

## مدل سازی ردپای آب گندم با استفاده از مدل های یادگیری ماشین در استان فارس

فاطمه بادروج<sup>۱</sup>، ام البنین بذرافشان<sup>۲\*</sup>

۱. دانش آموخته کارشناسی ارشد آبخیزداری، گروه مهندسی منابع طبیعی، دانشگاه هرمزگان، بندرعباس ایران
۲. هیئت علمی گروه مهندسی منابع طبیعی، دانشیار، دانشکده کشاورزی و منابع طبیعی، دانشگاه هرمزگان، بندرعباس ایران

(تاریخ دریافت ۱۴۰۱/۰۱/۱۱؛ تاریخ بازنگری ۱۴۰۱/۰۲/۱۰؛ تاریخ تصویب ۱۴۰۱/۰۳/۳۰)

### چکیده

این مطالعه با هدف تخمین، پیش بینی و مدل سازی ردپای آب سبز و آبی محصول زراعی گندم با استفاده از مدل های یادگیری ماشین در اراضی فاریاب در دوره آماری (۱۳۸۴ تا ۱۳۹۶) انجام شد. بر این اساس، با استفاده از داده های اقلیمی و گیاهی و روش فازی کلاستر، مناطق کشت گندم فاریاب در استان فارس به چهار منطقه همگن تقسیم شد. در هر منطقه براساس چارچوب اوکسترا، ردپای آب آبی، سبز و خاکستری برآورد شد. سپس، ردپای آب در هم اقلیم همگن به دو دسته آموزش (۷۰ درصد) و آزمون (۳۰ درصد) تقسیم شد و با استفاده از مدل شبکه عصبی و دو کرنل لوگ لجستیک و تانژانت هایپربولیک (۵۰ ترکیب ورودی)، مدل جنگل تصادفی و رگرسیون بردار پشتیبان (تابع کرنل سیگموئید) با متغیرهای اقلیمی و گیاهی، پیش بینی صورت گرفت و نتایج مدل ها با شاخص های ارزیابی خطا و دیاگرام تیلور مورد مقایسه قرار گرفت. نتایج نشان داد بهترین مدل برای برآورد ردپای آب گندم فاریاب در استان فارس مدل شبکه عصبی مصنوعی با تابع لوگ لجستیک با ضریب همبستگی بیش از ۰/۷۲ و میانگین خطای مطلق کمتر از ۰/۴۸ (مترمکعب بر تن) است و می تواند به ارتقای فرایند تصمیم گیری به مدیران آب و برنامه ریزان کمک کند.

**کلمات کلیدی:** ردپای آب، مناطق همگن، گندم فاریاب، مدل شبکه عصبی مصنوعی، تابع لوگ لجستیک.

## مقدمه

آب به عنوان یکی از عناصر اساسی حیات، امروزه با بحرانی جدی روبه‌رو است. رشد جمعیت و افزایش سرانه مصرف از یکسو و توسعه روزافزون فعالیت‌های صنعتی و کشاورزی از سوی دیگر و کاهش نزولات جوی موجب شده است تا این منبع حیاتی به طور فزاینده‌ای کاهش یابد [۱]. این امر و بروز خشکسالی‌های چند دهه اخیر موجب شده تا توسعه منابع آب بیشتر شده و فشار بر ذخایر آب‌های سطحی و زیرزمینی افزایش یابد و سیمای کشاورزی کشور، چهره ناپایداری به خود گیرد. در سال‌های اخیر دیدگاه ردپای آب به‌ویژه در محصولات کشاورزی، به عنوان راهکاری جدید جهت سازگاری با بحران آب مورد توجه متخصصان و مجامع علمی مرتبط با آب جهان قرار گرفته است [۲].

برای اینکه بتوان اثر الگوی مصرف مردم را بر منابع طبیعی نشان داد، می‌توان از مفهوم ردپای آب استفاده کرد که اولین بار در سال ۲۰۰۲ توسط هوکستر<sup>۱</sup> و هانگ<sup>۲</sup> [۲] معرفی شد. آن‌ها اظهار کردند که کل مصرف آب در داخل یک کشور به‌تنهایی معیار درستی از برداشت واقعی آب آن کشور از منابع آب جهانی نیست. در واقع، حجم آب مجازی وارداتی باید به کل مصرف آب داخلی افزوده شود تا تصویر درستی از نیاز واقعی یک کشور به منابع آب جهانی ترسیم شود. بنابراین، ردپای آب یک شاخص کلیدی در مدیریت منابع آب و کشاورزی است. ردپای آب با توجه به جایگاه کلیدی آن در مدیریت منابع آب و کشاورزی اهمیت زیادی دارد. با توجه به آنکه ردپای آب وابسته به پارامترهای مختلفی است، بنابراین پیش‌بینی آن موضوعی دشوار و پیچیده است. یکی از مدل‌های مؤثر برای پیش‌بینی و تخمین متغیرهای هیدرولوژیکی و کشاورزی، مدل‌های یادگیری ماشین است.

مدل‌های یادگیری ماشین از واحدهای پردازش اطلاعات به‌هم‌پیوسته، با شناسایی روابط و الگوهای موجود در داده‌ها برای تبدیل ورودی به خروجی استفاده می‌کنند [۳]. دقت مناسب، سرعت زیاد، همگرایی سریع و قابلیت انعطاف‌پذیری از ویژگی‌های این مدل‌هاست. پژوهش‌های قبلی از روش‌های مختلف محاسبات نرم برای پیش‌بینی

ردپای آب استفاده کرده‌اند. البلتقی و همکاران [۴] از پارامترهای اقلیمی و مدل (ANNs<sup>۳</sup>) برای پیش‌بینی ردپای آب استفاده کردند. آن‌ها گزارش دادند که مدل ANN با استفاده از داده‌های اقلیمی به طور دقیق ردپای آب را پیش‌بینی می‌کند. آماده کردن پارامترهای مدل ANN و انتخاب تعداد لایه‌های پنهان مدل ANN از چالش‌های تحقیق آن‌ها بود. مونی و همکاران [۵] از ANN و support vector machine (SVM) جهت پیش‌بینی ردپای آبی استفاده کردند. آن‌ها گزارش دادند که مدل ANN از مدل SVM دارای عملکرد بهتری است. علاوه بر این، بارندگی و تبخیر را به عنوان پارامترهای مؤثر روی ردپای آب اعلام کردند. آماده‌سازی و یافتن پارامترهای مدل‌های SVM و ANN با استفاده از الگوریتم‌های آموزشی یکی از چالش‌های اصلی بود. مختار و همکاران [۶] از روش جنگل تصادفی، درخت تصادفی و رگرسیون خطی چندگانه<sup>۴</sup> برای پیش‌بینی ردپای آب آبی و سبز استفاده کردند. آن‌ها پیش‌بینی خود را برای دوره آینده ۲۰۲۱-۲۰۵۰ تحت شرایط تغییر اقلیم انجام دادند. آن‌ها گزارش دادند که ردپای آب آبی و سبز برای دوره آینده یک روند افزایش خواهد داشت. آماده کردن سناریوی‌های اقلیمی، انتخاب بهترین مدل اقلیمی و جمع‌آوری داده‌ها از چالش‌های تحقیق آن‌ها بود.

عدم رعایت الگوی کشت بر مبنای مفهوم ردپای آب و آب مجازی در ایران در حال حاضر لطمات جبران‌ناپذیری به منابع آب زیرزمینی در دشت‌های مختلف در نقاط مختلف کشور وارد کرده است و ادامه روند الگوی کشت غیرعلمی و غیراصولی به تدریج طی دهه آینده سبب فرونشست بیش از پیش دشت‌ها بر اثر اضافه برداشت از سفره‌های آب زیرزمینی، تخریب اراضی، از بین بردن اراضی کشاورزی و باغی در روستاها، کمبود آب شرب و بهداشتی در شهرها، خالی از سکنه شدن بخش عظیمی از مناطق ایران و مهاجرت اجباری جمعیت به سایر نقاط دارای منابع آب، افزایش تنش‌های محلی در خصوص استفاده از منابع آبی بین کشاورزان و حتی واکنش‌ها در اعتراض به خشک شدن رودخانه‌ها و سدها و افزایش

3. Artificial Neural Networks  
4. MLR: Multiplier Linear Regression

1. Hoekstr  
2. Hung

سالانه در استان فارس به طور میانگین برابر با ۱۲۳۸۰۰۰ تن است. تقریباً ۱۰/۵ درصد از کل گندم کشور در استان فارس تولید می‌شود. به جهت تنوع آب‌وهوا در همه‌جای این استان گندم کشت می‌شود که بیشتر در دشت‌ها به صورت آبی و در نواحی کوهستانی به صورت دیم است. در این پژوهش آمار درازمدت داده‌های هواشناسی (جدول ۱) مناطق اصلی تولید گندم دیم و آبی از بانک داده‌های سازمان هواشناسی استان فارس تهیه شد. داده‌های مربوط به عملکرد گندم دیم و آبی، سطح زیر کشت، میزان تولید، میزان مصرف کود شیمیایی، راندمان آبیاری، تاریخ کشت و تیپ خاک استان فارس نیز از بانک اطلاعاتی وزارت جهاد کشاورزی فارس گردآوری شد.

### روش فازی کلاستر

الگوریتم‌های خوشه‌بندی از محبوب‌ترین روش‌های طبقه‌بندی اقلیمی هستند [۱۱]. در خوشه‌بندی، مناطق دارای ویژگی‌های نزدیک به هم در یک گروه قرار می‌گیرند [۱۲]. طبق الگوریتم‌های خوشه‌بندی هر منطقه فقط به یک خوشه تعلق دارد [۱۳]. به این معنا که مناطق عضو نیست [۱۱]. برای رفع این محدودیت، الگوریتم فازی توسعه یافت [۱۴].

داده‌های مورد استفاده در این بخش شامل داده‌های اقلیمی شامل متوسط دما، متوسط بارش، حداکثر دما، حداقل دما، ارتفاع از سطح دریا، مجموع ساعت‌های آفتابی، متوسط رطوبت از سازمان هواشناسی استان فارس و داده‌های گیاهی شامل طول دوره رشد، ماه شروع و پایان رشد، بارش مؤثر، تبخیر و تعرق گیاهی، نیاز آبی و متوسط عملکرد در واحد سطح در دوره آماری ۱۳۸۴ تا ۱۳۹۶ تهیه شد. درخور یادآوری است که هر شهرستان دارای یک ایستگاه سینوپتیک است که موقعیت آن در شکل ۱ نشان داده شده است.

بر اساس متغیرهای نامبرده در روش فازی کلاستر، منطقه مورد بررسی به چهار خوشه همگن تقسیم شد (شکل ۱). به این ترتیب، منطقه ۱ شامل ۵ شهرستان در نواحی جنوب و غرب فارس، منطقه ۲ شامل ۷ شهرستان (نواحی شمالی)، منطقه ۳ شامل ۷ شهرستان (نواحی شرقی) و منطقه ۴ شامل ۵ شامل نواحی مرکزی فارس است.

اعتراض‌های شکل گرفته نسبت به روند انتقال آب بین حوضه‌ای خواهد شد [۷].

در میان محصولات کشاورزی گندم یکی از غلات مهم است که در سراسر جهان رشد می‌کند و منبع کالری و تغذیه اولیه برای انسان است. بنابراین، به منظور تأمین امنیت غذایی نیاز به سیاست‌ها و استراتژی برای کاهش کمبود مواد غذایی است و برآورد به‌موقع و دقیق از تولید این محصول می‌تواند، بسیار مفید باشد [۸]. گندم به عنوان کهن‌ترین و بیشترین کشت در بین غلات مطرح است و تاریخچه کشت آن به ۱۰ هزار سال قبل از میلاد مسیح برمی‌گردد. تقریباً ۲۱۵ میلیون هکتار از اراضی زیر کشت در سطح جهان به کشت گندم مربوط می‌شود [۹]. محصول گندم در ایران یک محصول استراتژیک به شمار می‌رود و تولید آن از اهمیت بالایی برخوردار است و به طور متوسط با تولید ۱۳ میلیون تن در سال، بیش از ۲۰ درصد سطح اراضی زراعی تحت کشت را به خود اختصاص داده است. استان فارس به دلیل تنوع اقلیمی، در خصوص تولید گندم همواره دارای رتبه برتر در بین استان‌های تولیدکننده گندم در ایران است و سهم بالایی در تولید گندم در ایران ایفا می‌کند [۱۰].

بررسی مطالعات گذشته نشان می‌دهد محققان مختلف بیشتر بر ارزیابی ردپای آب در مقیاس ملی متمرکز بوده‌اند در حالی که مطالعات کمی روی تخمین ردپای آب با استفاده از مدل‌های یادگیری ماشین و هوش مصنوعی صورت گرفته است. بنابراین، هدف از این تحقیق برآورد ردپای آب در گندم فاریاب در استان فارس و پیش‌بینی آن با مدل‌های یادگیری ماشین است. در این تحقیق، از شبکه عصبی مصنوعی به عنوان یک ابزار برای مدل‌سازی ردپای آب استفاده شده و نتایج آن با روش‌های ماشین بردار پشتیبان و جنگل تصادفی مقایسه شده است.

### مواد و روش‌ها

#### منطقه مطالعه شده

بر اساس سرشماری اراضی، در سال ۱۳۹۶ مساحت اراضی کشاورزی ایران ۱۴۶۸۷۰۰۰ هکتار است که از این مقدار ۵۹۷۷۰۰۰ هکتار به صورت آبی و بقیه به صورت دیم است. از این میزان ۵۶ درصد اراضی تحت کشت یعنی ۸۲۲۵۰۰۰ هکتار زیر کشت گندم است. میزان تولید گندم

## برآورد ردپای آب

$$WF_{Blue} = \frac{(ET_c - P_e) * 10}{Y} \quad (2)$$

که  $WF_{Blue}$  ردپای آب آبی برحسب مترمکعب بر تن و  $ET_c$  تبخیر و تعرق گیاهی برحسب میلی متر است.  $WF_{grey}$  نیز به عنوان آب خاکستری تعریف می شود و به حجم آبی گفته می شود که طی تولید کالا یا محصولی آلوده می شود. در بخش کشاورزی، این آب ناشی از مصرف کود شیمیایی یا سموم دفع آفات است که به صورت رابطه ۳ محاسبه می شود.

$$WF_{Grey} = \frac{a * NAR}{C_{Max} - C_{Nat}} * \frac{1}{Y} \quad (3)$$

که  $a$  (٪) درصد تلفات کودهای نیتروژن که در شرایط فاریاب ۱۰ درصد در نظر گرفته می شود [۱۵ و ۱۶]،  $NAR$  مصرف کود نیتروژن بر حسب کیلوگرم بر هکتار،  $C_{Max}$  غلظت بحرانی نیتروژن بر حسب کیلوگرم بر متر مکعب و  $C_{Nat}$  غلظت طبیعی نیتروژن در خاک است، که به دلیل عدم اندازه گیری آن در استان فارس، براساس پیشنهاد اوکسترا و چاپاگین [۱۵ و ۱۶] این مقدار صفر در نظر گرفته شد. در این تحقیق، مقدار  $a$  در شرایط فاریاب ۱۰ درصد لحاظ شد. در نهایت، ردپای کل عبارت است از: مجموع کل اجزای ردپای آب (رابطه ۴).

$$WFT = WF_{green} + WF_{blue} + WF_{grey} \quad (4)$$

## شبکه عصبی پرسپترون چندلایه

شبکه عصبی MLP از شناخته ترین و محبوب ترین انواع شبکه عصبی است که در حل مسائل طبقه بندی و رگرسیونی کاربرد دارد [۲۰]. ساختار ساده این مدل دارای سه لایه شامل لایه ورودی، لایه پنهان یا میانی و لایه خروجی است. دریافت ورودی ها توسط لایه اول انجام می شود. لایه میانی پردازش اطلاعات دریافت شده از لایه اول را انجام می دهد. لایه میانی شامل توابع فعال و نرون های پنهان است. هر لایه به وسیله اتصالاتی به لایه قبل و بعد خود متصل می شود. هر اتصال دارای مقدار وزنی است. البته لایه نخست فقط به لایه های بعد خود متصل می شود. هر اتصال وزنی مدل MLP دارای دو پارامتر وزن و بایاس می شود. مقدار ورودی به نرون در لایه میانی در وزن اتصال آن نرون ضرب می شود و سپس، با

در این مطالعه، ردپای آبی و سبز محصول گندم با استفاده از چارچوب اصلی توصیف شده توسط هوکسترا و چاپاگین [۱۵ و ۱۶] محاسبه شد. قبل از محاسبه ردپای آب، ابتدا تبخیر و تعرق محصول، نیاز آبی، نیاز آبیاری و بارندگی مؤثر با استفاده از مدل CropWat در هر دشت محاسبه شد.

برای محاسبه تبخیر و تعرق محصول ابتدا با استفاده از معادله فائو- پنمن- مونتیت [۱۷] تبخیر و تعرق مرجع محاسبه شد و سپس، براساس ضرایب پیش فرض FAO-56، تبخیر و تعرق محصول برآورد شد. برای اطلاعات بیشتر می توانید به آلن و همکاران [۱۷] مراجعه کنید. از آنجا که برای محاسبه تبخیر و تعرق گیاهی به ضریب گیاهی<sup>۱</sup> نیاز است، ضرایب محصول (Kc) محاسبه شده توسط جهاد کشاورزی استان فارس برای هر منطقه تحت کشت گندم، استفاده شد و در محیط CropWat مقدار ETC کالیبره شد. در ادامه، بعد از محاسبه تبخیر و تعرق گیاهی، از اختلاف آن با بارش مؤثر نیاز خالص آبیاری محاسبه شد و با استفاده از راندمان آبیاری (گزارش شده توسط وزارت کشاورزی برای هر منطقه) نیاز آبی محاسبه شد و در نهایت، پس از تعیین متغیرهای مربوطه، اجزای ردپای آب سبز، آبی و خاکستری به صورت زیر محاسبه شد.

آب سبز سهمی از بارندگی مؤثر است که توسط گیاهان در یک محیط خاک غیر اشباع مصرف می شود [۱۸]. برای محاسبه ردپای آب سبز از رابطه ۱ استفاده شد.

$$WF_{Green} = \frac{(P_e) * 10}{Y} \quad (1)$$

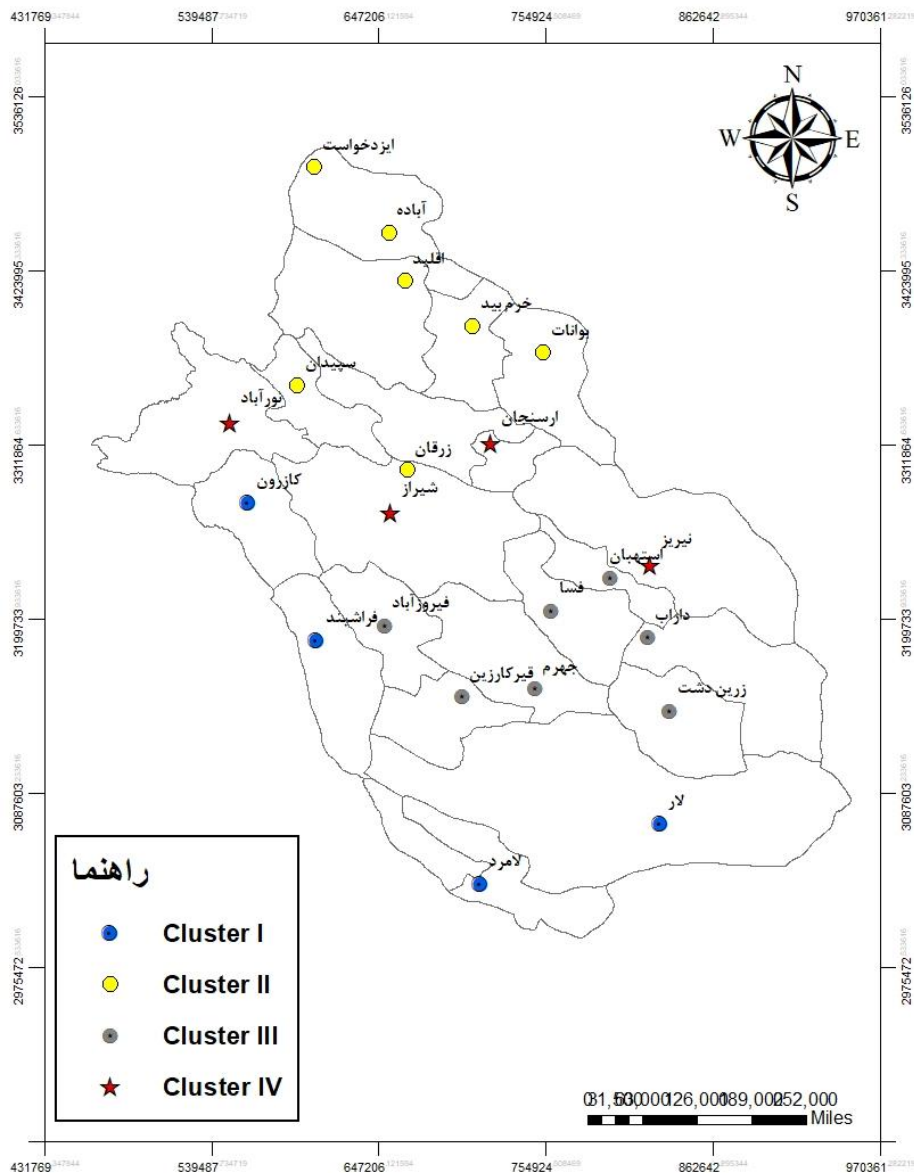
که  $WF_{Green}$  ردپای آب سبز بر حسب متر مکعب در هر تن؛  $P_e$  مجموع بارندگی مؤثر ( USDA S.C. using Method) طی دوره رشد گیاه به میلی متر و  $Y$  عملکرد محصول برحسب تن در هکتار است.

آب آبی عبارت است از: مجموع آب های زیرزمینی و رواناب های سطحی ناشی از نزولات جوی که ممکن است توسط مخازن جمع آوری آب ذخیره شده و در بخش کشاورزی برای آبیاری مورد استفاده قرار گیرند [۱۹]. ردپای آب آبی محاسبه شد.

1. Crop coefficient (Kc)

همگرایی و دقت زیاد نداشته باشند. بنابراین، الگوریتم‌های بهینه‌سازی به عنوان جایگزین مناسب جهت یافتن مقدار بهینه پارامترهای شبکه عصبی هستند. الگوریتم‌های بهینه‌سازی با استفاده از سرعت همگرایی زیاد، استفاده از اپراتورهای پیشرفته و دقت زیاد توان زیادی در آموزش مدل‌های محاسبات نرم دارند [۲۱]. در این تحقیق از دو تابع تانژانت سیگموئید و لوگ لوجستیک استفاده شد تا کارایی هر یک در پیش‌بینی ردپای آب مشخص شود.

بایاس به عنوان یک ترم ثابت جمع می‌شود. سپس، تابع فعال مقدار یادشده را دریافت می‌کند تا پردازش کند. خروجی‌های لایه میانی یا دوم به عنوان ورودی لایه آخر وارد می‌شوند. لایه آخر به عنوان لایه خروجی شناخته می‌شود. به هر حال، بایاس و وزن به عنوان پارامترهای ناشناخته مدل هستند. الگوریتم‌هایی نظیر گرادیان قابل استفاده جهت یافتن مقدار بهینه پارامترهای یادشده هستند. البته، الگوریتم‌های یادشده ممکن است سرعت



شکل ۱. موقعیت شهرستان‌ها و ایستگاه‌های سینوپتیک استان فارس در هر منطقه

کرنل رادیال  $k(x_i, x_{i'}) = \exp(-\gamma \sum_{j=1}^p (x_{ij} - x_{i'j})^2)$

کرنل سیگموئید  $k(x_i, x_{i'}) = \tanh(-\gamma \sum_{j=1}^p x_{ij} x_{i'j} + \alpha)$

### روش جنگل تصادفی

روش جنگل تصادفی در حال حاضر از بهترین روش‌های یادگیری است. این روش یک روش یادگیری ماشین ناپارامتری مبتنی بر دسته‌ای از درخت‌های تصمیم است. تعداد بسیاری درخت تصمیم در طبقه‌بندی روش RF رشد می‌کنند. پیکسل یا پدیده طبقه‌بندی نشده با توجه به ویژگی‌های همراه خود در یک کلاس گنجانده می‌شود. این امر با عبور دادن پیکسل از میان  $X$  درخت تصمیم انجام می‌شود. به گونه‌ای که تعلق پیکسل به یکی از کلاس‌ها ( $Y$  تعداد کلاس) میان درختان تصمیم به رأی گذاشته می‌شود و هر درخت تصمیم رأی خود را در مورد گنجاندن پیکسل در یک کلاس اعلام می‌کند. جنگل پیکسل به رأی گذاشته شده را به کلاسی که دارای بیشترین رأی درختان جنگل باشد، اختصاص می‌دهد. درختان تصمیم‌گیری به صورت منفرد از مجموعه نمونه آموزشی رشد می‌کنند. با  $N$  بار نمونه‌گیری با جایگزینی، دو سوم از مجموعه داده‌های اصلی برای آموزش یک درخت استفاده می‌شود،  $N$  تعداد نمونه‌ها در مجموعه داده‌های اصلی است. به این طریق با نمونه‌گیری با جایگزینی، یک سوم باقی‌مانده داده‌ها در آموزش درخت دخالت نخواهند داشت و از کیسه کنار گذاشته می‌شوند تا برای صحت‌سنجی درونی الگوریتم مورد استفاده قرار گیرند [۲۳]. در این روش به منظور اولویت‌بندی عوامل مؤثر بر ردپای آب در منطقه مطالعه از روش میانگین کاهش دقت و میانگین کاهش جینی برگرفته از الگوریتم جنگل‌های تصادفی استفاده شد [۲۴].

### معیار ارزیابی مدل‌ها

معمولاً سه معیار به منظور ارزیابی کارایی مدل‌ها استفاده می‌شود. ریشه مربع متوسط خطا (RMSE) که به صورت معادله ۹ محاسبه می‌شود:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Q_o - Q_e)^2}{N}} \quad (9)$$

### رگرسیون بردار پشتیبان<sup>۱</sup>

روشی است که در دهه ۱۹۹۰ به گسترش کامپیوترها توسعه یافت. ماشین بردار پشتیبان بسطی از یک روش طبقه‌بندی ساده است که به آن طبقه‌بندی بیشترین حاشیه<sup>۲</sup> گفته می‌شود و برای طبقه‌بندی گروه‌هایی با مرزبندی خطی به کار می‌رود. در همین راستا، طبقه‌بندی‌کننده بردار پشتیبان<sup>۳</sup> معرفی می‌شود که بسطی از طبقه‌بندی بیشترین حاشیه است که علاوه بر مرز خطی بین طبقات، برای طیف وسیع‌تری از مرز بین طبقات به کار برده می‌شود. ماشین بردار پشتیبان نیز بسطی از طبقه‌بندی بردار پشتیبان است که برای مرزهای غیرخطی بین طبقات به کار برده می‌شود [۲۲].

فرض کنید زوج متغیرهای  $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$  وجود داشته باشد به طوری که  $x_i \in \mathbb{R}^p$  و  $y_i \in \{-1, 1\}$ . حال ابرصفحه را در نظر می‌گیریم:

$$\{x: f(x) = x^T \beta + \beta_0 = 0\}, \quad (5)$$

در این صورت، ماشین بردار پشتیبان عبارت است از: جداسازی مشاهدات در فضای  $p$  بعدی بر پایه رابطه ۶

$$\min_{\beta, \beta_0} \beta^2 + C \sum_{i=1}^n \xi_i \quad (6)$$

$$\text{subject to } \xi_i \geq 0, y_i (x^T \beta + \beta_0) \geq 1 - \xi_i \forall i, \quad (7)$$

در رابطه یادشده  $\xi_i$  پارامتر کمکی و  $C$  پارامتر هزینه است (لین و لین، ۲۰۰۳)، اما در عمل معمولاً مرز بین طبقات غیرخطی است و  $f(x)$  به صورت رابطه ۸:

$$f(x) = \beta_0 + \sum_{i=s}^p \alpha_i k(x, x_i) \quad (8)$$

در نظر گرفته می‌شود که این تابع یک تابع غیرخطی بوده و  $k(x, x_i)$  را تابع کرنل گویند. برخی از توابع کرنل معروف که در این مقاله مورد استفاده قرار گرفته‌اند، عبارت‌اند از [۱۸].

$$k(x_i, x_{i'}) = \sum_{j=1}^p x_{ij} x_{i'j}$$

$$k(x_i, x_{i'}) = (\gamma \sum_{j=1}^p x_{ij} x_{i'j} + \alpha)^d$$

1. Support Vector regression: SVR
2. Maximal Margin Classifier
3. Support Vector Classifier

براساس نتایج، هر سه مؤلفه در حال کاهش است، به طوری که میزان تغییرات از سال ۱۳۹۴ به بعد به طور محسوسی قابل مشاهده است.

در منطقه ۲ (شکل ۲ ب)، متوسط کل ردپای آب ۲۶۴۲ متر مکعب بر تن است. بنابراین، سهم هر یک از ردپای آب آبی، سبز و خاکستری به ترتیب ۵۴، ۱۰ و ۳۶ درصد است. در منطقه ۳ (شکل ۲ ج)، متوسط کل ردپای آب ۲۰۱۴ متر مکعب در هر تن است. بنابراین، سهم هر یک از ردپای آب آبی، سبز و خاکستری به ترتیب ۴۵، ۱۷ و ۳۸ درصد است. در منطقه ۴ (شکل ۲ د)، میانگین ردپای آب آبی، سبز و خاکستری به ترتیب ۱۲۷۶/۲، ۳۷۳/۵ و ۸۷۲/۵ مترمکعب بر تن و کل ردپای آب ۲۵۲۲ مترمکعب بر تن است. براساس نتایج، تغییرات ردپای آب هر سه مؤلفه در حال کاهش است.

به طور متوسط ردپای آب گندم در استان فارس، ۲۴۴۸ مترمکعب در هر تن است. نتایج آبابایی و رضانی اعتدالی [۲۵]. ۲۳۱۹ مترمکعب در هر تن در استان فارس تخمین زدند. مهم‌ترین دلیل بالا بودن ردپا در تحقیق حاضر، جزئی‌تر بودن مقیاس مورد بررسی است. تفاوت بین نیاز آبی، بارش مؤثر و تبخیر و تعرق گیاهی در هر منطقه همگن اگر و کلماتیکی نشان‌دهنده تغییرات قابل توجهی در سراسر استان فارس است و این وضعیت، ضرورت استفاده از مقیاس‌های کوچک‌تر جهت ارزیابی را اثبات می‌کند. منطقی است که داده‌های برآورده شده در سطح استانی، دارای دقت کمتری از برآوردها در مقیاس شهرستان یا منطقه همگن است. از طرفی، انواع مفروضات در مورد داده‌های ورودی و مدل‌سازی ردپای آب بر میزان ردپای آب سبز نیز مؤثر است. در خصوص آب آبی نیز باید خاطر نشان کرد که برآورد آن با داده‌های آبیاری حقیقی، سبب برآورد دقیق‌تری از ردپای آب خواهد شد. بنابراین، در مطالعات منطقه‌ای به‌رغم مطالعات استانی یا ملی، برآورد پارامترهای مختلف، از اعتبار بالاتر و عدم قطعیت کمتری برخوردارند.

در رابطه یاد شده  $Q_i$  نتایج واقعی و  $\hat{Q}_i$  نتایج محاسباتی و  $N$  تعداد متغیرها است. هر چه مقدار RMSE کمتر باشد، صحت پیشگویی افزایش می‌یابد. همچنین،  $R$  نشان‌دهنده میزان بازده شبکه است که بر اساس معادله‌های ۱۰ و ۱۱ بیان می‌شود:

$$R^2 = 1 - \frac{\sum(Q_o - Q_e)^2}{\sum(Q_o - Q_e)^2} \quad (10)$$

$$MAE = \frac{|Q_o - Q_e|}{N} \quad (11)$$

و MAE که میانگین قدرمطلق خطاست. در این روابط،  $Q_o$  داده مشاهده‌ای،  $Q_e$  داده تخمین زده شده است. هر چه RMSE و MAE به صفر و  $R$  به یک نزدیک‌تر باشد، عملکرد بهتری برای مدل حاصل شده است.

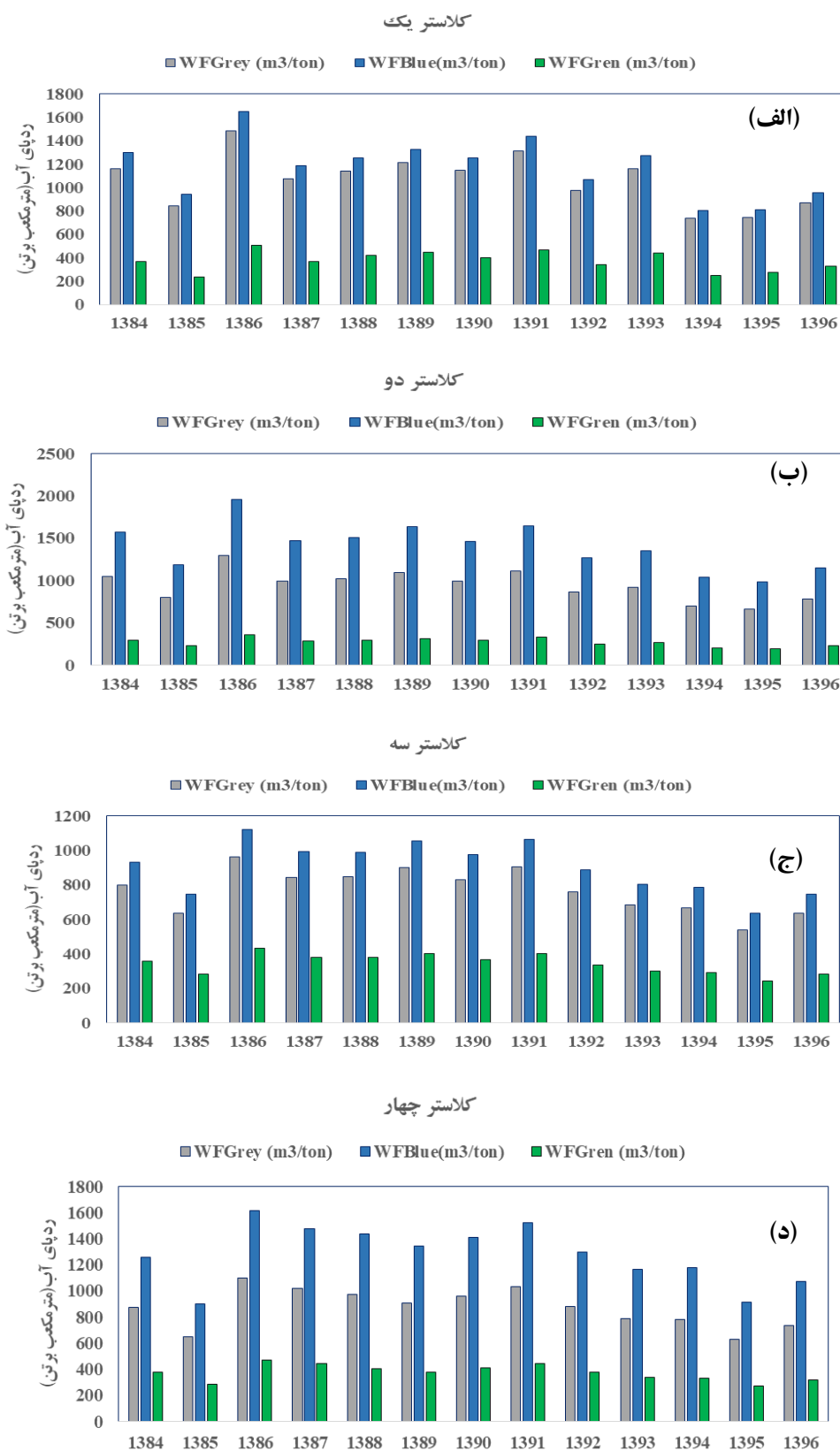
#### دیاگرام تیلور

دیاگرام تیلور به دو صورت نیم‌دایره (نمایش همبستگی منفی و مثبت) و ربع دایره (فقط نمایش همبستگی مثبت) ارائه می‌شود که در هر دو صورت، مقادیر ضریب همبستگی به صورت شعاع دایره روی قوس آن، مقادیر انحراف معیار به صورت دواير متحدالمرکز نسبت به مرکز دایره و مقادیر RMS به صورت دواير متحدالمرکز نسبت به نقطه مرجع روی محور افقی ترسیم می‌شود. هر چه موقعیت مکانی مدل‌ها به نقطه مرجع (مربع مشکی روی محور افقی) نزدیک‌تر باشند، عملکرد مدل بهتر است.

#### یافته‌ها

##### تغییرات ردپای آب در مناطق چهارگانه همگن استان فارس

تغییرات اجزای ردپای آب در هر منطقه در شکل ۲، نمایش داده شده است. شکل ۲ الف، تغییرات ردپای آب را در منطقه یک نشان می‌دهد. براساس نتایج میزان میانگین ردپای آب آبی، سبز و خاکستری به ترتیب ۱۱۷۴، ۳۷۳ و ۱۰۶۷/۵ مترمکعب بر تن برآورد شد. متوسط کل ردپای آب ۲۶۱۵ مترمکعب بر تن طی دوره آماری است. بنابراین، سهم هر یک از ردپاها به ترتیب ۴۵، ۱۴ و ۴۱ درصد است.



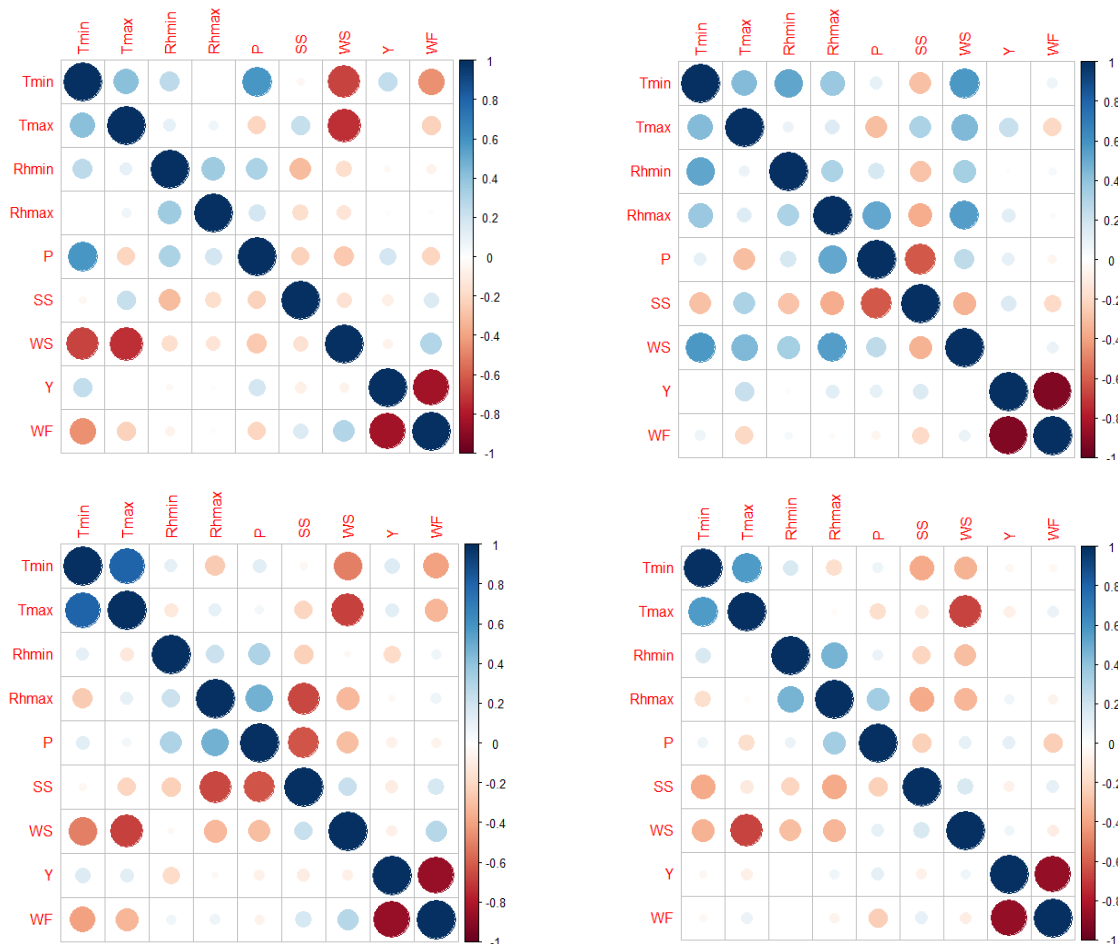
شکل ۲. تغییرات ردپای آب گندم فاریاب در مناطق مورد مطالعه (الف: کلاستر یک؛ ب: کلاستر ۲؛ ج: کلاستر ۳ و د: کلاستر ۴)



این کلاستر نشان می‌دهد افزایش دمای حداکثر، رطوبت حداکثر و ساعت‌های آفتابی سبب افزایش عملکرد می‌شود. عوامل اقلیمی یادشده مهم‌ترین عوامل مؤثر بر تبخیر و تعرق گیاهی و در نتیجه عملکرد است.

افزایش ساعت‌های آفتابی و در ادامه افزایش دما سبب افزایش تبخیر و تعرق و در نتیجه افزایش عملکرد شده [۱۷] و در نهایت میزان ردپا کاهش می‌یابد. در خصوص افزایش رطوبت هم می‌توان خاطر نشان کرد سبب کاهش تبخیر، کاهش نیاز آبی و در نتیجه، کاهش ردپای آبی و کاهش ردپای کل می‌شود [۱۸].

بررسی ارتباط بین متغیرهای مورد استفاده و ردپای آب شکل ۳ هیت مپ همبستگی بین متغیرها را به صورت دو به دو نشان می‌دهد. تقریباً در چهار کلاستر شرایط مشابهی از اثرات عوامل مؤثر بر ردپا دیده می‌شود. براساس شکل، بیشترین همبستگی بین عملکرد و WF اما با علامت منفی است. به این معنا که با افزایش عملکرد، ردپای آب کاهش می‌یابد. کمترین همبستگی بین رطوبت حداقل و حداکثر دیده می‌شود. در بررسی ارتباط عوامل اقلیمی بر ردپای آب، ساعت‌های آفتابی و دمای حداکثر ارتباط معکوس و دمای حداقل ارتباط مستقیم نسبتاً بالایی دارد. همچنین، در بررسی ارتباط عملکرد با سایر متغیرهای اقلیمی، نتایج در



شکل ۳. هیت مپ همبستگی بین متغیرهای مورد بررسی

(ANN-TH) نشان می‌دهد. براساس تعداد ورودی‌های مختلف و ترکیبات مختلف خطای شبکه و خطا در مرحله صحت‌سنجی ارائه شده است. در این تحقیق از تعداد ۱۰ متغیر شامل عملکرد (Y) محصول در واحد سطح، سرعت باد

نتایج حاصل از مدل‌سازی ردپای گندم با شبکه عصبی مصنوعی

جدول ۱، نتایج حاصل از کاربرد شبکه عصبی مصنوعی را با دو تابع لوگ لوجستیک (ANN-LL) و تانژانت هاپرپولیک

MAE ۰/۰۰۰۵ و ۰/۱۸ است در حالی که در ANN-TH، تنها با کاربرد چهارمتغیر عملکرد، سرعت باد، ساعت‌های آفتابی و بارش مؤثر، خطای شبکه افزایش، همبستگی افزایش و خطای مدل کاهش یافته است.

در کلاستر ۴ نیز هر دو شبکه از تمام متغیرهای ورودی استفاده کرده و ANN-LL دارای عملکرد بهتری در برآورد ردپای آب بوده است. شکل ۲، یک نمونه ساختار شبکه را در کلاستر ۱ نمایش می‌دهد.

نتایج البتقی و همکاران [۲۶] در پیش‌بینی عوامل مؤثر بر ردپای آب نشان داد، بهترین ترکیب داده، کاربرد تمام متغیرهای اقلیمی است. اگرچه همگی به صورت مستقیم در رابطه ردپای آب کاربرد ندارند، اما همگی بر نیاز آبی، بارش مؤثر، تبخیر و تعرق گیاهی و عملکرد مؤثرند.

#### نتایج حاصل از مدل‌سازی ردپای گندم با RF

در این روش با در نظر گرفتن ۵۰ ترکیب از متغیرهای اقلیمی و عملکرد گیاهی مورد استفاده، مؤثرترین متغیرها توسط میانگین کاهشی جینی و مجموع مربعات خطا و معیارهای ارزیابی خطا در مرحله صحت‌سنجی مدل در شکل ۵ و جدول ۲ ارائه شد. براساس نتایج در کلاستر ۱، مهم‌ترین متغیر، عملکرد بوده که دارای کمترین خطا و بیشترین میانگین کاهشی ضریب جینی بوده، میزان ضریب همبستگی برآوردشده توسط مدل RF در منطقه ۱، ۰/۹۷ است. در منطقه ۲، ترکیب تمام متغیرها بهترین عملکرد؛ در منطقه ۳، همه متغیرها جز دمای حداقل و در منطقه ۴ تمام ده متغیر در مدل‌سازی دارای بهترین ضریب همبستگی و کمترین خطا هستند.

(WS)، ساعت‌های آفتابی (SS)، بارش مؤثر (P)، دمای حداکثر (Tmax)، دمای حداقل (Tmin)، دمای متوسط (T)، رطوبت حداکثر (Rhmax)، رطوبت حداقل (Rhmin) و طوبت متوسط (R) به صورت ترکیب‌های مختلف از یک تا ۵۰ استفاده شد. همچنین، تعداد نوروها از یک تا پنج با ترکیب‌های مختلف ورودی و خروجی مورد مقایسه قرار گرفت. تعداد ۷۰ درصد داده‌ها در مرحله آموزش و ۳۰ درصد داده‌ها برای مرحله آزمایش مورد استفاده قرار گرفت.

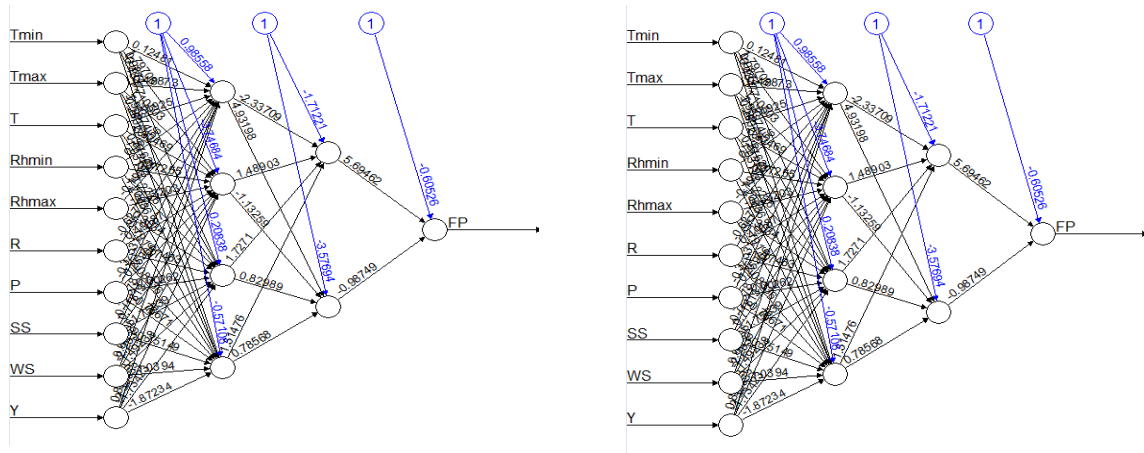
براساس نتایج، در کلاستر یک در مدل ANN-LL، با تعداد چهار نورون ورودی و دو نورون خروجی و ورودی‌های عملکرد، سرعت باد، ساعت‌های آفتابی، بارش مؤثر، رطوبت حداکثر و متوسط دارای کمترین خطای شبکه (Bias=0.26)، ضریب همبستگی ۰/۹۶، RMSE و MAE ۰/۰۰۵ و ۰/۲۳ (مترمکعب بر تن) است. در همین کلاستر، مدل ANN-TH نیز با همان تعداد نوروها، متغیرهای دمای حداکثر و رطوبت حداقل را نیز اضافه کرده و سبب کاهش خطای شبکه (Bias=0.11)، ولی کاهش ضریب همبستگی و افزایش خطا شده است که شاید یکی از دلایل آن تفاوت ساختار دو شبکه در شبیه‌سازی متغیر پاسخ باشد.

در کلاستر ۲ با کاربرد ده متغیر ورودی با تعداد چهار نورون ورودی و دو نورون خروجی، خطای شبکه ۰/۰۸ دارای ضریب همبستگی ۰/۸۱ در مدل ANN-LL و در مدل ANN-TH با همین تعداد نوروها، خطای شبکه کاهش و ضریب همبستگی کاهش می‌یابد.

در کلاستر ۳، مدل ANN-LL با کاربرد تمامی متغیرهای ورودی، با ترکیب دو و چهار نورون، دارای خطای شبکه ۰/۳۱، ضریب همبستگی ۰/۹۶ و RMSE

جدول ۱. نتایج شبکه عصبی مصنوعی

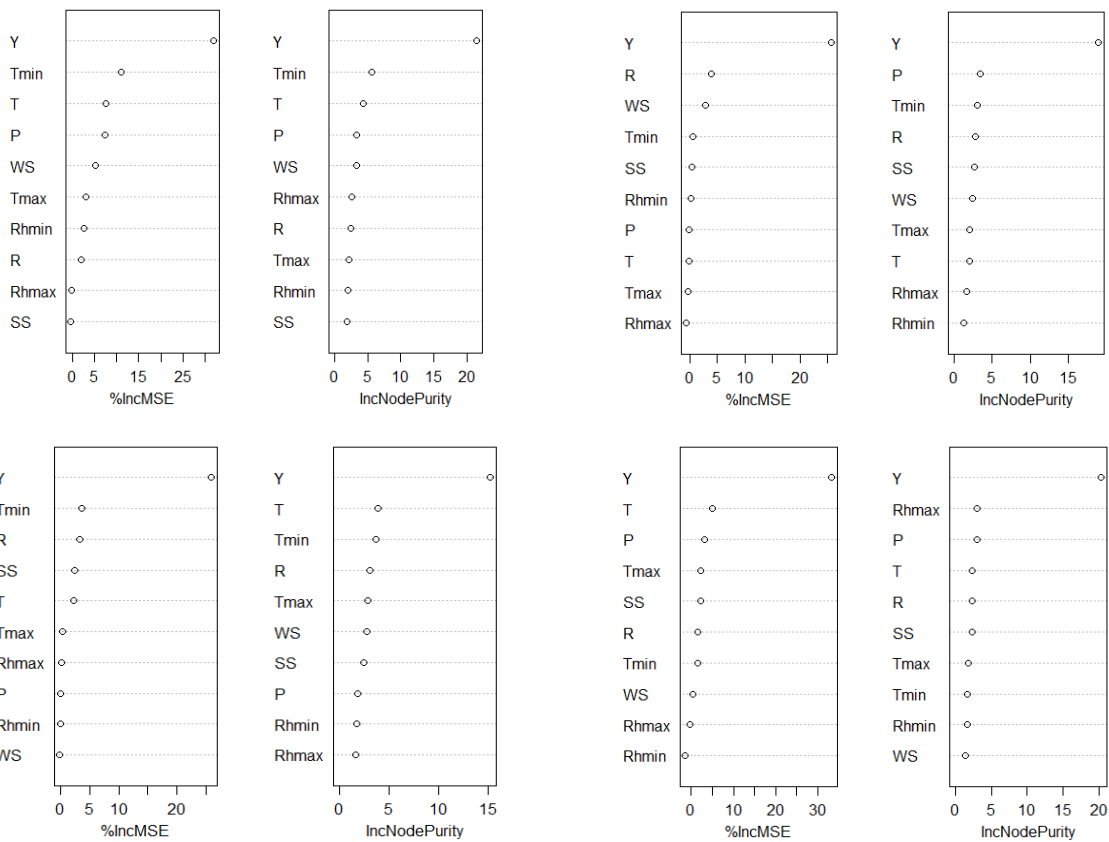
نام شبکه	منطقه	متغیرهای ورودی	نورون‌های ورودی - خروجی	خطای شبکه	RMSE	R	MAE
ANN-LL	۱	Y, WS, SS, P, R, Rhmax	۲-۴	۰/۲۶	۰/۰۰۵	۰/۹۶	۰/۲۳
	۲	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax, Tmin	۲-۴	۰/۰۸	۰/۰۰۳	۰/۸۱	۰/۴
	۳	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax, Tmin	۴-۲	۰/۳۱	۰/۰۰۰۵	۰/۹۶	۰/۱۸
	۴	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax, Tmin	۳-۳	۰/۳	۰/۰۰۰۴	۰/۹۴	۰/۳۰
ANN-TH	۱	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax	۲-۴	۰/۱۱	۰/۰۰۰۱	۰/۹۳	۰/۲۴
	۲	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax, Tmin	۴-۲	۰/۰۸	۰/۰۰۰۷	۰/۸۶	۰/۴۸
	۳	Y, WS, SS, P	۳-۳	۰/۳۸	۰/۰۰۰۶	۰/۹۷	۰/۱۶
	۴	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax	۵-۱	۰/۱۲	۰/۰۰۱	۰/۷۲	۰/۴۸



شکل ۴. نمایی از شبکه عصبی مصنوعی با تابع لوگ لجستیک (راست) و تانژانت هایپربولیک (چپ) در کلاستر ۱ با ۴ نرون ورودی و ۲ نرون خروجی

جدول ۲. نتایج مدل جنگل تصادفی در کلاستر ۱

منطقه	متغیرهای ورودی	R	RMSE	MAE
۱	Y	۰/۹۷	۰/۰۰۲	۰/۱۲
۲	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax, Tmin	۰/۸۷	۰/۰۰۴	۰/۴۰
۳	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax	۰/۹۱	۰/۰۱	۰/۳۰
۴	Y	۰/۹۳	۰/۰۰۴	۰/۲۷



شکل ۵. اهمیت نسبی متغیرها براساس میانگین کاهشی جینی و دقت در منطقه مورد بررسی منطقه ۱ (الف)؛ منطقه ۲ (ب)؛ منطقه ۳ (ج)؛ منطقه ۴ (د)

جدول ۳. ترکیب مدل بهینه در مدل ماشین بردار پشتیبان در مناطق مورد بررسی

منطقه	متغیرهای ورودی	R	RMSE	MAE
۱	Y, WS, SS, P, R, Rhmax, Rhmin, T	۰/۳۷	۰/۰۰۵	۰/۸۶
۲	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax, Tmin	۰/۸۹	۰/۰۱	۰/۳۷
۳	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax, Tmin	۰/۸	۰/۰۰۸	۰/۳۲
۴	Y, WS, SS, P, R, Rhmax, Rhmin, T, Tmax, Tmin	۰/۸۳	۰/۰۱	۰/۴۳

در منطقه ۳، مدل ANN-LL با در نظر گرفتن تمامی متغیرهای اقلیمی مؤثر بر ردپا، با تعداد ۲ نورون ورودی و چهار نورون خروجی با خطای شبکه ۰/۳۱، MAE و RMSE ۰/۱۸ و ۰/۰۰۵ و همبستگی بالای بین مقادیر مشاهده‌ای و پیش‌بینی شده (۰/۹۶) بهترین مدل و در نهایت در منطقه ۴، نیز همین مدل، با تعداد دو نورون ورودی و چهار نورون خروجی با خطای شبکه ۰/۳۱، MAE و RMSE ۰/۱۸ و ۰/۰۰۵ و همبستگی بالای بین مقادیر مشاهده‌ای و پیش‌بینی شده (۰/۹۶) بهترین مدل معرفی شد.

فتحی‌زاده و حکیم‌زاده اردکانی [۲۷] در پیش‌بینی پارامترهای کیفیت آب؛ نوروزی قوشبلاق و همکاران [۲۸] در پیش‌بینی قابلیت آبخوان؛ مختار و همکاران [۲۹] در پیش‌بینی ردپای آب کارایی شبکه عصبی پرسپترون چندلایه را نسبت به مدل‌های RF و SVR اعلام کردند. در خصوص کارایی شبکه عصبی مصنوعی در مدل‌سازی واضح است که توانایی‌های این تکنیک با توجه به ساختارهای مختلف آن و نیز طبیعت مسئله مورد نظر، متفاوت است. شبکه عصبی پرسپترون چندلایه با انتخاب نوع و تعداد مناسب عوامل ورودی و نیز سبب ارائه بهترین پاسخ می‌شود و در پیش‌بینی متغیرهای هیدرومتئورولوژیکی ابزاری کارآمد محسوب می‌شود.

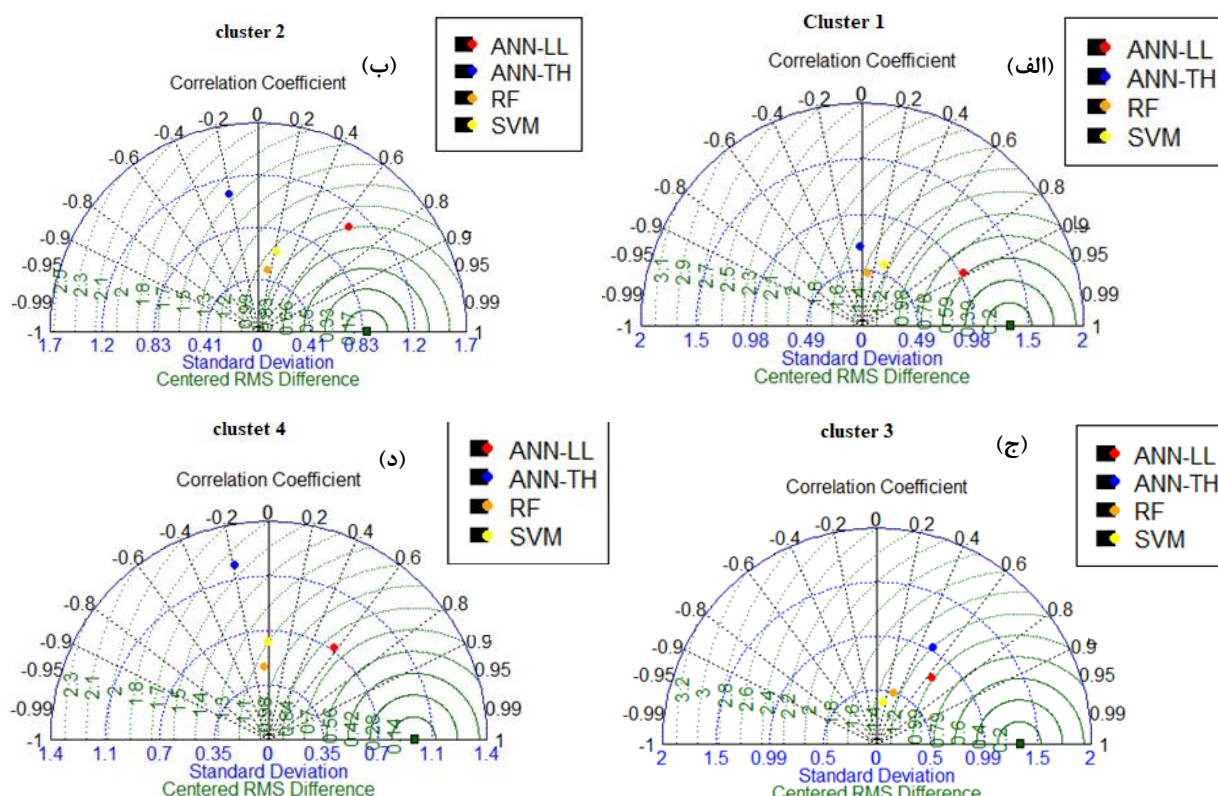
محمدی و همکاران [۳۰] در پیش‌بینی بارش-رواناب؛ سیاسر و هنر [۳۱] در پیش‌بینی تبخیر و تعرق مرجع؛ قاسمی و همکاران [۳۲] در پیش‌بینی تاج پوشش مانگرو اعلام کردند، مدل FR از کارایی بهتری نسبت به SVR برخوردار است. از مهم‌ترین دلایل برتری آن مکانیزم انتخاب داده تصادفی با کاربرد روش بوت استرپ است، چرا که این روش ترکیبی از چندین درخت تصمیم است که در ساخت آن چندین نمونه بوت استرپ از داده‌ها شرکت دارند و در ساخت هر درخت به طور تصادفی تعدادی از متغیرهای ورودی شرکت می‌کنند.

### نتایج حاصل از مدل‌سازی ردپای گندم با SVM

در این روش نیز تعداد ۵۰ ترکیب متغیرهای اقلیمی وارد مدل شد و در هر کلاستر براساس معیارهای ارزیابی خطا، بهترین ترکیب در مرحله صحت‌سنجی مدل در جدول ۳ ارائه شد. براساس نتایج در منطقه ۱، متغیرهای عملکرد سرعت باد، ساعت‌های آفتابی، بارش، رطوبت و دمای متوسط مهم‌ترین متغیرهای مؤثر بر ردپای آب هستند که دارای همبستگی ۰/۳۷۰، RMSE و MAE ۰/۰۰۵ و ۰/۸۶ هستند. در سه منطقه دیگر تمام ۱۰ متغیر مورد بررسی بهترین ترکیب ورودی با همبستگی بالا و خطای پایین بوده است. در این مناطق محدوده تغییرات همبستگی ۰/۸ تا ۰/۸۹ است.

### ارزیابی کارایی مدل‌ها در کلاسترهای مختلف با دیگرام تیلور

شکل ۶ دیگرام تیلور در کلاستر ۱ را نشان می‌دهد. همان‌طور که پیش‌تر گفته شد، در دیگرام تیلور، هر قدر نقطه مدل به رفرنس پوینت یا نقطه مرجع نزدیک‌تر باشد، نشان‌دهنده دقت بالاتر مدل است. در کلاستر یک، مدل ANN-LL نزدیک‌ترین نقطه به نقطه مرجع است. به این مفهوم که بیشترین دقت و شباهت را به مقادیر مشاهده‌ای داشته است. این مدل با در نظر گرفتن تمامی متغیرهای اقلیمی مؤثر بر ردپا، با تعداد سه نورون ورودی و سه نورون خروجی با خطای شبکه ۰/۳، MAE و RMSE ۰/۰۰۴ و ۰/۰۰۴ و همبستگی بالای بین مقادیر مشاهده‌ای و پیش‌بینی شده (۰/۹۴) بهترین مدل معرفی شد. در مدل ANN-TH با تعداد ۲ نورون ورودی و ۴ نورون خروجی میزان خطای شبکه ۰/۱۱، MAE و RMSE ۰/۲۴ و ۰/۰۰۱ و همبستگی ۰/۹۳ بوده است. در منطقه ۲، مدل ANN-TH با در نظر گرفتن تمامی متغیرهای اقلیمی مؤثر بر ردپا، با تعداد چهار نورون ورودی و دو نورون خروجی با خطای شبکه ۰/۰۸، MAE و RMSE ۰/۴ و ۰/۰۰۳ و همبستگی بالای بین مقادیر مشاهده‌ای و پیش‌بینی شده (۰/۸۱) بهترین مدل معرفی شد.



شکل ۶. دیاگرام تیلور در مناطق مورد بررسی: منطقه ۱ (الف); منطقه ۲ (ب); منطقه ۳ (ج); منطقه ۴ (د)

### بحث و نتیجه‌گیری

این تحقیق با هدف پیش‌بینی ردپای آب در گندم فاریاب در استان فارس صورت گرفته است. بر این اساس، از ۳ مدل یادگیری ماشین و برآورد ردپای آب با چهارچوب هوکسترا و چاپاگین [۱۵] و [۱۶] استفاده شد. ورودی‌های مدل، شامل ۵۰ ترکیب از متغیرهای اقلیمی و عملکرد گیاهی بودند. پارمترهای اقلیمی شامل متوسط دما، متوسط بارش، حداکثر دما، حداقل دما، ارتفاع از سطح دریا، ساعات آفتابی، متوسط رطوبت، و متوسط عملکرد در واحد سطح در استان فارس در پهنه ۲۲ شهرستان برای خوشه‌بندی مناطق مشابه استفاده شد. بنابراین، براساس روش خوشه‌بندی فازی شهرستان‌ها در استان فارس به چهار خوشه همگن تقسیم شدند و پیش‌بینی در این چهار منطقه صورت پذیرفت.

ردپای آب، بر اجزای مصرف آب شامل آب سبز، آبی، سفید و خاکستری دلالت دارد که شامل فرایند مصرف آب از زمان کاشت تا برداشت است. علاوه بر آن، کل مصرف آب یک محصول شامل آب آبی و آب سبز است که هر یک اثرات و هزینه‌های متفاوتی بر آب در دسترس دارند. مصرف آب سبز در تولید محصولات کشاورزی از آب باران

است که روی محیط زیست اثرات سوئی ندارد و در مقابل آب آبی از آب‌های سطحی و یا زیرزمینی است که اثرات آن بر محیط زیست مشهود است [۳۳]. بنابراین، WF به مصرف مستقیم و غیرمستقیم آب در تولید محصولات کشاورزی و اثرات زیست‌محیطی ناشی از آن دلالت دارد که کاربرد آن می‌تواند ارزیابی جامع‌تری از بهره‌وری آب زراعی، تحلیل توزیع زمانی و مکانی کشت داشت [۳۴].

نتایج بررسی ردپای آب در اجزای آن نشان داد، در تمام مناطق همگن مورد بررسی در استان فارس، میزان ردپای آب آبی بیش از ۵۰ درصد ردپای کل و ردپای سبز کمتر از ۱۱ درصد را به خود اختصاص می‌دهد. ردپای آبی، ناشی از استخراج آب‌های زیرزمینی و یا جمع‌آوری رواناب‌های سطحی است و ردپای سبز ناشی از بارش مؤثر. بنابراین، به‌کارگیری روش‌های افزایش راندمان آبیاری و افزایش عملکرد در واحد سطح می‌تواند سبب کاهش ردپای آب شود. سهم ردپای آب خاکستری در تولید گندم در فارس به طور متوسط بیش از ۳۴ درصد برآورد شد که بیشترین سهم مربوط به خوشه ۱ است. عمده‌ترین دلیل سهم قابل توجه آب خاکستری در این خوشه، عملکرد پایین، مصرف

- [5]. Moni, S., Aziz, E., Majeed, A. P. A., & Malek, M. The prediction of blue water footprint at Semambu water treatment plant by means of Artificial Neural Networks (ANN) and Support Vector Machine (SVM) models. *Physics and Chemistry of the Earth, Parts A/B/C*, 2021; 123, 103052.
- [6]. Mokhtar, A., Jalali, M., He, H., Al-Ansari, N., Elbeltagi, A., Alsafadi, K.,... & Rodrigo-Comino, J. Estimation of SPEI Meteorological Drought Using Machine Learning Algorithms. *IEEE Access*, 2021a; 9, 65503-65523.
- [7]. deghghan pir, S., bazrafshan, O., hlizadeh, A. Virtual Water Trade and Use in Watershed: (Case study: Baraftab-E Hajiabad and Payab- E Rudan watersheds, Hormozgan Province). *Journal of Range and Watershed Management*, 2017; 70(3), 647-660. [Persian]
- [8]. Shiff, S., Lensky, I. M., & Bonfil, D. J. Using satellite data to optimize wheat yield and quality under climate change. *Remote Sensing*, 2021; 13(11), 2049.
- [9]. Green, H. Wheat importance, history and adaptation. *Theoretical and practical new approach in cereal science and technology*, 2021; 3.
- [10]. Tavanpour, N., & Ghaemi, A. A. Zoning of Fars Province in terms of rain-fed winter wheat cultivation based on precipitation and morphological factors. *Iranian Journal of Irrigation & Drainage*, 2016; 10(4), 544-555.
- [11]. Srinivas, C. V., , Hariprasad, D. , Bhaskar Rao, D. V. , Anjaneyulu, Y. , Baskaran, R. , Venkatraman B. Simulation of the Indian summer monsoon regional climate using advanced research WRF model. *International Journal of Climatology*, 2013; 33(5), 1057-1320.
- [12]. Bhatia, V S and Singh, P and Wani, S P and Rao, A V R K and Srinivas, K. Yield Gap Analysis of Soybean, Groundnut, Pigeonpea and Chickpea in India Using Simulation Modeling: Global Theme on Agroecosystems Report no. 31. Monograph. International Crops Research Institute for the Semi-Arid Tropics, 2006.
- [13]. Žalik, K. R., Žalik, B. Validit index for clusters of different sizes and densities. *Pattern Recognition Letters*, 2011; 32(2), 221-234, ISSN 0167-8655,
- [14]. Bezdek, J. C., Coray, C., Gunderson, R. and Watson J. Detection and Characterization of Cluster Substructure I. Linear Structure: Fuzzy c-Lines. *SIAM Journal on Applied Mathematics*, 1981; 40(2) 10.1137/0140029.

بالای کود شیمیایی و آب‌شویی بالای کود در این منطقه است. بنابراین، با مصرف بهینه کود شیمیایی از نظر زمانی و مقدار آن و همچنین انتخاب ژنوتیپ‌های سازگارتر با شرایط اقلیمی، کاهش سهم آب خاکستری در این منطقه قابل حصول است.

نتایج حاصل از کاربرد مدل‌های یادگیری ماشین نشان داد مدل ANN-LL در تمام کلاسترها دارای کمترین خطا و بیشترین ضریب همبستگی بین مقادیر مشاهده‌شده و پیش‌بینی دارد و بعد از آن مدل ANN-TH در مرتبه بعدی است. بعد از شبکه عصبی مصنوعی مدل RF دارای کارایی قابل قبولی است، اما مدل SVM در این تحقیق فاقد نتایج مناسب بود. به دلیل ماهیت غیرخطی و رفتار پیچیده عوامل مؤثر بر ردپای آب، شبکه عصبی مصنوعی پرسپترون چندلایه به دلیل ساختار قدرتمندی که دارد، دارای عملکرد مناسب‌تری است.

نتایج یافته‌های این تحقیق می‌تواند، یک ابزار تصمیم‌گیری برای برنامه‌ریزان و بهره‌برداران برای رسیدن به مدیریت پایدار آب در بخش کشاورزی، به‌ویژه در مناطق خشک و نیمه‌خشک باشد. از مهم‌ترین چالش‌های این تحقیق کمبود داده‌های اقلیمی و کشاورزی به‌ویژه از نظر طول دوره آماری است. بی‌شک، افزایش طول دوره، برای درک بهتر مکانیسم تغییرات در آینده کمک شایانی خواهد کرد.

#### منابع

- [1]. Oveisi F, Fattahi Ardakani A, Fehresti Sani M. Investigation of Virtual Water and Ecological Footprints of Water in Wheat Fields of Isfahan Province. *JWSS*, 2019; 23 (1) :87-99. [Persian]
- [2]. Bazrafshan, O., Zamani, H., Etedali, H. R., & Dehghanpir, S. Assessment of citrus water footprint components and impact of climatic and non-climatic factors on them. *Scientia Horticulturae*, 2019; 250, 344-351.
- [3]. Elbeltagi, A., Aslam, M. R., Malik, A., Mehdinejadiani, B., Srivastava, A., Bhatia, A. S., & Deng, J. The impact of climate changes on the water footprint of wheat and maize production in the Nile Delta, Egypt. *Science of the Total Environment*, 2020a; 743, 140770.
- [4]. Elbeltagi, A., Deng, J., Wang, K., & Hong, Y. Crop Water footprint estimation and modeling using an artificial neural network approach in the Nile Delta, Egypt. *Agricultural Water Management*, 2020b; 235, 106080.

- [15]. Hoekstra, A. Y. Water Neutral: Reducing and of Setting the Impacts of Water Footprints. Value of Water Research Report Series, NO. 28, Delft, the Netherlands: Unesco-IHE Institute for Water Education, 2008.
- [16]. Chapagain, A., & Orr, S. UK Water Footprint: the impact of the UK's food and fibre consumption on global water resources Volume two: appendices. WWF-UK, Godalming, 2008; 31-33.
- [17]. Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage, 1998; 300(9), 56.D05109.
- [18]. Bazrafshan, o., ramezani etedali, h., garkani nezhad moshizi, z., shamili, m. Virtual water trade and water footprint accounting of Saffron production in Iran. *Agricultural Water Management*, 2019; 213 368–374.
- [19]. Bazrafshan, o., Zamani, h., Ramezani etedali, h., garkani nezhad moshizi, z., shamili, m., ismaelpour, y., gholami, h. Improving water management in date palms using economic value of water footprint and virtual water trade concepts in Iran. *Agricultural Water Management*, 2020; 229, 10594.
- [20]. Ravi, V., Pradeepkumar, D., Deb, K. Financial time series prediction using hybrids of chaos theory, multi-layer perceptron and multi-objective evolutionary algorithms. *Swarm and Evolutionary Computation*, 2017; 36, 136-149
- [21]. Khashei M., Hajirahimi Z. A comparative study of series arima/mlp hybrid models for stock price forecasting, 2019; 2625-2640.
- [22]. Friedman, J. H. Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 2001; 29(5), 1189–1232.
- [23]. Jahanbakhshi F, Ekhtesasi M R. Performance Evaluation of Three Image Classification Methods (Random Forest, Support Vector Machine and the Maximum Likelihood) in Land Use Mapping. *JWSS*, 2019; 22 (4) :235-247.
- [24]. Ahmadpour, H., Bazrafshan, O., Rafiei-Sardooi, E., Zamani, H., & Panagopoulos, T. Gully Erosion Susceptibility Assessment in the Kondoran Watershed Using Machine Learning Algorithms and the Boruta Feature Selection. *Sustainability*, 2021; 13(18), 10110.
- [25]. Ababaei, B., Ramezani Etedali, H. Estimation of Water Footprint Compartments in National Wheat Production. *Water and Soil*. 2016; 29(6), 1458-1468.. [Persian]
- [26]. Elbeltagi, A., Azad, N., Arshad, A., Mohammed, S., Mokhtar, A., Pande, C., R.Etedali, H., A.Bhat, S., Abu Reza, Md., Islam, T., Deng, J. Applications of Gaussian process regression for predicting blue water footprint: Case study in Ad Daqahliyah, Egypt. *Science Direct*, 2021; 107052.
- [27]. Fathizad, H., Hakimzadeh Ardakani, M. Evaluation of the Efficiency of Artificial Neural Network and Random Forest Models in Predicting Groundwater Quality Parameters of Yazd-Ardakan Plain. *Integrated Watershed Management*, 2022; 1(2), 1-19. [Persian]
- [28]. Norouzi, H., Nadiri, A., Asghari Moghaddam, A., Norouzi, M. Comparing Performans of Fuzzy Logic, Artificial Neural Network and Random Forest Models in Transmissivity Estimation of Malekan Plain Aquifer. *Iranian journal of Ecohydrology*, 2018; 5(3), 739-751. [Persian]
- [29]. Mokhtar, A., Elbeltagi, A., Maroufpoor, S., Azad, N., He, H., Alsafadi, K.,... & He, W. Estimation of the rice water footprint based on machine learning algorithms. *Computers and Electronics in Agriculture*, 2021b; 191, 106501.
- [30]. Mohammadi, M., Vagharfard, H., Mahdavi Najafabadi, R., Daneshkar Arasteh, P., Nazemosadat, M. Rainfall-runoff Modelling of Coastal Watersheds near Hormuz Strait Using Data Mining. *Iranian Journal of Soil and Water Research*, 2021; 52(2), 313-327. [Persian]
- [31]. siasar, H., honar, T. Application of Support vector machine, CHAID and Random forest models, in estimated daily Reference evapotranspiration in northern Sistan and Baluchestan province. *Iranian Journal of Irrigation & Drainage*, 2019; 13(2), 378-388.
- [32]. Ghasemi, A., Fallah, A., Shataee Joibari, S. Evaluation of four algorithms for estimation of canopy cover of mangrove forests by using aerial imagery. *Journal of RS and GIS for Natural Resources*, 2016; 7(2), 1-16. [Persian]
- [33]. Ridoutt, B. G. and Pfister S. Reducing humanity's water footprint. *Environmental Science & Technology*, 2010; 44 (16), 6019-6021.
- [34]. Zhao, P., Lu, D., Wang, G., Liu, L., Li, D., Zhu, J., Yu, S. (2016). Forest aboveground biomass estimation in Zhejiang Province using the integration of Landsat TM and ALOS PALSAR data. *International Journal of Applied Earth Observation and Geoinformation*, 2016; 53(10), 1-15.