

Research Article

Agricultural Engineering., 45(3) (2022) 299-315  
DOI: 10.22055/AGEN.2023.41754.1644

ISSN (E): 2588-526X

ISSN (P): 2588-5944

## Investigation of the most Important Factors in Prediction the Soil Exchangeable Sodium Percentage by Neural-Fuzzy Constant Neural Model

A. Barikloo<sup>1</sup>, M. Servati<sup>2</sup> and P. Alamdari<sup>3\*</sup>

1. PhD Student of the Department of Soil Science and Engineering, University of Zanjan, Iran
2. Assistant Professor of Shahid Bakri Miandoab Higher Education Center, Urmia University, Iran
3. Assistant Professor, Department of Soil Science and Engineering, University of Zanjan, Iran

Received: 29 August 2022

Accepted: 17 December 2022

### Abstract

**Introduction:** A variety of precise farming practices in arid and semi-arid regions such as Iran require periodic information on soil salinity and sodium content. Sodic soils have unfavorable physical and chemical properties due to the high percentage of exchangeable sodium (ESP). This problem reduces the capacity of available water and growth of plants. To measure the percentage of exchangeable sodium, it is necessary to measure the amount of cation exchange capacity (CEC). The measurement of CEC is time consuming, therefore it is appropriate and economical to develop a method for indirectly determination of ESP from easy-measured properties. One of the methods to study the relationships and correlations between different soil properties and their quantitative expression is the use of some statistical models. These models, called transfer functions, including data mining, regression models, artificial neural networks, and the coherent neural-fuzzy integrated system (CANFIS).

**Materials and Methods:** The aim of this study was to develop an intelligent model (CANFIS) for predicting soil ESP from soil easy-measured properties in approximately 1450 ha of salt affected soils, South East of Urmia Lake, Bonab region, East Azarbaijan Province. For this purpose, 209 soil samples were taken by grid survey method from surface (0-25 cm) and then carried out laboratory for measure necessary soil properties. Soil acidity and electrical conductivity of samples were measured in a ratio of 1: 5 soils to water, soil texture by hydrometric method, sodium by flame photometer and calcium and magnesium by returned titration method.

**Results and Discussion:** Pearson correlation method showed that the accuracy of estimating intelligent models depends on the correct choice of first layer input information. Therefore, using the correlation matrix, the relationship between soil parameters (independent variable) and the percentage of exchangeable sodium (dependent variable) was determined. Sodium absorption ratio (0.961), electrical conductivity (0.808), pH (0.638), clay content (0.524), sand (0.482) and silt (0.389) have the greatest effect on estimation Percentage of exchangeable sodium in soil. Also, the positive relationship between soil reaction and the percentage of exchangeable sodium on the one hand and the high correlation between the percentage of exchangeable sodium and electrical conductivity indicate the importance of the fine soil. In this study, the relationship of linear regression between the percentage of exchangeable sodium and the ratio of sodium uptake with an explanation coefficient of 0.91 was calculated, which is significant at the level of 5% probability. Two important targets were



designed in this paper. First target is determining performance of Fuzzy Neural Networks (CANFIS) in predicting ESP by sand, clay, pH, SAR, EC as input variable. The second target is evaluation of performance of CANFIS model by selected variable of PCA model. Results showed that the performance of second model was acceptable Model 1 justifies 88% of the changes in the percentage of exchangeable sodium by entering all inputs. But CANFIS model with higher inputs selected by PCA model (principal component analysis) including sodium adsorption ratio, electrical conductivity, soil reaction has higher accuracy. So that the values of root mean square error and correlation coefficient in the test stage for the first model were 0.88 and 3.25 and the second model was 0.96 and 1.0, respectively .

**Conclusion:** These results demonstrated the superiority of intelligent models in explanation of the relationship between ESP and other soil easily-measured properties. In order to model the soil retrieval properties such as cation exchange capacity and to achieve the most suitable model, it is necessary to pay attention to the number and most effective input variables. Because the main goal is to provide a model with a minimum number of inputs as well as inputs that are easy to measure and in a short time. The results of quantification of the importance of variables in the CANFIS model confirm the use of three characteristics of sodium adsorption ratio, electrical conductivity of soil saturated extract and soil acidity in modeling the percentage of exchangeable sodium. The results of this study can be generalized to soils of similar arid and semi-arid regions. Also, due to the ambiguity of soil-related phenomena or the approximate values of the measured values of different soil properties and the uncertainty in the data, the use of hybrid models such as CANFIS that use fuzzy sets, It can be useful in fitting soil transfer functions.

**Keywords:** *Easily-measured properties, hard-measured properties, PCA all right reserved*

## تأثیر پیش پردازش متغیرهای ورودی مدل عصبی - فازی همراستا با استفاده از روش تجزیه به مؤلفه اصلی جهت برآورد درصد سدیم تبادلی خاک

علی باریکلو<sup>۱</sup>، مسلم ثروتی<sup>۲</sup> و پریسا علمداری<sup>۳\*</sup>

۱- دانشجوی دکتری گروه علوم و مهندسی خاک دانشگاه زنجان، ایران

۲- استادیار مرکز آموزش عالی شهید باکری میاندوآب، دانشگاه ارومیه، ایران

۳- استادیار گروه علوم و مهندسی خاک دانشگاه زنجان، ایران

### تاریخچه مقاله

دریافت: ۱۴۰۱/۰۶/۰۷

پذیرش نهایی: ۱۴۰۱/۰۹/۲۶

### چکیده

خاک‌های سدیمی به دلیل درصد بالای سدیم تبادلی ویژگی‌های فیزیکی و شیمیایی نامطلوبی دارند. این مشکل منجر به کاهش ظرفیت آب قابل استفاده و کاهش رشد گیاهان می‌شود. برای اندازه‌گیری درصد سدیم تبادلی نیاز به اطلاع از مقدار ظرفیت تبادل کاتیونی است. اما اندازه‌گیری آن پرهزینه بوده و زمان بر نیز می‌باشد. بنابراین، اندازه‌گیری آن با استفاده از ویژگی‌های زودیافت خاک ضروری است. هدف از انجام این پژوهش توسعه یک مدل هوشمند بر اساس هوش مصنوعی و با استفاده از مدل تلفیقی عصبی - فازی همراستا (CANFIS) برای تخمین درصد سدیم تبادلی در جنوب شرقی استان آذربایجان شرقی است. در این رابطه ۲۰۹ نمونه خاک به صورت شبکه‌بندی منظم (۲۵۰×۲۵۰ متر) از عمق صفر تا ۲۵ سانتی‌متری برداشت و برای اندازه‌گیری درصد سدیم تبادلی و برخی دیگر از عوامل تأثیرگذار بر روی آن به آزمایشگاه منتقل شد. نتایج نشان داد که به ترتیب نسبت جذب سدیم (۰/۹۶۱)، هدایت الکتریکی (۰/۸۰۸)، pH (۰/۶۳۸)، مقدار رس (۰/۵۲۴)، شن (۰/۴۸۲) و سیلت (۰/۳۸۹) بیشترین تأثیر را در تخمین درصد سدیم تبادلی خاک دارند. در این پژوهش رابطه بین درصد سدیم تبادلی و نسبت جذب سدیم با ضریب تبیین ۰/۹۱ محاسبه شد. مدل CANFIS با ورودی‌های انتخاب شده از آنالیز به مؤلفه‌های اصلی مشتمل بر نسبت جذب سدیم، هدایت الکتریکی، واکنش خاک دارای کارایی بیشتری نسبت به مدل CANFIS با پنج ورودی نشان داد. طوری که آماره‌های جذرمیانگین مربعات خطا و ضریب تبیین برای مدل مناسب‌تر و به ترتیب ۱/۰ و ۰/۹۶ محاسبه شد. نتایج موید کارایی بالای شبکه‌های عصبی - فازی در تخمین درصد سدیم تبادلی است.

### کلمات کلیدی:

تجزیه به مؤلفه‌های اصلی،  
ویژگی‌های دیر یافت،  
ویژگی‌های زود یافت،

\* عهده دار مکاتبات:

Email: p\_alamdari@znu.ac.ir

### مقدمه

انواع شیوه‌های کشاورزی دقیق در مناطق خشک و نیمه خشک مانند ایران نیاز به اطلاعات دوره‌ای در مورد شوری و سدیمی بودن خاک دارد. این داده‌ها برای مدیریت کشاورزی پایدار بسیار مهم است (۳۷ و ۲۴). از جمله شناخت تغییرات درصد سدیم تبادلی<sup>۱</sup> و اطلاع از مقدار عددی آن در خاک‌های سدیمی یا شور و سدیمی به ویژه در اراضی کشاورزی برای برآورد مقدار مواد اصلاح کننده و مدیریت آن‌ها امری ضروری است (۱۰).

سدیمی بودن یکی از مهم‌ترین ویژگی‌های خاک‌های متأثر از نمک است. برای ارزیابی این ویژگی از معیارهای نسبت جذب سدیم<sup>۲</sup> و درصد سدیم تبادلی (ESP) استفاده می‌شود. اگرچه سدیمی بودن خاک به طور مستقیم با ESP در ارتباط است، با این حال استفاده از SAR متداول تر است (۴۲).

اندازه‌گیری ESP خاک، وقت گیر، پرهزینه و همراه با خطاهای دستگاهی است. عدم خارج شدن مازاد کاتیون شاخص طی مرحله شستشو در اندازه‌گیری ظرفیت تبادل کاتیونی و وجود کانی زئولیت در خاک منجر به بیش‌برآوردی ظرفیت تبادل کاتیونی و کم-برآوردی درصد سدیم تبادلی می‌شود. همچنین عدم اشباع کامل مکان‌های تبادلی با کاتیون شاخص، هدر رفتن خاک و هیدرولیز کاتیون شاخص تبادلی طی مرحله شستشو، عدم جایگزینی کامل آمونیوم به جای سدیم و انحلال گچ، ظرفیت تبادل کاتیونی را کمتر از مقدار واقعی اندازه‌گیری کرده و ESP بیش‌برآوردی خواهد داشت (۴۰). اندازه‌گیری این ویژگی در خاک‌های شور با هدایت-الکتریکی بالای ۱۰ دسی‌زیمنس بر متر با خطای اثر دفع آنیونی همراه است و ESP کمتر از مقدار واقعی برآورد می‌شود (۲۲). لذا برخی پژوهشگران برای اطلاع سریع از درجه سدیمی بودن خاک‌ها سعی کرده‌اند تا ESP را از ویژگی‌های زودیافت خاک برآورد کنند (۴۲).

یکی از روش‌های بررسی روابط و همبستگی بین ویژگی‌های مختلف خاک و بیان کمی آن‌ها استفاده از برخی مدل‌های آماری است. این مدل‌ها که اصطلاحاً توابع انتقالی نامیده می‌شوند، شامل مدل‌های رگرسیونی و هوش محاسباتی می‌باشند (۳۳). چنین مدل-

هایی ویژگی‌های دیریافت خاک را که اندازه‌گیری آن‌ها پرهزینه و زمان‌بر است، به صورت تابعی از ویژگی‌هایی که به سادگی و با هزینه کمتری قابل اندازه‌گیری هستند، ارائه می‌کند. پژوهشگران متعددی رابطه بین ESP و SAR خاک را با استفاده از روابط رگرسیونی خطی و لگاریتمی ارائه نموده‌اند (۱۵ و ۴۱).

داده‌کاوی یکی از پیشرفت‌های اخیر در زمینه فن‌آوری داده‌ها است که قابلیت زیادی برای کشف اطلاعات مهم از پایگاه داده را دارد. این فن‌آوری جزئی از تکنولوژی کشف دانش بوده که شامل مجموعه‌ای از روش‌های تخمین و گروه‌بندی است که رابطه‌ها، الگوها و دانش پنهان شده در درون انبوهی از داده‌ها را استخراج می‌کند. به طوری که داده‌کاوی را مترادف با تحلیل آماری پیشرفته برای کشف الگوهای مفید، قلمداد می‌کنند. تفاوت اصولی که بین داده-کاوی و آمار وجود دارد این است که داده‌کاوی یک رهیافت بدون پیش فرض است. درحالی که بیشتر فنون آماری معمول نیاز به پیش فرض دارند و پژوهشگران در جستجوی روابطی برای تطبیق آن پیش فرض‌ها هستند. در مقابل الگوریتم‌های داده‌کاوی، می‌توانند این روابط را به صورت خودکار از اطلاعات موجود در مجموعه داده‌ها توسعه دهند (۱۱).

اخیراً شبکه‌های عصبی مصنوعی<sup>۳</sup> به طور موفقیت‌آمیزی در تخمین بسیاری از ویژگی‌های دیریافت خاک مورد استفاده قرار گرفته است (۳۱). یکی از مزیت‌های استفاده از ANN نسبت به مدل‌های قدیمی این است که نیازمند تعیین یک تابع خاص برای بیان رابطه میان داده‌های ورودی و خروجی نیست و رابطه بین داده‌های ورودی و خروجی از طریق فرآیند آموزش به دست می‌آید (۴۶).

از دیگر روش‌های مناسب در تخمین ویژگی‌های خاک استفاده از سامانه‌های نروفازی (تلفیق شبکه‌های عصبی با منطق فازی) است. سامانه نروفازی با استفاده از الگوریتم آموزش شبکه عصبی پارامترهای سامانه فازی را تعیین می‌کند (۱۹). استفاده از سامانه‌های مبتنی بر قوانین فازی به دلیل پیوسته بودن تغییرات خاک‌ها و وجود عدم قطعیت در داده‌ها رو به گسترش است. در این زمینه پژوهشگران برای ایجاد توابع انتقالی خاک، دو روش رگرسیون و فازی را مورد مقایسه قرار دادند (۳۵). نتایج پژوهش محدث (۲۰۱۰) نشان داد که روش رگرسیون فازی، در شرایطی که بین متغیرها خطاهای ناشی از

1- Exchangeable Sodium Percentage (ESP)  
2- Sodium Adsorption Ratio (SAR)

3- Artificial Neural network (ANN)

داده‌ها پرداخت (۱۸). به‌طور مثال برای عملکرد بهتر سامانه‌های هوشمند می‌توان با استفاده از روش تجزیه به مؤلفه اصلی<sup>۵</sup> تعداد ورودی‌ها را کاهش داد تا بتوان مدل قوی‌تری ارائه کرد (۲۸). هدف از PCA، کشف ابعاد یا سازه‌های اصلی آن حوزه است. هدف بیشتر مطالعات PCA، خلاصه کردن ماتریس همبستگی به شیوه‌ای است که بتوان آنها را بر حسب چند عامل زیربنایی تبیین کرد (۲۶). پژوهش‌های انجام شده در این راستا حاکی از اهمیت توسعه روش‌های مقرون به صرفه برای تعیین مقدار ESP در خاک با دقت قابل اعتماد است. این پژوهش با هدف ارائه مدلی برای برآورد ESP خاک با استفاده از سایر ویژگی‌های خاک انجام شد. از آنجا که PCA یک روش وزن‌دهی به ویژگی‌ها براساس اهمیت آن‌ها است، در این پژوهش نیز با اعمال الگوریتم PCA تصمیم به کاهش واریانس ابعاد داده‌های ورودی گرفته شد. در واقع این فرایند ضمن کاهش داده‌های کم‌اهمیت‌تر و زمان محاسبات، باعث افزایش دقت مدل‌ها شد (۱۴). همچنین کارایی مدل‌های CANFIS در داده-کاوی و تخمین ویژگی‌های دیرپافت و قابلیت الگوریتم PCA برای انتخاب ترکیب ورودی‌های مناسب مدل مورد بررسی و مقایسه قرار گرفت.

### مواد و روش‌ها

محدوده مورد مطالعه در استان آذربایجان شرقی و شهرستان بناب (ساحل جنوب شرق دریاچه ارومیه) در منطقه قره چیق با وسعت ۱۴۵۰ هکتار واقع شده است. از نظر جغرافیایی نیز منطقه بین ۳۰' ۵۸' ۴۵" تا ۴۰' ۴۰" ۴۶°، طول شرقی و ۳۷' ۱۷' ۳۷" تا ۱۵' ۱۹' ۳۷° عرض شمالی قرار دارد. کاربری منطقه مورد مطالعه مرتع، زراعت دیم و اراضی کشت نشده بوده و رژیم‌های رطوبتی و حرارتی خاک در منطقه مورد مطالعه با استفاده از نرم افزار نیوهال<sup>۶</sup> به ترتیب اریدیک هم‌مرز با زیریک و مزیک برآورد گردید. برای نیل به اهداف تحقیق نمونه‌های خاک از محل هر گره از شبکه-بندی منظم با فواصل ۲۵۰×۲۵۰ متر و به تعداد ۲۰۹ نمونه از خاک سطحی (۰-۳۰ سانتی‌متری) در فصل تابستان سال ۱۳۹۹ برداشت و برای تجزیه و تحلیل به آزمایشگاه منتقل شد (شکل ۱). سپس واکنش خاک و هدایت الکتریکی نمونه‌ها در نسبت

ابهام در ساختار روابط وجود داشته باشد، مکمل و جایگزین مناسبی برای روش رگرسیون آماری است. قابلیت منطق فازی در علوم خاک مانند نقشه‌برداری و ارزیابی و مدل‌سازی فرآیندهای فیزیکی توسط پژوهشگران متعددی به اثبات رسیده است (۳۰ و ۳۲).

سامانه تلفیقی عصبی-فازی هم‌راستا<sup>۱</sup> یکی از مدل‌های تلفیقی عصبی و فازی است که به‌صورت توأم مزایای فازی شامل توجه به تغییرات پیوسته ویژگی‌های خاک و در نظرگیری عدم قطعیت در اندازه‌گیری ویژگی‌ها و محاسن ANN شامل بی‌نیازی از یک تابع ویژه برای بیان رابطه میان داده‌های ورودی و خروجی دارد. در این مدل کاربر قادر است تا مشخصات شبکه عصبی مصنوعی و مدل فازی را بر خلاف ساختار سامانه تلفیقی عصبی-فازی<sup>۲</sup> انتخاب و ساختار مطلوب را به روش آزمون و خطا به‌دست آورد (۶).

سارانی و همکاران<sup>۳</sup> (۴۴) در پژوهشی به تعیین رابطه بین ESP و SAR با استفاده از مدل‌های رگرسیونی و هوش محاسباتی اقدام کردند. نتایج این پژوهش نشان داد برای برآورد ESP خاک مدل ANFIS نتایج مطلوب‌تری نسبت به مدل‌های رگرسیونی ارائه می‌دهد. سرمدیان و کشاورزی<sup>۴</sup> (۴۵) از مدل CANFIS برای تخمین ظرفیت تبادل کاتیونی استفاده کردند. نتایج این پژوهش نشان داد که این روش دارای قابلیت و کارایی زیادی در تخمین ظرفیت تبادل کاتیونی خاک با استفاده از ویژگی‌های زودپافت خاک دارد. کشاورزی و همکاران (۲۶) از دو مدل ANN و ANFIS برای تخمین ESP در دشت مشهد استفاده کردند. مقادیر  $R^2$  و MSE از مدل شبکه‌های عصبی مصنوعی در مرحله آزمون به ترتیب ۰/۸۴ و ۵/۳۶ درصد بود، درحالی‌که این مقادیر برای مدل ANFIS به ترتیب ۰/۸۹ و ۵/۰۹ درصد بود. همچنین مدل ANFIS به دلیل فازی‌سازی ورودی‌ها از طریق توابع عضویت مختلف نسبت به رگرسیون چند متغیره در تخمین مقادیر ESP عملکرد بهتری داشت.

در مواجهه با حجم زیادی از داده‌ها در تحقیقات می‌توان با استفاده از روش‌های آماری متعددی به کاهش اندازه و دسته‌بندی

- 1- Co-active Neuro Fuzzy Inference system (CANFIS)
- 2- Neuro Fuzzy Inference system (ANFIS)
- 3- Sarani *et al.*
- 4- Sarmadian and Keshavarzi

5- Principal Component Analysis (PCA)  
6- Newhall

فازی را بر خلاف ساختار مدل ANFIS به تفکیک انتخاب و به ساختار بهینه به روش آزمون و خطا دست باید (۵۲).

در این تحقیق، مدل CANFIS انتخاب شده شامل پنج لایه است (شکل ۲). لایه اول، گره‌های عصبی و سری‌های مرتبط با قوانین فازی است که بخش مرکزی و اصلی آن‌ها در این لایه کنترل می‌شود. خروجی‌های این لایه مقادیر اولیه توابع عضویت فازی متناسب با ورودی‌های سامانه هستند. در لایه دوم، هر گروه موظف به محاسبه نتایج ورودی‌های خود است و دامنه فعال-سازی قوانین فازی تعیین می‌شود. در لایه سوم، درجه فعالیت مربوط به هر یک از قوانین، نرمال می‌شود. در لایه چهارم، گره-های موجود به صورت تطبیق‌پذیر در آمده و هر تابع یک رابطه درجه اول با پارامترهای حاصله ارائه می‌دهد و لایه پنجم، لایه خروجی است. به منظور بهینه‌نمودن مدل، لازم است که پارامترهای حاصله در لایه چهارم کاملاً صحیح تعریف شده و پارامترهای لایه اول، به عنوان بخش مرکزی دامنه تابع عضویت مطرح شوند.

در ساختار CANFIS توسعه یافته، قسمت عصبی ساختار پرسپترون چند لایه<sup>۵</sup> (MLP) و ساختار فازی دارای ۳ بخش اصلی فازی‌ساز ورودی<sup>۶</sup>، موتور استنتاج<sup>۷</sup> و غیرفازی‌ساز خروجی<sup>۸</sup> است. فازی‌ساز ورودی مقادیر قطعی متغیرهای ورودی را به مجموعه‌های فازی تبدیل می‌نماید و غیرفازی‌ساز خروجی مقادیر به دست آمده از موتور استنتاج فازی را به مقادیر قطعی تبدیل می‌نماید. سیستم کنترل فازی مورد استفاده در این پژوهش از نوع گوسین<sup>۹</sup> است که توانایی کاهش اغتشاشات احتمالی در داده‌های ورودی را دارد (۱۲). موتور استنتاج فازی از نوع سوگو<sup>۱۰</sup> است که از جمله موتورهای پر کاربرد در سامانه‌های فازی است (۷). شکل ۳ ساختار ترکیب شده شبکه‌های عصبی فازی توسعه یافته در این مدل را نشان می‌دهد.

۱: خاک به آب (۴۳)، بافت خاک به روش هیدرومتری (۹)، مقدار کاتیون سدیم توسط دستگاه فلیم فتومتر و کلسیم و منیزیم به روش تیتراسیون (۵۱) اندازه‌گیری شد.

### انتخاب متغیرها با PCA

در یک پژوهش، بررسی و شناسایی مؤلفه‌های اصلی یکی از ارزشمندترین نتایج جبر خطی است. زیرا روشی ساده و غیر پارامتری برای استخراج اطلاعات مورد نیاز از آن مجموعه می‌باشد. در الگوریتم تجزیه به مؤلفه‌های اصلی داده‌ها از مکانی با بعد زیاد به مکانی با بعد کم برده می‌شود. در واقع این شیوه یکی از روش‌های انتخاب ویژگی‌ها<sup>۱</sup> بوده و برای افزایش واریانس داده‌ها بکار برده می‌شود تا بررسی مجموعه‌ای از ویژگی‌ها در مکانی با بعد کم به سهولت انجام شود (۱۸). به طور کلی کاربرد عمده روش تجزیه به مؤلفه‌های اصلی عبارت است از کاهش تعداد متغیرها و یافتن ساختار ارتباطی بین آن‌ها (۳۴). اولین مؤلفه (PC1) دارای بزرگترین مقدار ویژه<sup>۲</sup> بوده و بیشترین درصد واریانس بین داده‌ها را نشان می‌دهد. مؤلفه‌های بعدی مقادیر باقی مانده واریانس بین داده‌ها را تبیین می‌نماید. معیار انتخاب مهم‌ترین مؤلفه‌ها در تجزیه به مؤلفه‌های اصلی مقدار ویژه بزرگتر از یک است (۲۱).

در این پژوهش از ماتریس واریانس-کوواریانس برای نشان دادن میزان تغییرات در نمونه و میزان همبستگی P متغیر با هم استفاده شد. این ماتریس، ماتریسی متقارن است که عضوهای قطر اصلی ماتریس، واریانس متغیرهای ورودی و بقیه درایه‌ها، کوواریانس بین متغیرهای ورودی است.

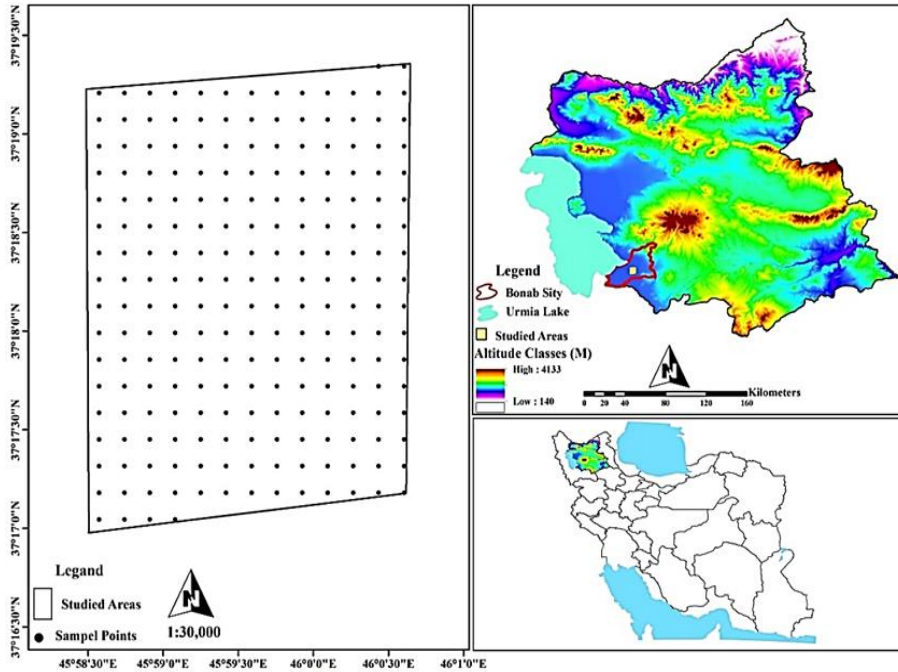
### مدل CANFIS

در این پژوهش با فرض موفقیت ANN و منطق فازی در مدل‌سازی و تخمین فرآیندهای پیچیده غیرخطی، یک مدل تلفیقی عصبی-فازی (CANFIS) برای برآورد ESP بسط داده شد (شکل ۲). این مدل از ویژگی‌های شبکه عصبی<sup>۳</sup> (ANN) و سامانه استنتاج فازی<sup>۴</sup> (FIS) بهره می‌گیرد (۴۵). در این مدل، کاربر قادر است تا مشخصات شبکه عصبی مصنوعی و مدل

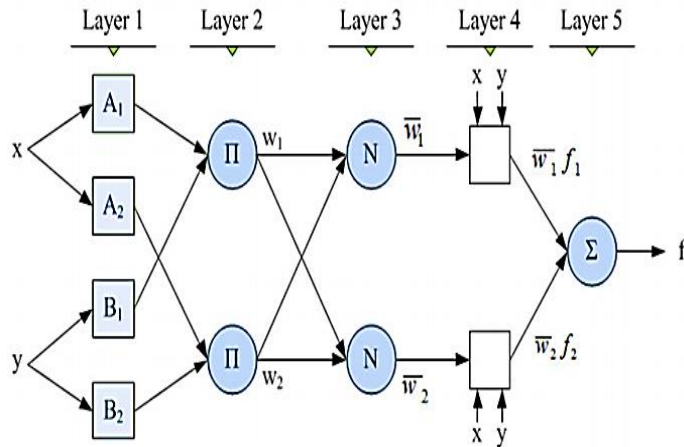
5- Multi-Layer Perceptron  
6- Fuzzifier Input  
7- Inference Engine  
8- Defuzzifier Output  
9- Gaussian  
10- Sugeno

1- Feature selection  
2- Eigen Value  
3- Artificial Neural Network  
4- Fuzzy Inference System

باریکلو و همکاران: تأثیر پیش پردازش متغیرهای ورودی مدل...



شکل (۱) محدوده مورد مطالعه و نقاط نمونه برداری  
Figure (1) Location of study area



شکل (۲) مدل CANFIS (۲۰)  
Figure (2) CANFIS model (Jang, 1993)

**معیارهای ارزیابی مدل‌ها**

برای این منظور از شاخص‌های ضریب همبستگی<sup>۱</sup>، جذر میانگین مربعات خطا<sup>۲</sup> به ترتیب رابطه‌های ۲ و ۳ (۵۳) و ضریب تطبیق همبستگی<sup>۳</sup> (CCC) از رابطه ۴ محاسبه شد (۳۸).

**نرمال کردن داده‌ها**

برای توسعه مدل CANFIS داده‌ها با استفاده از رابطه ۱ نرمال شدند.

$$y = 0.5 \left[ \frac{x - \bar{x}}{x_{max} - x_{min}} \right] 0.5 \tag{1}$$

که در آن  $\bar{x}$  میانگین داده‌های نرمال نشده،  $x$  مقدار داده نرمال نشده،  $x_{max}$  و  $x_{min}$  حداکثر داده‌های نرمال نشده و  $y$  داده نرمال شده است.

- 1-Correlation Coefficient (R2)
- 2- Root Mean Square Error (RMSE)
- 3 Concordance Correlation Coefficient (CCC)

باریکلو و همکاران: تأثیر پیش پردازش متغیرهای ورودی مدل...

همبستگی را با ESP داشتند. همچنین رابطه مثبتی بین واکنش خاک و ESP از یکسو (۱۷) و همبستگی بالا بین ESP و ECE از سوی دیگر توسط برخی پژوهشگران گزارش شده است. همبستگی مثبت ESP با محتوی رس و همبستگی منفی با محتوی شن نیز بیانگر اهمیت بخش ریز خاک است (۵۲). رویترز و می‌یر<sup>۱</sup> (۴۳) ESP را با استفاده از ویژگی‌های pH و EC در خاک‌های سدیمی و مرطوب استرالیا تخمین زدند و گزارش کردند که انتخاب درست متغیرهای ورودی باعث افزایش دقت مدل‌سازی می‌شود.

### انتخاب متغیرهای ورودی با استفاده از روش تجزیه به مؤلفه‌های اصلی (PCA)

برای عملکرد بهتر سامانه هوشمند CANFIS ابتدا با استفاده از الگوریتم PCA، سعی شد تعداد ورودی‌هایی که تأثیر چندانی در توسعه مدل نداشتند، کاهش یابند (۲۹). اگر تمام مؤلفه‌های اصلی ماتریس استخراج شود، می‌توان همبستگی‌های بین متغیرها را به‌طور کامل محاسبه نمود (۴۹). این مؤلفه‌ها باید قادر باشند جداگانه همبستگی‌ها را تبیین کنند (۲۶). در این پژوهش مقدار پارامتر KMO برابر با ۰/۷ بدست آمد که نشان‌دهنده امکان اجرای الگوریتم PCA روی داده‌ها است. لذا ماتریس همبستگی از مرتبه ۶ (معادل با تعداد متغیرهای ورودی) برای متغیرهای اولیه تشکیل شد (جدول ۲).

جدول ۳ مقادیر ویژه حاصل از اجرای الگوریتم PCA را نشان می‌دهد. در مرحله انتخاب، عامل‌هایی که حداقل ۷۰ درصد واریانس را توجیه کنند، انتخاب می‌شود. مطابق با جدول ۳ مقدار ویژه اول ( $\lambda_1$ ) برابر با ۳/۶۳ بوده و ۶۰/۵۵ درصد از کل واریانس داده‌ها را شامل می‌شود. دومین و سومین مقدار ویژه ۱/۳۰ و ۰/۶۴ است و به ترتیب ۲۱/۷۳ و ۱۰/۷۴ درصد از کل واریانس داده‌ها را نشان می‌دهد. بنابراین ۳ مؤلفه اول در مجموع، ۹۳/۰۳ درصد از واریانس متغیرها را دربر می‌گیرد. در نهایت از بین شش عامل، سه عامل انتخاب شد که نتایج نمودار واریزه‌ای (شکل ۴) نیز مؤید این مطلب است و برای مؤلفه‌های بزرگ‌تر از ۳ شیب

$$R^2 = \frac{[\sum_{i=1}^n (q_i - \bar{q}_i)(\hat{q}_i - \bar{\hat{q}}_i)]^2}{\sum_{i=1}^n (q_i - \bar{q}_i)^2 \sum_{i=1}^n (\hat{q}_i - \bar{\hat{q}}_i)^2} \quad (۲)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_i - \hat{q}_i)^2} \quad (۳)$$

که در آن  $q_i$  مقدار ESP اندازه‌گیری شده،  $\hat{q}_i$  مقدار ESP برآورد شده،  $\bar{q}_i$  میانگین مقادیر ESP اندازه‌گیری شده،  $\bar{\hat{q}}_i$  میانگین ESP محاسباتی و  $n$  تعداد داده‌ها است.

(۴)

$$CCC = \frac{2r\hat{a}\hat{b}}{\hat{\sigma}_a^2 + \hat{\sigma}_b^2 + (\bar{a} + \bar{b})^2}$$

که در آن،  $r$  ضریب همبستگی،  $\hat{\sigma}_a$  واریانس مقادیر مشاهده شده،  $\hat{\sigma}_b$  واریانس مقادیر برآورد شده،  $\bar{a}$  میانگین مقادیر مشاهده شده و  $\bar{b}$  میانگین مقادیر برآورد شده است.

### نتایج و بحث

#### خلاصه آماری توصیف داده‌ها

خلاصه آماری ویژگی‌های خاک‌های مورد استفاده در جدول ۱ ارائه شده است. به‌طور کلی ضریب تغییرات، معیاری از تغییرپذیری نسبی داده‌ها است. برخی از محققین ضریب تغییرات ویژگی‌های خاک را به دو دسته پایین (۰-۱۵٪) و بالا (۱۵-۷۵٪) طبقه‌بندی نمودند (۱۳). در این مطالعه مقادیر ضریب تغییرات بین ۰/۰۲ و ۱۶/۵۹ به ترتیب برای واکنش خاک و ESP بود (جدول ۱). میانگین درصد سدیم تبدالی در خاک‌های مورد مطالعه ۲۷/۳۱ بود که جزء خاک‌های سدیمی محسوب می‌شوند. همچنین میانگین واکنش خاک نیز بالاتر از ۸/۵ است.

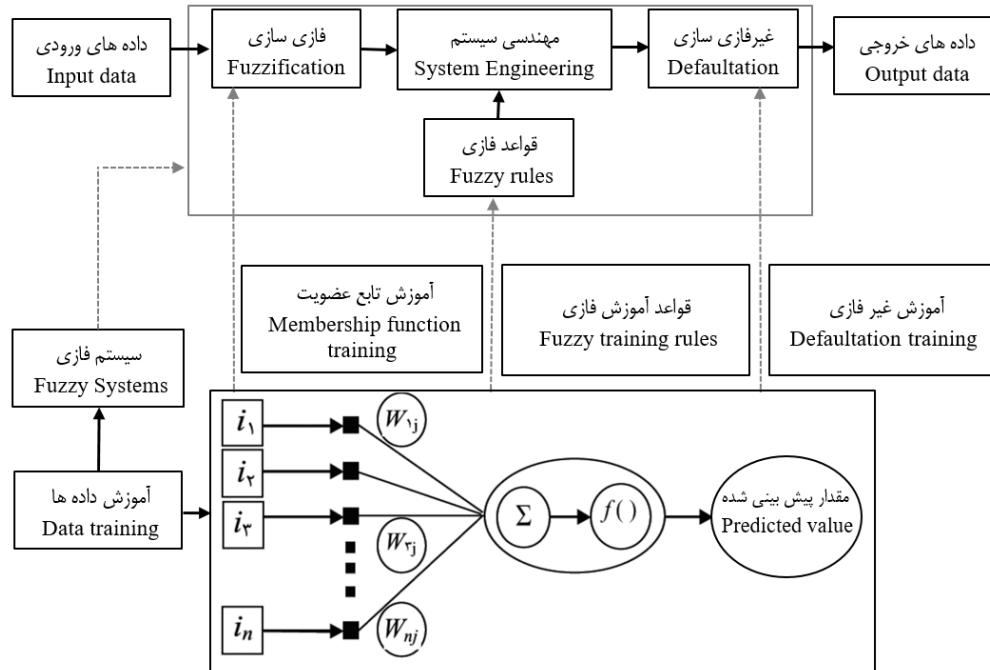
#### همبستگی پیرسون بین درصد سدیم تبدالی و ویژگی‌های خاک

جدول ۲ همبستگی پیرسون بین متغیرها را نشان می‌دهد. دقت تخمین مدل‌های هوشمند به انتخاب صحیح اطلاعات ورودی لایه اول بستگی دارد. بنابراین با استفاده از ماتریس همبستگی ارتباط بین پارامترهای خاک (متغیر مستقل) و ESP (متغیر وابسته) مشخص شد. نتایج نشان داد، SAR و سیلت به ترتیب بیشترین و کمترین



انتخاب شده با استفاده از الگوریتم PCA نسبت به هفت ویژگی ورودی اولیه اشاره شده است (۲۳).

خط کاهش زیادی را نشان می دهد. در پژوهشی به منظور مدل سازی اندازه قطعات سنگ، طی انفجار در معدن سنگ آهن سیرجان، به برتری مدل ANFIS با دو ویژگی ورودی



شکل (۳) ساختار عصبی فازی تلفیقی (۵۳)

Figure (3) Involved structure of fuzzy neural network (ZareAbyaneh & BayatVarkeshi, 2013)

جدول (۱) شاخص های آماری برخی از ویژگی های خاک در مدل سازی درصد سدیم تبدلی

Table (1) Statistical indicators of some soil properties in the ESP modeling

pH	Sand شن	Silt سیلت	Clay رس	EC هدایت الکتریکی	SAR	ESP	پارامتر آماری Statistical parameters
					نسبت جذب سدیم (mmol L <sup>-1</sup> )	درصد سدیم تبدلی (%)	
-		(%)		(dS m <sup>-1</sup> )			
۸/۵۲	۲۷/۹۶	۳۲/۶۶	۳۹/۳۷	۲۵/۴۷	۲۲/۵۲	۲۷/۳۱	میانگین Mean
-۰/۰۱	۰/۲۹	-۰/۰۱۵	-۰/۰۶	۱/۰۵	۰/۸۱	۰/۷۹	چولگی Skewness
-۰/۸۸	۰/۱۱	۰/۱۰	۰/۴۰	۰/۲۶	-۰/۲۶	-۰/۳۳	کشیدگی Kurtosis
۰/۰۲	۴/۷۱	۱/۴۴	۱/۹۸	۱۴/۶۷	۱۴/۰۴	۱۶/۵۹	ضریب تغییرات Coefficient of Variation (CV)
۷/۱۷	۲/۰۰	۱۵/۰۰	۱۶/۰۰	۱/۳۷	۰/۹	۱/۲	حداقل Min
۸/۸۸	۶۱/۰۰	۴۹/۰۰	۶۳/۰۰	۸۱/۵۵	۷۰/۳	۸۴/۷	حداکثر Max

### تعیین داده‌های آموزش و آزمون مدل

از آنجایی که شبکه عصبی در ساختار مدل نروفازی فاقد قدرت برون‌یابی است و توانایی تعمیم‌دهی آن فقط در چارچوب درون‌یابی مطرح است، لذا داده‌های آموزشی طوری باید انتخاب شوند که نماینده کل داده‌ها باشند. به عبارت دیگر تمامی حالات ممکن در داده‌ها از قبیل مقادیر حدی (حداقل و حداکثر) را در خود داشته باشند. در این پژوهش، تقسیم داده‌ها به وسیله تصادفی کردن نمونه‌ها و سپس انتخاب تصادفی ۸۰ درصد از داده‌ها برای آموزش مدل و ۲۰ درصد باقیمانده برای آزمون مدل انجام شد (۳). شایان ذکر است که خاک‌های مورد مطالعه با توجه به تنوع زیاد از نظر هدایت الکتریکی و SAR، بر اساس طبقه‌بندی آزمایشگاه شوری آمریکا در کلاس‌های نرمال، شور، سدیمی و شور و سدیمی طبقه‌بندی می‌شوند.

### برآورد درصد سدیم تبادلی خاک با استفاده از مدل CANFIS

برای این منظور، ابتدا مدل با متغیرهای ورودی مشتمل بر نسبت جذب سدیم، هدایت الکتریکی، عصاره اشباع، مقدار رس، شن، سیلت و واکنش خاک توسعه یافت (مدل ۱). سپس با اعمال الگوریتم PCA بر ورودی‌ها، تعداد آنها کاهش پیدا کرد (مدل ۲).

برای تعیین تعداد مطلوب نرون‌ها از روش سعی و خطا استفاده شد. نتایج مناسب‌ترین مدل با آماره‌های RMSE،  $R^2$  و CCC در جدول ۴ ارائه شده است.

مدل ۱ با وارد کردن تمام ورودی‌ها، ۸۸ درصد تغییرات و مدل ۲، ۹۶ درصد تغییرات مربوط به درصد سدیم تبادلی را توجیه می‌کند. به طوری که مقادیر آماره‌های RMSE،  $R^2$  و CCC در مرحله آزمون برای مدل ۱ و ۲/۲۵، ۰/۸۸ و ۰/۹۴ و مدل ۲، ۱/۰۱، ۰/۹۶ و ۰/۹۶ برآورد شد.

شکل ۵ (الف) نشان می‌دهد که مدل ۱ دارای دقت خوبی است. به طوری که در مرحله آزمون، ۸۸ درصد تغییرات ESP بستگی به متغیرهای انتخابی و بقیه مربوط به عوامل کنترل نشده دارد. میانگین انحراف مربعات خطا ۲/۲۵ درصد تخمین زده شد که با توجه به تعداد داده و نوع متغیر مناسب است. همچنین شکل ۶ (ج) تغییرات درصد سدیم تبادلی برآورد شده و اندازه‌گیری شده را نشان می‌دهد. این شکل نشان می‌دهد مدل ۱ در برخی نقاط دارای برآورد کمتر نسبت به داده‌های مشاهداتی بوده است.

جدول (۲) ماتریس همبستگی پیرسون بین ویژگی‌های خاک

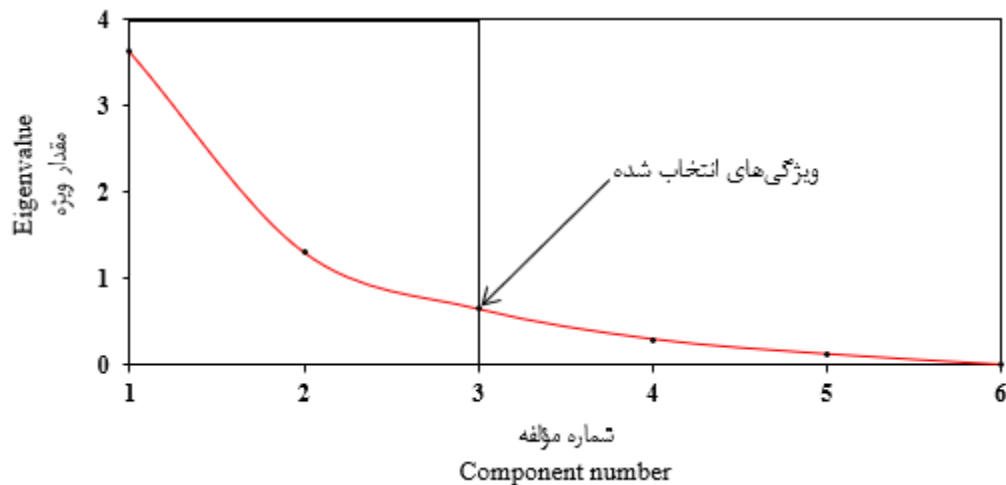
Table (2) Pearson correlation matrix among soil properties

pH	Sand	Silt	Clay	EC	SAR	ESP
اسیدیته	شن	سیلت	رس	هدایت الکتریکی	نسبت جذب سدیم	درصد سدیم تبادلی
						۱
					۰/۹۶۱**	۰/۸۰۸**
				۱	۰/۸۱۹**	۰/۵۲۴**
		۱	۰/۷۸۹	۰/۶۲۱**	۰/۵۱۱**	۰/۳۸۹*
	۱	-۰/۶۴۰**	-۰/۸۰۲**	-۰/۳۲۸*	-۰/۴۹۸**	-۰/۴۸۲**
۱	-۰/۱۳۴	-۰/۱۵۹	۰/۳۷۹*	۰/۴۱۷*	۰/۶۱۲**	۰/۶۳۸**

جدول (۳) مقادیر ویژه به دست آمده از ویژگی های خاک در مدل سازی درصد سدیم تبادلی

Table (3) Obtained eigenvalues values from some soil properties in the ESP modeling

مقدار ویژه اولیه Initial Eigenvalue			مؤلفه Component
درصد تجمعی Percentage of Cumulative	درصد توجیه واریانس Percentage of variance	مقدار ویژه Eigenvalue	
۶۰/۵۵	۶۰/۵۵	۳/۶۳	۱
۸۲/۲۸	۲۱/۷۳	۱/۳۰	۲
۹۳/۰۳	۱۰/۷۴	۰/۶۴	۳
۹۷/۹۵	۴/۹۲	۰/۲۹	۴
۹۹/۹۹	۲/۰۴	۰/۱۲	۵
۱	۰/۰۰۰۱	$۲/۴۵ \times 10^{-5}$	۶



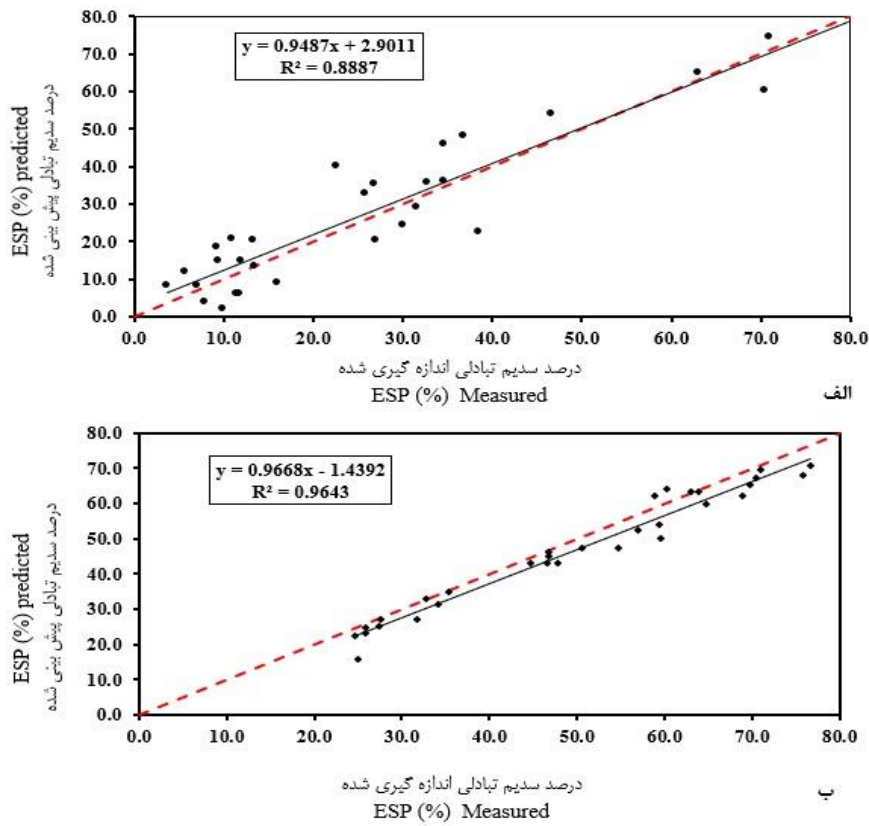
شکل (۴) نمودار واریزه ای برای متغیرهای ورودی  
Figure (4) Deposit diagram for input variables

جدول (۴) ویژگی های مدل های مختلف CANFIS برای تخمین درصد سدیم تبادلی

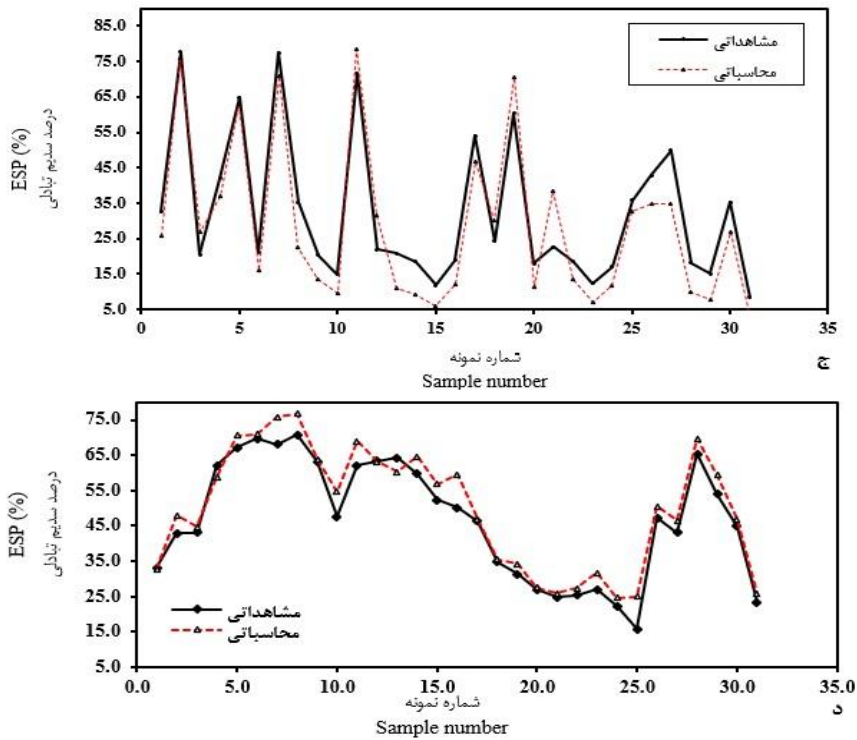
Table (4) Characteristics of different CANFIS models for estimating the exchangeable sodium percentage

CCC	RMSE (%)	R <sup>2</sup>	تعداد ورودی Number of Inputs	تکرار Repetition	تعداد توابع عضویت Number of Membership Functions	تابع عضویت Membership Function	یادگیری Learning	روش دفاژی کردن Defuzzification	ویژگی Property
۰/۹۴	۲/۲۵	۰/۸۸	۶	۱۱۸	۳	گوسین Gaussian	هیبرید Hybrid	میانگین وزنی Weighted average	مدل ۱ Model 1
۰/۹۶	۱/۰۱	۰/۹۶	۳	۱۲۴	۳	گوسین Gaussian	هیبرید Hybrid	میانگین وزنی Weighted average	مدل ۲ Model 2

باریکلو و همکاران: تأثیر پیش پردازش متغیرهای ورودی مدل...



شکل (۵) ضریب تبیین درصد سدیم تبادل برآورد شده و اندازه‌گیری شده مدل ۱ (شکل الف) و مدل ۲ (شکل ب)  
 Figure (5) Coefficient of determination for estimated ESP and measured ESP of model 1 (Figure A) and model 2 (Figure B)



شکل (۶) تغییرات درصد سدیم تبادل برآورد شده و اندازه‌گیری شده برای مدل ۱ (شکل ج) و مدل ۲ (شکل د)  
 Figure (6) Estimated and measured ESP variation for Model 1 (Figure J) and Model 2 (Figure D)

دقت بیشتری دارند. آموتا و پورشلوان<sup>۵</sup> (۴) به منظور پیش-بینی سطح آب زیرزمینی در حوزه آبخیز مالانار، عملکرد دو مدل RBFN و ANFIS را مورد ارزیابی قرار دادند. ایشان مدل‌ها را با سه ورودی اجرا نمودند و گزارش کردند که عملکرد مدل ANFIS با این تعداد ورودی نسبت به شبکه RBF مطلوب‌تر است.

### نتیجه‌گیری

در این پژوهش مدل CANFIS و روش تجزیه به مولفه‌های اصلی (PCA) برای تخمین درصد سدیم تبادلی با استفاده از ویژگی‌های زودیافت خاک ارزیابی و مورد استفاده قرار گرفت.

نتایج اعتبارسنجی برآورد درصد سدیم تبادلی خاک با استفاده از مدل ۱ که تمام متغیرهای اندازه‌گیری شده خاک در آن استفاده شده بود (  $R^2=0.88$ ;  $RMSE=2.25$  ) بیانگر عملکرد ضعیف‌تر نسبت به مدل ۲ که متغیرهای انتخابی با استفاده از روش PCA (  $R^2=0$ ;  $RMSE=1.01$ ;  $CCC=0.96$  ) انتخاب شد بود. نتایج نشان می‌دهد برای مدل‌سازی ویژگی‌های دیریافت خاک مانند درصد سدیم تبادلی خاک و برای دست‌یابی به مناسب‌ترین مدل، توجه به تعداد و مؤثرترین متغیرهای ورودی بسیار ضروری است. زیرا هدف اصلی، ارائه مدلی با حداقل تعداد ورودی و همچنین ورودی‌هایی که اندازه‌گیری آنها آسان و در زمان کوتاه باشد، است. نتایج کمی-سازی اهمیت متغیرها در مدل ۲ تأییدی بر استفاده از سه ویژگی نسبت جذب سدیم، هدایت الکتریکی عصاره اشباع خاک و اسیدیته خاک در مدل‌سازی درصد سدیم تبادلی است. با در نظر گرفتن شاخص‌های ارزیابی و تعداد متغیرهای بکار گرفته شده، مدل ۲ را می‌توان کارآمدترین مدل ارائه شده در پژوهش حاضر معرفی نمود. با کاهش تعداد ورودی و انتخاب ورودی‌های مؤثرتر با استفاده از الگوریتم PCA عملکرد مدل بهبود یافت. این مدل می‌تواند درصد سدیم تبادلی را با قدرت اطمینان بالایی تخمین

مدل ۲ در بیشتر نواحی توانسته به خوبی درصد سدیم تبادلی را تخمین بزند و به دنبال کم‌ترین خطای تخمین است (شکل ۵ ب و شکل ۶ د). این مدل با داشتن  $R^2=0.96$ ،  $RMSE=1/0.1$  و  $CCC=0.96$  تخمینگر بهتری نسبت به مدل ۱ است. دلیل کارایی کمتر مدل ۱، افزایش زمان یادگیری، پیچیده‌تر شدن ساختار و ایجاد اختشاش در مدل به دلیل حجم بالای اطلاعات ورودی است. همچنین داده‌ها در محیط خاک دارای عدم قطعیت هستند. لذا هرچه تعداد داده‌های ورودی غیر ضروری مدل کاهش یابد، نتایج حاصله بهبود می‌یابد. یکی از نقاط قوت در ساختار مدل CANFIS، حضور شبکه عصبی است. برخی از پژوهشگران معتقدند در طراحی شبکه‌های عصبی مصنوعی نوع خاصی از روابط رگرسیونی لازم نبوده و با ایجاد رابطه مناسب بین داده‌های ورودی و خروجی می‌توان به نتایج مناسب دست پیدا کرد (۴۶ و ۱). همچنین به علت وجود روابط غیر خطی میان متغیرهای مستقل و برآوردی، مدل ۲ عملکرد خوبی را نشان داده است. نکته دیگر این که انتخاب صحیح مجموعه داده‌های آموزش و آزمون که جزئی از فرآیند داده‌کاوی است، می‌تواند منجر به افزایش دقت مدل انتخابی در برآورد ESP خاک شود. بطوریکه برخی پژوهشگران نشان دادند که اگر ناپایداری داده‌ها بالا باشد، شبکه عصبی عملکرد خوبی نخواهد داشت. اما زمانی که داده‌هایی با دقت بالا به کار برده شود، شبکه عصبی کارایی بهتری را نشان می‌دهد (۵۰). در مدل‌سازی CANFIS کاربر قادر است تا مشخصات شبکه عصبی مصنوعی و مدل فازی را برخلاف ساختار ANFIS به تفکیک انتخاب و به ساختار بهینه آن به روش آزمون و خطا دست یابد (۵۵).

سینگ و دو<sup>۱</sup> (۴۹) با پیش‌بینی جریان روزانه رودخانه نارمادا در هند با استفاده از سامانه CANFIS و شبکه‌های عصبی<sup>۲</sup> GRNN،<sup>۳</sup> MLP و<sup>۴</sup> RBF، به این نتیجه رسیدند که سامانه CANFIS نسبت به شبکه‌های عصبی مصنوعی

- 1- Singh and Deo
- 2- General Regression Neural Network
- 3- Multi-Layer Perceptron
- 4- Radial Basis Function

<sup>5</sup> Amutha and Porchelvan

CANFIS که از مجموعه‌های فازی استفاده می‌کنند، در برازش توابع انتقالی خاک می‌تواند مفید واقع گردد.

### سپاس‌گزاری

بدین وسیله از آب منطقه‌ای استان اذربایجان شرقی که داده‌های پایه پژوهش مشتمل بر ویژگی‌های بافت، ESP، SAR، EC و pH مربوط به ۲۰۹ نمونه خاک را در اختیار مؤلفین مقاله قرار دادند، تشکر و قدردانی می‌شود.

بزند و محدودیت نیاز به صرف وقت و هزینه‌های زیاد در برآورد مستقیم درصد سدیم تبدیلی را مرتفع می‌سازند. نتایج این پژوهش را می‌توان به خاک‌های مناطق خشک و نیمه-خشک مشابه نیز تعمیم داد. همچنین با توجه به ماهیت ابهامی پدیده‌های مرتبط خاک یا تقریبی بودن مقادیر اندازه‌گیری شده خصوصیات مختلف خاک و عدم قطعیت موجود در داده‌ها، استفاده از مدل‌های ترکیبی مانند

### References

1. Abdulshahed A.M., Longstaff A.P., Fletcher S., and Myers A. 2015. Thermal error modelling of machine tools based on anfis with fuzzy c-means clustering using a thermal imaging camera. *Applied Mathematical Modelling*, 39(7):1837-1852.
2. Adhikari P., Shukla M.K., and Mexal J.G. 2011. Spatial variability of electrical conductivity of desert soil irrigated with treated wastewater: implications for irrigation management. *Applied and Environmental Soil Science*, 2011:1-11.
3. Amini M., Abbas Pour K.C., Khademi H., Fathianpour N., Afyuni M. and Schulin R. 2005. Neural network models to predict cation exchange capacity in arid regions of iran. *European Journal of Soil Science*, 53:748-757.
4. Amutha R., and Porchelvan P. 2011. Seasonal prediction of groundwater levels using anfis and radial basis neural network. *International Journal of Geology, Earth and Environmental Sciences*, 1:98-108.
5. Auerswald K., Simon S., Stanjek H. 2001. Influence of soil properties on electrical conductivity under humid water regimes. *Soil Sci*, 166(6):382-390.
6. Aytak A. 2008. Co-active neuro-fuzzy inference system for evapotranspiration modeling: soft computing, a fusion of foundations method. *Journal of Applied Science*, 13(7):691-700.
7. Bayat Varkeshi M., Zare Abyaneh H., Marofi S., Sabziparvar A.A. and Soltani F. 2009. Simulation of reference evapotranspiration using artificial neural method and empirical methods and comparison with experimental lysimeter data in cold semi-arid climate of hamedan. *Journal of Water and Soil Conservation*, 16(4):79-100. (In Parsian with English abstract)
8. Bhargava G.P., and Abrol I.P. 1978. Characteristics of some typical salt affected soils of uttar pradesh. division of soils and agronomy, Central Soil Salinity Research Institute: Yazd.
9. Bouyoucos G.J. 1962. Hydrometer method improved for making particle size analysis of soils. *Agronomy Journal*, 56: 464-466.
10. Bower C.A., Reitemeier R.F., and Fireman M. 1952. Exchangeable cation analysis of saline and alkali soils. *Soil Science* 73:251-261.
11. Cabena P.H., Stadler R., Verhees J., and Zanasi A. 1998. *Discovering data mining: from concept to implementation*, IBM, New Jersey, 195 Pp.
12. Chiu S.L. 1994. Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, 2(3):267-278.

13. Dahiya I. S., Richter J., and Malik R. S. 1984. Soil spatial variability: a review. *International Journal of Tropical Agriculture*, 11:91-102.
14. Ebrahimi F., Afsar A., Rezaeinour J., and Ghanbari Sorkhi A. 2017. An efficient intelligent approach for short-term load forecasting. *Industrial engineering & management sharif*. 33(1):43-54. (in Persian with English abstract)
15. Emerson W.W., and Bakker A.C. 1973. The comparative effects of exchangeable calcium, magnesium and sodium on some physical properties of red-brown earth sub-soils. II. The spontaneous dispersion of aggregates in water. *Australian Journal of Soil Research*, 11(2):151 – 157.
16. 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. *Journal of Soil and Water*, 22:1-15. (in Persian)
17. Fireman M., and Wadleigh C.H. 1951. A statistical study of the relation between pH and the exchangeable-sodium-percentage of western soils. *Journal of Soil Science*, 71(4):273-286.
18. Hemmatpour S., and Hashemi H. 2012. Using pca and rda feature reduction techniques for ranking seismic attributes. *Journal of the Earth and Space Physics*, 37(4):217-227. (In Persian with English abstract)
19. Jang J.S., Sun C.T., and Mizutani E. 1997. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence [Book Review]. *IEEE Transactions on automatic control*, 42(10):1482-1484.
20. Jang J.S. 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3):665-685
21. Johnson R.A., and Wichern D.W. 2014. *Applied multivariate statistical analysis*. London, UK Pearson Prentice-hall, 6:590p.
22. Jurinak J., and Suarez D.L. 1990. The Chemistry of Salt-Affected Soils. In K. K. Tanji (Ed.). *Agricultural Salinity Assessment and Management*, 71:42-63. American Society of Civil Engineers, New York: N.Y.
23. Karami A., and Afiuni-Zadeh S. 2012. Sizing of rock fragmentation modeling due to bench blasting using adaptive neuro-fuzzy inference system and radial basis function. *International Journal of Mining Science and Technology*, 22:459–463.
24. Keshavarzi A., and Sarmadian F. 2012. Mapping of spatial distribution of soil salinity and alkalinity in a semi-arid region. *Ann Warsaw Agricult Univ SGGW Land Reclam*, 44(1):3–14.
25. Keshavarzi A., Bagherzadeh A., Omran E.S.E., and Iqbal M. 2016. Modeling of soil exchangeable sodium percentage using easily obtained indices and artificial intelligence-based models. *Modeling Earth Systems and Environment*, 2(3):1-10.
26. Kline P. 2014. *An easy guide to factor analysis*. New York: Routledge.
27. Liu C.W., Lin K.H., and Kuo Y.M. 2003. Application of factor analysis in the assessment of groundwater quality in a black foot disease area in taiwan. *Science of the Total Environment*, 313:77-89.
28. Zhou S., Lin M., and Ma L. 2008. Support vector machine optimized with genetic algorithm for short-term load forecasting. *International Symposium on Knowledge Acquisition and Modeling*. IEEE, 654-657.
29. Mahdavi Tabatabaei N., Mortezaei S.R., Taghizadegan N., and Nayir A. 2008. Short term load forecasting by neural network in mashhad (iran) power system. *University of Pitesti Electronics and Computers Science*, 2(8):12-18.

30. Mcbratney A.B., and Odeh I.O.A. 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma*, 77(2-4):85-113.
31. Merdun H., Meral O.C., and Apan R.M. 2006. Comparison of artificial neural network and regression pedotransfer functions for predict of water retention and saturated hydraulic conductivity. *Soil and Tillage Research*, 90:108-116.
32. Metternicht G., and Gonzalez S. 2004. Foundations of a fuzzy exploratory model for soil erosion hazard prediction. *Environmental Modelling & Software*, 20(6):715-728.
33. Minasny B., Hopman J., Harter W.T., Eching S.O., Toli A., and Denton M.A. 2004. Neural networks prediction of soil hydraulic functions for alluvial soils using multistep outflow data. *Soil Science Society of America Journal*, 68:417-429.
34. Mohaddes F. 2010. Principal component and factor analysis case study: assets price evaluation and inflation impacts. Economic Research and Policy Department Central Bank of the Islamic Republic of Iran.
35. Mohammadi J., and Taheri M. 2005. Estimation of pretransfer function using fuzzy regression. *Journal of Agriculture Science and Technology*, 2:51-60.
36. Monem M.J., and Kiapasha M.S. 2009. Development and evaluation of an upstream fuzzy control system model for irrigation canals. *Journal of Iranian Water Research*, 3(4):41-50.
37. Mousavi S.R., Sarmadian F., Omid M., and Bogaret p. 2021. Application of Machine Learning Models in Spatial Estimation of Soil Phosphorus and Potassium in Some Parts of Abyek Plain. *Journal of Iranian Spoi Research*, 35(4):397-411.
38. Omran E.S. 2008. Is soil science dead and buried Future image in the world of 10 billion people. *Catrina: The International Journal of Environmental Sciences*, 3(2):59-68.
39. Rashidi M., and Seilsepour M. 2008. Modeling of soil exchangeable sodium percentage based on soil sodium adsorption ratio. *Journal of Agricultural and Biological Science*, 3(4):22-26.
40. Rhoades J. D. 1982. Cation exchange capacity. In A. L. Page Et Al (Ed.). (Part. 2). (Pp. 149-157). *Methods Of Soil Analysis, Agron. Monogr. 9, American Society Of Agronomy, Madison, WI, USA.*
41. Richards L.A. 1954. Diagnosis and improvement of saline alkali soils, *Agriculture*, 160, Handbook 60. US Department of Agriculture, Washington DC.
42. Robbins C.W., and Meyer W.S. 1990. Calculating ph from ec and sar values in salinity models and sar from soil and bore water pH and EC data. *Soil Research*, 28(6):1001-1011.
43. Rowell D.L. 1994. *Soil science: methods and application*. England: longman group, harlow, UK.
44. Sarani F., Gholamalizadeh A., and Shabani A. 2016. Comparing regression and artificial intelligence models for estimating soil exchangeable sodium percentage from sodium absorption ratio (case study: miankangi region soils, sistan. *Journal Water and Soil Science*. 26(2):125-137. (in Parsian with English abstract)
45. Sarmadian F., and Keshavarzi A. 2016. Application of canfis model in prediction of soil cation exchange capacity in some arid and semi-arid regions of iran. *Journal of range and watershed management. Iranian Journal of Natural Resources*, 69(2):397-410. (In Persian)
46. Schaap M.G., and Bouten W. 1996. Modelling water retention curves of sandy soils using neural networks. *Water Resources Research*, 32:3033-3040.
47. Seilsepour M., and Rashidi M. 2008a. Modelling of soil sodium adsorption ratio based on soil electrical conductivity. *ARPN J Agric Biol Sci* 3:27-31.



48. Singh P., and Deo M.C. 2007. Suitability of different neural networks in daily flow forecasting. *Applied Soft Computing*, 7:968-978.
49. Spearman C. 1994. General intelligence: objectively determined and measured. *American Journal of Psychology*, 15:92-201.
50. Tamari S., Wosten J.H.M., and Ruz-Suarez J.C. 1996. Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Science Society of America Journal*, 60:1732-1741.
51. USDA. 1996. Soil survey laboratory methods manual. Soil survey investigations republic, washington: united states government print.
52. Wiegand C. L., Lyles L., and Carter D.L. 1986. Interspersed salt-affected and unaffected dryland soils of the lower rio grande valley: ii. Occurrence of salinity in relation to infiltration rates and profile characteristics. *Soil Science Society of America Journal*, 30(1):106-110.
53. Wosten J.H.M., Pachepsky Y.A., and Rawls W.J. 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251:123-150.
54. Zare Abyaneh H., and Bayat Varkeshi M. 2013. Development and application of statistical and neural, fuzzy, genetic algorithm models in estimation of spatial distribution of water table level. *Journal of water and soil conservation. Journal of Agricultural Sciences and Natural Resources*, 20(4):1-25. (in Parsian with English abstract)
55. Zare Abyaneh H., Bayat Varkeshi M., and Marofi S. 2011. Forecasting of garlic (*allium sativum* L.) Evapotranspiration by using multiple modeling. *Journal of Agricultural Science and Natural Resource*, 18(2):89-107. (In Parsian with English abstract)