



Using Machine Learning to Improve Precipitation Estimation by ERA5 Data at Meteorological Stations in West Azerbaijan Province

Moslem Mohammadi^{1,*} 

1. Assistant Professor, Department of Computer Engineering, Engineering Faculty, Payame Noor University, Tehran, Iran.

*Corresponding Author: mo_mohammadi@pnu.ac.ir

Keywords:

Precipitation estimation; ERA5 reanalysis; Machine learning; Artificial neural networks; Support vector regression; Shepard interpolation; West Azerbaijan Province.

Introduction

Precipitation is one of the most critical components of the hydrological cycle, directly influencing water resources, agriculture, and climate-related risk management. Accurate precipitation estimation is essential for hydrological modeling, flood forecasting, drought monitoring, and agricultural planning. In mountainous and data-scarce regions, however, reliable precipitation information is difficult to obtain due to the limited density of ground-based meteorological stations and the inherent uncertainties of remote sensing products. In recent years, reanalysis datasets such as ERA5, produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), have provided long-term, globally consistent climate variables. Despite their advantages, these datasets often suffer from systematic biases, particularly in regions with complex topography, such as West Azerbaijan Province in northwestern Iran. Machine learning (ML) methods, capable of capturing nonlinear and multivariate relationships, have emerged as effective tools for improving precipitation estimation by integrating satellite or reanalysis data with ground observations. The objective of this study was to assess the potential of two machine learning techniques—Artificial Neural Networks (ANN) and Support Vector Regression (SVR)—in reducing ERA5 precipitation biases and improving estimation accuracy in West Azerbaijan Province. The performance of these approaches was also compared with a traditional interpolation method, Shepard's technique, to evaluate the relative benefits of modern ML-based approaches.

Received:

20 April 2025

Revised:

27 August 2025

Accepted:

29 September 2025

How to cite this article:

Mohammadi, M. (2026). Using machine learning to improve precipitation estimation by ERA5 data at meteorological stations in West Azerbaijan Province. *Journal of Drought and Climate Change Research (JDCR)*, 3 (Spatial Issue), 37-56. [10.22077/jdcr.2025.9231.1137](https://doi.org/10.22077/jdcr.2025.9231.1137).



Materials and Methods

The study area, West Azerbaijan Province, is characterized by diverse topography, ranging from lowland plains to high mountain ranges, leading to significant spatial variability in precipitation patterns. Ground-based daily precipitation records from 16 synoptic meteorological stations were collected for the period 2010–2023. ERA5 reanalysis precipitation data with a spatial resolution of 0.25° and hourly temporal resolution were aggregated to daily and monthly scales to match station data. Three modeling approaches were employed: Shepard interpolation, SVR, and ANN (Multilayer Perceptron). For the machine learning models, several predictor variables were tested, including ERA5 precipitation values, station elevation, latitude, longitude, and temporal lag features (previous-day and next-day ERA5 estimates for daily scale; previous-month and next-month ERA5 estimates for monthly scale). Data were divided into training and testing subsets using k-fold cross-validation to ensure generalizability. Model evaluation was based on four statistical metrics: Root Mean Square Error (RMSE), Normalized RMSE (NRMSE), Nash–Sutcliffe Efficiency (NSE), and Pearson Correlation Coefficient (CC). These indicators provided complementary insights into model accuracy, relative error, predictive skill, and correlation strength.

Results and Discussion

The results clearly demonstrated the superiority of machine learning models over the conventional Shepard interpolation method. Among the ML approaches, ANN consistently produced the best performance at both daily and monthly scales. At the daily scale, the ANN achieved the lowest RMSE (2.73 mm), the highest correlation coefficient ($CC = 0.71$), and the highest NSE (0.50), while maintaining a low NRMSE (0.05). In contrast, the SVR model achieved a moderate improvement over Shepard interpolation but was less accurate than ANN, particularly in capturing extreme precipitation events. Shepard interpolation, by relying solely on spatial proximity without considering atmospheric predictors, failed to adequately reproduce precipitation variability. At the monthly scale, all models showed improved performance due to reduced short-term variability; however, ANN again outperformed alternatives. The ANN model achieved an RMSE of 10.20 mm, CC of 0.84, and NSE of 0.67. SVR also provided reasonable accuracy, but its performance was less stable, particularly in stations located in mountainous terrain. Shepard interpolation continued to underperform compared with ML-based corrections. Feature analysis revealed that including temporal lag variables significantly enhanced predictive performance, especially at the daily scale. This highlights the importance of accounting for temporal persistence in precipitation processes. Elevation and spatial coordinates also contributed to improved accuracy, confirming the influence of topography on precipitation distribution.

Conclusion

The findings emphasize the effectiveness of machine learning approaches, particularly ANN, in correcting ERA5 precipitation estimates in regions with complex terrain and limited station density. By capturing nonlinear dependencies and incorporating temporal and spatial features, ANN substantially reduced systematic biases and improved alignment with ground observations. The comparison with SVR showed that while kernel-based methods are useful, they are less flexible than neural networks in modeling high-dimensional feature interactions. Shepard interpolation, though simple and computationally efficient, was unable to address the structural biases of reanalysis data, underscoring the need for more advanced methods. This research has both methodological and practical implications. From a methodological perspective, it demonstrates the importance of combining multiple predictors—including ERA5 estimates, station metadata, and temporal lags—for effective bias correction. From a practical standpoint, the improved precipitation datasets can support hydrological modeling, water resource management, agricultural planning, and climate change impact assessments in West Azerbaijan and other regions with similar geographic and data constraints. Future research could expand on these results by exploring deep learning architectures such as Long Short-Term Memory (LSTM) networks, which may better capture temporal dynamics, or by incorporating additional predictors such as temperature, humidity, and atmospheric circulation indices. Moreover, applying

spatial downscaling techniques in combination with ML approaches could further refine precipitation estimates at local scales. In conclusion, this study confirms that machine learning, and especially ANN-based methods, offer a robust and reliable approach to enhancing reanalysis precipitation datasets, thereby contributing to more accurate climate and hydrological analyses in data-sparse mountainous regions.



استفاده از یادگیری ماشین برای بهبود تخمین بارش توسط داده‌های ERA5 در ایستگاه‌های هواشناسی استان آذربایجان غربی

مسلم محمدی^۱ *

۱- استادیار گروه کامپیوتر، دانشکده مهندسی، دانشگاه پیام نور، تهران، ایران.

* نویسنده مسئول: mo_mohammadi@pnu.ac.ir

چکیده

تخمین دقیق بارش نقشی اساسی در پیش‌بینی آب‌وهوا، مدیریت منابع آبی و کاهش اثرات بلایای طبیعی ایفا می‌کند. روش‌های سنتی مبتنی بر مدل‌های فیزیکی و داده‌های راداری دارای محدودیت‌هایی در دقت و مقیاس‌پذیری هستند. در سال‌های اخیر، روش‌های یادگیری ماشین به‌عنوان ابزاری توانمند برای بهبود تخمین بارش مطرح شده‌اند. این مقاله به بررسی روش‌های یادگیری ماشین برای تخمین بارش، چالش‌های موجود و پیشرفت‌های اخیر در این حوزه می‌پردازد. در این مقاله، روش‌های رگرسیون شبکه‌های عصبی مصنوعی و رگرسیون بردار پشتیبان برای نگاهت میزان بارش تخمین زده شده توسط داده‌های ERA5 به بارش‌های سنجش شده در ایستگاه‌های هواشناسی مورد ارزیابی قرار می‌گیرد. دقت داده‌های ماهواره‌ای در مقایسه با داده‌های زمینی نیاز به ارزیابی و بهبود دارد. در این مطالعه، روش‌های مختلفی از جمله روش‌های آماری و یادگیری ماشین با استفاده از ترکیب ویژگی‌های مختلف همچون مقدار تخمین ماهواره‌ای روز/ماه جاری، ارتفاع ایستگاه هواشناسی، طول جغرافیایی، عرض جغرافیایی، مقدار تخمین ماهواره‌ای روز/ماه قبل و مقدار تخمین ماهواره‌ای روز / ماه بعد برای بهبود دقت داده‌های ERA5 مورد بررسی قرار می‌گیرند. داده‌های بارش از طول جغرافیایی ۴۴ درجه تا ۴۷ درجه شرقی و عرض جغرافیایی ۳۶ تا ۴۰ درجه شمالی با وضوح مکانی ۰/۲۵ درجه از سال ۲۰۱۰ تا ۲۰۲۳ مورد بررسی قرار گرفته‌اند. ناحیه مورد مطالعه شامل ۱۶ ایستگاه هواشناسی می‌باشد. نتایج نشان می‌دهد که روش رگرسیون شبکه‌های عصبی مصنوعی در مقایسه با روش درون‌یابی شپرد و رگرسیون ماشین بردار پشتیبان دقت تخمین بارش را تقریباً متناسب با تعداد ویژگی‌های ورودی افزایش داده و در بهترین حالت منجر به دستیابی به کمترین RMSE (mm 73/2) برای روزانه و mm 10/20 برای ماهانه)، بالاترین CC (۰/۷۱) برای روزانه و ۰/۸۴ برای ماهانه)، بالاترین NSE (۰/۵۰) برای روزانه و ۰/۶۷ برای ماهانه) و کمترین NRMSE (۰/۰۵) برای روزانه و ۰/۱۰ برای ماهانه) شد.

واژه‌های کلیدی:

تخمین بارش، شبکه‌های عصبی مصنوعی، ماشین بردار پشتیبان، یادگیری ماشین، داده‌های ERA5.

تاریخ دریافت:

۱۴۰۴/۰۱/۳۰

تاریخ بازنگری:

۱۴۰۴/۰۶/۰۵

تاریخ پذیرش:

۱۴۰۴/۰۷/۰۷

مقدمه

بارش یک جزء اساسی از چرخه هیدرولوژیکی زمین است که بر کشاورزی، مدیریت منابع آب و مدل سازی آب و هوا تأثیر می‌گذارد. داده‌های دقیق بارش برای برنامه ریزی و تصمیم‌گیری مؤثر در این حوزه‌ها ضروری است. به‌طور سنتی، اندازه‌گیری‌های بارش بر ایستگاه‌های هواشناسی زمینی تکیه می‌کنند که دقت بالایی ارائه می‌دهند، اما در پوشش مکانی، به‌ویژه در مناطق دورافتاده یا از نظر توپوگرافی پیچیده، محدود هستند. تخمین‌های بارش مبتنی بر ماهواره، مانند آن‌هایی که توسط مجموعه داده‌های تحلیل مجدد^۱ ERA5 از مرکز اروپا برای پیش‌بینی‌های هوای متوسط ارائه شده است، راه‌حلی را با ارائه پوشش مکانی و زمانی گسترده ارائه می‌دهد. با این حال، این تخمین‌ها اغلب در مقایسه با مشاهدات زمینی، سوگیری‌های را نشان می‌دهند که به روش‌های اصلاحی برای افزایش قابلیت اطمینان آنها نیاز دارد (Beck et al., 2020). در دهه‌های اخیر، استفاده از یادگیری ماشین در علوم جوی به‌عنوان یک ابزار کارآمد برای مدل‌سازی پدیده‌های پیچیده‌ای مانند بارش باران مطرح شده است. الگوریتم‌های یادگیری ماشین می‌توانند روابط غیرخطی بین متغیرهای جوی را کشف کرده و دقت پیش‌بینی بارش را افزایش دهند. یادگیری ماشین به‌عنوان یکی از شاخه‌های پویا و روبه‌رشد هوش مصنوعی، به مطالعه و توسعه الگوریتم‌هایی می‌پردازد که قادرند از داده‌ها یاد بگیرند و پیش‌بینی‌های دقیقی انجام دهند. این حوزه با بهره‌گیری از روش‌های آماری و محاسباتی، به ماشین‌ها امکان می‌دهد تا الگوهای پنهان در داده‌ها را شناسایی کرده و از آن‌ها برای حل مسائل پیچیده استفاده کنند. یکی از کاربردهای مهم یادگیری ماشین، تخمین بارش در هواشناسی است که به‌دلیل نقش حیاتی بارش در مدیریت

منابع آبی، کشاورزی، و برنامه‌ریزی شهری، از اهمیت ویژه‌ای برخوردار است (Saha et al., 2024). با توجه به پیچیدگی الگوهای بارش و تأثیر عوامل متعدد بر آن، روش‌های سنتی اغلب با محدودیت‌هایی مواجه هستند. در این راستا، یادگیری ماشین با ارائه روش‌های پیشرفته مانند شبکه عصبی پرسپترون چندلایه^۲ و ماشین بردار پشتیبان^۳، ابزارهای قدرتمندی را برای بهبود دقت و کارایی در تخمین بارش فراهم می‌کند (Houénafa et al., 2025).

بارش یکی از مهم‌ترین پارامترهای هواشناسی است که نقش کلیدی در مطالعات هیدرولوژی، کشاورزی و مدیریت منابع آب ایفا می‌کند (Jiang et al., 2023). داده‌های بارش معمولاً از دو منبع اصلی به‌دست می‌آیند: ایستگاه‌های هواشناسی زمینی و ماهواره‌ها. داده‌های زمینی به‌دلیل دقت بالا، معیار مناسبی برای ارزیابی داده‌های ماهواره‌ای هستند. جمع‌آوری مجموعه داده‌های بارش به‌موقع و قابل اعتماد در مناطق کم‌داده و کوهستانی ضروری و چالش‌برانگیز است (Zhou et al., 2020). روش‌های مختلفی مانند باران‌سنج، رادارهای زمینی و سنسور از دور ماهواره‌ای برای جمع‌آوری داده‌های بارش استفاده شده است (Kalhori et al., 2025). همه این روش‌های تخمین بارندگی مزایا و معایب خاص خود را دارند (Yu et al., 2020). روش‌های تخمین بارندگی دارای عدم قطعیت‌های متنوعی از منابع داده، نمونه‌برداری، الگوریتم‌های بازبازی، و شرایط آب و هوایی طبیعی و جغرافیایی هستند (Wang et al., 2021). بنابراین، ارزیابی کمی از دقت تخمین بارش می‌تواند در طیف گسترده‌ای از کاربردهای تحقیقاتی هیدرولوژیکی و اقلیمی، از جمله مدیریت منابع آب (Bagheri Khanghani et al., 2025; Espinosa et al., 2024; Donat et al.,

² MLP (Multi-Layer Perceptron)

³ SVM (Support Vector Machine)

¹ ECMWF Reanalysis v5

آن به طور گسترده در تحقیقات هواشناسی و هیدرولوژیکی استفاده می‌شود (Hersbach et al., 2020; Amjad et al., 2020). علی‌رغم مزایای آن، مطالعات نشان داده‌اند که تخمین‌های بارش ERA5 می‌تواند با اندازه‌گیری‌های زمین، به‌ویژه در مناطقی با توپوگرافی پیچیده یا شبکه‌های ایستگاهی پراکنده اختلاف زیادی داشته باشد (Tang et al., 2020). این اختلاف به‌ویژه در مناطقی مانند استان آذربایجان غربی که در شمال غربی ایران قرار دارد، قابل توجه است. آذربایجان غربی دارای توپوگرافی متنوعی از جمله رشته کوه‌های زاگرس، دشت میان‌دوآب و دریاچه ارومیه است که بر الگوهای بارش محلی تأثیر می‌گذارد و تخمین مبتنی بر ماهواره را پیچیده می‌کند.

روش‌های یادگیری ماشین، مانند ماشین‌های بردار پشتیبان (SVM) و پرسپترون‌های چند لایه (MLP)، به‌عنوان ابزار قدرتمندی برای تصحیح داده‌های ماهواره‌ای با مدل‌سازی روابط پیچیده و غیرخطی بین ویژگی‌های ورودی و متغیرهای هدف ظاهر شده‌اند (Kumar et al., 2019; Komasi et al., 2025). این مطالعه به بررسی کاربرد رگرسیون‌های SVM و MLP برای نقشه‌برداری تخمین بارش ERA5 در اندازه‌گیری‌های زمینی از ۱۶ ایستگاه هواشناسی آذربایجان غربی در بازه زمانی ۲۰۱۰ تا ۲۰۲۳ می‌پردازد. مطالعات متعددی به مقایسه داده‌های ERA5 با داده‌های زمینی پرداخته‌اند. به‌عنوان مثال سوسی و همکاران در مطالعه‌ای جامع نشان دادند که داده‌های ERA5 در مقایسه با نسخه‌های قبلی مانند ERA-Interim، بهبود قابل توجهی در دقت تخمین بارش دارند (Soci et al., 2024). با این حال، هنوز اختلافاتی بین داده‌های ERA5 و داده‌های زمینی وجود دارد که نیاز به اصلاح دارد. یکی از چالش‌های اصلی در استفاده از داده‌های ماهواره‌ای، ناهمگونی مکانی و زمانی داده‌ها است در مطالعه‌ای نشان دادند که ناهمگونی مکانی داده‌های ماهواره‌ای می‌تواند منجر به خطاهای قابل توجهی در

(Yuan et al., 2025;) و پیش بینی خشک‌سالی (Yang et al., 2024; Huang et al., 2024) مفید باشد. به‌طور سنتی، محصولات بارش مبتنی بر سنج شبکه‌ای، به‌دلیل منبع داده‌های قابل اعتمادشان (یعنی ایستگاه‌های جهانی) تخمین واقع‌بینانه‌ای را برای سوابق تاریخی بلندمدت ارائه می‌کنند. با این حال، کیفیت آنها به‌دلیل تراکم شبکه باران سنج زمینی و عدم دسترسی به موقع محدود شده است (Yousefi-Kebria et al., 2023). در حال حاضر، چندین محصول بارشی مبتنی بر چند ماهواره‌ای شبه جهانی در دهه‌های اخیر ظهور کرده‌اند که روش‌های جدیدی را برای ثبت سریع اطلاعات بارش مداوم و ارائه مؤثرترین ابزار برای مشاهده بارش در مقیاس جهانی ارائه می‌کنند (Sun et al., 2018). محصولات بارشی مبتنی بر ماهواره اکنون از دوران مأموریت اندازه‌گیری بارندگی استوایی^۱ (TRMM) به عصر اندازه‌گیری بارش جهانی^۲ (GPM) منتقل شده‌اند (Huffman et al., 2014). به‌طور کلی، طی چند دهه گذشته، محصولات بارش مبتنی بر ماهواره و سنج شامل بازیابی‌های یکپارچه چندماهواره‌ای برای مأموریت اندازه‌گیری باران جهانی^۳ (IMERG) برای بازیابی‌های چندماهواره‌ای یکپارچه برای GPM، تحلیل بارش چندماهواره‌ای TRMM^۴ (TMPA) برای بارش چند ماهواره‌ای TRMM، نسل پنجم تحلیل مجدد جوی توسط مرکز اروپایی^۵ (ERA و ECMWF) و داده‌های مرکز پیش بینی آب و هوا^۶ (CPC) می‌باشد.

مجموعه داده ERA5، با وضوح مکانی ۰/۲۵ درجه و وضوح زمانی ساعتی، به‌دلیل پوشش جهانی و دسترسی

¹ Tropical Rainfall Measuring Mission

² Global Precipitation Measurement

³ Integrated Multi-satellite Retrievals for GPM

⁴ TRMM Multi-satellite Precipitation Analysis

⁵ European Centre for Medium-Range Weather Forecasts

⁶ Climate Prediction Center

لایه خروجی تولید می‌شود. هر نورون با استفاده از یک تابع فعال‌سازی (مانند سیگموئید یا ReLU)، ورودی‌های وزن‌دار خود را به خروجی تبدیل می‌کند. فرآیند یادگیری در MLP از طریق الگوریتم‌هایی مانند پس‌انتشار^۱ خطا انجام می‌شود که وزن‌های اتصالات بین نورون‌ها را تنظیم می‌کند تا خطای پیش‌بینی کاهش یابد. در تخمین بارش، شبکه عصبی MLP می‌تواند از داده‌های هواشناسی مانند دما، رطوبت، فشار هوا، و سرعت باد به‌عنوان ورودی استفاده کند و میزان بارش را به‌عنوان خروجی پیش‌بینی کند. یکی از مزایای کلیدی این روش، توانایی آن در مدیریت داده‌های دارای نوفه و ناقص است که در هواشناسی به‌دلیل خطاهای اندازه‌گیری یا محدودیت‌های ابزارها رایج است. علاوه‌بر این، MLP قادر است روابط غیرخطی پیچیده بین متغیرها را مدل‌سازی کند، که در پدیده‌هایی مانند بارش، که تحت تأثیر تعاملات متعدد عوامل محیطی قرار دارد، بسیار مفید است. رگرسیون ماشین بردار پشتیبان (SVM) یک روش یادگیری ماشین است که بر اساس تئوری یادگیری آماری توسعه یافته و برای مسائل رگرسیون و طبقه‌بندی به کار می‌رود. در حالت رگرسیون، SVM به‌دنبال یافتن تابعی است که داده‌های آموزشی را با کمترین خطا پیش‌بینی کند، در حالی که از پیچیدگی بیش از حد مدل جلوگیری می‌کند (Nouhani et al., 2024). این روش از مفهوم حاشیه استفاده می‌کند و سعی دارد فاصله بین داده‌ها و ابر صفحه تصمیم را حداکثر کند، که این امر به بهبود قابلیت تعمیم مدل کمک می‌کند. SVM همچنین با استفاده از توابع کرنل (مانند کرنل گاوسی یا RBF)، می‌تواند روابط غیرخطی بین متغیرها را مدل‌سازی کند.

در تخمین بارش، SVM از داده‌های تاریخی بارش و متغیرهای هواشناسی برای ایجاد مدلی استفاده می‌کند که

تخمین بارش شود (Kidd et al., 2017). بنابراین، نیاز به روش‌هایی برای کاهش این ناهمگونی وجود دارد. روش‌های مختلفی برای بهبود دقت داده‌های ماهواره‌ای پیشنهاد شده است (Chen et al., 2018). از جمله روش‌های یادگیری ماشین، مانند شبکه‌های عصبی مصنوعی برای بهبود دقت داده‌های بارش ماهواره‌ای استفاده کردند. نتایج این مطالعه نشان دادند که استفاده از روش‌های یادگیری ماشین می‌تواند دقت تخمین بارش را به‌طور قابل توجهی افزایش دهد.

این مقاله به بررسی روش‌های نگاشت داده‌های بارش ماهواره ERA5 به داده‌های زمینی پرداخته و نشان می‌دهد که استفاده از روش‌های ترکیبی می‌تواند دقت تخمین بارش را افزایش دهد. در این مقاله، یک مدل ترکیبی برای بهبود تخمین بارش ارائه می‌شود که از داده‌های ERA5 و ایستگاه‌های زمینی به‌همراه الگوریتم‌های یادگیری ماشین بهره می‌برد. تعریف و استفاده از ویژگی‌های ترکیبی در رگرسیون‌های مختلف برای ارضای اهداف تعریف‌شده، نوآوری این تحقیق می‌باشد. اهداف این پژوهش سه مورد است:

- ارزیابی دقت برآورد بارش ERA5 در برابر اندازه‌گیری‌های زمینی در آذربایجان غربی.
- توسعه و ارزیابی رگرسیون‌های SVM و MLP برای بهبود دقت تخمین بارش ERA5.
- مقایسه عملکرد این مدل‌های یادگیری ماشینی با روش درونیابی سنتی شپرد.

مبانی نظری

شبکه عصبی MLP یکی از انواع شبکه‌های عصبی مصنوعی است که از چندین لایه نورون تشکیل شده است: لایه ورودی، لایه‌های پنهان، و لایه خروجی. این ساختار سلسله‌مراتبی به شبکه امکان می‌دهد تا الگوهای پیچیده و غیرخطی را در داده‌ها یاد بگیرد. در یک شبکه MLP، داده‌ها از لایه ورودی وارد شده، در لایه‌های پنهان پردازش می‌شوند، و در نهایت پیش‌بینی (مانند میزان بارش) از

^۱ Backpropagation

تحقیق، هم‌بستگی پیرسون) داده‌های نگاشت شده ERA5 با داده‌های واقعی ایستگاه‌های هواشناسی، میزان مطابقت دو مجموعه داده اندازه‌گیری می‌شود. از معیار خطای جذر میانگین مربعات^۲ (RMSE) و معیار خطای جذر میانگین مربعات نرمال شده^۳ (NRMSE) برای تعیین کمیت تفاوت بین برآوردهای ERA5 و معیار نش-ساتکلیف^۴ (NSE) برای ارزیابی دقت اندازه‌گیری‌های واقعی استفاده شده است. مقادیر بالاتر NSE نشان‌دهنده عملکرد بهتر مدل است و مقادیر نزدیک به ۱، تطابق کامل را نشان می‌دهد.

منطقه مورد مطالعه

منطقه مورد مطالعه، ایستگاه‌های هواشناسی استان آذربایجان غربی است. این منطقه بین طول جغرافیایی ۴۴ درجه تا ۴۷ درجه شرقی و عرض جغرافیایی ۳۶ تا ۴۰ درجه شمالی مطابق با شکل ۱، واقع شده است. از نظر توپوگرافی، دارای تنوع ارتفاعی از ارتفاع ۱۱۰۳/۴ تا ۱۸۸۸/۵ متر می‌باشد.

در این مطالعه، داده‌های بارش ماهواره ERA5 و داده‌های بارش اندازه‌گیری شده در ایستگاه‌های هواشناسی «ارومیه»، «اشنویه»، «بوکان»، «پیرانشهر»، «نکاب»، «چالدران»، «خوی»، «سردشت»، «سلماس»، «شاهیندژ»، «چاپاره»، «کهریز»، «ماکو»، «مهاباد»، «میاندوآب» و «نقده» برای یک دوره ۱۴ ساله از سال ۲۰۱۰ تا ۲۰۲۳، از سازمان هواشناسی استان آذربایجان غربی جمع‌آوری شده است. داده‌های ERA5 با رزولوشن زمانی ساعتی و رزولوشن مکانی ۰/۲۵ درجه در دسترس هستند که رزولوشن زمانی به داده‌های روزانه و ماهانه بارش تجمیع شد تا با داده‌های ایستگاه‌های هواشناسی هم مقیاس شود.

بتواند بارش آینده را پیش‌بینی کند. یکی از ویژگی‌های برجسته SVM، توانایی آن در مدیریت داده‌های با ابعاد بالا است، که در مواردی که تعداد ویژگی‌های ورودی زیاد باشد (مانند داده‌های هواشناسی)، بسیار کاربردی است. این روش در برابر یادگیری بیش از حد مقاوم است، به‌خصوص زمانی که تعداد نمونه‌های آموزشی محدود باشد، که در مطالعات هواشناسی، جایی که داده‌های دقیق همیشه در دسترس نیستند، مزیت بزرگی محسوب می‌شود.

مواد و روش‌ها

برای انجام رگرسیون غیرخطی می‌توان از مدل‌های مختلفی استفاده کرد که قادر به مدل‌سازی روابط پیچیده‌تر و غیرخطی بین ویژگی‌ها و متغیر وابسته هستند. رگرسیون ماشین بردار پشتیبان، رگرسیون درخت تصمیم، مدل جنگل تصادفی، روش مبتنی بر تقویت گرادیان، مدل KNN برای رگرسیون و رگرسیون با شبکه‌های عصبی مصنوعی برخی از روش‌های رگرسیون غیرخطی هستند (Modaresi et al., 2018). برای مدل‌های ساده‌تر و زمانی که داده‌ها به‌صورت سلسله‌مراتبی یا گروه‌بندی شده هستند، درخت تصمیم یا جنگل تصادفی مناسب است؛ ولی برای داده‌های پیچیده‌تر و زمانی که ارتباطات غیرخطی بسیار پیچیده هستند، مدل‌های مبتنی بر تقویت گرادیان، SVR یا شبکه‌های عصبی مصنوعی می‌تواند کارایی بیشتری داشته باشند.

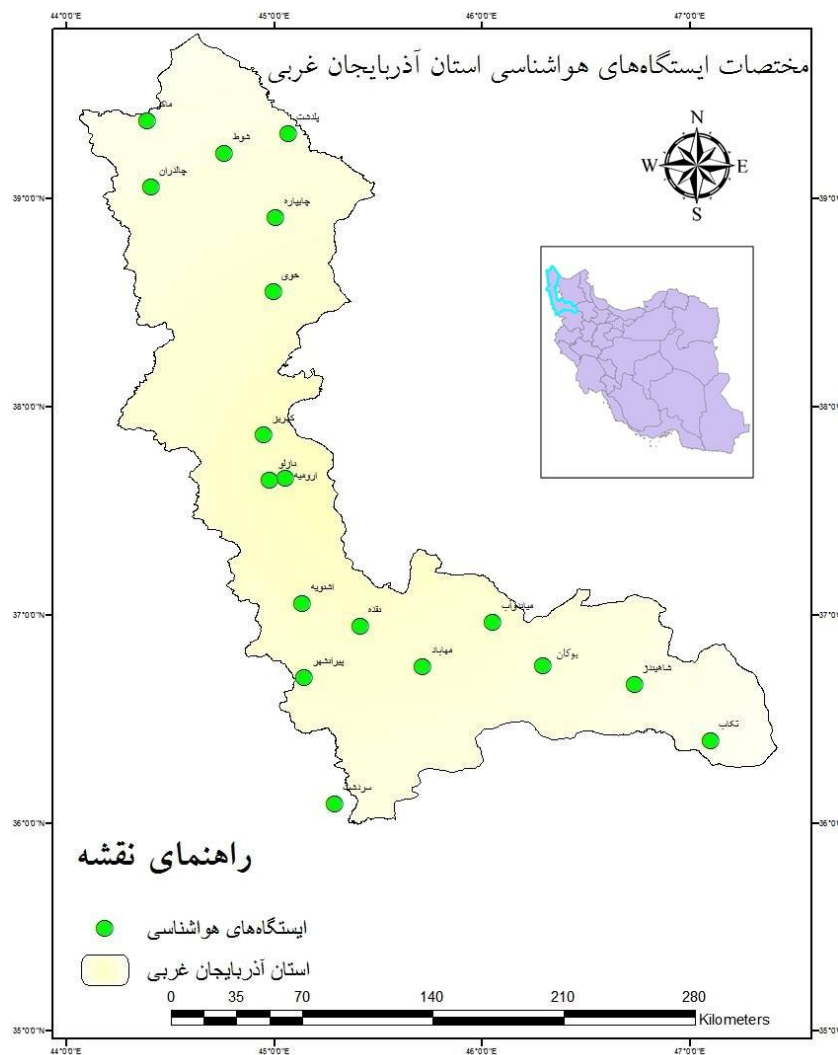
در این تحقیق از دو روش ماشین بردار پشتیبان و شبکه عصبی مصنوعی برای بهبود نتایج استفاده شده است. ضریب هم‌بستگی^۱ (CC) نشان‌دهنده میزان ارتباط خطی بین داده‌های بارش ERA5 و داده‌های بارش مشاهده‌شده است. با محاسبه ضرایب هم‌بستگی (در این

² Root Mean Squared Error

³ Normalized RMSE

⁴ Nash-Sutcliffe Efficiency

¹ Correlation Coefficient



شکل ۱. پراکندگی ایستگاه‌های هواشناسی مورد مطالعه در استان آذربایجان غربی

Fig 1. Distribution of studied meteorological stations in West Azerbaijan Province

درونیابی شپرد یک روش درونیابی با وزن دهی معکوس فاصله^۱ است که در آن مقدار یک نقطه مجهول بر اساس نقاط معلوم اطرافش و با وزن دهی معکوس فاصله آنها محاسبه می‌شود (Iacono et al., 2008).

بعد از تراز کردن مقیاس‌های زمانی و مکانی با درونیابی دوبعدی شپرد، دو گروه داده سنجش زمینی و تخمین ماهواره‌ای میزان بارش در هر ایستگاه هواشناسی برای بازه‌های زمانی مختلف وجود دارد. در گام نخست، ضریب هم‌بستگی و میزان RMSE, NRMSE و NSE

وضوح فضایی هر دو مجموعه داده باید تراز باشد. از آنجایی که داده‌های ERA5 اغلب در قالب‌های شبکه‌ای ارائه می‌شوند، اگر داده‌های حقیقی زمینی مبتنی بر نقطه (از یک ایستگاه هواشناسی) باشد، باید داده‌های ERA5 را در همان منطقه جغرافیایی به‌طور میانگین محاسبه کرده یا از روش‌های درونیابی برای استخراج مقادیری از شبکه ERA5 که با موقعیت ایستگاه زمینی مطابقت دارد، استفاده گردد (Mianabadi et al., 2024). برای تراز کردن اولیه داده‌های محدوده مورد مطالعه از روش درون یابی شپرد برای درونیابی تخمین مقادیر بارندگی در ایستگاه‌های هواشناسی جداگانه استفاده شده است.

¹ Inverse distance weighting (IDW)

الگوریتم‌های یادگیری ماشین شامل ترکیبی از ۶ ویژگی جدول ۱ می‌باشد که از بین حالت‌های مختلف ۵ اجرا انتخاب شده و نتایج گزارش شده است. نتایج به‌صورت مجزا برای داده‌های ماهانه بارش و روزانه با استفاده از روش اعتبارسنجی متقابل، سنجش و گزارش شده است.

داده‌های سنجش زمینی و تخمین ماهواره‌ای میزان بارش تجمیع شده در تمامی ایستگاه‌ها و زمان‌های مختلف محاسبه می‌شود. در گام بعدی، از دو روش یادگیری ماشین بردار پشتیبان و شبکه‌های عصبی مصنوعی برای بهبود درونیایی شپرد استفاده می‌شود. ورودی

جدول ۱. ویژگی‌های مورد استفاده در مدل‌های یادگیری ماشین
Table 1. Features used in machine learning models

نام ویژگی Feature Name	ویژگی‌های ورودی Input Features	نام ویژگی Feature Name	ویژگی‌های ورودی Input Features
CP	مقدار تخمین ماهواره‌ای روز / ماه جاری Satellite estimate value for the current day/month	عرض جغرافیایی latitude	LA
SA	ارتفاع ایستگاه هواشناسی station elevation	مقدار تخمین ماهواره‌ای روز / ماه قبل Satellite estimate value for the previous day/month	PE
LO	طول جغرافیایی Longitude	مقدار تخمین ماهواره‌ای روز / ماه بعد Satellite estimate value for the next day/month	NE

نتایج

نتایج حاصل از ۱۰ اجرای متفاوت بر روی داده‌های روزانه و ماهانه با ورودی‌های مختلف در جداول ۲ الی ۱۱ آورده شده است. هر ستون که با Fold نشان داده شده است بیانگر یک اجرا با ۹۰ درصد داده‌های آموزشی و ۱۰ درصد داده‌های آزمایشی می‌باشد. ستون آخر میانگین تمام ۱۰ اجرای متفاوت هست. هر سطر جداول مذکور نشان دهنده معیارهای ارزیابی مختلف بر اساس سه روش Raw Shepard¹، SVR² و MLPregressor³ است. Raw Shepard مدل پایه است که فقط از درونیایی مکانی ساده (شپرد) روی داده‌های ERA5 استفاده می‌کند و هیچ یادگیری ماشینی در آن انجام نشده است. این مدل معیار

مقایسه برای سنجش بهبود حاصل از روش‌های SVR و MLPregressor است. به دلیل عدم استفاده از یادگیری ماشین برای پارامترهای مختلف، مقدار RMSE، NRMSE، NSE و CC مربوط به سطر Raw Shepard در تمامی جدول‌ها یکسان و ثابت می‌باشد.

مدل‌های یادگیری با یک ویژگی ورودی

در این مدل فقط از مقدار تخمین ماهواره‌ای روز / ماه جاری (CP) استفاده شده است. یعنی در این مدل مقدار تخمین ماهواره‌ای محاسبه شده توسط درونیایی شپرد به مدل داده می‌شود و مقدار اصلاح شده متناسب با داده‌های آموزشی به‌عنوان خروجی برگردانده می‌شود.

در جدول ۲، MLPregressor با RMSE متوسط ۳/۵۲mm و NSE متوسط ۰/۱۶ بهترین عملکرد را دارد و نسبت به مدل پایه بهبود نشان می‌دهد.

¹ روش مبتنی بر داده‌های درونیایی شده با استفاده از روش درونیایی

شپرد بدون پردازش اضافی

² روش بهبود نتایج شپرد با استفاده از رگرسیون بردار پشتیبان

³ روش بهبود نتایج شپرد با استفاده از رگرسیون شبکه‌های عصبی مصنوعی

جدول ۲. نتایج تحلیل داده‌های روزانه بارش با یک پارامتر ورودی (CP)

Table 2. Results of daily data analysis with one input parameter (CP)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	4.03	3.92	3.95	4.18	4.18	3.91	4.12	3.98	3.91	4.20	4.04
RMSE: SVR (mm)	3.76	3.69	3.61	3.90	3.85	3.63	3.72	3.54	3.66	3.73	3.71
RMSE: MLPregressor (mm)	3.98	3.42	3.34	3.68	3.60	3.40	3.55	3.33	3.39	3.54	3.52
CC: Raw Shepard	0.47	0.46	0.43	0.40	0.37	0.46	0.44	0.38	0.45	0.37	0.42
CC: SVR	0.43	0.40	0.33	0.34	0.25	0.41	0.43	0.31	0.35	0.30	0.35
CC: MLPregressor	-0.26	0.46	0.44	0.41	0.38	0.46	0.44	0.39	0.46	0.38	0.36
NRMSE: Raw Shepard	0.06	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.05	0.06	0.06
NRMSE: SVR	0.07	0.06	0.06	0.07	0.07	0.06	0.07	0.06	0.06	0.07	0.07
NRMSE: MLPregressor	0.07	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
NSE: Raw Shepard	-0.14	-0.14	-0.17	-0.40	-0.42	-0.12	-0.19	-0.30	-0.16	-0.34	-0.24
NSE: SVR	0.10	0.09	0.06	0.06	0.02	0.10	0.11	0.04	0.07	0.04	0.07
NSE: MLPregressor	0.00	0.21	0.19	0.17	0.14	0.21	0.19	0.15	0.20	0.14	0.16

جدول ۳. نتایج تحلیل داده‌های ماهانه بارش با یک پارامتر ورودی (CP)

Table 3. Results of monthly data analysis with one input parameter (CP)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	28.22	26.24	27.32	26.06	33.19	31.31	28.45	28.52	33.73	23.99	28.70
RMSE: SVR (mm)	21.37	25.38	23.72	20.06	24.26	26.81	22.52	34.04	30.89	18.85	24.79
RMSE: MLPregressor (mm)	22.37	23.23	22.81	18.99	24.57	25.93	20.93	52.18	28.23	19.40	25.86
CC: Raw Shepard	0.73	0.84	0.77	0.79	0.74	0.74	0.76	0.83	0.76	0.82	0.78
CC: SVR	0.76	0.79	0.75	0.78	0.75	0.73	0.75	0.70	0.71	0.83	0.75
CC: MLPregressor	0.73	0.83	0.77	0.79	0.74	0.74	0.76	0.00	0.76	0.81	0.69
NRMSE: Raw Shepard	0.15	0.14	0.15	0.14	0.18	0.17	0.15	0.15	0.18	0.13	0.15
NRMSE: SVR	0.10	0.12	0.11	0.10	0.12	0.13	0.11	0.16	0.15	0.09	0.12
NRMSE: MLPregressor	0.11	0.11	0.11	0.09	0.12	0.13	0.10	0.25	0.14	0.09	0.13
NSE: Raw Shepard	0.42	0.61	0.45	0.53	0.42	0.45	0.49	0.60	0.47	0.56	0.50
NSE: SVR	0.56	0.61	0.56	0.58	0.55	0.52	0.49	0.47	0.50	0.68	0.55
NSE: MLPregressor	0.52	0.67	0.60	0.62	0.54	0.55	0.56	-0.25	0.58	0.66	0.51

بیش از حد^۱ روی داده‌های آموزشی در آن اجرا است. این موضوع اهمیت گزارش میانگین چندین اجرا را نشان می‌دهد. افزودن روش‌های یادگیری ماشین حتی با یک ویژگی می‌تواند مفید باشد، اما بهبود آن محدود است.

مدل‌های یادگیری با دو ویژگی ورودی

در این مدل از مقدار تخمین ماهواره‌ای روز / ماه جاری (CP) و مقدار تخمین ماهواره‌ای روز/ماه قبل (PE) استفاده شده است. تحلیل داده روزانه مطابق جدول ۴، نشان می‌دهد هر دو مدل SVR و MLPregressor به‌طور چشمگیری بهبود یافته‌اند. RMSE هر دو به حدود mm ۳/۲ کاهش یافته و CC مدل SVR به ۰/۵۴ رسیده است.

با این حال، CC پایین و NSE منفی یا نزدیک به صفر نشان می‌دهد که داده روزانه بسیار پر نوفه و غیرقطعی است و حتی مدل‌های یادگیری ماشین با یک ورودی قادر به مدل کردن کامل الگوهای آن نیستند.

عملکرد ضعیف MLPregressor در برخی Foldها (مثلاً CC برابر با ۰/۲۶- در Fold1) نشان از ناپایداری اولیه مدل با داده‌های کم دارد. تحلیل داده‌های ماهانه جدول ۳ نشان می‌دهد که SVR با RMSE متوسط mm ۲۴/۷۹ بهترین عملکرد را دارد. داده ماهانه با حذف نوفه‌های روزانه، اثبات‌تر است و رابطه خطی‌تری بین ERA5 و داده زمینی برقرار است که SVR در مدل کردن آن موفق‌تر عمل می‌کند. MLPregressor در Fold8 عملکرد بسیار ضعیفی دارد که احتمالاً به دلیل یادگیری

¹ Overfitting

جدول ۴. نتایج تحلیل داده‌های روزانه بارش با ۲ پارامتر ورودی (CP, PE)

Table 4. Results of daily data analysis with 2 input parameters (CP, PE)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	4.03	3.92	3.95	4.18	4.18	3.91	4.12	3.98	3.91	4.20	4.04
RMSE: SVR (mm)	3.29	3.36	3.13	3.37	3.33	3.31	3.44	3.10	3.12	3.24	3.27
RMSE: MLPregressor (mm)	3.22	3.38	3.09	3.15	3.24	3.22	3.27	3.06	3.31	3.26	3.22
CC: Raw Shepard	0.40	0.34	0.41	0.41	0.39	0.41	0.39	0.35	0.41	0.36	0.39
CC: SVR	0.57	0.51	0.55	0.56	0.55	0.52	0.51	0.54	0.58	0.55	0.54
CC: MLPregressor	0.44	0.41	0.45	0.44	0.43	0.45	0.43	0.41	0.44	0.41	0.43
NRMSE: Raw Shepard	0.06	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.05	0.06	0.06
NRMSE: SVR	0.06	0.06	0.05	0.06	0.06	0.06	0.06	0.05	0.05	0.06	0.06
NRMSE: MLPregressor	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
NSE: Raw Shepard	-0.13	-0.32	-0.12	-0.17	-0.25	-0.16	-0.18	-0.25	-0.20	-0.25	-0.20
NSE: SVR	0.32	0.24	0.29	0.30	0.27	0.25	0.24	0.27	0.33	0.28	0.28
NSE: MLPregressor	0.19	0.16	0.19	0.19	0.18	0.20	0.19	0.16	0.19	0.17	0.18

مدل‌های یادگیری با چهار ویژگی ورودی

در این مدل از مقدار تخمین ماهواره‌ای روز / ماه جاری (CP)، ارتفاع ایستگاه هواشناسی (SA)، طول جغرافیایی (LO) و عرض جغرافیایی (LA) استفاده شده است. این مدل بر مکان^۲ تمرکز دارد. ویژگی‌های مکانی (ارتفاع، طول و عرض جغرافیایی) به مدل اضافه شده‌اند. طبق جدول ۸، بهبود نسبی نسبت به مدل پایه مشاهده می‌شود اما این بهبود به خوبی حالت استفاده از توالی زمانی (مدل ۲ و ۳ ویژگی) نیست. این نشان می‌دهد برای پیش‌بینی در یک روز خاص، مکان، عامل مهمی است اما داده‌های تاریخچه زمانی مهم‌تر است. در تحلیل بارش ماهانه (جدول ۹)، یک نقطه عطف وجود دارد به نحوی که هر دو مدل SVR و MLPregressor به طور قابل توجهی بهبود یافته‌اند. RMSE MLPregressor به شگفت‌انگیز mm ۲۰/۸۶ و CC برابر ۰/۸۳ دست یافته است. این نشان می‌دهد برای میانگین‌گیری در مقیاس ماهانه، تأثیر توپوگرافی و موقعیت مکانی به شدت تعیین‌کننده است و مدل به خوبی این رابطه غیرخطی پیچیده را یاد گرفته است. (عملکرد بد MLPregressor در Fold9 احتمالاً به دلیل یک داده پرت خاص است).

این نشان می‌دهد که توالی زمانی^۱ یک عامل بسیار مهم در تخمین بارش روزانه است. در داده‌های ماهانه (جدول ۵)، بهبود نسبت به حالت یک ویژگی محسوس اما کمتر از داده روزانه است. MLPregressor عملکرد بهتری دارد (RMSE=23.54 mm, NSE=0.59). این نشان می‌دهد برای داده‌های ماهانه، اطلاعات ماه قبل نیز مفید است اما تأثیر آن به اندازه داده روزانه تعیین‌کننده نیست.

مدل‌های یادگیری با سه ویژگی ورودی

در این مدل از مقدار تخمین ماهواره‌ای روز / ماه جاری (CP)، مقدار تخمین ماهواره‌ای روز / ماه قبل (PE) و مقدار تخمین ماهواره‌ای روز / ماه بعد (NE) استفاده شده است. این مدل به اطلاعات آینده (ماه / روز بعد) دسترسی دارد. این سناریو بیشتر برای اصلاح داده‌های گذشته کاربرد دارد. در این مدل‌سازی هم، برای داده‌های روزانه (جدول ۶)، MLPregressor با RMSE متوسط mm ۲/۹۴ و CC متوسط ۰/۶۵، برتری قاطع خود را نشان می‌دهد.

MLPregressor به خوبی توانسته از اطلاعات آینده برای درک بهتر الگوهای جوی و اصلاح دقیق‌تر بارش استفاده کند.

² Geography

¹ Temporal Sequence

نتیجه را ارائه داده و به اوج عملکرد خود می‌رسد. در داده‌های ماهانه (جدول ۱۱) هم، MLPregressor بهترین Fold6 در RMSE، CC، NSE و NRMSE را دارد. اگرچه در Fold6 عملکرد ضعیفی دارد (NSE=-0.15)، اما به‌طور کلی برتری خود را نشان می‌دهد SVR. نیز عملکرد بسیار خوب و باثباتی دارد. ترکیب همه ویژگی‌ها نتیجه بهینه را تولید می‌کند. این موضوع فرضیه اصلی مقاله مبنی بر اینکه استفاده از ترکیب ویژگی‌های مختلف دقت را افزایش می‌دهد، به‌طور کامل تأیید می‌کند. MLPregressor در استفاده از این ترکیب ویژگی‌ها از SVR پیشی می‌گیرد.

مدل‌های یادگیری با شش ویژگی ورودی

در این مدل از مقدار تخمین ماهواره‌ای روز / ماه جاری (CP)، ارتفاع ایستگاه هواشناسی (SA)، طول جغرافیایی (LO)، عرض جغرافیایی (LA)، مقدار تخمین ماهواره‌ای روز / ماه قبل (PE) و مقدار تخمین ماهواره‌ای روز / ماه بعد (NE) استفاده شده است. این حالت کامل‌ترین مدل است که هم اطلاعات توالی زمانی (حال، گذشته، آینده) و هم اطلاعات مکانی (مکان و ارتفاع) را یکجا دارد. در داده‌های روزانه (جدول ۱۰)، MLP با RMSE متوسط ۲/۷۳ mm و CC متوسط ۰/۷۱، بهترین

جدول ۵. نتایج تحلیل داده‌های ماهانه بارش با ۲ پارامتر ورودی (CP, PE)

Table 5. Results of monthly data analysis with 2 input parameters (CP, PE)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw											
Shepard (mm)	28.22	26.24	27.32	26.06	33.19	31.31	28.45	28.52	33.73	23.99	28.70
RMSE: SVR (mm)	21.04	30.42	23.27	20.00	26.06	27.06	25.43	33.85	30.02	19.92	25.71
RMSE:											
MLPregressor (mm)	22.29	23.36	23.05	19.09	24.88	26.10	20.95	28.20	28.24	19.25	23.54
CC: Raw Shepard	0.73	0.84	0.77	0.79	0.74	0.74	0.76	0.83	0.76	0.82	0.78
CC: SVR	0.76	0.67	0.76	0.78	0.71	0.73	0.66	0.70	0.73	0.80	0.73
CC: MLPregressor	0.73	0.83	0.77	0.79	0.74	0.74	0.76	0.82	0.76	0.82	0.78
NRMSE: Raw											
Shepard	0.15	0.14	0.15	0.14	0.18	0.17	0.15	0.15	0.18	0.13	0.15
NRMSE: SVR	0.10	0.15	0.11	0.10	0.13	0.13	0.12	0.16	0.15	0.10	0.12
NRMSE:											
MLPregressor	0.11	0.11	0.11	0.09	0.12	0.13	0.10	0.14	0.14	0.09	0.11
NSE: Raw Shepard	0.42	0.61	0.45	0.53	0.42	0.45	0.49	0.60	0.47	0.56	0.50
NSE: SVR	0.57	0.44	0.58	0.58	0.48	0.51	0.35	0.47	0.53	0.64	0.52
NSE:											
MLPregressor	0.52	0.67	0.59	0.62	0.53	0.55	0.56	0.64	0.58	0.67	0.59

جدول ۶. نتایج تحلیل داده‌های روزانه بارش با ۳ پارامتر ورودی (CP, PE, NE)

Table 6. Results of daily data analysis with 3 input parameters (CP, PE, NE)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	4.03	3.92	3.95	4.18	4.18	3.91	4.12	3.98	3.91	4.20	4.04
RMSE: SVR (mm)	3.30	3.44	3.20	3.35	3.56	3.30	3.33	3.13	3.17	3.31	3.31
RMSE: MLPregressor (mm)	3.02	3.06	2.86	2.98	3.00	2.88	3.04	2.81	2.85	2.95	2.94
CC: Raw Shepard	0.47	0.46	0.43	0.40	0.37	0.46	0.44	0.38	0.45	0.37	0.42
CC: SVR	0.57	0.50	0.53	0.57	0.49	0.53	0.55	0.54	0.57	0.53	0.54
CC: MLPregressor	0.65	0.62	0.64	0.68	0.64	0.66	0.64	0.64	0.66	0.64	0.65
NRMSE: Raw Shepard	0.06	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.05	0.06	0.06
NRMSE: SVR	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.06	0.06	0.06
NRMSE: MLPregressor	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
NSE: Raw Shepard	-0.14	-0.14	-0.17	-0.40	-0.42	-0.12	-0.19	-0.30	-0.16	-0.34	-0.24
NSE: SVR	0.31	0.21	0.26	0.31	0.16	0.26	0.29	0.25	0.30	0.25	0.26
NSE: MLPregressor	0.42	0.37	0.41	0.45	0.41	0.43	0.41	0.40	0.44	0.40	0.41

جدول ۷. نتایج تحلیل داده‌های ماهانه بارش با ۳ پارامتر ورودی (CP, PE, NE)

Table 7. Results of monthly data analysis with 3 input parameters (CP, PE, NE)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	28.22	26.24	27.32	26.06	33.19	31.31	28.45	28.52	33.73	23.99	28.70
RMSE: SVR (mm)	21.66	32.07	21.72	19.87	27.95	26.16	24.61	34.76	32.57	19.52	26.09
RMSE: MLPregressor (mm)	22.30	23.43	23.12	19.06	25.18	26.25	36.09	28.58	28.26	19.29	25.16
CC: Raw Shepard	0.73	0.84	0.77	0.79	0.74	0.74	0.76	0.83	0.76	0.82	0.78
CC: SVR	0.75	0.63	0.80	0.79	0.69	0.75	0.70	0.68	0.68	0.81	0.73
CC: MLPregressor	0.73	0.82	0.77	0.79	0.74	0.74	-0.94	0.81	0.76	0.82	0.60
NRMSE: Raw Shepard	0.15	0.14	0.15	0.14	0.18	0.17	0.15	0.15	0.18	0.13	0.15
NRMSE: SVR	0.10	0.16	0.11	0.10	0.14	0.13	0.12	0.17	0.16	0.09	0.13
NRMSE: MLPregressor	0.11	0.11	0.11	0.09	0.12	0.13	0.17	0.14	0.14	0.09	0.12
NSE: Raw Shepard	0.42	0.61	0.45	0.53	0.42	0.45	0.49	0.60	0.47	0.56	0.50
NSE: SVR	0.55	0.38	0.64	0.58	0.40	0.54	0.39	0.45	0.44	0.66	0.50
NSE: MLPregressor	0.52	0.67	0.59	0.62	0.52	0.54	-0.31	0.63	0.58	0.66	0.50

جدول ۸. نتایج تحلیل داده‌های روزانه بارش با چهار پارامتر ورودی (CP, SA, LO, LA)

Table 8. Results of daily data analysis with four input parameters (CP, SA, LO, LA)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	4.03	3.92	3.95	4.18	4.18	3.91	4.12	3.98	3.91	4.20	4.04
RMSE: SVR (mm)	3.88	3.77	3.66	3.96	3.84	3.74	3.86	3.57	3.72	3.77	3.78
RMSE: MLPregressor (mm)	3.46	3.35	3.32	3.58	3.52	3.26	3.43	3.29	3.30	3.44	3.39
CC: Raw Shepard	0.47	0.46	0.43	0.40	0.37	0.46	0.44	0.38	0.45	0.37	0.42
CC: SVR	0.49	0.48	0.44	0.42	0.40	0.50	0.47	0.40	0.47	0.40	0.45
CC: MLPregressor	0.49	0.50	0.46	0.46	0.43	0.53	0.50	0.43	0.50	0.44	0.47
NRMSE: Raw Shepard	0.06	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.05	0.06	0.06
NRMSE: SVR	0.07	0.07	0.06	0.07	0.07	0.07	0.07	0.06	0.07	0.07	0.07
NRMSE: MLPregressor	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
NSE: Raw Shepard	-0.14	-0.14	-0.17	-0.40	-0.42	-0.12	-0.19	-0.30	-0.16	-0.34	-0.24
NSE: SVR	0.05	0.04	0.03	0.03	0.03	0.04	0.04	0.03	0.04	0.02	0.04
NSE: MLPregressor	0.24	0.25	0.21	0.21	0.18	0.27	0.25	0.18	0.25	0.19	0.22

جدول ۹. نتایج تحلیل داده‌های ماهانه بارش با چهار پارامتر ورودی (CP, SA, LO, LA)

Table 9. Results of monthly data analysis with four input parameters (CP, SA, LO, LA)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	28.22	26.24	27.32	26.06	33.19	31.31	28.45	28.52	33.73	23.99	28.70
RMSE: SVR (mm)	20.90	21.80	21.53	17.67	22.22	24.16	19.50	26.44	25.94	18.63	21.88
RMSE: MLPregressor (mm)	17.23	18.38	17.32	15.66	17.67	19.38	16.73	21.45	47.84	16.89	20.86
CC: Raw Shepard	0.73	0.84	0.77	0.79	0.74	0.74	0.76	0.83	0.76	0.82	0.78
CC: SVR	0.76	0.85	0.81	0.82	0.79	0.79	0.79	0.85	0.81	0.83	0.81
CC: MLPregressor	0.85	0.89	0.88	0.87	0.87	0.87	0.86	0.91	0.41	0.86	0.83
NRMSE: Raw Shepard	0.15	0.14	0.15	0.14	0.18	0.17	0.15	0.15	0.18	0.13	0.15
NRMSE: SVR	0.10	0.11	0.10	0.09	0.11	0.12	0.09	0.13	0.13	0.09	0.11
NRMSE: MLPregressor	0.08	0.09	0.08	0.08	0.09	0.09	0.08	0.10	0.23	0.08	0.10
NSE: Raw Shepard	0.42	0.61	0.45	0.53	0.42	0.45	0.49	0.60	0.47	0.56	0.50
NSE: SVR	0.58	0.71	0.64	0.67	0.62	0.61	0.62	0.68	0.65	0.69	0.65
NSE: MLPregressor	0.71	0.80	0.77	0.74	0.76	0.75	0.72	0.79	-0.20	0.74	0.66

جدول ۱۰. نتایج تحلیل داده‌های روزانه بارش با ۶ پارامتر ورودی (CP, SA, LO, LA, PE, NE)

Table 10. Results of daily data analysis with 6 input parameters (CP, SA, LO, LA, PE, NE)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	4.03	3.92	3.95	4.18	4.18	3.91	4.12	3.98	3.91	4.20	4.04
RMSE: SVR (mm)	3.35	3.28	3.18	3.37	3.26	3.19	3.35	3.02	3.20	3.19	3.24
RMSE: MLPregressor (mm)	2.86	2.81	2.75	2.80	2.71	2.62	2.75	2.65	2.64	2.72	2.73
CC: Raw Shepard	0.47	0.46	0.43	0.40	0.37	0.46	0.44	0.38	0.45	0.37	0.42
CC: SVR	0.64	0.61	0.61	0.68	0.65	0.67	0.64	0.65	0.65	0.65	0.64
CC: MLPregressor	0.70	0.68	0.68	0.72	0.72	0.73	0.72	0.68	0.72	0.70	0.71
NRMSE: Raw Shepard	0.06	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.05	0.06	0.06
NRMSE: SVR	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.06	0.06	0.06
NRMSE: MLPregressor	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
NSE: Raw Shepard	-0.14	-0.14	-0.17	-0.40	-0.42	-0.12	-0.19	-0.30	-0.16	-0.34	-0.24
NSE: SVR	0.29	0.28	0.27	0.30	0.30	0.30	0.28	0.30	0.29	0.30	0.29
NSE: MLPregressor	0.48	0.47	0.46	0.52	0.51	0.53	0.52	0.46	0.52	0.49	0.50

جدول ۱۱. نتایج تحلیل داده‌های ماهانه بارش با ۶ پارامتر ورودی (CP, SA, LO, LA, PE, NE)

Table 11. Results of monthly data analysis with 6 input parameters (CP, SA, LO, LA, PE, NE)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	average
RMSE: Raw Shepard (mm)	28.22	26.24	27.32	26.06	33.19	31.31	28.45	28.52	33.73	23.99	28.70
RMSE: SVR (mm)	21.02	21.86	21.61	17.74	22.37	24.28	19.56	26.42	26.01	18.62	21.95
RMSE: MLPregressor (mm)	16.97	18.60	17.17	15.65	17.73	41.54	16.23	21.78	18.71	16.65	20.10
CC: Raw Shepard	0.73	0.84	0.77	0.79	0.74	0.74	0.76	0.83	0.76	0.82	0.78
CC: SVR	0.76	0.85	0.81	0.82	0.79	0.78	0.79	0.85	0.81	0.83	0.81
CC: MLPregressor	0.85	0.89	0.88	0.87	0.87	0.48	0.87	0.90	0.91	0.87	0.84
NRMSE: Raw Shepard	0.15	0.14	0.15	0.14	0.18	0.17	0.15	0.15	0.18	0.13	0.15
NRMSE: SVR	0.10	0.11	0.10	0.09	0.11	0.12	0.09	0.13	0.13	0.09	0.11
NRMSE: MLPregressor	0.08	0.09	0.08	0.08	0.09	0.20	0.08	0.11	0.09	0.08	0.10
NSE: Raw Shepard	0.42	0.61	0.45	0.53	0.42	0.45	0.49	0.60	0.47	0.56	0.50
NSE: SVR	0.58	0.71	0.64	0.67	0.62	0.61	0.61	0.68	0.64	0.69	0.64
NSE: MLPregressor	0.72	0.79	0.77	0.74	0.76	-0.15	0.73	0.78	0.82	0.75	0.67

تحلیل نتایج داده‌های کلی

و در داده‌های ماهانه بارش کمتر می‌باشد. همچنین، در داده‌های روزانه بارش اختلاف بهبود نسبت به روش SVR هم قابل توجه هست. در مدل‌های مختلف اجرا، کمترین RMSE داده‌های روزانه و ماهانه مربوط به MLPregressor با ۶ ورودی با مقادیر ۲/۷۳ mm و ۲/۱۰ mm می‌باشد. بیشترین بهبود ضریب هم‌بستگی هم در این مدل اجرا اتفاق افتاده است. Raw Shepard همچنان ضعیف‌ترین مدل به لحاظ معیارهای ارزیابی شده می‌باشد. شکل ۲ و شکل ۳، نتایج کلی را به صورت بصری نشان می‌دهد.

در جدول ۱۲ و ۱۳ نتایج تجمیع شده ۵ مدل متفاوت اجرا از جدول‌های قبلی برای تحلیل بیشتر آورده شده است. همان طور که در جدول تجمیعی روزانه مشاهده می‌شود، مقدار NSE روش MLPregressor نسبت به روش Raw Shepard در تمامی مدل‌های ورودی افزایش قابل توجهی را نشان می‌دهد. به عنوان مثال در ستون ۶ ورودی مقدار CC: Raw Shepard از ۰/۲۴- به مثبت ۵۰ تغییر یافته است که بهبود ۰/۷۴ را نشان می‌دهد. همین نتیجه در مورد داده‌های ماهانه بارش هم صدق می‌کند ولی با این تفاوت که نسبت بهبود در داده‌های روزانه بارش بیشتر

جدول ۱۲. نتایج تحلیل داده‌های روزانه بارش کلی

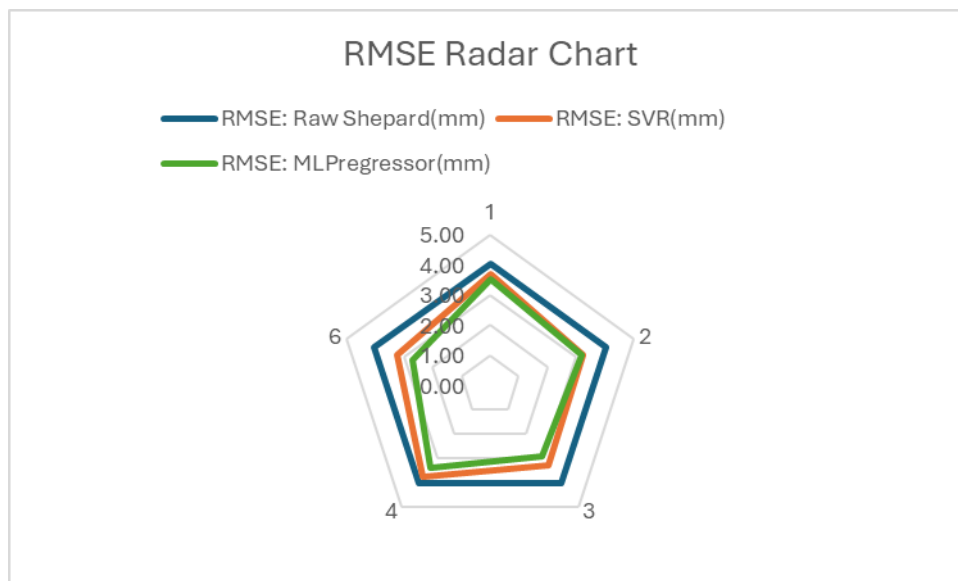
Table 12. Results of overall daily data analysis

Model	1	2	3	4	6
RMSE: Raw Shepard (mm)	4.04	4.04	4.04	4.04	4.04
RMSE: SVR (mm)	3.71	3.27	3.31	3.78	3.24
RMSE: MLPregressor (mm)	3.52	3.22	2.94	3.39	2.73
CC: Raw Shepard	0.42	0.39	0.42	0.42	0.42
CC: SVR	0.35	0.54	0.54	0.45	0.64
CC: MLPregressor	0.36	0.43	0.65	0.47	0.71
NRMSE: Raw Shepard	0.06	0.06	0.06	0.06	0.06
NRMSE: SVR	0.07	0.06	0.06	0.07	0.06
NRMSE: MLPregressor	0.06	0.06	0.05	0.06	0.05
NSE: Raw Shepard	-0.24	-0.20	-0.24	-0.24	-0.24
NSE: SVR	0.07	0.28	0.26	0.04	0.29
NSE: MLPregressor	0.16	0.18	0.41	0.22	0.50

جدول ۱۳. نتایج تحلیل داده‌های ماهانه بارش کلی

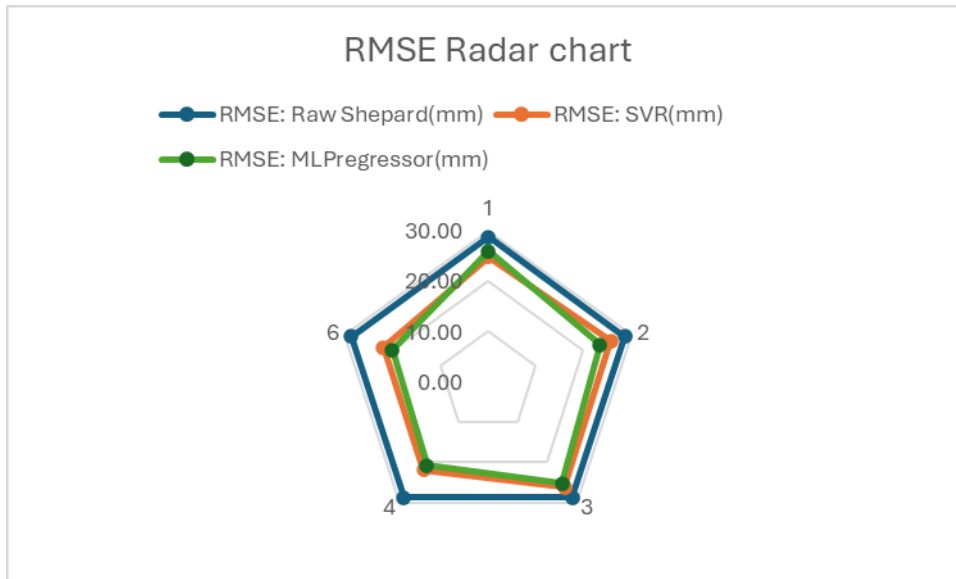
Table 13. Results of overall monthly data analysis

Model	1	2	3	4	6
RMSE: Raw Shepard (mm)	28.70	28.70	28.70	28.70	28.70
RMSE: SVR (mm)	24.79	25.71	26.09	21.88	21.95
RMSE: MLPregressor (mm)	25.86	23.54	25.16	20.86	20.10
CC: Raw Shepard	0.78	0.78	0.78	0.78	0.78
CC: SVR	0.75	0.73	0.73	0.81	0.81
CC: MLPregressor	0.69	0.78	0.60	0.83	0.84
NRMSE: Raw Shepard	0.15	0.15	0.15	0.15	0.15
NRMSE: SVR	0.12	0.12	0.13	0.11	0.11
NRMSE: MLPregressor	0.13	0.11	0.12	0.10	0.10
NSE: Raw Shepard	0.50	0.50	0.50	0.50	0.50
NSE: SVR	0.55	0.52	0.50	0.65	0.64
NSE: MLPregressor	0.51	0.59	0.50	0.66	0.67



شکل ۲. نمودار راداری معیار RMSE برای داده‌های بارش روزانه

Fig 2. Radar plot of RMSE criterion for daily precipitation data



شکل ۳. نمودار راداری معیار RMSE برای داده‌های بارش ماهانه

Fig 3. Radar plot of RMSE criterion for monthly precipitation data

نتیجه‌گیری

یادگیری ماشین با ارائه روش‌هایی مانند رگرسیون شبکه عصبی و رگرسیون ماشین بردار پشتیبان، افق‌های جدیدی را در تخمین بارش و مطالعات هواشناسی گشوده است. این روش‌ها با توانایی مدل‌سازی روابط پیچیده و غیرخطی، دقت پیش‌بینی‌های بارش را بهبود می‌بخشند و به تصمیم‌گیرندگان در زمینه‌هایی مانند مدیریت منابع آبی و کشاورزی کمک می‌کنند تا برنامه‌ریزی مؤثرتری داشته باشند. نتایج به وضوح نشان داد که کاربرد هر دو روش SVR و MLPregressor منجر به بهبود چشمگیر دقت تخمین بارش در مقایسه با روش درونیایی ساده شپرد (Raw Shepard) می‌شود. کاهش معنادار معیار RMSE و افزایش معیارهای CC و NSE در کلیه مدل‌ها و برای هر دو مقیاس روزانه و ماهانه، گواه بر توانایی ذاتی الگوریتم‌های یادگیری ماشین در تصحیح خطاها و کشف روابط غیرخطی پیچیده بین داده‌های ماهواره‌ای و مشاهدات زمینی است. نتایج حاصل از به‌کارگیری روش‌های پیشنهادی بر روی داده‌های تخمینی ماهواره‌ای ERA5 و داده‌های واقعی بارش ایستگاه‌های هواشناسی

استان آذربایجان غربی نشان دهنده بهبود قابل توجهی در محاسبات مقدار دقیق بارش ایستگاه‌های هواشناسی بود. با توجه به اینکه مجموعه ویژگی‌های متفاوتی برای ورودی الگوریتم‌های یادگیری استفاده شد، به‌طور کلی برای SVR و MLPregressor، ویژگی‌های مقدار تخمین ماهواره‌ای روز جاری (CP)، مقدار تخمین ماهواره‌ای روز قبل (PE) و مقدار تخمین ماهواره‌ای روز بعد (NE) در داده‌های روزانه بارش خیلی تعیین‌کننده هستند. مقاومت مدل‌ها در داده‌های ماهانه بارش نسبت به تغییرات ویژگی‌های ورودی بیشتر هست. مهم‌ترین یافته این پژوهش، وابستگی شدید دقت مدل به تعداد و به‌ویژه نوع ویژگی‌های ورودی بود: اگرچه هر دو الگوریتم SVR و MLPregressor بهبود قابل توجهی نسبت به روش پایه ایجاد کردند، اما عملکرد آنها متفاوت بود. الگوریتم SVR در شرایطی که داده‌ها از پیچیدگی کمتری برخوردار بودند (مانند مدل‌های با تعداد ویژگی کم) یا در داده‌های ماهانه، پایداری و عملکرد قابل اطمینان‌تری از خود نشان داد. ولی MLPregressor در شرایط بهینه (استفاده از تمامی ۶ ویژگی)، برتری خود را نشان داد. توانایی بالاتر این الگوریتم در یادگیری روابط غیرخطی بسیار پیچیده، هنگامی که حجم و تنوع

(1981/1982–2022/2023). *Water*, 16(5), 628.
<https://doi.org/10.3390/w16050628>

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J.,... & Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly journal of the royal meteorological society*, 146(730), 1999-2049.
<https://doi.org/10.1002/qj.3803>

Houénafa, S. E., Johnson, O., Ronoh, E. K., & Moore, S. E. (2025). Hybridization of Stochastic Hydrological Models and Machine Learning Methods for Improving Rainfall-Runoff Modelling. *Results in Engineering*, 104079.
<https://doi.org/10.1016/j.rineng.2025.104079>

Huang, S., Wang, S., Chen, J., Wang, C., Zhang, X., Wu, J., Chen, N. (2024). Urbanization-induced spatial and temporal patterns of local drought revealed by high-resolution fused remotely sensed datasets. *Remote Sensing of Environment*, 313, 114378. <https://doi.org/10.1016/j.rse.2024.114378>

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Xie, P. (2014). NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theoretical Basis Document (ATBD) Version 4*. NASA technical report.

Iacono, M.J.; Delamere, J.S.; Mlawer, E.J.; Shephard, M.W.; Clough, S.A.; Collins, W.D (2008).. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.* 2008, 113, D13103.
<https://doi.org/10.1029/2008JD009944>

Jiang, S. H., Wei, L. Y., Ren, L. L., Zhang, L. Q., Wang, M. H., & Cui, H. (2023). Evaluation of IMERG, TMPA, ERA5, and CPC precipitation products over mainland China: Spatiotemporal patterns and extremes. *Water Science and Engineering*, 16(1), 45-56.
<https://doi.org/10.1016/j.wse.2022.05.001>

Kalhari M, Tadayon M, Kahrizi E, Ghiasvand M. Analysis and monitoring of water resources and drought using a combination of GRACE, MODIS, and Landsat 8 satellite images (Case study: Hamedan City). *Water Resources and Climate Change*. (2025); 1(1), 62-74.
<https://doi.org/10.22091/wrcc.2025.11390.1007>

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, how much of the Earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, 98(1), 69-78.
<https://doi.org/10.1175/BAMS-D-14-00283.1>

داده‌های ورودی کافی باشد، منجر به دستیابی به کمترین RMSE (mm 10/20 برای روزانه و mm 73/2 برای ماهانه) و بالاترین CC (0/71 برای روزانه و 0/84 برای ماهانه) شد. برای مطالعات آینده، پیشنهاد می‌شود که از داده‌های ماهواره‌ای با رزولوشن بالاتر و روش‌های پیشرفته‌تر یادگیری ماشین و به‌کارگیری الگوریتم‌های پیچیده‌تر یادگیری عمیق مانند LSTM یا CNN برای مدل‌سازی بهتر وابستگی‌های مکانی-زمانی استفاده شود. همچنین می‌توان تأثیر افزودن سایر متغیرهای جوی (مانند رطوبت، فشار سطح دریا) به‌عنوان ویژگی‌های کمکی را بررسی کرد.

منابع

Amjad, M., Yilmaz, M. T., Yucel, I., & Yilmaz, K. K. (2020). Performance evaluation of satellite-and model-based precipitation products over varying climate and complex topography. *Journal of Hydrology*, 584, 124707.
<https://doi.org/10.1016/j.jhydrol.2020.124707>

Bagheri Khanghahi M., Hazar Jaribi A., Kamali Mohammad I., Zamani F. Forecasting Rainfall in Different Climatic Regions of Iran Using the LARS WG7 Climate Model. *Water Resources and Climate Change*. (2025); 1(1), 28-39.
<https://doi.org/10.22091/wrcc.2025.11744.1008>

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I., Weedon, G. P.,... & Wood, E. F. (2020). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Satellite precipitation measurement: Volume 2*, 625-653. https://doi.org/10.1007/978-3-030-35798-6_9

Chen, F., Gao, Y. (2018). Evaluation of precipitation trends from high-resolution satellite precipitation products over Mainland China. *Climate Dynamics*, 51, 3311-3331.
<https://doi.org/10.1007/s00382-018-4080-z>

Donat, M. G., Lowry, A. L., Alexander, L. V., O’Gorman, P. A., & Maher, N. (2016). More extreme precipitation in the world’s dry and wet regions. *Nature Climate Change*, 6(5), 508-513.
<https://doi.org/10.1038/nclimate2941>

Espinosa, L. A., Portela, M. M., & Gharbia, S. (2024). Assessing changes in exceptional rainfall in Portugal using ERA5-land reanalysis data

- Komasi M, Dalvand R. Evaluation of nonparametric decision tree models for predicting scour depth of bridges. *Water Resources and Climate Change*. (2025); 1(1), 40-50. <https://doi.org/10.22091/wrcc.2025.11363.1005>
- Kumar, A., Ramsankaran, R. A. A. J., Brocca, L., & Munoz-Arriola, F. (2019). A machine learning approach for improving near-real-time satellite-based rainfall estimates by integrating soil moisture. *Remote Sensing*, 11(19), 2221. <https://doi.org/10.3390/rs11192221>
- Mianabadi, A., Omidvar, J., & Pourreza-Bilandi, M. (2024). Development of intensity–duration–frequency curves at the basin scale using the ERA5 reanalysis product. *Journal of Drought and Climate Change Research*, 2(4), 121–140 [in Persian]. <https://doi.org/10.22077/jdcr.2025.8636.1098>
- Modaresi, F., Araghinejad, S., & Ebrahimi, K. (2018). A comparative assessment of artificial neural network, generalized regression neural network, least-square support vector regression, and K-nearest neighbor regression for monthly streamflow forecasting in linear and nonlinear conditions. *Water resources management*, 32, 243-258. <https://doi.org/10.1007/s11269-017-1807-2>
- Nouhani, E., Babaali, H. R., & Dehghani, R. (2024). Estimation of suspended sediments in the coastal areas of the Caspian Sea using machine learning techniques. *Journal of Drought and Climate Change Research*. Advance online publication. [in Persian]. <https://doi.org/10.22077/jdcr.2025.8983.1121>
- Saha, A., & Pal, S. C. (2024). Application of machine learning and emerging remote sensing techniques in hydrology: A state-of-the-art review and current research trends. *Journal of Hydrology*, 632, 130907. <https://doi.org/10.1016/j.jhydrol.2024.130907>
- Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., Thépaut, J. N. (2024). The ERA5 global reanalysis from 1940 to 2022. *Quarterly Journal of the Royal Meteorological Society*, 150(764), 4014-4048. <https://doi.org/10.1002/qj.4803>
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K. L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1), 79-107. <https://doi.org/10.1002/2017RG000574>
- Tang, W., Qin, J., Yang, K., Zhu, F., & Zhou, X. (2021). Does ERA5 outperform satellite products in estimating atmospheric downward longwave radiation at the surface?. *Atmospheric Research*, 252, 105453. <https://doi.org/10.1016/j.atmosres.2021.105453>
- Wang, Q., Xia, J., She, D., Zhang, X., Liu, J., & Zhang, Y. (2021). Assessment of the four latest long-term satellite-based precipitation products in capturing the extreme precipitation and streamflow across a humid region of southern China. *Atmospheric Research*, 257, 105554. <https://doi.org/10.1016/j.atmosres.2021.105554>
- Yang, L., Shi, Z., Liu, R., & Xing, M. (2024). Evaluating the performance of global precipitation products for precipitation and extreme precipitation in arid and semiarid China. *International Journal of Applied Earth Observation and Geoinformation*, 130, 103888. <https://doi.org/10.1016/j.jag.2024.103888>
- Yousefi-Kebria, A., & Nadi, M. (2023). Evaluation of the accuracy of GPM satellite precipitation estimates: A case study in Mazandaran Province. *Journal of Drought and Climate Change Research*, 1(3), 1–14. [in Persian]. <https://doi.org/10.22077/jdcr.2023.6232.1022>
- Yu, C., Hu, D., Liu, M., Wang, S., & Di, Y. (2020). Spatio-temporal accuracy evaluation of three high-resolution satellite precipitation products in the China area. *Atmospheric Research*, 241, 104952. <https://doi.org/10.1016/j.atmosres.2020.104952>
- Yuan, Y., & Liao, B. (2025). Evaluation of multi-source precipitation products for monitoring drought across China. *Frontiers in Environmental Science*, 13, 1524937. <https://doi.org/10.3389/fenvs.2025.1524937>
- Zhou, Z., Guo, B., Xing, W., Zhou, J., Xu, F., & Xu, Y. (2020). Comprehensive evaluation of the latest GPM era IMERG and GSMaP precipitation products over mainland China. *Atmospheric Research*, 246, 105132. <https://doi.org/10.1016/j.atmosres.2020.105132>

