

## Evaluation of Different Supervised Learning Smart Methods and Response Surface Method to Optimize Factors Affecting Erosion (Case Study: Emamzadeh Watershed of Baghmalek)

MOJTABA SHIRAZI<sup>1</sup>, ATAALLAH KHADEMALRASOUL<sup>\*1</sup>, SEYED MOHAMMAD SAFIEDDIN ARDEBILI<sup>2</sup>

1. Soil Science Department, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

2. Bio System Engineering Department, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

(Received: Jun. 27, 2020- Revised: Apr. 20, 2020- Accepted: Apr. 30, 2020)

### ABSTRACT

Evaluation of soil erosion control factors is important regarding the application of management practices. In this study, the effects of non-structural management practices including revision of crop cover (RC) and enclosure (EX) were simulated using WEPP model in Emamzadeh Abdullah watershed of Baghmalek, located in the northeast of Khuzestan Province. Optimization of physical and hydraulic parameters affecting erosion including soil texture and components of the Van Genuchten equation was performed using response surface methodology (RSM), random forest (RF), support vector machine (SVM) and artificial neural network (ANN). Also, the soil erosion rate before and after management practices was defined as the first response ( $R_1$ ) and the second response ( $R_2$ ), respectively. Optimization results by Orange software including random forest methods, support vector machine and artificial neural network showed that the random forest method with MSE, RMSE and  $R^2$  equal to 0.991, 0.995 and 0.963 respectively, for the first response and equal to 0.095, 0.307 and 0.974 respectively, for the second response is the most proper method. Also, optimization by response surface method is the most appropriate method with MSE, RMSE and  $R^2$  equal to 28.7, 5.37 and 0.999 respectively, for the first response and equal to 4.16, 2.03 and 0.998 respectively, for the second response. Generally, using optimization techniques is a convenient method for evaluating management practices and finally selecting the best one for critical areas. Appropriate management practices based on optimal conditions leading to water and soil loss reduction.

**Keywords:** Soil Erosion, Physical and Hydraulic Variables, RETC Software, Management Scenarios, Independent and Response Variables.

## ارزیابی روش‌های یادگیری نظارتی هوشمند و سطح پاسخ برای بهینه‌سازی عوامل مؤثر بر فرسایش خاک (مطالعه‌ی موردی حوزه آبخیز امامزاده عبدالله باغملک)

مجتبی شیرازی<sup>۱</sup>، عطاالله خادم الرسول<sup>۱\*</sup>، سید محمد صفی‌الدین اردبیلی<sup>۲</sup>

۱. گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران.

۲. گروه مهندسی بیوسیستم، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران.

(تاریخ دریافت: ۱۳۹۸/۱۱/۷ - تاریخ بازنگری: ۱۳۹۹/۲/۱ - تاریخ تصویب: ۱۳۹۹/۲/۱)

### چکیده

ارزیابی عوامل مؤثر بر کنترل فرسایش خاک در قالب شیوه‌های مدیریتی از اهمیت شایانی برخوردار است. در این پژوهش تأثیرات شیوه‌های مدیریتی غیر سازه‌ای شامل قرق و اصلاح پوشش گیاهی توسط مدل WEPP در حوزه آبخیز امامزاده عبدالله باغملک واقع در شمال شرقی استان خوزستان، شبیه‌سازی شد. بهینه‌سازی متغیرهای فیزیکی و هیدرولیکی مؤثر بر فرسایش شامل بافت خاک و اجزاء معادله ون‌گنوختن توسط روش‌های سطح پاسخ (RSM)، جنگل تصادفی (RF)، ماشین بردار پشتیبان (SVM) و شبکه‌ی عصبی مصنوعی (ANN) صورت پذیرفت. همچنین مقدار فرسایش خاک قبل از اعمال شیوه‌های مدیریتی به‌عنوان پاسخ اول ( $R_1$ ) و مقدار فرسایش خاک پس از اعمال شیوه‌های مدیریتی تحت عنوان پاسخ دوم ( $R_2$ ) تعریف شد. نتایج بهینه‌سازی توسط نرم‌افزار Orange شامل روش‌های جنگل تصادفی، ماشین بردار پشتیبان و شبکه عصبی مصنوعی نشان داد که روش جنگل تصادفی با  $MSE$ ،  $RMSE$  و  $R^2$  به ترتیب برابر  $۰/۹۶۳$  و  $۰/۹۹۵$ ،  $۰/۹۹۱$  و  $۰/۹۷۴$ ، به‌عنوان مناسب‌ترین روش است. همچنین بهینه‌سازی به روش سطح پاسخ با نتایج آماری  $MSE$ ،  $RMSE$  و  $R^2$  برای پاسخ اول به ترتیب  $۵/۳۷$ ،  $۲۸/۷$  و  $۰/۹۹۹$  و برای پاسخ دوم به ترتیب برابر  $۲/۰۳$ ،  $۴/۱۶$  و  $۰/۹۹۸$  مناسب‌ترین روش محسوب می‌شود. در مجموع استفاده از روش‌های بهینه‌سازی، ابزاری مناسب برای ارزیابی شیوه‌های مدیریتی و در نهایت انتخاب بهترین آن‌ها در مناطق بحرانی می‌باشد. متناسب بودن شیوه‌های مدیریتی بر پایه‌ی شرایط بهینه، منجر به کاهش هدررفت منابع آب و خاک می‌شود.

**واژه‌های کلیدی:** فرسایش خاک، متغیرهای فیزیکی و هیدرولیکی، نرم‌افزار RETC، شیوه‌های مدیریتی، متغیرهای مستقل و پاسخ.

### مقدمه

طبیعی، در نتیجه ضرورت دارد که از نرم‌افزارها، شیوه‌ها و روش‌های پیشرفته‌ای با قابلیت پایش فرایندهای پیچیده برای شبیه‌سازی رخدادهای فرسایشی و بهینه‌سازی آن‌ها در راستای اجرای سناریوهای مدیریتی پایدار در حوزه‌های آبخیز استفاده شود (Box and Draper, 2007; Tripathi et al., 2015). مقوله مدیریت حوزه‌های آبخیز اعم از بالادستی و پایین‌دستی از اهمیت بالایی در راستای تحقق اهداف آمایش سرزمینی برخوردار است (Nath and Lal, 2017). به همین دلیل ضرورت دارد که در ارتباط با مدل‌های برآورد فرسایش، از روش‌ها و شیوه‌های مختلفی استفاده شود که ضمن ساده‌سازی مدل‌های شبیه‌ساز بتوان با بهره‌گیری از ابزارها و شیوه‌های بهینه‌سازی گامی مؤثر را در راستای عملیاتی نمودن اهداف کلان مدیریتی و حفاظتی در عرصه‌های فرسایشی برداشت (Bagdi et al., 2015). با توجه به وسعت عرصه‌های فرسایشی ضروری است که از یک‌سری شیوه‌ها

فرسایش خاک یکی از بارزترین اشکال تخریب خاک است که طی فرایندی تدریجی و در بستر زمان منجر به هدررفت لایه سطحی خاک می‌شود و خاک را از مکان اصلی آن به مکانی دیگر منتقل می‌نماید؛ بدین ترتیب دارای مجموع اثرات درون‌حوزه‌ای و برون-حوزه‌ای می‌باشد (Brady and Weil, 2014; Lal, 2019; Wilson, 2019). همچنین فرایند فرسایش خاک می‌تواند تأثیرات محیطی و خسارت‌های اقتصادی جبران‌ناپذیری را بر اکوسیستم و شرایط مدیریتی حاکم بر آن وارد نماید که نتیجه‌ی آن تأثیر بر ویژگی‌های خاک منطقه و تولید بار رسوب است؛ لذا آگاهی داشتن از وضعیت فرسایش و رسوب برای اعمال مدیریت‌های نتیجه‌بخش در حوزه‌های آبخیز امری ضروری است (Nezhadafzali et al., 2019). به همین دلیل با توجه به اثرات گسترده و مخرب فرسایش خاک و پیچیدگی سازوکارهای حاکم بر آن در عرصه‌های منابع

حوضه‌ها فراهم می‌سازد. پژوهش انجام‌شده توسط Baigorria and Romero (2007) نشان می‌دهد که تلفیق سیستم‌های اطلاعات جغرافیایی و مدل فرسایشی WEPP امکان پهنه‌بندی مکان‌های مستعد فرسایش آبی را فراهم می‌نماید و در نتیجه ابزاری مدیریتی برای کنترل رخدادهای فرسایشی محسوب می‌شود. نتایج پژوهش‌های Naik and Samant (2016) بر روی همبستگی الگوریتم‌های طبقه‌بندی توسط نرم‌افزارهای مختلف داده‌کاوی و مقایسه‌ی آماری آن‌ها نشان داد که نرم‌افزار Orange از بیشترین میزان همبستگی با داده‌های واقعی برخوردار است، در نتیجه به‌عنوان بهترین ابزار داده‌کاوی معرفی شد. استفاده از الگوریتم جنگل تصادفی در داده‌های سنسجش از دور برای طبقه‌بندی پوشش اراضی از کارایی بالایی برخوردار است و می‌تواند به‌عنوان ابزاری مدیریتی برای مدیریت پایدار اراضی استفاده شود (Gislason *et al.*, 2004). همچنین در مقایسه صورت‌گرفته بین الگوریتم‌های مختلف داده‌کاوی شامل رگرسیون لجستیک، درخت تصمیم‌گیری و جنگل تصادفی توسط Booker and Snelder (2012)، نتایج نشان داد که در بین الگوریتم‌های ارزیابی‌شده، روش جنگل تصادفی از بهترین عملکرد برخوردار است و در نتیجه می‌توان از این روش برای ارائه‌ی راهکارهای مدیریتی استفاده نمود.

در پژوهشی Heung *et al.* (2014) با هدف طبقه‌بندی نقشه‌های رقومی از مدل یادگیری ماشین و الگوریتم جنگل تصادفی استفاده نمودند و نتایج حاصله حاکی از این بود که الگوریتم جنگل تصادفی به دلیل سرعت در پردازش‌های متغیری و سهولت در تفسیر خروجی‌ها به‌عنوان مناسب‌ترین الگوریتم برای تحقق اهداف طبقه‌بندی عوامل توپوگرافیکی محسوب می‌شود. در پژوهش دیگری توسط Yoon *et al.* (2011) برای پیش‌بینی تغییرات سطح آب زیرزمینی از دو روش شبکه عصبی مصنوعی و ماشین بردار پشتیبان استفاده شد که نتایج نشان داد، روش ماشین بردار پشتیبان ابزاری کارآمد برای پیش‌بینی تغییرات سطح آب زیرزمینی می‌باشد. در پژوهشی با استفاده از روش سطح پاسخ، بهینه‌سازی متغیرهای هیدرولیکی در ارتباط با فرسایش خاک انجام شد و نتایج نشان داد که روش سطح پاسخ از کارایی بالایی برای تعیین حدود بهینه‌ی متغیرهای مؤثر بر فرسایش خاک برخوردار است. بدین ترتیب می‌توان از این روش برای ارزیابی عوامل مؤثر بر فرسایش خاک استفاده نمود و حدود بهینه آن‌ها را مشخص نمود (Shirazi *et al.*, 2019). در پژوهشی دیگر،

که به نوعی نواقص و کاستی‌های موجود در مدل‌های شبیه‌ساز را برطرف می‌نمایند، استفاده شود. از جمله این شیوه‌ها، می‌توان به استفاده از روش‌های بهینه‌سازی اشاره نمود که استفاده از این روش‌ها به انتخاب بهترین شیوه‌های مدیریتی در حوزه‌های آبخیز و عرصه‌های فرسایشی کمک می‌نماید و نتیجه‌ی نهایی آن کاهش میزان رخدادهای فرسایش خاک و تولید بار رسوب می‌باشد. از جمله این روش‌ها می‌توان به روش‌های سطح پاسخ (RSM)، ماشین بردار پشتیبان (SVM)، جنگل تصادفی (RF) و شبکه‌ی عصبی مصنوعی (ANN) برای بهینه‌سازی و ارزیابی تأثیرگذاری عوامل مختلف بر مقادیر فرسایش اشاره نمود.

روش RSM، یکی از پرکاربردترین روش‌های بهینه‌سازی است که ارتباط بین چند متغیر مستقل با متغیر پاسخ را بررسی نموده و جزئیات فرایند را به‌صورت روابط ریاضی معین نمایش می‌دهد (Box and Draper, 2007). مهمترین ویژگی روش RSM برای ارزیابی متغیرهای چندگانه و روابط متقابل آن‌ها، کاهش تعداد آزمایش‌ها و در نتیجه کاهش هزینه‌ها است. در روش RSM، متغیر وابسته به‌عنوان پاسخ و متغیر مستقل به‌عنوان عامل‌های رابطه در نظر گرفته شده و در نهایت یک رابطه ریاضی بین  $x_i$ ها (که متغیرهای مستقل هستند) و  $y$ ها ارائه می‌دهد. با استفاده از این ضرایب می‌توان رابطه مربوط به فرایند پیش‌بینی را تکمیل کرد؛ به‌گونه‌ای که بین داده‌های واقعی و داده‌هایی که این رابطه ریاضی پیش‌بینی می‌کند، بهترین حالت ممکن باشد و در نهایت در راستای اهداف مدیریتی استفاده شود (Raymond *et al.*, 2009). استفاده از شیوه‌های داده‌کاوی نیز می‌تواند کمک شایانی به ارزیابی وضعیت فرسایش و رسوب در حوزه‌های آبخیز بنماید. از جمله نرم‌افزارهایی که قابلیت انجام روش‌های داده‌کاوی مذکور را داراست، نرم‌افزار Orange می‌باشد که کدنویسی آن بر اساس زبان پایتون صورت پذیرفته است (Naik and Samant, 2016). نرم‌افزار Orange در سال ۱۹۹۷ توسط Demsar and Zupan توسعه یافت. این نرم‌افزار یک کتابخانه پایتون و دارای محیطی با برنامه‌نویسی بصری است که بخش خط فرمان آن نیاز به کدنویسی نداشته و از طریق اسکریپت‌های موجود در خود نرم‌افزار مجموعه‌ای غنی از الگوریتم‌های استخراج، واکاوی و یادگیری ماشینی را برای پیش‌پردازش، طبقه‌بندی، مدل‌سازی و خوشه‌بندی داده‌ها ارائه می‌کند (Demšar and Zupan, 2013).

استفاده‌های ترکیبی و تلفیقی از نرم‌افزارهای شبیه‌سازی و بهینه‌سازی امکان انتخاب و اجرای راهکارهای مدیریتی را در

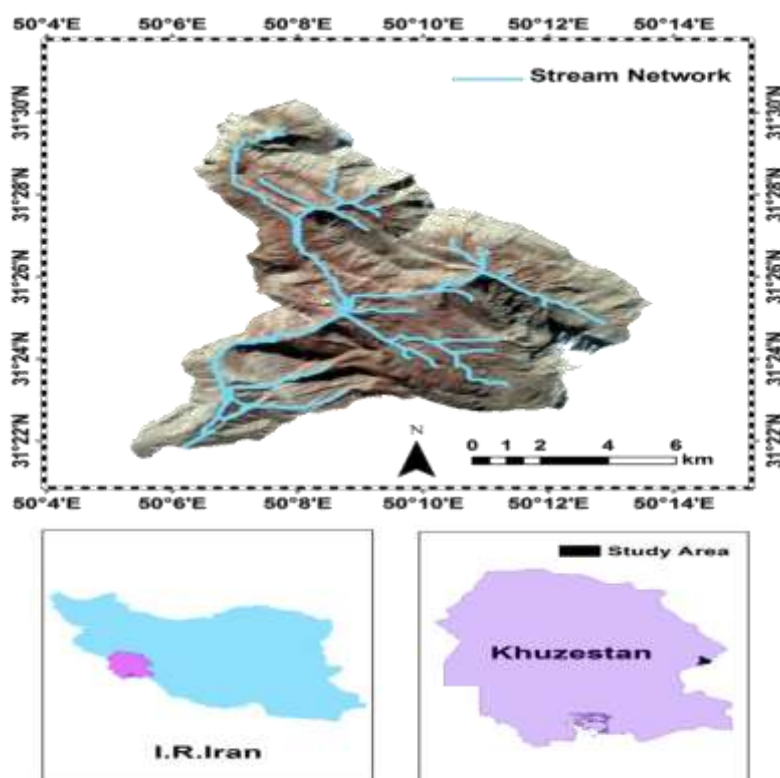
## مواد و روش‌ها

### معرفی منطقه مطالعاتی

منطقه‌ی مورد مطالعه حوزه آبخیز امامزاده عبدالله با وسعت تقریبی ۱۰۴ کیلومترمربع واقع در شمال شرقی شهرستان باغملک است (شکل ۱). محدوده تقریبی جغرافیایی حوزه مورد مطالعه بین ۳۱°۱۸' تا ۳۱°۳۳' شمالی و نیز ۵۰°۵' تا ۵۰°۱۳' شرقی می‌باشد. این حوزه دارای رژیم رطوبتی یوستیک و رژیم حرارتی ترمیک است. خاک‌های منطقه مطالعاتی بر اساس سیستم طبقه‌بندی خاک آمریکایی در رده‌های آنتی‌سول و اینسپتی‌سول قرار دارند و بافت غالب این خاک‌ها لوم رسی سیلتی می‌باشد. همچنین مهمترین کاربری‌های اراضی منطقه از نوع جنگل و ترکیب جنگل-مرتع است که با توجه به بهره‌برداری نامناسب اعم از جنگل‌تراشی و نیز چرای مفرط دام، در معرض تخریب می‌باشند و ضرورت دارد شیوه‌های مدیریتی مطلوب و متناسبی برای آنها اتخاذ شود.

نوروزی و همکاران با استفاده از روش جنگل تصادفی مناطق آسیب‌پذیر به نیترات در دشت ملکان را ارزیابی نموده و با استفاده از آماره‌ی MSE بهترین مدل را برای پیش‌بینی مناطق بحرانی تعیین نمودند (Nouruzi et al., 2015).

تعیین حدود بهینه به مفهوم اتخاذ و اعمال شیوه‌های مدیریتی کارآمد در عرصه‌های مطالعاتی است، که حاصل آن بالا رفتن بازدهی راه‌کارها و سناریوهای مدیریتی در مناطق بحرانی می‌باشد. در مجموع پژوهش‌های انجام‌شده با محوریت استفاده از روش سطح پاسخ، قابلیت‌ها و ظرفیت‌های این روش را به منظور مدیریت بهینه عوامل مؤثر بر یک فرایند نشان می‌دهد و با توجه به تأثیرات معنی‌دار این روش در کاهش هزینه‌ها، در نتیجه می‌تواند به‌عنوان یک ابزار مدیریتی مورد توجه و استفاده قرار گیرد. با توجه مطالب فوق‌الذکر و جایگاه استفاده از شیوه‌های مدیریتی، پژوهش حاضر با هدف ارزیابی روش‌های سطح پاسخ، ماشین بردار پشتیبان و جنگل تصادفی در راستای شرایط موجود در حوزه آبخیز مطالعاتی و نیز تأثیرات راهکارهای مدیریتی شبیه‌سازی شده مشتمل بر قرق و اصلاح پوشش گیاهی صورت پذیرفت.



شکل ۱- موقعیت جغرافیایی منطقه مطالعاتی (حوزه آبخیز امامزاده عبدالله باغملک)

صفر تا ۲۰ سانتی‌متری، عملیات نمونه‌برداری در هر سایت نمونه-برداری با برداشت خاک از ۵ نقطه و سپس ترکیب نمونه‌ها صورت گرفت تا نمونه‌هایی که نماینده‌ی خاک منطقه می‌باشند، تهیه شوند و در دمای ۲۰ درجه سانتی‌گراد در آزمایشگاه هوا خشک

نمونه‌برداری خاک و اندازه‌گیری متغیرهای مدل شبیه‌ساز فرسایش

برای انجام این پژوهش و اندازه‌گیری هر کدام از متغیرهای مستقل (متغیرهای ورودی)، در حوزه آبخیز مطالعاتی از عمق

دامنه‌ها در ارتباط با شبکه‌ی هیدروگرافی حوضه‌ی مطالعاتی مشخص و در نهایت مدل اجرا شد. واسنجی و اعتبارسنجی مدل با بهره‌گیری از مقادیر اندازه‌گیری شده برای تمامی لایه‌های اطلاعاتی مدل WEPP صورت پذیرفت. قابل ذکر است که حوزه آبخیز مطالعاتی دارای ۶ پارسل است که در هر پارسل، دو دامنه واقع در جبهه راست و جبهه‌ی چپ آبراهه تعریف شد که این دامنه‌ها در برگیرنده لایه‌های اطلاعاتی ورودی مدل اعم از خاک، پوشش گیاهی و مدیریت، توپوگرافی و اقلیم می‌باشند. همان‌گونه که جدول (۱) نشان می‌دهد، علاوه بر متغیرهای مستقل که مشتمل بر متغیرهای فیزیکی و هیدرولیکی اندازه‌گیری شده هستند، دسته دیگر، متغیرهای وابسته محسوب می‌شوند که با نماد  $R_1$ ، یعنی قبل از اعمال شیوه‌های مدیریتی و  $R_2$ ، به مفهوم پس از اعمال راهکارهای مدیریتی است، نمایش داده شده‌اند. شیوه‌های مدیریتی ارزیابی شده از نوع غیر سازه‌ای می‌باشند و عبارت از اجرای قرق و اصلاح پوشش گیاهی هستند.

همچنین در این پژوهش عملیات نرمال‌سازی داده‌ها با بهره‌گیری از نرم‌افزار اکسل صورت پذیرفت و داده‌ها در دامنه‌ی بین ۱ تا ۱- قرار داده شدند. پس از نرمال‌سازی داده‌ها، فرایندهای بهینه‌سازی داده‌ها انجام شد. دیگرام شبیه‌سازی و بهینه‌سازی در شکل (۲) نمایش داده شده است.

### روش سطح پاسخ

در روش سطح پاسخ، ورودی‌ها (شامل ویژگی‌های فیزیکی و هیدرولیکی خاک) در قالب متغیرهای مستقل  $(x_1, x_2, \dots, x_k)$  و وابسته یا پاسخ  $(Y)$  (رابطه ۱) در دو سطح +۱ و -۱ تعریف شد (جدول ۱). عملکرد نرم‌افزار بدین صورت است که بین متغیرهای مستقل و وابسته یک سری روابط ریاضی برقرار می‌نماید که در نهایت این روابط را به صورت نمودارهای دو بعدی و سه بعدی، خطوط تراز، تک‌متغیر، اثر متقابل و یک بخش مختص به آنالیز واریانس (ANOVA) ارائه می‌دهد (Oehlert, 2000).

$$y = f(x_1, x_2, \dots, x_k) + \varepsilon \quad (\text{رابطه ۱})$$

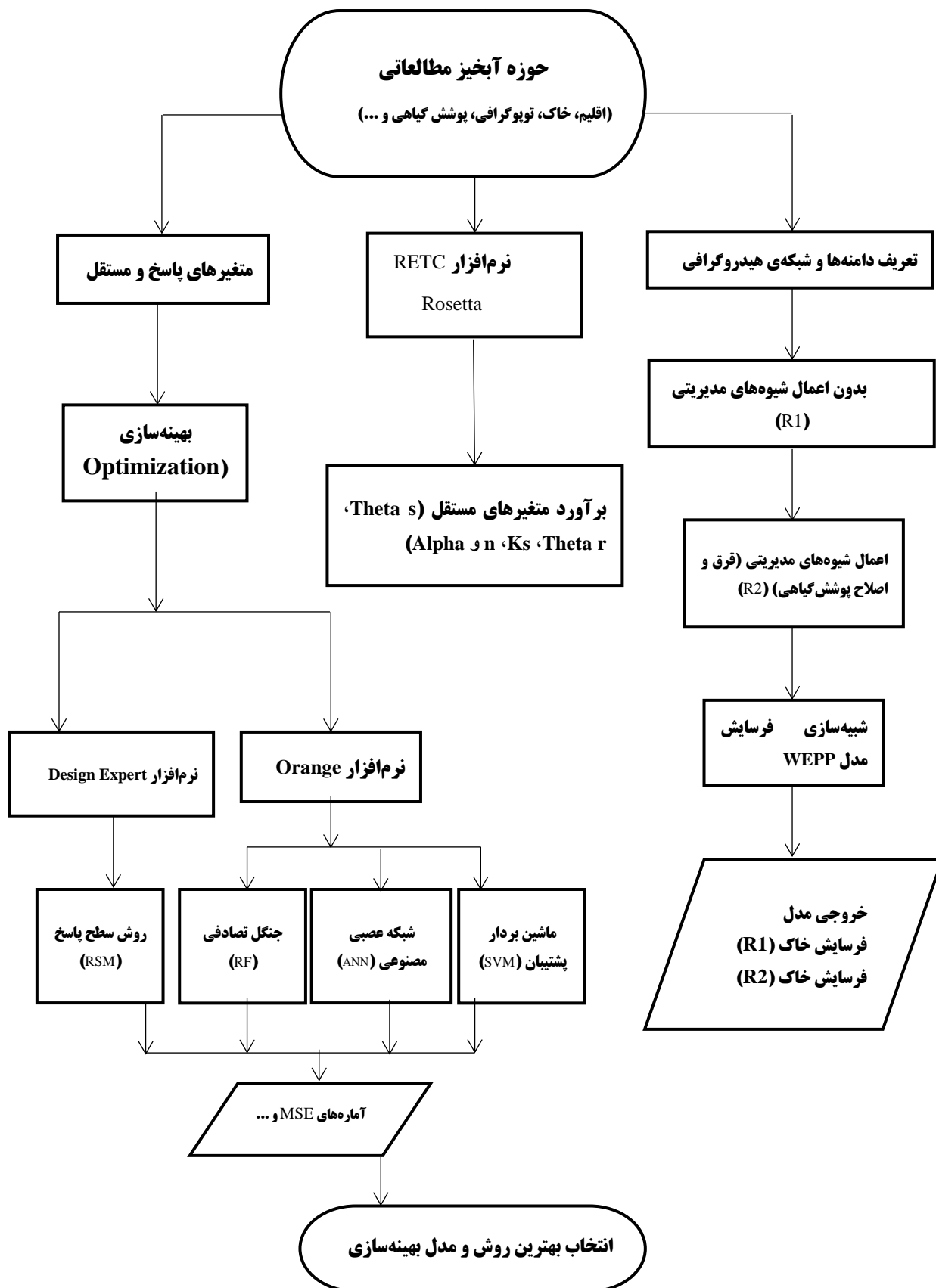
در این رابطه  $f$  تابع عملکرد،  $\varepsilon$  متغیر خطا،  $y$  متغیر وابسته یا پاسخ و  $x$  متغیر مستقل است.

در این پژوهش طرح‌های آزمایشی مختلفی شامل Box-Behnken, Central Composite, One factor, D-optimal, Split plot, Historical data و Supersaturated و مدل‌های متنوعی نظیر Mean, Linear, 2FI, Quadratic و Cubic برای بهینه‌سازی مورد ارزیابی قرار گرفت. در نهایت بر اساس آماره‌های به‌دست‌آمده، بهترین طرح آزمایشی و مدل انتخاب شد.

شدند تا متغیرهای فیزیکی و شیمیایی خاک مورد مطالعه با استفاده از روش‌های استاندارد اندازه‌گیری شود.

به منظور اندازه‌گیری درصد سنگریزه سطحی در قسمت‌های مختلف و به‌صورت پراکنده بروی سطح زمین پلات‌هایی با ابعاد ۱ متر در ۱ متر ایجاد شد و مقدار سنگ و سنگریزه‌های موجود در پلات جمع‌آوری و توزین شد و میانگین وزنی آن به‌عنوان سنگریزه سطحی آن ناحیه محاسبه شد. ظرفیت تبادل کاتیونی (CEC) به روش استات آمونیوم (Thomas, 1982)، درصد ماده آلی به روش والکی-بلک (Walkley and Black, 1934) و بافت خاک به روش هیدرومتری (Bouyucos, 1962) اندازه‌گیری شد. همچنین درصد اشباع شاخصی از بافت خاک است که از طریق اختلاف وزن حالت اشباع و خشک محاسبه شد. ضریب آلبو متغیری است که بیانگر وضعیت بازتاب تابش خورشید توسط سطوح است و با استفاده از رابطه‌ی بومر (Baumer, 1990) که ضریب بازتاب یا آلبو خاک را متأثر از میزان مواد آلی خاک در نظر می‌گیرد، محاسبه شد. همچنین متغیرهای  $\alpha$ ,  $\theta_s$ ,  $\theta_r$ ,  $n$  و  $K_s$  که جزء متغیرهای دیریافت خاک محسوب می‌شوند، با استفاده از بسته نرم‌افزاری RETC و بخش Rosetta و بهره‌گیری از متغیرهای زودریافت برآورد شدند. متغیرهای  $\alpha$ ,  $\theta_s$ ,  $\theta_r$  و  $K_s$  همگی متغیرهای معادله ون‌گونوختن و جزء متغیرهای دیریافت هستند. از آنجایی که به دست آوردن متغیرهای دیریافت کاری دشوار و زمان‌بر است لذا با استفاده از دو بسته نرم‌افزاری RETC و Rosetta هر کدام از این متغیرها محاسبه شد. بسته‌های نرم‌افزاری RETC و Rosetta برای ارزیابی و برآورد ویژگی‌های هیدرولیکی خاک استفاده می‌شود و اساس کار آن بدین صورت است که بر اساس شبکه‌های عصبی مصنوعی ویژگی‌های هیدرولیکی خاک را تخمین می‌زند. در این پژوهش برای دستیابی به متغیرهای معادله ون‌گونوختن که عبارت از:  $\theta_r$  (رطوبت باقی‌مانده)،  $\theta_s$  (رطوبت اشباع)،  $\alpha$  و  $n$  (ضرایب شکل) و  $K_s$  (هدایت هیدرولیکی اشباع) می‌باشند، از مدل SSC استفاده شد.

مدل SSC از نوع توابع انتقالی شبکه عصبی مصنوعی است که متغیرهای هیدرولیکی مورد نظر را بر اساس دیگر متغیرهای زودریافت همچون درصد شن، سیلت و رس برآورد می‌نماید. همچنین با استفاده از مدل شبیه‌ساز WEPP<sup>۱</sup> و متغیرهای فیزیکی و هیدرولیکی، فرسایش خاک شبیه‌سازی و مقدار فرسایش خاک با اعمال راهکارهای مدیریتی پیش‌بینی شد. شایان ذکر است که در مجموع لایه‌های اطلاعاتی شامل اقلیم، خاک، پوشش گیاهی و توپوگرافی برای هر دامنه تعریف شد و این



شکل ۲- نمودار روندنمای شبیه‌سازی فرسایش و بهینه‌سازی عوامل فیزیکی و هیدرولیکی مؤثر بر فرسایش خاک

جدول ۱- مقادیر بیشینه و کمینه متغیرهای مستقل (فیزیکی و هیدرولیکی) و وابسته (سطوح پاسخ)

مقادیر کدبندی شده		نوع متغیر	نوع متغیر
Min (-1)	Max (+1)		
۶	۶۴	Sand%	فیزیکی
۱۴	۷۶	Silt%	
۶	۴۶	Clay%	
۳۱	۶۸	SP%	
۰/۰۹	۵/۲	Stone%	
۰/۰۳۳۹	۰/۰۹۷۴	Theta r	هیدرولیکی
۰/۳۸۵۶	۰/۴۸۷۴	Theta s	
۰/۰۰۵۲	۰/۰۲۹۱	Alpha	
۱/۲۷۶۷	۱/۶۷۹۹	N	
۶/۸۴	۴۶/۰۱	K <sub>S</sub> (cm/day)	
۱۴/۵۹	۲۶/۸۹	R <sub>1</sub>	وابسته
۶/۰۱	۱/۴۸	R <sub>2</sub>	

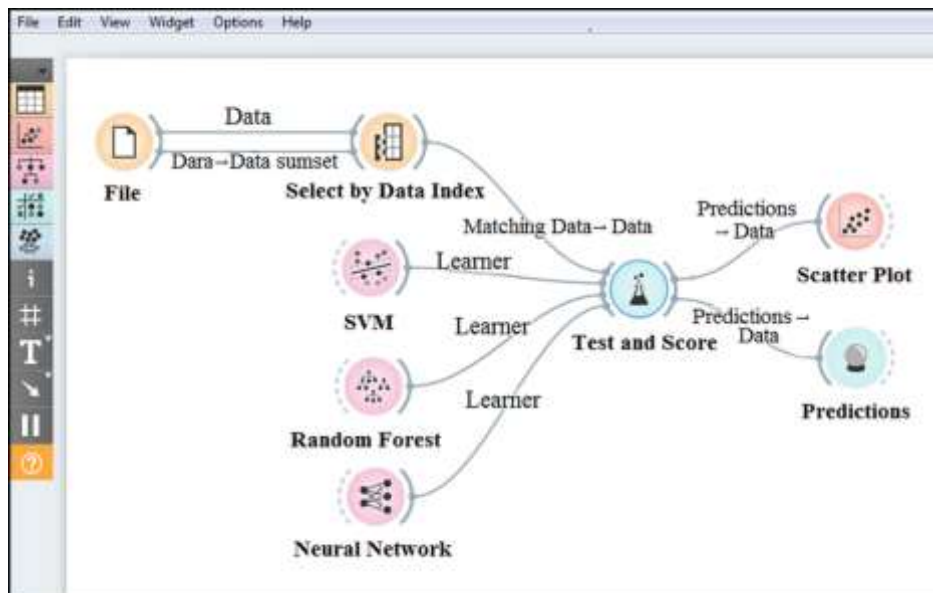
R<sub>1</sub> (پاسخ اول): مقدار فرسایش خاک قبل از اعمال سناریوهای مدیریتی

R<sub>2</sub> (پاسخ دوم): مقدار فرسایش خاک بعد از اعمال سناریوهای مدیریتی

گام بعدی با استفاده از گزینه Test and Score، میزان داده‌های آموزش و آزمون مشخص شد و همزمان پس از وارد نمودن اطلاعات لازم، فرایند شبیه‌سازی و بهینه‌سازی صورت گرفت و نتایج آن در دو بخش گزینه‌های Scatter Plot و Predictions ثبت شد (شکل ۳).

### بهینه‌سازی توسط نرم‌افزار Orange

در این بخش ابتدا داده‌ها در قالب یک فایل اکسل در نرم‌افزار بارگذاری شد، سپس برنامه‌نویسی بصری مدل‌های مورد ارزیابی شامل جنگل تصادفی، ماشین بردار پشتیبان و شبکه عصبی مصنوعی با استفاده از فن کشیدن و رهاکردن<sup>۱</sup> صورت گرفت. در



شکل ۳- برنامه‌نویسی بصری توسط روش‌های ماشین بردار پشتیبان (SVM)، جنگل تصادفی (RF) و شبکه‌ی عصبی مصنوعی (ANN) در نرم‌افزار Orange

مجموعه‌ای از درخت‌های تصمیم‌گیری است که با جایگزین کردن و تغییر مداوم متغیرهای مرتبط با هدف، سبب ایجاد تعداد زیادی درخت تصمیم‌گیری شده که در نهایت برای پیش‌بینی با یکدیگر

جنگل تصادفی، ماشین بردار پشتیبان و شبکه عصبی الگوریتم جنگل تصادفی در فرایندهای طبقه‌بندی جزء طبقه‌بندی نظارت‌شده و نامتغیر یک محسوب می‌شود و شامل

متغیرهای گزینه تعریف شده در این پژوهش در جدول (۴) ارائه شده است.

جدول ۳- متغیرهای گزینه تعریف شده در روش ماشین بردار پشتیبان

نوع متغیر	نوع/ مقدار
تابع کرنل	خطی
ارزش	۱/۰۰
رگرسیون اپسیلون	۰/۱
تحمل عددی	۰/۰۰۱
حد مجاز	۱۰۰

جدول ۴- متغیرهای گزینه تعریف شده در روش شبکه عصبی

نوع متغیر	نوع/ مقدار
تابع عملکرد	Logistic
تابع هدف	SGD
تکرار	۲۰
ضریب مقررات	۱
تعداد نورون‌ها در لایه پنهان	۱۰۰۰
حداکثر تکرارها	۲۰۰۰

### تجزیه و تحلیل آماری

به منظور ارزیابی و مقایسه عملکرد روش‌های سطح پاسخ، جنگل تصادفی، ماشین بردار پشتیبان و شبکه عصبی مصنوعی از شاخص‌های آماری شامل ضریب تعیین ( $R^2$ )، ضریب همبستگی ( $r$ )، جذر میانگین مربعات خطا (RMSE) و میانگین مربعات خطا (MSE) استفاده شد. شاخص ضریب تعیین که مقدار آن بین ۰ و ۱ می‌باشد، نشان‌دهنده مقدار تغییرات متغیر وابسته تحت تأثیر متغیر مستقل است. شاخص ضریب همبستگی (بین ۱ و -۱)، نشان‌دهنده همبستگی متغیر وابسته به متغیرهای مستقل است. اگر مقدار این متغیر برابر با صفر باشد، نشان‌دهنده عدم وجود همبستگی است ولی در مقابل نزدیک بودن مقدار این شاخص به عدد ۱ نشان‌دهنده وجود همبستگی مثبت و به عدد -۱ بیانگر وجود همبستگی منفی بین دو متغیر است. شاخص جذر میانگین مربعات خطا، تفاوت بین مقادیر پیش‌بینی شده با مقادیر واقعی را نشان می‌دهد و هر چه به صفر نزدیک‌تر باشد، نتایج قابل قبول‌تری را خواهد داشت (Parsafar and Marofi, 2011). طریقه محاسبه هر کدام از این شاخص‌ها به شرح رابطه (۶) تا (۸) می‌باشد.

$$R^2 = \frac{\left( \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \right)^2}{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2} \quad \text{(رابطه ۶)}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - X_i)^2} \quad \text{(رابطه ۷)}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i - Y_i)^2 \quad \text{(رابطه ۸)}$$

ترکیب می‌شوند (Vorpahl et al., 2012; Chen et al., 2018). پس از برنامه‌نویسی بصری برای روش جنگل تصادفی، متغیرهای فیزیکی و هیدرولیکی خاک به‌عنوان ورودی به مدل داده شد و متغیرهای گزینه تعریف شد (جدول ۲). در این الگوریتم، ۶۶ درصد داده‌ها برای آموزش و ۳۳ درصد از داده‌های باقی‌مانده نیز برای آزمون استفاده شد (Breiman, 2001; Partopour et al., 2018).

جدول ۲- متغیرهای گزینه تعریف شده در روش جنگل تصادفی

نوع متغیر	مقدار
تعداد درخت	۱۰
تعداد مؤلفه‌های در نظر گرفته در هر تقسیم	۵
محدوده عمق هر تک درخت	۳
محدوده عدم تقسیم زیرمجموعه‌ها	۲

الگوریتم SVM جزء الگوریتم‌های تشخیص الگو به شمار می‌آید و از آن می‌توان در هر پروژه‌ای که نیاز به تشخیص الگو یا دسته‌بندی در کلاس‌های خاص باشد، استفاده کرد. این الگوریتم یکی از روش‌های یادگیری با نظارت است که کاربران از آن در طبقه‌بندی و رگرسیون‌ها استفاده می‌کنند (Cortes and Vapnik, 1995; Sun et al., 2019). از مهمترین ویژگی‌های SVM می‌توان به طبقه‌بندی داده‌ها بر اساس حداقل خطای آزمایش اشاره کرد (AzimiPour et al., 2020). مفهوم اصلی رگرسیون بردار پشتیبانی ابتدا با یک مدل خطی مورد بحث قرار می‌گیرد و سپس با انتخاب یکی از توابع کرنل (Linear, RBF, Polynomial و Sigmoid) موجود در گزینه‌ها، SVM به یک مدل غیرخطی تبدیل می‌شود که روند محاسبه هر کدام از این توابع به شرح زیر می‌باشد (Vapnik, 2013):

Linear:  $k_{(x_i, x)} = x_i \cdot x$  (رابطه ۲)

Polynomial:  $k_{(x_i, x)} = (\gamma x_i \cdot x + C)^d$  (رابطه ۳)

RBF:  $k_{(x_i, x)} = \exp(-\gamma |x_i - x|^2)$  (رابطه ۴)

Sigmoid:  $k_{(x_i, x)} = \tan(\gamma x_i \cdot x + C)$  (رابطه ۵)

در این روابط  $x_i$  و  $x$  به ترتیب الگوهای آموزش و آزمون، d ابعادی از بردار ورودی است.

برای انجام فرایند پردازش در ابتدا مقدار هر کدام از متغیرهای Cost(c) و Regression loss epsilon ( $\epsilon$ )، Numerical tolerance، Iteration limit و نوع تابع کرنل مشخص و تنظیم شد (جدول ۳).

شبکه عصبی مصنوعی جزء روش‌های محاسباتی و سامانه‌های پویا است که در علم یادگیری ماشین و سایر علوم مرتبط کاربرد فراوانی دارد و متشکل از لایه‌ی ورودی، میانی (مخفی) و خروجی است (Podvalny and Vasiljev, 2017).



( $\delta$ ) نشان می‌دهد، روش سطح پاسخ دارای بالاترین  $R^2$  می‌باشد، لذا این روش می‌تواند به‌عنوان مناسب‌ترین روش برای بهینه‌سازی متغیرهای مؤثر بر فرسایش خاک مورد استفاده قرار گیرد. همچنین بین روش‌های محاسباتی در نرم‌افزار Orange، روش جنگل تصادفی دارای بالاترین  $R^2$  است، در نتیجه بین روش‌های ماشین بردار پشتیبان، جنگل تصادفی و شبکه عصبی مصنوعی، روش جنگل تصادفی به‌عنوان بهترین روش محسوب می‌شود.

در تمامی روابط بالا  $X_i$  و  $Y_i$  به ترتیب نشان‌دهنده  $i$  امین داده واقعی و برآوردشده،  $\bar{X}$  و  $\bar{Y}$  به ترتیب نشان‌دهنده میانگین داده‌های  $X_i$  و  $Y_i$  و  $n$  تعداد نمونه‌ها است.

## نتایج و بحث

شاخص‌های آماری مدل‌های ارزیابی‌شده شامل سطح پاسخ، ماشین بردار پشتیبان، جنگل تصادفی و شبکه‌ی عصبی مصنوعی ارائه شده است (جدول ۵). همان‌گونه که نتایج مندرج در جدول

جدول ۵- نتایج تحلیل آماری مدل‌ها برای  $R_1$  (پاسخ اول: مقدار فرسایش خاک قبل از اعمال مدیریت) و  $R_2$  (پاسخ دوم: مقدار فرسایش خاک بعد از اعمال مدیریت)

شاخص‌های آماری						روش بهینه‌سازی
$R^2$		RMSE		MSE		
$R_2$	$R_1$	$R_2$	$R_1$	$R_2$	$R_1$	
۰/۹۹۸	۰/۹۹۹	۲/۰۳	۵/۳۷	۴/۱۶	۲۸/۷	RSM
۰/۲۵۸	۰/۴۹۳	۱/۶۳۸	۳/۶۶۶	۲/۶۸۴	۱۴/۴۳۷	SVM
۰/۹۷۴	۰/۹۶۳	۰/۳۰۷	۰/۹۹۵	۰/۰۹۵	۰/۹۹۱	Random Forest
۰/۴۳۸	۰/۵۷۵	۱/۴۲۵	۳/۳۵۹	۲/۰۳۲	۱۱/۲۸۶	Neural Network

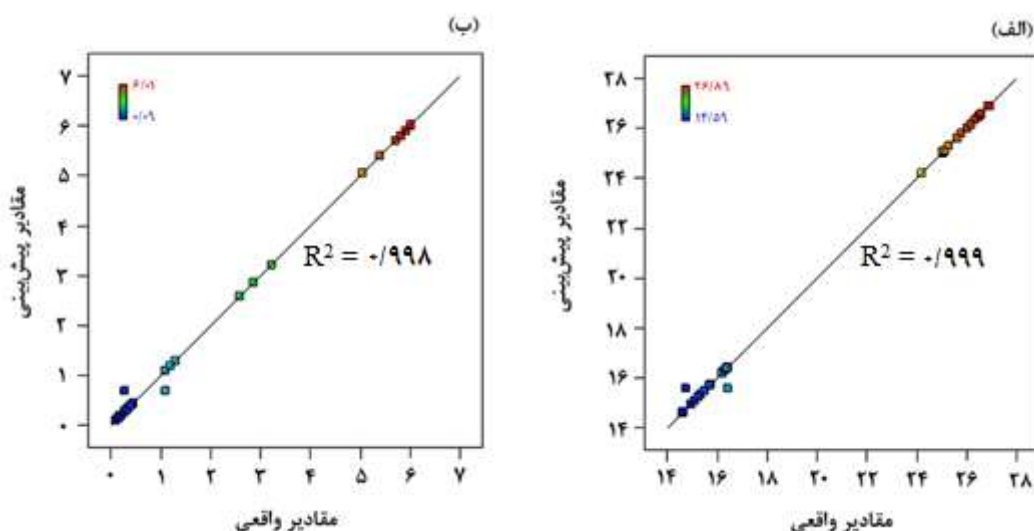
نمودارهای خروجی است و یکی از آن‌ها نمودار همبستگی است، که نقاط رنگی موجود در این نمودار نشان‌دهنده مقادیر واقعی و پیش‌بینی است (شکل ۴-الف و ب). همان‌گونه که شکل (۴-الف) نشان می‌دهد، دامنه تغییرات فرسایش خاک قبل از اعمال راهکارهای مدیریتی شامل قرق و اصلاح پوشش‌های گیاهی، بین ۱۴/۵۹ تن در هکتار در سال تا ۲۶/۸۹ تن در هکتار در سال بوده است. اعمال سناریوهای مدیریتی در دامنه‌ها سبب شده است تا اثرات فرایندهای تخریبی خاک با گذشت زمان کاهش یابند و همین امر زمینه را برای بهبود و احیاء خاک فراهم نموده است. تأثیرات اعمال شیوه‌های مدیریتی در قالب راه‌کارهای مدیریتی غیر سازه‌ای در شکل (۴-ب) نمایان است که مقدار فرسایش در دامنه‌ی ۰/۰۹ تا ۶/۰۱ تن در هکتار در سال قرار گرفته است. این کاهش معنی‌دار به خوبی تأثیرات اعمال شیوه‌های مدیریتی مطلوب و متناسب را تأیید می‌نماید (Eggers et al., 2017; Andersen et al., 2019). همان‌گونه که نتایج این پژوهش نیز نشان می‌دهد، اتخاذ و به‌کارگیری شیوه‌های مدیریتی مناسب و متناسب با شرایط منطقه منجر به کاهش معنی‌دار رخدادهای فرسایش خاک می‌شود. در واقع با توجه به توپوگرافی پیچیده حاکم بر منطقه و عدم تناسب پوشش‌های گیاهی موجود در زیر حوزه‌های مطالعاتی، در نتیجه اصلاح پوشش‌های گیاهی منجر به کاهش معنی‌دار هدررفت خاک شده است.

## مدل‌سازی با استفاده از روش سطح پاسخ

بر اساس تجزیه‌های انجام‌شده و خروجی‌های حاصله از تمامی مدل‌های ارزیابی‌شده و توابع انتقالی موجود در بانک اطلاعاتی نرم‌افزار Design Expert، در نهایت طرح Box-Behnken و مدل Quadratic به‌عنوان مناسب‌ترین طرح و مدل برای هر دو پاسخ  $R_1$  و  $R_2$  انتخاب شد. همچنین نتایج آماری نشان می‌دهد که مقدار شاخص میانگین مربعات خطا برای پاسخ اول و دوم به ترتیب ۲۸/۷ و ۴/۱۶ و شاخص جذر میانگین مربعات خطا برای پاسخ اول و دوم به ترتیب برابر با ۵/۳۷ و ۲/۰۳ است که در مقایسه با سایر روش‌های ارزیابی‌شده، دارای میانگین مربعات خطا و جذر میانگین مربعات خطای بالاتری است. مهم‌ترین دلیل این اختلاف در روش سطح پاسخ با سایر روش‌های ارزیابی‌شده این است، که در روش سطح پاسخ روند پردازش بدین صورت است که تک‌تک متغیرها را به دو صورت تک عامل و اثر متقابل مورد تجزیه و تحلیل قرار می‌دهد و این مقدار از MSE و RMSE برابر مجموع تمامی حالت‌ها یا به عبارتی برابر مدل (Quadratic) می‌باشد و همین امر سبب می‌شود تا مقدار بالایی داشته باشد. ضریب تعیین ( $R^2$ ) دیگر متغیر آماری است که برای پاسخ اول و دوم به ترتیب برابر ۰/۹۹۹ و ۰/۹۹۸ است (جدول ۲). شاخص همبستگی دیگر شاخصی است که با استفاده از آن می‌توان تشخیص داد که بین مقادیر واقعی و پیش‌بینی‌شده چه مقدار تفاوت وجود دارد. نرم‌افزار Design Expert دارای یک بخش گرافیکی مختص به

بالایی وجود دارد. همچنین قبل از اعمال شیوه‌های مدیریتی میزان فرسایش خاک از مقادیر بالایی برخوردار است که ضرورت اعمال شیوه‌های مدیریتی را تأیید می‌نماید. شاخص‌های آماری روش جنگل تصادفی برای پاسخ اول برای متغیرهای MSE، RMSE و  $R^2$  به ترتیب برابر ۰/۹۹۵، ۰/۹۶۳ و همچنین برای پاسخ دوم به ترتیب برابر ۰/۳۰۷، ۰/۹۷۴ است (جدول ۲). بعلاوه نمودار همبستگی برای پاسخ اول و دوم به خوبی نشان می‌دهد که بین متغیرهای مستقل و وابسته همبستگی بالایی وجود دارد و ضریب همبستگی برای هر دو پاسخ برابر ۰/۹۹ می‌باشد (شکل ۵ و ۶).

مدل‌سازی با استفاده از الگوریتم جنگل تصادفی شایان ذکر است که برازش داده‌های دارای اختلال<sup>۱</sup> در روش جنگل تصادفی در مقایسه با روش‌های ماشین بردار پشتیبان و شبکه‌ی عصبی مصنوعی ساده‌تر و مناسب‌تر است (Norouzi et al., 2019). فرایند پیش‌بینی با استفاده از ترکیب نمودن متغیرها صورت گرفت و توسط شاخص‌های آماری MSE، RMSE و  $R^2$  ارزیابی شد و مدل بهینه بر اساس مقایسه شاخص‌های آماری، مشخص شد. همان‌گونه که شکل (۵) نشان می‌دهد، بین مقادیر فرسایش اندازه‌گیری شده و برآوردی در شرایط پاسخ اول که نمایانگر عدم اعمال راهکارهای مدیریتی می‌باشد، همبستگی



شکل ۴- همبستگی بین مقادیر واقعی و پیش‌بینی شده فرسایش خاک با استفاده از روش سطح پاسخ (الف) قبل از اعمال شیوه‌های مدیریتی ( $R_1$ ) و (ب) پس از اعمال شیوه‌های مدیریتی ( $R_2$ )



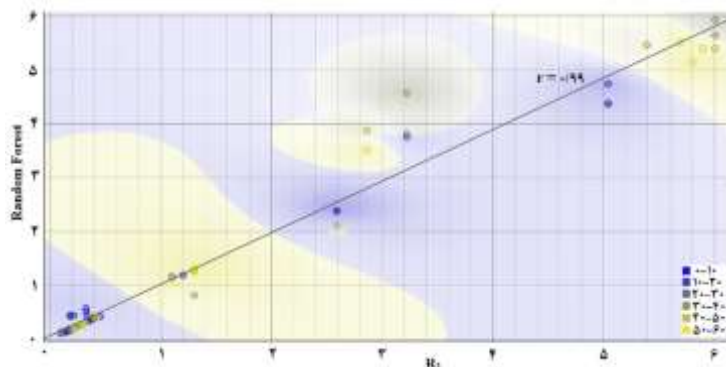
شکل ۵- نمودار همبستگی روش جنگل تصادفی برای پاسخ اول ( $R_1$ )

برخوردار است و در صورت اعمال متناسب این شیوه‌ها، خاک در برابر عوامل فرساینده محافظت شده و از دست رفتن خاک کاهش می‌یابد. آنچه که در این ارتباط از اهمیت قابل توجهی برخوردار است، مناسب

همان‌گونه که شکل (۶) نمایش می‌دهد، متأثر از اعمال شیوه‌های مدیریتی، مقدار فرسایش خاک به شکل معنی‌داری کاهش پیدا کرده است. شیوه‌های مدیریتی در حوزه‌های آبخیز از اهمیت بالایی

دارد که با پژوهش‌های پیشین انطباق دارد (Rouzies et al., 2019; Shirmohammadi et al., 2020).

بودن این شیوه‌های مدیریتی و انطباق آن با شرایط منطقه است. ضمناً همبستگی بالایی بین مقادیر اندازه‌گیری شده و برآوردی وجود



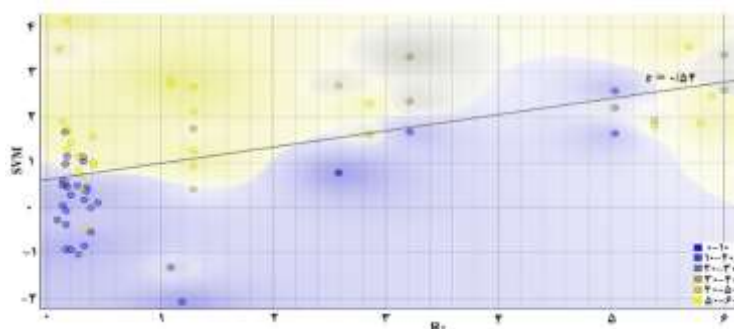
شکل ۶- نمودار همبستگی روش جنگل تصادفی برای پاسخ دوم ( $R_2$ )

ضریب همبستگی برابر  $0.73$  و برای پاسخ دوم برابر  $0.54$  است و این اختلاف نشان می‌دهد که متناسب با نوع راهکار مدیریتی نیز کارایی روش‌های مورد استفاده متفاوت می‌باشد. همان‌گونه که نتایج همبستگی بین متغیرهای مستقل و وابسته نشان می‌دهد، ضریب همبستگی حاصله از روش ماشین بردار پشتیبان، کمتر از روش جنگل تصادفی می‌باشد و برای پاسخ اول با پاسخ دوم ضریب همبستگی متفاوت است. شایان ذکر است که این راستا، Yoon et al (2011) نیز وجود اختلاف در دقت بهینه‌سازی توسط روش‌ها و الگوریتم‌های مختلف را گزارش نمودند.

مدل‌سازی با استفاده از الگوریتم ماشین بردار پشتیبان با انتخاب تابع خطی به‌عنوان تابع کرنل و عامل‌های  $C$  و  $\epsilon$  به ترتیب برابر  $1$  و  $0.1$ ، مقدار شاخص‌های آماری  $RMSE$ ،  $MSE$  و  $R^2$  برای پاسخ اول به ترتیب برابر  $14/437$ ،  $3/666$  و  $0/493$ ، برای پاسخ دوم به ترتیب برابر  $2/684$ ،  $1/638$  و  $0/258$  حاصل شد. روش ماشین بردار پشتیبان در مقایسه با روش جنگل تصادفی دارای شاخص‌های آماری ضعیف‌تری بوده و همین امر سبب می‌شود تا به نسبت دیگر روش‌ها از کارایی پایین‌تری برای پیش‌بینی وضعیت فرسایش خاک برخوردار باشد. شکل‌های (۷) و (۸) همبستگی بین متغیرها را نشان می‌دهد که برای پاسخ اول



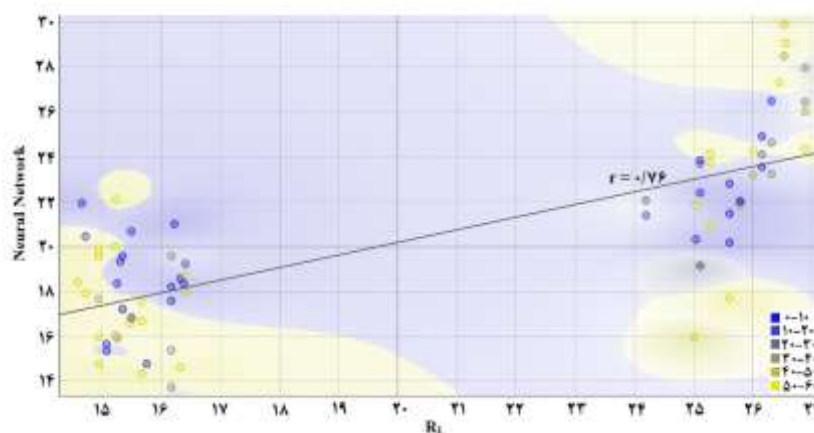
شکل ۷- نمودار همبستگی روش ماشین بردار پشتیبان برای پاسخ اول ( $R_1$ )



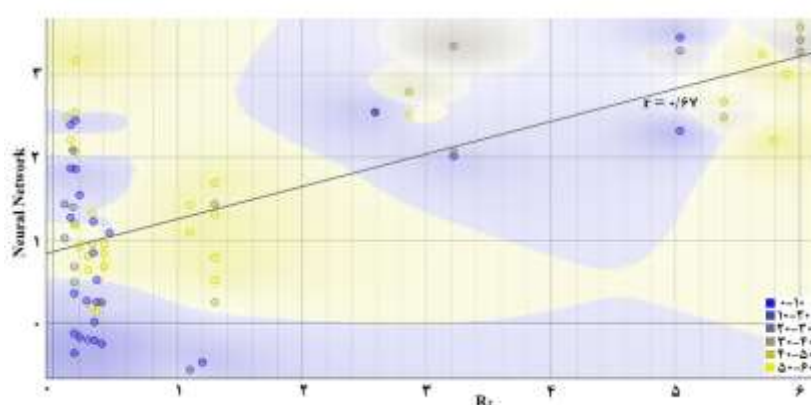
شکل ۸- نمودار همبستگی روش ماشین بردار پشتیبان برای پاسخ دوم ( $R_2$ )

نسبت به روش ماشین بردار پشتیبان و شبکه عصبی مصنوعی است. همچنین با توجه شکل های (۹) و (۱۰) متغیرهای مستقل و وابسته دارای همبستگی پایینی بوده و مقدار ضریب همبستگی ( $r$ ) آن برای هر دو پاسخ اول و دوم به ترتیب برابر با ۰/۷۶ و ۰/۶۷ است. از نظر ضریب همبستگی، روش شبکه عصبی مصنوعی در مقایسه با روش ماشین بردار پشتیبان دارای مقدار ضریب همبستگی بالاتری و نسبت به روش جنگل تصادفی مقدار ضریب همبستگی پایین تری برای هر دو پاسخ ( $R_1$  و  $R_2$ ) دارد.

مدل سازی با استفاده از الگوریتم شبکه‌ی عصبی مصنوعی نتایج آماری این بخش برای پاسخ اول شامل  $MSE = 11/286$ ،  $RMSE = 3/359$  و  $R^2 = 0/576$  و برای پاسخ دوم شامل  $MSE = 1/425$ ،  $RMSE = 1/438$  و  $R^2 = 0/438$  می باشد. با توجه به نتایج آماری به دست آمده، روش شبکه عصبی مصنوعی نسبت به روش ماشین بردار پشتیبان نتایج مطلوب تری را ارائه داده است؛ در حالی که نسبت به روش جنگل تصادفی نتایج آن نامطلوب تر است، بدین ترتیب روش جنگل تصادفی دارای بهترین نتایج



شکل ۹- نمودار همبستگی روش شبکه عصبی برای پاسخ اول ( $R_1$ )



شکل ۱۰- نمودار همبستگی روش شبکه عصبی برای پاسخ دوم ( $R_2$ )

## نتیجه گیری

ویژگی‌های فیزیکی و هیدرولیکی خاک متأثر از شیوه‌های حفاظتی غیر سازه‌ای مشتمل بر اصلاح پوشش گیاهی و اجرای قرق به عنوان راهکارهای مدیریتی بر روی رخدادهای فرسایش خاک، بهینه سازی شد. دستیابی به این حدود بهینه کمک می کند تا بتوان استراتژی‌های متناسب را با شرایط حاکم بر هر منطقه انتخاب نمود و از میزان فرسایش رخ داده و رسوب تولید شده به شکل معنی داری کاست. نتایج حاصله از این پژوهش حاکی از وجود اختلاف بین آماره‌های روش‌های ارزیابی شده است. همان گونه که نتایج نشان می دهد، روش سطح پاسخ در بین

با توجه به پیچیدگی فرسایش خاک و گستردگی عوامل تأثیرگذار بر آن در نتیجه ضرورت دارد که از روش‌ها و شیوه‌های مختلفی برای ارزیابی عوامل مؤثر بر آن در مقیاس حوزه آبخیز استفاده شود. در صورت بهره گیری صحیح از روش‌های برآوردی و نیز تحلیل آن‌ها می توان راهکارهای مدیریتی مطلوب و متناسبی را اتخاذ و عملیاتی نمود تا بتوان میزان هدررفت خاک را کاهش داد. در این پژوهش با بهره گیری از نرم افزارهای Design expert و Orange به عنوان ابزارهای بهینه سازی-مدلسازی، تأثیرات

تمامی روش‌های ارزیابی شده از بهترین آماره‌ها برخوردار است و لذا به‌عنوان مناسب‌ترین روش بهینه‌سازی پیشنهاد می‌شود. در بین روش‌های بهینه‌سازی ارزیابی شده توسط نرم‌افزار Orange نیز بهترین روش، الگوریتم جنگل تصادفی است. یکی از مزایای روش جنگل تصادفی آن است که هم برای رگرسیون و هم برای دسته‌بندی قابل استفاده است و راهکاری مناسب برای پی بردن به اهمیت نسبی ویژگی‌های ورودی مدل است. یکی از بزرگ‌ترین مشکلات در روش یادگیری ماشین، بیش‌برازش است اما اغلب اوقات این مسئله به آن آسانی که برای دسته‌بندی جنگل تصادفی به وقوع می‌پیوندد، اتفاق نمی‌افتد. محدودیت اصلی جنگل تصادفی آن است که تعداد زیاد درخت‌ها می‌توانند الگوریتم را برای پیش‌بینی‌های جهان واقعی کند و غیر مؤثر کنند. به طور کلی، آموزش دادن این الگوریتم‌ها سریع انجام می‌شود اما فرایند پیش‌بینی پس از آموزش مدل، اندکی کند به وقوع می‌پیوندد. یک پیش‌بینی صحیح‌تر نیازمند درختان بیشتری است که منجر به کندتر شدن مدل نیز می‌شود. در اغلب کاربردهای جهان واقعی،

تمامی روش‌های جنگل تصادفی به اندازه کافی سریع عمل می‌کند اما امکان دارد شرایطی نیز وجود داشته باشد که در آن کارایی زمان اجرا حائز اهمیت است و در چنین شرایطی استفاده از دیگر رویکردها ترجیح داده می‌شود. همچنین کارایی روش سطح پاسخ به‌عنوان یک روش توانمند جهت بهینه‌سازی عوامل فیزیکی و هیدرولیکی تأثیرگذار بر روی فرسایش خاک را تأیید نمودند (Shirazi et al., 2020). در مجموع بهره‌گیری از این روش‌های آماری، امکان پایش هر چه بهتر و دقیق‌تر عوامل مؤثر بر رخدادهای فرسایشی را فراهم می‌سازد و بستری را فراهم می‌کند تا تصمیم‌گیری‌های مدیریتی با دقت بیشتری صورت پذیرد. همان‌گونه که نتایج این پژوهش نیز به خوبی نشان می‌دهد، در راستای پیشرفت‌های صورت گرفته در علوم مختلف، در نتیجه استفاده از شیوه‌های نوین نیز امری اجتناب‌ناپذیر است و مجموعه‌ی این روش‌ها به‌عنوان ابزاری مدیریتی محسوب می‌شوند که تصمیم‌گیری‌ها را غنی‌تر نموده و زیربنایی علمی برای آن‌ها تعریف و فراهم می‌نمایند.

## REFERENCES

- Andersen, P. S., Andersen, E., Graversgaard, M., Christensen, A. A., Vejre, H., and Dalgaard, T. (2019). Using landscape scenarios to improve local nitrogen management and planning. *Journal of Environmental Management*, 232, 523–530.
- Azimi-Pour, M., Eskandari-Naddaf, H., and Pakzad, A. (2020). Linear and non-linear SVM prediction for fresh properties and compressive strength of high volume fly ash self-compacting concrete. *Construction and Building Materials*, 230: 117021.
- Bagdi, G. L., Mishra, P. K., Kurothe, R. S., Arya, S. L., Patil, S. L., Singh, A. K., ...and Sundarambal, P. (2015). Post-adoption behaviour of farmers towards soil and water conservation technologies of watershed management in India. *International Soil and Water Conservation Research*, 3(3): 161–169.
- Baigorria, G. A., and Romero, C. C. (2007). Assessment of erosion hotspots in a watershed: Integrating the WEPP model and GIS in a case study in the Peruvian Andes. *Environmental Modelling & Software*, 22(8): 1175–1183.
- Booker, D. J., and Snelder, T. (2012). Comparing methods for estimating flow duration curves at ungauged sites. *Journal of Hydrology*, 434: 78-94.
- Bouyucos, C.J. (1962). Hydrometer method improved for making particle-size analysis of soil. *Journal of Agrometer*, 54: 464-465.
- Box, G. E. P., and Draper, N. R. (2007). Response surfaces, mixtures and ridge analyses, *John Wiley and Sons*, New York, 857.
- Brady, N.C. and Weil, R.R. (2014). The Nature and Properties of Soils, Revised 14th edition. *Pearson, Prentice Hall, Upper Saddle River*, New Jersey and Columbus, Ohio, USA (pp 975).
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1): 5-32.
- Chen, G., Wang, Y., Li, S., Cao, W., Ren, H., Knibbs, L. D and Guo, Y. (2018). Spatiotemporal patterns of PM10 concentrations over China during 2005–2016: a satellite-based estimation using the random forests approach. *Environmental pollution*, 242: 605-613.
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3): 273-297.
- Demšar, J., and Zupan, B. (2013). Orange: Data mining fruitful and fun-a historical perspective. *Informatica*, 37(1): 56-50.
- Eggers, J., Holmgren, S., Nordström, E.-M., Lämås, T., Lind, T., and Öhman, K. (2017). Balancing different forest values: Evaluation of forest management scenarios in a multi-criteria decision analysis framework. *Forest Policy and Economics*, 103, 55-69.
- Heung, B., Bulmer, C. E., and Schmidt, M. G. (2014). Predictive soil parent material mapping at a regional-scale: a random forest approach. *Geoderma*, 214: 141-154.
- Lal, R. (2019). Accelerated soil erosion as a source of atmospheric CO<sub>2</sub>. *Soil and Tillage Research*, 188: 35-40.
- Naik, A., and Samant, L. (2016). Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime. *Procedia Computer Science*, 85: 662-668.
- Nath, A. J., and Lal, R. (2017). Managing tropical wetlands for advancing global rice production: Implications for land-use management. *Land Use*

- Policy, 68: 681-685.
- Nezhadafzali, K., Shahrokhi, M., and Bayatani, F. (2019). Assessment soil erosion using RUSLE model and identification the most effective factor in Dekhan watershed basin of southern Kerman. *Journal of Natural Environmental Hazards*, 20 (8): 21-38.
- Norouzi, H., Azgharimoghaddam, A., and Nadiri, A. (2019). Determination of vulnerable areas of Dasht-e malekan watershed to nitrate using random forest method. *Ecology*, 41(4):923-942.
- Oehlert, W. Gary. (2000). Design and analysis of experiments: Response surface design, New York: W.H. Freeman and Company.
- Parsafar, N and Marofi, S. (2011). Estimation of soil temperature from air temperature using regression models, artificial neural network and adaptive neuro-fuzzy inference system (Case Study: Kermanshah Region). *Journal of Water and Soil Science*, 21(4): 141-152. (In Farsi).
- Partopour, B., Paffenroth, R. C., and Dixon, A. G. (2018). Random forests for mapping and analysis of microkinetics models. *Computers & Chemical Engineering*, 115: 286-294.
- Podvalny, S. L., and Vasiljev, E. M. (2017). The principle of multi-alternativity in intelligent systems. Active neural network models. *Procedia Computer Science*, 103: 410-415.
- Raymond, H. M., Douglas C, Montgomery and Christine M, Anderson-cook. (2009). Response surface methodology. *Published by John Wiley & Sons, Inc., Hoboken, New Jersey published simultaneously in Canada*, 705.
- Rouzies, E., Lauvernet, C., Barachet, C., Morel, T., Branger, F., Braud, I., and Carluer, N. (2019). From agricultural catchment to management scenarios: A modular tool to assess effects of landscape features on water and pesticide behavior. *Science of the Total Environment*, 671, 1144-1160.
- Shirazi, M., Khadearasoul, A., and Saffieddin Ardebili, S. M. (2019). Evaluation of hydraulic parameters on water erosion using response surface methodology (RSM). 16th Iranian Soil Science Congress, University of Zanjan.
- Shirazi, M., Khadomalrasoul, A. & Safieddin Ardebili, S.M. (2020). Multi-objective optimization of soil erosion parameters using response surface method (RSM) in the Emamzadeh watershed. *Acta Geophys.* **68**, 505-517 (2020). <https://doi.org/10.1007/s11600-020-00404-5>.
- Shirmohammadi, B., Malekian, A., Salajegheh, A., Taheri, B., Azarnivand, H., Malek, Z., and Verburg, P. H. (2020). Scenario analysis for integrated water resources management under future land use change in the Urmia Lake region, Iran. *Land Use Policy*, 90, 104299.
- Sun, L., Zou, B., Fu, S., Chen, J., and Wang, F. (2019). Speech emotion recognition based on DNN-decision tree SVM model. *Speech Communication*, 115: 29-37.
- Thomas, G. W. (1982). "Exchangeable Cations. Methods of Soil Analysis, Part 2, Chemical and Microbiological Properties", Second Edition. A.L. Page (editor). Agronomy, No. 9, Part 2, American Society of Agronomy, *Soil Science Society of America*, Madison, WI: 159-165.
- Tripathi, R., Nayak, A. K., Shahid, M., Lal, B., Gautam, P., Raja, R. and Sahoo, R. N. (2015). Delineation of soil management zones for a rice cultivated area in eastern India using fuzzy clustering. *Catena*, 133: 128-136.
- Vapnik, V. (2013). The nature of statistical learning theory. *Springer science & business media*.
- Vorpahl, P., Elsenbeer, H., Märker, M., and Schröder, B. (2012). How can statistical models help to determine driving factors of landslides? *Ecological Modelling*, 239: 27-39.
- Walkley, A. and Black, I.A. (1934). An examination of the degtjareff method for determining soil organic matter and proposed modification of chromic acid titration method. *Soil SCIENCE*, 37: 29-38.
- Wilson, M. J. (2019). The importance of parent material in soil classification: A review in a historical context. *Catena*, 182: 104131.
- Yoon, H., Jun, S. C., Hyun, Y., Bae, G. O., and Lee, K. K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*, 396(1-2): 128-138.