مجله پژوهش آب ایران
جلد ۱۸/ شماره ۱/ پیاپی ۵۲/ بهار ۱۴۰۳ (؟؟-؟؟)
https://dx.doi.org/10.22034/IWRJ.2023.14295.2506

مقاله پژوهشی

ارزیابی کمی- کیفی داده‌های بارش Era5 در تخمین رخداد و مقدار بارش‌های به وقوع پیوسته در استان خراسان رضوی

علیرضا نوری[۱]، سمیرا نوری[۲]، جواد امیدوار[۳]، حسین بانژاد[۴*] و فرشته مدرسی[۵]

چکیده

اندازه‌گیری دقیق مقدار بارش به عنوان یکی از مهم‌ترین متغیرهای هواشناسی در مطالعات هیدرولوژیکی و کشاورزی است که تأثیر به سزایی در مدیریت بهینه منابع آبی، کشت محصولات کشاورزی، مدیریت شهری و شناسایی و دسته بندی مناطق از نظر میزان ریسک وقوع سیلاب یا قابلیت استحصال آب دارد. عدم پراکندگی مناسب ایستگاه‌های اندازه‌گیری بارش در مناطق مختلف، سبب گردیده است تا محققان و پژوهشگران به دنبال ایجاد مدل‌هایی باشند که بتوانند مقدار بارش را در مناطق فاقد یا دچار کمبود ایستگاه باران‌سنجی برآورد نمایند. در پژوهش حاضر از داده‌های مدلسازی شده Era5 که جدیدترین محصول مرکز اروپایی ECMWF می‌باشد و دقت خروجی‌های آن در منطقه خاورمیانه و ایران چندان مورد بررسی قرار نگرفته، استفاده شده است. به منظور ارزیابی داده‌های بارش Era5، از داده‌های بارش ۵۱ ایستگاه باران‌سنجی در استان خراسان رضوی طی سال‌های ۱۳۷۶ الی ۱۳۹۶ استفاده شد. برای بررسی خروجی‌های مدل مذکور، دو گروه شاخص ارزیابی کیفی (POD، FAR و CSI) و کمی (RMSE، NSE و R^2) مورد استفاده قرار گرفتند. نتایج نشان داد که خروجی‌های Era5 در مقیاس روزانه دارای خطای نسبتاً زیادی هستند. در حالی‌که این داده‌ها در مقیاس‌های ماهانه و فصلی به خصوص فصلی در مناطق شرق، جنوب، جنوب غربی و جنوب شرقی استان که از اقلیم گرم‌تر و بارش‌های کمتری برخوردارند، عملکرد بهتری داشته است.

واژه‌های کلیدی: ارزیابی بارندگی، مدلسازی بارندگی،ECMWF ، NSE، POD.

ارجاع: نوری ع. نوری س. امیدوار ج. بانژاد ح. و مدرسی ف. ۱۴۰۳ ارزیابی کمی- کیفی داده‌های بارش Era5 در تخمین رخداد و مقدار بارش‌های به وقوع پیوسته در استان خراسان رضوی. مجله پژوهش آب ایران. ۵۲: ؟؟-؟؟. https://dx.doi.org/10.22034/IWRJ.2023.14295.2506

۱- فارغ‌التحصیل کارشناسی ارشد اگرواکولوژی، دانشکده کشاورزی، دانشگاه فردوسی مشهد.
۲- فارغ‌التحصیل دکتری هواشناسی کشاورزی، دانشکده کشاورزی، دانشگاه فردوسی مشهد.
۳- دانشجوی دکتری مهندسی آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد.
۴- دانشیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد.
۵- استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد.
*** نویسنده مسئول: Banejad@um.ac.ir**
تاریخ دریافت: ۱۴۰۲/۱۲/۱۸ تاریخ پذیرش: ۱۴۰۲/۰۴/۰۵

مقدمه

برآورد دقیق داده‌های بارش به عنوان یکی از مهم‌ترین متغیرهای آب و هواشناسی که خود ورودی انواع مدل‌های هیدرلوژیکی و کشاورزی است از اهمیت به‌سزایی برخوردار است. علی رغم اهمیت بارش، ایستگاه‌های اندازه‌گیری این متغیر مهم هواشناسی از تعداد و پراکندگی مناسب در اکثر مناطق برخوردار نیستند. همین امر محققان را بر آن داشته است تا به سمت تولید داده‌های مدلسازی شده با استفاده از ایستگاه‌های زمینی و بعضاً داده‌های ماهواره‌ای سوق یابند. منابع داده‌های تولید شده توسط سازمان‌های مختلف جهانی متعدد هستند که از آن جمله می‌توان به داده‌های GPM[1]، PERSIANN[2]، TRMM[3]، CMORPH[4] و GPCC[5] اشاره کرد. داده‌های آنالیز مجدد[6] Era5 نیز محصول جدید از ECMWF[7] می‌باشد که در ادامه داده‌های Era-Interim و به منظور بهبود و جایگزینی آن ارائه شده است. هر کدام از مدل‌های مذکور دارای دقت متفاوتی بوده و بنابراین در مناطق مختلف، نتایج مختلفی را ارائه می‌کنند. همین امر سبب شده که نتوان داده‌های یک منبع را برای تمام مناطق جهان توصیه نمود. از طرفی از آنجا که خروجی‌های مدل‌ها می‌توانند در سایر مدل‌های هیدرولوژیکی و کشاورزی مورد استفاده قرار گیرند، انجام پژوهش‌های متعدد در مناطق مختلف می‌تواند از ورود خطای مضاعف به مدل‌ها جلوگیری نماید. در همین راستا صادقی و همکاران (۱۳۹۷) داده‌های بارش روزانه سنجنده GPM را طی ۳ سال در ۳۱ ایستگاه سینوپتیک مورد بررسی قرار دادند. طبق نتایج به دست آمده توسط ایشان، داده‌های روزانه سنجنده GPM در غرب میانه ایران از دقت مناسبی برخوردار نیستند به طوری که ضریب تبیین بیش از ۷۷ درصد ایستگاه‌ها کمتر از ۰/۵ بود. این در حالی است که در مقیاس‌های زمانی ماهانه و سالانه دقت خروجی‌ها افزایش یافته بود. در تحقیق دیگری، مقایسه‌ای که بین داده‌های بارش TRMM و GPCC توسط میری و همکاران (۱۳۹۵) انجام شد. نتایج نشان داد که داده‌های این دو منبع در مقیاس ماهانه نسبت به روزانه عملکرد بهتری دارند. سیابی و همکاران (۱۳۹۶) عملکرد داده‌های TRMM را در فصول گرم بهتر از فصول سرد بیان کردند. یائو و همکاران (۲۰۲۰) نیز نتایج مشابهی را در این زمینه ارائه کردند. پژوهش‌هایی که در زمینه Era5 نیز انجام شده نشان می‌دهد این داده‌ها علی رغم بهبود، در همه مناطق از دقت یکسانی برخوردار نیستند (نوگیرا، ۲۰۲۰ و جیانگ و همکاران ، ۲۰۲۱). داده‌های Era5 با استفاده از منابع داده‌ای بیشتر و روش‌هایی طراحی و تولید شده اندکه منجر به بهبود قدرت تفکیک مکانی (۳۱ کیلومتر) و قدرت تفکیک زمانی (ساعتی) نسبت به نسخه قبلی خود ERA-Interim گردیده است (تارک و همکاران، ۲۰۲۰). در پژوهشی که توسط گلکسینر و همکاران (۲۰۲۰) در آفریقا انجام شد، ایشان بیان کردند که علی‌رغم بهبود خروجی‌های مدل Era5 نسبت Era-Interim، تولید داده‌هایی با دقت بیشتر مورد نیاز است. شین و همکاران (۲۰۲۱) نیز با استفاده از ۳۰۰۰ ایستگاه باران‌سنجی در چین، عملکرد داده‌های Era5 را طی سال ۲۰۱۸ مورد بررسی قرار دادند. آنها بیان کردند که بهترین عملکرد محصولات Era5 در دشت‌های ساحلی جنوبی در طول فصل مرطوب و در مناطق کوهستانی شمالی در طول فصل خشک یافت می‌شود. در ادامه‌ی پژوهش‌های انجام شده در زمینه‌ی بررسی دقت داده‌های Era5 ، لاورس و همکاران (۲۰۲۲)، دقت داده‌های روزانه مدل مذکور را طی سال‌های ۲۰۰۱ الی ۲۰۲۰ با استفاده از ۵۶۳۷ ایستگاه زمینی در کل مناطق جهان مورد بررسی قرار دادند. نتایج این تحقیق نشان داد که مدل Era5 بارش‌های فصلی را بهتر از سایر بازه‌های زمانی برآورد می‌کند و این خطا در فصل زمستان در حداقل خود می‌باشد در حالی که بیشترین خطا در مناطق استوایی مشاهده شده است. عزیزی و همکاران (۱۳۹۹) نیز به بررسی عملکرد داده‌های Era5 در اردبیل در دو مقیاس زمانی روزانه و ماهانه پرداختند. ایشان بیان کردند که ضریب تبیین در مقیاس روزانه بیش از ۰/۷۵ و مقدار RMSE نیز کمتر از ۳ میلی‌متر بوده و در مقیاس ماهانه نیز ضریب همبستگی بیش از ۰/۸ با مقدار RMSE کمتر از ۲۰ میلی‌متر وجود داشته است. ایشان استفاده از داده‌های مذکور را در هر دو بازه روزانه و ماهانه پس از اعمال تصحیحات اریبی، برای منطقه مورد مطالعه پیشنهاد کردند.

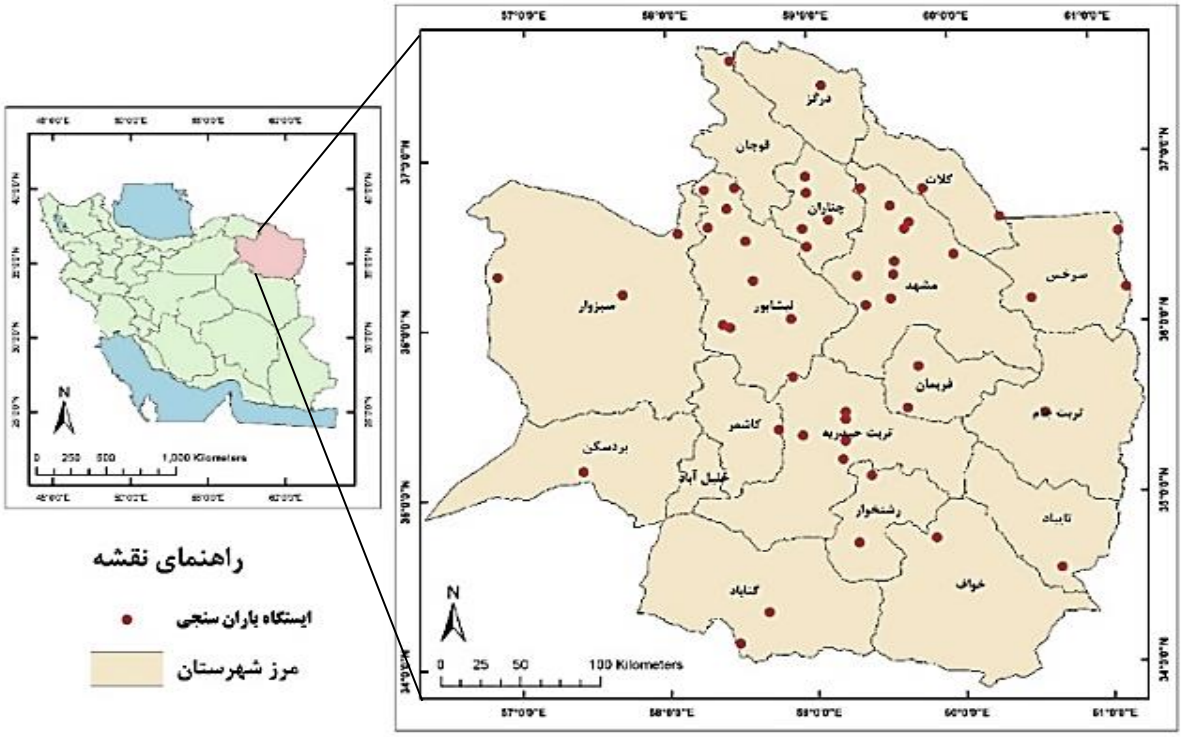
1- Global Precipitation Measurement
2- Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
3- Tropical Rainfall Measuring Mission
4- Climate Prediction Center (CPC) Morphing Technique
5- Global Precipitation Climatology Center
6- Reanalysis
7- European Centre for Medium-Range Weather Forecasts

مطالعات صورت گرفته نشان می‌دهد که مدلEra5 از دقت قابل قبولی نسبت به مدل‌های پیشین برخوردار است؛ با این وجود، دقت آن باید در نقاط مختلف با اقلیم‌های متفاوت مورد ارزیابی قرار گیرد. در این راستا، بکارگیری تعداد هر چه بیشتر از داده‌های مشاهداتی ایستگاهی می‌تواند این مدل را با دقت بیشتر مورد ارزیابی قرار دهد. با توجه به قرارگیری استان خراسان رضوی در اقلیم خشک و نیمه خشک و اهمیت این استان از لحاظ تولید محصولات کشاورزی، تخمین دقیق وقوع رخداد بارش و مقدار آن می‌تواند برای برنامه‌ریزی‌های تخصیص آب و انتخاب نوع کشت مورد استفاده قرار گیرد. از این‌رو در تحقیق حاضر، خروجی‌های مدل Era5 در مقیاس‌های زمانی روزانه، ماهانه و فصلی در استان خراسان رضوی با استفاده از داده‌های بارش ۵۱ ایستگاه باران‌سنجی در دوره آماری سال‌های ۱۳۷۶ الی ۱۳۹۶ مورد ارزیابی کمی و کیفی قرار گرفته است.

مواد و روش‌ها

منطقه مورد مطالعه

استان خراسان رضوی با مساحتی بالغ بر ۱۲۷ هزار کیلومترمربع بین عرض جغرافیایی ۳۳ درجه و ۵۲ دقیقه تا ۳۷ درجه و ۴۲ دقیقه شمالی و طول جغرافیایی ۵۶ درجه و ۱۹ دقیقه تا ۶۱ درجه و ۱۶ دقیقه شرقی قرار گرفته است (شکل ۱). میانگین مجموع بارش سالانه استان معادل ۲۵۱/۹ میلی‌متر است. میانگین دمای استان نیز برابر با ۱۵/۹ درجه سانتی‌گراد است (سازمان هواشناسی خراسان رضوی، ۱۴۰۱).



شکل ۱- منطقه مورد مطالعه و ایستگاه‌های باران‌سنجی (منبع اطلاعات ایستگاه‌ها: شرکت مدیریت منابع آب ایران)

داده‌های مشاهداتی زمینی

در مطالعه حاضر به منظور بررسی و ارزیابی دقت داده‌های مدل Era5، از داده‌های ایستگاه‌های باران‌سنجی شرکت مدیریت منابع آب ایران استفاده شده است. ایستگاه‌های مذکور نسبت به سایر انواع ایستگاه، دارای پراکنش و طول دوره آماری مناسبی می‌باشند. از طرفی این ایستگاه‌ها محدود به مراکز اطراف فرودگاه‌ها نبوده و در مناطق وسیع‌تری نسبت به سایر انواع ایستگاه‌ها را پوشش می‌دهند. با توجه به بازه زمانی مشترک داده‌ها در ایستگاه‌های زمینی، در پژوهش حاضر برای دوره آماری ۱۳۷۶ الی ۱۳۹۶ صورت گرفته است. موقعیت ایستگاه‌های مورد استفاده در شکل ۱ ارائه شده است.

داده‌های بارندگی مدل Era5

داده‌های Era5 نسل پنجم از داده‌های تحلیل مجدد ECMWF برای آب و هوا در چند دهه گذشته است. این داده‌ها تخمین ساعتی از متغیرهای مهم جوی در سطح زمین و سایر سطوح فشاری ارائه می‌کند. داده‌های Era5 با تأخیر ۵ روزه به روزرسانی می‌شوند و به صورت شبکه‌بندی شده با دقت ۰/۲۵ درجه (معادل ۳۱ کیلومتر) در دسترس قرار دارند. در پژوهش حاضر از داده¬های ساعتی با فرمت NetCDF طی بازه زمانی ۱۹۹۷ الی ۲۰۱۸ استفاده شد که از سایت cds.climate.copernicus.eu قابل دریافت می‌باشند.

معیارهای ارزیابی نتایج

به منظور ارزیابی داده‌های مدل Era5 ابتدا با استفاده از برنامه‌نویسی پایتون و روش نزدیکترین همسایگی (هیو و همکاران، ۲۰۲۱) داده‌های NetCDF پردازش و مقادیر بارندگی مربوط به هر ایستگاه استخراج گردید. پس از تبدیل داده‌های ساعتی به روزانه، مقادیر مدلسازی شده در مقایسه با مقادیر مشاهداتی با استفاده از دو گروه شاخص کمی و کیفی مورد ارزیابی قرار گرفتند. گروه اول شاخص‌ها که برای ارزیابی کمی، مقادیر بارش مشاهداتی و مدلسازی شده مورد استفاده قرار گرفتند شامل شاخص میانگین مجذور خطا (RMSE)، ضریب تبیین (R^2) (جیانگ و همکاران ، ۲۰۲۱) و معیار نش-ساتکلیف (NSE[1]) مطابق با معادله (۱) استفاده شده است.

$$NSE = \sum_{i=1}^{n} \frac{(Y - \widehat{Y})^2}{(Y - \bar{Y})^2} \qquad (۱)$$

در معادله فوق، مقدار بارش مشاهداتی ایستگاه، $\bar{Y}$ میانگین بارش مشاهداتی ایستگاه و مقدار بارش مدلسازی شده توسط Era5 است. شاخص نش- ساتکلیف از منفی بی‌نهایت تا ۱ برای بدترین تا بهترین شرایط تغییر می‌کند. مقدار ۱ برای این شاخص به معنای تطابق کامل بین داده‌های مشاهداتی و مدلسازی شده است (مدرسی و همکاران، ۲۰۱۸). مقادیر بزرگتر از صفر این شاخص نشان می‌دهد که مقادیر تخمین زده شده توسط مدل دارای دقتی بیشتر از میانگین مقادیر مشاهداتی است در حالی‌که مقادیر منفی نشان می‌دهد که دقت خروجی مدل کمتر از میانگین مقادیر مشاهداتی است و مقدار صفر این شاخص نشان میدهد که پیش بینی های مدل در حد میانگین داده‌های مشاهداتی است. شاخص RMSE نیز از صفر تا مثبت بی‌نهایت به ترتیب برای بهترین تا بدترین حالت تغییر می‌کند. هر چه مقدار این شاخص به صفر نزدیک‌تر باشد، دقت مدل‌سازی بهتر است. شاخص ضریب تبیین از صفر تا ۱+ تغییر می‌نماید و نشان دهنده درصدی از داده‌های مشاهداتی و شبیه‌سازی شده است که به یکدیگر ارتباط دارند. بر اساس طبقه‌بندی ارائه شده توسط موریاسی و همکاران (۲۰۰۷) کارایی مدل‌ها را بر اساس این سه شاخص می‌توان به مطابق جدول ۱ طبقه‌بندی کرد.

گروه دوم شاخص‌ها برای ارزیابی کیفی خروجی‌های مدل Era5 مورد استفاده قرار گرفت. با استفاده از این شاخص‌ها، وقوع یا عدم وقوع بارش (بدون در نظر گرفتن مقدار بارش ثبت شده) در هر دو گروه داده مورد بررسی قرار گرفت. شاخص‌های گروه دوم شامل احتمال آشکارسازی (POD)[2]، نرخ هشدار اشتباه (FAR)[3] و شاخص موفقیت بحرانی (CSI)[4] می‌باشد. در این گروه از شاخص‌ها وقوع یا عدم وقوع بارش ثبتی توسط داده‌های مشاهداتی و مدلسازی به صورت بله یا خیر ثبت می‌گردد. با استفاده از این شاخص‌ها می‌توان به توانایی تشخیص وقوع یا عدم وقوع بارش توسط مدل دست یافت. شاخص POD نشان‌دهنده میزان دقت مدل در تشخیص وقایع بارندگی در ایستگاه است و مقدار مطلوب این شاخص ۱ میباشد. شاخص FAR بیانگر وقایع بارندگی است که در ایستگاه رخ داده و ثبت شده اما مدل قادر به تشخیص آن نبوده است و مقدار مطلوب آن، صفر است. شاخص CSI نشان‌دهنده نسبتی از وقوع بارش است که توسط مدل به درستی شناسایی شده است. مقدار این شاخص نیز هر چه به ۱ نزدیک‌تر باشد مطلوبتر است. معادلات (۲) تا (۴) بیان‌کننده معادلات سه شاخص مذکور می‌باشند.

$$POD = \frac{RR}{RR + RN} \qquad (۲)$$

$$FAR = \frac{NR}{RR + NR} \qquad (۳)$$

$$CSI = \frac{RR}{RR + RN + NR} \qquad (۴)$$

1- Nash- Sutcliffe Efficiency

2- Probability of Detection
3- False Alarm Ratio
4- Critical Success Index

در معادلات فوق، RR بیانگر تعداد حالاتی است که هم مدل و هم ایستگاه وقوع بارش را نشان می‌دهند. NR بیانگر تعداد حالاتی است که در ایستگاه بارش رخ نداده اما مدل داده بارندگی ثبت کرده است. RN تعداد حالاتی است که ایستگاه بارش ثبت کرده اما مدل مقدار بارندگی را صفر نشان می‌دهد و NN تعداد حالاتی است که هم مدل و هم ایستگاه بارشی ثبت نکرده‌اند (عبدالهی و همکاران، ۱۳۹۶؛ عرفانیان و همکاران، ۱۳۹۵).

جدول ۱- طبقه بندی نرخ عملکرد مدل برای شاخص‌های NSE، RMSE و R^2 (موریاسی و همکاران، ۲۰۰۷)

رتبه بندی عملکرد	محدودیت کارایی مدل		
	NSE	RMSE	R^2
خیلی خوب	۰/۷۵< NSE ≤ ۱	۰≤RMSE≤۰/۵ SD*	۰/۸۶< $R^2$۱۰۰ <
خوب	۰/۶۵<NSE ≤ ۰/۷۵	۰/۵ SD ≤RMSE≤۰/۶ SD	۰/۷۷< $R^2$۰/۸۵ <
قابل قبول	۰/۵<NSE ≤ 0.65	۰/۶ SD ≤RMSE≤۰/۷ SD	۰/۶۵< $R^2$۰/۷۶ <
غیر قابل قبول	NSE ≤ 0.5	۰/۷ SD <RMSE	< ۰/۶۴

*SD انحراف معیار

نتایج و بحث

ارزیابی داده‌های روزانه Era5 در مقایسه با مشاهدات ایستگاهی

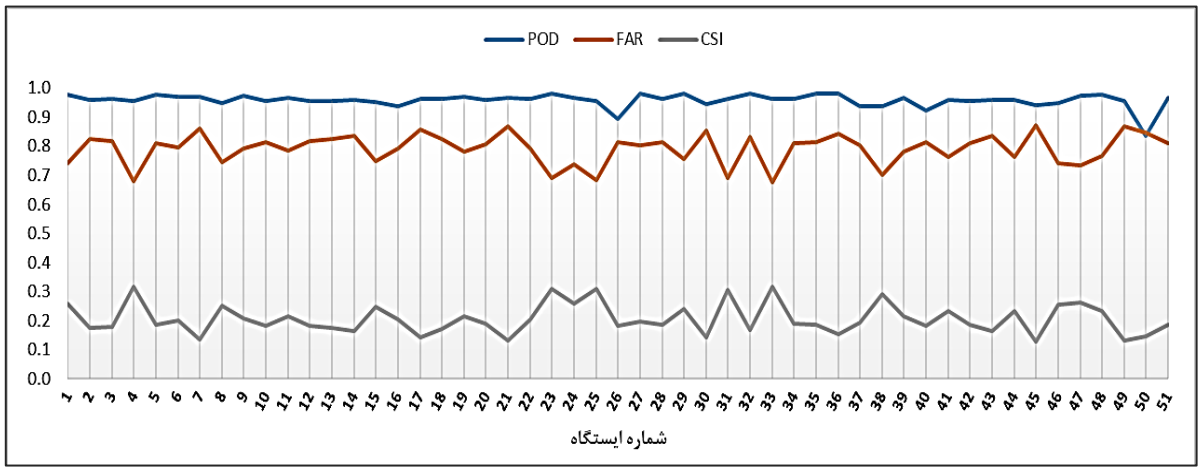
با توجه به شاخص‌های کیفی داده‌های بارشی Era5 در مقیاس روزانه، نتایج نشان داد که داده‌های مذکور از توانایی بالایی در شناسایی روزهایی که در ایستگاه بارش به وقوع پیوسته است برخوردارند. همانطور که در شکل ۲ نیز مشخص است مقادیر شاخص POD تقریباً در تمام ایستگاه‌ها به مقدار مطلوب ۱ نزدیک می‌باشد. این در حالی است که شاخص FAR که نشان‌دهنده وقایع بارشی است که مدل به اشتباه تشخیص داده است و در واقع ایستگاه بارشی را ثبت نکرده است، به مقدار نامطلوب ۱ نزدیک است و این امر نشان از خطای بالای مدل در شناسایی روزهای فاقد بارندگی به عنوان روز دارای بارندگی می‌باشد. شاخص FAR در شرایط مطلوب باید صفر باشد یعنی هیچ روز فاقد بارشی در ایستگاه، توسط مدل دارای بارندگی ثبت نشده باشد. اما مقادیر شکل ۲ خلاف این امر را نشان می‌دهد. این خطا خود را در شاخص CSI که مؤید کسری از کل بارش‌هایی است که مدل به درستی تشخیص داده است به خوبی نمایان می‌سازد. مقدار شاخص CSI هر چه به ۱ نزدیک‌تر باشد نشان می‌دهد که مدل در برآورد روزهای دارای بارش و فاقد بارش موفق عمل کرده است. شایان ذکر است که شاخص‌های کیفی فقط برای مقیاس روزانه قابل محاسبه و ارزیابی هستند. نتایج ارزیابی کمّی مقادیر بارش‌های برآوردی مدل Era5 در مقیاس روزانه هم نشان می‌دهد مقادیر ضریب تبیین در بازه ۰/۱۲ در ایستگاه هندل آباد تا ۰/۴۸در ایستگاه سرخس در نوسان است. این نتایج با نتایج عزیزی و همکاران (۱۳۹۹) مغایرت دارد. ایشان بیان کردند که داده‌های Era5 در اردبیل در مقیاس روزانه از عملکرد بالایی برخوردارند و در تمامی ایستگاه‌های آن منطقه مقداری بیش از ۰/۵۶ دارد. این امر می‌تواند به دلیل تفاوت در شرایط محیطی و همچنین تفاوت سیستم‌های بارشی حاکم بر یک منطقه با در مقایسه با منطقه‌ای دیگر تفاوت دارد باشد. پایین بودن دقت داده‌های روزانه مدلسازی و یا ماهواره‌ای در مطالعات زیادی در ایران و دنیا جهان بیان شده است (عبدالهی و همکاران، ۱۳۹۶صادقی و همکاران، ۱۳۹۷؛ علی بخشی و همکاران، ۱۳۹۷؛ ژیانگ و همکاران، ۲۰۲۱). لازم به ذکر است که دقت نتایج در مقیاس زمانی روزانه نمی‌تواند سبب نتیجه‌گیری از چگونگی عملکرد این مدل‌ها در سایر مقیاس‌های زمانی باشد. همانطور که پیش‌تر نیز ذکر شد، به منظور ارزیابی شبیه‌سازی مدل از شاخص‌های RMSE و NSE استفاده شده است. مقادیر شاخص‌های RMSE و NSE نشان می‌دهد علی رغم ناتوانی مدل در تشخیص روزهای فاقد بارندگی، مقدار خطا در تعیین میزان بارندگی اگرچه ایده‌آل نیست اما در دامنه قابل قبولی قرار دارد (شکل ۳).

به این معنا که در روزهایی که توسط مدل دارای بارش ثبت شده است نیز به طور کلی مقدار اندکی از بارش به

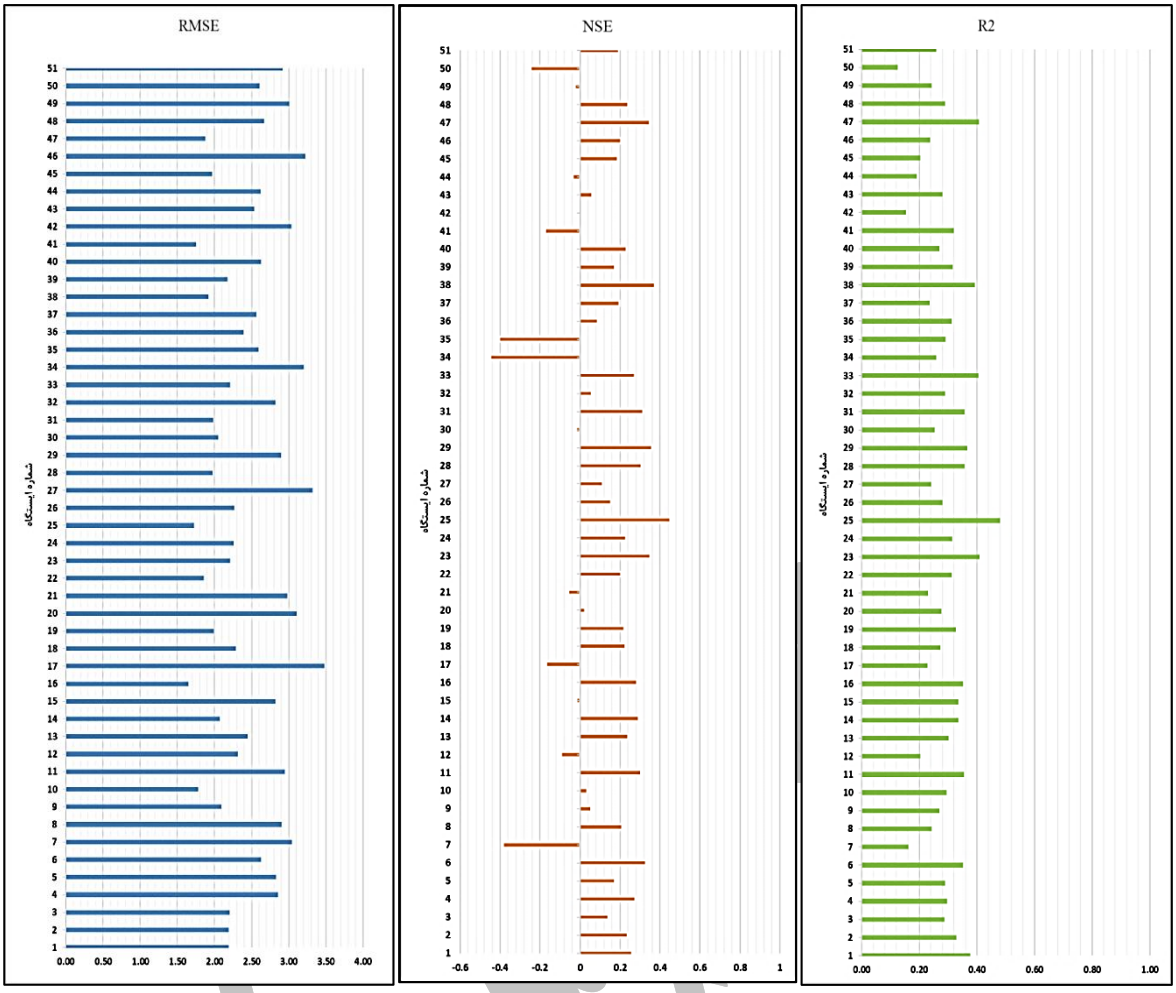
اشتباه برآورد شده است. مقدار شاخص RMSE از ۱/۶۵ میلی‌متر در روز در ایستگاه درونه تا ۳/۴۸ میلی‌متر در روز در ایستگاه دهنه اخلمد متغیر است. معیار کارایی NSE نیز از ۰/۴۵ در ایستگاه سرخس تا ۰/۴۵- در ایستگاه فریزی متغیر است. اگرچه مقادیر شاخص نش بر اساس طبقه‌بندی موریاسی و همکاران (۲۰۰۷) در بخش غیرقابل قبول قرار دارد، ولی بر اساس ضریب تبیین، در مقیاس ماهانه و فصلی به ترتیب در طبقه قابل قبول و خوب قرار دارد. به منظور بررسی بهتر داده‌ها اقدام به رسم نمودار شاخص‌های کمی (شکل ۳) و نمودار جعبه‌ای برای ۱۳ ایستگاه به صورت نمونه (شکل ۴) در منطقه مورد مطالعه در سه مقیاس زمانی روزانه، ماهانه و فصلی شد.از آنجا که میانگین داده‌های بارندگی روزانه حدود ۰/۵ میلی‌متر می‌باشد و بارش‌های حدی زیادی طی سال در مقیاس روزانه رخ داده است نمودار جعبه‌ای اطلاعات مناسبی ارائه نمی‌دهد. ولی به‌طور کلی مدل Era5 مقادیر بارش‌های حدی را در مقیاس روزانه با اختلاف نسبتاً زیادی کمتر از مقادیر مشاهداتی ثبت نموده است.

جدول ۲- نام و شماره ایستگاه های باران سنجی مورد بررسی در این پژوهش

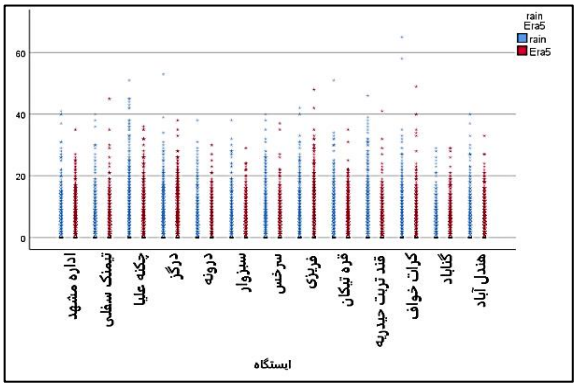
نام ایستگاه	شماره ایستگاه	نام ایستگاه	شماره ایستگاه	نام ایستگاه	شماره ایستگاه	نام ایستگاه	شماره ایستگاه
اداره مشهد	۱	حسین آباد جنگل	۱۴	شمخال	۲۷	گردونه فردوس	۴۰
برزنون	۲	درگز	۱۵	شیخ ابوالقاسم	۲۸	گناباد	۴۱
اندرکه	۳	درونه	۱۶	غنچی	۲۹	گوش بالا	۴۲
بلغور	۴	دهانه اخلمد	۱۷	فدک	۳۰	ماروسک	۴۳
بهمن علیا	۵	دهنه شور	۱۸	فدیشه	۳۱	مزدوران	۴۴
تروسک	۶	روح آباد	۱۹	فدیهه	۳۲	مزینان	۴۵
تلبخش	۷	جاغرق	۲۰	فرهادگرد	۳۳	مغان	۴۶
تلغور	۸	ساق بیک	۲۱	فریزی	۳۴	ملک آباد	۴۷
تیمنک سفلی	۹	سبزوار	۲۲	قدیرآباد	۳۵	ناری	۴۸
جنت آباد	۱۰	سد طرق	۲۳	قره تیکان	۳۶	نامقه	۴۹
چکنه علیا	۱۱	سد کارده	۲۴	قند تربت حیدریه	۳۷	هندل آباد	۵۰
چناران	۱۲	سرخس	۲۵	قند نیشابور	۳۸	ینگجه	۵۱
چهچه	۱۳	سنگر سرخس	۲۶	کرات خواف	۳۹		



شکل ۲- نمودار شاخص‌های کیفی (POD، FAR و CSI) در مقیاس روزانه در ایستگاه‌های باران‌سنجی منطقه مورد مطالعه



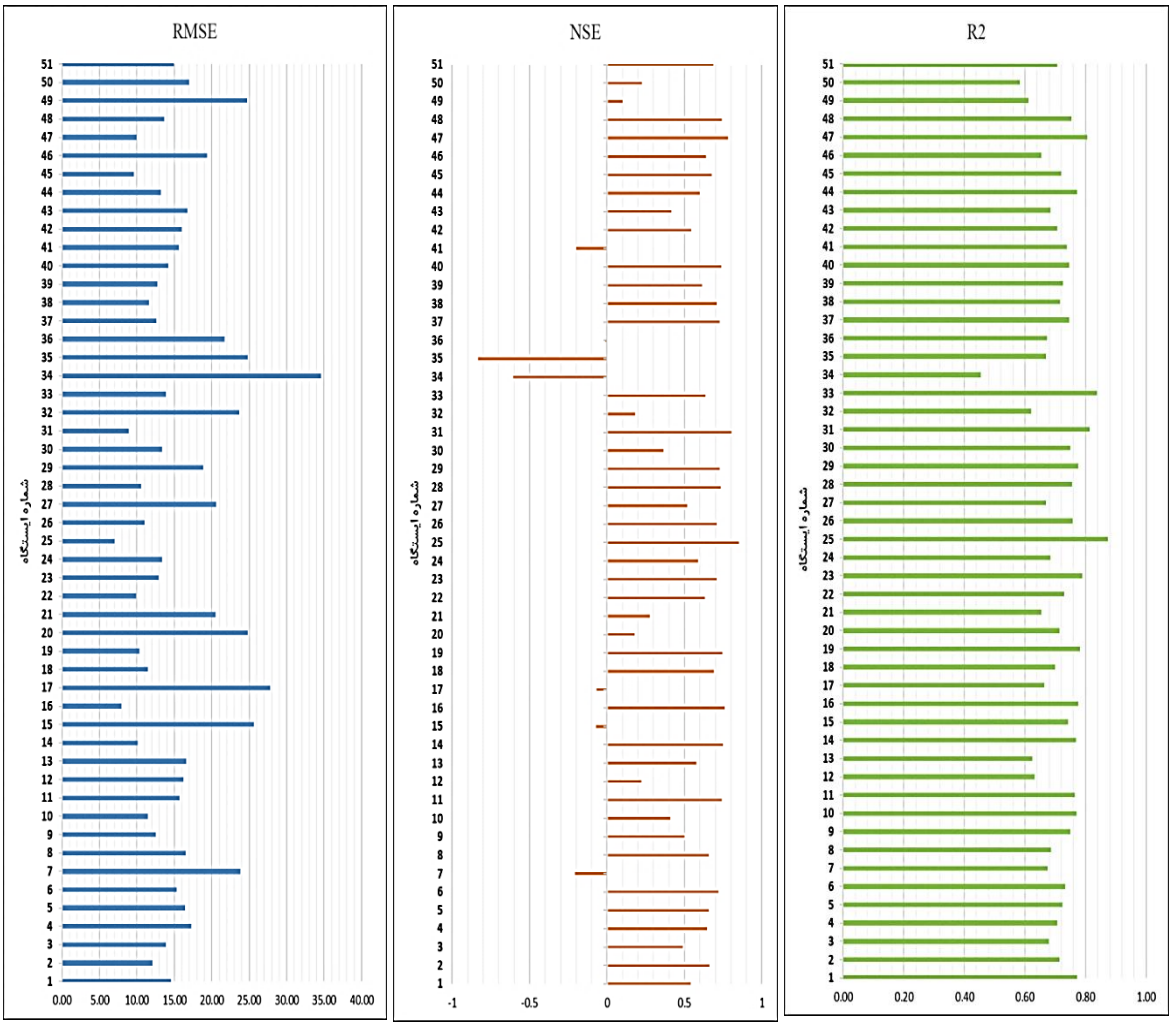
شکل ۳- نمودار شاخص‌های کمی (RMSE، NSE و R^2) در مقیاس روزانه در ایستگاه‌های باران‌سنجی منطقه مورد مطالعه



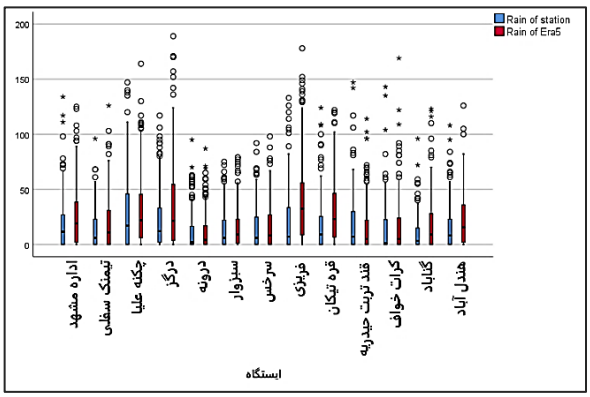
شکل ۴- نمودار جعبه‌ای داده‌های مشاهداتی و مدلسازی ۱۳ ایستگاه نمونه در مقیاس زمانی روزانه

ارزیابی داده‌های ماهانه Era5 در مقایسه با مشاهدات ایستگاهی

نتایج ارزیابی آماری بارش‌های ماهانه ایستگاه‌ها در مقایسه با مقادیر برآوردی مدل نشان می‌دهد که برخلاف ضریب تبیین پایین داده‌های مشاهداتی و مدل‌سازی شده در مقیاس روزانه، در مقیاس ماهانه ضریب تبیین افزایش قابل توجهی داشته است به طوری که دامنه نوسان آن از ۰/۴۵ در ایستگاه فریزی تا ۰/۸۶ در ایستگاه سرخس متغیر است. شاخص RMSE بیشترین مقدار را در ایستگاه سرخس با مقداری برابر با ۲۲/۵۷ و کمترین مقدار را در ایستگاه فریزی با مقداری معادل ۳/۵۷ به خود اختصاص داده است و شاخص NSE با بیشترین مقدار ۰/۸۶ در سرخس تا کمترین مقدار ۰/۸۴- در قدیرآباد مشاهده شد (شکل ۵). نمودار جعبه‌ای ایستگاه‌های نمونه در مقیاس ماهانه نشان می‌دهد داده‌های مشاهداتی و مدلسازی در اکثر ایستگاه‌ها از تطابق خوبی برخوردارند. در ایستگاه‌هایی مانند سبزوار، چکنه علیا و سرخس مقادیر میانه، چارک¬ها و داده¬های پرت با هم مطابقت داشته و نزدیک به هم است. این در حالی است که در ایستگاه‌های فریزی، قره تیکان و درگز اختلاف میان داده‌های ثبت شده در ایستگاه و داده‌های مدل‌سازی شده زیاد بوده و مدل با اختلاف نسبتاً زیادی در همه آماره‌ها عموماً دارای بیش برآورد است (شکل ۶).



شکل ۵- نمودار شاخص‌های کمی (RMSE، NSE و R^2) در مقیاس ماهانه در ایستگاه‌های باران‌سنجی منطقه مورد مطالعه



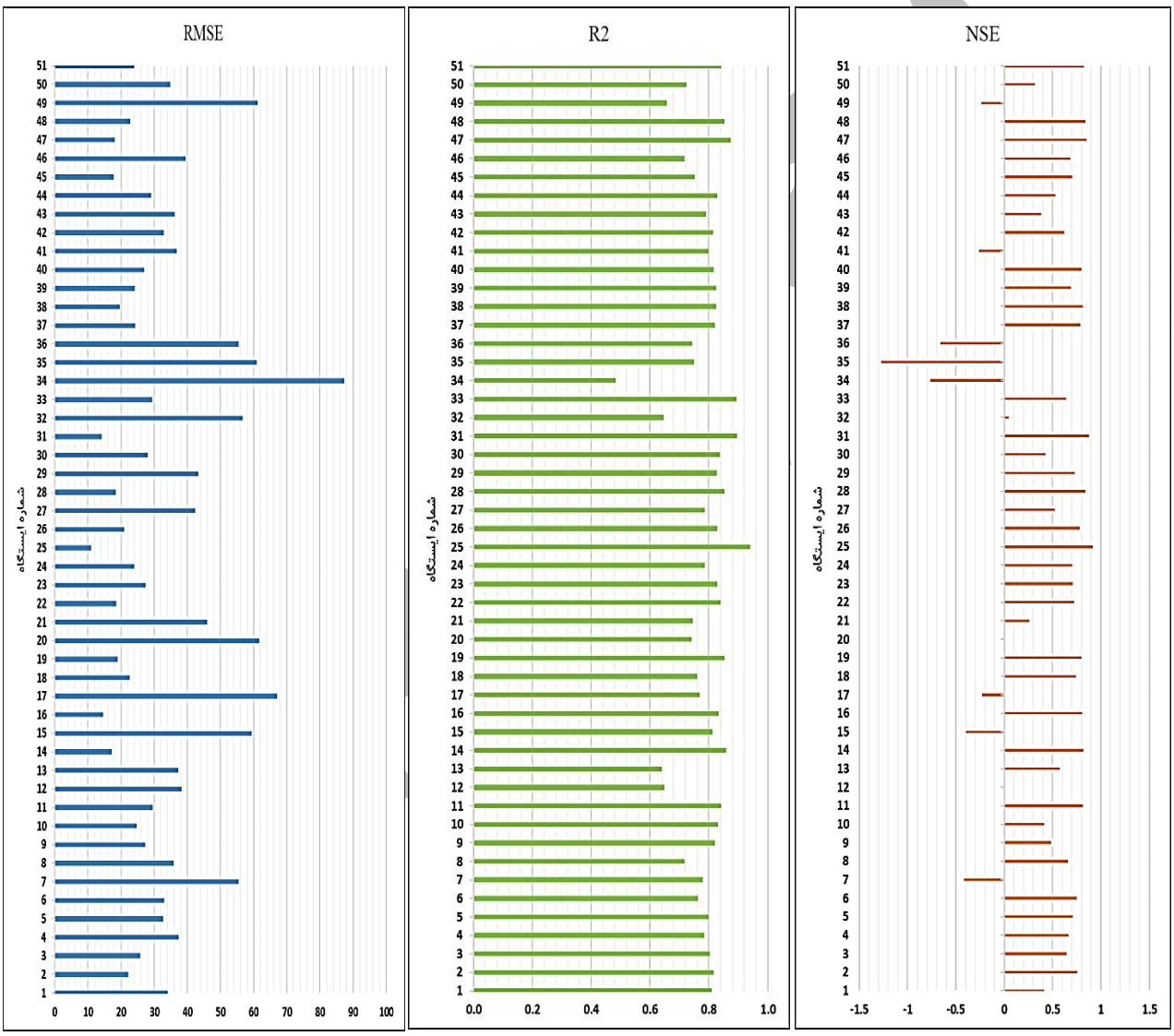
شکل ۶- نمودار جعبه‌ای داده‌های مشاهداتی و مدلسازی ۱۳ ایستگاه نمونه در مقیاس زمانی ماهانه

نتایج مناسب خروجی ماهانه مدل Era5 نشان می‌دهد که می‌توان از این داده‌ها در مدل‌هایی که نیازمند مقادیر بارش ماهانه هستند، با اطمینان خوبی استفاده کرد. البته لازم به ذکر است پیش از استفاده از داده‌های مذکور بهتر است با روش‌هایی مانند رگرسیون خطی، حذف بایاس انجام شود تا بتوان با دقت بیشتری از نتایج خروجی استفاده نمود. از آنجا که گستره ارتفاعی استان و رژیم‌های بارشی مؤثر بر آن در زمان‌های مختلف، متفاوت است، رفتارهای متفاوتی از مدل در ایستگاه‌های مختلف مشاهده می‌شود. این نتایج نشان می‌دهد مدل مذکور علی‌رغم دقت کلی مطلوب، نیازمند به ادامه بهبود در روابط عددی و مدلسازی دینامیکی است تا بتواند در آینده نتایج دقیق‌تری را متناسب با هر منطقه ارائه نماید.

ارزیابی داده‌های فصلی Era5 در مقایسه با مشاهدات ایستگاهی

مطابق با شکل ۷، با افزایش مقیاس زمانی، دقت برآوردهای مدل Era5 نیز افزایش می‌یابد. مقادیر ضریب تبیین ایستگاهی از ۰/۴۸ در ایستگاه فریزی تا ۰/۹۴ در ایستگاه سرخس تغییر داشته است. مقادیر شاخص RMSE به ترتیب ۸۷/۳۶ در ایستگاه فریزی و ۱۱/۰۳ در

ایستگاه سرخس متغیر بوده است. شاخص NSE نیز دارای مقادیر ۰/۹۳ در سرخس و ۱/۲۸- در قدیر آباد بوده است. مشابه با نتایج ارائه شده در نمودار جعبه‌ای داده‌های ماهانه، نمودار جعبه‌ای داده‌های فصلی (شکل ۸) در سرخس و سبزوار دارای مقادیر میانه، کرانه‌های بالا و پایین و چارک‌های نزدیک به هم است در حالیکه در ایستگاه هایی همچون فریزی و درگز اختلاف زیادی با هم دارند. در اکثر ایستگاه‌ها نیز داده‌های مدل دچار بیش برآورد بوده است. این نتایج نشان می‌دهد روش‌های دینامیکی و عددی مورد استفاده در مدل Era5 توانایی دارد تا تخمین‌هایی نزدیک به واقعیت و با کمترین خطا ارائه دهد و قابلیت بهبود در ارائه نتایج بهینه‌تر در سایر مناطق را نیز دارا است. عملکرد خوب و مناسب مدل Era5 در مقیاس فصلی در مقالات بسیاری در نواحی مختلف گزارش شده است (جیانگ و همکاران، ۲۰۲۱؛ گلکسینر و همکاران، ۲۰۲۲؛ کوروست و همکاران، ۲۰۲۰). مقادیر ضریب تبیین مقیاس فصلی (شکل ۶)، نشان دهنده برتری عملکرد Era5 در مقیاس‌های زمانی بلندمدت است.

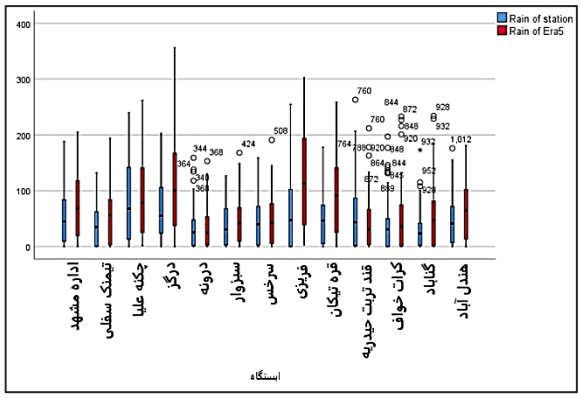


شکل ۷- نمودار شاخص‌های کمی (RMSE، NSE و R^2) در مقیاس فصلی در ایستگاه‌های باران‌سنجی منطقه مورد مطالعه

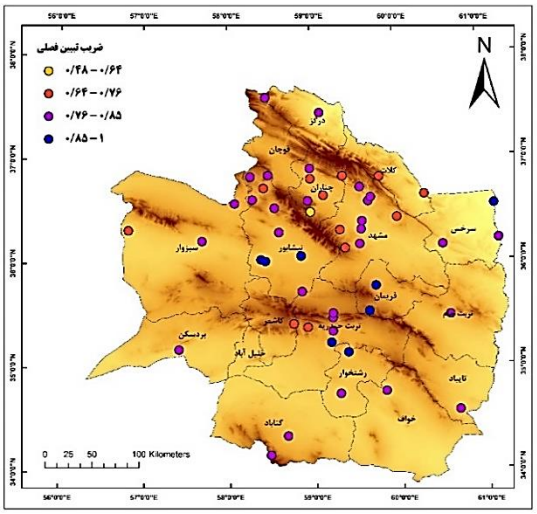
لاورس و همکاران (۲۰۲۲) نیز در یافته‌های خود اشاره کردند که این مدل در نواحی فراگرمسیری بهتر از نواحی گرمسیری استوایی عمل می‌کند و در این مناطق می‌توان با اطمینان بیشتری از این داده‌ها استفاده کرد. به منظور بررسی مکانی میزان صحت و دقت داده‌های مدل Era5 نقشه‌های توزیع مکانی ضریب تبیین، ضریب نش و همچنین شاخص RMSE در مقیاس فصلی در شکل‌های ۹ تا ۱۱ ارائه شده است. لازم به ذکر است که طبقه‌بندی

داده‌ها براساس جدول ۱ صورت گرفته است.

همانطور که در شکل ها مشاهده می‌شود اکثر ایستگاه‌ها از نظر معیارهای ضریب نش و شاخص RMSE و مطابق جدول ۱ در محدوده قابل قبول قرار دارند و در تعدادی از ایستگاه‌ها نیز مدل به خوبی عمل نکرده است. بنابراین نمی‌توان از داده‌های مدل Era5 برای تمامی مناطق یک محدوده استفاده کرد و بررسی صحت و دقت داده‌ها در هر منطقه مورد نیاز است.



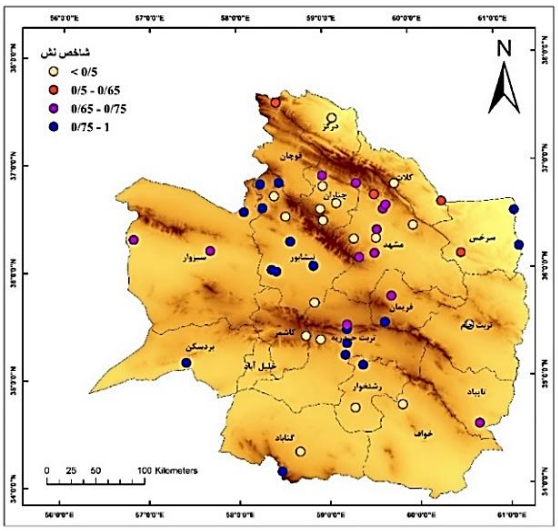
شکل ۸- نمودار جعبه‌ای داده‌های مشاهداتی و مدلسازی ۱۳ ایستگاه نمونه در مقیاس زمانی فصلی



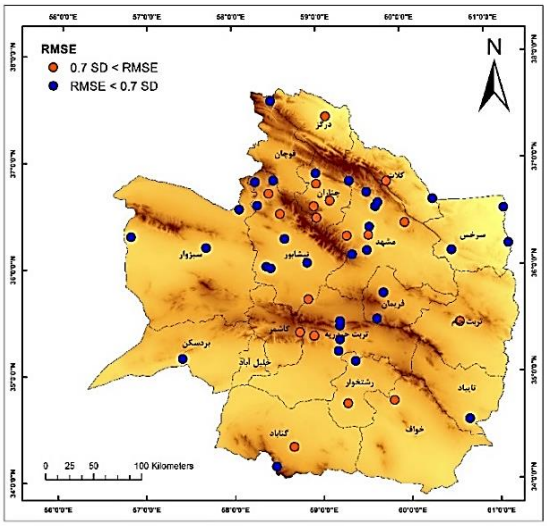
شکل ۹- نقشه توزیع ضریب تبیین در مقیاس فصلی در ایستگاه‌های باران‌سنجی منطقه مورد مطالعه

ضریب تبیین نیز از الگوی نحوه توزیع و پراکندگی ضریب نش و شاخص RMSE پیروی می‌کند (شکل ۱۰).

با توجه به چگونگی توزیع شاخص‌های آماری مورد استفاده، مشاهده می‌شود مدل Era5 در مناطق شرق، جنوب، جنوب غربی و جنوب شرقی استان که از اقلیم گرم‌تر و بارش‌های کمتری برخوردارند، عملکرد بهتری داشته است.



شکل ۱۰- نقشه توزیع ضریب نش در مقیاس فصلی در ایستگاه‌های باران‌سنجی منطقه مورد مطالعه



شکل ۱۱- نقشه توزیع شاخص RMSE در مقیاس فصلی در ایستگاه‌های باران‌سنجی منطقه مورد مطالعه

نتیجه‌گیری

پژوهش حاضر به منظور بررسی قابلیت استفاده داده‌های بارندگی مدل Era5 در استان خراسان رضوی انجام شد. از آنجا که پراکنش ایستگاه‌ها در همه نقاط استان از توزیع یکنواختی برخوردار نیست، استفاده از مدل‌هایی که بتوانند پوشش یکنواخت‌تری از داده‌های مهم و کلیدی از جمله بارش ارائه دهند در مباحث هیدرولوژیکی و کشاورزی بسیار ضروری است.

مقایسه خروجی‌های مدل Era5 در مقیاس روزانه در مقایسه با بارش‌های مشاهداتی در ۵۱ ایستگاه استان نشان داد که خطاهای این مدل برای استفاده در بازه روزانه در منطقه مورد مطالعه مناسب نیست و اگرچه مدل، روزهای بارندگی واقعی در ایستگاه‌ها را به خوبی نشان می‌دهد اما برآوردهایی از میزان بارش در روزهایی که ایستگاه بارندگی ثبت نکرده است منجر به افزایش خطا در این مقیاس زمانی شده است. علی‌رغم عملکرد نسبتاً ضعیف این مدل در بازه روزانه، در مقیاس‌های زمانی ماهانه و فصلی دقت برآوردهای مدل به خصوص در نواحی گرمتر استان افزایش داشته است. این نتایج، نشان می‌دهد برآوردهای مدل Era5 در مقیاس‌های ماهانه و فصلی در مناطق دارای اقلیم گرمتر و بارش کمتر استان خراسان رضوی بهترین عملکرد را داشته است و در صورت رفع خطای اریبی می‌توانند به عنوان داده ورودی به مدل‌های کشاورزی و هیدرولوژیکی مورد استفاده قرار گیرند.

منابع

۱. سازمان هواشناسی خراسان رضوی. ۱۴۰۰. www.razavimet.ir

۲. سیابی ن. ثنایی نژاد ح. و قهرمان ب. ۱۳۹۶. ارزیابی داده‌های بارندگی حاصل از ماهواره TRMM، مدل MM5 و مشاهدات زمینی به صورت مکانی-زمانی در مناطق خشک و نیمه‌خشک کوهستانی. نشریه جغرافیا و مخاطرات طبیعی. ۳(۲۳): ۱۶۳-۱۷۹.

۳. صادقی ح. معصوم‌پور ج. و میری م. ۱۳۹۷. ارزیابی داده‌های بارش دور سنجی GPM در مقابل داده‌های مشاهده‌ای (مورد مطالعه: غرب میانه ایران). نشریه سنجش از دور و GIS ایران. ۱۱(۲): ۱۱۵-۱۲۴.

۴. عبدالهی ب. حسینی موغاری م. ابراهیمی ک. ۱۳۹۶. ارزیابی داده‌های ماهواره‌ایTRMM RTV42B3 و CMORPH به منظور تخمین بارش در حوضه گرگانرود. نشریه علوم و مهندسی آبخیزداری ایران. ۱۱(۳۶): ۵۵-۶۸.

۵. عرفانیان م. کاظم‌پور س. و حیدری ح. ۱۳۹۵. واسنجی داده های باران سری B42۳ و B43۳ ماهواره TRMM در زون‌های اقلیمی ایران. پژوهش‌های جغرافیای طبیعی. ۴۸(۲)-۹۶: ۲۸۷-۳۰۳.

۶. عزیزی ج. رسول‌زاده ع. رحمتی ا. شایقی ا. و باختر آ. ۱۳۹۹. ارزیابی عملکرد داده‌های بازتحلیل شده Era-5 در تخمین بارش روزانه و ماهانه در استان اردبیل. نشریه تحقیقات آب و خاک ایران. ۵۱: ۲۹۳۷-۲۹۵۱.

۷. علی‌بخشی م. فرید حسینی ع. داوری ک. و علیزاده ا. ۱۳۹۷. ارزیابی داده‌های بارش زمینی، ماهواره GPM و MERRA (مطالعه موردی: حوضه آبریز کشف رود). پژوهشنامه مدیریت حوزه آبخیز. ۹(۱۸): ۱۱۱-۱۲۲.

۸. میری م. رضیئی ط. و رحیمی م. ۱۳۹۵. رزیابی و مقایسه آماری داده‌های بارش TRMM و GPCC با داده‌های مشاهده‌ای در ایران. نشریه فیزیک زمین و فضا. ۲(۳): ۶۵۷-۶۷۲.

9. Crossett C. C. Betts A. K. Dupigny-Giroux L. A. L. and Bomblies A. 2020. Evaluation of daily precipitation from the ERA5 global reanalysis against GHCN observations in the northeastern United States. Climate. 8(12): 148.
10. Gleixner S. Demissie T. and Diro G. T. 2020. Did ERA5 improve temperature and precipitation reanalysis over East Africa? Atmosphere. 11(9): 996.
11. Hu X. and Yuan W. 2021. Evaluation of ERA5 precipitation over the eastern periphery of the Tibetan plateau from the perspective of regional rainfall events. International Journal of Climatology. 41(4): 2625-2637.
12. Jiang Q. Li W. Fan Z. He X. Sun W. Chen S. Wen J. Gao J. and Wang J. 2021. Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland. Journal of hydrology. 595: 125660.
13. Lavers D. A. Simmons A. Vamborg F. and Rodwell M. J. 2022. An evaluation of ERA5 precipitation for climate monitoring. Quarterly Journal of the Royal Meteorological Society. 148(748): 3152-3165.
14. Modaresi F. Araghinejad S. and 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(1): 243-258.

15. Moriasi D. N. Arnold J. G. Van Liew M. W. Bingner R. L. Harmel R. D. and Veith T. L. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE. 50(3): 885-900.
16. Nogueira M. 2020. Inter-comparison of ERA-5, ERA-interim and GPCP rainfall over the last 40 years: Process-based analysis of systematic and random differences. Journal of Hydrology. 583: 124632.
17. Tarek M. Brissette F. P. and Arsenault R. 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrology and Earth System Sciences. 24(5): 2527-2544.
18. Xin Y. Lu N. Jiang H. Liu Y. and Yao L. 2021. Performance of ERA5 reanalysis precipitation products in the Guangdong-Hong Kong-Macao greater Bay Area, China. Journal of Hydrology. 602: 126791.
19. Yao J. Chen Y. Yu X. Zhao Y. Guan X. and Yang L. 2020. Evaluation of multiple gridded precipitation datasets for the arid region of northwestern China. Atmospheric Research. 236: 104818.

Research paper

A qualitative- quantitative evaluation of Era5 rainfall data in identifying the occurrence and amount of rainfall in Razavi Khorasan province

A. Nouri[1], S. Nouri[2], J. Omidvar[3], H. Banejad[4*] and F. Modaresi[5]

Extended Abstract

Accurate measurement of the rainfall, as one of the essential meteorological variables in hydrological and agricultural studies, has a significant impact on the optimal management of water resources, cultivation of agricultural products, urban management, and identification and classification of areas in terms of the occurrence of floods risk or the ability to extract water. The inappropriate distribution of rainfall measurement stations in different regions has caused researchers to seek to create models that can estimate the amount of rainfall in areas that lack rain gauge stations or have Insufficient of them.

In the current research, the data from Era5, the newest product of the ECMWF, has been used. Era5 data is the fifth generation of ECMWF reanalysis data for climate over the past several decades. These data provide hourly estimates of important atmospheric variables at the surface of the earth and other pressure levels. Era5 data is updated with a delay of 5 days and is available in a gridded form with an accuracy of 0.25 degrees (equivalent to 31 km). In the present study, hourly data in NetCDF format were used during the period from 1997 to 2018. The accuracy of Era5 in the Middle East and Iran has yet to be investigated much. In order to evaluate the accuracy of Era5 model data, the data of 51 rain gauge stations of Iran Water Resources Management Company in Razavi Khorasan province from 1997 to 2017 were used were used. The mentioned stations have a suitable distribution and length of statistical period compared to other types of stations. To evaluate the data of Era5 model, first using Python programming and the nearest neighbor method, NetCDF data was processed and rainfall values for each station were extracted. Then the hourly data were converted to daily. The modeled values compared to the observed values were evaluated using two groups of quantitative and qualitative indicators. The first group of indicators that were used for quantitative evaluation of observed and modeled rainfall values include root mean square error (RMSE), coefficient of explanation (R^2) and Nash-Sutcliffe Efficiency (NSE). The second group of indicators was used to evaluate the quality of Era۵ model outputs. Using these indicators, the occurrence or non-occurrence of precipitation (regardless of the amount of recorded precipitation) was investigated in both data groups. The indicators of the second group include probability of detection (POD), false alarm ratio (FAR) and critical success index (CSI). In this group of indicators, the occurrence or non-occurrence of recorded precipitation is recorded as yes or no by observational data and modeling. Using these indicators, it is possible to determine the occurrence or non-occurrence of precipitation by the model.

1- M.Sc. Graduated in Agroecology, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.
2- Ph.D. Graduated in Agrometeorology, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.
3- Ph.D. Student in Water Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.
4- Associate Professor, Water Engineering Department, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.
5- Assistant Professor, Water Engineering Department, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.
*** Corresponding author: Banejad@um.ac.ir**
Received: 2023/03/09 Accepted: 2023/06/26
https://dx.doi.org/10.22034/IWRJ.2023.14295.2506

According to the quality indicators of Era5 rainfall data on a daily scale, the results showed that the said data have a high ability to identify the days that rainfall occurred at the station. It is worth noting that qualitative indicators can only be calculated and evaluated for the daily scale. The results of the investigations show that the Era5 model recorded the maximum rainfall values on a daily scale with a relatively large difference lower than the observed values. The results of the statistical evaluation of the monthly rainfall of the stations in comparison with the estimated values of the model show that unlike the low coefficient of explanation of the observed and modeled data on a daily scale, the coefficient of explanation has increased significantly on a monthly scale so that the range Its fluctuation ranges from 0.45 in Ferizi station to 0.86 in Sarakhs station. As the time scale increases, the accuracy of Era۵ model estimates also increases. The results show that the dynamic and numerical methods used in the Era۵ model have the ability to provide estimates that are close to reality with the least error and have the ability to improve in providing more optimal results in other areas. These results show that Era5 model estimates have performed best on monthly and seasonal scales in regions with a warmer climate and less rainfall in Razavi Khorasan province. These results show that Era5 model estimates performed best in monthly and seasonal scales in areas with warmer climate and less rainfall in Razavi Khorasan Province. Therefore, if the skew error is fixed, the data of this model can be used as input data to agricultural and hydrological models.

Keywords: Precipitation assessment, Precipitation modeling, ECMWF, NSE, POD.

Citation: Nouri A. Nouri S. Omidvar J. Banejad H. and Modaresi F. 2024. A qualitative- quantitative evaluation of Era5 rainfall data in identifying the occurrence and amount of rainfall in Razavi Khorasan province. Iranian Water Research Journal. 52: ??-??. https://dx.doi.org/10.22034/IWRJ.2023.14295.2506