

ارزیابی مدل‌های داده محور مبتنی بر ریزمقیاس نمایی مقادیر دمای روزانه

حسین خزیمه‌نژاد^{۱*}، مریم صفوی گردینی^۲، مهدی امیرآبادی زاده^۳، محمد ناظری تهرودی^۴

تاریخ ارسال: ۱۳۹۹/۱۱/۲۰

تاریخ پذیرش: ۱۴۰۰/۰۲/۱۳

مقاله پژوهشی

چکیده

در این پژوهش با استفاده از شش مدل شبکه عصبی (ANN)، انفیس (ANFIS)، ماشین بردار پشتیبان (SVM)، برنامه‌ریزی ژنتیک (GP)، رگرسیون بردار پشتیبان (SVR) و رگرسیون چندمتغیره (Reg)، متوسط دمای روزانه ایستگاه‌های کرمان و بم طی دوره آماری ۱۹۶۱-۲۰۰۵ بررسی و شبیه‌سازی شد. نتایج نشان داد که متوسط دمای روزانه طی دوره‌های یاد شده افزایش چشمگیری برای هر دو منطقه خواهند داشت. نتایج کلی به دست آمده حاکی از برتری نتایج مدل رگرسیونی SVR (کرمان: $R=0.992$ و $RMSE=1.105$ و بم: $R=0.99$ و $RMSE=1.01$) می‌باشد. نتایج بررسی‌ها نشان داد که مدل SVR، میزان خطای شبیه‌سازی را نسبت به مدل‌های شبکه عصبی (ANN)، انفیس (ANFIS)، برنامه‌ریزی ژنتیک (GP) و رگرسیون چندمتغیره (Reg) در ایستگاه کرمان به ترتیب ۳۲، ۴۲، ۳۰ و ۱۱ درصد و در ایستگاه بم به ترتیب ۶۲، ۵۹، ۲۷ و ۲۷ درصد بهبود می‌بخشد. نتایج بررسی جذر میانگین مربعات خطا نشان داد که از بین شش مدل مورد بررسی، مدل رگرسیون بردار پشتیبان و برنامه‌ریزی ژنتیک برای ایستگاه بم و مدل رگرسیون بردار پشتیبان برای ایستگاه کرمان از دقت بالاتری برخوردار می‌باشد. همچنین نتایج نشان داد که برآورد میزان دمای متوسط ایستگاه سینوپتیک بم دارای کارایی و دقت بیشتری نسبت به ایستگاه سینوپتیک کرمان هستند. در این پژوهش هر چند آنالیز نتایج خروجی مدل‌ها منجر به نتایج یکسانی نشد ولی نتایج مدل‌ها بیانگر افزایش در متغیرهای دمایی در دو شهر کرمان و بم در دوره‌های آتی می‌باشد.

واژه‌های کلیدی: پیش‌بینی، تغییر اقلیم، کرمان، گردش عمومی جو.

^۱ - نویسنده مسئول و دانشیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند Email: hkhozayemeh@birjand.ac.ir
^۲ - دانشجوی دکتری منابع آب، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند Email: maryam70safavi@gmail.com
^۳ - استادیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند Email: m.amirabadi@birjand.ac.ir
^۴ - دانش‌آموخته دکتری منابع آب، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند Email: m_nazeri2007@yahoo.com

مقدمه

صنعتی شدن جوامع و افزایش گازهای گلخانه‌ای در دهه‌های گذشته باعث افزایش دمای کره زمین و تغییر در دیگر پارامترهای اقلیمی شده است که به آن پدیده تغییر اقلیم گفته می‌شود (IPCC, 1995). تغییر الگوی بارش در کنار تغییرات دمایی از جمله مهم‌ترین تأثیرات تغییر اقلیم به شمار می‌رود. مطمئناً برای یک برنامه‌ریزی درست برای آینده، پیش‌یابی دقیق این پارامترهای اقلیمی امری کاملاً ضروری به حساب می‌آید (Hay et al., 2000).

تغییر اقلیم و گرم شدن کره زمین پیامدهای زیان‌باری بر منابع مختلف از جمله آب، پوشش گیاهی و سرانجام زندگی انسان دارد. اولین اثر تغییر اقلیم روی عناصر اتمسفری خصوصاً دما و بارش می‌باشد (آبکار و همکاران، ۱۳۹۲). مهم‌ترین پیامدهای تغییر اقلیم، تأثیرگذاری آن بر پدیده‌های جوی و اقلیمی نظیر طوفان، سیل، تگرگ، خشک‌سالی، امواج گرمایی و سرماهای نابهنگام خواهد بود. اهمیت بررسی و مدل‌سازی هرچه دقیق‌تر اقلیم از نگاه دیگری نیز قابل طرح است و آن اهمیت پارامترهای اقلیمی در طرح‌های آمایش سرزمین است، که با این وجود داده‌های مورد استفاده اقلیمی در این پژوهش‌ها که نتایج آن زمینه‌ساز شکل‌گیری توسعه پایدار در حوضه مورد مطالعه خواهد بود شامل؛ نقشه‌ها و داده‌های مربوط به آمارهای ثبت شده دوره دیده‌بانی منطقه بوده، در صورتی که، اقلیم برخلاف پارامترهای اکولوژیکی پایداری چون؛ زمین‌شناسی و خاکشناسی سرزمین، دارای ماهیتی ناپایدار بوده و تحت تأثیر تغییرات جهانی ناگهانی و دوره‌ای اقلیم می‌باشد. در نظر نگرفتن پویایی پارامترهای اقلیمی، سبب کاهش ضمانت پایدار ماندن توسعه در سالیان آتی سرزمین خواهد شد (امیری و همکاران، ۱۳۹۵).

نتایج حاصل از بررسی‌های انجام شده از روند تغییرات دمای کره زمین در قرن گذشته توسط هیئت بین‌الدول تغییر اقلیم نشان می‌دهد که پدیده تغییر اقلیم، به صورت افزایش دمای کره زمین در اکثر نقاط جهان نمود یافته که در جوامع علمی از آن تحت عنوان تغییر اقلیم یاد می‌شود. طبق تحقیقات جامعی که توسط IPCC در مقیاس قاره‌ای انجام یافته است، پدیده تغییر اقلیم نه تنها بر دمای کره زمین تأثیرگذار بوده است، بلکه تغییراتی را نیز در

خصوصیات سیستم‌هایی که با جو زمین در تعادل هستند، به وجود آورده است.

این هیئت در گزارش‌های خود در سال‌های ۱۹۹۵ و ۲۰۰۱ میلادی به تغییرات جهانی اقلیمی حاکی از افزایش دمای کره زمین حدود ۰/۷ درجه سانتی‌گراد از نیمه قرن نوزدهم اشاره نمودند. بر طبق گزارش منتشر شده این هیئت در سال ۲۰۰۷، بیش‌ترین افزایش مشاهده شده در دمای میانگین جهانی از نیمه قرن بیستم به احتمال زیاد به‌دلیلی افزایش گازهای گلخانه‌ای ناشی از فعالیت‌های بشر می‌باشد (IPCC, 2007). تغییر اقلیم و پیامدهای ناشی از آن خصوصاً گرم شدن کره زمین، به‌عنوان بزرگ‌ترین معضل قرن بیست و یکم شناخته شده است. به‌طوری‌که طی چند سال اخیر همواره یکی از محورهای اصلی مباحث سران کشورهای صنعتی جهان، مسئله تغییر اقلیم بوده است (صمدی و همکاران، ۱۳۸۸). مهم‌ترین ابزار جهت شبیه‌سازی وضعیت آینده پارامترهای اقلیمی استفاده از مدل‌های عمومی گردش جوی اقیانوسی است یکی از بزرگ‌ترین مشکلات مدل‌های اقلیمی بزرگ مقیاس بودن خروجی‌های این مدل‌ها می‌باشد. تفاوت دقت روش‌های ریزمقیاس‌نمایی می‌تواند باعث اختلاف در نتایج شبیه‌سازی شود. که این امر موجبات بررسی دقت روش‌های ریزمقیاس‌نمایی را فراهم می‌نماید. با توجه به اهمیت پیش‌یابی‌های هر چه دقیق‌تر تغییر در شرایط جوی و نقش مدل‌های مورد استفاده در شبیه‌سازی و پیش‌یابی پارامترهای مؤثر بر چرخه هیدرولوژی پژوهش‌های مختلفی در سطح جهان و ایران در زمینه ریزمقیاس‌نمایی و شبیه‌سازی و پیش‌یابی پارامتر-های اقلیمی صورت گرفته است.

Cheema et al (2013) عملکرد روش ریزمقیاس‌نمایی مدل SDSM را روی روند داده‌های کمینه دمای ایستگاه‌های کشور پاکستان در دوره ۲۰۱۰-۱۹۹۱ ارزیابی کردند. بر اساس نتایج این پژوهش، طبق آزمون من‌کنندال روند افزایشی کمینه دمای سالانه معنادار است. علاوه بر آن، تحلیل‌ها نشان داد که تطابق خوبی بین داده‌های دمای الگو شده و داده‌های واقعی وجود دارد؛

ضمن آنکه ضریب همبستگی پیرسون برای بیش‌تر نواحی بیش از نود درصد بود. این پژوهشگران روش-



استفاده می‌شوند. نتایج نشان داد اگرچه هر دوی این دو مدل نتایج یکسانی ارائه نمی‌دهند ولی هر دوی این روش‌ها روند افزایش عمومی در میانگین دمای روزانه را نشان می‌دهد. رضایی و همکاران (۱۳۹۳)، به بررسی کارایی مدل ریزمقیاس نمایی آماری SDSM در پیش‌یابی پارامترهای دمایی در دو اقلیم خشک و فرا خشک پرداخته‌اند. نتایج این بررسی نشان داد که این مدل در مناطق خشک کارایی بالاتر و دقت بیشتری را برای پیش‌یابی دما نسبت به مناطق فراخشک ارائه می‌دهد. عینی و همکاران (۱۳۹۵) نیز کمینه دمایی شرق استان کرمانشاه را مدل‌سازی کرده‌اند. نتایج حاصل از تحلیل و مدل‌سازی دمای ماهانه ایستگاه‌های مورد مطالعه با استفاده از مدل آریمای و پیش‌یابی کمینه دمای ماهانه تا سال ۲۰۱۶ نشان داد که کمینه دمایی دو ایستگاه سینوپتیک کرمانشاه و کنگاور از یک مدل نهایی ARIMA (1,0,1)(1,0,0) و در واقع از یک روند مشابه در تغییرات کمینه دمایی برخوردار می‌باشند که با روندی نسبتاً ملایم در حال افزایش است. سلگی (۱۳۹۷)، برای پیش‌یابی بارش ماهانه با استفاده از داده‌های دما و رطوبت نسبی ماهانه در یک دوره ۳۲ ساله (۱۳۶۲-۱۳۹۳) از دو مدل GP، SVM بهره برد. نتایج وی نشان داد که ضریب همبستگی هر دو مدل ۰/۹۲ و RMSE این دو مدل به ترتیب ۰/۰۴۸۶ و ۰/۰۴۷۸ بوده است. کدخدا حسینی و همکاران (۱۳۹۷)، عملکرد دو مدل GP و SVM را در بازسازی داده‌های گمشده بارندگی در دوره (۱۳۷۰-۱۳۸۹) مورد بررسی قرار دادند. نتایج آن‌ها نشان داد که ضریب همبستگی در مدل GP و SVM به ترتیب ۰/۹۳ و ۰/۹۵ بود. ارزیابی خطا بر اساس معیار RMSE برای دو مدل به ترتیب ۱۳ و ۱۲/۲ میلی‌متر محاسبه شده است. شاهی نژاد و دهقانی (۱۳۹۷)، به مقایسه مدل‌سازی با استفاده از SVM و ANN در تخمین اکسیژن محلول آب رودخانه در یک دوره ۱۰ ساله (۲۰۱۰-۲۰۱۶) پرداختند. نتایج آن‌ها نشان داد که ANN با ضریب همبستگی ۰/۹۶ و RMSE برابر ۰/۶۶۸ دارای بیشترین ضریب همبستگی و کمترین خطا بوده است.

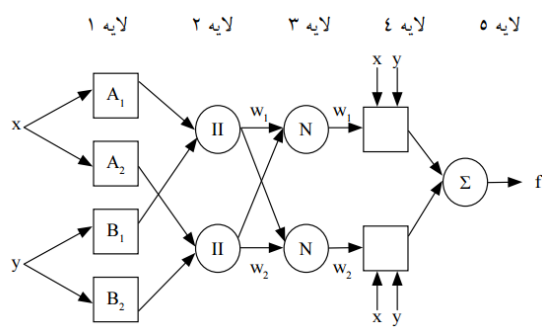
به طور کلی می‌توان گفت تاکنون در نقاط مختلف جهان تحقیقات متعددی در خصوص مقایسه و کارایی مدل‌ها صورت گرفته است که در هیچ کدام از این تحقیقات مزیت هیچ مدلی نسبت به دیگری به صورت قطعی به اثبات

های آماری مختلفی برای بررسی روند به کار گرفتند. یافته‌های آن‌ها نشان داد تغییر معنادار اقلیمی بیش‌تر در نواحی شمالی کشور پاکستان رخ داده است. نوری و علم (۲۰۱۴) الگوی SDSM را برای داده‌های مشاهداتی دما و بارش روزانه از سال ۱۹۸۱ تا ۲۰۰۶ با استفاده از الگوی جهانی HadCM3 به کار بردند. در این تحقیق، از شاخص درصد اریبی (PBIAS)، شاخص کفایت نش - ساتکلیف و شاخص اصلاح شده تطابقی برای ارزیابی داده‌های بارش و دمای ریزمقیاس شده استفاده شده است. مقدار شاخص PBIAS دمای ریزمقیاس شده کمینه بوده (۰/۳۰ - درصد)، شاخص نش - ساتکلیف (۰/۸۰) و شاخص اصلاح شده تطابقی (۰/۸۳) بیش‌ترین حد برای دمای بیشینه روزانه در ایستگاه سیلحت بنگلادش بوده است. در میان پنج ایستگاه باران‌سنجی شاخص PBIAS بارش ریزمقیاس شده دارای مقدار کمینه (۱/۳۱٪)، نش - ساتکلیف (۰/۷۶) و شاخص اصلاح شده (۰/۷۹) بیشترین حد را در ایستگاه کانایرات داشته است. داده‌های دما و بارش ریزمقیاس شده با داده‌های مشاهده شده تا حدودی مطابقت داشته‌اند. Kazemi et al (2014) روش SDSM را برای خروجی دمای روزانه الگوی جهانی ECHAM5 به کار بردند. نتایج بررسی آن‌ها نشان داد داده‌های ریزمقیاس شده بسیار دقیق‌تر از داده‌های الگوی ECHAM5 است. ضریب همبستگی داده‌های ریزمقیاس شده با داده‌های مشاهده شده بین ۸۱ تا ۹۴ درصد است؛ در حالی که همین همبستگی برای داده‌های الگوی جهانی بین ۷۳ تا ۸۷ درصد است. Hassan et al (2014) به مطالعه مدل ریزمقیاس نمایی آماری برای تولید متغیرهای هواشناسی با ارزش از روش SDSM و LARS-WG برای شبیه‌سازی و ریزمقیاس نمایی بارش و دما استفاده کردند. به عنوان مثال برای بارش و دما در سایت‌های انتخاب شده در شبه جزیره پنین سولار مالزی از این دو مدل استفاده کردند که عبارتند از: ۱) مدل ریزمقیاس نمایی آماری SDSM که برای مدل رگرسیون و ژنراتور آب و هوای تصادفی استفاده می‌شود، ۲) ژنراتورهای آب و هوای ایستگاه تحقیقاتی لانگ اشتون LARS-WG که فقط برای تولید آب و هوای تصادفی استفاده می‌شود. این دو مدل به عنوان ابزارهایی برای اندازه‌گیری کمی ویژگی‌های محیطی در مقیاس محلی

روش سیستم استنتاج فازی- عصبی تطبیقی روشی ترکیبی است که در آن بخش فازی رابطه بین متغیرهای ورودی و خروجی را برقرار می‌نماید و مشخصه‌های مربوط به توابع عضویت بخش فازی نیز به‌وسیله شبکه عصبی تعیین می‌شود. سیستم استنتاج فازی- عصبی تطبیقی بیشتر با استفاده از سیستم فازی سوگنو و به‌صورت ساختار شبکه‌ای پیش‌رونده استفاده می‌شود که دارای ۵ لایه مطابق شکل ۲ به شرح زیر می‌باشد (Karakus and Tutmez, 2006).



شکل (۱): موقعیت ایستگاه‌های مورد مطالعه در کشور



شکل (۲): ساختار شبکه ANFIS

سامانه فازی، سامانه‌ای منطبق بر قوانین فازی است. بنابراین، انفیس مبتنی بر سامانه‌های فازی است که ساختار

نرسیده است. تحقیق حاضر به بررسی میزان کارایی مدل‌های شبکه عصبی، انفیس، برنامه‌ریزی ژنتیک، ماشین بردار پشتیبان، SVR و رگرسیون چند متغیره در شبیه‌سازی میانگین دمای روزانه در ایستگاه‌های سینوپتیک کرمان و بم پرداخته است. در این مطالعه از بهینه‌سازی مقادیر پارامترهای مدل رگرسیون بردار پشتیبان جهت رسیدن به نتیجه بهتر استفاده شده است. در این مطالعه، ارزیابی مدل‌ها در دو مرحله آموزش و آزمایش استفاده شده است. پیش‌یابی کننده‌های CanESM2 متناسب با هر ایستگاه نیز به عنوان ورودی‌های مدل در نظر گرفته شد.

مواد و روش‌ها

مطالعه حاضر در دو ایستگاه سینوپتیک کرمان و بم که قدیمی‌ترین آمار دما در استان کرمان را دارا می‌باشند، انجام گرفته است. ایستگاه‌های مذکور به عنوان نماینده دو منطقه اقلیمی خشک و فراخشک در نظر گرفته شده است. منطقه کرمان در طول جغرافیایی ۵۶ درجه و ۳۸ دقیقه و عرض جغرافیایی ۳۰ درجه و ۱۵ دقیقه و منطقه بم در طول جغرافیایی ۵۸ درجه و ۲۱ دقیقه و عرض جغرافیایی ۲۶ درجه و ۹ دقیقه قرار دارند و دوره آماری برای هر دو ایستگاه از سال ۱۹۶۱ تا ۲۰۰۵ میلادی بوده است. موقعیت ایستگاه‌های مورد مطالعه به شرح شکل ۱ می‌باشد. در این مطالعه روند تغییرات سری‌های زمانی مقادیر دما موجود در منطقه مورد مطالعه با آزمون ناپارامتری من-کندال اصلاح شده با حذف اثرات خودهمبستگی داخلی مورد بررسی قرار می‌گیرد (Kendall, 1975). این آزمون به‌طور وسیعی در مطالعات هیدرولوژی، کلیماتولوژی و متئورولوژی مورد استفاده قرار گرفته است. شرط لازم برای استفاده از این آزمون عدم وجود خودهمبستگی در سری زمانی داده‌ها می‌باشد، با این حال ممکن است داده‌ها دارای خودهمبستگی معنی‌دار باشند. بنابراین باید ابتدا اثر خودهمبستگی داده‌ها حذف گردد تا بتوان از آزمون من-کندال استفاده نمود (Khalili et al., 2016). در این تحقیق از روش‌های ANFIS، GP، SVM، ANN، SVR و رگرسیون چند متغیره برای شبیه‌سازی میانگین دمای روزانه برای دو منطقه مورد مطالعه استفاده شد و سپس به ارزیابی دقت این روش‌ها پرداخته شده است.

روش سیستم استنتاج فازی- عصبی تطبیقی (ANFIS)



می‌شود. $O_{2,k}$ نمایانگر خروجی گره k ام در لایه دوم و برابر حاصل ضرب درجه عضویت هر ورودی است.

$$O_{2,k} = \mu_{A_i}(x) \times \mu_{B_j}(y) \quad (5)$$

لایه سوم: شامل گره‌های نرمال شده است که نسبت درجه مشارکت هر قانون به مجموع درجه مشارکت همه قوانین را محاسبه می‌کند. در نتیجه، این لایه به شکل رابطه (۶) تعریف می‌شود:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{\sum_{k=1}^4 w_k} \quad (6)$$

لایه چهارم: مشتمل بر گره‌های تطبیقی است که با استفاده از پارامترهای نتیجه، خروجی هر گره را محاسبه می‌کند و به شکل رابطه (۷) تعریف می‌شود:

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (7)$$

که w_i خروجی i امین گره از لایه قبلی است و $\{p_i, r_i, q_i\}$ پارامترهای تطبیقی خطی هستند. لایه پنجم: مشتمل بر گره خروجی است که مقدار نهایی خروجی را به شکل مجموع خروجی گره‌های لایه ماقبل بیان می‌کند:

$$O_{5,i} = \bar{w}_i f_i = \frac{\sum_{i=1}^4 w_i f_i}{\sum_{i=1}^4 w_i} \quad (8)$$

الگوریتم یادگیری شبکه انفیسی یک الگوریتم یادگیری هیبریدی مشتمل بر الگوریتم گرادیان نزولی و روش حداقل مربع بازگشتی است. از الگوریتم گرادیان نزولی برای به روز رسانی پارامترهای غیرخطی شبکه و از تخمین زن حداقل مربع بازگشتی برای تنظیم وزن‌های شبکه استفاده می‌شود. خطای آموزش شبکه به شکل رابطه (۹) تعریف شده است:

$$E = \sum_{i=1}^N (f_i - \hat{f}_i)^2 \quad (9)$$

که در آن f_i و \hat{f}_i به ترتیب خروجی مطلوب و تخمین زده شده شبکه به ازای i امین ورودی هستند. همچنین، N تعداد کل زوج داده‌های ورودی - خروجی (داده‌های آموزشی) شبکه است (Jang, 1992).

برنامه‌ریزی بیان ژن

روش برنامه‌ریزی بیان ژن، در سال ۱۹۹۹ توسط فی‌یرا (Ferreira, 2001). این روش ترکیبی از روش‌های

آن در شکل ۱ نشان داده شده است. در این ساختار، گره‌های ورودی و خروجی به ترتیب مقادیر پیش‌یابی شده را نشان می‌دهند. برای ساده‌سازی، ورودی و مقادیر یک شبکه دو ورودی - تک خروجی در نظر گرفته شده است. برای نمونه دو قانون در زیر آورده شده است:

قانون اول: اگر x برابر A_1 و y برابر A_1 باشد، آنگاه:

$$f_1 = p_1 x + q_1 y + r_1 \quad (1)$$

قانون اول: اگر x برابر A_2 و y برابر A_2 باشد، آنگاه:

$$f_2 = p_2 x + q_2 y + r_2 \quad (2)$$

همان‌طور که است در شکل ۲ مشاهده می‌شود، انفیسی شبکه‌ای پنج لایه است و هر لایه گره‌های مختلف دارد و هر گره در یک لایه ثابت یا تطبیقی است. به عبارت دیگر، خروجی هر گره در این لایه درجه عضویت تخصیص داده شده به متغیرهای ورودی در مجموعه‌های فازی است که به شکل زیر بیان می‌شود:

$$i = 1, 2 \quad O_{1,i} = \mu_{A_i}(X) \quad (3)$$

$$i = 3, 4 \quad O_{2,i} = \mu_{B_{i-2}}(X) \quad (4)$$

در روابط ۳ و ۴، x و y ورودی‌های غیر فازی گره i ام و A_i و B_i توابع عضویت فازی هستند. همچنین، مقدار عضویت هر ورودی به عنوان خروجی اول با $(\mu_{A_i})x$ و $(\mu_{B_i})y$ مشخص می‌شود. بنابراین، باید پارامترهای توابع عضویت که به عنوان پارامترهای بخش مقدم قوانین فازی شناخته می‌شوند و از دسته پارامترهای غیرخطی هستند، مشخص شوند. در ساختار انفیسی معمولاً از فازی‌سازی گوسی استفاده می‌شود. فازی‌سازی گوسی یکی از انواع فازی‌سازها در سامانه‌های فازی است. به دلیل اینکه سامانه‌های واقعی و کاربردهای مهندسی همواره با متغیرهای عددی سروکار دارند، بنابراین برای این که بتوان از سامانه فازی برای توصیف یا مدل‌سازی هر سامانه واقعی استفاده کرد، باید نگاشتی بین ورودی‌های عددی (غیر فازی) سامانه و یک مجموعه فازی ایجاد کرد. به همین منظور از فازی‌سازها استفاده می‌شود. به عبارت دیگر، فازی‌ساز قابلیت آن را دارد تا به شکل یک رابط، هر ورودی غیرفازی (عددی) را به یک ورودی فازی تبدیل کند. لایه دوم: شامل گره‌های قوانین است، هر گره در این لایه درجه فعالیت یک قانون را محاسبه می‌کند. در این لایه از عملگر "و" برای محاسبه درجه مشارکت هر قانون استفاده

است. ϕ تابع کرنل، N تعداد نمونه‌ها و دو مشخصه ξ_i و ξ_i^* متغیر کمبود هستند که حد بالا و پایین خطای آموزش مرتبط با مقدار خطای مجاز ε را مشخص می‌کنند. در مسائل پیش‌بینی 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 و بیش تخمینی را با کمینه کردن هم‌زمان دو ترم $W^T W / 2$ و خطای آموزشی، یعنی $\sum_{i=1}^N (\xi_i + \xi_i^*)$ را در رابطه ۱۲ حل می‌کند. بنابراین با معرفی ۲ ضریب لاگرانژ a_i و a_i^* مسئله بهینه‌سازی با حداکثر سازی عددی تابع درجه دوم زیر (معادله ۱۴) با شرایط معادله ۱۵ حل خواهد شد (اسکندری و همکاران ۱۳۹۱).

$$\sum_{i=1}^N y_i (a_i + a_i^*) - \varepsilon \sum_{i=1}^N (a_i + a_i^*) - 0.5 \cdot \sum_{i,j=1}^N (a_i + a_i^*)(a_j + a_j^*) \phi(x_i)^t \phi(x_j) \quad (14)$$

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

$$0 \leq a_i \leq C, 0 \leq a_i^* \leq C, \quad (15)$$

$$i = 1, 2, \dots, N$$

تابع هدف بالا در رابطه ۱۶ تابع محدب است و بنابراین جواب رابطه ۱۶ یکتا و بهینه خواهد بود. پس از تعریف ضرایب لاگرانژ در رابطه ۱۶ مشخصه‌های w و b در مدل SVM رگرسیونی با استفاده از تئوری کراش - کوهن - تاکر محاسبه می‌شود. که در آن $W = \sum_{i=1}^N (a_i + a_i^*) \phi(x_i)$ است. در نتیجه برای مدل SVM رگرسیونی خواهیم داشت (ناظری تهرودی و همکاران، ۱۳۹۵):

$$W = \sum_{i=1}^N (a_i + a_i^*) \phi(x_i) + b \quad (16)$$

باید توجه داشت که ترم‌های لاگرانژ $(a_i + a_i^*)$ می‌تواند صفر، و یا غیر صفر باشند. بنابراین فقط مجموعه داده‌هایی که ضرایب a_i آن‌ها غیر صفر است در معادله رگرسیون نهایی وارد می‌شوند و این مجموعه داده‌ها به‌عنوان بردارهای پشتیبان شناخته می‌شوند. به‌طور ساده،

برنامه‌ریزی ژنتیک و الگوریتم ژنتیک بوده که در آن، کروموزوم‌های خطی و ساده با طول ثابت، مشابه با آنچه که در الگوریتم ژنتیک استفاده می‌شود و ساختارهای شاخه‌ای با اندازه‌ها و اشکال متفاوت، مشابه با درختان تجزیه در برنامه‌ریزی ژنتیک، ترکیب می‌شوند. از آنجایی که در این روش تمام ساختارهای شاخه‌ای با اندازه و اشکال متفاوت، در کروموزوم‌های خطی با طول ثابت کدگذاری می‌شوند، سبب گردیده که در این روش فنوتیپ و ژنوتیپ از هم جدا شوند و سیستم بتواند از تمام مزایای تکاملی به سبب وجود آن‌ها بهره‌مند شود.

ماشین بردار پشتیبان

اولین کاربرد این روش در مسائل آب توسط دیباک و همکاران در سال ۲۰۰۱ با مدل‌سازی بارندگی - رواناب ارائه شد (Hofman, 2001). 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 $f(x)$ به‌علاوه مقداری اغتشاش (خطای مجاز ε) مشخص شود (رابطه ۱۱) (اسکندری و همکاران ۱۳۹۱).

$$F(x) = W^T \cdot \phi(x) + b \quad (10)$$

$$Y = f(x) + noise \quad (11)$$

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 ۱۲) در مدل SVM با در نظر گرفتن شرایط مندرج در رابطه ۱۳ بهینه شود.

$$\frac{1}{2} W^T W + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^* \quad (12)$$

$$w^t \cdot \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^*$$

$$y_i - W^T \cdot \phi(X_i) - b \leq \varepsilon + \xi_i \quad (13)$$

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

در روابط فوق C عددی صحیح مثبت است که عامل تعیین جرمه در هنگام رخ دادن خطای آموزش مدل



ضریب کارایی نش- ساتکلیف

$$NSE = 1 - \frac{\sum_{i=1}^T (Q_i - Q'_i)^2}{\sum_{i=1}^T (Q_i - \bar{Q}_i)^2} \quad (19)$$

که Q_i : داده‌های واقعی یا مشاهده‌ای، Q'_i : داده‌های پیش‌یابی، \bar{Q}_i : میانگین داده‌ها، n تعداد داده‌ها و $\sigma_{Q'_i}$ و σ_{Q_i} به ترتیب برابر با واریانس داده‌های پیش‌یابی و مشاهداتی می‌باشد (Nash and Sutcliffe, 1970; Swinnow and Campbell, 2002; Salas et al., 1980; Raji et al., 2022).

نتایج و بحث

در این مطالعه جهت بررسی روند تغییرات مقادیر متوسط دما ایستگاه‌های سینوپتیک کرمان و بم از آزمون من - کندال اصلاح شده با حذف اثرات خودهمبستگی استفاده شده است. نتایج بررسی روند تغییرات مقادیر متوسط دمای سالانه در ایستگاه سینوپتیک کرمان و بم نشان داد که روند تغییرات این پارامتر کاهش‌ی ولی غیرمعنی‌دار می‌باشد و نتایج آماره کندال و شیب سن برای ایستگاه کرمان به ترتیب $-0/45$ و $-0/02$ - برای ایستگاه بم به ترتیب برابر با $0/5$ و $0/000593$ می‌باشد. همان‌طور که ذکر گردید، در این تحقیق جهت ریزمقیاس‌نمایی مقادیر متوسط دمای روزانه ایستگاه سینوپتیک کرمان و بم با استفاده از داده‌های پیش‌یابی کننده CanESM2 دقت و کارایی مدل‌های مختلفی نظیر ANN، SVM، SVR، ANFIS، GP و رگرسیون چند متغیره مورد بررسی قرار گرفت.

یکی از مهم‌ترین مراحل در مدل‌سازی، انتخاب ترکیب مناسبی از متغیرهای ورودی است، لذا باید همبستگی متقابل بین متغیرهای ورودی و خروجی محاسبه گردید. قبل از بررسی و شبیه‌سازی داده‌های مورد بررسی، ابتدا با استفاده از روش همبستگی پیرسون، همبستگی بین داده‌ها مورد بررسی قرار گرفت. از آنجا که یکی از مهم‌ترین مراحل در مدل‌سازی، انتخاب ترکیب مناسبی از متغیرهای ورودی است، لذا باید همبستگی متقابل بین متغیرهای ورودی و

بردارهای پشتیبان آن داده‌هایی هستند که به ساخته شدن تابع رگرسیونی کمک می‌کنند.

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 (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 (Braddock et al., 1998). 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 و در لایه خروجی از تابع محرکه خطی استفاده می‌نماید و تابع یادگیری وزن‌ها تابع و تابع آموزش آن تابع لونیبرگ مارکوات 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. شکل ۳ نمونه‌ای از یک شبکه پرسپترون چندلایه را نشان می‌دهد.

ارزیابی دقت مدل‌های مورد بررسی

به منظور ارزیابی دقت و کارایی مدل‌های مورد بررسی از ضریب تبیین، جذر میانگین مربعات خطا و نش- ساتکلیف به شرح زیر استفاده شد.

جذر میانگین مربعات خطا (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_i - Q'_i)^2}{n}} \quad (17)$$

آزمون ضریب همبستگی

$$r = \frac{cov(Q_i, Q'_i)}{\sigma_{Q_i} \sigma_{Q'_i}} \quad (18)$$

مشاهده می‌باشد. بعد از انتخاب پارامترهای پیش‌یابی کننده، با استفاده از این ۱۵ پارامتر پیش‌یابی کننده برای بم و ۸ پارامتر پیش‌یابی کننده برای کرمان، مقادیر دمای روزانه ایستگاه سینوپتیک کرمان و بم و کرمان در دوره آماری ۲۰۰۵-۱۹۶۱ با استفاده از مدل‌های داده محور ذکر شده شبیه‌سازی گردید.

خروجی محاسبه گردید. قبل از بررسی و شبیه‌سازی داده-های مورد بررسی، ابتدا با استفاده از روش همبستگی پیرسون، همبستگی بین ۲۶ پارامتر پیش‌یابی کننده CanESM2 و مقادیر متوسط روزانه دمای ایستگاه‌های سینوپتیک کرمان و بم بررسی و برآورد گردید. نتایج بررسی همبستگی بین مقادیر دما و ۲۶ پارامتر پیش-یابی کننده به شرح جدول ۱ و ۲ برای هر دو ایستگاه قابل

جدول (۱): نتایج بررسی همبستگی بین مقادیر پیش‌بینی‌کننده‌ها و دمای روزانه ایستگاه سینوپتیک بم

همبستگی	توضیح متغیر پیش‌بینی‌کننده	پیش‌بینی‌کننده	همبستگی	عنوان پیش‌بینی‌کننده‌ها	پیش‌بینی‌کننده
-۰/۰۹	سرعت باد در فشار ۸۰۰ هکتوپاسکال	<i>p8-fgl</i>	-۰/۶۱	متوسط فشار سطح دریا	<i>MSLPGL</i>
-۰/۴۵	مولفه سرعت مداری در فشار ۸۰۰ هکتوپاسکال	<i>p8-ugl</i>	-۰/۰۶	سرعت باد در فشار ۱۰۰۰ هکتوپاسکال	<i>p1_FGL</i>
-۰/۴۲	مولفه سرعت نصف‌النهاری در فشار ۸۰۰ هکتوپاسکال	<i>p8-vgl</i>	-۰/۵۸	مولفه سرعت مداری در فشار ۱۰۰۰ هکتوپاسکال	<i>p1_Ugl</i>
۰/۰۵	تاوایی در فشار ۸۰۰ هکتوپاسکال	<i>p8-zgl</i>	-۰/۴۲	مولفه سرعت نصف‌النهاری در فشار ۱۰۰۰ هکتوپاسکال	<i>p1_Vgl</i>
-۰/۲۸	جهت باد در فشار ۸۰۰ هکتوپاسکال	<i>p8-thgl</i>	۰/۰۴۶	تاوایی در فشار ۱۰۰۰ هکتوپاسکال	<i>p1_zgl</i>
-۰/۰۷	واگرایی در فشار ۸۰۰ هکتوپاسکال	<i>p8-zhgl</i>	-۰/۰۵۵	جهت باد در فشار ۱۰۰۰ هکتوپاسکال	<i>p1thgl</i>
۰/۶۹	رطوبت نسبی در فشار ۵۰۰ هکتوپاسکال	<i>p500gl</i>	-۰/۱۰۶	واگرایی در فشار ۱۰۰۰ هکتوپاسکال	<i>p1zhgl</i>
-۰/۴۱	رطوبت نسبی در فشار ۸۵۰ هکتوپاسکال	<i>p850gl</i>	-۰/۰۷۳	سرعت باد در فشار ۵۰۰ هکتوپاسکال	<i>p5-fgl</i>
-۰/۵۱	بارش کل	<i>prcpgl</i>	-۰/۶۳	مولفه سرعت مداری در فشار ۵۰۰ هکتوپاسکال	<i>p5-ugl</i>
۰/۴۰	رطوبت ویژه در فشار ۵۰۰ هکتوپاسکال	<i>s500gl</i>	۰/۲۴۰	مولفه سرعت نصف‌النهاری در فشار ۵۰۰ هکتوپاسکال	<i>p5-vgl</i>
۰/۴۹	رطوبت ویژه در فشار ۸۵۰ هکتوپاسکال	<i>s850gl</i>	-۰/۲۸	تاوایی در فشار ۵۰۰ هکتوپاسکال	<i>p5-zgl</i>
۰/۴۲	رطوبت ویژه در نزدیک سطح	<i>shumgl</i>	۰/۴۳	جهت باد در فشار ۵۰۰ هکتوپاسکال	<i>5-thgl</i>
۰/۲۹	متوسط دما در ارتفاع ۲ متری	<i>tempgl</i>	۰/۵۳	واگرایی در فشار ۵۰۰ هکتوپاسکال	<i>5-zhgl</i>

جدول (۲): نتایج بررسی همبستگی بین مقادیر پیش‌بینی‌کننده‌ها و دمای روزانه ایستگاه سینوپتیک کرمان

همبستگی	عنوان پیش‌بینی‌کننده‌ها	پیش‌بینی‌کننده	همبستگی	عنوان پیش‌بینی‌کننده‌ها	پیش‌بینی‌کننده
۰/۰۰۱	سرعت باد در فشار ۸۰۰ هکتوپاسکال	<i>p8-fgl</i>	-۰/۸۸۴	متوسط فشار سطح دریا	<i>MSLPGL</i>
-۰/۱۵۱	مولفه سرعت مداری در فشار ۸۰۰ هکتوپاسکال	<i>p8-ugl</i>	۰/۱۰۸	سرعت باد در فشار ۱۰۰۰ هکتوپاسکال	<i>p1_FGL</i>
-۰/۲۷۶	مولفه سرعت نصف النهاری در فشار ۸۰۰ هکتوپاسکال	<i>p8-vgl</i>	-۰/۰۴۷	مولفه سرعت مداری در فشار ۱۰۰۰ هکتوپاسکال	<i>p1_Ugl</i>
۰/۰۸۴	تاوایی در فشار ۸۰۰ هکتوپاسکال	<i>p8-zgl</i>	۰/۲۹	النهاری در فشار ۱۰۰۰ هکتوپاسکال	<i>p1_Vgl</i>
-۰/۰۶۹	جهت باد در فشار ۸۰۰ هکتوپاسکال	<i>p8-thgl</i>	۰/۱۲۴	تاوایی در فشار ۱۰۰۰ هکتوپاسکال	<i>p1_zgl</i>
-۰/۵۱۷	واگرایی در فشار ۸۰۰ هکتوپاسکال	<i>p8-zhgl</i>	۰/۰۱۴	جهت باد در فشار ۱۰۰۰ هکتوپاسکال	<i>p1thgl</i>
۰/۸۳۹۰	رطوبت نسبی در فشار ۵۰۰ هکتوپاسکال	<i>p500gl</i>	-۰/۳۶	واگرایی در فشار ۱۰۰۰ هکتوپاسکال	<i>p1zhgl</i>
-۰/۶۷۸	رطوبت نسبی در فشار ۸۵۰ هکتوپاسکال	<i>p850gl</i>	-۰/۶۴۸	سرعت باد در فشار ۵۰۰ هکتوپاسکال	<i>p5-fgl</i>
-۰/۱۲۴	بارش کل	<i>prcpgl</i>	-۰/۷۱۱	مولفه سرعت مداری در فشار ۵۰۰ هکتوپاسکال	<i>p5-ugl</i>
-۰/۳۳۶	رطوبت ویژه در فشار ۵۰۰ هکتوپاسکال	<i>s500gl</i>	-۰/۲۴	مولفه سرعت نصف النهاری در فشار ۵۰۰ هکتوپاسکال	<i>p5-vgl</i>
۰/۲۴۲	رطوبت ویژه در فشار ۸۵۰ هکتوپاسکال	<i>s850gl</i>	-۰/۴۳۶	تاوایی در فشار ۵۰۰ هکتوپاسکال	<i>p5-zgl</i>
۰/۱۲۶	رطوبت ویژه در نزدیک سطح	<i>shumgl</i>	-۰/۲۱۸	جهت باد در فشار ۵۰۰ هکتوپاسکال	<i>5-thgl</i>
۰/۹۳۵	متوسط دما در ارتفاع ۲ متری	<i>tempgl</i>	۰/۱۲۷	واگرایی در فشار ۵۰۰ هکتوپاسکال	<i>5-zhgl</i>

جدول (۳): نتایج بررسی دقت مدل‌های مورد بررسی در برآورد مقادیر دمای متوسط روزانه ایستگاه سینوپتیک کرمان

مدل مورد بررسی	مرحله ارزیابی	همبستگی	نش - سانتکیلف	جذر میانگین مربعات خطا (سانتی‌گراد)
شبکه عصبی	آموزش	۱	۰/۳۷۳	۴/۸۷۶
	آزمایش	۱	۰/۷۷۲	۴/۷۱۶
انفیس	آموزش	۰/۸۹۴	۰/۷۹۹	۴/۲۳۹
	آزمایش	۰/۸۷	۰/۹۶	۴/۴۹۴
برنامه‌ریزی ژنتیک	آموزش	۰/۹	۰/۸۱۴	۴/۱۱۶
	آزمایش	۰/۸۸۳	۰/۷۷۴	۴/۲۶۲
ماشین بردار پشتیبان	آموزش	۰/۸۹۸	۰/۷۳۷	۴/۸۵۹
	آزمایش	۰/۸۸	۰/۹۳۹	۵/۲۵۶
رگرسیون بردار پشتیبان	آموزش	۰/۹۹۲	۰/۹۸۴	۱/۱۰۵
	رگرسیون چندگانه	۰/۹۰۱	۰/۸۱۲	۴/۰۶۹

*: واحد *RMSE* بر حسب درجه سانتی‌گراد می‌باشد.



دمای ایستگاه بـم در مرحله آموزش و آزمایش نشان داد که این مدل از دقت پایین تری نسبت به سایر روش ها برخوردار می باشد. همچنین میزان خطای مدل شبکه عصبی در تخمین متوسط دمای ایستگاه کرمان از دقت پایین تری برخوردار می باشد.

Sajjad Khan et al و Karamouz et al (2009) (2006) نیز دقت مدل های هوش مصنوعی را در ریزمقیاس نمایی مقادیر بیشینه دمای روزانه را ضعیف معرفی کردند. نتایج کلی به دست آمده حاکی از برتری نتایج مدل های بررسی شده برای منطقه بـم در مقایسه با منطقه کرمان می باشد. نتایج مدل سازی و برازش داده های مورد بررسی با بهترین مدل ها برای هر دو منطقه در شکل های ۴ و ۷ ارائه شده است.

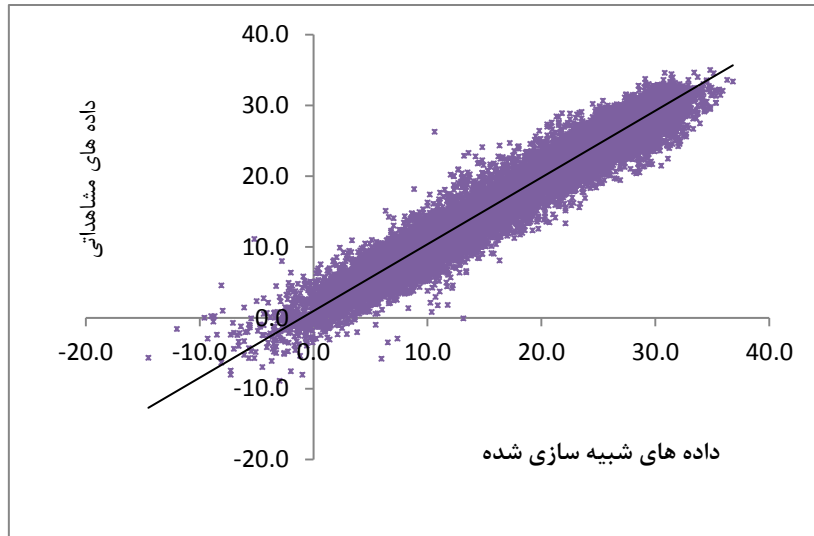
در پژوهش حاضر بعد از انجام واسنجی و ارزیابی مدل، پارامتر متوسط دمای روزانه توسط مدل های ذکر شده برای دوره ی (۱۹۶۱-۲۰۰۵) میلادی توسط داده های NCEP در دو شهر کرمان و بـم شبیه سازی شده و با داده های مشاهداتی مورد مقایسه قرار گرفته است.

در استفاده از مدل های مورد بررسی حدود ۸۰ درصد داده های متوسط دمای روزانه ایستگاه سینوپتیک کرمان و بـم به عنوان ورودی آموزشی مدل و ۲۰ درصد آن به عنوان ورودی آزمایشی مدل ها مورد بررسی قرار گرفت. نتایج بررسی دقت مدل سازی مدل های مورد بررسی به شرح جدول ۳ و ۴ ارائه شد.

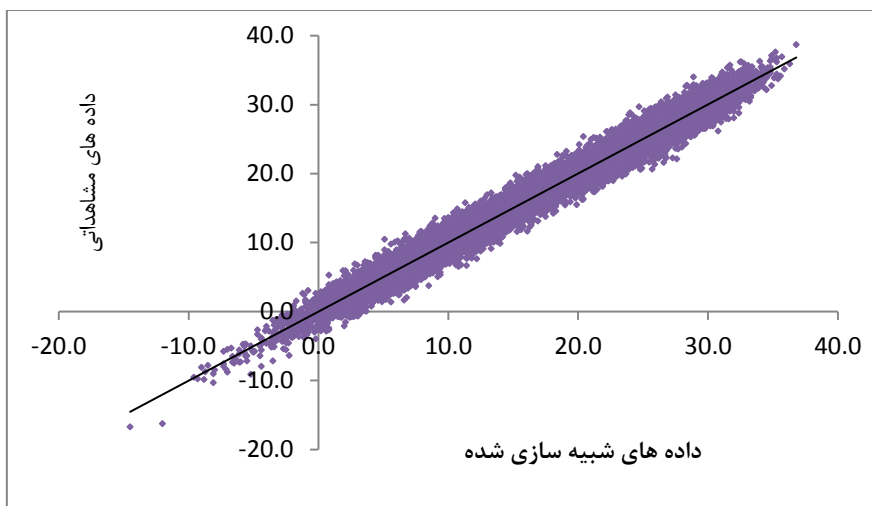
با توجه به نتایج بدست آمده، میزان خطای مدل رگرسیون چندگانه و انفیس در تخمین مقادیر متوسط

جدول (۴): نتایج بررسی دقت مدل های مورد بررسی در برآورد مقادیر دمای متوسط روزانه سینوپتیک بـم

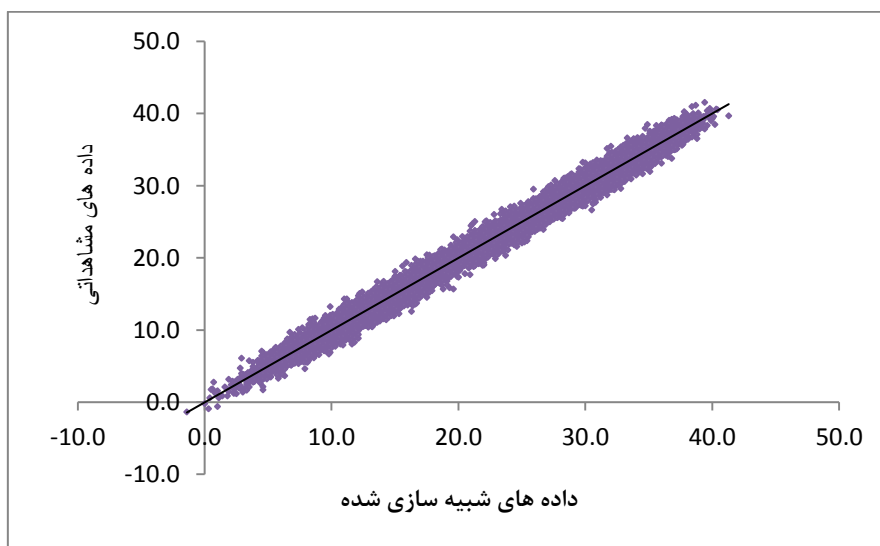
مدل مورد بررسی	مرحله ارزیابی	همبستگی	نش - ساتکلیف	جذر میانگین مربعات خطا (سانتی گراد)
شبکه عصبی	آموزش	۰/۹۵	۰/۹۷	۲/۷
	آزمایش	۰/۹۶	۰/۹۹	۲/۵۷
انفیس	آموزش	۰/۹۶	۰/۹۲	۲/۴۹
	آزمایش	۰/۹۷	۰/۹۹	۲/۹۷
برنامه ریزی ژنتیک	آموزش	۰/۹۸	۰/۹۹	۱/۴۷
	آزمایش	۰/۹۸	۰/۹۹	۱/۳۷
ماشین بردار پشتیبان	آموزش	۰/۹۸	۰/۹۷	۱/۴۰
	آزمایش	۰/۹۸	۰/۹۹	۱/۴۳
رگرسیون بردار پشتیبان	آموزش	۰/۹۹	۰/۹۸	۱/۰۱
	آزمایش	۰/۹۹	۰/۹۸	۱/۰۱
رگرسیون چندگانه		۰/۶۱	۰/۳۷	۷/۱۳



شکل (۴): نتایج تخمین مقادیر متوسط دمای روزانه ایستگاه هواشناسی کرمان با استفاده از مدل *Reg*

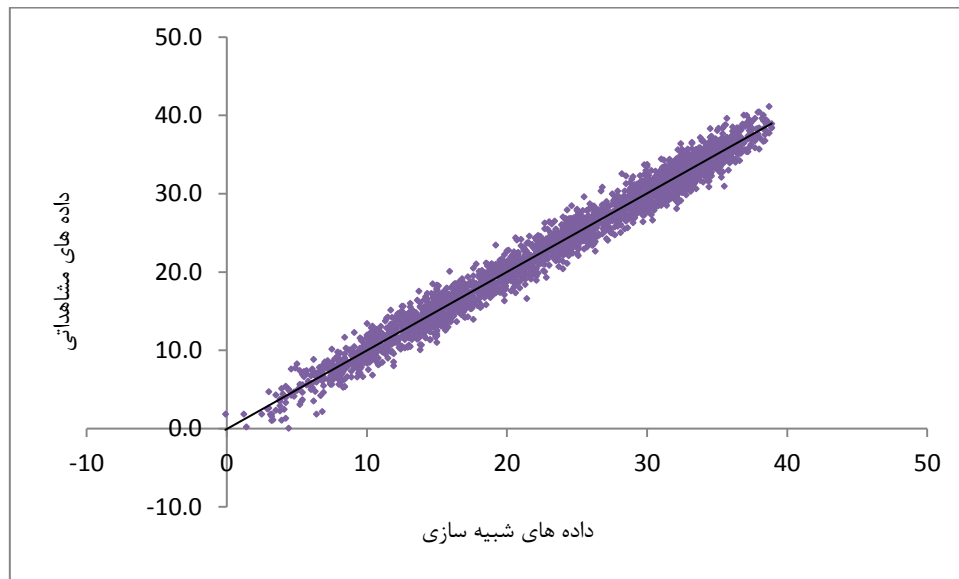


شکل (۵): نتایج تخمین مقادیر متوسط دمای روزانه ایستگاه هواشناسی کرمان با استفاده از مدل *SVR*





شکل (۶): نتایج تخمین مقادیر متوسط دمای روزانه ایستگاه هواشناسی بم با استفاده از مدل SVR



شکل (۷): نتایج تخمین مقادیر متوسط دمای روزانه ایستگاه هواشناسی بم با استفاده از مدل GP

نتیجه‌گیری

در این مطالعه دقت شش مدل مختلف در شبیه‌سازی مقادیر متوسط دمای روزانه ایستگاه‌های کرمان و بم در دوره آماری ۲۰۰۵-۱۹۶۱ مورد بررسی قرار گرفت. در این راستا از ۲۶ پارامتر پیش‌یابی شونده CanESM2 استفاده گردید. مدل‌های مورد بررسی در این تحقیق عبارتند از: رگرسیون بردار پشتیبان، رگرسیون چندمتغیره، برنامه‌ریزی ژنتیک، ماشین بردار پشتیبان، انفیس، و مدل شبکه عصبی.

بعد از بررسی و تصحیح و تکمیل داده‌های مورد بررسی، از بین ۲۶ پارامتر پیش‌یابی شونده مورد بررسی، ۱۵ پارامتر با توجه به همبستگی آن‌ها با مقادیر متوسط دما روزانه برای ایستگاه بم و ۸ پارامتر برای ایستگاه کرمان انتخاب شدند. با استفاده از این پارامترها و مقادیر متوسط دما ایستگاه کرمان و بم اقدام به مدل‌سازی و تخمین مقادیر یاد شده با استفاده از مدل‌های موجود شد. نتایج بررسی دقت مدل‌های یاد شده در دو مرحله آموزش و آزمایش مورد بررسی قرار گرفت. نتایج بررسی جذر میانگین مربعات خطا نشان داد که از بین شش مدل مورد بررسی، مدل رگرسیون بردار پشتیبان و برنامه‌ریزی ژنتیک برای ایستگاه بم و مدل رگرسیون بردار پشتیبان برای ایستگاه کرمان از دقت

بالتری برخوردار می‌باشد. همچنین نتایج نشان داد که برآورد میزان دمای متوسط ایستگاه سینوپتیک بم دارای کارایی و دقت بیشتری نسبت به ایستگاه سینوپتیک کرمان هستند. با توجه به نتایج بدست آمده، میزان خطای مدل رگرسیون چندگانه و انفیس در تخمین مقادیر متوسط دمای ایستگاه بم در مرحله آموزش و آزمایش نشان داد که این مدل از دقت پایین‌تری نسبت به سایر روش‌ها برخوردار می‌باشد. همچنین میزان خطای مدل شبکه عصبی در تخمین متوسط دمای ایستگاه کرمان از دقت پایین‌تری برخوردار می‌باشد.

همچنین نتایج نشان داد که برآورد میزان دمای متوسط ایستگاه سینوپتیک بم دارای کارایی و دقت بیشتری نسبت به ایستگاه سینوپتیک کرمان هستند. در این پژوهش هر چند آنالیز نتایج خروجی مدل‌ها منجر به نتایج یکسانی نشد ولی نتایج مدل‌ها بیانگر افزایش در متغیرهای دمایی در دو شهر کرمان و بم در دوره‌های آتی می‌باشد.

تشکر و قدردانی

این تحقیق در قالب طرح پژوهشی به شماره ابلاغیه ۱۳۹۸/د/۱۴۸۱۷ مورخ ۱۳۹۸/۷/۱۶ و با استفاده از اعتبارات

پژوهشی دانشگاه بیرجند انجام شده است که بدینوسیله
تشکر و قدردانی می‌شود.

منابع

- آبکار، ح.، ن، سلیمانی، م، علیجان و ه، کریم. ۱۳۹۲. بررسی میزان کارایی مدل *SDSM* در شبیه‌سازی شاخصه‌های دمایی در مناطق خشک و نیمه خشک، فصلنامه علمی پژوهشی مهندسی آبیاری و آب، شماره ۱۴.
- اسکندری، ع.، ر، نوری، ح، معراجی و ا، کیاقدی. ۱۳۹۱. توسعه مدلی مناسب بر مبنای شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای پیش‌بینی بهنگام اکسیژن‌خواهی بیوشیمیایی ۵ روزه. محیط‌شناسی، دوره ۳۸، شماره ۶۱، ص ۸۲-۷۱.
- جباریان امیری، ب.، ا، فاتحی و ن، محمدزاده. ۱۳۹۵. ریزمقیاس نمایی مدل گردش عمومی جو و کاربرد آن در شبیه سازی داده های هواشناسی استان گیلان، نشریه محیط زیست طبیعی، دوره ۶۹، شماره ۱، ص ۱۵۸-۱۴۳.
- رضایی، م.، م، نهتانی، ع، آبکار، م، رضایی و م، میرکازهی ریگی. ۱۳۹۳. بررسی کارایی مدل ریز مقیاس نمایی آماری (*SDSM*) در پیش‌نگری پارامترهای دمایی در دو اقلیم خشک و فراخشک (مطالعه موردی: کرمان و بم)، پژوهشنامه مدیریت حوزه آبخیز، دوره ۵، شماره ۱۰، ص ۱۱۷-۱۳۱.
- زینلی، م و س. ر، هاشمی. ۱۳۹۵. مقایسه توابع یادگیری شبکه عصبی در مدل‌سازی رواناب، اکوهیدرولوژی، دوره ۳، شماره ۴، ص ۶۵۹-۶۶۷.
- سلگی، ا.، ح، زراعی، م، شهنی دارابی و س، علی دادی. ۱۳۹۷. پیش‌نگری بارش ماهانه با استفاده از مدل‌های برنامه ریزی بیان ژن و ماشین بردار پشتیبان، نشریه تحقیقات کاربردی علوم جغرافیایی، دوره ۱۸، شماره ۵۰، ص ۱۰۳-۹۱.
- شاهی نژاد، ب و ر، دهقانی. ۱۳۹۷. مقایسه مدل های شبکه عصبی موجک، ماشین بردار پشتیبان و برنامه ریزی بیان ژن در تخمین میزان اکسیژن محلول در رودخانه ها، نشریه تحقیقات منابع آب ایران، دوره ۱۴، شماره ۳، ص ۲۷۷-۲۶۵.
- صمدی، م.، س، مهدوی و م، علیرضا. ۱۳۸۸. انتخاب متغیر پیش بینی کننده به منظور کوچک مقیاس کردن داده های دما و بارندگی در حوضه آبخیز کرخه، پنجمین همایش ملی علوم و مهندسی آبخیزداری ایران (مدیریت پایدار بلایای طبیعی)، دانشگاه علوم کشاورزی و منابع طبیعی گرگان.
- کدخدا حسینی، م.، ر، میرعباسی، ح، نودری و ع، رستمی. ۱۳۹۷. ارزیابی عملکرد رهیافت های برنامه ریزی ژنتیک و ماشین بردار پشتیبان در بازسازی داده های گم شده بارش. نشریه هواشناسی کشاورزی، دوره ۱، شماره ۶، ص ۴۹-۴۱.
- Braddock, R.D., M.L, Kremmer and L. Sanzogni. 1998. Feedforward artificial neural network model for forecasting rainfall-runoff, *Journal of Environmental Sciences*, 9: 419-432.
- Cheema, S.B., G, Rasul., G, Ali and D.H, Kazmi. 2013. A Comparison of Minimum Temperature Trends with Model Projections, *Pakistan Journal of Meteorology*, 8(15): 39-52.
- Ferreira, C. 2001. Gene Expression Programming: A New Adaptive Algorithm for Solving Problems, *Complex Systems*, forthcoming.
- Hassan, A., S, Shamsudin and S, Harun. 2014. Application of SDSM and LARS-WG for simulating and downscaling of rainfall and temperature. *Theor Appl Climatol*, 116: 243-257.
- Hay, L., R, Wilby and G, Leavesley. 2000. A comparison of delta change and downscaled GCM scenarios for three mountain.
- Hofmann, T. 2001. Unsupervised learning by probabilistic latent semantic analysis. *Machine learning*, 42.1: 177-196.
- IPCC. 1995. In: Watson, R.T., Zinyowera, M.C., Moss, R.H. Eds., *Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses*, Cambridge University Press, UK p. 878.
- IPCC. 2001. In: Watson, R.T., Zinyowera, M.C., Moss, R.H., Dokken, D.J., Eds., *Special Report on The Regional Impacts of Climate Change, An Assessment of Vulnerability*, Cambridge University Press, UK.



- Jang J.S.R. 1992. Fuzzy controllers based on temporal back propagation, IEEE Trans. Neural Netw, 3: 714–723.
- Karakus, M and B, Tutmez. 2006. Fuzzy and multiple regression modeling for evaluation of intact rock strength based on point load, Schmidt hammer and sonic velocity, Rock Mech, Rock Eng, 39: 45–57.
- Karamouz, M., M, Fallahi., S, Nazif and M, Rahimi Farahan. 2009. Long Lead Rainfall Prediction Using Statistical Downscaling and Arti_cial Neural Network Modeling, Transaction A: Civil Engineering, 16(2): 165-172.
- Kazmi, D.H., G, Rasul., J, Li and S.B, Cheema. 2014. Comparative Study for ECHAM5 and SDSM in Downscaling Temperature for a Geo-Climatically Diversified Region, Pakistan, Applied Mathematics, 5: 137-143.
- Kendall, M.G. 1975. Rank Correlation Measures, Charles Griffin, London.
- Khalili, K., M.N, Tahoudi., R, Mirabbasi and F, Ahmadi. 2016. Investigation of spatial and temporal variability of precipitation in Iran over the last half century, Stochastic environmental research and risk assessment, 30(4): 1205-1221.
- Nash, J.E and V, Sutcliffe. 1970. River flow forecasting through conceptual models part I A discussion of principles, Journal of Hydrology, 10(3): 282–290
- Raji, M., M. N, Tahroudi., F, Ye and J, Dutta. 2022. Prediction of heterogeneous Fenton process in treatment of melanoidin-containing wastewater using data-based models. Journal of Environmental Management, 307: 114518.
- Sajjad Khan, M., P, Coulibaly and Y, Dibike. 2006. Uncertainty analysis of statistical downscaling methods, Journal of Hydrology, 319: 357–382.
- Salas, J.D., J.W, Delleur., V, Yevjevich W.L, Lane. 1980. Applied Modeling of Hydrologic Time Series, Water resource Publications, P. O. Box 2841. Littleton, Colorado .80161, U.S.A, 484 P.
- Swinscow, T.D.V and M.J, Campbell. 2002. Statistics at Square One London: BMJ Publication, 106 P.



Evaluation of Data-Driven Models Based on Downscaling of Daily Temperature Values

Hossein Khozayemehzad^{*1}, Maryam Safavi², Mehdi Amirabadizadeh³, Mohammad Nazeri Tahroudi⁴

Abstract

In this study, using six models of neural network (ANN), ANFIS, support vector machine (SVM), genetic programming (GP), support vector regression (SVR) and multivariate regression (Reg), the mean daily temperature at Kerman and Bam stations, Iran were studied and simulated during the period of 1961-2005. The results showed that the mean daily temperature during the mentioned periods will increase significantly for both stations. The overall results indicate the superiority of the results of the SVR model (Kerman: RMSE = 1.105 °C and R = 0.992) and (Bam: RMSE = 1.01 °C and R = 0.99). The results showed that the SVR model improved the simulation error rate compared to the neural network (ANN), ANFIS, genetic programming (GP) and multivariate regression (Reg) models in Kerman station about 32, 42, 30 and 11 percent respectively and 62, 59, 27 and 27 percent respectively in Bam station. The results of the root mean square error showed that among the six studied models, the support vector regression model and genetic planning for Bam station and the support vector regression model for Kerman station have higher accuracy. The results also showed that estimating the mean temperature of Bam station has more efficiency and accuracy than Kerman station. In this study, although the analysis of the output results of the models did not lead to the same results, but the results of the models indicate an increase in temperature variables in the two stations of Kerman and Bam in future periods.

Keywords: Climate change, General circulation model, Kerman, Prediction.

¹ - Associate Professor, Department of Water Engineering, University of Birjand, Birjand, Email: Hkhozeymeh@birjand.ac.ir

² - Ph.D Student of Water Resources Management, University of Birjand, Birjand, Email: maryam70safavi@gmail.com

³ - Assistant Professor, Department of Water Engineering, University of Birjand, Birjand, Email: m.amirabadi@birjand.ac.ir

⁴ - Ph.D Graduate of Water Resources Management, University of Birjand, Birjand, Email: m_nazeri2007@yahoo.com

Research Paper

Evaluation of data-driven models based on downscaling of daily temperature values

Hossein Khozaymehnezhad^{1*},Maryam Safavi²,Mehdi Amirabadizadeh³Mohammad Nazeri Tahroudi⁴

¹ Associate Professor, Department of Water Engineering, Faculty of agriculture, University of Birjand, Birjand, Iran

² Ph.D Student, Department of Water Engineering, Faculty of agriculture, University of Birjand, Birjand, Iran

³ Assistant Professor, Department of Water Engineering, Faculty of agriculture, University of Birjand, Birjand, Iran

⁴ Ph.D Graduate, Department of Water Engineering, Faculty of agriculture, University of Birjand, Birjand, Iran



10.22125/IWE.2021.272470.1485

Received:

February 8, 2021

Accepted:

May 3, 2021

Available online:

June.01.2022**Keywords:****Climate change,
General circulation
model, Kerman,
Prediction****Abstract**

In this study, using six models of neural network (ANN), ANFIS, support vector machine (SVM), genetic programming (GP), support vector regression (SVR) and multivariate regression (Reg), the mean daily temperature at Kerman and Bam stations, Iran were studied and simulated during the period of 1961-2005. The results showed that the mean daily temperature during the mentioned periods will increase significantly for both stations. The overall results indicate the superiority of the results of the SVR model (Kerman: RMSE = 1.105 oC and R = 0.992) and (Bam: RMSE = 1.01 oC and R = 0.99). The results showed that the SVR model improved the simulation error rate compared to the neural network (ANN), ANFIS, genetic programming (GP) and multivariate regression (Reg) models in Kerman station about 32, 42, 30 and 11 percent respectively and 62, 59, 27 and 27 percent respectively in Bam station. The results of the root mean square error showed that among the six studied models, the support vector regression model and genetic planning for Bam station and the support vector regression model for Kerman station have higher accuracy. The results also showed that estimating the mean temperature of Bam station has more efficiency and accuracy than Kerman station. In this study, although the analysis of the output results of the models did not lead to the same results, but the results of the models indicate an increase in temperature variables in the two stations of Kerman and Bam in future periods.

1. Introduction

Predicting global climate change using the recorded values in the current statistical period requires an accurate method to be able to identify the fluctuations of these changes well. By modeling these changes,

* **Corresponding Author:** Hossein Khozaymehnezhad

Address: Department of Water Engineering, University of Birjand, Iran

Email: Hkhozaymeh@birjand.ac.ir

Tel: 09929974414

the desired parameter can be predicted for future years or periods. Due to climate change and global warming, simulation of mean temperature, which is one of the most important climate parameter, provides a good opportunity for planners to plan and make the necessary arrangements.

2. Materials and Methods

The purpose of this study is to investigate the accuracy of six data-driven models in downscaling and simulation the mean temperature values of Kerman and Bam synoptic stations in Iran in the statistical period of 1961-2005. In this regard, different data-driven models such as artificial neural network (ANN), ANFIS, genetic programming (GP), multivariate regression, support vector regression (SVR) and support vector machine (SVM) were used.

This study was performed using 26 predictive parameters resulting from the fifth IPCC report and was evaluated and compared. In all the mentioned methods, using Pearson correlation test, among 26 predictor parameters, the parameters that had a high correlation with the mean daily temperature values were selected for each region.

3. Results

To investigate the error values due to modeling, three methods of root mean square error (RMSE), correlation coefficient (R) and Nash-Sutcliffe efficiency coefficient were used. The results of accuracy and error rate of the models showed that among the intelligent models, the support vector machine model for Kerman station and the genetic programming model for Bam station had the lowest error and among the regression based models (multivariate regression and support vector regression (SVR) method) have the lowest error rate and the highest accuracy in simulating the mean daily temperature values of Kerman and Bam synoptic stations.

4. Discussion and Conclusion

The results of the error rate of the studied data showed that after the support vector regression model, the genetic programming model and the support vector machine model have high and acceptable accuracy. In general, the results of simulating the mean daily temperature values show that regression methods are more accurate than intelligent methods. The general results show the superiority of the results of SVR regression model (Kerman: RMSE = 1.411 °C and $R > 0.98$) and (Bam: RMSE = 1.01 °C and $R > 0.98$). Also, the results indicate the superiority of the results of Bam station with the mentioned models in comparison with Kerman station.

5. Six important references

1. Braddock, R.D., M.L, Kremmer and L. Sanzogni. 1998. Feedforward artificial neural network model for forecasting rainfall-runoff, *Journal of Environmental Sciences*, 9: 419-432.
2. Ferreira, C. 2001. Gene Expression Programming: A New Adaptive Algorithm for Solving Problems, *Complex Systems*, forthcoming.
3. IPCC.1995. In: Watson, R.T., Zinyowera, M.C., Moss, R.H. Eds., *Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses*, Cambridge University Press, UK p. 878.
4. IPCC. 2001. In: Watson, R.T., Zinyowera, M.C., Moss, R.H., Dokken, D.J., Eds., *Special Report on The Regional Impacts of Climate Change, An Assessment of Vulnerability*, Cambridge University Press, UK.
5. Khalili, K., M.N, Tahoudi., R, Mirabbasi and F, Ahmadi. 2016. Investigation of spatial and temporal variability of precipitation in Iran over the last half century, *stochastic environmental research and risk assessment*, 30(4): 1205-1221.
6. Raji, M., M. N, Tahroudi., F, Ye and J, Dutta. 2022. Prediction of heterogeneous Fenton process in treatment of melanoidin-containing wastewater using data-based models. *Journal of Environmental Management*, 307: 114518.

Conflict of Interest

Authors declared no conflict of interest.

Acknowledgments

The authors would like to acknowledge the financial support of University of Birjand for this research under contract number 1398/D/14817.