

## پیش‌بینی تراز سطح آب زیرزمینی با شبکه‌های عصبی مصنوعی بر مبنای انتخاب متغیرهای مؤثر ورودی با الگوریتم اطلاعات متقابل جزئی

فرشته ظریف<sup>۱</sup>، علی عصاره<sup>۱\*</sup>، مهدی اسدی لورا<sup>۱</sup>، حسین فتحیان<sup>۲</sup> و داوود خدادادی دهکردی<sup>۱</sup>

(تاریخ دریافت: ۱۳۹۹/۹/۴؛ تاریخ پذیرش: ۱۴۰۰/۶/۲)

### چکیده

پیش‌بینی دقیق و قابل اعتماد از سطح آب زیرزمینی در یک منطقه برای استفاده پایدار و مدیریت منابع آب بسیار مهم است. این پژوهش با هدف ارزیابی شبکه‌های عصبی مصنوعی (ANNs)؛ پیش‌رونده عمومی (GFF) و تابع پایه شعاعی (RBF) در پیش‌بینی ماهانه تراز سطح آب زیرزمینی در دشت دزفول- اندیشک در جنوب غربی ایران انجام شد. برای تعیین متغیرهای مؤثر ورودی در ANN از الگوریتم اطلاعات متقابل جزئی (PMI) استفاده شد. نتایج به‌کارگیری الگوریتم PMI نشان می‌دهد که متغیرهای ورودی مؤثر بر پیش‌بینی ماهانه تراز سطح آب زیرزمینی برای پیرومترهای تحت تأثیر برداشت و تغذیه آب، فقط شامل تراز سطح آب در ماه فعلی است. همچنین متغیرهای ورودی مؤثر بر پیش‌بینی تراز سطح آب برای پیرومترهای تحت تأثیر برداشت آب، به ترتیب شامل تراز سطح آب در ماه فعلی، تراز سطح آب در یک ماه قبل، تراز سطح آب در دو ماه قبل، مختصات عرضی پیرومتر به UTM، تراز سطح آب در سه ماه قبل، تراز سطح آب در چهار ماه قبل، تراز سطح آب در پنج ماه قبل و مختصات طولی پیرومتر به UTM است. علاوه بر این متغیرهای ورودی مؤثر بر پیش‌بینی ماهانه تراز سطح آب زیرزمینی برای پیرومترهای نه تحت تأثیر برداشت و نه تغذیه آب، به ترتیب شامل تراز سطح آب در ماه فعلی، تراز سطح آب در یک ماه قبل، تراز سطح آب در دو ماه قبل، تراز سطح آب در سه ماه قبل، تراز سطح آب در چهار ماه قبل، تراز سطح آب در پنج ماه قبل، تراز سطح آب در شش ماه قبل، مختصات عرضی پیرومتر به UTM و مختصات طولی پیرومتر به UTM است. نتایج نشان می‌دهد که شبکه GFF از دقت بیشتری نسبت به شبکه RBF، در پیش‌بینی ماهانه تراز سطح آب زیرزمینی برای پیرومترهای شامل برداشت و تغذیه آب و پیرومترهای شامل فقط برداشت آب برخوردار است. علاوه بر این شبکه RBF دقت بیشتری در پیش‌بینی ماهانه تراز سطح آب زیرزمینی برای پیرومترهای شامل نه برداشت و نه تغذیه آب نسبت به شبکه GFF برخوردار است.

واژه‌های کلیدی: پیش‌بینی ماهانه، سطح آب زیرزمینی، شبکه‌های عصبی مصنوعی، الگوریتم PMI

۱. گروه علوم و مهندسی آب، واحد اهواز، دانشگاه آزاد اسلامی، اهواز، ایران.

۲. گروه مهندسی منابع آب، واحد اهواز، دانشگاه آزاد اسلامی، اهواز، ایران.

\*: مسئول مکاتبات: پست الکترونیکی: ali\_assareh\_2003@yahoo.com

## مقدمه

آب‌های زیرزمینی یکی از منابع عمده آبی در مناطق خشک و نیمه خشک است؛ که نقش مهمی در تأمین آب شرب، کشاورزی و صنعتی دارد. رشد سریع جمعیت و افزایش تقاضای آب می‌تواند منجر به بهره‌برداری بیش از حد از منابع آب زیرزمینی شود (۷). از طرف دیگر، به دلیل تأثیر تغییر اقلیم، کاهش بارندگی در این مناطق باعث افزایش فشار بر منابع آب زیرزمینی شده است. بنابراین برای چنین مناطقی کاهش سریع سطح آب زیرزمینی پیش‌بینی می‌شود (۱۳). امروزه کنترل و مدیریت در سفره‌های زیرزمینی آب به ابزاری مدیریتی برای بهره‌وری آب تبدیل شده است. تخمین تراز سطح ایستابی از مسائل مهم و اساسی است که در برنامه‌ریزی کشاورزی، مدیریت منابع آب و تعیین نیاز آبی گیاهان به‌ویژه در مواردی که از راهکارهای کم آبیاری بهره برده می‌شود، دارای اهمیت فراوانی است (۲۲). آگاهی از تراز سطح ایستابی می‌تواند در شوری و ماندابی شدن زمین و حتی زهکشی اراضی مفید باشد. با توجه به مسائل زیادی که در بررسی آب زیرزمینی وجود دارد، استفاده از مدل به‌عنوان ابزاری کارآمد مطرح شده است. تاکنون مدل‌ها و روش‌های مختلفی برای پیش‌بینی سطح آب‌های زیرزمینی ارائه شده است. مدل‌های مفهومی با مبنای فیزیکی به‌طور مرسوم، معمولاً برای مدل‌سازی هیدرولوژیکی درک فرایندهای فیزیکی رخ داده در یک سیستم خاص و رابطه بین متغیرهای هیدرولوژیکی به‌کار برده می‌شوند. با این حال، چنین مدل‌هایی به تعداد زیادی پارامتر از طریق فرایند مدل‌سازی نیاز دارند؛ که دسترسی و اندازه‌گیری همه پارامترها کار ساده‌ای نیست. علاوه بر این، برخی از روابط غیرخطی و پیچیده در میان برخی از این متغیرها وجود دارد که به‌راحتی قابل درک نیستند. تکنیک‌های متداول آماری، مانند اتورگرسیو (AR)، ARMA و ARIMA معمولاً در مدل‌سازی منابع آب زیرزمینی به‌کار برده شده‌اند (۲). اشکال عمده مرتبط با روش‌های گفته شده این است که آنها ویژگی‌های ناپایدار و غیرخطی

ساختار داده را در نظر نمی‌گیرند (۱۶). به‌تازگی، تکنیک‌های هوش مصنوعی توانایی زیادی در مدل‌سازی مسائل غیرخطی و ناپایدار ایجاد کرده‌اند (۱۴ و ۲۵). استفاده از الگوریتم‌های شبکه‌های عصبی مصنوعی (Artificial Neural Networks) برای پیش‌بینی متغیرهای هیدرولوژیکی در طی بیست سال گذشته محبوبیت یافته است (۱ و ۷). تعدادی از مطالعات کارایی مدل‌های مبتنی بر فرایند و روش‌های جدید مبتنی بر داده مانند مدل‌های شبکه عصبی مصنوعی (ANN) برای پیش‌بینی سطح آب زیرزمینی را بررسی کرده و عملکرد پیش‌بینی آنها را با یکدیگر مقایسه کرده‌اند (۴ و ۲۷). میتی و تیواری (۱۵) یک مطالعه تطبیقی از شبکه‌های عصبی مصنوعی، شبکه‌های عصبی بیزی و سیستم استنتاج عصبی فازی تطبیقی (ANFIS) در پیش‌بینی سطح آب زیرزمینی انجام دادند. نتایج نشان داد که مدل شبکه عصبی بیزی عملکرد بهتری نسبت به ANFIS و ANN دارد. راج و همکاران (۲۱) از روش ANN برای پیش‌بینی عمق سطح آب در منطقه کانیاکوماری کشور هند استفاده کردند. آنها برای از بین بردن نویز داده‌های ورودی، از تبدیل موجک استفاده کردند. دوره ۱۰ ساله بارندگی به‌عنوان داده ورودی استفاده و عمق سطح آب به‌عنوان خروجی مدل در نظر گرفته شد. رخشنده‌رو و همکاران (۲۲) نوسانات طولانی مدت روزانه سطح آب زیرزمینی را به ترتیب در چاه‌های کم عمق و عمیق واقع در فلوریدا و آرکانزاس پیش‌بینی کردند. آنها از شبکه‌های عصبی موجک استفاده شده توسط الگوریتم جستجوی هارمونی بهبود یافته استفاده کردند. نتایج نشان داد که مدل پیشنهادی در هر دو چاه، عملکرد بهتری در مقایسه با مدل‌های دارای الگوریتم‌های آموزش جایگزین دارد. وونش و همکاران (۲۶) سطح آب‌های زیرزمینی را با استفاده از شبکه‌های غیرخطی اتورگرسیو با ورودی برون‌ی (NARX) پیش‌بینی کردند. نتایج نشان‌دهنده مناسب بودن NARX برای پیش‌بینی سطح آب زیرزمینی با مجموعه کوچکی از ورودی‌ها بود. لی و همکاران (۱۳) ANN را برای پیش‌بینی سطح آب

## مواد و روش‌ها

### منطقه مورد مطالعه

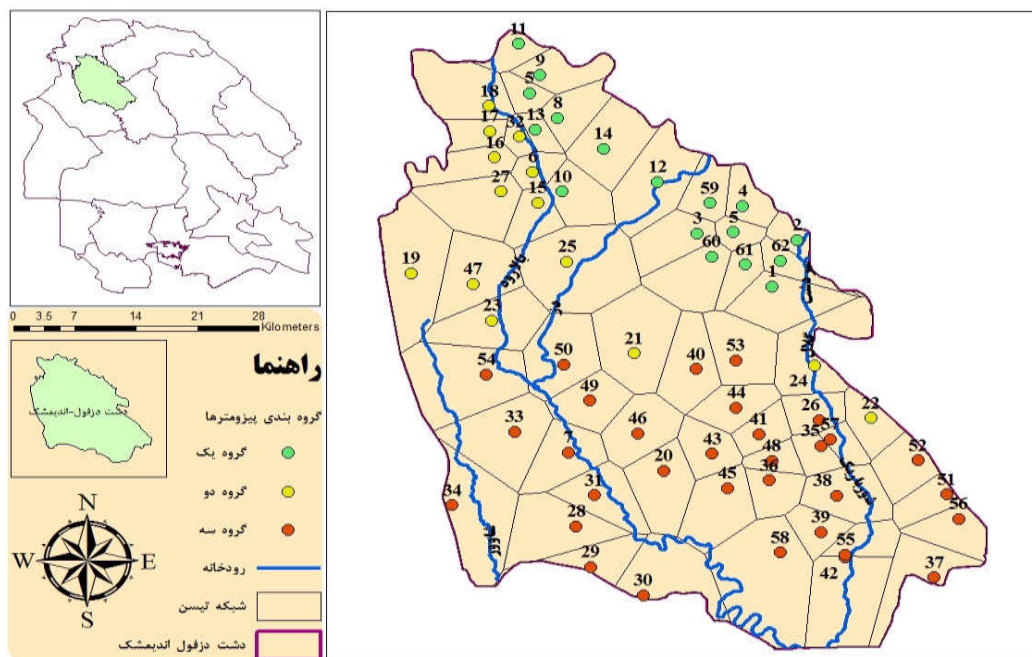
محدود مطالعاتی دزفول - اندیمشک با مساحتی حدود ۲۴۷۸ کیلومتر مربع، در ارتفاع ۱۴۳ متری از سطح دریا و مختصات جغرافیایی ۳۱°۵۱' تا ۳۳°۳۵' عرض شمالی و ۴۵°۵۰' تا ۴۸°۱۴' طول شرقی در شمال - شمال غرب استان خوزستان واقع شده است. این محدوده، بخشی از حوضه آبریز رودخانه دز است که از شمال به سراب جلدون از غرب به اوان از جنوب به آهودشت و از شرق به گتوند - عقیلی منتهی می‌شود. تراز سطح ایستابی در گستره این آبخوان بین ۴۵ تا ۲۶۰ متر در محدوده آبخوان تغییر می‌کند. برای رسیدن به اهداف این تحقیق ابتدا با رسم پلیگون‌های تیسن مساحت مربوط به هر پیزومتر مشخص شد. با توجه به موقعیت پیزومترها نسبت به رودخانه و چاه‌های بهره‌برداری از آب زیرزمینی، پیزومترها به ۳ گروه شامل: پیزومترهایی که تحت تأثیر تغذیه از رودخانه قرار می‌گیرند، پیزومترهایی که در محدوده شعاع تأثیر چاه‌های بهره‌برداری قرار دارند و پیزومترهایی که در محدوده شعاع تأثیر چاه‌های بهره‌برداری قرار ندارند، تقسیم‌بندی شدند. با توجه به بازده‌های میدانی، پیزومترهای مورد مطالعه در سه طبقه ۱ (شامل پیزومترهای تحت تأثیر برداشت و تغذیه آب)، طبقه ۲ (شامل پیزومترهای تحت تأثیر فقط برداشت آب) و طبقه ۳ (شامل پیزومترهای نه تحت تأثیر برداشت و نه تغذیه آب) دیده شدند. شکل ۱ موقعیت پیزومترها و رودخانه و جدول ۱ اسامی و مختصات پیزومترهای مورد استفاده را نشان می‌دهد. در جدول ۱، طبقه اول پیزومترها شامل شماره ۱ تا ۱۴ هستند. همچنین پیزومترهای شماره ۱۵ تا ۲۷، مربوط به طبقه دوم هستند و پیزومترهای از شماره ۲۸ تا ۵۸، مربوط به طبقه سوم هستند.

### انتخاب متغیرهای ورودی مؤثر در شبکه عصبی مصنوعی با

#### الگوریتم PMI

تنها الگوریتم غیرخطی انتخاب متغیرهای ورودی، برای تعیین

زیرزمینی در سیستم‌های آب زیرزمینی غیرقطعی و نامشخص در منطقه رودخانه یانگ یونگ، کره جنوبی به کار بردند. نتایج مطالعه آنها نشان داد که سطح رودخانه به طور غالب بر نوسان سطح آب زیرزمینی تأثیر می‌گذارد. آنها همچنین نتیجه گرفتند که مدل‌های شبکه ANN توانایی پیش‌بینی نوسانات سطح آب زیرزمینی را دارند. به‌تازگی الگوریتم‌هایی برای انتخاب متغیرهای مؤثر ورودی در مدل‌های بر مبنای پردازش داده‌ها، بر اساس روش‌های آماری، ارائه شده است. تنها الگوریتم غیرخطی برای تعیین متغیرهای ورودی مؤثر در مدل‌های بر مبنای پردازش داده‌ها، الگوریتم اطلاعات متقابل جزئی (Partial Mutual Information) است. الگوریتم PMI از شدت روابط بین ورودی‌ها و خروجی‌ها و بین متغیرهای اضافی برای انتخاب متغیرهای ورودی مؤثر استفاده می‌کند. فانگ و همکاران (۸) به پیش‌بینی تبخیر و تعرق مرجع ( $ET_0$ ) در حوضه‌های رودخانه جینگ و بیولو در چین پرداختند. آنها ابتدا متغیرهای مؤثر ورودی با الگوریتم PMI را مشخص کردند و با نتایج روش همبستگی خطی جزئی مقایسه کردند. نتایج نشان داد که الگوریتم PMI از نظر حذف متغیرهای اضافی و دستیابی به تعداد کمتر متغیرهای پیش‌بینی‌کننده، بهتر از روش همبستگی خطی جزئی است. مطالعات بالا نشان می‌دهد؛ با توجه به اهمیت منابع آب زیرزمینی در تأمین نیازهای آبی، پیش‌بینی سطح آب زیرزمینی یکی از ابزارهای مهم برای مدیریت و تصمیم‌گیری در مورد نحوه استفاده و تخصیص آب به بخش‌های مختلف است. اگرچه مدل‌های عددی و جعبه‌های سیاه توانسته‌اند نتایج پیش‌بینی قابل قبولی را ارائه دهند، اما انتخاب یک روش خاص برای پیش‌بینی در مناطق مختلف، همیشه در معرض عدم قطعیت زیادی است. لذا این تحقیق با هدف ارزیابی شبکه‌های عصبی مصنوعی بر مبنای پردازش داده‌ها در پیش‌بینی تراز سطح آب در پیزومترهای مختلف از نظر برداشت و تغذیه آب صورت گرفت. همچنین متغیرهای مؤثر ورودی در شبکه‌های عصبی مصنوعی با الگوریتم PMI انتخاب شد.



شکل ۱. نقشه موقعیت منطقه مورد مطالعه

جدول ۱. اسامی و مختصات پیزومترهای مورد مطالعه

UTMy	UTMx	نام پیزومتر	شماره پیزومتر	UTMy	UTMx	نام پیزومتر	شماره پیزومتر
۳۵۴۸۶۳۶	۲۵۷۷۰۸	Ht-5	۳۰	۳۵۸۱۵۷۰	۲۷۰۹۴۳	Gr-5	۱
۳۵۵۹۵۰۵	۲۵۳۲۰۷	Ht-9	۳۱	۳۵۸۶۴۵۶	۲۷۳۴۸۷	Gr-7	۲
۳۵۹۸۱۹۸	۲۴۶۹۲۷	Pz-4	۳۲	۳۵۸۷۳۸۳	۲۶۳۸۴۲	Sa2	۳
۳۵۶۶۵۱۲	۲۴۵۶۳۴	SU-17	۳۳	۳۵۹۰۲۶۰	۲۶۸۳۲۵	Sa4	۴
۳۵۵۸۷۹۸	۲۳۹۳۹۳	SU-6	۳۴	۳۶۰۲۸۴۶	۲۴۷۹۸۲	AN-38	۵
۳۵۶۴۳۳۹	۲۷۵۳۲۵	DM6	۳۵	۳۵۹۴۳۵۶	۲۴۸۰۹۴	Pz-2	۶
۳۵۶۰۷۲۴	۲۷۰۲۳۰	N10	۳۶	۳۵۶۴۱۵۳	۲۵۰۷۸۰	SU-14	۷
۳۵۴۹۹۲۶	۲۸۵۹۹۳	N28	۳۷	۳۶۰۰۱۱۹	۲۵۰۶۵۷	AN-39	۸
۳۵۵۸۹۳۳	۲۷۶۷۶۷	N25	۳۸	۳۶۰۴۷۳۷	۲۴۹۰۸۷	AN-2	۹
۳۵۵۵۰۳۳	۲۷۵۱۰۸	N24	۳۹	۳۵۹۲۲۱۸	۲۵۰۹۱۹	AN-34	۱۰
۳۵۷۲۸۴۴	۲۶۳۳۷۷	N19	۴۰	۳۶۰۸۱۶۹	۲۴۷۱۲۲	An-1	۱۱
۳۵۶۵۶۷۱	۲۶۹۳۳۲	Dz2	۴۱	۳۵۹۳۰۱۳	۲۶۰۱۵۰	AN-6	۱۲
۳۵۵۲۲۶۰	۲۷۷۴۲۲	DM2	۴۲	۳۵۹۸۹۰۹	۲۴۸۴۸۲	AN-37	۱۳

ادامه جدول ۱.

UTMy	UTMx	نام پیزومتر	شماره پیزومتر	UTMy	UTMx	نام پیزومتر	شماره پیزومتر
۳۵۶۳۷۰۷	۲۶۷۵۵	N11	۴۳	۳۵۹۶۷۱۱	۲۵۴۹۷۸	AN-36	۱۴
۳۵۶۸۵۷۴	۲۶۷۱۸۹	N13	۴۴	۳۵۹۱۰۸۸	۲۴۸۴۹۲	Pz-1	۱۵
۳۵۵۹۹۵۲	۲۶۶۱۳۶	N9	۴۵	۳۵۹۶۰۳۷	۲۴۴۴۸۲	Pz-3	۱۶
۳۵۶۶۰۳۸	۲۵۷۵۶۲	N15	۴۶	۳۵۹۸۸۲۷	۲۴۴۰۳۶	Pz-5	۱۷
۳۵۸۲۵۳۴	۲۴۲۰۲۱	DZ28	۴۷	۳۶۰۱۵۷۶	۲۴۴۰۹۴	Pz-8	۱۸
۳۵۶۲۸۳۹	۲۷۰۵۲۰	DZ1	۴۸	۳۵۸۳۸۴۱	۲۳۶۰۶۷	AN-27	۱۹
۳۵۶۹۶۹۹	۲۵۳۰۲۷	N16	۴۹	۳۵۶۱۹۷۵	۲۵۹۹۹۷	N12	۲۰
۳۵۷۳۶۶۰	۲۵۰۶۱۲	N21	۵۰	۳۵۷۴۶۷۴	۲۵۷۴۴۴	N20	۲۱
۳۵۵۸۸۳۶	۲۸۷۳۷۶	N4	۵۱	۳۵۶۷۲۰۵	۲۸۰۲۳۹	N2	۲۲
۳۵۶۲۵۳۹	۲۸۴۶۹۶	N3	۵۲	۳۵۷۸۵۲۰	۲۴۳۷۶۴	DZ25	۲۳
۳۵۷۳۶۹۲	۲۶۷۲۵۸	N18	۵۳	۳۵۷۲۹۸۳	۲۷۴۹۰۰	DM1	۲۴
۳۵۷۲۷۴۲	۲۴۳۰۷۴	DZ21	۵۴	۳۵۸۴۶۹۶	۲۵۱۱۸۷	DZ-42	۲۵
۳۵۵۲۵۷۶	۲۷۷۴۳۳	29N	۵۵	۳۵۶۷۱۰۹	۲۷۵۲۰۶	DM5	۲۶
۳۵۵۶۱۴۸	۲۸۸۵۳۴	DM4	۵۶	۳۵۹۲۳۶۱	۲۴۴۹۸۲	AN-44	۲۷
۳۵۶۴۹۸۸	۲۷۶۲۲۵	N5	۵۷	۳۵۵۶۱۷۹	۲۵۱۳۶۶	Ht-10	۲۸
۳۵۵۲۹۳۶	۲۷۱۱۱۹	N23	۵۸	۳۵۵۱۷۷۷	۲۵۲۶۳۰	Ht-3	۲۹

شده است. اگر ورودی معنی‌دار باشد، Cs به S (مجموعه متغیرهای ورودی انتخاب شده) اضافه شده و انتخاب کردن ادامه می‌یابد، تا جایی که هیچ ورودی معنی‌داری باقی نماند، سپس متعاقباً الگوریتم متوقف می‌شود.

#### برآورد اطلاعات متقابل جزئی (PMI)

با توجه به یک متغیر خروجی تصادفی Y، مقداری عدم قطعیت در مورد یک مشاهده y که عضوی از Y است، وجود دارد؛ که می‌توان آن را با توجه به آنتروپی شانون h (۲۳) تعریف کرد. ولی با فرض یک متغیر ورودی تصادفی X که Y وابسته به آن است، مشاهدات متقابل (x,y) این عدم قطعیت را کاهش می‌دهد، زیرا اطلاع از x،

متغیرهای ورودی مؤثر در مدل‌های بر مبنای پردازش داده‌ها، الگوریتم PMI است. الگوریتم انتخاب ورودی بر مبنای PMI (PMIS)، در ابتدا توسط شارما (۲۴) برای شناسایی متغیرهای ورودی مؤثر در مدل‌های هیدرولوژیکی توسعه داده شد. الگوریتم PMI هر تکرار را با در نظر گرفتن یک ورودی (C) و یک خروجی (Y) و با پیدا کردن یک Cs (فرض می‌شود که Cs متفاوت از C است) که مقدار PMI را با توجه به متغیر خروجی پیشینه می‌کند (با توجه به ورودی‌هایی که از قبل انتخاب شده‌اند)، انجام می‌دهد. مفهوم آماری که PMI برای Cs برآورد می‌کند بر مبنای محدوده‌های اطمینان است که از توزیع تشکیل شده توسط یک حلقه خود راه‌انداز (Bootstrap loop) تعیین

مورد ضریب همبستگی خطی (R)، توزیع تخمین نمونه از یک توزیع t پیروی می‌کند. جدول‌ها مقادیر بحرانی ضریب همبستگی (R) بر مبنای توزیع t با در نظر گرفتن تعداد نمونه‌ها و سطح اعتماد معین به سادگی تهیه می‌شوند (۵). با این حال برخلاف ضریب همبستگی خطی، یک تعریف تحلیلی معادل برای I را نمی‌توان بر طبق رابطه (۲) نتیجه‌گیری کرد (۹). از این رو پژوهشگران باید از خودراه انداز کردن، برای محاسبه  $f(\hat{I})$  کمک بگیرند (۱۰ و ۲۴). به جای استفاده از مقادیر تحلیلی، یک روش برای ساختن جدول‌های مقادیر بحرانی محاسبه شده برای I، استفاده از شبیه‌سازی مونت‌کارلو که توسط می و همکاران (۱۷) شرح داده شده، است. شبیه‌سازی مونت‌کارلو برای تعیین توزیع تخمین‌گر MI در گام اول در توسعه معیار متوقف کننده بر مبنای مقادیر بحرانی تقریبی، به کار گرفته شد. در هر شبیه‌سازی، MI برای یک سری داده‌ها در مقایسه با داده‌های نویز-سفید (White-noise) گوسین محاسبه می‌شود (با اندازه نمونه n، که مقداری بین ۵۰ تا ۵۰۰۰ دارد)، که این محاسبه برای به دست آوردن داده‌هایی با مقادیر بحرانی استفاده می‌شود و این داده‌ها می‌تواند برای آزمون مستقل بودن بر مبنای MI به کار گرفته شود. برای هر اندازه نمونه، ابتدا یک سری  $\varepsilon_y \sim N(0,1)$  ساخته می‌شود؛ و سپس توابع چگالی احتمال حاشیه‌ای  $f_{\varepsilon_y}$  محاسبه می‌شود. مجموع یکصد هزار تکرار از سری  $\varepsilon_x \sim N(0,1)$ ، مستقل از  $\varepsilon_y \sim N(0,1)$  تولید می‌شود. برای هر نمونه از  $\varepsilon_x$ ، توابع چگالی احتمال حاشیه‌ای  $f_{\varepsilon_x}$  و  $f_{\varepsilon_x \varepsilon_y}$  تخمین زده و متعاقباً  $\hat{I}(\varepsilon_x \varepsilon_y)$  ارزیابی می‌شود. نتایج مقادیر بحرانی I در جدول ۲ برای سطوح اعتماد مختلف درج شده است. دو ضابطه جایگزین متوقف کننده فرمول‌بندی شده که به وسیله آن در هر تکرار،  $I_{CSYS}$  با مقادیر بحرانی مربوطه  $I_b^{(95)}$  و  $I_b^{(99)}$  که از جدول ۲ به دست آمده مقایسه می‌شود، که از این دو معیار به جای محاسبات مستقیم از سیستم خودراه انداز، برای تعیین اینکه کدام متغیر باید انتخاب شود یا اینکه الگوریتم متوقف شود، استفاده می‌شود. حذف محاسباتی حلقه سیستم خودراه انداز باعث می‌شود تا انتخاب متغیرهای ورودی بسیار سریع‌تر انجام شود (۱۸).

اجازه می‌دهد تا مقدار y استنتاج شود و برعکس. بنا به تعریف، اطلاعات متقابل  $I(x;y)$ ، کاهش در عدم قطعیت متغیر Y به دلیل مشاهده X است (۳). این مسئله به صورت یک قسمت مشترک بین دو دایره در شکل ۲ نمایش داده شده است. این قسمت مشترک جایی است که عدم قطعیت کاهش یافته حول X و Y به ترتیب توسط آنتروپی شرطی  $H(Y|X)$  و  $H(X|Y)$  مشخص شده است. اطلاعات متقابل (MI) را می‌توان با رابطه (۱) به صورت مستقیم محاسبه کرد (۱۸):

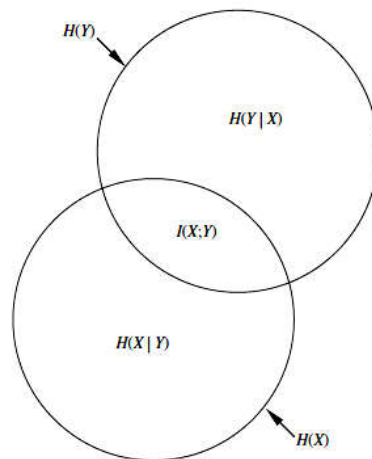
$$I(X; Y) = \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (1)$$

$p(x)$  و  $p(y)$  به ترتیب توابع چگالی احتمال حاشیه‌ای (pdfs) X و Y هستند؛ و  $p(x, y)$  تابع چگالی احتمال توأم است. به هر حال به صورت عملی، فرم صحیح توابع چگالی احتمال در رابطه (۱) مجهول است. از این رو، از برآورد چگالی‌های احتمال به جای آن استفاده می‌شود. با جایگذاری برآوردهای چگالی احتمال با تقریب عددی انتگرال در رابطه (۱) خواهیم داشت (۱۸):

$$I(X; Y) \approx \frac{1}{n} \sum_{i=1}^n \log \left[ \frac{f(x_i, y_i)}{f(x_i)f(y_i)} \right] \quad (2)$$

که در آن:  $f$  نمایانگر چگالی برآورد شده بر مبنای نمونه‌ای از n مشاهده از  $(x, y)$  است. از مبنای مختلفی برای لگاریتم در رابطه (۲) می‌توان استفاده کرد؛ ولی معمولاً از ۲ یا e استفاده می‌شود. اگر مبنای لگاریتم مشخص نشده باشد، لگاریتم طبیعی در نظر گرفته می‌شود. با فرض رابطه (۲) می‌توان گفت که برآورد دقیق و مؤثر (اطلاعات متقابل) بستگی زیادی به روش به کار گرفته شده در برآورد توابع چگالی احتمال حاشیه‌ای و توأم دارد.

در مجموع سه معیار توقف الگوریتم PMI وجود دارد که عبارتند از مقادیر بحرانی جدول‌بندی شده (Tabulated critical values)، معیار بر مبنای اطلاعات اکائیکه (Akaike information criterion) و معیار آزمون همپل (Hampel). test criterion جدول‌های مقادیر بحرانی ضریب همبستگی (R) به سهولت در دسترس هستند، که بر پایه فرمول تحلیلی برای توزیع خطای یک تخمین برای اندازه نمونه فرض شده استوار است. در



شکل ۲. دیاگرام وین (Venn)، نشان‌دهنده رابطه بین اطلاعات متقابل (MI) و آنتروپی برای خروجی  $Y$  و متغیر ورودی منفرد (۱۸)

جدول ۲. مقادیر بحرانی تخمینگر اطلاعات متقابل بر مبنای برآورد چگالی هسته (Kernel density estimation) (۱۷)

$I^{(99)}$	$I^{(95)}$	$I^{(90)}$	$\bar{I}$	n	$I^{(99)}$	$I^{(95)}$	$I^{(90)}$	$\bar{I}$	n
۰/۱۰۶۲	۰/۰۹۲۸	۰/۰۸۶۲	۰/۰۶۶۱	۲۸۰	۰/۲۷۰۵	۰/۲۲۲۴	۰/۱۹۹۰	۰/۱۳۲۳	۵۰
۰/۱۰۲۲	۰/۰۸۹۶	۰/۰۸۳۴	۰/۰۶۴۲	۳۰۰	۰/۲۴۵۲	۰/۲۰۳۱	۰/۱۸۲۵	۰/۱۲۳۶	۶۰
۰/۰۸۷۶	۰/۰۷۷۵	۰/۰۷۲۴	۰/۰۵۶۷	۴۰۰	۰/۲۲۵۴	۰/۱۸۷۹	۰/۱۶۹۴	۰/۱۱۶۶	۷۰
۰/۰۷۷۵	۰/۰۶۸۹	۰/۰۶۴۶	۰/۰۵۱۳	۵۰۰	۰/۲۰۹۱	۰/۱۷۵۶	۰/۱۵۹۲	۰/۱۱۰۶	۸۰
۰/۰۷۰۲	۰/۰۶۲۷	۰/۰۵۸۹	۰/۰۴۷۳	۶۰۰	۰/۱۸۵۸	۰/۱۵۷۲	۰/۱۴۲۹	۰/۱۰۱۳	۱۰۰
۰/۰۶۴۴	۰/۰۵۷۸	۰/۰۵۴۴	۰/۰۴۴۱	۷۰۰	۰/۱۶۸۸	۰/۱۴۳۴	۰/۱۳۰۹	۰/۰۹۴۳	۱۲۰
۰/۰۵۹۷	۰/۰۵۳۹	۰/۰۵۰۹	۰/۰۴۱۵	۸۰۰	۰/۱۵۴۶	۰/۱۳۲۱	۰/۱۲۱۱	۰/۰۸۸۳	۱۴۰
۰/۰۵۶۳	۰/۰۵۰۷	۰/۰۴۷۹	۰/۰۳۹۳	۹۰۰	۰/۱۴۴۴	۰/۱۲۳۷	۰/۱۱۳۸	۰/۰۸۳۹	۱۶۰
۰/۰۵۳۱	۰/۰۴۸۱	۰/۰۴۵۵	۰/۰۳۷۵	۱۰۰۰	۰/۱۳۵۶	۰/۱۱۶۶	۰/۱۰۷۲	۰/۰۷۹۸	۱۸۰
۰/۰۳۶۱	۰/۰۳۳۳	۰/۰۳۱۸	۰/۰۲۷۰	۲۰۰۰	۰/۱۲۷۶	۰/۱۱۰۳	۰/۱۰۱۹	۰/۰۷۶۳	۲۰۰
۰/۰۲۸۹	۰/۰۲۶۸	۰/۰۲۵۷	۰/۰۲۲۲	۳۰۰۰	۰/۱۲۱۵	۰/۱۰۵۵	۰/۰۹۷۵	۰/۰۷۳۵	۲۲۰
۰/۰۲۴۷	۰/۰۲۳۰	۰/۰۲۲۱	۰/۰۱۹۲	۴۰۰۰	۰/۱۱۵۸	۰/۱۰۰۵	۰/۰۹۳۲	۰/۰۷۰۷	۲۴۰
۰/۰۲۱۸	۰/۰۲۰۴	۰/۰۱۹۶	۰/۰۱۷۲	۵۰۰۰	۰/۱۱۰۸	۰/۰۹۶۵	۰/۰۸۹۴	۰/۰۶۸۲	۲۶۰

### معيار آزمون همپل (Z)

این، پنهان شدن داده‌های پرت از اهمیت زیادی برخوردار است، زیرا احتمال دارد که مجموعه داده‌ها دارای بیش از یک متغیر وابسته باشند. از این‌رو، یک مقدار  $Z$  اصلاح شده، که از فاصله همپل استفاده می‌کند، به جای افزایش کارایی روش بکار گرفته شده است. فاصله همپل بر پایه میانه مجموعه ورودی‌ها بنا

در فرمول‌بندی کردن یک ضابطه متوقف‌کننده بر پایه تشخیص داده‌های پرت برای الگوریتم PMIS، فرض اساسی بر این است که مجموعه داده‌ها ابتدا دارای برخی متغیرهای اضافی و غیروابسته بوده و متغیرهای مهم شناسایی خواهند شد. با وجود

غیرخطی هستند و دوم، آنها بیش از حد به هم متصل شده‌اند به طوری که یک عنصر از یک لایه، تمام عناصر دیگر لایه بعدی را تغذیه می‌کند (۱۲). شبکه‌های RBF نوعی شبکه عصبی پیشخور بوده که ساختار آنها شبیه شبکه‌های MLP است. شبکه‌های RBF نیازمند نرون‌های بیشتری نسبت به شبکه‌های استاندارد پیشخور با الگوریتم پس‌انتشار خطا هستند؛ ولی این شبکه‌ها را می‌توان در زمانی کوتاه‌تر نسبت به شبکه‌های پیشخور آموزش داد (۱۱). این نوع شبکه الگوی ورودی  $n$  بعدی را با استفاده از گره‌های واقع در لایه مجاور به یک الگوی خروجی  $z$  بعدی نگاشت می‌کند. لایه ورودی مقادیر ورودی را به هر یک از گره‌های لایه پنهان ارسال می‌کند. هر یک از گره‌های واقع در لایه پنهان با یک تابع انتقال (f) مشخص می‌شوند که سیگنال ورودی را انتقال می‌دهند. برای  $p$  امین الگوی ورودی ( $X_p$ ) پاسخ  $p$  امین گره پنهان ( $y_j$ ) برابر است با:

$$y_j = f\left(\frac{\|X^p - U_j\|}{2\sigma_j^2}\right) \quad (5)$$

که در آن  $\|X^p - U_j\|$  فاصله هندسی،  $U_j$  مرکز  $p$  امین تابع با پایه شعاعی و  $\sigma_j$  واریانس توابع با پایه شعاعی است که توسط کاربر تعریف می‌شود (۱۱). خروجی شبکه با استفاده از مجموع وزن‌دار خطی پاسخ‌های لایه پنهان در هر یک از گره‌های خروجی به دست می‌آید. خروجی  $k$  امین گره در لایه خروجی ( $Z_{kj}$ ) از رابطه (۶) محاسبه می‌شود:

$$Z_k = \sum_{j=1}^{N_H} y_j w_{kj} \quad (6)$$

که در آن:  $w_{kj}$  وزن پیوند بین گره‌های پنهان و خروجی و  $N_H$  تعداد گره لایه پنهان است. از متداول‌ترین توابع انتقال RBF، می‌توان دو تابع گاوسی و سیگموئیدی را نام برد.

#### شاخص‌های ارزیابی شبکه‌ها

برای مقایسه تراز سطح آب مشاهداتی با تراز سطح آب پیش‌بینی شده در مراحل آموزش و آزمون شبکه‌ها از شاخص

نهاده شده است (۶). نقطه از کارافتادگی آزمون همپل  $\frac{n}{2}$  است (۶ و ۲۰). آزمون همپل با محاسبه انحراف مطلق از میانه PMI برای همه ورودی‌ها به صورت رابطه (۳) شروع می‌شود (۱۸):

$$d_j = \left| I_{C_j Y.S} - I_{C_j Y.S}^{(50)} \right| \quad (3)$$

که در آن:  $d_j$  نمایانگر انحراف مطلق،  $I_{C_j Y.S}$  برابر با مقدار PMI برای متغیر  $x$  و  $I_{C_j Y.S}^{(50)}$  نشان‌دهنده میانه PMI برای مجموعه ورودی‌های  $C$  است. سپس فاصله همپل را می‌توان به صورت رابطه (۴) تعیین کرد (۱۸):

$$Z_j = \frac{d_j}{1.4826d_j^{(50)}} \quad (4)$$

که  $Z_j$  نشان‌دهنده فاصله همپل برای مجموعه ورودی  $C_j$  و  $d_j^{(50)}$  نشان‌دهنده میانه انحراف مطلق  $d_j^{(50)}$  (MAD) است. ضریب  $1/4826$  فاصله را به نحوی تغییر داده تا قانون  $Z > 3$  را بتوان اعمال کرد، همان طوری که در مورد  $Z$ -test متعارف به کار برده شده است (۲۰). با استفاده از این معیار متوقف کننده، الگوریتم PMI بر پایه انتخاب ورودی، دیگر حاوی حلقه خودراه انداز نیست، و نیز PMI با هیچ مقدار بحرانی از  $I$  مقایسه نمی‌شود. در عوض، مقدار  $Z_s$  برای برگزیدن  $C_s$  تعیین می‌شود و اگر  $Z_s > 3$  باشد، ورودی، انتخاب شده و به  $S$  اضافه می‌شود، در غیر این صورت ادامه عملیات الگوریتم انتخاب متغیرهای ورودی متوقف می‌شود.

#### شبکه‌های عصبی مصنوعی

در این تحقیق دو نوع شبکه پیش‌رونده عمومی (General Feed Forward) و شبکه تابع پایه شعاعی (Radial basis function) برای پیش‌بینی تراز سطح آب در پیزومترها مورد استفاده قرار گرفت. شبکه (GFF) تعمیم یافته شبکه پرسپترون چندلایه (MLP) است که در آن اتصالات در یک یا چند لایه می‌تواند برقرار باشد. یک GFF از نظر تئوری مسائل را به روشی کارآمدتر نسبت به شبکه MLP حل می‌کند. GFF دو ویژگی قابل توجه دارد. اول، عناصر پردازش آن (PE)



تابع انتقال برای لایه میانی و خروجی، تعداد نرون‌ها در لایه میانی و تعداد تکرار آموزش در نظر گرفته شد. سپس اقدام به آموزش و آزمون هر یک از شبکه‌ها شد و برای هر شبکه شاخص‌های آماری NSE، RMSE و MAE بین تراز سطح آب مشاهداتی و پیش‌بینی شده محاسبه شد. برای این منظور در گروه اول چاهک‌های پیژومتری که مربوط به چاهک‌های شامل برداشت و تغذیه آب بود، برای آموزش شبکه‌های عصبی مصنوعی از داده‌های تراز سطح آب فروردین ۱۳۸۰ تا اسفند ۱۳۹۳ و برای آزمون (تست) شبکه‌ها از داده‌های تراز سطح آب فروردین ۱۳۹۴ تا آذر ۱۳۹۵ استفاده شد. برای گروه دوم چاهک‌های پیژومتری که مربوط به چاهک‌های شامل برداشت آب بود؛ برای آموزش شبکه‌های عصبی مصنوعی از داده‌های شهریور ۱۳۸۰ تا اسفند ۱۳۹۳ و برای آزمون شبکه‌ها از داده‌های فروردین ۱۳۹۴ تا آبان ۱۳۹۵ استفاده شد. برای گروه سوم چاهک‌های پیژومتری که مربوط به چاهک‌های شامل نه برداشت و نه تغذیه آب بود، برای آموزش شبکه‌های عصبی مصنوعی از داده‌های مهر ۱۳۸۰ تا اسفند ۱۳۹۳ و برای آزمون شبکه‌ها از داده‌های فروردین ۱۳۹۴ تا بهمن ۱۳۹۵ استفاده شد. شکل ۳ روند اجرای پژوهش را نشان می‌دهد.

## نتایج و بحث

تعیین متغیرهای ورودی مؤثر بر متغیر خروجی با به‌کارگیری

### الگوریتم PMI

جدول‌های ۳ تا ۵ نتایج الگوریتم PMI برای ۱۰ متغیر مؤثر به ترتیب اولویت برای طبقه اول، دوم و سوم پیژومترها را نشان می‌دهد. نتایج نشان داد با در نظر گرفتن ضابطه همپل، متغیرهای ورودی مؤثر بر پیش‌بینی تراز سطح آب در یک ماه آینده برای پیژومترهای تحت تأثیر برداشت و تغذیه آب، فقط تراز سطح آب در ماه فعلی است (جدول ۳).

نتایج تحقیق نشان داد با در نظر گرفتن ضابطه همپل، متغیرهای ورودی مؤثر بر پیش‌بینی تراز سطح آب در یک ماه آینده در

های آماری ضریب ناش - ساتکالیف (NSE) و ریشه میانگین مربعات خطا (RMSE)، میانگین قدرمطلق خطا (MAE) استفاده شد، که به ترتیب در روابط (۷) تا (۹) آمده است. ضریب ناش - ساتکالیف بیانگر راندمان مدل بوده که به تازگی در مسائل هیدرولوژی به وفور مورد استفاده قرار گرفته است. ضریب ناش - ساتکالیف می‌تواند مقادیری از منفی بی‌نهایت تا یک را به خود اختصاص دهند که عدد یک نشان‌دهنده برآزش کامل بین مقادیر مشاهداتی و پیش‌بینی شده است (۱۹):

$$NSE = 1 - \left[ \frac{\sum_{i=1}^n (S_i - O_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \right] \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (S_i - O_i)^2} \quad (8)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |S_i - O_i| \quad (9)$$

که در آنها: n برابر با تعداد داده‌ها،  $O_i$  و  $S_i$  به ترتیب داده‌های تراز سطح آب مشاهداتی و پیش‌بینی شده در گام زمانی  $t$ ام و  $\bar{O}$  میانگین تراز سطح آب مشاهداتی است.

برای تهیه مجموعه پتانسیل متغیرهای ورودی، هر یک از متغیرهای ورودی شامل بارش، دمای هوا، تبخیر و تعرق، برداشت آب از چاه و تغذیه چاه نیز تا ۶ ماه قبل در نظر گرفته شد. همچنین با توجه به اینکه تراز سطح آب در یک پیژومتر وابسته به تراز سطح آب در ماه‌های قبل است، تأخیر زمانی تا ۶ ماه برای تراز سطح آب در نظر گرفته شد. سپس متغیرهای ورودی مؤثر بر متغیر خروجی (تراز سطح آب در یک ماه آینده) با استفاده از الگوریتم PMI به دست آمد. این الگوریتم به صورت پیش فرض در نظر می‌گیرد که متغیرهای ورودی دارای چولگی صفر هستند و از یک توزیع نرمال پیروی می‌کنند. در همین راستا برای استفاده از این الگوریتم بایستی ابتدا داده‌ها را نرمال کرد. برای نرمال‌سازی داده‌ها از تبدیل لگاریتمی استفاده شد و داده‌ها به صورت تبدیل شده به الگوریتم معرفی شدند. برای تعیین مشخصات بهترین شبکه‌های GFF و RBF، ساختارهای مختلف از لحاظ نوع الگوریتم آموزش شبکه، نوع



شکل ۳. روندنمای مراحل تعیین شبکه برتر در پیش‌بینی سطح آب زیرزمینی برای هر طبقه از پیزومترها

جدول ۳. نتایج الگوریتم PMI برای طبقه اول پیزومترها شامل برداشت و تغذیه آب

Hampel	MC-I*(99)	MC-I*(95)	I(x,y)	متغیر	تکرار
۶/۰۳	۰/۰۳۲	۰/۰۲۹	۲/۴۸	logWT(t)	۰
۱/۹۲	۰/۰۳۲	۰/۰۲۹	۰/۴۴	logWT(t-1)	۱
۱/۰۸	۰/۰۳۲	۰/۰۲۹	۰/۲۱	logWT(t-2)	۲
۲/۲۴	۰/۰۳۲	۰/۰۲۹	۰/۱۴	logWT(t-3)	۳
۱/۳۲	۰/۰۳۲	۰/۰۲۹	۰/۱۱	logWT(t-4)	۴
۱/۷۲	۰/۰۳۲	۰/۰۲۹	۰/۱	logWT(t-6)	۵
۱/۳۸	۰/۰۳۲	۰/۰۲۹	۰/۰۹	logWT(t-5)	۶
۱/۳۵	۰/۰۳۲	۰/۰۲۹	۰/۰۸	UTMx	۷
۰/۹۹	۰/۰۳۲	۰/۰۲۹	۰/۰۲	logDis(t-1)	۸
۲/۹۷	۰/۰۳۲	۰/۰۲۹	۰/۰۳	UTMy	۹

ماه قبل و مختصات طولی پیزومتر به UTM است (جدول ۴). همچنین با در نظر گرفتن ضابطه همپل، متغیرهای ورودی مؤثر بر پیش‌بینی تراز سطح آب در یک ماه آینده برای پیزومترهای نه تحت تأثیر برداشت و نه تغذیه آب، به ترتیب شامل تراز سطح آب

پیزومترهای تحت تأثیر فقط برداشت آب، به ترتیب شامل تراز سطح آب در ماه فعلی، تراز سطح آب در یک ماه قبل، تراز سطح آب در دو ماه قبل، مختصات عرضی پیزومتر به UTM، تراز سطح آب در سه ماه قبل، تراز سطح آب در چهار ماه قبل، تراز سطح آب در پنج

جدول ۴. نتایج الگوریتم PMI برای طبقه دوم پیزومترها شامل فقط برداشت آب

تکرار	متغیر	I(x;y)	MC-I*(95)	MC-I*(99)	Hampel
۰	WT(t)	۳/۲	۰/۰۳۲	۰/۰۳۵	۶۸/۴۳
۱	WT(t-1)	۰/۴۴	۰/۰۳۲	۰/۰۳۵	۹/۸۸
۲	WT(t-2)	۰/۲۱	۰/۰۳۲	۰/۰۳۵	۴/۹۹
۳	UTMy	۰/۱۸	۰/۰۳۲	۰/۰۳۵	۴/۶
۴	WT(t-3)	۰/۲۴	۰/۰۳۲	۰/۰۳۵	۵/۹۳
۵	WT(t-4)	۰/۱۹	۰/۰۳۲	۰/۰۳۵	۵/۴۲
۶	WT(t-5)	۰/۱۶	۰/۰۳۲	۰/۰۳۵	۴/۲
۷	UTMx	۰/۱۲	۰/۰۳۲	۰/۰۳۵	۳/۳
۸	logDis(t)	۰/۰۷	۰/۰۳۲	۰/۰۳۵	۱/۹۹
۹	logDis(t-5)	۰/۰۴	۰/۰۳۲	۰/۰۳۵	۲/۳۴

در ماه فعلی، تراز سطح آب در یک ماه قبل، تراز سطح آب در دو ماه قبل، تراز سطح آب در سه ماه قبل، تراز سطح آب در چهار ماه قبل، تراز سطح آب در پنج ماه قبل، تراز سطح آب در شش ماه قبل، مختصات عرضی پیزومتر به UTM و مختصات طولی پیزومتر به UTM است (جدول ۵).

**پیش‌بینی تراز سطح آب زیرزمینی با شبکه‌های عصبی مصنوعی**  
مشخصات بهترین ساختار شبکه GFF و RBF در پیش‌بینی تراز سطح آب زیرزمینی برای سه طبقه از پیزومترها به ترتیب در جدول‌های ۶ و ۷ نشان داده شده است. با مقایسه مقادیر RMSE و  $R^2$  دو شبکه GFF و RBF برای طبقه اول و دوم پیزومترها، می‌توان دریافت که شبکه GFF از RMSE کمتر و  $R^2$  بیشتری نسبت به شبکه RBF برخوردار است. این در حالی است که در طبقه سوم پیزومترها، شبکه RBF از RMSE کمتر و  $R^2$  بیشتری نسبت به شبکه GFF برخوردار است. بنابراین می‌توان نتیجه گرفت که در پیزومترهای طبقه اول و دوم، شبکه GFF، و در پیزومترهای طبقه سوم شبکه RBF از دقت بیشتری در پیش‌بینی تراز سطح آب زیرزمینی برخوردار است.

نتایج شاخص‌های آماری در پیش‌بینی تراز سطح آب با شبکه

عصبی GFF و RBF برای پیزومترهای مختلف شامل برداشت و تغذیه آب (جدول ۸)، برای پیزومترهای شامل فقط برداشت آب (جدول ۹) و برای پیزومترهای شامل نه برداشت و نه تغذیه آب (جدول ۱۰) نشان داده شده است. با توجه به مقادیر شاخص‌های آماری، شبکه عصبی مصنوعی GFF از دقت بیشتری نسبت به شبکه RBF در پیش‌بینی ماهانه تراز سطح آب زیرزمینی برای پیزومترهای شامل برداشت و تغذیه آب و پیزومترهای شامل فقط برداشت آب برخوردار است. به طوری که مقدار ضریب NSE در اکثر پیزومترها برای شبکه GFF بیشتر از شبکه RBF است. همچنین مقدار RMSE در اکثر پیزومترها برای شبکه GFF کمتر از شبکه RBF است. نتایج نشان داد که شبکه عصبی RBF، از دقت بیشتری نسبت به شبکه GFF، در پیش‌بینی ماهانه تراز سطح آب زیرزمینی برای پیزومترهای شامل نه برداشت و نه تغذیه آب برخوردار است (جدول ۱۰). به طوری که مقدار ضریب NSE در اکثر پیزومترها برای شبکه RBF بیشتر از شبکه GFF است. همچنین مقدار RMSE در اکثر پیزومترها برای شبکه RBF کمتر از شبکه GFF است.

شکل ۴ مقایسه تراز سطح آب زیرزمینی پیش‌بینی شده با شبکه عصبی GFF و RBF برای پیزومترهای شماره ۴ و ۷ شامل

جدول ۵. نتایج الگوریتم PMI برای طبقه سوم پیزومترها شامل نه برداشت و نه تغذیه آب

Hampel	MC-I*(99)	MC-I*(95)	I(x;y)	متغیر	تکرار
۵۴/۷۲	۰/۰۲۱	۰/۰۱۹	۳/۲	WT(t)	۰
۸/۹۳	۰/۰۲۱	۰/۰۱۹	۰/۴۷	WT(t-1)	۱
۵/۸۹	۰/۰۲۱	۰/۰۱۹	۰/۲۹	WT(t-2)	۲
۲۰/۱۷	۰/۰۲۱	۰/۰۱۹	۰/۲۴	WT(t-3)	۳
۸/۷۹	۰/۰۲۱	۰/۰۱۹	۰/۲۲	WT(t-4)	۴
۲۳/۰۱	۰/۰۲۱	۰/۰۱۹	۰/۲۱	WT(t-5)	۵
۱۸/۵۸	۰/۰۲۱	۰/۰۱۹	۰/۲	WT(t-6)	۶
۱۱/۶۶	۰/۰۲۱	۰/۰۱۹	۰/۰۷	UTMy	۷
۱۲/۰۷	۰/۰۲۱	۰/۰۱۹	۰/۰۸	UTMx	۸
۰/۹۸	۰/۰۲۱	۰/۰۱۹	۰/۰۱	ET(t)	۹

جدول ۶. مشخصات بهترین ساختار شبکه GFF در پیش‌بینی تراز سطح آب زیرزمینی برای سه طبقه از پیزومترها

Group	ساختار	تابع انتقال در		نام الگوریتم	epoch	Train				test			
		لایه میانی	لایه خروجی			R <sup>2</sup>	MAE (m)	RMSE (m)	NSE	R <sup>2</sup>	MAE (m)	RMSE (m)	NSE
۱	۱-۱-۱	tansig	purelin	trainbfg	۱۰۰۰	۰/۹۹۹	۰/۷۲۹	۰/۴۶۵	۰/۹۹۹	۰/۹۹۹	۰/۵۴۴	۰/۳۷۱	۰/۹۹۹
۲	۱-۴-۸	tansig	purelin	Traincgp	۱۰۰۰	۰/۹۹۸	۰/۶۴۰	۰/۴۰۴	۰/۹۹۸	۰/۹۹۸	۰/۷۰۶	۰/۴۹۰	۰/۹۹۷
۳	۱-۹-۹	tansig	purelin	Trainlm	۱۰۰۰	۰/۹۹۹	۰/۴۷۶	۰/۳۲۰	۰/۹۹۸	۰/۹۹۸	۰/۵۶۱	۰/۳۷۹	۰/۹۹۸

جدول ۷. مشخصات بهترین ساختار شبکه RBF در پیش‌بینی تراز سطح آب زیرزمینی برای سه طبقه از پیزومترها

Group	ساختار	epoch	spread	train				test			
				R <sup>2</sup>	MAE (m)	RMSE (m)	NSE	R <sup>2</sup>	MAE (m)	RMSE (m)	NSE
۱	۱-۴-۱	۱۰۰۰	۰/۹	۰/۹۹۹	۰/۷۲۲	۰/۴۶۵	۰/۹۹۹۳	۰/۹۹۹	۰/۵۴۳	۰/۳۷۱	۰/۹۹۹۱
۲	۸-۲۴-۱	۱۰۰۰	۰/۷	۰/۹۹۸	۰/۶۱۵	۰/۳۸۹	۰/۹۹۸۴	۰/۹۹۷	۰/۷۳۶	۰/۵۰۸	۰/۹۹۷۴
۳	۹-۵۰-۱	۱۰۰۰	۱	۰/۹۹۹	۰/۴۹۴	۰/۳۳۰	۰/۹۹۸۶	۰/۹۹۸	۰/۵۵۳	۰/۳۷۴	۰/۹۹۸۲

جدول ۸. مقادیر شاخص‌های آماری در پیش‌بینی تراز سطح آب زیرزمینی با شبکه‌های GFF و RBF برای پیژومترهای شامل برداشت و تغذیه آب

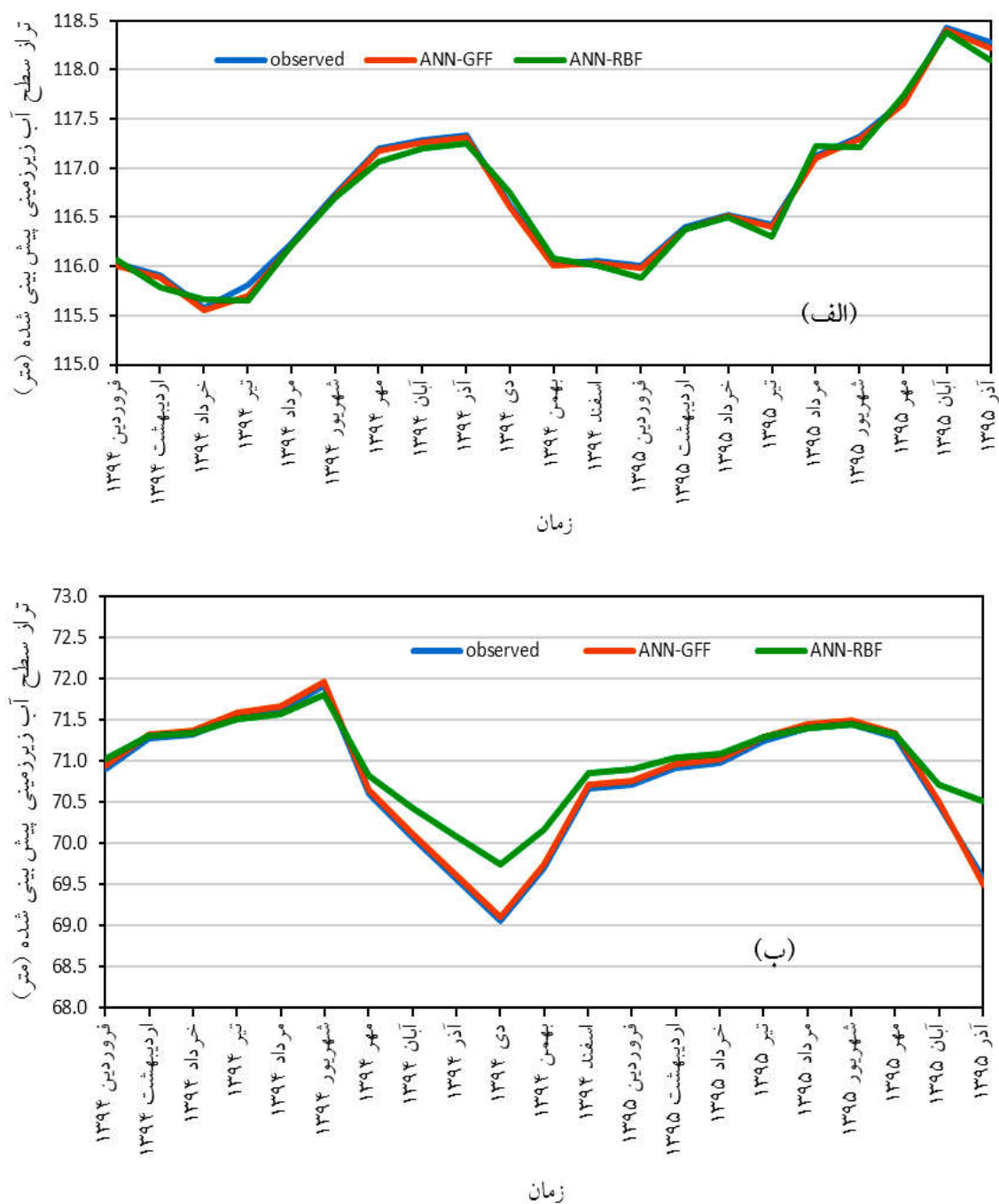
ANN- RBF		ANN-GFF		شماره پیژومتر
RMSE (m)	NSE	RMSE (m)	NSE	
۰/۳۱۷	۰/۷۱۱	۰/۳۱۷	۰/۷۱۱	۱
۰/۴۱۶	۰/۷۸۴	۰/۴۱۴	۰/۷۸۷	۲
۰/۱۲۹	۰/۴۷۵	۰/۱۲۹	۰/۴۷۴	۳
۰/۳۹۵	۰/۷۴۶	۰/۳۹۴	۰/۷۴۷	۴
۰/۹۶۶	۰/۵۹۳	۰/۹۶۸	۰/۵۹۲	۵
۰/۶۵۵	۰/۶۱۸	۰/۶۴۸	۰/۶۲۷	۶
۰/۵۵۴	۰/۴۷۴	۰/۵۴	۰/۵	۷
۰/۷۱۹	۰/۲۹۹	۰/۷۱۵	۰/۳۰۷	۸
۰/۵۹	۰/۹	۰/۵۹۵	۰/۸۹۹	۹
۰/۵۲	۰/۵۲۷	۰/۵۱۷	۰/۵۳۳	۱۰
۰/۶۰۲	۰/۸۰۵	۰/۶۱۹	۰/۷۹۴	۱۱
۰/۷۶۶	۰/۳۴۵	۰/۷۷۱	۰/۳۳۷	۱۲
۰/۳۶۸	۰/۳۵	۰/۳۷۱	۰/۳۳۹	۱۳
۰/۵۹۸	۰/۴۶۱	۰/۶	۰/۴۵۷	۱۴

جدول ۹. مقادیر شاخص‌های آماری در پیش‌بینی تراز سطح آب زیرزمینی با شبکه‌های GFF و RBF برای پیژومترهای شامل فقط برداشت آب

ANN- RBF		ANN-GFF		شماره پیژومتر
RMSE (m)	NSE	RMSE (m)	NSE	
۰/۳۴۸	۰/۹۰۸	۰/۴۳۵	۰/۸۵۶	۱
۰/۵۴۷	۰/۸۳۴	۰/۵۴۹	۰/۸۳۲	۲
۰/۷۱۰	۰/۷۸۳	۰/۶۹۲	۰/۷۹۳	۳
۰/۶۸۵	۰/۸۱۴	۰/۷	۰/۸۰۶	۴
۱/۱۳۲	۰/۰۷۱	۱/۱۱	۰/۱۰۵	۵
۰/۴۱۱	۰/۵۵۶	۰/۴۱۶	۰/۵۴۵	۶
۰/۴۹۲	۰/۵۸۸	۰/۵۱۸	۰/۵۴۳	۷
۰/۳۰۵	۰/۰۰۳	۰/۳۳۵	۰/۱۹۹	۸
۰/۴۰۱	۰/۲۴۹	۰/۳۹۴	۰/۲۰۵	۹
۱/۴۶۹	۰/۱۳۱	۱/۲۵۷	۰/۱۷۱	۱۰
۰/۹۳۳	۰/۶۰۴	۰/۹۲۲	۰/۶۱۴	۱۱
۰/۴۳۹	۰/۳۶۷	۰/۴۳۳	۰/۳۸۵	۱۲
۰/۶۸۶	۰/۸۷۷	۰/۶۷۱	۰/۸۸۳	۱۳

جدول ۱۰. مقادیر شاخص‌های آماری در پیش‌بینی تراز سطح آب زیرزمینی با شبکه‌های GFF و RBF برای پیزومترهای شامل نه برداشت و نه تغذیه آب

ANN- RBF		ANN-GFF		شماره پیزومتر
RMSE (m)	NSE	RMSE (m)	NSE	
۰/۶۵۴	۰/۱۷۳	۰/۸۳۲	۰/۳۱۱	۱
۰/۳۷۶	۰/۷۹۶	۰/۴۲۱	۱/۲۵	۲
۰/۲۹۲	۰/۵۸۳	۰/۲۴۷	۰/۷	۳
۰/۶۴۵	۰/۳۵۷	۰/۶۴۳	۰/۳۵	۴
۱/۰۳	۰/۶۳۷	۱/۰۵	۰/۶۲	۵
۱/۳۴	۰/۱۱۳	۱/۳۳	۰/۱۲۶	۶
۰/۴۷۲	۰/۰۲۹	۰/۵۲	۰/۲۶۹	۷
۰/۳۳۹	۰/۵۶۵	۰/۳۳۷	۰/۵۶۹	۸
۰/۱۰۸	۰/۷۸۱	۰/۱۱۴	۰/۹۸۱	۹
۰/۴۱۲	۰/۱۷۶	۰/۴۱۸	۰/۲۰۸	۱۰
۰/۲۳۴	۰/۳۸۸	۰/۲۳۸	۰/۳۶۸	۱۱
۰/۳۲	۰/۰۲۱	۰/۲۸۶	۰/۲۱۶	۱۲
۰/۷۹۵	۰/۴۹۴	۰/۸۸۵	۰/۳۷۳	۱۳
۰/۷۶	۰/۵۳۹	۰/۸۱۳	۰/۴۷۲	۱۴
۰/۱۴۴	۰/۷۸۳	۰/۱۵۳	۰/۷۵۵	۱۵
۰/۷۲۲	۰/۴۵	۰/۶۷۴	۰/۵۲۱	۱۶
۰/۳۵۷	۰/۶۰۳	۰/۳۵۷	۰/۶۰۶	۱۷
۰/۷۰۹	۰/۳۳۸	۰/۶۹۷	۰/۳۶۱	۱۸
۰/۶۵	۰/۵۱۰	۰/۶۲	۰/۵۵۵	۱۹
۰/۵۱۴	۰/۲۴۳	۰/۵۵	۰/۱۲۸	۲۰
۰/۰۸۱	۰/۳۷۱	۰/۰۹	۰/۹۶۶	۲۱
۰/۵۷۴	۰/۴۷۵	۰/۵۶	۰/۴۳۴	۲۲
۰/۳۰۷	۰/۱۹۸	۰/۳۰۶	۰/۲۰۵	۲۳
۰/۵۱۴	۰/۴۵۶	۰/۵۵۹	۰/۳۵۸	۲۴
۰/۲۲۵	۰/۴۷۵	۰/۲۲۵	۰/۴۷۴	۲۵
۰/۵۷۸	۰/۴۶۹	۰/۵۱۴	۰/۵۸	۲۶
۰/۳۸	۰/۳۷۴	۰/۳۹۴	۰/۴۷۳	۲۷
۰/۲۸۴	۰/۳۳۲	۰/۲۸۱	۰/۳۴۸	۲۸
۰/۳۴۹	۰/۱۵۵	۰/۳۴۷	۰/۱۶۵	۲۹
۰/۴۶۶	۰/۷۴۹	۰/۴۶۴	۰/۷۵۲	۳۰
۰/۴۱۸	۰/۲۳	۰/۴۳۹	۰/۱۵۱	۳۱

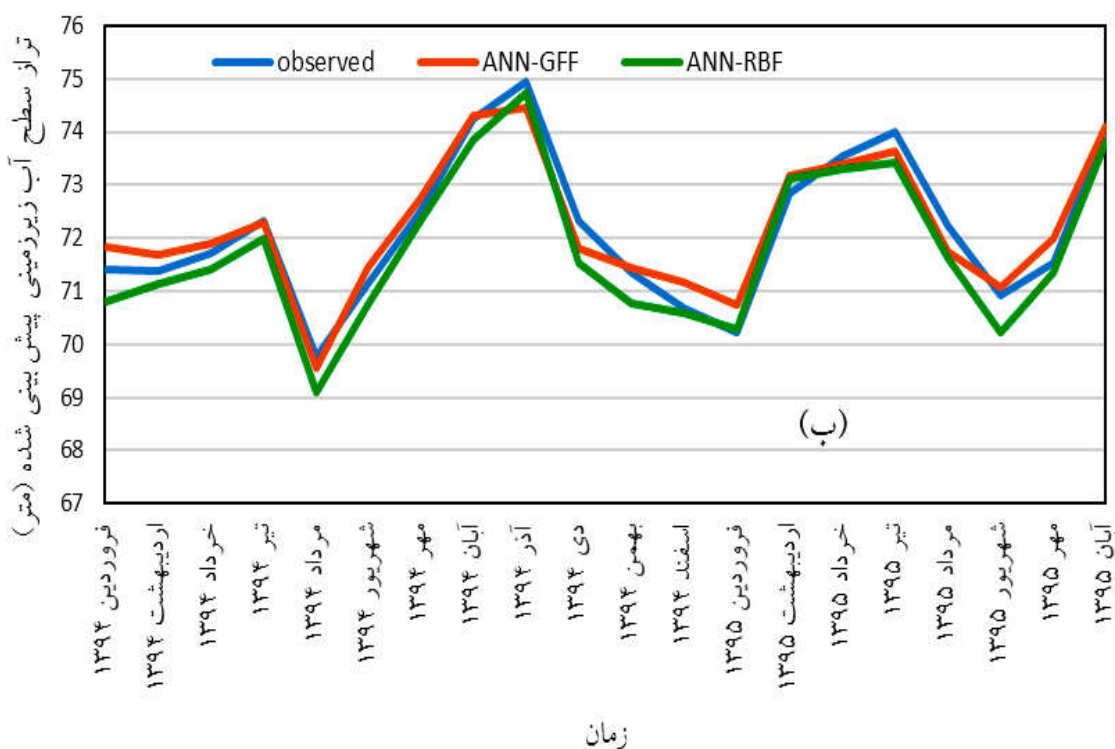
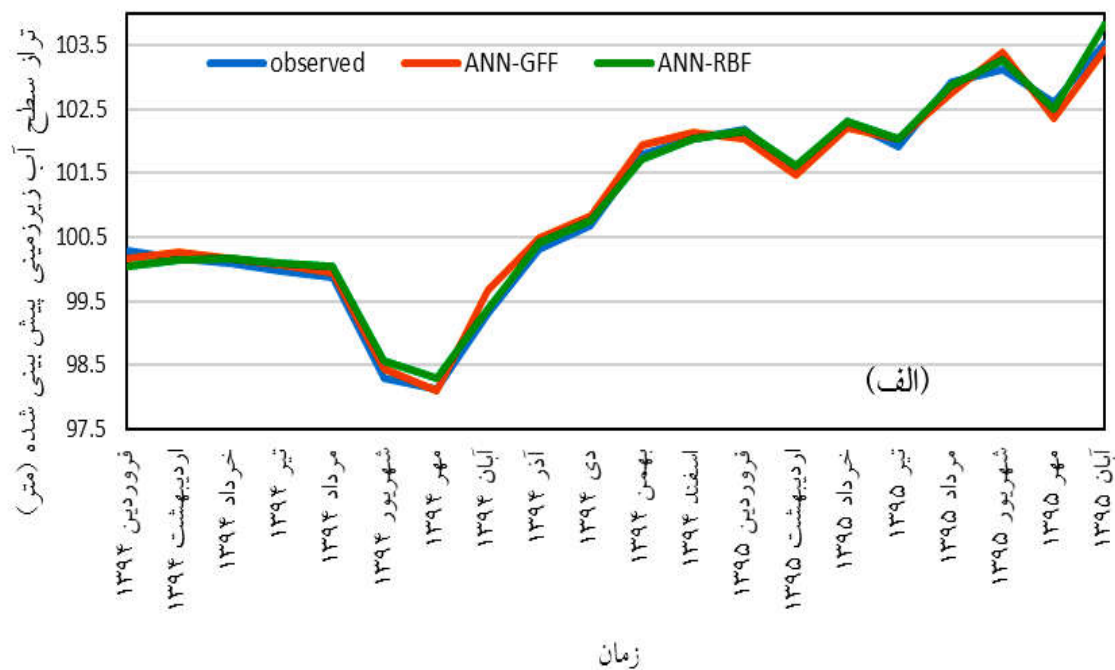


شکل ۴. مقایسه تراز سطح آب زیرزمینی پیش‌بینی شده با شبکه عصبی GFF و RBF برای پیژومترهای: الف) شماره ۴ و ب)

شماره ۷ شامل پیژومترهای برداشت و تغذیه آب (رنگی در نسخه الکترونیکی)

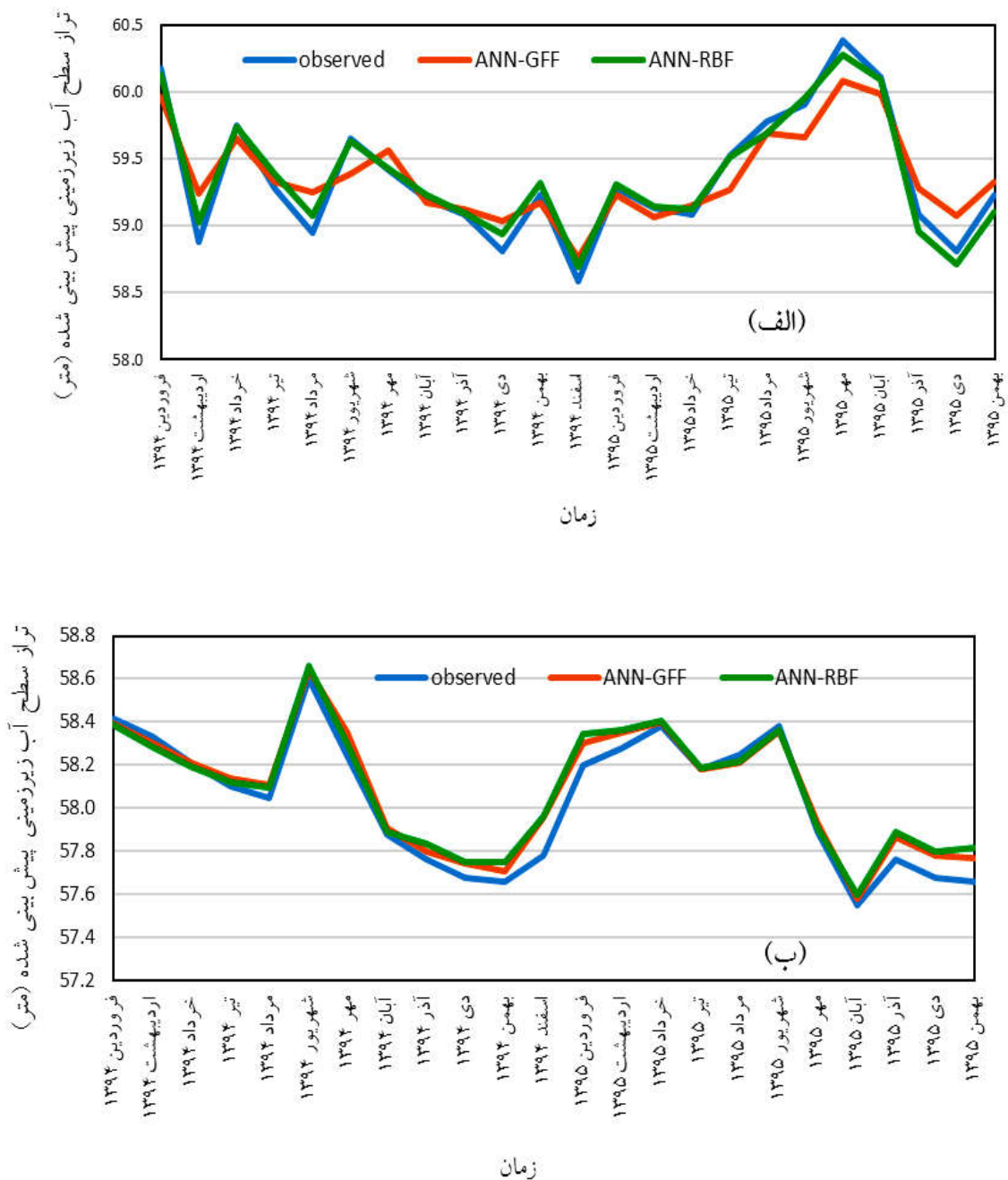
پیژومترهای فقط برداشت آب را نشان می‌دهند. شکل ۶ مقایسه تراز سطح آب زیرزمینی پیش‌بینی شده با شبکه عصبی GFF و RBF برای پیژومترهای شماره ۳۴ و ۴۱ شامل پیژومترهای نه

پیژومترهای برداشت و تغذیه آب را نشان می‌دهند. شکل ۵ مقایسه تراز سطح آب زیرزمینی پیش‌بینی شده با شبکه عصبی GFF و RBF برای پیژومترهای شماره ۱۷ و ۲۴ شامل



شکل ۵. مقایسه تراز سطح آب زیرزمینی پیش‌بینی شده با شبکه عصبی ANN-GFF و ANN-RBF برای پیزومترهای: الف) شماره ۱۷ و ب) شماره ۲۴ شامل پیزومترهای فقط برداشت آب (رنگی در نسخه الکترونیکی)





شکل ۶. مقایسه تراز سطح آب زیرزمینی پیش‌بینی شده با شبکه عصبی GFF و RBF برای چاهک‌های (الف) شماره ۳۴ و (ب) شماره ۴۱ شامل چاهک‌های نه برداشت و نه تغذیه آب (رنگی در نسخه الکترونیکی)

نزدیک به تغییرات تراز سطح آب مشاهداتی نسبت به شبکه RBF است. همچنین با توجه به شکل ۶، تغییرات تراز سطح آب پیش‌بینی شده با شبکه عصبی RBF، در پیزومترهای شماره ۳۴ و ۴۱ نسبت به شبکه GFF نزدیک به تغییرات تراز سطح

برداشت و نه تغذیه آب را نشان می‌دهند. پیزومترهای شماره ۴، ۷، ۱۷، ۲۴، ۳۴ و ۴۱ به‌عنوان نمونه در نظر گرفته شده‌اند. با توجه به شکل ۴ و ۵، تغییرات تراز سطح آب پیش‌بینی شده با شبکه عصبی GFF در پیزومترهای شماره ۴، ۷، ۱۷ و ۲۴

آب مشاهداتی است.

تأثیرات عوامل مختلف شامل تغذیه، برداشت و بارش است، بنابراین تراز سطح آب زیرزمینی در گام‌های زمانی قبل به‌عنوان متغیر مؤثر ورودی در شبکه عصبی مصنوعی برای پیش‌بینی تراز سطح آب زیرزمینی در ماه آینده است. مقایسه تراز سطح آب زیرزمینی پیش‌بینی شده با شبکه عصبی GFF و RBF نشان می‌دهد که شبکه عصبی مصنوعی GFF از دقت بیشتری نسبت به شبکه RBF در پیش‌بینی ماهانه تراز سطح آب زیرزمینی برای پیرومترهای شامل برداشت و تغذیه آب و پیرومترهای شامل فقط برداشت آب برخوردار است. علاوه بر این شبکه عصبی RBF از دقت بیشتری نسبت به شبکه GFF، در پیش‌بینی ماهانه تراز سطح آب زیرزمینی برای پیرومترهای شامل نه برداشت و نه تغذیه آب برخوردار هستند. بنابراین نتایج این تحقیق نشان می‌دهد که به‌کارگیری الگوریتم PMI برای تعیین متغیرهای ورودی مؤثر در شبکه‌های عصبی مصنوعی موجب کاهش قابل توجه در زمان مورد نیاز برای انتخاب متغیرهای ورودی مؤثر و در نتیجه توسعه مدل می‌شود.

## نتیجه‌گیری

در این تحقیق به ارزیابی شبکه‌های عصبی مصنوعی برای پیش‌بینی ماهانه تراز سطح آب زیرزمینی پرداخته شده است. با توجه به موقعیت چاهک‌ها نسبت به رودخانه و همچنین برداشت یا عدم برداشت آب، چاهک‌ها به سه طبقه به‌ترتیب شامل چاهک‌های تحت تأثیر برداشت و تغذیه آب، پیرومترهای تحت تأثیر فقط برداشت آب و چاهک‌های نه تحت تأثیر برداشت و نه تغذیه آب تقسیم‌بندی شدند. برای انتخاب متغیرهای مؤثر ورودی در شبکه‌های عصبی مصنوعی از الگوریتم PMI استفاده شد. نتایج به‌کارگیری الگوریتم PMI نشان می‌دهد که با توجه به ضابطه همپل، متغیرهای ورودی مؤثر بر پیش‌بینی تراز سطح آب در یک ماه آینده برای چاهک‌های مختلف شامل برداشت و تغذیه آب، فقط برداشت آب و نه تحت تأثیر برداشت و نه تغذیه آب، تراز سطح آب زیرزمینی در گام‌های زمانی قبل است. با توجه به اینکه تراز سطح آب زیرزمینی در نتیجه

## منابع مورد استفاده

1. Antonopoulos, V. Z., P. E. Georgiou and Z. V. Antonopoulos. 2015. Dispersion coefficient prediction using empirical models and ANNs. *Environmental Processes* 2(2): 379-394.
2. Bidwell, V. J. 2005. Realistic forecasting of groundwater level, based on the eigenstructure of aquifer dynamics. *Mathematics and Computers in Simulation* 69(1-2): 12-20.
3. Cover, T. M. and J. A. Thomas. 1991. Elements of Information Theory. In: Wiley Series in Telecommunications. John Wiley & Sons, Inc., New York.
4. Daliakopoulos, I. N., P. Coulibaly and I. K. Tsanis. 2005. Groundwater level forecasting using artificial neural networks. *Journal of Hydrology* 309(1-4): 229-240.
5. David, F. N. 1966. Tables of the correlation coefficient. In: Pearson ES, Hartley HO (Eds.) *Biometrika tables for statisticians*, third ed., vol. 1. Cambridge University Press, Cambridge.
6. Davies, L. and U. Gather. 1993. The identification of multiple outliers. *Journal of the American Statistical Association* 88(423): 782-792.
7. Emamgholizadeh, S., K. Moslemi and G. Karami. 2014. Prediction the groundwater level of bastam plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS). *Water Resources Management* 28(15): 5433-5446.
8. Fang, W., S. Huang, Q. Huang, G. Huang, E. Meng and J. Luan. 2018. Reference evapotranspiration forecasting based on local meteorological and global climate information screened by partial mutual information. *Journal of Hydrology* 561:764-779.
9. Goebel, B., Z. Dawy, J. Hagenauer and J. C. Mueller. 2005. An approximation to the distribution of finite sample size mutual information estimates. In: IEEE International Conference on Communications (ICC-05), Seoul, South Korea.
10. Granger, C. W., E. Maasoumi and J. Racine. 2004. A dependence metric for possibly nonlinear processes. *Journal of Time Series Analysis* 25(5): 649-669.

11. Kalra, R., M. C. Deo, R. Kumar and V. K. Agarwal. 2005. RBF network for spatial mapping of wave heights. *Marine Structures* 18(3): 289-300.
12. Kişi, Ö. 2009. Neural networks and wavelet conjunction model for intermittent streamflow forecasting. *Journal of Hydrologic Engineering* 14(8): 773-782.
13. Lee, S., K. K. Lee and H. Yoon. 2019. Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors. *Hydrogeology Journal* 27(2): 567-579.
14. Luo, J., W. Lu, Y. Ji and D. Ye. 2016. A comparison of three prediction models for predicting monthly precipitation in Liaoyuan city, China. *Water Science and Technology: Water Supply* 16(3): 845-854.
15. Maiti, S. and R. K. Tiwari. 2014. A comparative study of artificial neural networks, Bayesian neural networks and adaptive neuro-fuzzy inference system in groundwater level prediction. *Environmental Earth Sciences* 71(7): 3147-3160.
16. Maity, R. and D. Nagesh Kumar. 2008. Probabilistic prediction of hydroclimatic variables with nonparametric quantification of uncertainty. *Journal of Geophysical Research: Atmospheres* 113(D14).
17. May, R. J., G. C. Dandy, H. R. Maier and T. G. Fernando. 2006. Critical values of a kernel density-based mutual information estimator. In *The 2006 IEEE International Joint Conference on Neural Network Proceedings* 4898-4903.
18. May, R. J., H. R. Maier, G. C. Dandy and T. G. Fernando. 2008. Non-linear variable selection for artificial neural networks using partial mutual information. *Environmental Modelling & Software* 23(10-11): 1312-1326.
19. Nash, J. E. and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models; part I: A discussion of principles. *Journal of Hydrology* 10: 282-290.
20. Pearson, R. K. 2002. Outliers in process modeling and identification. *IEEE Transactions on Control Systems Technology* 10(1): 55-63.
21. Raj, A. S., D. H. Oliver, Y. Srinivas and J. Viswanath. 2017. Wavelet based analysis on rainfall and water table depth forecasting using Neural Networks in Kanyakumari district, Tamil Nadu, India. *Groundwater for Sustainable Development* 5: 178-186.
22. Rakhshandehroo, G., H. Akbari, M. Afshari Igder and E. Ostadzadeh. 2018. Long-term groundwater-level forecasting in shallow and deep wells using wavelet neural networks trained by an improved harmony search algorithm. *Journal of Hydrologic Engineering* 23(2): 1094.
23. Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27: 379-423.
24. Sharma, A. 2000. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: part I: A strategy for system predictor identification. *Journal of Hydrology* 239: 232-239.
25. Singh, N. and G. J. Chakrapani. 2015. ANN modelling of sediment concentration in the dynamic glacial environment of Gangotri in Himalaya. *Environmental Monitoring and Assessment* 187(8): 494.
26. Wunsch, A., T. Liesch and S. Broda. 2018. Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX). *Journal of Hydrology* 567: 743-758.
27. Yang, L., X. Zhao, S. Peng and G. Zhou. 2015. Integration of Bayesian analysis for eutrophication prediction and assessment in a landscape lake. *Environmental Monitoring and Assessment* 187(1): 1-19.

## Prediction of Groundwater Level Using Artificial Neural Networks Based on Efficient Input Variables Selection by Partial Mutual Information Algorithm

F. Zarif<sup>1</sup>, A. Asareh<sup>1\*</sup>, M. Asadi Lour<sup>1</sup>, H. Fathian<sup>2</sup>, and D. Khodadadi Dehkordi<sup>1</sup>

(Received: November 24-2020; Accepted: August 24-2021)

### Abstract

An accurate and reliable prediction of groundwater level in a region is very important for sustainable use and management of water resources. In this study, the generalized feedforward (GFF) and radial basis function (RBF) of artificial neural networks (ANNs) have been evaluated for monthly predicting groundwater levels in the Dezful-Andimeshk plain in southwestern Iran. The partial mutual information (PMI) algorithm was used to determine efficient input variables in ANNs. The results of using the PMI algorithm showed that efficient input variables for monthly predicting groundwater level for piezometers affected by water discharge and recharge include only water level in the current month. Also, efficient input variables for predicting the water level for piezometers affected only by water discharge include the water level in the current month, the water level in the previous month, the water level in the previous two months, transverse coordinates of piezometers to UTM, the water level in the previous three months, the water level in the previous four months, the water level in the previous five months and longitudinal coordinates of piezometers to UTM. In addition, efficient input variables of monthly predicting groundwater level for piezometers neither affected by water discharge nor water recharge, respectively, include the water level in the current month, the water level in the previous month, the water level in the previous two months, the water level in the previous three months, the water level in the previous four months, the water level in the previous five months, the water level in the previous six months, transverse coordinates of piezometer to UTM and longitudinal coordinates of piezometer to UTM. The results indicated that the GFF network is more accurate than the RBF network for monthly predicting groundwater level for piezometers including water discharge and recharge and piezometers including only water discharge. Also, the RBF network is more accurate for monthly predicting groundwater levels for piezometers that include neither water discharge nor recharge than the GFF network.

**Keywords:** Monthly predicting, Groundwater level, Artificial Neural Networks, PMI algorithm.

---

1. Department of Water Sciences and Engineering, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran.

2. Department of Water Resources Engineering, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran.

\*: Corresponding author, Email: ali\_assareh\_2003@yahoo.com