



دانشگاه گوارن و منابع طبیعی

نشریه پژوهش‌های حفاظت آب و خاک

جلد بیست و ششم، شماره دوم، ۱۳۹۸

۲۲۵-۲۳۸

<http://jwsc.gau.ac.ir>

DOI: 10.22069/jwsc.2019.15927.3113

کارایی روش‌های حداقل مربعات معمولی (OLS) و شبکه بیزین در تخمین درصد سدیم تبادلی (ESP) خاک

علی باریکلو^۱، * مسلم ثروتی^۲ و محمدصادق علیانی^۳

^۱دانش‌آموخته کارشناسی ارشد گروه علوم و مهندسی خاک، دانشگاه زنجان، آستادیار مرکز آموزش عالی شهید باکری میاندوآب،

دانشگاه ارومیه، آستادیار حوزه معاونت پژوهشی و فناوری، وزارت علوم، تحقیقات و فناوری

تاریخ دریافت: ۹۷/۹/۱؛ تاریخ پذیرش: ۹۷/۱۱/۲۱

چکیده

سابقه و هدف: در علوم خاک، نسبت جذب سدیم و درصد سدیم تبادلی، دو معیار برای ارزیابی درجه سدیمی بودن خاک هستند. برای اندازه‌گیری درصد سدیم تبادلی، اندازه‌گیری ظرفیت تبادل کاتیونی ضروری است که اندازه‌گیری آن وقت‌گیر و پرهزینه است. بنابراین توسعه یک مدل که بتواند از ویژگی‌های زودیافت خاک درصد سدیم تبادلی را تخمین بزند، بسیار بهینه و اقتصادی می‌باشد. پژوهش‌های نشان داده‌اند که بین درصد سدیم تبادلی و نسبت جذب سدیم رابطه وجود دارد. بنابراین، از نسبت جذب سدیم می‌توان برای پیش‌بینی درصد سدیم تبادلی استفاده کرد. به‌همین دلیل تلاش‌های زیادی برای تخمین درصد سدیم تبادلی با استفاده از پارامترهای زودیافت خاک انجام شده است. هدف از این پژوهش، توسعه مدل کم‌ترین مربعات معمولی و شبکه بیزین برای تعیین درصد سدیم تبادلی با استفاده از نسبت جذب سدیم در خاک‌های شهرستان بناب در استان آذربایجان شرقی است.

مواد و روش‌ها: برای نیل به اهداف پژوهش، ۲۰۹ نمونه خاک با استفاده از روش شبکه‌بندی منظم (۲۵۰ × ۲۵۰) از شهرستان بناب برداشته شد. این محدوده ۱۳۰۰ متر بالاتر از سطح آزاد دریاها آزاد بوده و در آب و هوای نیمه‌خشک در شمال‌غرب ایران واقع شده است. برای نیل به اهداف، برخی از ویژگی‌های شیمیایی خاک مانند Na^+ ، Ca^{2+} ، Mg^{2+} ، SAR و ESP نمونه‌های خاک با استفاده از روش‌های آزمایشگاهی اندازه‌گیری شد. سپس، دو مدل کم‌ترین مربعات معمولی و شبکه بیزین برای تخمین درصد سدیم تبادلی از نسبت جذب سدیمی توسعه داده شد. تخمین‌گر OLS، تابع خطی از مقادیر متغیر وابسته بوده که با استفاده از وزن‌های یک تابع غیرخطی از مقادیر متغیرهای مستقل عمل می‌کند. بنابراین، تخمین‌گر OLS وابسته به نحوه استفاده از مقادیر متغیر وابسته است، صرف‌نظر از این‌که چگونه از مقادیر مستقل استفاده می‌کند. شبکه بیزین نیز یک مدل گرافیکی احتمالی است که مجموعه‌ای از متغیرها و وابستگی‌های مشروط را از طریق یک نمودار آسیکلک هدایت‌شده (DAG) نشان می‌دهد.

* مسئول مکاتبه: m.sarvati@urmia.ac.ir

این شبکه یک گراف جهت‌دار غیرمدور است که مجموعه‌ای از متغیرهای تصادفی و نحوه ارتباط مستقل آن‌ها را نشان می‌دهد. در واقع شبکه‌های بیزین می‌توانند مشکلات تصمیم‌گیری را در حین عدم اطمینان حل کنند.

یافته‌ها: ضریب تبیین و جذر میانگین مربعات خطا مدل ESP-SAR با روش‌های کم‌ترین مربعات معمولی و شبکه بیزین به ترتیب ۰/۹۹، ۰/۷۱ و ۰/۹۸، ۱/۶۳ برآورد گردید. بر اساس نتایج آماری فوق هر دو مدل قابل قبول هستند. برای مقایسه مقادیر ESP خاک با استفاده از مدل خاک ESP-SAR با مقادیر ESP اندازه‌گیری شده با استفاده از روش آزمایشگاهی، از آزمون t استفاده شد. نتایج آزمون t نشان داد که تفاوت بین مقادیر ESP خاک تخمینی توسط مدل‌ها و مقادیر اندازه‌گیری شده با روش‌های آزمایشگاهی از لحاظ آماری معنی دار نبود ($P > 0/05$). بنابراین، مدل‌های ESP-SAR توسعه‌یافته می‌تواند یک روش آسان، اقتصادی و کوتاه برای ارزیابی ESP خاک ارائه دهد. شاخص GMER همچنین تخمین کم دو روش را مورد ارزیابی قرار داد.

نتیجه‌گیری: نتایج مطالعه حاضر نشان داد که مدل‌های کم‌ترین مربعات معمولی و شبکه بیزین به‌عنوان مدل‌های ریاضی بین متغیرهای ورودی و خروجی می‌توانند درصد سدیم تبدلی را با محدودیت‌های قابل قبول پیش‌بینی کنند. همچنین روش OLS نسبت به BS به دلیل ضریب تبیین بالا و ریشه میانگین مربعات خطای کم از کارایی بیشتری برخوردار است.

واژه‌های کلیدی: بناب، خاک‌های سدیمی، مدل‌سازی، ویژگی‌های زود یافت

مقدمه

در خاک‌های شور یا شور و سدیمی، به‌علت افزایش غلظت نمک‌های محلول در خاک، یون‌های سدیم، کلر و بور باعث ایجاد مسمومیت در گیاهان می‌شود (۱). همچنین با فزونی سدیم در مکان‌های تبدلی خاک، آثار سوء آن بر رشد و نمو گیاهان به‌صورت مسمومیت‌های گیاهی و عدم توازن تغذیه‌ای ظاهر می‌شود (۱۲).

سدیمی بودن از مهم‌ترین ویژگی‌های خاک‌های متأثر از نمک است که ویژگی‌های فیزیکی و شیمیایی آن‌ها را تحت تأثیر قرار می‌دهد (۷). هدایت هیدرولیکی پایین، تهویه نامناسب، نفوذپذیری کم، pH بالا و ساختمان تخریب‌یافته از ویژگی‌های مهم خاک‌های سدیمی به‌شمار می‌رود (۹). با این تفاسیر مدیریت این خاک‌ها نیازمند اطلاعات و داده‌های به‌هنگام در

راستای کشاورزی دقیق می‌باشد. برای ارزیابی درجه سدیمی بودن خاک از دو ویژگی مهم نسبت جذب سدیم^۱ با حد آستانه‌ای ۱۳ (رابطه ۱) و درصد سدیم تبدلی^۲ (ESP) (رابطه ۲) با حد آستانه‌ای ۱۵ درصد استفاده می‌شود (۱۹).

$$SAR = \frac{Na^+}{\sqrt{\left(\frac{Ca^{2+} + Mg^{2+}}{2}\right)}} \quad (1)$$

$$ESP = \frac{Na^+}{CEC} \quad (2)$$

در رابطه ۱، Ca^{2+} ، Na^+ و Mg^{2+} به ترتیب یون‌های کلسیم، سدیم و منیزیم محلول (میلی‌اکی‌والان بر

1- Sodium absorption ratio (SAR)

2- Exchangeable sodium percentage (ESP)

برآورد پارامترهای مدل، روش حداقل مربعات معمولی^۳ به عنوان پرکاربردترین و غالبترین روش شناخته شده است. این روش که به کارل فردریک گوس، ریاضیدان نامی آلمان نسبت داده می شود از طریق حداقل کردن مجموع مربعات جملات تلاش می کند که بهترین خط رگرسیونی برای داده ها را برازش نماید (۱۳ و ۲۰). شبکه بیزین نیز یک مدل گرافیکی برای نمایش احتمالات بین متغیرهای مورد نظر بوده که اجازه محاسبات احتمالی را می دهد. از طرفی این مدل به استدلال در شرایط نامشخص (عدم قطعیت) کمک شایانی می نماید (۸). با توجه به اهمیت تخمین فوق و کارایی روش های انتخابی انتظار می رود نتایج مناسبی به دست آید، بنابراین در این مطالعه رابطه بین نسبت جذب سدیم با مقدار درصد سدیم تبادلی با استفاده از شبکه بیزین و کمترین مربعات معمولی مدل سازی و کاراترین مدل برای تخمین درصد سدیم تبادلی از روی نسبت جذب سدیمی مشخص شد. کاربرد شبکه بیزین و کمترین مربعات معمولی در برآورد درصد سدیم تبادلی خاک، از نوآوری های پژوهش فوق محسوب می شود.

مواد و روش ها

منطقه مورد مطالعه قره چبق (شهرستان بناب) با وسعت ۱۴۵۰ هکتار در جنوب غرب استان آذربایجان شرقی و ساحل جنوب شرق دریاچه ارومیه واقع شده است. از نظر جغرافیایی منطقه بین ۴۵ درجه، ۵۸ دقیقه، ۳۰ ثانیه تا ۴۶ درجه، ۵۸ دقیقه و ۴۰ ثانیه طول شرقی و ۳۷ درجه، ۱۷ دقیقه، ۵ ثانیه تا ۳۷ درجه، ۱۹ دقیقه و ۱۵ ثانیه عرض شمالی قرار دارد (شکل ۱). کاربری منطقه مرتع و زراعت بوده، رژیم رطوبتی خاک های منطقه اریدیک هم مرز با زیریک و رژیم حرارتی آن مزیک است.

لیتر)، در رابطه ۲ یون های تبادلی اندازه گیری شده (سانتی مول بر کیلوگرم) و CEC ظرفیت تبادل کاتیونی^۱ (سانتی مول بر کیلوگرم) می باشد.

برای اندازه گیری ESP، اندازه گیری CEC ضروری است که اندازه گیری آن وقت گیر، پرهزینه و همراه با خطا است (۱۵)، بنابراین تخمین غیرمستقیم ESP از ویژگی های زودیافت خاک بسیار بهینه و اقتصادی است. یکی از این پارامترها SAR بوده که همبستگی آن با ESP برای خاک های مختلف جهان گزارش شده است، ولی این روابط با تغییر درصد و نوع رس غالب در خاک، کانی شناسی، شوری محلول تعادلی و درصد رطوبت اشباع، تغییر می یابند (۴). آزمایشگاه شوری ایالات متحده^۲ (USSL) برای تخمین ESP از روی SAR رابطه ای را برای خاک های ایالات متحده آمریکا پیشنهاد کرد (۱۶). پژوهش ها نشان داد که این رابطه ثابت نبوده و با قدرت یونی محلول و تغییر کانی غالب تغییرات اساسی می کند (۶).

برای تخمین ESP از SAR روش های دیگری نیز از جمله رگرسیون خطی چندمتغیره و هوش مصنوعی وجود دارد (۱۲). پژوهش های متعددی در این راستا توسط زارع و همکاران (۲۳) در استان فارس، لک و همکاران (۱۱) در استان گیلان و سیلس پور و همکاران (۱۸) در استان تهران صورت گرفته و روابط متعددی با توجه به شرایط خاک گزارش شده است. فرهمند و همکاران (۲۰۱۱) نیز رابطه لگاریتمی با دقت بالا بین درصد سدیم تبادلی و نسبت جذب سدیم با ۲۹ نمونه از خاک های متأثر از نمک دشت تبریز توسعه دادند (۷). ژائو و همکاران (۲۰۱۱) نیز از هوش مصنوعی برای تخمین ESP از SAR استفاده نمودند و گزارش کردند که هوش مصنوعی با دقت و صحت بیشتری این تخمین را انجام می دهد (۲۵). با این حال از میان روش های مختلف خطی برای

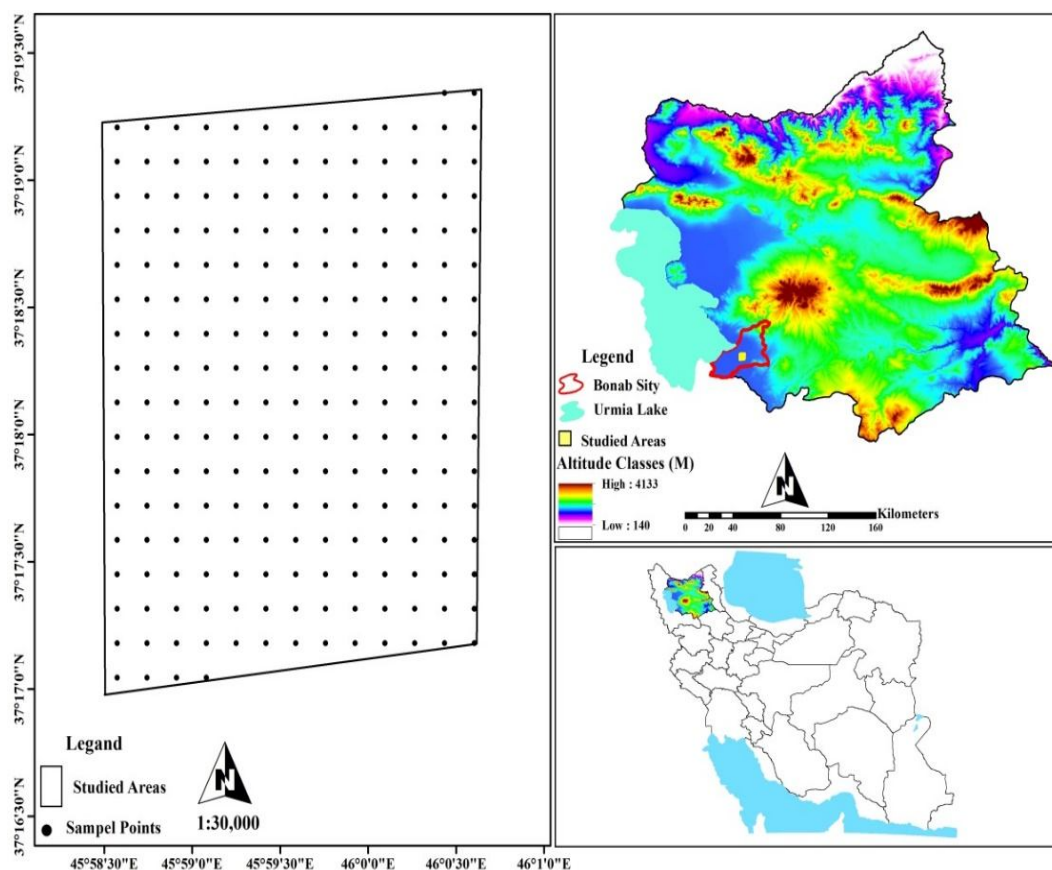
1- Cation Exchange Capacity

2- United states salinity laboratory

3- Ordinary least Squares (OLS)

الکتریکی نمونه‌ها در نسبت ۵ به ۱ آب به خاک (۱۷). بافت خاک به روش هیدرومتری (۲)، Na^+ به روش فیلم‌فوتومتری و Ca^{2+} و Mg^{2+} به روش تیتراسیون برگشتی (۳ و ۲۱) اندازه‌گیری شدند.

جهت رسیدن به اهداف پژوهش، نمونه‌های خاک از محل هر گره از شبکه‌بندی منظم با فواصل ۲۵۰ متر (۲۰۹ نمونه) از خاک سطحی (۰-۲۵ سانتی‌متری) در تابستان ۹۶ برداشته شد و جهت آنالیزهای خاک به آزمایشگاه منتقل گردید. سپس pH و هدایت

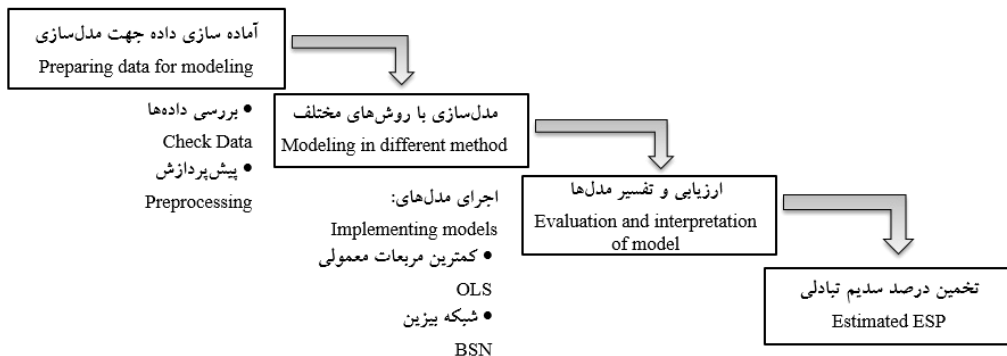


شکل ۱- موقعیت اراضی مورد مطالعه و نقاط نمونه‌برداری.

Figure 1. Location of studied area and sampling points.

بیزین انجام و سپس این مدل‌ها بر اساس آماره‌های آماری مورد ارزیابی قرار گرفته و مناسب‌ترین مدل برای تخمین ESP انتخاب شد. در این پژوهش از نرم‌افزار MATLAB R2016b برای مدل‌سازی استفاده شد.

به‌منظور ارائه مدل مناسب داده‌ها به سه دسته تقسیم شدند، ۷۰ درصد سری داده‌ها به آموزش، ۱۵ درصد به اعتبارسنجی و ۱۵ درصد به آزمون شبکه اختصاص یافت. شکل ۲ روند اجرای مطالعه حاضر را نشان می‌دهد که داده‌ها بعد از آماده‌سازی، با استفاده از مدل‌های حداقل مربعات معمولی و شبکه



شکل ۲- روند اجرای پژوهش.

Figure 2. The process of conducting research.

$$y = B_0(u_i + v_i) + \sum_{i=1}^p B_i(u_i + v_i)X_i + E \quad (۱)$$

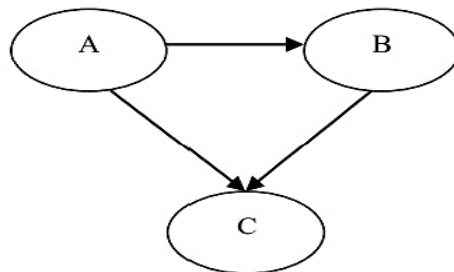
که در آن، u_i و v_i مختصات هر موقعیت برای i ، $\beta_0(u_i, v_i)$ محل تقاطع برای موقعیت i ، $\beta_i(u_i, v_i)$ یک ضریب محلی است که متغیر مستقل X_i را در موقعیت i تخمین می زند و E نیز خطای تصادفی با فرض نرمال بودن داده ها می باشد (۲۰).

شبکه بیزین: شبکه های بیزین^۲ مدل های گرافیکی هستند که برای استدلال در مواردی که پیچیدگی و عدم قطعیت وجود دارد به کار می روند (۱۰). در این گراف گره ها نشانگر متغیرهای تصادفی گسسته یا پیوسته هستند و کمان های جهت داری که هر جفت از گره ها را به یکدیگر متصل می کند (شکل ۳) وابستگی بین متغیرها را نمایش می دهد (۸).

روش حداقل مربعات معمولی (OLS): مدل OLS مناسب ترین تخمین گر خطی ناریب^۱ (BLUE) است. در واقع یک الگوی جهانی است که با فرض ارتباطات ثابت در فضای نمونه گیری استوار است و متغیرها را با استفاده از متغیرهای موجود از مناطق نمونه گیری شده تخمین می زند (رابطه ۳).

$$y = B_0 + \sum_{i=1}^p B_i X_i + E \quad (۳)$$

که در آن، y متغیر وابسته، B_0 عرض از مبدا، B_i ضرایب تخمینی برای متغیر مستقل x_i ، p تعداد متغیرهای مستقل و E جزء خطاست. برای این که در ضرایب محلی تخمین در تمامی موقعیت های نمونه انجام شود، رابطه ۳ به صورت رابطه ۴ قابل بسط است.



شکل ۳- گراف شبکه بیزین.

Figure 3. Bayesian Network Diagram.

1- Best Linear Unbiased Estimator

2- Bayesian Network

که در آن‌ها، q_i نسبت سدیم تبادلی اندازه‌گیری شده، \hat{q}_i درصد سدیم تبادلی تخمینی، \bar{q}_i میانگین نسبت سدیم تبادلی اندازه‌گیری شده، $\bar{\hat{q}}_i$ میانگین نسبت سدیم تبادلی تخمینی و n تعداد داده‌ها می‌باشد. هرچه مقدار RMSE کوچک‌تر و مقدار R^2 به ۱ نزدیک‌تر باشد، نمایانگر دقت بالاتر مدل توسعه‌یافته خواهد بود. GMER نیز بیان‌کننده وجود هماهنگی و تطابق بین مقادیر مشاهده‌ای و تخمینی می‌باشد. چنانچه مقدار ضریب GMER برابر با ۱ باشد، بین مقادیر اندازه‌گیری شده و تخمینی مطابقت کامل وجود دارد. بنابراین هرچه مقدار این آماره به یک نزدیک‌تر باشد مدل از دقت بالاتری برخوردار است. از طرفی، بیشتر و کم‌تر بودن این پارامتر نسبت به یک، به ترتیب نشان‌دهنده کم‌برآوردی و بیش‌برآوردی مدل می‌باشد (۲۲). برای تعیین وجود و عدم وجود اختلاف معنی‌دار بین داده‌های تخمینی و اندازه‌گیری شده از t-test استفاده شد.

نتایج و بحث

ضریب تغییرات معیاری از تغییرپذیری نسبی است. داهیا و همکاران (۱۹۸۴) ضریب تغییرات ویژگی‌های خاک را به دو دسته پایین (۰ تا ۱۵ درصد) و بالا (۱۵ تا ۷۵ درصد) تقسیم‌بندی نمودند (۵). در این مطالعه مقادیر ضریب تغییرات برای ویژگی‌های خاک در محدوده بین ۰/۰۲ درصد برای pH (کم‌ترین) و ۱۶/۵۹ درصد برای درصد سدیم تبادلی (بیش‌ترین) قرار داشت (جدول ۱). محتوای درصد سدیم تبادلی در خاک‌های مورد مطالعه با میانگین ۲۷/۳۱ بالا بوده و این مسأله نشان‌دهنده ضرورت پایش و بررسی تغییرات درصد سدیم تبادلی، به‌عنوان شاخص سدیمی بودن خاک، در منطقه مورد مطالعه می‌باشد. حداقل درصد سدیم قابل تبادل ۱/۲ و حداکثر آن ۸۴/۷ درصد می‌باشد که می‌تواند به‌دلیل وابستگی به فرآیندهای ذاتی و مدیریتی خاک باشد.

در مورد ساختار شبکه بیزین اگر مجموعه‌ای شامل n متغیر به صورت $\{X_1, X_2, \dots, X_n\}$ باشد، هر X_i یک متغیر تصادفی بوده و گره‌های والد این متغیر به صورت $\text{parent}(x_i)$ باشد، احتمال هر رویداد مانند $X = (X_1, X_2, \dots, X_n)$ به صورت زیر محاسبه می‌گردد. سپس مقادیر بین گره‌های متصل به هم به دست می‌آید (رابطه ۵). هر گره یک جدول احتمال شرطی دارد که برای هر ترکیب از مقادیر والد هایش، مقدار مشخصی دارد. نهایتاً توزیع‌های احتمال بعدی برای هر گره محاسبه می‌شود که این توزیع احتمال امکان^۱ نامیده می‌شود (۱۴).

$$(X_1, \dots, X_n) = \prod_{i=1}^n P(x_i | \text{Parent}(X_i)) \quad (5)$$

مقایسه عملکرد مدل‌ها: برای مقایسه کمی عملکرد مدل‌ها می‌توان از تجزیه و تحلیل خطاهای باقی‌مانده و اختلاف بین مقادیر اندازه‌گیری و برآورد شده، بهره گرفت. در این پژوهش، از معیارهای ضریب تبیین (R^2) (رابطه ۶)، ریشه میانگین مربعات خطا^۲ (RMSE) (رابطه ۷) و میانگین هندسی نسبت خطا^۳ (GMER) (رابطه ۸) استفاده شد.

$$R^2 = \frac{[\sum_{i=1}^n (q_i - \bar{q}_i)(\hat{q}_i - \bar{\hat{q}}_i)]^2}{\sum_{i=1}^n (q_i - \bar{q}_i) \sum_{i=1}^n (\hat{q}_i - \bar{\hat{q}}_i)^2} \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_i - \hat{q}_i)^2} \quad (7)$$

$$GMRE = \sqrt{\frac{1}{n} \sum_{i=1}^n -Ln(q_i / \hat{q}_i)} \quad (8)$$

- 1- Probability of possibility
- 2- Root mean square error
- 3- Geometric Mean Error Ratio

جدول ۱- آماره‌های توصیفی ویژگی‌های خاک.

Table 1. Descriptive statistics of soil properties.

pH	Sand	Silt	Clay	EC	SAR (%)	ESP (%)	
8.5	28	32.7	39.4	25.5	22.52	27.31	میانگین Average
8.9	61	49	63	81.6	70.3	84.7	حداکثر Maximum
7.2	2	15	16	1.4	0.9	1.2	حداقل Minimum
0.02	4.7	1.4	2	14.7	14.04	16.59	ضریب تغییرات Coefficient of Variation

اطلاعات ورودی در لایه اول بستگی دارد، بنابراین با ارائه ماتریس همبستگی ارتباط بین متغیر مستقل و متغیر وابسته مشخص شد.

جدول ۲ ارتباط بین متغیرهای درصد سدیم تبدلی و نسبت جذب سدیم را بر اساس ضریب همبستگی پیرسون نشان می‌دهد. از آنجایی که توانمندی پیش‌بینی مدل‌های هوشمند در انتخاب صحیح نوع

جدول ۲- ضرایب همبستگی پیرسون بین ESP و SAR.

Table 2. Pearson Correlation Coefficients between ESP and SAR.

SAR	ESP	
	1	ESP
1	0.961**	SAR

** همبستگی در سطح احتمال ۰/۰۱

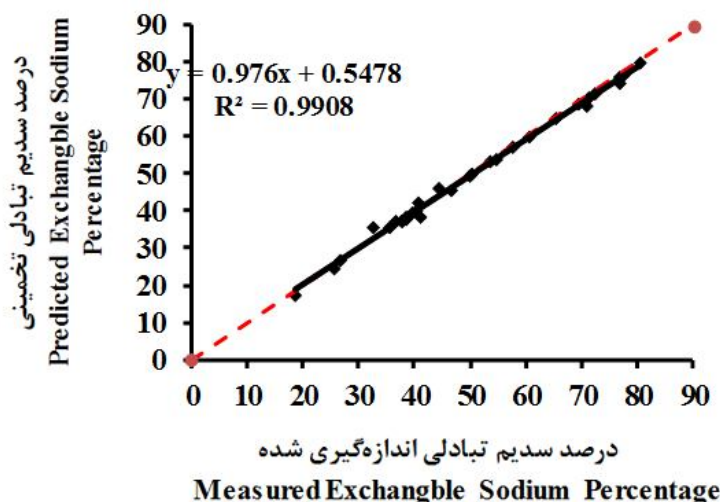
RMSE=۰/۷۱ و میانگین هندسی نسبت خطا GMRE=۱/۱ بوده و مدل از دقت بالایی برخوردار است. نمودار پراکنش کم‌ترین مربعات در مرحله صحت‌سنجی نیز نشان‌دهنده برازش مناسب مقادیر تخمینی و اندازه‌گیری شده است، به طوری که مقادیر تخمینی و اندازه‌گیری شده به غیر از چند نقطه که روی خط نیمساز قرار نگرفته در سایر نقاط بر روی خط برازش قرار دارد (شکل ۴).

تخمین درصد سدیم تبدلی با استفاده از روش حداقل مربعات معمولی: تخمین‌گر OLS دارای کم‌ترین واریانس در بین تخمین‌گرهای خطی نارایب است. مدلی که کم‌ترین واریانس را در بین سایر تخمین‌گرهای نارایب خطی داشته باشد، تخمین‌گر کاراست. نتایج تخمین ESP با مدل کم‌ترین مربعات معمولی (جدول ۳)، بیانگر این مطلب است که ضریب تبیین $R^2=۹۹$ ، ریشه میانگین مربعات خطا

جدول ۳- تحلیل نتایج مدل OLS برای تخمین ESP.

Table 3. Analysis of OLS model results for ESP estimation.

صحت‌سنجی			آموزش			ساختار
GMRE	RMSE	R ²	GMRE	RMSE	R ²	
1.1	0.71	0.992	1.3	1.28	0.981	OLS



شکل ۴- رابطه بین درصد سدیم تبادلی تخمینی و اندازه‌گیری شده با مدل حداقل مربعات معمولی.

Figure 4. Correlation of predicted and measured ESP by ordinary least squares model.

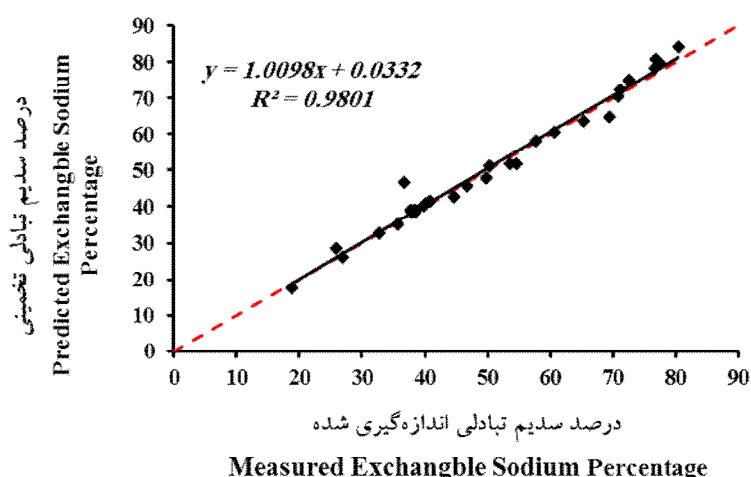
قابل قبول (۱/۶۳) در مرحله صحت‌سنجی درصد سدیم تبادلی را از روی نسبت جذب سدیمی تخمین بزند (جدول ۴). شکل ۵ نیز نمودار پراکنش شبکه بیزین در مرحله صحت‌سنجی را نشان می‌دهد. خط برازش مقادیر تخمینی و اندازه‌گیری شده به‌غیر از چند نقطه که روی خط نیمساز قرار نگرفته، در سایر نقاط بر روی خط برازش قرار داشته که این امر دلالت بر دقت بالای مدل دارد.

تخمین درصد سدیم تبادلی با استفاده از شبکه بیزین: در این پژوهش، الگوریتم تخمینی K₂ و الگوریتم یادگیری Thinning Thick Greedy با روش سعی و خطا برای مدل‌سازی SAR-ESP با شبکه‌های بیزین انتخاب شد. نتایج بیانگر این مطلب است که SAR تأثیر مستقیم و بالا بر روی درصد سدیم تبادلی دارد. در واقع مدل انتخابی توانسته با دقت بالایی (۰/۹۸) و ریشه میانگین مربعات خطا

جدول ۴- نتایج شبکه بیزین برای تخمین ESP.

Table 4. Results of Bayesian network of ESP prediction.

صحت‌سنجی			آموزش			ساختار
GMRE	RMSE	R ²	GMRE	RMSE	R ²	
1.3	1.63	98	0.98	2.32	91	Bayesian network



شکل ۵- مقادیر درصد سدیم تبادلی تخمینی و اندازه‌گیری شده با استفاده از مدل شبکه بیزین.
Figure 5. Correlation of predicted and measured ESP by Bayesian network model.

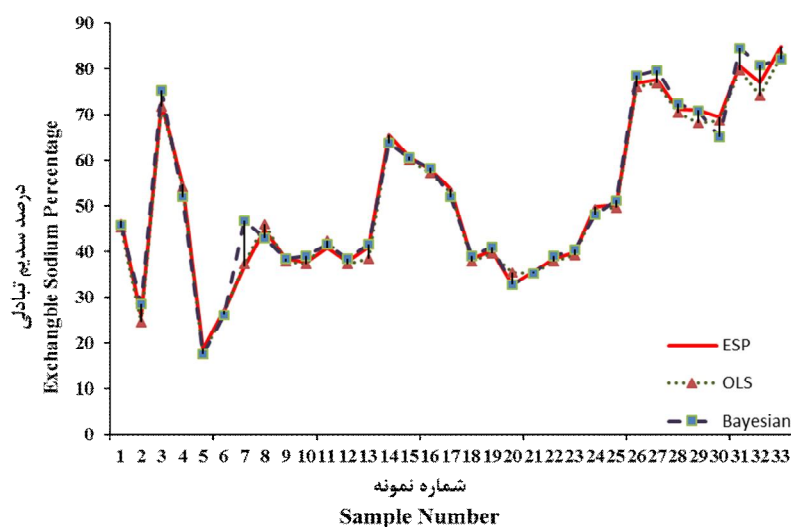
مدل‌هاست. مدل کم‌ترین مربعات خطا، بیش‌ترین ضریب تبیین $R^2 = 0/99$ و کم‌ترین ریشه میانگین مربعات خط $RMSE = 0/71$ و میانگین هندسی نسبت خطا $GMRE = 1/1$ در مرحله صحت‌سنجی دارد که به‌عنوان کاراترین مدل معرفی می‌گردد. لک و همکاران (۲۰۰۹) به کمک روش‌های رگرسیون چندمتغیره و شبکه عصبی مصنوعی، ESP را با استفاده از SAR در سواحل استان گیلان، شبیه‌سازی نمودند (۱۲). در مطالعه ایشان مقادیر ضریب تبیین و ریشه میانگین مربعات خطا را با روش رگرسیون خطی به‌ترتیب برابر با $0/66$ و $1/13$ محاسبه نمودند. همچنین با به‌کارگیری شبکه عصبی MLP به‌ترتیب مقادیر ضریب تبیین و ریشه میانگین مربعات خطای $0/94$ و $0/18$ به‌دست آمد. نتایج بیانگر برتری شبکه‌های عصبی مصنوعی نسبت به روش رگرسیون خطی در پیش‌بینی ESP می‌باشد.

مقایسه دقت مدل‌ها: نتایج آزمون t نشان داد که در سطح احتمال ۵ درصد تفاوت معنی‌داری بین مقادیر تخمینی درصد سدیم تبادلی با دو مدل انتخابی و مقادیر اندازه‌گیری‌شده وجود ندارد. به‌منظور ارزیابی و مقایسه عملکرد مدل‌ها در تخمین درصد سدیم تبادلی، مقادیر درصد سدیم تبادلی برای ۱۵ درصد از داده‌ها که به‌صورت تصادفی قبل از انجام تجزیه و تحلیل‌ها کنار گذاشته شده بودند، به‌وسیله روش‌های کم‌ترین مربعات معمولی و شبکه بیزین، تخمین و ضرایب تبیین و ریشه میانگین مربعات خطا بین مقادیر تخمینی و اندازه‌گیری‌شده به‌دست آمد (جدول ۵). با مقایسه مدل‌ها از طریق آماره‌های ذکرشده مشخص شد که هر دو مدل با دقت خوبی و خطای کم‌تری می‌توانند درصد سدیم تبادلی را شبیه‌سازی کنند. شکل ۶ نتایج دو مدل را نسبت به مقادیر مشاهداتی نشان می‌دهد که انطباق مقادیر تخمینی با مقادیر اندازه‌گیری‌شده نیز بیانگر صحت بالای

جدول ۵- مقادیر ضریب تبیین و ریشه میانگین مربعات خطا برای تخمین ESP از SAR.

Table 5. Correlation coefficients and root mean square errors for predicted ESP from SAR.

GMRE	صحت‌سنجی accuracy		R ²	ساختار
	RMSE			Structure
1.1	0.71		0.99	OLS
3.1	63.1		0.98	Bayesian Network



شکل ۶- تغییرات درصد سدیم تبادلی تخمینی و اندازه‌گیری شده با مدل‌های مختلف.

Figure 6. Variations of predicted and measured ESP by different models.

سدیم برای خاک‌های متأثر از نمک دشت تبریز ارائه دادند (۷). زارع و همکاران (۲۰۱۴) نیز رابطه $ESP = 0.941 + 1.11 \times SAR$ را برای خاک‌های شور مرودشت استان فارس پیشنهاد کردند (۲۳).

نتیجه‌گیری کلی

در پژوهش حاضر امکان پیش‌بینی درصد سدیم تبادلی با استفاده از نسبت جذب سدیم با روش‌های کم‌ترین مربعات معمولی و شبکه بیزین مورد بررسی قرار گرفت. نتایج این پژوهش نشان داد مدل‌های هوش مصنوعی نیز به خوبی قادر به تخمین درصد سدیم تبادلی هستند. همچنین هر دو روش به خوبی

چی و همکاران (۲۰۱۱) با مطالعه ۱۱۷ نمونه خاک بخش‌های غربی دشت سنگان چین، مقدار درصد سدیم تبادلی را با استفاده از SAR برآورد نمودند (۴). ایشان در پژوهش‌های خود رابطه بین درصد سدیم تبادلی و SAR را نسبتاً خوب با ضریب تبیین ۰/۷۶ گزارش نمودند. گزارش کردند. سیلس‌پور و همکاران (۲۰۰۹) ضمن تخمین درصد سدیم تبادلی به کمک نسبت جذب سدیم در خاک‌های خشک منطقه ورامین، رابطه رگرسیونی خطی $ESP = 1.95 + 1.03 \times SAR$ را برای این خاک‌ها ارائه دادند (۱۸). فرهمند و همکاران (۲۰۱۱) نیز رابطه لگاریتمی بین درصد سدیم تبادلی و نسبت جذب

مربعات معمولی در مقایسه با شبکه بیزین از دقت و صحت بالاتری برخوردار بود و نتایج مطلوب‌تری ارائه داد. نهایتاً می‌توان چنین عنوان کرد که مدل‌های هوش محاسباتی می‌توانند درصد سدیم تبادلی را با قدرت اطمینان بالایی تخمین بزنند و نیاز به صرف وقت و هزینه‌های زیاد در برآورد مستقیم درصد سدیم تبادلی را مرتفع سازند.

توانست تغییرات ESP را با استفاده از SAR توجیه نماید که به ترتیب ضریب همبستگی و جذر میانگین مربعات خطا ۰/۷۱؛ ۰/۹۸ و ۱/۶۳ برای مدل‌های حداقل مربعات معمولی و شبکه‌های بیزین محاسبه شد. همچنین خط برازش مقادیر تخمینی و اندازه‌گیری‌شده با مدل‌های انتخابی تقریباً بر روی خط برازش قرار داشتند که دلالت بر دقت بالای مدل‌ها دارد. بنابراین در مقام مقایسه، روش کم‌ترین

منابع

1. Banaei, M.H., Momeni, A., Baybordi, M., and Malakouti, J. 2004. Iranian Soils. Sana Press, Tehran, Iran.
2. Bouyoucos, G.J. 1962. Hydrometer method improved for making particle size analysis of soils. *Agron. J.* 56: 464-466.
3. Bower, CA., Reitemeier, RF., and Fireman, M. 1952. Exchangeable cation analysis of saline and alkali soils. *Soil Science.* 73: 251-261.
4. Chi, Ch.M., Zhao, Ch.W., Sun, X.J., and Wang, Z.C. 2011. Estimating exchangeable sodium percentage from sodium adsorption ratio of salt-affected soil in The Songnen plain of Northeast China. *Soil Science Society of China Pedosphere* 21: 2. 271-276.
5. Dahiya, I.S., Richter, J., and Malik, R.S. 1984. Soil spatial variability: A review. *Inter. J. Trop. Agric.* 11: 91-102.
6. Evangelou, V.P., and Marsi, M. 2003. Influence of ionic strength on sodium-calcium exchange of two temperate climate soils. *Plant and Soil.* 250: 307-313.
7. Farahmand, A., Oustan, S.H., Jafarzadeh, A.J., and Asgarzad, A.N. 2011. The parameters of sodium and salinity in some salt affected soils of the Tabriz Plain. *J of Soil and Water*, 19: 2. 22. 1-15. (In Persian)
8. Heckerman, D. 1997. Bayesian networks for data mining. *Data Mining and Knowledge Discovery.* 1: 1. 79-119.
9. Jurinak, J.J., and Suarez, D.L. 1990. The chemistry of salt-affected soils, P 42-63. In: Tanji, K.K. (ed). *Agricultural Salinity Assessment and Management*, No, 71. American Society of Civil Engineers, New York, N.Y.
10. Kevin, B., and Nicholson, E. 2010. Bayesian artificial intelligence. Second Edition, United states. 3: 1. 370-450.
11. Lake, H.R., Akbarzadeh, A., and Mehrjardi, R.T. 2009. Development of pedotransfer functions (PTFs) to predict soil physico-chemical and hydrological characteristics in southern coastal zones of the Caspian Sea. *J. Ecol. Natur. Environ.* 1: 7. 160-172.
12. Lal, P., Chippa, B.R., and Arvind, K. 2003. Salt affected soils and crop production, a modern synthesis, AGROBIS (India). Pp: 42-61.
13. Lesch, S.M., Strauss, D.J., and Rhoades, J.D. 1995. Spatial prediction of soil salinity using electromagnetic induction techniques 1. Statistical prediction models: A comparison of multiple linear regression and cokriging. *Water Resources Research*, 31: 373-386.
14. Nguyen, R.T., Prentiss, D., and Shively, J.E. 1998. Rainfall interpolation for Santa Barbara County. UCSB, Department Geography. USA.
15. Rhoades, J.D. 1982. Cation exchange capacity. P 149-157. In: Page, A.L., Miller, R.H. and Keeney, D.R. (eds). *Methods of Soil Analysis. Part 2. Agron. Monogr. 9*, American Society of Agronomy, Madison, WI, USA.
16. Richards, L.A. 1954. *USDA Handbook 60*. U.S. Department of Agriculture, Washington DC. USA

17. Rowell, D.L. 1994. Soil Science: Methods and Application. Longman Group, Harlow, England, 345p.
18. Seilsepour, M., Rashidi, M., and Khabbaz, B.G. 2009. Prediction of soil exchangeable sodium percentage based on soil sodium adsorption ratio. *Amer.-Euras. J. Agric. Environ. Sci.* 5: 1. 1-4.
19. Sumner, M.E. 1993. Sodic soils: New perspectives. *Austr. J. Soil Res.* 31: 683-750.
20. Tamari, S., WoËsten, J.H.M., and Ruiz-SuaËrez, J.C. 1996. Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Sci. Soc. Amer. J.* 60: 1732-1741.
21. Tu, J. 2011. Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. *Applied Geography*, 31: 1. 376-392.
22. USDA. 1996. Soil Survey Laboratory Methods Manual. Soil Survey Investigations Republic, Washington: United States Government Print.
23. Wagner, B., Tarnawski, V.R., Hennings, V., Muller, U., Wessolek, G., and Plagge, R. 2001. Evaluation of pedotransfer function for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma*. 102: 275-297.
24. Zare, M., Ordoorkhani, K., Emadi, A., and Azarpanah, A. 2014. Relationship between soil exchangeable sodium percentage and soil sodium adsorption ratio in Marvdasht plain, Iran. *Inter. J. Adv. Biol. Biom. Res.* 2: 12. 2934-2939.



Gorgan University of Agricultural
Sciences and Natural Resources

J. of Water and Soil Conservation, Vol. 26(2), 2019

<http://jwsc.gau.ac.ir>

DOI: 10.22069/jwsc.2019.15927.3113

Performance of Ordinary least Squares (OLS) and Bayesian network (BN) in Exchange sodium percentage prediction based on Sodium adsorption Ratio

A. Barikloo¹, *M. Servati² and M.S. Oliaei³

¹M.Sc. Graduate, Dept. of Soil Science and Engineering, University Zanjan,

²Assistant Prof., Shahid Bakeri High Education Center of Miandoab, Urmia University,

³Assistant Prof., Adjutancy of Research and Technology, Ministry of Science, Research and Technology

Received: 11.22.2018; Accepted: 02.10.2019

Abstract

Background and Objectives: In the soil science, exchangeable sodium percentage and sodium adsorption ratio are two different criteria to evaluate of soil alkalinity. For measured of ESP, it is essential to have soil Cation Exchange Capacity (CEC). But, CEC determined by using laborious method is expensive and time consuming. Developing a model that predicts ESP indirectly from a easily-measured properties to be more appropriate and economical. Researches showed a relationship between ESP and SAR. So, SAR can be allocated to predict of ESP. For this reason, many attempts have been made to predict ESP from soil. The specific goal of the research develop model to determining ESP based on SAR by OLS and BN models for Bonab soils in East Azarbaijan province, Iran.

Materials and Methods: For arrived presented research, 209 soil samples were taken by grid survey (250×250) of Bonab, Iran. The site is located at mean 1300 m above mean sea level, in semiarid climate in the Northwest of Iran. Then, some soil chemical properties such as Sodium, calcium, magnesium, SAR and ESP of the soil samples were measured using laboratory experiments. Then, two model was developed by OLS and BN. OLS estimators are linear functions of the values of the dependent variable which are linearly combined using weights that are a non-linear function of the values of the explanatory variables. So the OLS estimator is respect to how it uses the values of the dependent variable only and irrespective of how it uses the values of the explanatory. So A Bayesian network is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph (DAG). Given symptoms, the network can be used to compute the probabilities of the presence of various diseases. Efficient algorithms can perform inference and learning in Bayesian networks. Generalizations of Bayesian networks that can represent and solve decision problems under uncertainty.

Results: The Coefficient of Determination (R^2) and Root Mean Square error (RMSE) of the soil ESP-SAR model is reported 0.99, 0.71 and 0.98, 1.63 by OLS and BN respectively. Based on the statistical result, both of soil ESP-SAR model was judged acceptable. T-test were used to compare the soil ESP values predicted using the soil ESP-SAR model with the soil ESP values measured by laboratory tests. The paired samples t-test results indicated that the difference between the soil ESP values predicted by the model and measured by laboratory tests were not

* Corresponding Author; Email: m.sarvati@urmia.ac.ir

statistically significant ($P > 0.05$). Therefore, the soil ESP-SAR model can provide an easy, economic and brief methodology to estimate soil ESP. The GMER index also indicated low estimation of two selected land evaluation method.

Conclusion: The results of present study illustrated that OLS and BN models can predict ESP with acceptable limits. OLS and BN are mathematical models between input and output variables and have the ability of modeling between ESP and SAR.

Keywords: Bonab, Easily-Measured properties, Modeling, Sodic soil