

ارزیابی عملکرد روش‌های داده محور در مدل‌سازی بارش ماهانه مشهد

علی رضازاده جودی^{۱*} و محمدتقی ستاری^۲

چکیده

بارش همواره از مهم‌ترین اجزاء چرخه آب شناخته شده است و نقش بسیار مهمی در تأمین منابع آبی مورد نیاز دارد؛ از این رو پیش‌آگاهی از میزان بارش در برنامه‌ریزی مدیریت منابع آب تأثیر بسزایی دارد. در این تحقیق میزان کارایی روش‌های رگرسیون بردار پشتیبان و رگرسیون فرآیند گاوسی در پیش‌بینی بارش ماهانه شهر مشهد با استفاده از پارامترهای هواشناسی مختلف بررسی شد و نتایج نشان داد بهترین ترکیب پارامتر ورودی شامل شاخص ماهانه، میانگین رطوبت نسبی، میانگین بیشینه رطوبت نسبی، اختلاف میانگین دمای کمینه و بیشینه و بارش در ماه قبل بود. مدل‌ها توسط آماره‌های ضریب همبستگی، ریشه میانگین مربعات خطا و میانگین خطای مطلق ارزیابی شدند. نتایج نشان‌دهنده کارایی بالای هر دو روش بررسی شده در این زمینه و برتری نسبی روش رگرسیون فرآیند گاوسی با ارائه مقادیر $(R = 0/87, N_S = 0/74)$ (mm) $RMSE = 12/37$ و $MAE = 7/85$ (mm) بود. با بررسی نمودار سری زمانی مقادیر محاسباتی و مشاهداتی بارش ماهانه مشهد مشاهده شد که روش رگرسیون فرآیند گاوسی توانایی بیشتری در پیش‌بینی مقادیر بیشینه بارش ماهانه داشته و در اکثر موارد پیش‌بینی‌های دقیق‌تری در مقادیر بارش ماهانه بیشینه ارائه کرده که این مورد در پیش‌بینی وقوع سیلاب‌ها بسیار مهم و کاربردی است.

واژه‌های کلیدی: پیش‌بینی بارش ماهانه، رگرسیون فرآیند گاوسی، رگرسیون بردار پشتیبان، مشهد.

ارجاع: رضازاده جودی ع. و ستاری م. ت. ۱۳۹۶. ارزیابی عملکرد روش‌های داده محور در مدل‌سازی بارش ماهانه مشهد. مجله پژوهش آب ایران. ۲۷: ۱۰۵-۹۷.

۱- دانش‌آموخته کارشناسی ارشد عمران - آب، باشگاه پژوهشگران جوان و نخبگان، واحد مراغه، دانشگاه آزاد اسلامی، مراغه.
۲- استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز.

* نویسنده مسئول: aligovdi66@gmail.com

تاریخ پذیرش: ۱۳۹۵/۰۷/۱۸

تاریخ دریافت: ۱۳۹۵/۰۲/۲۲

مقدمه

بارش پدیده غیرخطی بوده که تحت تأثیر عوامل متعددی از جمله عوامل اقلیمی و جغرافیایی، تغییرات زمانی و مکانی زیادی را نشان می‌دهد. از عوامل اقلیمی می‌توان رطوبت، فشار، دما، پوشش ابر، سرعت باد، عوامل فصلی مؤثر بر تغییرات سینوسی بارش، همچنین سیگنال‌های بزرگ مقیاس اقلیمی و به خصوص دوره‌های النینو و لانینا را نام برد. همچنین از عوامل جغرافیایی می‌توان به دوری و نزدیکی از مرکز تولید جبهه (مثلاً دریاها یا صحراهای بزرگ) و همچنین ارتفاع اشاره کرد (کویندارجو، ۲۰۰۰). گرچه تا به حال مدل‌های مفهومی و آماری زیادی برای پیش‌بینی و پیش‌آگاهی از متغیرهای اقلیمی ارائه شده است، امروزه نیز سیستم‌های هوشمندی همچون روش شبکه عصبی مصنوعی و مدل‌های درختی در مدل‌سازی فرایندهای هیدرولوژیکی و مهندسی آب مورد توجه قرار گرفته‌اند. روش‌های مذکور، به عنوان روش‌هایی معتبر و شناخته شده در مدل‌سازی فرایندهای پیچیده غیرخطی، قادرند با استفاده از داده‌های هواشناسی ثبت شده در ایستگاه‌های هواشناسی تخمین دقیق‌تری از پارامترهای هواشناسی منطقه ارائه دهند (ستاری و همکاران، ۱۳۹۳). با توجه به توانایی بسیار بالای روش‌های هوش مصنوعی شامل شبکه‌های عصبی مصنوعی و مدل‌های درختی و اهمیت آگاهی دقیق از میزان نزولات جوی در مناطق مختلف همواره پژوهش‌گران بسیاری نسبت به ارزیابی کارایی این روش‌ها در این زمینه پرداخته‌اند. برای مثال ترافلیس و همکاران (۲۰۰۵) از روش شبکه عصبی مصنوعی و چند روش دیگر برای تخمین بارش استفاده کردند که در نهایت روش رگرسیون بردار پشتیبانی^۱ را به‌عنوان روشی که بهترین نتیجه را می‌دهد، معرفی کردند. چتوپادیای (۲۰۰۷) با استفاده از شبکه عصبی پیش‌خور^۲ به پیش‌بینی میانگین بارش‌های موسمی تابستانی در هند پرداخت. نتایج نشان داد مدل شبکه عصبی پیش‌خور برتری بالایی نسبت به روش رگرسیون چندمتغیره خطی دارد. موهانتی و موهاپاترا (۲۰۰۷) با استفاده از رگرسیون گام به گام و آنالیز چندگانه افتراقی^۳ به پیش‌بینی وقوع و مقدار بارش‌های موسمی روزانه در تابستان در هند

پرداختند و توانایی بالای این روش را در پیش‌بینی احتمال بارش اعلام کردند. تاکوترا (۲۰۰۸) بارش‌های موسمی دراز مدت را برای ۳۶ زیربخش هواشناسی هند با استفاده از مدل شبکه عصبی مصنوعی بررسی کرد. ونکاتا رامانا و همکاران (۲۰۱۳) با استفاده از روش تحلیل شبکه عصبی موجکی^۴ به پیش‌بینی بارش ماهانه پرداختند. نتایج نشان‌دهنده برتری روش شبکه عصبی موجکی نسبت به شبکه عصبی معمولی بود. آلتین‌کاپناک و نیگوسیه (۲۰۱۵) با استفاده از روش ترکیبی شبکه عصبی فصلی - موجکی به پیش‌بینی بارش روزانه پرداختند و کارایی بالای این روش را در مقایسه با کاربرد منفرد هر کدام از این روش‌ها اعلام کردند. میسلان و همکاران (۲۰۱۵) با استفاده از روش شبکه عصبی مصنوعی بارش ماهانه را در ایستگاهی در اندونزی پیش‌بینی کردند. نتایج نشان‌دهنده کارایی بالای شبکه عصبی پس انتشار خطا^۵ در این زمینه بود. فلاح قاهره‌ری و همکاران (۱۳۸۸) کارکرد دو مدل سیستم استنباط فازی و شبکه عصبی مصنوعی در پیش‌بینی بارش در منطقه خراسان بزرگ را مقایسه کردند و نشان دادند که شبکه عصبی مصنوعی در پیش‌بینی بارش موفق‌تر از مدل فازی است. فلاح قاهره‌ری و خوشحال (۱۳۸۸) میزان بارش را در استان خراسان رضوی با استفاده از شبکه‌های عصبی مصنوعی پیش‌بینی کردند و به این نتیجه رسیدند که شبکه عصبی مصنوعی در تمامی سال‌ها می‌تواند بارش را با دقت قابل قبولی پیش‌بینی کند. حلبیان (۱۳۸۸) مقادیر بارش ماهانه یزد را با استفاده از شبکه عصبی مصنوعی پیش‌بینی کردند. قلی‌زاده و دارند (۱۳۸۹) میزان بارش ماهانه را در تهران با استفاده از شبکه عصبی مصنوعی تخمین زدند. فاتحی مرج و مهدیان (۱۳۸۹) ارتباط بین پدیده انسو^۶ را با بارش در حوضه دریاچه ارومیه بررسی و برای پیش‌بینی بارش پاییزه در دو ایستگاه تبریز و ارومیه از مدل شبکه‌های عصبی مصنوعی استفاده کردند و نشان دادند که مدل غیرخطی با استفاده از شاخص‌های اقلیمی مورد استفاده، بارش پاییزه را با دقت بیشتری پیش‌بینی می‌کند. حلبیان (۱۳۹۱) با استفاده از روش شبکه عصبی مصنوعی به پیش‌بینی بارش ماهانه اصفهان پرداخت و اعلام کرد که استفاده از شبکه

4- Wavelet

5- Backpropagation

6- El Nino Southern Oscillation

1 -Support vector regression

2- Feed forward

3- Multiple discriminant analysis

(۱۳۸۸)، در این پژوهش داده‌های ماهانه مربوط به پارامترهای هواشناسی شامل میانگین رطوبت نسبی (rh_{mean})، بیشینه رطوبت نسبی (rh_{max})، میانگین کمینه دما ($T_{minmean}$)، میانگین بیشینه دما ($T_{maxmean}$)، اختلاف میانگین دمای کمینه و بیشینه ($T_{maxmean} - T_{minmean}$)، کمینه دما (T_{min})، میانگین فشار بخار آب ($q_{ff_{mean}}$)، میانگین متوسط فشار سطح ایستگاه ($q_{fe_{mean}}$)، میانگین بیشینه فشار سطح ایستگاه ($q_{fe_{maxmean}}$)، بارش در ماه قبل (P_{t-1}) و بارش ماهانه (P_t) در ایستگاه سینوپتیک شهر مشهد بین سال‌های ۱۳۲۹ تا ۱۳۸۸ بررسی شده است. همچنین برای هر ماه پارامتری به عنوان شاخص ماهانه تعریف شد که بین ۱ تا ۱۲ تغییر می‌کند. در این پژوهش ۸۰٪ داده‌ها برای قسمت آموزش مدل (به صورت تصادفی) و ۲۰٪ داده‌ها برای قسمت آزمون مدل (به صورت تصادفی) در نظر گرفته شد. جدول ۱ نشان‌دهنده مشخصات آماری داده‌های آماری مورد استفاده در این پژوهش است.

رگرسیون بردار پشتیبان

ماشین بردار پشتیبان، یک از روش‌های یادگیری با نظارت است که در سال ۱۹۹۵ میلادی توسط واپنیک بر پایه تئوری یادگیری آماری معرفی شد. منظور از رگرسیون بردار پشتیبان، به دست آوردن یک ابرصفحه است که بر داده‌های مورد نظر برازش داده می‌شود. فاصله هر نقطه از این ابرصفحه، نشان‌دهنده خطای آن نقطه خاص است. در رگرسیون بردار پشتیبان می‌بایست یک برآوردکننده نیرومند را یافت که نسبت به تغییرات کوچک در مدل حساس نباشد (شهرابی و شجاعی، ۱۳۹۰). در واقع، یک تابع زیان e - غیرحساس به صورت زیر تعریف می‌شود:

$$L^E(x, y, f) = |y - f(x)|_E = \max(0, |y - f(x)| - \varepsilon) \quad (1)$$

اگر $|y - f(x)| \leq \varepsilon$ باشد، آن‌گاه تابع زیان برابر صفر و در غیر این صورت برابر $|y - f(x)| - \varepsilon$ خواهد بود. مجموعه داده‌های آموزشی به صورت $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ و کلاس تابع به صورت $f(x) = \{w^T x + b, w \in R^m, b \in R\}$ است. در صورتی که داده‌ها از مقدار e تخطی کنند، متناسب با مقدار تخطی، متغیر کمبود تعریف می‌شود. مطابق با تابع

عصبی مصنوعی دارای دو لایه پنهان با ضریب یادگیری ۰/۴ نتایج بهتری ارائه می‌کند. ستاری و نهرین (۱۳۹۲) با استفاده از روش‌های هوشمند شامل برنامه‌ریزی ژنتیک، شبکه عصبی مصنوعی و مدل درختی M5 به پیش‌بینی مقادیر بیشینه بارش روزانه در اهر و جلفا پرداختند و کارایی بالای هر سه روش مورد مطالعه را در زمینه مذکور اعلام کردند. گلابی و همکاران (۱۳۹۳) با استفاده از مدل‌های باکس-جنکینز به پیش‌بینی بارندگی فصلی در ایستگاه‌های منتخب استان خوزستان پرداختند و دقت بالای این مدل‌ها را گزارش دادند. ستاری و همکاران (۱۳۹۳) با استفاده از روش‌های شبکه عصبی مصنوعی و مدل درختی M5 به پیش‌بینی مقادیر بارش ماهانه در شهر اهر پرداختند و با توجه به ارائه روابط خطی ساده و قابل فهم کاربرد روش مدل درختی M5 را توصیه کردند. همچنین ماه‌آورپور (۱۳۹۳) با استفاده از مدل زنجیره مارکوف به پیش‌بینی احتمال وقوع بارش‌های روزانه ایران پرداخت. رحیمی و همکاران (۱۳۹۳) با استفاده از مدل GSDM به بیشینه بارش محتمل در حوضه‌های آبریز را برآورد کردند که ارتباطی مستقیم به موضوع دارد. بررسی منابع صورت گرفته، نشان از کاربرد بالای روش شبکه عصبی مصنوعی در زمینه پیش‌بینی بارش است؛ اما بررسی مطالعات پیشین حاکی از این است که مطالعات چندانی در زمینه کاربرد ماشین بردار پشتیبان و رگرسیون فرآیند گاوسی در زمینه پیش‌بینی میزان بارش ماهانه انجام شده است. هدف از این مطالعه ارزیابی کارایی روش‌های رگرسیون بردار پشتیبان و رگرسیون فرآیند گاوسی در زمینه پیش‌بینی میزان بارش ماهانه شهر مشهد و همچنین انجام تحلیل حساسیت میزان بارش نسبت به پارامترهای هواشناسی است.

مواد و روش‌ها

منطقه مورد مطالعه

شهر مشهد در حوضه رودخانه‌ی کشف‌رود و در دشت مشهد، بین کوه‌های هزارمسجد و بینالود گسترده شده است. شهر مشهد دارای آب و هوای متغیر، اما معتدل و متمایل به سرد و خشک است و تابستان‌های گرم و خشک و زمستان‌های سرد و مرطوب دارد. با توجه به این‌که برای تشکیل بارش سه پارامتر دما، رطوبت نسبی و فشار بخار آب تأثیرگذار و پراهمیت است (شریفیان و هزارجریبی،

تابع کرنل

از جمله روش‌های متداول برای حل مسائل غیرخطی، استفاده از توابع کرنل است. از مهم‌ترین توابع کرنل می‌توان به کرنل خطی، چند جمله‌ای (رابطه (۵))، چند جمله‌ای نرمال شده (رابطه (۶))، کرنل تابع شعاع محور (رابطه (۷)) و کرنل مبتنی بر تابع پیرسون (رابطه (۸)) اشاره کرد که در این مطالعه، از این توابع کرنل استفاده شده است.

$$k(x_i, x_j) = (x_i^T x_j + 1)^p \quad (۵)$$

$$k(x_i, x_j) = \frac{(x_i^T x_j + 1)^p}{\sqrt{(x_i^T x_i)(x_j^T x_j)}} \quad (۶)$$

$$k(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2) \quad (۷)$$

$$k(x_i, x_j) = \frac{1}{1 + \left(2\sqrt{\|x_i, x_j\|^2} \sqrt{2^{1/w} - 1/\sigma}\right)^{2w}} \quad (۸)$$

هر کدام از این توابع پارامترهای خاص کرنل خود را دارند که به آنها هایپارامتر^۲ گفته می‌شود؛ برای مثال تابع پایه شعاعی دارای هایپارامتر گاما (γ) و تابع پیرسون دارای هایپارامترهای سیگما (σ) و امگا (ω) است. در کاربرد روش‌های مبتنی بر تابع کرنل علاوه بر انتخاب تابع کرنل بهینه، تعیین مقدار بهینه هایپارامترهای مربوط به هر تابع بسیار حائز اهمیت است. علاوه بر انتخاب پارامترهای خاص هر کرنل (هایپارامتر)، در رگرسیون فرآیند گاوسی به مشخص کردن مقادیر بهینه سطح نویز گاوسی (ϵ) و در رگرسیون بردار پشتیبان به مشخص کردن ضریب هموارساز (C) نیاز است. برای انتخاب پارامترهای انتخاب شده توسط کاربر (یعنی $C, \gamma, \sigma, \omega, \epsilon$) در مطالعه حاضر، از روش سعی و خطا استفاده شده است. در این مطالعه برای پیش‌بینی بارش ماهانه توسط روش‌های رگرسیون فرآیند گاوسی و رگرسیون بردار پشتیبان از نرم‌افزار وکا^۳ که در دانشگاه ویکاتو نیوزیلند توسعه یافته، استفاده شده است.

زیان اشاره شده، کمینه‌سازی مسأله اولیه، نقطه زینی و مسأله دوگان مربوطه را با استفاده از توابع کرنل تشکیل و شرایط روش کان تاکر^۱ بررسی می‌شود:

subject to:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i^+ + \xi_i^-) \quad (۲)$$

$$f(x) - y_i \leq \epsilon + \xi_i^+, \quad (۳)$$

$$f(x) - y_i \leq \epsilon + \xi_i^-, \quad \xi_i^+ \geq 0, \xi_i^- \geq 0,$$

$$i = 1, 2, 3, \dots, n$$

در رابطه فوق، $\|w\|^2$ نرم بردار وزن، x_i^+, x_i^- متغیرهای کمبود کمکی هستند و پارامتر C ضریب تعادل پیچیدگی میان ماشین و تعداد نقاط تفکیک‌ناپذیر است که با سعی و خطا به دست می‌آید (شهرابی و شجاعی، ۱۳۹۰).

رگرسیون فرآیند گاوسی

در نظریه احتمال و آمار، فرآیند گاوسی یک فرآیند تصادفی است که بر پای، توزیع نرمال (معمولاً توزیع گاوسی نیز نامیده می‌شود) استوار است؛ به گونه‌ای که هر یک از این متغیرهای تصادفی دارای توزیع نرمال باشد. علاوه بر این، هر مجموعه محدود از این متغیرهای تصادفی دارای توزیع نرمال چند متغیره است. در واقع از یک نکته نظر فرآیند گاوسی، یک تعمیم دادن نامحدود ابعادی برای توزیع‌های نامحدود چندمتغیره است. فرآیندهای گاوسی به دلیل برخورداری از خصوصیات نرمال در مدل‌سازی آماری بسیار پراهمیت‌اند (نیل، ۱۹۹۷).

$$f \sim \text{GP}(m, k) \quad (۴)$$

مدل‌های رگرسیون فرآیند گاوسی بر این فرض استوار هستند که مشاهدات باید حامل اطلاعاتی درباره همدیگر باشد. فرایندهای گاوسی، راهی برای مشخص کردن اولییتی به صورت مستقیم روی فضای تابع هستند. این کار تعمیم طبیعی توزیع گاوسی است که میانگین و کواریانس آن به ترتیب بردار و ماتریس هستند (پال و دسوال، ۲۰۱۰).

جدول ۱- مشخصات آماری داده‌های هواشناسی مورد استفاده

انحراف استاندارد	میانگین	بیشینه	کمینه	
۱۷/۶۴۲	۵۴/۵۶۲	۹۰	۱۸	میانگین رطوبت نسبی (%)
۱۷/۹۱۶	۷۳/۶۲۶	۹۹	۳۰	میانگین بیشینه رطوبت نسبی (%)
۷/۸۷۵	۷/۱۸۲	۲۱/۸	-۱۵/۹	میانگین کمینه دمای (°C)
۹/۷۹۳	۲۱/۲۶۳	۳۶/۹	-۱/۹	میانگین بیشینه دمای (°C)
۳/۲۱۸	۱۴/۰۸۲	۲۶/۱	۵/۵	اختلاف میانگین دمای کمینه و بیشینه (°C)
۹/۵۶	۰/۶۴۷	-۱۸/۵	-۲۸	کمینه دما (°C)
۲/۶۴۱	۸/۲۰۵	۱۷/۷	۲/۸	میانگین فشار بخار آب (HPA)
۳/۳۹۴	۹۰۳/۰۲	۹۱۱/۶	۸۹۰/۹	میانگین متوسط فشار سطح ایستگاه (HPA)
۵/۲۰۴	۹۱۱/۷۹۷	۹۳۳/۴	۸۹۶/۶	میانگین بیشینه فشار سطح ایستگاه (HPA)
۲۵/۳۸۵	۲۰/۹۴۷	۱۳۲	۰	بارش ماهانه (mm)

یافته‌های پژوهش

شده در بخش مواد و روش‌ها به عنوان ورودی مدل مورد توجه و آزمون قرار گرفتند. با استفاده از روش آزمون و خطا، آزمون گاما و همچنین ماتریس همبستگی، برای کاستن از پیچیدگی مدل، پارامترهایی حذف شد که تأثیر چندانی نداشتند؛ در آزمون گاما چنانچه با حذف یک پارامتر مقدار گاما نسبت به حالتی که تمامی پارامترها حضور دارند، افزایش یابد، نشان‌دهنده تأثیر مثبت آن پارامتر در مدل‌سازی است (مقصود و همکاران، ۱۳۹۵). در این مطالعه برای اجرای آزمون گاما از نرم‌افزار Wingamma استفاده شده است. نتایج آزمون گاما در جدول ۲ ارائه شده است.

اصلی‌ترین قسمت روش رگرسیون بردار پشتیبان و رگرسیون فرآیند گاوسی انتخاب تابع کرنل بهینه برای خطی کردن ساختار و ماهیت غیرخطی پارامتر بارش و پارامترهای وابسته به آن است. در این زمینه پس از بررسی توابع کرنل مختلف مشخص شد، نتایج بهینه مربوط به زمانی است که از تابع کرنل پیرسون استفاده می‌شود. پس از انجام تکرار و آزمون و خطا مقادیر بهینه پارامترهای این تابع در روش رگرسیون بردار پشتیبان به صورت $(C=0/4, \sigma=1/5, \omega=2)$ و در روش رگرسیون فرآیند گاوسی به صورت $(\epsilon=0/7, \sigma=1/2, \omega=0/6)$ به دست آمد. سپس پارامترهای متنوع هواشناسی مطرح

جدول ۲- نتایج آزمون گاما برای بررسی تأثیر متغیرهای ورودی مختلف در مدل‌سازی بارش

پارامتر حذف شده	مقدار گاما	پارامتر حذف شده	مقدار گاما
بدون حذف	۱۸۴/۱۶	$T_{min}mean - T_{max}mean$	۲۰۱/۸۶
Monthly Index	۲۱۸/۹۸	T_{min}	۱۹۳/۹۴
rh_{mean}	۲۱۳/۱	qff_{mean}	۲۰۱/۰۵
rh_{max}	۲۰۲/۷۱	qfe_{mean}	۱۹۱/۲۱
$T_{min} mean$	۲۲۳/۷	qfe_{max}	۲۰۴/۴۲
$T_{max} mean$	۱۸۹/۰۷	P_{t-1}	۲۰۴/۲۲

همچنان که در جدول ۳ مشاهده می‌شود هر دو روش بررسی شده در تمامی حالات دقت عمل و کارایی قابل قبولی از خود نشان داده‌اند که نشان‌دهنده توانایی بالای این روش‌های داده‌کاوی در زمینه تخمین مقادیر بارش است. همچنین با بررسی آماره‌های ارزیابی دقت مدل‌ها، مشاهده می‌شود در تمامی ترکیب پارامترها روش رگرسیون فرآیند گاوسی دقت بیشتر و خطای کمتری داشته است. همچنین ترکیب پارامتر شماره ۶ شامل

نتایج جدول ۲ نشان می‌دهد کلیه پارامترهای بررسی شده در این پژوهش در مدل‌سازی بارش مؤثر هستند. بر این اساس شش ترکیب ورودی مختلف که ترکیبی از مهم‌ترین پارامترهای تأثیرگذار در بارش بودند، به عنوان ورودی مدل رگرسیون بردار پشتیبان و رگرسیون فرآیند گاوسی در نظر گرفته شد. این ترکیب پارامترهای ورودی و نتایج مربوطه در دو مدل بررسی شده بر اساس آماره‌های R ، $RMSE$ ، N_S و MAE در جدول ۳ ارائه شده است.

تمامی پارامترهای ورودی اجرا شد. سپس با حذف یک به یک پارامترهای ورودی نسبت به ارزیابی دقت مدل با آماره‌های مطرح شده اقدام گردید. نتایج مربوط به تحلیل حساسیت مدل به کاهش پارامترهای ورودی در جدول ۴ ارائه شده است.

نتایج جدول ۴ نشان می‌دهد که بیشترین حساسیت مدل به ترتیب نسبت به حذف پارامترهای بارش در ماه قبل، شاخص ماهانه و میانگین رطوبت نسبی است که نشان‌دهنده تأثیر مستقیم و بیشتر این پارامترها در پدیده بارش و به تبع آن پیش‌بینی مقادیر بارش ماهانه است.

پارامترهای شاخص ماهانه، میانگین رطوبت نسبی، میانگین بیشینه دمای ماهانه، اختلاف میانگین دمای کمینه و بیشینه، بیشینه رطوبت نسبی، کمینه دمای ماهانه و بارش در ماه قبل دارای بیشترین دقت و کمترین خطاست. این امر نشان‌دهنده تأثیر بیشتر و مستقیم این پارامترها در میزان بارش است.

تحلیل حساسیت مدل به پارامترهای ورودی

برای تحلیل حساسیت، ابتدا مدل رگرسیون فرآیند گاوسی با ساختار بهینه ($\varepsilon=0/7$, $\omega=0/55$ و $\sigma=1/2$) با استفاده از

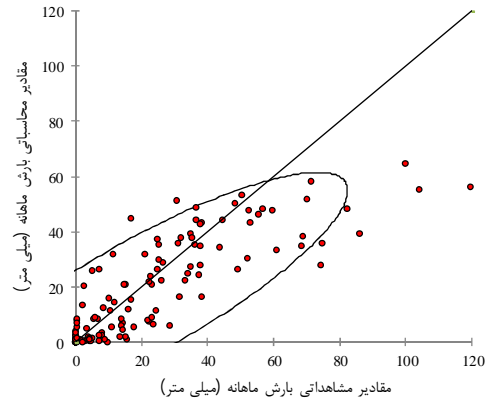
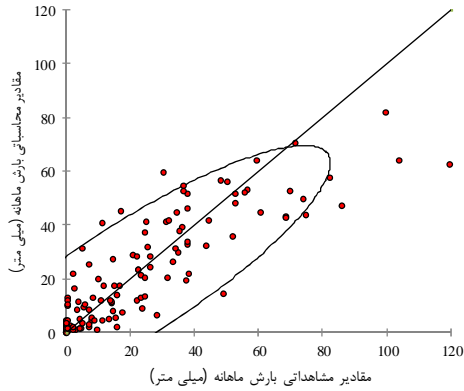
جدول ۳- نتایج آماری به دست آمده از روش‌های رگرسیون بردار پشتیبان و رگرسیون فرآیند گاوس

رگرسیون فرآیند گاوسی				رگرسیون بردار پشتیبان				ترکیب پارامترهای ورودی	ترکیب ورودی
R	N-S	RMSE (mm)	MAE (mm)	R	N_S	RMSE (mm)	MAE (mm)		
۰/۸۳۳	۰/۶۸۹	۸/۳۹	۱۳/۸۲	۰/۸۱۴	۰/۶۲۵	۸/۹۷	۱۵/۱۸	Monthly index, mean (rh, T _{min} , T _{max} , qff, qfe), T _{min} , (T _{max} mean-T _{min} mean), rh _{max} , qfe _{max} , P _{t-1}	(۱)
۰/۸۳۰	۰/۶۸۵	۸/۴۲	۱۳/۹۰	۰/۸۳۰	۰/۶۴۷	۸/۶۸	۱۴/۷۱	Monthly index, mean (rh, T _{min} , T _{max}), T _{min} , T _{max} mean-T _{min} mean, rh _{max} , qfe _{max} , P _{t-1}	(۲)
۰/۸۲۰	۰/۶۶۲	۸/۸۵	۱۴/۴۰	۰/۸۱۲	۰/۶۰۷	۹/۴۰	۱۵/۵۳	mean (T _{max}), T _{min} , (T _{max} mean-T _{min} mean), rh _{max} , qfe _{max} , P _{t-1}	(۳)
۰/۸۳۳	۰/۶۹۲	۸/۳۹	۱۳/۷۷	۰/۸۲۷	۰/۶۴۱	۸/۷۷	۱۴/۸۰	Monthly index, mean (rh, T _{min} , T _{max} , qff), T _{min} , (T _{max} mean-T _{min} mean), rh _{max} , qfe _{max} , P _{t-1}	(۴)
۰/۸۷۰	۰/۷۳۴	۷/۸۷	۱۲/۳۸	۰/۸۵۱	۰/۷۱۴	۸/۴۲	۱۳/۸۷	Monthly index, mean (rh), (T _{max} mean-T _{min} mean), rh _{max} , P _{t-1}	(۵)
۰/۸۷۰	۰/۷۳۶	۷/۸۵	۱۲/۳۷	۰/۸۵۴	۰/۷۲۰	۸/۳۶	۱۳/۸۵	Monthly index, mean (rh, T _{max}), (T _{max} mean-T _{min} mean), rh _{max} , T _{min} , P _{t-1}	(۶)

با بررسی شکل‌های ۱ و ۲ مشخص شد که در هر دو روش داده‌کاوی بررسی شده در این پژوهش تعداد نقاطی که خارج از باند اطمینان قرار دارند برابر نه نقطه (۶/۲۵٪ کل داده‌های قسمت آزمون) است؛ اما باند اطمینان به دست آمده از روش رگرسیون فرآیند گاوسی تطابق بیشتری با نیمساز ربع اول دارد که نشان‌دهنده همخوانی و پراکنش بهتر داده‌های به دست آمده از این روش نسبت به داده‌های مشاهداتی است. همچنین با بررسی دقیق نتایج ارایه شده در شکل ۳ مشاهده می‌شود روش رگرسیون فرآیند گاوسی توانایی بیشتری در پیش‌بینی مقادیر بیشینه بارش دارد و در اکثر موارد پیش‌بینی‌های دقیق‌تری در مواردی ارائه کرده است که مقادیر بارش بیشینه می‌شود که این مورد، در پیش‌بینی وقوع سیلاب‌ها بسیار مهم و کاربردی است.

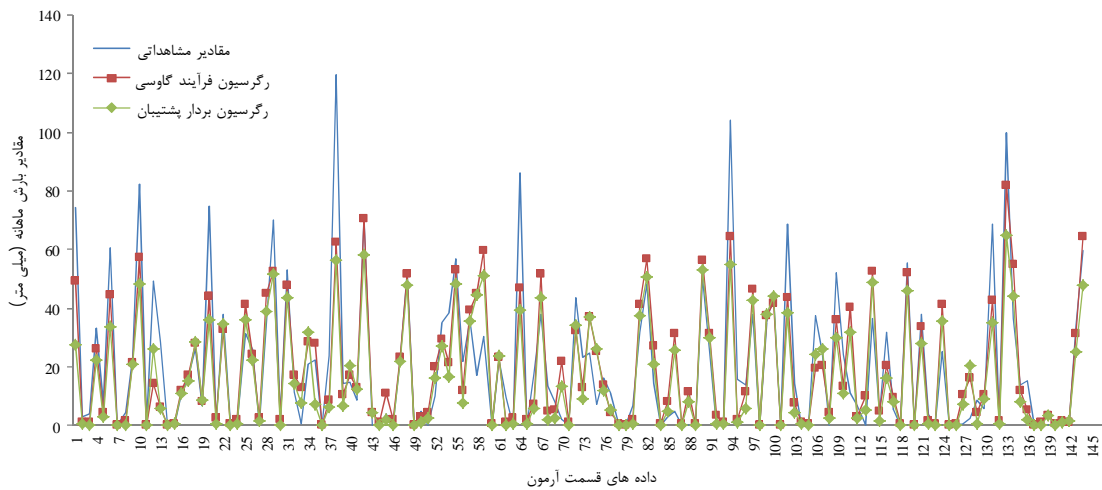
روش رگرسیون فرآیند گاوسی با ترکیب پارامتر ورودی شماره ۶ با ارائه مقادیر $R=0/87$, $N_S=0/736$, $RMSE=12/37$ (mm) و $MAE=7/85$ (mm) به عنوان روش مناسب برای پیش‌بینی مقادیر بارش ماهانه معرفی شد؛ زیرا با کمترین تعداد پارامتر ورودی قابل قبول‌ترین نتیجه را ارائه کرده است. شکل ۱ و ۲ نشان‌دهنده نمودار پراکنش مقادیر محاسباتی بارش به دست آمده توسط روش رگرسیون بردار پشتیبان و رگرسیون فرآیند گاوسی نسبت به مقادیر مشاهداتی بارش هستند. همچنین در شکل ۳ نمودار سری زمانی مقادیر محاسباتی بارش توسط روش‌های داده‌کاوی بررسی شده نسبت به مقادیر واقعی بارش ارائه شد.

باند اطمینان ارائه شده در شکل‌های ۱ و ۲ توسط نرم‌افزار آماری XLSTAT در سطح معنی‌داری ۹۵ درصد رسم شده‌اند که نشان‌دهنده این است که ۹۵ درصد داده‌ها با درجه اطمینان ۹۵ درصد در آن محدوده قرار دارند.



شکل ۲- پراکنش مقادیر محاسباتی به دست آمده از رگرسیون فرآیند گاوسی نسبت به مقادیر مشاهده‌ای

شکل ۱- پراکنش مقادیر محاسباتی به دست آمده از رگرسیون بردار پشتیبان نسبت به مقادیر مشاهده‌ای



شکل ۳- سری زمانی مقادیر محاسباتی به دست آمده نسبت به مقادیر مشاهده‌ای بارش مشهد

جدول ۴- نتایج تحلیل حساسیت مدل به کاهش پارامترهای ورودی

شماره سناریو	پارامتر ورودی حذف شده	R	RMSE (mm)	MAE (mm)
(۱)	بدون حذف	۰/۸۳۷	۱۱/۹۵	۸/۴۲
(۲)	میانگین رطوبت نسبی	۰/۸۲۳	۱۴/۲۱	۸/۷۶
(۳)	بیشینه رطوبت نسبی	۰/۸۳۱	۱۳/۹۴	۸/۶۶
(۴)	میانگین کمینه دماها	۰/۸۳۷	۱۳/۷۴	۸/۵۲
(۵)	میانگین بیشینه دماها	۰/۸۳۳	۱۳/۸۵	۸/۵۴
(۶)	اختلاف میانگین بیشینه و کمینه دما	۰/۸۲۹	۱۴/۰۲	۸/۵۹
(۷)	کمینه دما	۰/۸۲۵	۱۴/۱۴	۸/۷۹
(۸)	میانگین فشار بخار آب	۰/۸۳۰	۱۳/۹۷	۸/۶۸
(۹)	میانگین فشار سطح ایستگاه	۰/۸۳۶	۱۳/۷۷	۸/۴۲
(۱۰)	بیشینه فشار سطح ایستگاه	۰/۸۳۲	۱۳/۸۶	۸/۵۵
(۱۱)	بارش در ماه قبل	۰/۸۱۶	۱۴/۴۳	۸/۸۹
(۱۲)	شاخص ماهانه	۰/۸۱۶	۱۴/۴۲	۸/۷۸

نتیجه‌گیری

در این پژوهش تلاش شد میزان بارش شهر مشهد با استفاده از پارامترهای اقلیمی و بارش در ماه قبل پیش‌بینی شود. نتایج پژوهش نشان‌دهنده دقت بیشتر و خطای کمتر پیش‌بینی هنگامی است که از پارامترهای شاخص ماهانه، میانگین رطوبت نسبی، میانگین بیشینه رطوبت نسبی، اختلاف میانگین دمای کمینه و بیشینه و بارش در ماه قبل استفاده شد. همچنین نتایج نشان‌دهنده کارایی بالای روش‌های داده‌کاوی رگرسیون بردار پشتیبان و رگرسیون فرآیند گاوسی در زمینه پیش‌بینی بارش است. روش رگرسیون فرآیند گاوسی با ارائه مقادیر ضریب همبستگی برابر $0/870$ ، ضریب نش- ساتکلیف برابر با $0/736$ ، ریشه میانگین مربعات خطا برابر $12/37$ (میلی‌متر) و میانگین خطای مطلق برابر $7/85$ (میلی‌متر) به‌عنوان بهترین روش برای پیش‌بینی مقادیر بارش برای کاربرد در موارد مشابه توصیه می‌شود. همچنین نتایج، نشان می‌دهند روش رگرسیون فرآیند گاوسی توانایی بیشتری در پیش‌بینی مقادیر بیشینه بارش ماهانه دارد و در اکثر موارد پیش‌بینی‌های دقیق‌تری در مواردی که مقادیر بارش بیشینه می‌شود، ارائه کرده است که این مورد در پیش‌بینی وقوع سیلاب‌ها بسیار مهم و کاربردی است. انجام تحلیل حساسیت مدل‌ها به پارامترهای ورودی نشان‌دهنده این مطلب است که پارامتر بارش ماهانه بیشترین تأثیر را از میزان بارش در ماه قبل، شاخص ماهانه و سپس کمینه دما می‌پذیرد. با در نظر گرفتن مطالعات سایر پژوهش‌گران در این زمینه مشاهده می‌شود که پژوهش حاضر از دقت مطلوبی نسبت به سایر تحقیقات برخوردار است؛ برای مثال، ترافالیس و همکاران (۲۰۰۵) بارش را با استفاده از روش شبکه عصبی مصنوعی با $(R=0/447)$ و با استفاده از ماشین بردار پشتیبان $(R=0/582)$ تخمین زدند و یا فلاح قاله‌ری و همکاران $(R=0/616)$ را برای روش شبکه عصبی مصنوعی در پیش‌بینی بارش به دست آوردند. همچنین حلبیان (۱۳۸۸) میزان بارش ماهانه یزد را با استفاده از شبکه عصبی مصنوعی تخمین زده و مقدار $(R=0/743)$ را به دست آوردند؛ در حالیکه در این پژوهش مقدار $(R=0/870)$ با استفاده از رگرسیون فرآیند گاوسی به دست آمد. همچنین نتایج به دست آمده به نتایج تحقیقات ستاری و همکاران (۱۳۹۳) که برای پیش‌بینی

بارش ماهانه اهر مقدار $(R=0/84)$ را با استفاده از روش شبکه عصبی مصنوعی و $(R=0/87)$ مدل درختی M5 به دست آورده بودند، نزدیک است. مقایسه کلی نتایج این پژوهش با سایر پژوهش‌های نسبتاً مشابه بیانگر این نکته است که روش‌های معرفی شده قابلیت رقابت با سایر روش‌های هوش مصنوعی را دارند و می‌توان از آن‌ها در مدل‌سازی بارش استفاده کرد.

منابع

۱. حلبیان ا. ۱۳۸۸. پیش‌آگاهی و برآورد بارش یزد با استفاده از شبکه عصبی مصنوعی. نشریه تحقیقات کاربردی علوم جغرافیایی. ۱۱: ۷-۲۸.
۲. حلبیان ا. ۱۳۹۱. پیش‌بینی بارش اصفهان با استفاده از شبکه‌های عصبی مصنوعی. نشریه تحقیقات کاربردی علوم جغرافیایی. ۱۲(۲۶): ۴۷-۶۳.
۳. رحیمی د. خالقی د. و آنتشی ن. ۱۳۹۳. برآورد حداکثر بارش محتمل در حوضه‌های آبی با استفاده از مدل GSDM (حوضه آبی جهان‌بین). فصلنامه تحقیقات جغرافیایی. ۳۰(۲): ۲۱۵-۲۲۶.
۴. رضازاده جودی ع. و ستاری م. ۱۳۹۴. تخمین عمق چاله آبستستگی پایه پل در سازه‌های رودخانه‌ای با روش رگرسیون فرآیند گاوسی. نشریه تحقیقات کاربردی مهندسی سازه‌های آبیاری و زهکشی. ۱۶(۶۵): ۱۹-۳۶.
۵. ستاری م. رضازاده جودی ع. و نهرین ف. ۱۳۹۳. پیش‌بینی بارش ماهانه با استفاده از شبکه عصبی مصنوعی و مدل درختی M5 (مطالعه موردی: ایستگاه اهر)، مجله پژوهش‌های جغرافیای طبیعی. ۴۶(۲): ۲۴۷-۲۶۰.
۶. ستاری م. و نهرین ف. ۱۳۹۲. پیش‌بینی مقادیر حداکثر بارش روزانه با استفاده از سیستم‌های هوشمند و مقایسه آن با مدل درختی M5؛ مطالعه موردی ایستگاه‌های اهر و حلفا، فصلنامه علمی پژوهشی مهندسی آبیاری و آب. ۴(۱۴): ۸۳-۹۸.
۷. شریفیان ح. و هزارجریبی ا. ۱۳۸۸. اصول و عملیات هوا و اقلیم‌شناسی از دیدگاه کشاورزی. انتشارات آژند.
۸. شهرابی ج. و ذوالقدر شجاعی ع. ۱۳۹۰. داده‌کاوی

16. Altunkaynak A. and Nigussie T. 2015. Prediction of daily rainfall by a hybrid wavelet-season-neuro technique. *Journal of Hydrology*. 529: 287-301.
17. Chattopadhyay S. 2007. Feed forward Artificial Neural Network model to predict the average summer-monsoon rainfall in India, *Acta Geophys*. 55(3): 369-382.
18. Govindaraju R. S. 2000. Artificial neural network in hydrology. I: Preliminary Concepts. *Journal of Hydrologic Engineering*. 5(2): 115-123.
19. Mislan H. Hardwinarto Sigit S. and Aipassa M. 2015. Rainfall Monthly Prediction Based on Artificial Neural Network: A Case Study in Tenggara Station, East Kalimantan-Indonesia. *Procedia Computer Science*. 59: 142-151.
20. Mohanty U. C. and Mohapatra M. 2007. Prediction of occurrence and quantity of daily summer monsoon precipitation over Orissa (India). *Meteorological Applications*. 14: 95-103.
21. Neal R. M. 1997. Monte carlo implementation of gaussian process models for bayesian regression and classification. University of Toronto. Toronto: Department of Statistics and Department of Computer Science. Technical report no, 9702.
22. Pal M. and Deswal S. 2010. Modelling pile capacity using Gaussian process regression. *COMPUT GEOTECH*, 37: 942-947.
23. Thakurta P. 2008. Long lead monsoon rainfall prediction for meteorological subdivisions of India using deterministic artificial neural network model. *Meteorology and Atmospheric Physics*. 101: 93-108.
24. Trafalis T. B. Santosa B. and Richman M. B. 2005. Learning networks in rainfall estimation. *CMS*. 2: 229-251.
25. Vapnik V. N. 1995. *The Nature of Statistical Learning Theory*. Springer, New York. 314 p.
26. VenkataRamana R. Krishna B. Kumar S. R. and Pandey N. G. 2013. Monthly Rainfall Prediction Using Wavelet Neural Network Analysis, *Water Resources Management*. 27: 3697-3711.
- پیشرفته (مفاهیم و الگوریتم‌ها)، انتشارات جهاد دانشگاهی واحد صنعتی امیرکبیر. ۴۵۷ ص.
۹. فاتحی مرج ا. و مهدیان م. ۱۳۸۸. پیش‌بینی بارش پاییزه با استفاده از شاخص‌های انسو به روش شبکه عصبی در حوضه دریاچه ارومیه. پژوهش‌های آب‌خیزداری. ۲۲: ۴۲-۵۲.
۱۰. فلاح قاله‌ری غ. موسوی بایگی م. و حبیبی نوخندان م. ۱۳۸۸. مقایسه نتایج به دست آمده از کاربرد سیستم استنباط فازی ممدانی و شبکه‌های عصبی مصنوعی در پیش‌بینی بارش فصلی، مطالعه موردی: منطقه خراسان، تحقیقات منابع آب ایران. ۴۰-۵۲: ۵.
۱۱. فلاح قاله‌ری غ. و خوشحال دستجردی ج. ۱۳۸۸. پیش‌بینی بارش بهاره استان خراسان رضوی بر اساس سیگنال‌های بزرگ مقیاس اقلیمی با استفاده از شبکه‌های عصبی مصنوعی. فصلنامه پژوهش‌های جغرافیای طبیعی. ۴۱(۶۹): ۱۱۵-۱۳۳.
۱۲. قلی‌زاده م. و دارند م. ۱۳۸۹. پیش‌بینی بارش ماهانه با استفاده از شبکه‌های عصبی مصنوعی (مورد: تهران). پژوهش‌های جغرافیای طبیعی. ۶۳: ۵۱-۷۱.
۱۳. گلابی م. آخوندعلی ع. رادمنش ف. کاشفی‌پور م. ۱۳۹۳. مقایسه دقت پیش‌بینی مدل‌های باکس-جنکینز در مدل‌سازی بارندگی فصلی (مطالعه موردی: ایستگاه‌های منتخب استان خوزستان). فصلنامه تحقیقات جغرافیایی. ۲۹(۱۱۴): ۶۱-۷۲.
۱۴. ماه‌آورپور ز. ۱۳۹۳. احتمال وقوع بارش‌های روزانه ایران و پیش‌بینی آن با مدل زنجیره مارکوف. فصلنامه تحقیقات جغرافیایی. ۲۹(۴): ۲۲۹-۲۴۰.
۱۵. مقصود ف. یزدانی م. رحیمی م. ملکیان آ. و ذوالفقاری علی ا. ۱۳۹۵. مقایسه کارایی شبکه عصبی مصنوعی، سری زمانی تجمعی مدل ترکیبی ANN-ARIMA در مدل‌سازی و پیش‌بینی شاخص منبع آب زیرزمینی (GRI) (مطالعه موردی: جنوب استان قزوین). علوم و مهندسی آب‌خیزداری ایران. ۱۰(۳۳): ۴۷-۵۷.

