

صص ۱۹-۱

ارتقاء مدل سازی هیبرید با استفاده از مدلی کارآمد جهت پیش بینی بارندگی

لاله پرویز*

دانشیار دانشکده کشاورزی دانشگاه شهید مدنی آذربایجان، تبریز، ایران

تاریخ دریافت: ۱۴۰۰/۵/۳۱ تاریخ پذیرش: ۱۴۰۰/۸/۴

چکیده

پیش بینی بارندگی بادقت کم منجر به ضررهای قابل توجه در بخش های مختلف مانند کشاورزی، محیط زیست می شود. در این راستا تأثیر مدل های رگرسیون بردار پشتیبان (SVR)، برنامه ریزی بیان ژن (GEP) و روش گروهی مدل سازی داده ها (GMDH) در ارتقاء عملکرد مدل هیبرید مورد بررسی قرار گرفت که از داده های بارندگی ایستگاه های ارومیه و اصفهان با دو اقلیم متفاوت در بازه زمانی ۱۹۶۴-۲۰۱۹ استفاده شد. در مدل سازی بخش غیرخطی، ترکیب سوم با ترکیب بخش خطی، باقی مانده ها و داده های مشاهداتی در گام زمانی پیشین خطای کمتری داشت، به عنوان نمونه در ایستگاه اصفهان میزان کاهش RMSE از ترکیب ۱ به ۶۲/۳،۷۳ و میزان کاهش SMAPE از ترکیب ۲ به ۳ برابر با ۶۲/۷۹ درصد بود. مدل هیبرید نسبت به مدل استوکستیک دارای عملکرد بهتری بود، به طوری که مقدار RMSE از مدل استوکستیک به مدل هیبرید با SVR، GEP و GMDH در ایستگاه ارومیه به ترتیب ۷۹/۴۶، ۶۸/۳۴ و ۷۵/۷۷ درصد کاهش داشت. مدل برنامه ریزی بیان ژن نسبت به سایر مدل های مورد مطالعه دقت کمتری داشت (در ایستگاه ارومیه میزان کاهش UII از مدل GEP به SVR به ترتیب برابر با ۳۲/۵ و ۱۵/۶۲ درصد و در ایستگاه اصفهان میزان افزایش ضریب نش-ساتکلیف از مدل GEP به GMDH برابر با ۲۲/۳۸ بود). میزان ضریب نش-ساتکلیف در هر سه مدل در ایستگاه ارومیه بیشتر از اصفهان بود (متوسط میزان کاهش ضریب نش-ساتکلیف از ایستگاه ارومیه به اصفهان ۶/۲۲ درصد بود) ولی مقدار ضریب در هر دو ایستگاه در محدوده قابل قبول است؛ بنابراین انتخاب مدل کارا با ترکیب درست در مدل سازی بخش غیرخطی تأثیر چشمگیری در افزایش کارایی مدل هیبرید خواهد داشت.

واژگان کلیدی: بارندگی، هیبرید، ترکیب، غیرخطی.

مقدمه

بارندگی از اساسی ترین پدیده های اتمسفری است که بر فعالیت های انسانی در همه مقیاس ها تأثیر می گذارد مانند رفت و آمد روزمره، ساخت و سازهای شهری، کنترل سیل، آبیاری، فرآیندهای علوم زمینی و مدیریت نیروی برق آبی. گرم شدن کره زمین به دلیل افزایش اثر گازهای گلخانه ای بر تغییرات روند بارش مؤثر است؛ بنابراین داشتن اطلاعات مربوط به

بارش آینده در اجرای برنامه‌ریزی موفقیت‌آمیز نقش مهمی دارد (Ponnoprat, 2021). در این راستا استفاده از مدل‌های کارآمد که بتوانند ماهیت پیچیده بارش را شبیه‌سازی کنند، از اهمیت چشمگیری برخوردار هستند. پیش‌بینی بارندگی به‌عنوان یک مسئله دشوار در بین جامعه مهندسين هیدرولوژی که با مدل‌سازی بارش سر و کار دارند، شناخته شده است (Kalteh, 2017). ساختارهای گوناگونی از بارندگی توسط مدل‌های مختلف شبیه‌سازی شده‌اند مانند استفاده از مدل‌های استوکستیک در بخش خطی، مدل‌های هوش مصنوعی در بخش غیرخطی (Papalaskaris and Panagiotidis, 2017-Mislan et al, 2015). با توجه به این که ترکیب مدل‌ها می‌تواند دقت پیش‌بینی را افزایش دهد، مدل هیبرید معرفی شد که در واقع ترکیب مدل‌های مختلف برای بخش خطی و غیرخطی است. سیستم‌های هیبرید که مدل‌های کلاسیک و هوش مصنوعی را ترکیب می‌کنند، توجه زیادی را به خود جلب کرده‌اند که دلیل آن عملکرد بهتر مدل‌های هیبرید نسبت به مدل‌های انفرادی می‌باشد. تا به امروز مدل‌سازی هیبرید متداول‌ترین روش برای ساخت سیستم‌های پیش‌بینی ترکیبی بوده است. این روش هر یک از اجزاء را به‌صورت جداگانه در یک فرآیند متوالی مدل می‌کند (Chen et al, 2021). اولین سیستم ترکیبی که به‌عنوان سیستم هیبرید معرفی شد شامل مدل ARIMA و ANN به ترتیب برای مدل‌سازی بخش خطی و غیرخطی بود (Zhang, 2003). سپس بسیاری از مطالعات، رویکرد هیبرید را برای حل مشکلات مختلف پیش‌بینی عملی بکار گرفتند، به‌عنوان مثال تلفیق مدل‌های خود رگرسیون خطی تعمیم یافته (GLAR) و شبکه عصبی برای پیش‌بینی نرخ ارز خارجی (Yu et al, 2005)، ترکیب مدل‌های ARIMA و حداقل مربعات ماشین بردار پشتیبان (LSSVM) برای پیش‌بینی قیمت کربن (Zhu and Chevallier, 2017)، ترکیب یک شبکه باور عمیق (DBN)، مدل ARIMA و بهینه‌سازی ازدحام ذرات (PSO) برای پیش‌بینی جزر و مد (Qin et al, 2017). در تمام موارد افزایش عملکرد مدل‌های هیبرید نسبت به مدل‌های انفرادی قابل توجه بوده است، به‌عنوان نمونه میزان خطای MAE در پیش‌بینی حجم بازرسی در مدل SARIMA با مقدار ۱۲/۷۲ و شبکه عصبی مصنوعی با مقدار ۱۴/۰۲ به مقدار ۱۲/۱۹ در مدل هیبرید رسید. در ترکیب ARIMA و شبکه عصبی مصنوعی، مقادیر MSE برای مدل ANN، ARIMA و مدل هیبرید در گام زمانی ۱ ماهه به ترتیب برابر با ۳/۶۸، ۲/۷۶ و ۲/۶۷ و در گام زمانی ۶ ماهه به ترتیب برابر با ۵/۶۵، ۵/۷۱ و ۵/۶۵ بود (Ruiz-Aguilar et al, 2014). در سال‌های اخیر از روش جدید در قسمت خطای سری زمانی در مدل‌های هیبرید استفاده شده است که منجر به افزایش عملکرد سیستم‌های هیبرید شد (Chen et al, 2021)؛ بنابراین استفاده از مدل هوش مصنوعی با کارایی بالا در پیش‌بینی بخش غیرخطی دقت مدل هیبرید را بالا خواهد برد. مدل‌هایی مانند رگرسیون بردار پشتیبان، برنامه‌ریزی بیان ژن و روش گروهی مدل‌سازی داده‌ها جهت پیش‌بینی مسائل هیدرولوژیکی استفاده شده‌اند. با استفاده از مدل برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان به پیش‌بینی بارش ماهانه ایستگاه ورابه در شهرستان نهاوند پرداخته شد. داده‌های مورد استفاده شامل بارش، دما و رطوبت نسبی بودند. با توجه به ضریب همبستگی ۰/۹۲، کارایی دو مدل تقریباً مشابه بود (سلگی و همکاران، ۱۳۹۷). با دو مدل رگرسیون بردار پشتیبان و برنامه‌ریزی بیان ژن به پیش‌بینی بارندگی ایستگاه شیراز پرداخته شد. در این پژوهش از

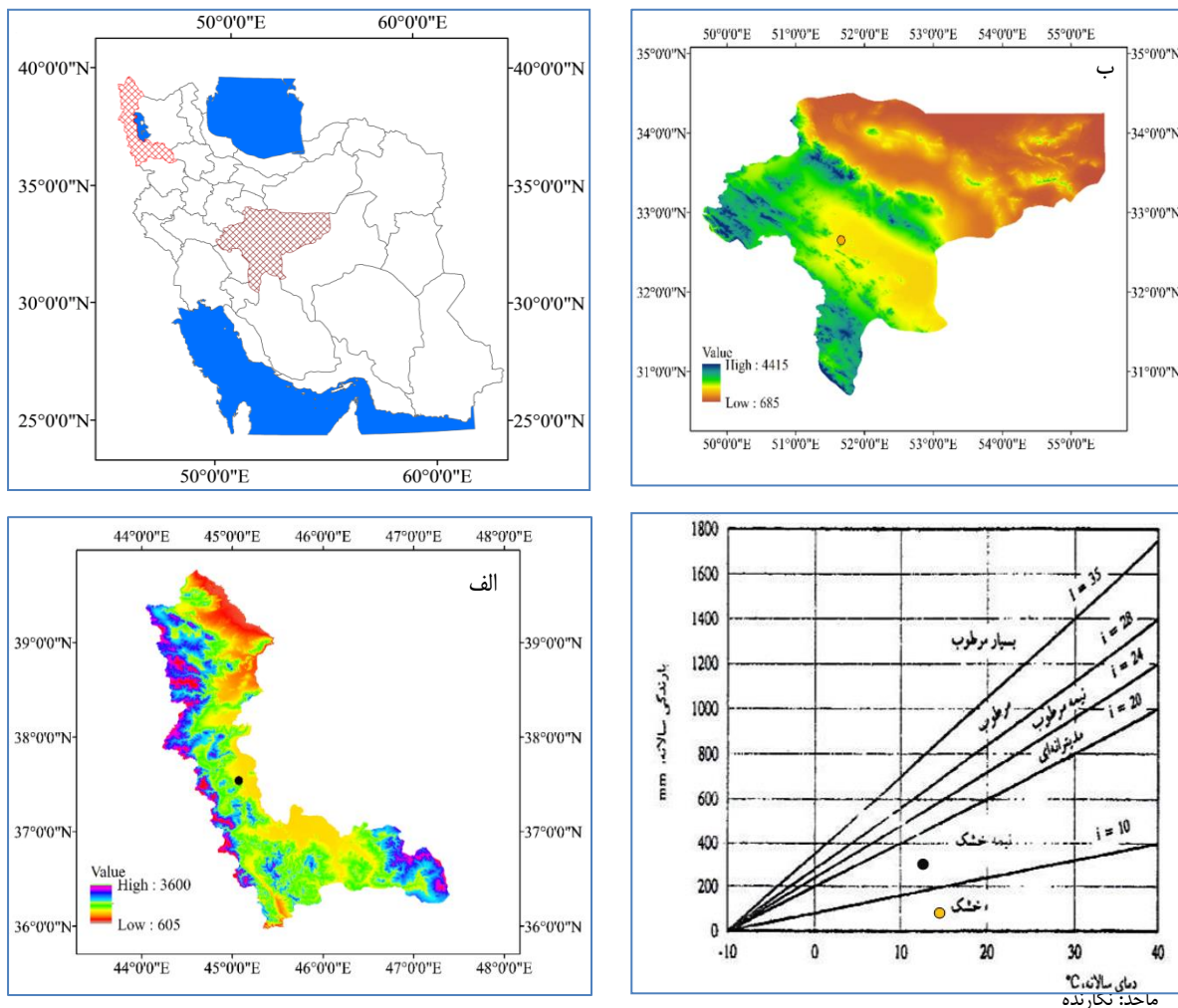
روش‌های آزمون گاما و آنالیز همبستگی جهت پیش‌پردازش داده‌های ورودی استفاده شد. نتایج بیانگر آن است که مدل ترکیبی گاما- رگرسیون بردار پشتیبان در پیش‌بینی بارندگی بهتر عمل کرده است (برجی حسن گاوپار و همکاران، ۱۳۹۶). از روش گروهی مدل‌سازی داده‌ها و شبکه عصبی مصنوعی جهت مدل‌سازی بارش-رواناب در حوضه آبخیز پلرود استفاده شد. روش گروهی مدل‌سازی داده‌ها نسبت به شبکه عصبی مصنوعی دارای کارایی بالایی بود (پورنعمت رودسری و همکاران، ۱۳۹۳). از مدل‌های برنامه‌ریزی ژنتیک و ماشین بردار پشتیبان جهت پیش‌بینی بارش ماهانه استان اردبیل استفاده شد که مدل رگرسیون بردار پشتیبان نتایج بهتری داشت (شرفی و همکاران ۱۳۹۹).

بنابراین کارایی مدل هیبرید در توسعه پیش‌بینی سری زمانی به اثبات رسیده است، در این میان افزایش کارایی مدل هیبرید گزینه مناسبی در افزایش دقت پیش‌بینی‌ها خواهد بود، بخصوص در مورد سری زمانی بارندگی با ماهیت پیچیده آن ضروری به نظر می‌رسد. هدف این پژوهش ارتقاء بخش مدل‌سازی غیرخطی مدل هیبرید جهت پیش‌بینی بارندگی است. برای این منظور ترکیب مدل استوکستیک با مدل‌های رگرسیون بردار پشتیبان، برنامه‌ریزی بیان ژن و روش گروهی مدل‌سازی داده‌ها مورد ارزیابی قرار گرفت تا اثرات تنوع اقلیمی در روند پیش‌بینی هم مورد بررسی قرار گرفت. همچنین تحلیل حساسیت جهت تخمین پارامترهای مدل‌های مورد بررسی از گام‌های اساسی در مدل‌سازی بود.

داده‌ها و روش‌ها

منطقه مورد مطالعه

داده‌های ایستگاه‌های اصفهان و ارومیه جهت بررسی تأثیر مدل‌های مختلف بر بخش غیرخطی مدل هیبرید مورد استفاده قرار گرفت که موقعیت ایستگاه‌ها در شکل ۱ آورده شده است. دوره آماری مورد بررسی بازه زمانی ۱۹۶۴-۲۰۱۹ بود. بر اساس اقلیم‌نمای دوماتن، اقلیم ایستگاه ارومیه نیمه‌خشک و اصفهان خشک می‌باشد. انتخاب ایستگاه‌ها طوری صورت گرفته است که تأثیرات تفاوت اقلیمی ایستگاه‌ها نیز در نظر گرفته شود. با توجه به تغییرات اقلیمی اخیر، پیش‌بینی بارندگی با روشی کارآمد از اهمیت بالایی برخوردار است.

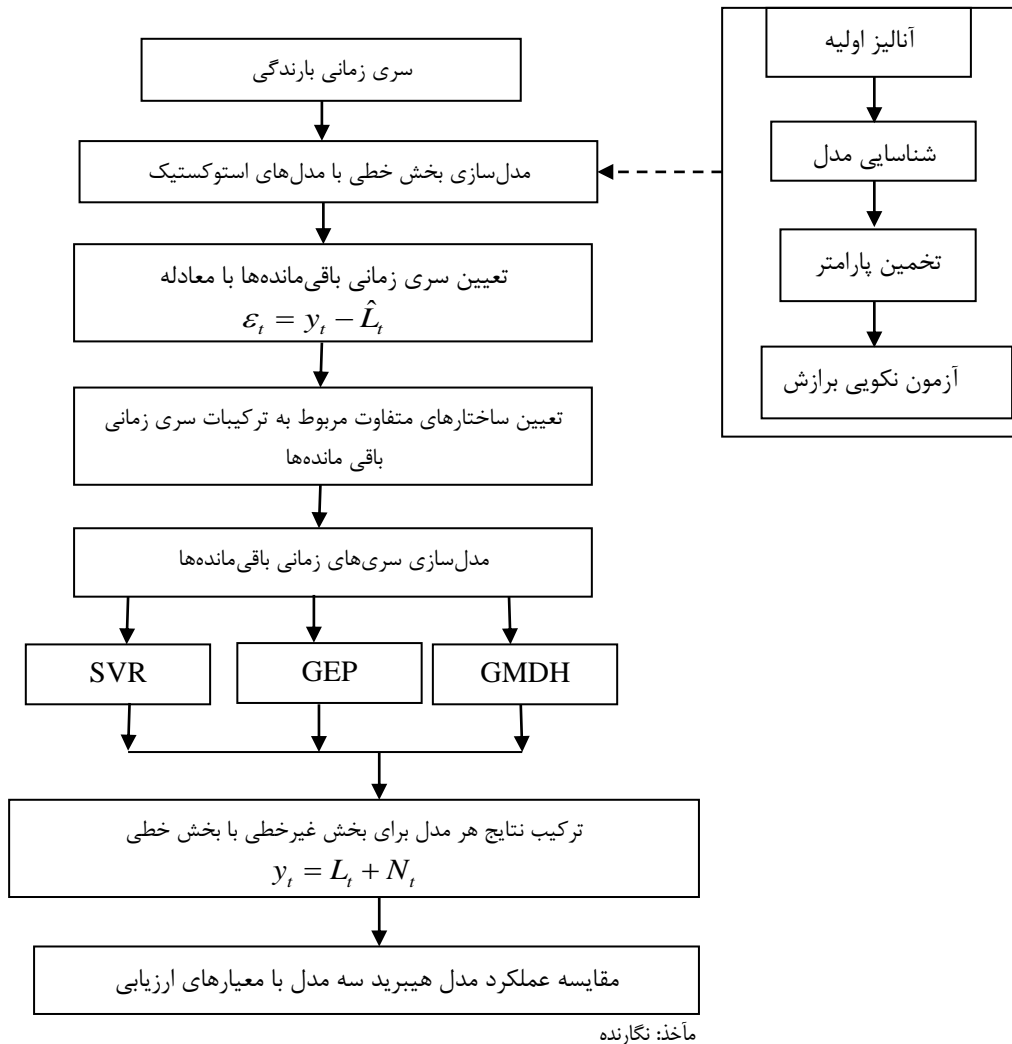


شکل ۱: موقعیت مکانی ایستگاه های ارومیه (الف)، اصفهان (ب) همراه با اقلیم نمای دومازن ایستگاه ها (ج)

روش پژوهش

مدل هیبرید

سیستم هیبرید ترکیبی از مدل های استوکستیک مانند ARIMASARIMA و مدل های هوش مصنوعی مانند شبکه عصبی مصنوعی می باشد که به ترتیب برای مدل سازی بخش خطی و غیرخطی سری های زمانی مورد استفاده قرار می گیرد. در این پژوهش از مدل های رگرسیون بردار پشتیبان، برنامه ریزی بیان ژن و روش گروهی مدل سازی داده ها جهت مدل سازی بخش غیرخطی مدل هیبرید استفاده شد. روند نمای پژوهش در شکل ۲ آورده شده است.



شکل ۲: روندنمای انجام پژوهش جهت ارتقاء عملکرد مدل هیبرید

رگرسیون بردار پشتیبان (SVR¹)

ماشین بردار پشتیبان یک روش طبقه بندی است که از تئوری یادگیری آماری برگرفته شده است. در مسائل رگرسیون یادگیری یک تابع غیرخطی بر اساس یک ماشین یادگیری خطی (تابع کرنل) است، در حالی که ظرفیت سیستم توسط پارامتری کنترل می شود که به بعد فضا بستگی ندارد. با در نظر گرفتن داده های بخش آموزشی که x و y به عنوان داده های ورودی و خروجی می باشند $[(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)]$ ، اگر تابع رگرسیون به صورت $f(x) = w^T \cdot \Phi(x) + b$ بیان شود، مسئله رگرسیون بردار پشتیبان بر اساس مسئله بهینه سازی (رابطه ۱) فرمول بندی می شود.

¹Support Vector Regression

$$\min \frac{1}{2} w^T \cdot w + C \sum_{i=1}^l (\xi_i + \xi_i^*)$$

$$\text{subject to } y_i - [w^T \cdot \Phi(x_i) + b] \leq \varepsilon + \xi_i$$

$$[w^T \cdot \Phi(x_i) + b] - y_i \leq \varepsilon + \xi_i^*$$

ξ_i و ξ_i^* متغیرهای کمکی، (ε) یک ثابت مثبت که درجه تلفات جریمه‌ای را هنگام بروز خطای آموزشی تعیین می‌کند (Chen et al, 2010-Acosta et al. 2021).

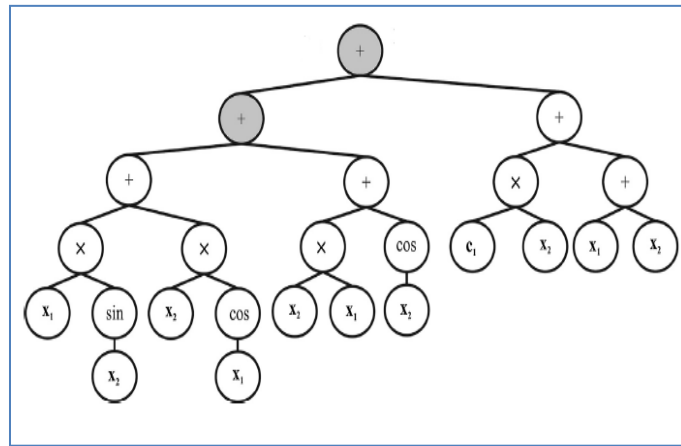
برنامه‌ریزی بیان ژن (GEP¹)

برنامه‌ریزی بیان ژن یک روش الگوریتم ژنتیک پیشرفته است. این مدل بر اساس نظریه تکامل داروین استوار است. هر کروموزوم شامل یک یا چند ژن است که به‌عنوان زیر درخت بیان می‌شود. در برنامه‌ریزی بیان ژن بهترین راه حل با پیوند دادن زیردرخت‌ها با توابع جبری یا بولی (OR-NOT-AND) بدست می‌آید، به‌عنوان نمونه کروموزوم برنامه‌ریزی بیان ژن در شکل ۳ نشان داده شده است که با یک مدل که ترکیب سه زیردرخت با تابع پیوند مثبت می‌باشد را نشان می‌دهد. فرم ریاضی مدل به‌صورت رابطه ۲ است (Danandeh Mehr, 2018).

$$(x_1 \times \sin x_2 + x_2 \times \cos x_1) + (x_1 \times x_2 + \log x_2) + (c_1 \times x_2 + x_1 + x_2)$$

(۲)

x_1 و x_2 بردارهای متغیرهای ورودی و تابع مثبت تابع پیوند است.



ماخذ: نگارنده

شکل ۳. نمونه‌ای از برنامه‌ریزی بیان ژن با سه زیر درخت (Danandeh Mehr, 2018)

شروع این الگوریتم تولید جمعیت اولیه از راه‌حل‌ها است که این امر با فرآیند تصادفی و با استفاده از اطلاعات ورودی در مورد مسئله انجام‌پذیر است. سپس قالب بیان درختی جهت نمایش کروموزوم‌ها مورد استفاده قرار می‌گیرد و ارزیابی بر

¹Gene Expression Programming

اساس تابع برازش انجام می‌گیرد. در صورت تأمین شرایط همگرایی (به‌عنوان نمونه رسیدن به تعداد نسل معین) تکامل متوقف می‌شود. در صورت عدم تأمین شرایط همگرایی، نخبه‌گرایی انجام خواهد گرفت؛ بنابراین مراحل انجام الگوریتم شامل الف: انتخاب مجموعه ترمینال ب: انتخاب مجموعه توابع ج: شاخص اندازه‌گیری دقت مدل د: مولفه‌های کنترل (سلگی و همکاران، ۱۳۹۶).

روش گروهی مدل‌سازی داده‌ها (GMDH¹)

این روش بر اساس مفاهیم خودسازمانده فراکوشی است همچنین بر اساس تکنیک ماشین یادگیری و محاسبات تکاملی می‌باشد. در این روش برای یک بردار ورودی داده شده تابع f طوری بدست می‌آید که به مقدار واقعی نزدیک باشد. بدین منظور روش گروهی مدل‌سازی داده‌ها رابطه‌ای بین ورودی و خروجی را ایجاد می‌کند. توصیف رابطه بین ورودی و خروجی معمولاً در فرم گسسته به صورت چند جمله‌ای کلموگروف-گابور (K-G) بیان می‌شود (رابطه ۳).

$$\hat{y} = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_{ijk} x_i x_j x_k + \dots \quad (3)$$

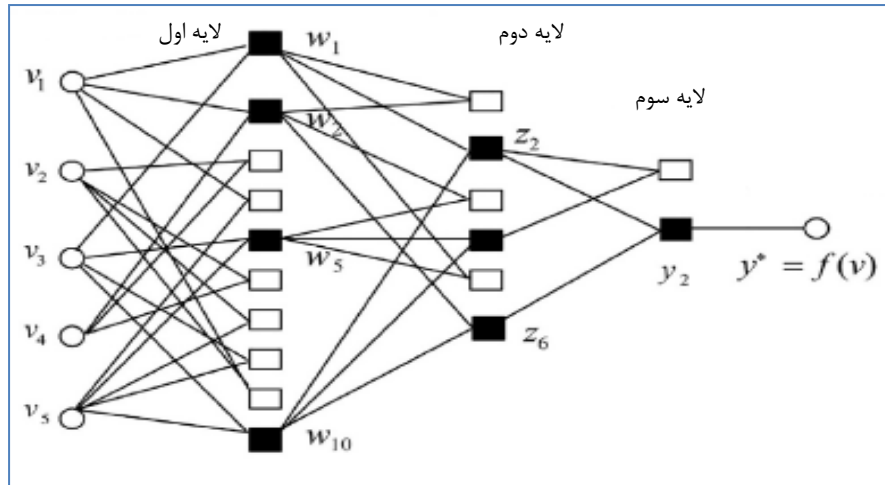
Y: خروجی، a: وزن‌های مربوطه (Razzaghi et al. 2021).

به‌عنوان نمونه چند جمله‌ای K-G مرتبه اول از درجه ۱ شامل n نرون به صورت رابطه ۴ است.

$$\hat{y} = f(x_1, x_2, \dots, x_n) = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n \quad (4)$$

روش گروهی مدل‌سازی داده‌ها دارای ساختار چند لایه‌ای است. اگر معادله ۴ انتخاب شود، زیر مجموعه‌های معادله به صورت Π مدل اولیه از روش گروهی مدل‌سازی داده‌ها در نظر گرفته می‌شود: $v_1 = a_1 x_1, \dots, v_n = a_n x_n$ در مدل‌سازی، داده‌های آموزشی به صورت تصادفی به دو گروه تقسیم می‌شود (گروه A و B). این مسئله با مجموعه مدل اولیه که شامل $v_1 = a_1 x_1, \dots, v_n = a_n x_n$ است، شروع می‌شود. برآورد پارامترها با معیارهای داخلی اجرا می‌شود. مدل‌های کاندید در مجموعه A بدست می‌آیند و مدل‌های بدست آمده را با معیارهای خارجی در مجموعه B ارزیابی می‌کند و بهترین مدل کاندید برای ورود به لایه بعدی انتخاب می‌شود. این روند تا زمان رسیدن به معیار همگرایی ادامه پیدا می‌کند. با افزایش پیچیدگی مدل مقدار معیار خارجی در ابتدا افزایش و سپس کاهش می‌یابد (Xu et al, 2020). شکل ۴ فرآیند مدل‌سازی را نشان می‌دهد.

¹Group Method of Data Handling



مأخذ: نگارنده

شکل ۴: نمایش ساختار روش گروهی مدل‌سازی داده‌ها (et al, 2020Xu)

در شکل ۴ لایه اول مدل کاندید را تولید می‌کند. لایه دوم مدل‌های کاندید را بطور مکرر تولید و انتخاب می‌کند تا به مدل بهینه برسد.

معیارهای ارزیابی

جهت بررسی عملکرد مدل‌های برنامه‌ریزی بیان ژن، روش گروهی مدل‌سازی داده‌ها و رگرسیون بردار پشتیبان در مدل هیبرید از برخی معیارهای ارزیابی استفاده شد که معادلات معیارها در روابط ۵ تا ۱۰ آورده شده است.

$$SMAPE = \frac{1}{N} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{\frac{y_t + \hat{y}_t}{2}} \quad (5)$$

$$RMSE = \frac{1}{N} \sqrt{\sum_{i=1}^N (O_i - P_i)^2} \quad (6)$$

$$UII = \frac{\left[\sum_{i=1}^N (P_i - O_i)^2 \right]^{0.5}}{\left[\sum_{i=1}^N (O_i)^2 \right]^{0.5}} \quad (7)$$

$$RRMSE = \frac{RMSE}{O} \quad (8)$$

$$RPD = \frac{SD}{RMSE} \quad (9)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (10)$$

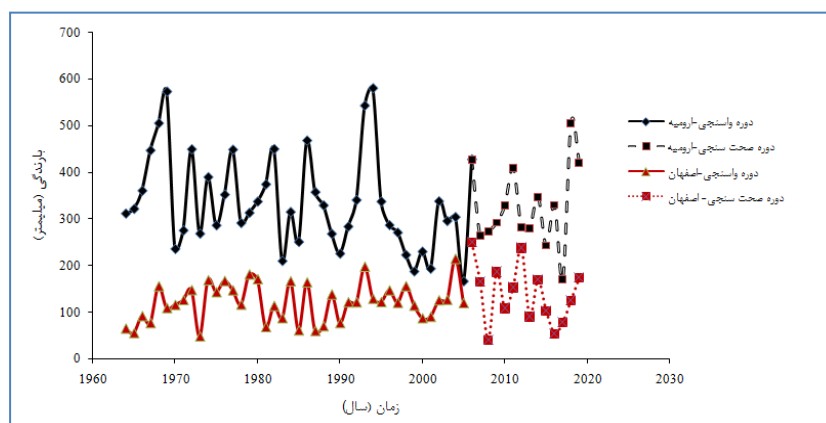
O: داده‌های مشاهداتی، P: داده‌های شبیه‌سازی شده، N: تعداد داده.

کمینه مقادیر SMAPE-UII-RMSE-RRMSE بیانگر عملکرد بهتر مدل است. مقادیر RPD بزرگ‌تر از ۱/۴ نشان‌دهنده عملکرد قابل قبول مدل است. مقادیر ضریب نش-ساتکلیف (NSE) نزدیک یک بیانگر عملکرد خوب مدل

است (Ruiz-Aguilar et al, 2014-Wang et al, 2020)

بحث و یافته‌ها

داده‌های مورد استفاده در این پژوهش مربوط به داده‌های سالانه بارندگی ایستگاه‌های ارومیه و اصفهان در طی بازه زمانی ۱۹۶۴-۲۰۱۹ است که دوره ۲۰۰۶-۲۰۱۹ به‌عنوان دوره صحت‌سنجی در نظر گرفته شد. از ضریب هرست برای آزمون کفایت طول داده‌ها استفاده شد. ضریب هرست برای سری زمانی بارندگی ارومیه و اصفهان به ترتیب ۰/۷۲ و ۰/۶۴ برآورد شد که چون از ۰/۵ بیشتر است، بیانگر آن است که حافظه طولانی مدت در سری زمانی بیشتر است و نیازی به تطویل داده‌ها در سری زمانی نمی‌باشد. نمودار سری‌های زمانی بارندگی در شکل ۵ و خصوصیات آماری سری‌ها در جدول ۱ آورده شده است.



مأخذ: نگارنده

شکل ۵: نمودار سری زمانی بارندگی ایستگاه‌های ارومیه و اصفهان در دو دوره واسنجی و صحت‌سنجی

جدول ۱: خصوصیات آماری سری زمانی بارندگی در دو دوره واسنجی و صحت‌سنجی در دو ایستگاه مطالعاتی

| دوره | دوره | دوره | دوره واسنجی | خصوصیات آماری |
|----------|----------|--------|-------------|-------------------|
| صحت‌سنجی | صحت‌سنجی | واسنجی | اصفهان | |
| ارومیه | اصفهان | ارومیه | | |
| ۳۲۶/۷۸ | ۱۳۸/۵۲ | ۳۳۳/۴۴ | ۱۲۲/۰۸ | میانگین (mm) |
| ۵۰۶ | ۲۵۰ | ۵۷۹/۶۱ | ۲۱۵/۷ | بیشینه (mm) |
| ۱۶۹/۹۱ | ۴۰/۸ | ۱۶۷/۲۴ | ۴۹ | کمینه (mm) |
| ۳۱۰/۶۵ | ۱۳۹/۴۶ | ۳۱۴/۴ | ۱۲۲/۵ | میانه (mm) |
| ۸۸/۶۲ | ۶۳/۷ | ۱۰۲/۳۷ | ۴۰/۸۴ | انحراف معیار (mm) |
| ۰/۲۷ | ۰/۴۵ | ۰/۳ | ۰/۳۳ | ضریب تغییرات |
| ۰/۳۹ | ۰/۲۲ | ۰/۷۸ | ۰/۱ | ضریب چولگی |

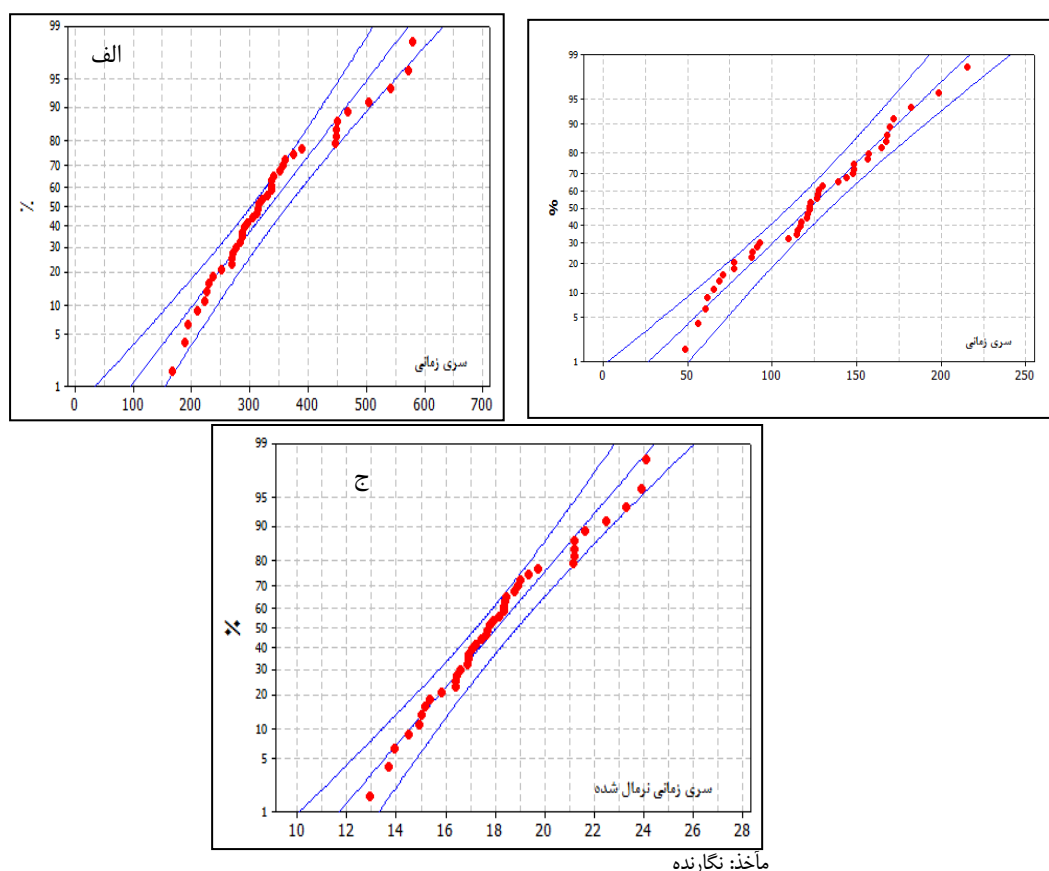
مأخذ: نگارنده

با توجه به مقادیر به شکل ۵ نمودار سری زمانی اصفهان پایین‌تر از سری زمانی ارومیه است و مقادیر آماری محاسبه شده در جدول ۱ مانند میانگین و کمینه مقدار نیز از این روند تبعیت می‌کنند. مقادیر کمینه، میانه و ضریب تغییرات در دوره‌های واسنجی و صحت‌سنجی ایستگاه ارومیه دارای تشابه زیادی هستند. در این پژوهش جهت مدل‌سازی بارش از

مدل هیبرید استفاده شد. ساختار مدل هیبرید در دو قسمت خلاصه می‌شود ۱- بخش خطی برای برآورد روند خطی سری زمانی ۲- بخش غیرخطی جهت مدل‌سازی سری زمانی خطای مدل.

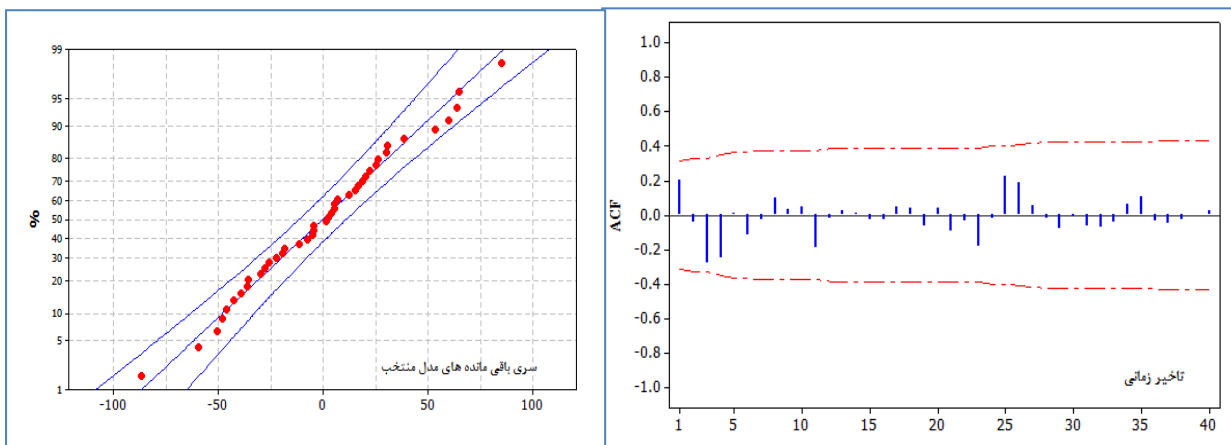
مدل‌سازی بخش خطی مدل هیبرید با مدل استوکستیک

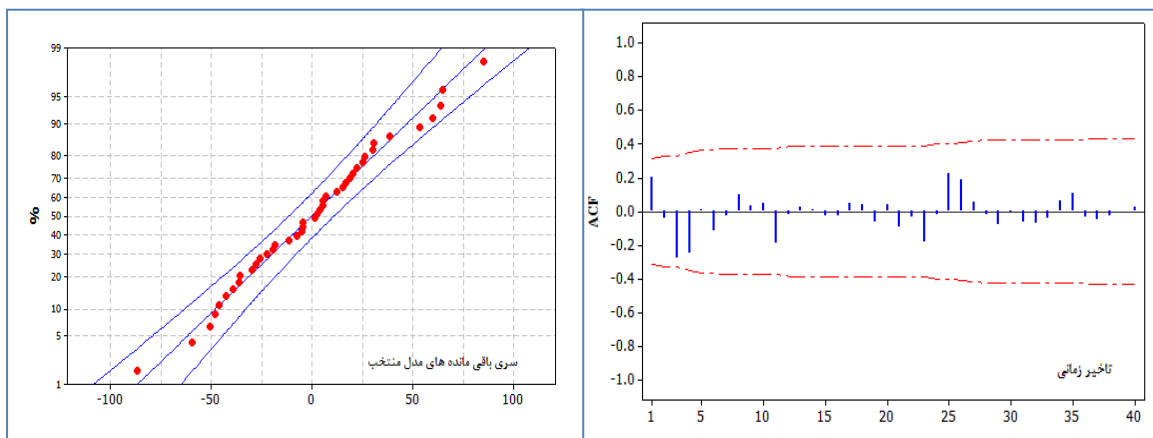
مدل‌سازی استوکستیک بر اساس متدولوژی باکس-جنکینز استوار است. اولین مرحله از آنالیز اولیه در مدل‌سازی استوکستیک مربوط به بررسی نرمال بودن سری زمانی است که در این پژوهش از سه روش برای بررسی این موضوع استفاده شد: ۱. رسم نمودار احتمالاتی توزیع نرمال ۲. آزمون کولموگروف-اسمیرنوف ۳. آزمون شاپیرو-ویلک. رسم نمودار احتمالاتی در شکل ۶ حاکی از نرمال نبودن سری زمانی بارندگی ارومیه (P-Value برابر با ۰/۰۲۴) و نرمال بودن سری زمانی اصفهان (P-Value برابر با ۰/۴۳) است. در نمودار احتمالاتی ایستگاه ارومیه برخی نقاط در خارج از محدوده‌های اطمینان (۹۵٪) هستند که حاکی از نرمال نبودن سری زمانی است. مقادیر مربوط سطح معنی داری در آزمون کولموگروف-اسمیرنوف و شاپیرو-ویلک در ایستگاه ارومیه به ترتیب ۰/۰۴۳ و ۰/۰۲۸ و در ایستگاه اصفهان به ترتیب برابر با ۰/۲ و ۰/۴۸ تخمین زده شدند و نتایج دو آزمون با نمودار احتمالاتی همخوانی دارد؛ بنابراین نیاز به نرمال سازی سری زمانی ارومیه است که در این پژوهش از روش ریشه دوم استفاده شد.



شکل ۶: نمودار احتمال سری زمانی بارندگی ارومیه (الف)، اصفهان (ب) و سری زمانی بارندگی نرمال شده ایستگاه ارومیه (ج)

با توجه به شکل ۶ - ج تمام نقاط در سری زمانی نرمال شده در محدوده اطمینان قرار گرفته‌اند و مقدار P-Value برابر با ۰/۲۲۹ برآورد شد که بیانگر نرمال بودن سری زمانی است. مقادیر آزمون‌های کولموگروف-اسمیرنوف و شاپرو-ویلک به ترتیب ۰/۲ و ۰/۲۸ تخمین زده شدند که با نتایج نمودار احتمالاتی همخوانی دارند. بررسی دیگر که مربوط به آنالیز اولیه در مدل‌سازی استوکستیک است، بررسی ایستایی و نا ایستایی سری زمانی است که بدین منظور از آزمون ADF استفاده شد. مقادیر آماره t_α برای سری زمانی ارومیه برابر با ۰/۸۵- برآورد شد که مقادیر بحرانی در سطح اطمینان ۰/۱، ۰/۵ و ۰/۱۰ برابر با ۳/۶-، ۲/۹۵- و ۲/۶۱- بودند که حاکی از عدم رد فرض صفر است و بیانگر سری نا ایستا است. مقادیر آماره t_α برای سری زمانی اصفهان برابر با ۶/۶۳- و مقادیر بحرانی در سطح اطمینان ۰/۱، ۰/۵ و ۰/۱۰ برابر با ۳/۶، ۲/۹- و ۲/۶ بودند که بیانگر سری ایستا است. در این حالت آماره آزمون من- کندال نیز محاسبه شد که مقدار آن برای ایستگاه‌های ارومیه و اصفهان به ترتیب برابر با ۲/۲۷- و ۱/۱ بود و که روند مربوط به ایستگاه ارومیه معنی‌دار است؛ بنابراین یک‌مرتبه تفاضل بر روی سری زمانی بارندگی ارومیه اعمال شد و با آزمون ADF ایستایی سری دوباره بررسی شد که در این حالت مقادیر آماره t_α برابر با ۸/۳۱- بود که از مقادیر بحرانی کوچک‌تر است، بنابراین سری ایستا است. مرحله بعد مربوط به تعیین مرتبه‌های مدل ARIMA است که در این حالت مرتبه‌های مختلف به مدل داده شد (مرتبه‌های مورد بررسی $p=0, 1, 2, 3, 4, 5, 6$ - $q=0, 1, 2, 3, 4, 5, 6$) و در ایستگاه ارومیه مدلی با مرتبه‌های خود همبسته و میانگین متحرک به ترتیب برابر با ۰ و ۱، کمینه SBC را داشت (SBC=۸۱/۲۸) و در ایستگاه اصفهان مدلی با خود همبسته و میانگین متحرک به ترتیب برابر با ۱ و ۳ کمینه SBC را داشت (SBC=۳۱۸/۴۷). البته در مدل انتخاب شده باید معنی‌داری پارامترها، نرمال و مستقل بودن باقی‌مانده‌های مدل هم مورد بررسی قرار گیرد. از نسبت t معنی‌داری پارامترها در سطح اطمینان ۰/۹۵ بررسی شد که این مرحله نیز برای مدل منتخب تأیید شد، به‌عنوان نمونه مقادیر P-Value مربوط به پارامتر خود همبسته در اصفهان برابر با ۰/۰۱۹ بود. در مورد نرمال بودن باقی‌مانده‌های مدل منتخب از نمودار احتمالاتی در مورد استقلال باقی‌مانده‌های مدل از نمودار همبستگی نگار (ACF) استفاده شد که نتایج در شکل ۷ آورده شده است.





ماخذ: نگارنده

شکل ۷: نمودار احتمالاتی نرمال اصفهان (الف) و همبستگی نگار ارومیه (ب) مربوط به سری زمانی باقی‌مانده‌های مدل منتخب مقادیر P-Value شکل ۷-الف برابر با ۰/۹۳ است که حاکی از نرمال بودن سری زمانی است. همچنین در همبستگی نگار ترسیم‌شده چون ضرایب خودهمبستگی از حدود بیان‌شده تجاوز نکردند، سری زمانی نسبت به زمان مستقل خواهد بود.

مدل‌سازی بخش غیرخطی مدل هیبرید

در مدل‌سازی بخش غیرخطی سری زمانی خطا دو نکته باید مورد توجه قرار بگیرد:

الف- نوع ترکیب مورد استفاده: ترکیبات مورد استفاده در این پژوهش شامل:

$$e_t = f(e_{t-1}, e_{t-2}, e_{t-3}) + \varepsilon_t - 1$$

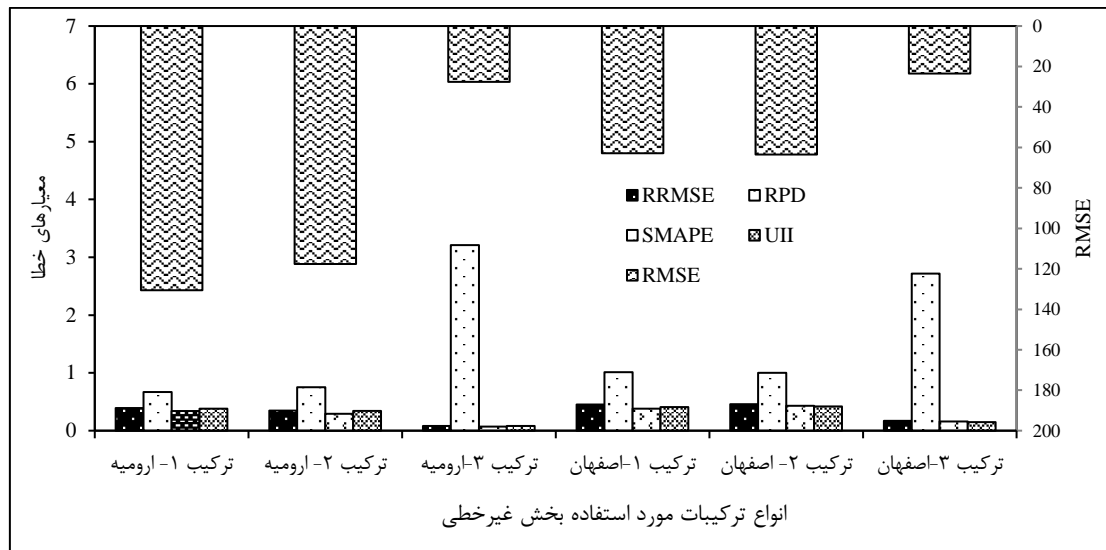
$$e_t = f(e_{t-1}, e_{t-2}) + \varepsilon_t - 2$$

$$y_t = f(y_{t-1}, y_{t-2}, \hat{L}_t, e_t, e_{t-1}) + \varepsilon_t - 3$$

ب- نوع مدل مورد استفاده: در این پژوهش مدل‌های GEP, GMDH و SVR استفاده شدند. جهت بررسی مورد اول

به تأثیر عملکرد ترکیبات مختلف با استفاده از مدل رگرسیون بردار پشتیبان پرداخته شد که نتایج در شکل ۸ آورده شده

است. نوع رگرسیون بردار پشتیبان مورد استفاده Nu SVR بود.



مأخذ: نگارنده

شکل ۸: بررسی ساختارهای متفاوت ترکیبات در بخش غیر خطی بر عملکرد مدل هیبرید

با توجه شکل ۸ در هر دو سری زمانی در ترکیب ۳ میزان $RMSE$ ، $RRMSE$ ، UII ، $SMAPE$ کاهش و مقدار RPD افزایش یافته است، به عنوان نمونه درصد کاهش $RMSE$ در ایستگاه ارومیه از ترکیب ۱ به ۲ برابر با $9/95$ و از ترکیب ۱ به ترکیب ۳ برابر با $78/88$ درصد و از ترکیب ۲ به ۳ برابر با $76/54$ بود. درصد کاهش $RRMSE$ از ترکیب ۱ به ۲ برابر با $10/25$ و از ترکیب ۱ به ترکیب ۳ برابر با $78/46$ و از ترکیب ۲ به ۳ برابر با 76 بود. درصد کاهش $SMAPE$ در ایستگاه اصفهان از ترکیب ۲ به ۱ برابر با $11/62$ و از ترکیب ۲ به ۳ برابر با $62/79$ و از ترکیب ۱ به ۳ برابر با $57/89$ بود. درصد کاهش UII از ترکیب ۲ به ۱ برابر با $2/38$ و از ترکیب ۲ به ۳ برابر با $64/28$ و از ترکیب ۱ به ۳ برابر با $63/41$ بود. مقادیر RPD در هر دو ایستگاه در ترکیب ۳ افزایش یافته است و در این ترکیب در محدوده قابل قبول از نظر عملکرد قرار گرفته است. مسئله قابل توجه در مورد عملکرد ترکیب ۱ و ۲ در دو ایستگاه است که مقایسه دو ترکیب حاکی از عملکرد بهتر ترکیب ۱ در ایستگاه اصفهان و ترکیب ۲ در ایستگاه ارومیه است. در بررسی معیارهای ارزیابی در ترکیب ۳، ارومیه نسبت به اصفهان وضعیت بهتری دارد. در مورد بخش دوم که مربوط به نوع مدل‌های منتخب است، تخمین پارامترهای مدل‌ها در افزایش کارایی مدل‌ها نقش بسزایی دارد، به عنوان نمونه در مدل‌سازی سری زمانی ارومیه در ترکیب ۲ با مدل رگرسیون بردار پشتیبان میزان کاهش $RMSE$ از تابع کرنل شعاعی با C برابر با ۲ به تابع تحریک خطی با C برابر با $34/71$ ، ۱ درصد است. مقادیر پارامترهای تخمینی هر مدل در جدول ۲ آورده شده است.

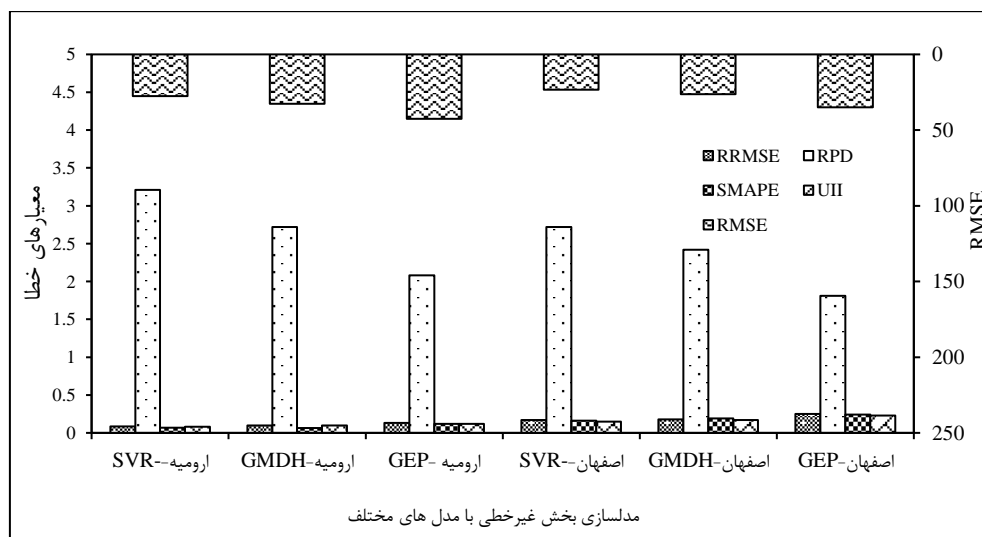
جدول ۲: پارامترهای تخمینی مدل‌های مورد استفاده در بخش غیرخطی مدل هیبرید- ترکیب ۳

| | | | | | | | |
|-----|-------------|---|----------------------|----------|------------------------------|--------|----------------------------|
| ۵ | تعداد ژن | ۸ | سر | ۲۰۰ | تعداد نسل | اصفهان | برنامه‌ریزی بیان ژن |
| | | | | | | ارومیه | |
| ۳ | تعداد ژن | ۷ | سر | ۵۰ | تعداد نسل | اصفهان | روش گروهی مدل‌سازی داده‌ها |
| | | | | | | ارومیه | |
| ۰/۵ | فشار انتخاب | ۲ | بیشینه مقدار لایه‌ها | ۹ | بیشینه مقدار نرون در یک لایه | اصفهان | رگرسیون بردار پشتیبان |
| | | | | | | ارومیه | |
| ۰/۲ | فشار انتخاب | ۵ | بیشینه مقدار لایه‌ها | ۱۰ | تابع کرنل | اصفهان | |
| | | | | | | ارومیه | |
| | | ۱ | C | سیگموئید | تابع کرنل | اصفهان | |
| | | ۲ | C | سیگموئید | تابع کرنل | ارومیه | |

مأخذ: نگارنده

برخی از پارامترهای مدل برنامه‌ریزی بیان ژن به این صورت بودند: تابع پیوند در ارومیه (+) در اصفهان (-)، نرخ جهش برابر با ۰/۰۶، نرخ ترکیب تک نقطه‌ای برابر با ۰/۲، نرخ ترکیب دو نقطه‌ای برابر با ۰/۳، نرخ ترکیب ژن برابر با ۰/۲. در مدل رگرسیون بردار پشتیبان تابع کرنل سیگموئید، بیشترین فراوانی را دارد. در روش گروهی مدل‌سازی داده‌ها، پارامتر فشار انتخاب پارامتر حساسی بود.

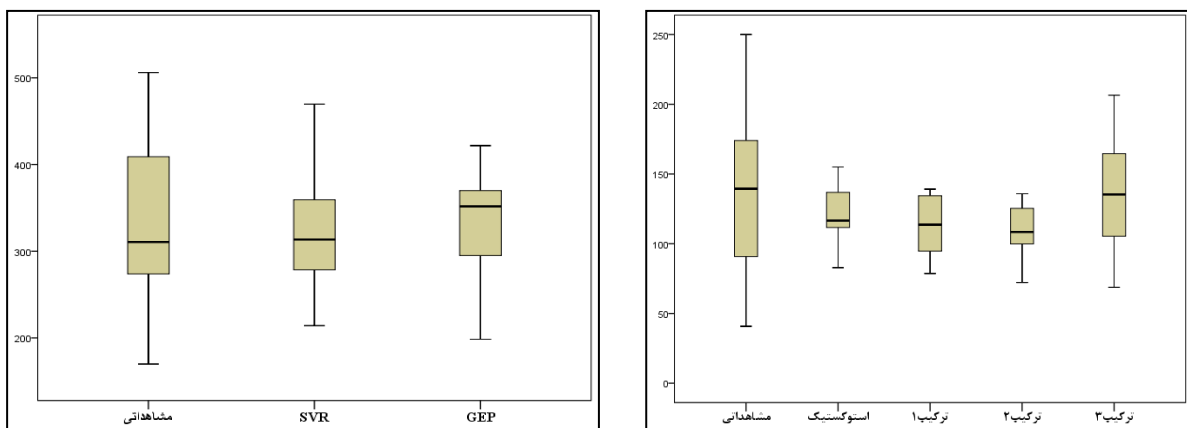
میزان کاهش RMSE از مدل استوکستیک به مدل هیبرید با رگرسیون بردار پشتیبان، برنامه‌ریزی بیان ژن و روش گروهی مدل‌سازی داده‌ها در ایستگاه ارومیه به ترتیب برابر با ۷۹/۴۶، ۶۸/۳۴ و ۷۵/۷۷ درصد و در ایستگاه اصفهان به ترتیب برابر با ۴۱/۶۰، ۰۶/۶۴ و ۵۵/۸۷ درصد بود و این مسائل نشان‌دهنده عملکرد مدل بهتر مدل هیبرید نسبت به مدل استوکستیک است. عملکرد قابل قبول مدل هیبرید در برابر مدل استوکستیک در تحقیقات (Chen and Wang, 2007)، (Ruiz-Aguilar et al, 2014) و (Mo et al, 2018) هم مشاهده شد. در بررسی دیگر به تأثیر عملکرد سه مدل برنامه‌ریزی بیان ژن، رگرسیون بردار پشتیبان و روش گروهی مدل‌سازی داده‌ها بر کارایی مدل هیبرید پرداخته شد که نتایج در شکل ۹ آورده شده است.



مأخذ: نگارنده

شکل ۹: بررسی تأثیر مدل‌های مختلف بر عملکرد مدل‌سازی هیبرید در ترکیب ۳

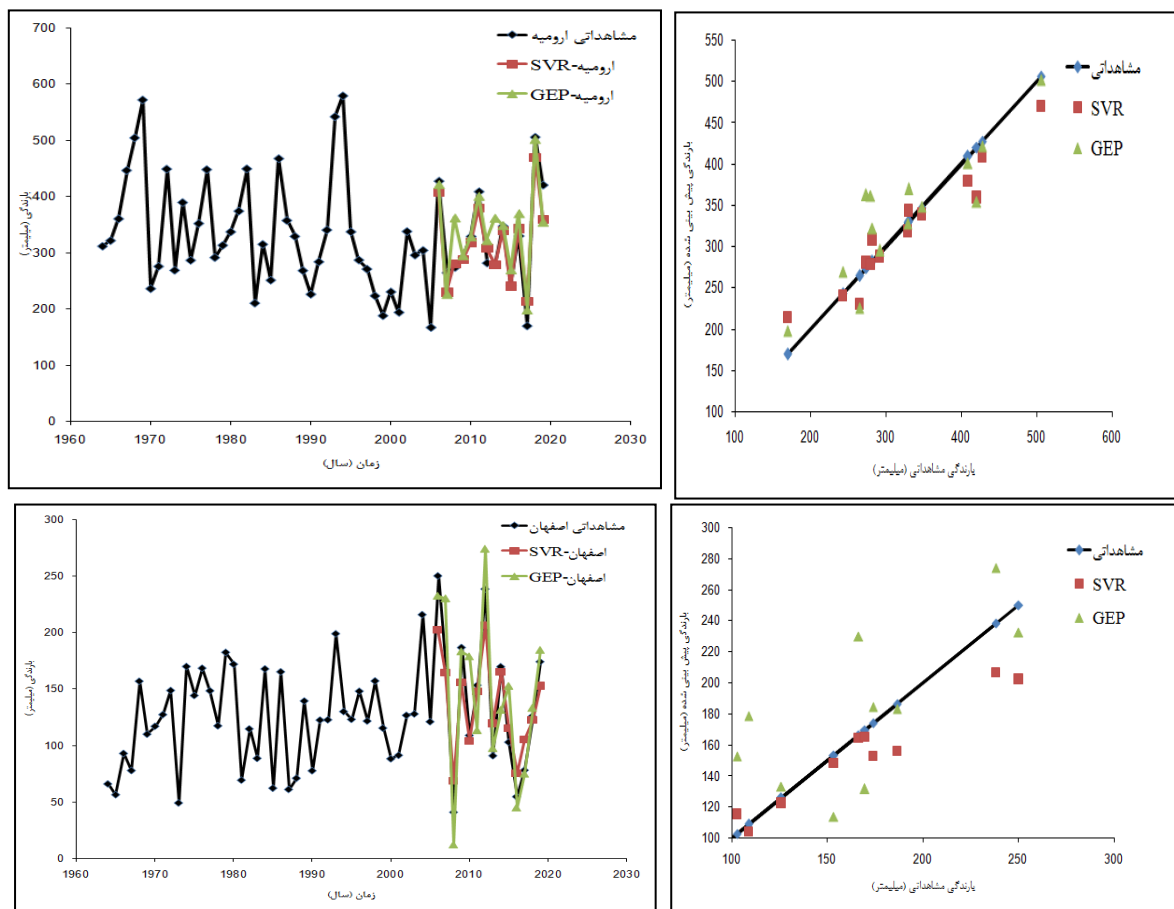
بر اساس شکل ۹ میزان کاهش آماره UII در ایستگاه ارومیه از مدل برنامه‌ریزی بیان ژن به روش گروهی مدل‌سازی داده‌ها برابر با ۲۰ درصد و از مدل برنامه‌ریزی بیان ژن به رگرسیون بردار پشتیبان برابر با ۳۲/۵ درصد بود. میزان کاهش $RMSE$ در این ایستگاه و در حالت بیان‌شده به ترتیب برابر با ۲۳/۸۸ و ۳۵/۱۲ درصد بود. میزان کاهش $RRMSE$ در ایستگاه اصفهان از مدل برنامه‌ریزی بیان ژن به روش گروهی مدل‌سازی داده‌ها برابر با ۲۸ درصد و از مدل برنامه‌ریزی بیان ژن به رگرسیون بردار پشتیبان برابر با ۳۲/۴ درصد بود. میزان افزایش RPD از مدل برنامه‌ریزی بیان ژن به رگرسیون بردار پشتیبان در ایستگاه‌های ارومیه و اصفهان به ترتیب ۵۴/۳۲ و ۵۰/۲۷ درصد بود. مدل برنامه‌ریزی بیان ژن نسبت به سایر مدل‌ها عملکرد ضعیفی داشت. مقادیر RPD در سه مدل در ایستگاه ارومیه نسبت به اصفهان در وضعیت بهتری است. میزان ضریب نش-ساتکلیف در مدل هیبرید با رگرسیون بردار پشتیبان، روش گروهی مدل‌سازی داده‌ها و برنامه‌ریزی بیان ژن در ایستگاه ارومیه به ترتیب برابر با ۰/۸۹، ۰/۸۵ و ۰/۷۵ و در ایستگاه اصفهان برابر با ۰/۸۵، ۰/۸۲ و ۰/۶۷ بود. بر اساس مقادیر ضریب نش-ساتکلیف اختلاف بین رگرسیون بردار پشتیبان و روش گروهی مدل‌سازی داده‌ها کم است. در پژوهش برجی حسن گاویار و همکاران (۱۳۹۶) در پیش‌بینی بارندگی، مدل رگرسیون بردار پشتیبان نسبت به برنامه‌ریزی بیان ژن عملکرد خوبی داشت. در تخمین تبخیر- تعرق مرجع، مدل رگرسیون بردار پشتیبان نسبت به مدل برنامه‌ریزی بیان ژن عملکرد بهتری داشت (Ahmadi et al, 2021). در مقایسه چندین مدل مانند حداقل مربعات رگرسیون بردار ماشین، برنامه‌ریزی بیان ژن و شبکه عصبی مصنوعی در تخمین بارندگی ماهانه، رگرسیون بردار پشتیبان عملکرد بهتری داشت (Mirabbasi et al, 2019). در ادامه جهت بیان تأثیر نوع ترکیب مورد استفاده و نوع مدل بکار رفته در مدل هیبرید بر نتایج پیش‌بینی بارندگی از نمودار جعبه‌ای (شکل ۱۰) استفاده شد.



مأخذ: نگارنده

شکل ۱۰: نمودار جعبه‌ای داده‌های مشاهداتی، استوکستیک و مدل هیبرید با انواع ترکیبات با مدل SVR در ایستگاه اصفهان (الف)، باکس پلات داده‌های مشاهداتی و مدل هیبرید با SVR و GEP در ایستگاه ارومیه (ب)

در نمودار جعبه‌ای ترسیم شده نکاتی قابل حصول است. در شکل ۱۰- الف تشابه بین نمودار جعبه‌ای پیش‌بینی شده با نمودار جعبه‌ای مشاهداتی از ترکیب ۱ به ترکیب ۳ افزایش می‌یابد که شامل تشابه خطوط Whikeres و میانه نمودار جعبه‌ای است و این مسئله در مورد معیارهای ارزیابی هم مشاهده شد. در شکل ۱۰- ب نیز خطوط Whikeres و میانه رگرسیون بردار پشتیبان نسبت به مدل برنامه‌ریزی بیان ژن به حالت مشاهداتی نزدیک‌تر است که این مسئله هم با نتایج معیارهای ارزیابی تطابق دارد. در ادامه نمودار بارندگی مشاهداتی - پیش‌بینی و دیاگرام پراکنش در شکل ۱۱ نشان داده شده است.



مأخذ: نگارنده

شکل ۱۱: نمودار بارندگی مشاهداتی و شبیه‌سازی به همراه دیاگرام پراکنش بارندگی مشاهداتی و پیش‌بینی در ایستگاه‌های ارومیه و اصفهان

بر اساس شکل ۱۱، همخوانی نمودار مدل رگرسیون بردار پشتیبان با نمودار داده‌های مشاهداتی بیشتر از برنامه‌ریزی بیان ژن است. ضریب همبستگی خط برازشی بر دیاگرام پراکنش از برنامه‌ریزی بیان ژن به رگرسیون بردار پشتیبان در ایستگاه‌های ارومیه و اصفهان به ترتیب ۲۰/۷۷ و ۲۳/۳۳ درصد افزایش داشته است. کمترین مقدار بارندگی در طی دوره صحت‌سنجی در ارومیه و اصفهان به ترتیب در سال‌های ۲۰۱۷ و ۲۰۰۸ اتفاق افتاده است که این روند در سه مدل حفظ

شده است. بیشترین مقدار بارندگی در طی دوره صحت‌سنجی در ارومیه و اصفهان به ترتیب در سال‌های ۲۰۱۸ و ۲۰۰۶ اتفاق افتاده است که این روند فقط در ایستگاه ارومیه حفظ شده است. اگر قدر مطلق اختلاف متوسط بارندگی بین داده‌های مشاهداتی و شبیه‌سازی محاسبه شود، نتایج به این صورت خواهد بود: در ایستگاه ارومیه رگرسیون بردار پشتیبان برابر با $۸/۳۵$ میلی‌متر، برنامه‌ریزی بیان ژن برابر با $۱۳/۳۷$ میلی‌متر و در ایستگاه اصفهان رگرسیون بردار پشتیبان برابر با $۲/۳۳$ میلی‌متر، برنامه‌ریزی بیان ژن برابر با $۷/۵۶$ میلی‌متر است، در این حالت هم برنامه‌ریزی بیان ژن عملکرد خوبی نداشته است. همچنین مجموع اختلاف بین داده‌های مشاهداتی و پیش‌بینی با رگرسیون بردار پشتیبان (مشاهداتی - پیش‌بینی) در هر دو ایستگاه دارای علامت مثبت بود که نشانگر کم برآورد در داده‌های پیش‌بینی شده است.

نتیجه‌گیری

پیش‌بینی بارندگی با مدلی کارآمد به دلیل ماهیت پیچیده آن از موضوعات چالشی است. در این راستا مدل هیبرید به دلیل ترکیب چند مدل می‌توانند دقت پیش‌بینی ساختار حاکم بر سری زمانی بارندگی را افزایش دهد. در این پژوهش مدل هیبرید نسبت به مدل استوکستیک دارای عملکرد بهتری بود؛ بنابراین مدل هیبرید به‌عنوان گزینه قابل قبول در پیش‌بینی بارش مطرح است که جهت نیل به مقادیر پیش‌بینی بارندگی با دقت بالا باید مدل هیبرید توسعه یابد. انتخاب نوع ترکیب در مدل‌سازی بخش غیرخطی نیز در افزایش کارایی مدل هیبرید نقش مهمی دارد. یکی از نکاتی که می‌تواند دقت مدل هیبرید را افزایش دهد، مدل‌سازی سری زمانی خطا با مدلی قوی است که در این پژوهش از مدل‌های رگرسیون بردار پشتیبان، برنامه‌ریزی بیان ژن و روش گروهی مدل‌سازی داده‌ها استفاده شد. مدل برنامه‌ریزی بیان ژن نسبت به دو نوع دیگر مدل از عملکرد ضعیفی برخوردار بود؛ بنابراین مدل رگرسیون بردار پشتیبان و روش گروهی مدل‌سازی داده‌ها در ارتقاء عملکرد مدل هیبرید نقش دارند. ماشین بردار پشتیبان با کاهش مسئله برازش با در نظر گرفتن تئوری کمینه‌سازی ریسک ساختاری دارای قدرت بالایی در مسائل پیش‌بینی است. در واقع این مدل دارای یک پارامتر تنظیم‌کننده است که مانع از بیش برازش می‌شود. همچنین ماشین بردار پشتیبان توسط یک مسئله بهینه‌سازی محدب تعریف می‌شود که برای حل آن روش‌های کارآمدی وجود دارد. مسئله قابل توجه در ماشین بردار پشتیبان، تابع کرنل است که از طبقه‌بندی‌کننده خطی برای حل مسئله غیرخطی استفاده می‌کند. تابع کرنل بر روی هر نمونه اعمال می‌شود تا مشاهدات غیرخطی اصلی را در یک فضای با ابعاد بالا قرار دهد که در آن قابل تفکیک هستند. روش مدل‌سازی گروهی داده‌ها با تعریف تابع محرک به‌صورت چند جمله‌ای خطی یا غیرخطی نیز روش مناسبی برای پیش‌بینی است. روش مدل‌سازی گروهی داده‌ها با روش استقرایی مشخص می‌شود که مرتب‌سازی از مدل‌های چند جمله‌ای پیچیده را شروع و در نهایت به بهترین راه‌حل می‌رسد. در بررسی تأثیر تنوع اقلیمی بر روند مدل‌سازی اثر قابل توجهی مشاهده نشد ولی به‌طور کلی مقادیر آماره‌های ارزیابی در ارومیه از وضعیت مطلوب‌تری نسبت به اصفهان

برخوردار بودند. البته در این مسئله انتخاب نوع ترکیب مناسب بر اساس اقلیم ایستگاه مهم است چرا که تغییرات معیارهای خطا در ترکیب دوم و اول در دو ایستگاه متفاوت بود ولی در ترکیب سوم در هر دو ایستگاه هماهنگی وجود داشت. به‌طور کلی روند ارتقاء مدل‌سازی هیبرید در هر دو ایستگاه با تنوع اقلیمی اثبات شد؛ بنابراین به نظر می‌رسد مدل هوش مصنوعی کارآمد یا استفاده از ترکیب نتایج با مدل‌های مختلف هوش مصنوعی می‌تواند به‌عنوان راهکار مؤثری در ارتقاء عملکرد مدل‌های هیبرید نقش داشته باشد.

منابع

- ۱- برجی حسن گاویار، م. مقدم نیا، ع؛ و ساجدی، ف (۱۳۹۶): بررسی کارایی دو روش داده محور در پیش‌بینی بارندگی ماهانه، پژوهش‌های دانش زمین، شماره سی و یکم، صص ۴۲-۶۱.
- ۲- پورنعمت رودسری، ع. قادری، ک؛ و کریمی گوگری، ش (۱۳۹۳): مدل‌سازی فرآیند بارش-رواناب با استفاده از روش کنترل گروهی داده‌ها (GMDH) و شبکه‌های عصبی مصنوعی (ANN) در حوضه آبخیز پلرود، پژوهشنامه مدیریت حوضه آبخیز، شماره دهم، صص ۸۴-۶۸.
- ۳- سلگی، ا. زارعی، ح. شهنی دارابی، م؛ و علیدادی ده کهنه، ص (۱۳۹۷): پیش‌بینی بارش ماهانه با استفاده از مدل‌های برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان، نشریه تحقیقات کاربردی علوم جغرافیایی، شماره پنجاه، صص ۹۱-۱۰۳.
- ۴- سلگی، ا. زارعی، ح. و گلابی، م.ر (۱۳۹۶): بررسی عملکرد مدل برنامه‌ریزی بیان ژن با روش‌های پیش‌پردازش داده‌ها جهت مدل‌سازی جریان رودخانه، پژوهش‌های حفاظت آب و خاک، شماره دو، صص ۲۰۱-۱۸۵.
- ۵- شرفی، م. صمدیان فرد، س؛ و هاشمی، س. (۱۳۹۹): پیش‌بینی بارش ماهانه با استفاده از مدل‌های برنامه‌ریزی ژنتیک و ماشین بردار پشتیبان، سامانه‌های سطوح آبگیر باران، دوره هشت، صص ۶۳-۷۱.

- 6- Acosta, S.M. Amoroso, A.L. Sant' Anna, Â.M.O. And Junior, O.C. (2021): Predictive Modeling In A Steelmaking Process Using Optimized Relevance Vector Regression And Support Vector Regression. *Annals Of Operations Research*, P.1-22.
- 7- Ahmadi, F. Mehdizadeh, S. Mohammadi, B. Pham, Q. B. Doan, T. N. C. And Vo. N. D. (2021): Application Of An Artificial Intelligence Technique Enhanced With Intelligent Water Drops For Monthly Reference Evapotranspiration Estimation: *Agricultural Water Management*, V. 244, 106622.
- 8- Chen, W. Xu, H. Chen, Z. And Jiang, M. (2021): A Novel Method For Time Series Prediction Based On Error Decomposition And Nonlinear Combination Of Forecasters: *Neurocomputing*, V. 22, P. 85-103.
- 9- Chen, S.T. Yu, P.S. And Tang, Y.H. (2010): Statistical Downscaling Of Daily Precipitation Using Support Vector Machines And Multivariate Analysis: *Journal Of Hydrology*, V. 385, P. 13-22.
- 10- Chen, K.Y. And Wang, C.H. (2007): A Hybrid SARIMA And Support Vector Machines In Forecasting The Production Values Of The Machinery Industry In Taiwan: *Expert Systems With Applications*, V. 32, P. 254-64.
- 11- Danandeh Mehr, A. (2018): An Improved Gene Expression Programming Model For Streamflow Forecasting In Intermittent Streams: *Journal Of Hydrology*, V.563, P. 669-78.
- 12- Kalteh, A.M. 2017: Enhanced Monthly Precipitation Forecasting Using Artificial Neural Network And Singular Spectrum Analysis Conjunction Models: *INAE Letters*, V.2, P.73-81.

- 13- Mislán, M. Haviluddin, H. Hardwinarto, S. Sumaryono, S. And Aipassa, M. (2015): Rainfall Monthly Prediction Based On Artificial Neural Network: A Case Study In Tenggara Station, East Kalimantan-Indonesia. The International Conference On Computer Science And Computational Intelligence (ICCSICI 2015)-Procedia Computer Science 59.
- 14- Mirabbasi, R. Kisi, O. Sanikhani, H. And Meshram, S. G. (2019): Monthly Long-Term Rainfall Estimation In Central India Using M5Tree, MARS, LSSVR, ANN And GEP Models: Neural Computing And Applications, V. 31,P. 6843-6862.
- 15- Mo, L. Xie. L. Jiang. X. Teng. G. Xu. L. And Xiao, J. (2018): GMDH-Based Hybrid Model For Container Throughput Forecasting: Selective Combination Forecasting In Nonlinear Subseries: Applied Soft Computing, V.62, P.478-490.
- 16- Papalaskaris, T. And Panagiotidis, T. (2017): Stochastic Monthly Rainfall Time Series Analysis, Modeling And Forecasting In Karyes, Chios Island, Greece, Central-Eastern Mediterranean Basin. In Proceedings Of The 15th International Conference On Environmental Science & Technology.
- 17- Ponnoprat, D. (2021): Short-Term Daily Precipitation Forecasting With Seasonally-Integrated Autoencoder: Applied Soft Computing, V.11,107083.
- 18- Qin, M. Li, Z. And Du, Z. (2017): Red Tide Time Series Forecasting By Combining ARIMA And Deep Belief Network: Knowledge-Based Systems, V.1,P.39-52.
- 19- Razzaghi, H. Madandoust, R. Aghabarti, H. (2021): Point-Load Test And UPV For Compressive Strength Prediction Of Recycled Coarse Aggregate Concrete Via Generalized GMDH-Class Neural Network. Construction And Building Materials.276.122143:1-9.
- 20- Ruiz-Aguilar, J.J. Turias, I.J, And Jiménez-Come, M.J. 2014:Hybrid Approaches Based On SARIMA And Artificial Neural Networks For Inspection Time Series Forecasting: Transportation Research Part E: Logistics And Transportation Review,V.67, P.67: 1-3.
- 21- Wang, J. Zhou, H. Hong, T. Li, X. And Wang, S. (2020): A Multi-Granularity Heterogeneous Combination Approach To Crude Oil Price Forecasting: Energy Economics, V.91, 104790.
- 22- Xu, L. Wang, X. Bai, Lu. Xiao, J. Liu, Q. Chen, E. Jiang, X. And Luo, B. (2020): Probabilistic SVM Classifier Ensemble Selection Based On GMDH-Type Neural Network: Pattern Recognition, V.106, 107373.
- 23- Yu, L. Wang, S. And Lai, K.K. (2005): A Novel Nonlinear Ensemble Forecasting Model Incorporating GLAR And ANN For Foreign Exchange Rates; Computers & Operations Research,V.32,P.2523-41.
- 24- Zhang, G.P. (2003): Time Series Forecasting Using A Hybrid ARIMA And Neural Network Model: Neurocomputing, V. 50, P.159-75.
- 25- Zhu, B. And Chevallier, J. (2017): Carbon Price Forecasting With A Hybrid ARIMA And Least Squares Support Vector Machines Methodology: In Pricing And Forecasting Carbon Markets, mV. 41, P. 517-524.