سانتی‌متر بر ثانیه) (Saha et al., 2021). برای هر نوع خاک، مقدار K_s با تغییر مقدار رطوبت اولیه خاک، خصوصیات شیمیایی سیال عبوری و دما نیز تفاوت می‌کند (Russo et al., 2018).

در اراضی کشاورزی، مقدار K لایه‌ی سطحی خاک یک ویژگی ذاتی است؛ با این حال، عملیات خاک‌ورزی، کارهای حفاظتی، روش‌های مدیریت بقایا و تناوب زراعی می‌توانند با تغییر شکل منافذ خاک، تراکم توده خاک و پایداری خاک، مقدار K را تغییر دهند (Alletto and Coquet, 2009). هدایت هیدرولیکی نقشی مؤثر در مدیریت مزارع کشاورزی، از جمله در آبیاری (توزیع آب بر سطح یا درون خاک) و مهندسی منابع آب دارد (Fashi et al., 2016). رفتار جریان آب در خاک نقش مهمی در تعیین سرنوشت آب دارد — اینکه آیا تبخیر می‌شود، به داخل خاک نفوذ می‌کند، جذب ریشه‌ی گیاهان می‌شود و یا به زهکش‌ها و

چاه‌ها منتقل می‌شود (Jarvis et al., 2013).

تغییرات مکانی و زمانی K_s در سطح مزرعه، یکی از محدودیت‌های اساسی در دستیابی به کارایی بالای آبیاری است (Romano and Palladino, 2002). پژوهش‌های متعددی نشان داده‌اند که هدایت هیدرولیکی نقشی تعیین‌کننده در مدیریت آبیاری و درک فرایندهای هیدرولوژی در مقیاس حوضه‌های آبخیز به ویژه در مباحث فرسایش خاک این نواحی دارد (Motaghian and Mohammadi, 2011; Sihag, 2018; Tian et al., 2018). بر این اساس، هدایت هیدرولیکی به عنوان یکی از پارامترهای ورودی‌های پایه در طراحی سیستم‌های جمع‌آوری رواناب، طراحی شبکه‌های زهکشی و مدل‌سازی جابه‌جایی مواد محلول در خاک، مورد توجه قرار می‌گیرد (Khodaverdiloo et al., 2017).

درک عمیق ما از توزیع مکانی هدایت هیدرولیکی و ظرفیت نفوذ لایه‌ی سطحی خاک، نقشی حیاتی در کاهش مؤثر رواناب سطحی و هدرفت نفوذ عمقی در خاک دارد. با این حال، در بسیاری از مناطق، به دلیل چالش‌های ذاتی در اندازه‌گیری مستقیم هدایت هیدرولیکی در شرایط مزرعه و نیاز به جمع‌آوری تعداد زیادی نمونه، داده‌های مکانی مربوط به K_s به صورت محدود و پراکنده در دسترس می‌باشد (Yang et al., 2026). افزون بر این، تهیه نقشه‌های مکانی هدایت هیدرولیکی برای مناطق وسیع، با استفاده از روش‌های متداول میدانی، کاری دشوار، پرهزینه، زمان‌بر و با عدم قطعیت همراه است (Schaap et al., 1999).

طی سالیان محققین علوم خاک مقادیر هدایت هیدرولیکی را برای انواع مختلف خاک‌ها ثبت کرده‌اند و این مقادیر را، هرچند با محدودیت‌هایی، در مدل‌سازی‌های مربوط به حوزه‌های کشاورزی و هیدرولوژیک مورد استفاده قرار داده‌اند. به عنوان مثال، Mustafa et al. (2019) برای نشان دادن لایه سطحی خاک در مدل آب زیرزمینی، از یک مقدار میانگین برای هدایت هیدرولیکی کل حوضه آبخیز استفاده کردند. به دلیل تغییرپذیری بالای ویژگی‌های هیدرولیکی خاک (Agyare et al., 2007)، بسیاری از پژوهش‌ها در خصوص کمی کردن K_s به استفاده از روابط تجربی و فیزیکی در قالب توابع انتقالی خاک (PTFs) (Fashi et al., 2016) محدود می‌شود. افزون بر این، در برخی روش‌های زمین‌آماري مانند وزن‌دهی معکوس فاصله (IDW)، کریجینگ (Kriging)، کوکریجینگ (coKriging) و اسپلاین (Spline) برای تهیه نقشه‌های مکانی K_s از داده‌های اندازه‌گیری شده استفاده می‌شوند (Motaghian and Mohammadi, 2011). این در حالی است که این روش‌های درون‌یابی، مقادیر کمی K_s را بر اساس الگوهای داده‌ها پیش‌بینی می‌کنند و فرایندهای پیوسته‌ی خاک (عوارض توپوگرافی، توسعه، تکامل، تشکیل و منشأ پدایش خاک) را نادیده می‌گیرند (Moore et al., 1993).

از این رو، پژوهشگران مدل‌هایی مبتنی بر شبکه عصبی مصنوعی (ANN) (Agyare et al., 2007; Motaghian and)، ماشین بردار پشتیبان (SVM) (Saha et al., 2021)، جنگل تصادفی (RF) (Mohammadi, 2011, Qaderi et al., 2018)، منطق فازی (Sihag, 2018)، درخت تصمیم (DT) (Moncada et al., 2014)، رگرسیون درختی (RT) (دهقانی و همکاران، ۱۳۹۰؛ McKenzie and Jacquier, 1997) و رگرسیون‌های خطی چندگانه (MLR) (Fashi et al., 2016;)، (Saha et al., 2021) را توسعه داده‌اند تا هدایت هیدرولیکی را بر اساس ویژگی‌های مختلف خاکی و زمین‌ریختی (داده‌های توپوگرافی و ژئومورفولوژیکی) و داده‌های سنجش از دور پیش‌بینی کنند (Moore et al., 1993; Agyare et al., 2007;)، استفاده از چنین رویکردهایی در مدل‌سازی، می‌تواند به مدل‌سازان حوزه‌های کشاورزی و هیدرولوژیک در تهیه نقشه مکانی هدایت هیدرولیکی در مقیاس مزرعه، منطقه یا حوضه آبخیز کمک کند. ویژگی‌های زمین‌ریختی نقشی بنیادین در تعدیل فرایندهای سطح زمین و اتم‌سفر دارد، زیرا شکل و ساختار عوارض زمین معمولاً حرکت مواد و آب را در چشم‌انداز کنترل می‌کند (Romano and Palladino, 2002); در نتیجه هیدرولوژی حوضه آبخیز را در مقیاس توپوگرافی تحت تأثیر قرار می‌دهد (Wilson and Gallant, 2000). استفاده از ویژگی‌های زمین‌ریختی (McBratney et al., 2000) و همچنین داده‌های سنجش از دور (Odeh and McBratney, 2000) برای مدل‌سازی K_s می‌تواند گزینه‌ای مناسب باشد، زیرا گردآوری این داده‌ها در مقایسه با نمونه‌برداری‌های گسترده خاک، به نسبت آسان‌تر است (Agyare et al., 2007).

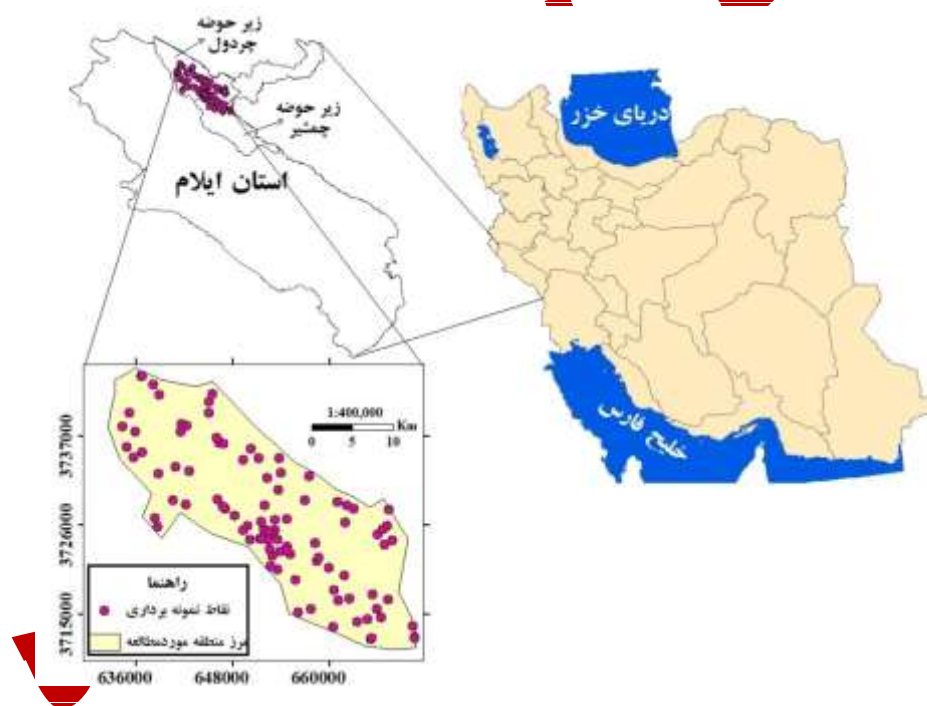
نقشه مکانی هدایت هیدرولیکی اشباع خاک به عنوان یکی از پارامترهای کلیدی ورودی در نرم‌افزارهای شبیه‌سازی رشد محصولات زراعی و هیدرولوژی و همچنین برخی مدل‌های فرسایش خاک به کار می‌رود تا تأثیر آن بر رفتار خاک و گیاه را در

مقیاس مزرعه و حوضه آبخیز برر سی کند (Motaghian and Mohammadi, 2011; Gumiere et al., 2014). در همین راستا، هدف این پژوهش اشتقاق مدل‌های MLR، DT و ANN مبتنی بر داده‌های خاکی و متغیرهای کمکی محیطی (داده‌های عوارض زمین و سنجش از دور) و پیش‌بینی نقشه تغییرات مکانی هدایت هیدرولیکی اشباع خاک در بخشی از حوضه‌های آبخیز چرداول و چمشیر در استان ایلام است.

مواد و روش‌ها

منطقه مورد مطالعه، نمونه‌برداری و تعیین ویژگی‌های خاک

محدوده مورد مطالعه در استان ایلام، در زیر حوضه‌های آبخیز چرداول و چمشیر، بین عرض‌های جغرافیایی $33^{\circ} 31' 23''/11'$ تا $33^{\circ} 50' 25''/59'$ شمالی و طول‌های جغرافیایی $46^{\circ} 25' 53''/08'$ تا $46^{\circ} 50' 13''/52'$ شرقی واقع شده است (شکل ۱). این منطقه بر اساس ویژگی‌های توپوگرافی، زمین‌شناسی، کاربری اراضی و قابلیت اراضی به واحدهای کاری مجزایی تفکیک گردید. پس از تعیین این واحدهای کاری، ۹۵ نقطه محل به روش تصادفی طبقه‌بندی شده و با توجه به مساحت هر واحد انتخاب گردید. در همه مکان‌های نمونه‌برداری، هدایت هیدرولیکی اشباع خاک (Ks) در عمق ۰-۳۰ سانتی‌متری با استفاده از دستگاه نفوذسنج گولف اندازه‌گیری شد. همزمان با اندازه‌گیری هدایت هیدرولیکی، نمونه‌هایی دست‌خورده جهت انجام تجزیه‌های آزمایشگاهی از همان محل برداشت شدند. نمونه‌ها پس از انتقال به آزمایشگاه، ابتدا هوا خشک شدند، سپس به نرمی کوبیده و در نهایت با عبور از الک ۲ میلی‌متری برای تجزیه‌های فیزیکی و شیمیایی آماده شدند. برای کسب جزئیات بیشتر راجع به توصیف منطقه مورد مطالعه به مقاله حکم‌زاد و همکاران (۱۳۹۷) رجوع شود.



شکل ۱. موقعیت منطقه مورد مطالعه و مکان‌های نمونه‌برداری در استان ایلام و ایران

در این پژوهش برای تعیین توزیع اندازه‌ی ذرات خاک از روش هیدرومتری و الک خشک، کربن آلی خاک از روش والکلی - بلک، جرم ویژه ظاهری خاک از روش کلوخه، جرم ویژه حقیقی خاک از روش پیکنومتر، رطوبت معادل ظرفیت زراعی با دستگاه صفحات فشاری (Pressure plate) و تخلخل کل خاک نیز از طریق نسبت جرم ویژه ظاهری (ρ_b) به جرم ویژه حقیقی (ρ_s) و استفاده از رابطه $n=1-\rho_b/\rho_s$ به دست آمد. همچنین، تخلخل مؤثر (θ_e) با استفاده از تخلخل کل و رطوبت معادل ظرفیت زراعی (θ_{FC}) و بهره‌گیری از رابطه $\theta_e = n - \theta_{FC}$ محاسبه شد (Dane and Topp, 2002). به منظور محاسبه‌ی میانگین هندسی و انحراف معیار هندسی قطر ذرات خاک، از روابط پیشنهادی شیرازی و بورسما (Shirazi and Borsema, 1984) بهره گرفته شد.

استخراج متغیرهای کمکی محیطی

مدل رقومی ارتفاع (DEM) با قدرت تفکیک مکانی $12/5 \times 12/5$ متر از پایگاه داده ALOS PALSAR (وبگاه سازمان زمین‌شناسی ایالات متحده، USGS) بارگیری و برای استخراج ویژگی‌های عوارض زمین استفاده شد. این مدل در قالبی شبکه‌ای به محیط نرم‌افزار SAGA-GIS V. 9.3.2 وارد شد و مشخصه‌های مختلف عوارض زمین شامل ارتفاع، درصد و جهت شیب، تحلیل سایه روشن، سطح کل حوضه، عمق دره، انحنای مسطح و پروفیلی، موقعیت نسبی شیب، شاخص همگرایی، شاخص رطوبت توپوگرافیک، شاخص قدرت جریان، فاکتور LS، فاصله از شبکه آبراهه، شاخص صافی کف دره با وضوح چندگانه (MrVBF) و شاخص همواری بالای پشته با درجه تفکیک بالا (MrRTF) استخراج گردید (Wilson and Gallant, 2000).

در مرحله بعد، برای استخراج متغیرهای سنجش از دور، از تصاویر ماهواره لندست ۸ مربوط به سنجنده‌های OLI/TIRS با قدرت تفکیک مکانی 30×30 متر (وبگاه USGS) استفاده شد. از باندهای B2, B3, B4, B5, B6, B7 و B10 مقادیر انعکاس طیفی محاسبه و براساس آن، شاخص‌های تفاوت پوشش گیاهی نرمال شده (NDVI) و تعدیل خاک و پوشش گیاهی (SAVI) تعیین شدند. در استفاده از داده‌های خاک و متغیرهای کمکی محیطی، برخی متغیرها ممکن است اهمیت کمتری در مدل‌سازی ایفا کرده و موجب افزایش پیچیدگی مدل و زمان پردازش آن شوند. از این‌رو، پس از اندازه‌گیری ویژگی‌های مبنایی خاک و آماده‌سازی متغیرهای کمکی محیطی، مرحله‌ی انتخاب متغیرها با هدف کاهش ابعاد داده و افزایش کارایی الگوریتم‌های مدل‌سازی انجام شد. برای استخراج و رتبه‌بندی متغیرهای مؤثر در مدل‌سازی، می‌توان از روش‌های متعددی بهره برد؛ از جمله آن‌ها می‌توان به انتخاب ویژگی مبتنی بر همبستگی (Correlation Based Feature Selection, CFS)، تحلیل مؤلفه‌های اصلی (Principal Component Analysis, PCA) و تحلیل عاملی (Factor Analysis, FA) اشاره کرد. در این پژوهش، برای شناسایی و گزینش متغیرهای مؤثر، از روش PCA با بهره‌گیری از جعبه ابزار XLSTAT در محیط نرم‌افزار EXCEL 2019 استفاده شد.

مدل‌های برآوردگر

برای برآورد مقدار هدایت هیدرولیکی اشباع خاک (K_s) رویکردهای رگرسیون خطی چندگانه (Multiple Linear Regression, MLR)، درخت تصمیم (Decision Tree, DT) و شبکه عصبی مصنوعی (Artificial Neural Network, ANN) در محیط نرم‌افزار R نسخه 4.5.1 توسعه داده شدند. در مدل رگرسیونی، با استفاده از روش کلاسیک حداقل مربعات و بهره‌گیری از روش گام‌به‌گام، رابطه‌ای میان متغیر هدف و مجموعه‌ای از متغیرهای مستقل برقرار شد. مدل اشتقاق یافته برای پیش‌بینی مقادیر K_s در نقطه محل‌هایی که داده‌های متغیرهای مستقل در آن نقاط قابل اندازه‌گیری بودند، مورد استفاده قرار گرفت. درخت تصمیم‌گیری (DT) یک ساختار سلسله‌مراتبی شبیه درخت است که برای انجام پیش‌بینی‌ها به کار می‌رود. روند ساخت آن بدین صورت است که الگوریتم در ابتدا با تمامی نمونه‌های آموزشی شروع می‌کند، سپس متغیر را انتخاب می‌کند که بهترین توانایی تفکیک یا پیش‌بینی را دارد و بر اساس آن، داده‌ها را به چند زیرمجموعه تقسیم می‌کند. شاخه‌های درخت نشان‌دهنده‌ی خروجی آزمون‌هایی هستند که الگوریتم در هر گره‌ی میانی انجام می‌دهد و پیش‌بینی نهایی در گره‌های برگ درخت حاصل می‌شود. فرایند هرس کردن با حذف گره‌های غیرضروری، از بیش‌برازش مدل بر داده‌های آموزشی جلوگیری می‌کند. در مرحله‌ی پایانی نیز یک درخت مدل هموار ساخته می‌شود تا ناپیوستگی‌های ناگزیر میان مدل‌های خطی قرار گرفته در برگ‌های درخت هرس شده کاهش یابد و یک ساختار پیش‌بینی پایدارتر و منسجم‌تری ایجاد شود (مصری و همکاران، ۱۳۹۹). برای مقایسه با مدل‌های رگرسیونی و درخت تصمیم، از شبکه عصبی مصنوعی نوع پرسپترون چندلایه (Multilayer Perceptron; MLP) با الگوریتم آموزشی پس انتشار خطا (Back-Propagation)، که به‌عنوان یکی از متداول‌ترین مدل‌ها برای پیش‌بینی ویژگی‌های خاک محسوب می‌شود، استفاده شد. شبکه MLP از مجموعه‌ای از واحدهای پردازشی (نورون‌های پایه) تشکیل شده است که شامل یک لایه ورودی، یک یا چند لایه پنهان و یک لایه خروجی می‌باشد (Motaghian and Mohammadi, 2011). در این پژوهش، شبکه MLP توسعه یافته شامل چهار لایه اصلی، یعنی یک لایه ورودی، دو لایه پنهان و یک لایه خروجی است. در معماری شبکه‌های عصبی، تعداد نورون‌های لایه‌های ورودی و خروجی معمولاً با تعداد پارامترهای ورودی و خروجی داده‌ها متناظر است. در این پژوهش، برای تعیین تعداد بهینه نورون‌های لایه‌های پنهان، مقادیر ریشه میانگین مربعات خطا (RMSE) شبکه نسبت به تعداد نورون‌های پنهان مورد ارزیابی قرار گرفت (نمودار ارائه نشده است). تعداد لایه‌های پنهانی که در آن، مقدار RMSE داده‌های آموزش کمترین باشد، به‌عنوان پیکربندی شبکه بهینه در نظر گرفته شد.

شایان ذکر است که پیش از آموزش شبکه، داده‌های ورودی نرمال شدند. هدف از نرمال‌سازی داده‌ها، تبدیل آن‌ها به مقادیری در بازه‌ی صفر تا یک است. زیرا در این پژوهش، تابع فعال‌سازی برای نورون‌های لایه پنهان، از نوع سیگموئیدی بوده که خروجی آن بین صفر تا یک می‌باشد. از این رو لازم است داده‌های ورودی به این تابع نیز اعدادی بین صفر و یک باشند (Motaghian and Mohammadi, 2011). بنابراین در این پژوهش نرمال‌سازی داده‌ها بر اساس رابطه زیر انجام گردید:

$$y_{norm} = \left[\frac{y - y_{min}}{y_{max} - y_{min}} \right] \quad (1)$$

که در آن، y_{norm} مقدار نرمال شده ورودی y ، y_{max} بیشینه و y_{min} کمینه داده‌های اولیه می‌باشد. پس از بدست آوردن بهترین مدل برای برآورد هدایت هیدرولیکی اشباع خاک (K_s)، نقشه‌های ویژگی‌های خاک، داده‌های عوارض زمین و تصاویر ماهواره‌ای وارد مدل شده و خروجی مدل که همان نقشه K_s باشد به دست آمد. کلیه فرایندهای مدل‌سازی با استفاده از کتابخانه‌های caret، rpart، neuralnet و تولید نقشه K_s توسط بهترین مدل با بهره‌گیری از کتابخانه Raster در محیط نرم افزار 4.5.1 R انجام شد. همچنین میان‌یابی و تولید نقشه‌های مرتبط با ویژگی‌های خاک با استفاده از روش کریجینگ معمولی (Ordinary Kriging) در محیط نرم افزار ArcGIS 10.8 انجام شد.

شاخص‌های ارزیابی مدل‌های پیشنهادی

به منظور ارزیابی نتایج حاصل از مدل‌های رگرسیونی و شبکه عصبی مصنوعی و مقایسه آن با داده‌های واقعی اندازه‌گیری شده، از آماره‌های ضریب همبستگی تطابق (ρ_c : Concordance correlation coefficient) ریشه میانگین مربعات خطا (RMSE: Root mean square error)، میانگین خطا (ME: Mean error) و ضریب تبیین (R^2) استفاده شد. در زیر بیان ریاضی این شاخص‌های آماری ارائه شده است:

$$\rho_c = \frac{2 \cdot \rho \sigma_M \sigma_E}{\sigma_M^2 + \sigma_E^2 + (\bar{M}_i + \bar{E}_i)^2} \quad (2)$$

$$RMSE = \left[\frac{\sum_{i=1}^n (E_i - M_i)^2}{n} \right]^{1/2} \quad (3)$$

$$ME = \left[\frac{\sum_{i=1}^n (E_i - M_i)}{n} \right] \quad (4)$$

$$R^2 = \left[1 - \frac{\sum_{i=1}^n (M_i - E_i)^2}{\sum_{i=1}^n (M_i - \bar{M}_i)^2} \right] \quad (5)$$

که در آن‌ها، M_i مقادیر اندازه‌گیری شده‌ی متغیر، E_i مقادیر پیش‌بینی شده متغیر، \bar{E}_i میانگین مقادیر پیش‌بینی شده متغیر، \bar{M}_i میانگین مقادیر اندازه‌گیری شده متغیر، ρ ضریب همبستگی پیرسون بین مقادیر اندازه‌گیری شده و پیش‌بینی شده متغیر، σ_M انحراف معیار مقادیر اندازه‌گیری شده متغیر، σ_E انحراف معیار مقادیر پیش‌بینی شده متغیر و n تعداد کل نمونه‌ها می‌باشد. با توجه به این آماره‌ها، مدلی دارای نتایج بهتری است که کمترین مقدار RMSE و ME (مقادیر نزدیک به صفر) و ρ_c و R^2 نزدیک به ۱ داشته باشد.

نتایج و بحث

نتایج حاصل از تجزیه و تحلیل نمونه‌های

نتایج آماره‌های توصیفی ویژگی‌های خاکی اندازه‌گیری شده و همچنین متغیرهای کمکی محیطی در جدول (۱) ارائه شده است. Wilding (1985) یک سیستم طبقه‌بندی برای شناسایی میزان تغییرپذیری ویژگی‌های خاک بر اساس مقادیر ضریب تغییرات (CV) ارائه کرد که در آن مقادیر CV در محدوده ۰-۱۵٪، ۱۵-۳۵٪ و بیش از ۳۵٪ به ترتیب نشان‌دهنده تغییرپذیری کم، متوسط و زیاد هستند. بالا بودن ضریب تغییرات بیان‌گر وجود تغییرات مکانی زیاد ویژگی‌های تعیین شده در منطقه مورد مطالعه می‌باشد. با توجه به جدول (۱)، CV برای هدایت هیدرولیکی اشباع خاک (K_s)، درصد شن، میانگین هندسی قطر ذرات، درصد ماده آلی، تخلخل مؤثر، ارتفاع، درصد شیب، فاکتور LS، شاخص انحنای سطح، شاخص صافی کف دره با وضوح چندگانه (MrVBF)، شاخص همواری بالای پشته با درجه تفکیک بالا (MrRTF)، شاخص قدرت جریان، فاصله از شبکه آبراهه، سطح

کل حوضه، شدت انعکاس در B4 و B5، در کلاس بالا قرار دارند. این در حالی است که درصد رس، درصد سیلت، انحراف معیار هندسی قطر ذرات، شاخص رطوبت توپوگرافیک (TWI) و شاخص‌های تفاوت پوشش گیاهی نرمال شده (NDVI) و تعدیل خاک و پوشش گیاهی (SAVI) در کلاس متوسط است و جرم ویژه ظاهری و حقیقی خاک، تخلخل کل و تحلیل سایه-روشن در کلاس کم قرار می‌گیرد. آستانه‌ی کمینه و بیشینه K_s به ترتیب ۰/۰۱۴ و ۱/۳۳۲ با میانگینی معادل ۰/۲۳۳ متر بر روز می‌باشد. برای کسب جزئیات بیشتر راجع به همبستگی بین ویژگی‌های خاکی و متغیرهای کمکی محیطی به مقاله حکمت‌زاد و همکاران (۱۳۹۷) رجوع شود.

جدول ۱. آماره‌های توصیفی برخی از ویژگی‌های خاکی اندازه‌گیری شده و متغیرهای کمکی محیطی در منطقه مورد مطالعه

ویژگی‌های مورد بررسی	کمینه	بیشینه	میانگین	انحراف معیار	درصد ضریب تغییرات
هدایت هیدرولیکی اشباع خاک (m/d)	۰/۰۱۴	۱/۳۳۲	۰/۲۳۳	۰/۲۹۱	۱۲۵/۱
رس (%)	۸	۶۱	۳۵/۶	۱۰/۲	۲۸/۷
سیلت (%)	۲۷	۶۶	۴۶/۶	۷/۱	۱۵/۲
شن (%)	۱/۷۳	۶۳	۱۷/۷۴	۹/۴۴	۵۲/۸
میانگین هندسی قطر ذرات (mm)	۰/۰۰۱۸	۰/۱۹	۰/۰۲	۰/۰۲	۱۰۴/۷
انحراف معیار هندسی قطر ذرات	۶/۲	۱۵/۵	۱۰/۶	۲/۲	۲۰/۸
جرم ویژه ظاهری (g/cm^3)	۱/۰۱۵	۱/۴۹۳	۱/۳۲۷	۰/۰۸۶	۶/۴۶
جرم ویژه حقیقی (g/cm^3)	۲/۴۰	۲/۸۰	۲/۶۰	۰/۰۷	۲/۶۴
ماده آلی (%)	۰/۵۰	۵/۳۸	۱/۸۱	۰/۹۴	۵۱/۶
تخلخل کل (cm^3/cm^3)	۰/۴۴۳	۰/۵۸۴	۰/۴۸۹	۰/۰۲۷	۵/۶
تخلخل موثر (cm^3/cm^3)	۰/۰۱	۰/۳۲	۰/۱۲۷	۰/۰۵۶	۴۴/۰۲
ارتفاع (m)	۷۳۴/۴	۱۶۳۴/۵	۱۰۸۵/۸	۲۰۱/۴	۱۸۵/۵
درصد شیب	۰	۵۱/۷۹	۱۵/۲۱	۱۰/۷۱	۷۹/۸
شاخص رطوبت (-)	۴/۴۰	۱۶/۰۴	۷/۱۹	۲/۱۸	۳۰/۳
فاکتور LS (-)	۰/۰۰۰۲	۳۰/۱۴۷	۶/۰۱۰	۶/۳۰۰	۱۰۳/۲
شاخص انحنای مسطح (1/m)	-۰/۰۰۲۴	۰/۰۰۵۰	-۰/۰۰۰۳	۰/۰۰۱۳	۴۶۵/۲
شاخص انحنای پروفیلی (1/m)	-۰/۰۰۴۹	۰/۰۰۴۳	-۰/۰۰۰۵	۰/۰۰۱۵	*
شاخص صافی کف دره با وضوح چندگانه (MrVBF)	6×10^{-9}	۴/۸۲	۰/۸۲	۰/۹۵	۱۱۶/۳
شاخص همواری بالای پشته با درجه تفکیک بالا (MrRTF)	$6/00 \times 10^{-9}$	۱/۴۷	۰/۰۸	۰/۲۱	۲۷۴/۸
شاخص قدرت جریان (-)	۰/۰۱۳	۹۱/۱۵۸	۲۳/۶۴۴	۱۸/۸۶۵	۷۹/۸
فاصله از شبکه آبراهه (m)	۷۳۳/۰	۱۶۲۵/۲	۱۰۷۷/۸	۱۹۷/۲	۱۸۳/۰
تحلیل سایه روشن (-)	۰/۴۰۲	۰/۹۶۶	۰/۷۸۳	۰/۱۰۰	۱۲/۷
سطح کل حوضه (m^2)	۱۷۸/۸	۱۸۵۵۴۸۰/۰	۳۳۹۲۴/۰	۲۵۲۵۹۱/۱	۵۷۵/۰
شدت انعکاس در B4 (-)	۰/۰۷۵	۰/۲۷۵	۰/۱۴۸	۰/۰۴۴	۲۹۴/۶
شدت انعکاس در B5 (-)	۰/۲۵۲	۰/۴۵۰	۰/۳۱۹	۰/۰۴۳	۱۲۴/۱
شاخص تفاوت پوشش گیاهی نرمال شده (NDVI) (-)	۰/۱۱۲	۰/۶۸۹	۰/۳۷۳	۰/۱۳۰	۳۴/۹

* به دلیل منفی بودن میانگین ضریب تغییرات قابل محاسبه نیست.

انتخاب ویژگی‌های خاکی و متغیرهای کمکی بهینه

به منظور کاهش تعداد متغیرهای ورودی کم‌اهمیت، از روش تجزیه مولفه‌های اصلی (PCA) استفاده شد. این روش امکان خلاصه‌سازی مجموعه‌ای بزرگ از متغیرهای هم‌بسته را در قالب تعدادی محدود از متغیرهای معرف و مستقل فراهم می‌کند. متغیرهای معرف قادرند بیشترین میزان تغییرپذیری را در داده‌های اولیه تبیین کنند و در شرایطی که تعداد زیادی متغیر پیش‌بینی کننده‌ی وابسته وجود دارد، رویکردی کارآمد و مؤثر محسوب می‌شوند (Shahabi et al., 2022).

بر اساس نتایج حاصل از اجرای PCA، از میان ۴۰ ویژگی خاک و متغیر کمکی محیطی، تعداد ۱۲ متغیر مؤثر به عنوان ورودی نهایی برای مدل‌سازی هدایت هیدرولیکی اشباع خاک (K_s) انتخاب شدند. این متغیرها شامل درصد شیب، شاخص رطوبت توپوگرافیک، فاکتور LS، تحلیل سایه روشن - روشن، سطح کل حوضه و فاصله از شبکه آبراهه بوده که به عنوان عوامل توپوگرافی در نظر گرفته می‌شوند (جدول ۲). از میان ویژگی‌های خاک، درصد سیلت، درصد شن، d_{60} و جرم مخصوص حقیقی خاک به عنوان

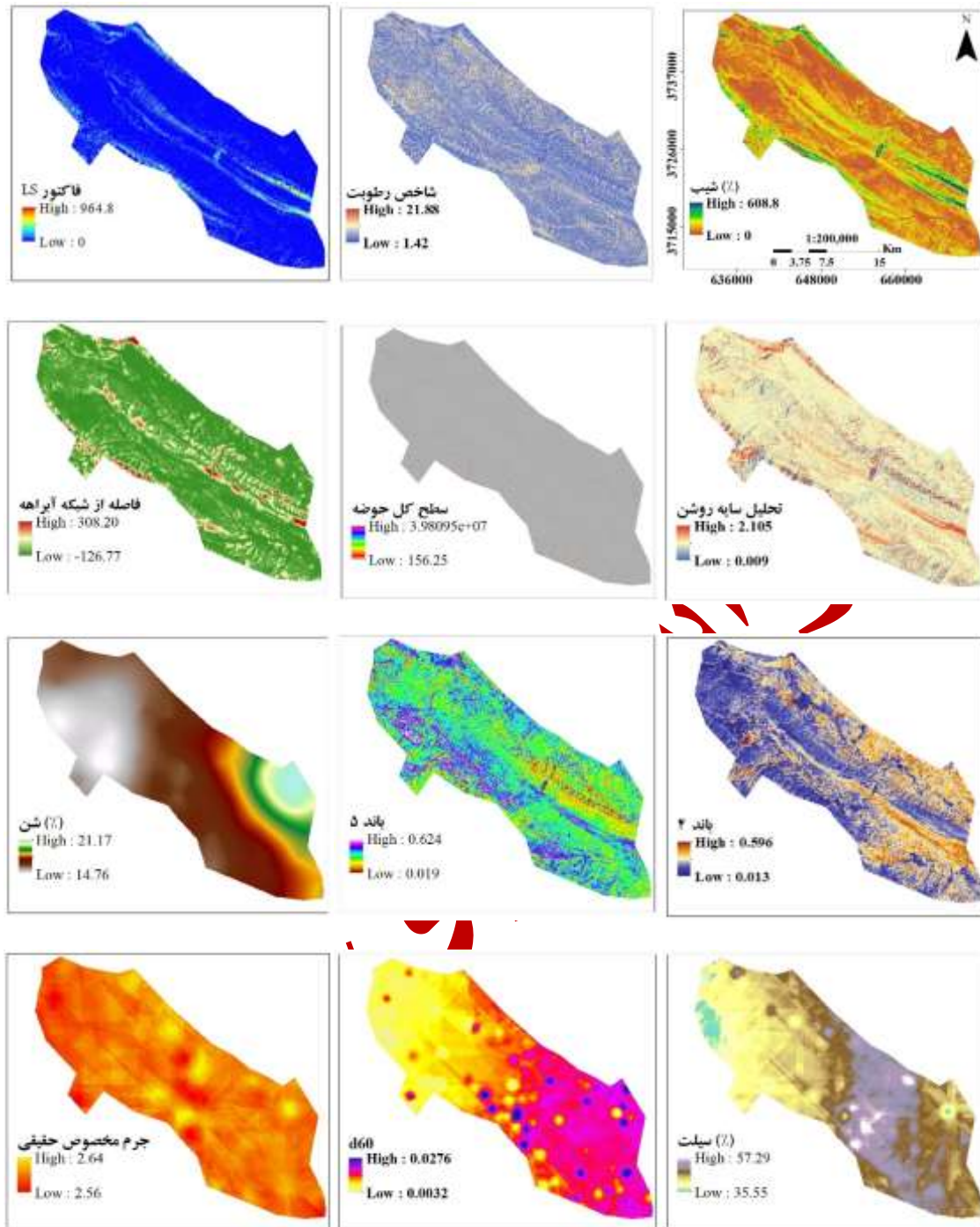
مؤثرترین ویژگی‌های خاکی (جدول ۳) و از میان شاخص‌های سنجش از دور، باندهای ۴ و ۵ به‌عنوان عوامل بیوفیزیکی (داده‌های حاصل از PCA این بخش ارائه نشده است) تشخیص داده شدند. در شکل ۲ نقشه مربوط به متغیرهای خاکی و کمکی محیطی مورد استفاده برای مدل‌سازی مکانی هدایت هیدرولیکی اشیاع خاک ارائه شده است.

جدول ۲. نتایج حاصل از تجزیه به مؤلفه‌های اصلی داده‌های عوارض زمین منطقه مورد مطالعه (شاخص‌های پرننگ به عنوان متغیرهای مؤثر انتخاب شده اند)

ارزش ویژه	۴/۶۳۲	۲/۵۸۳	۱/۸۷۹	۱/۵۰۰	۱/۲۹۹
واریانس (%)	۹۵۴/۲۸	۱۴۲/۱۶	۷۴۴/۱۱	۳۷۵/۹	۱۱۷/۸
واریانس تجمعی (%)	۹۵۴/۲۸	۰۹۶/۴۵	۸۴۰/۵۶	۲۱۵/۶۶	۳۳۲/۷۴
مؤلفه ۱	مؤلفه ۲	مؤلفه ۳	مؤلفه ۴	مؤلفه ۵	
عمق دره	-۰/۶۰۲	۰/۲۵۷	-۰/۱۳۵	-۰/۱۱۸	-۰/۳۵۶
سطح کل حوضه	-۰/۲۴۵	۰/۱۵۲	۰/۳۱۸	۰/۷۵۸	۰/۱۰۸
شاخص رطوبت توپوگرافیک	-۰/۵۰۵	۰/۱۴۳	۰/۵۱۱	۰/۵۱۷	۰/۲۵۱
شاخص قدرت جریان	۰/۰۸۲	۰/۴۳۸	-۰/۱۶۷	۰/۰۱۷	-۰/۲۳۵
شیب (%)	۰/۸۳۵	۰/۴۶۰	-۰/۰۹۱	۰/۰۴۵	-۰/۱۷۴
موقعیت نسبی شیب	۰/۶۷۶	-۰/۳۱۳	۰/۰۴۴	۰/۱۰۳	۰/۲۵۵
انحنای پروفیل	۰/۰۷۵	-۰/۶۳۲	-۰/۳۸۴	۰/۰۵۵	۰/۰۶۲
انحنای مسطح	۰/۵۸۱	-۰/۴۲۲	-۰/۲۹۹	۰/۳۴۷	-۰/۱۷۱
MRVBF	-۰/۷۲۳	-۰/۲۴۰	۰/۱۵۱	-۰/۰۹۲	-۰/۰۶۴
MRRTF	-۰/۱۱۵	-۰/۴۲۲	-۰/۴۳۵	۰/۱۵۴	۰/۰۸۱
فاکتور LS	۰/۵۵۱	۰/۷۰۳	۰/۰۷۵	۰/۲۴۲	-۰/۰۴۵
ارتفاع	۰/۶۴۱	-۰/۳۰۵	۰/۵۰۴	-۰/۲۳۸	۰/۲۷۵
شاخص همگرایی	۰/۳۱۶	-۰/۳۴۴	-۰/۳۵۶	۰/۴۸۵	۰/۰۷۱
فاصله از شبکه آبراهه	۰/۶۱۲	-۰/۲۹۶	۰/۵۲۵	-۰/۲۵۱	۰/۲۶۹
جهت شیب	-۰/۲۳۰	۰/۳۱۱	-۰/۴۴۶	-۰/۰۹۵	۰/۶۱۷
تحلیل سایه روشن	۰/۰۴۸	-۰/۵۱۹	۰/۳۸۸	۰/۱۰۸	-۰/۶۰۶

جدول ۳. نتایج حاصل از تجزیه به مؤلفه‌های اصلی ویژگی‌های خاک منطقه مورد مطالعه (شاخص‌های پرننگ به عنوان متغیرهای مؤثر انتخاب شده اند)

ارزش ویژه	۶/۰۱۱	۲/۶۲۴	۱/۶۴۸	۱/۱۱۹
واریانس (%)	۴۰/۰۷۴	۱۷/۴۹۵	۱۰/۹۸۶	۷/۴۶۰
واریانس تجمعی (%)	۴۰/۰۷۴	۵۷/۵۷۰	۶۸/۵۵۶	۷۶/۰۱۶
مؤلفه ۱	مؤلفه ۲	مؤلفه ۳	مؤلفه ۴	
رس (%)	-۰/۸۴۴	-۰/۴۱۸	-۰/۱۳۴	-۰/۱۲۵
سیلت (%)	-۰/۰۱۰	۰/۸۲۲	-۰/۳۳۰	-۰/۱۷۲
شن (%)	۰/۹۱۴	-۰/۱۶۳	۰/۲۷۰	-۰/۰۰۵
d ₁₀ (mm)	۰/۳۹۰	۰/۴۵۲	-۰/۳۸۴	-۰/۴۴۱
d ₂₀ (mm)	۰/۶۵۷	۰/۴۰۸	-۰/۳۷۴	۰/۱۱۸
d ₅₀ (mm)	۰/۸۰۳	۰/۲۷۳	-۰/۱۳۱	۰/۲۹۸
d ₆₀ (mm)	۰/۵۰۱	۰/۱۰۰	-۰/۱۱۰	۰/۵۵۴
میانگین هندسی قطر ذرات (mm)	۰/۷۳۶	-۰/۰۷۰	۰/۴۰۹	-۰/۳۰۸
انحراف معیار هندسی قطر ذرات (mm)	۰/۶۵۷	-۰/۴۲۰	۰/۱۱۶	۰/۳۳۰
جرم ویژه حقیقی (gr.cm ⁻³)	-۰/۱۱۰	۰/۲۵۱	۰/۷۵۴	-۰/۲۴۱
جرم ویژه ظاهری (gr.cm ⁻³)	۰/۸۶۳	-۰/۲۸۲	۰/۱۰۹	-۰/۱۴۲
تخلخل	-۰/۸۴۷	۰/۳۴۷	۰/۱۶۹	۰/۰۴۰
تخلخل مؤثر	۰/۰۲۵	۰/۷۳۱	۰/۴۲۹	۰/۲۳۰
ماده آلی خاک (%)	-۰/۱۳۵	-۰/۵۳۹	-۰/۴۱۷	-۰/۲۲۲



شکل ۲. متغیرهای خاکی و کمکی محیطی مورد استفاده برای مدل سازی مکانی هدایت هیدرولیکی اشباع خاک (K_s)

برآورد هدایت هیدرولیکی اشباع خاک (K_s) با استفاده از رگرسیون خطی چندگانه (MLR)

تحلیل رگرسیون خطی چندگانه (MLR) به روش گام به گام ورود پارامترها برای برآورد هدایت هیدرولیکی اشباع خاک (K_s) به کار گرفته شد. در این تحلیل، از تمامی ویژگی‌های خاک و متغیرهای کمکی محیطی، که در مرحله قبل و با استفاده از تحلیل مؤلفه‌های اصلی (PCA) انتخاب شده بودند، استفاده گردید. نتایج نشان داد که مابین این پارامترها، درصد شن و جرم ویژه حقیقی خاک، مهمترین ویژگی‌های زودیافت خاک در برآورد K_s هستند. هیچ یک از پارامترهای عوارض زمین ورودی به مدل نهایی گام به گام، در مدل پیشنهادی باقی نماندند. در زیر مدل رگرسیونی چندگانه پیشنهادی، بین هدایت هیدرولیکی اشباع با پارامترهای زودیافت خاکی ارائه شده است:

$$\ln K_s = 9.491 + 0.058 Sa - 4.90 Pd$$

(۶)

در این رابطه، Sa شن (% و Pd جرم ویژه حقیقی خاک (gr/cm^3) است. با توجه به نتایج، مدل پیشنهاد شده برای خاک‌های مورد مطالعه، از دقت پیش‌بینی نسبتاً پایینی برخوردار است ($R^2 = 0/35$). این مقدار ضریب تبیین بدین معنی است که این مدل تنها توانسته حدود یک سوم از تغییرات هدایت هیدرولیکی اشباع خاک را در منطقه مورد مطالعه توجیه کند و حدود ۶۵ درصد از تغییرپذیری K_s توسط این مدل تبیین نگردیده است. از دیگر شاخص‌های ارزیابی مدل، جذر میانگین مربعات خطا (RMSE)، معادل $0/25 \text{ m/d}$ و میانگین خطا (ME) نیز معادل $-0/091 \text{ m/d}$ به دست آمد. مقدار منفی ME بیان‌گر تمایل مدل پیشنهادی به کم‌برآوردی هدایت هیدرولیکی اشباع خاک است. ضریب همبستگی تطابق (ρ) نیز بیان‌گر تطابق ضعیف برآوردهای مدل پیشنهادی می‌باشد (Viscarra Rossel and Hicks, 2015).

دقت نسبتاً پایین مدل پیشنهادی را شاید بتوان به وجود روابط غیرخطی بین هدایت هیدرولیکی اشباع و ویژگی‌های زود یافت خاک و همچنین داده‌های عوارض زمین نسبت داد که در مدل‌سازی MLR لحاظ نشده است. (Leiz et al. (2004) بین K_s و ویژگی‌های مبنایی خاک رابطه خطی معنی‌داری مشاهده نکردند. این محققین دلیل این موضوع را انجام نمونه برداری در نزدیکی سطح خاک عنوان کردند؛ جایی که K_s بیشتر تحت تأثیر فشردگی و بهم‌خوردگی خاک و همچنین فعالیت‌های بیولوژیک قرار می‌گیرد. یک دلیل دیگر پایین بودن R^2 و بالا بودن مقدار RMSE شاید ناشی از آن باشد که ویژگی‌های عوارض زمین استخراج شده، بر تغییرپذیری مکانی K_s در منطقه مورد مطالعه تأثیر معنی‌داری نداشته است. (Mosleh et al. 2016 در پژوهشی مشابه در استان چهارمحال و بختیاری، عملکرد MLR را در برآورد برخی از ویژگی‌های خاک ضعیف گزارش کردند.

هرچند مدل رگرسیون پیشنهادی از عملکرد نسبتاً پایینی برخوردار بود، اما نتایج حاصل از مرحله انتخاب متغیرها با استفاده از تحلیل مؤلفه‌های اصلی (PCA)، با یافته‌های سایر محققین همخوانی داشت. این نتایج با یافته‌های (Khaledian et al. (2017 که از PCA برای تفکیک شاخص‌های خاک با در نظرگیری متغیرهای مکانی استفاده کردند و همچنین با پژوهش (Zeraatpisheh et al. (2019 که این روش را با درون‌یابی و رن‌دهی معکوس فاصله برای پیش‌بینی ویژگی‌های خاک در مناطق خشک ایران تلفیق کردند، همخوانی دارد. افزون بر این، هم‌سو با این پژوهش، (Souza et al. (2018 از PCA برای کاهش ابعاد داده‌ها و مدل‌سازی ویژگی‌های خاک‌های آلتی‌سول بهره‌گرفتند.

بر اساس آماره‌های ارزیابی ارائه شده در جدول ۴، تابع رگرسیونی پیشنهادی در برآورد K_s از دقت متوسطی برای داده‌های آزمون برخوردار است. مقادیر ضریب تبیین (R^2)، ریشه میانگین مربعات خطا (RMSE)، میانگین خطای (ME) و ضریب همبستگی تطابق این تابع به ترتیب $0/51$ ، $0/26 \text{ m/d}$ و $-0/12$ و $0/43$ محاسبه شد. همان‌گونه که در قبل بیان شد مقدار منفی ME نشان‌دهنده تمایل مدل به کم‌برآوردی مقدار K_s است. در مجموع، این شاخص‌ها حاکی از کارایی نسبتاً متوسط مدل در مرحله آزمون می‌باشد.

جدول ۴. ارزیابی تابع رگرسیونی خطی پیشنهادی به منظور برآورد هدایت هیدرولیکی اشباع خاک

آماره‌ها				مدل پیشنهادی
ρ	ME	RMSE	R^2	
$0/31$	$-0/091$	$0/25$	$0/35$	آموزش ($n=66$)
$0/43$	$-0/12$	$0/26$	$0/51$	آزمون ($n=29$)

واحد RMSE و ME همان واحد K_s یعنی (m/d) است.

برآورد هدایت هیدرولیکی اشباع خاک با استفاده از درخت تصمیم (DT)

پس از اجرای الگوریتم درخت تصمیم (DT) بر روی متغیرهای ورودی استخراج شده با استفاده از تحلیل مؤلفه‌های اصلی (PCA) و متغیر هدف هدایت هیدرولیکی اشباع خاک (K_s)، ساختار درختی نسبتاً ساده‌ای تشکیل شد (نمودار مربوطه ارائه نشده است). به‌طور کلی در ساختار درخت تصمیم، انشعابات بالاتر بیان‌گر اهمیت بیشتر متغیرهای ورودی در فرایند مدل‌سازی می‌باشند. بر اساس نتایج، برای متغیر هدف K_s ، پارامتر جرم مخصوص حقیقی خاک (PD) به‌عنوان مؤثرترین پارامتر شناخته شد و در گره ریشه درخت، یعنی در بالاترین سطح طبقه‌بندی قرار گرفت. پس از PD، متغیرهای شاخص رطوبت توپوگرافیک و درصد شن، به‌ترتیب بیشترین تأثیر را بر K_s داشتند. این مشاهدات با نتایج حاصل از رگرسیون خطی چندگانه (SMLR) نیز همخوانی دارد؛

به گونه‌ای که در آن نیز درصد شن و جرم ویژه حقیقی خاک، به عنوان مهمترین پارامترهای خاکی ورودی در برآورد K_s تعیین شدند. شایان ذکر است که فرزاد مهر و همکاران (۱۳۹۷) در مطالعه‌ای مشابه، جرم مخصوص ظاهری خاک را مؤثرترین متغیر در دسته‌بندی K_s با استفاده از درخت تصمیم گزارش کردند.

نتایج ارزیابی الگوریتم طبقه‌بندی درخت تصمیم در برآورد هدایت هیدرولیکی اشباع خاک (K_s) در دو مرحله آموزش و آزمون در جدول ۵ ارائه شده است. طبق این نتایج، ضریب تبیین (R^2) مدل اشتقاق یافته برای داده‌های مورد مطالعه در هر دو مرحله کمتر از ۰/۳۹ است. همچنین مقادیر جذر میانگین مربعات خطا (RMSE) برای مراحل آموزش و آزمون به ترتیب ۰/۲۵ و ۰/۲۶ m/d و میانگین خطا (ME) نیز به ترتیب ۰/۰۹۱- و ۰/۱۲- m/d به دست آمد (جدول ۵). همان گونه که پیش‌تر اشاره شد، مقادیر منفی میانگین خطا نشان دهنده تمایل مدل به کم‌برآوردی مقادیر K_s می‌باشد. ضریب ρ نیز بیان‌گر تطابق ضعیف مدل در هر دو مرحله آموزش و آزمون می‌باشد ($\rho < 0/65$). در مجموع، نتایج بیان‌گر آن است که مدل اشتقاق یافته در برآورد هدایت هیدرولیکی اشباع خاک‌های مورد مطالعه از دقت پیش‌بینی نسبتاً پایین برخوردار است. شایان ذکر است که در استفاده از داده‌های خاکی به تنهایی به عنوان ورودی به مدل DT، دقت پیش‌بینی پایین‌تری به دست آمد (داده‌ها ارائه نشده است). این نتایج با یافته‌های پژوهش فرزاد مهر و همکاران (۱۳۹۷) همخوانی دارد؛ آنان نیز عملکرد نسبتاً پایین الگوریتم درخت تصمیم را در پیش‌بینی هدایت هیدرولیکی اشباع خاک گزارش کرده‌اند. نتایج جدول ۵ همچنین نشان می‌دهد که برآورد K_s به وسیله DT برای داده‌های جدید مرحله آزمون به خوبی مرحله آموزش نبوده و دقت مدل کاهش یافته است. این یافته‌ها با نتایج حاصل از پژوهش مصری و همکاران (۱۳۹۹) نیز سازگار است.

جدول ۵. نتایج حاصل از ارزیابی الگوریتم تصادفی درخت تصمیم در برآورد هدایت هیدرولیکی اشباع خاک

مدل پیشنهادی	R^2	RMSE	ME	ρ	آماره‌ها
CT	آموزش (n= ۶۶)	۰/۳۹	۰/۲۵	-۰/۰۹۱	۰/۵۶
	آزمون (n= ۲۹)	۰/۲۵	۰/۲۷	-۰/۰۴۴	۰/۴۲

واحد RMSE و ME همان واحد K_s یعنی (m/d) است.

برآورد هدایت هیدرولیکی اشباع با استفاده از شبکه عصبی مصنوعی (ANN)

معماری شبکه با کمترین حاشیه خطا شامل دوازده نورون در لایه ورودی و دو لایه پنهان است و هر لایه پنهان نیز دو نورون دارد. لایه‌های پنهان از توابع فعال‌سازی سیگموئیدی (Sigmoid) استفاده می‌کردند که به دلیل خواص غیرخطی و مشتق‌پذیر بودن آن‌ها، برای مدل‌سازی روابط پیچیده بین داده‌ها مناسب است. این در حالی است که لایه خروجی نیز دارای تابع فعال‌سازی خطی (Linear) بود تا دامنه خروجی را برای برآورد هدایت هیدرولیکی اشباع (K_s) محدود نکند. الگوریتم آموزشی مورد استفاده انتشار رو به عقب (Backpropagation) با نسخه «rprop+» بود. این الگوریتم برای مجموعه داده‌هایی با ابعاد متوسط و برای توابع سیگموئیدی، به دلیل تطبیق‌پذیری نرخ یادگیری در هر پارامتر، کارایی بالایی دارد (Igel and Hüsken, 2003). در جدول (۶)، بهترین معماری شبکه عصبی گزینش شده به همراه شاخص‌های آماری مورد استفاده برای تعیین کارایی این شبکه جهت برآورد K_s در دو مرحله آموزش و آزمون ارائه شده است.

جدول ۶. بهترین شبکه عصبی گزینش شده برای برآورد K_s با استفاده از داده‌های زودیاخت خاک و داده‌های کمکی محیطی و نتایج شاخص‌های آماری مورد استفاده برای ارزیابی کارایی این شبکه در دو مرحله آموزش (n= ۶۶) و آزمون (n= ۲۹)

نوع شبکه	معماری شبکه	توابع آستانه	R^2	RMSE	ME	ρ
rprop+	۱-۲-۱۲	Sig-lin	آموزش (n= ۶۶)	۰/۸۰	۰/۰۰۸	$6/21 \times 10^{-5}$
			آزمون (n= ۲۹)	۰/۴۹	۰/۲۵	۰/۷۰

واحد RMSE و ME همان واحد K_s یعنی m/d است.

همان گونه که در این جدول مشاهده می‌شود ضریب تبیین (R^2) مدل شبکه عصبی پی‌ریزی شده برای برآورد K_s در هر دو مرحله آموزش و آزمون، به ترتیب معادل ۰/۸۰ و ۰/۴۹ است. این نتایج بیان‌گر آن است که مدل شبکه عصبی در برآورد هدایت هیدرولیکی اشباع خاک در مرحله آموزش، از دقت پیش‌بینی بالایی برخوردار است. حال آن که کاهش R^2 در مرحله آزمون به حدود

۰/۴۹، نشان‌دهنده توان تعمیم‌پذیری متوسط مدل به داده‌های اندازه‌گیری شده می‌باشد. با توجه به نتایج، مقدار آماره RMSE در مرحله آموزش (معادل ۰/۰۰۸ m/d) به‌طور قابل‌توجهی کمتر از مقدار آن در مرحله آزمون (معادل ۰/۲۵ m/d) می‌باشد. مقادیر میانگین خطای مدل شبکه عصبی پی‌ریزی شده در برآورد هدایت هیدرولیکی اشباع خاک نیز، در مرحله آموزش و آزمون به‌ترتیب معادل $10^{-5} \times 6/21$ و ۰/۰۱ می‌باشد. مقادیر مثبت این آماره بیان‌گر تمایل مدل به اندکی بیش‌برآوردی مقدار K_s در هر دو مرحله آموزش و آزمون است. هرچند شدت این بیش‌برآوردی در مرحله آزمون بیشتر از آموزش است که بر چالش تعمیم‌پذیری مدل تأکید دارد. ضریب همبستگی تطابق (ρ) نیز نشان‌دهنده تطابق قابل‌توجه مدل در مرحله آموزش (معادل ۰/۸۹) و تطابق متوسط آن در مرحله آزمون (معادل ۰/۷۰) دارد (Viscarra Rossel and Hicks, 2015).

در مجموع، با وجود کارایی بالای مدل شبکه عصبی پی‌ریزی شده در مرحله آموزش، عملکرد آن در مرحله آزمون از دقت و اطمینان کمتری برخوردار است. این امر لزوم احتیاط در استفاده از این مدل برای پیش‌بینی K_s در شرایط جدید را متذکر می‌شود. با این حال، این نتایج به وضوح نشان داد که استفاده توأمان از داده‌های خاک و داده‌های عوارض زمین به عنوان ورودی به شبکه عصبی، در مقایسه با حالتی که تنها از داده‌های خاک استفاده می‌شود (داده‌ها ارائه نشده است)، منجر به برآوردهای دقیق‌تری از K_s می‌گردد. این نتایج با یافته‌های (Agyaree et al., 2007) که K_s را با استفاده از داده‌های زودیافت خاک و پارامترهای سرزمین توسط شبکه‌های عصبی با دقت خوبی برآورد کردند، سازگاری دارد.

تعیین متغیرهای مؤثر ورودی به شبکه‌های عصبی مصنوعی

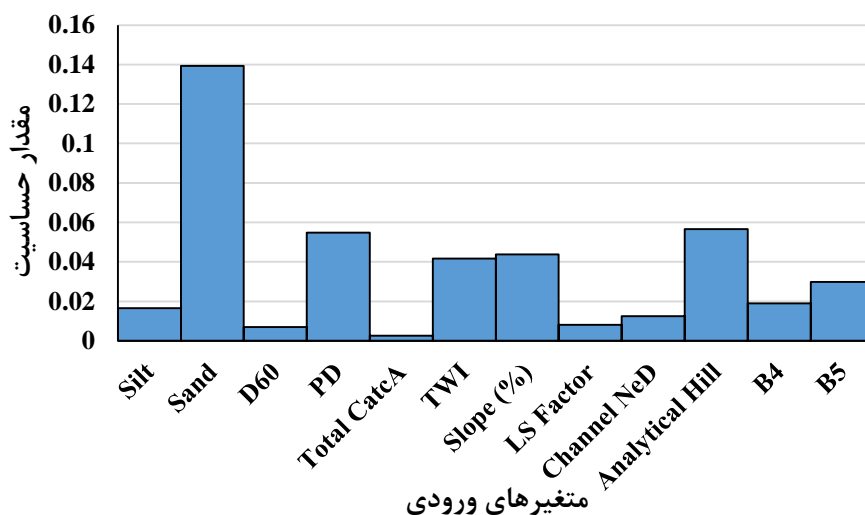
آنالیز حساسیت به‌عنوان یک روش آماری، نقش متغیرهای مستقل در پیش‌بینی متغیر وابسته را کمی می‌کند. این تحلیل مشخص می‌کند که همبستگی متغیر وابسته با متغیرهای مستقل به علت اثر مستقیم، یا غیرمستقیم آن‌ها بر K_s از طریق ویژگی‌های دیگر است. اگر همبستگی بین متغیر وابسته و یک متغیر مستقل به علت اثر مستقیم متغیر مستقل باشد، نشان‌گر وجود یک رابطه علی واقعی بین آن‌ها می‌باشد. اما اگر این وابستگی اصولاً به‌علت اثر غیرمستقیم متغیر مستقل از طریق متغیرهای همراه باشد، در آن صورت بایستی به متغیرهای همراه که موجب اثر غیرمستقیم شده‌اند، توجه کرد.

در این پژوهش به‌منظور تعیین مؤثرترین متغیرهای مستقل در تخمین K_s با استفاده از شبکه عصبی مصنوعی، از آنالیز حساسیت بهره گرفته شد. فرآیند مذکور با استفاده از کتابخانه NeuralSens در محیط نرم افزار R 4.5.1 صورت پذیرفت. این کتابخانه با محاسبه مشتق‌های جزئی خروجی شبکه نسبت به ورودی‌ها، اطلاعاتی ارزشمند راجع به میزان حساسیت مدل به متغیرهای ورودی در اختیار طراح و معمار مدل قرار می‌دهد. با شناسایی میزان تأثیر متغیرهای ورودی بر دقت پیش‌بینی مدل، می‌توان متغیرهای کم‌اثر را از شبکه حذف و مدلی ساده‌تر، کارآمدتر و با خطر کمتر برازش بیش‌ازحد (Overfitting) بسط و توسعه داد.

برای انجام آنالیز حساسیت معمولاً بدین صورت اقدام می‌شود که مقادیر ضریب حساسیت متغیرهای ورودی از تقسیم نمودن خطای کل شبکه در غیاب یک متغیر ورودی به خطای کل شبکه بدست می‌آید. براساس این روش، هرچه مقدار ضریب حساسیت یک متغیر بیشتر باشد، آن متغیر در تبیین متغیر وابسته دارای سهم بیشتری است (نورزویان عزیز و همکاران، ۱۳۹۶).

نتایج آنالیز حساسیت شبکه عصبی مصنوعی پی‌ریزی شده به متغیرهای ورودی خاک و عوارض زمین در شکل (۴) ارائه شده است. با توجه به شکل (۴)، درصد شن و جرم ویژه حقیقی، مهم‌ترین پارامترهای ورودی خاک به شبکه عصبی برای برآورد هدایت هیدرولیکی اشباع (K_s) در منطقه مورد مطالعه هستند. شایان ذکر است که هر دو این دو متغیرها، پارامترهای درگیر در برآورد K_s در مدل‌های رگرسیون خطی و درخت تصمیم بودند. از نظر فیزیکی، رابطه‌ای ذاتی بین مقدار شن و هدایت هیدرولیکی اشباع خاک وجود دارد؛ به‌این صورت که برای یک خاک معین، با افزایش مقدار شن، هدایت هیدرولیکی اشباع خاک نیز افزایش می‌یابد. پارامتر Pd که نمایه‌ی بیان‌کننده جرم ویژه حقیقی خاک است، نشان‌دهنده وضعیت کانی‌شناسی خاک می‌باشد. Pd به تنهایی و به‌طور مستقیم بر جریان آب در خاک تأثیر نمی‌گذارد، زیرا این پارامتر صرفاً نشان‌دهنده چگالی ذرات جامد خاک است. با این حال، تأثیر غیرمستقیم این پارامتر از طریق ارتباط آن با سایر ویژگی‌های مؤثر بر هدایت هیدرولیکی اشباع خاک - همچون تخلخل، توزیع اندازه ذرات خاک و ساختمان خاک - قابل‌توجه است.

در بین پارامترهای محیطی کمکی، شبکه مصنوعی پی‌ریزی شده بیشترین حساسیت را به پارامترهای تحلیل سایه روشن، شیب و شاخص رطوبت توپوگرافیک دارد. پارامترهای سایه‌اندازی با تأثیر بر مقدار تشعشع دریافتی از خورشید، به‌طور غیرمستقیم بر میزان رطوبت خاک و همچنین دمای خاک اثر می‌گذارد. شیب زمین نیز با کنترل کردن سرعت و حجم رواناب‌های سطحی و جریان‌های زیرسطحی، بر مقدار رطوبت خاک و به‌تبع آن قابلیت‌گذردهی آب در خاک تأثیر مستقیم دارد. به‌طوری‌که با افزایش شیب، پتانسیل نفوذ آب کاهش و رواناب سطحی افزایش می‌یابد. شاخص رطوبت نیز بیان‌گر فرآیندهای تجمع و تمرکز آب در سطح زمین‌نما بوده و به‌عنوان شاخصی از توزیع مکانی رطوبت خاک است.



شکل ۴. آنالیز حساسیت متغیرهای ورودی خاکی و عوارض زمین به شبکه عصبی مصنوعی (Silt): سیلت، Sand: شن، d60: قطری که ۶۰ درصد ذرات به‌صورت وزنی از آن کوچک‌تر است، PD: جرم مخصوص حقیقی خاک، Total CatchA: سطح کل حوضه آبخیز، TWI: شاخص رطوبت توپوگرافی، Slope (%): شیب (درصد)، LS Factor: فاکتور LS، Channel NeD: فاصله از شبکه آبره‌ها، Analytical Hill: تحلیل سایه روشن، B4 و B5: باندهای ۴ و ۵ تصویر ماهواره (LandSat ETM+).

مقایسه نتایج حاصل از روش‌های رگرسیونی خطی چندگانه (SMLR)، درخت تصمیم (DT) و شبکه‌های عصبی مصنوعی (ANN) پی‌ریزی شده و انتخاب بهترین مدل

نتایج ارزیابی مدل‌ها نشان می‌دهد هرچه ضریب تبیین (R^2) و ضریب همبستگی تطابق (ρ) به یک نزدیک‌تر و مقادیر خطای میانگین (ME) و ریشه میانگین مربعات خطا (RMSE) به صفر نزدیک‌تر باشند، دقت و کارایی مدل در برآورد هدایت هیدرولیکی اشباع خاک (K_s) بیشتر است. بر این اساس، مدل شبکه عصبی مصنوعی پیشنهادی (ANN)، با دارا بودن مقدار R^2 و ρ بالاتر و مقادیر ME و RMSE پایین‌تر، در مقایسه با مدل‌های رگرسیونی خطی چندگانه (MLR) و درخت تصمیم (DT)، از عملکردی بهتر و دقتی بالاتر در برآورد K_s خاک‌های منطقه مورد مطالعه برخوردار است (جدول ۴ تا ۶).

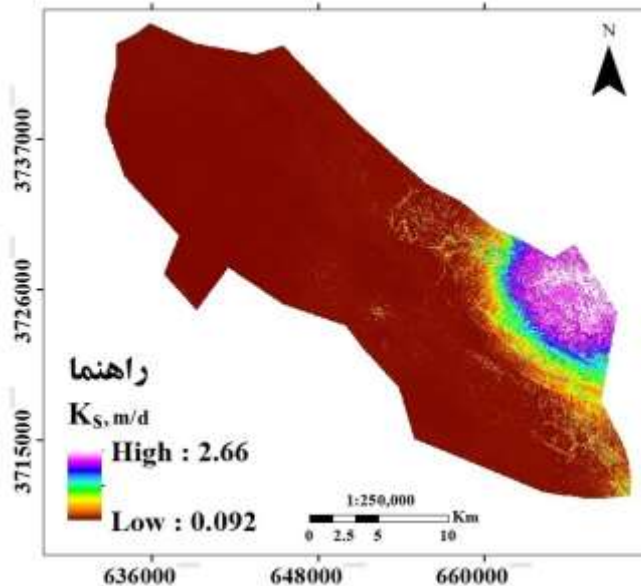
برتری عملکرد ANN را شاید بتوان به عوامل زیر نسبت داد. نخست، در این پژوهش داده‌های مربوط به عوارض زمین در کنار ویژگی‌های زودیافت خاک به‌عنوان متغیرهای ورودی استفاده شده‌اند که این خود موجب ارائه تو صیفی جامع‌تر از فضای مؤثر بر جریان آب در خاک می‌شود. از آنجا که K_s تابعی از مجموعه عوامل فیزیکی خاک (همچون بافت، تخلخل و ساختمان خاک) و همچنین عوامل توپوگرافی و فرایندهای ژئومورفولوژیک مؤثر بر توزیع رطوبت و رسوب‌گذاری است، ترکیب این دو گروه از داده‌ها می‌تواند اطلاعات غنی‌تری از محیط هیدرولوژیک خاک فراهم کند. عامل مهم دیگر به توانایی شبکه‌های عصبی مصنوعی در مدل‌سازی روابط غیرخطی و پیچیده میان متغیرها مربوط می‌شود. هدایت هیدرولیکی اشباع ماهیتی ناهمگن و غیرخطی دارد و تحت تأثیر برهم‌کنش‌های چندگانه میان ویژگی‌های خاک و عوامل محیطی قرار می‌گیرد. شبکه‌های عصبی مصنوعی به‌دلیل ساختار لایه‌ای و توابع فعال‌ساز غیرخطی، قادرند الگوهای پنهان و روابط پیچیده بین متغیرهای مستقل و وابسته را بدون نیاز به در نظر گرفتن شکل تابعی مشخص، استخراج کنند. در مقابل، مدل رگرسیونی خطی چندگانه به روابط خطی محدود است و درخت تصمیم نیز اگرچه قادر به مدل‌سازی برخی روابط غیرخطی است، اما به دلیل ساختار گسسته و قطعه‌ای خود ممکن است نتواند پیوستگی و پیچیدگی کامل روابط موجود در این داده‌ها را کمی کند. نتایج به‌دست‌آمده در این

پژوهش با یافته‌های (Agyaree et al. 2007) همخوانی دارد؛ به گونه‌ای که آنان نیز گزارش کردند شبکه‌های عصبی مصنوعی با استفاده از داده‌های زودیافت خاک و پارامترهای سرزمین‌قارند K_s را با دقت بیشتری نسبت به روش‌های رگرسیون برآورد کنند. همچنین (Motaghian and Mohammadi 2011) نشان دادند که ادغام داده‌های عوارض زمین با داده‌های خاکی می‌تواند دقت پیش‌بینی K_s را به‌طور معنی‌داری افزایش دهد.

در مجموع می‌توان این‌گونه بیان کرد که استفاده از داده‌های عوارض زمین و سنجش از دور به‌عنوان ورودی مکمل ویژگی‌های زودیافت خاک در چارچوب معماری شبکه عصبی مصنوعی، راهبردی مؤثر برای افزایش دقت برآورد K_s در منطقه مورد مطالعه است. با در نظر گرفتن عملکرد مناسب مدل ANN و پتانسیل بالای آن در استخراج الگوهای پیچیده از داده‌ها، توسعه این مدل‌ها با به‌کارگیری الگوریتم‌های بهینه‌سازی پیشرفته‌تر (مانند الگوریتم‌های فراابتکاری یا شبکه‌های عصبی عمیق) و افزودن متغیرهای محیطی جدید (مانند شاخص‌های سنجش از دور یا داده‌های اقلیمی)، می‌تواند به عنوان رویکردی امیدبخش برای پژوهش‌های آتی این حوزه مورد توجه قرار گیرد.

تغییرپذیری مکانی هدایت هیدرولیکی اشباع خاک

در این پژوهش برای تهیه نقشه هدایت هیدرولیکی اشباع خاک (K_s) در حوضه آبخیز مورد مطالعه، از مدل شبکه عصبی مصنوعی (ANN) پی‌ریزی شده استفاده شد. داده‌های ورودی مدل از نقشه‌های ویژگی‌های مبنایی خاک و همچنین نقشه‌های عوارض زمین و باندهای سنجش از دور استخراج شدند (شکل ۲). نقشه تغییرپذیری مکانی سیلت، شن، d_{60} و جرم ویژه حقیقی خاک با استفاده از داده‌های اندازه‌گیری شده و به روش کریجینگ معمولی در محیط نرم‌افزار ArcGIS 10.8 با وضوح ۱۲/۵ متر تهیه شد. همچنین، نقشه‌های سطح حوضه آبخیز، شاخص رطوبت توپوگرافی، درصد شیب، فاکتور LS، فاصله از شبکه آبراهه‌ها و تحلیل سایه روشن در محیط SAGA GIS 9.3.2 با وضوح ۱۲/۵ متر تولید شدند. افزون بر این، باندهای ۴ و ۵ (B4 و B5) تصاویر ماهواره‌ای LandSat ETM+ با وضوح ۳۰ متر برش داده شده و مورد استفاده قرار گرفتند. اندازه پیکسلی ۱۲/۵ متر برای غلبه بر عدم تطابق وضوح بین ورودی‌های مختلف مورد استفاده انتخاب شد. نقشه تغییرپذیری مکانی K_s پیش‌بینی شده در شکل (۵) ارائه شده است. همان‌طور که در این شکل مشاهده می‌شود، بخش‌های شرقی و جنوب شرقی منطقه مورد مطالعه دارای بیشترین مقدار هدایت هیدرولیکی اشباع و بخش‌های غربی و شمال غربی دارای کمترین مقادیر آن هستند. این الگوی توزیع مکانی، همخوانی و مطابقت بالایی با تغییرپذیری مکانی پارامترهای خاکی مورد استفاده در مدل دارد. این نتایج خاکی از آن است که تأثیر پارامترهای خاکی بر هدایت هیدرولیکی اشباع، بسیار بیشتر از داده‌های عوارض زمین است. واقعیت این است که تأثیر ویژگی‌های ژئومورفولوژیک بر K_s غیرمستقیم است و بخش عمده‌ای از این تغییرپذیری که توسط این ویژگی‌ها تبیین می‌شود، به‌طور همزمان با ویژگی‌های خاک نیز قابل توجیه است. افزون بر این، در وضوح مکانی و میزان دقت این دو گروه از داده‌ها نیز، تفاوت‌هایی وجود دارد. ویژگی‌های خاک اغلب روی همان نمونه‌ای که برای تعیین هدایت هیدرولیکی خاک استفاده شده‌اند، اندازه‌گیری شده‌اند. این در حالی است که داده‌های عوارض زمین از یک مدل رقومی ارتفاع (DEM) 12.5×12.5 متری استخراج شده‌اند که مقادیر ارتفاع در نقاط شبکه‌ای آن تحت تأثیر خطاهای ناشی از عملیات ریاضی (همچون مشتق‌گیری و درون‌یابی) و نیز خطاهای مکان‌یابی نقاط نمونه‌برداری قرار داشته است.



شکل ۵. توزیع مکانی هدایت هیدرولیکی اشباع خاک (K_s) با استفاده از شبکه عصبی مصنوعی

نتیجه گیری

در این پژوهش، سه مدل رگرسیونی خطی چندگانه (MLR)، درخت تصمیم (DT) و شبکه عصبی مصنوعی (ANN) با بهره‌گیری از پارامترهای زود یافت خاک و متغیرهای کمکی محیطی (شامل داده‌های عوارض زمین و داده‌های سنجش از دور) برای پیش‌بینی هدایت هیدرولیکی اشباع خاک (K_s) مورد استفاده قرار گرفتند. همچنین، به منظور کاهش ابعاد داده‌ها و حذف متغیرهای کم‌اهمیت، از تحلیل مؤلفه‌های اصلی (PCA) استفاده شد. نتایج نشان داد که مدل‌های مبتنی بر MLR و DT از دقت پیش‌بینی نسبتاً پایینی در برآورد K_s برخوردار می‌باشند. افزون بر این، بکارگیری داده‌های کمکی محیطی در کنار داده‌های زود یافت خاک برای اشتقاق توابعی نوین مبتنی بر رگرسیون خطی، تأثیر معنی‌داری در بهبود دقت برآورد K_s نداشت. در مدل رگرسیونی خطی چندگانه، پارامترهای جرم مخصوص حقیقی خاک و درصد شن به عنوان مؤثرترین پارامترهای ورودی شناخته شدند. در اجرای الگوریتم درخت تصمیم نیز پارامترهای جرم مخصوص حقیقی خاک، شاخص رطوبت توپوگرافیک و درصد شن، به ترتیب بیشترین تأثیر را در برآورد K_s داشتند. در مقابل، نتایج نشان داد که استفاده از شبکه‌های عصبی مصنوعی همراه با تلفیق داده‌های خاک و متغیرهای کمکی محیطی می‌تواند منجر به دستیابی به برآوردهایی دقیق از K_s در خاک‌های منطقه مورد مطالعه گردد. این موضوع بیان‌گر آن است که استفاده همزمان از داده‌های زود یافت خاک و متغیرهای کمکی محیطی می‌تواند کارایی شبکه عصبی مصنوعی را در پیش‌بینی K_s بهبود دهد. براساس نتایج تحلیل اهمیت متغیرها، از میان ویژگی‌های خاک، درصد شن و جرم مخصوص حقیقی و از میان متغیرهای کمکی محیطی، تحلیل سایه روشن، شیب و شاخص رطوبت توپوگرافیک به عنوان مؤثرترین پارامترهای ورودی به مدل شبکه عصبی مصنوعی شناسایی شدند. در مجموع، نتایج این پژوهش نشان داد مدل مبتنی بر شبکه‌های عصبی چندلایه پرسپترون (MLP) در مقایسه با مدل‌های رگرسیونی خطی چندگانه و درخت تصمیم، از کارایی بالاتری در برآورد K_s برخوردار است. همچنین نقشه تهیه شده از هدایت هیدرولیکی اشباع خاک (K_s) در منطقه مورد مطالعه از دقت و قابلیت مناسبی برای کاربردهای مدیریتی و پژوهشی برخوردار است.

منابع

حکمت‌زاد، علی، داوری، مسعود، محمودی، محمدعلی، نبی‌اللهی، کمال. (۱۳۹۷). تخمین هدایت هیدرولیکی اشباع در برخی از خاک‌های استان ایلام با استفاده از شبکه‌های عصبی مصنوعی و روش‌های رگرسیونی، نشریه حفاظت منابع آب و خاک، ۷(۳): ۱۳۱-۱۴۹.

دهقانی بانیانی، سیروس، قربانی دشتکی، شجاع، محمدی، جهانگرد، خداوردی‌لو حبیب. (۱۳۹۰). مقایسه کارایی رگرسیون خطی چندگانه و رگرسیون درختی در برآورد هدایت آبی اشباع و پارامتر عکس طول درشت موئینگی خاک، مجله پژوهش آب ایران، ۹(۹): ۱۹۳-۲۰۴.

فرزادمهر، مهنوش، دستورانی، مهدی و خاشعی سیوکی، عباس. (۱۳۹۷). مقایسه مدل‌های درخت تصمیم و یادگیری بر پایه نمونه در برآورد هدایت هیدرولیکی اشباع خاک. نشریه پژوهش‌های حفاظت خاک، ۲۵(۵): ۱۶۷-۱۸۴.

مصری، سمیرا، قربانی دشتکی، شجاع، شیرانی، حسین، کامکار روحانی، ابوالقاسم و متقیان، حمید رضا. (۱۳۹۹). تفکیک ویژگی‌های مؤثر بر منحنی مشخصه رطوبتی با استفاده از درخت تصمیم‌گیری. نشریه مدیریت خاک و تولید پایدار، ۱۰(۳): ۱۳۵-۱۵۰.

نوروزیان عزیز، زهرا، قاجار سپانلو، مهدی، عمادی، سید مصطفی و صادق زاده، فردین. (۱۳۹۶). ارزیابی مدل‌های رگرسیونی و شبکه عصبی مصنوعی در تخمین هدایت هیدرولیکی اشباع خاک در مازندران، پژوهش‌های خاک، ۳۱(۱): ۷۵-۸۷.

References

- Agyare, W., Park, S., & Vlek, P. (2007). Artificial neural network estimation of saturated hydraulic conductivity. *Vadose Zone Journal*, 6, 423-431. <https://doi.org/10.2136/vzj2006.0131>.
- Alletto, L., & Coquet, Y. (2009). Temporal and spatial variability of soil bulk density and near-saturated hydraulic conductivity under two contrasted tillage management systems. *Geoderma*, 152:85-94. <https://doi.org/10.1016/j.geoderma.2009.05.023>
- Cheng, Q., Xu, Q., Cheng, X., Li, C., & Wang, X. (2019). In-situ estimation of unsaturated hydraulic conductivity in freezing soil using improved field data and inverse numerical modeling. *Agricultural and Forest Meteorology*, 279, 107746. <https://doi.org/10.1016/j.agrformet.2019.107746>.
- Dane, J.H. & Topp, C.G., (eds.). 2002. *Methods of Soil Analysis: Part 4 Physical Methods* Madison, WI: Soil Science Society of America, Soil Science Society of America Book Series Number 5, ISBN 0-89118-810-X, 1692 p.
- Dehghani, B.S., Ghorbani Dashtaki, Sh., Mohammadi, J. & Khodaverdilou, H. (2012). Comparing the performance of multiple linear regression and regression tree to predict saturated hydraulic conductivity and the inverse of macroscopic capillary length. *Iranian Water Research Journal*, (2): 193-203. (In Persian)
- Farzadmehr, M., Dastourani, M., & Khashei-Siuki, A. (2019). Comparing decision tree and instance-based learning models to estimate soil saturated hydraulic conductivity. *Journal of Water and Soil Conservation*, 25(5), 167-184. <http://doi.org/10.22069/jwsc.2018.12334.2691>. (In Persian)
- Fashi, F.H., Gorji, M., & Shorafa, M. (2016) Estimation of soil hydraulic parameters for different land-uses. *Modeling Earth Systems and Environment*, 2, 1-7. <https://doi.org/10.1007/s40808-016-0229-0>.
- Fitts, C.R. (2002). *Groundwater science*. Academic Press, New York.
- Gabriel, J.L., Quemada, M., Martín-Lammerding, D., & Vanclouster, M. (2019). Assessing the cover crop effect on soil hydraulic properties by inverse modelling in a 10-year field trial. *Agricultural Water Management*, 222:62-71. <https://doi.org/10.1016/j.agwat.2019.05.034>
- Gumiere, S. J., Lafond, J. A., Hallema, D. W., Périard, Y., & Caron, J. (2014). Mapping soil hydraulic conductivity and matric potential for water management of cranberry: Characterization and spatial interpolation methods. *Biosystems Engineering*, 128, 29-40. <https://doi.org/10.1016/j.biosystemseng.2014.09.002>
- Hekmatzad, A., Davari, M., Mahmoudi, M.A., and Nabiollahi, K. (2018). Estimation of saturated hydraulic conductivity in some soils of Ilam province using artificial neural networks and regression methods. *Journal of Water and Soil Resources Conservation*, 7(3), 131-149. (in Persian)
- Igel, C., & Hüskens, M. (2005). Empirical evaluation of the improved Rprop learning algorithms. *Neurocomputing*, 50, 105-123. [https://doi.org/10.1016/S0925-2312\(01\)00700-7](https://doi.org/10.1016/S0925-2312(01)00700-7).
- Jarvis, N., Koestel, J., Messing, I., Moeys, J., & Lindahl, A. (2013). Influence of soil, land use and climatic factors on the hydraulic conductivity of soil. *Hydrology and Earth System Sciences*, 17, 5185-5195. <https://doi.org/10.5194/hess-17-5185-2013>.
- Khaledian, Y., Kiani, E., Ebrahimi, S., Brevik, E.C., & Aitkenhead-Peterson, J. (2017). Assessment and monitoring of soil degradation during land use change using multivariate analysis. *Land Degradation and Development*, 28, 128-141. <https://doi.org/10.1002/ldr.2541>.
- Khodaverdilou, H., Cheraghabdol, H.K., Bagarello, V., Iovino, M., Asgarzadeh, H., Ghorbani Dashtaki, Sh. (2017). Ring diameter effects on determination of field-saturated hydraulic conductivity of different loam soils. *Geoderma*, 303, 60-69. <https://doi.org/10.1016/j.geoderma.2017.04.031>.
- Kirkham, M. B. (2005). *Principles of Soil and Plant Water Relations*. Elsevier Academic Press: Burlington, MA. pp. 145 - 172.
- Leij, F. J., Romano, N., Palladino, M., Schaap, M. G., & Coppola, A. (2004). Topographical attributes to predict soil hydraulic properties along a hillslope transect. *Water Resources Research*, 40, W02407. <https://doi.org/10.1029/2002WR001641>
- McBratney, A. B., Odeh, I. O. A., Bishop, T. F. A., Dunbar, M. S., & Shatar, T. M. 2000. An overview of pedometric techniques for use in soil survey. *Geoderma*, 97, 293-327. [https://doi.org/10.1016/S0016-7061\(00\)00043-4](https://doi.org/10.1016/S0016-7061(00)00043-4)
- McKenzie N.J., & Jacquier D.W. (1997). Improving the Field Estimation of Saturated Hydraulic Conductivity in Soil Survey. *Australian Journal of Soil Research*, 35, 803-825. <https://doi.org/10.1071/S96093>
- Mesri, S., Ghorbani Dashtaki, Sh., Shirani, H., Kamkar Rohani, A., & Motaghian, H.R. (2020). Separating effective soil properties on moisture characteristic curve using Decision Tree. *Journal of Soil Management and Sustainable Production*, 10(3), 135-150. <http://doi.org/10.22069/ejsms.2021.17673.1928>. (In Persian).

- Moncada, M.P., Gabriels, D., and Cornelis, W.M. (2014). Data-driven analysis of soil quality indicators using limited data. *Geoderma*, 235, 271-278. <https://doi.org/10.1016/j.geoderma.2014.07.014>.
- Moore, I.D., Gessler, P.E., Nielsen, G.A.E., & Peterson, G.A. (1993) Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal*, 57, 443–452. <https://doi.org/10.2136/sssaj1993.03615995005700020026x>
- Mosleh, Z., Salehi, M.H., Jafari, A., and Esfandiarpour Borujeni, I., & Mehnatkesh, A. (2016). The effectiveness of digital soil mapping to predict soil properties over low-relief areas. *Environmental Monitoring and Assessment*, 188, 195. <https://doi.org/10.1007/s10661-016-5204-8>.
- Motaghian, H.R., & Mohammadi, J. (2011) Spatial estimation of saturated hydraulic conductivity from terrain attributes using regression, kriging, and artificial neural networks. *Pedosphere*, 21, 170–177. [https://doi.org/10.1016/S1002-0160\(11\)60115-X](https://doi.org/10.1016/S1002-0160(11)60115-X).
- Mustafa, S. M. T., Hasan, M. M., Saha, A. K., Ahmed, S., Kabir, M. R., & Rahman, M. M. (2019). Multi-model approach to quantify groundwater-level prediction uncertainty using an ensemble of global climate models and multiple abstraction scenarios. *Hydrology and Earth System Sciences*, 23, 2279–2303. <https://doi.org/10.5194/hess-23-2279-2019>
- Neuzil, C.E. (1994) How permeable are clays and shales? *Water Resources Research*, 30, 145–150. <https://doi.org/10.1029/93WR02930>.
- Norouzian Azizi, Z., Ghajar Sepanlou, M., Emadi, S.M., & Sadeqzade, F. (2017). Evaluation of regression and Artificial Neural Network models to estimate the saturated hydraulic conductivity in Mazandaran Province. *Journal of Soil Research*, 31(1), 75-87. <http://doi.org/10.22092/ijsr.2017.110346>. (In Persian)
- Odeh, I. O. A., & McBratney, A. B. (2000). Using AVHRR images for spatial prediction of clay content in the lower Namoi Valley of eastern Australia. *Geoderma*, 97, 237–254. [https://doi.org/10.1016/S0016-7061\(00\)00041-0](https://doi.org/10.1016/S0016-7061(00)00041-0).
- Pan, Y., Wang, D., Lyu, Sh., Li, X., Li, S., & Su, Y. (2026). Evaluating saturated hydraulic conductivity schemes: impacts on soil moisture simulations and soil texture-dependent applicability. *Journal of Hydrology*, 671, 135129. <https://doi.org/10.1016/j.jhydrol.2026.135129>.
- Park, S.J., & Vlek, P.L.G. 2002. Environmental correlation of three-dimensional spatial soil variability: A comparison of three adaptive techniques. *Geoderma*, 109, 117–140. [https://doi.org/10.1016/S0016-7061\(02\)00146-5](https://doi.org/10.1016/S0016-7061(02)00146-5).
- Qaderi, K., Jalali, V., Etmnan, S., Masoumi Shahr-babak, M. & Homaei, M. (2018). Estimating soil hydraulic conductivity using different data-driven models of ANN, GMDH and GMDH-HS. *Paddy and Water Environment*, 16, 823–833. <https://doi.org/10.1007/s10333-018-0672-9>.
- Rezaei, M., Mousavi, S. R., Rahmani, A., Zeraatpisheh, M., Rahmati, M., Pakparvar, M. & Cornelis, W. (2023). Incorporating machine learning models and remote sensing to assess the spatial distribution of saturated hydraulic conductivity in a light-textured soil. *Computers and Electronics in Agriculture*, 209, 107821. <https://doi.org/10.1016/j.compag.2023.107821>.
- Romano, N., & Palladino, P. (2002). Prediction of soil water using soil physical data and terrain attributes. *Journal of Hydrology*, 265, 36–75. [https://doi.org/10.1016/S0022-1694\(02\)00094-X](https://doi.org/10.1016/S0022-1694(02)00094-X).
- Russo, S.L., Taddia, G., & Abdin, E.C. (2018) Modeling the effects of the variability of temperature-related dynamic viscosity on the thermal affected zone of groundwater heat-pump systems. *Hydrogeology Journal*, 26, 1239–1247. <https://doi.org/10.1007/s10040-017-1714-x>.
- Saha, A. K., Hasan, M. M., Khan, M. R. A., & Al-Amin, M. (2021). Prediction of spatial saturated hydraulic conductivity at the upper soil layer using soil class and terrain attributes. *Modeling Earth Systems and Environment*. <https://doi.org/10.1007/s40808-021-01317-y>
- Schaap, M.G., Leij, F.J., & van Genuchten, M.Th. (1999). A bootstrap-neural networks approach to predict soil hydraulic parameters. p. 1237–1250. In M.Th. van Genuchten et al. (ed.) Proc. Int. Workshop on characterization and measurement of the hydraulic properties of unsaturated porous media. Univ. of California, Riverside.
- Shahabi, A., Nabiollahi, K., Davari, M., Zeraatpisheh, M., Heung, B., Scholten, T. and Taghizadeh-Mehrjardi, R. (2022). Spatial prediction of soil properties through hybridized random forest model and combination of reflectance spectroscopy and environmental covariates. *Gecarto International*, 37(27), 18172–18195. <https://doi.org/10.1080/10106049.2022.2138565>.
- Shirazi, M.A., & Boersma, L. (1984). A unifying quantitative analysis of soil texture. *Soil Science Society of America Journal*, 48, 142-147. <https://doi.org/10.2136/sssaj1984.03615995004800010026x>
- Sihag, P. (2018). Prediction of unsaturated hydraulic conductivity using fuzzy logic and artificial neural network. *Modeling Earth Systems and Environment*, 4, 189–198. <https://doi.org/10.1007/s40808-018-0434-0>
- Souza, C.M.P.D., Thomazini, A., Schaefer, C.E.G.R., Veloso, G.V., Moreira, G.M., & Fernandes Filho, E.I. (2018) Multivariate analysis and machine learning in properties of ultisols (Argissolos) of Brazilian Amazon. *Revista Brasileira de Ciência do Solo*, 42, 1–20. <https://doi.org/10.1590/18069657rbc20170419>

- Tian, Z., Kool, D., Ren, T., Horton, R., & Heitman, J.L. (2018). Determining in-situ unsaturated soil hydraulic conductivity at a fine depth scale with heat pulse and water potential sensors. *Journal of Hydrology*, 564, 802–810. <https://doi.org/10.1016/j.jhydrol.2018.07.052>
- Wilson, J.P., & J.C. Gallant. (2000). Digital terrain Analysis. p. 1–27. In J.P. Wilson and J.C. Gallant (ed.) *Terrain analysis: Principles and application*. John Wiley & Sons, New York.
- Yang, T., Jia, X., Niu, L., Huang, L., Zhao, Ch., Li, X., Ren, X., Wang, R., & Shao, M. (2026). High-resolution mapping of saturated soil hydraulic conductivity across China's drylands. *International Journal of Applied Earth Observation and Geoinformation*, 148, 105176. <https://doi.org/10.1016/j.jag.2026.105176>.
- Zeraatpisheh, M., Ayoubi, S., Sulieman, M., & Rodrigo-Comino, J. (2019). Determining the spatial distribution of soil properties using the environmental covariates and multivariate statistical analysis: a case study in semi-arid regions of Iran. *Journal of Arid Land*, 11(4), 551–566. <https://doi.org/10.1007/s40333-019-0059-9>

Spatial Prediction of Soil Saturated Hydraulic Conductivity by Integrating Soil Properties and Environmental Covariates

Introduction:

Soil and water are two vital natural resources on which human life strongly depends. Soil hydraulic properties, particularly saturated hydraulic conductivity (K_s), play a crucial role in addressing problems related to soil and water management in agricultural, ecological, and environmental systems. K_s directly affects key hydrological processes such as surface runoff, soil erosion, and deep percolation. However, soil and biotic factors influencing K_s vary greatly in both space and time. Hence, it is also expected that K_s will vary in both space and time. Moreover, K_s is one of the most difficult, costly, and time-consuming soil hydraulic properties to measure directly; Therefore, indirect estimation methods are commonly employed.

Terrains play a fundamental role in modulating the Earth's surface and atmospheric processes. Landform features generally control the movement of water and materials across the landscape, thereby influencing watershed hydrology at the topographic scale. Using terrain attributes to model K_s may serve as a suitable alternative, as terrain data are relatively easy to obtain compared to intensive soil sampling. In previous soil and landscape studies, the relationship between terrain attributes, remote sensing data, and K_s has not been thoroughly investigated. The main question is whether integrating these data with soil properties can enhance the accuracy of K_s estimation. Accordingly, this study evaluated the efficiency of multiple linear regression (MLR), decision tree (DT), and artificial neural network (ANN) methods for estimating and mapping the spatial distribution of K_s in parts of the Chardawel and Chamshir sub-basins (Ilam Province, Iran), using readily measurable soil properties along with environmental covariates (terrain attributes and remote sensing data).

Materials and Methods

The study area extends from 46° 25' 53.08" to 46° 50' 13.52" E longitude and from 33° 31' 23.11" to 33° 50' 25.59" N latitude, located in the Chardawel and Chamshir sub-basins, Ilam Province, Iran. A total of 95 sampling sites were selected using a stratified random sampling approach, ensuring representation of land capability classes, land use categories, and geological and topographical features. At each site, K_s was measured using a Guelph permeameter. Disturbed and undisturbed soil samples to a depth of 30 cm were collected from the same sites for laboratory analysis. All samples were air-dried, gently crushed, and sieved through a 2-mm

opening. The samples were analyzed for particle size distribution (PSD), particle density (D_p), soil moisture at field capacity (FC), total porosity, effective porosity, and soil organic carbon (SOC). The geometric mean diameter (d_g) and geometric standard deviation (σ_g) of soil particles were calculated based on the method proposed by Shirazi and Boersma (1984). Undisturbed soil samples were analyzed for bulk density (D_b) using the cylindrical core method.

A set of 16 terrain features was extracted from a Digital Elevation Model (DEM) with 12.5×12.5 m spatial resolution (ALOS PALSAR database) using SAGA GIS software (Version 9.3.2). Additionally, Landsat 8 satellite imagery (OLI/TIRS sensors) with a 30×30 m spatial resolution (sourced from the USGS) was employed to extract remote sensing variables. Principal Component Analysis (PCA) was performed using the XLSTAT toolbox in Microsoft Excel 2019 to identify and select the most influential variables. To estimate K_s values, MLR, DT, and ANN models were developed using R software (version 4.5.1). The validity of the derived models was assessed using the coefficient of determination (R^2), root mean square error (RMSE), and mean error (ME).

Results and Discussion

Based on the PCA results, among the 40 soil properties and environmental covariates used for spatial prediction of K_s , 12 variables including percentage slope, topographic wetness index (TWI), LS factor, analytical hillshade, total catchment area, and channel network distance were selected. From the remote sensing indices and reflectance values, bands 4 and 5 of Landsat 8 OLI/TIRS (as the biophysical variables), along with silt percentage, sand percentage, d_{60} , and soil particle density, were identified as the most effective covariates.

The results demonstrated that the ANN model, with a greater coefficient of determination ($R^2_{\text{train}} = 0.80$, $R^2_{\text{test}} = 0.49$) and the lower root mean square error ($\text{RMSE}_{\text{train}} = 0.008$ m/d, $\text{RMSE}_{\text{test}} = 0.25$ m/d), outperformed both the MLR and DT models in estimating K_s . While the MLR and DT models tended to underestimate K_s , the ANN model produced more accurate and reliable predictions. Overall, the ANN model yielded reasonable estimates of K_s for the study area's soils.

A spatial variability map of K_s was also successfully generated by integrating soil properties with environmental covariates, suggesting the usefulness of this approach for developing agro-hydrological models in data-limited regions.

Sensitivity analysis of the ANN model highlighted the importance of two soil properties, sand percentage and soil particle density, along with slope percentage, analytical hillshade, and topographic wetness index terrain variables in the spatial modeling and distribution of K_s in the studied lands.

Conclusion

Terrain and remote sensing data, along with soil properties, were integrated using stepwise multiple linear regression (SMLR), decision tree (DT), and artificial neural network (ANN) models to predict the spatial variability of soil saturated hydraulic conductivity (K_s) in parts of the Chardawel and Chamshir sub-basins, Ilam Province, Iran. Our findings indicated that incorporating environmental covariates with soil properties can significantly enhance the accuracy of K_s estimation, particularly when using the ANN model; thus, this model can be used to develop a digital K_s map with acceptable accuracy to support sustainable land planning.