مقایسه‌ای بر روش‌های آماری کنترل کیفیت داده‌های رواناب؛ آزمون‌های تشخیص داده‌های پرت و بررسی همگنی

پویا اللهویردی‌پور[۱*]، یعقوب دین‌پژوه[۲]

۱- دانشجوی دکتری تخصصی، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

۲- استاد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

*آدرس رایانامه نویسنده رابط (a.pouya1996@gmail.com)

خلاصه

کنترل کیفیت داده‌های رواناب برای مدل‌سازی دقیق هیدرولوژیکی ضروری است. این مطالعه به بررسی کیفیت داده‌های رواناب رودخانه آجی‌چای با استفاده از آزمون‌های آماری می‌پردازد. تحلیل داده‌های رواناب با استفاده از آزمون‌های Grubbs و Dixon برای تشخیص داده‌های پرت و آزمون‌های SNHT، Buishand و Pettitt برای بررسی همگنی انجام شد. آزمون‌های Grubbs و Dixon حدود ۴/۳ درصد از داده‌ها را به‌عنوان پرت تشخیص دادند (سطح ۹۵ درصد). آزمون همگنی Pettitt نقطه تغییر را در تاریخ ۷ فروردین ۱۳۹۵ با میانگین قبل و بعد به‌ترتیب ۰/۹۹۴ و ۰/۴۳۷ مترمکعب بر ثانیه، SNHT نقطه تغییر را در ۲۹ خرداد ۱۳۹۲ با میانگین قبل و بعد به‌ترتیب ۱/۰۵۴ و ۰/۵۷۹ مترمکعب بر ثانیه و Buishand نقطه تغییر را در ۱۴ خرداد ۱۳۹۲ با میانگین قبل و بعد به‌ترتیب ۱/۰۵۴ و ۰/۵۸۱ مترمکعب بر ثانیه شناسایی کردند. این نتایج بر لزوم کنترل کیفیت داده‌ها برای بهبود مدل‌سازی هیدرولوژیکی و مدیریت منابع آب تأکید دارند.

کلمات کلیدی: داده‌های پرت، همگنی، رواناب، هیدرولوژی، آزمون‌های آماری.

۱. مقدمه

رواناب رودخانه، به عنوان جریان آب سطحی ناشی از بارش و ذوب برف، نقش کلیدی در چرخه آب، تأمین آب شرب، کشاورزی و تولید انرژی دارد. با این حال، در دهه‌های اخیر، با تشدید تغییرات اقلیم، افزایش رویدادهای حدی مانند سیلاب‌های ناگهانی و خشکسالی‌های طولانی، کیفیت داده‌های رواناب بیش از پیش مورد توجه قرار گرفته است [۱]. کنترل کیفیت داده‌های هیدرولوژیکی و بررسی همگنی سری‌های زمانی رواناب رودخانه‌ها پیش‌نیاز اساسی برای هرگونه تحلیل آماری، مدل‌سازی سیلاب و طراحی سازه‌های هیدرولیکی است. داده‌های نامناسب یا حاوی مقادیر پرت می‌تواند برآورد پارامترهای آماری، نتایج آزمون‌های همگنی و در پی آن، طراحی‌ها و تصمیم‌گیری‌های مدیریتی را به‌طور چشمگیری مغشوش سازد؛ بنابراین شناسایی و برخورد مناسب با داده‌های پرت و تأیید یا رد فرض همگنی سری زمانی، از گام‌های کلیدی در پژوهش‌های هیدرولوژی و مهندسی آب است. افزایش فراوانی و شدت رویدادهای حدی (مانند سیلاب‌های ناگهانی) ، هم نیاز به داده‌های با کیفیتِ بیشتر و هم حساسیت تحلیل‌ها نسبت به مقادیر پرت را بالا برده است [۲].

پیامدهای ناشی از تصمیم‌گیری مبتنی بر داده‌های نادرست می‌تواند به خسارات اقتصادی و زیست‌محیطی جدی منجر شود؛ بنابراین توسعه و مقایسه روش‌های موثر کنترل کیفیت برای اطمینان از قابل اتکا بودن نتایج ضروری است. سیلاب‌های مخرب ایران مانند سیل سال ۱۳۹۸ در استان‌های شمالی و جنوبی، اهمیت دقت در داده‌های رواناب را برجسته کرده و نشان داده که داده‌های نادرست می‌تواند مدل‌سازی هیدرولوژیکی را مختل کند [۳].

در مطالعات کنترل کیفیت، تشخیص داده‌های پرت مرحله نخست است. داده‌های پرت، مقادیری هستند که به طور غیرعادی از سایر داده‌ها فاصله دارند و اغلب ناشی از خطاهای اندازه‌گیری، نقص تجهیزات هیدرومتری، یا رویدادهای طبیعی نادر مانند طوفان‌های شدید هستند. اگر این داده‌ها شناسایی و اصلاح نشوند، می‌توانند تحلیل‌های آماری را تحریف کرده و پیش‌بینی‌های هیدرولوژیکی را نامعتبر سازند. آزمون‌های آماری مختلفی برای

شناسایی مقادیر غیرعادی پیشنهاد شده‌اند که از میان آن‌ها آزمون Grubbs و آزمون Dixon به‌دلیل سادگی، کارایی در نمونه‌های کوچک تا متوسط و کاربرد گسترده در ادبیات هیدرولوژی، اغلب استفاده می‌شوند. آزمون Grubbs مبتنی بر بررسی فاصله بزرگترین (یا کوچکترین) مشاهده نسبت به میانگین داده‌ها و انحراف معیار آن است و برای شناسایی یک یا چند مقدار پرت در توزیع نرمال مناسب است؛ در حالی که آزمون Dixon با تکیه بر نسبت اختلافات ترتیبی در نمونه‌های مرتب‌شده، به‌ویژه در نمونه‌های کوچک، توانایی شناسایی مقادیر پرت تکی یا چندتایی را دارد. انتخاب بین این آزمون‌ها بستگی به اندازه نمونه، فرضیات توزیع و هدف تحلیل (حذف یا پر کردن مقادیر پرت) دارد، و بررسی تطبیقی عملکرد آن‌ها در داده‌های رواناب که معمولاً توزیع‌های چولگی دارند، از اهمیت ویژه‌ای برخوردار است [۴].

پس از برخورد با داده‌های پرت، گام اساسی بعدی بررسی همگنی سری زمانی است. ناهمگنی در سری‌های زمانی رواناب، به معنای وجود تغییرات ساختاری یا جهش‌های ناگهانی در داده‌ها است که ممکن است به دلیل جابجایی ایستگاه‌های اندازه‌گیری، تغییرات کاربری اراضی، یا تأثیرات انسانی مانند ساخت سدها ایجاد شود. کنترل کیفیت داده‌ها شامل مراحل تشخیص، اصلاح و اعتباربخشی است تا اطمینان حاصل شود که سری‌های زمانی برای مدل‌سازی‌های آینده مناسب هستند [۵]. وجود تغییرات ناگهانی یا روندهای ساختاری در داده‌ها می‌تواند فرض همگنی را نقض کند و کاربرد مدل‌ها و پارامترهای آماری را نامعتبر سازد. برای ارزیابی همگنی، آزمون‌های متعددی وجود دارند که هر یک به جنبه‌ای متفاوت از تغییرپذیری زمانی توجه دارند. آزمون Standard Normal Homogeneity Test (SNHT) به‌طور گسترده برای شناسایی شکست‌های ناگهانی در میانگین استفاده می‌شود و حساسیت مناسبی نسبت به تغییرات گسسته دارد. آزمون Buishand بر مبنای دامنه تجمعی انحرافات از میانگین، برای تشخیص جابه‌جایی‌های سطح پایه مفید است. آزمون Pettitt به‌عنوان یک آزمون ناپارامتریک، بدون نیاز به فرض نرمال بودن، تغییر نقطه‌ای در میانه سری را شناسایی می‌کند و برای داده‌های هیدرولوژیکی که اغلب از توزیع‌های نرمال تبعیت نمی‌کنند مناسب است. ترکیب این آزمون‌ها می‌تواند تصویری جامع‌تر از انواع ناپیوستگی‌ها و روندهای موجود در سری رواناب ارائه دهد [۶-۴].

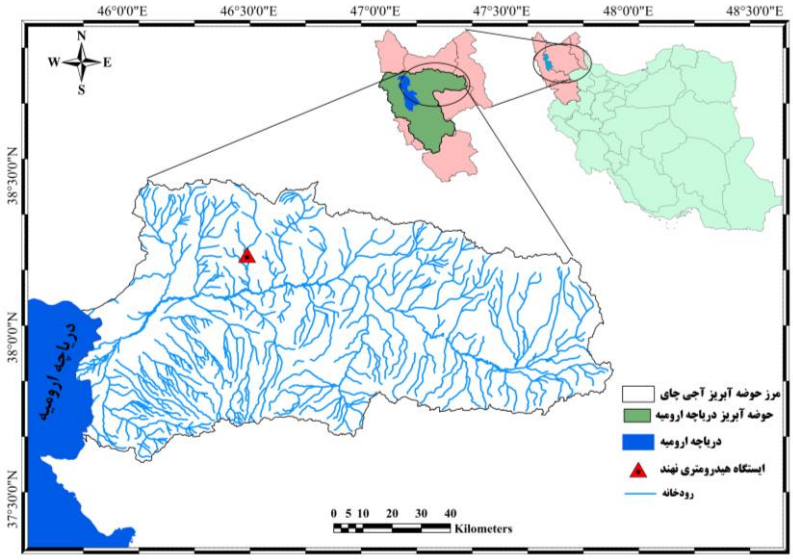
محققان متعددی در جهان و ایران موضوع کنترل کیفیت داده‌های هیدرولوژیکی را بررسی کرده‌اند. از جمله می‌توان تحقیقات بیکیجی و همکاران در ترکیه [۷]، کوشیس و همکاران در مجارستان [۵]، وو و همکاران در چین [۶]، کابیلاش و همکاران در هند و بیلفرنیچ و همکاران در کشورهای بورکینافاسو، غنا، بنین و توگو [۸] را نام برد. همچنین در ایران نیز می‌توان به تحقیقات اللهویردی پور و ستاری در اهواز [۴]، رستم پور در استان خراسان جنوبی [۹] و شیرمحمدی و همکاران در استان‌های چهارمحال و بختیاری و خوزستان [۳] اشاره کرد.

با توجه به بررسی پیشینه تحقیق و اهمیت کنترل کیفیت در داده‌های رواناب رودخانه، به‌ویژه در شرایط تغییرات اقلیم که داده‌ها را ناپایدارتر می‌کند، هدف از این مطالعه مقایسه روش‌های آماری کنترل کیفیت داده‌های رواناب است. تمرکز بر آزمون‌های Grubbs و Dixon برای تشخیص داده‌های پرت و آزمون‌های SNHT، Buishand و Pettitt برای بررسی همگنی خواهد بود. این مقایسه بر اساس داده‌های واقعی از حوضه رودخانه آجی‌چای در شمال‌غرب ایران انجام می‌شود تا کارایی، حساسیت و محدودیت‌های هر روش ارزیابی شود. نتایج این پژوهش برای کاربردهای عملی در پایش جریان رودخانه‌ها، تحلیل روندهای بلندمدت، مدیریت پایدار منابع آب، مدل‌سازی سیلاب و طراحی سازه‌های مهندسی آب می‌تواند راهگشا باشد.

۲. مواد و روش‌ها

۱.۲. منطقه مورد مطالعه و داده‌های مورد استفاده

حوضه آبریز آجی‌چای، یکی از زیرحوضه‌های دریاچه ارومیه در شمال‌غرب ایران و استان آذربایجان شرقی، در محدوده جغرافیایی '۳۰ °۴۵ تا '۴۵ °۴۷ طول شرقی و '۴۵ °۳۷ تا '۳۰ °۳۸ عرض شمالی قرار دارد. این حوضه با مساحت ۱۱۴۵۷ کیلومترمربع، مرتفع‌ترین نقطه ۳۸۵۰ متر و پست‌ترین نقطه ۱۲۷۴ متر از سطح دریا، در غرب دریاچه ارومیه واقع شده است (شکل ۱). بر اساس طبقه‌بندی دومارتن، اقلیم این منطقه نیمه‌خشک است. میانگین رواناب ایستگاه هیدرومتری نهند واقع در این حوضه ۰/۹۳۵ مترمکعب بر ثانیه است. در این مطالعه از داده‌های رواناب این ایستگاه در یک دوره ۲۱ ساله (۱۳۷۶-۱۳۹۷) استفاده شده است.



شکل ۱ – موقعیت منطقه مورد مطالعه، حوضه آبریز آجی‌چای در شمال‌غرب ایران

۲.۲. آزمون‌های شناسایی داده‌های پرت

۱.۲.۲. آزمون Grubbs

آزمون Grubbs که توسط فرانک گرابز در دهه ۱۹۵۰ معرفی شد، یک آزمون پارامتریک است که فرض می‌کند داده‌ها از توزیع نرمال پیروی می‌کنند. اساس آزمون بر مقایسه دوری بیشترین (یا کمترین) مشاهده از میانگین نمونه نسبت به انحراف معیار نمونه است. فرض صفر آزمون گرابز این است که داده‌ها فاقد مقدار پرت هستند (تمام مشاهدات از یک توزیع نرمال یکسان نشأت می‌گیرند) و فرض مقابل آن، حداقل یک مشاهده پرت وجود دارد. این آزمون می‌تواند به‌صورت یک‌طرفه (برای شناسایی یک پرت در یک انتهای توزیع) یا دوطرفه (برای شناسایی پرت در هر دو انتها) اجرا شود. در هیدرولوژی، آزمون گرابز برای شناسایی مقادیر غیرعادی در داده‌های دبی، بارش یا سایر متغیرهای هیدرولوژیکی کاربرد دارد. برای مثال، یک مقدار دبی غیرعادی ناشی از نقص دستگاه هیدرومتری می‌تواند با این آزمون تشخیص داده شود [۱۰].

۲.۲.۲. آزمون Dixon

آزمون Dixon که به آزمون Q نیز شناخته می‌شود، در سال ۱۹۵۳ توسط Dixon ارائه شد. این آزمون ساده و غیرپارامتریک، بر پایه نسبت فاصله داده مشکوک به نزدیک‌ترین داده‌ها عمل می‌کند و در داده‌های هیدرولوژیکی کاربرد وسیعی دارد. هدف اصلی آن تشخیص یک مقدار پرت منفرد در هر دو انتهای توزیع (بزرگترین یا کوچکترین مقدار) است. این آزمون در هیدرولوژی برای داده‌هایی با توزیع نامعلوم یا مجموعه‌های کوچک، مانند داده‌های دبی در ایستگاه‌های هیدرومتری با سوابق محدود، بسیار کاربردی است. مزایای آزمون دیکسون شامل سادگی، عدم وابستگی به توزیع داده‌ها و مناسب بودن برای نمونه‌های کوچک است. هر دو آزمون Grubbs و Dixon در استانداردهای بین‌المللی مانند ISO برای کنترل کیفیت داده‌های محیطی توصیه شده‌اند. ترکیب این دو آزمون، با توجه به ویژگی‌های پارامتریک و غیرپارامتریک آن‌ها، رویکرد جامعی برای تشخیص داده‌های پرت ارائه می‌دهد. [۱۱].

۳.۲. آزمون‌های همگنی

۱.۳.۲. آزمون Standard Normal Homogeneity Test (SNHT)

آزمون Standard Normal Homogeneity Test (SNHT)، توسعه‌یافته توسط Alexandersson در سال ۱۹۸۶، یک آزمون پارامتریک است که سری زمانی را به دو بخش تقسیم کرده و میانگین‌ها را مقایسه می‌کند تا جهش را شناسایی کند. SNHT آزمونی شناخته‌شده برای تشخیص شکست ساختاری (changepoint) در میانگین یک سری زمانی است؛ به‌ویژه در بررسی همگنی متغیرهای اقلیمی و هیدرولوژیکی. مزیت SNHT عملکرد بسیار خوب آن در تشخیص جهش‌های میانی است [۱۲].

۲.۳.۲.آزمون Buishand Range Test

آزمون Buishand ارائه شده در سال ۱۹۸۲، بر مبنای آنالیز دامنه تجمعی انحرافات از میانگین (cumulative deviations) طراحی شده و برای تشخیص جابه‌جایی سطح پایه (shift) در میانگین سری زمانی کاربرد دارد. این آزمون اغلب در مطالعات اقلیمی و هیدرولوژیکی برای شناسایی نقاط شکست مورد استفاده قرار می‌گیرد [۱۳].

۳.۳.۲.آزمون Pettitt Homogeneity Test

آزمون Pettitt Homogeneity Test، یک روش غیرپارامتریک است که نقطه تغییر را با استفاده از آمار رتبه‌بندی تشخیص می‌دهد و در داده‌های با توزیع نامعلوم مفید است. مزیت اصلی آن بی‌نیازی از فرض نرمال بودن و حساسیت خوب به تغییرات نقطه‌ای در مرکز توزیع است [۱۴].

۳. نتایج و بحث

۱.۳. نتایج آزمون‌های تشخیص داده‌های پرت

در جدول ۱ نتایج آزمون‌های Grubbs و Dixon برای تشخیص داده‌های پرت در سری زمانی رواناب رودخانه آجی‌چای در دوره ارائه شده است. فرض صفر این آزمون‌ها مبنی بر عدم وجود داده پرت در سری داده‌ها است و فرض مقابل، به معنی وجود حداقل یک داده‌پرت می‌باشد. این آزمون‌ها در سطح معنی‌داری ۵ درصد انجام شده‌اند.

همانطور که در جدول ۱ مشاهده می‌شود، در آزمون Grubbs مقدار آماره G مشاهداتی ($G_{Observed}$ = 13.345) به‌طور قابل‌توجهی بیشتر از مقدار بحرانی ($G_{Critical}$ = 4.506) است، و در آزمون Dixon مقدار آماره R مشاهداتی ($R_{Observed}$ = 0.164) بیشتر از R بحرانی ($R_{Critical}$ = -0.479) است. بنابراین، فرض صفر در هر دو آزمون Grubbs و Dixon رد شده و نتایج حاکی از وجود داده‌های پرت در سری رواناب رودخانه آجی‌چای می‌باشد. همچنین در هر دو آزمون، مقدار p-value < alpha است، بنابراین نتایج آزمون‌ها مبنی بر وجود داده پرت در سری رواناب، در سطح ۵ درصد معنی‌دار می‌باشد.

هر دو آزمون Grubbs و Dixon وجود داده‌های پرت در سری زمانی رواناب رودخانه آجی‌چای را تأیید می‌کنند، که نشان‌دهنده انسجام نتایج در سطح معنی‌داری ۵ درصد است. آزمون Grubbs به دلیل استفاده از انحراف استاندارد و میانگین، حساسیت بالایی به داده‌هایی با انحراف زیاد دارد، اما وابستگی آن به توزیع نرمال می‌تواند محدودیت ایجاد کند. در مقابل، آزمون Dixon به دلیل غیرپارامتریک بودن، انعطاف‌پذیری بیشتری در داده‌های هیدرولوژیکی با توزیع نامشخص دارد، اما برای مجموعه‌های داده بزرگ یا در حضور پرت‌های متعدد ممکن است حساسیت کمتری داشته باشد. نتایج نشان می‌دهند که سری زمانی رواناب رودخانه آجی‌چای شامل حداقل یک داده پرت است که می‌تواند به دلایل مختلفی مانند خطای اندازه‌گیری، نقص تجهیزات یا رویدادهای هیدرولوژیکی غیرعادی (مانند سیلاب‌های شدید) ایجاد شده باشد.

جدول ۱- نتایج آزمون‌های Grubbs و Dixon در تشخیص داده‌های پرت

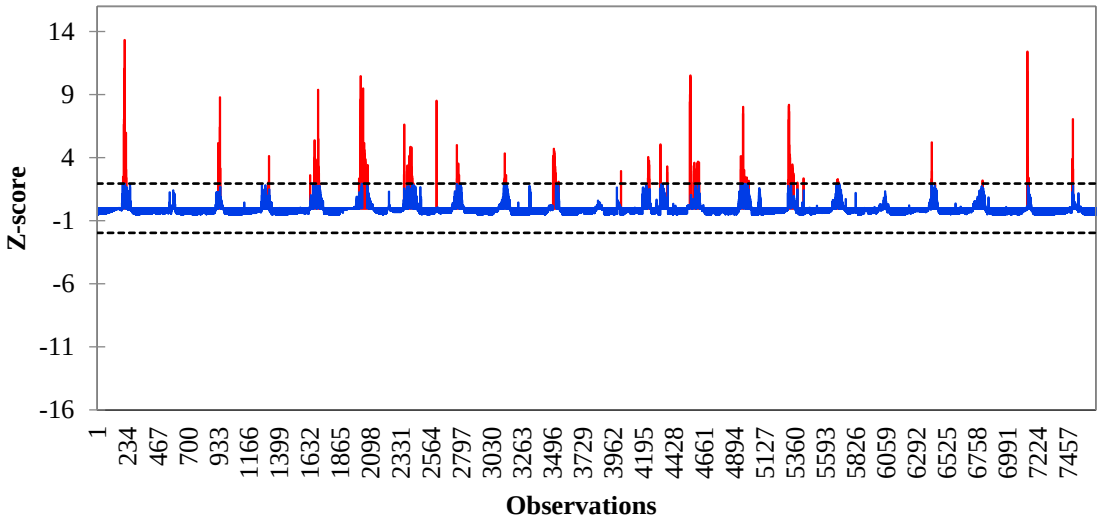
Grubbs test for outliers / Two-tailed test	Value	Dixon test for outliers / Two-tailed test	Value

G (Observed value)	13.345	R (Observed value)	0.164
G (Critical value)	4.506	R (Critical value)	-0.479
p-value (Two-tailed)	< 0.0001	p-value (Two-tailed)	< 0.0001
alpha	0.05	alpha	0.05

تحلیل داده‌های رواناب رودخانه آجی‌چای با استفاده از نتایج ارائه‌شده در شکل ۲، که نمودار سری زمانی مقادیر Z (اماره استانداردشده) و نتایج آزمون‌های Grubbs و Dixon را نشان می‌دهد، اطلاعات مهمی درباره وجود داده‌های پرت در این سری زمانی فراهم می‌کند. در این تحلیل، مقادیر Z بزرگتر از حد معنی‌داری ۵ درصد ($Z = 1.96$) به‌عنوان معیاری برای شناسایی داده‌های پرت در نظر گرفته شده‌اند. این حد، که از توزیع نرمال استاندارد استخراج می‌شود، نشان‌دهنده داده‌هایی است که با احتمال ۹۵ درصد در محدوده نرمال قرار نمی‌گیرند و بنابراین به‌عنوان پرت تلقی می‌شوند.

بر اساس شکل ۲، از مجموع ۷۶۷۰ رویداد ثبت‌شده طی ۲۱ سال (دوره زمانی روزانه)، ۳۲۹ رویداد دارای مقادیر Z بزرگتر از ۱.۹۶ هستند، که نشان‌دهنده شناسایی این تعداد به‌عنوان داده‌های پرت است. این داده‌های پرت با رنگ قرمز در نمودار مشخص شده‌اند و نشان‌دهنده انحرافات قابل‌توجه از رفتار معمول رواناب رودخانه آجی‌چای هستند. همچنین، مشاهده شده است که مقادیر رواناب بزرگتر از ۴.۵۵ مترمکعب بر ثانیه همگی دارای مقادیر Z بیشتر از ۱.۹۶ بوده و به‌عنوان پرت تشخیص داده شده‌اند. این آستانه (۴.۵۵ مترمکعب بر ثانیه) می‌تواند به‌عنوان یک معیار تجربی برای شناسایی رویدادهای غیرعادی در این حوضه در نظر گرفته شود، که احتمالاً با رویدادهای هیدرولوژیکی شدید مانند سیلاب‌های ناگهانی یا خطاهای اندازه‌گیری مرتبط است. این یافته‌ها بر لزوم کنترل کیفیت داده‌ها تأکید دارند و نشان می‌دهند که ترکیبی از مقادیر Z و آزمون‌های آماری می‌تواند ابزار قدرتمندی برای شناسایی داده‌های پرت باشد.

از منظر آماری، شناسایی ۳۲۹ داده پرت از ۷۶۷۰ رویداد (حدود ۴/۳ درصد) نشان‌دهنده وجود تعداد قابل‌توجهی انحراف در سری زمانی است. این نسبت ممکن است به ویژگی‌های حوضه آجی‌چای، مانند تغییرات فصلی شدید، رویدادهای اقلیمی حدی یا خطاهای سیستماتیک در اندازه‌گیری (مانند نقص تجهیزات هیدرومتری) مرتبط باشد. آستانه ۴/۵۵ مترمکعب بر ثانیه به‌عنوان معیاری برای پرت بودن، نشان‌دهنده وجود رویدادهای هیدرولوژیکی با شدت بالا یا خطاهای ثبت داده است. برای مثال، رواناب‌های بیش از این مقدار ممکن است با سیلاب‌های ناشی از بارش‌های شدید یا ذوب ناگهانی برف در حوضه آجی‌چای مرتبط باشند. در نهایت، استفاده از روش‌های مکمل مانند تحلیل همگنی (مانند آزمون‌های SNHT یا پتیت) می‌تواند به شناسایی تغییرات ساختاری در سری زمانی کمک کند که ممکن است با پرت‌ها مرتبط باشند. بنابراین در بخش بعدی، نتایج آزمون‌های همگنی ارائه شده است.



شکل ۲ – نمودار سری زمانی مقادیر Z آزمون‌های Grubbs و Dixon مربوط به داده‌های رواناب آجی‌چای

۲.۳. نتایج آزمون‌های همگنی

در جدول ۲ نتایج آزمون‌های همگنی شامل آزمون‌های SNHT، Buishand و Pettitt ارائه شده است. مشاهده می‌شود که طبق نتایج هر سه آزمون، در سری داده‌های رواناب آجی‌چای ناهمگنی وجود دارد. از آنجا که p-value هر سه آزمون کمتر از alpha = 0.05 است، فرض صفر مبنی بر همگن بودن سری زمانی رد شده و وجود ناهمگنی در داده‌های رواناب آجی‌چای تأیید می‌شود. این نتایج از نظر آماری در سطح ۵ درصد معنی‌دار هستند، که نشان‌دهنده وجود تغییرات ساختاری قابل‌توجه در سری زمانی است. p-valueهای بسیار پایین (به‌ویژه در Pettitt و Buishand) نشان‌دهنده اطمینان آماری بالا در وجود ناهمگنی هستند، در حالی که در آزمون SNHT (۰/۰۰۲) کمی بالاتر است، اما همچنان معنی‌دار. نزدیکی تاریخی نقاط تغییر SNHT و Buishand (حدود دو هفته اختلاف) نشان‌دهنده شناسایی یک تغییر ساختاری مشابه در حوالی سال ۱۳۹۲ است، در حالی که آزمون Pettitt نقطه تغییر متفاوتی (۱۳۹۵) را گزارش کرده است. این نتایج بر لزوم کنترل کیفیت داده‌ها قبل از استفاده در مدل‌سازی‌های هیدرولوژیکی تأکید دارند و می‌توانند به مدیریت بهتر منابع آب در حوضه آجی‌چای کمک کنند.

جدول ۲- نتایج آزمون‌های SNHT، Buishand و Pettitt برای بررسی همگنی سری داده‌های رواناب آجی‌چای

Homogeneity tests	Date	p-value (Two-tailed)	alpha
Pettitt's test	07-04-1395	< 0.0001	0.05
SNHT test	29-03-1392	0.002	0.05
Buishand's test	14-03-1392	< 0.0001	0.05

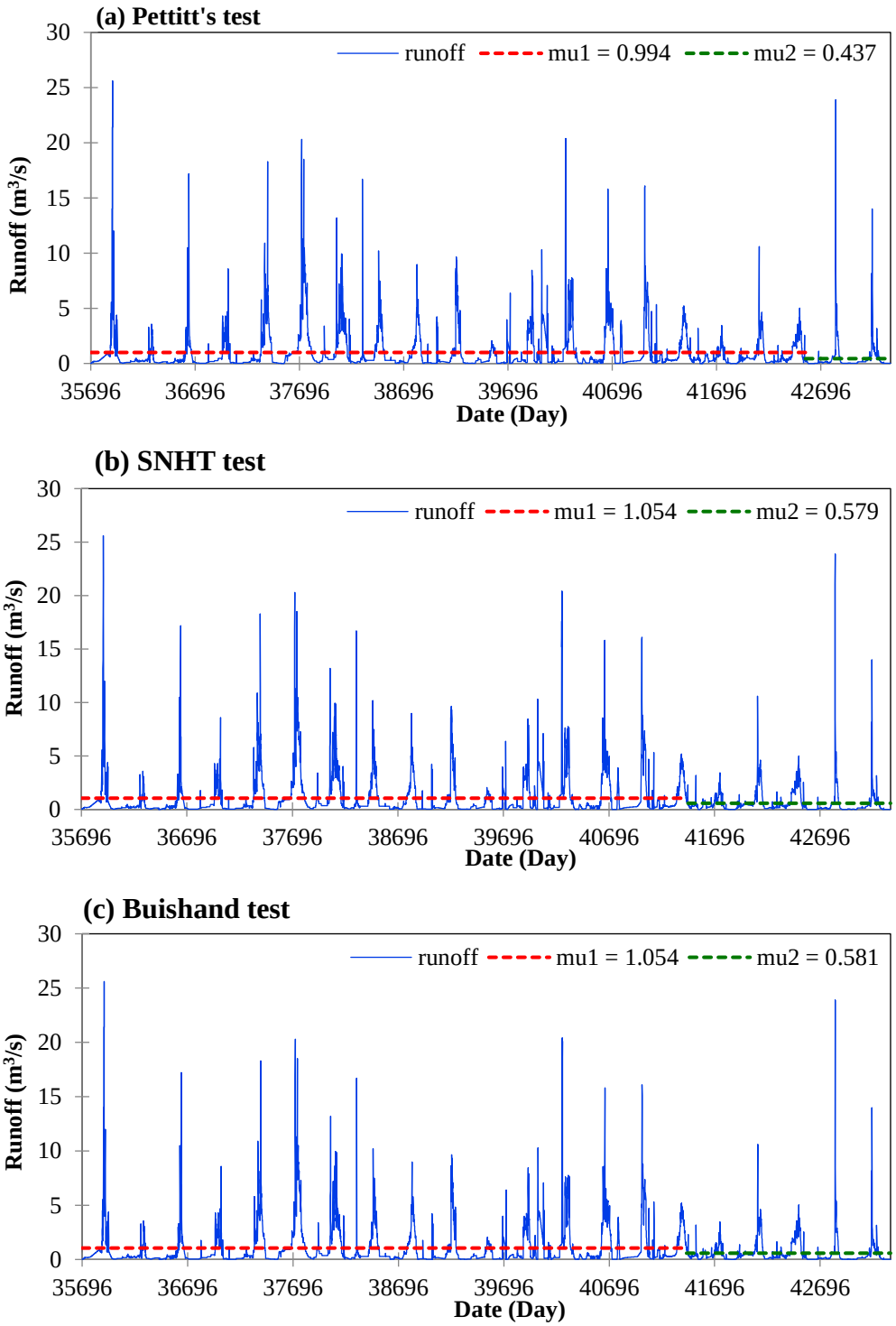
در شکل ۳ نمودارهای سری زمانی رواناب آجی‌چای، به همراه نتایج آزمون‌های همگنی که وجود نقطه شکست و میانگین رواناب قبل و بعد از نقطه شکست را نشان می‌دهد، ارائه شده است.

آزمون Pettitt نقطه شکست را در روز ۶۸۵۳ (۷ تیر ۱۳۹۵) شناسایی کرده است و دو میانگین متفاوت برای سری رواناب قبل از نقطه شکست (۰/۹۹۴ مترمکعب بر ثانیه) و بعد از آن (۰/۴۳۷ مترمکب بر ثانیه) محاسبه شده است. این کاهش چشمگیر (حدود ۵۶ درصد) در میانگین رواناب نشان‌دهنده یک تغییر ساختاری قابل‌توجه است که می‌تواند به عواملی مانند کاهش بارش، تغییر الگوی فصلی، احداث سازه‌های آبی یا تغییرات کاربری اراضی در حوضه آجی‌چای مرتبط باشد. آزمون پتیت به دلیل غیرپارامتریک بودن، برای سری‌های زمانی هیدرولوژیکی که ممکن است توزیع غیرنرمال داشته باشند (مانند رواناب با تغییرات فصلی یا رویدادهای حدی) بسیار مناسب است و حساسیت بالایی به تغییرات در انتهای سری زمانی دارد.

در آزمون SNHT نیز نقطه شکست در روز ۵۷۴۹ معادل ۲۹ خرداد ۱۳۹۲ در سری رواناب آجی‌چای تشخیص داده شده است. میانگین رواناب قبل از این نقطه ۱/۰۵۴ مترمکعب بر ثانیه و بعد از آن ۰/۵۷۹ مترمکعب بر ثانیه است، که کاهش حدود ۴۵ درصدی را نشان می‌دهد.

همچنین آزمون Buishand نقطه شکست را در روز ۵۷۳۴ (۱۴ خرداد ۱۳۹۲) با میانگین رواناب قبل از نقطه شکست مشابه SNHT (۱/۰۵۴ مترمکعب بر ثانیه) و میانگین بعد از نقطه شکست ۰/۵۸۱ مترمکعب بر ثانیه گزارش کرده است. نزدیکی تاریخی نقاط شکست SNHT و Buishand (با اختلاف ۱۵ روز) نشان‌دهنده شناسایی یک تغییر ساختاری مشابه در حوالی سال ۱۳۹۲ است، در حالی که آزمون پتیت نقطه شکست متفاوتی در سال ۱۳۹۵ را نشان می‌دهد. این تفاوت‌ها به ویژگی‌های متفاوت آزمون‌ها مربوط است: آزمون‌های SNHT و Buishand به تغییرات میانی سری زمانی حساس‌ترند، در حالی که Pettitt ممکن است به تغییرات در انتهای سری یا انواع دیگر ناهمگنی (مانند تغییر در واریانس) واکنش نشان دهد.

از منظر هیدرولوژیکی، کاهش میانگین رواناب پس از نقاط شکست (۴۵ تا ۵۶ درصد) می‌تواند پیامدهای مهمی برای مدیریت منابع آب در حوضه آجی‌چای داشته باشد. این کاهش ممکن است به عواملی مانند کاهش بارش، افزایش تبخیر به دلیل گرمایش جهانی، احداث سدها یا بندها، یا تغییرات در کاربری اراضی (مانند گسترش کشاورزی یا شهرسازی) مرتبط باشد. این نتایج بر لزوم کنترل کیفیت داده‌ها و بررسی دقیق عوامل محیطی و انسانی تأکید دارند. ترکیب این آزمون‌ها، با توجه به ویژگی‌های پارامتریک و غیرپارامتریک آن‌ها، تحلیل جامعی از ناهمگنی‌ها ارائه می‌دهد و می‌تواند به بهبود مدیریت منابع آب و پیش‌بینی سیلاب در حوضه آجی‌چای کمک کند.



شکل ۳ – نتایج آزمون‌های همگنی نشان‌دهنده وجود نقطه شکست و میانگین رواناب قبل و بعد از آن در سری زمانی رواناب

۴. نتیجه‌گیری

تحلیل داده‌های رواناب رودخانه آجی‌چای با استفاده از آزمون‌های Grubbs، Dixon، SNHT، Buishand و Pettitt نشان‌دهنده وجود داده‌های پرت و ناهمگنی‌های معنی‌دار در سری زمانی است. آزمون‌های Grubbs و Dixon داده‌های پرت را شناسایی کردند که می‌توانند ناشی از خطاهای اندازه‌گیری یا رویدادهای هیدرولوژیکی شدید باشند. آزمون‌های همگنی نیز نقاط شکست را تأیید کردند که احتمالاً به تغییرات اقلیمی، احداث سازه‌های آبی یا تغییر کاربری اراضی مرتبط هستند. تفاوت در نقاط شکست شناسایی‌شده توسط آزمون‌ها به ویژگی‌های پارامتریک و غیرپارامتریک آن‌ها مربوط است، که نشان‌دهنده لزوم استفاده ترکیبی از این روش‌ها برای تحلیل جامع است.

این نتایج بر اهمیت کنترل کیفیت داده‌ها قبل از مدل‌سازی هیدرولوژیکی تأکید دارند. بررسی منشأ پرت‌ها و ناهمگنی‌ها با داده‌های هواشناسی و سوابق انسانی، همراه با تحلیل نرمال بودن داده‌ها، می‌تواند به بهبود مدیریت منابع آب و پیش‌بینی دقیق‌تر سیلاب در حوضه‌های مختلف کمک کند.

۵. مراجع

1. Pandžić, K., Kobold, M., Oskoruš, D., Biondić, B., Biondić, R., Bonacci, O., ... & Curić, O. (2020). Standard normal homogeneity test as a tool to detect change points in climate-related river discharge variation: case study of the Kupa River Basin. Hydrological sciences journal, 65(2), 227-241.
2. Hofmeister, F., Graziano, F., Marcolini, G., Willems, W., Disse, M., & Chiogna, G. (2023). Quality assessment of hydrometeorological observational data and their influence on hydrological model results in Alpine catchments. Hydrological Sciences Journal, 68(4), 552-571.

۳. شیرمحمدی فرادنبه، ف. بختیاری، ب. و حجابی، س. (۱۴۰۲). کاربرد آزمون های پتی و بیشاند در کنترل کیفیت داده های سالانه بارش استان های چهارمحال و بختیاری و خوزستان. *سامانه‌های سطوح آبگیر باران*، ۱۱(۴)، ۳۰-۱۶.

۴. اللهویردی پور، پ. و ستاری، م.ت. (۱۴۰۲). مقایسه عملکرد روش کلاسیک رگرسیون خطی چندگانه و روش‌های داده‌کاوی نوین در مدل‌سازی بارش سالانه (مطالعه موردی: شهر اهواز). *مدل‌سازی و مدیریت آب و خاک*، ۳(۲)، ۱۴۲-۱۲۵.

5. Kocsis, T., Kovács-Székely, I., & Anda, A. (2020). Homogeneity tests and non-parametric analyses of tendencies in precipitation time series in Keszthely, Western Hungary. *Theoretical and Applied Climatology*, *139*(3), 849-859.
6. Wu, Y., Mu, H., Wu, S.Y., Xu, Y., Xu, M., Wu, H., ... & Gao, C. (2023). Changes in mean and extreme homogeneous precipitation in China during 1960–2020. *Atmospheric Research*, *292*, 106891.
7. Bickici Arikan, B., & Kahya, E. (2019). Homogeneity revisited: analysis of updated precipitation series in Turkey. *Theoretical and Applied Climatology*, *135*(1), 211-220.
8. Bliefernicht, J., Salack, S., Waongo, M., Annor, T., Laux, P., & Kunstmann, H. (2022). Towards a historical precipitation database for West Africa: Overview, quality control and harmonization. *International Journal of Climatology*, *42*(7), 4001-4023.

۹. رستم پور، م. (۱۴۰۱). مقایسه روش های شناسایی داده های پرت و تاثیر آنها در مطالعات اندازه گیری و ارزیابی مراتع. *نشریه علمی - پژوهشی مرتع و آبخیزداری*، ۷۵(۴)، ۶۶۰-۶۳۹.

10. Grubbs, F.E. (1950). Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*, 21, 27–58.
11. Dixon, W.J. (1953). Processing data for outliers. Biometrics, 9(1), 74-89.
12. Alexandersson H. (1986). A homogeneity test applied to precipitation data. *Journal of Climatology*, 6, 661–675.
13. Buishand, T.A. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of hydrology*, 58(1-2), 11-27.
14. Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society*, 28(2), 126-135.