

تحلیل فراوانی حداکثر بارش روزانه و تعیین مناسب‌ترین توابع توزیعی در حوزه آبخیز سد بوستان، استان گلستان

ثریا یعقوبی^۱، چوقی بایرام کمکی^{۱*}، محسن حسینعلی زاده^۱، علی نجفی نژاد^۲، حمیدرضا پورقاسمی^۳
و مرزبان فرامزی^۴

(تاریخ دریافت: ۱۴۰۱/۴/۲؛ تاریخ پذیرش: ۱۴۰۱/۶/۶)

چکیده

تحلیل فراوانی بارش‌های روزانه و یا دوره بازگشت رخداد‌های بارش و سیلاب به دلیل پیچیدگی رفتاری فراوان در مدیریت منابع آب، از اهمیت زیادی برخوردار است. به عبارتی، عدم توجه به آن ممکن است سیل‌های مخرب را به دنبال داشته باشد. به همین منظور در پژوهش حاضر جهت بررسی و انتخاب مناسب‌ترین تابع توزیع، با استفاده از داده‌های هواشناسی و ماهواره CHIRPS در هفت ایستگاه موجود در حوزه آبخیز سد بوستان، سه تابع توزیع پیرسون ۳، بتا و گاما مورد مقایسه قرار گرفتند. آنالیز آماری نشان داد که داده‌های ماهواره‌ای در برآورد بارش روزانه، به دلیل بالا بودن مقدار خطای MAD, RMSE و NASH ناکارآمد هستند، به همین دلیل تنها از داده‌های ایستگاه‌های هواشناسی برای تعیین بهترین توزیع استفاده شد. برای این منظور زبان برنامه‌نویسی گوگل ارث انجین و پایتون مورد استفاده قرار گرفت. سپس از تابع توزیع منتخب برای تعیین حداکثر بارش روزانه، احتمال فراوانی در دوره‌های بازگشت ۲، ۱۰، ۵۰، ۱۰۰ و ۲۰۰ ساله استفاده شد. نتایج به‌دست‌آمده از آزمون نکویی برازش مجموع مربعات خطا، معیارهای ارزیابی آکائیک، بیزین و معیار واگرایی کولبک لیبلر نشان داد که در پنج ایستگاه کلانه، قرناق، پارک ملی گلستان، سد گلستان و گلیداغ تابع پیرسون ۳ مناسب‌ترین تابع توزیعی است. همچنین، در دو ایستگاه دیگر (گنبد و تمر)، تابع بتا به‌عنوان تابع مناسب تشخیص داده شد. این درحالی است که طبق نتایج به‌دست‌آمده توزیع گاما در منطقه مورد مطالعه دارای کارایی مناسبی نیست. پس می‌توان نتیجه گرفت که بارش‌های شدید و نامنظم از نظر زمانی و مکانی می‌تواند در انتخاب مناسب‌ترین تابع توزیع آماری در هر ایستگاه مؤثر باشد. بنابراین، توصیه می‌شود بارش‌های حداکثر محتمل و در نتیجه وقوع سیلاب‌های محتمل در نظر گرفته شوند تا با مدیریتی اصولی و دقیق از خسارات جانی و مالی در مناطق مستعد بخصوص در منطقه مورد مطالعه جلوگیری شود.

واژه‌های کلیدی: حداکثر بارش روزانه، تحلیل فراوانی، توزیع آماری، داده‌های هواشناسی و ماهواره‌ای، حوزه آبخیز سد بوستان

۱. گروه مدیریت مناطق بیابانی، دانشکده مرتع و آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان

۲. گروه مدیریت مرتع، دانشکده مرتع و آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان

۳. گروه مهندسی منابع طبیعی و محیط زیست، دانشکده کشاورزی، دانشگاه شیراز

۴. گروه مرتع و آبخیزداری، دانشکده کشاورزی، دانشگاه ایلام

*: مسئول مکاتبات: پست الکترونیکی: bkomaki@gmail.com, komaki@gau.ac.ir

مقدمه

بارش به عنوان کنترل کننده طبیعت نقش مهمی در بیلان جرم و انرژی در کره زمین دارد و به دلیل پیچیدگی رفتاری فراوان آن از اهمیت زیادی برخوردار است (۴ و ۱۶)، زیرا بارش‌های سیل آسا جزء عوامل تخریب سرزمین به حساب می‌آیند که منجر به خسارات زیادی می‌شوند. به همین منظور دانش کافی در مورد زمان وقوع این بارش‌ها کمک شایانی به آمادگی در مقابل خطرات و در نهایت کاهش خسارات احتمالی منطقه می‌کند (۲۲). از طرفی شدت بارش از نظر ساختار مکانی و زمانی روی فرایندهای هیدرولوژیکی از جمله تولید رواناب و میزان رطوبت خاک تأثیر می‌گذارد؛ بنابراین، تجزیه و تحلیل تغییرات بارندگی مورد مطالعه در مقیاس‌های مختلف از محلی تا جهانی بسیار مهم بوده و مورد توجه مدیران و ذینفعان مدیریت ریسک سیل است (۱۸ و ۲۷). منحنی‌های توزیعی بارندگی‌های شدید اغلب برای طراحی زیرساخت‌های شهری و ارزیابی ریسک استفاده می‌شوند (۷)، زیرا این منحنی‌ها روابط بین شدت (یا مقادیر) بارندگی‌های مختلف، مدت زمان بارندگی (معمولاً چند دقیقه تا چند روز) و دفعات مختلف وقوع بارش را به صورت دوره بازگشت مشخص می‌کنند (۲۴). به دلیل اهمیت بالای بارش‌های روزانه که منجر به سیل‌های مخرب شهری می‌شود، این داده‌ها مورد تجزیه و تحلیل قرار می‌گیرند. به عقیده سانیر و همکاران (۳۰) این سیل‌ها ممکن است بیشتر تحت تأثیر آب‌وهوا باشند تا بارش‌های روزانه (۳۴). در واقع حداکثر بارش روزانه، بیشترین مقدار بارندگی است که در طول یک روز یا ۲۴ ساعت، طی یک دوره زمانی معین مانند یک ماه و یا یک سال اتفاق می‌افتد. مقدار بارش روزانه، شدت بارندگی را نشان می‌دهد. این مقدار در کشوری مانند ایران که از نظر توزیع بارش سالانه، نوسان فراوانی دارد، از اهمیت بالایی برخوردار است (۷). با توجه به تغییرات مکانی و زمانی بارش‌ها و پراکندگی ایستگاه‌های زمینی، در زمینه برآورد بارندگی‌ها با استفاده از تصاویر ماهواره‌ای و سنسور از دوری پیشرفت‌های شایانی صورت گرفته است که از مهم‌ترین این داده‌های اقلیمی می‌توان

به The Tropical Rainfall Measuring Mission (TRMM)، Global Precipitation Measurement (GPM)، Climate Hazards Group Infrared Precipitation with Precipitation Estimation (PERSIANN)، Station Data from Remotely Sensed Information using Artificial Neural Networks اشاره کرد. این داده‌ها به عنوان مدل، جایگزین مناسبی از اطلاعات بارندگی در نظر گرفته شده‌اند (۱۵). از دلایل ضرورت این پژوهش می‌توان به تفاوت رژیم و الگوی مکانی و زمانی بارندگی در ایران به خصوص در استان گلستان به دلیل تغییرات توپوگرافی زیاد، کاربری اراضی متفاوت، دوری و نزدیکی نسبت به منبع رطوبت (دریای خزر) با تغییرات گسترده اشاره کرد. با توجه به تغییرات مکانی و زمانی بارش و همچنین پتانسیل سیل‌خیزی در کشور اطلاع از احتمال وقوع و یا دوره بازگشت سیلاب‌ها و همچنین اطلاع از محدوده گسترش سیل می‌تواند برنامه‌ریزان و متخصصان کشور را در زمینه راه‌کارهای کنترل این پدیده مخرب یاری نماید (۳). مرور ادبیات پژوهش نشان می‌دهد که مطالعات متعددی در زمینه داده‌های بارش ماهواره‌ای و تعیین بهترین توزیع‌های آماری در مناطق مختلف دنیا و ایران انجام شده است:

تافته و همکاران (۳۲) به بررسی نتایج داده‌های روزانه، ده روزه و ماهانه داده‌های بارشی تصاویر ماهواره‌ای با استفاده از سامانه گوگل ارث انجین در استان خوزستان پرداختند. در این پژوهش از محصولات بارشی CHIRPS، TRMM و GPM طی سال‌های ۱۳۹۳-۱۳۹۴ و ۱۳۹۶-۱۳۹۷ استفاده شد و با اطلاعات نه ایستگاه صحت‌سنجی شدند؛ نتایج آنها نشان دادند که سامانه در بازه ماهانه از شاخص کارایی بالایی برخوردار است و مقدار خطای نرمال شده آن حدود ۳۶ درصد است. در پژوهش دوان و همکاران (۱۰) که در حوزه Adgie واقع در ایتالیا انجام شد از ۸ منبع بارش استفاده شد. منابع بارش شامل داده‌های CHIRPS، PGE، PERSIANN-CDR، CMORPH، TRMM-3B42V7 و GSMaP_MVK بودند. ارزیابی‌ها در مقیاس زمانی روزانه، ماهانه و سالانه و در مقیاس مکانی در سطح حوزه صورت گرفت. نتایج حاکی از آن بود که سه داده

ماهواره‌های CHIRPS، TRMM و CMORPH_BLD عملکرد بهتری نسبت به بقیه داده‌ها دارند.

دی کاروالها و همکاران (۹) به مدل‌سازی توزیع کلی مقادیر شدت و احتمال وقوع سطح خاصی از حداکثر بارندگی روزانه در سه منطقه برزیل (غرب، جنوب شرقی و جنوبی) و رفتار آن‌ها در ۷۱ سال گذشته پرداختند. بر اساس نتایج پژوهش، منطقه غرب میانه در مقایسه با سایر مناطق دارای دوره بازگشت با حداکثر بارندگی روزانه کمتر از ۳۰۰ میلی‌متر بود. لیمساکول و سینفراک (۱۹) به بررسی روندهای بلندمدت و تغییرپذیری رویدادهای بارشی در تایلند پرداختند و به این نتیجه رسیدند که در بیشتر مناطق تایلند فراوانی بارش کمتر ولی شدت آن بیشتر شده است. علاوه بر این، شاخص اندازه‌گیری میزان حوادث بارشی نشان داد که گرایش به سمت شرایط مرطوب‌تر با بارش سنگین کمک بیشتری به کسری بارش سالانه می‌کند. یکی از نتایج این تغییرات افزایش فراوانی و شدت سیل است که به‌تازگی در بیشتر مناطق تایلند مشاهده شده است. موکیا و همکاران (۲۱) در دو منطقه لاتزیو و سیسیل ایتالیا به مطالعه حداکثر مقدار بارندگی روزانه پرداختند. آنها از شش تابع توزیع احتمال برای برازش نمونه‌های بارش شدید استفاده کردند. بهترین توزیع برازش برای نمونه‌های تجربی مقادیر شدید از طریق روش مجموع مربعات خطا (RSS) و آزمون کولموگروف-اسمیرنوف (K-S) ارزیابی شد. فوفانا و همکاران (۱۲) در مطالعه خود با استفاده از پنج شاخص شدت بارندگی که توسط تیم متخصص برای نظارت تخمین زده شده بود به تجزیه و تحلیل شاخص‌های شدت و فراوانی بارش پرداختند. جهت تخمین دوره بازگشت سیل و بارندگی در دوره زمانی ۵ تا ۱۰۰ ساله بر اساس حداکثر بارندگی روزانه از توزیع گامبل استفاده شد. نتایج نشان داد که دوره بازگشت حدود ۵۸ درصد از حوادث سیل در باماکو ناشی از بارندگی عادی است که ۳۳/۳ درصد ناشی از بارندگی شدید است. فرزندلی و پزند (۱۱) به مقایسه روند نیکویی برازش توزیع‌های احتمالی با حجم نمونه ۳۱ تا ۱۱۹ سال در شهر مشهد پرداختند. سپس

هفت قانون احتمالی گامبل نوع ۱، گامای ۲ پارامتری، لوگ نرمال ۲ و ۳ پارامتری، پارتوی تعمیم‌یافته، حدی تعمیم‌یافته و پیرسون نوع ۳ با پنج روش برآوردی بر داده‌ها برازش داده شد. نتایج نشان داد که دو تابع پیرسون نوع ۳ و پارتوی تعمیم‌یافته بهترین برازش را برای نمونه‌های فوق دارند، و قانون گامای ۲ پارامتری در اولویت بعدی قرار دارد. داوری‌زاده و وفاخواه (۸) در بررسی تعیین حداکثر بارش روزانه در حوزه آبخیز امامه از توزیع‌های احتمال استفاده کردند. نتایج نشان داد که توزیع آماری مقدار حدی تعمیم‌یافته مناسب‌ترین توزیع آماری در ایستگاه‌های منطقه در بازه زمانی ۴۷ سال می‌باشد، در مقابل تابع پارتو کمترین تطابق را در پیش‌بینی حداکثر بارش روزانه در دوره بازگشت‌های مختلف دارد. جهانگیری پیرمردای و همکاران (۱۴) به بررسی انتخاب بهترین تابع توزیع آماری برای دبی‌های لحظه‌ای سیل در حوزه‌های شهرچای (بند)، نازلوچای (تپیک)، باراندوزچای (دیزج) در غرب دریاچه ارومیه و حوزه‌های گادارچای (پی‌قطعه) و مهابادپای (بیطاس) در سال آبی ۱۳۹۱-۱۳۴۴ پرداختند. این داده‌ها با توزیع‌های احتمالی رایج مثل نرمال، گامبل، لوگ‌پیرسون تیپ سه و ویکبای پنج پارامتره برازش داده‌شده و پس از انجام آزمون‌های نکویی برازش بهترین تابع توزیع برازش تابع توزیع ویکبای برآورد بهتر و نتایج قابل قبولی برای داده‌های لحظه‌ای سیل برای ایستگاه‌های منتخب داشتند. ارجمند و همکاران (۵) جهت برآورد دوره بازگشت بارندگی سالانه و ارائه بهترین توزیع آماری از داده‌های بارندگی سالانه ایستگاه سینوپتیک کرج با دوره بازگشت‌های ۵، ۲۵، ۵۰، ۱۰۰ و ۲۰۰ ساله استفاده کردند. آزمون $K-S$ ، RSS ، $CHI2$ جهت انتخاب بهترین توزیع آماری در نرم‌افزار SMADA استفاده شد. نتایج حاصل از پژوهش نشان داد که توزیع لوگ-نرمال مناسب‌ترین توزیع است.

مرور منابع بیانگر این واقعیت است که تصاویر ماهواره‌ای استخراج‌شده از گوگل ارث انجین به مراتب داده‌های اقلیمی مناسبی برای تجزیه و تحلیل آزمون‌های آماری هستند و

ناقص و پرت، ایستگاه‌های مورد بررسی به هفت ایستگاه با آمار کافی برای تجزیه و تحلیل کاهش یافتند. شکل ۲ نمودار مربوط به طول دوره آماری ایستگاه‌های منطقه مورد مطالعه است. در این مطالعه جهت مقایسه داده‌های بارش هواشناسی با داده‌های ماهواره‌ای، از داده ماهواره CHIRPS با حداکثر بارش روزانه در دوره زمانی ۴۰ ساله و همین طور سه تابع توزیع احتمال شامل بتا، پیرسون نوع ۳ و گاما استفاده شد. در انتها جهت تخمین پارامترها نیز آزمون مجموع مربعات خطا مورد استفاده قرار گرفت.

ماهواره CHIRPS

داده اقلیمی CHIRPS یک پایگاه داده جهانی با طول دوره آماری بیش از ۳۰ سال بوده که از سال ۱۹۸۱ تاکنون دارای داده است. این پایگاه داده، داده‌های بارش ایستگاهی و ماهواره‌ای را به منظور بررسی روند و پایش خشک‌سالی تلفیق کرده و یک پایگاه داده بارش شبکه‌بندی شده به صورت سری زمانی ارائه می‌دهد. در این پژوهش نیز از نسخه CHIRPS 2.0 (<http://chg.geog.ucsb.edu/data/chirps>) و تفکیک مکانی ۰/۲۵ * ۰/۲۵ استفاده شده است (۲۸).

ارزیابی رابطه بین داده‌های ماهواره‌ای و زمینی

برای ارزیابی روابط بین داده‌های واقعی و داده‌های مستخرج از سامانه گوگل ارث‌انجین، شاخص‌های عملکردی مختلفی وجود دارد که عبارت‌اند از ضریب تعیین (R^2)، مجذور میانگین مربع خطا (RMSE)، میانگین قدر مطلق خطا (MAD) و ضریب نش - ساتکلیف (NASH) (۲۳) استفاده شد روابط (۱، ۲ و ۳). با توجه به بررسی‌های صورت گرفته در زمینه دقت روابط بین داده‌های واقعی و ماهواره‌ای مشخص شده است که هرچقدر از کمترین مقدار MAD و RMSE برخوردار باشند دارای بیشترین دقت هستند.

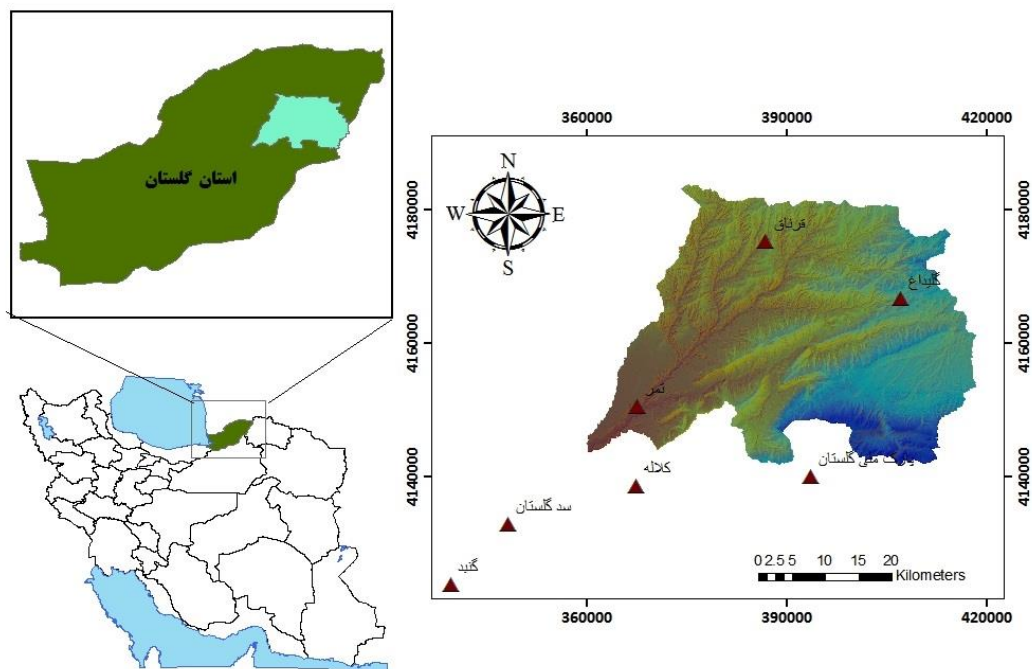
$$MAD = \sum \frac{|e_i|}{N} \quad (1)$$

نتایج سامانه نشان می‌دهد که عملکرد مناسبی در تحلیل داده‌های بارشی در بازه‌های زمانی (ماهانه و سالانه) دارد. به همین منظور، یکی از اهداف ما در این مطالعه نه تنها مقایسه داده‌های ماهواره‌ای و ایستگاه‌های هواشناسی در بازه زمانی ماهانه و سالانه است بلکه بررسی عملکرد این داده‌ها در بازه زمانی روزانه است. همچنین، هدف دیگر تحلیل فراوانی و تعیین حداکثر بارش روزانه حوزه آبخیز سد بوستان در بارش سیلابی سال ۹۷ است. زیرا تا کنون مطالعه‌ای در خصوص تحلیل فراوانی بارش سیلابی استان گلستان به‌ویژه در مناطقی که بیشتر در معرض خسارات مالی قرار گرفته بودند، انجام نشده است. از این رو پژوهش حاضر درصدد آن است که نه تنها ناتوانی عملکرد داده‌های ماهواره‌ای را در بررسی بارش‌های روزانه به نمایش بگذارد بلکه با استفاده از داده‌های زمینی حداکثر بارش روزانه با دوره بازگشت مشخص را تعیین و بهترین تابع برازش یافته را انتخاب نماید.

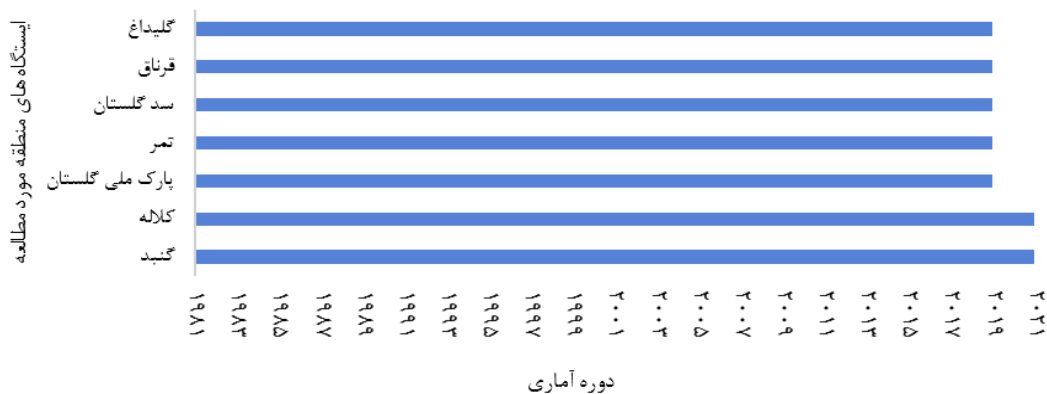
مواد و روش‌ها

حوزه آبخیز سد بوستان یکی از زیرحوزه‌های حوزه آبخیز گرگان‌رود واقع در شرق استان گلستان و شمال شرقی شهرستان کلاله است. این حوزه در محدوده جغرافیایی عرض $37^{\circ} 23' 00''$ تا $37^{\circ} 46' 00''$ شمالی و طول $55^{\circ} 26' 26''$ تا $56^{\circ} 04' 00''$ شرقی واقع شده است. شکل ۱ موقعیت بالادست حوزه آبخیز سد بوستان به همراه ایستگاه‌های مطالعاتی اطراف آن را نشان می‌دهد. مساحت این حوزه آبخیز حدود ۱۵۶۲ کیلومتر مربع، متوسط بارش سالانه آن حدود ۴۶۵ میلی‌متر و اقلیم آن نیمه‌خشک تا نیمه‌مرطوب است.

با توجه به اینکه هدف از مطالعه حاضر، بررسی دوره بازگشت و احتمال فراوانی بارش در ایستگاه‌های بالادست حوزه آبخیز سد بوستان است، نیاز است تا داده‌های بارش از ایستگاه‌های محدوده مورد مطالعه جمع‌آوری و مورد استفاده قرار گیرند. طبق بررسی‌های صورت گرفته، در محدوده حوزه آبخیز سد بوستان با حذف ایستگاه‌های باران‌سنجی دارای داده



شکل ۱. موقعیت جغرافیایی و ایستگاه‌های هواشناسی منطقه مورد مطالعه (سیستم مختصات UTM زون ۴۰) (رنگی در نسخه الکترونیکی)



شکل ۲. نمودار طول آماری ایستگاه‌های حوزه آبخیز سد بوستان

تناسب مناسبی برقرار است. اگر مقدار آن بزرگ‌تر از ۰/۷۵ باشد نتایج داده‌های ماهواره‌ای خوب توصیف می‌شود. چنانچه این ضریب به عدد یک نزدیک تر مدل از عملکرد بهتری برخوردار است و هرچقدر به سمت منفی گرایش داشته باشد از کارایی مدل کاسته می‌شود.

توابع توزیع آماری در تحلیل داده‌های بارشی زمینی توزیع آماری گاما جزء توابع توزیع دو پارامتری است. تابع

$$RMSE = \sqrt{\frac{\sum e^2}{N}} \quad (2)$$

$$NASH = 1 - \frac{\sum_{i=1}^N [Q_0 - Q_s]^2}{\sum_{i=1}^N [Q_0 - \bar{Q}_0]^2} \quad (3)$$

که در آن Q_0 بارش مشاهده‌ای، Q_s بارش داده ماهواره‌ای، \bar{Q}_0 میانگین بارش مشاهده‌ای و N تعداد داده‌ها است. ضریب نش- ساتکلیف از منفی بی‌نهایت الی یک متغیر است. اگر مقدار آن برابر یک باشد بین داده‌های مشاهده‌ای و ماهواره‌ای

همچنین گفتنی که با برابر بودن مقدار x با 0 یا 1 ، تابع چگالی برابر با صفر می‌شود. ولی برای بقیه نقاط در دامنه 0 تا 1 مقدار تابع چگالی غیرصفر خواهد بود. اگر پارامترهای شکل تابع توزیع بتا از 1 کوچکتر باشند، نمودار در نقطه 0 و 1 دارای شکست خواهد شد و به نظر پیوسته نخواهد بود. ولی در بقیه حالات پیوستگی تابع چگالی مشخص است، از طرف دیگر با بزرگ شدن پارامترهای شکل در توزیع بتا، نمودار تابع چگالی به نرمال نزدیک خواهد شد.

تابع پیرسون نوع ۳، همان توزیع گامای تعمیم یافته است و یکی از پرکاربردترین توزیع‌های آماری برای تجزیه و تحلیل فراوانی هیدرولوژیکی است. برای استفاده از این تابع نیاز به محاسبه چولگی و انحراف معیار با استفاده از رابطه ۸ است. بر اساس محاسبات، انحراف معیار یا s برابر با مقدار 0.089 و C_s برابر با 0.836 است.

$$\lambda = \frac{s_x}{\sqrt{\beta}} B = \left(\frac{2}{C_s} \right)^2; \varepsilon = \bar{x} - s_x \sqrt{\beta} F(x) \quad (8)$$

برای انتخاب توزیع برتر آزمون‌های آماری مختلفی وجود دارند؛ توزیعی که در مجموع برازش بهتری نسبت به بقیه توزیع‌ها در منطقه نشان می‌دهد انتخاب می‌شود. آزمون‌هایی که امروزه از آن‌ها برای انتخاب توزیع استفاده می‌شود که شامل آزمون مربع کای دو، آزمون کولموگروف-اسمیرنوف، آزمون اندرسون-دالینگ و آزمون مجموع مربعات خطا است (۲). به همین منظور در این پژوهش از آزمون نکویی برازش مجموع مربعات خطا استفاده شد. مجموع مربعات خطا (RSS) اختلاف مجموع مربعات بین هر داده مشاهده شده و میانگین آن گروه است. بنابراین، می‌تواند به عنوان ابزار اندازه‌گیری اختلاف در یک گروه مورد استفاده قرار گیرد. اگر همه موارد در یک گروه یکسان باشند، RSS برابر با 0 خواهد بود. با فرض اینکه y_i مقدار متغیر خروجی مدل و x_i مقدار متغیر ورودی مدل، در مدل رگرسیونی خطی ساده استاندارد، $y_i = (w_0 + w_1 x_i)_i$ است، که w_0 و w_2 عرض از مبدأ و ضریب مدل بوده و ε مقدار خطا یا تفاوت مدل با واقعیت است (رابطه ۹):

چگالی احتمال برای این متغیر تصادفی براساس مقادیر بارش x و پارامترهای α و β (به ترتیب پارامتر شکل و مقیاس توزیع آماری) دارای ارزش‌های مثبت هستند و $\Gamma(\alpha)$ نیز مقدار تابع گاما در نقطه α است (رابطه ۴):

$$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad (4)$$

با توجه به این تعریف مشخص است که تکیه‌گاه این متغیر تصادفی مقدارهای مثبت خواهد بود (رابطه ۵):

$$f(x; k, \theta) = \frac{x^{k-1} e^{-x/\theta}}{\theta^k \Gamma(k)} \quad (5)$$

در تابع توزیع چگالی احتمال، k و θ (به ترتیب پارامتر شکل و مقیاس توزیع آماری) بوده و گاهی تابع توزیع گاما با پارامتر سوم یعنی موقعیت جابجایی loc بیان می‌کنند که در این حالت تابع توزیع چگالی احتمال رابطه ۵ به تابع $\frac{f(y; k)}{\theta}$ تبدیل می‌شود که در آن y برابر $\frac{x - loc}{\theta}$ است؛ در این پژوهش نیز از افزونه `fitter` و `scipy.stats` استفاده شد (۱۷).

تابع توزیع بتا (Beta Distribution) یکی از توزیع‌های احتمالی است که مقدار متغیر تصادفی در دامنه عددی بین 0 تا 1 تعریف شده است. این توزیع به خصوص در بحث مربوط به آماره‌های ترتیبی و توزیع آن‌ها برای متغیرهای تصادفی یکنواخت اهمیت پیدا می‌کند. همچنین، در بسیاری از موارد از جمله در استنباط آماری بر پایه بیزین (Bayesian Inference) از توزیع بتا به عنوان توزیع پیشین توزیع دو جمله‌ای برای برآورد بیزی پارامتر احتمال موفقیت استفاده می‌شود. پارامترهای این توزیع با α و β به عنوان پارامترهای شکل (Shape Parameter) مشخص می‌شود. اگر متغیر تصادفی X دارای توزیع بتا با دامنه ارزشی $0 \leq X \leq 1$ باشد و تابع چگالی احتمال برای این متغیر تصادفی با رابطه ۶ بیان می‌شود:

$$f(x; \alpha, \beta) = \frac{1}{\text{Beta}(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad (6)$$

منظور از $\text{Beta}(\alpha, \beta)$ نسبت دو مقدار تابع گاما است (رابطه ۷):

$$\text{Beta}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \quad (7)$$

جایی که K تعداد متغیرهای مستقل است L ارزش احتمال و n تعداد اندازه گیری های ثبت شده است.

معیار واگرایبی کولبک لیبلی (Kullback-Leibler Divergence) KLD روشی برای اندازه گیری تفاوت بین دو توزیع احتمال بر روی یک متغیر x است که به طور رایج در ادبیات داده کاوی استفاده شده است (۳۵). این مفهوم در تئوری احتمالات و نظریه اطلاعات سرچشمه گرفته است. از جمله کاربردهای این معیار می توان به ارتباط نزدیک آن به آنتروپی نسبی در سیستم های اطلاعاتی اشاره کرد (رابطه ۱۲):

$$D_{KL}(p \parallel q) = \sum_{i=1}^N p(x_i) \cdot \log\left(\frac{p(x_i)}{q(x_i)}\right) \quad (12)$$

که در آن: $p(x)$ توزیع واقعی و $q(x)$ توزیع تقریبی برای داده ها است. این شاخص نشان می دهد که توزیع تقریبی چقدر نسبت به توزیع واقعی دور است. از طرفی هر چه مقدار شاخص KL کوچک تر باشد، میزان مطابقت بین توزیع $p(x)$ با $q(x)$ بیشتر است. به منظور اجرای برخی مراحل پژوهش از زبان برنامه نویسی گوگل ارث انجین و پایتون استفاده شد. برای استخراج داده های بارشی ۴۰ ساله از داده ماهواره CHIRPS از برنامه گوگل ارث انجین استفاده شد و اطلاعات ماهانه و سالانه در طول دوره آماری در منطقه مورد مطالعه استخراج شد. زبان برنامه نویسی پایتون با دارا بودن افزونه های مثل Fitter و Scipy.Stats توانایی لازم برای اجرای توابع توزیع آماری از جمله گاما، بتا، پیرسون نوع ۳ و انواع دیگر توابع را دارد (۱۷). به همین منظور، ورودی برنامه شامل اطلاعات بارش هواشناسی مربوط به ایستگاه های موجود در منطقه مورد مطالعه در طول دوره آماری است. همچنین، داده های ورودی توسط برنامه مرتب، تصحیح، ذخیره و اجرا شد، در نهایت داده ها با استفاده از سه تابع توزیع چگالی احتمال برازش داده شدند.

نتایج و بحث

از آنجا که هدف پژوهش حاضر ارزیابی بارندگی سیل آسای

$$RSS = \sum_{i=1}^n (y_i - f(x_i))^2 = \sum_{i=1}^n (\varepsilon_i)^2 = \sum_{i=1}^n (y_i - (w_0 + w_1 x_i))^2 \quad (9)$$

معیار ارزیابی آکائیک (AIC) و معیار اطلاعات بیزین (BIC) برای ارزیابی مدل هایی که بر مبنای استنباط آماری ایجاد می شوند، ابزارها و معیارهای مختلف وجود دارند؛ از جمله AIC (Akaike Information Criterion) که معیاری برای سنجش نیکویی برازش است. به ویژه که انتخاب مدل AIC به این دلیل است که در مقایسه با مدل واقعی کمترین میزان قابل انتظار انحراف KL دارد (۲ و ۶). در واقع AIC (رابطه ۱۰) یک تکنیک براساس نمونه مناسب برای برآورد احتمال یک مدل برای پیش بینی یا برآورد ارزش ها است (۶).

$$AIC = 2k - 2 \log(L) \quad (10)$$

یکی دیگر از ابزارهای ارزیابی، معیار BIC است که میزان اطلاع از دست رفته توسط مدل را مشخص می کند. معیار BIC که مخفف Bayesian Information Criterion است، بر مبنای تابع درست نمایی محاسبه شده و ارتباط نزدیکی با معیار ارزیابی آکائیک (AIC) دارد. برای انتخاب مدل، معیارهای اطلاعات آکائیک (AIC) یا معیار اطلاعات بیزین (BIC) اغلب استفاده می شود. پایین ترین مقدار در مدل ها، بهترین حالت را نشان می دهند (۳۱). هرچه تعداد پارامترهای مدل بیشتر باشد، مقدار AIC نیز بزرگتر، در نتیجه میزان اطلاعاتی که توسط مدل نادیده گرفته شده است، بیشتر خواهد بود. بنابراین، مدلی که بتواند کمترین میزان پیچیدگی و در عین حال بیشترین میزان برازش را داشته باشد، مدل مناسب تشخیص داده خواهد شد. در BIC نسبت به AIC میزان جریمه برای تعداد پارامترها در مدل بیشتر است (۱). همچنین معیار انتخاب مدل جایگزین، معیار ارزیابی اطلاع آکائیک است، از طرفی BIC (رابطه ۱۱) مبادله بین مدل مناسب و پیچیدگی مدل را اندازه گیری می کند (۲۶ و ۳۳).

$$BIC = 2 \log(n)K - 2 \ln(L) \quad (11)$$

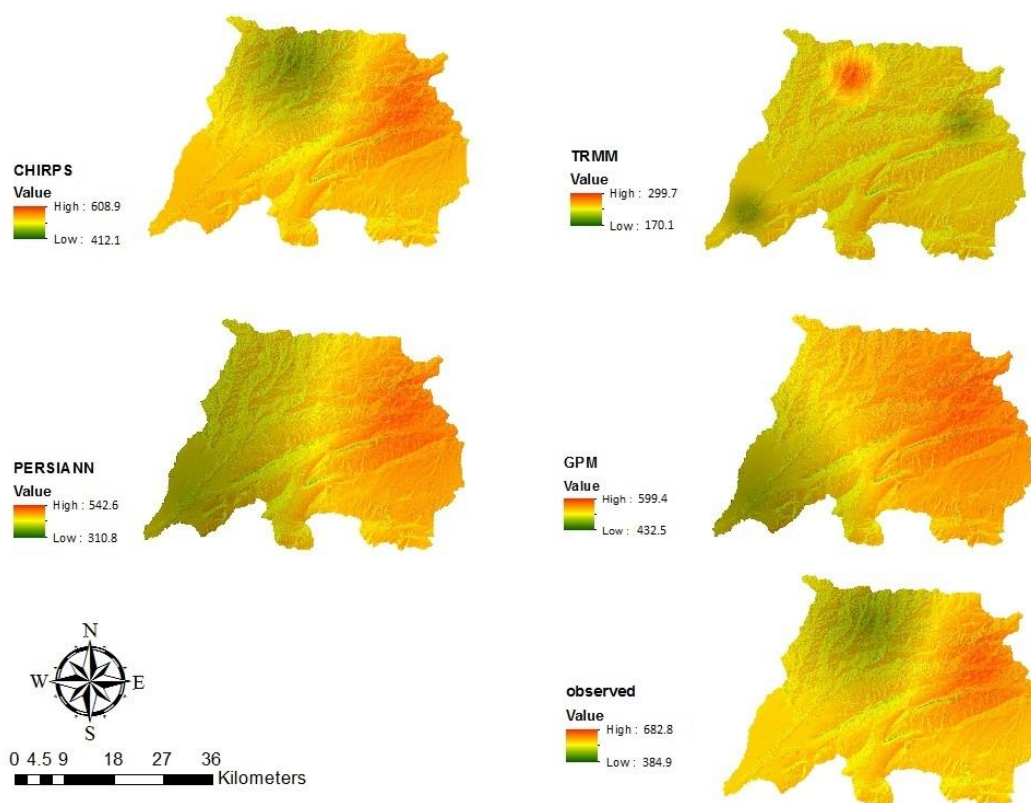
توجه به اینکه هر چقدر مقدار خطا کمتر و ضریب تعیین بالاتر باشد به معنی کارایی بالای روش مورد نظر است. در این پژوهش نیز نتایج حاصل نشان داد که R^2 در داده ماهواره‌ای CHIRPS با ۸۱ درصد و مقدار خطای مجذور مربعات ۶/۴ نسبت به بقیه داده‌های ماهواره‌ای در تعیین میزان بارش ماهانه نتایج قابل قبول تری را ارائه داده است. همچنین، میزان ضریب نش-ساتکلیف با ۹۲ درصد عملکرد بالایی در تطبیق داده‌های ماهواره‌ای و مشاهداتی داشته است. در مقابل داده ماهواره‌ای TRMM با میزان ضریب تعیین ۱۸ درصد، خطای مجذور مربعات ۳۴/۳۹، میانگین قدر مطلق خطا ۱۷/۵ و ضریب نش به میزان ۱۴ درصد کمترین رابطه را با داده‌های مشاهداتی دارد. سپس نمودار میزان بارندگی ماه‌های مختلف سال در دوره زمانی ۴۰ ساله ترسیم شد (شکل ۶) که نشان می‌دهد میانگین بارندگی با میزان ۲۶۰ میلی‌متر در سه‌ماهه اول نسبت به بقیه ماه‌های سال بسیار بیشتر است. نتایج نمودار ماهانه داده‌های CHIRPS نشان‌دهنده آن است که در سال ۲۰۱۹ در استان گلستان بیشترین میزان بارش در ماه‌های ژانویه، فوریه و مارس به ترتیب ۱۵۱/۱، ۱۳۸/۶ و ۱۶۷/۳ میلی‌متر است که به داده‌های واقعی اندازه‌گیری شده نزدیکتر است (شکل ۷).

نتایج بررسی مقادیر چهار داده ماهواره‌ای استخراج شده از گوگل‌ارث‌انجین در بازه زمانی روزانه با خطای بالایی همراه بودند (شکل ۸). با توجه به نمودارهای بدست آمده، میزان ضریب تعیین بین داده‌های روزانه ماهواره‌ای و مشاهداتی بسیار کم برآورد شده است. هرچند در داده ماهواره‌ای CHIRPS، میزان خطای MAD، RMSE و NASH به ترتیب ۷/۷، ۱۴/۳۷ و ۰/۳۹ بدست آمد اما میزان R^2 یا ضریب تعیین (۰/۱) معقول ارزیابی نشده است. حتی نمودار بارش روزانه گویای این موضوع است که داده‌های ماهواره‌ای مقدار بارش روزانه را بسیار کمتر از میزان اندازه‌گیری شده برآورد کرده است که لازم است در آینده با به کارگیری محصولاتی با الگوریتم دقیق‌تر دقت داده‌های ماهواره‌ای افزایش چشمگیری داشته باشد.

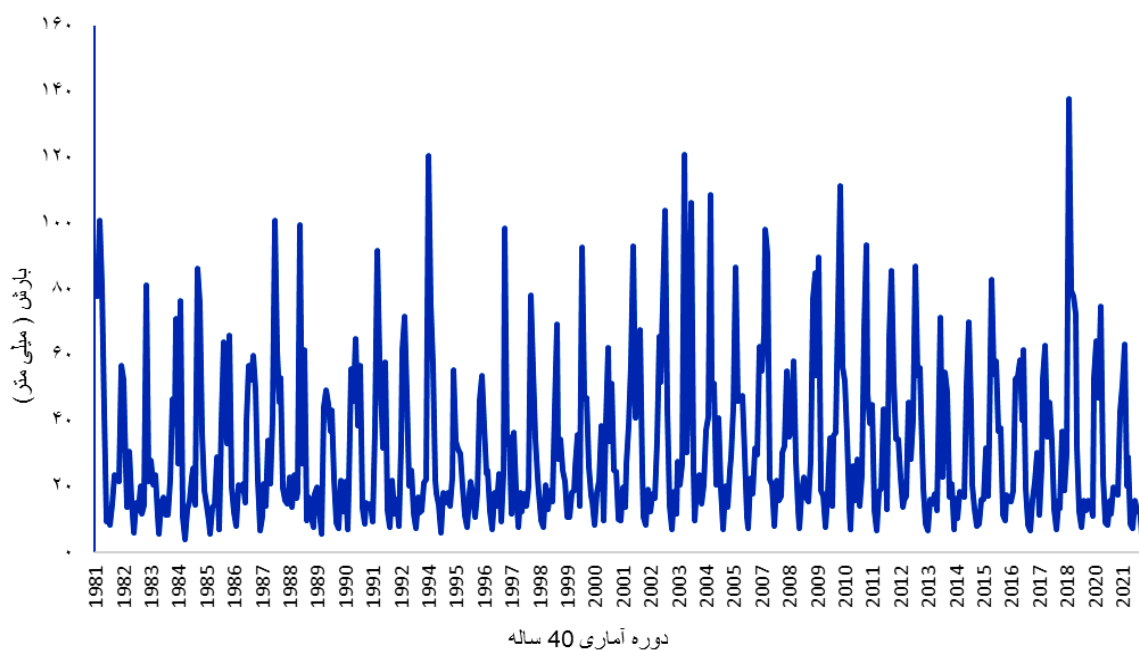
به همین منظور با استفاده از داده‌های CHIRPS، میزان

سال ۲۰۱۹ در استان گلستان است سعی شده که داده‌های بارشی این سال مورد بررسی قرار گیرد، چراکه بخش قابل توجهی از بارندگی‌ها در استان گلستان طی فصل زمستان (ژانویه تا مارس) ۲۰۱۹ اتفاق افتاده است. به همین منظور پس از آماده‌سازی داده‌های بارش هواشناسی ایستگاه‌های موجود، داده‌های بارش منطقه مورد مطالعه نیز با استفاده از داده‌های اقلیمی ماهواره CHIRPS، GPM، TRMM و PERSIANN در گوگل‌ارث‌انجین استخراج شد سپس به منظور برآورد منطقه‌ای مقادیر بارش در سال ۲۰۱۹ بین ایستگاه‌های موجود در منطقه به روش کریجینگ معمولی درون‌یابی انجام شد، که نتایج پهنه‌بندی در شکل ۳ آورده شد. با توجه به نتایج حاصل مشخص شد که داده‌های CHIRPS در مقایسه با داده‌های دیگر نتایج مشابه‌ای با داده‌های مشاهداتی به نمایش گذاشته است. حداکثر و حداقل بارندگی سال ۲۰۱۹ در نقشه واقعیت زمینی به ترتیب ۶۸۲/۸ و ۳۸۴/۹ میلی‌متر بوده که با نتایج بارش سالانه داده‌های CHIRPS با حداکثر و حداقل بارندگی ۶۰۸/۹ و ۴۱۲/۱ میلی‌متر نزدیک است. همینطور میزان بارش به خوبی در تمامی ایستگاه‌های موجود در منطقه توزیع شده است. به‌طوری که بیشترین بارندگی سالانه در ایستگاه گلیداغ و کمترین بارش سالانه در ایستگاه قرناق اتفاق افتاده است. از طرفی با وجود نزدیک‌بودن مقدار داده‌های بارش GPM و PERSIANN به نقشه مشاهداتی، میزان پراکنش و توزیع بارش در منطقه به خوبی نمایش داده نشده است، چرا که نواحی جنوبی منطقه برخلاف نقشه مشاهداتی کمترین بارش را به خود اختصاص داده‌اند. این مطلب نشانگر این موضوع است که داده‌های ماهواره‌ای CHIRPS در بازه زمانی سالانه نتایج قابل قبول‌تری را ارائه می‌دهد. درنهایت با استفاده از داده‌های CHIRPS به‌عنوان مناسب‌ترین داده اقلیمی از لحاظ بازه زمانی سالانه، نمودار بارش (۱۹۸۱-۲۰۲۱) در منطقه مورد مطالعه رسم شد (شکل ۴).

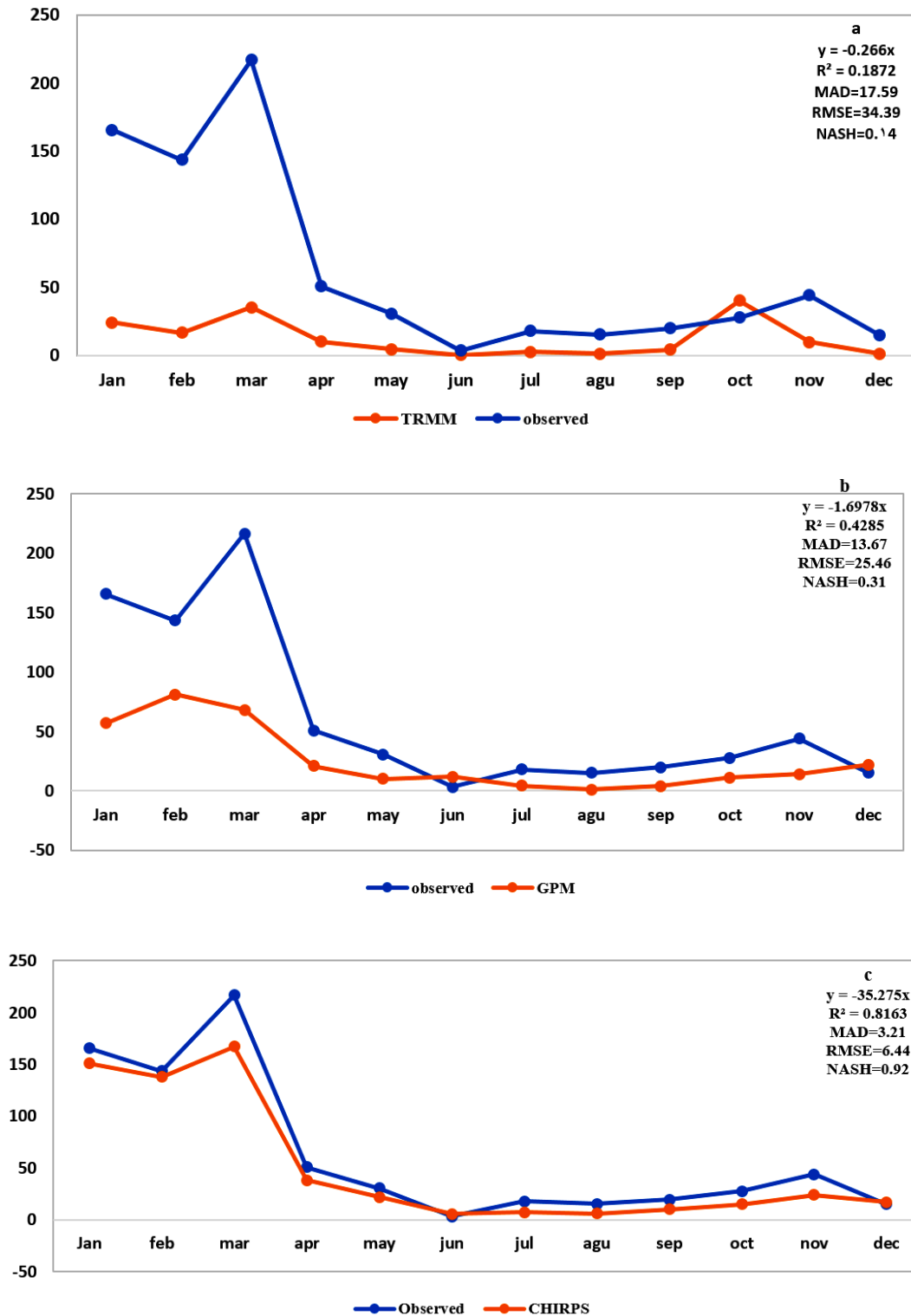
در مرحله بعد جهت مقایسه بارش ماهانه داده‌های مشاهداتی و ماهواره‌ای نمودار همبستگی رسم شد (شکل ۵). با



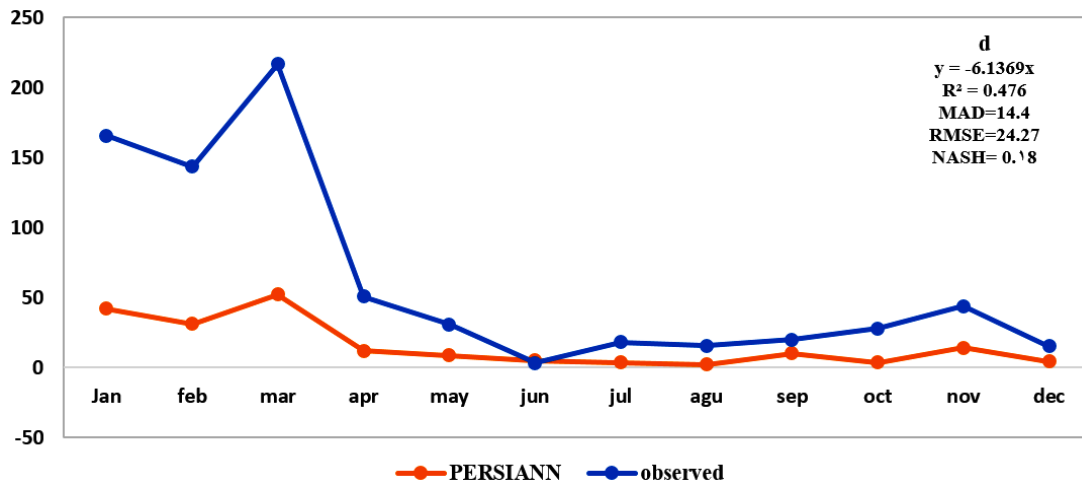
شکل ۳. نقشه پهنه‌بندی بارش سالانه ۲۰۱۹ داده‌های ماهواره‌ای و ایستگاه‌های هواشناسی منطقه مورد مطالعه (رنگی در نسخه الکترونیکی)



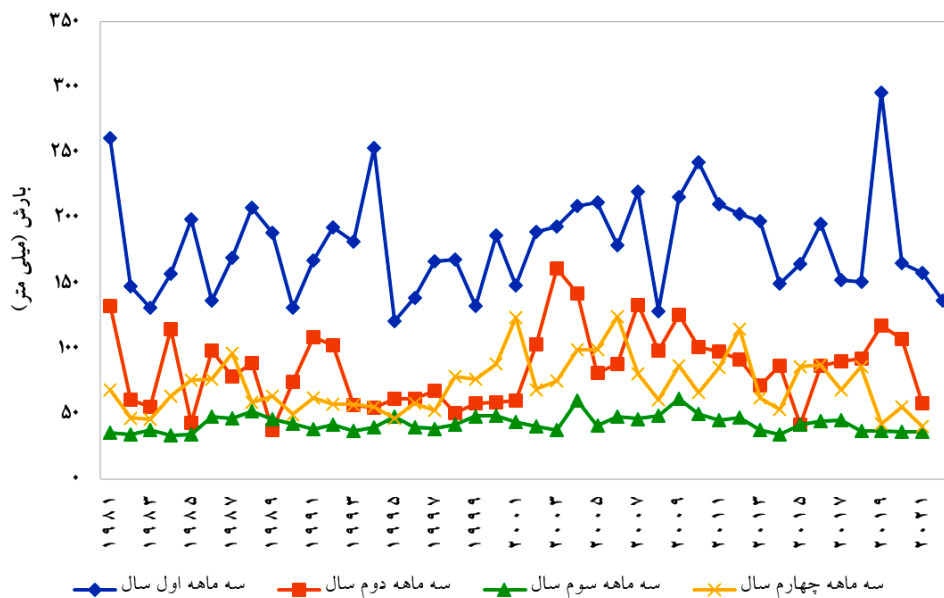
شکل ۴. داده‌های بارش ماهواره CHIRPS (۱۹۸۱ تا ۲۰۲۱) در حوزه آبخیز مورد مطالعه



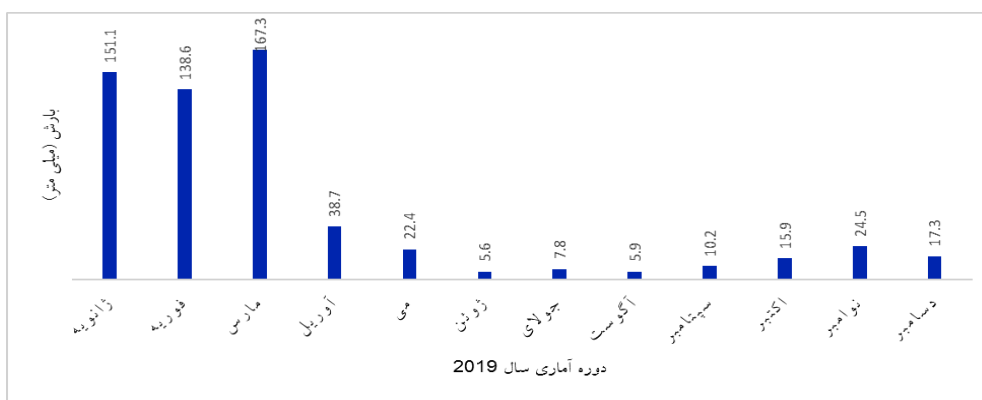
شکل ۵. نمودار مقایسه داده‌های ماهواره‌ای (a) TRMM، (b) GPM، (c) CHIRPS، (d) PERSIANN و داده‌های مشاهده‌ای منطقه مورد مطالعه در بازه زمانی ماهانه (رنگی در نسخه الکترونیکی)



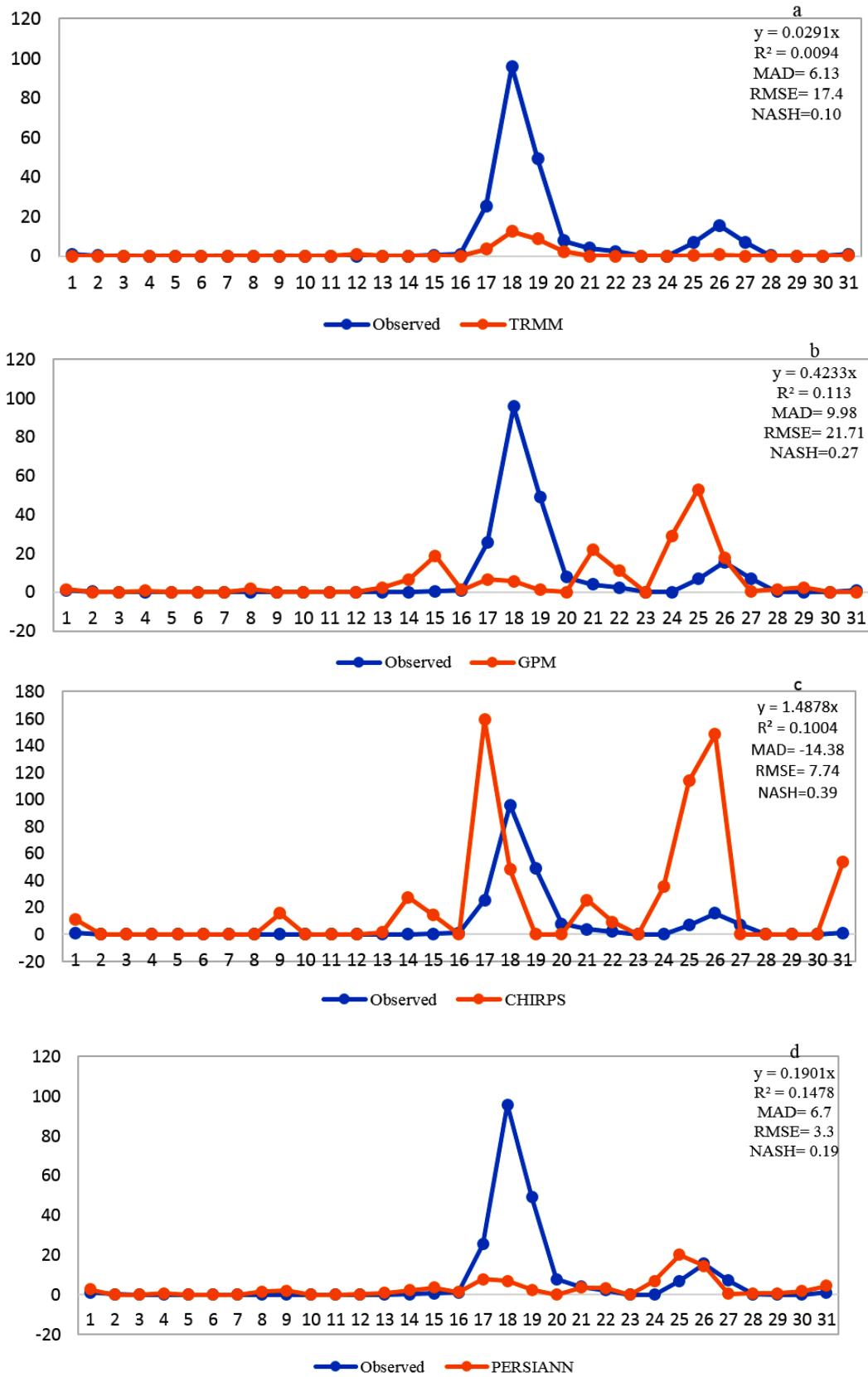
ادامه شکل ۵



شکل ۶. داده‌های بارش ۴۰ ساله ماهواره CHIRPS در ماه‌های مختلف سال در حوزه آبخیز مورد مطالعه (رنگی در نسخه الکترونیکی)



شکل ۷. داده‌های بارش ماهانه (ژانویه تا دسامبر ۲۰۱۹) در حوزه آبخیز مورد مطالعه



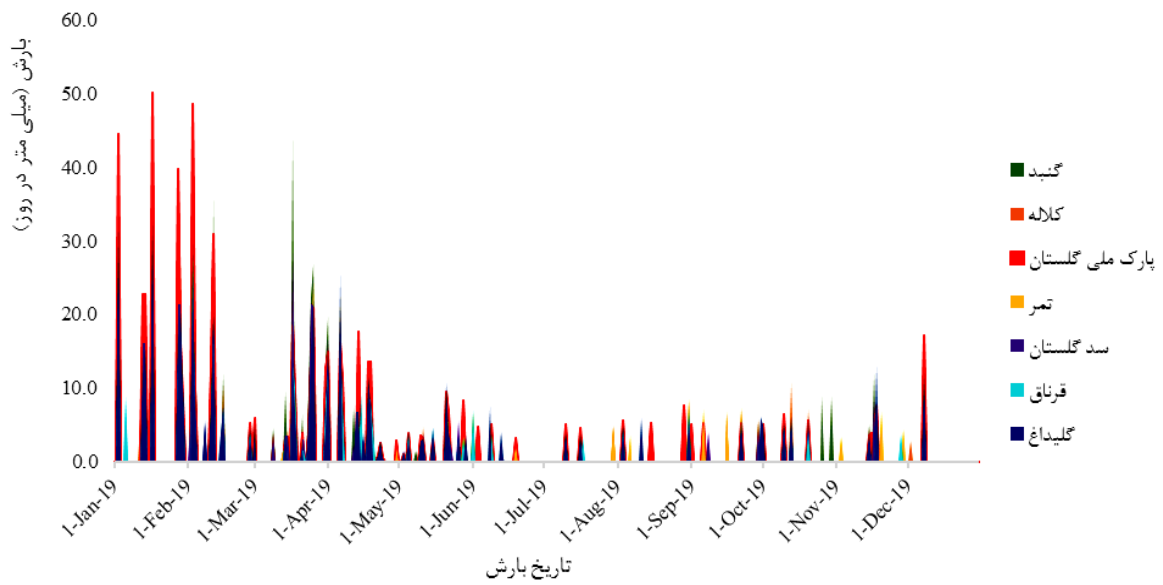
شکل ۸. مقایسه داده‌های ماهواره‌ای (a) TRMM، (b) GPM، (c) CHIRPS، (d) PERSIANN و داده‌های مشاهده‌ای منطقه مورد مطالعه در بازه زمانی روزانه (رنگی در نسخه الکترونیکی)

اطلاعات مربوط به میزان خطای مجموع مربعات، معیار اطلاعات ارزیابی آکائیک (AIC)، معیار اطلاعات بیزین (BIC) و معیار واگرایی به دست آمد. با توجه به نتایج به دست آمده، مشخص شد که تنها در ایستگاه‌های گنبد و تمر (شکل‌های ۱۰ و ۱۱) تابع بتا به عنوان بهترین توزیع آماری، برازش مناسبی بر روی داده‌های بارش داشته است، زیرا در این دو ایستگاه میزان مربعات خطا، معیار اطلاعاتی آکائیک، بیزین و میزان واگرایی از میزان کمتری نسبت به توابع دیگر برخوردار است. میزان مجموع مربعات خطا در ایستگاه گنبد و تمر به ترتیب ۰/۰۳۳ و ۰/۰۲۴ به دست آمد که نسبت به توابع دیگر از خطای کمتری برخوردار بود. از طرف دیگر میزان معیار آکائیک (AIC) و معیار بیزین (BIC) در ایستگاه گنبد به ترتیب برابر با ۹۹۲/۹۵۴۵ و ۱۱۶۱/۸۴۴- و در ایستگاه تمر به ترتیب برابر با ۱۱۶۱/۸۴۴ و ۴۰۰/۷۱۵- به دست آمد. با توجه به این که بیشترین میزان بارندگی در سه ماهه اول سال ۲۰۱۹ به وقوع پیوسته حداکثر بارش روزانه در هر دو ایستگاه گنبد و تمر در تاریخ ۲۷ اسفند ۱۳۹۷ (۱۸-۰۳-۲۰۱۹) اتفاق افتاده که میزان بارش آن روز در ایستگاه گنبد ۱۰۷/۸ و در ایستگاه تمر ۱۱۱ میلی‌متر ثبت شده است. دوره بازگشت بارش در این روز در ایستگاه گنبد و تمر به ترتیب برابر با ۴۲۰ و ۳۳۶ بوده است (جدول‌های ۱ و ۲).

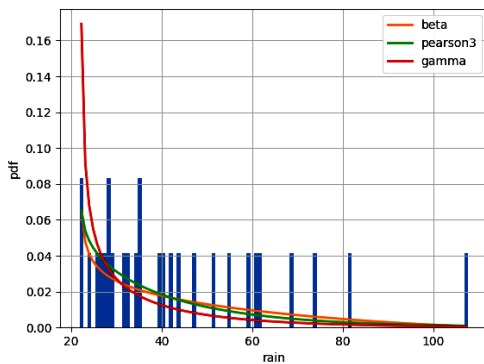
در ایستگاه‌های پارک ملی گلستان، کلاله، سد گلستان، قرقاق و گلیداغ نیز با توجه به یافته‌های پژوهش، بهترین تابع توزیع برازش شده تابع پیرسون ۳ می‌باشد (شکل‌های ۱۲ الی ۱۶) که با نتایج نیکزاد و همکاران (۲۵) و ملکی نژاد و همکاران (۲۰) همخوانی دارد. این تابع با کمترین میزان خطا و کمترین معیارهای AIC و BIC و همین‌طور بیشترین واگرایی مناسب‌ترین تابع انتخاب شد که نتایج آن در جدول‌های (۳ الی ۷) قابل مشاهده است. حداکثر بارش روزانه در ایستگاه پارک ملی گلستان مربوط به تاریخ ۲۰ مرداد (۰۱-۲۰-۱۱) با میزان ۴۵۰ میلی‌متر با دوره بازگشت ۸۵۴ سال است. در مقابل در ایستگاه‌های کلاله و قرقاق حداکثر بارش روزانه در تاریخ ۲۷ اسفند ۱۳۹۷ (۱۹-۰۳-۲۰۱۸) به ترتیب با بارش ۱۵۸/۱ و

حداکثر بارش روزانه در ماه‌های مختلف سال در ایستگاه‌های موردنظر بررسی شدند. همان‌طور که نتایج نشان می‌دهد حداکثر بارش روزانه در ایستگاه‌های کلاله، پارک ملی گلستان و سد گلستان اتفاق افتاده است (شکل ۹) که در مقایسه با داده‌های اندازه‌گیری شده ایستگاه‌های زمینی، مقدار بارندگی بسیار کمتر برآورد شده است. در داده‌های مشاهداتی حداکثر و حداقل میزان بارندگی روزانه در ایستگاه کلاله و ایستگاه قرقاق اتفاق افتاده است که به ترتیب ۱۵۸/۱ و ۶۰/۵ میلی‌متر اندازه‌گیری شده است در مقابل در داده ماهواره CHIRPS میزان حداکثر و حداقل بارندگی روزانه این ایستگاه‌ها به ترتیب برابر با ۵۰/۲ و ۴/۶ ارزیابی شد.

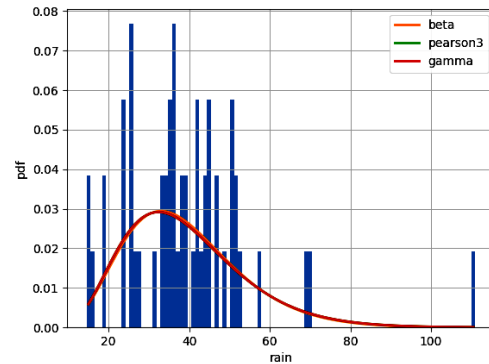
با توجه به نتایج حاصل از تصاویر ماهواره‌ای CHIRPS مشخص شد که در ۶ ماهه اول سال ۲۰۱۹ به خصوص در ماه‌های ژانویه، فوریه و مارس، بیشترین میزان بارندگی در منطقه اتفاق افتاده است به گونه‌ای که هر چه به سمت ماه‌های گرم‌تر سال پیش می‌رویم از میزان بارندگی در منطقه مورد مطالعه کاسته می‌شود. در مقابل میزان حداکثر و حداقل بارش روزانه در تصاویر CHIRPS به درستی برآورد نشده است که نشان‌دهنده عدم کارایی این تصاویر در بررسی بارش‌های روزانه است. در واقع داده‌های بارشی CHIRPS تنها برای تخمین بارندگی‌های ماهانه و سالانه کاربرد دارد. پژوهش‌های انجام شده در ایران در رابطه با بررسی عملکرد تخمین بارندگی به کمک تولیدات CHIRPS نسبت به داده‌های زمینی در طی بازه زمانی ۲۰۱۴-۲۰۰۴ نشان داد که این داده‌ها اغلب در ماه‌های همراه با بارندگی همرفتی و در نواحی پست بهترین عملکرد را از خود نشان می‌دهند (۲۹). سپس با استفاده از داده‌های بارش هواشناسی دوره زمانی ۱۹۹۸-۲۰۱۹ برای ایستگاه‌های منتخب گراف‌ها و آزمون نکویی برازش توسط زبان برنامه‌نویسی پایتون رسم شد. نمودار توزیع‌های مورد نظر برای هفت ایستگاه موجود در شکل‌های (۱۰ الی ۱۶) ارائه شده است. پس از بررسی توابع توزیع احتمال، نه تنها مناسب‌ترین توزیع بر مبنای مجموع مربعات خطا برای هر ایستگاه مشخص شد، بلکه



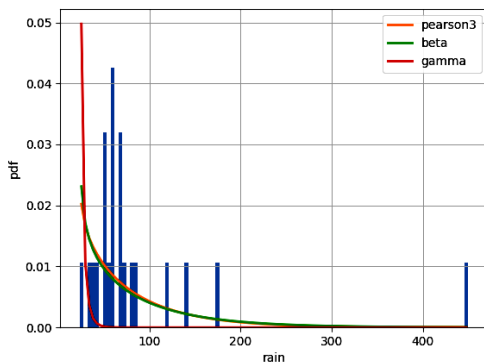
شکل ۹. داده‌های حداکثر بارش روزانه (ماه‌های مختلف سال ۲۰۱۹) در ایستگاه‌های منطقه مورد مطالعه (رنگی در نسخه الکترونیکی)



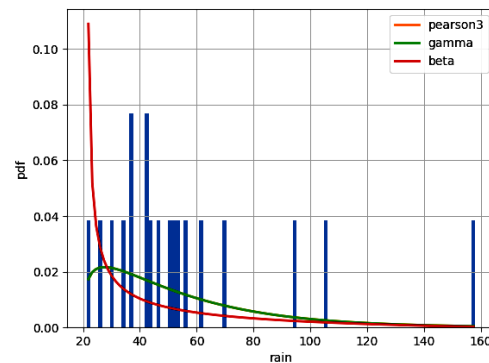
شکل ۱۱. نمودار توزیع آماری در ایستگاه تمر (رنگی در نسخه الکترونیکی)



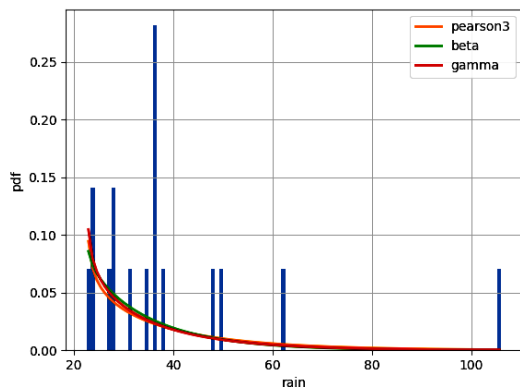
شکل ۱۰. نمودار توزیع آماری در ایستگاه گنبد (رنگی در نسخه الکترونیکی)



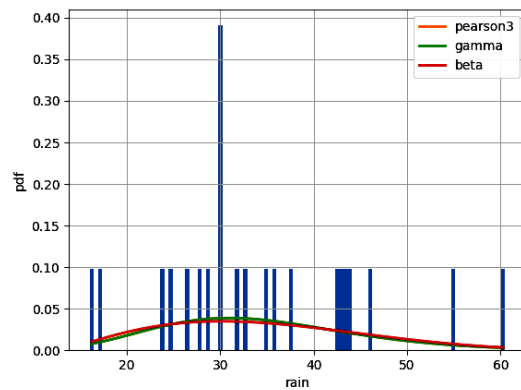
شکل ۱۳. نمودار توزیع آماری در ایستگاه کلانه (رنگی در نسخه الکترونیکی)



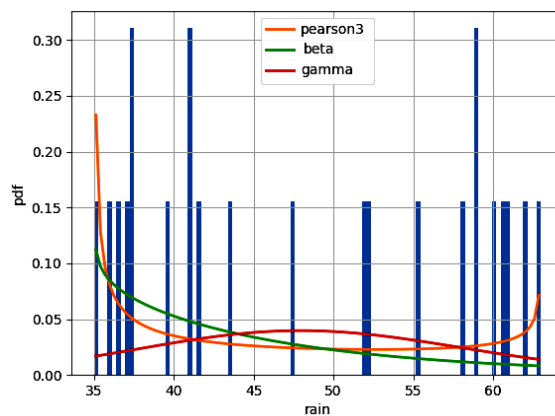
شکل ۱۲. نمودار توزیع آماری در ایستگاه پارک ملی گلستان (رنگی در نسخه الکترونیکی)



شکل ۱۵. نمودار توزیع آماری در ایستگاه قرناق (رنگی در نسخه الکترونیکی)



شکل ۱۴. نمودار توزیع آماری در ایستگاه سد گلستان (رنگی در نسخه الکترونیکی)



شکل ۱۶. نمودار توزیع آماری در ایستگاه گلیداغ (رنگی در نسخه الکترونیکی)

جدول ۱. نتایج آزمون نکویی مربوط به ایستگاه گنبد

ایستگاه گنبد	مجموع مربعات خطا	AIC	BIC	KI-div
بتا	۰/۰۳۳۱	۹۹۲/۹۵	-۱۷۵/۳۹	۰/۰۲۷۴
پیرسون نوع ۳	۰/۰۳۳۷	۱۰۲۷/۰۶	-۱۷۸/۲۱	۰/۰۲۶۶
گاما	۰/۰۴۹۴	۱۱۱۶/۱۸	-۱۶۷/۴۸	۰/۰۲۴۱

حداکثر بارش روزانه ایستگاه گنبد

تاریخ	بارش	احتمال	دوره بازگشت
۲۰۱۹-۰۳-۱۸	۱۰۷/۸	۰/۹۹۷	۴۲۰

جدول ۲. نتایج آزمون نکویی مربوط به ایستگاه تمر

KI-div	BIC	AIC	مجموع مربعات خطا	ایستگاه تمر
۰/۰۱۴۲۳	-۴۰۰/۷۱	۱۱۶۱/۸۴۴	۰/۰۲۴۰	بتا
۰/۰۱۵۳۲	-۴۰۴/۲۲	۱۱۴۱/۸۶۳	۰/۰۲۴۲	پیرسون نوع ۳
۰/۰۱۵۴۴	-۴۰۴/۲۲	۱۱۴۱/۸۶۲	۰/۰۲۴۲	گاما

حداکثر بارش روزانه ایستگاه تمر

دوره بازگشت	احتمال	بارش	تاریخ
۳۳۶	۰/۹۹۹	۱۱۱	۲۰۱۹-۰۳-۱۸

جدول ۳. نتایج آزمون نکویی مربوط به ایستگاه پارک ملی گلستان

KI-div	BIC	AIC	مجموع مربعات خطا	ایستگاه پارک ملی گلستان
۰/۲۵۱۴	-۱۸۳/۶۰	۱۵۲۵/۰۰	۰/۰۰۳۴	پیرسون نوع ۳
۰/۲۷۱۶	-۱۷۹/۵۵	۱۴۸۸/۶۰	۰/۰۰۳۵	بتا
۰/۲۹۹۲	-۱۶۹/۱۱	۷۴۸۷/۴۱	۰/۰۰۶۶	گاما

حداکثر بارش روزانه ایستگاه پارک ملی گلستان

دوره بازگشت	احتمال	بارش	تاریخ
۸۵۴	۰/۹۹۸	۴۵۰	۲۰۰۱-۰۸-۱۱

جدول ۴. نتایج آزمون نکویی مربوط به ایستگاه کلاله

KI-div	BIC	AIC	مجموع مربعات خطا	ایستگاه کلاله
۰/۱۳۷۹	-۱۱۸/۰۸۶	۱۱۱۳/۵۱	۰/۰۲۳۸۶	پیرسون نوع ۳
۰/۱۴۵۲	-۱۱۸/۰۸۶	۱۱۱۳/۵۵	۰/۰۲۳۸۶	گاما
۰/۱۹۳۳	-۱۱۸/۵۲۲	۱۱۸۹/۴۴	۰/۰۳۳۸۰	بتا

حداکثر بارش روزانه ایستگاه کلاله

دوره بازگشت	احتمال	بارش	تاریخ
۷۶	۰/۹۸۶	۱۵۸/۱	۲۰۱۹-۰۳-۱۸

جدول ۵. نتايج آزمون نكوبى مربوط به ايستگاه سد گلستان

KI-div	BIC	AIC	مجموع مربعات خطا	ايستگاه سد گلستان
۰/۲۰۰۳	-۷۴/۹۲	۱۰۶۷/۷۰	۰/۱۲۵۶	پيرسون نوع ۳
۰/۲۲۱۱	-۷۴/۷۷	۱۱۵۶/۲۶	۰/۱۲۶۰	بتا
۰/۲۵۹۷	-۷۲/۰۳	۱۱۶۰/۵۶	۰/۱۲۶۷	گاما
حداكثر بارش روزانه ايستگاه سد گلستان				
دوره بازگشت	احتمال	بارش	تاريخ	
۹۶	۰/۹۸۹	۱۰۶	۲۰۱۹-۰۳-۱۷	

جدول ۶. نتايج آزمون نكوبى مربوط به ايستگاه قرقان

KI-div	BIC	AIC	مجموع مربعات خطا	ايستگاه قرقان
۰/۰۲۴۲	-۹۲/۷۴۶	۸۱۶/۳۴	۰/۰۲۷۰۹	پيرسون نوع ۳
۰/۰۲۴۴	-۹۲/۷۴۴	۸۱۶/۳۳	۰/۰۲۷۰۹	بتا
۰/۰۲۹۹	-۸۹/۴۴۲	۸۲۴/۶۹	۰/۰۲۷۲۹	گاما
حداكثر بارش روزانه ايستگاه قرقان				
دوره بازگشت	احتمال	بارش	تاريخ	
۵۰	۰/۹۸۱	۶۰/۵	۲۰۱۹-۰۳-۱۸	

جدول ۷. نتايج آزمون نكوبى مربوط به ايستگاه گلیداغ

KI-div	BIC	AIC	مجموع مربعات خطا	ايستگاه گلیداغ
۰/۲۳۱	-۱۰۶/۶۰	۸۵۴/۹۷	۰/۱۲۹۳	پيرسون نوع ۳
۰/۲۵۶	-۱۰۹/۶۸	۸۷۱/۷۰	۰/۱۲۹۶	بتا
۰/۲۸۵	-۱۰۹/۶۹	۸۷۱/۷۰	۰/۱۲۹۶	گاما
حداكثر بارش روزانه ايستگاه گلیداغ				
دوره بازگشت	احتمال	بارش	تاريخ	
۱۲۲	۰/۹۸۱	۸۲/۵	۲۰۱۹-۰۳-۱۹	

داده‌های ماهواره‌ای نشان داد که ماهواره CHIRPS در تحلیل داده‌های سالانه و ماهانه نتایج خوبی از خود نشان داده است و نتایج حاصله به داده‌های زمینی و هواشناسی بسیار نزدیک است، زیرا همان‌طور که داده‌های هواشناسی نشان داد در سه‌ماهه اول سال بیشترین میزان بارندگی به وقوع پیوسته است. اما داده‌های ماهواره‌ای در تحلیل بارش روزانه کمترین دقت را برای انتخاب بهترین توزیع داشتند و به‌خوبی نتوانستند نتایج بارش روزانه را نشان دهند، به همین منظور از داده‌های واقعی بارش ایستگاه‌های هواشناسی برای انتخاب بهترین توزیع آماری استفاده شد. پس از انجام تحلیل‌های آماری مناسب، عملکرد سه توزیع پرکاربرد از جمله پیرسون ۳، گاما و بتا برای توصیف بارندگی روزانه ارزیابی شد. برای برآزش این سه توزیع احتمال از آزمون مجموع مربعات خطا برای محاسبه عددی پارامترها و هم‌زمان شناسایی بهترین توزیع برآزش استفاده شد. با توجه به اینکه در بین متخصصین آمار توافقی در مورد استفاده از یک تابع توزیع خاص وجود ندارد، از این‌رو می‌توان توزیعی که نتایج آن دارای کمترین خطا و نزدیک‌تر به واقعیت است را به‌عنوان مناسب‌ترین تابع توزیع انتخاب کرد. نتایج حاصل از آزمون نکویی برآزش داده‌های هواشناسی نشان داد که مناسب‌ترین تابع توزیع احتمالی در هر یک از آزمون‌ها متفاوت است. همچنین، نتایج محاسبه دوره‌های بازگشت (۲، ۱۰، ۵۰، ۱۰۰ و ۲۰۰ ساله) و مقادیر بارندگی در ایستگاه‌های مختلف نیز نشان داد که تفاوت مقادیر بارش قابل توجه است، چراکه برای دوره‌های بازگشت بالا میزان بارندگی افزایش پیدا می‌کند. از طرف دیگر آزمون مجموع مربعات خطا نیز نتایج قابل قبولی را ارائه داد. این آزمون دارای دو مزیت اصلی است: نه‌تنها امکان مقایسه و تعریف بهترین توزیع برآزش و پارامترهای آن را فراهم می‌کند بلکه می‌تواند تا حد زیادی به متخصصان در هنگام برخورد با تعیین بهترین توزیع مناسب کمک کند. از آنجاکه هدف از این پژوهش ارائه بهترین تابع توزیع است، نتایج حاصله نشان داد که تابع پیرسون ۳ توزیع مناسب‌تری نسبت به توابع بتا و گاما در ایستگاه‌های منطقه

۶۰/۵ میلی‌متر و دوره بازگشت ۷۶ و ۵۰ سال اتفاق افتاده در ایستگاه سد گلستان نیز در تاریخ ۲۶ اسفند ۱۳۹۷ (۱۷-۰۳-۲۰۱۹) با بارش ۱۰۶ میلی‌متر و دوره بازگشت ۹۶ سال و در ایستگاه گلیداغ بارش ۸۲/۵ میلی‌متری با دوره بازگشت ۱۲۲ سال در تاریخ ۲۸ اسفند ۱۳۹۷ (۱۹-۰۳-۲۰۱۹) به وقوع پیوسته است.

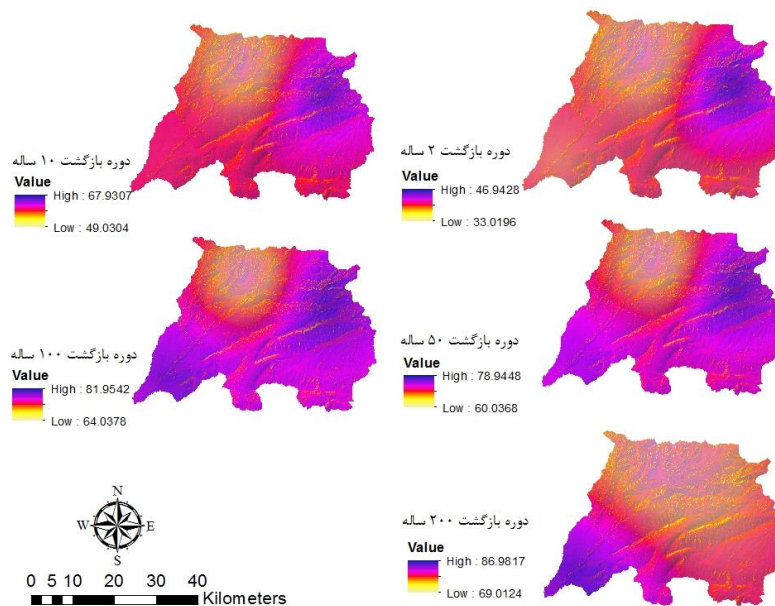
با توجه به نتایج تحلیل توابع مختلف در جداول فوق به تفکیک در هر ایستگاه و انتخاب بهترین توزیع از طریق آزمون نکویی برآزش به روش مجموع مربعات خطا، مقادیر حداکثر بارش با دوره‌های بازگشت (۲، ۱۰، ۵۰، ۱۰۰ و ۲۰۰) از طریق مناسب‌ترین توزیع در هر ایستگاه برآورد شد که نتایج آن در جدول ۸ و شکل ۱۷ ارائه شده است. همین‌طور نقشه پهنه‌بندی حداکثر بارش روزانه در دوره‌های بازگشت مورد نظر با استفاده از روش کریجینگ معمولی به‌دست آمد که نتایج آن در شکل ۱۸ ارزیابی شد. با توجه به نقشه پهنه‌بندی بارش حداکثر روزانه، در دوره‌های بازگشت مختلف مشخص شد که بیشترین بارش روزانه در دوره بازگشت ۲ و ۱۰ ساله در بخش شرقی منطقه در نواحی ایستگاه گلیداغ اتفاق افتاده است. همین‌طور در دوره‌های بازگشت ۵۰ و ۱۰۰ ساله نیز بیشترین بارش حداکثر روزانه، نواحی شرقی و جنوبی منطقه (ایستگاه گلیداغ و تمر) را در برگرفته است. در نهایت در دوره بازگشت ۲۰۰ ساله نیز قسمت جنوبی منطقه بیشترین بارش روزانه و قسمت مرکزی حداقل بارش روزانه را تجربه کرده‌اند. همان‌طور که از نقشه‌های درون‌یابی پیداست، تغییر در دوره‌های بازگشت بر الگوی مکانی بارش هرچند به صورت جزئی تأثیر داشته است به‌ویژه اینکه همه منطقه در شرایط اقلیمی مشابه و همگنی قرار گرفته است. به‌طور کلی با افزایش دوره بازگشت، مقادیر حداکثر بارش روزانه افزایش یافته است.

نتیجه‌گیری

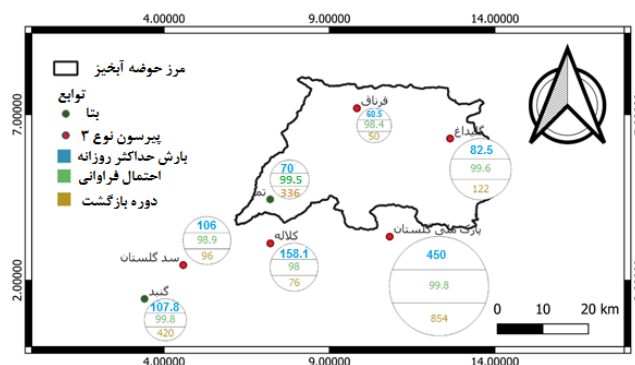
تجزیه و تحلیل داده‌های بارش روزانه، موضوعی مورد علاقه در بسیاری از زمینه‌های هیدرولوژیکی است. نتایج تجزیه و تحلیل

جدول ۸. برآورد بارش حداکثر روزانه با دوره‌های بازگشت مختلف با استفاده از توزیع‌های منتخب به تفکیک ایستگاه

ایستگاه	دوره بازگشت (احتمال وقوع)				
	۲۰۰ (۰/۹۹۵)	۱۰۰ (۰/۹۹)	۵۰ (۰/۹۸)	۱۰ (۰/۹)	۲ (۰/۵)
گنبد	۱۰۴	۱۰۰	۹۵	۷۴	۳۹
تمر	۸۷	۸۱	۷۵	۵۸	۳۷
پارک ملی گلستان	۳۵۵	۳۱۰	۲۶۵	۱۶۲	۶۱
کلاله	۱۸۶	۱۶۶	۱۴۶	۹۷	۴۷
سد گلستان	۱۲۱	۱۰۷	۹۳	۶۲	۳۲
قرناق	۶۹	۶۴	۶۰	۴۹	۳۳
گلیداغ	۸۴	۸۲	۷۹	۶۸	۴۷



شکل ۱۷. نقشه پهنه‌بندی دوره‌های بازگشت ۲، ۱۰، ۵۰، ۱۰۰ و ۲۰۰ ساله در منطقه مورد مطالعه (رنگی در نسخه الکترونیکی)



شکل ۱۸. نقشه بهترین توابع توزیعی با دوره‌های بازگشت اندازه‌گیری شده برای ایستگاه‌های هواشناسی منطقه مورد مطالعه (رنگی در نسخه الکترونیکی)

مدیران و مسئولان ذی‌ربط با استفاده از این نتایج از وقوع سیلاب‌های مخرب در این مناطق سیل‌خیز در مدیریت بحران استفاده کنند. چراکه در کشورهای پیشرفته مدیریت بر مبنای خطر و استفاده از نتایج پژوهش‌ها و شناخت علمی پدیده‌هایی همچون بارش صورت می‌گیرد. با توجه به نتایج این پژوهش که تداوم وقوع بارش روزانه باعث وقوع سیل شده است، بنابراین انتظار می‌رود در آینده با توجه به نتایج به‌دست آمده، سیلاب‌های شدیدتری در منطقه به وقوع بپیوندد. بنابراین، توصیه می‌شود با در نظر گرفتن بارش‌های حداکثر محتمل و در پی آن وقوع سیلاب‌های محتمل با مدیریت صحیح از خسارات جانی و مالی در مناطق مستعد به‌ویژه منطقه مورد مطالعه جلوگیری به عمل آید.

سپاسگزاری

این مقاله از نتایج رساله دکتری در قالب یک طرح پژوهشی با شناسه طرح ۹۹۰۲۷۵۳۲ استخراج شده است؛ بنابراین، نویسندگان این مقاله از همکاری‌های ارزشمند بنیاد ملی علم ایران (صندوق حمایت از پژوهشگران کشور) جهت حمایت مالی در اجرای این طرح کمال تشکر و قدردانی را دارند.

مورد مطالعه است. همچنین، برای برآورد مقادیر حداکثر بارش روزانه در دوره‌های بازگشت مختلف با این تابع توزیع صورت گیرد. با بررسی میزان خطای مجموع مربعات و معیارهای آکائیک، بیزین و واگرایی لیبیل در برآورد حداکثر بارش روزانه در دوره بازگشت‌های مختلف در ایستگاه‌های منطقه مشخص شد که بارش مربوط به ایستگاه‌های زمینی در منطقه مورد مطالعه با وجود داشتن ویژگی‌های مشترک تا حدودی از توزیع‌های آماری متفاوتی پیروی می‌کنند. به همین دلیل است که بارش‌های شدید و نامنظم از نظر زمانی و مکانی می‌تواند در انتخاب مناسب‌ترین تابع توزیع آماری در هر ایستگاه مؤثر باشد. بدین منظور حداکثر بارش روزانه به‌عنوان یک متغیر مستقل، به دلیل اهمیت آن در بحث مخاطرات سیل در این پژوهش مورد ارزیابی قرار گرفت. بنابراین، به خاطر وقوع سیل در منطقه مورد مطالعه که به دلیل افزایش بارش‌های روزانه اتفاق افتاد و خسارات فراوانی از جمله تخریب پل‌ها، جاده‌ها، زمین‌های کشاورزی، بندها، لوله‌های آب و گاز و منازل مسکونی به بار آورد، باید توجه بیشتری به چنین بارش‌های سیل‌آسایی شود. به همین دلیل تحلیل فراوانی بارش از روش‌های مؤثر برای برآورد میزان سیلاب محسوب می‌شود. چراکه شرایط لازم، از جمله آمار در دسترس و امکان برآزش توابع توزیع مناسب، می‌تواند نتایج مطلوبی را ارائه دهد، تا

منابع مورد استفاده

1. Aho, K., D. Derryberry and T. Peterson. 2014. Model selection for ecologists: The world views of AIC and BIC. *Journal of Ecology* 95(3):631-636.
2. Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. *In: Second International Symposium on Information Theory*. Akademiai Kiado, Budapest.
3. Akbari, G. H., M. SHadmehr and M. Reihani. 2021. Statistical processing the best distribution of floods with different return periods using the commercial code SAMADA and Easyfit (study of Atrak River Watershed Basin). *In: 1st National Conference on Civil Engineering, Intelligent Development and Sustainable Systems*, Golestan university, Gorgan, Iran (In Farsi).
4. Alibakhshi, S. M., A. Farid Hossini, K. Davari, A. Alizadeh and H. Munyka. 2018. Assessment of Ground Station, GPM Satellite and MERRA Precipitation Products in Kashafrud Basin. *Journal of Watershed Management Research* 18(9):111-122. (In Farsi).
5. Arjomand, M., N. Sadeghian Pirmahaleh and S. Amini. 2020. Select the best statistical distribution in estimating the rainfall return period using Smada software (Case study: Karaj city). *Journal of Civil and Project* 2(1): 11-17 (In Farsi).

6. Burnham, K. P., D. R. Anderson and K. P. Huyvaert. 2011. AIC model selection and multi model inference in behavioral ecology: some background, observations, and comparisons. *Journal of Behavioral Ecology and Sociobiology* 65(1):23-35.
7. CSA, 2019. Technical guide: development, interpretation and use of rainfall intensity-duration-frequency (IDF) information. Guideline for Canadian Water Resources Practitioners. Canadian Standards Association, Toronto, Ontario, Canada.
8. Davarizadeh, A. and M. Vafakhah. 2019. Determining the maximum daily rainfall by using the best of regional possible distribution in the Imameh watershed. *In: 9th National conference on Watershed Management of Water and Soil Resources, Shahid Bahonar University of Kerman, Kerman, Iran (In Farsi)*.
9. De Carvalho, J. R. P., E. D. Assad, A. F. de Oliveira and H. S. Pinto. 2014. Annual maximum daily rainfall trends in the Midwest, southeast and southern Brazil in the last 71 years. *Weather and Climate Extremes* 5(6): 7-15.
10. Duan, Z., J. Liu, Y. Tuo, G. Chiogna and M. Disse. 2016. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. *Journal of Science of the Total Environment* 573:1536-1553
11. Farzandi, M. and H. Rezaee Pajand. 2017. Comparison of goodness of fit of probably distributions with the increase in sample size of the case study: 119 year Mashhad rainfall. *Journal of Climate Research* 8(31-32):123-136 (In Farsi).
12. Fofana, M., J. Adoukpe, I. Larbi, J. Hounkpe, H. D. N. Koubodana, A. Toure and A. M. Limantol. 2022. Urban flash flood and extreme rainfall events trend analysis in Bamako, Mali. *Environmental Challenges* (6):100449.
13. Funk, C.C., P.J. Peterson, M.F. Landsfeld, D.H. Pedreros, J.P. Verdin, J.D. Rowland, B.E. Romero, G.J. Husak, J.C. Michaelsen, A.P. Verdin and P. Pedreros. 2014. A quasi-global precipitation time series for drought monitoring. US Geological Survey Data Series 832.
14. Jahangiri Pirmorad, F., H. Rezaee and K. Khalili. 2019. Comparison distribution methods of peak discharge statistical with emphasis on 5 parameter Wakeby. *In: 5th National Conference of Irrigation and Drainage of Iran, Birjand University, Birjand, Iran (In Farsi)*.
15. Katsanos, D., A. Retalis and S. Michaelides. 2016. Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. *Atmospheric Research* 169: 459-464.
16. Kidd, C. and G. Huffman. 2011. Global precipitation measurement. *Journal of Applied Meteorology and Climatology* 18: 334-353
17. Komaki, C. B., M. Afzalzadeh, and H. Najafi. 2020. Python Scripting Language for Environmental Sciences. Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran (In Farsi).
18. Kumar, V., S. K. Jain and Y. Singh. 2010. Analysis of long-term rainfall trends in India. *Journal of Hydrological and Sciences* 55 (4): 484-496.
19. Limsakul, A. and P. Singhruck. 2016. Long-term trends and variability of total and extreme precipitation in Thailand. *Journal of Atmospheric Research* 169: 301-317.
20. Malekinejad, H., M. Soleimani Motlagh, A. Jaidari and S. Shater Abshori. 2015. Comparison of the limited and minor amounts of daily maximum rainfall to determine the best probabilities distribution (case study of Tehran and Alborz provinces). *Journal of Nivar* 90(39): 3-14 (In Farsi).
21. Moccia, B., C. Mineo, E. Ridolfi, F. Russo and F. Napolitano. 2021. Probability distributions of daily rainfall extremes in Lazio and Sicily, Italy, and design rainfall inferences. *Journal of Hydrology: Regional Studies* 33:100771.
22. Mozafari, G. A., K. Omidvar and N. Hassani Sadi. 2019. Review of return period heavy rainfall in Kerman province (Case Study: Kerman). *In: 4th International Conference of Agriculture Development, Natural Resources, Environment and Tourism of Iran, Tabriz, Iran*.
23. Nash, J. E. and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models I: a discussion of principles. *Journal of Hydrology* 10: 282-290.
24. Nguyen, T. H. and V. T. V. Nguyen. 2019. Decision-support tool for constructing robust rainfall IDF relations in consideration of model uncertainty. *Journal of Hydrological Engineering* 24 (7): 06019004.
25. Nikzad, M., M. Rad and S. Mirzaee. 2016. Frequency analysis of rainfall maximum of daily in North Karun area. *In: 11th National Conference on Watershed Management and Engineering of Iran. Iran Watershed Management Association, Yasuj University, Yasuj, Iran (In Farsi)*.
26. Pham, T. D., K. Yoshino and D.T. Bui. 2017. Biomass estimation of *Sonneratia caseolaris* (L.) Engler at a coastal area of Hai Phong city (Vietnam) using ALOS-2 PALSAR imagery and GIS-based multi-layer perceptron neural networks. *Journal of GIS Science and Remote Sensing* 54: 329-353
27. Phuong, D.N.D., V. T. Linh, T. T. Nhat, H. M. Dung and N. K. Loi. 2019. Spatiotemporal variability of annual and seasonal rainfall time series in Ho Chi Minh city, Vietnam. *Journal of Water and Climate Change* 10 (3): 658-670.

28. Rahmati, A. and A. R. Massah Bavani. 2019. Comparative evaluation of global rainfall datasets with observation rainfall value (Case Study: Karoun Basin). *Journal of Iran Water Resources Research* 15(1): 178-192 (In Farsi).
29. Saeidizand, R., S. Sabetghadam, E. Tarnavsky and A. Pierleoni. 2018. Evaluation of CHIRPS rainfall estimates over Iran. *Quarterly Journal of the Royal Meteorological Society* 144: 282-291.
30. Sunyer, M. A., I. B. Gregersen, D. Rosbjerg, H. Madsen, J. Luchner and K. Arnbjerg-Nielsen. 2015. Comparison of different statistical downscaling methods to estimate changes in hourly extreme precipitation using RCM projections from ENSEMBLES. *Journal of Climatology* 35: 2528-2539.
31. Symonds, M. R. and A. Moussalli. 2011. A brief guide to model selection, multi model inference and model averaging in behavioral ecology using Akaike's information criterion. *Journal of Behavioral Ecology and Sociobiology* 65(1):13-21.
32. Tafteh, A., S. Malah and N. A. Ebrahimi Pak. 2020. Evaluation of daily, decade and monthly data satellite images to estimate of precipitation using Google Earth engine in Khuzestan Province. *Journal of Water and Soil Resources Conservation* 9(3): 93-104 (In Farsi).
33. Vrieze, S. I. 2012. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Journal of Psychology and Methods* 17: 228-243.
34. Westra, S., H. Fowler, J. Evans, L. Alexander, P. Berg, F. Johnson, E. Kendon, G. Lenderink and N. Roberts. 2014. Future changes to the intensity and frequency of short-duration extreme rainfall. *Journal of Geophysics* 52: 522-555.
35. Xie, L., J. Zeng, U. Kruger, X. Wang and J. Geluk. 2015. Fault detection in dynamic systems using the Kullback-Leibler divergence. *Journal of Control Engineering Practice* 43:39-48.

Frequency Analysis of Maximum Daily Rainfall and Determining the Best Appropriate Distribution Functions in the Bostan Watershed, Golestan Province

S. Yaghobi¹, C. B. Komaki^{1*}, M. Hosseinalizadeh¹, A. Najafinejad²,
H. R. Pourghasemi³ and M. Faramarzi⁴

(Received: June 23-2022 ; Accepted: August 28-2022)

Abstract

Frequency analysis of daily rainfall or return period of rainfall and flooding events is very important considering the behavioral complexity in water resources management; because ignoring it can lead to urban destructive floods. In the present research, three distribution functions of Pearson, Beta, and Gamma were compared to investigate and select the most appropriate distribution function for the precipitation data acquired from meteorology stations and CHIRPS satellite in seven stations in the watershed of Bustan Dam. Statistical analyses showed that satellite data were ineffective to estimate daily precipitation due to high errors in RMSE, MAD, and NASH. Meteorological data were used to spot the best distribution. Google Earth Engine and Python programming language were used. Then, the selected distribution function was used to determine the maximum daily rainfall, frequency probability, and return period of 2, 10, 50, 100, and 200 years. The results of the goodness of fit test, Error Sum of Squares, Bayesian Information Criterion, Akaike Information Criteria well as Kullback-Leibler Divergence showed that in five stations of Kalaleh, Qarnaq, Golestan National Park, Golestan Dam, and Glidagh, the Pearson function is the most suitable distribution function. Also, in the other two stations (Gonbad and Tamar), the Beta function was recognized as a suitable function. However, Gamma distribution in the study area is not efficient. So, it can be concluded that heavy and irregular rainfall can be effective in choosing the best distribution function at each station. Therefore, it is recommended to consider the maximum possible rainfall and as a result of the possible occurrence of floods with principled and accurate management to prevent human and financial losses in susceptible areas, especially in the study area.

Keywords: Maximum daily rainfall, Frequency analysis, Statistical distribution, Meteorological and satellite data, Bostan watershed

1. Department of Desert Zone Management, Faculty of Rangeland and Watershed Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran.

2. Department of Rangeland Management, Faculty of Rangeland and Watershed Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran.

3. Department of Natural Resources and Environmental Engineering, Shiraz University, Shiraz, Iran.

4. Department of Range and Watershed Management, Ilam University, Ilam, Iran.

*: Corresponding author, Email: komaki@gau.ac.ir, bkomaki@gmail.com