

Comparison of the Performance of Neuro-fuzzy, Gene Expression Programming and Random Forest Models in Estimating Soil Penetration Resistance

H. Ali Abdoraza¹, Sh. Asghari^{1*}, M. Hasanpour Kashani², H. Shahab Arkhazloo¹

1- Department of Soil Sciences and Engineering, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran

(*- Corresponding Author Email: shasghari@uma.ac.ir)

2- Water Engineering Department, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran

Received: 09 June 2026

Revised: 15 July 2026

Accepted: 21 July 2026

Available Online: 21 July 2026

How to cite this article:

Ali Abdoraza, H., Asghari, Sh., Hasanpour Kashani, M., & Shahab Arkhazloo, H. (2026). Comparison of the performance of neuro-fuzzy, gene expression programming and random forest models in estimating soil penetration resistance. *Journal of Soil and Water*, 40(2), 227-244. (In Persian). <https://doi.org/10.22067/jsw.2026.99264.1555>

Introduction

Soil compaction is a significant component that has a detrimental impact on soil structure, inhibits plant development, lowers water penetration rate, diminishes crop production, and raises machine (e.g., tractors, cultivators, etc.) usage costs. In general, compaction is the result of applied pressure that results in a rise in bulk density or decrease in soil porosity. One of the most used indices for the investigation and evaluation of soil compaction is based on penetration resistance (PR) of the soil. However, identifying soil compaction by the records of soil penetrometer equipment in the field is time-consuming, expensive, and may produce unreliable results due to instrumental errors; thus, it is useful to apply different intelligent models to predict PR through easily accessible and low-cost soil parameters. The aim of this study was to compare the performance of three intelligent models, namely neuro-fuzzy (NF), gene expression programming (GEP), and random forest (RF), in estimating PR from readily available soil variables.

Materials and methods

Disturbed and undisturbed soil samples (n=105) were collected from 0-10 cm depth of agricultural lands in Ardabil plain, Iran. The values of sand, silt, clay, CaCO₃, bulk (BD) and particle (PD) density, organic carbon (OC), field water content (FWC) and saturated water content were measured in the soil samples in the laboratory according to the standard methods. Mean geometric diameter (d_g) and geometric standard deviation (σ_g) of soil particles were calculated from sand, silt and clay percentages. Total porosity (n) was computed using BD and PD data. The penetration resistance (PR) of the soil was obtained in the field using cone penetrometer (analog model) with 5 replicates. Data were randomly divided in to two sets: 78 data points for training and 27 for testing the models. Fifteen different combinations of readily available soil variables were selected as inputs for the neuro-fuzzy (NF), gene expression programming (GEP), and random forest (RF) models to estimate PR. The triangular, sigmoid, trapezoid, gaussian and bell shape membership functions in the input layer and constant membership function in the output layer, were applied in the neuro-fuzzy (NF) modeling using MATLAB software through a trial-and-error method. A set of optimal parameters was chosen before developing a best GEP model in the Gene Xpro Tools 4.0 software. The number of chromosomes and genes, head size and linking function were selected by the trial and error method, and they are 30, 3, 8, and +, respectively. The rates of genetic operators were chosen according to the literature studies. Weka software was used for planning random forest (RF) models. The accuracy of NF, GEP and RF models in estimating PR was evaluated by coefficient of determination (R^2), normalized root mean square error (NRMSE), mean error (ME) and Nash-Sutcliffe coefficient (NS) statistics.



Authors retain the copyright. This is an open access article distributed under [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

<https://doi.org/10.22067/jsw.2026.99264.1555>

Results and discussion

The values of sand (26.26 to 87.43%), silt (5.99 to 67.18%), clay (3.99 to 17.34%), OC (0.30 to 2.41%), FWC (4.56 to 33.18%), BD (1.02 to 1.63 g cm⁻³) and PR (1.10 to 6.60 MPa) indicated good variations in the soils of the studied area. Significant correlations were found between PR and FWC ($r = -0.57^{**}$), sand ($r = -0.21^*$), OC ($r = -0.47^{**}$) and BD ($r = 0.66^{**}$). Previous researchers also reported a negative and significant correlation between PR and FWC and a positive and significant correlation between PR and BD. The results of NF, GEP and RF models showed that the most suitable variables to predict PR were field water content (FWC), sand, bulk density (BD), total porosity and mean geometric diameter (d_g) of soil particles. Values of coefficient of determination (R^2), normalized root mean square error (NRMSE), mean error (ME), and Nash-Sutcliffe coefficient (NS) were calculated for the best models based on the test data as follows: 0.50, 0.19, 0.03 MPa, 0.51 for NF; 0.51, 0.20, 0.13 MPa, 0.48 for GEP; and 0.58, 0.20, 0.31 MPa, 0.50 for RF.

Conclusion

The results showed that according to the lowest values of normalized root mean square error (NRMSE) and the highest values of Nash-Sutcliffe coefficient (NS), the accuracy of the neuro-fuzzy (NF) model to estimate soil PR was higher than that of the gene expression programming (GEP) and random forest (RF) models in this study; The input variables of the best NF model in estimating soil penetration resistance (PR) were field water content (FWC), sand and bulk density (BD).

Keywords: Agricultural lands, Intelligent models, Readily available parameters, Soil penetration resistance

مقایسه عملکرد مدل‌های نیرو فازی، برنامه‌ریزی بیان ژن و جنگل تصادفی در برآورد مقاومت فروروی خاک

حسین علی عبد الرضا^۱ - شکراله اصغری^{۱*} - مهسا حسینیور کاشانی^۲ - حسین شهاب آرخازلو^۱

تاریخ دریافت: ۱۴۰۵/۰۳/۱۹

تاریخ پذیرش: ۱۴۰۵/۰۴/۳۰

چکیده

مقاومت فروروی (PR) خاک یکی از ویژگی‌های کلیدی و مهم در ارزیابی میزان فشردگی خاک در برابر تردد ادوات کشاورزی در اراضی زراعی به شمار می‌رود. اندازه‌گیری مستقیم PR در خاک وقت‌گیر و پرهزینه بوده و به دلیل خطای ابزاری ممکن است چندان قابل اعتماد نباشد. هدف اصلی این پژوهش، ارزیابی و مقایسه عملکرد سه مدل هوشمند نروفازی (NF)، برنامه‌ریزی بیان ژن (GEP) و جنگل تصادفی (RF) در برآورد PR از روی ویژگی‌های زود یافت خاک می‌باشد. برای این منظور، داده‌های مربوط به ۱۰۵ نمونه خاک زراعی دشت اردبیل شامل رطوبت خاک مزرعه، رطوبت اشباع، شن، سیلت، رس، میانگین و انحراف معیار هندسی قطر ذرات، جرم مخصوص ظاهری و حقیقی، تخلخل کل، کربن آلی، آهک و PR به کار گرفته شد. برای آموزش مدل‌ها از ۷۸ داده و برای آزمون مدل‌ها از ۲۷ داده استفاده گردید. پانزده ترکیب متفاوت از متغیرهای زود یافت خاک به عنوان ورودی مدل‌های NF، GEP و RF برای برآورد PR انتخاب شد. رطوبت خاک مزرعه به عنوان یک متغیر سهل الوصول و زود یافت در بهترین مدل NF، GEP و RF برای برآورد PR خاک مورد استفاده قرار گرفت. مقادیر آماره‌های ضریب تبیین (R^2)، ریشه میانگین مربعات خطای نرمال شده (NRMSE)، میانگین خطا (ME) و ضریب نش-ساتکلیف (NS) براساس داده‌های آزمون برای ۰/۱۹، ۰/۵۰، ۰/۰۳ MPa و ۰/۵۱ و ۰/۲۰، ۰/۴۸ و ۰/۵۸ MPa، ۰/۲۰ و ۰/۳۱ به ترتیب برای بهترین مدل NF، GEP و RF در برآورد PR تعیین گردید. نتایج نشان داد که مدل NF با سه متغیر ورودی شامل جرم مخصوص ظاهری، شن و رطوبت خاک مزرعه در مقایسه با دو مدل هوشمند دیگر (GEP و RF) به دلیل داشتن NRMSE پایین و NS بالا توانست مقاومت فروروی خاک را با دقت نسبتاً بالایی در اراضی زراعی دشت اردبیل برآورد نماید.

واژه‌های کلیدی: اراضی زراعی، مقاومت خاک، مدل‌های هوشمند، ویژگی‌های زود یافت

مقدمه

روش قابل اعتمادی برای تشخیص علل اصلی فشردگی خاک، به ویژه در خاک‌های در معرض رطوبت، وجود ندارد (Elaoud et al., 2023). مقاومت فروروی^۳ (PR) خاک یکی از شاخص‌های اصلی فشردگی محسوب می‌شود که مقاومت مکانیکی خاک را در برابر نفوذ هوا، آب و ریشه در خاک نشان می‌دهد؛ اگر تخلخل تهویه‌ای خاک به زیر ۱۰ درصد کاهش یابد و یا مقاومت فروروی (PR) خاک به بالای ۲

بررسی و ارزیابی میزان فشردگی خاک پس از عبور ادوات کشاورزی (تراکتور، کولتیواتور و...) به عنوان یک راه حل پیشگیرانه برای مراقبت از خاک‌های کشاورزی در برابر تخریب و نیز حفاظت از کشاورزی پایدار در نظر گرفته می‌شود. در واقع، در حال حاضر هیچ

۱- گروه علوم و مهندسی خاک، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران
(*) نویسنده مسئول: Email: shasghari@uma.ac.ir

۲- گروه مهندسی آب، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران

مگا پاسکال افزایش یابد رشد و گسترش ریشه در خاک متوقف می‌گردد (Warrick, 2002). دشت اردبیل یکی از قطب‌های تولید سیب‌زمینی در ایران به شمار می‌رود بنابراین اطلاع از وضعیت فشردگی خاک‌های زراعی دشت اردبیل و ارائه راهکار مناسب برای مدیریت این مساله ضروری به نظر می‌رسد.

اندازه‌گیری مستقیم مقاومت فروروی خاک به دلیل تغییرات مکانی و زمانی بسیار بالای آن کاری بسیار طاقت‌فرسا، وقت‌گیر و پرهزینه می‌باشد لذا اخیراً محققان به برآورد این ویژگی دیرپافت خاک از روی ویژگی‌های زودپافت و سهل‌الوصول خاک (مانند شن، سیلت، رس، کربن آلی، جرم مخصوص ظاهری و رطوبت خاک مزرعه) روی آورده اند. برای این منظور لازم است توابع انتقالی^۱ (PTF) خاک ایجاد گردد تا مقاومت فروروی خاک را به عنوان یک متغیر دیرپافت به متغیرهای زودپافت خاک مرتبط سازد. توابع انتقالی رگرسیونی چندگانه خطی^۲ (MLR) و غیر خطی متعددی برای برآورد مقاومت فروروی خاک توسط محققان مختلف مانند بیات و همکاران (Bayat et al., 2008) و شیخ‌زاده و همکاران (Sheykhzadeh et al., 2016) ارائه گردیده است همچنین در دهه‌های اخیر، استفاده از توابع انتقالی خاک مبتنی بر مدل‌های هوشمند مانند شبکه عصبی مصنوعی^۳ (ANN)، نرو فازی^۴ (NF)، برنامه‌ریزی بیان ژن^۵ (GEP) و جنگل تصادفی^۶ (RF) برای برآورد متغیرهای دیرپافت خاک رایج گردیده است.

بیات و همکاران (Bayat et al., 2008) از توابع انتقالی رگرسیون خطی و غیرخطی و نیز مدل شبکه عصبی مصنوعی برای برآورد مقاومت فروروی در خاک‌های منتخب از دو منطقه در ایران شامل مراغه و همدان استفاده کردند. نتایج نشان داد که شبکه عصبی مصنوعی ($R^2=0.85$) در مقایسه با توابع رگرسیونی ($R^2=0.70$) دارای دقت بالا و خطای پایین در برآورد مقاومت فروروی خاک بود. آن‌ها گزارش نمودند مدل شبکه عصبی مصنوعی در مقایسه با توابع رگرسیونی قادر است مقاومت فروروی خاک را از روی متغیرهای رطوبت وزنی، تخلخل کل، درجه اشباع نسبی (نسبت رطوبت اشباع به تخلخل کل) با دقت بالایی برآورد نماید. نتایج پژوهش سانتوس و همکاران (Santos et al., 2011) نشان دادند مقاومت فروروی خاک رابطه همبستگی مثبت و معنی‌دار با جرم مخصوص ظاهری خاک و رابطه همبستگی منفی و معنی‌دار با رطوبت وزنی خاک دارد. همچنین آن‌ها گزارش کردند توابع رگرسیونی دارای دقت کم ($R^2=0.92$) و خطای زیاد ($RMSE = 0.95 \text{ MPa}$) نسبت به مدل شبکه عصبی مصنوعی ($R^2=0.98$ و $RMSE = 0.08 \text{ MPa}$) در برآورد مقاومت فروروی

خاک بودند. نتایج مطالعه شیخ‌زاده و همکاران (Sheykhzadeh et al., 2016) در دشت اردبیل نشان داد که مدل ANN در مقایسه با مدل MLR دارای دقت بالاتری در برآورد PR خاک بود و براساس نتایج آنالیز حساسیت، متغیرهای انحراف معیار هندسی قطر ذرات، جرم مخصوص ظاهری، رطوبت خاک مزرعه و سیلت به ترتیب بیشترین تأثیر را در برآورد PR خاک توسط بهترین مدل ANN از خود نشان دادند. اخیراً از سایر روش‌های هوشمند مانند مدل نرو فازی (NF)، مدل برنامه‌ریزی بیان ژن (GEP) و مدل جنگل تصادفی (RF) نیز برای برآورد مقاومت فروروی (PR) از روی ویژگی‌های زودپافت خاک استفاده شده است؛ حسینی و همکاران (Hosseini et al., 2018) از الگوریتم‌های آماری، سیستم‌های استنتاج تطبیقی عصبی-فازی (NF)، شبکه‌های عصبی مصنوعی (ANN) و سیستم‌های استنتاج فازی (FIS) برای پیش‌بینی مقاومت فروروی خاک استفاده کردند و آن‌ها را با مدل‌های آماری سنتی مانند رگرسیون خطی چندگانه (MLR) مقایسه کردند. برای دستیابی به این هدف، جرم مخصوص ظاهری و رطوبت حجمی خاک (به عنوان متغیرهای ورودی مدل) و مقاومت فروروی خاک (به عنوان متغیر دیرپافت) در ۲۰۰ نمونه خاک برداشته شده از عمق ۰ تا ۲۵ سانتی‌متری مزرعه دانشگاه کشاورزی و منابع طبیعی گرگان اندازه‌گیری شد. نتایج نشان داد که روش‌های هوشمند ابزار مناسبی برای به حداقل رساندن عدم قطعیت‌ها در پروژه‌های مهندسی خاک هستند. مدل NF توانست مقاومت فروروی خاک را با دقت بیشتری نسبت به سایر مدل‌ها با $R^2 = 0.93$ و $RMSE = 299.41$ برآورد نماید. الایود و همکاران (Elaoud et al., 2023) از مدل ANN برای برآورد مقاومت فروروی خاک در کشور تونس استفاده کردند. ریشه میانگین مربعات خطا (RMSE) برابر با ۰/۵۱ مگا پاسکال نشان داد که برآوردهای حاصل از مدل‌های ANN با داده‌های PR اندازه‌گیری شده میدانی مطابقت دارند. آن‌ها نتیجه گرفتند که مدل‌های ANN توسعه‌یافته به طور مؤثر می‌توانند برای پیش‌بینی‌های دقیق‌تر در مورد فشردگی خاک در شرایط رطوبتی مختلف استفاده شوند. این کار به کشاورزان کمک می‌کند تا استفاده از ماشین را بهینه کنند و از این‌رو تولید را افزایش داده و بازده را با حداقل هزینه بهبود بخشند. نتایج پژوهش اصغری و همکاران (Asghari et al., 2024) در اراضی جنگلی، مرتعی و زراعی منطقه فندقلوی اردبیل نشان داد که مدل GEP در مقایسه با مدل رگرسیون خطی به علت استفاده از روابط غیر خطی، از دقت بیشتر و خطای کمتر در برآورد PR خاک برخوردار بود هر چند دقت آن در مقایسه با مدل شبکه عصبی مصنوعی پایین بود؛ مقادیر

5- Gene expression programming
6- Random forest

1- Pedotransfer function
2- Multiple linear regression
3- Artificial neural network
4- Neuro-fuzzy

دقت توابع انتقالی رگرسیونی ارائه شده توسط شیخ‌زاده و همکاران (Sheykhzadeh et al., 2016) با مدل‌های هوشمند به کار رفته در پژوهش حاضر از نظر برآورد مقاومت فروروی خاک.

مواد و روش‌ها

در تحقیق حاضر از اطلاعات مربوط به ۱۰۵ نمونه خاک (عمق صفر تا ۱۰ سانتی‌متر) بخشی از اراضی زراعی دشت اردبیل با مساحت حدود ۱۰۰۰ کیلومتر مربع (۲۴ × ۴۲) و کاربری غالب زیر کشت سیب‌زمینی، گندم، چغندرقد و یونجه که در فواصل تقریباً منظم (۳۰۰۰ × ۳۰۰۰) طی بهار و تابستان ۱۳۹۲ برداشته شده بود، استفاده گردید (Sheykhzadeh et al., 2016). متغیرهای زودیاخت و سهل الوصول خاک شامل ذرات شن، سیلت و رس به روش هیدرومتری چهار قرائته، پارامترهای میانگین (d_g) و انحراف استاندارد (σ_g) هندسی قطر ذرات خاک از روی درصد ذرات، کربن آلی به روش اکسیداسیون تر، آهک به روش تیتراسیون، جرم مخصوص ظاهری (BD) در استوانه های استیل به قطر و ارتفاع ۵ سانتی‌متر، جرم مخصوص حقیقی (PD) به روش پیکنومتر، تخلخل کل از روی جرم مخصوص ظاهری و حقیقی، رطوبت اشباع (SWC) و رطوبت خاک مزرعه (FWC) در لحظه نمونه‌برداری هر دو به روش وزنی در استوانه‌های استیل براساس دستورالعمل‌های رایج و استاندارد تعیین گردیدند. مقاومت فروروی (PR) به‌عنوان متغیردیاریافت و پرهزینه خاک به طور درجا و در موقع نمونه‌برداری خاک دست‌نخورده در مزرعه با استفاده از دستگاه فروسج مخروطی مدل عقربه‌ای ساخت شرکت اکل کمپ کشور هلند در ۱۰ تکرار اندازه‌گیری شد (Lowery & Morrison, 2002). برخی ویژگی‌های نمونه‌های خاک در جدول ۱ و نیز ضریب همبستگی پیرسون بین متغیرها در جدول ۲ آورده شده است. خاک‌های مورد مطالعه دارای کلاس بافتی درشت تا متوسط (۸ نمونه شن لومی، ۷۰ نمونه لوم شنی، ۶ نمونه لوم و ۲۱ نمونه لوم سیلتی) بودند (شکل ۱). برای اشتقاق توابع رگرسیونی خطی چندگانه (MLR) بین متغیرهای زودیاخت (شن، سیلت، رس، جرم مخصوص ظاهری و حقیقی، تخلخل کل، SWC و FWC، CaCO_3 ، کربن آلی، d_g و σ_g) با متغیر دیاریافت (PR) خاک از روش گام به گام^۴ در نرم‌افزار SPSS 22 استفاده گردید. براساس روش گام به گام، ۱۵ مدل رگرسیونی (جدول ۳) با ترکیب‌های متفاوت از متغیرهای زودیاخت ورودی برای برآورد متغیر دیاریافت (PR) تعیین شدند (Sheykhzadeh et al., 2016). به‌منظور بررسی و مقایسه دقیق نتایج پژوهش حاضر، در فرآیند مدل‌سازی PR با استفاده از سه روش هوشمند نرو فازی (NF)، برنامه

آماره‌های ضریب تبیین (R^2)^۱ و ریشه میانگین مربعات خطا (RMSE)^۲، میانگین خطا^۳ (ME) و ضریب نش-ساتکلیف^۴ (NS) براساس داده‌های آزمونی) به‌ترتیب برای بهترین مدل ANN، MLR، و GEP برابر با ۰/۴۴، ۱/۱۹ MPa، ۰/۱۹ MPa و ۰/۳۶، ۰/۹۲، ۰/۴۱ MPa، ۰/۴۴ MPa، ۰/۰۵ - و ۰/۹۲، ۰/۷۹، ۰/۹۱ MPa، ۰/۱۳ MPa و ۰/۶۳ محاسبه گردید. آن‌ها گزارش نمودند رطوبت خاک مزرعه، سیلت و جرم مخصوص ظاهری نسبی، مناسب‌ترین متغیرهای ورودی در برآورد PR خاک هستند. اونال و همکاران (Unal et al., 2024) رابطه بین رطوبت وزنی، جرم مخصوص ظاهری، نسبت تخلخل، عمق و مقاومت فروروی خاک را با کمک ANN در یک تحقیق مزرعه‌ای در آنتالیای ترکیه مدل‌سازی نمودند. برای این منظور، مدل‌های شبکه عصبی رگرسیون تعمیم‌یافته (GRNN) و تابع پایه شعاعی (RBF) برای تخمین مقادیر مقاومت فروروی خاک مقایسه گردیدند. نتایج آماره‌های ارزیابی نشان داد که مدل GRNN نتایج بهتری نسبت به مدل RBF در برآورد مقاومت فروروی خاک ارائه می‌دهد. پارسایی و همکاران (Parsaie et al., 2025) نقشه رقومی مقاومت فروروی خاک را با استفاده از مدل جنگل تصادفی (RF) و درخت تصمیم تقویت شده با گرادیان (XGBoost) در حوزه آبخیز کیلانه استان کردستان تهیه نمودند. مقاومت فروروی خاک با استفاده از فروسج دستی در ۱۵۰ نقطه و در عمق ۰ تا ۱۰ سانتی‌متری خاک اندازه‌گیری شد. برای برآورد مقاومت فروروی خاک از داده‌های طیفی و متغیرهای کمکی مستخرج از مدل رقومی ارتفاع و تصاویر ماهواره‌ای و نیز متغیرهای خاک شامل شن، سیلت، رس، کربن آلی، آهک، جرم مخصوص ظاهری و میانگین هندسی قطر خاکدانه‌ها استفاده گردید. نتایج نشان داد مدل XGBoost از دقت بالا ($R^2 = 0.61$) و خطای پایین (RMSE) نرمال شده برابر با ۰/۱۱) در برآورد مقاومت فروروی خاک برخوردار بود.

بررسی منابع نشان داد تحقیقات انجام گرفته در زمینه برآورد مقاومت فروروی از روی متغیرهای زودیاخت خاک با استفاده مدل‌های هوشمند، انگشت‌شمار است همچنین تاکنون دقت توابع انتقالی رگرسیونی و مدل‌های هوشمند در برآورد مقاومت فروروی خاک به‌ویژه در دشت اردبیل مورد مقایسه قرار نگرفته است. بنابراین عمده‌ترین اهداف این پژوهش عبارتند از ۱) برآورد مقاومت فروروی خاک با استفاده از سه مدل هوشمند نرو فازی (NF)، برنامه‌ریزی بیان ژن (GEP) و جنگل تصادفی (RF). ۲) ارزیابی عملکرد سه مدل هوشمند مذکور در برآورد مقاومت فروروی خاک با استفاده از محک‌های آماری شامل ضریب تبیین (R^2)، ریشه میانگین مربعات خطای نرمال شده^۵ (NRMSE)، میانگین خطا (ME) و نش-ساتکلیف (NS). ۳) مقایسه

4- Nash-Sutcliffe coefficient

5- Normalized root mean square error

6- Stepwise

1- Determination coefficient

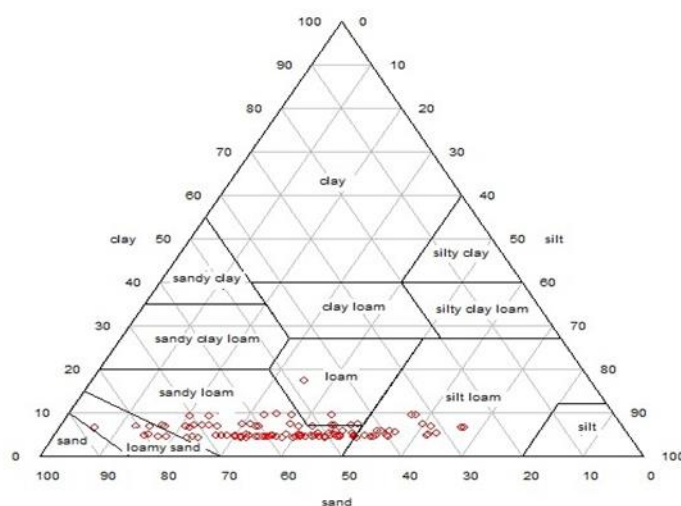
2- Root mean square error

3- Mean error

عصبی مصنوعی جهت آسان کردن فرآیند یادگیری و انطباق ترکیب می‌کند به این صورت که مشکل اصلی در طراحی سیستم‌های فازی که همان به‌دست آوردن قواعد "اگر- آنگاه" فازی است را به‌وسیله استفاده مؤثر از قابلیت یادگیری شبکه عصبی مصنوعی جهت تولید خودکار این قواعد و بهینه‌سازی پارامترها مرتفع می‌کند (Shiri et al., 2017).

ریزی بیان ژن (GEP) و جنگل تصادفی (RF) نیز از همان ۱۵ ترکیب متفاوت متغیرهای ورودی در مدل‌های رگرسیونی ارائه شده توسط شیخ‌زاده و همکاران (Sheykhzadeh et al., 2016) استفاده شد. لازم به ذکر است که مطابق تشخیص روش گام به گام، از بین ۱۲ متغیر ورودی (زودیافت)، دو متغیر CaCO_3 و SWC در برآورد متغیر دیریافت (PR) حائز اهمیت شناخته نشده و لذا وارد هیچ‌کدام از مدل‌های پانزده-گانه رگرسیونی و هوشمند نگردید.

سیستم استنتاجی نروفازی (NF) منطق فازی را با شبکه‌های



شکل ۱- توزیع کلاس بافت در خاک‌های مورد مطالعه

Figure 1- Distribution of textural class in the studied soils

جدول ۱- آماره‌های توصیفی متغیرهای خاک در منطقه مورد مطالعه

Table 1- Descriptive statistics of soil variables in the studied area

متغیر Variable	آموزش Training (n= 78)				آزمون Testing (n= 27)			
	Min.	Max.	Mean	Sd.	Min.	Max.	Mean	Sd.
BD (g cm^{-3})	1.02	1.63	1.27	0.11	1.03	1.44	1.24	0.11
PD (g cm^{-3})	2	2.57	2.27	0.13	2.04	2.46	2.27	0.11
TP (%)	26.12	55.62	44	175	37.22	56.78	45.46	5.16
Sand (%)	26.87	87.43	54.96	13.21	26.26	80.22	56.37	14.21
Silt (%)	5.99	66.8	39.24	13.17	15.21	67.18	38.31	14.49
Clay (%)	3.99	17.34	5.8	2.02	4.07	9.1	5.32	1.45
d_g (mm)	0.05	0.56	0.2	0.1	0.07	0.48	0.21	0.11
σ_g (-)	4.69	12.15	6.83	0.95	4.86	8.11	6.55	0.8
OC (%)	0.3	2.41	1.09	0.41	0.47	1.82	1.14	0.37
CaCO_3 (%)	3.45	26.99	12.77	5.17	5.31	15.76	10.58	3
SWC ($\% \text{ g g}^{-1}$)	25.23	73.67	50.01	9.07	32.83	71.57	52.26	9.71
FWC ($\% \text{ g g}^{-1}$)	4.79	31.34	10.68	5.15	4.56	33.18	11.96	6.55
PR (MPa)	1.1	6.6	2.86	1.4	1.23	2.07	2.7	1.33

BD: جرم مخصوص ظاهری (bulk density); PD: جرم مخصوص حقیقی (particle density); TP: تخلخل کل (total porosity); clay: رس، silt: سیلت، sand: شن، d_g : میانگین هندسی قطر ذرات (geometric mean diameter)، σ_g : انحراف استاندارد هندسی ذرات (geometric standard deviation)، OC: کربن آلی (organic carbon)، CaCO_3 : آهک، SWC: رطوبت اشباع (saturated water content)، FWC: رطوبت مزرعه (field water content)، PR: مقاومت فروروی (penetration resistance). Sd: انحراف استاندارد، Mean: میانگین، Max: حداکثر، Min: حداقل.

جدول ۲- ضریب همبستگی پیرسون (r) بین مقاومت فروروی خاک (PR) و متغیرهای ورودی مدل‌ها (n=105)

Table 2- Pearson correlation coefficient (r) between soil available water and input variables of models (n=102)

Variable متغیر	Sand	Silt	Clay	dg	σ_g	OC	CaCO ₃	BD	PD	TP	FWC	SWC
PR	-0.21*	0.45**	0.19	-0.26*	0.33**	-0.47**	0.29**	0.66**	0.27*	-0.59**	-0.57**	-0.33**

* و **: به ترتیب معنی دار در سطح احتمال ۱ و ۵ درصد. توصیف علائم متغیرها در زیر جدول ۱ آورده شده است.

Variable symbols description is available under Table 1. *, **: Significant at $P < 0.05$ and $P < 0.01$, respectively.

جدول ۳- مدل‌های رگرسیون خطی چندگانه (MLR) مقاومت فروروی خاک (Sheykhzadeh et al., 2016)

Table 3- MLR models of soil penetration resistance (PR) (Sheykhzadeh et al., 2016)

شماره Number	مدل Model
1	$PR = -0.572 + 0.974 BD - 0.005 Sand$
2	$PR = 0.675 BD - 0.004 Sand - 0.302 FWC$
3	$PR = 0.588 BD - 0.003 Sand - 0.309 FWC + 0.033 \sigma_g$
4	$PR = -1.023 + 0.997 BD + 0.004 Silt$
5	$PR = 0.695 BD - 0.004 Silt - 0.305 FWC$
6	$PR = 0.585 BD - 0.003 Silt - 0.307 FWC + 0.039 \sigma_g$
7	$PR = 1.066 - 0.541 FWC - 0.608 d_g$
8	$PR = 0.725 - 0.527 FWC - 0.404 d_g + 0.042 \sigma_g$
9	$PR = 0.718 - 0.428 FWC - 0.404 d_g + 0.042 \sigma_g - 0.085 OC$
10	$PR = 1.521 - 0.476 FWC - 0.014 TP$
11	$PR = 1.429 - 0.407 FWC - 0.012 TP - 0.459 d_g$
12	$PR = 1.477 - 0.018 TP - 0.005 Sand$
13	$PR = 0.705 - 0.019 TP - 0.004 Sand + 0.346 PD$
14	$PR = 0.929 - 0.006 Sand - 0.176 OC$
15	$PR = 0.785 - 0.006 Sand - 0.162 OC + 0.018 Clay$

توصیف علائم متغیرها در زیر جدول ۱ آورده شده است

Variable symbols description is available under Table 1.

Tools 4.0 استفاده گردید. در موقع اجرای برنامه، توابع مختلف حسابی، مثلثاتی و نیز توابع تعریف شده توسط کاربر به روش سعی و خطا به کار برده شد. عملگرهای ریاضی پر کاربرد در فرآیند اجرای مدل، چهار عمل اصلی (+، -، × و /)، لگاریتم طبیعی، تانژانت، کتانژانت، توان ۲ و تعداد ۵۰۰۰ تکرار در نظر گرفته شد. پارامترهای مورد استفاده در اجرای مدل GEP به طور خلاصه در جدول ۴ و نیز درخت‌های توصیفی به دست آمده در پایان فرآیند بهینه‌سازی برای بهترین مدل GEP (مدل شماره ۱ با دو متغیر ورودی جرم مخصوص ظاهری و شن) در شکل ۲ آورده شده است.

اولین بار بریمن، مدل جنگل تصادفی (RF) را ارائه نمود (Breiman, 2001). این روش یک الگوریتم ساده و انعطاف‌پذیر برای یادگیری ماشین می‌باشد. مدل جنگل تصادفی یک روش گروهی بر پایه درختی است که در آن همه درختان به مجموعه‌ای از متغیرهای تصادفی وابسته هستند و جنگل از درختان رگرسیونی زیاد واقع در کنار یکدیگر رشد می‌یابد (Singh et al., 2020).

در این پژوهش، مدل‌های نروفازی (NF) لازم برای برآورد مقاومت فروروی (PR) خاک، با استفاده از سیستم استنتاجی نروفازی موجود در جعبه ابزار نرم‌افزار MATLAB به روش سعی و خطا و با به کارگیری توابع عضویت مثلثی (trimf)، سیگموئیدی (sigmf)، دوزنقه‌ای (trapmf)، گوسی (gaussmf) و زنگوله‌ای شکل (bellmf) در لایه ورودی و تابع عضویت ثابت (constant) در لایه خروجی و با تکرارهای مختلف ایجاد شدند (Ghorbani et al., 2019). همچنین تعداد تابع عضویت ۳ و تعداد تکرار بهینه ۵۰ تا ۱۵۰ تا برای اجرای ۱۵ مدل NF به کار رفته در روش نروفازی تعیین گردید. روش اصلی آموزش در سیستم فازی، روش پس انتشار خطا می‌باشد.

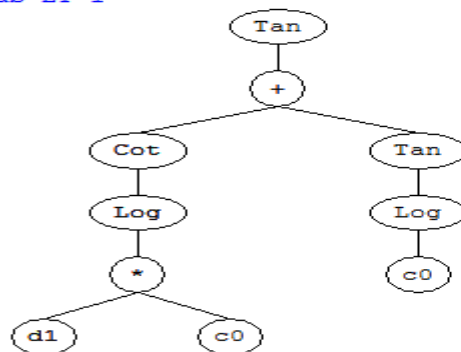
برنامه‌ریزی بیان ژن (GEP) توسط فریرا در سال ۱۹۹۹ ابداع شد. در این برنامه کروموزوم‌های خطی و ساده با طول ثابت مشابه با الگوریتم ژنتیک و نیز ساختارهای شاخه‌ای با اندازه‌ها و اشکال متفاوت مشابه با نمودار درختی در برنامه‌ریزی ژنتیک استفاده می‌گردد (Ferreira, 2006). در این پژوهش برای برآورد مقاومت فروروی (PR) خاک به روش برنامه‌ریزی بیان ژن (GEP)، از نرم‌افزار Gene Xpro

جدول ۴- مقادیر پارامترهای مدل برنامه‌ریزی بیان ژن (GEP)

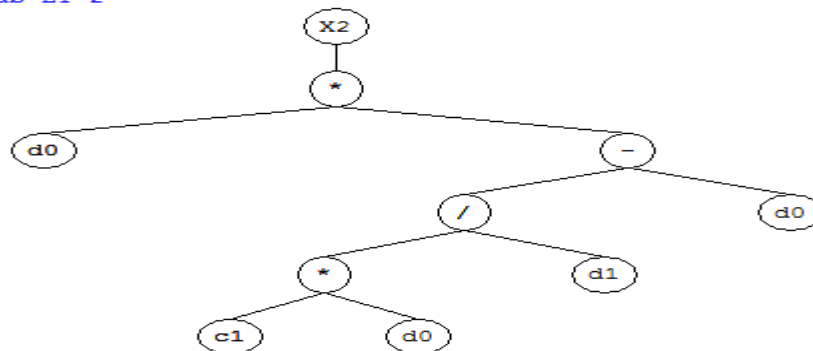
Table 4- The values of GEP model parameters

مقادیر Values	عوامل ژنتیک Genetic factors	مقادیر Values	تنظیمات عمومی General adjustment
0.044	Mutation rate نرخ جهش	30	تعداد کروموزم‌ها Number of chromosome
0.1	Inversion rate نرخ وارون‌سازی	8	اندازه سر Head size
0.3	one point recombination rate نرخ ترکیب تک نقطه‌ای	3	تعداد ژن‌ها Number of genes
0.3	two point recombination rate نرخ ترکیب دو نقطه‌ای	جمع (+)	تابع پیوند Linking function
0.1	Gene recombination rate نرخ ترکیب ژن	RMSE	معیار خطای تابع Fitness function error type

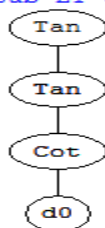
Sub-ET 1



Sub-ET 2



Sub-ET 3



شکل ۲- درخت‌های توصیفی ایجاد شده توسط بهترین مدل برنامه‌ریزی بیان ژن (GEP)

d0 و d1 به ترتیب شن و جرم مخصوص ظاهری را نشان می‌دهد.

Figure 2- The expression trees developed by the best gene expression programming (GEP) model
d0 and d1 indicates sand and bulk density, respectively.

مدل می‌باشد. از میان مدل‌های ایجاد شده برای برآورد PR، مناسب ترین مدل NF، GEP و RF با ترکیب متغیرهای ورودی در داده‌های آزمونی بر اساس معیارهای آماری مذکور تعیین شد. آزمون مدل‌ها در تشخیص دقت مدل به ازای داده‌هایی که در آموزش مدل استفاده نشده‌اند به کار می‌رود.

نتایج و بحث

ارزیابی مدل‌های نروفازی (NF) در برآورد مقاومت فروروی (PR) خاک

از ۱۵ مدل نروفازی (NF) با همان ترکیب متغیرهای ورودی به کار رفته در مدل‌های رگرسیون خطی چندگانه (جدول ۳) برای تخمین مقاومت فروروی (PR) خاک استفاده شد. مقادیر آماره‌های ضریب تبیین (R^2) ، ریشه میانگین مربعات خطای نرمال شده (NRMSE)، میانگین خطا (ME) و ضریب نش - ساتکلیف (NS) برای ارزیابی دقت تخمین مدل‌های NF در تخمین PR در جدول ۵ آورده شده است. کاربرد R^2 به تنهایی شاید از نظر مقایسه اعتبار مدل‌ها، گمراه‌کننده بوده و الزاماً R^2 نزدیک به یک همیشه بیانگر برازش دقیق مقادیر تخمینی و اندازه‌گیری یک متغیر دیر یافت و در نتیجه پایین بودن خطا نیست (Kozak et al., 1996) لذا در این پژوهش، برای رتبه‌دهی هر سه مدل هوشمند مورد استفاده (NF، GEP و RF) از نظر دقت تخمین PR، از آماره‌های NRMSE (بر اساس رتبه‌دهی Wan et al., 2023) و NS در داده‌های آزمون استفاده گردید. واسیم و همکاران (Waseem et al., 2017) نیز گزارش نمودند اگر مقدار آماره NS بین ۱ و ۰/۷۵ باشد عملکرد مدل بسیار خوب است؛ اگر مقدار NS بین ۰/۷۵ و ۰/۶۵ باشد عملکرد مدل خوب است؛ اگر مقدار NS بین ۰/۶۵ و ۰/۵۰ باشد عملکرد مدل قابل قبول است ولی اگر مقدار آماره NS کوچکتر ۰/۵۰ باشد، عملکرد مدل نامطلوب است. بر اساس پایین بودن آماره NRMSE و بالا بودن آماره NS در داده‌های آزمون، می‌توان گفت که مدل NF شماره ۲ (جدول ۵) با سه متغیر ورودی شامل جرم مخصوص ظاهری، شن و رطوبت خاک مزرعه دارای بالاترین دقت از نظر برآورد PR نسبت به سایر مدل‌های NF بود. مقادیر R^2 ، NRMSE، ME و NS این مدل در سری آموزش به ترتیب برابر ۰/۶۷، ۰/۱۴، ۰/۰۰ و ۰/۶۷ و برای داده‌های آزمونی برابر ۰/۵۱، ۰/۱۹، ۰/۰۳ و ۰/۵۰ تعیین گردید (جدول ۵). مطابق آماره‌های NRMSE و NS در داده‌های آزمون، عملکرد این مدل در تخمین PR به ترتیب دارای کارکرد مناسب و قابل قبول می‌باشد. اگرچه مقدار اندک مثبت ME (۰/۰۳ MPa) در سری آزمون بیان‌گر برآورد جزئی بیشتر از مقدار واقعی PR خاک توسط مدل شماره ۲ در روش نروفازی می‌باشد ولی در مقایسه با مقادیر به دست آمده توسط شیخ‌زاده و همکاران (Sheykhzadeh et al., 2016) برای

در روش جنگل تصادفی برای هر درخت یک زیر مجموعه‌ای از متغیرهای ورودی اعمال می‌شود. تعداد ورودی‌ها در هر درخت و نیز تعداد درختان بهینه، به روش سعی و خطا تعیین می‌گردد. در این پژوهش برای برآورد مقاومت فروروی (PR) خاک به روش RF از نرم افزار Weka استفاده شد؛ بهترین اجرای مدل RF برای برآورد PR، با ۲ متغیر کمکی موجود در زیر مجموعه تصادفی هر گره ($mtree = 2$) و تعداد ۷۰ تا ۱۰۰ درخت به دست آمد. کل داده‌ها (۱۰۵ داده) در اکسل به طور تصادفی به دو مجموعه آموزشی (۷۸ داده) و آزمونی (۲۷ داده) تقسیم گردیدند سپس برای اجرای هر سه مدل نروفازی (NF)، برنامه‌ریزی بیان ژن (GEP) و جنگل تصادفی (RF) به کار گرفته شدند.

برای مقایسه و ارزیابی دقت مدل‌های NF، GEP و RF در برآورد PR، از آماره‌های ضریب تبیین (R^2)، ریشه میانگین مربعات خطای نرمال شده (NRMSE)، میانگین خطا (ME) و نش - ساتکلیف (NS) (Ghorbani et al., 2019) استفاده شد. به طور کلی، هرچه مقدار R^2 به یک و میزان NRMSE به صفر نزدیک باشد، دقت مدل بیشتر بوده و مقادیر اندازه‌گیری و برآورد شده PR باهم برابر هستند. مقادیر NRMSE کمتر از ۰/۱ بیانگر کارکرد عالی، مقادیر ۰/۱ تا ۰/۲ بیانگر کارکرد مناسب، مقادیر ۰/۲ تا ۰/۳ بیانگر کارکرد متوسط و مقادیر بیش از ۰/۳ NRMSE بیانگر کارکرد ضعیف مدل می‌باشد (Wan et al., 2023). مقادیر مثبت ME بیش‌برازشی و مقادیر منفی آن کم‌برازشی مدل را نشان می‌دهد. اگر معیار NS برابر یک باشد بیانگر آن است که مدل‌سازی با دقت بالایی انجام گرفته و مقادیر اندازه‌گیری و برآورد شده PR مطابقت کامل دارند. میزان NS صفر بیانگر آن است که برآورد‌های مدل، دقتی برابر با دقت میانگین داده‌های اندازه‌گیری شده دارند.

$$NRMSE = \frac{1}{Y_{\max(m)} - Y_{\min(m)}} \sqrt{\frac{\sum_{i=1}^n (Y_{i(p)} - Y_{i(m)})^2}{n}} \quad (1)$$

$$ME = \frac{\sum_{i=1}^n Y_{i(p)} - Y_{i(m)}}{n} \quad (2)$$

$$NS = 1 - \frac{\sum_{i=1}^n (Y_{i(p)} - Y_{i(m)})^2}{\sum_{i=1}^n (Y_{i(m)} - \bar{Y}_m)^2} \quad (3)$$

در معادله‌های ۱، ۲ و ۳، $\bar{Y}_m$ ، $Y_{i(p)}$ ، $Y_{i(m)}$ ، $Y_{\max(m)}$ و $Y_{\min(m)}$ به ترتیب مقادیر اندازه‌گیری، برآورد شده توسط مدل، میانگین، حداکثر و حداقل مقادیر اندازه‌گیری شده PR و n تعداد داده‌ها برای PR در هر

همکاران (Asghari et al., 2024) نیز گزارش کردند مدل‌های هوشمند نروفازی و برنامه‌ریزی بیان ژن در مقایسه با مدل سنتی رگرسیون خطی دارای دقت بیشتر و خطای کمتر در برآورد PR خاک بودند همچنین رطوبت خاک مزرعه، سیلت و جرم مخصوص ظاهری نسبی به‌عنوان مناسب‌ترین متغیرهای ورودی در تخمین PR خاک تعیین گردید.

ارزیابی مدل‌های برنامه‌ریزی بیان ژن (GEP) در برآورد مقاومت فروروی (PR) خاک

مدل‌های برنامه‌ریزی بیان ژن (GEP) برای تخمین مقاومت فروروی (PR) خاک نیز با همان ترکیب متغیرهای ورودی در مدل‌های رگرسیون خطی چندگانه (جدول ۳) و نروفازی (جدول ۵) ایجاد گردیدند. مقادیر آماره‌های ضریب تبیین (R^2)، ریشه میانگین مربعات خطای نرمال شده (NRMSE)، میانگین خطا (ME) و ضریب نش-ساتکلیف (NS) برای ارزیابی دقت تخمین مدل‌های NF در برآورد PR در جدول ۶ آورده شده است.

همین مدل در روش‌های رگرسیون خطی چندگانه ($ME = 0.64$) و شبکه عصبی مصنوعی ($ME = 0.21$ MPa)، به مراتب دارای بیش‌برآوردی کمتری است؛ معادله رگرسیونی این مدل در جدول ۳ نیز بیانگر آن است که جرم مخصوص ظاهری با علامت مثبت و رطوبت خاک مزرعه و شن با علامت منفی وارد مدل گردیدند و اهمیت متغیر جرم مخصوص ظاهری در برآورد PR به علت داشتن ضریب رگرسیون بالاتر به مراتب بیشتر از متغیر رطوبت خاک مزرعه و شن می‌باشد. همبستگی معنی‌دار و مثبت جرم مخصوص ظاهری و همبستگی معنی‌دار و منفی رطوبت خاک مزرعه با PR نیز مؤید این قضیه می‌باشد (جدول ۲). در رتبه‌های بعدی، مدل شماره ۱ با دو متغیر ورودی جرم مخصوص ظاهری و شن ($NRMSE = 0.20$ و $NS = 0.49$) و مدل شماره ۴ با دو متغیر ورودی جرم مخصوص ظاهری و سیلت ($NRMSE = 0.20$ و $NS = 0.49$) توانستند PR را با دقت نسبتاً بالایی برآورد نمایند (جدول ۵). نتایج پژوهش بیات و همکاران (Bayat et al., 2008) نیز نشان داد که مقاومت فروروی خاک با دقت بالا و خطای پایین توسط مدل هوشمند شبکه عصبی مصنوعی در مقایسه با مدل سنتی رگرسیون خطی مورد تخمین قرار گرفت. اصغری و

جدول ۵- معیارهای ارزیابی مدل‌های نروفازی (NF) مقاومت فروروی (PR) خاک
Table 5- Evaluation criteria for neuro-fuzzy (NF) models of PR

شماره Number	متغیرهای ورودی Input variables	آزمون Testing set				آموزش Training			
		R^2	NRMSE	ME (MPa)	NS	R^2	NRMSE	ME (MPa)	NS
1	BD, Sand	0.50	0.20	0.12	0.49	0.58	0.15	0.00	0.58
2	BD, Sand, FWC	0.50	0.19	0.03	0.51	0.67	0.14	0.00	0.67
3	BD, Sand, FWC, σ_g	0.07	0.52	0.01	-2.41	0.78	0.11	0.00	0.77
4	BD, Silt	0.50	0.20	0.16	0.49	0.58	0.15	0.00	0.58
5	BD, Silt, FWC	0.43	0.21	0.01	0.41	0.68	0.13	0.00	0.68
6	BD, Silt, FWC, σ_g	0.18	0.34	0.12	-0.47	0.78	0.11	0.00	0.78
7	FWC, d_g	0.39	0.22	0.18	0.37	0.61	0.15	0.00	0.61
8	FWC, d_g , σ_g	0.19	0.26	0.27	0.11	0.64	0.14	0.00	0.64
9	FWC, d_g , σ_g , OC	0.001	0.38	0.18	-0.85	0.81	0.11	0.00	0.81
10	FWC, TP	0.43	0.21	0.00	0.41	0.57	0.16	0.00	0.57
11	FWC, TP, d_g	0.32	0.23	0.19	0.30	0.72	0.13	0.00	0.72
12	TP, Sand	0.29	0.24	0.38	0.23	0.49	0.17	0.00	0.49
13	TP, Sand, PD	0.02	0.44	0.72	-1.52	0.72	0.13	0.00	0.72
14	Sand, OC	0.30	0.23	0.50	0.30	0.38	0.19	0.00	0.38
15	Sand, OC, Clay	0.11	0.29	0.08	-0.11	0.60	0.15	0.00	0.60

NS: ضریب نش-ساتکلیف، NRMSE: ریشه میانگین مربعات خطای نرمال شده، ME: میانگین خطا، R^2 : ضریب تبیین. توصیف متغیرها در زیر جدول ۱ آورده شده است.

R^2 : coefficient of determination; NRMSE: normalized root mean square error; ME: mean error; NS: Nash-Sutcliffe coefficient.

Variables description is available under Table 1.

همین مدل در روش رگرسیون خطی چندگانه ($ME = 0.40 \text{ MPa}$)، به مراتب دارای بیش‌برآوردی کمتری است؛ معادله رگرسیونی این مدل در **جدول ۳** نیز بیانگر آن است که جرم مخصوص ظاهری با علامت مثبت و شن با علامت منفی وارد مدل گردیدند و اهمیت متغیر جرم مخصوص ظاهری در برآورد PR به علت داشتن ضریب رگرسیون بالاتر به مراتب بیشتر از متغیر شن می‌باشد. در درجات بعدی، به‌ترتیب مدل GEP شماره ۶ با چهار متغیر ورودی شامل جرم مخصوص ظاهری، سیلت، رطوبت خاک مزرعه و انحراف معیار هندسی ذرات و مدل GEP شماره ۹ با چهار متغیر ورودی شامل رطوبت خاک مزرعه، کربن آلی، میانگین قطر و انحراف معیار هندسی ذرات خاک قرار دارد (**جدول ۶**). اصغری و همکاران (*Asghari et al., 2024*) نیز گزارش نمودند مدل GEP با متغیرهای ورودی رطوبت خاک مزرعه، جرم مخصوص ظاهری نسبی و سیلت توانست PR خاک را در اراضی جنگلی، مرتعی و زراعی منطقه فندقلوی اردبیل با دقت بیشتری نسبت به سایر مدل‌های GEP برآورد نماید و مقادیر R^2 ، RMSE، ME و NS مدل مذکور به‌ترتیب برابر با ۰/۷۹، ۰/۹۱ MPa، ۰/۱۳ MPa و ۰/۶۳ در داده‌های آزمونی به‌دست آمد.

رتبه‌بندی مدل‌های GEP نیز همانند مدل NF (**جدول ۵**)، بر اساس پایین بودن NRMSE و بالا بودن NS انجام شد. بر مبنای آماره‌های مذکور در داده‌های آزمونی (**جدول ۶**)، دقیق‌ترین مدل GEP برای برآورد PR، مدل شماره ۱ با دو متغیر ورودی شامل جرم مخصوص ظاهری و شن از میان ۱۵ مدل GEP تعیین گردید. معادله ۴، رابطه ریاضی به‌دست آمده برای محاسبه PR توسط مدل GEP شماره ۱ را نشان می‌دهد:

$$PR = \tan(\cot(\log 4.898BD) + \tan(0.69)) + \left(\frac{-1.461sand^2}{BD} - sand^2 \right)^2 + \tan(\tan(\cot(sand))) \quad (4)$$

مقادیر آماره‌های R^2 ، NRMSE، ME و NS برای این مدل بر اساس داده‌های آموزشی به‌ترتیب برابر ۰/۵۰، ۰/۱۷، ۰/۰۱ MPa و ۰/۵۰ و داده‌های آزمونی برابر ۰/۵۱، ۰/۲۰، ۰/۱۳ MPa و ۰/۴۸ به دست آمد (**جدول ۶**) و این مدل دارای بالاترین ضریب تبیین در مقایسه با سایر مدل‌های GEP بود؛ اگرچه مقدار مثبت ME در داده‌های آزمونی (۰/۱۳ مگاپاسکال) بیانگر اندکی بیش‌برآوردی در تخمین PR خاک توسط مدل GEP شماره ۱ می‌باشد ولی در مقایسه با مقدار گزارش شده توسط شیخ‌زاده و همکاران (*Sheykhzadeh et al., 2016*) برای

جدول ۶- معیارهای ارزیابی مدل‌های برنامه‌ریزی بیان ژن (GEP) مقاومت فروری (PR) خاک
Table 6- Evaluation criteria for gene expression programming (GEP) models of PR

شماره Number	Input variables متغیرهای ورودی	R ²	آزمون Testing set			آموزش Training			
			NRMSE	ME (MPa)	NS	R ²	NRMSE	ME (MPa)	NS
1	BD, Sand	0.51	0.20	0.13	0.48	0.50	0.17	-0.01	0.50
2	BD, Sand, FWC	0.36	0.25	0.04	0.19	0.58	0.15	-0.02	0.58
3	BD, Sand, FWC, σ _g	0.38	0.23	0.25	0.31	0.58	0.16	0.02	0.57
4	BD, Silt	0.001	0.40	0.51	-1.03	0.50	0.17	-0.01	0.50
5	BD, Silt, FWC	0.34	0.24	0.07	0.26	0.55	0.16	0.01	0.55
6	BD, Silt, FWC, σ _g	0.48	0.21	0.27	0.45	0.60	0.15	0.02	0.60
7	FWC, d _g	0.33	0.23	0.33	0.28	0.56	0.16	-0.01	0.56
8	FWC, d _g , σ _g	0.34	0.24	0.45	0.24	0.61	0.15	0.03	0.61
9	FWC, d _g , σ _g , OC	0.49	0.21	0.41	0.41	0.58	0.15	0.03	0.58
10	FWC, TP	0.33	0.43	-0.23	-1.35	0.58	0.15	-0.04	0.58
11	FWC, TP, d _g	0.38	0.33	0.07	-0.36	0.62	0.15	0.01	0.62
12	TP, Sand	0.08	0.28	0.21	-0.05	0.50	0.17	0.13	0.49
13	TP, Sand, PD	0.34	0.23	0.22	0.30	0.45	0.18	-0.02	0.45
14	Sand, OC	0.22	0.25	0.21	0.19	0.31	0.20	-0.01	0.31
15	Sand, OC, Clay	0.20	0.25	0.29	0.17	0.45	0.18	0.01	0.45

NS: ضریب نش-ساتکلِف، NRMSE: ریشه میانگین مربعات خطای نرمال شده، ME: میانگین خطا، R^2 : ضریب تبیین. توصیف متغیرها در زیر **جدول ۱** آورده شده است.
 R^2 : coefficient of determination; NRMSE: normalized root mean square error; ME: mean error; NS: Nash-Sutcliffe coefficient.
Variables description is available under Table 1.

ارزیابی مدل‌های جنگل تصادفی (RF) در برآورد مقاومت فروروی (PR) خاک

پانزده مدل جنگل تصادفی (RF) با ترکیب متغیرهای ورودی به کار رفته در مدل‌های MLR (جدول ۳)، NF (جدول ۵) و GEP (جدول ۶)، برای تخمین مقاومت فروروی (PR) خاک طراحی گردید. مقادیر آماره‌های ضریب تبیین (R^2)، ریشه میانگین مربعات خطای نرمال شده (NRMSE)، میانگین خطا (ME) و ضریب نش-ساتکلیف (NS) برای ارزیابی عملکرد مدل‌های RF از نظر دقت برآورد PR در جدول ۷ آورده شده است. برای درجه‌بندی و مقایسه دقت مدل‌های RF در تخمین PR، همانند دو مدل هوشمند قبلی از معیار پایین بودن آماره NRMSE و بالا بودن آماره NS استفاده شد؛ بر این اساس، مدل RF شماره ۱۱ در داده‌های آزمونی (جدول ۷) با سه متغیر ورودی شامل رطوبت خاک، مزرعه، تخلخل کل و میانگین هندسی قطر ذرات خاک توانست PR را با خطای پایین و دقت بالا در مقایسه با سایر مدل‌های RF برآورد

نماید؛ مقادیر R^2 ، NRMSE، ME و NS این مدل برای داده‌های آموزشی به ترتیب برابر ۰/۹۴، ۰/۰۶، ۰/۰۱ MPa و ۰/۹۳ و برای داده‌های آزمونی به ترتیب برابر ۰/۵۸، ۰/۲۰، ۰/۳۱ MPa و ۰/۵۰ به دست آمد (جدول ۷). مطابق آماره‌های NRMSE و NS در سری آزمون، عملکرد مدل RF شماره ۱۱ به ترتیب با کارکرد مناسب و قابل قبول ارزیابی می‌شود (Waseem et al., 2017). معادله رگرسیونی مدل ۱۱ در جدول ۳ نیز بیانگر آن است که هر سه متغیر ورودی با علامت منفی وارد مدل گردیدند و اهمیت متغیر میانگین هندسی قطر ذرات خاک در برآورد PR به علت داشتن ضریب رگرسیون بالاتر به مراتب بیشتر از متغیرهای رطوبت خاک مزرعه و تخلخل کل می‌باشد. در درجات بعدی، به ترتیب مدل RF شماره ۷ (NRMSE = 0.20 و NS = 0.47) با دو متغیر ورودی شامل رطوبت خاک مزرعه و میانگین قطر هندسی ذرات خاک و مدل RF شماره ۹ (NRMSE = 0.20 و NS = 0.47) با چهار متغیر ورودی شامل رطوبت خاک مزرعه، کربن آلی، میانگین قطر و انحراف معیار هندسی ذرات قرار دارد (جدول ۷).

جدول ۷- معیارهای ارزیابی مدل‌های جنگل تصادفی (RF) مقاومت فروروی (PR) خاک

Table 7- Evaluation criteria for random forest (RF) models of PR

شماره Number	Input variables متغیرهای ورودی	آزمون Testing set				آموزش Training			
		R^2	NRMSE	ME (MPa)	NS	R^2	NRMSE	ME (MPa)	NS
1	BD, Sand	0.26	0.25	0.11	0.21	0.94	0.07	-0.03	0.92
2	BD, Sand, FWC	0.46	0.21	0.18	0.44	0.94	0.06	-0.01	0.93
3	BD, Sand, FWC, σ_g	0.43	0.29	0.28	-0.04	0.95	0.06	-0.02	0.94
4	BD, Silt	0.33	0.24	0.08	0.30	0.94	0.07	-0.02	0.93
5	BD, Silt, FWC	0.43	0.22	0.18	0.41	0.94	0.06	-0.02	0.93
6	BD, Silt, FWC, σ_g	0.41	0.22	0.25	0.38	0.95	0.06	-0.01	0.94
7	FWC, d_g	0.51	0.20	0.27	0.47	0.93	0.07	-0.02	0.92
8	FWC, d_g , σ_g	0.45	0.22	0.35	0.39	0.94	0.06	-0.04	0.93
9	FWC, d_g , σ_g , OC	0.53	0.20	0.26	0.47	0.95	0.06	-0.04	0.94
10	FWC, TP	0.42	0.23	0.38	0.35	0.93	0.07	-0.01	0.92
11	FWC, TP, d_g	0.58	0.20	0.31	0.50	0.94	0.06	-0.01	0.93
12	TP, Sand	0.33	0.24	0.30	0.29	0.92	0.08	-0.03	0.90
13	TP, Sand, PD	0.33	0.24	0.27	0.29	0.94	0.07	-0.02	0.92
14	Sand, OC	0.08	0.28	0.30	0.01	0.89	0.09	-0.06	0.87
15	Sand, OC, Clay	0.27	0.25	0.30	0.23	0.92	0.08	-0.02	0.90

NS: ضریب نش-ساتکلیف، NRMSE: ریشه میانگین مربعات خطای نرمال شده، ME: میانگین خطا، R^2 : ضریب تبیین. توصیف متغیرها در زیر جدول ۱ آورده شده است.

R^2 : coefficient of determination; NRMSE: normalized root mean square error; ME: mean error; NS: Nash-Sutcliffe coefficient. Variables description is available under Table 1.

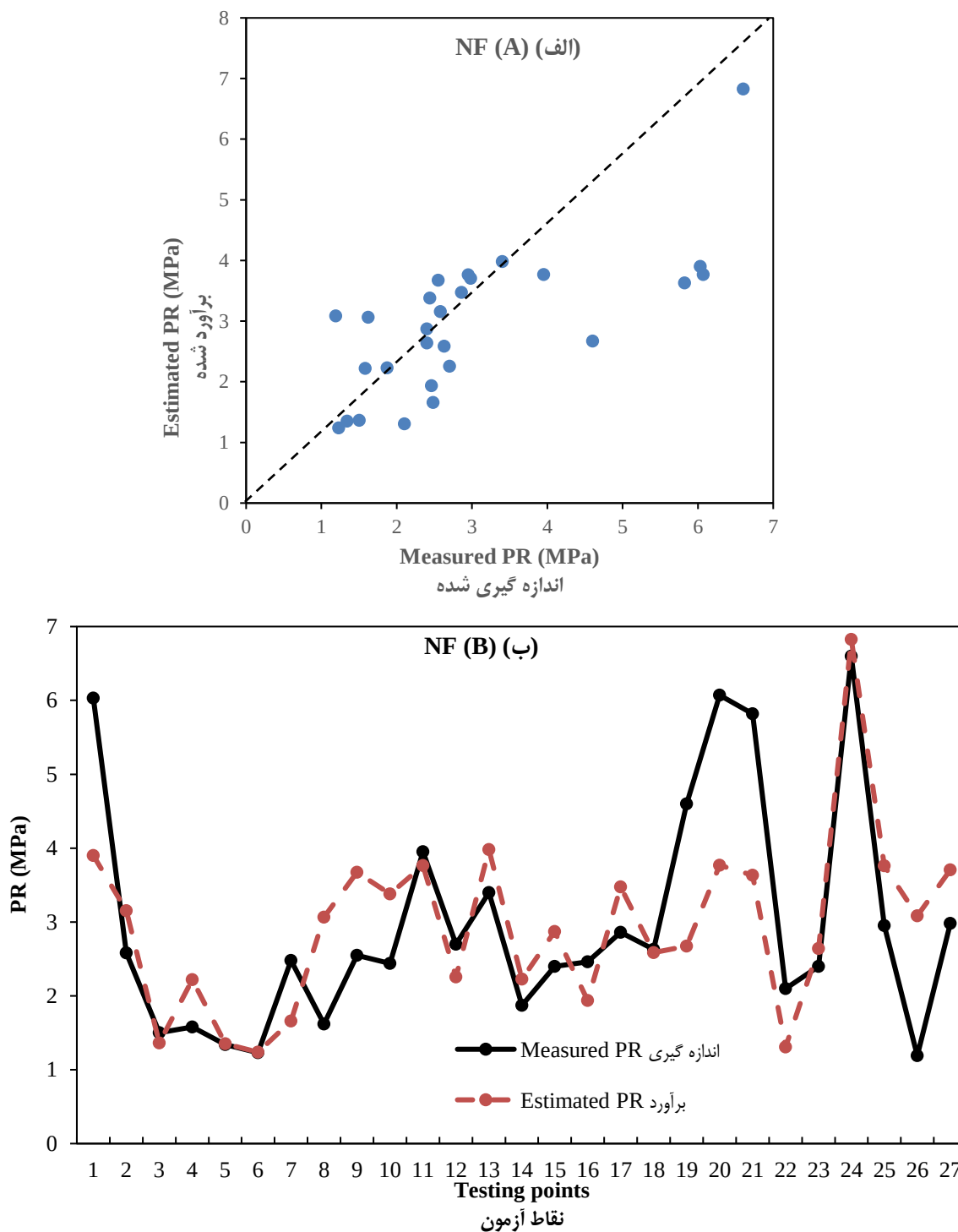
شکل ۳ (الف)، شکل ۴ (الف) و شکل ۵ (الف) به ترتیب پراکنش داده های اندازه‌گیری و برآورد شده PR حول محور ۱:۱ بر اساس بهترین مدل NF (مدل شماره ۲ در جدول ۵)، بهترین مدل GEP (مدل شماره ۱ در جدول ۶) و بهترین مدل RF (مدل شماره ۱۱ در جدول ۷) در سری آزمون نشان داده شده است. با توجه به این نمودارها مشاهده می‌شود که برای مقادیر بزرگتر مقاومت فروروی خاک، خطای (اختلاف مقادیر اندازه‌گیری و برآورد شده) هر سه روش هوشمند به کار رفته در پژوهش، بالاتر است؛ شاید دلیل این قضیه را بتوان به تعداد اندک داده های با مقادیر بزرگ PR در این تحقیق نسبت داد که موجب گردید آموزش مدل‌ها در این محدوده ($PR > 4 \text{ MPa}$) به خوبی انجام نگیرد. مقادیر بزرگ PR معمولاً در رطوبت‌های پایین و یا فشردگی بالای خاک اتفاق می‌افتد که ممکن است باعث بروز خطای ابزاری در سنجش میزان مقاومت فروروی خاک نیز گردد. همچنین مشاهده می‌شود که پراکندگی داده‌ها اطراف خط ۱:۱ در بهترین مدل NF نسبت به بهترین مدل‌های GEP و RF کمتر است. بنابراین می‌توان چنین استنباط نمود که در پژوهش حاضر، مدل نروفازی در برآورد PR از روی متغیرهای زود یافت خاک به علت استفاده از روابط غیر خطی پیچیده، در مقایسه با دو مدل هوشمند دیگر نسبتاً دقیق‌تر عمل نموده است. حسینی و همکاران (Hosseini et al., 2018) نیز به برتری نسبی روش نروفازی در مقایسه با سایر روش‌های هوشمند در برآورد مقاومت فروروی خاک دست یافتند. اصغری و همکاران (Asghari et al., 2024) نیز گزارش نمودند روش شبکه عصبی مصنوعی به علت به کارگیری روابط غیرخطی پیچیده در مقایسه با مدل هوشمند برنامه‌ریزی بیان ژن و مدل رگرسیون سستی توانست مقاومت فروروی خاک را در منطقه فندقلوی اردبیل با دقت بالایی برآورد نماید. با عنایت به اندک بودن تعداد تحقیقات انجام شده در زمینه به کارگیری مدل‌های هوشمند در برآورد مقاومت فروروی خاک، امکان مقایسه دقیق نتایج میسر نشد.

به طور کلی با توجه به نتایج جداول ۳، ۵، ۶ و ۷ می‌توان گفت که متغیرهای ورودی BD، FWC، Sand، Silt، TP و dg از مهمترین متغیرهای زود یافت در افزایش کارایی همه مدل‌ها از نظر برآورد مقاومت فروروی خاک در منطقه مورد مطالعه هستند؛ در مقابل، متغیرهای ورودی OC، PD، Clay و σ_g اهمیت کمتری از خود نشان داده‌اند. این نتایج با یافته‌های محققان قبلی مانند بیات و همکاران (Bayat et al., 2008) و اصغری و همکاران (Asghari et al., 2024) در یک راستا می‌باشد.

مقایسه نتایج نشان می‌دهد که بهترین مدل ($R^2 = 0.58$) در روش هوشمند جنگل تصادفی (مدل شماره ۱۱ در جدول ۷) در تحقیق حاضر نسبت به همان مدل ($R^2 = 0.49$) در روش رگرسیون سنتی (Sheykhzadeh et al., 2016) از دقت بالاتری برخوردار است همچنین براساس آماره ME، روش جنگل تصادفی ($ME = 0.31$) در مقایسه با روش رگرسیون خطی چندگانه ($ME = 0.34$) دارای بیش‌برآوردی کمتری می‌باشد.

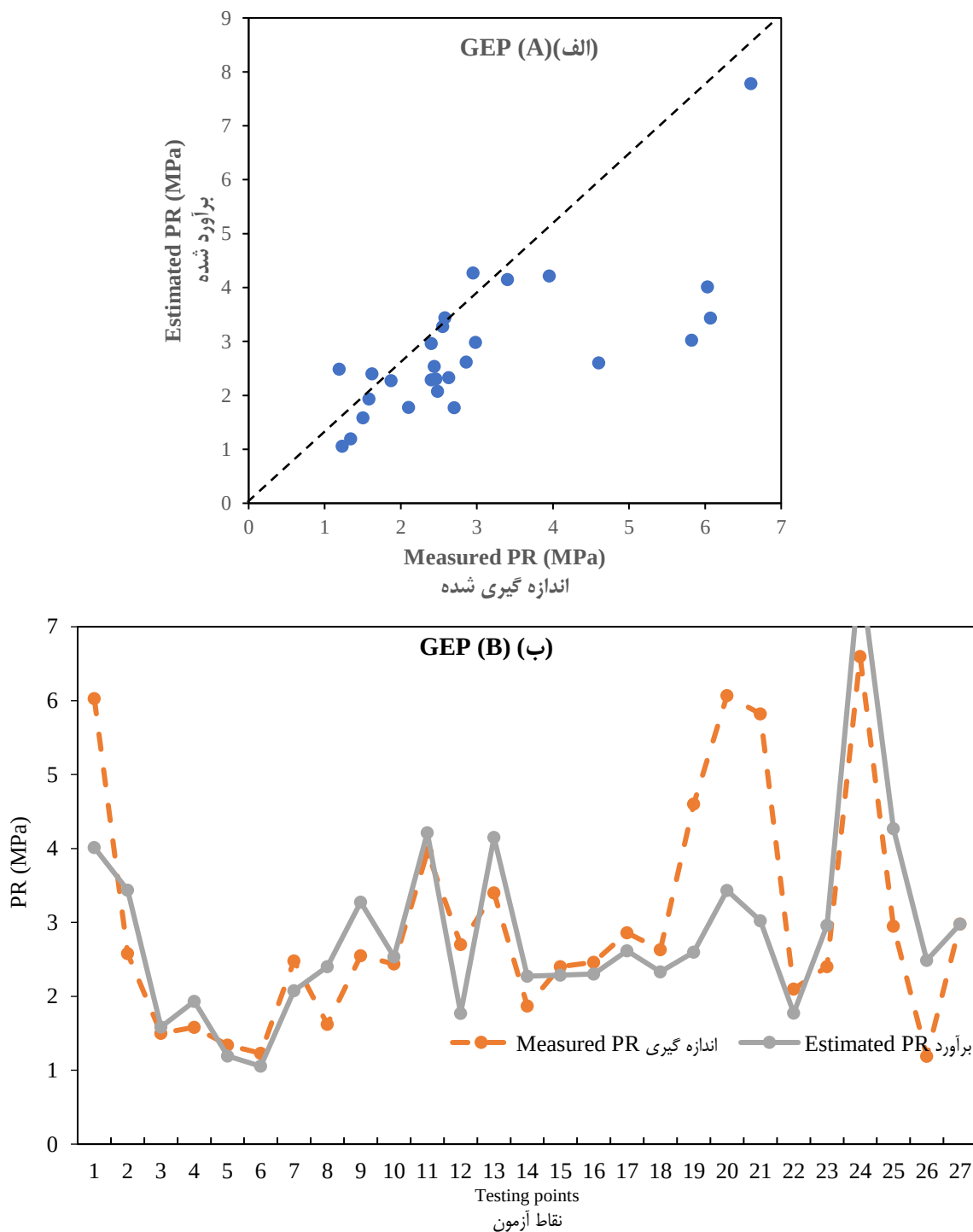
مقایسه عملکرد مدل‌های نروفازی، برنامه‌ریزی بیان ژن و جنگل تصادفی در برآورد مقاومت فروروی خاک

مقادیر اندازه‌گیری و برآورد شده مقاومت فروروی (PR) خاک بر اساس ۲۷ داده در مرحله آزمون برای بهترین مدل نروفازی (NF)، برنامه‌ریزی بیان ژن (GEP) و جنگل تصادفی (RF) به ترتیب در شکل ۳ (ب)، شکل ۴ (ب) و شکل ۵ (ب) به طور مقایسه‌ای آورده شده است. با توجه به مقادیر آماره‌های ارزیابی ریشه میانگین مربعات خطای نرمال شده (NRMSE) و ضریب نش-ساتکیف (NS) برای بهترین مدل NF (مدل شماره ۲ در جدول ۵) با سه متغیر ورودی شامل جرم مخصوص ظاهری، شن و رطوبت خاک مزرعه، بهترین مدل GEP (مدل شماره ۱ در جدول ۶) با دو متغیر ورودی شامل جرم مخصوص ظاهری و شن و بهترین مدل RF (مدل شماره ۱۱ در جدول ۷) با سه متغیر ورودی شامل رطوبت خاک مزرعه، تخلخل کل و میانگین هندسی قطر ذرات خاک، از نظر برآورد PR در سری آزمون مشاهده می‌شود که بهترین مدل NF به دلیل داشتن مقدار NS بالا (۰/۵۱) و NRMSE پایین (۰/۱۹) در مقایسه با بهترین مدل GEP ($NRMSE = 0.20$ و $NS = 0.48$) و بهترین مدل RF ($NRMSE = 0.20$ و $NS = 0.50$) دارای دقت بیشتر و خطای کمتر در برآورد PR می‌باشد. همچنین می‌توان گفت بهترین مدل RF نسبت به بهترین مدل GEP (به دلیل داشتن NS نسبتاً بالاتر) قادر است PR را با دقت مناسبی تخمین بزند. مطابق درجه‌بندی واسیم و همکاران (Waseem et al., 2017) و وان و همکاران (Wan et al., 2023) می‌توان بیان نمود که در این پژوهش در بین مدل‌های پانزده‌گانه NF (جدول ۵)، GEP (جدول ۶) و RF (جدول ۷)، مدل شماره ۳ (شامل ۴ متغیر ورودی رطوبت خاک مزرعه، جرم مخصوص ظاهری، شن و میانگین هندسی قطر ذرات خاک) در روش NF به دلیل داشتن کوچکترین مقدار آماره NS و بالاترین مقدار آماره NRMSE، دارای ضعیف‌ترین عملکرد در برآورد مقاومت فروروی (PR) خاک در اراضی دشت اردبیل می‌باشد. در



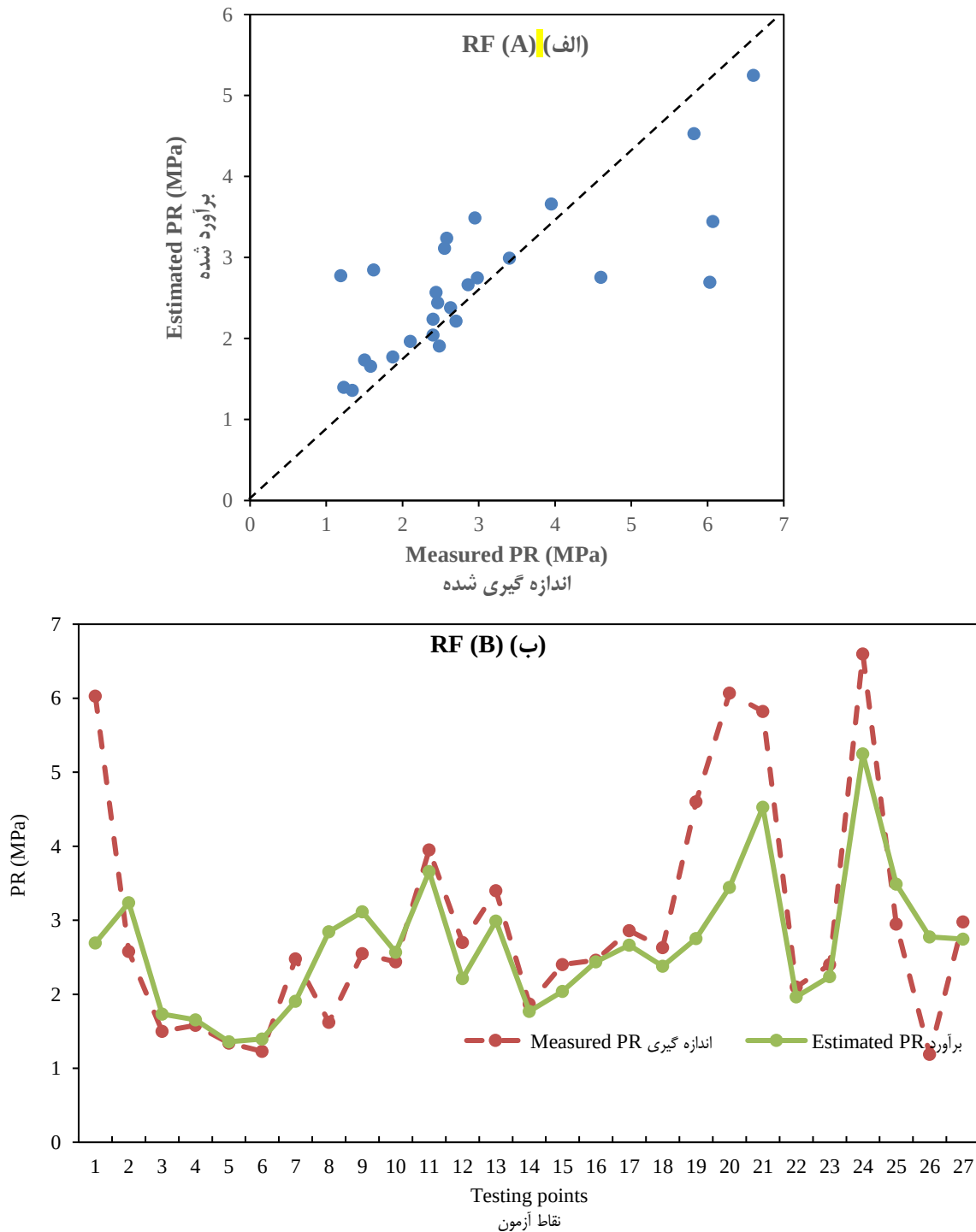
شکل ۳- پراکنش حول محور ۱:۱ (الف) و مقایسه (ب) مقادیر اندازه‌گیری و برآورد شده مقاومت فروروی خاک (PR) با استفاده از بهترین مدل نروفازی (NF) بر اساس داده‌های آزمونی (n=27)
اطلاعات مربوط به مدل و متغیرهای ورودی در جدول ۵ آورده شده است.

Figure 3- Distribution around the 1:1 axis (A) and comparison (B) of estimated and measured values of soil penetration resistance (PR) using the best NF model based on testing data (n=27)
Table 5 provides detailed information about model and input variables.



شکل ۴- پراکنش حول محور ۱:۱ (الف) و مقایسه (ب) مقادیر اندازه‌گیری و برآورد شده مقاومت فروروی خاک (PR) با استفاده از بهترین مدل برنامه‌ریزی بیان ژن (GEP) بر اساس داده‌های آزمونی (n=27)
اطلاعات مربوط به مدل و متغیرهای ورودی در جدول ۵ آورده شده است.

Figure 4- Distribution around the 1:1 axis (A) and comparison (B) of estimated and measured values of soil penetration resistance (PR) using the best gene expression programming (GEP) model based on testing data (n=27)
Table 5 provides detailed information about model and input variables.

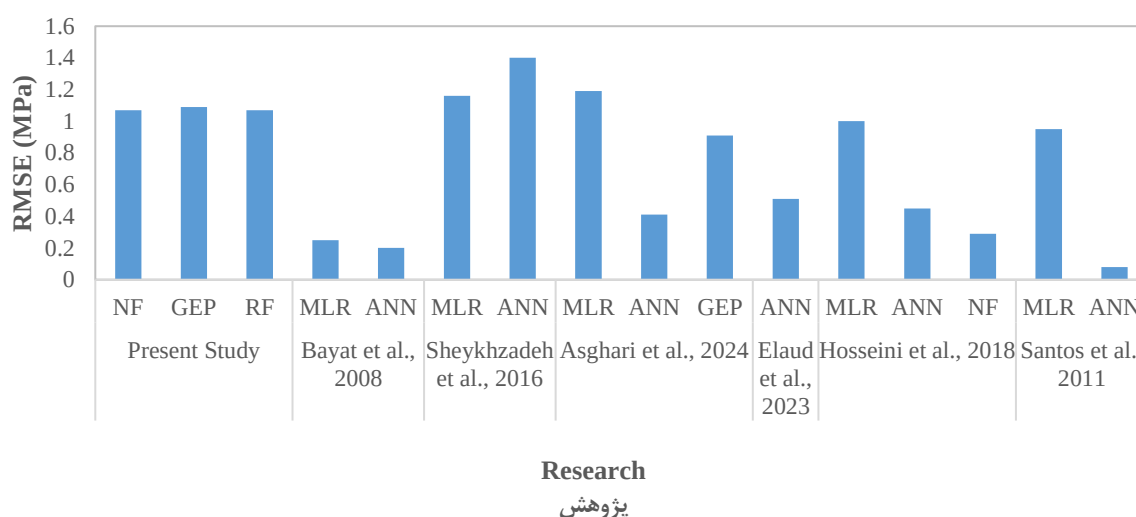


شکل ۵- پراکنش حول محور ۱:۱ (الف) و مقایسه (ب) مقادیر اندازه‌گیری و برآورد شده مقاومت فروروی خاک (PR) با استفاده از بهترین مدل جنگل تصادفی (RF) بر اساس داده‌های آزمونی (n=27)

اطلاعات مربوط به مدل و متغیرهای ورودی در جدول ۵ آورده شده است.

Figure 5- Distribution around the 1:1 axis (A) and comparison (B) of estimated and measured values of soil penetration resistance (PR) using the best random forest (RF) model based on testing data (n=27)

Table 5 provides detailed information about model and input variables.



شکل ۶- مقایسه مقادیر ریشه میانگین مربعات خطا (RMSE) در چند پژوهش گذشته با مطالعه حاضر برای برآورد مقاومت فروروی خاک (PR) با استفاده از مدل‌های مختلف پژوهش

Figure 6- Comparison of root mean square error (RMSE) values in some pervious researches with present study for estimating soil penetration resistance (PR) using different models

متغیرهای ورودی ۱۵ مدل رگرسیون خطی چندگانه ایجاد شده توسط شیخزاده و همکاران (Sheykhzadeh et al., 2016) شامل شن، سیلت، رس، میانگین و انحراف معیار هندسی قطر ذرات، جرم مخصوص ظاهری و حقیقی، تخلخل کل، رطوبت خاک مزرعه و کربن آلی خاک به کار گرفته شد. نتایج نشان داد که مقاومت فروروی (PR) دارای همبستگی مثبت و معنی‌دار با جرم مخصوص ظاهری و سیلت و همبستگی منفی و معنی‌دار با رطوبت خاک مزرعه و کربن آلی می‌باشد. ارزیابی و مقایسه مدل‌ها بر پایه آماره RMSE در داده‌های آزمونی نشان داد که هر سه روش هوشمند (مقادیر RMSE از ۱/۰۷ تا ۱/۰۹ مگاپاسکال) به کار رفته در این پژوهش در مقایسه با روش رگرسیون خطی چندگانه (مقدار RMSE برابر با ۱/۱۶ مگاپاسکال) مورد استفاده توسط شیخزاده و همکاران (Sheykhzadeh et al., 2016) توانستند مقاومت فروروی (PR) خاک را با خطای نسبتاً کمتری برآورد نمایند. نتایج بهترین مدل‌ها در هر سه روش هوشمند نیز بیانگر آن است که جرم مخصوص ظاهری، شن، رطوبت خاک مزرعه و میانگین هندسی قطر ذرات خاک، مناسب‌ترین متغیرهای ورودی در برآورد PR خاک در منطقه مورد مطالعه هستند. همچنین جرم مخصوص ظاهری به علت داشتن بیشترین مقدار ضریب رگرسیون، از اهمیت نسبی بالاتری نسبت به سایر متغیرهای ورودی در تخمین مقاومت فروروی خاک منطقه مورد مطالعه برخوردار می‌باشد. در بین سه مدل هوشمند مورد استفاده، مدل NF به علت داشتن آماره نش ساتکلیف (NS) بالاتر و ریشه میانگین مربعات خطای نرمال شده (NRMSE) پایین‌تر، دارای دقت بیشتری در برآورد PR می‌باشد. همچنین مدل RF در مقایسه با مدل GEP توانست عملکرد نسبتاً بهتری در برآورد PR از خود نشان دهد.

با توجه به شکل ۶ مشاهده می‌شود میزان خطای مدل‌های رگرسیونی و هوشمند به کار رفته در پژوهش‌های مختلف برای برآورد مقاومت فروروی (PR) خاک به علت تعداد متفاوت نمونه‌ها، شرایط اقلیمی و توپوگرافیکی منطقه مورد مطالعه و نیز خطای ابزاری در سنجش مقاومت فروروی خاک یکسان نمی‌باشد. مقایسه مقادیر RMSE (براساس داده‌های مرحله آزمون) در برآورد مقاومت فروروی (PR) خاک نشان می‌دهد (شکل ۶) که بهترین مدل نروفازی (NF)، برنامه‌ریزی بیان ژن (GEP) و جنگل تصادفی (RF) به کار رفته در تحقیق حاضر در مقایسه با مدل رگرسیونی خطی چندگانه (MLR) در پژوهش‌های شیخزاده و همکاران (Sheykhzadeh et al., 2016) و نیز اصغری و همکاران (Asghari et al., 2024) از خطای کمتری برخوردار بوده و مناسب‌تر است در حالی که هر سه مدل هوشمند به کار رفته در پژوهش حاضر نسبت به مدل‌های رگرسیونی و هوشمند استفاده شده در سایر پژوهش‌ها مانند بیات و همکاران (Bayat et al., 2008)، سانتوس و همکاران (Santos et al., 2011)، حسینی و همکاران (Hosseini et al., 2018) و الایود و همکاران (Elaoud et al., 2023) دارای خطای (RMSE) بالاتری از نظر برآورد PR بوده و مناسب‌تر نیستند.

نتیجه‌گیری

در پژوهش حاضر از سه مدل هوشمند نروفازی (NF)، برنامه‌ریزی بیان ژن (GEP) و جنگل تصادفی (RF) برای برآورد مقاومت فروروی (PR) خاک‌های زراعی دشت اردبیل استفاده گردید. برای این منظور،

پیشنهاد می‌شود در پژوهش‌های بعدی از سایر مدل‌های مبتنی بر یادگیری ماشین مانند یادگیری عمیق و ماشین بردار پشتیبان نیز در برآورد PR در منطقه مورد مطالعه استفاده گردد.

References

- 1- Asghari, Sh., Hasanpour Kashani, M., & Shahab Arkhazloo, H. (2024). Modeling soil penetration resistance using regression, artificial neural network and gene expression programming. *Journal of Soil and Water*, 38(2), 269-283. (In Persian). <https://doi.org/10.22067/JSW.2024.86792.1385>
- 2- Bayat, H., Neyshabouri, M.R., & Hajabbasi, M. (2008). Comparing neural networks, linear and nonlinear regression techniques to model penetration resistance. *Turkish Journal of Agriculture and Forestry*, 32, 425-433.
- 3- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.
- 4- Elaoud, A., Hassen, H.B., Jalel, R., Salah, N.B., Masmoudi, A., & Masmoudi, A. (2023). Machine learning approach for predicting soil penetration resistance under different moisture conditions. *Journal of Terramechanics*, 110, 39-45. <https://doi.org/10.1016/j.jterra.2023.08.002>
- 5- Ferreira, C. (2006). Gene Expression Programming: Mathematical modeling by an artificial intelligence. Springer, Berlin, Heidelberg, New York, p. 478.
- 6- Ghorbani, M.A., Deo, R.C., Kashani, M.H., Shahabi, M., & Ghorbani, S. (2019). Artificial intelligence-based fast and efficient hybrid approach for spatial modeling of soil electrical conductivity. *Soil and Tillage Research*, 186, 152-164. <https://doi.org/10.1016/j.still.2018.09.012>
- 7- Hosseini, M., Movahedi Naeini, S.A.R., Dehghani, A.A., & Zeraatpisheh, M. (2018). Modeling of soil mechanical resistance using intelligent methods. *Journal of Soil Science and Plant Nutrition*, 18(4), 939-951. <https://doi.org/10.4067/s0718-95162018005002702>
- 8- Kozak, E., Pachepsky, Y.A., Sokolowski, S., Sokolowska, Z., & Stepniewski, W. (1996). A modified number-based method for estimating fragmentation fractal dimensions of soils. *Soil Science Society of America Journal*, 60, 1291-1297. <https://doi.org/10.2136/sssaj1996.03615995006000050002x>
- 9- Lowery, B., & Morrison, J.E. (2002). *Soil penetrometer and penetrability*. In: Dane J.H., & Topp GC (eds.). *Methods of soil analysis, part 4. Physical methods*. Madison (WI): Soil Science Society of America, 363-388. <https://doi.org/10.2136/sssabookser5.4.c16>
- 10- Parsaie, F., Farrokhan Firouzi, A., Davari, M., & Taghizadeh-Mehrjardi, R. (2025). Digital mapping of soil penetration resistance and shear strength using machine learning algorithms in the Kilane watershed, Kurdistan province. *Journal of Range & Watershed Management*, 78(1), 124-143. <https://doi.org/10.22059/jrwm.2024.381733.1783>
- 11- Santos, F.L., De Jesus, V.A.M., & Valente, D.S.M. (2012). Modeling of soil penetration resistance using statistical analyses and artificial neural networks. *Acta Scientiarum. Agronomy*, 34, 219-224. <https://doi.org/10.4025/actasciagron.v34i2.11627>
- 12- Sheykhzadeh, G.R., Asghari, Sh., & Mesri Gundoshmian, T. (2016). Estimating penetration resistance in agricultural soils of Ardabil plain using artificial neural network and regression methods. *Water and Soil*, 30(3), 941-954. (In Persian). <https://doi.org/10.22067/JSW.V30I3.42235>
- 13- Shiri, J., Keshavarzi, A., Kisi, O., & Karimi, S. (2017). Using soil easily measured parameters for estimating soil water capacity: Soft computing approaches. *Computers and Electronics in Agriculture*, 141, 327-339. <https://doi.org/10.1016/j.compag.2017.08.012>
- 14- Singh, V.K., Kumar, D., Kashyap, P.S., Singh, P.K., Kumar, A., & Singh, S.K. (2020). Modelling of soil permeability using different data driven algorithms based on physical properties of soil. *Journal of Hydrology*, 580(124223). <https://doi.org/10.1016/j.jhydrol.2019.124223>
- 15- Unal, I., Kabaş, O., & Sozer, S. (2024). Comparison of two different artificial neural network models for prediction of soil penetration resistance. *Journal of Agricultural Engineering*, 55(1), 79-86. <https://doi.org/10.4081/jae.2023.1550>
- 16- Warrick, A.W. (2002). *Soil Physics Companion*. CRC Press. 247p.
- 17- Wan, H., Qi, H., & Shang, S. (2023). Estimating soil water and salt contents from field measurements with time domain reflectometry using machine learning algorithms. *Agricultural Water Management*, 285, 108364. <https://doi.org/10.1016/j.agwat.2023.108364>
- 18- Waseem, M., Mani, N., Andiego, G., & Usman, M. (2017). A review of criteria of fit for hydrological models. *International Research Journal of Engineering and Technology*, 4(11), 1765-1772.