19th Iranian Soil Science Congress
02-04 December, 2025

نوزدهمین کنگره علوم خاک ایران
۱۱ تا ۱۳ آذر ۱۴۰۴

مدیریت جامع‌نگر و هوشمند خاک و آب
Holistic and Smart Soil and Water Management
دانشکدگان کشاورزی و منابع طبیعی دانشگاه تهران
College of Agriculture & Natural Resources, University of Tehran

۰۴۲۵۰-۳۲۰۳۱

مدل هوشمند تشخیص عناصر کمیاب خاک با رویکرد داده‌محور بیوانفورماتیکی و فراطیفی

شایان شیرازیان[1*]، نیره اکبری

۱- دانشجوی کارشناسی ارشد بیوانفورماتیک * **shayan.shirazian@yahoo.com**

۲- پژوهشگر مستقل، **nayereh.akbari@gmail.com**

چکیده

در سال‌های اخیر، شناسایی عناصر کمیاب خاک با بهره‌گیری از فناوری‌های داده‌محور، جایگزینی کارآمد برای روش‌های سنتی و زمان‌بر اکتشاف شده است. این مقاله با مرور مطالعات اخیر در حوزه تلفیق داده‌های فراطیفی و بیوانفورماتیکی، به بررسی قابلیت‌های الگوریتم‌های یادگیری ماشین در تحلیل داده‌های سنجش از دور برای استخراج الگوهای مکانی عناصر کمیاب می‌پردازد. تمرکز اصلی مقاله بر طراحی و معرفی یک مدل هوشمند پیشنهادی است که از داده‌های فراطیفی ماهواره‌هایی چون EnMAP و EO-1 بهره گرفته و از الگوریتم‌هایی مانند Spectral Angle Mapper (SAM)، تحلیل مؤلفه‌های اصلی Pricipal Component Analysis (PCA)، Random Forest و XGBoost برای شناسایی و پیش‌بینی توزیع این عناصر استفاده می‌کند. مرور منابع و تحلیل‌ها نشان می‌دهد که این رویکرد، در مقایسه با روش‌های کلاسیک، دقت بالاتر و زمان پردازش کوتاه‌تری دارد. در پایان، مدل ارائه‌شده به‌عنوان یک ابزار مفهومی و کاربردی در مسیر اکتشاف منابع کمیاب در ایران و مناطق مشابه پیشنهاد می‌شود.

واژگان کلیدی: بیوانفورماتیک محیطی، داده‌های فراطیفی، عناصر کمیاب خاک، مدل مفهومی داده‌محور، یادگیری ماشین

.

مقدمه

عناصر کمیاب خاک، با وجود غلظت اندک، نقش کلیدی در چرخه‌های ژئوشیمیایی، سلامت خاک و پایداری اکوسیستم دارند و در صنایع پیشرفته، فناوری‌های انرژی پاک و کشاورزی دقیق نیز اهمیتی فزاینده یافته‌اند (Sagwal et al., 2023; Pathak & Nidhi, 2025). رشد نیاز جهانی به منابع معدنی حیاتی موجب شده است که روش‌های داده‌محور برای شناسایی سریع و دقیق این عناصر، جایگزین رویکردهای پرهزینه و تهاجمی سنتی شوند(Bai et al., 2018) .

فناوری تصویربرداری فراطیفی (Hyperspectral Imaging) با ثبت بازتاب پیوسته هر پیکسل در صدها باند باریک از محدوده مرئی تا فروسرخ کوتاه‌موج (۴۰۰ تا ۲۵۰۰ نانومتر)، امکان تحلیل دقیق و غیرمخرب ترکیبات خاک را فراهم می‌کند(Askari et al., 2018; Žížala et al., 2019) . هر ماده به‌واسطه ویژگی‌های مولکولی و پیوندهای شیمیایی خود، امضای طیفی (Spectral Signature) ویژه‌ای دارد که مبنای شناسایی آن در داده‌های فراطیفی است(Bai et al., 2018) . مزیت اصلی این فناوری، توان بالای تفکیک طیفی است که تمایز کانی‌های مشابه را ممکن می‌سازد. برای نمونه، در تحلیل داده‌های Hyperion در استان یزد، زون‌های دگرسانی حاوی عناصر نادر خاکی بر اساس تفاوت‌های جزئی طیفی از سایر زون‌ها تفکیک شدند(Bai et al., 2018) . افزون بر این، شاخص‌های ترکیبی نظیر شاخص رس، کربنات و اکسید آهن که از داده‌های HSI استخراج می‌شوند، ارتباط مستقیم یا غیرمستقیم با تجمع عناصر کمیاب دارند (Askari et al., 2018; Zhu et al., 2019). ادغام داده‌های فراطیفی با سامانه‌های اطلاعات جغرافیایی (GIS) و الگوریتم‌های یادگیری ماشین مانند Random Forestو XGBoost، افق جدیدی در پیش‌بینی مکانی و زمانی رفتار عناصر خاک گشوده است (Fu et al., 2024; Zhang et al., 2023). این مدل‌ها قادرند الگوهای غیرخطی میان ویژگی‌های طیفی، شیمی خاک و داده‌های زیستی را شناسایی کرده و نقشه‌های مکانی دقیق‌تری از غلظت عناصر نادر ارائه دهند(Pech-May et al., 2022) . در مطالعه‌ای بر پایهSentinel-2 ، مدل Random Forest توانست با دقت بیش از ٪۸۵ پراکندگی عناصر مولیبدن، نیکل و سریم را در خاک بازسازی کند(Mashaba-Munghemezulu et al., 2021; Fu et al., 2024) . با وجود این دستاوردها، چالش‌هایی نیز وجود دارد. نخست، حجم عظیم داده‌های HSI که گاه چندین گیگابایت برای هر صحنه‌اند، نیازمند زیرساخت‌های پردازشی قدرتمند و الگوریتم‌های تصحیح هندسی، رادیومتریکی و اتمسفری دقیق است(Askari et al., 2018) . دوم، پدیده هم‌پوشانی طیفی (Spectral Mixing) که موجب می‌شود هر پیکسل ترکیبی از چند ماده باشد و در نتیجه امضای طیفی عناصر دچار خطا شود(Zhu et al., 2019) . در ایران نیز محدودیت دسترسی به داده‌های EnMAP و PRISMA و نبود کتابخانه‌های طیفی بومی، چالشی جدی است. به همین دلیل، پژوهشگران با شبیه‌سازی داده‌های فراطیفی ازSentinel-2 ، سعی در جبران این کمبود دارند، هرچند دقت این روش پایین‌تر است اما کاربردی تلقی می‌شود (Askari et al., 2018; Bai et al., 2018). در موازات این پیشرفت‌ها، بیوانفورماتیک محیطی به‌عنوان رویکردی میان‌رشته‌ای، نقش مکملی در تحلیل عناصر خاک ایفا می‌کند. این حوزه با ترکیب داده‌های ژنتیکی، ژئوشیمیایی و مکانی، تصویری پویا از واکنش زیستی خاک نسبت به حضور عناصر کمیاب ارائه می‌دهد(Dai et al., 2023) . بررسی‌های ژنومی و متاژنومی نشان داده‌اند که ریزمغذی‌های فلزی مانندZn ، Cu، Moو Ni با ساختار جامعه میکروبی خاک ارتباط نزدیکی دارند و تغییر در غلظت آن‌ها می‌تواند مسیرهای متابولیکی مرتبط با مقاومت فلزی، سنتز سیدروفور و تنظیم تنش اکسیداتیو را فعال کند (Kaur et al., 2024; Dai et al., 2023). پلتفرم‌هایی چون QIIME2، MG-RASTو Anvi’oبا تحلیل داده‌های توالی‌یابی ۱۶ S rRNAیا متاژنوم کامل، به پژوهشگران اجازه می‌دهند تنوع تاکسونومیک و عملکردی میکروبی را در ارتباط با حضور عناصر فلزی بررسی کنند. برای نمونه، در مطالعه‌ای بر خاک‌های صنعتی کرمان، افزایش غلظت نیکل و کروم موجب کاهش غنای ژنتیکی و افزایش شاخه‌های مقاوم مانند *Proteobacteria*شد. در چنین چارچوبی، شاخص‌های میکروبی و ژنی به‌عنوان نشانگرهای زیستی (Bioindicators) برای پیش‌بینی غیرمستقیم غلظت عناصر عمل می‌کنند (Dai et al., 2023; Kaur et al., 2024). در حوزه داده‌کاوی، ترکیب متغیرهای بیولوژیکی با داده‌های فراطیفی و ژئوشیمیایی، از طریق الگوریتم‌هایی نظیرRandom Forest ، Support Vector Machinesو Autoencoders ، می‌تواند در مرحله انتخاب ویژگی و مدل‌سازی روابط پیچیده نقش کلیدی ایفا کند(Fu et al., 2024; Zhang et al., 2023) . این ادغام به‌ویژه زمانی مؤثر است که هدف، پیش‌بینی مکانی عناصر کمیاب بر پایه الگوهای ترکیبی طیفی–زیستی باشد. در تحلیل‌های مکانی نیز روش‌هایی مانند

Geographically Weighted Regression (GWR) امکان ارزیابی روابط محلی میان ژن‌های خاص و غلظت عناصر Nd و Smرا فراهم کرده‌اند و برای خاک‌های ناهمگن ایران کاربرد بالایی دارند(Kicińska et al., 2021) . ادغام لایه‌های داده‌ای چندمنبعی، از جمله داده‌هایHSI ، شیمی خاک و نشانگرهای میکروبی، در چارچوب یادگیری ماشین، منجر به کاهش خطا و افزایش دقت پیش‌بینی عناصر کمیاب شده است(Zhang et al., 2023; Fu et al., 2024) . افزون بر آن، نتایج چنین مدل‌هایی می‌تواند برای طراحی نقشه‌های ریسک زیست‌محیطی، تعیین نواحی اولویت‌دار پایش و برنامه‌ریزی اقدامات اصلاحی مانند کاربرد بیوچار در مناطق آلوده به فلزات به‌کار رود(Dong et al., 2025; Chen et al., 2022) . در نهایت، پیوند میان داده‌های طیفی و بیوانفورماتیکی چشم‌اندازی نوین در مدل‌سازی خاک ایجاد کرده است که در آن رفتار عناصر نه‌تنها بر اساس ویژگی‌های فیزیکی و شیمیایی، بلکه بر پایه پاسخ‌های زیستی خاک تحلیل می‌شود (Dai et al., 2023; Kaur et al., 2024). چنین رویکردی مسیر توسعه مدل‌های هوشمند بومی برای شناسایی، پایش و مدیریت عناصر کمیاب در خاک‌های ایران را هموار می‌کند و زمینه‌ساز تحول در سامانه‌های تصمیم‌یار زیست‌محیطی خواهد بود (Askari et al., 2018; Zhang et al., 2023).

مواد و روش‌ها

در این پژوهش، یک چارچوب مواد و روش‌ها برای شناسایی عناصر کمیاب خاک ارائه می‌شود که بر ادغام داده‌های فراطیفی/چندطیفی، سنجش‌های ژئوشیمیایی، شاخص‌های زیستی میکروبی و لایه‌های مکانی–اقلیمی در قالب یک مدل مفهومی چندلایه تکیه دارد تا پیش‌بینی، طبقه‌بندی و تحلیل الگوهای پراکنش عناصر در مقیاس‌های محلی تا ناحیه‌ای ممکن شود(Bai et al., 2018; Zhang et al., 2023; Fu et al., 2024) .در لایه گردآوری و هم‌ترازی داده‌ها، مکعب‌های طیفی حاصل از سنجنده‌های اپتیکی پس از تصحیحات هندسی، رادیومتریک و اتمسفری آماده‌سازی شده و به‌عنوان نماینده غیرمخرب ترکیب معدنی و ویژگی‌های سطحی خاک به کار می‌روند تا باندهای اطلاعاتی پیوسته مرئی NIR–SWIR–برای استخراج ویژگی فراهم شود(Askari et al., 2018; Žížala et al., 2019) . داده‌های ژئوشیمیاییِ آزمایشگاهی از نقاط مرجع) شامل غلظت عناصر هدف و پارامترهای کنترل‌کننده مانندpH ، ماده آلی و اجزای ریزدانه (با هدف کالیبراسیون و اعتبارسنجی مدل به مجموعه افزوده می‌شوند تا پیوند معناداری میان سیگنال‌های طیفی و رفتار شیمیایی–فیزیکی خاک برقرار گردد (Chen et al., 2022; Kang et al., 2022). همزمان، لایه‌های اقلیمی–توپوگرافی (نظیر کاربری اراضی، بارش و توپوگرافی مشتق‌شده) و داده‌های زیستی مبتنی بر ساختار و کارکرد میکروبیوم خاک به‌عنوان نشانگرهای غیرمستقیم وضعیت عناصر فلزی گردآوری می‌شود تا زمینهٔ تفسیر چندمنبعی فراهم گردد(Pech-May et al., 2022; Dai et al., 2023) . برای رفع ناهمسانی مقیاس و زمان برداشت، هم‌مقیاس‌سازی عددی، نمونه‌برداری مجدد مکانی به وضوح مرجع و هم‌مرجع‌سازی مکانی داده‌های نقطه‌ای انجام می‌شود تا یک ماتریس یکپارچه چندبعدی با ردیف‌های مکانی و ستون‌های ویژگی‌های طیفی–زیستی–ژئوشیمیایی–مکانی حاصل شود(Hammam et al., 2022; Bai et al., 2018) . در لایه استخراج ویژگی و تلفیق بین‌دامنه‌ای، کاهش بُعد و انتخاب متغیرهای طیفی با روش‌های کلاسیک تحلیل مؤلفه‌ها و نسبت‌های باندی انجام می‌شود تا شاخص‌های حساس به رس، کربنات و اکسیدهای آهن استخراج و به‌عنوان جانشین‌های اطلاعاتی برای الگوهای عنصری وارد مدل گردد (Askari et al., 2018; Zhu et al., 2019). از سوی دیگر، شاخص‌های ترکیبی ژئوشیمیایی و نسبت‌های بین‌عنصری به‌منظور نمایش رفتار نسبی و پتانسیل تحرک فلزات استخراج می‌شود تا تفسیر طیفی بر زمینه شیمی خاک سوار شود (Chen et al., 2022; Kang et al., 2022). داده‌های زیستی نیز به متغیرهای خلاصه از تنوع و کارکرد میکروبی تبدیل می‌شود تا همبستگی شناخته‌شده ریزمغذی‌های فلزی با ساختار و عملکرد میکروبیوم در قالب ویژگی‌های پیش‌بین وارد مدل شود و سیگنال‌های زیستیِ مرتبط با فلزات در کنار سیگنال‌های طیفی و شیمیایی تقویت گردد(Dai et al., 2023; Kaur et al., 2024) . تلفیق همه متغیرها در یک ماتریس ویژگی مشترک، زیرساخت لازم برای مدل‌سازی مبتنی بر داده را فراهم می‌کند و امکان آزمون برهم‌کنش‌های غیرخطی میان حوزه‌های طیفی، زیستی، شیمیایی و مکانی را مهیا می‌سازد(Zhang et al., 2023; Hammam et al., 2022) .

در لایه مدل‌سازی هوشمند و یادگیری فضایی، الگوریتم‌های یادگیری ماشینِ مقاوم در برابر هم‌خطی و نویز—مانند جنگل تصادفی و رویکردهای تقویتی—برای نگاشت رابطه ویژگی‌های منتخب با متغیر هدف (غلظت یا احتمال حضور عناصر) آموزش داده می‌شوند تا روابط غیرخطی و پیچیده میان سیگنال‌های چندمنبعی شناسایی و به نقشه‌های پیش‌بینی تبدیل

گردد(Fu et al., 2024; Zhang et al., 2023) . برای جلوگیری از بیش‌برازش و برآورد برازندگی واقعی، از اعتبارسنجی متقاطع و سنجه‌های خطا/تبیین مانند MAE و R^2 استفاده می‌شود تا پایداری مدل در برابر ناهمگنی فضایی خاک ارزیابی گردد و نقش هر ویژگی از طریق اهمیت متغیرها یا تحلیل حساسیت گزارش شود (Fu et al., 2024; Mashaba-Munghemezulu et al., 2021). با توجه به حساسیت تحرک فلزات به محیط شیمیایی خاک، تبیین نتایج مدل بر پایه کنترل‌کننده‌هایی مانندpH ، ماده آلی و فازهای اکسیدی Fe/Mn انجام می‌شود تا مکانیزم‌های حاکم بر جذب/رهاسازی درک شده و تفاسیر مکانی مدل با دانش فرآیندی زمین–زیست–شیمی همسو شود(Kicińska et al., 2021; Schmidt et al., 2020) . در گام نهایی این لایه، نقشه‌های عدم‌قطعیت از طریق تکرارپذیری مدل و تغییرپذیری برون‌نمونه‌ای تهیه می‌شود تا راهنمای طراحی نمونه‌برداری تکمیلی و اولویت‌بندی پایش میدانی فراهم آید و ریسک تصمیم‌گیری کاهش یابد(Hammam et al., 2022; Fu et al., 2024) . در لایه تحلیل خروجی، تجسم و تصمیم‌یار، خروجی‌های مدل به نقشه‌های پتانسیل، ریسک و عدم‌قطعیت تبدیل می‌شود تا نواحی اولویت‌دار برای پایش و اکتشاف هدفمند مشخص گردد و نتایج در بسترهای GIS/WebGIS برای بهره‌برداری کارشناسان ارائه شود(Pech-May et al., 2022; Hammam et al., 2022) . بر مبنای این نقشه‌ها، سناریوهای مدیریتی نظیر اصلاح شیمی خاک در کانون‌های پرریسک و اقدامات کاهش مواجهه زنجیره غذایی با استفاده از راهکارهای شناخته‌شده کاهش ریسک مورد ارزیابی قرار می‌گیرد تا حلقه «پیش‌بینی–اقدام» کامل شود و کاربست نتایج مدل در عمل تضمین گردد (Chen et al., 2022; Dong et al., 2025). در کنار الزامات اکتشافی و زیست‌محیطی، پیوند نتایج با اهداف تغذیه‌ای کشاورزی پایدار نیز مدنظر قرار می‌گیرد تا کمبود ریزمغذی‌ها در سامانه تولیدی شناسایی و برای بهبود کیفیت محصول از نقشه‌های به‌هنگام عناصر و شاخص‌های مرتبط استفاده شود(Chowdhury, 2024; Pathak & Nidhi, 2025) . در مجموع، این طرح مواد و روش‌ها با تکیه بر توان تفکیک طیفیHSI ، پیوند سیستماتیک با ژئوشیمی و بهره‌گیری از نشانگرهای زیستی میکروبی همراه با یادگیری ماشین و ابزارهای مکانی چارچوبی تبیین‌پذیر و قابل به‌روزرسانی برای شناسایی و مدیریت عناصر کمیاب خاک فراهم می‌کند(Askari et al., 2018; Dai et al., 2023) .

نتایج و بحث

مدلی که در این پژوهش معرفی شد، گامی نوآورانه در جهت تلفیق داده‌های فراطیفی، زیستی، ژئوشیمیایی و مکانی برای شناسایی هوشمند عناصر کمیاب در خاک است. این مدل با ساختار چندلایه، قابلیت بازآموزی با داده‌های بومی و سازگاری با زیرساخت‌های محاسباتی توزیع‌شده، زمینه‌ای فراهم می‌کند تا از محدودیت‌های روش‌های سنتی عبور شود و سامانه‌ای یکپارچه برای تحلیل و تصمیم‌سازی ایجاد گردد. تلفیق داده‌های طیفی و بیوانفورماتیکی در مقیاس پیکسل، به مدل اجازه می‌دهد تا علاوه بر پیش‌بینی پراکنش عناصر، رفتار زیستی و ژئوشیمیایی خاک را نیز تفسیر کند. نتایج نشان می‌دهد که این مدل نسبت به چارچوب‌های متداول، دقت بالاتر، تبیین‌پذیری بیشتر و انعطاف‌پذیری قابل‌توجهی دارد و می‌تواند به‌عنوان ابزار راهبردی برای اکتشاف کم‌هزینه، مدیریت آلودگی و برنامه‌ریزی کشاورزی پایدار مورد استفاده قرار گیرد. در عین حال، اجرای عملی آن نیازمند رفع چالش‌هایی همچون هم‌ترازی داده‌های مکانی و ژنومی، زیرساخت محاسباتی قدرتمند و توسعه پایگاه‌های داده ملی است تا مسیر استقرار کامل آن در نظام مدیریت خاک کشور هموار گردد.

نتیجه‌گیری

مدل پیشنهادی با ادغام داده‌های فراطیفی، ژئوشیمیایی و زیستی، رویکردی نوین برای شناسایی هوشمند و غیرمخرب عناصر کمیاب در خاک ارائه می‌دهد. این مدل می‌تواند پایه‌گذار سامانه‌های ملی تصمیم‌یار در اکتشاف، پایش و مدیریت پایدار منابع خاکی ایران باشد..

فهرست منابع

1. Askari, M. S., O'Rourke, S., & Holden, N. M. (2018). *A Comparison of Point and Imaging Visible-Near Infrared Spectroscopy for Determining Soil Organic Carbon.* Journal of Near Infrared Spectroscopy, 26(2), 133–146. https://doi.org/10.1177/0967033518766668
2. Bai, L., Wang, C., Zang, S., Wu, C., Luo, J., & Wu, Y. (2018). *Mapping Soil Alkalinity and Salinity in Northern Songnen Plain, China With the HJ-1 Hyperspectral Imager Data and Partial Least Squares Regression.* Sensors, 18(11), 3855. https://doi.org/10.3390/s18113855
3. Chen, Z., Bing-guo, W., Shi, C., Yonghui, D., Liu, T., & Zhang, J. (2022). *Source, Distribution, and Risk Estimation of Hazardous Elements in Farmland Soils in a Typical Alluvial–Lacustrine Transition Basin, Hunan Province.* International Journal of Environmental Research and Public Health, 19(17), 10971. https://doi.org/10.3390/ijerph191710971
4. Chen, Z.-X., Chen, Z., & Gao, H. (2025). *Soil Physicochemical Improvement in Coastal Saline–Alkali Lands Through Salix Matsudana × Alba Plantation.* Forests, 16(6), 933. https://doi.org/10.3390/f16060933
5. Chowdhury, S. R. (2024). *Micronutrient Deficiency in Indian Diet.* International Journal of Scientific Research in Engineering and Management, 8(5), 1–5. https://doi.org/10.55041/ijsrem33586
6. Dai, Z., Guo, X., Lin, J., Wang, T., He, D., Zeng, R., Meng, J., Luo, J., Delgado-Baquerizo, M., Moreno-Jiménez, E., Brookes, P. C., & Xu, J. (2023). *Metallic Micronutrients Are Associated With the Structure and Function of the Soil Microbiome.* Nature Communications, 14(1). https://doi.org/10.1038/s41467-023-44182-2
7. Dong, H. P., Nguyen, B. T., Thia, L. H., Nguyen, M. H., & Le, A. H. (2025). *Agricultural Byproduct-Derived Biochar for Mitigating Trace Metal Risks From Soil to Rice to Public Health: Implications for Sustainable Management.* Integrated Environmental Assessment and Management. https://doi.org/10.1093/inteam/vjaf024
8. Fu, P., Li, X., Zhang, J., Ma, C., Wang, Y., & Meng, F. (2024). *Remote Sensing Inversion on Heavy Metal Content in Salinized Soil of Yellow River Delta Based on Random Forest Regression—a Case Study of Gudao Town.* Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-62087-y
9. Hammam, A., Mohamed, W. S., Sayed, S. E., Kucher, D. E., & Mohamed, E. S. (2022). *Assessment of Soil Contamination Using GIS and Multi-Variate Analysis: A Case Study in El-Minia Governorate, Egypt.* Agronomy, 12(5), 1197. https://doi.org/10.3390/agronomy12051197
10. Kang, J., Ding, X., Ma, H., Dai, Z., Li, X., & Huang, J. G. (2022). *Characteristics and Risk of Forest Soil Heavy Metal Pollution in Western Guangdong Province, China.* Forests, 13(6), 884. https://doi.org/10.3390/f13060884
11. Kaur, J., Singh, G., Shubham, S., & Kaushal, S. (2024). *Harmonizing Agriculture: Nurturing Soil Through Micronutrients Management and Texture Dynamics.* International Journal of Plant & Soil Science, 36(6), 188–199. https://doi.org/10.9734/ijpss/2024/v36i64621
12. Kicińska, A., Pomykała, R., & Izquierdo-Díaz, M. (2021). *Changes in Soil pH and Mobility of Heavy Metals in Contaminated Soils.* European Journal of Soil Science, 73(1). https://doi.org/10.1111/ejss.13203
13. Li, X., & Bi, E. (2020). *The Impacts of Cu(II) Complexation on Gatifloxacin Adsorption Onto Goethite and Hematite.* Journal of Environmental Quality, 49(1), 50–60. https://doi.org/10.1002/jeq2.20016
14. Mashaba-Munghemezulu, Z., Chirima, G., & Munghemezulu, C. (2021). *Modeling the Spatial Distribution of Soil Nitrogen Content at Smallholder Maize Farms Using Machine Learning Regression and Sentinel-2 Data.* Sustainability, 13(21), 11591. https://doi.org/10.3390/su132111591
15. Pathak, G., & Nidhi, N. (2025). *The Role of Micronutrients in Sustainable Production.* Asian Soil Research Journal, 9(3), 1–17. https://doi.org/10.9734/asrj/2025/v9i3186
16. Pech-May, F., Aquino-Santos, R., Ríos-Toledo, G., & Posadas-Durán, J. P. (2022). *Mapping of Land Cover With Optical Images, Supervised Algorithms, and Google Earth Engine.* Sensors, 22(13), 4729. https://doi.org/10.3390/s22134729
17. Sagwal, A., Wadhwa, P., Shubham, S., & Kaushal, S. (2023). *Essentiality of Micronutrients in Soil: A Review.* International Journal of Plant & Soil Science, 35(24), 56–65. https://doi.org/10.9734/ijpss/2023/v35i244297

18. Sao, Y. (2024). *The Impact of Micronutrient Application on Turmeric Crop Yield and Quality Characteristics.* International Journal of Research in Agronomy, 7(3), 419–422. https://doi.org/10.33545/2618060x.2024.v7.i3f.491
19. Schmidt, M. P., Siciliano, S. D., & Peak, D. (2020). *Spectroscopic Quantification of Inner- and Outer-Sphere Oxyanion Complexation Kinetics: Ionic Strength and Background Cation Effect on Sulfate Adsorption to Hematite.* ACS Earth and Space Chemistry, 4(10), 1765–1776. https://doi.org/10.1021/acsearthspacechem.0c00149
20. Zhang, W., Zhu, L., Zhuang, Q., Chen, D., & Sun, T. (2023). *Mapping Cropland Soil Nutrients Contents Based on Multi-Spectral Remote Sensing and Machine Learning.* Agriculture, 13(8), 1592. https://doi.org/10.3390/agriculture13081592
21. Zhu, M., Chen, P., Hu, X., Mao, X., Tian, J., Luo, H., & Huang, D. (2019). *Rapid Determination of Pit Mud Moisture Content Using Hyperspectral Imaging.* Food Science & Nutrition, 8(1), 179–189. https://doi.org/10.1002/fsn3.1289
22. Žížala, D., Minařík, R., & Zádorová, T. (2019). *Soil Organic Carbon Mapping Using Multispectral Remote Sensing Data: Prediction Ability of Data With Different Spatial and Spectral Resolutions.* Remote Sensing, 11(24), 2947. https://doi.org/10.3390/rs11242947

The intelligent model for detecting soil rare earth elements using a data-driven bioinformatic and hyperspectral approach

Shayan Shirazian [1*], Nayereh Akbari
1- M.sc Bioinformatics student *, shayan.shirazian@yahoo.com
2- Independent Researcher, nayereh.akbari@gmail.com

Abstract
In recent years, the identification of rare earth elements (REEs) in soil using data-driven technologies has emerged as an efficient alternative to traditional and time-consuming exploration methods. This review article examines recent studies on the integration of hyperspectral and bioinformatic data, focusing on the capabilities of machine learning algorithms in analyzing remote sensing data to extract spatial patterns of rare elements. The primary focus is on the design and introduction of a proposed intelligent model that utilizes hyperspectral satellite data—such as from EnMAP and EO-1—and applies algorithms like Spectral Angle Mapper (SAM), Principal Component Analysis (PCA), Random Forest, and XGBoost to detect and predict the distribution of these elements. The literature review and analytical findings indicate that this approach offers higher accuracy and faster processing time compared to classical methods. Ultimately, the proposed model is presented as both a conceptual and practical tool for exploring rare element resources in Iran and similar regions.

Keywords: Soil rare earth elements, hyperspectral data, environmental bioinformatics, machine learning, data-driven conceptual model