

پیش‌بینی سطح آب زیرزمینی دشت بوکان با استفاده از مدل‌های منطق فازی، جنگل تصادفی و شبکه عصبی

❖ حسین نوروزی قوشبلاغ*؛ دانشجوی دکتری هیدروژئولوژی، دانشکده علوم طبیعی، دانشگاه تبریز، تبریز، ایران

❖ عطاالله ندیری؛ دانشیار هیدروژئولوژی، دانشکده علوم طبیعی، دانشگاه تبریز، تبریز، ایران

چکیده

مطالعات یک سیستم آب زیرزمینی به منظور شناخت رفتار آن، نیازمند حفر تعدادی زیادی چاه اکتشافی و انجام عملیات پمپاژ و آزمایش‌های ژئوفیزیک می‌باشد که با صرف هزینه‌های فراوان عملی می‌گردد. به همین دلیل، امروزه شبیه‌سازی جریان آب زیرزمینی توسط مدل‌های ریاضی و کامپیوتری که یک روش غیرمستقیم مطالعه آب زیرزمینی می‌باشد، با صرف هزینه کمتر صورت می‌گیرد. در این تحقیق کارایی مدل‌های شبکه‌های عصبی مصنوعی، منطق فازی و جنگل تصادفی (RF) در تخمین سطح آب زیرزمینی آبخوان دشت بوکان مورد بررسی قرار گرفت. پارامترهای بارندگی، دما، دبی جریان و تراز سطح ایستابی در دوره زمانی ماه قبل به‌عنوان ورودی و تراز سطح ایستابی در دوره موردنظر به‌عنوان خروجی مدل‌ها در مقیاس زمانی ماهانه در طی دوره آماری (۱۳۸۳-۱۳۹۵) انتخاب گردید. معیارهای ضریب همبستگی، ریشه میانگین مربعات خطا و میانگین قدر مطلق خطا برای ارزیابی و نیز مقایسه عملکرد مدل‌ها مورد استفاده قرار گرفت. نتایج حاصله نشان داد که مدل‌های منطق فازی و RF می‌توانند تراز سطح ایستابی را با دقت قابل قبولی پیش‌بینی نمایند. سطح آب زیرزمینی در منطقه مورد مطالعه توسط مدل جنگل تصادفی با دقت بیشتر و با همبستگی بالایی بین داده‌های مشاهداتی و محاسباتی پیش‌بینی گردید. بر اساس یافته‌های تحقیق مدل جنگل تصادفی که برای اولین بار در این زمینه مورد استفاده قرار گرفته است، توانایی بیشتری در پیش‌بینی پارامترهای هیدروژئولوژیکی را دارد.

کلید واژگان: سطح ایستابی، شبکه‌های عصبی، منطق فازی، جنگل تصادفی

۱. مقدمه

بهره‌برداری بی‌رویه از منابع آب زیرزمینی در سال‌های اخیر باعث به هم خوردن تعادل طبیعی آن شده و تراز آب زیرزمینی در آبخوان‌های بسیاری از نقاط کشور منفی شده است. کاهش میزان آب‌های زیرزمینی، به عبارت دیگر استفاده بیش از حد از منابع آب‌های زیرزمینی به دلیل مدیریت نامناسب در منطقه عمدتاً ناشی از روش‌های آبیاری غرقابی با کمبود کارایی و مصرف بیشتر آب، حفاری چاه‌های عمیق به دلیل افزایش در سطح زیر کشت می‌تواند باعث تخریب سرزمین گردد [۳۳].

به‌منظور آگاهی از وضعیت این منابع و مدیریت بهینه آن‌ها، لازم است پیش‌بینی دقیقی از نوسانات تراز آب زیرزمینی صورت گیرد. بیشتر سری‌های زمانی هیدرولوژیکی نظیر تغییرات سطح آب زیرزمینی همواره شامل فرایندهای گذرا و پیچیده هستند که با استفاده از مدل‌های خطی متداول و کلاسیک به‌خوبی قابل توصیف و مدل‌سازی نیستند؛ بنابراین به‌منظور مدل‌سازی این پدیده‌های هیدرولوژیکی لازم است از مدل‌های غیرخطی استفاده کرد. امروزه سیستم‌های هوشمند به‌طور گسترده برای پیش‌بینی پدیده‌های غیرخطی مورد استفاده قرار می‌گیرد.

عدم وجود داده‌های کافی برای توسعه مدل‌های ریاضی به علت محدودیت‌های زمانی و هزینه‌ای و همچنین زمان زیاد لازم جهت اجرای مکرر این مدل‌ها خصوصاً در شرایط ترکیب با مدل‌های بهینه‌سازی به منظور یافتن سناریوی بهینه مدیریت آب زیرزمینی، کاربرد مدل‌های ریاضی را با محدودیت مواجه می‌کند [۲۶]. در چنین شرایطی، مدل‌های مبتنی بر هوش مصنوعی، ماشین‌های یادگیری^۱ و داده‌کاوی^۲ می‌توانند جایگزین مناسبی به شمار آیند. در این نوع شبکه‌ها، با انجام پردازش‌های موازی روی داده‌های تجربی، دانش یا قانون نهفته در داده‌ها به ساختار شبکه منتقل شده و بر

اساس محاسبات روی داده‌های عددی یا مثال‌ها، قوانین کلی آن‌ها فرا گرفته می‌شود [۲۸]. پژوهش‌های زیادی برای شبیه‌سازی آبخوان به کمک روش‌های هوش مصنوعی صورت گرفته است. دالیاکوپولوس و همکاران [۱۲]، با استفاده از شبکه‌های عصبی مصنوعی با معماری پیشرو و الگوریتم آموزش لونبرگ-مارکوارت، مقادیر سطح آب زیرزمینی دره مسارا در جزیره کرت در جنوب یونان را تا ۱۸ ماه آینده پیش‌بینی کردند. فنگ و همکاران [۱۷] با کمک شبکه‌های عصبی مصنوعی پیشرو با الگوریتم پس انتشار گرادیان مزدوج، تغییرات سطح آب زیرزمینی منطقه مین‌کین در شمال غربی چین را شبیه‌سازی و پیش‌بینی کردند. پورمحمدی و همکاران [۲] کارایی تکنیک شبکه عصبی مصنوعی و سری‌های زمانی (آریمما) را در پیش‌بینی سطح ایستابی آب زیرزمینی در یکی از زیرحوضه‌های بختگان مورد بررسی قرار دادند. نتایج بیان گر کارایی و دقت بالای هر دو تکنیک در پیش‌بینی سطح ایستابی چاه‌های منطقه بود. نتایج پژوهش خاشعی سیوکی و همکاران [۳] نشان داد که دقت مدل‌های شبکه‌های عصبی مصنوعی و ANFIS در پیش‌بینی سطح آب زیرزمینی دشت نیشابور نسبت به مدل‌های رگرسیونی بالاتر بود. نتایج شبیه‌سازی و پیش‌بینی سطح آب زیرزمینی با دو مدل عددی (مادفلو) و شبکه‌های عصبی مصنوعی در منطقه اودیشا هندوستان نشان داد که پیش‌بینی‌های مدل شبکه‌های عصبی مصنوعی در بازه زمانی کوتاه مدت از مادفلو بهتر بود [۲۲]. ندیری و همکاران [۵] در پژوهشی با کمک مدل‌های منطق فازی، شبکه‌های عصبی مصنوعی و نروفازی مقادیر قابلیت انتقال سفره آب زیرزمینی دشت تسوج را برآورد کردند. ایشان در ادامه پژوهش با طراحی مدلی ترکیبی از مدل‌های هوش مصنوعی، مقادیر خروجی از هر یک از مدل‌های فوق را با در نظر گرفتن وزن بهینه شده با روش الگوریتم بهینه‌سازی شده با جامعه ذرات وارد مدل ترکیبی کردند. نتایج نشان داد که مقادیر قابلیت انتقال برآورد شده با مدل ترکیبی دقیق تر

¹ Learning Machine² Data Mining

مقادیر پیش‌بینی شده تا ۷ روز آینده بود ولی با طولانی‌تر شدن زمان پیش‌بینی دقت مقادیر کاهش یافت. همچنین در سال‌های اخیر ابزارهای یادگیری^۱ و روش‌های جدید، با بهره‌گیری از رگرسیون‌های گروهی ارائه شده و در حال توسعه می‌باشند. جنگل تصادفی^۲ (RF) یکی از انواع روش‌های یادگیری است که از الگوریتم‌های پایه برای پیش‌بینی‌های چندگانه تکراری^۳ استفاده می‌کند. جنگل تصادفی هزاران متغیر ورودی را پذیرا بوده، همچنین می‌تواند اهمیت نسبی متغیرها را ارزیابی و مهم‌ترین متغیر را در پیش‌بینی مدل تشخیص دهد [۲۹]. این روش پتانسیل لازم را به‌عنوان یک ابزار مدل مکانی برای ارزیابی مسائل زیست‌محیطی و منابع آب دارا می‌باشد [۲۹]. در سال ۲۰۱۲، رودریگز و همکاران [۲۹] عملکرد رگرسیون منطقی، درخت کلاس‌بندی رگرسیونی و RF را بررسی نمودند که نتایج نشان‌دهنده برتری مدل RF نسبت به سایر مدل‌ها می‌باشد. این مدل برای تعیین مناطق آسیب‌پذیر آبخوان دشت ملکان نیز به‌کار برده شد و نتایج نشان داد که مدل موردنظر می‌تواند با شناسایی مناطق مستعد به آلودگی، برای مدیریت و نظارت صحیح آب‌های زیرزمینی مورد استفاده قرار گیرد [۲۵]. در مجموع با توجه به پژوهش‌های انجام‌شده و همچنین اهمیت دشت بوکان آذربایجان غربی از لحاظ شرب و کشاورزی، بررسی تغییرات تراز سطح ایستابی به‌منظور پیش‌بینی و اقدامات مدیریتی بیش‌ازپیش ضروری است. هدف کلی از این تحقیق بررسی کارایی روش‌های هوش مصنوعی و ماشین‌های یادگیری، به‌خصوص روش جنگل تصادفی که برای نخستین بار در زمینه پیش‌بینی سطح آب زیرزمینی مورد استفاده قرار گرفته است، می‌باشد.

از تک‌تک مدل‌های اولیه بود. نتایج مقایسه عملکرد سه روش هوش مصنوعی (شبکه‌های عصبی مصنوعی، سامانه استنتاج تطبیقی فازی-عصبی و برنامه ریزی بیان ژن) برای پیش‌بینی نوسانات تراز آب زیرزمینی در حوضه زرينگل در استان گلستان نشان داد هر سه روش به‌طور موفقیت‌آمیزی برای پیش‌بینی نوسانات سطح ایستابی مورد استفاده قرار گیرند اما بهترین عملکرد مربوط به مدل برنامه ریزی بیان ژن بود [۱]. ندیری و همکاران [۲۳] طی پژوهشی از ۳ مدل هوش مصنوعی شامل منطق فازی-تاکاگی-سوگنو، شبکه‌های عصبی مصنوعی و نرو-فازی، برای تخمین هدایت هیدرولیکی آبخوان تسوج-آذربایجان شرقی استفاده کردند. بر اساس نتایج، با وجود مقادیر هدایت هیدرولیکی تخمینی کاملاً متفاوت از طریق دو مدل شبکه عصبی و منطق فازی، ولی این دو مدل اهمیت یکسانی نشان دادند. به‌منظور تخمین مقادیر هدایت هیدرولیکی آبخوان تسوج-آذربایجان شرقی، تیفور و همکاران [۳۲] از مدل‌های هوش مصنوعی شامل منطق فازی سوگنو، منطق فازی مامدانی، شبکه‌های عصبی پرسپترون چندلایه ترکیبی با تابع لونیبرگ-مارکوارت و نرو-فازی، استفاده کردند. بر اساس نتایج، کارایی مدل‌های نرو-فازی و منطق فازی سوگنو قابل قبول بود در حالی که تخمین‌های شبکه‌های عصبی و منطق فازی مامدانی، ضعیف بود. ژورویچ و همکاران [۱۴] طی پژوهشی جهت پیش‌بینی سطح آب زیرزمینی در ساحل چپ رودخانه دانوب در استان وجودینا در شمال صربستان از دو مدل ANFIS و ANN استفاده کردند. نتایج نشان داد که از هر دو مدل می‌توان به‌عنوان ابزاری مفید برای مدل‌سازی سطح آب زیرزمینی با دقت بالا استفاده کرد. سان و همکاران [۳۱] به‌منظور پیش‌بینی سطح آب زیرزمینی جنگل‌های باتلاقی سنگاپور با کمک شبکه عصبی از معماری پرسپترون چندلایه با الگوریتم لونیبرگ-مارکوارت و فقط مقادیر سطح آب برکه‌ها و میزان بارندگی استفاده کردند. نتایج حاکی از صحت

¹ Machine learning

² Random Forest

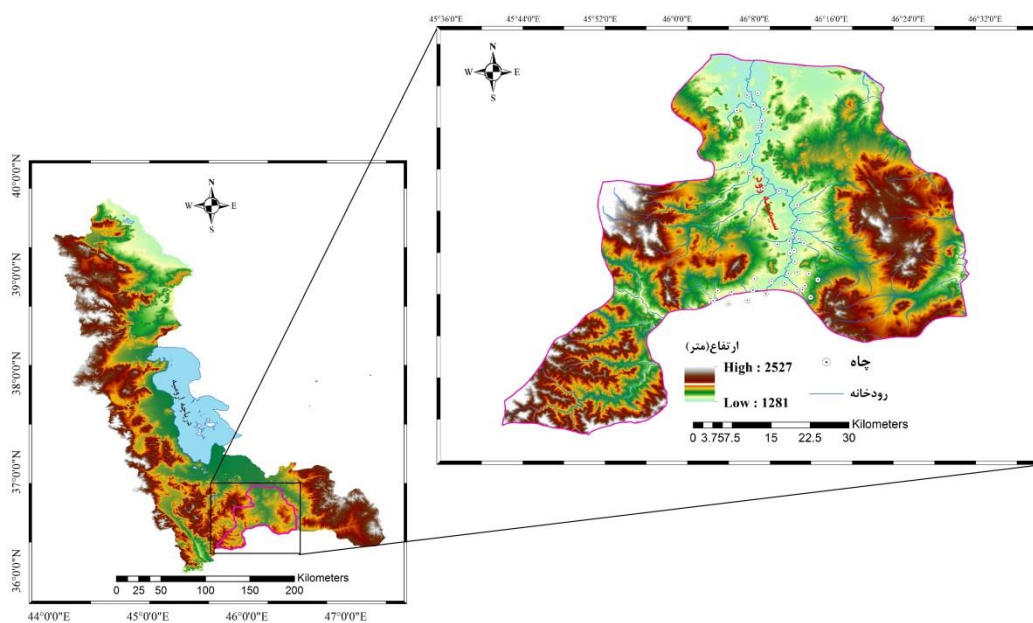
³ Repeated Multiple Prediction

۲. روش شناسی

۱،۲. معرفی منطقه مورد مطالعه

۱۳۸۹ میلی‌متر در سال می‌باشد. سیمینه‌رود مهم‌ترین زهکش سطحی دشت بوکان می‌باشند که منابع آب سطحی این رودخانه‌ها پس از عبور از دشت بوکان و میاندوآب به دریاچه ارومیه می‌ریزند. سرشاخه‌های رودخانه سیمینه‌رود شامل رودخانه گولان، چم آغوتمان، چم فالوچه و رودخانه کوله‌پز در این محدوده مطالعاتی شده و جریان سطحی از محدوده‌های مطالعاتی مجاور وارد این محدوده مطالعاتی نمی‌گردد. جریان سطحی خروجی از این محدوده به محدوده مطالعاتی میاندوآب تخلیه می‌گردد. سازندهای زمین‌شناسی حاشیه دشت عمدتاً مربوط به تشکیلات کرتاسه و میوسن می‌باشند [۴]. در قسمت غربی دشت سنگ‌های آذرین و دگرگونی و به مقدار کمتر سنگ‌های رسوبی مشاهده می‌گردد. قسمت شرقی به‌طور عمده از سنگ‌های رسوبی و عمدتاً آبرفتی است. وسعت رسوبات ناپیوسته در سطح محدوده مطالعاتی مذکور ۴۱۳/۵ کیلومتر مربع و وسعت سازندهای سخت در این محدوده مطالعاتی ۲۴۹۰/۴ کیلومتر مربع است.

محدوده مطالعاتی بوکان، در جنوب استان آذربایجان غربی و جنوب دریاچه ارومیه واقع شده است. مساحت محدوده مطالعاتی بوکان ۲۹۰۰ کیلومتر مربع می‌باشد که از این مقدار سهم ارتفاعات ۲۶۳۸ کیلومتر مربع و دشت بوکان تنها ۲۶۲ کیلومتر مربع از مساحت کل محدوده را به خود اختصاص داده است. از نظر تقسیمات طبیعی این منطقه حدود ۵/۵۴ درصد کل حوضه آبریز دریاچه ارومیه را در بر می‌گیرد (شکل ۱). بر اساس اطلاعات حاصل از ایستگاه سینوپتیک بوکان و گزارش هواشناسی منطقه مربوط به سال ۱۳۹۶، این حوزه بر اساس روش تجربی آمبرژه [۱۶] دارای اقلیم سرد و نیمه‌خشک می‌باشد. میزان بارندگی سالیانه منطقه بر اساس آمار سی ساله (۱۳۶۶-۹۵) ایستگاه بوکان در دشت و بخش کوهستانی به ترتیب ۳۱۱ و ۴۲۰ میلی‌متر و متوسط بارندگی ۳۲۶ میلی‌متر در سال می‌باشد. همچنین میزان تبخیر و تعرق پتانسیل و تبخیر از تشتک به ترتیب برابر با ۶۵۰ و



شکل ۱- موقعیت منطقه مورد مطالعه

۲.۲. مدل‌های هوش مصنوعی

هوش مصنوعی برای داده‌هایی که دارای رفتار پیچیده و غیرخطی و یا دارای ویژگی عدم قطعیت هستند، دارای اهمیت بوده و کاربردی می‌باشد. در بین روش‌های هوش مصنوعی، می‌توان روش شبکه عصبی مصنوعی و منطق فازی را نام برد که برای چنین داده‌هایی کاربرد دارند. در این مطالعه از مدل‌های هوش مصنوعی مانند فازی^۱ و شبکه‌های عصبی مصنوعی (ANN^۲) برای پیش‌بینی سطح آب زیرزمینی استفاده شده است. شبکه‌های عصبی مصنوعی یک سیستم پردازشگر اطلاعات توده‌ای هستند که به صورت موازی قرار گرفته‌اند و عملکردی شبیه شبکه عصبی مغز انسان دارند [۲۰]. در این تحقیق، از شبکه پرسپترون چند لایه^۳ (MLP) که یکی از شبکه‌های پیشرو می‌باشد، استفاده شده است. در این شبکه‌ها هر نرون دارای یک تابع محرک هست و هر ورودی $P(p_1, p_2, \dots, p_n)$ توسط یک بردار وزن $W(w_1, w_2, \dots, w_n)$ به نرون مورد نظر مرتبط می‌گردد (شکل ۲). خروجی هر نرون که a نامیده می‌شود طبق رابطه (۱) محاسبه می‌شود:

$$a = f(P \cdot w + b) \quad (1)$$

در این رابطه؛ f تابع فعال‌سازی^۴ و b مقدار بایاس^۵ می‌باشند.

بنابراین خروجی یک نرون در یک لایه به سیگنالی که از لایه قبلی دریافت می‌کند، وزن تعیین شده برای آن و نوع تابع تبدیل‌گر بستگی دارد. نرون‌ها می‌توانند از توابع فعال‌سازی متفاوتی جهت تولید خروجی استفاده کنند. از رایج‌ترین آن‌ها می‌توان به توابع لگاریتم سیگموئیدی، تانژانت سیگموئیدی و تابع فعال‌سازی

خطی اشاره کرد [۱۳]. روش‌های آموزش در شبکه‌های پرسپترون چند لایه جهت آموزش شبکه‌های عصبی توسط الگوریتم‌های ریاضی متفاوت صورت می‌گیرد که از جمله مهم‌ترین آن‌ها الگوریتم انتشار خطا به عقب^۶ (BP)، الگوریتم گرادینان مزدوج^۷ (CG)، الگوریتم لونیبرگ-مارکوارت^۸ (LM) و غیره می‌باشند. از این میان الگوریتم لونیبرگ-مارکوارت به‌عنوان کارا ترین و سریع‌ترین روش آموزش شبکه‌های عصبی شناخته شده است که جزئیات آموزش و فرایند محاسبه الگوریتم لونیبرگ-مارکوارت توسط انجمن مهندسیین عمران آمریکا [۶] ارائه شده است. در این تحقیق برای اجرای مدل شبکه‌های عصبی مصنوعی، داده‌های ورودی و هدف به دو دسته آموزش و آزمایش تقسیم گردید. به طوری که ۸۰ درصد داده‌ها برای مرحله آموزش و ۲۰ درصد داده‌ها برای مرحله آزمایش در نظر گرفته شده است. روش فازی روشی مناسب جهت کاهش خطای برآوردی و انسانی نسبت به تئوری‌های قابل اعتماد دیگر است [۱۰ و ۲۴]. در منطق فازی دو بحث مدل‌سازی و دسته‌بندی فازی مطرح است که از مدل‌سازی فازی جهت تخمین پارامترهای عددی مانند تخلخل، هدایت هیدرولیکی، تراوایی و قابلیت انتقال استفاده می‌شود. مدل‌سازی فازی به سه روش فازی ممدانی (MFL)، تاگائی - سوگونو - کانگ که به‌طور خلاصه روش سوگونو (SFL) نامیده می‌شود و لارسن (LFL) قابل اجراست. تفاوت روش سوگونو با دو روش دیگر در خروجی آن‌ها است، در روش سوگونو تابع عضویت داده‌های خروجی سیستم فازی به صورت رابطه خطی یا ثابت است که توسط روش دسته‌بندی به دست می‌آید. اولین مرحله در ایجاد یک مدل فازی دسته‌بندی داده‌هاست که بسته به نوع مدل فازی استفاده شده می‌توان از روش‌های دسته‌بندی مختلف استفاده کرد. از

¹ Fuzzy Logic

² Artificial Neural Network

³ Multilayer perceptron

⁴ Activation Function

⁵ Bias

⁶ Back Propagation Algorithm

⁷ Conjugate Gradient Algorithm

⁸ Levenberg-Marquard

ضعیف، قابل قبول هستند [۳۰]. در این تحقیق برای فائق آمدن بر مشکلات یادگیرنده‌های پایه، الگوریتم جنگل تصادفی که یک روش یادگیری مبتنی بر دسته‌ای از درخت‌های تصمیم است، پیشنهاد شده است.

۱،۳،۲. رگرسیون و الگوریتم جنگل تصادفی

در الگوریتم RF برای تشکیل هر درخت، دسته‌های متفاوتی از الگوهای موجود با در نظر گرفتن جایگزینی دوباره هر الگوی منتخب، انتخاب می‌شوند. اندازه این دسته نمونه برداری شده، برابر تعداد کل الگوهای موجود خواهد بود. RF در سال ۲۰۰۱ توسط بریمن به‌عنوان یک روشی که از توسعه جدید درخت‌های تصمیم‌گیری ایجاد شده، ارائه گردید که پیش‌بینی چندین الگوریتم منفرد را با هم با استفاده از قوانین مبتنی ترکیب می‌کند. اصول کلی تکنیک‌های آموزش گروهی بر پایه این فرض است که دقت آن‌ها از دیگر الگوریتم‌های آموزشی بالاتر است چون ترکیبی از چند مدل پیش‌بینی، دقیق‌تر از یک مدل می‌باشد و گروه‌ها قدرت مجموعه‌های منفرد و منحصر به فرد از طبقه‌ها را بیشتر می‌کنند در حالی که نقاط ضعف طبقه‌ها را در همان زمان کاهش می‌دهند [۲۱]. درخت رگرسیونی^۴ (RT) مجموعه‌ای از شرایط یا محدودیت‌ها را بیان می‌کند که به‌صورت سلسله مراتبی سازمان یافته‌اند و به حالت متوالی از گره ریشه به سمت پایین رشد می‌کنند و به گره‌های پایانی یا گره‌های برگ می‌رسند [۲۷ و ۸]. به‌منظور به وجود آوردن درخت رگرسیونی از پارتیشن‌بندی بازگشتی و رگرسیون‌های چندگانه استفاده می‌شود. فرایند تصمیم از گره ریشه، در هر گره داخلی، طبق قانون درختی تا زمانی که شرط توقف قبلی تعیین شده به دست آید، تکرار می‌شود. هریک از گره‌های نهایی یا برگ‌ها به یک مدل رگرسیونی ساده که فقط در گره بکار برده می‌شوند، متصل می‌شود. زمانی که فرایند فراخوانی

روش دسته‌بندی کاهشی^۱ برای مدل ساگنو و از روش دسته‌بندی میان مرکز^۲ (FCM) برای روش ممدانی و لارسن استفاده می‌شود. هر مدل فازی شامل سه مرحله اصلی است: الف) فازی کردن داده‌ها: که این عمل با تعریف تابع عضویت انجام می‌شود، ب) ایجاد ارتباط بین ورودی و خروجی: که این کار نیز با یک سری قوانینی مانند اگر آنگاه (if-then) انجام می‌شود و ج) مرحله آخر که مرحله بررسی سیستم، تجمیع نتایج و غیرفازی سازی است. در این مطالعه از روش فازی سوگنو استفاده شده است.

۳،۲. مدل جنگل تصادفی

یکی از ابزارهای کارآمد در مسائل مربوط به تخمین متغیرهای هدف و یا طبقه‌بندی الگوها، درخت تصمیم است. جنگل تصادفی (RF) یک روش گروهی است که چند الگوریتم درختی را برای تولید یک پیش‌بینی مکرر از هر پدیده ترکیب می‌کند. به‌طور کلی، درخت تصمیم منفرد مستعد برآزش اضافی^۳ بوده و قدرت تعمیم‌پذیری کمی دارد. در هنگام تشکیل یک درخت تصمیم، تغییرات کم در الگوهای یادگیری می‌تواند باعث تغییرات اساسی در ساختار آن درخت گردد [۲۷]. RF می‌تواند الگوهای پیچیده را یاد بگیرد و ارتباط غیرخطی بین متغیرهای توضیحی و متغیرهای وابسته را در نظر بگیرد. همچنین می‌تواند انواع مختلف داده‌ها را به علت عدم وجود توزیع نرمال داده‌های استفاده شده، در تجزیه و تحلیل گنجانده و ترکیب کند. ترکیب درخت‌های تصمیم را روش‌های دسته جمعی می‌گویند. یک مجموعه دسته جمعی شامل تعدادی یادگیرنده است که به آن‌ها یادگیرنده‌های پایه گفته می‌شود. یادگیرنده‌های پایه معمولاً از داده‌های آموزشی به وسیله یک الگوریتم یادگیری پایه ساخته می‌شوند که این الگوریتم یادگیری پایه می‌تواند درخت تصمیم، شبکه عصبی یا الگوریتم‌های یادگیری دیگر باشد. قابلیت تعمیم یک مجموعه، اغلب قوی‌تر از یادگیرنده‌های پایه است و در واقع روش‌های مجموعه‌ای بیشتر به دلیل داشتن توانایی در تقویت یادگیرنده‌های

¹ Subtractive

² Fuzzy C-Means

³ Over fitting

⁴ Regression Tree

می‌شوند که الگوهای خارج از کیسه^۳ (OOB) نامیده می‌شوند. این قسمت در روش RF می‌تواند برای ارزیابی عملکرد مدل مورد استفاده قرار گیرد [۸]. به این ترتیب RF می‌تواند تخمین غیر مرتبط داخلی از خطای تعمیم را محاسبه کند بدون اینکه از زیرمجموعه‌های داده‌های خارجی استفاده کند [۹]. روند کلی الگوریتم RF به صورت ساده در شکل (۲) نشان داده شده است. اگر بخواهیم روش RF برای طبقه‌بندی را به طور خلاصه بیان کنیم، بدین صورت است که: در ابتدا T نمونه خود راه‌انداز^۴ از داده آموزشی بیرون کشیده می‌شود و سپس از هر نمونه خودراه‌انداز^۵ یک درخت طبقه‌بندی و رگرسیون^۵ (CART) هرس نشده ایجاد می‌شود که برای انشعاب در هر گره CART، تنها یک متغیر انتخاب شده به صورت تصادفی استفاده می‌شود. در نهایت خروجی طبقه‌بندی بر اساس یک نتیجه میانگین، از پیش‌بینی‌های تمام درخت‌های منفرد آموزش دیده به دست می‌آید. یک مجموعه داده خود راه‌انداز، مجموعه‌ای از نقاط انتخابی به طور تصادفی است که از مجموعه داده آموزشی بیرون کشیده شده است. برای اینکه همیشه اندازه نمونه آموزشی اولیه ثابت بماند مجموعه داده خود راه‌انداز یک نسخه کپی از نقاط را به کار می‌برد. لازم به ذکر است که الگوریتم RF به دلیل ارزیابی درونی نتایج هر طبقه‌بندی درختی که در داخل خود دارد و با وزن‌دهی به نتایج هر درخت، می‌تواند نتایج صحیحی را تولید کند.

گرچه اطلاعات زیاد برای مدل‌سازی ممکن است مفید باشند اما افزایش تعداد پارامترهای ورودی پیچیدگی‌های اضافی و افزایش زمان محاسبات و مشکلات ابعادی را به سیستم تحمیل می‌کند [۷].

درخت به پایان برسد هرس کردن یا پرونینگ^۱ با هدف بهبود ظرفیت تعمیم درخت‌ها به وسیله کاهش پیچیدگی ساختار بکار برده می‌شود و تعداد موردها یا نمونه‌ها در گره‌ها می‌تواند به عنوان معیار پرونینگ در نظر گرفته شود. برای جلوگیری از تطابق RTsهای مختلف، RF تنوع درختان را از طریق درست کردن زیرمجموعه‌های مختلف از داده‌های آموزشی کم می‌کند که اصطلاحاً کیسه‌بندی^۲ نامیده می‌شود. اصطلاح Bagging از مخفف Bootstrap Aggregating به دست آمده است [۹]. Bagging و تکنیکی است که از طریق نمونه‌برداری دوباره تصادفی از مجموعه داده‌های اصلی و همراه با جایگزینی، برای ایجاد داده‌های آموزشی استفاده می‌شود و در این مرحله هیچ کدام از داده‌های انتخاب شده از نمونه‌های ورودی را برای تولید زیرمجموعه بعدی، حذف نمی‌کند و بدین ترتیب واریانس نیز کاهش می‌یابد. از این رو برخی از داده‌ها ممکن است بیش از یک بار در شاخه‌های آموزشی استفاده شود در حالی که برخی از داده‌های دیگر که در مدل‌سازی مؤثر نیستند ممکن است هرگز استفاده نشوند؛ بنابراین ثبات بیشتری برای مدل به دست می‌آید و مدل را در برابر تغییرات جزئی در داده‌های ورودی قابل اعتمادتر می‌سازد و دقت پیش‌بینی آن را افزایش می‌دهد [۱۰]. از سوی دیگر هنگامی که RF، یک درخت رشد ایجاد می‌کند از بهترین متغیرها یا نقاط تقسیم در داخل زیرمجموعه‌های متغیرها استفاده می‌کند که به صورت تصادفی از بین مجموعه‌های کلی متغیرهای ورودی انتخاب می‌کند. در نتیجه قدرت هر RT منفرد را کاهش داده و میزان تطابق را پایین می‌آورد و بدین صورت خطای کلی مدل را کاهش می‌دهد [۹]. یکی دیگر از ویژگی‌های خوب RF این است که درختان RF، بدون پرونینگ یا هرس کردن رشد می‌کنند و در این روش، آموزش بیش از اندازه بر دقت مدل تأثیری نمی‌گذارد. علاوه بر این آن دسته از نمونه‌هایی که در آموزش درختان در فرایند Bagging انتخاب نمی‌شوند شامل بخشی از زیر مجموعه‌هایی

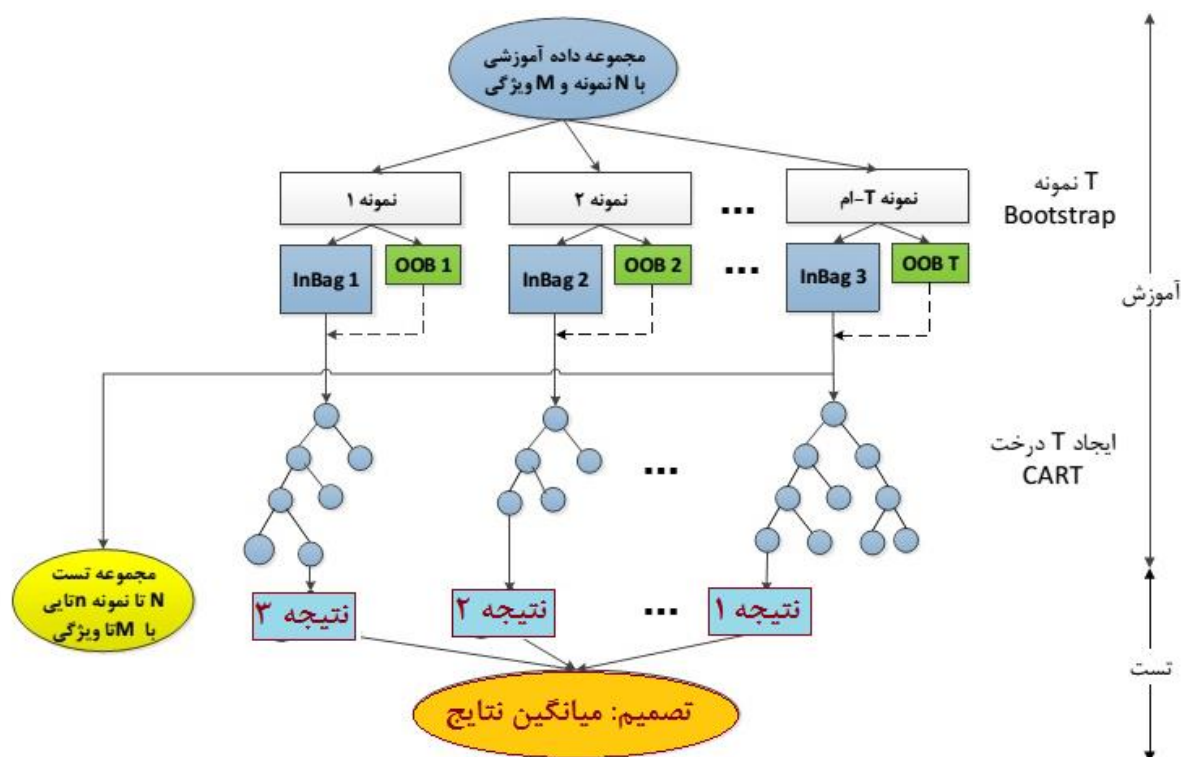
¹ Pruning

² Bagging

³ Out of Bag

⁴ Bootstrap

⁵ Classification and Regression Tree



شکل ۲. فلوچارت مدل جنگل تصادفی برای رگرسیون

باشند [۱۴]. FS با بالا بردن سرعت فرایند آموزش، افزایش قابلیت تعمیم، کاهش اثر از بین رفتن ابعاد و افزایش قابلیت تفسیر، دقت مدل‌های پیش‌بینی را افزایش می‌دهد. روش‌های زیادی برای FS بیان شده است ولی معمول‌ترین روش آن، روش آماری چند متغیره است که به خاطر کاهش ابعاد داده‌ها، مؤلفه‌های اصلی جایگزین پارامترهای اولیه می‌شوند [۱۱].

۳. نتایج

۳.۱. شبکه‌های عصبی مصنوعی

مرحله اول مدل‌سازی توسط شبکه‌های عصبی مصنوعی مربوط به شناخت بهترین ساختار شبکه‌های عصبی مصنوعی و طراحی شبکه جهت تخمین سطح ایستابی یا تراز آب زیرزمینی در منطقه مطالعاتی می‌باشد. به این منظور ساختارهای شبکه‌های عصبی

برای رسیدن به یک تقسیم بهینه، تصمیم‌ساز^۱ می‌تواند به وسیله مقدار بسیار عظیمی از داده‌های مختلف، اطلاعات و شرایطی که در منطقه موجود می‌باشند به سرعت تحت‌الشعاع قرار بگیرد. تعداد زیادی از متغیرهای وابسته به خواص و رفتار سیستم آب زیرزمینی و نیروهای راندگی، می‌توانند توانایی مدل را افزایش دهد. ابعاد بالا در مجموعه داده‌ها می‌تواند باعث کاهش دقت مدل شود و بنابراین برای جلوگیری از این خطاها و ابعاد بالای داده‌ها، انتخاب ویژگی^۲ (FS) بکار برده می‌شود. FS یک روشی برای انتخاب زیرمجموعه‌های پارامترهای مربوطه برای آموزش بهتر مدل می‌باشد [۱۸]. در مطالعات آب زیرزمینی، از تعداد زیادی متغیرهای وابسته به خصوصیات فیزیکی و شیمیایی آبخوان استفاده می‌شود که برخی از آنها می‌توانند مرتبط و یا غیرمرتبط

^۱ Decision-Maker

^۲ Feature Selection

لایه سوم مورد استفاده قرار گرفت [۲۳]. توابع انتقال با آزمون و خطا به گونه‌ای انتخاب شد که میانگین مربعات خطا (MSE) در دو مجموعه آموزش و آزمایش کمینه گردید، مقادیر ضریب همبستگی و میانگین مربعات خطا برای مرحله آموزش و آزمایش در جدول (۱) ارائه شده است. در نهایت به منظور پیش‌بینی سطح آب زیرزمینی به کمک شبکه‌های عصبی مصنوعی، شبکه چندلایه پرسپترون با الگوریتم آموزشی لونبرگ-مارکوارت به عنوان کاراترین شبکه شناخته شد. به طوری که با کمترین MSE و بیشترین مقدار R^2 به عنوان بهترین مدل شبکه عصبی مصنوعی شناخته شده در صورتی که سایر ساختارهای شبکه عصبی مصنوعی به طور میانگین از MSE و R^2 بالاتری برخوردار بودند.

مصنوعی که متشکل از شبکه‌های (FFN) و (RNN) است با الگوریتم‌های مختلف از جمله الگوریتم‌های انتشار خطا به عقب (BP)، الگوریتم گرادیان مزدوج (CG) و الگوریتم لونبرگ-مارکوارت (LM) مورد بررسی و مطالعه قرار گرفتند. پارامترهای سطح ایستابی در دوره زمانی ماه قبل، دبی جریان، دما و بارش در دوره زمانی t به عنوان لایه‌های ورودی شبکه استفاده شدند. با توجه به اینکه تعداد نرون‌ها در لایه مخفی در راستای کمینه کردن خطا و بیشینه کردن دقت مدل، بهینه‌سازی می‌شوند بنابراین با استفاده از روش آزمون و خطا، تعداد نرون‌ها در لایه میانی سه بدست آمد. تابع تانژانت هیپربولیک که به عنوان بهترین تابع برای مدل‌سازی سیستم‌های طبیعی غیرخطی تشخیص داده شده‌اند، به عنوان تابع تبدیلگر نرون‌های پردازشگر در لایه دوم و تابع خطی (Purlin) در

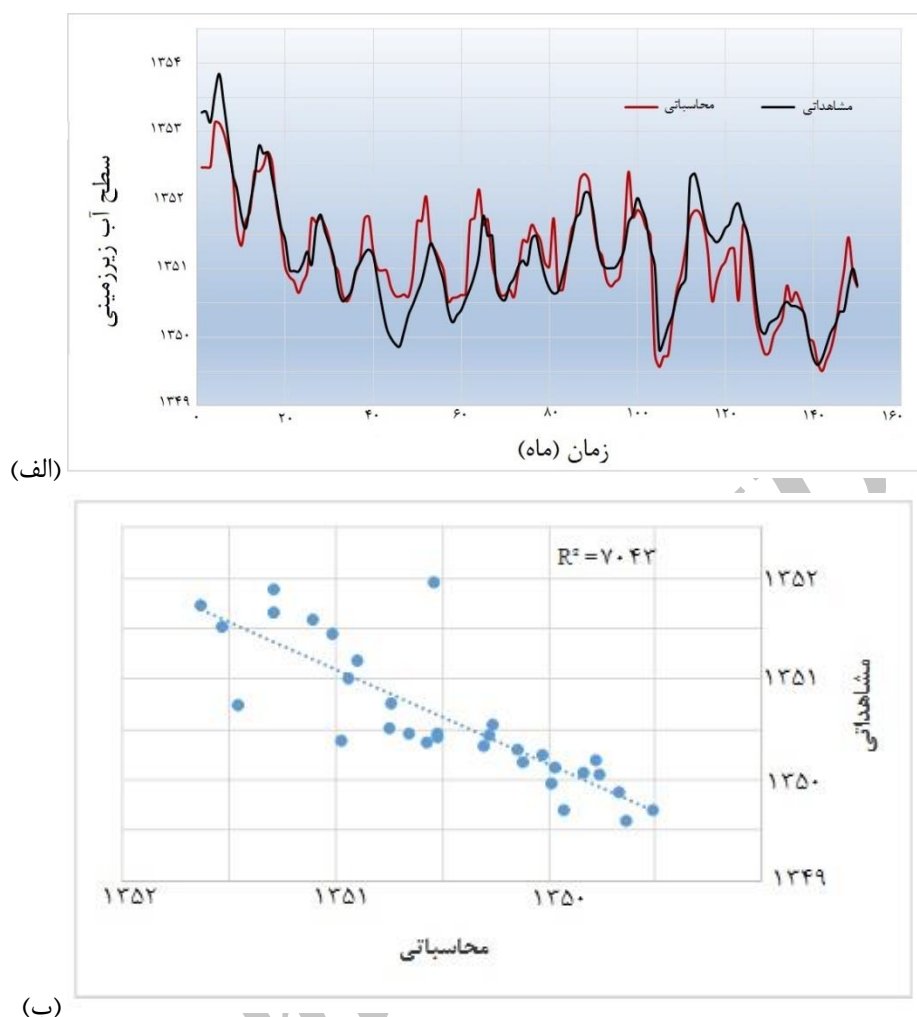
جدول ۱. نتایج ارزیابی عملکرد شبکه عصبی مصنوعی با ساختارهای مختلف در مراحل آموزش و آزمایش

RNN		FFN		نوع شبکه		
BP	GC	LM	BP	GC	LM	
MSE	R^2	MSE	R^2	MSE	R^2	معیار ارزیابی
۰/۵۱۶	۰/۵۸۲	۰/۷۱۲	۰/۶۶۱	۰/۲۱۴	۰/۷۳۳	مرحله آموزش
۰/۶۷۱	۰/۵۹۳	۰/۷۹۴	۰/۵۹۹	۰/۱۹۹	۰/۷۱۴	مرحله آزمایش

۲.۳. مدل فازی

برای دسته‌بندی داده‌ها در روش فازی از روش کاهشی که روشی کارا و مفید در دسته‌بندی داده‌ها، تعیین توابع عضویت و قوانین در تحقیقات اخیر بوده است استفاده گردید. در این سیستم استنتاج فازی از یکسری قوانین اگر-آنگاه فازی جهت برقراری ارتباط بین توابع عضویت داده‌های ورودی و توابع عضویت داده‌های خروجی استفاده شده است. پارامتر اصلی در این روش دسته‌بندی برای تعیین دسته‌ها و قوانین اگر آنگاه شعاع دسته‌بندی است.

در شکل ۳ نمودار مقادیر مشاهداتی و محاسباتی و نمودار پراکنش نشان داده شده است که داده‌های مشاهداتی مربوط به دشت بوکان است. همان‌طور که در شکل (۴.الف) مشخص است تطابق مقادیر محاسباتی میزان سطح ایستابی مدل شبکه عصبی مصنوعی با مقادیر مشاهداتی در برخی نقاط وجود دارد. به گونه‌ای که در شکل (۴.ب) این امر مشهود است که مدل مذکور در تخمین نقاط بیشینه و کمینه ضعیف عمل کرده است که این نقاط را نزدیک به مقداری واقعی‌شان تخمین نزنده است.



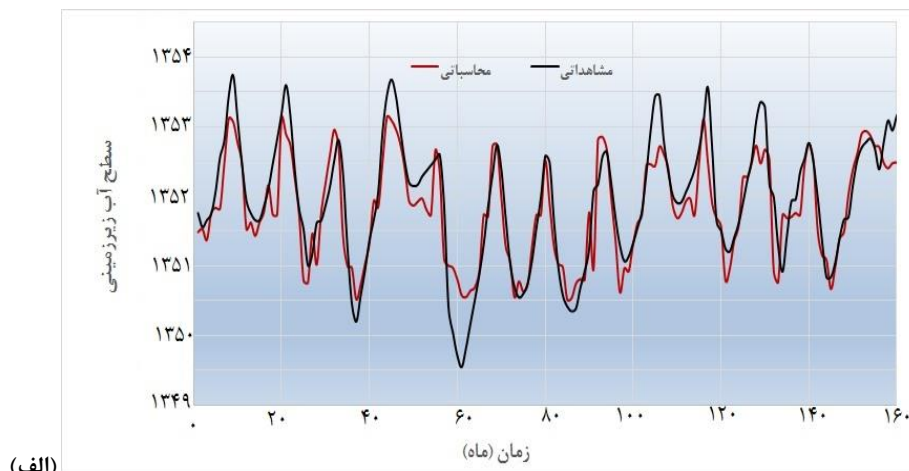
شکل ۳. الف) مقادیر محاسباتی و مشاهداتی نسبت به زمان ب) نمودار پراکنش میان مقادیر مشاهداتی و محاسباتی

بود. بازده مدل فازی بستگی به تعیین شعاع بهینه دسته‌بندی دارد که در برآورد سطح ایستابی مقدار بهینه آن به روش‌آزمون و خطا $0/7$ تعیین شد. توابع عضویت استفاده شده برای مدل‌سازی فازی سطح ایستابی، تابع عضویت گوسی بود که به داده‌های دسته‌بندی شده برازش داده شد و نیز تابع عضویت خروجی مدل سوگنو تابعی است خطی که بر اساس ورودی‌ها ساخته می‌شود. جدول (۲) نتایج ارزیابی عملکرد مدل‌سازی فازی در مراحل آموزش و آزمایش را نشان می‌دهد. بر اساس نتایج مدل منطق فازی ساگنو در بین روش‌های هوش مصنوعی دارای بیشترین بازده، کمترین درصد خطا و توانایی بالایی در پیش‌بینی سطح ایستابی این دشت می‌باشند.

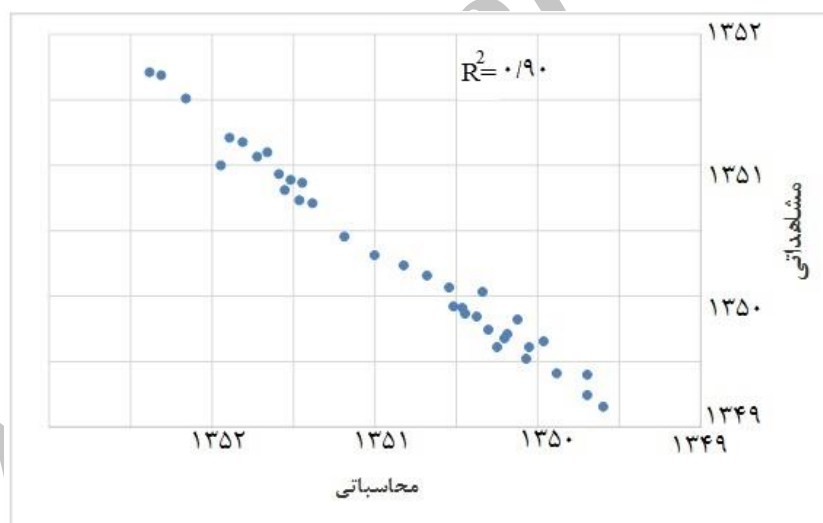
توابع عضویت با استفاده از دسته‌بندی کاهشی توسط دستور $\text{genfis2}(\text{datin}, \text{datout}, r)$ تعیین گردید که در آن، datin داده‌های ورودی، datout داده‌های خروجی و r شعاع دسته‌بندی است که بین صفر تا یک متغیر است. تعریف مقدار بزرگ شعاع دسته‌بندی منجر به تولید تعداد کمی دسته می‌شود و تعریف مقادیر کوچک، تعداد زیادی دسته کوچک تولید خواهد کرد. هر یک از دسته‌ها، نماینده یک تابع عضویت می‌باشد و هر تابع عضویت نیز نشانگر یک مجموعه فازی در داده‌ها می‌باشد که برای استخراج قوانین به کار می‌رود. کاهش این پارامتر باعث افزایش دسته‌ها و قوانین می‌شود و افزایش آن به سمت یک باعث کاهش دسته‌ها و قوانین خواهد

جدول ۲. نتایج ارزیابی عملکرد مدل منطق فازی در مراحل آموزش و آزمایش

معیار ارزیابی	R^2	MSE
مرحله آموزش	۰/۹۶	۰/۰۶۸
مرحله آزمایش	۰/۹۰۶	۰/۰۵۷



(الف)



(ب)

شکل ۴. الف) مقادیر محاسباتی و مشاهداتی نسبت به زمان ب نمودار پراکنش میان مقادیر مشاهداتی و محاسباتی برای داده‌ها

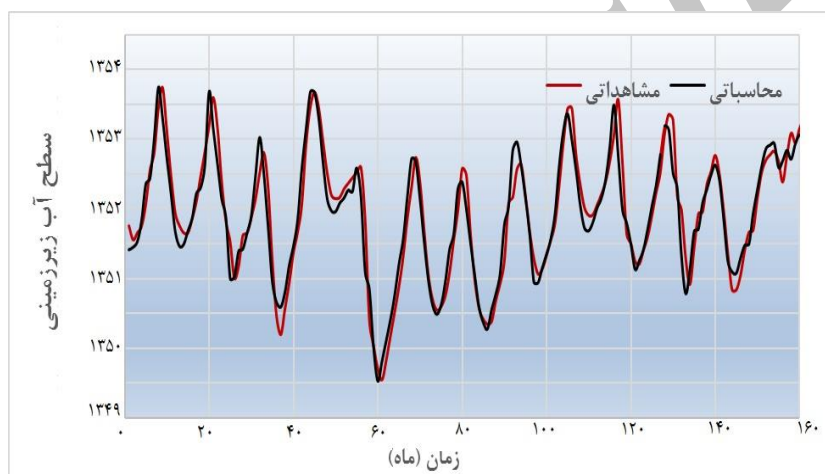
تعداد درخت‌ها (k) به طوری که مقدار خطا همگرا شود و تخمین قابل اعتمادتر باشد، مدلی از ۱ تا ۵۰۰ درخت ساخته شد پارامتر m نیز به وسیله تغییر تعداد متغیرها بین یک و ماکزیمم متغیرهای هر زیرمجموعه، بهینه شد. علاوه بر این، برای کاهش ابعاد و افزایش دقت و قابل

۳.۳. مدل جنگل تصادفی

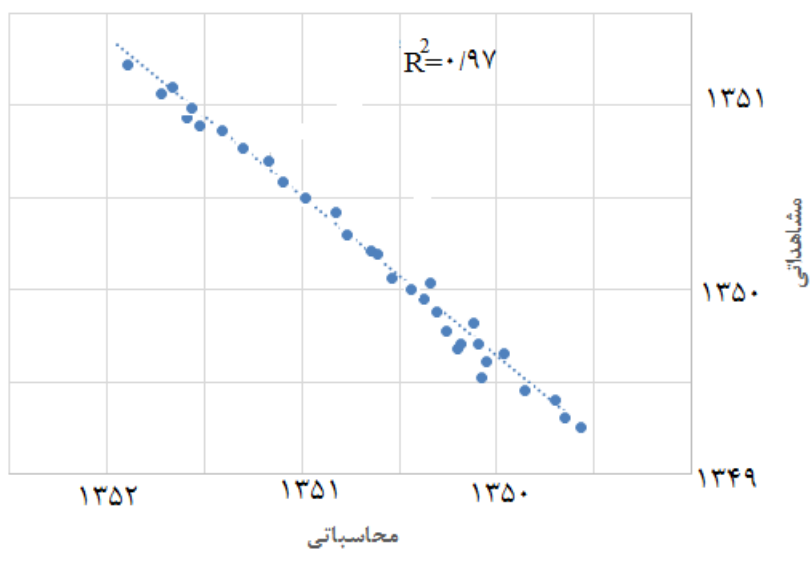
در مطالعه حاضر روش RF، که بر اساس درخت‌های گروهی است ترجیح داده شده است. اولین مرحله در ایجاد مدل پیش‌بینی قابلیت انتقال به وسیله مدل جنگل تصادفی، تنظیم پارامترهای مدل است. به منظور تنظیم

شکل (۵. الف) نتایج پیش‌بینی مدل RF (داده‌های محاسباتی) را در مقابل داده‌های مشاهداتی برای پارامتر سطح ایستابی را نشان می‌دهد. از آنجا که طبقه‌بندی در RF از میانگین‌گیری و شمارش نتایج طبقه‌بندی کننده‌های پایه ساده و متنوع حاصل می‌شود، بنابراین RF قادر است با داده‌های دارای نویز نیز به گونه‌ای صحیح‌تر، در مقایسه با دیگر الگوریتم‌ها از جمله SVM و شبکه‌های عصبی مصنوعی که در آن‌ها احتمال برازش اضافی در هنگام آموزش با داده‌های دارای نویز بالا وجود دارد، آموزش ببیند.

تفسیر بودن مدل، روش FS بکار برده شد. برای فراخوانی مدل پیش‌بینی سطح ایستابی، داده‌های ورودی همراه با متغیر هدف (سطح ایستابی) به‌عنوان ورودی مدل استفاده شدند. متغیرهای ورودی یا پیش‌بینی کننده و متغیر پاسخ یا سطح ایستابی در مجموعه‌ای از بردارهای ویژگی ورودی ترکیب شدند. این بردارها که ورودی الگوریتم RF را تشکیل می‌دهند، به‌عنوان بردارهای Input-feature شناخته شده است. متغیر پاسخ دوتایی (متغیر ورودی و متغیر پاسخ) به‌عنوان مقادیر هدف برای آموزش الگوریتم مورد استفاده قرار گرفتند.



(الف)



(ب)

شکل ۵. الف) مقادیر محاسباتی و مشاهداتی نسبت به زمان، ۵. ب) نمودار پراکنش میان مقادیر مشاهداتی و محاسباتی

می‌تواند در کاهش خطای موجود در مدل ارائه شده مؤثر باشد.

همانطور که از نتایج قابل مشاهده است نتایج حاکی از برتری مدل پیش‌بینی فازی نسبت به مدل‌های شبکه عصبی در پیش‌بینی سطح آب زیرزمینی می‌باشد. با این حال مدل جنگل تصادفی یا RF نیز عملکرد بهتری در پیش‌بینی سطح ایستابی داشته و به‌عنوان یک روش مؤثر و کارا در بحث پیش‌بینی شناخته شد.

۴. بحث و نتیجه‌گیری

از جمله شاخص‌های بیابان‌زایی مرتبط با موضوع تخلیه آب‌های زیرزمینی می‌توان به شور شدن آبخوان و نشست زمین اشاره کرد که متأسفانه در چند سال گذشته در نقاط زیادی نمایان شده است. به‌منظور آگاهی از این وضعیت و مدیریت بهینه منابع آب، لازم است پیش‌بینی دقیقی از نوسانات تراز آب زیرزمینی صورت گیرد. عدم وجود داده‌های کافی برای توسعه مدل‌های ریاضی به علت محدودیت‌های زمانی و هزینه‌ای و همچنین زمان زیاد لازم جهت اجرای مکرر این مدل‌ها خصوصاً در شرایط ترکیب با مدل‌های بهینه‌سازی به منظور یافتن سناریوی بهینه مدیریت آب زیرزمینی، کاربرد مدل‌های ریاضی را با محدودیت مواجه می‌کند. در چنین شرایطی، مدل‌های مبتنی بر هوش مصنوعی، ماشین‌های یادگیری و داده‌کاوی می‌توانند جایگزین مناسبی به شمار آیند. در پژوهش حاضر عملکرد مدل‌های شبکه عصبی، منطق فازی و RF برای مدل‌سازی سطح آب زیرزمینی دشت بوکان با استفاده از پارامترهای بارش، دما، دبی جریان و تراز سطح ایستابی در طی دوره آماری (۱۳۹۵-۱۳۸۳) مورد ارزیابی قرار گرفت و مقادیر سطح آب زیرزمینی مشاهده‌ای با سطح آب زیرزمینی تخمین زده شده این مدل با استفاده از معیارهای ارزیابی بررسی گردید.

پیش‌بینی‌ها به‌واسطه ترکیب تعداد زیادی از پارامترهای ممکن انجام می‌گیرد و به وسیله میانگین مربعات خطا (MSE) ارزیابی می‌گردند و در نهایت مدلی که کمترین میزان (MSE) را داشته باشد به‌عنوان مدل بهینه، برای پیش‌بینی انتخاب می‌گردد. روش دیگر تجزیه و تحلیل مدل بر پایه منحنی‌های مشخصه عملکرد سیستم^۱ (ROC) می‌باشد. منحنی‌های ROC به نحوی مشابه نرخ موفقیت هستند که می‌توانند به وسیله نرخ مثبت واقعی^۲ (TPR) کنترل شوند. به‌طور کلی در منحنی‌های ROC نتایج^۳ FPR (محور X) در مقابل TPR (محور Y) رسم می‌شوند. به‌عنوان حساسیت^۴ و (1-FPR) به‌عنوان اختصاصیت^۵ شناخته می‌شوند. هر یک از نتایج در جفت (TPR- FPR) و یک سری از چنین جفت‌هایی برای رسم منحنی‌های ROC استفاده می‌شوند. مساحت زیر منحنی^۶ (AUC) در منحنی‌های ROC نیز به‌عنوان یکی از روش‌های ارزیابی خطا بکار برده می‌شود که هرچه میزان AUC به یک نزدیک باشد، مدل از صحت بیشتری برخوردار است. مطابق شکل (الف) میزان خطای مدل‌سازی با افزایش درخت‌های تصمیم کاهش می‌یابد به نحوی که در درخت تصمیم اولی میزان خطای MSE برابر با ۰/۱۰۸ بوده و هنگامی که تعداد درخت‌های تصمیم ۵۰۰ می‌گردد میزان خطا به ۰/۰۷۱ کاهش می‌یابد. همچنین مقدار AUC برابر با ۰/۹۱ می‌باشد که نمایانگر دقت بالای شبیه‌سازی است (شکل ب). البته خطای موجود در مدل نهایی به عوامل مختلفی از جمله تعداد محدود داده‌ها، ناهمگنی آبخوان، خطای ذاتی موجود در داده‌های ورودی و حتی داده‌های خروجی بستگی دارد. افزایش دقت داده‌های ورودی

¹ Receiver Operating Characteristic

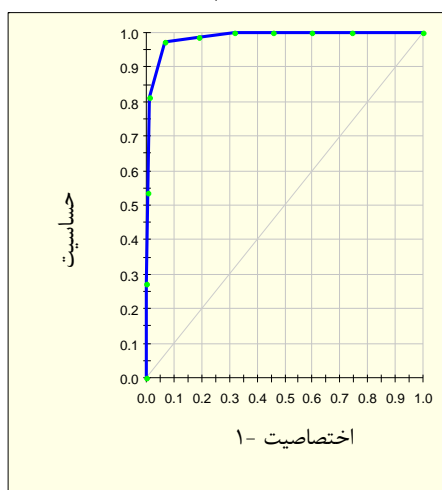
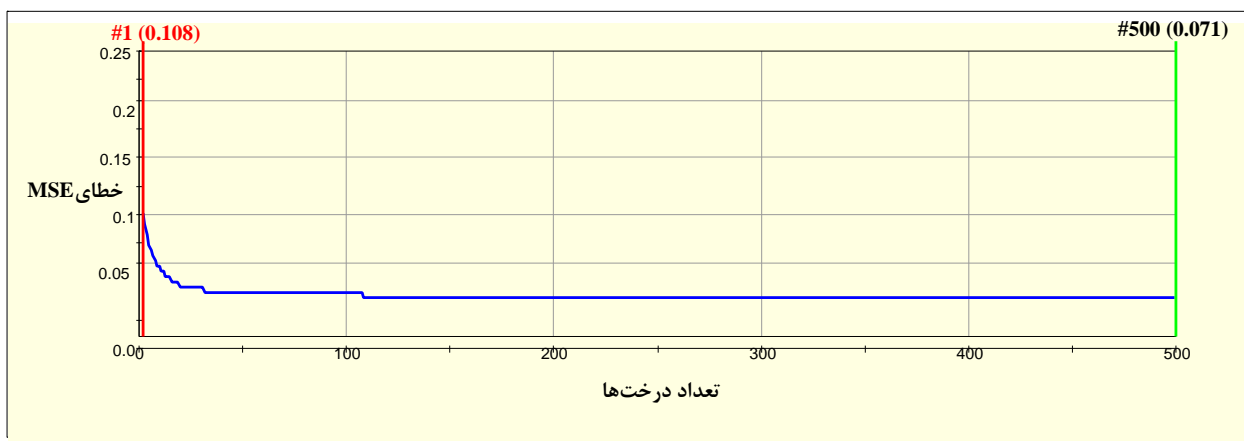
² True Positive Rate (TPR)

³ False positive rate

⁴ Sensitivity

⁵ Specificity

⁶ Area under Curve



(ب)

شکل ۷. الف) خطای MSE مدل و کاهش آن با افزایش درخت‌های تصمیم، ۵. ب) منحنی ROC حاصل از مدل پیش‌بینی

در تخمین نقاط بیشینه و کمینه ضعیف عمل کرده است و این نقاط را نزدیک به مقداری واقعی‌شان تخمین نزده است. بر اساس نتایج مدل منطق فازی ساگنو دارای بیشترین بازده، کمترین درصد خطا و توانایی بالایی در پیش‌بینی سطح ایستابی این دشت نسبت به شبکه عصبی می‌باشند، در صورتی که در مطالعه [۲۲] شبکه عصبی نسبت به منطق فازی کارایی بهتری داشته است. همچنین روش RF با داشتن مزایایی مثل یادگیری روابط غیرخطی، توانایی مقابله با داده‌های پرت و ساختگی، برآورد خطای غیرمرتبط داخلی، اجرای هزاران داده ورودی و داشتن حساسیت کمتر در برابر گیر افتادن در کمینه محلی، به عنوان یک روش دقیق در پیش‌بینی

نتایج تحقیق نشان داد که مدل جنگل تصادفی دقت بالا و خطای ناچیزی در تخمین سطح آب زیرزمینی داشته و قابلیت بسیار بالایی در تخمین برخی مقادیر کمینه، بیشینه و میانی نسبت به سایر مدل‌ها دارد. به‌منظور پیش‌بینی سطح آب زیرزمینی به کمک شبکه‌های عصبی مصنوعی، شبکه چندلایه پرسپترون با الگوریتم آموزشی لونیگ-مارکوارت به‌عنوان کاراترین شبکه شناخته شد. به‌طوری که با کمترین MSE و بیشترین مقدار R^2 به‌عنوان بهترین مدل شبکه عصبی مصنوعی شناخته شده در صورتی که سایر ساختارهای شبکه عصبی مصنوعی به‌طور میانگین از MSE و R^2 بالاتری برخوردار بودند. نتایج این مدل نشان می‌دهد که

روش فازی داشته است. در مجموع بیان می‌گردد که استفاده از مدل‌های منطق فازی و RF می‌تواند در زمینه تخمین سطح آب زیرزمینی مؤثر باشند و می‌توانند به نوبه خود برای تسهیل، توسعه و پیاده‌سازی استراتژی‌های مدیریت آب‌های زیرزمینی مفید باشد. با توجه به نتایج مشاهده شده در این پژوهش و صرف وقت زیاد برای دستیابی به مقادیر دقیق‌تر، می‌توان با ترکیب چند روش هوش مصنوعی با یکدیگر و با روش‌های داده‌کاوی در زمان انجام محاسبات صرفه‌جویی کرده و سریع‌تر به مقادیر دقیق‌تر دست یافت.

سطح ایستابی شناخته شد. نتایج صحیح مدل RF در پیش‌بینی سطح ایستابی و همچنین در مقایسه با کارهای انجام شده قبلی در این زمینه مانند مطالعات [۳]، [۱۲]، [۱۷]، [۲۲] و [۱۸]، نشان از برتری مدل ارائه شده نسبت مدل شبکه عصبی در بحث پیش‌بینی، می‌باشد. آنچه که در بیشتر مطالعات قبلی برای پیش‌بینی سطح آب زیرزمینی مورد استفاده قرار نگرفته است، استفاده از کارایی روش‌های داده‌کاوی مانند جنگل تصادفی می‌باشد که در این مطالعه از این روش استفاده گردید و نتایج بهتری هم نسبت به سایر روش‌ها مانند شبکه عصبی و

References

- [1] Abarashi, F., Mofteh Heghi, M., Sani Khani, H. and Dehghani, A.A. (2014). Comparison of performance of three intelligent methods in order to predict static level fluctuations (Case study: Zerling plain). *Water and Soil Conservation Studies*, 21(1), 163-180.
- [2] Pourmohammadi, S., Malekinejad, H. and Pourshriyati, R. (2013). Comparison of the Efficiency of Neural Network Techniques and Time Series in Groundwater Forecasting (Case Study: Bakhtegan Basin of Fars Province). *Water and Soil Conservation*, 20(4), 251-262.
- [4] Arom, Z., Forghani Tehrani, G. and Kazemi, Gh. (2014). Investigating the Factors Affecting the Qualitative Characteristics of the Abkhazian Resources of the Bukan Plain, the First National Conference on Geography, Tourism, Natural Resources and Sustainable Development, Tehran, Iran Institute of Science, Tehran University of Science and Technology.
- [3] Khasheiy Siyuki, A., Ghahreman, B. and Koochakzade, M. (2012). Comparison of Artificial Neural Network Models, ANFIS and Regression in Estimating the Staging Level of the Aquifer in Neishabour Plain. *Irrigation and drainage of Iran*, 7, 10-22.
- [5] Nadiri, A., Asgharimghdam, A., Abghari, H. and Fijani, A. (2012). Development of composite artificial intelligence models in the estimation of aquifer transferability, Case study: Tasuj plain. *Iranian Water Resources Research*, 9, 1-14.
- [6] ASCE. (2000). Task Committee on Application of Artificial Neural Networks in Hydrology, Part I and II. *Journal of Hydrology*, 5(2), 115-137.
- [7] Bellman, R. (2003). *Dynamic programming*. Mineola, NY: Dover Publications. 366 pp.
- [8] Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and regression trees*, Chapman & Hall/CRC, New York. pp.744.
- [9] Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 23-40.
- [10] Chiu, S. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, 2(4), 267-278.

- [11] Critto, A., Carlon, C. and Marcomini, A. (2003). Characterization of contaminated soil and groundwater surrounding an illegal landfill by principal component analysis and kriging. *Journal of Environmental Pollution*, 122(2), 235–44.
- [12] Daliakopoulos, I.N., Coulibaly, P. and Tsanis, I.K. (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, 309(4), 229-240.
- [13] Demuth, H., Beale, M. (2000). *Neural Network Toolbox User, s Guide*, By the Math Works. Inc Version, 4, 840pp.
- [14] Dixon, B.A. (2010). Case study using support vector machines, neural networks and logistic regression in a GIS to identify wells contaminated with nitrate-N. *Journal of Hydrogeology*, 17(6), 1507–20.
- [15] Djurovic, N., Domazet, N., Stricevic, R., Pocuca, V., Spalevic, V., Pivic, R., Gregoric, E. and Domazet, U. (2015). Comparison of groundwater level models based on artificial neural networks and ANFIS. *The Scientific World Journal*, 2, 1-14.
- [16] Emberger, L. (1952). Sur le quotient pluviothermique. *C.R. Sciences*, 234, 2508-2511.
- [17] Feng, S., Kang, S., Huo, Z., Chen, S. and Mao, X. (2008). Neural networks to simulate regional groundwater levels affected by human activities. *Groundwater*, 46(1), 80-90.
- [18] Guyon, I., Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Res*, 3, 1157–82.
- [19] Harb, N., Haddad, K. and Farkh, S. (2010). Calculation of transverse resistance to correct aquifer resistivity of groundwater saturated zones, implications for estimating its hydrogeological properties. *Lebanese science journal*, 11(1), 105-115.
- [20] Hopfield, J.J. (1982). Neural network and physical systems with emergent collective computational abilities. *Proc. Nat. Academy of scientists*, 79, 2554-2558.
- [21] Kotsiantis, S., Pintelas, P. (2004). Combining bagging and boosting. *International Journal of Computational Intelligence*, 1(4), 324–33.
- [22] Mohanty, S., Jha, M.K., Kumar, A. and Panda, D.K. (2013). Comparative evaluation of numerical model and artificial neural network for simulating groundwater flow in Kathajodi–Surua Inter-basin of Odisha, India. *Journal of Hydrology*, 495, 38–51.
- [23] Nadiri, A., Chitsazan, N., Tsai, F.T.C. and Asghari Moghaddam, A. (2014). Bayesian artificial intelligence model averaging for hydraulic conductivity estimation. *Journal of Hydrologic Engineering*, 19(3), 520-532.
- [24] Nikravesh, M., Aminzadeh, F. (2003). *Soft Computing and Intelligent Data Analysis in Oil Exploration. Part1: Introduction: Fundamentals of Soft Computing*. Elsevier, Berkeley, USA. pp.744.
- [25] Norouzi, H., Asghari Mogaddam, A. and Nadiri, A.A. (2015). Determining vulnerable areas of Malikan Plain Aquifer for Nitrate, Using Random Forest method. *Journal of Environmental Studies*, 41(4), 923-94.
- [26] Philip Plumb, A., Rowe, R.C., York, P. and Brown, M. (2005). Optimization of the predictive ability of artificial neural network (ANN) models: A comparison of three ANN programs and four classes of training algorithm. *European Journal of Pharmaceutical Sciences*, 25(5), 395–405.
- [27] Quinlan, J.R. (1993). *C4.5 programs for machine learning*. San Mateo, CA: Morgan Kaurmann. 303 pp.
- [28] Rani Sethi, R., Kumar, A., Sharma, S.P. and Verma, H.C. (2010). Prediction of water table depth in a hard rock basin by using artificial neural network. *International Journal of Water Resources and Environmental Engineering*, 2(4), 95-102.
- [29] Rodriguez, V., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sánchez, J.P. (2012d). An assessment of the effectiveness of a Random Forest classifier for land-cover classification. *ISPRS Journal of Photogram Remote Sens*, 67, 9 -104.
- [30] Schapire, R. (1990). The strength of weak learnability. *Journal of Machine learning*, 5, 197-227.

- [31] Sun, Y., Wendi, D., Kim, D.E. and Liang, S.Y. (2015). Application of artificial neural networks in groundwater table forecasting: a case study in Singapore swamp forest. *Hydrology and Earth System Science*, 12, 9317–9336.
- [32] Tayfur, G., Nadiri, A. and Asghari Moghaddam, A. (2014). Supervised intelligent committee machine method for hydraulic conductivity estimation. *Water Resources Management*, 28, 1173–1184.
- [33] Sadeghravesh, M. H., Khosravi, H. and Ghasemian, S. (2015). Application of fuzzy analytical hierarchy process for assessment of combating-desertification alternatives in central Iran. *Natural Hazards*, 75(1), 653-667.

Archive of SID

Archive of SID