

Research Article

Prediction of Copper and Manganese Concentrations in Citrus Leaves Based on Easily Measured Soil Characteristics: Comparison of Stepwise Regression Models and Gene Expression Programming

Saber Heidari¹, Seyed Ali Ghaffari Nejad², Javad Sarhadi³

- 1-Soil and Water Research Department, South Kerman Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization (AREEO), Jiroft, Iran. E-mail: s.heydari@areeo.ac.ir
- 2-Soil and Water Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. E-mail: ma_ghaffari51@yahoo.com
- 3-Corresponding Author, Soil and Water Research Department, South Kerman Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization (AREEO), Jiroft, Iran. E-mail: javad.sarhadi2009@gmail.com

Received: November 25, 2025 **Revised:** January 06, 2026
Accepted: January 06, 2026 **Published:** January 07, 2026

Extended Abstract

Background and Objectives

The availability of micronutrients, particularly copper (Cu) and manganese (Mn), is critical for sustainable citrus production. Cu serves as a vital component of enzymes involved in photosynthesis and respiration, while Mn is crucial for enzyme activation and chlorophyll synthesis. Deficiencies in these nutrients can severely compromise plant vitality and yield, posing a significant economic threat to agricultural regions. This challenge is particularly acute in arid areas like southern Kerman, Iran, where the calcareous soils (characterized by high pH and low organic matter content) severely limit the bioavailability of Cu and Mn. Consequently, there is an urgent need to develop accurate predictive methods for assessing the nutritional status of citrus trees to guide precision fertilization strategies. Traditionally, researchers have relied on linear statistical methods, such as Stepwise Regression (SWR), to model the relationship between soil properties and nutrient uptake. However, these models are fundamentally limited by their assumption of linearity, whereas soil-plant systems are inherently non-linear and interactive. This limitation has prompted a shift toward more sophisticated machine learning approaches, such as Gene Expression Programming (GEP). GEP is an evolutionary algorithm that offers a unique advantage by producing explicit, transparent mathematical equations, thereby combining high predictive accuracy with interpretability. This study aims to fill a critical research gap by directly comparing the performance of traditional stepwise regression against GEP for predicting citrus leaf Cu and Mn concentrations under the specific soil conditions of southern Kerman. The central hypothesis was that the non-linear GEP model would significantly outperform its linear counterpart.

Materials and Method

This study was conducted across 40 commercial Valencia orange orchards in southern Kerman, Iran. From each orchard, composite soil samples were collected from two depths (0–30 cm and 31–60 cm), alongside leaf samples from 4- to 6-month-old spring shoots. In the laboratory, leaf Cu and Mn concentrations were quantified using atomic absorption spectrometry. Soil samples were analyzed for key physicochemical properties using standard protocols, including texture, pH, electrical conductivity (EC), organic carbon (OC), total neutralizing value (TNV), and available phosphorus (P). Two distinct modeling approaches were developed and compared: (1) Stepwise Multiple Linear Regression (SWR) was employed to generate linear predictive models, and (2) GEP was implemented using GeneXproTools software to generate non-linear models. To ensure robust model validation,

the full dataset (n=40) was randomly partitioned into a training set (70%, 28 samples) and a testing set (30%, 12 samples). The performance of all developed models was rigorously evaluated using the Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE).

Results

Descriptive statistics showed that the soils were predominantly neutral to alkaline (mean pH 7.75). Pearson correlation analysis revealed that citrus leaf Cu concentration was negatively correlated with soil pH ($r=-0.41^*$) and positively with OC ($r=0.39^*$). Similarly, citrus leaf Mn concentration showed negative correlations with both pH ($r=-0.33^*$) and EC ($r=-0.29^*$).

The SWR models yielded limited predictive accuracy on the test data. For leaf Cu, the SWR model (predictors: pH, clay, P) resulted in an R^2 of 0.36 and an RMSE of 1.30 mg/kg. The model for leaf Mn concentration (predictors: EC, pH) exhibited lower performance, with an R^2 of 0.28 and an RMSE of 5.31 mg/kg. In stark contrast, the GEP models demonstrated vastly superior predictive power. The optimal GEP model for Cu (GEP3; inputs: pH, OC, and clay) achieved an R^2 of 0.60 and an RMSE of 0.82 mg/kg on the test set. The best model for Mn (GEP2; inputs: pH and EC) yielded an R^2 of 0.53 and an RMSE of 3.67 mg/kg. These results represent a 67% improvement in Cu concentration and a near-doubling for Mn concentration compared to the SWR models. Furthermore, the prediction error (RMSE) was reduced by 37% for Cu concentration and 31% for Mn concentration.

Conclusions

The findings of this study demonstrate that Gene Expression Programming (GEP) is a significantly more powerful and accurate tool than conventional stepwise regression for predicting citrus leaf Cu and Mn concentrations in calcareous soils. The results confirm that the relationships between soil properties and nutrient uptake are fundamentally non-linear, and GEP's capacity to model these complex interactions is the primary driver of its superior performance. The models consistently identified soil pH as a master variable controlling micronutrient availability. Moreover, the GEP model for Cu highlighted the crucial, non-linear role of organic carbon in protecting Cu from precipitation a nuance that the linear model failed to capture. The practical implication of this research is twofold. First, the explicit mathematical equations generated by the GEP models can be directly embedded into spreadsheet software or mobile applications to serve as a preliminary screening tool. This allows growers to input routine soil test data and receive instant, reliable predictions of potential nutrient deficiencies. Second, on a broader scale, this work provides a methodological framework for other regions, encouraging the adoption of non-linear modeling to develop locally calibrated decision-support systems. Ultimately, this study offers a tangible pathway toward empowering farmers with data-driven tools for enhancing citrus productivity in challenging agricultural environments.

Keywords: Bioavailability, Calcareous soils, *Citrus sinensis*, Evolutionary algorithms, Micronutrients.

Author Contributions

S. Heidari and S.A. Ghaffari Nejad conceived and planned the experiments. J. Sarhadi carried out the experiments. S. Heidari and S.A. Ghaffari Nejad analyzed data. S. Heidari and S.A. Ghaffari Nejad wrote the first manuscript. All authors contributed to the interpretation of the results. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The authors would like to thank the research council of the Soil and Water Research Institute, Iran for the financial support of this research.

Ethical considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Conflict of interest

The author declares no conflict of interest.

Cite this article: Heidari, S., Ghaffari Nejad, S.A. & Sarhadi, J. (2025). Prediction of copper and manganese concentrations in citrus leaves based on easily measured soil characteristics: Comparison of stepwise regression models and gene expression programming. *Journal of Soil and Plant Science*, 35(4), 1-22.

<https://doi.org/10.22034/sps.2026.70403.1026>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.



Copyright © 2025 The Authors.
Publisher: The University of Tabriz





پیش‌بینی غلظت مس و منگنز برگ مرکبات با استفاده از ویژگی‌های زودیافت خاک: مقایسه مدل‌های رگرسیون گام‌به‌گام و برنامه‌ریزی بیان ژن

صابر حیدری^۱، سید علی غفاری‌نژاد^۲، جواد سرحدی^۳

۱- بخش تحقیقات خاک و آب، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی جنوب استان کرمان، سازمان تحقیقات، آموزش و ترویج کشاورزی، جیرفت، ایران. رایانامه: s.heydari@areeo.ac.ir
۲- مؤسسه تحقیقات خاک و آب، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران. رایانامه: ma_ghaffari51@yahoo.com
۳- نویسنده مسئول، بخش تحقیقات خاک و آب، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی جنوب استان کرمان، سازمان تحقیقات، آموزش و ترویج کشاورزی، جیرفت، ایران. رایانامه: javad.sarhadi2009@gmail.com

تاریخ دریافت: ۱۴۰۴/۰۹/۰۴

تاریخ پذیرش: ۱۴۰۴/۱۰/۱۶

تاریخ بازنگری: ۱۴۰۴/۱۰/۱۶

تاریخ انتشار: ۱۴۰۴/۱۰/۱۷

چکیده

مدیریت بهینه تغذیه مرکبات در باغ‌های جنوب کرمان، با توجه به خاک‌های آهکی منطقه و بروز کمبودهای عنصرهای غذایی کم‌مصرف، عاملی تعیین‌کننده در اقتصاد کشاورزی منطقه است. هدف این پژوهش، توسعه و مقایسه مدل‌های پیش‌بینی‌کننده برای غلظت مس و منگنز در برگ مرکبات بر اساس ویژگی‌های خاک، با مقایسه یک روش خطی سنتی و یک الگوریتم تکاملی غیرخطی بود. تعداد ۴۰ نمونه خاک و نمونه برگ متناظر از باغ‌های پرتقال (والنسیا) جمع‌آوری و تجزیه شدند. مدل‌ها با استفاده از رگرسیون خطی گام‌به‌گام و برنامه‌ریزی بیان ژن (GEP) و با ورودی‌های حاکی شامل بافت خاک، pH، قابلیت هدایت الکتریکی و غلظت کربن آلی و فسفر ساخته شدند. نتایج به‌طور قاطع برتری عملکرد رویکرد غیرخطی را تأیید کرد. مدل بهینه GEP برای پیش‌بینی غلظت مس برگ بر اساس pH، کربن آلی و رس (بر روی داده‌های آزمون) دارای ضریب تبیین (R^2) ۰/۶۰ بود که به‌طور معناداری از مدل رگرسیون گام‌به‌گام با R^2 برابر ۰/۳۶ بهتر عمل کرد. به‌طور مشابه، برای غلظت منگنز برگ، بهترین مدل GEP بر اساس pH و قابلیت هدایت الکتریکی ضریب تبیین (R^2) ۰/۵۳ داشت که قدرت تبیین‌کنندگی آن تقریباً دو برابر مدل رگرسیون گام‌به‌گام با R^2 برابر ۰/۲۸ بود. مدل‌های GEP نه‌تنها با درک برهم‌کنش‌های پیچیده و غیرخطی خاک-گیاه، دقت پیش‌بینی بالاتری داشتند، بلکه معادلات ریاضی صریحی را نیز تولید کردند. به‌طور کلی، مدل GEP ابزاری به مراتب قدرتمندتر و قابل‌اطمینان‌تر از رگرسیون گام‌به‌گام برای این کار است و مسیری ملموس برای ایجاد سامانه‌های پشتیبان تصمیم در مدیریت دقیق مصرف کودها در باغ‌های مرکبات ارائه می‌دهد.

واژه‌های کلیدی: الگوریتم‌های تکاملی، پرتقال والنسیا، خاک‌های آهکی، عنصرهای غذایی کم‌مصرف، زیست‌فراهمی.

استناد به این مقاله: حیدری، ص.، غفاری‌نژاد، س.ع. و سرحدی، ج. (۱۴۰۴). پیش‌بینی غلظت مس و منگنز برگ مرکبات با استفاده از ویژگی‌های زودیافت خاک: مقایسه مدل‌های رگرسیون گام‌به‌گام و برنامه‌ریزی بیان ژن. نشریه دانش خاک و گیاه، ۳۵(۴)، ۱-۲۲.

<https://doi.org/10.22034/sps.2026.70403.1026>

مقدمه

فراهمی عنصرهای غذایی کم‌مصرف، یکی از ارکان اصلی کشاورزی پایدار و امنیت غذایی در جهان است. در میان این عنصرها، مس (Cu) و منگنز (Mn) برای رشد، نمو و بهره‌وری بهینه محصولات باغی مانند مرکبات حیاتی هستند (Vashisth & Kadyampakeni, 2020). کمبود این عنصرها می‌تواند به کاهش چشمگیر در شادابی گیاه، عملکرد و کیفیت میوه منجر شود و تهدیدی اقتصادی برای مناطقی محسوب می‌شود که به‌شدت به تولید مرکبات وابسته‌اند (Ahmad *et al.*, 2022). این چالش به‌ویژه در مناطق خشک و نیمه‌خشک، مانند جنوب استان کرمان در ایران، حادث است؛ جایی که خاک‌ها عمدتاً آهکی با pH بالا و مقدار ماده آلی کم هستند.

رابطه میان ویژگی‌های خاک و جذب عنصرهای غذایی به‌وسیله گیاه، به روابط بیورئوشیمیایی پیچیده بستگی دارد. درک این روابط چندوجهی برای پیش‌بینی وضعیت تغذیه‌ای گیاه از روی داده‌های خاک، کاری بنیادین است. برای آشکارسازی این روابط پیچیده خاک و گیاه، پژوهشگران طیف وسیعی از روش‌های مدل‌سازی را به کار گرفته‌اند. رویکردهای سنتی اغلب بر روش‌های آماری خطی، مانند تحلیل همبستگی و رگرسیون گام به گام، برای شناسایی تأثیرگذارترین متغیرهای خاک تکیه دارند (Yang *et al.*, 2022). این روش‌ها هرچند تفسیرپذیر هستند و می‌توانند بینش‌های اولیه ارزشمندی ارائه دهند، اما اساساً به‌دلیل فرض خطی بودن روابط، با محدودیت مواجه‌اند. نظام‌های خاک-گیاه ذاتاً غیرخطی هستند و مدل‌های خطی اغلب در ثبت اثرهای پیچیده، تعاملی و آستانه‌ای ویژگی‌های خاک بر فراهمی عنصرهای غذایی ناتوان‌اند (Heidari, *et al.*, 2024b). این نارسایی، منجر به گذار به سمت رویکردهای پیچیده‌تر یادگیری ماشین^۱ شده است که در مدل‌سازی پدیده‌های پیچیده و غیرخطی بدون پیش‌فرض در مورد روابط زیربنایی، برتری دارند (Ebrahimi *et al.*, 2022). اخیراً، مدل برنامه‌ریزی بیان ژن^۲، که یک الگوریتم تکاملی است، به عنوان یک جایگزین قدرتمند ظهور کرده است. مزیت کلیدی GEP نسبت به مدل‌های مانند ANN، توانایی آن در تولید معادلات ریاضی صریح و شفاف است که رابطه ورودی-خروجی را توصیف می‌کنند و ترکیبی ارزشمند از دقت پیش‌بینی بالا و تفسیرپذیری را ارائه می‌دهند (Ebrahimi *et al.*, 2021; Shiri *et al.*, 2014).

مطالعات نشان می‌دهند که در بسیاری از مناطق زیر کشت مرکبات، خاک‌ها دارای pH بالا و ماده آلی کم هستند؛ شرایطی که فراهمی عنصرهای غذایی کم‌مصرف را به‌شدت محدود می‌کند (Ahmad *et al.*, 2022; Li *et al.*, 2015). pH بالا در این مناطق، رسوب عنصرهای غذایی کم‌مصرف به شکل هیدروکسیدها و کربنات‌های نامحلول را تسریع کرده و عملاً آن‌ها را از دسترس ریشه‌های گیاه خارج می‌کند. این مطالعات به‌طور مداوم pH خاک را به عنوان یک متغیر اصلی و کلیدی معرفی می‌کنند و در عین حال به تأثیر قابل‌توجه کربن آلی (OC)، بافت خاک و حضور عنصرهای ناهم‌ساز^۳ مانند فسفر (P) نیز اشاره دارند. pH خاک حل‌پذیری و تحرک عنصرهای غذایی کم‌مصرف را کنترل می‌کند؛ به‌طوری که با افزایش pH، زیست‌فراهمی مس و منگنز به‌شدت کاهش می‌یابد (Chatzistathis *et al.*, 2017). Najafi *et al.* (2010) گزارش کردند که با افزایش pH محلول غذایی در کشت هیدروپونیک، غلظت مس و منگنز در بخش هوایی گیاه اسفناج کاهش یافت. (Salimi Tarazoj *et al.*, 2024) مشاهده کردند که در حضور کاه گندم، مصرف کود فسفر در خاک سبب کاهش غلظت مس و منگنز بخش هوایی گیاه کلزا شد.

در مقابل، کربن آلی خاک (OC) غالباً با تشکیل کمپلکس‌های آلی-فلزی محلول (کی‌لیت‌ها)، فراهمی عنصرهای غذایی کم‌مصرف را افزایش می‌دهد؛ این کی‌لیت‌ها از رسوب عنصرها جلوگیری کرده و انتقال آن‌ها را به سطح ریشه تسهیل می‌کنند (Fu *et al.*, 2004). سایر ویژگی‌های خاک، از جمله بافت (مقدار رس و سیلت)، قابلیت

¹-Machine learning (ML)

²-Gene expression programming (GEP)

³-Antagonistic

هدایت الکتریکی (EC) و وجود عنصرهای ناهم‌ساز مانند فسفر (P)، این پویایی را پیچیده‌تر می‌کنند. برای مثال، اثر ناهم‌سازی سطوح بالای فسفر بر جذب مس و روی که می‌تواند حتی در صورت وجود این عنصرها در خاک منجر به کمبود آن‌ها شود، یک موضوع تکراری در تحقیقات است (Moreno-Lora & Delgado, 2020)، در مقابل، نقش سودمند ماده آلی در تشکیل کمپلکس‌های کی‌لیت-فلز محلول که عنصرهای غذایی کم‌مصرف را برای جذب به‌وسیله گیاه در دسترس نگه می‌دارد، یک عامل کلیدی تعدیل‌کننده است (Feil *et al.*, 2020)

در گذشته، پژوهشگران (Ahmad *et al.*, 2022; Li *et al.*, 2015) برای مدل‌سازی این روابط از روش‌های آماری سنتی مانند رگرسیون خطی گام به گام استفاده می‌کردند که کارایی این مدل‌های خطی را در شناسایی پیش‌بینی‌کننده‌های مهم خاک نشان می‌دهد. با این حال، یک نقد عمده و مداوم به این رویکرد، محدودیت ذاتی آن در به تصویر کشیدن دینامیک‌های غیرخطی نظام‌های خاک-گیاه است. فرایندهای خاک شامل برهم‌کنش‌های پیچیده، آستانه‌ها و اثرهای هم‌افزایی یا بازدارندگی هستند که با مفروضات خطی نمی‌توان آن‌ها را به اندازه کافی نمایش داد. این محدودیت اغلب به مدل‌هایی با قدرت پیش‌بینی پایین منجر می‌شود و بر ضرورت استفاده از رویکردهای پیچیده‌تر تأکید می‌کند (Hosseinifard *et al.*, 2019; Sahoo & Jha, 2013)

این محدودیت‌ها منجر به تغییر رویکرد به سمت روش‌های پیچیده‌تر یادگیری ماشین شده است؛ روش‌هایی که اخیراً به عنوان ابزارهای تحول‌آفرین برای تحلیل بلادرنگ خاک و کشاورزی دقیق مورد بررسی و توجه قرار گرفته‌اند (Rahman & Nath Das, 2025, Gunasekaran & Sreevardhan, 2025). در مطالعات حیدری و همکاران (Heidari, *et al.*, 2024a; Heidari, *et al.*, 2024b)، از شبکه‌های عصبی مصنوعی^۱ (ANN) برای پیش‌بینی غلظت آهن و روی برگ مرکبات با موفقیت استفاده و نشان داده شد که شبکه‌های عصبی مصنوعی می‌توانند این روابط پیچیده و غیرخطی را با دقتی به مراتب بالاتر از رگرسیون سنتی مدل‌سازی کنند. به‌طور مشابه، مطالعه دیگری از شبکه‌های عصبی مصنوعی برای تخمین فعالیت میکروبی خاک بر اساس ویژگی‌های فیزیکی و شیمیایی آن استفاده کردند (Ebrahimi *et al.*, 2022). اگرچه شبکه‌های عصبی مصنوعی قدرتمند هستند، اما اغلب به دلیل عدم تفسیرپذیری، به عنوان مدل‌های جعبه سیاه^۲ مورد انتقاد قرار می‌گیرند.

برنامه‌ریزی بیان ژن (GEP)، به عنوان یک الگوریتم تکاملی و جایگزین مدل‌های جعبه سیاه معرفی شد (Ferreira, 2001). GEP دقت پیش‌بینی بالای یادگیری ماشین را با شفافیت رگرسیون سنتی ترکیب می‌کند. مطالعه بر روی مدل‌سازی فعالیت آنزیمی خاک نشان داد که GEP نه تنها عملکرد برتری دارد، بلکه معادلات ریاضی صریحی نیز تولید می‌کند (Ebrahimi *et al.*, 2022). این ویژگی منحصر به فرد، آن را به ابزاری بسیار ارزشمند برای کاربردهای عملی تبدیل می‌کند، زیرا هم پیش‌بینی‌های دقیق و هم مدلی قابل فهم از رفتار نظام ارائه می‌دهد.

علی‌رغم پیشرفت‌ها در مدل‌سازی یادگیری ماشین برای علوم خاک، یک ارزیابی مقایسه‌ای مستقیم بین یک روش آماری سنتی (رگرسیون گام به گام) و یک الگوریتم تکاملی پیشرفته (مدل برنامه‌ریزی بیان ژن) برای پیش‌بینی هم‌زمان غلظت مس و منگنز برگ در مرکبات، به‌طور کامل بررسی نشده است. اگرچه عوامل منفرد مؤثر بر زیست‌فراهمی مس و منگنز شناخته شده‌اند و مدل‌های یادگیری ماشین برای سایر عنصرهای غذایی به کار رفته‌اند، اما شکاف قابل‌توجهی در شناسایی قوی‌ترین و کاربردی‌ترین ابزار پیش‌بینی، به‌ویژه برای مدیریت مس و منگنز در شرایط منحصر به فرد خاک‌های آهکی جنوب کرمان، باقی مانده است. ویژگی‌های خاص خاک این منطقه، چالشی متمایز را ایجاد می‌کند که نیازمند مدل‌های پیش‌بینی متناسب برای تدوین راهبردهای مؤثر کوددهی است.

با لحاظ نمودن چالش‌های مذکور، پژوهش حاضر مسیری فراتر از مدل‌سازی‌های آماری محض را در پیش گرفته و تحقق اهداف زیر را دنبال می‌کند: ۱- عبور از محدودیت مدل‌های خطی در خاک‌های آهکی برای

¹-Artificial neural networks (ANN)

²-black box

آشکارسازی روابط غیرخطی و پیچیده‌ای که فراهمی مس و منگنز را کنترل می‌کنند. ۲- بهره‌برداری از ماهیت جعبه‌سفید الگوریتم GEP برای استخراج معادلات ریاضی صریح و قابل تفسیر. این رویکرد، بر خلاف مدل‌های جعبه‌سیاه رایج، به پژوهشگران این امکان را می‌دهد که منطق اکولوژیک نهفته در پیش‌بینی‌ها را درک نمایند. ۳- ارائه یک ابزار محاسباتی کاربردی و اقتصادی برای باغداران منطقه، که تخمین وضعیت تغذیه‌ای درختان را صرفاً با اتکا به داده‌های ارزان‌قیمت آزمون خاک امکان‌پذیر می‌سازد.

مواد و روش‌ها

منطقه مورد مطالعه

این پژوهش در باغ‌های مرکبات جنوب استان کرمان، واقع در جنوب شرقی ایران (طول جغرافیایی 56° و $17'$ و عرض جغرافیایی 26° و $43'$) انجام شد که به دلیل سطح زیر کشت گسترده و تولید بالا، یکی از قطب‌های راهبردی و رتبه سوم تولید مرکبات در ایران محسوب می‌شود. بر اساس طبقه‌بندی اقلیمی کوپن، این منطقه دارای اقلیم گرم و خشک^۱ تا نیمه‌خشک است. تابستان‌های طولانی و گرم و زمستان‌های کوتاه و معتدل از ویژگی‌های بارز این ناحیه می‌باشد (Sarhadi et al. 2020).

نمونه‌برداری از خاک و گیاه

در مجموع، ۴۰ باغ مرکبات رقم والنسیا (*Citrus sinensis*, cv. Valencia) با سن درختان بیش از ۱۰ سال به صورت تصادفی انتخاب شدند. در هر باغ، ۱۰ نمونه خاک از دو عمق ۳۰-۶۰ سانتی‌متر از ناحیه سایه‌انداز درختان (یک‌سوم انتهایی) در شهریور ماه برداشت گردید. پس از مخلوط کردن نمونه‌های هر عمق، یک نمونه مرکب یک کیلوگرمی برای هر باغ و هر عمق تهیه شد. از آنجا که جذب عنصرهای غذایی به‌وسیله ریشه مرکبات از کل پروفیل خاک (۰ تا ۶۰ سانتی‌متر) انجام می‌شود و غلظت عنصر در برگ بازتابی از شرایط کل محیط ریشه است، برای ورود داده‌ها به مدل‌های پیش‌بینی، مقادیر اندازه‌گیری شده در دو عمق ادغام شدند. برای این کار و جهت به دست آوردن یک شاخص واحد که نماینده کل منطقه توسعه ریشه گیاه باشد، از میانگین وزنی پارامترها در دو عمق استفاده گردید.

هم‌زمان با نمونه‌برداری خاک، نمونه‌برداری برگ نیز انجام شد. برگ‌های شاخه‌های غیربارده با رشد بهاره ۴ تا ۶ ماهه نیز از همان باغ‌ها و از قسمت میانی درخت نمونه‌برداری شدند. برگ‌ها در آزمایشگاه شسته و در آون با دمای 70° درجه سلسیوس خشک شدند. اندازه‌گیری غلظت مس و منگنز در برگ‌ها پس از هضم با اسید سولفوریک در دمای 280° درجه سلسیوس و با دستگاه جذب اتمی انجام شد (Kalra, 1997).

نمونه‌های خاک پس از خشک‌شدن در هوا و عبور از الک ۲ میلی‌متری، برای تجزیه آماده شدند. ویژگی‌های بافت خاک به روش هیدرومتری ۲ زمانه، pH و EC، به ترتیب در گل اشباع و عصاره اشباع خاک (Allison & Richards, 1954)، کربن آلی (OC) با روش اکسایش مرطوب والکی-بلک (Nelson & Sommers, 1983)، درصد کل مواد خنثی شونده (TNV) با روش خنثی‌سازی با اسید و تیتراژ کردن با سود (Rowell, 2014) و فسفر قابل‌دسترس (P) به روش اولسن (Olsen, 1954) با دستگاه اسپکتروفتومتر اندازه‌گیری شدند.

تحلیل رگرسیون گام به گام

¹-Arid

رگرسیون گام به گام یک ساختار تکرارشونده و گام به گام برای ایجاد یک مدل رگرسیونی است که متغیرهای مستقل را برای استفاده در مدل نهایی انتخاب می‌کند. این فرایند شامل افزودن یا حذف متوالی متغیرهای توضیحی احتمالی و بررسی نتایج مدل پس از هر تکرار است تا ترکیب بهینه مجموعه داده ورودی از متغیرهای مستقل با توجه به مجموع مربعات باقی‌مانده به دست آید (Hosseinpour et al., 2018) در روش رگرسیون گام به گام از معادله زیر استفاده شد (معادله ۱).

$$Y = a + bX_1 + cX_2 + \dots \quad (1)$$

در این رابطه، Y متغیر وابسته (خروجی مدل)، X_1 و X_2 متغیرهای مستقل (متغیرهای ورودی به مدل) و a ، b و c ثابت‌های مدل هستند.

مدل برنامه‌ریزی بیان ژن

مدل برنامه‌ریزی بیان ژن (GEP) یک الگوریتم تکاملی است که به‌وسیله فریرا (Ferreira, 2001) معرفی شد و قادر به مدل‌سازی روابط پیچیده و غیرخطی است. برخلاف مدل‌های جعبه سیاه، GEP یک مدل ریاضی صریح در قالب درختان بیان^۱ تولید می‌کند. فرایند GEP شامل مراحل زیر است:

- ۱- ایجاد جمعیت اولیه: مجموعه‌ای از راه‌حل‌های کاندید (کروموزوم‌ها) به صورت تصادفی تولید می‌شود.
- ۲- بیان کروموزوم: هر کروموزوم خطی به یک ساختار درختی غیرخطی (درخت بیان) ترجمه می‌شود.
- ۳- ارزیابی برازش^۲ هر درخت بیان (مدل) با استفاده از یک تابع برازش ارزیابی می‌شود. در این پژوهش، از ریشه میانگین مربعات خطا^۳ (RMSE) به عنوان تابع برازش استفاده شد.
- ۴- انتخاب و تکامل: مدل‌ها بر اساس میزان برازش خود انتخاب می‌شوند. بهترین مدل (نخبه) مستقیماً به نسل بعد منتقل شده و سایر مدل‌ها از طریق عملگرهای ژنتیکی (مانند جهش، نو ترکیبی و جابجایی) برای تولید نسل جدید تکامل می‌یابند. این فرایند تا رسیدن به شرط توقف (مثلاً تعداد نسل مشخص) تکرار می‌شود.

در این مطالعه، از نرم‌افزار GeneXpro Tools (نسخه ۵) برای پیاده‌سازی مدل GEP استفاده شد. مجموعه داده‌ها شامل ۴۰ نمونه، به صورت تصادفی به دو بخش آموزشی (۷۰ درصد، معادل ۲۸ نمونه) و آزمایشی (۳۰ درصد، معادل ۱۲ نمونه) تقسیم شد تا از ارزیابی مستقل و قوی مدل‌ها اطمینان حاصل شود. در این پژوهش، علاوه بر پارامترهای ورودی مختلف، توالی این پارامترها نیز در پیش‌بینی غلظت مس و منگنز در نظر گرفته شد. مجموعه پایانه‌ها^۴ برای مدل برنامه‌ریزی بیان ژن، ترکیبات مختلفی از متغیرهای خاک مانند سیلت، رس، pH، EC، TNV، OC و P، مطابق جدول ۵ و ۷ بودند. از ریشه میانگین مربعات خطا (RMSE) به عنوان تابع برازش استفاده شد. توابع به کار رفته عبارت بودند از $(-)$ ، $(+)$ ، (\times) ، (\div) ، \ln ، $\sqrt{\quad}$ ، \cos ، \sin ، \arctan ، $x^{1/3}$ ، x^2 و x^3 . متغیرهای مورد استفاده در توسعه مدل‌های مختلف GEP در جدول ۱ نشان داده شده‌اند. در این مطالعه، مشخصات خاک به عنوان ورودی مدل GEP برای ارزیابی جذب مس و منگنز به کار رفت.

1 - Expression Trees
2 - Fitness Evaluation
3 - Root Mean Square Error (RMSE)
4 - Terminals

جدول ۱- پارامترهای مورد استفاده در GEP.

پارامترها	مقادیر
Number of chromosomes	۳۰
Head size	۸
Number of genes	۳
Linking function	Addition
Mutation rate	۰/۰۴۴
Inversion rate	۰/۱
Gene recombination rate	۰/۱
Gene transposition rate	۰/۱
One-point recombination rate	۰/۳
Two-point recombination rate	۰/۳
IS transposition rate	۰/۱
RIS transposition rate	۰/۱
Functions set	+, -, ×, ÷, sqrt, exp., ln, x ² , x ³ , x ^{1/3} , sin, cos, arctan

معیارهای ارزیابی

برای ارزیابی کارایی و دقت مدل‌ها، از ضریب تبیین (R^2)، ریشه میانگین مربعات خطا (RMSE) و میانگین قدر مطلق خطا^۱ (MAE) استفاده شد (معادلات ۲، ۳ و ۴). R^2 ، که در معادله ۲ ارائه شده، میزان نزدیکی داده‌های مشاهده شده و شبیه‌سازی شده را نشان می‌دهد و مقادیر R^2 نزدیک به ۱، بیانگر مدلی با دقت بالاتر است. RMSE (معادله ۳) و MAE (معادله ۴) نشان‌دهنده خطای تخمین‌ها هستند و هرچه مقادیر آن‌ها به صفر نزدیک‌تر باشد، شبیه‌سازی دقیق‌تر است (Mehdizadeh et al., 2017).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

که در آن، y_i ، \hat{y}_i و \bar{y} به ترتیب مقادیر غلظت مس و منگنز اندازه‌گیری شده، تخمین زده شده به وسیله مدل‌ها و میانگین مقادیر اندازه‌گیری شده هستند. بهترین مدل، مدلی است که مقادیر R^2 و RMSE یا MAE آن به ترتیب به یک و صفر نزدیک‌تر باشد.

نتایج

آمار توصیفی و همبستگی‌های اولیه

ویژگی‌های فیزیکی و شیمیایی نمونه‌های خاک و غلظت متناظر مس و منگنز در برگ از ۴۰ باغ مرکبات در جدول ۲ خلاصه شده است. خاک‌ها عمدتاً خنثی تا قلیایی با میانگین pH برابر با ۷/۷۵ بودند و تنوع قابل‌توجهی در قابلیت هدایت الکتریکی، بافت و ارزش خنثی‌شوندگی کل (TNV) نشان دادند. بافت خاک‌ها از شنی تا لومی رسی متغیر بود. از نظر pH، ۱۰ خاک دارای pH کمتر از ۷/۵ و در ارتباط با قابلیت هدایت الکتریکی، ۹ نمونه خاک دارای EC ۴ دسی‌زیمنس بر متر و بالاتر بودند. غلظت مس برگ در بازه ۱/۵۰ تا ۶/۵۱ میلی‌گرم بر کیلوگرم و غلظت منگنز برگ در بازه گسترده‌تری از ۹/۵۰ تا ۳۶/۵۰ میلی‌گرم بر کیلوگرم متغیر بود. همچنین، در مورد غلظت مس و منگنز برگ، بر اساس معیارهای تفسیر نتایج آزمون برگ در مرکبات (Asadi Kangarshahi et al., 2019)، مقدار آن‌ها در برگ‌ها بین مقادیر کمبود تا مناسب متغیر بود.

جدول ۲- خلاصه ویژگی‌های نمونه‌های خاک و مس و منگنز برگ.

ویژگی	حداکثر	حداقل	میانگین	انحراف معیار	ضریب تغییرات
EC (dS/m)	۱۳/۳۵	۰/۵۵	۳/۰۳	۲/۴۹	۸۲/۱۷
pH	۸/۳	۷/۲۵	۷/۷۵	۰/۲۸	۳/۶۱
Silt (%)	۴۶/۶۵	۴/۲۰	۲۴/۲۲	۱۳/۲۱	۵۴/۵۴
Clay (%)	۲۶/۵۰	۳	۸/۸۹	۵/۶۷	۶۳/۷۷
TNV (%)	۱۵/۳۷	۰	۶/۳۷	۳/۴۷	۵۴/۴۷
OC (%)	۰/۵۱	۰/۰۹	۰/۳۱	۰/۱۱	۳۵/۴۸
P (mg/kg)	۲۲/۱۳	۰/۹۷	۶/۵۶	۳/۰۶	۴۶/۶۴
Leaf Mn conc. (mg/kg)	۳۶/۵۰	۹/۵۰	۲۰/۲۱	۵/۷۰	۲۸/۲۰
Leaf Cu conc. (mg/kg)	۶/۵۱	۱/۵۰	۳/۵۰	۱/۰۸	۳۰/۸۵

تحلیل همبستگی پیرسون برای شناسایی روابط خطی بین ویژگی‌های خاک و مقدار عنصرهای غذایی برگ انجام شد (جدول ۳). غلظت مس برگ همبستگی مثبت معناداری با کربن آلی ($r = 0.39^*$) و درصد رس ($r = 0.32^*$) و همبستگی منفی معناداری با pH خاک ($r = -0.41^*$) و فسفر قابل‌دسترس ($r = -0.31^*$) نشان داد. غلظت منگنز برگ رابطه مثبت معناداری با کربن آلی ($r = 0.28^*$) داشت و با pH خاک ($r = -0.33^*$) و با قابلیت هدایت الکتریکی ($r = -0.29^*$) همبستگی منفی معنادار داشت.

جدول ۳- ضریب همبستگی ویژگی‌های فیزیکی‌شیمیایی خاک و غلظت مس و منگنز در برگ‌ها.

ویژگی	غلظت منگنز برگ	غلظت مس برگ
EC (dS/m)	۰/۲۹*	-۰/۲۲
pH	-۰/۳۳*	-۰/۴۱*
Silt (%)	۰/۱۳	۰/۲۲
Clay (%)	۰/۱۲	۰/۳۲*
TNV (%)	۰/۰۰۷	-۰/۱۸
OC (%)	۰/۲۸*	۰/۳۹*

P (mg/kg)

۰/۰۱

-۰/۳۱*

تحلیل رگرسیون گام به گام

برای ساخت یک مدل پیش‌بینی خطی، تحلیل رگرسیون خطی گام به گام انجام شد. مدل رگرسیون برای مس برگ، متغیرهای pH، رس و P را به عنوان مهم‌ترین پیش‌بینی‌کننده‌ها شناسایی کرد. برای منگنز برگ، متغیرهای EC و pH را به عنوان تأثیرگذارترین متغیرها حفظ کرد. عملکرد پیش‌بینی این مدل‌ها، که بر روی مجموعه داده آزمون ارزیابی شد، در جدول ۴ خلاصه شده است. مدل مس برگ به ضریب تبیین (R^2) برابر با ۰/۳۶، ریشه میانگین مربعات خطا (RMSE) برابر با ۱/۳۰ میلی‌گرم بر کیلوگرم و میانگین قدر مطلق خطا (MAE) برابر با ۱/۱۶ میلی‌گرم بر کیلوگرم دست یافت. مدل برای منگنز برگ، R^2 برابر با ۰/۲۸، RMSE برابر با ۵/۳۱ میلی‌گرم بر کیلوگرم و MAE برابر با ۴/۵۲ میلی‌گرم بر کیلوگرم به دست آورد. در رگرسیون گام به گام، مقدار مس و منگنز برگ به عنوان متغیر وابسته و بقیه صفات خاک به عنوان متغیر مستقل در مدل وارد شد. معادلات ۵ و ۶، مدل رگرسیونی رابطه مس و منگنز برگ با ویژگی‌های خاک را نشان می‌دهند.

$$\text{Leaf Cu conc.} = 16.76661 - 1.73436\text{pH} + 0.06558\text{Clay} - 0.05867\text{P} \quad (5)$$

$$\text{Leaf Mn conc.} = 68.63872 - 0.66426\text{EC} - 6.05257\text{pH} \quad (6)$$

جدول ۴- معیارهای ارزیابی مدل‌های رگرسیون گام به گام برای رابطه غلظت مس و منگنز برگ با ویژگی‌های خاک.

مدل	R^2	RMSE (mg/kg)	MAE (mg/kg)
مس برگ	۰/۳۶	۱/۳۰	۱/۱۶
منگنز برگ	۰/۲۸	۵/۳۱	۴/۵۲

بر اساس نتایج حاصل از تحلیل رگرسیون چندمتغیره، مشخص گردید که سه پارامتر کلیدی خاک، شامل pH، درصد رس و میزان فسفر، در مجموع قادر به توجیه ۳۶ درصد از واریانس مشاهده‌شده در غلظت مس برگ‌های مرکبات هستند (جدول ۴). ضریب رگرسیونی منفی (۱/۷۳-) برای pH، نشان‌دهنده یک ارتباط معکوس است؛ به این معنا که با افزایش pH خاک، از میزان تجمع مس در بافت برگ کاسته می‌شود. در مقابل، درصد رس خاک با ضریب همبستگی مثبت (۰/۶۵) رابطه مستقیمی با غلظت مس برگ نشان داد که حاکی از نقش بافت‌های سنگین‌تر خاک در بهبود جذب این عنصر است. نهایتاً، فسفر خاک با ضریبی معادل ۰/۵۸-، یک رابطه منفی ضعیف با غلظت مس برگ برقرار نمود.

برای مدل‌سازی غلظت منگنز در برگ، روش رگرسیون گام‌به‌گام دو متغیر اصلی پیش‌بینی‌کننده یعنی هدایت الکتریکی و pH خاک را شناسایی کرد (معادله ۶). این مدل ترکیبی توانست ۲۸ درصد از کل واریانس غلظت منگنز برگ را تبیین نماید (جدول ۴). نکته حائز اهمیت، تأثیر بازدارنده هر دو متغیر بر جذب منگنز بود. ضریب منفی و بزرگ pH (۶/۰۵) بر اهمیت حیاتی pH خاک در فراهمی منگنز برای مرکبات تأکید دارد، به طوری که خاک‌های اسیدی‌تر (با pH پایین‌تر) به طور مستقیم با سطوح بالاتر منگنز در برگ مرتبط بودند. به همین ترتیب، شوری خاک که با شاخص EC ارزیابی می‌شود، با ضریب ۰/۶۶-، یک ارتباط معکوس با غلظت منگنز برگ نشان داد که مؤید تأثیر منفی شوری بر قابلیت جذب این عنصر به وسیله گیاه است.

مدل برنامه‌ریزی بیان ژن

برای ثبت روابط پیچیده‌تر و غیرخطی، از مدل برنامه‌ریزی بیان ژن (GEP) برای توسعه مجموعه‌ای از مدل‌های پیچیده‌تر استفاده شد. برای هر دو عنصر مس و منگنز، هفت مدل با افزودن تدریجی ویژگی‌های خاک به عنوان متغیرهای ورودی ساخته شد؛ این فرایند با یک پیش‌بینی‌کننده واحد آغاز و به مدلی با هر هفت متغیر ختم شد (برای ترکیب ورودی‌ها به جدول ۵ و جدول ۷ مراجعه کنید).

عملکرد این مدل‌ها در هر دو فاز آموزش و آزمون در جدول ۶ (برای مس) و جدول ۸ (برای غلظت منگنز) به تفصیل آمده است. برای پیش‌بینی غلظت مس برگ، عملکرد مدل بر روی داده‌های آزمون عموماً با افزودن متغیرهای مرتبط‌تر بهبود یافت، به طوری که مدل GEP3 (ورودی‌ها pH، OC، رس) به بالاترین R^2 آزمون برابر با ۰/۶۰ و کمترین RMSE آزمون برابر با ۰/۸۲ میلی‌گرم بر کیلوگرم دست یافت. اگرچه مدل GEP7، که از هر هفت ورودی استفاده می‌کرد، بهترین عملکرد را بر روی داده‌های آموزش داشت ($R^2=0.73$)، عملکرد آن بر روی داده‌های آزمون کمی پایین‌تر بود ($R^2=0.58$) که نشان می‌دهد مدل GEP3 راه‌حلی تعمیم‌پذیرتر ارائه می‌دهد.

برای پیش‌بینی غلظت منگنز برگ، روند مشابهی مشاهده شد. مدل GEP2، که تنها از pH و EC به عنوان ورودی استفاده می‌کرد، بهترین پیش‌بینی را برای مجموعه داده آزمون ارائه داد و به R^2 برابر با ۰/۵۳ و RMSE برابر با ۲/۶۷ میلی‌گرم بر کیلوگرم دست یافت. افزودن متغیرهای بیشتر در مدل‌های بعدی (GEP3-GEP7) به‌طور مداوم باعث بهبود دقت پیش‌بینی در مجموعه آزمون نشد و در برخی موارد آن را کاهش داد، که نشان می‌دهد pH و EC حیاتی‌ترین پیش‌بینی‌کننده‌های غیرخطی برای منگنز در این نظام بودند.

یک مزیت کلیدی GEP، توانایی آن در تولید فرمول‌های ریاضی صریح است که روابط شناسایی شده را نشان می‌دهند. این فرمول‌ها به صورت بصری به شکل درختان بیان^۱ نمایش داده می‌شوند. شکل‌های ۱ و ۲ به ترتیب درختان بیان را برای مدل‌های GEP3 (غلظت مس) و GEP2 (غلظت منگنز) نشان می‌دهند. این درختان ماهیت تفسیرپذیر الگوریتم GEP را به نمایش می‌گذارند و نشان می‌دهند که معادلات ریاضی نهایی (معادلات ۷ و ۸) دقیقاً از چه مسیری و با ترکیب کدام توابع غیرخطی از این ساختارها استخراج شده‌اند. هر درخت نشان می‌دهد که چگونه متغیرهای ورودی (مثلاً d_0 و d_1) از طریق توابع ریاضی مختلف (مانند $+$ ، \cos ، \ln ، x^2) ترکیب می‌شوند تا پیش‌بینی نهایی را تولید کنند. این ساختارهای بصری، پیچیدگی ذاتی نظام خاک-گیاه را آشکار می‌سازند که مدل‌های خطی قادر به ثبت آن نیستند. معادلات ریاضی متناظر برای مدل بهینه مس (GEP3) و مدل منگنز (GEP2) به ترتیب در معادلات ۷ و ۸ ارائه شده است:

$$\text{Leaf Cu conc.} = \text{OC} + \cos(\text{pH} + \cos(\ln(\text{pH} \times \text{Clay}) + \text{Clay})) + \text{Clay}^{0.5} \quad (7)$$

جدول ۵- پارامترهای ورودی با ترکیبات مختلف در مدل GEP برای غلظت مس برگ.

مدل	ترکیب پارامترهای ورودی
GEP 1	pH
GEP 2	pH, OC
GEP 3	pH, OC, Clay
GEP 4	pH, OC, Clay, P
GEP 5	pH, OC, Clay, P, EC
GEP 6	pH, OC, Clay, P, EC, Silt
GEP 7	pH, OC, Clay, P, EC, Silt, TNV

جدول ۶- معیارهای ارزیابی مدل‌های GEP در پیش‌بینی غلظت مس در برگ‌ها.

مدل	مس برگ	R ²	RMSE (mg/kg)	MAE (mg/kg)
GEP1	آموزش	۰/۵۱	۰/۷۵	۰/۶۴
	آزمون	۰/۵۳	۰/۷۵	۰/۶۴
GEP2	آموزش	۰/۶۲	۰/۶۷	۰/۵۹
	آزمون	۰/۵۶	۰/۷۹	۰/۶۶
GEP3	آموزش	۰/۶۶	۰/۶۵	۰/۵۶
	آزمون	۰/۶۰	۰/۸۲	۰/۶۶
GEP4	آموزش	۰/۵۳	۱/۱۶	۰/۹۴
	آزمون	۰/۴۹	۰/۸۴	۰/۶۶
GEP5	آموزش	۰/۷۱	۰/۶۰	۰/۴۹
	آزمون	۰/۵۷	۰/۹۸	۰/۸۱
GEP6	آموزش	۰/۷۰	۰/۶۰	۰/۵۱
	آزمون	۰/۵۰	۰/۹۳	۰/۷۳
GEP7	آموزش	۰/۷۳	۰/۵۶	۰/۴۶
	آزمون	۰/۵۸	۰/۸۷	۰/۶۹

$$\text{Leaf Mn conc.} = (\text{pH} + \ln(\text{EC}) + \text{EC}) \times \cos(\text{pH}) + (\text{pH} + \cos(((\text{pH} + \text{EC}) \times \text{pH})^2)) + (\sin(\text{EC}) + (2 \times \text{pH}) - \text{EC}) - \ln(9.792939 - \text{EC}) \quad (8)$$

جدول ۷- پارامترهای ورودی با ترکیبات مختلف در مدل GEP برای غلظت منگنز برگ.

مدل	ترکیب پارامترهای ورودی
GEP 1	pH
GEP 2	pH, EC
GEP 3	pH, EC, OC
GEP 4	pH, EC, OC, Silt
GEP 5	pH, EC, OC, Silt, Clay
GEP 6	pH, EC, OC, Silt, Clay, P
GEP 7	pH, EC, OC, Silt, Clay, P, TNV

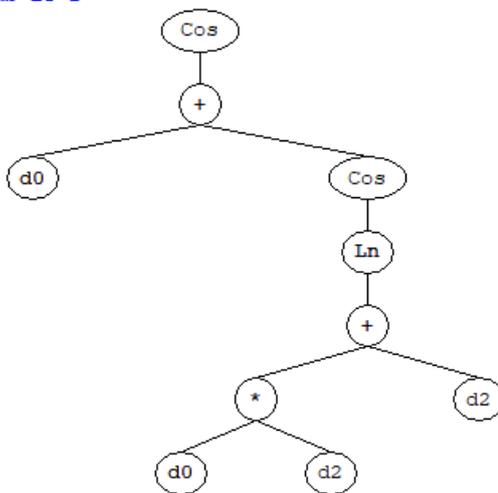
جدول ۸- معیارهای ارزیابی مدل‌های GEP در پیش‌بینی غلظت منگنز برگ.

مدل	غلظت منگنز برگ	R ²	RMSE (mg/kg)	MAE (mg/kg)
GEP1	آموزش	۰/۳۳	۴/۲۳	۳/۷۱
	آزمون	۰/۳۴	۴/۳۹	۳/۷۷
GEP2	آموزش	۰/۵۶	۲/۳۲	۲/۷۰
	آزمون	۰/۵۳	۲/۶۷	۳/۱۳
GEP3	آموزش	۰/۴۸	۳/۵۹	۲/۹۸
	آزمون	۰/۶۳	۴/۰۸	۳/۵۴
GEP4	آموزش	۰/۵۳	۳/۴۷	۳/۸۹
	آزمون	۰/۳۹	۴/۱۷	۳/۶۴
GEP5	آموزش	۰/۶۰	۲/۱۹	۲/۴۷
	آزمون	۰/۵۱	۲/۸۸	۳/۵۷
GEP6	آموزش	۰/۵۴	۳/۴۰	۲/۷۹
	آزمون	۰/۴۴	۳/۷۸	۳/۱۴
GEP7	آموزش	۰/۵۵	۳/۳۴	۲/۶۸
	آزمون	۰/۳۶	۴/۰۹	۳/۵۲

Sub-ET 1



Sub-ET 2

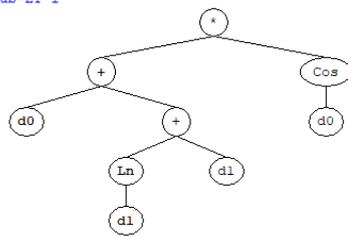


Sub-ET 3

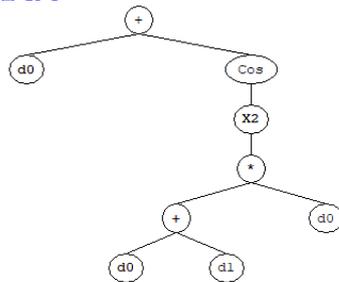


شکل ۱- درخت بیان مدل GEP3 برای تخمین غلظت Cu برگ.

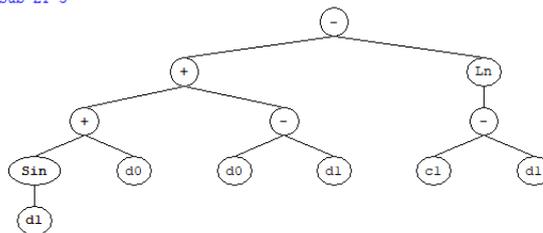
Sub-ET 1



Sub-ET 2



Sub-ET 3

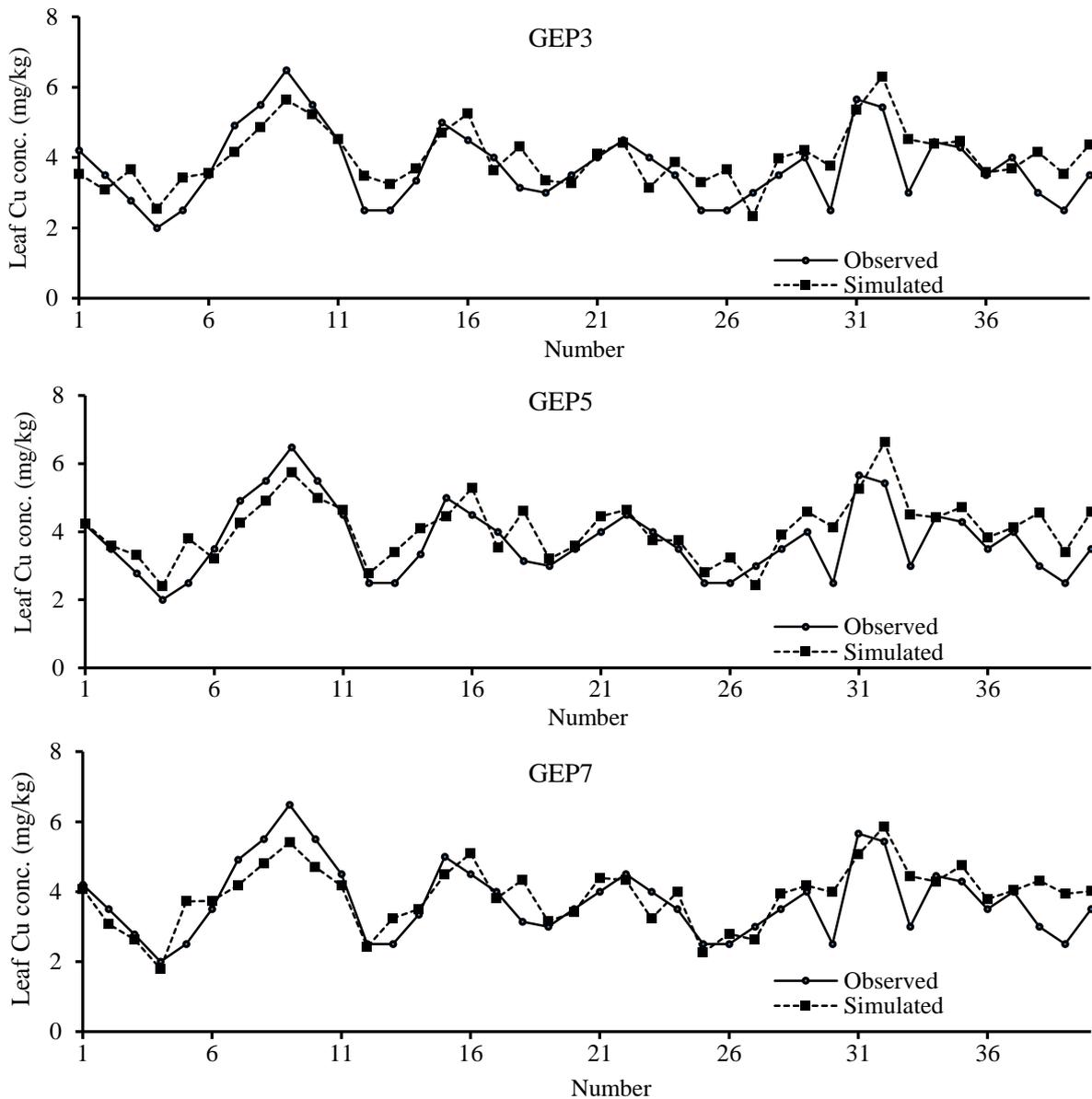


شکل ۲- درخت بیان مدل GEP2 برای تخمین غلظت Mn برگ.

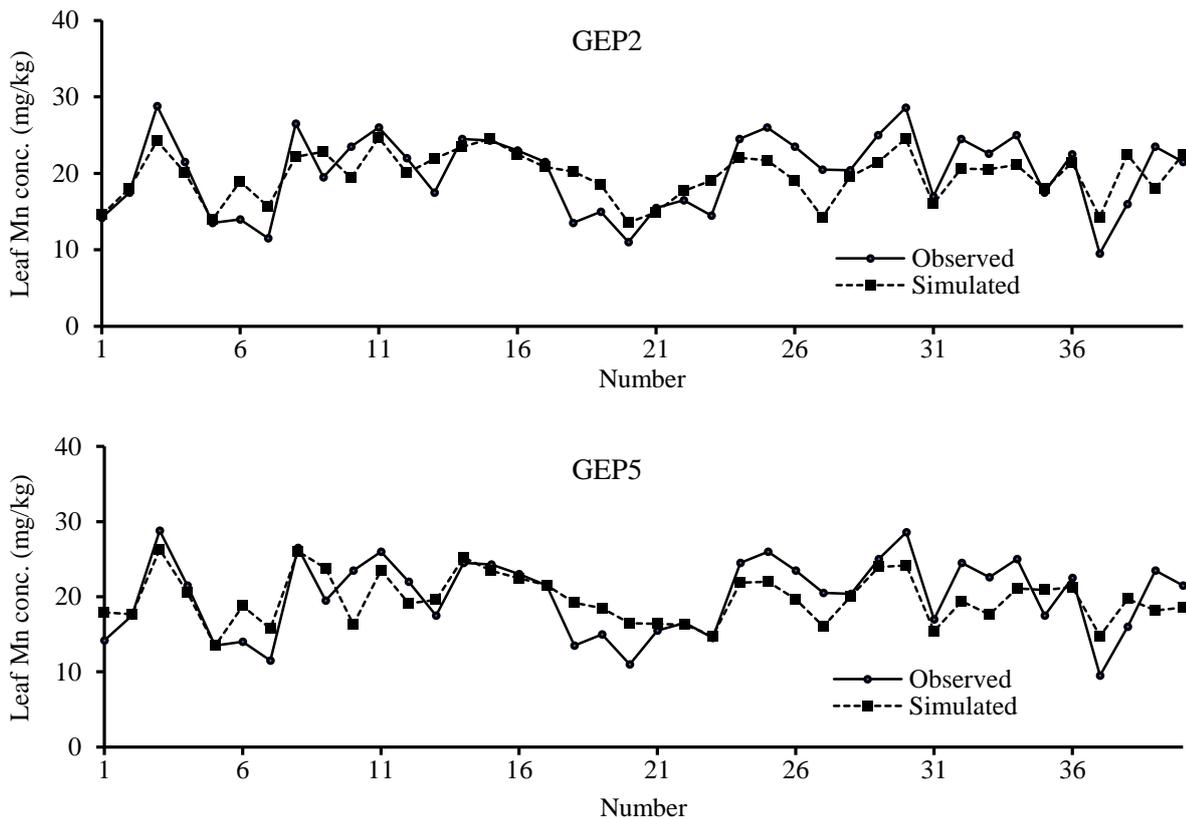
عملکرد مدل‌های بهینه به صورت بصری نیز نمایش داده شد. مقایسه بین مقادیر مشاهده‌شده و شبیه‌سازی‌شده غلظت مس و منگنز برگ مرکبات برای مدل‌های منتخب GEP به ترتیب در شکل‌های ۳ و ۴ نمایش داده شده است.

مقایسه مدل‌ها

برای پاسخ قطعی به این پرسش پژوهشی که کدام رویکرد مدل‌سازی برتر است، عملکرد بهترین مدل‌های رگرسیون و GEP بر روی داده‌های آزمون به‌طور مستقیم مقایسه شد. نتایج برتری مدل‌های GEP را نشان می‌دهد. برای پیش‌بینی غلظت مس برگ، مدل بهینه GEP3 ($R^2=0.60$) بهبودی ۶۷ درصدی در ضریب تبیین نسبت به مدل رگرسیون گام به گام ($R^2=0.36$) ارائه داد. علاوه بر این، خطای پیش‌بینی (RMSE) را ۳۷ درصد کاهش داد و از ۱/۳۰ میلی‌گرم بر کیلوگرم به ۰/۸۲ میلی‌گرم بر کیلوگرم رساند. برای پیش‌بینی غلظت منگنز برگ، اختلاف عملکرد حتی بارزتر بود. مدل GEP2 ($R^2=0.53$) ضریب تبیین را در مقایسه با مدل رگرسیون گام به گام ($R^2=0.28$) تقریباً دو برابر کرد و RMSE را ۳۱ درصد کاهش داد و از ۵/۳۱ میلی‌گرم بر کیلوگرم به ۳/۶۷ میلی‌گرم بر کیلوگرم رساند.



شکل ۳- غلظت مس برگ مرکبات مشاهده شده (Observed) و پیش‌بینی شده (Simulated) با استفاده از مدل‌های GEP برای ترکیب‌های ۳، ۵ و ۷ (خطوط برای تسهیل مقایسه بصری انطباق مدل‌ها رسم شده‌اند).



شکل ۴- غلظت منگنز برگ مرکبات مشاهده شده (Observed) و پیش‌بینی شده (Simulated) با استفاده از مدل‌های GEP برای ترکیب‌های ۲ و ۵ (خطوط برای تسهیل مقایسه بصری انطباق مدل‌ها رسم شده‌اند).

بحث

رگرسیون گام به گام

یکی از ملاحظات کلیدی در تفسیر نتایج حاضر، ضریب تبیین (R^2) نسبتاً پایین مدل رگرسیون، به‌ویژه در پیش‌بینی غلظت منگنز برگ مرکبات است. این موضوع به محدودیت‌های ذاتی مدل‌سازی خطی در به تصویر کشیدن کامل برهمکنش‌های پیچیده و پویای میان ویژگی‌های خاک و تغذیه گیاه اشاره دارد؛ پدیده‌هایی که اغلب ماهیتی غیرخطی و وابسته به شرایط محیطی دارند (Hosseini et al., 2019). افزون بر این، روش رگرسیون گام‌به‌گام، علی‌رغم اعتبار آماری، به دلیل تمرکز بر اثرهای خطی تجمعی، این ریسک را به همراه دارد که متغیرهایی با اثرهای متقابل و غیرخطی پیچیده را از مدل نهایی حذف کند (Sahoo & Jha, 2013).

با وجود این محدودیت‌ها، تحلیل انجام‌شده توانست چندین رابطه کلیدی را که با مبانی علمی و پژوهش‌های پیشین هم‌خوانی دارند، به خوبی روشن سازد. برای مثال، همبستگی منفی مشاهده‌شده میان غلظت فسفر و غلظت مس، مؤید پدیده ناهم‌سازی است. در این فرایند، سطوح بالای فسفر در خاک با تشکیل کمپلکس‌های نامحلول مس-فسفات، عملاً فراهمی و جذب مس به وسیله گیاه را کاهش می‌دهد (Feil et al., 2020). به علاوه، نقش مثبت درصد رس در افزایش فراهمی مس، با توجه به ظرفیت تبادل کاتیونی (CEC) بالای ذرات رس قابل‌توجیه است. این ویژگی به خاک امکان می‌دهد تا عنصرهای غذایی کم‌مصرف را به‌طور مؤثرتری حفظ کرده و به صورت تدریجی در اختیار ریشه گیاه قرار دهد (Fageria et al., 2011).

شاید برجسته‌ترین یافته این پژوهش، تأیید مجدد ارتباط معکوس و نیرومند میان pH خاک و غلظت مس و منگنز در گیاهان مرکبات بود. این نتیجه با بررسی منابع موجود که نشان می‌دهد شرایط اسیدی (pH پایین) حل‌پذیری این عناصر را به‌شدت افزایش داده و در نتیجه زیست‌فراهمی آن‌ها را برای گیاه بهبود می‌بخشد، کاملاً مطابقت دارد (Mhalla *et al.*, 2021). در جمع‌بندی، می‌توان گفت هرچند رگرسیون گام به گام بینش‌های مفید و قابل‌تفسیری، به‌ویژه در مورد اثرهای pH و غلظت فسفر، ارائه داد، اما ناتوانی آن در مدل‌سازی جامع وابستگی‌های متقابل و پیچیده میان ویژگی‌های خاک و دینامیک عنصرهای غذایی، نشان می‌دهد که برای دستیابی به قدرت پیش‌بینی بالاتر، به‌کارگیری یک رویکرد تحلیلی مکمل یا جایگزین، ضروری به نظر می‌رسد.

مدل برنامه‌ریزی بیان ژن

با مقایسه روش رگرسیون گام به گام با الگوریتم تکاملی مدل برنامه‌ریزی بیان ژن (GEP)، فرضیه ما این بود که قابلیت‌های غیرخطی GEP به پیش‌بینی‌های به مراتب دقیق‌تر و قابل‌اعتمادتری منجر خواهد شد. نتایج، این فرضیه را تأیید کرد و محدودیت‌های مدل‌های خطی در ثبت دینامیک پیچیده بیوژئوشیمیایی زیست‌فراهمی عنصرهای غذایی کم‌مصرف در این منطقه را آشکار ساخت. یافته اصلی این پژوهش، عملکرد بهتر مدل‌های GEP نسبت به رگرسیون گام به گام در پیش‌بینی غلظت هر دو عنصر مس و منگنز در برگ مرکبات است. برای غلظت مس برگ، مدل بهینه (GEP3) ضریب تبیین را ۶۷ درصد افزایش داد و خطای پیش‌بینی را ۳۷ درصد در مقایسه با مدل رگرسیون گام به گام کاهش داد. این تفاوت برای غلظت منگنز برگ حتی چشمگیرتر بود، به‌طوری که بهترین مدل (GEP2)، ضریب تبیین را تقریباً دو برابر کرد و خطا را ۳۱ درصد کاهش داد. این جهش کمی در عملکرد، مستقیماً به پرسش اصلی پژوهش پاسخ می‌دهد و تأیید می‌کند که روابط بین ویژگی‌های خاک و غلظت عنصرهای غذایی برگ در این نظام، اساساً غیرخطی است.

این نتیجه با مجموعه مقالات علمی که مزایای رویکردهای یادگیری ماشین و هوش مصنوعی را بر روش‌های آماری سنتی برای مدل‌سازی فرایندهای پیچیده خاک نشان می‌دهند، همخوانی دارد. مطالعاتی که از شبکه‌های عصبی مصنوعی (ANN) و GEP استفاده کرده‌اند، عملکرد برتری را در تخمین ضریب میکروبی خاک (Ebrahimi *et al.*, 2022)، فعالیت آنزیمی خاک (Ebrahimi *et al.*, 2021) و فراهمی سایر عنصرهای غذایی مانند آهن (Heidari, *et al.*, 2024a) و روی (Heidari, *et al.*, 2024b) در شرایط محیطی مشابه گزارش کرده‌اند. در حالی که مدل‌های خطی مانند رگرسیون با فرض اصلی خود یعنی جمع‌پذیری و خطی بودن محدود می‌شوند، نظام خاک-گیاه به اثرهای متقابل هم‌افزا یا متضاد وابسته است و GEP و سایر روش‌های هوش مصنوعی ذاتاً برای ثبت آن‌ها طراحی شده‌اند (Heidari, Vadiati, *et al.*, 2024). بنابراین، برتری آشکار GEP در این مطالعه، صرفاً یک تصادف آماری نیست، بلکه بازتابی از توانایی آن در مدل‌سازی واقع‌گرایانه‌تر پیچیدگی‌های زیربنایی است.

یک مزیت حیاتی GEP که آن را از سایر مدل‌های قدرتمند یادگیری ماشین مانند ANN متمایز می‌کند، تولید معادلات ریاضی شفاف است. در حالی که یک مدل ANN ممکن است به‌دقتی قابل‌مقایسه یا حتی کمی بالاتر دست یابد، منطق درونی آن غیرشفاف باقی می‌ماند (Patel *et al.*, 2022). در مقابل، فرمول‌های صریح تولید شده به‌وسیله مدل‌های GEP3 (معادله ۷) و GEP2 (معادله ۸) یک راه‌حل میانه را ارائه می‌دهند که هم قدرت پیش‌بینی بالا و هم درجه‌ای از تفسیرپذیری را فراهم می‌کند که برای کاربرد عملی و تولید فرضیه‌های جدید بسیار ارزشمند است (Ebrahimi *et al.*, 2021).

مدل‌های GEP نه‌تنها دقت بهتری از خود نشان دادند، بلکه تعداد معدودی از ویژگی‌های خاک را به عنوان حیاتی‌ترین عوامل مؤثر بر زیست‌فراهمی مس و منگنز شناسایی کردند. مدل بهینه برای غلظت مس (GEP3) بر pH، کربن آلی و رس تکیه داشت، در حالی که مؤثرترین مدل برای غلظت منگنز (GEP2) تنها از pH و هدایت الکتریکی

استفاده کرد. این موضوع در تضاد با مدل‌های رگرسیون گام به گام بود که ترکیب متفاوتی از متغیرها را انتخاب کردند (pH، رس و P برای غلظت مس، EC و pH برای غلظت منگنز) که نشان می‌دهد GEP چگونه می‌تواند روابط غیرخطی متفاوت و مسلماً قوی‌تری را شناسایی کند.

شناسایی pH خاک به عنوان یک پیش‌بینی‌کننده کلیدی تقریباً در تمام مدل‌های موفق (هم رگرسیون گام به گام و هم GEP) نقش آن را به عنوان یک متغیر اصلی و کلیدی در کنترل زیست‌فراهمی عنصرهای غذایی کم‌مصرف در منطقه مورد مطالعه تأیید می‌کند. همبستگی منفی قوی یافت‌شده بین pH و هر دو عنصر مس ($r = -0.41$) و منگنز ($r = -0.33$) کاملاً با اصول بیولوژیکی اثبات‌شده حاکم بر خاک‌های قلیایی و آهکی مطابقت دارد. در مقادیر pH بالاتر، حل‌پذیری عنصرهای غذایی کم‌مصرف مانند مس و منگنز به دلیل رسوب آن‌ها به شکل هیدروکسیدهای نامحلول و افزایش جذب سطحی بر روی کلسیت و سطوح رس، به شدت کاهش می‌یابد (Moreno-Lora & Delgado, 2020). این پدیده یک محدودیت مستند در تولید مرکبات در مناطق مشابه (Ahmad *et al.*, 2022) و در سایر نظام‌های کشاورزی است (Li *et al.*, 2015). معادلات GEP (معادلات ۷ و ۸) فراتر از یک همبستگی خطی ساده عمل کرده و این اثر را از طریق توابع پیچیده‌ای شامل عبارات مثلثاتی و لگاریتمی مدل‌سازی می‌کنند که حاکی از تأثیر بسیار غیرخطی و تعاملی pH است.

برجستگی کربن آلی در مدل برتر GEP برای مس (GEP3) نقش حیاتی آن را در تعدیل اثرهای منفی pH بالا نشان می‌دهد. همبستگی مثبت بین کربن آلی و مس برگ ($r = 0.39$) به توانایی ماده آلی در تشکیل کمپلکس‌های آلی-فلزی محلول (کی‌لیت‌ها) نسبت داده می‌شود. این کی‌لیت‌ها از یون‌های مس در برابر رسوب محافظت کرده و آن‌ها را برای جذب به‌وسیله گیاه، فراهم می‌کند (Fu *et al.*, 2004). عدم انتخاب کربن آلی به عنوان یک پیش‌بینی‌کننده معنادار به‌وسیله مدل خطی رگرسیون گام به گام، در حالی که مدل قدرتمندتر GEP آن را انتخاب کرد، قویاً نشان می‌دهد که اثر مفید کربن آلی غیرخطی بوده و به‌ویژه در بستر تعاملات با pH و درصد رس اهمیت می‌یابد.

نقش فسفر (P) قابل‌جذب گیاه در خاک تصویر پیچیده‌تری را ارائه می‌دهد. مدل رگرسیون گام به گام برای غلظت مس، فسفر قابل‌جذب گیاه در خاک را به عنوان یک پیش‌بینی‌کننده منفی معنادار کرد که با تضاد شناخته‌شده بین فسفر و مس مطابقت دارد. سطوح بالای فسفر قابل‌جذب گیاه در خاک می‌تواند با تشکیل رسوبات نامحلول در خاک یا از طریق رقابت فیزیولوژیکی درون گیاه، باعث کمبود عنصرهای غذایی کم‌مصرف شود (Zhang *et al.*, 2012). با این حال، مدل بهینه GEP برای مس (GEP3)، فسفر را حفظ نکرد و در عوض اثرهای تعاملی pH، کربن آلی و رس را در اولویت قرار داد. این نشان می‌دهد که اگرچه یک تضاد خطی برای فسفر وجود دارد، اما اثر آن ممکن است به‌وسیله دینامیک پیچیده‌تر و غیرخطی که به‌وسیله GEP ثبت می‌شود، به‌ویژه نقش محافظتی کربن آلی، تحت‌الشعاع قرار گرفته یا تعدیل شود.

در نهایت، گنجاندن EC به عنوان یک پیش‌بینی‌کننده اصلی برای غلظت منگنز برگ در مدل GEP2، اما نه برای غلظت مس برگ، به حساسیت‌های ویژه هر عنصر به شوری اشاره دارد. رابطه منفی نشان می‌دهد که تنش اسمزی یا عدم تعادل یونی ناشی از شوری بالاتر ممکن است به‌طور نامتناسبی مانع جذب منگنز در مقایسه با مس شود (Ahmad *et al.*, 2022). مدل GEP برای غلظت مس ($R^2 = 0.60$) به‌دقت بالاتری نسبت به غلظت منگنز ($R^2 = 0.53$) دست یافت. این تفاوت احتمالاً به دینامیک ژئوشیمیایی پیچیده‌تر منگنز بازمی‌گردد. فراهمی منگنز نه‌تنها به pH، بلکه به‌شدت به پتانسیل اکسایش-کاهش (ردوکس) خاک نیز وابسته است؛ عاملی که در این مطالعه اندازه‌گیری نشد. شیمی مس پایدارتر و بیشتر تحت کنترل pH و OC است که هر دو در مدل ما گنجانده شده بودند (Koukoulakis *et al.*, 2013).

یافته‌های این پژوهش با مطالعات اخیر انجام شده بر روی سایر عنصرهای غذایی کم‌مصرف در همین منطقه همسو است. حیدری و همکاران (Heidari *et al.*, 2024a) در مدل‌سازی غلظت آهن (Fe) برگ مرکبات گزارش

دادند که مدل‌های غیرخطی (ANN) با ضریب همبستگی $0/81$ به مراتب دقیق‌تر از رگرسیون خطی ($0/20$) عمل کردند و کربن آلی و pH را مهم‌ترین عوامل مؤثر بر جذب آهن معرفی نمودند. همچنین، در مطالعه دیگری بر روی عنصر روی (Zn)، حیدری و همکاران (Heidari et al., 2024b) نشان دادند که مدل‌های هوشمند توانستند ۷۷ درصد از تغییرات روی برگ را بر اساس ویژگی‌های بافت خاک و فسفر پیش‌بینی کنند. مجموع این نتایج در کنار یافته‌های حاضر (برای مس و منگنز)، تأیید می‌کند که رویکردهای داده‌کاوی (ANN و GEP) به دلیل توانایی در ثبت اثر متقابل pH و ماده آلی، راهبردی برتر برای مدیریت تغذیه باغ‌ها در خاک‌های آهکی هستند.

پیامد عملی اولیه این پژوهش دو جنبه دارد. نخست، برای منطقه مورد مطالعه، معادلات صریح (معادلات ۷ و ۸) می‌توانند در نرم‌افزارهای صفحه گسترده تعبیه شوند تا به عنوان یک ابزار غربالگری اولیه برای باغداران عمل کنند. این امر امکان ارزیابی سریع کمبودهای احتمالی مس و منگنز را بر اساس آزمون‌های معمول خاک (pH، کربن آلی، رس، شوری) فراهم کرده و به بهینه‌سازی برنامه‌های کوددهی پیش از ظهور نشانه‌های کمبود کمک می‌کند. دوم، در مقیاسی وسیع‌تر، این مطالعه یک چارچوب روش‌شناختی برای سایر مناطق کشت مرکبات فراهم می‌کند. با اثبات برتری GEP نسبت به رگرسیون خطی در ثبت تعاملات پیچیده خاک-گیاه، این پژوهش، پژوهشگران سایر نقاط را تشویق می‌کند تا رویکردهای مدل‌سازی غیرخطی مشابهی را برای توسعه سامانه‌های پشتیبان تصمیم‌گیری که برای شرایط خاکی خاص خودشان واسنجی شده‌اند، اتخاذ نمایند.

محدودیت‌ها و پژوهش‌های آینده

با وجود بینش‌های ارزشمندی که این مطالعه فراهم کرده است، تبیین محدودیت‌های آن برای ترسیم مسیر تحقیقات آینده ضروری است. نخست، مدل‌های پیش‌بینی حاضر صرفاً بر پایه ویژگی‌های فیزیکی و شیمیایی معمول خاک استوار هستند و پارامترهای تأثیرگذاری همچون زیست‌توده، فعالیت میکروبی و پتانسیل اکسایش-کاهش (ردوکس) را نادیده گرفته‌اند. عدم لحاظ کردن پتانسیل ردوکس، دلیل احتمالی دقت پایین‌تر مدل در پیش‌بینی غلظت منگنز ($R^2=0.53$ برای آزمون) در مقایسه با غلظت مس ($R^2=0.60$) بود.

دوم، توسعه و اعتبارسنجی مدل‌ها محدود به داده‌های منطقه جغرافیایی خاص (جنوب کرمان) و یک رقم معین از مرکبات بوده است. با توجه به اینکه عملکرد مدل‌های برنامه‌ریزی بیان ژن GEP می‌تواند وابسته به مکان باشد، تعمیم‌پذیری این معادلات به سایر مناطق، خاک‌ها و ارقام دیگر نیازمند اعتبارسنجی تجربی است (Shiri et al., 2014). سوم، اگرچه مدل‌ها عملکرد مطلوبی روی زیرمجموعه آزمون داشتند، اما هنوز در برابر یک مجموعه داده کاملاً مستقل (از بازه زمانی یا باغ‌ها متفاوت) سنجیده نشده‌اند، که پیش‌نیاز حیاتی برای پیاده‌سازی عملیاتی در مقیاس گسترده است.

بر مبنای این محدودیت‌ها، چهار پیشنهاد کلیدی برای پژوهش‌های آتی ارائه می‌شود: ۱- توسعه ورودی‌های مدل: مطالعات آینده باید با افزودن داده‌های بیولوژیکی خاک (نظیر زیست‌توده میکروبی و فعالیت‌های آنزیمی) و پتانسیل ردوکس به عنوان متغیرهای ورودی، پتانسیل بهبود دقت پیش‌بینی برای غلظت مس و منگنز را ارزیابی کنند. ۲- اعتبارسنجی خارجی: معادلات بهینه GEP استخراج‌شده در این مطالعه (معادلات ۷ و ۸) باید با استفاده از داده‌های مستقل از سایر مناطق کشت مرکبات (با خاک‌های آهکی) راستی‌آزمایی شوند تا قابلیت تعمیم آن‌ها اثبات گردد. ۳- ارزیابی میدانی: کارآزمایی‌های کنترل‌شده میدانی برای مقایسه راهبردهای کوددهی مبتنی بر GEP با روش‌های رایج کشاورزان ضروری است. این آزمون نهایی مشخص خواهد کرد که آیا این مدل‌ها عملاً منجر به بهبود جذب عنصرهای غذایی، عملکرد و کیفیت میوه می‌شوند یا خیر؟ ۴- ارتقای ساختار مدل: اگرچه GEP استاندارد در این مطالعه عملکرد موفقی داشت، اما پیشنهاد می‌شود که پژوهش‌های آتی برای دستیابی

به دقت‌های بالاتر، از روش‌های پیشرفته‌تری بهره ببرند. به‌کارگیری تنظیم نرخ جهش دینامیکی^۱ می‌تواند از همگرایی زودرس مدل و افتادن در دام بهینه‌های محلی جلوگیری کند. همچنین، استفاده از الگوریتم‌های ترکیبی^۲، مانند ادغام GEP با الگوریتم‌های فراابتکاری دیگر (نظیر بهینه‌سازی ازدحام ذرات^۳)، برای تنظیم دقیق‌تر ضرایب و ثابت‌های عددی در معادلات استخراج‌شده توصیه می‌شود.

نتیجه‌گیری

این پژوهش نشان داد که مدل برنامه‌ریزی بیان ژن ابزاری به مراتب قدرتمندتر و دقیق‌تر از رگرسیون گام به گام سنتی برای پیش‌بینی غلظت مس و منگنز برگ در خاک‌های خشک و آهکی جنوب کرمان است. GEP با مدل‌سازی مؤثر اثرهای حیاتی و غیرخطی ویژگی‌های خاک (به‌ویژه pH، کربن آلی و قابلیت هدایت الکتریکی)، چارچوبی قوی برای غلبه بر محدودیت‌های فرضیات خطی فراهم می‌کند. معادلات صریح و تفسیرپذیر تولید شده به‌وسیله مدل‌های بهینه GEP، زیربنای فنی لازم برای توسعه نرم‌افزارهای کاربردی و ارزان‌قیمت را فراهم می‌کنند. این معادلات می‌توانند به سادگی در قالب برنامه‌های کاربردی در موبایل یا فایل‌های محاسباتی در اختیار باغداران قرار گیرند تا آن‌ها بتوانند تنها با وارد کردن نتایج آزمون خاک، غلظت مس و غلظت منگنز برگ مرکبات در باغ خود را پیش‌بینی کنند. این قابلیت، پژوهش حاضر را از یک مقایسه نظری صرف فراتر برده و آن را به راهکاری عملی برای اصلاح الگوی مصرف کود، کاهش هزینه‌های آزمایشگاهی و مدیریت پایدار باغ‌های مرکبات در این مناطق تبدیل می‌کند.

منابع مورد استفاده

References

- Ahmad, N., Hussain, S., Ali, M. A., Minhas, A., Waheed, W., Danish, S., Fahad, S., Ghafoor, U., Baig, K. S., Sultan, H., Hussain, M. I., Ansari, M. J., Marfo, T. D., & Datta, R. (2022). Correlation of soil characteristics and citrus leaf nutrients contents in current scenario of Layyah District. *Horticulturae*, 8(1), 61. <https://doi.org/10.3390/horticulturae8010061>
- Allison, L., & Richards, L. (1954). *Diagnosis and improvement of saline and alkali soils*. Agriculture Handbook No. 60, Soil and Water Conservative Research Branch, Agricultural Research Service, USDA, Washington, USA.
- Asadi Kangarshahi, A., Fallah Nosratabad, A., & Akhlaghi Amiri, N. (2019). *Guide for sampling and interpretation of soil and leaf analysis results of citrus trees*. Technical Paper No. 561, Soil and Water Research Institute, Karaj, Iran. (in Persian with English abstract)
- Chatzistathis, T., Papaioannou, A., Gasparatos, D., & Molassiotis, A. (2017). From which soil metal fractions Fe, Mn, Zn and Cu are taken up by olive trees (*Olea europaea* L., cv. 'Chondrolia Chalkidikis') in organic groves? *Journal of Environmental Management*, 203, 489-499. <https://doi.org/10.1016/j.jenvman.2017.07.079>
- Ebrahimi, M., Sarikhani, M. R., & Shiri, J. (2022). Application of artificial neural network and gene expression programming to estimate soil microbial metabolic quotient. *Applied Soil Ecology*, 175, 104465. <https://doi.org/10.1016/j.apsoil.2022.104465>
- Ebrahimi, M., Sarikhani, M. R., Shiri, J., & Shahbazi, F. (2021). Modeling soil enzyme activity using easily measured variables: Heuristic alternatives. *Applied Soil Ecology*, 1(157), 103753. <https://doi.org/10.1016/j.apsoil.2020.103753>
- Fageria, N. K., Gheyi, H. R., & Moreira, A. (2011). Nutrient bioavailability in salt affected soils. *Journal of Plant Nutrition*, 34(7), 945-962. <https://doi.org/10.1080/01904167.2011.555578>
- Feil, S. B., Pii, Y., Valentinuzzi, F., Tiziani, R., Mimmo, T., & Cesco, S. (2020). Copper toxicity affects phosphorus uptake mechanisms at molecular and physiological levels in *Cucumis*

¹-Dynamic Mutation

²-Hybrid Models

³-Particle Swarm Optimization (PSO)

- sativus* plants. *Plant Physiology and Biochemistry*, 157, 138-147.
<https://doi.org/10.1016/j.plaphy.2020.10.023>
- Ferreira, C. (2001). Gene expression programming: A new adaptive algorithm for solving problems. *Complex Systems*, 13. <https://doi.org/10.48550/arXiv.cs/0102027>
- Fu, B. J., Liu, S. L., Ma, K. M., & Zhu, Y. G. (2004). Relationships between soil characteristics, topography and plant diversity in a heterogeneous deciduous broad-leaved forest near Beijing, China. *Plant and Soil*, 261(1), 47-54. <https://doi.org/10.1023/B:PLSO.0000035567.97093.48>
- Gunasekaran K, A. K and Sreevardhan P (2025) Real-time soil fertility analysis, crop prediction, and insights using machine learning and deep learning algorithms. *Frontiers in Soil Science*, 5, 1652058. <https://doi.org/10.3389/fsoil.2025.1652058>
- Heidari, S., Ghaffari Nejad, S. A., Sarhadi, J., & Sharif, M. (2024). Modeling the relationship between iron concentration in citrus leaves and some soil properties using artificial neural network (case study of southern Kerman province). *Iranian Journal of Soil and Water Research*, 55(2), 285-296. (in Persian with English abstract)
<https://doi.org/10.22059/ijswr.2024.369507.669619>.
- Heidari, S., Vadiati, M., Ghaffari Nejad, S. A., Sarhadi, J., & Kisi, O. (2024). Modeling Zn availability and uptake by citrus plants using easily measured soil characteristics. *Environmental Modeling & Assessment*, 29(5), 883-900. <https://doi.org/10.1007/s10666-024-09962-0>
- Hosseinfard, S. J., Shirani, H., & Hashemipour, H. (2019). Modeling the relationship between cadmium and some soil physical and chemical properties in pistachio orchards using regression and artificial neural network. *Environmental Sciences*, 17(3), 177-188. (In Persian)
<https://doi.org/10.29252/envs.17.3.177>
- Hosseinpour, M., Sharifi, H., & Sharifi, Y. (2018). Stepwise regression modeling for compressive strength assessment of mortar containing metakaolin. *International Journal of Modelling and Simulation*, 38(4), 207-215.
- Koukoulakis, P., Chatzissavvidis, C., Papadopoulos, A., & Pontikis, D. (2013). Interactions between leaf macro, micronutrients and soil properties in pistachio (*Pistacia vera* L.) orchards. *Acta Botanica Croatica*, 72(2), 295-310.
- Li, Y., Han, M.-Q., Lin, F., Ten, Y., Lin, J., Zhu, D.-H., Guo, P., Weng, Y., & Chen, L.-S. (2015). Soil chemical properties, 'Guanximiyu' pummelo leaf mineral nutrient status and fruit quality in the southern region of Fujian province, China. *Journal of Soil Science and Plant Nutrition*, 15(3), 615-628. <http://dx.doi.org/10.4067/S0718-95162015005000029>
- Mehdizadeh, S., Behmanesh, J., & Khalili, K. (2017). Application of gene expression programming to predict daily dew point temperature. *Applied Thermal Engineering*, 112, 1097-1107.
<https://doi.org/10.1016/j.applthermaleng.2016.10.181>
- Mhalla, B., Ahmed, N., Datta, S. P., Golui, D., Singh, M., & Shrivastava, M. (2021). Solubility relationship of metals in acid soils of kumaon himalaya region of India. *Communications in Soil Science and Plant Analysis*, 52(19), 2373-2387.
<https://doi.org/10.1080/00103624.2021.1928170>
- Moreno-Lora, A., & Delgado, A. (2020). Factors determining Zn availability and uptake by plants in soils developed under Mediterranean climate. *Geoderma*, 376, 114509.
<https://doi.org/10.1016/j.geoderma.2020.114509>
- Najafi N., Parsazadeh M., Tabatabaei S.J., & Oustan S. (2010). Effect of nitrogen form and pH of nutrient solution on the uptake of Fe, Zn, Cu and Mn by spinach plant in hydroponic culture. *Iranian Journal of Soil and Water Research*, 41(2), 283-295. (in Persian with English abstract) <https://dor.isc.ac/dor/20.1001.1.2008479.1389.41.2.16.7>
- Nelson, D. a., & Sommers, L. E. (1983). Total carbon, organic carbon, and organic matter. Pp. 539-579. In: *Methods of soil analysis: Part 2. Chemical and microbiological properties*. ASA, SSSA, USA.
- Olsen, S. R. (1954). *Estimation of available phosphorus in soils by extraction with sodium bicarbonate*. US Department of Agriculture, USA.
- Patel, M. B., Patel, J. N., & Bhilota, U. M. (2022). Comprehensive modelling of ANN. Pp. 31-40. In: *Research anthology on artificial neural network applications*. IGI Global.
<https://doi.org/10.4018/978-1-6684-2408-7.ch002>

- Rahman, R. & Nath Das, K. (2025). Artificial intelligence and machine learning in soil analysis for precision agriculture: a review. *Journal of Experimental Agriculture International*, 47(5), 511–524. <https://doi.org/10.9734/jeai/2025/v47i53440>
- Rengel, Z. (2015). Availability of Mn, Zn and Fe in the rhizosphere. *Journal of Soil Science and Plant Nutrition*, 15(2), 397-409. <http://dx.doi.org/10.4067/S0718-95162015005000036>
- Rowell, D. L. (2014). *Soil science: Methods & applications (1st ed.)*. University of Reading, Routledge, London. <https://doi.org/10.4324/9781315844855>
- Sahoo, S., & Jha, M. K. (2013). Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. *Hydrogeology Journal*, 21(8), 1865-1887. <https://doi.org/10.1007/s10040-013-1029-5>
- Salimi Tarazoj, S., Reyhanitabar A., & Najafi N. (2024) Effects of biochar and phosphorus on dry matter and uptake of calcium, magnesium, iron, zinc, copper, and manganese by rapeseed in a calcareous soil. *Journal of Soil and Plant Science*, 34(4), 91–113. (in Persian with English abstract) <https://doi.org/10.22034/sps.2024.19185>
- Sarhadi, J., heidari, S., & Sharif, M. (2020). The effect of organic, chemical fertilizer and superabsorbant on nutritional status of sure orange rootstock (*Citrus aurantium*). *Horticultural Plants Nutrition*, 2(2), 198-212 (in Persian with English abstract) . <https://doi.org/10.22070/hpn.2020.4840.1047>
- Shiri, J., Sadraddini, A. A., Nazemi, A. H., Kisi, O., Landeras, G., Fard, A. F., & Marti, P. (2014). Generalizability of gene expression programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran. *Journal of hydrology*, 508, 1-11. <https://doi.org/10.1016/j.jhydrol.2013.10.034>
- Vashisth, T., & Kadyampakeni, D. (2020). Diagnosis and management of nutrient constraints in citrus. Pp. 723-737. In: *Fruit crops*. Elsevier.
- Yang, J., Wang, J., Xu, C., Liao, X., & Tao, H. (2022). Modeling the spatial relationship between rice cadmium and soil properties at a regional scale considering confounding effects and spatial heterogeneity. *Chemosphere*, 287, 132402. <https://doi.org/10.1016/j.chemosphere.2021.132402>
- Zhang, Y.-Q., Deng, Y., Chen, R.-Y., Cui, Z.-L., Chen, X.-P., Yost, R., Zhang, F.-S., & Zou, C.-Q. (2012). The reduction in zinc concentration of wheat grain upon increased phosphorus-fertilization and its mitigation by foliar zinc application. *Plant and Soil*, 361(1), 143-152. <https://doi.org/10.1007/s11104-012-1238-z>