

Document Type

Applied Article

پژوهش کاربردی

نوع مقاله

Enhancing Meteorological Drought Monitoring in Arid Regions: A Machine Learning-Based Fusion of Satellite Precipitation Data

M. Hajiani^{1*}, M. Mardani², M. Tosan^{3,4}

1- Associate Professor, Department of Civil Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran.

2- Ph.D. Candidate in Water Resources Engineering and Management, Department of Civil Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran.

3- Department of Irrigation and Reclamation Engineering, Faculty of Agriculture, University of Tehran, Karaj, Iran.

4- Center of Excellence in Hydroinformatics, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

* Corresponding, Author Email: hajiani@birjand.ac.ir

Received ▶ 28-11-2025
Revised ▶ 28-12-2025
Accepted ▶ 28-01-2026
Available online ▶ 16-03-2026

Abstract

The accurate monitoring of drought and the management of water resources in arid and semi-arid regions consistently face the challenge of insufficient and unevenly distributed ground-based stations, the potential of gridded precipitation products (GPPs) to fill data gaps is significant. However, systematic errors and inherent uncertainties in estimating extreme values limit their efficacy in hydrological modeling. The present research aims to enhance the accuracy of meteorological drought monitoring in South Khorasan Province by developing an integrated model based on a machine learning algorithm. To this end, four widely used products, including TRMM, CHIRPS, GSMaP, and ERA5, along with spatio-temporal auxiliary variables, were considered as inputs for a Random Forest model. This model was trained and validated using daily precipitation data from synoptic stations over a 10-year period (2010-2019). The evaluation of the model's performance on a daily scale demonstrated that the merged product achieved a median correlation coefficient of 0.87, a root mean square error of 0.78 mm/day, and a critical success index of 0.545, exhibiting a statistically significant superiority over all raw products. In addition to improving quantitative accuracy, the hydrological evaluation based on the Standardized Precipitation Index indicated a high agreement between the corrected product and the ground reality at both short-term and long-term scales (with $R^2 > 0.81$ and $RMSE < 0.40$). The findings of this research prove that the intelligent fusion approach is capable of providing reliable climate data for the development of early warning systems for drought in data-scarce regions by effectively reducing the uncertainties of precipitation products.

Keywords: Random Forest, Standardized Precipitation Index, Bias Correction, Gridded Precipitation Products.

How to cite this article:

Hajiani, M., Mardani, M., & Tosan, M. (2026). Enhancing Meteorological Drought Monitoring in Arid Regions: A Machine Learning-Based Fusion of Satellite Precipitation Data. *Journal of Water and Sustainable Development*, 12(4), 190-203. <https://doi.org/10.22067/jwsd.v12i4.2511-1474>

ارتقاء پایش خشکسالی هواشناسی در مناطق خشک: ادغام داده‌های بارش ماهواره‌ای مبتنی بر یادگیری ماشین

محمود حاجیانی^{۱*}، مهسا مردانی^۲، معین توسان^{۳،۴}

۱- دانشیار گروه مهندسی عمران، دانشکده مهندسی، دانشگاه بیرجند، بیرجند، ایران.

۲- دانشجوی دکتری مهندسی و مدیریت منابع آب، گروه مهندسی عمران، دانشکده مهندسی، دانشگاه بیرجند، بیرجند، ایران.

۳- گروه مهندسی آبیاری و آبادانی، دانشکده کشاورزی، دانشگاه تهران، کرج، ایران.

۴- قطب هیدروانفورماتیک دانشکده مهندسی عمران دانشگاه تبریز، تبریز، ایران.

* رایانامه نویسنده مسئول: hajiani@birjand.ac.ir

تاریخ دریافت ▶ ۱۴۰۴/۰۹/۰۷
تاریخ بازنگری ▶ ۱۴۰۴/۱۰/۰۷
تاریخ پذیرش ▶ ۱۴۰۴/۱۱/۰۸
تاریخ انتشار ▶ ۱۴۰۴/۱۲/۲۵

چکیده

پایش دقیق خشکسالی و مدیریت منابع آب در مناطق خشک و نیمه‌خشک، همواره با چالش کمبود و توزیع نامناسب ایستگاه‌های زمینی مواجه است. اگرچه محصولات بارش شبکه‌بندی شده پتانسیل بالایی برای پر کردن خلأهای داده‌ای دارند، اما وجود خطاهای سیستماتیک و عدم قطعیت‌های ذاتی در برآورد مقادیر حدی، کارایی آن‌ها را در مدل‌سازی هیدرولوژیکی محدود می‌سازد. پژوهش حاضر با هدف ارتقای دقت پایش خشکسالی هواشناسی در استان خراسان جنوبی، به توسعه یک مدل ادغام شده مبتنی بر الگوریتم یادگیری ماشین پرداخته است. به این منظور چهار محصول پرکاربرد شامل TRMM، CHIRPS، GSMaP و ERA5 به همراه متغیرهای کمکی مکانی-زمانی، به عنوان ورودی‌های مدل جنگل تصادفی در نظر گرفته شدند و با استفاده از داده‌های بارش روزانه ایستگاه‌های سینوپتیک طی دوره ۱۰ ساله (۲۰۱۰-۲۰۱۹) آموزش و اعتبارسنجی شدند. ارزیابی عملکرد مدل در مقیاس روزانه نشان داد محصول ادغام شده با دستیابی به ضریب همبستگی میانه ۰/۸۷، ریشه میانگین مربعات خطای ۰/۷۸ میلی‌متر در روز و شاخص موفقیت بحرانی ۰/۵۴۵، برتری آماری معناداری نسبت به تمامی محصولات خام دارد. علاوه بر بهبود دقت کمی، ارزیابی هیدرولوژیکی بر اساس شاخص استاندارد بارش بیانگر انطباق بالای محصول اصلاح شده با واقعیت زمینی در هر دو مقیاس کوتاه‌مدت و بلندمدت بود (با $R^2 > 0.81$ و $RMSE < 0.40$). یافته‌های این پژوهش اثبات می‌کند که رویکرد همجوشی هوشمند قادر است با کاهش مؤثر عدم قطعیت‌های محصولات بارشی، داده‌های اقلیمی قابل‌اتکایی را جهت توسعه سیستم‌های هشدار زودهنگام خشکسالی در مناطق فاقد داده فراهم آورد.

واژه‌های کلیدی: جنگل تصادفی، شاخص استاندارد بارش، اصلاح خطا، محصولات بارش شبکه‌بندی شده.

دارند، خطاهای به ظاهر کوچک در تخمین شدت بارش یا تشخیص نادرست روزهای بارانی (اثر تورم صفر)، می تواند منجر به تغییرات فاحش در پارامترهای توزیع آماری و در نتیجه جابه جایی طبقات خشکسالی (مثلاً نمایش ترسالی کاذب در دوره خشک) شود (توسن و همکاران، ۲۰۲۴؛ Negri و همکاران، ۲۰۲۵). افزون بر این، هر یک از محصولات ماهواره ای و بازتحلیل، بسته به نوع حسگرهای فعال/غیرفعال و الگوریتم های پردازشی، نقاط قوت و ضعف منحصربه فردی دارند؛ به عنوان مثال، ممکن است یک محصول در تشخیص بارش های همرفتی تابستانه قوی تر باشد و دیگری در برآورد بارش های جبهه ای زمستانه عملکرد بهتری ارائه دهد (Batista و همکاران، ۲۰۲۵؛ Shiru و همکاران، ۲۰۲۵). بنابراین، تکیه بر یک محصول واحد یا استفاده از روش های اصلاح خطای خطی ساده، نمی تواند پیچیدگی های رفتار مکانی-زمانی بارش را در حوضه های ناهمگون پوشش دهد و رویکرد همجوشی چندمنبعی^۳ به عنوان راهکاری کارآمد برای بهره گیری هم زمان از نقاط قوت محصولات مختلف و کاهش عدم قطعیت ساختاری ضروری به نظر می رسد.

در سال های اخیر، استفاده از روش های یادگیری ماشین (ML)^۴ به عنوان جایگزینی توانمند برای روش های آماری سنتی در اصلاح خطای محصولات ماهواره ای مطرح شده است (Oyarzabal و همکاران، ۲۰۲۵). الگوریتم هایی مانند جنگل تصادفی (RF)^۵، ماشین بردار پشتیبان (SVM) و روش های بوسستینگ (XGBoost) به دلیل قابلیت بالا در مدل سازی روابط غیرخطی، قدرت تفکیک ویژگی ها و کاهش بیش برآزش، در بسیاری از مطالعات بارش سازی و همجوشی چندمنبعی عملکرد موفقی داشته اند (Kossieris و همکاران، ۲۰۲۴؛ Zhang و همکاران، ۲۰۲۴). این روش ها با استفاده هم زمان از چند محصول شبکه بندی شده و متغیرهای کمکی مکانی-زمانی، می توانند الگوهای خطای وابسته به توپوگرافی و فصلی را شناسایی کرده و محصول بارشی دقیق تر ارائه دهند. به عنوان نمونه نشان داده شده است که ترکیب داده های سنجش اذدور با ML موجب بهبود چشمگیر دقت پایش خشکسالی در مناطق نیمه خشک می شود (Zhang و همکاران، ۲۰۲۳). همچنین بر اهمیت بهره گیری از الگوریتم های داده کاوی برای پایش مخاطرات اقلیمی در سرزمین های خشک ایران تأکید شده است (Hashemi و همکاران، ۲۰۲۴).

با وجود این پیشرفت ها، خلأ علمی ارزیابی جامع تأثیر مدل های همجوشی ML بر پایش مستقیم خشکسالی (به ویژه SPI) در مناطق فراخشک و کوهستانی همچنان باقی است. اغلب پژوهش های پیشین صرفاً به سنجش دقت کمی بارش پرداخته اند و میزان اثرگذاری اصلاحات انجام شده بر شاخص های خشکسالی را بررسی نکرده اند. این موضوع از آن جهت اهمیت دارد که حتی خطاهای کوچک در داده های روزانه بارش می تواند موجب خطای انباشته در زمان محاسبه

خشکسالی یکی از پیچیده ترین و مخرب ترین مخاطرات اقلیمی در جهان است که آثار عمیقی بر سامانه های هیدرولوژیکی، تولیدات کشاورزی، امنیت غذایی و پایداری اقتصادی-اجتماعی جوامع انسانی بر جای می گذارد (Panu و Sharma، ۲۰۲۴). اهمیت این پدیده در مناطق خشک و نیمه خشک چند برابر است؛ زیرا کوچک ترین تغییر در بارش می تواند پیامدهای قابل توجهی بر بیلان آب، تغذیه آب زیرزمینی و وابستگی شدید به منابع آب محدود ایجاد نماید (Lalmuanzuala و همکاران، ۲۰۲۵). مدیریت پایدار این مناطق نیازمند دسترسی به داده های بارش دقیق، پیوسته، درازمدت و با وضوح مکانی بالا است؛ اما در بسیاری از کشورهای درحال توسعه، شبکه های پایش زمینی هم از نظر تعداد ایستگاه و هم از نظر پراکنش مکانی، با کمبودهای جدی مواجه می باشند (Bhaga و همکاران، ۲۰۲۰). این مشکل در ایران و به ویژه در نواحی شرقی مانند استان خراسان جنوبی، به صورت بارزی مشاهده می شود، جایی که شرایط فراخشک، ارتفاعات ناهمگون و گستردگی جغرافیایی باعث شده داده های زمینی نتوانند نماینده واقعی تغییرات مکانی بارش باشند (Nikoo و همکاران، ۲۰۲۶).

در چنین شرایطی، محصولات بارش شبکه بندی شده (GPPs)^۶ حاصل از سنجش اذدور، مانند CHIRPS، TRMM، GSMaP و محصولات بازتحلیل جوی مانند ERA5، به عنوان منابع مکمل ارزشمند جهت پر کردن خلأ داده های مشاهداتی مطرح شده اند (Mahmud و Osawa، ۲۰۲۵). این داده ها با پوشش مکانی یکنواخت و دسترسی طولانی مدت، ابزار مهمی برای ارزیابی خشکسالی، مدل سازی هیدرولوژیکی و توسعه سامانه های هشدار زود هنگام به شمار می روند (Nourani و همکاران، ۲۰۲۵). با این حال، مطالعات متعدد نشان داده اند که این محصولات با وجود مزایای فراوان، دارای خطاهای سیستماتیک ناشی از محدودیت های سنجنده ها، الگوریتم های بازیابی بارش، تأثیر توپوگرافی پیچیده و شرایط اقلیمی خاص مناطق خشک می باشند (Michibata، ۲۰۲۴؛ Tong و همکاران، ۲۰۲۴). وجود روزهای بدون بارش، شدت پایین بارش ها، رخداد های حدی و اثر بادهای منطقه ای، باعث تشدید عدم قطعیت این داده ها در مناطق بیابانی می شود (Rahat و همکاران، ۲۰۲۴). این ضعف ها به طور مستقیم بر دقت شاخص های خشکسالی مانند SPI^۲ تأثیر می گذارند و لزوم انجام فرایندهای اصلاح خطا و ادغام داده ها را آشکار می سازند (Tosan و همکاران، ۲۰۲۶).

از سوی دیگر، محاسبه نمایه های خشکسالی استاندارد شده مانند SPI که مبتنی بر برآزش تابع توزیع احتمالاتی بر سری های زمانی بارش هستند، حساسیت بسیار بالایی به کیفیت داده های ورودی دارد (Mwinjuma و همکاران، ۲۰۲۶). در مناطق فراخشک که بارش ها ماهیت رگباری و نامنظم

اقلیمی قابل اعتماد برای سیستم‌های پایش و هشدار خشکسالی فراهم کنند؛ موضوعی که برای مدیریت منابع آب در شرق ایران اهمیت حیاتی دارد.

مواد و روش‌ها

منطقه مورد مطالعه

استان خراسان جنوبی با مساحتی بالغ بر ۱۵۱,۰۰۰ کیلومتر مربع با مختصات جغرافیایی ۳۰ درجه و ۳۱ دقیقه تا ۳۴ درجه و ۵۳ دقیقه شمالی و ۵۷ درجه و ۳ دقیقه تا ۶۱ درجه شرقی در شرق فلات ایران واقع شده است (شکل ۱). این منطقه دارای توپوگرافی پیچیده‌ای شامل رشته‌کوه‌های شمالی-جنوبی (با حداکثر ارتفاع ۳۶۱۵ متر) و دشت‌های پست نظیر دشت لوت (با حداقل ارتفاع ۶۵۰ متر) است که منجر به تغییرات مکانی شدید در رژیم بارش می‌شود (رضوانی مقدم و همکاران، ۱۳۹۵). اقلیم منطقه عمدتاً از نوع خشک و بیابانی بوده و متأثر از بادهای ۱۲۰ روزه سیستان می‌باشد. کمبود و توزیع نامناسب ایستگاه‌های باران‌سنجی در این پهنه وسیع، پایش دقیق هیدرولوژیکی را با چالش مواجه ساخته است (Feizi و Tosan, ۲۰۱۷).

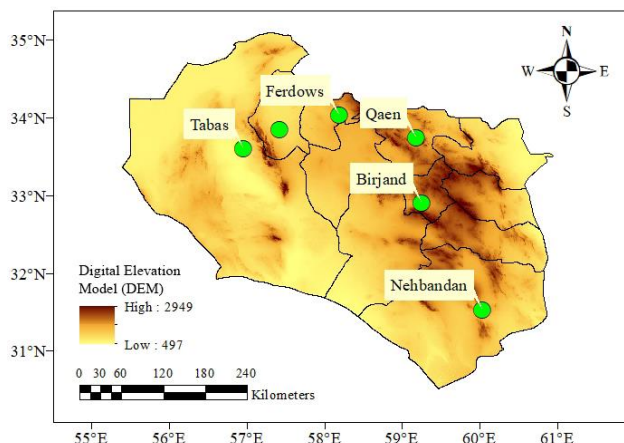
SPI شود و تصمیم‌گیری مدیران منابع آب را به‌شدت تحت تأثیر قرار دهد. استان خراسان جنوبی، با اقلیم بیابانی، تنوع ارتفاعی و فقر ایستگاه‌های زمینی، نمونه‌ای شاخص از مناطقی است که بررسی این خلأ در آن ضرورت ویژه دارد. بر این اساس، پژوهش حاضر با هدف پر کردن این شکاف علمی، یک چارچوب همجوشی مبتنی بر یادگیری ماشین را برای تولید یک محصول بارش اصلاح‌شده با قابلیت اعتماد بالا توسعه می‌دهد. اهداف اصلی این مطالعه عبارت‌اند از:

۱. توسعه مدل جنگل تصادفی برای ادغام چهار محصول کمکی مکانی-زمانی، جهت تولید سری زمانی بارش روزانه دقیق‌تر.

۲. ارزیابی جامع محصول ادغام شده با بهره‌گیری از شاخص‌های آماری پیوسته و طبقه‌بندی شده و مقایسه آن با محصولات خام.

۳. سنجش کارایی محصول اصلاح شده در پایش خشکسالی هواشناسی از طریق محاسبه شاخص SPI در مقیاس‌های کوتاه‌مدت و بلندمدت و مقایسه آن با مقادیر زمینی.

این پژوهش علاوه بر ارائه یک چارچوب کاربردی برای مناطق کم‌داده، نشان می‌دهد که رویکردهای ML-Fusion قادرند با کاهش عدم قطعیت محصولات شبکه‌بندی شده، داده‌های



شکل ۱- نقشه منطقه مورد مطالعه (استان خراسان جنوبی) و موقعیت ایستگاه‌های سینوپتیک زمینی

جدول ۱- مشخصات جغرافیایی ایستگاه‌های سینوپتیک مورد استفاده در پژوهش

شناسه ایستگاه	نام ایستگاه	طول جغرافیایی (°E)	عرض جغرافیایی (°N)	ارتفاع (متر)	میانگین بارش سالانه (mm/year)
۴۰۸۰۹	بیرجند	۵۹/۲۵	۳۲/۹	۱۴۹۱	۱۲
۹۹۴۰۷	بشروییه	۵۷/۴۲	۳۳/۸۵	۸۷۹	۷۳
۴۰۷۹۲	فردوس	۵۸/۱۸	۳۴/۰۳	۱۲۹۳	۱۱۸
۴۰۸۲۷	نهبندان	۶۰/۰۳	۳۱/۵۳	۱۱۸۸	۸۹
۴۰۷۹۳	قائن	۵۹/۱۸	۳۳/۷۴	۱۴۳۲	۱۵۰
۴۰۷۹۱	طبس	۵۹/۹۵	۳۳/۶	۷۱۱	۶۶

مجموعه داده‌ها

داده‌های بارش روزانه ۶ ایستگاه سینوپتیک (جدول ۱) طی بازه زمانی ۱۰ ساله (۲۰۱۹-۲۰۱۰) از سازمان هواشناسی کشور دریافت و پس از اعمال کنترل کیفی، به‌عنوان متغیر هدف جهت آموزش و اعتبارسنجی مدل استفاده شد. به‌منظور فراهم‌سازی داده‌های بارش قابل اعتماد برای تحلیل هیدرولوژیک، چهار مجموعه داده گسترده‌کاربرد شامل TRMM 3B42 نسخه ۷، CHIRPS نسخه ۲، GSMaP نسخه ۸ و ERA5 به‌عنوان ورودی مدل استفاده شدند. این محصولات نماینده طیفی از رویکردهای ماهواره‌ای و بازتحلیل بوده و از نظر سازوکار برآورد بارش، پوشش مکانی، وضوح زمانی و بازه زمانی موجود تفاوت‌های قابل توجهی دارند. مشخصات کلیدی هر محصول در جدول (۲) ارائه شده است.

جدول ۲- مشخصات فنی و آماری محصولات بارش شبکه‌بندی شده و متغیرهای ورودی مورد استفاده در پژوهش

محصول	نسخه	نوع داده	وضوح زمانی	وضوح مکانی	دوره آماری مورد استفاده	منبع / سنجنده‌ها	منبع
TRMM 3B42	نسخه ۷ (نهایی)	ماهواره‌ای (Merged Microwave-IR)	روزانه	۰/۲۵°	۲۰۱۹-۲۰۱۰	ماهواره + TRMM سنجنده‌های مایکروویو چندمنبعه	(Maggion و همکاران، ۲۰۱۶)
CHIRPS	نسخه ۲	شبه‌ماهواره‌ای (Satellite + Gauge Blended)	روزانه	۰/۰۵°	۲۰۱۹-۲۰۱۰	داده‌های ماهواره‌ای IR به‌همراه باران‌سنج‌ها	(Bouaziz و همکاران، ۲۰۲۱)
GSMaP	نسخه ۸	ماهواره‌ای مایکروویو با اصلاح باران‌سنج	روزانه	۰/۱°	۲۰۱۹-۲۰۱۰	چندین سنجنده مایکروویو + کالیبراسیون محلی	(Bui و همکاران، ۲۰۱۹)
ERA5	آخرین نسخه عملیاتی	بازتحلیل	روزانه	۰/۲۵°	۲۰۱۹-۲۰۱۰	مدل اقلیمی ECMWF (IFS) با ادغام مشاهدات زمین، ماهواره و جو	(Hersbach و همکاران، ۲۰۲۰)

روش‌شناسی

جهت همگام‌سازی مکانی-زمانی، استخراج داده‌های پیکسل-به-نقطه برای موقعیت ایستگاه‌ها انجام شد. داده‌های کمتر از روزانه به مقیاس روزانه گردآوری شدند. همچنین، به‌منظور حذف اثر نویزهای پس‌زمینه و تمرکز بر رخداد‌های بارش مؤثر، آستانه بارش ۱ میلی‌متر در روز اعمال گردید.

جنگل تصادفی یک الگوریتم یادگیری گروهی قدرتمند است که بر پایه گردآوری نتایج تعداد زیادی درخت تصمیم^۶ عمل می‌کند. در این پژوهش، از تکنیک بگینگ^۷ استفاده شد؛ به‌این‌صورت که هر درخت با استفاده از یک زیرمجموعه تصادفی از داده‌های آموزشی و زیرمجموعه‌ای تصادفی از ویژگی‌ها (GPPs و متغیرهای مکانی-زمانی) آموزش می‌بیند. این فرآیند باعث کاهش واریانس مدل و جلوگیری از بیش‌برازش می‌شود، مشکلی که اغلب در مدل‌های شبکه‌عصبی منفرد دیده می‌شود (Tosan و همکاران، ۲۰۲۶). در نهایت خروجی مدل از طریق میانگین‌گیری خروجی تمام

لازم به ذکر است اگرچه مأموریت ماهواره TRMM در سال ۲۰۱۵ پایان یافته است، اما محصول استاندارد پژوهشی آن (TMPA 3B42 V7) تا اواخر سال ۲۰۱۹ کالیبره و منتشر شده است. در این پژوهش، از داده‌های TRMM به‌عنوان یک مرجع تاریخی و بنچ‌مارک برای توسعه متدولوژی در دوره پایه (۲۰۱۹-۲۰۱۰) استفاده شده است. ساختار مدل پیشنهادی کاملاً آنعطاف‌پذیر بوده و برای دوره‌های پس از ۲۰۱۹، محصول GPM IMERG می‌تواند بدون تغییر در ساختار مدل جایگزین TRMM شود. همچنین به‌منظور بهبود عملکرد مدل در شرایط توپوگرافی پیچیده و در نظر گرفتن تغییرات فصلی، چهار متغیر کمکی شامل مؤلفه‌های مکانی (طول و عرض جغرافیایی) و زمانی (ماه و روز سال) به‌عنوان ورودی‌های تکمیلی به مدل معرفی شدند.

درخت‌ها به‌دست می‌آید. تقسیم داده‌ها به صورت تصادفی با نسبت ۷۰ درصد برای آموزش جهت یادگیری الگوهای خطا و ۳۰ درصد برای آزمون جهت ارزیابی تعمیم‌پذیری مدل بر روی داده‌های دیده نشده انجام پذیرفت. مدل RF با هدف یادگیری تابع f_{RF} جهت برآورد بارش مشاهداتی (R_{obs}) بر اساس بردار ویژگی‌های ورودی (X) توسعه یافت (رابطه ۱):

$$R_{pred} = f_{RF}(TRMM, CHIRPS, GSMaP, ERA5, Lat, Lon, Time) \quad (1)$$

که در آن؛ R_{pred} ، مقدار بارش برآورد شده توسط مدل، f_{RF} ؛ تابع برازش‌یافته توسط الگوریتم جنگل تصادفی، ERA5، GSMaP، CHIRPS، TRMM، محصولات بارش شبکه‌ای ماهواره‌ای و بازتحلیل (به‌عنوان ویژگی‌های ورودی)، Lat و Lon موقعیت مکانی (به‌ترتیب عرض و طول جغرافیایی) و Time، متغیر زمان است.

برای ارزیابی کارایی محصول تولیدی در پایش خشکسالی، شاخص SPI در مقیاس‌های ۳ و ۱۲ ماهه محاسبه شد.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_i - O_i)^2} \quad (7)$$

میانگین خطای مطلق (MAE^{۱۲}): متوسط قدرمطلق خطاها را نمایش می‌دهد (رابطه ۸).

$$MAE = \frac{1}{n} \sum_{i=1}^n |E_i - O_i| \quad (8)$$

جهت ارزیابی توانایی مدل در تشخیص رخدادهای بارانی (با آستانه T=1.0 میلی‌متر)، جدول توافقی تشکیل شد که در آن H تعداد رخدادهای صحیح، M بارش‌های از دست رفته و F هشدارهای کاذب می‌باشند. بر این اساس شاخص‌های احتمال تشخیص (POD^{۱۳})، نسبت هشدار کاذب (FAR^{۱۴}) و شاخص موفقیت بحرانی (CSI^{۱۵}) محاسبه شد (روابط ۹ تا ۱۱):

$$POD = \frac{H}{H + F} \quad (9)$$

$$FAR = \frac{F}{H + F} \quad (10)$$

$$CSI = \frac{H}{H + M + F} \quad (11)$$

که در این روابط، H، تعداد رخدادهای بارانی که به‌درستی توسط مدل تشخیص داده شده‌اند، M، تعداد بارش‌های مشاهده‌ای که مدل در تشخیص آن‌ها ناکام بوده است (خطای از دست‌رفته)، F، تعداد هشدارهای کاذب (مدل بارش پیش‌بینی کرده اما در واقعیت بارشی رخ نداده است).

دقت شاخص SPI حاصل از مدل (SPI_{est}) در مقایسه با مقدار مشاهده‌ای (SPI_{obs}) با استفاده از ضریب تعیین (R²) و خطای RMSE ویژه SPI سنجیده شد (روابط ۱۲ و ۱۳):

$$R^2 = 1 - \frac{\sum_{i=1}^n (SPI_{est} - SPI_{obs})^2}{\sum_{i=1}^n (SPI_{obs} - \overline{SPI_{obs}})^2} \quad (12)$$

$$RMSE_{SPI} = \sqrt{\frac{1}{n} \sum_{i=1}^n (SPI_{est} - SPI_{obs})^2} \quad (13)$$

که در آن:

SPI_{est}: مقدار شاخص خشکسالی برآورد شده توسط مدل.
SPI_{obs}: مقدار شاخص خشکسالی محاسبه‌شده از داده‌های مشاهده‌ای (ایستگاه).
SPI_{obs}: میانگین مقادیر شاخص خشکسالی مشاهده‌ای.

نتایج و بحث

تحلیل ویژگی‌های آماری و ساختار داده‌های ورودی

پیش از ارزیابی عملکرد مدل همجوشی، شناخت ماهیت آماری داده‌های ورودی برای درک چالش‌های مدل‌سازی در مناطق خشک ضروری است. جدول (۳) آماره‌های توصیفی داده‌های بارش مشاهده‌ای (ایستگاهی) و محصولات شبکه‌بندی شده را در طول دوره مطالعه نشان می‌دهد. بررسی داده‌های زمینی نشان‌دهنده میانگین بارش روزانه بسیار پایین (۰/۲۸ میلی‌متر) و انحراف معیار بالا (۱/۶۹ میلی‌متر) است که بیانگر ناهمگنی شدید زمانی و مکانی بارش در منطقه خراسان جنوبی می‌باشد. نکته حائز اهمیت،

باتوجه به اقلیم فراخشک منطقه و کثرت روزهای بدون بارش، برازش توزیع گاما با خطا مواجه می‌شود (Wu و همکاران، ۲۰۲۵). ازاین‌رو، در این پژوهش از روش ناپارامتری مبتنی بر احتمال تجربی استفاده شد. در این روش، ابتدا احتمال رخداد بارش صفر (P₀) با استفاده از موقعیت ترسیمی گرینگورت^{۱۸} (Raziei و Pereira، ۲۰۲۴) محاسبه می‌شود (رابطه ۲):

$$P_0 = \frac{m_{zero} - 0.44}{n + 0.12} \quad (2)$$

که در آن n تعداد کل مشاهدات، m_{zero} تعداد رخداد بارش صفر و P₀ احتمال رخداد بارش صفر است. سپس برای مقادیر بارش غیرصفر، رتبه‌بندی انجام شده و احتمال تجمعی آن‌ها F_{pos} با استفاده از موقعیت ترسیمی کینین^۹ محاسبه می‌شود (رابطه ۳):

$$F_{pos} = \frac{R_i - 0.4}{n_{pos} + 0.2} \quad (3)$$

که در آن، F_{pos}، احتمال تجمعی برای مقادیر بارش غیرصفر، R_i رتبه مقدار بارش در بین داده‌های غیر صفر (به‌ترتیب صعودی) و n_{pos} تعداد مشاهدات با بارش غیرصفر است. درنهایت، تابع توزیع تجمعی نهایی (F(x)) از رابطه (۴) به دست آمده و درنهایت مقادیر نهایی SPI از طریق تبدیل معکوس توزیع نرمال استاندارد ϕ^{-1} استخراج می‌شود (رابطه ۵):

$$F(x) = P_0 + (1 - P_0) \times F_{pos} \quad (4)$$

$$SPI = \phi^{-1}(F(x)) = Z \quad (5)$$

که در آن: F(x)، تابع توزیع تجمعی بارش، x، مقدار بارش SPI مقدار شاخص استاندارد بارش (نرمال‌شده) و ϕ^{-1} تابع وارون توزیع نرمال استاندارد می‌باشند.

۲-۴- معیارهای ارزیابی

عملکرد مدل در دو سطح برآورد کمی بارش و تشخیص رخداد، با استفاده از معیارهای آماری ارزیابی شد. پارامترهای مشترک در ارزیابی‌های کمی (روابط ۶ تا ۸) عبارتند از: n: تعداد کل داده‌های مورد ارزیابی.
O_i: مقدار بارش مشاهده‌ای در گام i.
E_i: مقدار بارش برآوردی (مدل) در گام i.
 $\bar{O}$: میانگین مقادیر بارش مشاهده‌ای.
 $\bar{E}$: میانگین مقادیر بارش برآوردی.

معیارهای پیوسته دقت کمی مقادیر بارش برآوردی (E_i) را نسبت به مقادیر مشاهده‌ای O_i می‌سنجند:

ضریب همبستگی پیرسون (CC^{۱۰}): بیانگر میزان همبستگی خطی است (رابطه ۶).

$$CC = \frac{\sum_{i=1}^n (O_i - \bar{O})(E_i - \bar{E})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (E_i - \bar{E})^2}} \quad (6)$$

ریشه میانگین مربعات خطا (RMSE^{۱۱}): حساسیت مدل به خطاهای بزرگ را نشان می‌دهد (رابطه ۷).

پدیده تورم صفر در داده‌ها است؛ به‌طوری‌که مقادیر میانه و صدک هفتاد و پنجم برابر با صفر می‌باشد. این ساختار نامتوازن داده‌ای که ویژگی ذاتی مناطق فراخشک است، چالش اصلی الگوریتم‌های رگرسیونی محسوب می‌شود، زیرا مدل‌ها تمایل دارند به سمت کلاس اکثریت (روزهای خشک) بایاس پیدا کنند (Li و همکاران، ۲۰۲۳). مقایسه آماره‌های محصولات ماهواره‌ای با داده‌های مشاهداتی، وجود خطای سیستماتیک در تمام GPP ها را آشکار می‌سازد. محصولات TRMM و ERA5 با میانگین بارش روزانه به ترتیب ۰/۳۵ و

۰/۳۳ میلی‌متر، تمایل آشکاری به بیش‌برآورد بارش دارند. در مقابل، عدم تطابق مقادیر حداکثر بارش ثبت‌شده (به‌عنوان مثال ۶۵/۳ میلی‌متر در TRMM در برابر ۴۲/۶ میلی‌متر مشاهداتی) حاکی از عدم قطعیت بالای این محصولات در برآورد رخدادهای حدی^{۱۶} است. این یافته‌ها با نتایج سایر پژوهشگران که بر لزوم اصلاح خطای محصولات ماهواره‌ای پیش از استفاده در مطالعات هیدرولوژیکی تأکید داشتند، همخوانی دارد (Iqbal و همکاران، ۲۰۲۲).

جدول ۳- آماره‌های توصیفی متغیرهای ورودی (GPPs، مکانی و زمانی) و متغیر هدف (بارش مشاهداتی) مورد استفاده در مدل

متغیر	میانگین	انحراف معیار	حداکثر	صدک ۷۵	صدک ۵۰ (میانه)	صدک ۲۵	حداقل
	Mm	Mm	Mm	Mm	Mm	mm	mm
بارش مشاهداتی	۰/۲۸	۱/۶۹	۴۲/۶	۰	۰	۰	۰
TRMM	۰/۳۵	۲/۲۳	۶۵/۳	۰	۰	۰	۰
CHIRPS	۰/۳۲	۱/۴۳	۳۰	۰	۰	۰	۰
GSMaP	۰/۲۵	۱/۳۳	۳۵	۰	۰	۰	۰
ERA	۰/۳۳	۱/۶۲	۳۴/۳	۰	۰	۰	۰
Latitude	۳۳/۲۸	۰/۸۶	۳۴	۳۳/۹	۳۳/۶	۳۲/۹	۳۱/۵
Longitude	۵۸/۵	۱/۰۸	۶۰	۵۹/۳	۵۹/۲	۵۷/۵	۵۷
Month	۶/۵۲	۳/۴۵	۱۲	۱۰	۷	۴	۱
روز از سال	۱/۱۸۳	۱۰۵/۴۳	۳۶۶	۲۷۴	۱۸۳	۹۲	۱

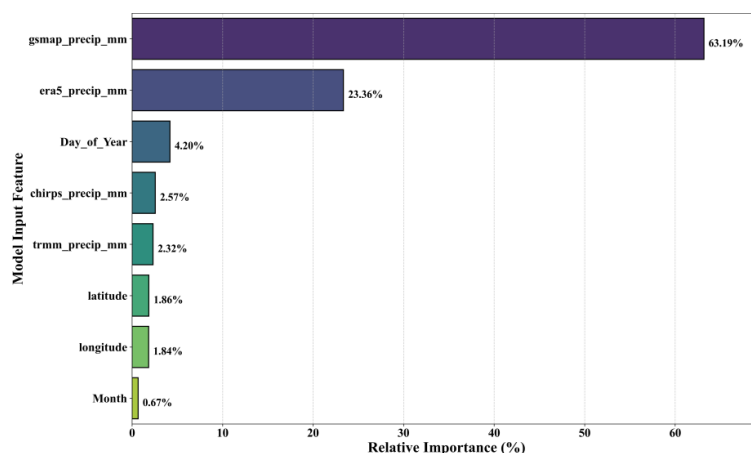
ارزیابی عملکرد مدل و واکاوی اهمیت ویژگی‌ها

عملکرد مدل RF بر روی مجموعه داده آزمون (۳۰٪ داده‌های دیده نشده)، توانایی این الگوریتم را در استخراج الگوهای غیرخطی بارش تأیید می‌کند. جدول (۴) معیارهای ارزیابی مدل را ارائه می‌دهد.

جدول ۴- معیارهای ارزیابی عملکرد مدل بر روی مجموعه داده آزمون (۳۰ درصد)

معیار	مقدار	واحد
R ²	۰/۵۱	-
RMSE	۱/۱۹۴	میلیمتر در روز
MAE	۰/۲۶۳	میلیمتر در روز

مدل پیشنهادی با ضریب تبیین R² برابر با ۰/۵۱ و خطای میانگین مطلق (MAE) پایین (۰/۲۶۳ میلی‌متر)، عملکرد قابل‌قبولی در برآورد بارش روزانه داشته است. اگرچه مقدار RMSE (۱/۱۹۴ میلی‌متر) به دلیل حساسیت به خطاهای بزرگ در مقادیر حدی، بالاتر به نظر می‌رسد، اما اختلاف قابل‌توجه آن با MAE نشان‌دهنده دقت بالای مدل در برآورد رخدادهای نرمال و روزهای بدون بارش است. این رفتار مدل RF که ناشی از میانگین‌گیری روی درخت‌های تصمیم است، منجر به کاهش واریانس خطا می‌شود (Feng و همکاران، ۲۰۱۹). یکی از مزایای الگوریتم RF، قابلیت تعیین اهمیت نسبی ویژگی‌ها^{۱۷} است که در شکل (۲) نمایش داده شده است.

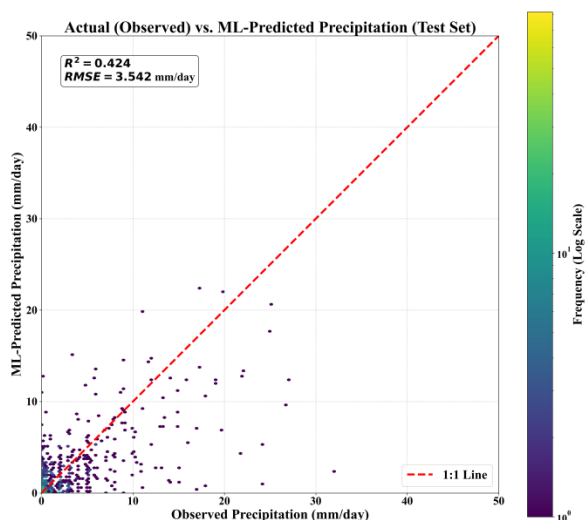


شکل ۲- نمودار اهمیت نسبی ویژگی‌های ورودی در مدل جنگل تصادفی.

۳-۳- تحلیل پراکندگی و رفتار مدل در رژیم‌های مختلف بارش

برای بررسی دقیق‌تر عملکرد مدل در بازه‌های مختلف بارش، نمودار پراکندگی چگال^{۱۸} در شکل (۳) نشان داده شده است. تراکم بالای داده‌ها در نزدیکی مبدأ مختصات (لکه زرد رنگ) نشان‌دهنده توانایی مدل در بازتولید صحیح فراوانی بالای روزهای خشک است که برای پایش دقیق خشکسالی در مناطق خشک بسیار حیاتی است. توزیع داده‌ها حول خط ۱:۱ بیانگر همبستگی مناسب مدل است؛ با این حال، با افزایش شدت بارش، پراکندگی نقاط افزایش می‌یابد. تمایل نقاط به قرارگیری در زیر خط برازش در بارش‌های سنگین (کم‌برآورد) و بالای خط در بارش‌های خفیف (بیش‌برآورد)، یک رفتار رایج در مدل‌های رگرسیونی یادگیری ماشین است که ناشی از ماهیت میانگین‌گیر توابع هدف (مانند MSE) می‌باشد (Chen و همکاران، ۲۰۲۵). با وجود این محدودیت ذاتی، مدل توانسته است تعادل مناسبی میان تشخیص رخداد و برآورد مقدار برقرار نماید که برای مطالعات بیلان آب بلندمدت کافی به نظر می‌رسد.

شکل (۲) نشان می‌دهد که محصول GSMap با سهم ۶۳/۱۹٪، مؤثرترین متغیر ورودی در مدل بوده است. این یافته با نتایج سایر محققان که عملکرد بهتر محصولات مایکروویو غیرفعال (مانند GSMap) را در مناطق خشک گزارش کرده‌اند، مطابقت دارد (Zhang و همکاران، ۲۰۲۱). پس از آن، محصول بازتحلیل ERA5 با سهم ۲۳/۳۶٪ در رتبه دوم قرار دارد. نکته کلیدی و متمایز این پژوهش، نقش متغیرهای کمکی مکانی و زمانی (طول و عرض جغرافیایی، روز سال و ماه) با مجموع سهم حدود ۸ درصد است. اگرچه این سهم در ظاهر اندک به نظر می‌رسد، اما نقش حیاتی در تصحیح خطاهای محلی ناشی از توپوگرافی پیچیده منطقه و خطاهای فصلی ایفا می‌کند. این نتیجه مؤید یافته‌های موجود است که نشان دادند ادغام متغیرهای محیطی نظیر مختصات جغرافیایی در مدل‌های یادگیری ماشین، دقت ریزمقیاس‌نمایی بارش را به‌طور معناداری افزایش می‌دهد (Zhu و همکاران، ۲۰۲۳).

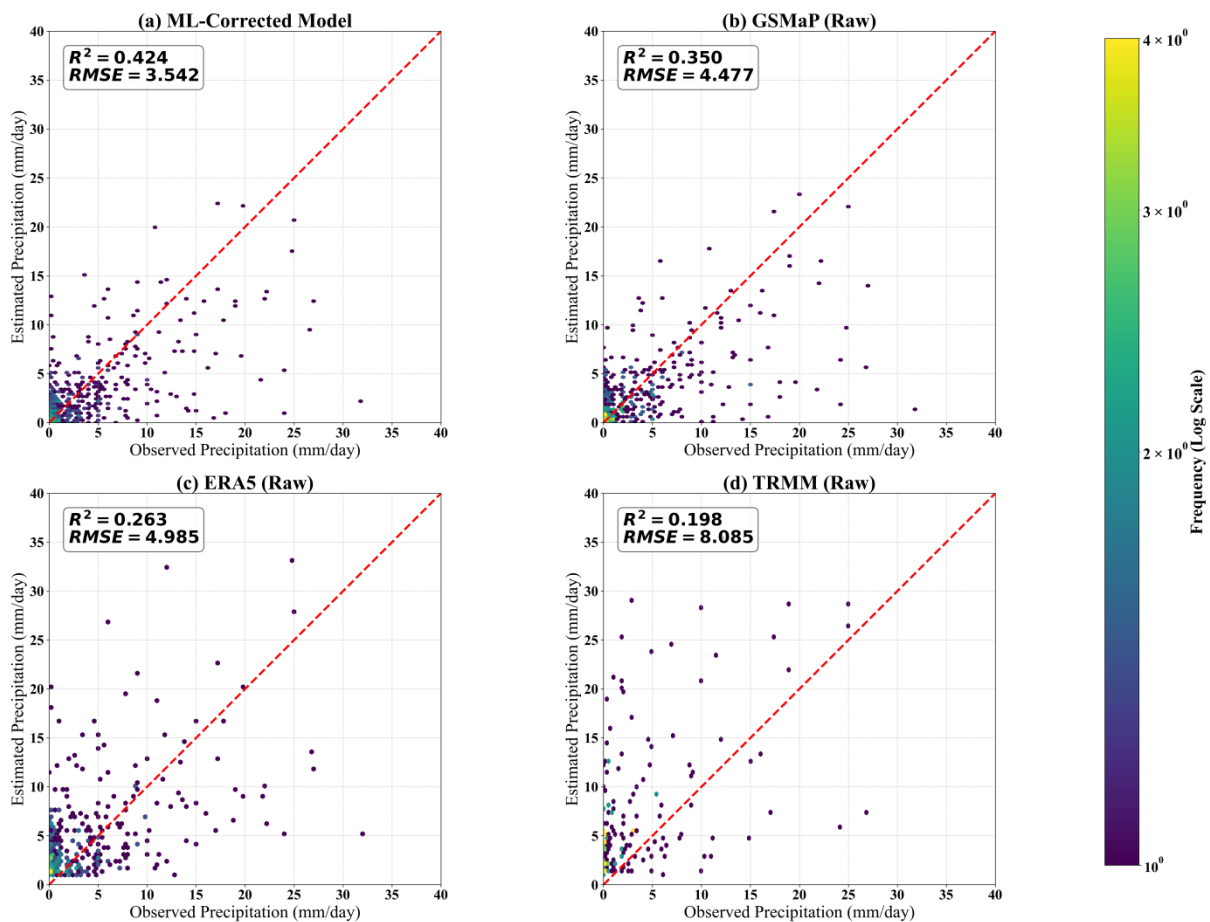


شکل ۳- نمودار پراکندگی چگال مقایسه بارش مشاهداتی (محور افقی) و بارش پیش‌بینی شده توسط مدل

تحلیل تطبیقی؛ اثربخشی مدل ادغام‌شده در برابر محصولات خام

برای درک بهتر ارزش افزوده مدل پیشنهادی و اثبات فرضیه اصلی پژوهش مبنی بر برتری روش‌های همجوشی، عملکرد مدل در مقایسه با محصولات ورودی خام در شکل (۴) ارائه شده است. شکل (۴) نشان‌دهنده بهبود چشمگیر دقت برآورد بارش در محصول ML-Corrected (پانل a) نسبت به تمامی محصولات خام است. مدل پیشنهادی با $R^2=0.63$ و $RMSE=1/34$ میلی‌متر، عملکرد بهتری نسبت به بهترین محصول خام (GSMaP با $R^2=0.51$) و ضعیف‌ترین محصول (TRMM با $R^2=0.28$) ارائه داده است. کاهش پراکندگی نقاط

حول خط ۱:۱ در محصول اصلاح شده نشان می‌دهد که مدل همجوشی صرفاً یک میانگین‌گیری وزنی نیست، بلکه با شناسایی و یادگیری الگوهای خطای غیرخطی محصولات مختلف، ترکیبی بهینه با کمترین واریانس خطا تولید کرده است. این نتیجه همسو با یافته‌های سایر پژوهشگران است که نشان دادند الگوریتم‌های یادگیری ماشین در تلفیق داده‌های چندمنبعی بارش، برتری آماری معناداری نسبت به تک‌محصول‌ها دارند و قادر به کاهش عدم قطعیت‌های ذاتی محصولات ماهواره‌ای می‌باشند (Ghosh و همکاران ۲۰۲۴؛ Das و همکاران، ۲۰۲۴).



شکل ۴- مقایسه نمودارهای پراکندگی چگال بارش مشاهداتی در برابر (a) محصول ML-Corrected، (b) محصول خام GSmAP، (c) محصول خام ERA5، و (d) محصول خام TRMM

مدل RF به‌صورت دینامیک و بر اساس شرایط مکانی، وزن بیشتری به محصولات دقیق‌تر (GSMaP) اختصاص داده و نویز محصولات ضعیف‌تر را فیلتر کرده است.

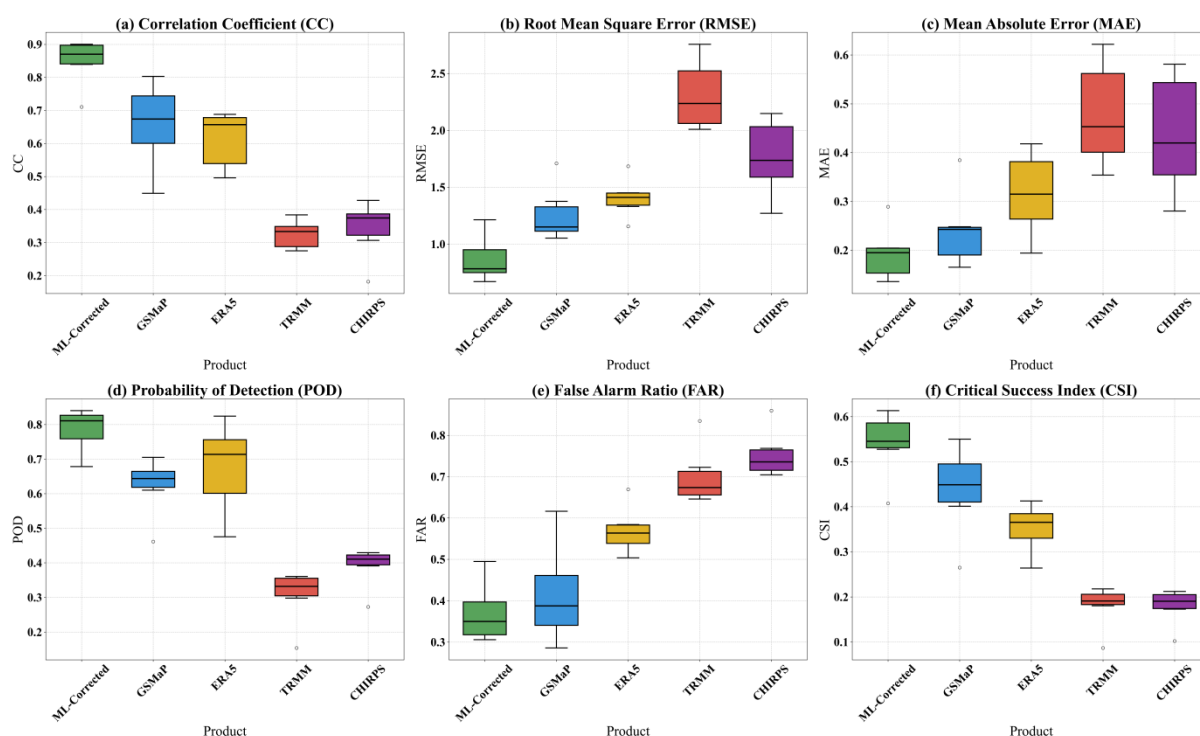
ارزیابی جامع آماری و قابلیت اعتماد مکانی

به‌منظور بررسی پایداری^{۱۹} عملکرد مدل در شرایط مختلف جغرافیایی و اطمینان از اینکه بهبود عملکرد محدود به نقاط خاصی نیست، توزیع شاخص‌های آماری در تمام ایستگاه‌های مورد مطالعه در شکل (۵) و جدول (۵) خلاصه شده است.

تحلیل کمی نتایج نشان‌دهنده نرخ بهبود قابل‌توجهی در مدل ادغام شده است. مدل پیشنهادی توانسته است خطای $RMSE$ را نسبت به دقیق‌ترین محصول خام (GSMaP) حدود ۳۲ درصد و نسبت به ضعیف‌ترین محصول (TRMM) بیش از ۶۰ درصد کاهش دهد. دلیل اصلی این اختلاف عملکرد و برتری مدل همجوشی، توانایی یادگیری ماشین در شناسایی روابط غیرخطی بین متغیرهای کمکی (مانند ارتفاع و فصل) و خطای محصولات است. درحالی‌که روش‌های سنتی (مانند میانگین‌گیری ساده) وزن یکسانی به محصولات می‌دهند،

یافته است. پژوهشگران در مطالعه‌ای بر اهمیت کاهش هشدارهای کاذب در کاربردهای هواشناسی تأکید کرده‌اند (Liu و همکاران، ۲۰۱۹). این بهبود همزمان در معیارهای تشخیص^{۲۰} و تخمین مقدار بارش، نشان‌دهنده کارایی بالای مدل در تفکیک روزهای تر و خشک و حذف نویزهای ناشی از خطاهای حسگرهای ماهواره‌ای است. این ویژگی برای پایش خشکسالی در مناطق خشک که وقوع بارش‌های کاذب می‌تواند منجر به خطای فاحش در محاسبه شاخص‌های تجمعی نظیر SPI شود، بسیار حائز اهمیت است.

تحلیل نمودارهای جعبه‌ای نشان می‌دهد که محصول ML-Corrected در تمامی معیارهای پیوسته (CC, RMSE, MAE) بهبود قابل توجهی داشته و توزیع خطای فشرده‌تری را نسبت به محصولات خام نشان می‌دهد (میان‌همبستگی ۰/۸۷). علاوه بر دقت کمی، یکی از دستاوردهای مهم این پژوهش، بهبود توانایی تشخیص رخداد بارش است. شاخص CSI برای محصول اصلاح‌شده به ۰/۵۴۵ افزایش یافته است که نسبت به GSMaP (۰/۴۴۹) و ERA5 (۰/۳۶۶) برتری مطلق دارد. همچنین، نرخ هشدار کاذب (FAR) به ۰/۳۵ کاهش



شکل ۵- مقایسه آماری توزیع خطا محصول ML-Corrected در برابر GPP‌های خام در تمام ایستگاه‌ها برای (a) CC، (b) RMSE، (c) MAE، (d) POD، (e) FAR و (f) CSI

جدول ۵- خلاصه میانه معیارهای آماری ارزیابی بارش روزانه در تمام ایستگاه‌ها

محصول	CC	RMSE (mm/day)	MAE (mm/day)	CSI	FAR	POD
ML-Corrected	۰/۸۷	۰/۷۸	۰/۱۹	۰/۵۴	۰/۳۵	۰/۸۱
GSMaP	۰/۶۷	۱/۱۵	۰/۲۴	۰/۴۵	۰/۳۹	۰/۶۴
ERA5	۰/۶۵	۱/۴۱	۰/۳۱	۰/۳۷	۰/۵۶	۰/۷۱
TRMM	۰/۳۳	۲/۲۳	۰/۴۵	۰/۱۹	۰/۶۷	۰/۳۳
CHIRPS	۰/۳۷	۱/۷۳	۰/۴۲	۰/۱۹	۰/۷۴	۰/۴۱

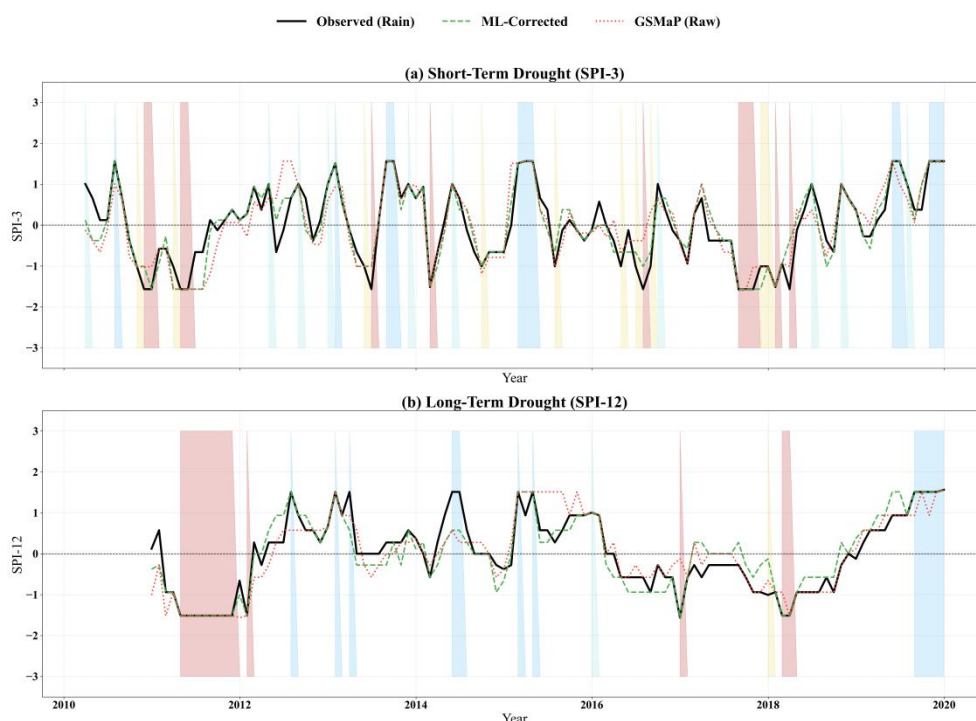
Corrected انطباق بسیار بالایی با مقادیر مشاهداتی دارد. مدل پیشنهادی توانسته است نوسانات زمانی، شدت و مدت زمان رخدادهای خشکسالی و ترسالی را با دقت بالا بازتولید کند. به‌عنوان مثال در دوره خشکسالی شدید سال ۲۰۱۱، محصول اصلاح شده دقیقاً هم‌فاز با داده‌های زمینی عمل کرده است، درحالی‌که محصول خام GSMaP دچار تأخیر زمانی و بیش‌برآورد شدت خشکسالی بوده است. این

ارزیابی هیدرولوژیکی: پایش خشکسالی (SPI)

هدف نهایی این پژوهش، ارزیابی قابلیت محصول اصلاح شده در کاربردهای واقعی هیدرولوژیکی، به‌ویژه پایش خشکسالی است. شکل (۶) سری زمانی شاخص SPI را برای یک ایستگاه نمونه در دو مقیاس زمانی کوتاه‌مدت (SPI-3) و بلندمدت (SPI-12) نمایش می‌دهد. تحلیل بصری شکل (۶) نشان می‌دهد که سری زمانی SPI حاصل از محصول ML-

همچوشی پیشنهادی موفق به رفع آن شده است. ارزیابی آماری شاخص SPI در تمام ایستگاه‌ها در شکل (۷) و جدول (۶) ارائه شده است.

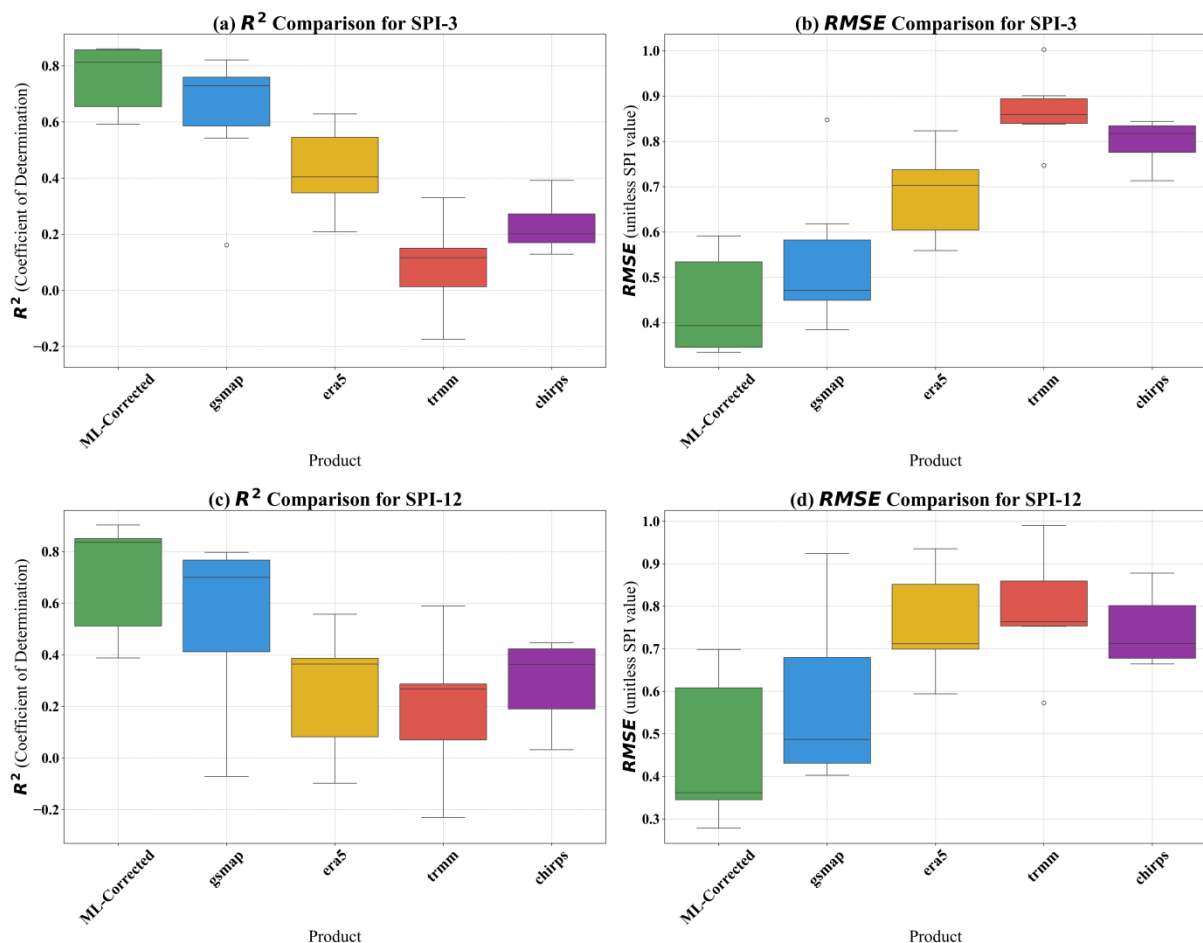
تأخیر زمانی در محصولات ماهواره‌ای که توسط کاشانی و صفوی (Kashani و Safavi, ۲۰۲۵) نیز در پایش خشکسالی‌های آب زیرزمینی گزارش شده است، می‌تواند منجر به تصمیم‌گیری‌های مدیریتی نادرست شود که مدل



شکل ۶- مقایسه سری زمانی SPI در ایستگاه نماینده برای (a) خشکسالی کوتاه‌مدت (SPI-3) و (b) خشکسالی بلندمدت (SPI-12)

جدول ۶- خلاصه میانه معیارهای ارزیابی SPI در تمام ایستگاه‌ها

مقیاس	محصول	RMSE (خطای SPI)	R ²
SPI-3	ML-Corrected	۰/۳۹۳	۰/۸۱۲
	Gsmap	۰/۴۷۱	۰/۷۲۸
	era5	۰/۷۰۳	۰/۴۰۴
	Trmm	۰/۸۶	۰/۱۱۶
	Chirps	۰/۸۱۷	۰/۲۰۱
SPI-12	ML-Corrected	۰/۳۶۲	۰/۸۳۶
	Gsmap	۰/۴۸۷	۰/۷
	era5	۰/۷۱۲	۰/۳۶۴
	Trmm	۰/۷۶۴	۰/۲۶۷
	Chirps	۰/۷۱۳	۰/۳۶۲



شکل ۷- مقایسه آماری عملکرد SPI محصولات مختلف در برابر SPI مشاهداتی در تمام ایستگاه‌ها. (a) R^2 برای SPI-3، (b) RMSE برای SPI-3، (c) R^2 برای SPI-12، (d) RMSE برای SPI-12.

استراتژی یادگیری گروهی^{۲۱} برای ارزیابی فضایی-زمانی خشکسالی استفاده کردند، همخوانی دارد (Rezaie-Balf & Chua, ۲۰۲۵). آن‌ها نیز نشان دادند که تلفیق داده‌های چندمنبعی می‌تواند خطاهای ساختاری در سری‌های زمانی بارش را کاهش داده و پایش دقیق‌تری از خشکسالی ارائه دهد. بنابراین، مدل پیشنهادی در این پژوهش نه تنها یک ابزار آماری، بلکه یک راهکار عملیاتی برای غلبه بر محدودیت‌های ذاتی داده‌های ماهواره‌ای در مناطق خشک و نیمه‌خشک محسوب می‌شود.

نتیجه‌گیری

پژوهش حاضر با هدف غلبه بر چالش‌های پایش هیدرولوژیکی در مناطق فراخشک و فاقد داده، کارایی یک مدل همجوشی مبتنی بر یادگیری ماشین را برای تلفیق محصولات بارش ماهواره‌ای و بازتحلیل (TRMM, CHIRPS, GSMaP, ERA5) ارزیابی کردند. نتایج حاصل از این مطالعه منجر به استنتاج‌های کلیدی زیر شد:

۱. مدل پیشنهادی با بهره‌گیری از متغیرهای کمکی مکانی-زمانی، توانست خطاهای سیستماتیک موجود در محصولات

نتایج کمی نشان می‌دهد که محصول اصلاح شده در شاخص R^2 برای SPI-3 و SPI-12 به مقادیر ۰/۸۱۲ و ۰/۸۳۶ دست یافته است، درحالی‌که این مقادیر برای بهترین محصول خام (GSMaP) کمتر از ۰/۷۳ بوده است. نکته قابل توجه، عملکرد بسیار ضعیف محصولات TRMM و CHIRPS در پایش خشکسالی این منطقه است (با R^2 کمتر از ۰/۲۷)، که لزوم استفاده از روش‌های اصلاحی را دوچندان می‌کند. همچنین خطای RMSE شاخص SPI برای محصول پیشنهادی به کمتر از ۰/۴ واحد کاهش یافته است. این سطح از دقت، قابلیت اعتماد به این محصول را برای سیستم‌های هشدار زودهنگام خشکسالی و مدیریت منابع آب در مناطق فاقد داده تأیید می‌کند. دقت مدل در شاخص SPI-12 (بلندمدت) بالاتر از SPI-3 (کوتاه‌مدت) است. دلیل فیزیکی این امر، ماهیت «فیلتر پایین‌گذر» در مقیاس‌های زمانی طولانی است. در محاسبه SPI دوازده ماه، خطاهای تصادفی و نوسانات شدید بارش‌های روزانه یا ماهانه در یک بازه زمانی طولانی‌تر مستهلک می‌شوند و روند کلی دقیق‌تری باقی می‌ماند. در مقابل، SPI سه ماهه حساسیت بالایی به هرگونه خطا در تک‌رخدادهای بارش دارد که منجر به نوسانات بیشتر در همبستگی می‌شود. این نتایج با یافته‌های اخیر که از

Bhaga, T. D., Dube, T., Shekede, M. D., & Shoko, C. (2020). Impacts of Climate Variability and Drought on Surface Water Resources in Sub-Saharan Africa Using Remote Sensing: A Review. *Remote Sensing*, 12(24), 4184. <https://doi.org/10.3390/rs12244184>

Bouaziz, M., Medhioub, E., & Csaplovisc, E. (2021). A machine learning model for drought tracking and forecasting using remote precipitation data and a standardized precipitation index from arid regions. *Journal of Arid Environments*, 189, 104478. <https://doi.org/10.1016/j.jaridenv.2021.104478>

Bui, H. T., Ishidaira, H., & Shaowei, N. (2019). Evaluation of the use of global satellite-gauge and satellite-only precipitation products in stream flow simulations. *Applied Water Science*, 9(3), 53. <https://doi.org/10.1007/s13201-019-0931-y>

Chen, J., Chen, K., & Zhu, Y. (2025). Hybrid CNN-LSTM-Attention Model for Enhanced Reference Crop Evapotranspiration Estimation toward Optimized Irrigation Scheduling. *Journal of Agriculture and Food Research*, 102545. <https://doi.org/10.1016/j.jafr.2025.102545>

Das, P., Zhang, Z., Ghosh, S., & Hang, R. (2024). A hybrid ensemble learning merging approach for enhancing the super drought computation over Lake Victoria Basin. *Scientific Reports*, 14(1), 13870. <https://doi.org/10.1038/s41598-024-61520-6>

Feizi, H., & Tosan, M. (2017). Saffron yield variability by climatic factors in the northeast of Iran. *Acta Horticulturae*, 1184, 109-114. <https://doi.org/10.17660/ActaHortic.2017.1184.15>

Feng, P., Wang, B., Liu, D. L., & Yu, Q. (2019). Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. *Agricultural Systems*, 173, 303-316. <https://doi.org/10.1016/j.agsy.2019.03.015>

Ghosh, S., Lu, J., Das, P., & Zhang, Z. (2024). Machine learning algorithms for merging satellite-based precipitation products and their application on meteorological drought monitoring over Kenya. *Climate Dynamics*, 62(1), 141-163. <https://doi.org/10.1007/s00382-023-06893-6>

Hashemi, Z., Sodaeezadeh, H., Mokhtari, M. H., Ardakani, M. A. H., & Aliabadi, K. K. (2024). Monitoring and forecasting desertification and land degradation using remote sensing and machine learning techniques in Sistan plain, Iran.

خام را به‌طور مؤثر شناسایی و اصلاح نماید. دستیابی به محصول اصلاح شده (ML-Corrected) با ضریب همبستگی میانه ۰/۸۷ و کاهش خطای RMSE، نشان‌دهنده برتری مطلق این رویکرد نسبت به استفاده منفرد از دقیق‌ترین محصول خام (GSMaP) است. همچنین، بهبود شاخص‌های تشخیص (CSI و POD) بیانگر توانایی مدل در تفکیک صحیح روزهای تر و خشک و حذف نویزهای ناشی از خطاهای حسگر است.

۲. مهمترین دستاورد این پژوهش، ترجمه بهبود دقت کمی بارش به ارتقای شاخص‌های پایش خشکسالی است. انطباق بالای سری زمانی SPI محصول اصلاح شده با داده‌های مشاهداتی ($R^2 > 0.81$) در هر دو مقیاس کوتاه‌مدت و بلندمدت، ثابت کرد که روش همجوشی قادر به رفع چالش‌های تأخیر زمانی و بیش‌برآورد شدت در محصولات ماهواره‌ای است.

۳. یافته‌ها نشان می‌دهد مدل توسعه‌یافته فراتر از یک ابزار آماری، به‌عنوان یک سامانه اصلاح‌گر هوشمند عمل می‌کند. وزن‌دهی دینامیک به ورودی‌ها بر اساس دقت آن‌ها (مانند سهم بیشتر GSMaP و ERA5)، این مدل را به ابزاری قابل اتکا برای تولید داده‌های بارش در مناطق با شبکه پایش زمینی ضعیف تبدیل کرده است.

در مجموع، محصول تولیدی این پژوهش پتانسیل بالایی جهت به‌کارگیری در سیستم‌های هشدار زودبهنگام خشکسالی و مدل‌سازی بیلان آب در استان خراسان جنوبی و مناطق مشابه دارد. برای پژوهش‌های آینده، استفاده از معماری‌های یادگیری عمیق مانند شبکه‌های عصبی بازگشتی جهت درک بهتر وابستگی‌های زمانی سری‌های بارش و همچنین توسعه نقشه‌های بارش شبکه‌بندی شده با وضوح مکانی بالا در سطح حوضه پیشنهاد می‌شود.

منابع

رضوانی مقدم، پ، کرباسی، ع، توسن، م، قراری، ف، فیضی، ح، و محتشمی، ت. (۱۳۹۵). *زراعت و فناوری زعفران* (Book of Abstracts: 2013-2016). زراعت و فناوری زعفران، ۴(ویژه‌نامه)، ۷۸-۱. <https://doi.org/10.22048/jsat.2016.39250>

توسن، م، خاشعی سیوکی، ع، سنگری، م، و رضوانی مقدم، پ. (۱۴۰۳). *آنالیز روند جهانی تحقیقات زعفران* (Crocus sativus L) در بین سال‌های ۲۰۰۰ تا ۲۰۲۳. زراعت و فناوری زعفران، ۱۲ (۲)، ۱۱۵-۱۳۸. <https://doi.org/10.22048/jsat.2024.443037.1524>

Batista, F. F., Rodrigues, D. T., Santos e Silva, C. M., Andrade, L. d. M. B., Mutti, P. R., Potes, M., & Costa, M. J. (2025). Performance Assessment of IMERG V07 Versus V06 for Precipitation Estimation in the Parnaíba River Basin. *Remote Sensing*, 17(21), 3613. <https://doi.org/10.3390/rs17213613>

- Mahmud, H. B., & Osawa, T. (2025). Analysis of the social impact of meteorological drought on Peninsular Malaysia based on the integration of precipitation satellites and sentiment analysis. *International Journal of Disaster Risk Reduction*, 119, 105314. <https://doi.org/10.1016/j.ijdrr.2025.105314>
- Michibata, T. (2024). Radiative effects of precipitation on the global energy budget and Arctic amplification. *npj Climate and Atmospheric Science*, 7(1), 136. <https://doi.org/10.1038/s41612-024-00684-4>
- Mwinjuma, M., Wang, R., Mtupili, M., & Twaha, M. (2026). Comparisons of SPI and SPEI in capturing drought dynamics: A Global assessment across arid and humid regions. *Atmospheric Research*, 329, 108475. <https://doi.org/10.1016/j.atmosres.2025.108475>
- Negri, P., Acevedo, D., Ruiz, J., Gonzalez, S., Vidal, L., Silvarrey, A., & Nicora, M. G. (2025). Enhancing precipitation detection: A multi-sensor approach using conditional GANs and recurrent networks. *Pattern Recognition Letters*, <https://doi.org/10.1016/j.patrec.2025.05.022>
- Nikoo, M. R., Tosan, M., Nazari, R., Vahabzadeh, M., Karimi, M., & Zarrinabadi, E. (2026). Assessing the fidelity of multi-satellite precipitation estimates for drought monitoring in a mountain water tower to arid basin system. *Journal of Arid Environments*, 232, 105519. <https://doi.org/10.1016/j.jaridenv.2025.105519>
- Nourani, V., Tosan, M., Huang, J. J., Gebremichael, M., Kantoush, S. A., & Dastourani, M. (2025). Advances in multi-source data fusion for precipitation estimation :remote sensing and machine learning perspectives. *Earth-Science Reviews*, 270, 105253. <https://doi.org/10.1016/j.earscirev.2025.105253>
- Oyarzabal, R. S., Santos, L. B. L., Cunningham, C., Broedel, E., de Lima, G. R. T., Cunha-Zeri, G., Peixoto, J. S., Anochi, J. A., Garcia, K., Costa, L. C. O., Pampuch, L. A., Cuartas, L. A., Zeri, M., Guedes, M. R. G., Negri, R. G., Muñoz, V. A., & Cunha, A. P. M. A. (2025). Forecasting drought using machine learning: a systematic literature review. *Natural Hazards*, 121, (8). 9823-9851. <https://doi.org/10.1007/s11069-025-07195-2>
- Rahat, S. H., Saki, S., Khaira, U., Biswas, N. K., Dollan, I. J., Wasti, A., Miura, Y., Bhuiyan, M. A. E., & Ray, P. (2024). Bracing for impact: how shifting precipitation extremes may influence physical Journal of African Earth Sciences, 218, 105375. <https://doi.org/10.1016/j.jafrearsci.2024.105375>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., & Schepers, D. (2020). The ERA5 global reanalysis. *Quarterly journal of the royal meteorological society*, 146(730), 1999-2049. <https://doi.org/10.1002/qj.3803>
- Iqbal, Z., Shahid, S., Ahmed, K., Wang, X., Ismail, T., & Gabriel, H. F. (2022). Bias correction method of high-resolution satellite-based precipitation product for Peninsular Malaysia. *Theoretical and Applied Climatology*, 148(3), 1429-1446. <https://doi.org/10.1007/s00704-022-04007-6>
- Kashani, A., & Safavi, H. R. (2025). Assessing groundwater drought in Iran using GRACE data and machine learning. *Scientific Reports*, 15(1), 14671. <https://doi.org/10.1038/s41598-025-99342-9>
- Kossieris, P., Tsoukalas, I., Brocca, L., Mosaffa, H., Makropoulos, C., & Anghelea, A. (2024). Precipitation data merging via machine learning: Revisiting conceptual and technical aspects. *Journal of Hydrology*, 637, 131424. <https://doi.org/10.1016/j.jhydrol.2024.131424>
- Lalmuanzuala, B., Sathyamoorthy, N., Dheebakaran, G., Jagadeeswaran, R., Ravikumar, R., & Sumathi, C. (2025). A comprehensive review of meteorological and remote sensing indices for drought monitoring and their in situ application in India. *Environmental Reviews*, 33, 1-15. <https://doi.org/10.10139/er-2025-0077>
- Li, X., Jia, H., & Wang, L. (2023). Remote Sensing Monitoring of Drought in Southwest China Using Random Forest and eXtreme Gradient Boosting Methods. *Remote Sensing*, 15(19), 4840. <https://doi.org/10.3390/rs15194840>
- Liu, Q., Li, Y., Yu, M., Chiu, L. S., Hao, X., Duffy, D. Q., & Yang, C. (2019). Daytime Rainy Cloud Detection and Convective Precipitation Delineation Based on a Deep Neural Network Method Using GOES-16 ABI Images. *Remote Sensing*, 11(21), 2555. <https://doi.org/10.3390/rs11212555>
- Maggioni, V., Meyers, P. C., & Robinson, M. D. (2016). A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era. *Journal of Hydrometeorology*, 17(4), 1101-1117. <https://doi.org/10.1175/JHM-D-15-0190.1>

Zhang, Z., Wang, D., Qiu, J., Zhu, J., & Wang, T. (2021). Machine learning approaches for improving near-real-time IMERG rainfall estimates by integrating cloud properties from NOAA CDR PATMOS-x. *Journal of Hydrometeorology*, 22(10), 2767-2781. <https://doi.org/10.1175/JHM-D-21-0019.1>

Zhang, J., Ding, J., Wang, J., Lin, H., Han, L., Li, X., & Liu, J. (2023). Remote sensing drought factor integration based on machine learning can improve the estimation of drought in arid and semi-arid regions. *Theoretical and Applied Climatology*, 151(3), 1753-1770. <https://doi.org/10.1007/s00704-022-04305-z>

Zhang, X., Song, Y., Nam, W.-H., Huang, T., Gu, X., Zeng, J., Huang, S., Chen, N., Yan, Z & ,Niyogi, D. (2024). Data fusion of satellite imagery and downscaling for generating highly fine-scale precipitation. *Journal of Hydrology*, 631, 130665. <https://doi.org/10.1016/j.jhydrol.2024.130665>

Zhu, H., Zhou, Q., & Cui ,A. (2023). Comparison and evaluation of machine-learning-based spatial downscaling approaches on satellite-derived precipitation data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences.*, X-1/W1-2023, 919-924. <https://doi.org/10.5194/isprs-annals-X-1-W1-2023-919-2023>

پی‌نوشت‌ها

- 1- Gridded precipitation products
- 2- Standardized Precipitation Index
- 3- Multi-source fusion
- 4- Machine learning
- 5- Random forest
- 6- Decision Trees
- 7- Bagging
- 8- Gringorten
- 9- Cunnane
- 10- Correlation coefficient
- 11- Root mean square error
- 12- Mean absolute error
- 13- Probability of detection
- 14- False-alarm rate
- 15- Critical success index
- 16- Extreme events
- 17- Feature importance
- 18- Density scatter plot
- 19- Stability
- 20- Detection capability
- 21- Ensemble learning

climate risks in an uncertain future. *Scientific Reports*, 14(1), 17398. <https://doi.org/10.1038/s41598-024-65618-9>

Raziei, T., & Pereira, L. S. (2024). Relating the Drought Precipitation Percentiles Index to the Standardized Precipitation Index (SPI) under Influence of Aridity and Timescale. *Water Resources Management*, 38(14), 5739-5758. <https://doi.org/10.1007/s11269-024-03932-7>

Rezaie-Balf, M., & Chua, L. H. C. (2025). An ensemble learning strategy for spatiotemporal drought assessment through the fusion of multi-satellite and atmospheric reanalysis precipitation products. *Stochastic Environmental Research and Risk Assessment*, 39(9), 3875-3901. <https://doi.org/10.1007/s00477-025-03041-w>

Sharma, T. C., & Panu, U. S. (2024). Current State of Advances in Quantification and Modeling of Hydrological Droughts. *Water*, 16(5), 729. <https://doi.org/10.3390/w16050729>

Shiru, M. S., Chung, E.-S., & Sa'adi, Z. (2025). Performance assessment of satellite precipitation products over Nigeria: A compromise programming approach. *Earth Systems and Environment*, 9(4), 3289-3310. <https://doi.org/10.1007/s41748-024-00563-1>

Tong, X., Zhou, W., & Xia, J. (2024). Improving Boreal Summer Precipitation Predictions From the Global NMME Through Res34-Unet. *Geophysical Research Letters*, 51(2), e2023GL106391. <https://doi.org/10.1029/2023GL106391>

Tosan, M., Gharib, M. R., Mardani, M., & Sabbagh, A. (2026). Spatiotemporal performance and error analysis of satellite precipitation products over a topographically complex semi-arid region in Iran. *Journal of Mountain Science*, 23(1), 118-138. <https://doi.org/10.1007/s11629-025-9984-6>

Tosan, M., Nourani, V., Kisi, O., Zhang, Y., Kantoush, S.A. et al. (2026). The Transparency Revolution in Geohazard Science: A Systematic Review and Research Roadmap for Explainable Artificial Intelligence. *Computer Modeling in Engineering & Sciences*, 146(1), 3. <https://doi.org/10.32604/cmescs.2025.074768>

Wu, X., Liu, Y., & Wang, R. (2025). Assessment of non-parametric method for evapotranspiration estimation across extreme conditions. *Atmospheric Research*, 326, 108279. <https://doi.org/10.1016/j.atmosres.2025.108279>