

بررسی روش های هوش مصنوعی و رگرسیونی در شبیه سازی مقادیر دمای روزانه

مهدی امیرآبادی زاده^۱، محمد ناظری تهرودی^{۲*}، محمد جواد زینلی^۲

^۱ استادیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران
^۲ دانشجوی دکتری منابع آب، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران

تاریخ دریافت: ۱۳۹۶/۱۲/۱۸، تاریخ پذیرش: ۱۳۹۷/۰۳/۲۳

چکیده

پیش بینی تغییرات آب و هوایی کره زمین با استفاده از مقادیر ثبت شده در دوره آماری حاضر نیازمند روشی دقیق بوده که بتوان نوسانات این تغییرات را به خوبی شناسایی کرده و با الگوگیری از این تغییرات مقادیر پارامتر مورد نظر را برای سال ها و یا دوره های آینده پیش بینی نمود. در این مطالعه شش مدل رگرسیونی چند متغیره، ANN، SVR، ANFIS، SVM و GP جهت ریزمقیاس نمایی مقادیر متوسط دمای روزانه ایستگاه همدیدی ارومیه با استفاده از ۲۶ پارامتر پیش بینی کننده منتج از گزارش پنجم IPCC مورد بررسی و مقایسه قرار گرفت. مقادیر دمای متوسط روزانه ایستگاه مورد بررسی از تاریخ ۱۲ مارس ۱۹۶۱ (۲۹ اسفند ۱۳۸۴) تا تاریخ ۲۰ دسامبر ۲۰۰۵ (۲۹ آذر ۱۳۸۴) انتخاب گردید. در تمامی روش های ذکر شده با استفاده از آزمون پیرسون از بین ۲۶ پارامتر پیش بینی کننده، ۱۶ پارامتر که همبستگی بالایی با مقادیر دمای متوسط روزانه داشته انتخاب گردید. جهت بررسی مقادیر خطای ناشی از مدل سازی از سه معیار ضریب تبیین، مجذور میانگین مربعات خطا و معیار کارایی مدل استفاده شد. نتایج بررسی دقت و میزان خطای مدل ها نشان داد که در بین مدل های هوشمند ANFIS، ANN، GP و SVM، مدل برنامه ریزی ژنتیک کمترین مقدار خطا را داشته و در بین مدل های رگرسیونی (رگرسیون چندمتغیره و رگرسیون بردار پشتیبان) روش رگرسیون بردار پشتیبان، کمترین میزان خطا و بیشترین میزان دقت را در شبیه سازی مقادیر دمای روزانه ایستگاه همدیدی ارومیه داشته است. به طور کلی نتایج شبیه سازی مقادیر دما روزانه نشان دهنده دقیق تر بودن روش های رگرسیونی نسبت به روش های هوشمند می باشد. از آنجا که این مطالعه تنها با استفاده از داده های ایستگاه همدیدی ارومیه انجام گرفته است، لذا نتایج حاصله تنها برای ایستگاه مذکور اعتبار داشته و با اطمینان نمی توان نتایج را برای تمامی ایستگاه ها تعمیم داد.

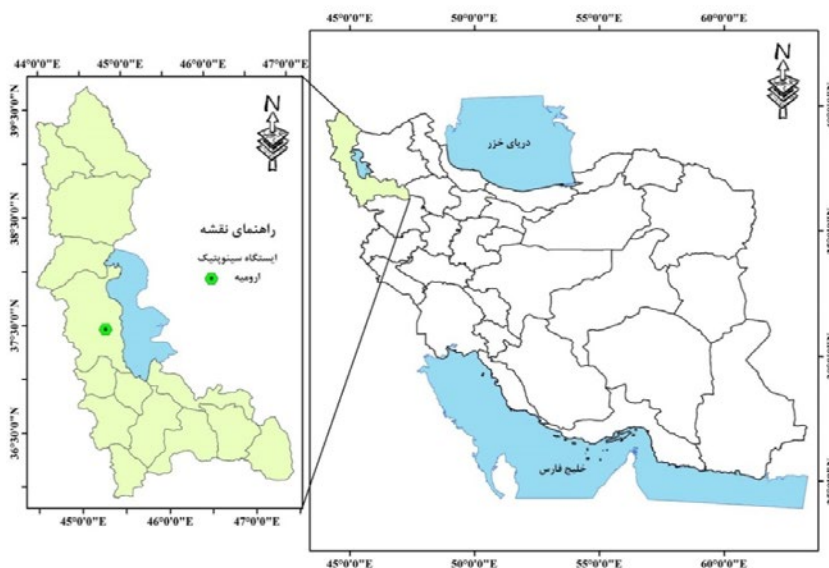
کلمات کلیدی: الگوریتم ژنتیک، دما، دریاچه ارومیه، شبکه عصبی، ماشین بردار پشتیبان، ریزمقیاس نمایی.

از روش ماشین بردار پشتیبان جهت بررسی مقدار بارش استفاده کردند. کامپوزانو و همکاران (۲۰۱۶) مدل های مختلف ریزمقیاس نمایی نظیر ANN، SVM و SDSM را در ریزمقیاس نمایی بارش ماهانه حوضه رودخانه پوته در جنوب اکوادور مورد بررسی و مقایسه قرار دادند. نتایج نشان داد که مدل های LS-SVM و ANN دقت بالاتری نسبت به مدل ریزمقیاس نمایی SDSM دارند. جهت ریزمقیاس نمایی مقادیر بارش و دمای روزانه روش های مختلفی در تحقیقات مختلف مورد بحث قرار گرفته است. اما آن چه مهم است این است که با توجه به ماهیت داده های دمای روزانه، چه مدلی می تواند به خوبی این داده ها را برازش دهد. لذا هدف از این تحقیق بررسی مدل های هوشمند و رگرسیونی موجود جهت بررسی دقت آن ها در ریزمقیاس نمایی مقادیر دمای روزانه می باشد. در این تحقیق شش مدل GP، ANFIS، ANN، SVM، SVR و رگرسیون چندمتغیره جهت ریزمقیاس نمایی و شبیه سازی مقادیر دمای

متوسط روزانه ایستگاه همدیدی ارومیه با استفاده از ۲۶ پارامتر پیش بینی کننده سناریوهای آب و هوایی گزارش پنجم هیئت بین دول تغییر اقلیم مورد بررسی قرار گرفته و دقت هر مدل برآورد گردید.

روش تحقیق

ایستگاه همدیدی ارومیه در سال ۱۳۳۹ تأسیس گردیده و دارای ارتفاع ۱۳۲۸ متری از سطح دریا است. این ایستگاه در طول جغرافیایی ۳۷°۴۵' و عرض جغرافیایی ۴۰°۰۰' قرار دارد (ناظری تهرودی و همکاران، ۲۰۱۸). منطقه مورد مطالعه در این تحقیق، ایستگاه هواشناسی همدیدی ارومیه و سری زمانی مورد مطالعه، داده های میانگین روزانه دما در ایستگاه همدیدی ارومیه در دوره آماری ۲۰۰۵-۱۹۶۱ (۱۳۸۴-۱۳۴۰) است. شکل ۱ منطقه مورد مطالعه و موقعیت ایستگاه مورد مطالعه را نشان می دهد. مشخصات آماری داده های مورد استفاده نیز به شرح جدول ۱ نشان داده شده است.



شکل ۱: موقعیت ایستگاه مورد مطالعه در سطح کشور و استان آذربایجان غربی

جدول ۱: مشخصات آماری سری های زمانی مورد مطالعه در مقیاس سالانه

| پارامتر | واحد | ضریب چولگی | انحراف معیار | میانگین دما | کمینه | بیشینه |
|-------------|------|------------|--------------|-------------|-------|--------|
| دمای روزانه | (°C) | -۰/۰۲ | ۹/۴ | ۱۱/۴ | -۱۵/۳ | ۲۹/۵ |

حال، شبکه های عصبی قابلیت استفاده از دانش بشری را ندارد و نمی تواند مانند سامانه های فازی با استفاده از عبارات زبانی استنتاج کند (کاراکاس ۲۰۰۶). بنابراین، برای دست یابی به قابلیت یادگیری بهتر، دقت تقریب بیشتر و ساختار ساده تر، با الهام از قابلیت یادگیری شبکه های عصبی و خواص استنباط سامانه های فازی در مدل فازی، شبکه های فازی - عصبی تطبیقی موسوم به انفیس در سال ۱۹۹۹ ارائه شد (مندانی و عصیلان ۱۹۷۵، جانگ ۱۹۹۲).

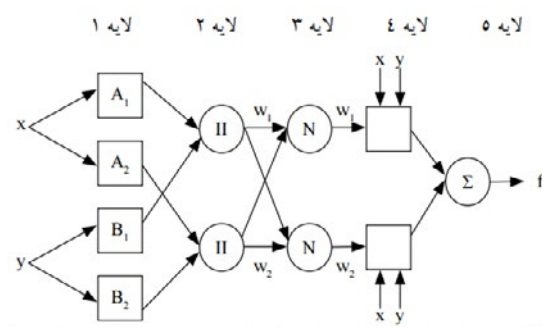
برنامه ریزی بیان ژن

روش برنامه ریزی بیان ژن، در سال ۱۹۹۹ توسط فی پرا ارایه شد (فی پرا، ۲۰۰۱). این روش ترکیبی از روش های برنامه ریزی ژنتیک و الگوریتم ژنتیک بوده که در آن، کروموزوم های خطی و ساده با طول ثابت، مشابه با آنچه که در الگوریتم ژنتیک استفاده می شود و ساختارهای شاخه ای با اندازه ها و اشکال متفاوت، مشابه با درختان تجزیه در برنامه ریزی ژنتیک، ترکیب می شوند. از آنجا که در این روش تمام ساختارهای شاخه ای با اندازه و اشکال متفاوت، در کروموزوم های خطی با طول ثابت کدگذاری می شوند، سبب گردیده که در این روش فنوتیپ و ژنوتیپ از هم جدا شوند و سیستم بتواند از تمام مزایای تکاملی به سبب وجود آن ها بهره مند شود.

اکنون باوجود اینکه فنوتیپ در GEP، همان نوع از ساختارهای شاخه ای مورد استفاده در GP را شامل می شود، اما ساختارهای شاخه ای که به وسیله GEP استنتاج می شوند مبین تمامی ژنوم های مستقل هستند. به طور خلاصه می توان گفت در GEP بهسازی ها در یک ساختار خطی اتفاق افتاده و سپس به صورت ساختار درختی بیان می شود و این موجب می شود تنها ژنوم اصلاح شده به نسل بعد منتقل شده و نیازی به ساختارهای سنگین برای تکثیر و جهش وجود نداشته باشد (فی پرا، ۲۰۰۱). در این روش پدیده های مختلف با استفاده از مجموعه ای از توابع و مجموعه ای از ترمینال ها، مدل سازی می شوند. مجموعه توابع، معمولاً

روش سیستم استنتاج فازی- عصبی تطبیقی (ANFIS) روش سیستم استنتاج فازی- عصبی تطبیقی روشی ترکیبی است که در آن بخش فازی رابطه بین متغیرهای ورودی و خروجی را برقرار می نماید و مشخصه های مربوط به توابع عضویت بخش فازی نیز به وسیله شبکه عصبی تعیین می شود. در سیستم استنتاج فازی- عصبی تطبیقی، ابتدا ساختار مدل با پارامترهای مشخص متناسب با ورودی ها، درجه عضویت و قوانین و توابع درجه عضویت خروجی است، انتخاب می شود. سپس بخشی از داده های موجود به عنوان ورودی- خروجی که قابل استفاده برای آموزش این سیستم باشد، انتخاب می شود. در مرحله آموزش با اصلاح مشخصه های درجه عضویت بر اساس میزان خطای قابل قبول، مشخصه های مدل به مقادیر واقعی نزدیک می شوند. سیستم استنتاج فازی- عصبی تطبیقی بیشتر با استفاده از سیستم فازی سوگنو و به صورت ساختار شبکه ای پیشرونده استفاده می شود که دارای ۵ لایه مطابق شکل ۲ به شرح زیر می باشد (کاراکاس ۲۰۰۶).

سامانه فازی، سامانه ای منطبق بر قوانین فازی است. هدف از منطق فازی، احتمال کلاسیک تحلیل های نظریه استخراج نتایج دقیق با استفاده از مجموعه ای از قوانین است که توسط افراد خبره و متخصص تعریف شده است. از طرفی، شبکه های عصبی قابلیت آموزش و یادگیری دارند و می توانند با استفاده از داده های مشاهده شده، پارامترهای شبکه را به نحوی معین کنند که به ازای ورودی دلخواه، خروجی مطلوب حاصل شود. در عین



شکل ۲: طرحواره یک شبکه ANFIS نوعی

در روابط فوق C عددی صحیح مثبت است که عامل تعیین جرمه در هنگام رخ دادن خطای آموزش مدل است. ϕ تابع کرنل^۷، N تعداد نمونه‌ها و دو مشخصه ξ_i^* و ξ_i متغیر کمبود^۸ هستند که حد بالا و پایین خطای آموزش مرتبط با مقدار خطای مجاز ε را مشخص می‌کنند. در مسائل پیش بینی می‌شود که داده‌ها، درون بازه مرزی ε قرار گیرند. حال اگر داده‌ای خارج از بازه ε قرار گرفت آنگاه یک خطا معادل ξ_i و ξ_i^* وجود خواهد داشت. ذکر این نکته لازم است که مدل SVM مشکلات ناشی از کم تخمینی^۹ و فوق برازشی^{۱۰} را با کمینه کردن هم‌زمان دو ترم $W^T W / 2$ و خطای آموزشی، یعنی $C \sum_{i=1}^N (\xi_i + \xi_i^*)$ را در رابطه ۵ حل می‌کند. بنابراین با معرفی ۲ ضریب لاگرانژ a_i و a_i^* مسئله بهینه سازی با حداکثر سازی عددی تابع درجه دوم زیر (معادله ۵) با شرایط معادله ۶ حل خواهد شد (اسکندری و همکاران ۱۳۹۱).

$$\sum_{i=1}^N y_i (a_i + a_i^*) - \varepsilon \sum_{i=1}^N (a_i + a_i^*) - \quad (5)$$

$$0.5 \sum_{i,j=1}^N (a_i + a_i^*)(a_j + a_j^*) \phi(x_i)^T \phi(x_j)$$

$$\sum_{i=1}^N (a_i + a_i^*) = 0 \quad (6)$$

$$0 \leq a_i \leq C, \quad 0 \leq a_i^* \leq C, \quad i = 1, 2, \dots, N$$

تابع هدف بالا در رابطه ۵ تابع محدب است و بنابراین جواب رابطه ۵ یکتا و بهینه خواهد بود. پس از تعریف ضرایب لاگرانژ در رابطه ۷ مشخصه‌های w و b در مدل SVM رگرسیونی با استفاده از تئوری کرش -کوهن- تاکر محاسبه می‌شود (فلتچر ۱۹۸۷) که در آن $w = \sum_{j=1}^N (a_j + a_j^*) \phi(x_j)$ است. در نتیجه برای مدل SVM رگرسیونی خواهیم داشت (ناظری تهرودی و همکاران، ۲۰۱۷):

$$W = \sum_{i=1}^N (a_i + a_i^*) \phi(x_i)^T \phi(x) + b \quad (7)$$

باید توجه داشت که ترم‌های لاگرانژ $(a_i + a_i^*)$ می‌تواند صفر، و یا غیر صفر باشد. بنابراین فقط

7- Kernel Function
8- Slack Variable
9- Under Fitting
10- Over Fitting

شامل توابع اصلی حسابی $\{+, -, \times, /\}$ ، توابع مثلثاتی یا هر نوع تابع ریاضی دیگر $\{\sqrt{\quad}, x^2, \exp, \log, \sin, \cos, \dots\}$ و یا توابع تعریف شده توسط کاربر است که برای تفسیر مدل مناسب می‌باشند. مجموعه ترمینال‌ها، از مقادیر ثابت و متغیرهای مستقل تشکیل شده اند (فی‌یرا، ۲۰۰۱).

ماشین بردار پشتیبان

اولین کاربرد این روش در مسائل آب توسط دیباک و همکاران در سال ۲۰۰۱ با مدل سازی بارندگی- رواناب ارائه شد (هافمن و همکاران ۲۰۰۲). ماشین بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه سازی مقید است که از اصل استقرای کمینه سازی خطای ساختاری استفاده کرده و منجر به یک جواب بهینه کلی می‌گردد. در مدل رگرسیون SVM تابعی مرتبط با متغیر وابسته Y که خود تابعی از چند متغیر مستقل x است، برآورد می‌شود. مشابه سایر مسائل رگرسیونی فرض می‌شود رابطه میان متغیرهای مستقل و وابسته با تابع جبری مانند $f(x)$ بعلاوه مقداری اغتشاش (خطای مجاز ε) مشخص شود (رابطه ۲) (اسکندری و همکاران ۱۳۹۱).

$$f(x) = W^T \cdot \phi(x) + b \quad (1)$$

$$y = f(x) + noise \quad (2)$$

چنانچه W (بردار ضرایب) و b (ثابت) مشخصه‌های تابع رگرسیونی و ϕ نیز تابع کرنل باشد، آنگاه هدف پیدا کردن فرم تابعی برای $f(x)$ است. این مهم با آموزش مدل SVM توسط مجموعه‌ای از نمونه‌ها (مجموعه آموزش) محقق می‌شود. بنابراین برای محاسبه w و b لازم است تابع خطا (رابطه ۳) در مدل SVM- ε با در نظر گرفتن شرایط مندرج در رابطه ۴ بهینه شود.

$$\frac{1}{2} W^T W + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^* \quad (3)$$

$$W^T \cdot \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \quad (4)$$

$$y_i - W^T \cdot \phi(x_i) - b \leq \varepsilon + \xi_i$$

$$\xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, N$$

یک نرون کوچک‌ترین واحد پردازشگر اطلاعات است، که اساس عملکرد شبکه عصبی را تشکیل می‌دهد. هر شبکه از یک لایه ورودی، یک لایه خروجی و یک یا چند لایه میانی تشکیل شده است. هر لایه توسط وزن‌ها به نرون‌های لایه بعد متصل می‌شوند.

هر نرون مقادیر ورودی را در وزن خود ضرب کرده و با بایاس جمع می‌کند و حاصل به تابع محرکه انتقال داده می‌شود تا به این وسیله به نرون‌های لایه بعد انتقال داده شود. این عمل در نرون‌های همه لایه‌های شبکه انجام شده تا در نهایت خروجی شبکه به دست آید (زینلی و هاشمی، ۱۳۹۵).

شبکه عصبی مصنوعی پرسپترون

۹۰ درصد شبکه‌های عصبی مصنوعی که در مسائل مربوط به آب استفاده می‌شود از نوع الگوریتم پس انتشار خطا هستند (براداک و همکاران، ۱۹۹۸). شبکه عصبی پرسپترون چند لایه، شبکه‌ای با الگوریتم پس انتشار خطا است که در تولباکس آن به صورت پیش فرض، تابع انتقال در لایه پنهان از تابع محرکه تانژانت سیگموئید (tansig) و در لایه خروجی از تابع محرکه خطی (purelin) استفاده می‌نماید و تابع یادگیری وزن‌ها تابع (learngdm) و تابع آموزش آن تابع لونبرگ مارکوات (trainlm) می‌باشد. در این شبکه‌ها اطلاعات ورودی توسط نرون‌های ورودی دریافت شده، سپس به نرون‌های لایه پنهان فرستاده می‌شود و پردازش مورد نظر روی آن‌ها انجام می‌گیرد و نتیجه به نرون‌های لایه خروجی فرستاده شده تا به عنوان خروجی شبکه نمایش داده شوند (زینلی و همکاران، ۱۳۹۲). در این تحقیق شبکه دارای یک لایه پنهان بوده که در آن ۵ نرون قرار داده شده است.

در این تحقیق از ۶۰ درصد داده‌های موجود به آموزش، ۲۰ درصد به صحت سنجی و ۲۰ درصد باقیمانده نیز به تست شبکه اختصاص داده شده است.

معیارهای ارزیابی

به منظور ارزیابی دقت و کارایی مدل‌های مورد بررسی از ضریب تبیین، جذر میانگین مربعات خطا و

مجموعه داده‌هایی که ضرایب \bar{a}_i آن‌ها غیر صفر است در معادله رگرسیون نهایی وارد می‌شوند و این مجموعه داده‌ها به عنوان بردارهای پشتیبان شناخته می‌شوند. به طور ساده، بردارهای پشتیبان آن داده‌هایی هستند که به ساخته شدن تابع رگرسیونی کمک می‌کنند. در میان بردارهای مذکور آن‌هایی که مقدار $|a_i|$ آن‌ها کمتر از C باشد بردارهای پشتیبان حاشیه‌ای^{۱۱} نامیده می‌شوند. هنگامی که مقدار $|a_i|$ بردارهای پشتیبان برابر مقدار C باشد، به عنوان بردار پشتیبان خطا^{۱۲} یا بردار پشتیبان کراندار شناخته می‌شود. بردارهای پشتیبان حاشیه‌ای در حاشیه مرز غیر حساس یافت می‌شوند، در حالی که بردارهای پشتیبان خطا خارج از بازه هستند. در نهایت تابع SVM رگرسیونی را می‌توان به فرم زیر بازنویسی کرد:

$$f(x) = \sum_{i=1}^N a_i \phi(x_i)^T \phi(x_j) + b \quad (8)$$

در رابطه ۸ محاسبه $\phi(x)$ در فضای مشخصه آن ممکن است بسیار پیچیده باشد. برای حل این مشکل روند معمول در مدل SVM رگرسیونی انتخاب یک تابع کرنل به صورت $K(x_i, x)$ برابر با $\phi(x_i)^T \phi(x_j) = \sqrt{b^2 - 4ac}$ است. می‌توان از توابع مختلف کرنل برای ساخت انواع مختلف مدل ϵ -SVM استفاده کرد. انواع رایج توابع کرنل قابل استفاده در مدل SVM رگرسیونی عبارت‌اند از: کرنل چندجمله‌ای با ۳ مشخصه هدف، کرنل سیگموئیدی شامل ۲ مشخصه هدف و کرنل توابع پایه شعاعی (RBF) با یک مشخصه هدف (اسکندری و همکاران ۱۳۹۱).

شبکه عصبی مصنوعی

شبکه عصبی مصنوعی ایده‌ای است برای پردازش اطلاعات که از سیستم عصبی زیستی الهام گرفته شده است و مانند مغز انسان به پردازش اطلاعات می‌پردازد این سیستم از عناصر زیادی به نام نرون تشکیل شده است که برای حل یک مسئله با هم هماهنگ عمل می‌کنند.

11- Margin Support Vector
12- Error Support Vector

جدول ۲: نتایج بررسی همبستگی بین مقادیر پیش‌بینی کننده‌ها و دمای روزانه ایستگاه همدیدی ارومیه

| پیش‌بینی کننده | عنوان پیش‌بینی کننده‌ها | همبستگی | پیش‌بینی کننده | عنوان پیش‌بینی کننده‌ها | همبستگی |
|----------------|---|---------|----------------|--|---------|
| MSLPGL | متوسط فشار سطح دریا | -0.684 | p8-fgl | سرعت باد در فشار ۸۰۰ هکتوپاسکال | -0.133 |
| p1_FGL | سرعت باد در فشار ۱۰۰۰ هکتوپاسکال | -0.118 | p8-ugl | مولفه سرعت مداری در فشار ۸۰۰ هکتوپاسکال | -0.415 |
| p1_Ugl | مولفه سرعت مداری در فشار ۱۰۰۰ هکتوپاسکال | -0.421 | p8-vgl | مولفه سرعت النهای در فشار ۸۰۰ هکتوپاسکال | -0.374 |
| p1_Vgl | مولفه سرعت نصف النهای در فشار ۱۰۰۰ هکتوپاسکال | -0.374 | p8-zgl | تاوایی در فشار ۸۰۰ هکتوپاسکال | 0.432 |
| p1_zgl | تاوایی در فشار ۱۰۰۰ هکتوپاسکال | 0.081 | p8-thgl | جهت باد در فشار ۸۰۰ هکتوپاسکال | -0.375 |
| p1thgl | جهت باد در فشار ۱۰۰۰ هکتوپاسکال | -0.367 | p8-zhgl | واگرایی در فشار ۸۰۰ هکتوپاسکال | -0.364 |
| p1zhgl | واگرایی در فشار ۱۰۰۰ هکتوپاسکال | -0.062 | p500gl | رطوبت نسبی در فشار ۵۰۰ هکتوپاسکال | 0.795 |
| p5-fgl | سرعت باد در فشار ۵۰۰ هکتوپاسکال | -0.171 | p850gl | رطوبت نسبی در فشار ۸۵۰ هکتوپاسکال | -0.285 |
| p5-ugl | مولفه سرعت مداری در فشار ۵۰۰ هکتوپاسکال | -0.205 | prcpgl | بارش کل | -0.135 |
| p5-vgl | مولفه سرعت نصف النهای در فشار ۵۰۰ هکتوپاسکال | 0.340 | s500gl | رطوبت ویژه در فشار ۵۰۰ هکتوپاسکال | 0.468 |
| p5-zgl | تاوایی در فشار ۵۰۰ هکتوپاسکال | -0.233 | s850gl | رطوبت ویژه در فشار ۸۵۰ هکتوپاسکال | 0.579 |
| 5-thgl | جهت باد در فشار ۵۰۰ هکتوپاسکال | -0.210 | shumgl | رطوبت ویژه در نزدیک سطح | 0.578 |
| 5-zhgl | واگرایی در فشار ۵۰۰ هکتوپاسکال | 0.323 | tempgl | متوسط دما در ارتفاع ۲ متری | 0.862 |

جدول ۳: نتایج بررسی دقت مدل‌های ANN، ANFIS، SVM و GP در برآورد مقادیر دمای متوسط روزانه

| GP | | SVM | | ANFIS | | ANN | | Test |
|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Train | Test | Train | Test | Train | Test | Train | Test | RMSE |
| 4.116 | 4.262 | 4.859 | 5.256 | 4.239 | 4.494 | 4.876 | 4.716 | RMSE |
| 0.9 | 0.883 | 0.898 | 0.88 | 0.894 | 0.87 | 1.00 | 1.00 | r |
| 0.814 | 0.774 | 0.737 | 0.939 | 0.799 | 0.96 | 0.737 | 0.773 | N-S |

*: واحد RMSE بر حسب درجه سلسیوس می‌باشد.

جدول ۴: نتایج بررسی دقت مدل‌های SVR و رگرسیون چندمتغیره در برآورد مقادیر دمای متوسط روزانه

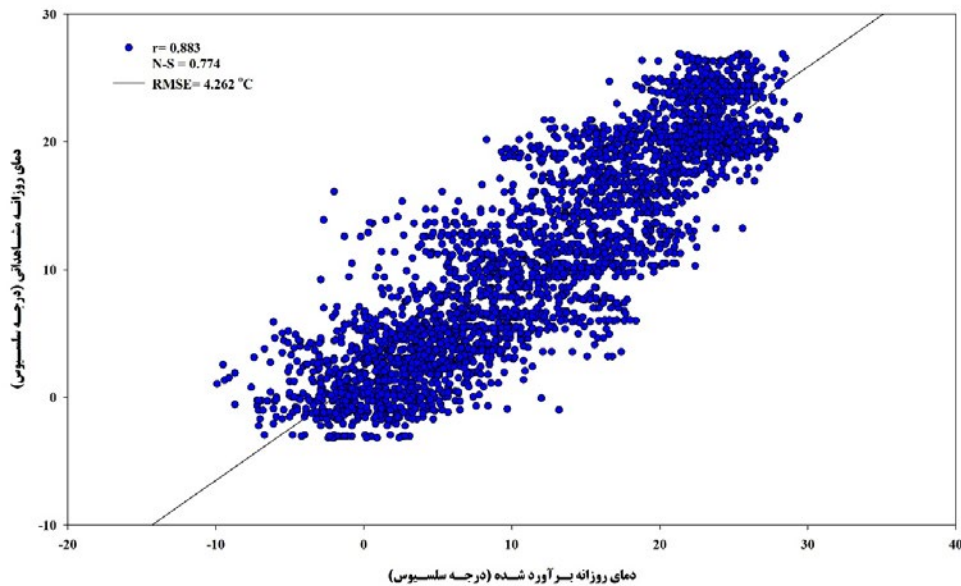
| SVR | Reg | Test |
|-------|-------|------|
| 1.105 | 4.069 | RMSE |
| 0.992 | 0.901 | r |
| 0.984 | 0.812 | N-S |

آزمایش بین ۴ تا ۶ درجه سلسیوس متغیر می‌باشد. از بین چهار مدل مذکور مدل GP با کمترین میزان خطا در مرحله آزمایش به عنوان بهترین مدل شناخته شد. در مرحله آموزش نیز دقت مدل GP از سایر مدل‌های مورد بررسی بهتر بود. بر اساس آماره همبستگی، مدل ANN در هر دو مرحله آزمایش و آموزش بهترین دقت را در شبیه‌سازی مقادیر دمای روزانه ارائه نمود. بعد از مدل ANN، مدل GP بر اساس ضریب همبستگی بهترین جواب‌ها را ارائه نمود. آزمون نش- ساتکلیف برخلاف نتایج دو آزمون همبستگی و جذر میانگین مربعات خطا، مدل ANFIS با به عنوان بهترین مدل در مرحله آزمایش معرفی کرد. با این حال در مرحله آزمایش همچنان مدل GP بر اساس آماره نش- ساتکلیف بهترین کارایی را

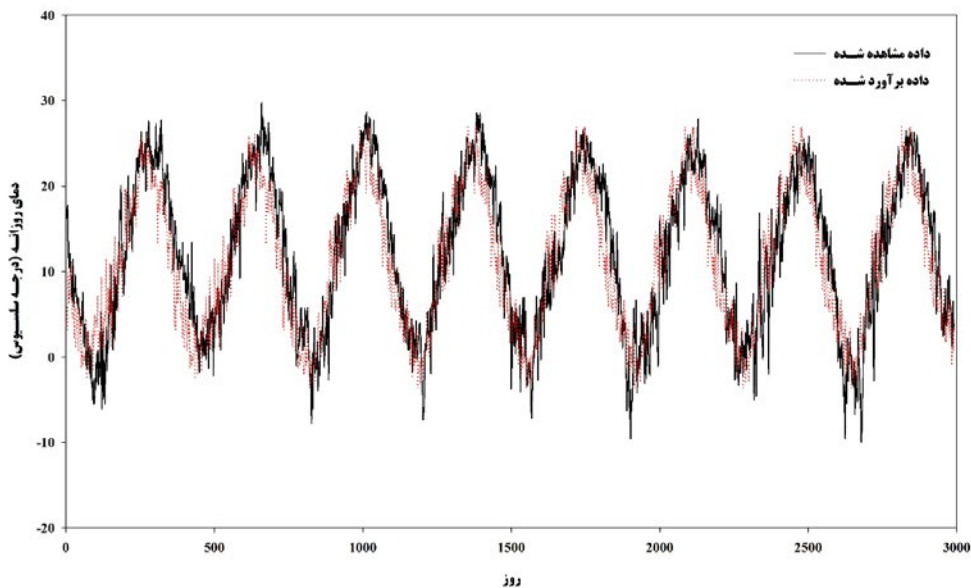
که دقت و میزان خطای این دو مدل نسبت به یکدیگر سنجیده می‌شود. دقت و میزان خطای مدل‌های هوشمند ANN، ANFIS، SVM و GP نیز نسبت به هم تحلیل می‌شود. نتایج بررسی دقت و میزان خطای مدل‌های هوشمند در دو مرحله آزمایش و آموزش مطابق جدول ۳ محاسبه گردید. همان‌طور که از جدول ۳ مشاهده می‌شود، میزان خطای هر چهار مدل مورد بررسی در مرحله آموزش بین ۴ تا ۵ درجه سلسیوس و در مرحله

نتایج نشان داد که از بین چهار مدل هوشمند SVM، ANFIS، ANN و GP، مدل برنامه ریزی ژنتیک بهترین مدل جهت ریزمقیاس نمایی مقادیر متوسط دمای روزانه ایستگاه همدیدی ارومیه می باشد و مدل های SVM و ANN بیشترین میزان خطا را در میان هر شش مدل مورد بررسی کسب کردند. از میان مدل های رگرسیونی نیز

لازم به ذکر است که جهت برآورد و تخمین پارامترهای مدل SVR از روش بهینه یابی الگوریتم جامع مورچگان پیوسته استفاده شد و سه پارامتر مدل SVR غیرخطی بهینه گردید. یکی از عواملی که باعث بهبود نتایج مدل SVR نسبت به سایر مدل های مورد بررسی می باشد، بهینه کردن پارامترهای آن است. به طور کلی



شکل ۵: نتایج بررسی همبستگی مقادیر مشاهداتی و محاسباتی دمای روزانه ایستگاه همدیدی ارومیه با استفاده از مدل GP



شکل ۶: نتایج شبیه سازی مقادیر دمای روزانه ایستگاه همدیدی ارومیه با استفاده از مدل

می باشد. با مقایسه هر شش مدل مورد بررسی نتایج نشان داد که بعد از مدل SVR، مدل رگرسیون چندمتغیره کمترین میزان خطا را در شبیه سازی متوسط دمای روزانه دارد. به طور کلی نتایج نشان داد که مدل های رگرسیونی از دقت بالاتری در شبیه سازی مقادیر متوسط دمای روزانه ایستگاه همدیدی ارومیه نسبت به مدل های هوشمند برخوردار می باشند.

منابع

- اسکندری، ع، نوری، ر، معراجی، ح. و کیاقدی، ا. ۱۳۹۱. توسعه مدلی مناسب بر مبنای شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای پیش بینی بهنگام اکسیژن خواهی بیوشیمیایی ۵ روزه، محیط شناسی، سال سی و هشتم، شماره ۶۱، ص ۸۲-۷۱.
- رضایی، م، نهانی، م، آبکار، ع، رضایی، م. و میرکازهی ریگی، م. ۱۳۹۳. بررسی کارایی مدل ریزمقیاس نمایی آماری (SDSM) در پیش بینی پارامترهای دمایی در دو اقلیم خشک و فراخشک (مطالعه موردی: کرمان و بم)، پژوهشنامه مدیریت حوزه آبخیز، ۵(۱۰)، ۱۳۱-۱۱۷.
- زینلی، م. ج، نیکبخت، ص. و محمدرضاپور، ا. ۱۳۹۲. پیش بینی جریان ورودی به سد ملاصدرا با استفاده از شبکه عصبی مصنوعی، پنجمین کنفرانس منابع آب ایران. ۲۹-۳۰ بهمن ماه، دانشگاه شهید بهشتی تهران، تهران.
- زینلی، م. ج. و هاشمی س. ر. ۱۳۹۵. مقایسه توابع یادگیری شبکه عصبی در مدل سازی رواناب، اکوهیدرولوژی، ۳(۴)، ۶۵۹-۶۶۷.
- ناظری تهرودی، م، هاشمی، س. ر، احمدی، ف. و ناظری تهرودی. ز. ۱۳۹۵. ارزیابی دقت مدل های ANFIS، SVM و GP در مدل سازی دبی جریان، اکوهیدرولوژی، ۳(۳)، ۳۶۱-۳۴۷.
- Amirabadizadeh, M., Ghazali, A., Wayayok, A. and Huang, Y.F., 2015. Climate Change and its Impacts on Hydrological Regime of The Langat River Basin, Malaysia (Doctor of philosophy). University Putra Malaysia.
- Braddock, R.D., Kremmer, M.L. and Sanzogni, L., 1998. Feed-forward artificial neural network model for forecasting rainfall run-off. *Environmetrics*, 9(4), pp.419-432.
- Campozano, L., Tenelanda, D., Sanchez, E., Samaniego, E. and Feyen, J., 2016. Comparison of statistical downscaling methods for monthly total precipitation: case study for the paute river basin in Southern Ecuador. *Advances in Meteorology*, 2016.
- Ferreira, C., 2001. Gene expression programming

مدل SVR بیشترین میزان دقت را به خود اختصاص دارد. تریپاتی و همکاران (۲۰۰۶) نیز در تحقیقات خود دقت روش رگرسیونی SVR را در ریزمقیاس نمایی داده های اقلیمی تأیید کردند. با توجه به نتایج حاصل شده می توان شش مدل رگرسیون بردار پشتیبان، رگرسیون چند متغیره، برنامه ریزی ژنتیک، سیستم های استنتاج فازی، ماشین بردار پشتیبان و شبکه عصبی مصنوعی را به ترتیب بر اساس میزان دقت و کارایی مدل در جهت برآورد و شبیه سازی مقادیر متوسط دمای روزانه ایستگاه همدیدی ارومیه معرفی نمود و در واقع مدل ANN کمترین میزان دقت و بیشترین میزان خطا را به خود اختصاص داد. کارآموز و همکاران (۲۰۰۹) و سجادخان و همکاران (۲۰۰۴) نیز دقت مدل ANN را در ریزمقیاس نمایی مقادیر بیشینه دمای روزانه ضعیف معرفی کردند.

جمع بندی

در این مطالعه شش روش برنامه ریزی ژنتیک، شبکه عصبی مصنوعی، سیستم استنتاج فازی، ماشین بردار پشتیبان، رگرسیون بردار پشتیبان و رگرسیون چندمتغیره جهت شبیه سازی و ریزمقیاس نمایی متوسط دمای روزانه ایستگاه همدیدی ارومیه در دوره آماری ۲۰۰۵-۱۹۶۱ (۱۳۸۴-۱۳۴۰) مورد بررسی قرار گرفته و دقت و میزان خطای هر مدل برآورد گردید. نتایج بررسی دقت مدل های مذکور بر اساس ضریب کارایی نش - ساتکلیف نشان داد که به ترتیب مدل های SVR، AFIS، SVM، Reg، GP و ANN بیشترین مقادیر آماره نش - ساتکلیف را بدست آورده و رتبه های یک تا شش را به خود اختصاص دادند. مدل SVR بر اساس معیار خطای جذر میانگین مربعات خطا نیز رتبه اول را کسب کرد. نتایج بررسی میزان خطای مدل های مورد مطالعه بر اساس آماره جذر میانگین مربعات خطا نشان داد که درصد بهبودی خطای مدل GP نسبت به مدل های SVM، ANFIS و ANN به ترتیب حدود ۲۳، ۶ و ۱۱ درصد می باشد. هم چنین نتایج نشان دهنده دقت ۳/۶ برابری مدل SVR نسبت به مدل رگرسیون چندمتغیره

- Impact of 30 Years Changing of River Flow on Urmia Lake Basin. *AUT Journal of Civil Engineering*, 2(1), pp. 115-122.
- Pour, S.H., Shahid, S. and Chung, E.S., 2016. A hybrid model for statistical downscaling of daily rainfall. *Procedia Engineering*, 154, pp.1424-1430.
- Khan, M.S., Coulibaly, P. and Dibike, Y., 2006. Uncertainty analysis of statistical downscaling methods. *Journal of Hydrology*, 319(1-4), pp.357-382.
- Salas, J.D., 1980. Applied modeling of hydrologic time series. Water Resources Publication, P. O. Box 2841. Littleton, Colorado .80161, U.S.A. 1980; 484 P.
- Swincow, T.D.W., 2002. *Statistics at Square One*. London. British Medical Journal.
- Tripathi, S., Srinivas, V.V. and Nanjundiah, R.S., 2006. Downscaling of precipitation for climate change scenarios: a support vector machine approach. *Journal of hydrology*, 330(3-4), pp.621-640.
- Wilby, R.L., Dawson, C.W. and Barrow, E.M., 2008. SDSM—a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling & Software*, 17(2), pp.145-157.
- Liu, Z., Xu, Z., Charles, S.P., Fu, G. and Liu, L., 2011. Evaluation of two statistical downscaling models for daily precipitation over an arid basin in China. *International Journal of Climatology*, 31(13), pp.2006-2020.
- in problem solving. In *Soft computing and industry* (pp. 635-653). Springer, London.
- Hofmann, T., Tsochantaridis, I. and Altun, Y., 2002. Learning over structured output spaces via joint kernel functions. In *Proceedings of the Sixth Kernel Workshop*.
- Jang, J.S., 1992. Self-learning fuzzy controllers based on temporal backpropagation. *IEEE Transactions on neural networks*, 3(5), pp.714-723.
- Karakus, M. and Tutmez, B., 2006. Fuzzy and multiple regression modelling for evaluation of intact rock strength based on point load, Schmidt hammer and sonic velocity. *Rock mechanics and rock engineering*, 39(1), pp.45-57.
- Karamouz, M., FALAHI, M., Nazif, S. and RAHIMI, F.M., 2009. Long lead rainfall prediction using statistical downscaling and artificial neural network modeling.
- Mamdani, E.H. and Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of man-machine studies*, 7(1), pp.1-13.
- Mander, W.J., 1994. *Dictionary of global climatic change*. 2nd edn, VCL Press, London, England, 1994: 262 pp
- Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. *Journal of hydrology*, 10(3), pp.282-290.
- Nazeri-Tahroudi, M., Ahmadi, F. and Khalili, K., 2018.

