

تهیه نقشه رقومی شکل‌های مختلف آهن خاک با استفاده از داده‌های سنجنده OLI ماهواره لندست در ساحل شرقی دریاچه ارومیه

امین موسوی^{۱*}، فرزین شهبازی^۲، شاهین اوستان^۳، علی اصغر جعفرزاده^۴، بودیمن میناسنی^۵

(تاریخ دریافت: ۱۳۹۸/۱۱/۲۴ تاریخ پذیرش: ۱۳۹۹/۰۶/۰۸)

چکیده

در این مطالعه با استفاده از دو روش داده‌کاوی شامل مدل درخت تصمیم‌گیری (DT) و مدل کیوبیست (Cu) نقشه رقومی مهم‌ترین شکل‌های آهن ۱۳۱ نمونه خاک سطحی (عمق صفر تا ۱۰ سانتی‌متری) شامل آهن معادل کل (Fe_t)، آهن پدوژنیک (Fe_d) و آهن بی‌شکل (Fe_o) در منطقه‌ای به مساحت ۵۰۰ کیلومترمربع از دو سایت جداگانه در بستر خشک شده ساحل شرقی دریاچه ارومیه تهیه شد. در این پژوهش در مجموع تعداد ۱۹ متغیر کمکی برگرفته از تصویر سنجنده OLI ماهواره لندست ۸ مربوط به تیرماه سال ۱۳۹۶ مورد استفاده قرار گرفت. نتایج نشان داد که مدل کیوبیست با داشتن مقادیر (RMSE = ۰/۸۹ و R² = ۰/۸۸) و (RMSE = ۰/۵۷ و R² = ۰/۸۵) برای پیش‌بینی Fe_t، (RMSE = ۲/۲۵ و R² = ۰/۸۹) برای پیش‌بینی Fe_o دارای دقت بالاتری نسبت به مدل درخت تصمیم‌گیری به‌منظور پیش‌بینی هر سه شکل آهن داشت. هم‌چنین نتایج میزان اهمیت و درصد مشارکت متغیرهای کمکی در هر دو مدل نشان‌دهنده اهمیت بالای برخی شاخص‌های طیفی از جمله شاخص نسبت رطوبتی نرمال شده (NDMI) و شاخص اصلاح شده گیاهی تعدیل‌کننده اثر خاک (MSAVI) در پیش‌بینی Fe_t، Fe_d و Fe_o می‌باشد. به‌طور کلی نتایج نشان داد که مدل کیوبیست در مقایسه با مدل درخت تصمیم‌گیری دارای توانایی و کارایی بالاتری در مدل‌سازی و تخمین پراکنش مکانی شکل‌های مختلف آهن خاک در منطقه مورد مطالعه بوده است.

واژه‌های کلیدی: داده‌کاوی، درخت تصمیم‌گیری، شکل‌های مختلف آهن، کیوبیست، متغیرهای کمکی

موسوی ا.، شهبازی ف.، اوستان ش.، جعفرزاده ع.ا.، میناسنی ب. ۱۴۰۰. تهیه نقشه رقومی شکل‌های مختلف آهن خاک با استفاده از داده‌های سنجنده OLI ماهواره لندست در ساحل شرقی دریاچه ارومیه. تحقیقات کاربردی خاک. جلد ۹، شماره ۲. صفحه: ۴۷-۶۰.

۱- دانشجوی دکترای گروه علوم و مهندسی خاک، دانشکده کشاورزی دانشگاه تبریز (مکاتبه کننده)

۲- دانشیار گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه تبریز

۳- استاد گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه تبریز

۴- استاد گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه تبریز

۵- استاد، دانشکده علوم محیطی و کشاورزی، دانشگاه سیدنی، استرالیا

* پست الکترونیک: amin_mousavi@alumni.ut.ac.ir

مقدمه

سانتی‌متر) گزارش کردند. در سال‌های اخیر، شهبازی و همکاران (Shahbazi *et al.*, 2019a) با استفاده از تکنیک‌های داده‌کاوی مانند مدل‌های جنگل تصادفی^۵ و کیوبیست^۶ و به کمک مشتقات مدل رقمی ارتفاع^۷ و تصاویر ماهواره‌ای لندست ۸ نقشه‌های پراکنش مکانی و عمودی برخی عناصر مهم غذایی مانند نیتروژن، فسفر و بور را برای منطقه‌ای از شمال غرب ایران تهیه کردند و ضمن گزارش ارجحیت مدل کیوبیست، مشاهده کردند که قسمت‌های غربی محدوده مطالعاتی مستعد به آلودگی بور می‌باشد. در طی تحقیقی دیگر، شهبازی و همکاران (Shahbazi *et al.*, 2019b) با استفاده از برخی تکنیک‌های داده‌کاوی و به کمک شاخص‌های دورسنجی نقشه رقمی آهن متبلور را در ساحل شرقی دریاچه ارومیه تهیه و گزارش نمودند که مدل رگرسیون خطی چندگانه برای پیش‌بینی مقدار رس خاک و مدل کیوبیست برای پیش‌بینی تراکم آهن کل خاک و اکسید آهن متبلور مناسب می‌باشد. سوماراتنا و همکاران (Somarathna *et al.*, 2016) در استرالیا در منطقه‌ای به وسعت ۸۱۰ هزار کیلومتر مربع، کربن آلی خاک را با چهار روش رگرسیون خطی چندگانه^۸، درخت تصمیم‌گیری، مدل کیوبیست و رگرسیون بردار پشتیبان^۹ در چهار عمق جداگانه مدل‌سازی و گزارش نمودند که مدل رگرسیون بردار پشتیبان اجرای بهتری نسبت به مدل‌های دیگر در همه عمق‌ها داشته است. مصلح و همکاران (Mosleh *et al.*, 2016) چهار تکنیک شبکه عصبی مصنوعی، درخت رگرسیون توسعه‌یافته^۱، مدل خطی تعمیم‌یافته^۲ و مدل خطی رگرسیون^۳ را برای تخمین خصوصیات خاک در مناطقی با پستی و بلندی کم استفاده کردند. آن‌ها مشاهده کردند که دقت مدل خطی تعمیم‌یافته برای تخمین اکثر خصوصیات خاک بیش‌تر بوده است. فاتحی و همکاران (Fatehi *et al.*, 2016) ریز مقیاس‌سازی نقشه کربن آلی خاک از اندازه پیکسل ۵۰ متری به اندازه پیکسل ۱۰ متری را با استفاده از نقشه رقمی خاک در استان-

برنامه‌ریزی دقیق و مدیریت اراضی به‌منظور بهره‌وری پایدار اراضی و هم‌چنین شناخت تغییرات مکانی و زمانی در خاک امری اجتناب‌ناپذیر است. در این میان، مطالعه توزیع و مقدار شکل‌های مختلف آهن در خاک که حدود پنج درصد از پوسته جامد زمین را تشکیل می‌دهد، می‌تواند شاخصی برای ارزیابی وضعیت تکامل خاک باشد (Stonehouse & Arnaud, 1971). در این میان به‌منظور شناسایی دقیق و تعیین الگوی پراکنش منابع خاک به شکل ساده و قابل فهم و تفسیر برای کاربران مختلف نقش نقشه‌برداری رقمی خاک^۱ بیش از پیش روشن می‌گردد. نقشه‌برداری رقمی خاک شامل روش‌ها و مدل‌هایی است که بین توزیع خاک، ویژگی‌های خاک و داده‌هایی که به آسانی و با صرف هزینه و وقت کم‌تر و با دقت بیش‌تر از تصاویر ماهواره‌ای به‌دست می‌آیند و تحت عنوان متغیرهای کمکی نامیده می‌شوند، ارتباط برقرار می‌کند. در این میان به‌دلیل برخی مشکلات و محدودیت‌های موجود در نقشه‌برداری سنتی (مرسوم) خاک از جمله داشتن سرعت و دقت اندک، هزینه بالا، عدم قطعیت و کارایی لازم و هم‌چنین ناتوانی داده‌ها و اطلاعات حاصل از نقشه‌های سنتی خاک در تشریح ساختار و ویژگی‌های دینامیکی و پیوسته اراضی (Zhu *et al.*, 2001) در سال‌های اخیر به‌منظور ساده‌سازی روابط پیچیده موجود در خاک تکنیک‌ها و مدل‌های مختلفی از نقشه‌برداری رقمی خاک مانند داده‌کاوی^۲ از جمله روش‌های ماشین یادگیری^۳ با تلفیق داده‌های مختلف برای پیش‌بینی و برآورد ویژگی‌های خاک بر اساس اطلاعات مکانی با تفکیک مکانی بالا مورد استفاده قرار گرفته است. تقی‌زاده مهرجردی و همکاران (Taghizadeh-Mehrjardi *et al.*, 2014) در منطقه اردکان با استفاده از روش درخت تصمیم‌گیری^۴ و نقشه ژئومورفولوژی به‌عنوان متغیر محیطی، نقشه شوری خاک در عمق‌های مختلف خاک را تهیه کرده و دقت این روش را بین ۷۸ درصد (برای عمق صفر تا ۱۰ سانتی‌متر) تا ۱۱ درصد (برای عمق ۶۰ تا ۱۰۰

8 . Multiple linear regression

9 . Support Vector Regression (SVR)

1 . Artificial Neural Network (ANN) 0

1 . Boosting Regression Tree (BRT) 1

1 . Generalized linear model 2

1 . Linear regression model 3

1 . Digital Soil Mapping (DSM)

2 . Data Mining

3 . Machine learning

4 . Decision Tree (DT)

5 . Random Forests (RF)

6 . Cubist (Cu)

7 . Digital Elevation Model (DEM)

نمونه‌برداری و تجزیه‌های متداول آزمایشگاهی

در این مرحله ابتدا محدوده دو سایت مورد نظر (سایت‌های یک و دو) به مساحت ۲۲۳ و ۲۷۵ کیلومترمربع واقع در ساحل شرقی دریاچه ارومیه در محیط سامانه اطلاعات جغرافیایی تعیین گردید. سپس به روش نمونه‌برداری تصادفی مرتب شده^۳ (McKenzie *et al.*, 2008) تعداد ۶۰ و ۷۱ نمونه به ترتیب از سایت‌های یک و دو و در مجموع تعداد ۱۳۱ نمونه خاک سطحی از عمق صفر تا ۱۰ سانتی‌متری از کل منطقه مورد مطالعه برداشت گردید. تمامی نمونه‌های مورد نظر به آزمایشگاه منتقل و پس از هواخشک کردن، از الک دو میلی‌متری عبور داده شده و شکل‌های مختلف آهن تمام نمونه‌های خاک شامل آهن معادل کل (Fe_t) توسط هضم نمونه‌های خاک به روش آکوارجیا طبق روش مک‌گراس و کواینلیفی (McGrath & Cunliffe, 1985)، آهن پدوژنیک (آزاد) قابل استخراج با سیترا-دی تیونات-بیکربنات (Fe_d) و آهن بی‌شکل قابل استخراج با اگزالات آمونیوم (Fe_o) به روش لوئپر و اینسکپ (Loeppert & Inskeep, 1996) تعیین گردید و سپس اندازه‌گیری غلظت آهن در نمونه‌ها با دستگاه جذب اتمی شیماتزو مدل (AA-۶۳۰۰) صورت پذیرفت.

تحلیل آماری داده‌ها

قبل از انجام تجزیه‌های آماری و مدل‌سازی، آزمون نرمال بودن توزیع داده‌های حاصل از شکل‌های مختلف آهن با استفاده از روش کولموگروف-اسمیرنوف‌گر نرم‌افزار SPSS انجام گرفت و به دلیل نداشتن توزیع نرمال، با استفاده از تبدیل لگاریتمی داده‌ها نرمال شدند.

داده‌های کمکی دورسنجی

گام بعدی برای تهیه نقشه رقومی پراکنش شکل‌های مختلف آهن خاک در منطقه مورد مطالعه جمع‌آوری مجموعه‌ای از داده‌های کمکی است که بدین منظور از داده‌های کمکی برگرفته شده از تصاویر ماهوره‌ای سنجنده OLI^۴ نلدست ۸ با قدرت تفکیک مکانی ۳۰ متر و مربوط به ماه جولای ۲۰۱۷ (تیرماه ۱۳۹۶) به‌عنوان داده‌های طیفی برای انجام مدل‌سازی و نقشه‌برداری رقومی شکل‌های مختلف آهن استفاده شد.

کرمانشاه انجام دادند و دریافتند که روش درخت تصمیم-گیری دارای صحت و کارایی بالاتری نسبت به مدل‌های خطی تصمیم‌یافته و شبکه عصبی مصنوعی بود. ور و همکاران (Were *et al.*, 2015) با مقایسه چهار روش رگرسیون بردار پشتیبان، شبکه عصبی مصنوعی، جنگل تصادفی و درخت تصمیم‌گیری را برای نقشه‌برداری رقومی ذخیره کربن آلی خاک استفاده نمودند. نتایج آن‌ها دقت بالاتر نقشه‌های تهیه شده با روش درخت تصمیم-گیری را نشان داد و تأکیدی بر امکان استفاده از روش‌های مختلف نقشه‌برداری رقومی در تهیه نقشه ویژگی‌های خاک بود. از آنجایی که تاکنون مطالعات چندانی در مورد استفاده از روش نقشه‌برداری رقومی خاک با استفاده از تکنیک‌های مختلف داده‌کاوی مانند درخت تصمیم‌گیری و کوبیست برای پیش‌بینی شکل‌های مختلف آهن در اراضی مسطح و ناهموار ایران صورت نگرفته است و با توجه به اهمیت روش‌های درخت تصمیم‌گیری و کوبیست این پژوهش با هدف مقایسه و ارزیابی قابلیت هر یک از این مدل‌ها برای پیش‌بینی مقدار و تغییرات شکل‌های مختلف آهن خاک شامل (Fe_o، Fe_d، Fe_t) در ساحل شرقی دریاچه ارومیه انجام گرفت.

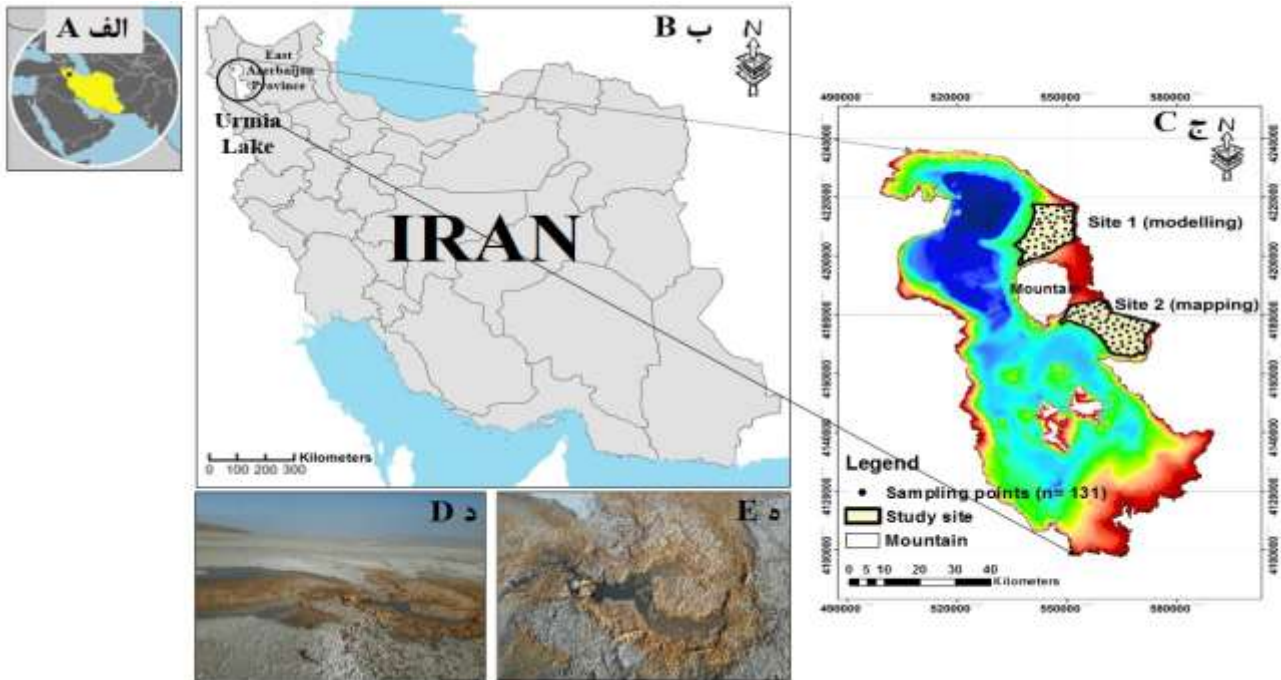
مواد و روش‌ها

مشخصات منطقه مورد مطالعه

این مطالعه در دو سایت مطالعاتی و در منطقه‌ای به وسعت حدود ۵۰۰ کیلومترمربع و بین مختصات جغرافیایی "۳۷° ۱۴' ۵۸" تا "۳۸° ۷' ۱۲" عرض شمالی و "۱۸° ۲۲' ۱۸" تا "۴۵° ۷' ۱۷" طول شرقی ساحل شرقی دریاچه ارومیه واقع شده در استان آذربایجان شرقی انجام شد (شکل ۱). اراضی این منطقه به‌طور گسترده متشکل از باتلاق و لجن‌زار بوده و به‌عنوان عرصه‌ای برای تجمعات اکسیدهای آهن محسوب می‌شوند. رژیم‌های رطوبتی و حرارتی خاک منطقه به ترتیب زریک و مزیک می‌باشد. هم‌چنین منطقه مورد مطالعه دارای ۳۴۱ میلی‌متر بارندگی در سال است و هم‌چنین حداقل و حداکثر دما در این منطقه بین صفر تا ۲۳- درجه سلسیوس در زمستان و در تابستان تا ۳۹ درجه سلسیوس نوسان دارد (IRIMO, 2012).

3 . Kolmogorov.Smirnov
4 . Operational Land Imagery (OLI)

1. Geographic Information System (GIS)
2. Stratified random sampling



شکل ۱- الف و ب- موقعیت دریاچه ارومیه و محدوده مطالعاتی در ایران، ج- محدوده مناطق مطالعاتی انتخاب شده برای نقشه- برداری و مدل‌سازی در ساحل شرقی دریاچه ارومیه همراه با نقاط نمونه‌برداری (n=131)، د و ه- تجمعات آهن در سطح منطقه مورد مطالعه

Figure 1. A and B: The location of Urmia Lake in Iran; C: The locations of the study area for mapping and modelling in the east shore of Urmia Lake (n= 131); D and E: Accumulated iron in the surface of the study area

کمی شدت سوختن اراضی یک^۸ (Parks *et al.*, 2014)، شاخص کمی شدت سوختن اراضی دو^۹ (Parks *et al.*, 2014) و شاخص‌های وابسته به زمین‌شناسی و آهن شامل شاخص نسبت شوری^{۱۰} (Metternicht & Zinck, 2003)، شاخص رس^{۱۱} (Boettinger *et al.*, 2008)، شاخص اکسیدهای آهن^{۱۲} (Drury, 2016) و شاخص کانی‌های آهن فروس (دو ظرفیتی)^{۱۳} (Drury, 2016) با استفاده از ابزار Raster calculator در نرم‌افزار ArcGIS 10.2 (ESRI, 2011) محاسبه و تهیه گردیدند.

مدل‌سازی شکل‌های مختلف آهن با استفاده از داده- کاوی

در این پژوهش از دو مدل کاربردی و متداول داده‌کاوی برای پیش‌بینی و نقشه‌برداری رقومی شکل‌های مختلف

در این بخش علاوه بر داده‌های حاصل از باندهای منفرد از چهار دسته از مهم‌ترین شاخص‌های دورسنجی و متغیرهای کمی مشتق شده از ترکیب باندهای منفرد ۲ تا ۷ سنجنده OLI ماهواره لندست ۸ و همچنین تجزیه به مؤلفه‌های اصلی تعداد شش باند منفرد^۱ مربوط به منطقه مورد مطالعه از جمله شاخص‌های وابسته به پوشش گیاهی و خاک شامل شاخص نسبت گیاهی نرمال شده^۲ (Boettinger *et al.*, 2008)، شاخص مرئی مقاومت اتمسفریک^۳ (Gitelson *et al.*, 2002)، شاخص گیاهی تعدیل‌کننده اثر خاک^۴ (Gilbert *et al.*, 2002)، شاخص اصلاح شده گیاهی تعدیل‌کننده اثر خاک^۵ (Qi *et al.*, 1994)، شاخص‌های وابسته به آب شامل شاخص نسبت رطوبتی نرمال شده^۶ (Skakun *et al.*, 2003)، شاخص اصلاح شده نسبت رطوبتی نرمال شده^۷ (Xu *et al.*, 2006)، شاخص‌های وابسته به زمین‌ما شامل شاخص

7. Modified Normalized Difference Water Index (MNDWI)
8. Normalized Burn Ratio (NBR)
9. Normalized Burn Ratio2 (NBR2)
1. Salinity Ratio (SR) 0
1. Clay Index (CI) 1
1. Iron Oxide Index (IOI) 2
1. Ferrous Minerals Index (FMI) 3

1. Principal component analysis of six individual bands (PCA)
2. Normalized Difference Vegetation Index (NDVI)
3. Visible Atmospherically Resistant Index (VARI)
4. Soil Adjusted Vegetation Index (SAVI)
5. Modified Soil Adjusted Vegetation Index (MSAVI)
6. Normalized Difference Moisture Index (NDMI)

اعتبارسنجی و ارزیابی کیفیت مدل‌ها

برای ارزیابی کیفیت هر مدل مجموعه داده‌های مشاهده-ای به یک مجموعه آموزشی (۷۵ درصد داده‌ها) و هم‌چنین یک مجموعه آزمون (۲۵ درصد داده‌ها) تقسیم گردیدند و به‌منظور واسنجی و اعتبارسنجی مدل و مقایسه میزان دقت و کارایی پیش‌بینی‌های صورت گرفته در مورد مقادیر شکل‌های مختلف آهن خاک توسط هر یک از مدل‌های کیوبیست و درخت تصمیم‌گیری از معیارهای آماری میانگین ریشه مربعات خط^۱ (رابطه ۱)، میانگین خطا^۲ (رابطه ۲)، ضریب تعیین^۳ (رابطه ۳) و هم‌چنین ضریب توافقی یا سازگاری^۴ (رابطه ۴) ارائه شده توسط لین (Lin, 1989) استفاده شد.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (obs_i - pred_i)^2}{n}}$$

$$ME = \frac{\sum_{i=1}^n (obs_i - pred_i)}{n}$$

$$r = \frac{\sum_{i=1}^n (obs_i - \overline{obs})(pred_i - \overline{pred})}{\sqrt{\sum_{i=1}^n (obs_i - \overline{obs})^2} \sqrt{\sum_{i=1}^n (pred_i - \overline{pred})^2}}$$

$$\rho_c = \frac{2\rho\sigma_{pred}\sigma_{obs}}{\sigma_{pred}^2 + \sigma_{obs}^2 + (\mu_{pred} - \mu_{obs})^2}$$

در هر یک از این رابطه‌ها obs_i و $pred_i$ به ترتیب مقادیر مشاهده شده (واقعی) و پیش‌بینی شده، \overline{obs} و \overline{pred} به ترتیب مقادیر مطلق مشاهده شده و پیش‌بینی شده، μ_{obs} و μ_{pred} به ترتیب میانگین مقادیر مشاهده شده و پیش‌بینی شده، σ_{obs}^2 و σ_{pred}^2 به ترتیب واریانس‌های مقادیر مشاهده شده و پیش‌بینی شده، ρ_c ضریب همبستگی بین مقادیر مشاهده شده و پیش‌بینی شده متغیر هدف توسط مدل‌های مورد نظر و یا برآوردگر آماری و n نیز تعداد مشاهدات است.

نتایج و بحث

توصیف آماری داده‌ها

نتایج توصیف آماری سه شکل مختلف آهن شامل (Fe_t ، Fe_d و Fe_o) در عمق صفر تا ۱۰ سانتی‌متری در جدول ۱ آمده است. مقادیر میانگین برای Fe_t ، Fe_d و Fe_o به ترتیب برابر ۱۴/۷۲ گرم بر کیلوگرم، ۲/۸۷ گرم بر کیلوگرم و ۰/۳۸ گرم بر کیلوگرم و هم‌چنین مقادیر میانه برای هر

آهن خاک در عمق صفر تا ۱۰ سانتی‌متری شامل مدل-های درخت تصمیم‌گیری و کیوبیست به‌منظور مقایسه و بررسی میزان کارایی هر یک از این مدل‌ها استفاده شد.

مدل درخت تصمیم‌گیری

در این مدل چندین متغیر مستقل در ارتباط با درک روابط مستقیم و یا غیر مستقیم و میزان همبستگی با متغیر هدف و جزءبندی داده‌ها به صورت بازگشتی نسبت به تعدادی از گروه‌ها وارد یک ساختار درختی می‌شوند (Breiman, 1984). در این روش مجموعه‌ای از شرط‌های منطقی با ساختار درختی برای طبقه‌بندی و پیش‌بینی کمی یک متغیر به‌کار می‌روند.

مدل کیوبیست

- (۱) مدل کیوبیست یک الگوریتم پیشرفته از روش درخت رگرسیون است که مدل‌های مختلفی را از داده‌های آموزشی ایجاد می‌کند (Quinlan, 1992). این مدل به صورت یکسری قواعد که با جمله شرطی اگر شروع می‌شوند ارائه می‌شود. در این مدل مجموعه داده‌ها بر اساس معیارهای متغیرهای مورد استفاده در مدل دسته‌بندی می‌شوند. در این مدل اطلاعاتی را در مورد شناخت شرایط هر قاعده در مدل، قواعد مدل‌های رگرسیونی، تشخیص بهترین مدل برآزش داده و هم‌چنین میزان شرایط، تناوب و تکرار متغیرهای به‌کار گرفته شده در مدل فراهم می‌آورد.

در ادامه برای مشخص نمودن رابطه بین متغیرهای هدف (شکل‌های مختلف آهن خاک) و متغیرهای کمکی مورد نظر و انتخاب بهترین مدل برآزش داده شده در مدل درخت تصمیم‌گیری با استفاده از بسته (rpart) Therneau (et al., 2017) و در مدل کیوبیست با استفاده از بسته Cubis (Kuhn et al., 2016) در نرم‌افزار R و هم‌چنین با استفاده از نتایج واسنجی هر یک از این مدل‌ها مقدار شکل‌های مختلف آهن خاک در ۱۳۱ نقطه مشاهداتی با داده‌ها و اطلاعات حاصل از ۱۹ متغیرهای کمکی در منطقه مورد مطالعه نقشه شکل‌های مختلف آهن خاک شامل (Fe_t ، Fe_d و Fe_o) در محیط نرم‌افزارهای R و ArcGIS تهیه شد.

3. Coefficient of Determination (R^2)

4. Concordance Correlation Coefficient (CCC)

1. Root mean square error (RMSE)

2. Mean Error (ME)

چولگی به سمت راست بوده‌اند در حالی که مقدار Fe_o در مجموعه داده‌ها هیچ گونه چولگی نشان نداد. در ادامه در شکل ۲ توزیع پراکنش داده‌ها به صورت نمودار جعبه‌ای نشان داده شده است.

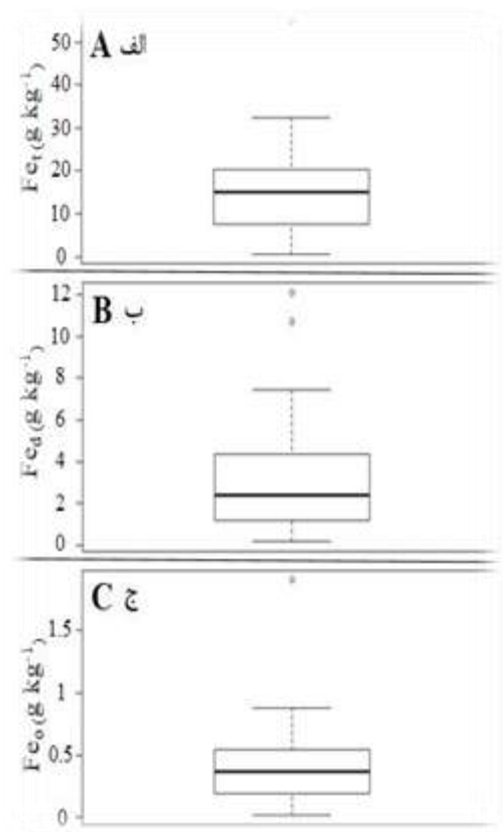
یک از این شکل‌های آهن به ترتیب برابر $۱۵/۰۸$ ، $۲/۴۱$ و $۰/۳۷$ گرم بر کیلوگرم به دست آمد. در بین شکل‌های آهن Fe_d دارای بیش‌ترین ضریب تغییرات ($۷۴/۵۶$ درصد) و Fe_t دارای کم‌ترین ضریب تغییرات ($۵۷/۲۷$ درصد) می‌باشد. همچنین نتایج نشان داد که مقادیر Fe_d و Fe_t دارای

جدول ۱- توصیف آماری مقادیر کربن آلی خاک برای نمونه‌های مورد استفاده در این مطالعه ($n=131$)
Table 1. Descriptive statistics of soil organic carbon for samples used in this study ($n=131$)

	Number	Min	Max	Mean	First Quintile	Median	Third Quintile	Standard deviation	Coefficient of Variation	Skewness	Kurtosis
Fe_t ($g\ kg^{-1}$)	131	0.62	54.94	14.72	7.30	15.08	20.54	8.43	57.27	0.75	2.54
Fe_d ($g\ kg^{-1}$)	131	0.19	12.09	2.87	1.17	2.41	4.45	2.14	74.56	1.29	2.52
Fe_o ($g\ kg^{-1}$)	131	0.02	1.90	0.38	0.19	0.37	0.54	0.24	63.16	1.92	10.27

ارزیابی کارایی و انتخاب بهترین مدل

جدول ۲ میزان کارایی هر یک از مدل‌های کیوبیست و درخت تصمیم‌گیری را با استفاده از معیارهای آماری مختلف مانند R^2 ، CCC، RMSE و ME در مجموعه داده‌های واسنجی و اعتبارسنجی را برای پیش‌بینی پراکنش مکانی شکل‌های مختلف آهن در منطقه مورد مطالعه بر اساس داده‌های آزمون (۲۵ درصد از داده‌ها) و داده‌های آموزش (۷۵ درصد از داده‌ها) نشان می‌دهد. مقدار ضریب R^2 در مجموعه داده‌های واسنجی در مدل درخت تصمیم‌گیری برای Fe_o ، Fe_d ، Fe_t از $۰/۷۳$ تا $۰/۷۸$ و در مدل کیوبیست از $۰/۸۵$ تا $۰/۸۹$ متغیر می‌باشد. به طوری که در واسنجی مجموعه داده‌ها میزان کارایی و عملکرد ضریب R^2 برای پیش‌بینی پراکنش Fe_d ، Fe_t و Fe_o به ترتیب در مدل درخت تصمیم‌گیری برابر با $(R^2=۰/۷۸)$ ، $(R^2=۰/۷۳)$ و $(R^2=۰/۷۶)$ و در مدل کیوبیست این مقادیر افزایش و به میزان بهتر و مناسب‌تری و برابر با $(R^2=۰/۸۹)$ ، $(R^2=۰/۸۵)$ و $(R^2=۰/۸۸)$ به دست آمد. علاوه بر این در مورد مقادیر CCC روند مشابهی مشاهده شد به طوری که بیش‌ترین و بالاترین مقدار CCC در مجموعه داده‌های اعتبارسنجی برای پیش‌بینی Fe_t و برابر با $۰/۳۵$ در مدل کیوبیست و پایین‌ترین و کم‌ترین مقدار CCC برای Fe_d و برابر با $۰/۱۱$ در مدل درخت تصمیم‌گیری ثبت شد.



شکل ۲- نمودار جعبه‌ای برای الف- مقادیر آهن معادل کل خاک ($n=131$)، ب- مقادیر آهن پدوژنیک خاک ($n=131$)، ج- مقادیر آهن بی‌شکل خاک در نمونه‌ها مورد استفاده در تحقیق حاضر ($n=131$)

Figure 2. Box plots of A: soil total equivalent iron ($n=131$); B: soil pedogenic iron ($n=131$); C: soil noncrystalline iron for samples used in the present study ($n=131$)

جدول ۲- کارایی مدل درخت تصمیم‌گیری و کیوبیست در مدل‌سازی و پیش‌بینی پراکنش مکانی شکل‌های مختلف آهن خاک در منطقه مورد مطالعه (n=۱۳۱)

Table 2. Performance of DT and Cubist models in modelling and predicting spatial distribution of iron forms in the study area study (n=131)

	Different forms of iron	R ²		CCC		RMSE (g kg ⁻¹)		ME	
		DT	Cu	DT	Cu	DT	Cu	DT	Cu
Calibration dataset (in the bag)	Fe _t (g kg ⁻¹)	0.78	0.89	0.85	0.91	2.31	2.25	0.09	0.07
	Fe _d (g kg ⁻¹)	0.73	0.85	0.82	0.87	0.61	0.57	0.08	0.05
	Fe _o (g kg ⁻¹)	0.76	0.88	0.87	0.89	0.12	0.09	0.06	0.02
Validation dataset (out of bag)	Fe _t (g kg ⁻¹)	0.31	0.35	0.46	0.57	3.22	3.08	0.19	0.16
	Fe _d (g kg ⁻¹)	0.11	0.13	0.35	0.38	1.35	1.28	0.17	0.12
	Fe _o (g kg ⁻¹)	0.28	0.31	0.49	0.53	0.23	0.19	0.13	0.06

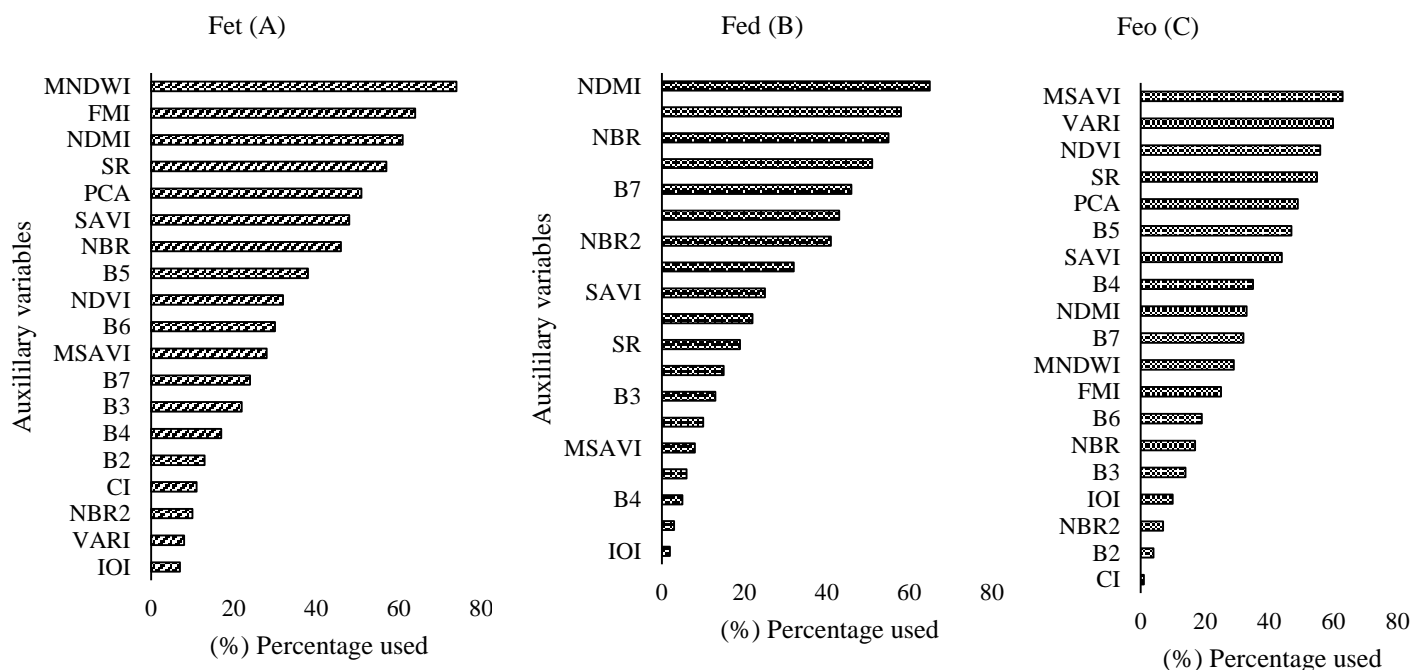
هم‌چنین زمین‌نما شامل VARI, SAVI, SR, NDVI و NBR از اهمیت بالاتر و شاخص‌های IOI و CI و برخی باندهای مستقل از جمله B2, B3, B4 و B5 از اهمیت کم‌تری نسبت به سایر شاخص‌ها ترکیبی (مرکب) برخوردار بودند. بنابراین می‌توان نتیجه گرفت که اهمیت شاخص‌های طیفی مستقل و مرکب برای پیش‌بینی پراکنش مکانی شکل‌ها و شاخص‌های مختلف آهن متفاوت است. جدول ۳ میزان اهمیت هر یک از متغیرهای کمکی را برحسب درصد در مدل کیوبیست به‌عنوان بهترین مدل برای برای پیش‌بینی شکل‌های مختلف آهن شامل Fe_t, Fe_d و Fe_o نشان می‌دهد. نتایج این جدول نشان داد که شاخص‌های NDMI, MNDWI, SR, FMI و SAVI با سهمی برابر ۶۰ درصد به‌عنوان پراهمیت‌ترین و پرتکرارترین متغیرهای مورد استفاده در مدل برای پیش‌بینی Fe_t هستند. هم‌چنین شاخص CI نیز با سهمی برابر ۱۰ درصد کم‌ترین سهم را در متغیرهای مورد استفاده در مدل دارا است. سایر متغیرهای کمکی نیز سهمی بین ۱۰ تا ۶۰ درصد را نشان دادند. هم‌چنین بر اساس نتایج به‌دست آمده از این جدول شاخص‌های NDMI و SR با سهمی برابر ۶۰ و ۱۴ درصد به‌ترتیب به‌عنوان پراهمیت و کم‌اهمیت‌ترین متغیر مورد استفاده در مدل برای پیش‌بینی Fe_d مشخص گردیدند. علاوه بر این بر اساس نتایج میزان سهم متغیرهای کمکی پیش‌بینی کننده Fe_o در مدل کیوبیست این متغیرهای کمکی سهمی به میزان ۱۲ تا ۵۴ درصد از خود نشان دادند.

از طرف دیگر دارا بودن مقادیر بالای CCC در مجموعه داده‌های اعتبارسنجی مدل کیوبیست و هم‌چنین وجود مقادیر بسیار اندک و ناچیز میانگین خطا و نزدیکی این معیار به صفر در این مدل نسبت به مدل درخت تصمیم‌گیری نشان داد که مدل کیوبیست با کارایی و دقت بالاتری نسبت به مدل درخت تصمیم‌گیری با بهره‌گیری از داده‌ها و متغیرهای کمکی شامل شاخص‌های طیفی سنسجش از دور به‌عنوان بهترین و مناسب‌ترین مدل برازش داده شده برای پیش‌بینی توزیع مکانی و تهیه نقشه رقومی شکل‌های مختلف آهن در سراسر منطقه مورد مطالعه است.

تعیین اهمیت و درصد مشارکت متغیرهای کمکی

در این مرحله با استفاده از نتایج به‌دست آمده توسط هر یک از مدل‌های درخت تصمیم‌گیری و کیوبیست به‌عنوان پیش‌بینی‌کننده‌های کمی برای مدل‌سازی پراکنش مکانی و روابط بین شکل‌های مختلف آهن و متغیرهای کمکی مورد استفاده، میزان اهمیت و درصد مشارکت متغیرهای بکار گرفته شده در مدل درخت تصمیم‌گیری برای پیش‌بینی شکل‌های مختلف آهن در شکل ۳ و هم‌چنین نتایج مدل کیوبیست به‌منظور پیش‌بینی شکل‌های مختلف آهن در جدول ۳ ارائه شده است.

بر اساس شکل ۳ شاخص‌های NDMI, MNDWI و FMI به‌عنوان سه شاخص مهم (دارای بیش‌ترین اهمیت) برای پیش‌بینی Fe_t و Fe_d تعیین شدند. هم‌چنین شاخص MSAVI به‌عنوان پیش‌بینی‌کننده قابل توجه و مهمی از مقدار Fe_o در منطقه مورد مطالعه بود. هم‌چنین طبق شکل ۳ در انتخاب متغیرهای کمکی مؤثر در مدل درخت تصمیم‌گیری و پیش‌بینی شکل‌های مختلف آهن خاک به‌ترتیب شاخص‌های مربوط به پوشش گیاهی و خاک و



شکل ۳- رتبه‌بندی اهمیت متغیرهای محیطی مورد استفاده در این مطالعه برای الف- پیش‌بینی آهن معادل کل خاک با استفاده از مدل درخت تصمیم‌گیری، ب- پیش‌بینی آهن پدوژنیک خاک با استفاده از مدل درخت تصمیم‌گیری، ج- پیش‌بینی آهن بی‌شکل خاک با استفاده از مدل درخت تصمیم‌گیری

Figure 3. Covariate importance rankings used in this study for A: predicting soil total equivalent iron using DT model; B: predicting soil pedogenic iron using DT model; C: predicting soil noncrystalline iron using DT model

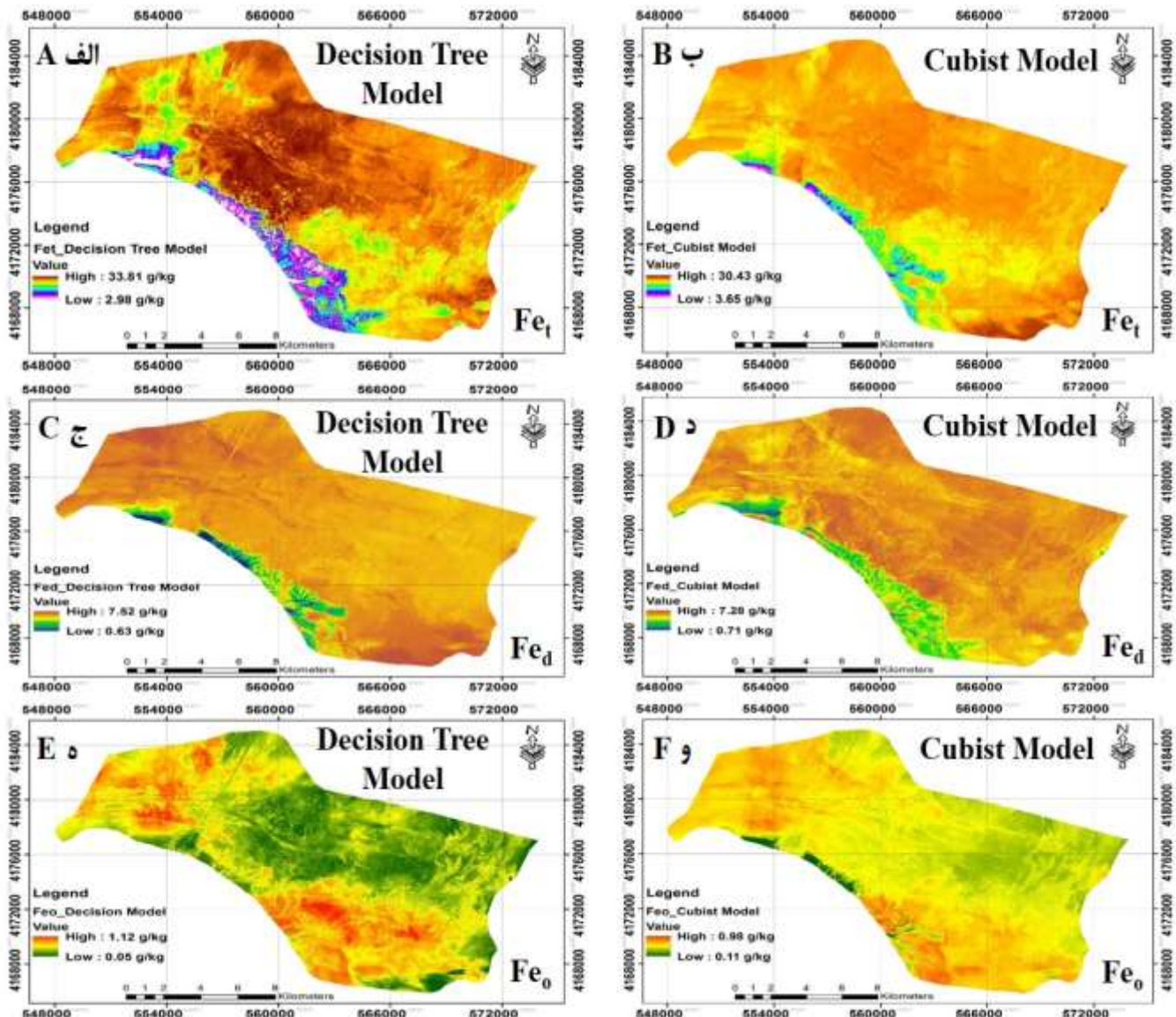
جدول ۳- سهم متغیرهای کمکی پیش‌بینی کننده در مدل کیوبیست

Table 3. Contributions of the environmental covariates predictors in the Cubist model

Covariates	Fe _t (g kg ⁻¹)	Fe _d (g kg ⁻¹)	Fe _o (g kg ⁻¹)
	Contribution in the model	Contribution in the model	Contribution in the model
B2	20%	20%	
B3	30%	20%	20%
B4	20%		20%
B5	52%		40%
B6	30%	20%	40%
B7	30%	20%	30%
PCA	0%		
NDVI	40%	40%	40%
VARI			40%
SAVI	60%	40%	30%
MSAVI	30%		54%
NDMI	60%	60%	20%
MNDWI	60%	40%	20%
NBR	40%	30%	20%
NBR2			
SR	60%	14%	40%
CI	10%	40%	12%
FMI	60%	40%	30%
IOI			

ArcGIS با قدرت تفکیک ۳۰ متر تهیه و ارائه شد (شکل ۴). در ادامه میانگین مقادیر پیش‌بینی شده شکل‌های مختلف آهن خاک با استفاده از میانگین مقادیر عددی پیکسل‌های نقشه‌های رقمی تهیه شده توسط هر یک از دو مدل در جدول ۴ ارائه گردید.

نقشه‌برداری رقمی و پهنه‌بندی شکل‌های مختلف آهن خاک در منطقه مورد مطالعه
در این مرحله نقشه‌های پراکنش و توزیع مکانی شکل‌های مختلف آهن شامل (Fe_o و Fe_d , Fe_t) در سراسر منطقه مورد مطالعه (سایت‌های یک و دو) با انجام مدل‌سازی و بکارگیری هر یک از مدل‌های درخت تصمیم‌گیری و کیوبیست در محیط نرم‌افزارهای R و



شکل ۴- پیش‌بینی پراکنش مکانی در این مطالعه برای الف- آهن معادل کل خاک با استفاده از مدل درخت تصمیم‌گیری، ب- آهن معادل کل خاک با استفاده از مدل کیوبیست،

ج- آهن پدوژنیک خاک با استفاده از مدل درخت تصمیم‌گیری، د- آهن پدوژنیک خاک با استفاده از مدل کیوبیست،

ه- آهن بی‌شکل خاک با استفاده از مدل درخت تصمیم‌گیری، و- آهن بی‌شکل خاک با استفاده از مدل کیوبیست

Figure 4. Predicted spatial distribution in this study of A: soil total equivalent iron using DT model; B: soil total equivalent iron Cu model;

C: soil pedogenic iron using DT model; D: soil pedogenic iron using Cu model;

E: soil noncystalline iron using DT model; F: soil noncystalline iron using Cu model

جدول ۴- میانگین مقادیر پیش‌بینی شده همراه با انحراف معیار شکل‌های مختلف آهن خاک با استفاده از مدل درخت تصمیم‌گیری و کیوبیست در سال ۲۰۱۷

Table 4. The mean predicted values associated with standard deviations of iron forms using Cu and DT models in 2017

	Fe _t (g kg ⁻¹)	Fe _d (g kg ⁻¹)	Fe _o (g kg ⁻¹)
DT model	15.19	3.08	0.59
*SD	3.70	2.28	0.32
Cu model	14.75	2.89	0.41
*SD	3.55	2.12	0.28

*: انحراف معیار برای هر میانگین مقادیر پیش‌بینی شده شکل‌های مختلف آهن خاک

*: Standard deviation (SD) for each predicted mean value of iron forms soil

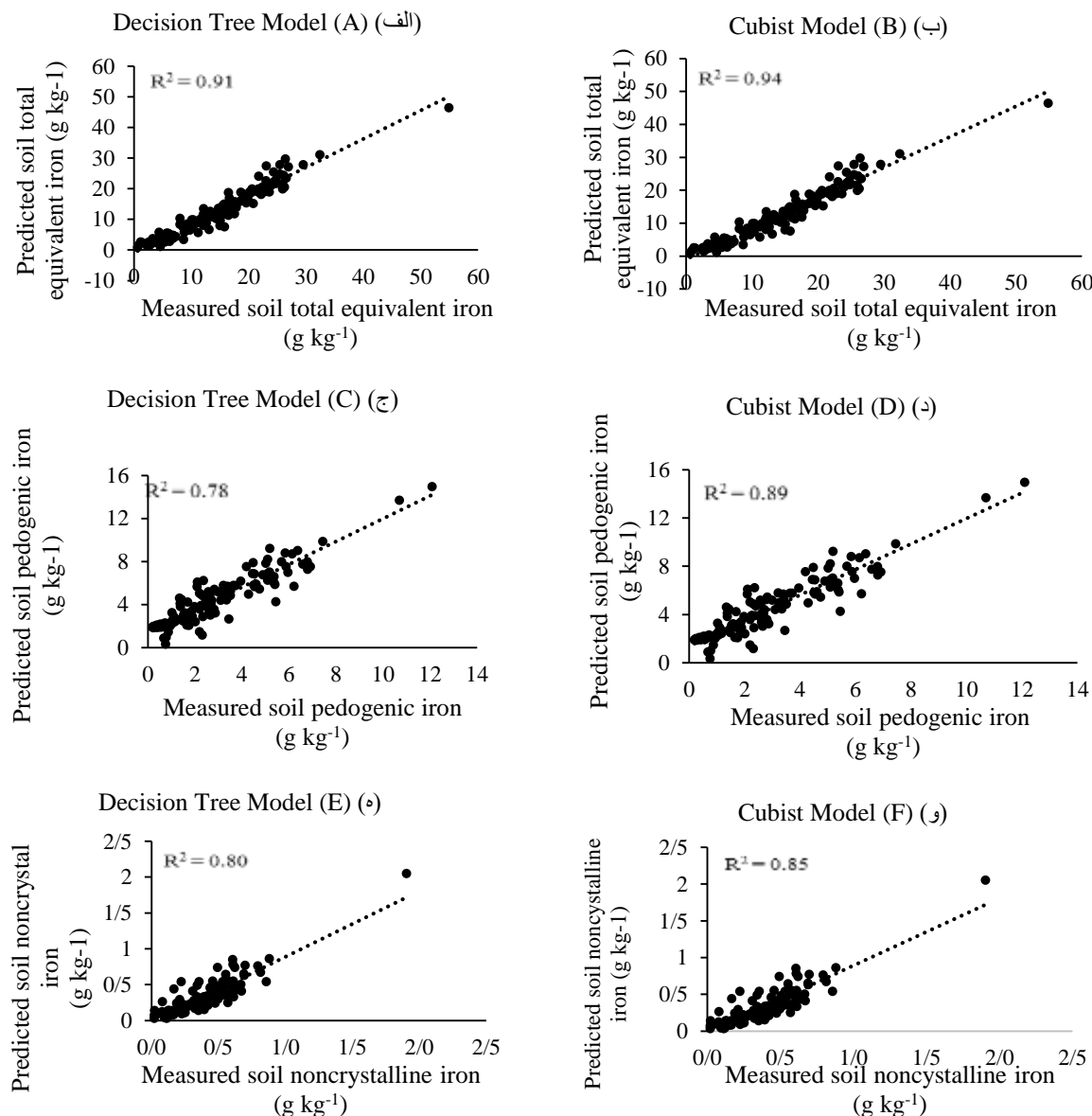
دارند. بررسی‌های بیش‌تر نقشه‌های تهیه شده حاکی از آن است که مقدار متوسط هر سه شکل آهن با استفاده از مدل کیوبیست با مقدار متوسط واقعی نمونه‌ها (جدول ۱) همخوانی بیش‌تری داشته است و هم‌چنین با استناد به معیارهای آماری مدل‌سازی مختلف و به‌دلیل دارابودن ضریب تبیین بالاتر و پایین بودن مقدار RMSE و مقدار ME نزدیک به صفر مدل کیوبیست به مراتب از قابلیت بهتر و دقت بالاتری در پیش‌بینی و برآورد شکل‌های مختلف آهن معرفی می‌گردد. علاوه بر این بر اساس نتایج به‌دست آمده از جدول ۴ و هم‌چنین با استناد به نقشه‌های رقمی تهیه شده غلظت Fe_t در سال ۲۰۱۷ با استفاده از مدل درخت تصمیم‌گیری و کیوبیست به‌ترتیب و به‌طور میانگین برابر با ۱۵/۱۹ و ۱۴/۷۵ گرم بر کیلوگرم بوده است که این امر نشان دهنده افزایش تدریجی آهن معادل کل در طول زمان و خشک شدن دریاچه ارومیه می‌باشد. هم‌چنین طبق نتایج جدول ۴ غلظت Fe_d در سال ۲۰۱۷ با استفاده از مدل درخت تصمیم‌گیری و کیوبیست به‌ترتیب و به‌طور میانگین برابر با ۳/۰۸ و ۲/۸۹ گرم بر کیلوگرم بود که بر این اساس همان‌طوری که انتظار می‌رفت تغییرات در مورد Fe_d در سراسر محدوده مورد مطالعه یکنواخت و یکسان نبوده است و این بدان معنی است که منطقه مورد مطالعه بسیار پویا بوده و هم‌چنین در طول زمان و در اثر خشک شدن دریاچه فرآیند تشکیل آغازین خاک در نتیجه خروج بیش‌تر اکسیدهای آهن از کانی‌ها در نتیجه هوازدگی در محدوده مورد مطالعه اتفاق افتاده است که این امر تکامل بالای خاک‌ها را نشان می‌دهد. علاوه بر این نتایج جدول ۴ نشان‌دهنده غلظت

بر اساس نتایج جدول ۴ و هم‌چنین با استناد به نقشه‌های تهیه شده Fe_t، Fe_d و Fe_o در شکل ۴ تغییرات شکل‌های مختلف آهن خاک در سراسر منطقه مورد مطالعه یکسان و یکنواخت نبوده و این بدان معنا است که منطقه مورد مطالعه بسیار پویا بوده و از بعد مکانی دچار تغییر و تحولات ناشی از خشک و مرطوب شدن متوالی خاک می‌باشد که یکی از دلایل این امر و تغییرات رخ داده شده در سایت‌های مورد مطالعه مربوط به فرآیند فرآیندهای خاکساختی و تشکیل خاک در طول زمان مانند فرآیند رسیدن خاک^۱ و یا قرار گرفتن در مسیر و جهت جریان رودخانه آبی‌چای در حوضه دریاچه ارومیه می‌باشد که سبب هوازدگی و هم‌چنین خروج و تحرک بیش‌تر اکسیدهای آهن از کانی‌ها در خاک‌های مورد مطالعه می‌گردد که پوشش‌دار شدن اکسیدهای آهن توسط کربنات‌ها می‌تواند این رویداد را توجیه کند (Tack & Verloo, 1997). از این لحاظ هائسه و همکاران (Haese et al., 1997) به اهمیت نقشه‌های پراکنش مکانی برای شناسایی و تشخیص فرآیندهای مختلف رخ داده شده در خاک از جمله فرآیند رسیدن اشاره نموده‌اند. شهبازی و همکاران (Shahbazi et al., 2019b) نیز فرآیند رسیدن خاک در ساحل شرقی دریاچه ارومیه را با استفاده از نقشه‌های رقمی آهن بلورین به‌خوبی نشان داده‌اند. آن‌ها بیان کردند که حضور بیش‌تر آهن بلورین در خاک‌ها و مناطق مطالعه شده بیانگر پیشرفت و گسترش فرآیند رسیدن در خاک است و خاک‌های مناطقی که مدت زمان طولانی در معرض هوای آزاد هستند بالاترین میزان آهن بلورین را به دلیل داشتن زمان کافی به‌منظور تشکیل آهن متبلور

1. Soil ripening process

اساس داده‌های آزمون و هم‌چنین نمودار پراکنش مقادیر Fe_o ، Fe_d ، Fe_t و Fe_e اندازه‌گیری شده (واقعی) در مقابل مقادیر پیش‌بینی شده نشان داد که مدل کیوبیست به ترتیب با داشتن ضریب تعیین $0/94$ ، $0/89$ و $0/85$ به‌عنوان بهترین مدل از قابلیت و دقت بالاتری در پیش‌بینی و تخمین تغییرات مقدار شکل‌های مختلف آهن خاک نسبت به مدل درخت تصمیم‌گیری در منطقه مورد مطالعه برخوردار است (شکل ۵).

Fe_o به‌ترتیب برابر با $0/59$ و $0/41$ گرم بر کیلوگرم در سال ۲۰۱۷ با استفاده از مدل درخت تصمیم‌گیری و کیوبیست می‌باشد. بر این اساس مقدار Fe_o تغییرات ناچیز و اندکی را نشان داد که غیر قابل تشخیص بودن اکسید-های آهن ضعیف بلور می‌تواند یکی از دلایل تغییرات ناچیز و اندک Fe_o باشد. در ادامه بررسی نقشه‌های پیش-بینی شکل‌های مختلف آهن خاک و نتایج حاصل از مقایسه مدل‌های درخت تصمیم‌گیری و کیوبیست بر



شکل ۵- نمودار پراکنش مقادیر اندازه‌گیری شده (واقعی) و مقادیر پیش‌بینی شده در این مطالعه برای الف- آهن معادل کل خاک با استفاده از مدل درخت تصمیم‌گیری، ب- آهن معادل کل خاک با استفاده از مدل کیوبیست،

ج- آهن پدوژنیک خاک با استفاده از مدل درخت تصمیم‌گیری، د- آهن پدوژنیک خاک با استفاده از مدل کیوبیست،

ه- آهن بی‌شکل خاک با استفاده از مدل درخت تصمیم‌گیری، و- آهن بی‌شکل خاک با استفاده از مدل کیوبیست

Figure 5. Scatter gram of measured and predicted values in this study of A: soil total equivalent iron using DT model; B: soil total equivalent iron using Cu model; C: soil pedogenic iron using DT model; D: soil pedogenic iron using Cu model; E: soil noncrystalline iron using DT model; F: soil noncrystalline iron using Cu model

نتیجه‌گیری کلی

نتایج این مطالعه نشان داد که اگرچه متغیرها و داده‌های کمکی اهمیت متفاوتی در پیش‌بینی پراکنش شکل‌های مختلف آهن خاک دارا بودند اما به‌طور کلی می‌توان با یافتن ارتباط بین آن‌ها و شکل‌های مختلف آهن خاک از طریق مدل‌هایی مانند مدل درخت تصمیم‌گیری و مدل کیوبیست نقشه رقومی پیوسته شکل‌های مختلف آهن خاک را با دقتی قابل قبول تهیه نمود. هم‌چنین این نتایج نشان داد که مدل درخت تصمیم‌گیری برای پیش‌بینی تمامی متغیرهای هدف شامل شکل‌های مختلف آهن مناسب نیست. درحالی که مدل کیوبیست از کارایی بالاتری نسبت به مدل درخت تصمیم‌گیری برای پیش‌بینی توزیع مکانی شکل‌های مختلف آهن برخوردار است.

بنابراین می‌توان بیان نمود که الگوریتم و مدل کیوبیست به عنوان راه‌حلی ساده، کم‌هزینه و با صرف زمانی اندک و با دقت پیش‌بینی بالا به‌خوبی توانسته تغییرات و پراکنش شکل‌های مختلف آهن خاک را در ساحل شرقی دریاچه ارومیه نشان دهد. به علاوه نقشه‌های رقومی و شبیه‌سازی شده پراکنش مکانی شکل‌های مختلف آهن خاک توسط مدل مذکور می‌تواند میزان توانایی و قابلیت فناوری سنجش از دور را برای تشخیص شکل‌های مختلف آهن نشان دهد و هم‌چنین سرعت عمل و کارآمدی نقشه‌های تهیه شده را در انتقال داده‌ها و اطلاعات آشکار سازد و قابلیت استفاده از آن‌ها را برای دسته‌ای وسیع از مطالعات و شاخه‌های علمی مرتبط با خاک و محیط زیست امکان‌پذیر نماید.

References

- Boettinger J.L., Ramsey R.D., Bodily J.M., Cole N.J., Kienast-Brown S., Nield S.J., Saunders A.M., and Stum A.K. 2008. Landsat spectral data for digital soil mapping, *In: Hartemink A.E., McBratney A.B., and Mendonca-Santos M. (Eds.), Digital Soil Mapping with Limited Data. Springer, Dordrecht, pp. 193-203.*
- Breiman L., Friedman J.H., Olshen R.A., and Stone C.J. 1984. Classification and Regression Trees. Chapman & Hall, New York.
- Drury S.A. 2016. Image Interpretation in Geology (2nd Ed.). Nelson Thornes, London.
- ESRI. 2011. ArcGIS Desktop: Release 10.2. Environmental Systems Research Institute. Redlands, CA.
- Fatehi S.H., Mohammadi J., Salehi M.H., Momeni M., Toomanian N., and Jafari N. 2016. Downscaling digital soil organic carbon map. *Journal of Water and Soil*, 30(1-2): 1142-1157. (In Persian)
- Gilabert M.A., Gonzalez-Piqueras J., Garcia-Haro F.J., and Melia J. 2002. A generalized soil adjusted vegetation index. *Remote Sensing of Environment*, 82(2): 303-310.
- Gitelson A.A., Kaufman Y.J., Stark R., and Rundquist D. 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, 80: 76-87.
- Haese R.R., Wallmann K., Dahmke A., Kretzmann U., Muller P.J., and Schulz H.D. 1997. Iron species determination to investigate early diagenetic reactivity in marine sediments. *Geochim Cosmochim Acta*, 61: 63-72.
- IRIMO. 2012. Islamic Republic of Iran Meteorological Organization.
- Kuhn M., Weston S., Keefer C., and Coulter N. 2016. C code for Cubist. Cubist: Rule- and Instance-based Regression Modeling. R Package Version 0.0.19. <https://CRAN.Rproject.org/package=Cubist>.
- Lin L.I. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45: 255-268.
- Loeppert R.L., and Inskeep W.P. 1996. Iron. *In: Sparks D.L. (Ed.), Methods of Soil Analysis-Part 3. (3rd Ed.). Agron. Monogr. vol. 9. ASA, CSSA, and SSSA, Madison, WI. pp. 639-664.*
- McGrath S.P., and Cunliffe C.H. 1985. A simplified method for the extraction of the metals Fe, Zn, Cu, Ni, Cd, Pb, Cr, Co and Mn from soils and sewage sludges. *Journal of the Science of Food and Agriculture*, 36: 794-798.
- McKenzie N.J., Webster R., and Ryan P.J. 2008. Sampling using statistical methods. *In: McKenzie N.J., Grundy M.J., Webster R., and Ringrose-Voase A.J. (Eds.), Guidelines for Surveying Soil and Land Resources, 2nd. CSIRO, Australia, pp. 319-326.*
- Metternicht G.I., and Zinck J.A. 2003. Remote sensing of soil salinity: potentials and constraints. *Remote Sensing of Environment*, 85: 1-20.

- Mosleh Z., Salehi M.H., Jafari A., Borujeni I.E., and Mehnatkesh A. 2016. The effectiveness of digital soil mapping to predict soil properties over low-relief areas. *Environmental Monitoring and Assessment*, 188(3): 195.
- Parks S.A., Dillon G.K., and Miller C. 2014. A new metric for quantifying burn severity: the relativized burn ratio. *Remote Sensing*, 6: 1827-1844.
- Qi J., Chehbouni A.R., Kerr Y.H., and Sorooshian S. 1994. A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48: 119-126.
- Quinlan J.R. 1992. Learning with continuous classes. In: *Proceedings of Australian, 5th Joint Conference on Artificial Intelligence*. World Scientific, Singapore, pp. 343-348.
- Shahbazi F., Hughes P., McBratney A., Minasny B., and Malone B.P. 2019a. Evaluating the spatial and vertical distribution of agriculturally important nutrients-nitrogen, phosphorous and boron-in North West Iran. *Catena*, 173: 71-82.
- Shahbazi F., McBratney A., Malone B.P., Oustan S., and Minasny B. 2019b. Retrospective monitoring of the spatial variability of crystalline iron in soils of the east shore of Urmia Lake, Iran using remotely sensed data and digital maps. *Geoderma*, 337: 1196-1207.
- Skakun R.S., Wulder M.A., and Franklin S.E. 2003. Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sensing of Environment*, 86: 433-443.
- Somarathna P.D.S.N., Malone B.P., and Minasny B., 2016. Mapping soil organic carbon content over New South Wales, Australia using local regression kriging. *Geoderma Regional*, 7: 38-48.
- Stonehouse H.B., and Arnaud R.J.St. 1971. Distribution of iron, clay and extractable iron and aluminum in some Skatchewan soils. *Canadian Journal of Soil Science*, 51(2): 283-292.
- Tack F.M.G., and Verloo M.G. 1997. Single extractions versus sequential extraction for the estimation of heavy metal fractions in reduced and oxidised dredged sediments. *Chemical Speciation & Bioavailab*, 11: 43-50.
- Taghizadeh-Mehrjardi R., Minasny B., Sarmadian F., and Malone B.P. 2014. Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma*, 213: 15-28.
- Therneau T., Atkinson B., and Ripley B. 2017. rpart: Recursive Partitioning and Regression Trees. R Package Version 4.1-11. <https://CRAN.R-project.org/package=rpart>. ULRP, 2015. Report of Lake Urmia Conditions. Urmia Lake Restoration Program. University of Sharif, Iran.
- Were K., Bui D.T., Dick Q.B., and Singh B.R. 2015. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecological Indicators*, 52: 394-403.
- Xu H. 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27: 3025-3033.
- Zhu A., Hudson B., Burt J., Lubich K., and Simonson D. 2001. Soil mapping using GIS, expert knowledge and fuzzy logic. *Soil Science Society of America Journal*, 65: 1463-1472.

Digital Mapping of Different forms of Soil Iron in the Eastern Shore of Urmia Lake by using Landsat-8 OLI Imagery

Amin Mousavi^{1*}, Farzin Shahbazi², Shahin Oustan³, Ali Asghar Jafarzadeh⁴, Budiman Minasny⁵

(Received: February, 2020 Accepted: August, 2020)

Abstract

In this study, digital mapping of the most important forms of soil Iron were done using two data mining techniques namely Decision Tree (DT) and Cubist (Cu) models. The study area includes 500 km² of lands from two different sites located in the eastern shore of dried bed of Urmia Lake, northwest of Iran. 131 surface soil samples were taken from depth of 0-10 cm and three different forms of Iron including i): total iron (Fe_t); ii) pedogenic iron (Fe_d); and iii) amorphous iron (Fe_o) were measured. A total of 19 environmental covariates (auxiliary variables) derived from the Landsat-8 OLI imagery related to July 2017 were used in this study. It was found that Cu model has a higher precision than that of the DT model for predicting all three forms of soil iron with the values R²=0.89 and RMSE= 2.25 g/kg , R²=0.85 and RMSE=0.57 g/kg and R²=0.88 and RMSE=0.09 g/kg for predicting Fe_t, Fe_d and Fe_o, respectively. In addition, the results of the importance and percentage of contribution of environmental covariates in both models indicated the high importance of some spectral indices such as Normalized Difference Moisture Index (NDMI) and Modified Soil Adjusted Vegetation Index (MSAVI) in the prediction of Fet, Fed and Feo. Generally, the Cu model has a higher ability and performance in modeling and predicting the spatial distribution of different forms of soil iron in the study area compared to the DT model.

Keywords: Cubist, Data mining, Decision tree, Different forms of iron, Environmental covariates

Mousavi A., Shahbazi F., Oustan S., Jafarzadeh A.A., and Minasny B. 2021. Digital Mapping of different forms of soil iron in the eastern shore of Urmia Lake by using Landsat-8 OLI imagery. *Applied Soil Research*, 9(2): 47-60.

1. Ph.D. Student, Department of Science and Soil Engineering, Faculty of Agriculture, University of Tabriz

2. Associate Professor, Department of Science and Soil Engineering, Faculty of Agriculture, University of Tabriz

3. Professor, Department of Science and Soil Engineering, Faculty of Agriculture, University of Tabriz

4. Professor, Department of Science and Soil Engineering, Faculty of Agriculture, University of Tabriz

5. Professor, Sydney Institute of Agriculture and School of Life and Environmental Sciences, Faculty of Agriculture, The University of Sydney, Australia

* Corresponding Author Email: amin_mousavi@alumni.ut.ac.ir