

مدل‌سازی توزیع سرب با استفاده از ماشین‌های بردار پشتیبان¹ در خاک سطحی زمین‌های حاشیه بزرگراه دزفول- هواز

ثمین نظریات²، مه‌ران هودجی و علی اصغر بسالت پور

دانشجوی کارشناسی ارشد، دانشگاه آزاد اسلامی واحد خوراسگان، اصفهان؛ Snazariyat@yahoo.com

دانشیار دانشگاه آزاد اسلامی واحد خوراسگان، اصفهان؛ mehnan.hoodajil@gmail.com

استادیار دانشگاه ولی عصر (عج) رفسنجان؛ a.besalatpour@vru.ac.ir

دریافت: 94/11/7 و پذیرش: 95/9/17

چکیده

بررسی توزیع غلظت فلزات سنگین، برای پایش آلودگی خاک و حفظ کیفیت آن امری ضروری است. در این پژوهش، تغییرات غلظت سرب در خاک سطحی زمین‌های حاشیه بزرگراه دزفول-هواز بررسی و ویژگی‌های خاک مؤثر بر آن تعیین شد. بدین منظور، نمونه‌برداری خاک از عمق 0 تا 10 سانتی‌متر بر روی خط مجاور با محور بزرگراه با سه نمونه یکسان به فواصل 15، 40 و 100 متری از حاشیه جاده انجام شد (مجموعاً 135 نمونه). سپس در هر نمونه خاک، غلظت کل سرب، ماده آلی، کربنات کلسیم معادل، اسیدیته، قابلیت هدایت الکتریکی (EC) و توزیع اندازه‌های ذرات خاک (درصد رس، سیلت، شن ریز و خیلی ریز و درصد کل شن) اندازه‌گیری گردید. مدل‌سازی با استفاده از روش ماشین‌های بردار پشتیبان (SVMs) و رگرسیون خطی چندمتغیره (MLR) انجام شد. برای بررسی کارایی مدل‌ها نیز از برخی شاخص‌های آماری نظیر ضریب همبستگی (r)، ریشه میانگین مربعات خطا (RMSE) و شاخص کارایی مدل (MEF) بین مقادیر اندازه‌گیری شده و برآورد شده، استفاده گردید. بررسی نتایج حاصل از انتخاب متغیر به روش SVMs نمایان‌گر آن بود که در فاصله 15 متری از جاده، پارامتر درصد کربنات کلسیم معادل، در فاصله 45 متری، درصد رس و در فاصله 45 تا 100 متری، pH و سپس میزان درصد شن بیشترین ضریب اهمیت را در جذب سطحی سرب داشتند. ضریب r و شاخص MEF برای برآورد سرب با استفاده از مدل SVM به ترتیب برابر 0/86 و 0/70 درصد بود، در حالی که مقدار آن‌ها برای مدل MLR برابر 0/62 و 0/38 درصد بود. بنابراین؛ بر اساس نتایج این پژوهش به نظر می‌رسد بتوان از ماشین‌های بردار پشتیبان برای تعیین عوامل مؤثر بر دریافت سرب در خاک حاشیه جاده‌ها و مدل‌سازی تغییرات غلظت آن در خاک استفاده نمود.

واژه‌های کلیدی: آلاینده‌های خودرویی، آلودگی خاک، انتخاب متغیر، فلزات سمی

¹ Support Vector Machines

² نویسنده مسئول. آدرس: دانشگاه آزاد اسلامی واحد خوراسگان، اصفهان - دانشجوی کارشناسی ارشد. گروه مهندسی علوم خاک

مقدمه

سرب از جمله آلاینده‌های خودرویی است که به دلیل سمیت زیاد برای طبیعت و انسان و ایجاد عوارضی متعدد همچون بیماری‌های خونی، عصبی و استخوانی، مورد توجه پژوهشگران بسیاری قرار گرفته است (سارکار، 2002). سرب عموماً در نتیجه استفاده از سوخت‌های بنزینی سرب‌دار وارد مناطق شهری می‌شود. پارامترهای گوناگونی بر انتشار آلاینده‌های خودرویی در محیط اثر دارند که در مقیاس بزرگ، دربرگیرنده پارامترهای مربوط به شرایط جاده‌ای، ترافیکی و به‌ویژه محیطی هستند. پارامترهای جاده‌ای و خیابانی مربوط به شرایط فیزیکی و معماری راه‌ها همچون طول، شیب، عرض تنگه و عمر جاده یا خیابان، ارتفاع متوسط، چپش و عمر ساختمان‌های اطراف، نوع پوشش سبز و ارتفاع گیاهان اطراف هر جاده یا خیابان و نسبت دیدگاهی هستند (وانگ، 2005). جریان ترافیکی، هوای اطراف راه‌ها را که حاوی ذرات آلاینده‌های فلزات سنگین است، به جریان می‌اندازد. از آن‌جا که این جریان عمده‌تاً عمود بر محور جاده است، در برخورد با هر یک از عوامل فیزیکی یاد شده، می‌تواند حالت‌های گوناگونی به خود بگیرد و نهایتاً روند توزیع آلاینده‌ها در حاشیه راه‌ها را متأثر کند. جریان ترافیکی که از مهمترین عوامل در انتشار آلاینده‌های خودرویی است، شامل عوامل اصلی سرعت، حجم ترافیک و چگالی جریان و نیز عواملی همچون رفتار رانندگی در زمان شروع و حرکت، ترکیب ناوگان عبوری و نحوه مدیریت ترافیکی و چپش علائم راهنمایی و رانندگی در خیابان‌ها هستند (کلهو و همکاران، 2005). علاوه بر پارامترهای خیابانی و جریان ترافیکی، بسیاری از ویژگی‌های خاک حاشیه جاده‌ها نیز بر دریافت آلاینده‌های خودرویی و نگهداشت آن‌ها در خاک مؤثرند. غلظت فلزات سنگین از جمله سرب، در خاک‌های کشاورزی اطراف شهرها به دلیل دوری نسبی از مناطق آلاینده در حد پایین تری قرار دارد. میزان این آلاینده‌ها در مناطق آلاینده مانند پمپ بنزین‌ها و باطری سازی‌ها، ناهنجاری بسیار بالایی را به وجود می‌آورد. این مناطق می‌تواند به وسیله باد یا آب، فلزات سنگین را به نقاط دیگر انتقال دهد. انتقال فلزات سنگین در خاک مناطق شهری به طور کلی بستگی به میزان رواناب سطحی داشته که منجر به حرکت این عناصر از محیط‌های آلاینده نقطه‌ای به صورت محلول بونی و یا بار معلق و جامد (که در آن فلزات سنگین جذب کلئیدهای خاک شده است)، می‌شود. همچنین مهاجرت این عناصر و جذب آن‌ها توسط گیاهان به ویژگی‌های فیزیکی، شیمیایی و زیستی

خاک مانند اسیدیت، شوری، اجزای هیومیک خاک و کانی‌های آن، ساختار افق‌های خاک، ترکیب و دانه‌بندی خاک، ماده آلی خاک، توانایی جذب، فراوانی ریزجانداران و از همه مهم‌تر آب مربوط است و در نهایت انتقال این عناصر سمی به وسیله انحلال‌پذیری اشکال مختلف شیمیایی آن‌ها و توانایی جذب این عناصر توسط خاک، کنترل می‌شود (حمزه 1385، سلیمی و همکاران 1390، ازکوتلو و همکاران 2007، چانگ و همکاران 2010). چانگ و همکاران (2010) گزارش نمودند که شوری خاک از طریق افزایش فشار اسمزی محلول خاک، مهاجرت عناصر سنگین به سمت ریشه گیاه، جذب آن و نهایتاً بروز سمیت برخی عناصر وارد شده به خاک را تشدید می‌کند. در واقع با افزایش شوری، امکان جابجایی فلزات سنگین وارد شده به خاک و جذب آن‌ها به وسیله گیاه افزایش می‌یابد (سلیمی و همکاران، 1390).

روش‌های آزمایشگاهی اندازه‌گیری مقدار فلزات سنگین (کل یا قابل جذب) در خاک عمدتاً وقت‌گیر و هزینه‌بر است (به‌ویژه در سطوح مطالعاتی وسیع که تعداد نمونه‌های فراوان برای در نظر گرفتن تغییرات زمانی و مکانی نیاز است). به همین سبب امروزه پژوهشگران تمایل فراوانی به استفاده از روش‌های غیرمستقیم برای تخمین ویژگی‌های خاک (نظیر توابع انتقالی) با استفاده از ویژگی‌های زود یافت دارند (شیرانی و همکاران، 2016). در همین راستا گزارش شده است که سیستم‌های هوش مصنوعی در طبقه‌بندی، شناسایی و برآورد این ویژگی‌های دیر یافت عملکرد مطلوبی داشته‌اند (بسالت پور و همکاران، 1394؛ کیم کیم، 2002؛ بسالت پور و همکاران، 2014؛ شیرانی و همکاران، 2016). مزیت هوش مصنوعی، قابلیت پیش‌بینی و مقرون به صرفه بودن آن بوده و به همین سبب از آن تاکنون در مباحث خاک-شناسی گوناگونی استفاده شده است. برای نمونه می‌توان به استفاده از هوش مصنوعی در تعیین ویژگی‌های مؤثر بر شاخص‌های فیزیکی خاک‌های آهکی (شیرانی و همکاران، 2016)، برآورد پایداری خاکدانه‌ها و مقاومت برشی خاک (بسالت پور و همکاران، 2013) پیش‌بینی میزان رس و پی-هاش خاک (هندرسون و همکاران، 2005) و نقشه برداری زمین‌های خشک و شور (اسپنسر و همکاران، 2004) اشاره کرد. با این وجود به نظر می‌رسد قابلیت هوش مصنوعی و روش‌های نوین مدل‌سازی همچون ماشین‌های بردار پشتیبان (Support Vector Machines, SVMs) در پیش‌بینی تغییرات مکانی غلظت فلزات سنگین در خاک حاشیه جاده‌ها و عوامل خاک‌شناسی مؤثر بر آن بررسی نشده باشد.

ماشین‌های بردارپشتیبان از جمله روش‌های یادگیری نظارت شونده می‌باشند که از آن‌ها برای طبقه‌بندی و رگرسیون استفاده می‌گردد (بسالت پور و همکاران، 1394). الگوریتم ماشین‌های بردار پشتیبان، نوع خاصی از مدل‌های خطی را می‌یابد که حاشیه‌ی ابر صفحه بیشینه شود. بیشینه کردن حاشیه‌ی ابر صفحه، منجر به حداکثر شدن تفکیک بین طبقات می‌شود. در واقع، نزدیک‌ترین نقاط آموزشی به حداکثر حاشیه‌ی ابر صفحه را بردارهای پشتیبان می‌نامند که از آن‌ها برای مشخص کردن مرز بین طبقات استفاده می‌شود. اگر داده‌ها به صورت خطی قابل جداسازی باشند، از ماشین‌های خطی برای تولید یک سطح بهینه که داده‌ها را بدون خطا و با حداکثر فاصله‌ی میان صفحه و نزدیکترین نقاط آموزشی (بردارهای پشتیبان) تفکیک می‌نماید، استفاده می‌شود (وپ نیک، 1998، بسالت پور و همکاران، 1394).

از جمله خصوصیات خاک، غلظت عناصر سنگین در خاک است که به دلیل غیرقابل تجزیه بودن و اثرات سوء آن‌ها بر موجودات زنده حائز اهمیت می‌باشند (یالکین و همکاران، 2007). عناصر سنگینی چون سرب به دلیل توانایی بالقوه در آسیب رسانی به سلامت انسان‌ها و حیوانات در چند دهه اخیر از نظر مسائل زیست محیطی بسیار مورد توجه قرار گرفته و سعی شده که از ورود آن به چرخه محیط زیست تا حد امکان جلوگیری شود. غلظت سرب در خاک‌ها بین 1 تا 200 میلی‌گرم بر کیلوگرم و به طور متوسط 15 میلی‌گرم بر کیلوگرم و حد بحرانی آن 50 میلی‌گرم بر کیلوگرم می‌باشد (کاباتا، 2011).

در این پژوهش با اندازه‌گیری غلظت فلز سرب در خاک سطحی حاشیه جاده دزفول- اهواز، به بررسی میزان اثرپذیری روند توزیع مکانی این فلز و نیز ارتباط غلظت آن با ویژگی‌های مختلف فیزیکی- شیمیایی خاک با استفاده از ماشین‌های بردار پشتیبان پرداخته شد. همچنین بررسی قابلیت استفاده از روش نوین ماشین‌های بردار پشتیبان برای مدل‌سازی تغییرات غلظت این عنصر در خاک سطحی حاشیه جاده دزفول- اهواز در مقایسه با روش مرسوم رگرسیون خطی از دیگر اهداف این پژوهش بود.

مواد و روش‌ها

منطقه مورد مطالعه، نمونه‌برداری و آزمایش‌های خاک

این پژوهش در بهار سال 1393 در اراضی حاشیه بزرگراه دزفول - اهواز انجام شد. شهرستان دزفول با مساحت 4762 کیلومتر مربع بین 48 درجه و 20 دقیقه تا 48 درجه و 31 دقیقه طول شرقی و 32 درجه و 75

دقیقه عرض شمالی قرار گرفته است. اهواز نیز یکی از شهرهای بزرگ ایران و مرکز استان خوزستان است که در جلگه خوزستان به ارتفاع 18 متر از سطح دریا واقع شده و از جمله کلان‌شهرهای آلوده کشور می‌باشد. در پژوهش حاضر، نمونه‌برداری از خاک به صورت مرکب و در یک خط مجاور بزرگراه دزفول- اهواز صورت گرفت. نمونه- برداری از خاک سطحی (عمق 0 تا 10 سانتی‌متری) در خطوط گوناگون شامل فاصله 0 تا 15، 15 تا 45 و 45 تا 100 متری از جاده، انجام شد. از هر خط، 3 نمونه خاک تهیه و موقعیت جغرافیایی هر نقطه پس از نمونه‌برداری با دستگاه GPS ثبت گردید. در مجموع، تعداد 135 نمونه خاک برداشت شد که با توجه به طول مسیر اتوبان (150 کیلومتر) و تعداد نمونه برداشته شده در سه فاصله از محور جاده (45 نقطه نمونه برداری)، فواصل نقاط نمونه برداری 4 کیلومتر بود که پوشش دهنده کل منطقه مورد مطالعه می‌باشد. پس از انتقال نمونه‌ها به آزمایشگاه، برخی ویژگی‌های فیزیکی و شیمیایی خاک اندازه‌گیری شد.

میزان ماده آلی خاک با اندازه‌گیری کربن آلی به روش اکسایش با دی‌کرومات پتاسیم (والکلی و بلک، 1935) و استفاده از رابطه [درصد کربن آلی $\times 1/724 =$ درصد ماده آلی] محاسبه شد. توزیع اندازه‌ای ذرات خاک (درصد رس، سیلت، شن ریز و خیلی ریز و درصد کل شن) به روش پیت اندازه‌گیری و تعیین گردیدند (جی و بودر، 1986). مقدار کربنات کلسیم معادل خاک به روش تیتراسیون با سود (آدرانو، 1998)، مقدار اسیدیته خاک در گل اشباع با استفاده از دستگاه pH متر (مدل MI151، Milwaukee ساخت کشور ایتالیا)، (کلوت، 1986) و قابلیت هدایت الکتریکی (EC) در عصاره گل اشباع با استفاده از دستگاه EC متر (مدل JENWAY 9500 Do₂ meter ساخت کشور انگلستان) تعیین شدند (پیک و همکاران، 2003). غلظت کل سرب در خاک نیز به روش هضم تر با اسید و قرائت با دستگاه جذب اتمی (مدل ContraAA300 ساخت آلمان) اندازه‌گیری شد (کاو و همکاران، 1984). دلیل استفاده از غلظت کل سرب در این پژوهش، قابلیت تبدیل آن به شکل‌های قابل جذب و بنابراین لحاظ نمودن تمامی مخاطرات زیست محیطی آن بود. پس از اندازه‌گیری و تعیین ویژگی‌های خاک، نقاط مورد مطالعه، برخی ویژگی‌های آماری داده‌ها نظیر بیشینه، کمینه، میانگین، واریانس و انحراف معیار با استفاده از نرم‌افزار SPSS تعیین شدند.

مدل‌سازی با رگرسیون خطی چند متغیره (MLR)

رگرسیون‌های خطی از جمله روش‌های ساده‌ی مدل‌سازی می‌باشند که از آن‌ها می‌توان برای پیش‌بینی یک

پشتیبان هستند. در این معادله، پارامترهای b و a_i تعیین کننده‌ی ابر صفحه هستند. اگر داده‌ها به صورت خطی قابل تفکیک نباشند، معادله‌ی 6 به معادله‌ی زیر تغییر می‌یابد:

$$y = \text{sign} \left(\sum_{i=1}^n y_i a_i K(X, X_i) + b \right) \quad (4)$$

تابع $K(X, X_i)$ تابع کرنلی¹ است که برای ایجاد ماشین‌هایی با انواع مختلفی از سطوح تصمیم‌گیری غیرخطی در فضای داده‌ها، ضرب‌های داخلی تولید می‌کند (وونگ و همکاران 2009، بسالت پور و همکاران 1394). شکل 1 فرآیند مدل SVM را نشان می‌دهد. مدل‌سازی با ماشین‌های بردار پشتیبان در این پژوهش در محیط نرم افزار Clementine انجام شد.

برای ساخت و آزمون هر دو مدل مورد استفاده در این پژوهش (مدل رگرسیون خطی چند متغیره و مدل SVM)، از 80 درصد داده برای ساخت و آموزش مدل، 10 درصد برای اعتبارسنجی و 10 درصد باقی‌مانده (14 نمونه) برای آزمون کارایی مدل‌ها استفاده شد.

شاخص‌های ارزیابی کارایی مدل‌ها

جهت ارزیابی نتایج مدل‌ها و بررسی کارایی و دقت آنها برای برآورد غلظت سرب در این پژوهش، از برخی شاخص‌های آماری نظیر ضریب همبستگی (r)، جذر میانگین مربعات خطا (RMSE)، شاخص کارایی مدل و درصد خطا (ERROR%) استفاده شد که معادله‌ی دو شاخص آخر به صورت زیر می‌باشند:

$$MEF = 1 - \frac{\sum_{i=1}^n |y_k - \hat{y}_k|^2}{\sum_{i=1}^n |y_k - \bar{y}_k|^2} \quad (5)$$

$$ERROR \% = \frac{\sum_{k=1}^n |y_k - \hat{y}_k|}{\sum_{k=1}^n y_k} \times 100 \quad (6)$$

که در این روابط، y_k و \hat{y}_k به ترتیب مقادیر مشاهده‌ای و برآوردشده، \bar{y} متوسط مقادیر مشاهده‌ای و n تعداد کل مشاهدات می‌باشند.

یا چند متغیر پاسخ (وابسته) از مجموعه‌ای از مقادیر متغیرهای پیش‌بینی کننده (مستقل) استفاده نمود. رگرسیون‌های خطی به صورت دو متغیره و چندمتغیره می‌باشند که در رگرسیون دو متغیره، یک متغیر مستقل و یک متغیر وابسته وجود دارد. در رگرسیون چندمتغیره نیز، یک متغیر وابسته و دو یا چند متغیر مستقل وجود دارد. معادله‌ی کلی یک رگرسیون خطی چندمتغیره به صورت رابطه زیر می‌باشد (بسالت پور و همکاران، 2014):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n \quad (1)$$

که در آن، Y متغیر وابسته، X متغیر مستقل و $\beta_1 \dots \beta_n$ ضرایب رگرسیونی می‌باشند. مدل‌سازی رگرسیونی در محیط نرم‌افزار SPSS انجام شد.

مدل‌سازی با ماشین‌های بردار پشتیبان (SVM)

ماشین‌های بردار پشتیبان از جمله روش‌های هوش مصنوعی می‌باشند که از آن‌ها برای طبقه‌بندی و رگرسیون استفاده می‌شود (بسالت پور، 1394). الگوریتم SVM نوع خاصی از مدل‌های خطی را می‌یابد که حاشیه‌ی ابر صفحه بیشینه شود. بیشینه کردن حاشیه‌ی ابر صفحه، منجر به حداکثر شدن تفکیک بین طبقات می‌شود. در واقع، نزدیک‌ترین نقاط آموزشی به حداکثر حاشیه‌ی ابر صفحه را بردارهای پشتیبان می‌نامند که از آن‌ها برای مشخص کردن مرز بین طبقات استفاده می‌شود. اگر داده‌ها به صورت خطی قابل جداسازی باشند، از ماشین‌های خطی برای تولید یک سطح بهینه که داده‌ها را بدون خطا و با حداکثر فاصله‌ی میان صفحه و نزدیکترین نقاط آموزشی (بردارهای پشتیبان) تفکیک می‌نماید، استفاده می‌شود. به منظور مفهوم‌سازی چگونگی عملکرد ماشین‌های بردار پشتیبان، نقاط آموزشی به صورت رابطه‌ی 5 در نظر گرفته شود (وپ نیک، 1998):

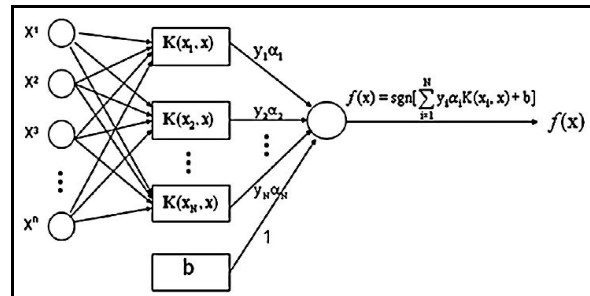
$$D = \{(x_i, y_i)\}_{i=1}^n \quad (2)$$

که در این رابطه، x_i بردار ورودی بوده که $x_i \in X$ و y_i مقدار مطلوب می‌باشد که $y_i \in R$ است و n تعداد کل الگوهای داده‌ای می‌باشد. اگر داده‌ها به صورت خطی قابل تفکیک باشند؛ آن‌گاه قواعد تصمیم‌گیری که توسط یک صفحه‌ی بهینه که طبقات تصمیم‌گیری باینری را تفکیک می‌کند به صورت معادله‌ی زیر خواهد بود:

$$y = \text{sign} \left(\sum_{i=1}^n y_i a_i (X \cdot X_i) + b \right) \quad (3)$$

که در آن، y خروجی معادله، y_i ارزش طبقه‌ی نمونه‌ی آموزشی، بردار $x = (x_1, x_2, \dots, x_n)$ نشان‌دهنده‌ی یک داده‌ی ورودی و بردارهای X_i ($i=1, \dots, N$) بردارهای

¹ Kernel function



شکل 1- فرآیند مدل SVM بردار ورودی، $X=[x^1, x^2, \dots, x^n]$ بردار $[y_1\alpha_1, y_2\alpha_2, \dots, y_n\alpha_n]$ بردار وزن‌ها و $f(x)$ معادله تصمیم‌گیری (بسالت پور و همکاران، 1394)

نتایج و بحث

ویژگی‌های فیزیکی - شیمیایی خاک منطقه مورد مطالعه

توصیف آماری ویژگی‌های فیزیکی و شیمیایی خاک منطقه مورد مطالعه در جدول 1 آمده است. بیشینه، کمینه و میانگین pH خاک منطقه مورد مطالعه در عمق 0 تا 10 سانتی‌متری به ترتیب برابر 9/10، 6/86، 7/76 بود (جدول 1). با توجه به نتایج بدست آمده، خاک سطحی منطقه مورد مطالعه دارای pH قلیایی است که می‌تواند بسیاری از ویژگی‌های شیمیایی، حاصلخیزی، رشد گیاه و فعالیت ریز جانداران خاک و جذب و انتقال عناصر در خاک تحت تأثیر قرار دهد. بسیاری از واکنش‌های شیمیایی موثر بر غلظت عناصر در خاک و فعالیت موجودات زنده نیز به اسیدیته خاک وابسته بوده و بر اساس پژوهش‌های انجام شده pH بین 5 تا 9 محدوده مناسبی جهت فعالیت‌های زیستی بیان شده است (آبرز، 1985). نتایج بررسی قابلیت هدایت الکتریکی خاک، نشان دهنده شور بودن خاک منطقه مورد بررسی بود (میانگین برابر 37/31 دسی‌زیمنس بر متر، جدول 1).

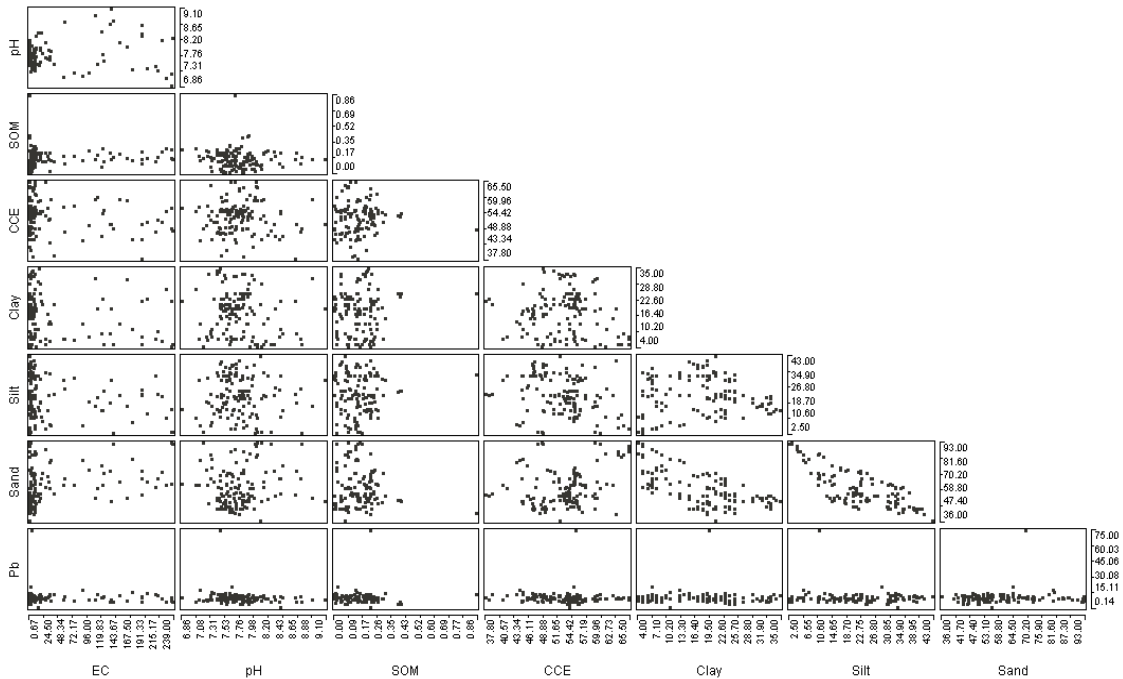
همچنین با بررسی شاخص چولگی محاسبه شده برای این ویژگی‌ها (EC و pH) مشخص شد که این پارامترها فاقد توزیع آماری نرمال بوده و بنابراین نرمال سازی با استفاده از روش‌های مرسوم نرمال‌سازی (لگاریتم طبیعی، جذر و غیره) انجام شد. بیشتر اراضی قابل کشت در استان خوزستان دارای میزان کربنات کلسیم معادل بالایی هستند. خاک منطقه مورد مطالعه در این پژوهش نیز دارای میزان کربنات کلسیم معادل نسبتاً زیادی بود به گونه‌ای که میانگین، کمینه و بیشینه کربنات کلسیم معادل در عمق 0-10 سانتی‌متری خاک به ترتیب 65/50، 37/80، 52/87 بود (جدول 1). خاک‌های کربنات کلسیم معادل بخش چشمگیری از اراضی قابل کشت کشور را به خود اختصاص داده‌اند که ویژگی این خاک‌ها نقش مؤثری بر فراهمی و نگهداشت عناصر در خاک دارند.

شناخت و تعیین این ویژگی می‌تواند نقش بسزایی در مدیریت خاک‌ها داشته باشد. ماده آلی در خاک منطقه مورد مطالعه در محدوده صفر تا 0/86 درصد متغیر و میانگین آن 0/15 درصد بود (جدول 1). ماده آلی یکی از مهم‌ترین شاخص برآورد کیفیت خاک است (دوران و پارکین، 1996) که نقش بسزایی در نگهداشت و جلوگیری از جابجایی عناصر در خاک دارد. برای این ویژگی نیز بررسی شاخص چولگی مشخص نمود که این پارامتر نیز نیاز به نرمال‌سازی دارد. میزان سرب خاک منطقه مورد مطالعه دارای دامنه تغییرات از 0/14 تا 75/00 میلی گرم بر کیلوگرم بود (جدول 1). مطابق نتایج حاصل از اندازه‌گیری بافت و تعیین درصد شن، سیلت و رس نمونه‌های خاک مورد بررسی، مقادیر میانگین آن‌ها به ترتیب 17/60، 22/39 و 60 درصد بود (جدول 1). با توجه به مقادیر حاصله بافت خاک بر اساس سیستم طبقه بندی امریکایی (2011)، لومی - شنی و لوم رس شن نامگذاری شد.

بررسی ماتریس پراکنش ویژگی‌های فیزیکی - شیمیایی خاک منطقه مورد مطالعه و ارتباط آن‌ها با یکدیگر (شکل 2)، نمایانگر آن بود که اگرچه به نظر می‌رسد که تغییرات غلظت سرب در خاک سطحی این منطقه با ویژگی‌های فیزیکی و شیمیایی اندازه‌گیری شده ارتباط داشته باشد ولی با این وجود تعیین این ارتباط با استفاده از روش‌های خطی مرسوم امکان‌پذیر نبوده و استفاده از روش‌های غیر خطی باید مورد توجه قرار گیرد.

جدول 1- توصیف آماری ویژگی‌های اندازه‌گیری شده

پارامتر	شاخص آماری				
	میانگین	کمینه	بیشینه	چولگی	انحراف معیار
رس	17/60	4/00	35/00	0/09	8/55
سیلت	22/39	2/50	43/00	-0/13	10/25
شن	60/00	36/00	93/00	0/84	13/78
ماده‌ی آلی	0/15	0/00	0/86	2/32	0/11
کربنات کلسیم	52/87	37/80	65/50	-0/10	5/58
pH	7/76	6/86	9/10	0/96	0/36
قابلیت هدایت الکتریکی	37/13	0/67	239/00	1/94	65/26
سرب کل	9/50	0/14	75/00	8/25	6/36



شکل 2- ماتریس پراکنش ویژگی‌های فیزیکی- شیمیایی خاک مورد مطالعه و ارتباط آن‌ها با یکدیگر (EC: قابلیت هدایت الکتریکی، SOM: درصد ماده آلی خاک، CCE: درصد کربنات کلسیم معادل، Pb: غلظت کل سرب خاک)

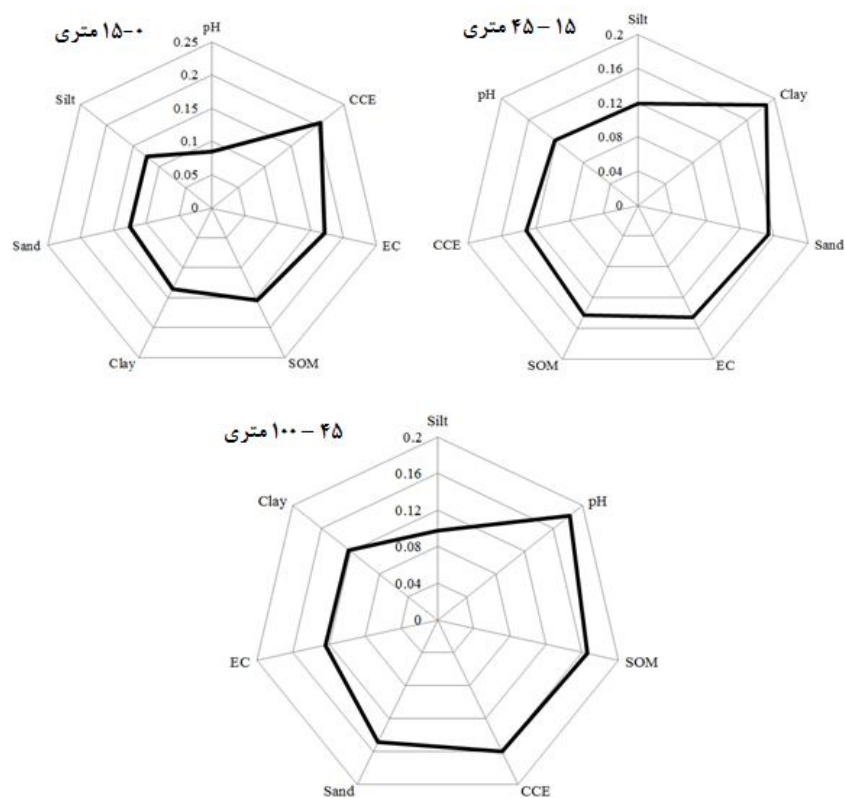
معادل خاک از طریق تشکیل کمپلکس با فلزات می‌تواند سبب افزایش جذب و نگهداری آن‌ها شود (بروکس و هرمان، 1998). همچنین pH خاک عامل مهمی در کنترل فراوانی و توزیع مکانی فلزات سنگین در خاک است (لینسای، 1979). بررسی نمودارهای حلالیت ترکیبات کربناتی و فسفوری برخی فلزات سنگین در خاک‌های دارای کربنات کلسیم معادل نیز نشان داد که حلالیت این ترکیبات با افزایش pH خاک (تا حدود 7/9) افزایش یافته

ویژگی‌های مؤثر بر توزیع مکانی سرب (انتخاب متغیر) نتایج نشان داد که در فاصله 0-15 متری، کربنات کلسیم معادل بیش‌ترین ضریب اهمیت را در جذب سرب داشته و بقیه فاکتورها از اهمیت کمتری برخوردار بودند. در فاصله 15-45 سانتی متری میزان رس خاک و در فاصله 45-100 متری نیز pH بیش‌ترین ضریب اهمیت را در جذب سرب داشت (شکل 3). مطالعات گوناگون نشان داده‌اند که میزان کربنات کلسیم

خوبی برای جذب عناصر سنگین داشته باشد (سینگارام و همکاران، 2006). بررسی همبستگی پارامترهای هم-دماهای جذب سرب با ویژگی‌های فیزیکی و شیمیایی خاک نیز نشان داده است که ویژگی‌های خاک شامل درصد رس، ماده آلی و گنجایش تبادل کاتیونی، مهم-ترین پارامترها در جذب سرب هستند (گومز و همکاران، 2001).

و بعد از این pH، مقدار حلالیت روندی کاهشی را نشان می‌دهد (کرکهام، 2002).

اجزاء بافت خاک نیز اهمیت بالایی در حرکت و تغییرات توزیع مکانی سرب در خاک دارد. رس‌ها (همچون رس پالیگوسکایت)، توانایی قابل توجهی در جذب عناصر، یون‌ها و مولکول‌های معدنی و آلی دارند. بنابراین انتظار می‌رود خاک‌های حاوی رس زیاد، پتانسیل



شکل 3- ویژگی‌های خاک مؤثر بر جذب سرب در خاک سطحی حاشیه جاده دزفول - اهواز (فواصل 0-15، 15-45 متر و 45-100) حاصل از ماشین‌های بردار پشتیبان

(EC: قابلیت هدایت الکتریکی، SOM: درصد ماده آلی خاک و CCE: درصد کربنات کلسیم معادل)

مشخص است که در بسیاری از نقاط به صورت غیریکنواخت، بیش‌برآوردی از غلظت سرب صورت گرفته است و در برخی دیگر از نقاط، غلظت سرب کمتر از مقادیر مشاهده‌ای آن‌ها تخمین زده شده است. با توجه به ضریب تبیین پایین ($R^2=0/38$) بین مقادیر مشاهده-ای و شبیه‌سازی شده برای داده‌های تست مدل رگرسیونی طراحی شده در این پژوهش (شکل 4) و نیز بر اساس مقادیر شاخص‌های آماری ارزیابی مدل مورد استفاده به نظر می‌رسد که مدل رگرسیونی خطی چندمتغیره فاقد دقت قابل قبول برای برآورد و مدل-سازی تغییرات غلظت سرب در منطقه مورد مطالعه

مدل‌سازی تغییرات غلظت سرب

مدل رگرسیونی خطی چند متغیره

مقادیر شاخص‌های آماری مورد استفاده برای ارزیابی کارایی مدل رگرسیونی طراحی شده برای برآورد غلظت سرب در جدول 2 آورده شده است. مقادیر r ، MEF ، $RMSE$ ، $ERROR\%$ برای مدل رگرسیونی طراحی شده برای برآورد سرب به ترتیب برابر با $0/25$ ، $0/38$ ، $0/62$ و $0/19$ درصد بودند. مقادیر مشاهده‌ای و شبیه‌سازی شده‌ی غلظت سرب برای داده‌های تست مدل نیز در شکل 4 مقایسه شده‌اند. با توجه به این شکل،

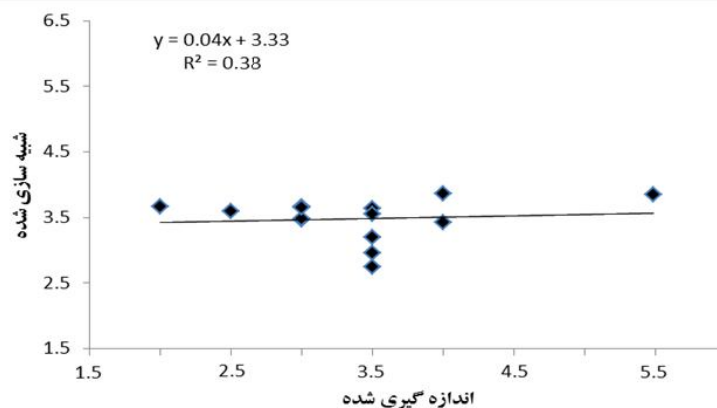
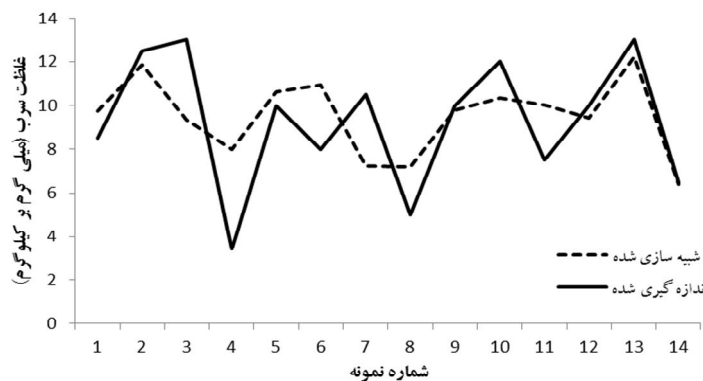
؛ بسالت پور و همکاران، 2013). بنابراین به نظر می‌رسد دلیل اصلی عملکرد ضعیف مدل رگرسیونی خطی در برآورد ویژگی مورد مطالعه در این پژوهش، روابط غیرخطی بین این ویژگی و ورودی‌های مورد استفاده در مدل‌سازی باشد.

است. در واقع مدل‌های رگرسیون خطی تنها قادر به تشخیص روابط خطی بین متغیرهای مستقل و وابسته هستند و چنانچه روابط غیرخطی بین این متغیرها وجود داشته باشد، کارایی این مدل‌ها به حد قابل ملاحظه‌ای کاهش می‌یابد (تاراکوی و همکاران، 2009).

جدول 2- شاخص‌های آماری مورد استفاده برای ارزیابی کارایی مدل‌های مورد مطالعه برای برآورد غلظت سرب

شاخص آماری				
مدل	RMSE	MEF	r	ERROR%
رگرسیون خطی چندمتغیره	2/25	0/38	0/62	0/19
ماشین‌های بردار پشتیبان	0/83	0/70	0/86	0/15

R: ضریب هم‌بستگی، RMSE: ریشه میانگین مربعات خطا، ERROR%: درصد خطا و MEF: فاکتور بازدهی مدل



شکل 4- مقایسه مقادیر مشاهده‌ای و شبیه‌سازی شده غلظت سرب برای داده‌های آزمون مدل رگرسیون خطی چندمتغیره

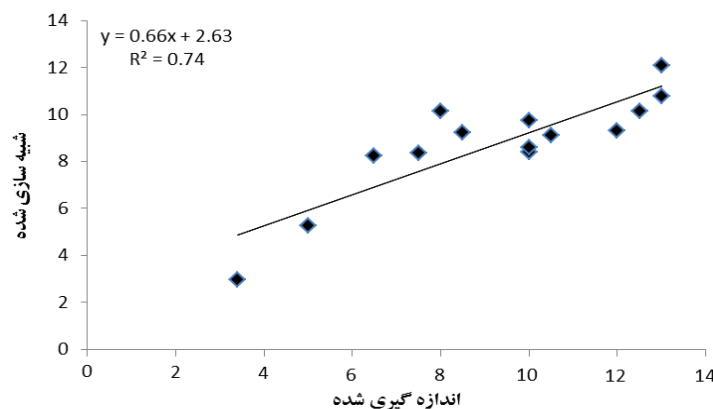
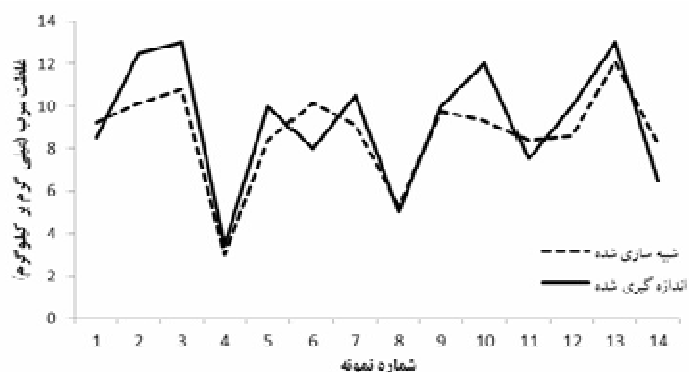
اساس مقادیر نسبتاً قابل قبول شاخص‌های آماری ارزیابی مدل، ماشین‌های بردار پشتیبان دقت بیشتری نسبت به مدل رگرسیونی در برآورد و مدل‌سازی غلظت سرب در منطقه‌ی مورد مطالعه داشته باشند. در واقع، ماشین‌های بردار پشتیبان ساخته‌شده با تابع کرنلی، دارای انعطاف پذیری زیاد و قابلیت بالایی در تشخیص روابط غیرخطی و پیچیده‌ی بین داده‌های ورودی و خروجی می‌باشند (لیو و همکاران، 2014). از سوی دیگر، روش‌های مبتنی بر

مدل ماشین‌های بردار پشتیبان

مقادیر r ، MEF ، $RMSE$ و $ERROR\%$ برای مدل SVM در برآورد سرب به ترتیب برابر با 0/83، 0/70، 0/86 و 0/15 درصد بودند. هم‌چنین ضریب تبیین نسبتاً مناسبی ($R^2=0/74$) بین مقادیر مشاهده‌ای و شبیه‌سازی شده برای داده‌های تست مدل SVM حاصل شد، اگرچه در برخی از نقاط، کم‌برآوردی از غلظت سرب وجود داشت (شکل 5). بنابراین به نظر می‌رسد که بر

این پژوهش می‌باشند که به نظر می‌رسد مدل ماشین‌های بردار پشتیبان در مقایسه با مدل رگرسیون خطی دارای انعطاف‌پذیری بیشتری بوده است. بنابراین به نظر می‌رسد که در مناطقی که پیچیدگی‌های محیطی زیاد و بین داده‌های ورودی و خروجی، روابط غیرخطی پیچیده وجود داشته و درعین حال تعداد داده‌های قابل دسترس کم می‌باشد، روش‌های هوش مصنوعی مانند ماشین‌های بردار پشتیبان نسبت به روش‌های مرسوم مدل‌سازی مانند رگرسیون‌های خطی برتری دارند (بسالت پور و همکاران، 1394).

آموزش در مقایسه با مدل‌های رگرسیون خطی از توانایی بیشتری در تشخیص روابط بین ورودی‌ها و خروجی برخوردار می‌باشند (بسالت پور و همکاران، 2013). همچنین روش‌های هوش مصنوعی مانند ماشین‌های بردار پشتیبان در مقایسه با روش‌های قدیمی و مرسوم مدل‌سازی مانند رگرسیون‌های خطی دارای قابلیت بالایی در تشخیص روابط پیچیده‌ی بین داده‌های ورودی و خروجی با تعداد نمونه‌ی کم هستند (بسالت پور و همکاران، 2014). پراکنش نمونه‌ها، تغییرات مکانی نقاط نمونه‌بردی و مقیاس منطقه‌ی مورد بررسی نیز از دیگر عوامل مؤثر بر قابلیت مدل‌ها در برآورد و مدل‌سازی غلظت سرب در



شکل 5- مقایسه مقادیر مشاهده‌ای و شبیه‌سازی شده غلظت سرب برای داده‌های آزمون مدل ماشین‌های بردار پشتیبان

در برآورد و مدل‌سازی غلظت سرب دارد. مدل‌های رگرسیون خطی مرسوم تنها قادر به تشخیص روابط خطی بین متغیرها می‌باشند در حالی که به سبب تنوع و پیچیدگی ویژگی‌های زیست محیطی همچون ویژگی‌های خاک، عمده روابط بین ورودی‌ها و خروجی‌ها به صورت غیرخطی است که این مدل‌ها توان تشخیص این روابط را ندارند. بنابراین می‌توان نتیجه گرفت که روش‌های نوین مدل‌سازی (روش‌های مبتنی بر هوش مصنوعی مانند ماشین‌های بردار پشتیبان) نسبت به روش‌های مرسوم و

نتیجه‌گیری

نتایج حاصل از این پژوهش نمایان‌گر آن بود که در فاصله 0-15 متری از حاشیه جاده دزفول - اهواز، کربنات کلسیم معادل بیش‌ترین ضریب اهمیت را در جذب سرب دارد. در فاصله 15-45 سانتی‌متری نیز میزان رس خاک و در فاصله 45-100 متری pH ویژگی‌های خاک مؤثر در جذب سرب بودند. بررسی قابلیت استفاده از روش رگرسیون خطی و مدل ماشین‌های بردار پشتیبان (SVMs) نیز بیان‌گر آن بود که مدل SVMs دقت بالاتری

قدیمی (رگرسیون خطی) که همچنان به شکل وسیعی در علوم مهندسی کشاورزی و محیط زیست استفاده می‌شوند، دارای قابلیت بالاتری در تشخیص روابط پیچیده بین متغیرهای ورودی و خروجی با تعداد نمونه کم هستند و استفاده از آن‌ها به منظور دستیابی به بیش‌ترین دقت و در عین حال صرف کمترین هزینه و زمان توصیه می‌گردد.

فهرست منابع:

1. بسالت پور، ع.ا.، ح. شیرانی، و ع. اسفندیارپور. مدل‌سازی پایداری خاکدانه‌ها با استفاده از ماشین‌های بردار پشتیبان و رگرسیون چند متغیره خطی. مجله آب و خاک. 1394. 29(2): 406-417.
2. حمزه، م. ع. نشان‌گرهای ژئوشیمیایی و زیست محیطی در محدوده شهری. پایان نامه کارشناسی ارشد. دانشگاه شهید باهنر کرمان. 1385. 387 ص.
3. سلیمی، م.، امین، م.، ا. ابراهیمی، ا.، قاضی فرد، ا.، نجفی، پ.، امینی، ح.، رزمجو، پ. و وحید دستجردی، ک. تأثیر شوری بر گیاه پلائی کادمیوم از خاک‌های آلوده. مجله تحقیقات نظام سلامت. 1390. 7(6): 1130-1137.
4. Adriano, D. C. Trace element in the terrestrial environment, Springer Vsley, New York, 1998. 520 pp.
5. Ayres, A. D., and D.W. Westcot. Water quality for agriculture. F.A.O. Irrigation and Drainage bulletin. 1985. No 29.
6. Besalatpour A.A., Ayoubi S., Hajabbasi M.A., Gharipour A., and A. Yousefian Jazi. Feature selection using parallel genetic algorithm for the prediction of geometric mean diameter of soil aggregates by machine learning methods. Arid Land Research and Management. 2014. 28: 383-394.
7. Besalatpour A.A., Hajabbasi M.A., Ayoubi S., Mosaddeghi M.R., and R. Schulin. Estimating wet soil aggregate stability from easily available data in a highly mountainous watershed. Catena. 2013. 111: 72-79.
8. Brooks S.C., and Herman J.S. Rate and extent of cobalt sorption to representative aquifer minerals in the presence of a moderately strong ligand. Applied Geochemistry. 1998. 13: 77-88.
9. Cao, H.F., Chang, A.C., and Page, A.L. Heavy metal contents of sludge treated soils as determined by three extraction procedures. Journal of Environmental Quality. 1984. 13: 632-634.
10. Chang C, Wang B, Shi L, Li Y, Duo L and Zhang W. Alleviation of salt stress-induced inhibition of seed germination in cucumber (*Cucumis sativus* L.) by ethylene and glutamate. Journal of Plant Physiology, 2010. 167: 1152-1156.
11. Coelho, M. C., faras, T. L. and Roupail, N.M. "Impact of speed control traffic signals on pollutant emissions", Tranportation Research Part D. 2005. 10: 323- 340.
12. Doran JW, Parkin TB. Quantitative indicators of soil quality: a minimum data ser. International Journal. Doran W, Jones AJ (Eds), Methods of Assessing Soil Quality. Soil Science society of America. Special Pub. 1996. 49: 25-37
13. Gee G.W., and Bauder J.W. Particle size analysis. In: Klute, A. (Ed.), Methods of Soil Analysis: Part 1., Agronomy Handbook No 9., American Society of Agronomy and Soil Science Society of America, Madison, WI. 1986. 383-411.
14. Gomes P.C., Fontes M.P., Dasilva A.G., Mendoca E.S., and Netto A.R. Selectivity sequence and competitive adsorption of heavy metal by Brazilin soils. Soil Science Society of America Journal. 2001. 65: 1115-1121.
15. Henderson BL, Bui EN, Moran CJ, Simon DAP. Australia-wide predictions of soil properties using decision trees. Geo derma, 2005. 124(3): 383-398.

16. Kabata A. Trace Element in Soils and Plants. CRC Press, Boca Raton Ann. Arbor. London, 2011.
17. Kim MJ, Kim TS. A neural classifier with fraud density map for effective credit card fraud detection. In Intelligent Data Engineering and Automated Learning-IDEAL, 2002. 378-383.
18. Kirkham, M.B. Cadmium in plants on polluted soils: Effects of soil factors, hyperaccumulation, and amendments. *Geoderma*. 2002. 137: 19-32.
19. Klute, A. Methods of soil analysis, part I, physical and mineralogical methods. Second edition. Soil Science Society of America INC. Wisconsin. USA. 1986.
20. Lindsay, W.L. Chemical equilibria in soils. John Wiley & Sons, New York. 1979; 412p.
21. Liao K., Xu S., Wu J, Zhu Q., and An L. Using support vector machines to predict cation exchange capacity of different soil horizons in Qingdao City, China. *J. Plant Nutrition Soil Science*. 2014. 177(5): 775-782.
22. Ozkutlu F, Ozturk L, Erdem H, McLaughlin M, CakmakI. Leaf-applied sodium chloride promotes cadmium accumulation in durum wheat grain. *Plant and soil*. 2007. 290(1-2): 323-31.
23. Peak, D., G. W. Luther, and D. L. Sparks. ATR-FTIR spectroscopic studies of boric acid adsorption on hydrous ferric oxide. *Geochimica Et Cosmochimica Acta*. 2003. 67(14): 2551-2560.
24. Sarkar, B. "Heavy Metals in environment" , NewYork: Marcel Dekker, 2002.
25. Shirani, H., Habibi, M., Besalatpour A.A., and Esfandiarpour I. Physical quality of calcareous agricultural soils in a semiarid region of Iran: tackling challenges with the affecting parameters using a hybrid PSO-DT algorithm. *Geoderma*. 2016. 259-260: 1-11.
26. Singaram, P., LaIsuna, K., and Mahimairaja, S. Metal contamination in urban soil, water environment and remediation strategies. 18th World congress of soil science July 9_15, 2006_ Philadelphia.
27. Spencer M, Whitfort T, McCullagh J. Mapping dry land salinity using neural networks, AI SpringerVerlag. 2004. 1233-1238.
28. Twarakavi N.K.C., Simunek J., and Schaap M.G. Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. *Soil Science Society American Journal*. 2009. 73: 1443-1452.
29. Vapnik V. *Statistical Learning Theory*, Wiley, New York. 1998.
30. Walk ley A, Black IA. An examination of Degtjaref method for determination of soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science*.1934. Soc. Am. J. 37: 27-29.
31. Wang L. *Support Vector Machines: Theory and Applications*. Springer-Ver lag, 2005, New York.
32. Wang W.C., Chau K.W, Cheng C.T., and Qiu L. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J. Hydrology*. 2009. 374: 294-306.

Modeling the Pb Distribution Using Support Vector Machines in Surface Soil of the Lands Surrounding the Dezful-Ahvaz Road

S. Nazariyat¹, M. Hodaji, and A. A. Besalatpour

M.Sc. student. Khorasgan Islamic Azad University of Esfahan. Iran;

E-mail: Snazariyat@yahoo.com

Associate Professor, Faculty member and Education Deputy, Islamic Azad University of Esfahan.

Iran; E-mail: mehran.hodaji1@gmail.com

Assistant Prof., Department of Soil Science, Vali-e-Asr University of Rafsanjan, Iran;

E-mail: a.besalatpour@vru.ac.ir

Received: January, 2016 & Accepted: December, 2016

Abstract

Investigation of heavy metal distributions in environment is vital for the remediation of contaminated soils and environmental monitoring. In this study, the Pb-concentration variation in surface soils surrounding the Dezful-Ahvaz highway and the soil parameters influencing it were investigated. For this purpose, soil sampling was performed from depth of 0-10 cm of adjacent lines with the highway by collecting three identical samples at intervals of 15, 40, and 100 m from the road side (a total of 135 samples). Then, the total Pb-concentration, organic matter, calcium carbonate equivalent (CCE), pH, electrical conductivity (EC), and soil particle size distribution (clay, silt, sand, fine sand, and very fine) were measured in each soil sample. Modeling analyses were performed using support vector machine (SVM) and multiple linear regression (MLR) methods. To investigate the model performances, some statistical indicators including the correlation coefficient (r), root mean square error (RMSE), and model efficiency factor (MEF) were calculated between the measured and the predicted values. The variable selection results using the SVMs method indicated that the CCE was the most effective factor on Pb-adsorption by soil particles at the distance of 0-15 m from the road side and the clay % and pH (followed by sand %) parameters had the highest importance coefficients at the distances of 15-45 and 45-100 m, respectively. The r coefficient and MEF values for estimation of Pb using the SVM model were 0.86 and 0.70 %, while they were 0.62 and 0.38 % for the MLR models. Therefore, according to the obtained results in this study, it appears that it would be possible to use support vector machines for determining the factors influencing the Pb- adsorption by soil surrounding the road sides and for modelling heavy metal concentration variations in soils.

Keywords: Car pollutants, Soil pollution, Variable selection, Heavy metals

¹ Corresponding author: Soil Science Department, Khorasgan Islamic Azad University of Esfahan- M.A student.