Iranian Journal of Soil and Water Research ISNN: 2423-7833
Homepage: http://ijswr.ut.ac.ir

Investigating the homogeneity of temperature and precipitation data using statistical and statistical-climatic approaches

Majid Cheraghalizadeh[1] | Ali Khalili[2] | Somayeh Hejabi[3] | Java Bazrafshan[4]✉

1. Department of Irrigation and Reclamation Engineering, Faculty of Agriculture, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran. E-mail: cheraghalizadeh@ut.ac.ir
2. Department of Irrigation and Reclamation Engineering, Faculty of Agriculture, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran. E-mail: akhalili@ut.ac.ir
3. Department of Water Engineering, Faculty of Agriculture, Urmia University, Urmia, Iran. E-mail: s.hejabi@urmia.ac.ir
4. Corresponding Author, Department of Irrigation and Reclamation Engineering, Faculty of Agriculture, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran . E-mail: jbazr@ut.ac.ir

Article Info

Article type: Research Article

Article history:

Received: Sep. 30, 2023

Revised: Nov. 29, 2023

Accepted: Dec. 20, 2023

Published online: March. 2024

Keywords:
Adjacent station,
Iran,
Statistical-climate,
Climate signal,
Homogeneity.

ABSTRACT

In meteorological and hydrological studies, the initial and crucial step is to assess the quality of the data under investigation. In cases where long-term metadata and prior knowledge regarding data quality are lacking, the reliability of data analysis results becomes questionable across various applications. Therefore, this research aims to conduct a comparative analysis and investigation of the homogeneity of two significant atmospheric variables: temperature (including average, minimum, and maximum) and precipitation. This assessment is carried out using both statistical and statistical-climatic approaches, covering the period from 1990 to 2019 and involving data from 70 synoptic stations in Iran. The findings obtained through absolute statistical approaches revealed that the majority of breaking years for temperature and precipitation variables occurred during the latter half of the 1990s, with a particular emphasis on the year 1997 for temperature and 2006 for precipitation (and 1999 for precipitation, respectively). In general, it was observed that more than 90% of the breaking years for temperature and precipitation variables exhibited heterogeneity, significantly limiting the utility of the data in various applications. However, it is worth noting that understanding the influence of climatic signals on a wide geographical area allows for the attribution of similar breaking years in a region to climatic factors, regardless of their specific causes. This perspective allows for the consideration of heterogeneity as conditional homogeneity, with the heterogeneity factor being seen as part of the climatic norm. The results from this study demonstrate that by employing a statistical-climatic approach based on the concept of adjacent stations (nearest neighbors) to check data homogeneity, it is possible to consider 75% (100%) of the heterogeneous temperature (precipitation) variables as conditionally homogeneous. This approach helps alleviate the limitations associated with using heterogeneous data. Nevertheless, it is advisable to conduct further thorough investigations into the statistical-climatic approach to ensure its robustness and reliability.

Cite this article: Cheraghizadeh, M., Khalili, A., Hejabi, S., & Bazrafshan, J. (2024) Investigating the homogeneity of temperature and precipitation data using statistical and statistical-climatic approaches, *Iranian Journal of Soil and Water Research*, 55 (1), 1-15. https://doi.org/10.22059/ijswr.2023.365950.669584

© The Author(s).

Publisher: The University of Tehran Press.

DOI: https://doi.org/10.22059/ijswr.2023.365950.669584

مجله تحقیقات آب و خاک ایران، دوره ۵۵، شماره ۱ شاپا: ۷۸۳۳-۲۴۲۳
Homepage: http://ijswr.ut.ac.ir

بررسی همگنی داده‌های دما و بارش با استفاده از رویکردهای آماری و آماری-اقلیمی

مجید چراغعلی‌زاده[۱] | علی خلیلی[۲] | سمیه حجابی[۳] | جواد بذرافشان[۴]✉
۱. گروه مهندسی آبیاری و آبادانی، دانشکده کشاورزی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران. رایانامه: Cheraghalizadeh@ut.ac.ir
۲. گروه مهندسی آبیاری و آبادانی، دانشکده کشاورزی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران. رایانامه: akhalili@ut.ac.ir
۳. گروه مهندسی آب، دانشکده کشاورزی، دانشگاه ارومیه، ارومیه، ایران رایانامه: s.hejabi@urmia.ac.ir
۴. نویسنده مسئول، گروه مهندسی آبیاری و آبادانی، دانشکده کشاورزی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران. رایانامه: jbazr@ut.ac.ir

اطلاعات مقاله

نوع مقاله: مقالهٔ پژوهشی

تاریخ دریافت: ۱۴۰۲/۷/۸
تاریخ بازنگری: ۱۴۰۲/۹/۸
تاریخ پذیرش: ۱۴۰۲/۹/۲۹
تاریخ انتشار: فروردین ۱۴۰۳

واژه‌های کلیدی:
نزدیک‌ترین ایستگاه ،
ایران،
آماری-اقلیمی،
سیگنال اقلیمی،
همگنی.

چکیده

اولین گام در مطالعات مرتبط با پدیده‌های هواشناسی و آبشناسی بررسی کیفیت داده‌های مورد مطالعه است. به دلیل در دسترس نبودن فراداده بلندمدت و پیش‌آگاهی نسبت به کیفیت داده‌های مورد مطالعه، نتایج حاصل از تحلیل داده‌ها در کاربردهای مختلف قابل اعتماد نخواهد بود. از این‌رو، این پژوهش با هدف بررسی همگنی دو متغیر مهم جوی شامل دما (متوسط، کمینه و بیشینه) و بارش در دوره آماری ۱۹۹۰ تا ۲۰۱۹ با دو رویکرد آماری و آماری-اقلیمی برای ۷۰ ایستگاه همدید در ایران به انجام رسید. نتایج این پژوهش مبتنی بر رویکردهای مطلق آماری نشان داد که بیشترین سال‌های شکست متغیرهای دما (بارش) در نیمه دوم دهه ۱۹۹۰ میلادی با تمرکز بر سال ۱۹۹۷ (۱۹۹۹، ۲۰۰۶) است. به طور کلی بیشتر (کمتر) از ۹۰ درصد از سال‌های شکست متغیر دما (بارش) معنی‌دار می‌باشند که این امر محدودیت‌های جدی را در استفاده از داده‌ها به وجود می‌آورد. از سوی دیگر، با دانش به اینکه سیگنال‌های اقلیمی محدوده وسیعی را تحت تاثیر قرار می‌دهند می‌توان سال‌های شکست مشابه در یک منطقه را به عوامل اقلیمی فارغ از علت آن نسبت داد و ناهمگنی را همگن مشروط و عامل ناهمگنی را یک هنجار اقلیمی در نظر گرفت. نتایج این مطالعه نشان داد که با بررسی همگنی داده‌ها با رویکرد آماری-اقلیمی مبتنی بر مفهوم ایستگاه مجاور (نزدیک‌ترین ایستگاه یا ایستگاهها به ایستگاه هدف) می‌توان ۷۵ (۱۰۰) درصد ایستگاه‌های ناهمگن متغیر دما (بارش) را همگن مشروط فرض کرد و محدودیت استفاده از داده‌ها را به حداقل رساند. با این حال توصیه می‌شود که در استفاده از رویکرد آماری-اقلیمی در تحلیل همگنی داده‌های اقلیمی دقت بیشتری صورت گیرد.

استناد: چراغعلی زاده؛ مجید، خلیلی؛ علی، حجابی؛ سمیه، بذرافشان؛ جواد، (۱۴۰۳) بررسی همگنی داده‌های دما و بارش با استفاده از رویکردهای آماری و آماری-اقلیمی، *مجله تحقیقات آب و خاک ایران*، ۵۵ (۱)، ۱۵-۱. https://doi.org/10.22059/ijswr.2023.365950.669584
ناشر: مؤسسه انتشارات دانشگاه تهران.
© نویسندگان.
DOI: https://doi.org/10.22059/ijswr.2023.365950.669584

مقدمه

مدیریت و سیاستگذاری در مطالعات مرتبط با مخاطرات طبیعی نظیر خشک‌سالی و سیل متاثر از کیفیت داده‌های مورد مطالعه می‌باشد. در نتیجه، مهمترین بخش در مباحث مدیریتی قابلیت اطمینان و کیفیت داده‌های مورد نیاز است (Hänsel et al., 2016; Elzeiny et al., 2019). اولین مرحله در کنترل کیفی داده‌های مورد مطالعه بررسی همگنی است. در پژوهش‌های مبتنی بر داده‌های هواشناسی و آبشناسی، تغییرات سری زمانی همگن تنها به وسیله تغییرات در آب و هوا ایجاد می‌شود (Conrad & Pollak, 1950; Enric Aguilar & Llanso, 2003; Zhang et al., 2014). به عنوان مثال، می‌توان اثرات اقلیمی شامل فوران آتشفشان و دورپیوندها (Reid et al., 2016 ; Ghebreyesus & Sharif, 2021; Roshan et al., 2021; Lemus-Canovas, 2022) و اثرات غیراقلیمی شامل نحوه دیدبانی (روش اندازه‌گیری)، جابجایی ایستگاه، تغییر کاربری اطراف ایستگاه (جزیره گرمایی[1]) را نام برد (Aguilar et al., 2003; Sahin & Cigizoglu, 2010; Suhaila & Yusop, 2018).

اغلب سری‌های زمانی بلندمدت تحت تاثیر یکی از عوامل غیراقلیمی شامل تغییر در دستورالعمل محاسبه متغیرهای دیدبانی، تعویض و یا تغییر سنجنده‌های هواشناسی، تغییر دیدبان و تغییر کاربری اطراف ایستگاه قرار می‌گیرند. در نتیجه، این فرض که داده‌های دهه‌های اخیر دارای کیفیت بالاتری هستند را با چالش همراه می‌سازد. از این جهت، پیشنهاد شده که در سری‌های زمانی بلندمدت می‌توان با کاهش طول دوره آماری، سری‌های همگن را استخراج نمود (Bazrafshan & Cheraghalizadeh, 2021). ناهمگنی به هر دلیلی که در سری زمانی داده‌های مورد مطالعه رخ داده باشد می‌تواند منجر به نتایج نادرست شود و از این جهت باید با یکی از روش‌های همگن‌سازی، سری ناهمگن را تصحیح نمود (Rafati & Karimi, 2018; Kolendowicz et al., 2019 ; Khalil, 2021) .

به طور کلی، روش‌های تشخیص ناهمگنی داده‌ها شامل روش‌های مستقیم (مبتنی بر فراداده) و غیرمستقیم (مبتنی بر روش‌های آماری) است. روش‌های غیرمستقیم شامل روش‌های مطلق (تک ایستگاه) و نسبی (چند ایستگاه) می‌باشد (Ribeiro et al., 2016). با توجه به اینکه اغلب ایستگاه‌های هواشناسی فاقد فراداده می‌باشند و یا امکان انتخاب ایستگاه مرجع میسر نیست، پژوهشگران به سمت استفاده از روش‌های غیرمستقیم سوق داده شدند. از مهمترین روش‌های تشخیص همگنی مطلق می‌توان آزمون‌های همگنی انحرافات تجمعی[2] (CDT) (Buishand, 1982)، نسبت احتمال ورسلی[3] (WLRT) (Worsley, 1979)، همگنی نرمال استاندارد[4] (SNHT) (Alexandersson, 1986)، پتیت[5] (PT) (Pettitt, 1979) و دامنه بیشاند[6] (BRT) (Buishand, 1982) را نام برد که با توجه به نوع آزمون می‌توانند پارامتری/ ناپارمتری و در محدوده ابتدا/انتها یا میانی سری زمانی از حساسیت بیشتری برخوردار باشند. از این رو، با توجه به توزیع مکانی متغیرها در مقیاس مورد مطالعه و پیش آگاهی از فرضیات و حساسیت آزمون‌ها باید در انتخاب آزمون دقت نمود. بررسی مطالعات قبلی نشان داد که نوع اقلیم تاثیری در کاربرد نوع آزمون همگنی ندارد (بازگیر و همکاران، ۱۳۹۸) اما عملکرد آن به توزیع داده‌ها بستگی دارد (Ahmed et al., 2018). همچنین، آگاهی از محدوده ناهمگنی می‌تواند در انتخاب آزمون همگنی مفید باشد (مرتضی‌پور و همکاران، ۱۳۹۹). به کارگیری یک آزمون همگنی در صورت نبود فراداده کامل و قابل اطمینان، بدون دخالت اطلاعات حاصل از ایستگاه های دیگر و دیدگاه‌های کارشناسی ممکن است با خطا همراه باشد (Rafati & Karimi, 2018). به دلیل محدودیت‌های موجود، بررسی همگنی داده‌های یک ایستگاه با استفاده از چند آزمون همگنی انجام می‌شود و با بررسی نتایج به‌دست آمده، متغیرها در طبقات همگن، مشکوک و ناهمگن تقسیم بندی می‌شود (Wijngaard et al., 2003) .

مطالعات متعددی در زمینه بررسی همگنی داده‌های هواشناسی در سطح کشور انجام شده است. به عنوان مثال، علی‌پور و ملکیان (۱۳۹۸) به بررسی همگنی داده‌های بارش هشت ایستگاه هواشناسی طی دوره ۲۰۱۳-۱۹۸۴ واقع در شمال غرب ایران با آزمون‌های همگنی (نسبت فن نیومن[7] (VNRT)، BRT، SNHT و PT) پرداختند. نتایج این پژوهش نشان داد که آزمون‌های BRT ، SNHT و PT مشابه یکدیگر می‌باشند و به طور متوسط ۲۵/۹ درصد از آستانه‌های بارشی در منطقه مورد مطالعه ناهمگن بوده است. مرتضی‌پور و همکاران (۱۳۹۹)با انجام آزمون‌های همگنی (CDT، WLRT، SNHT، BRT و PT) داده‌های اقلیمی دما، بارش و فشار ایستگاه

1 Heat Island
2 Cumulative Deviations Test
3 Worsley's Likelihood Ratio Test
4 Standard Normal Homogeneity Test
5 Pettitt Test
6 Buishand Range Homogeneity Test
7 Von Neumann ratio test

هواشناسی رشت در بازه زمانی سال های ۱۹۶۱ تا ۲۰۱۸ را ارزیابی کردند. نتایج بررسی داده‌های بارش نشان داد که آزمون‌های همگنی نتایجی مشابه با فراداده ندارند. با این حال، نتایج PT و BRT داده‌های دما و فشار همخوانی بیشتری در کشف محدوده تغییرات با فراداده داشتند. بهاروندی و همکاران (۱۴۰۰) با توجه به آزمون SNHT و فراداده سری‌های دمای حداقل ایستگاه خرم آباد طی دوره آماری ۲۰۱۳ – ۱۹۸۴ نشان دادند که سری زمانی مورد مطالعه همگن می‌باشد. بازگیر و همکاران (۱۳۹۸) همگنی سری‌های زمانی میانگین سالانهٔ دمای کمینه، دمای بیشینه و بارش ۳۶ ایستگاه هواشناسی کشور را با آزمون های مختلف (CDT، WLRT، SNHT و PT) طی دوره آماری ۱۹۶۶ تا ۲۰۱۵ مورد بررسی قرار دادند. نتایج این مطالعه نشان داد که ۹۱/۵ و ۸۸/۵ درصد از داده‌های کمینه و بیشینه دما ناهمگن هستند. درخصوص داده‌های بارش، هر چهار آزمون، همگنی ۲۸ ایستگاه را نشان دادند. نتایج نشان داد که سال شروع تغییرات دمای بیشینه، در اکثر ایستگاه‌ها سال ۱۹۹۸ بوده ولی در خصوص دمای کمینه و بارش شروع تغییرات در ایستگاه‌های مختلف، متفاوت بود.

اکثر مطالعات بین‌المللی به منظور بررسی همگنی داده‌های هواشناسی نیز همانند مطالعات داخلی متکی بر آزمون‌های همگنی مطلق می‌باشند. به عنوان مثال، Suhaila & Yusop (2018) در مطالعه تشخیص نقاط تغییر سری‌های زمانی دمای میانگین، بیشینه و کمینه به صورت سالانه و فصلی طی دوره آماری ۱۹۸۰ تا ۲۰۱۱ در شبه جزیره مالزی با استفاده از آزمون PT پرداختند. نتایج این پژوهش نشان داد که بیشترین نقاط تغییر در سال‌های ۱۹۹۶، ۱۹۹۷ و ۱۹۹۸ رخ داده است. Ahmed et al. (2018) به بررسی همگنی مجموعه داده های بارش ماهانه و سالانه ۱۴ ایستگاه هواشناسی طی دوره آماری ۱۹۶۱ تا ۲۰۰۹ واقع در منطقه خشک پاکستان با استفاده از آزمون‌های PT، SNHT ، CDT ، VNRT ، WLRT و Student's t پرداختند. نتایج حاکی از آن است که سری زمانی بارش در اکثر ماه‌ها در تمامی ایستگاه‌ها همگن است. سری زمانی بارندگی برای ماه ژوئن در دو ایستگاه، برای آوریل در یک ایستگاه مشکوک و برای نوامبر تنها در یک ایستگاه ناهمگن است. از سوی دیگر، در سری سالانه در ۱۲ ایستگاه همگن و در دو ایستگاه مشکوک بودند. مقایسه آزمون‌های مختلف همگنی نشان داد که SNHT و WLRT حساس‌ترین و آزمون CDT کمترین حساسیت را نسبت به تغییرات داده‌های بارش ماهانه دارند. در مورد سری‌های سالانه، VNRT در مقایسه با سایر آزمون‌ها حساس‌ترین بود. Elzeiny et al. (2019) به بررسی همگنی مجموعه داده‌های بارندگی سالانه ۳۰ ایستگاه هواشناسی در سراسر حوضه رودخانه Upper Blue Nile با استفاده از چهار آزمون همگنی SNHT، BRT ، PT و VNRT طی دوره ۱۹۰۱ تا ۲۰۱۳ پرداختند. نتایج آزمون‌های SNH و BR نشان داد که سری‌های سالانه ایستگاه‌های مورد بررسی همگن بوده، در حالی که آزمون نسبت فون نیومن ناهمگنی را در شش ایستگاه تشخیص داد. Salehi et al. (2020) به بررسی همگنی مکانی و زمانی و روند بارش‌های فصلی و سالانه ۲۵ ایستگاه هواشناسی طی دوره ۱۹۵۷ تا ۲۰۱۶ در ایران پرداختند. برای بررسی همگنی از آزمون‌های Pettitt-Whitney–Mann ، SNHT ، WLRT ، BRT و VNRT استفاده شد. آزمون‌های همگنی نشان دادند که بیشتر سری‌های زمانی بارش‌های فصلی و سالانه همگن هستند. Khalil, (2021) به بررسی همگنی سری زمانی بارش هشت ایستگاه هواشناسی واقع در حوضه رودخانه Mae Klong طی دوره ۱۹۷۱ تا ۲۰۱۵ تایلند پرداخت. برای این منظور از آزمون‌های PT ، SNHT ، BRT و VNRT بصورت گروهی استفاده شد. نتایج به‌دست آمده همگنی سری زمانی ماهانه داده‌های بارندگی در ۷۵ درصد ایستگاه‌ها را مفید نشان داد، در حالی که داده‌های دو ایستگاه به‌عنوان مشکوک و رد طبقه‌بندی شدند. در مقیاس سالانه، هفت ایستگاه از هشت ایستگاه به عنوان مفید طبقه‌بندی شدند، در حالی که یک ایستگاه به عنوان رد طبقه‌بندی شد.

کاربرد روش‌های غیرمستقیم بدون وجود فراداده یا ایستگاه‌های مرجع، تصمیم‌گیری در مورد همگنی را با عدم قطعیت‌هایی مواجه می‌سازد. به عنوان مثال، محدودیتی آشکار در ظرفیت روش‌های مطلق برای جداسازی علائم اقلیمی از غیراقلیمی در سری‌های زمانی است. با این حال با توجه به این که ایستگاه‌هایی که در مجاورت هم قرار دارند و دارای همبستگی معنی‌داری می‌باشند اغلب به پدیده‌های اقلیمی منطقه‌ای پاسخ یکسانی می‌دهند، از این رو استفاده از روش‌های غیرمستقیم می‌تواند موثر واقع شود. از این جهت در این مطالعه به منظور بررسی همگنی داده‌های دما و بارش از دو رویکرد آماری و آماری–اقلیمی استفاده می‌شود.

مواد و روش‌ها

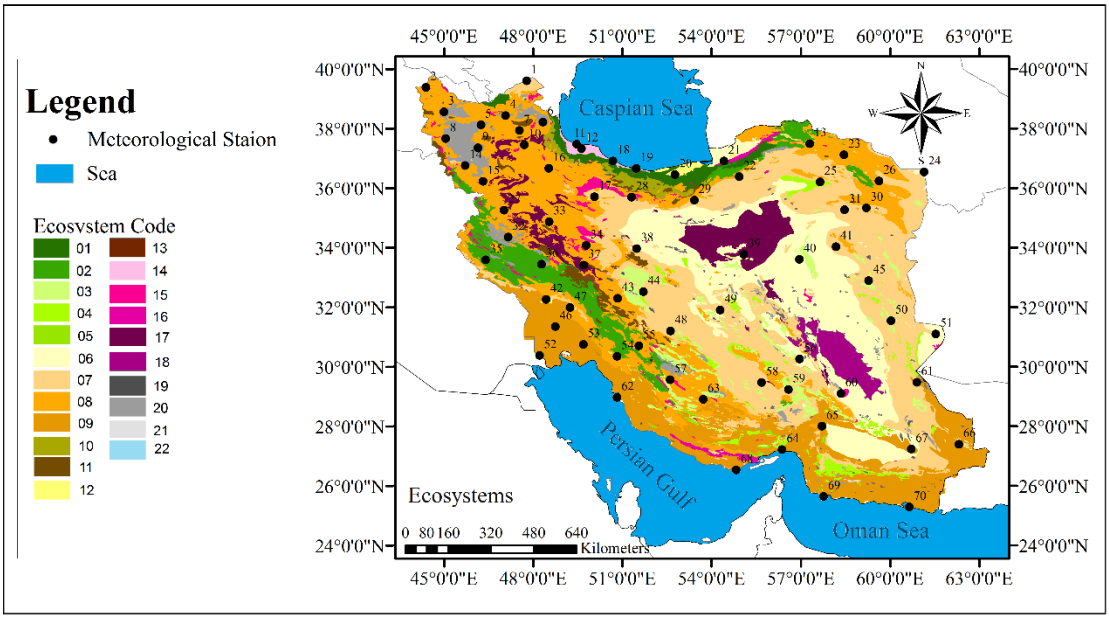
منطقه مطالعاتی

ایران با مساحت ۱/۶۵ میلیون کیلومتر مربع کشوری عمدتاً خشک و نیمه‌خشک محسوب می‌شود که محصور به عرض و طول جغرافیایی N° ۴۰ – ۲۵ و E° ۶۴ – ۴۴ می‌باشد (شکل ۱). ایران در جنوب منطقه‌ی معتدله و شمال منطقه‌ی گرمسیری با پستی و بلندی‌های شدید در جنوب‌غرب آسیا واقع شده است. این شرایط موجب شده تا اقلیم ایران شرایط متنوعی را تجربه کند. بیش از نیمی از مساحت منطقه

مورد مطالعه (متوسط ارتفاع ۱۲۰۰ متر از سطح دریا) را بلندی‌ها و کوهستان‌ها تشکیل می‌دهد. دمای متوسط ۱۸ درجه سلسیوس در کنار متوسط بارش سالانه ۲۵۰ میلیمتر با دامنه تغییرات بارش سالانه ۲۰۰۰ میلیمتری ثبت شده در ایستگاه‌های همدید نشان دهنده تنوع زیاد اقلیمی در گستره وسیع ایران است.

داده‌های مورد استفاده

داده‌های مورد استفاده در این پژوهش شامل داده‌های روزانه دما (بیشینه، کمینه و متوسط) و بارش مشاهداتی ۷۰ ایستگاه همدید برای دوره ۳۰ ساله در دوره آماری ۱۹۹۰ تا ۲۰۱۹ از سازمان هواشناسی کشور دریافت شد (https://www.irimo.ir). پراکنش ایستگاه‌های هواشناسی در شکل ۱ نشان داده شده است.



(1: Humid & Semi-Humid Forests, 2: Arid & Semi-Arid Forests, 3: Cold- Ecosystem Code: Ecosystems with Dominant Vegetation Desert Woodland & Shrublands, 4: Warm-Desert Woodland & Shrublands, 5: Moderate-Desert Woodland & Shrublands, 6: Arid Scrubland & Halophytes, 7: Cold-Desert Steppe Scrubland, 8: Cold & Arid Semi steppe Scrubland & Grasslands, 9: Warm & Arid Ecosystems Without Vegetation (12: Shrubland & Scrublands, 10: Cold & Humid Prairies, 11: Cold & Humid Cushion Scrublands), Moderate-Semi humid Bare Lowlands , 13: Warm & Arid Bare Hills , 14: Humid & Moderate Bare Hill Lowlands , 15: Semi arid & Cold Bare Mixed Plains ,16: Warm & Arid Bare Mountain Hills ,17: Moderate & Arid Bare Mountain Plains , 18: Moderate & Arid Bare diverse land forms, 19: Moderate & Arid Bare Sparse Mountains, 20: Cold & Arid Bare Mountains , 21: Warm-Desert Bare Hill Mountains , 22: Lake)

شکل ۱. نقشه مناطق اقلیمی، بوم‌سازگان (Azizi Jalilian et al., 2020) و پراکنش ایستگاه‌های هواشناسی

در این مطالعه از دو رویکرد آماری و آماری–اقلیمی به منظور بررسی همگنی داده‌ها استفاده شد که در ادامه، تشریح می‌شوند.

الف) رویکرد آماری

رویکرد آماری مورد بررسی در این مطالعه مبتنی بر Wijngaard et al., (2003) می‌باشد. در این رویکرد چهار آزمون همگنی شامل آزمون همگنی پتیت، آزمون همگنی انحرافات تجمعی، آزمون احتمال ورسلی و آزمون همگنی نرمال استاندارد مورد بررسی قرار گرفته و سپس طبقه‌بندی همگنی مطابق جدول ۱ انجام می‌شود.

جدول ۱. شرط همگنی داده‌های مورد مطالعه مبتنی بر روش Wijngaard et al. (2003)

طبقه	نتیجه آزمون های همگنی (در سطح اطینان ۵ درصد)
مفید (همگن)	حداقل سه آزمون از چهار آزمون همگنی را تایید کنند
مشکوک	فقط دو آزمون از چهار آزمون همگنی را تایید کنند
رد (ناهمگن)	حداقل سه ازمون از چهار آزمون ناهمگنی را تایید کنند

آزمون همگنی پتیت (PT)

آزمون همگنی پتیت (Pettitt, 1979) یک آزمون ناپارامتری است که نیازی به فرض نرمال بودن سری زمانی داده‌ها ندارد. کاربرد اصلی آزمون پتیت مقایسه میانگین زیر مجموعه‌های یک مجموعه یا مقایسه میانگین سری‌های زمانی اقلیمی مختلف است. فرض صفر این آزمون استقلال مقادیر مختلف سری زمانی از یکدیگر است. در مقابل، در صورتی که شکستگی و ناپیوستگی در مقدار میانگین سری‌های زمانی مختلف یا زیر–سری‌های یک جامعه آماری وجود داشته باشد، مطابق فرض یک این آزمون، سری زمانی ناهمگن است. حساسیت این آزمون در محدوده میانی سری زمانی می‌باشد. معادلات آزمون همگنی پتیت به صورت زیر است:

$$X_k = 2\sum_{i=1}^{k} r_i - k(n+1), \quad k = 1,2,3,\dots,n \qquad \text{رابطهٔ ۱)}$$

$$K = \max(|X_k|), k = 1,2,3,\dots,n \qquad \text{رابطهٔ ۲)}$$

$$p_k \approx 2\exp(\frac{-6k_*^2}{n^3+n^2}) \qquad \text{رابطهٔ ۳)}$$

که در آن، r_i رتبه متناظر با iامین داده در ترتیب صعودی، n تعداد داده‌ها، K نقطه شکست سری زمانی متناظر با بیشینه قدرمطلق آماره پتیت ، k_* نقطه شکست فرضی و p_k تقریبی از p-value آماره پتیت به منظور بررسی معنی‌داری سال شکست می‌باشد.

آزمون همگنی انحرافات تجمعی (CDT)

این آزمون مبتنی بر انحرافات تجمعی از میانگین است و توسط بیشاند ارایه شد (Buishand, 1982). فرض صفر این آزمون ثابت بودن میانگین و در مقابل، فرض یک آن تغییر در میانگین را نشان می‌دهد. حساسیت این آزمون به نقطه شکست در میانه یک سری زمانی است که به شرح زیر محاسبه می‌شود.

$$S_k^* = \sum_{i=1}^{k}(x_i - \bar{x}) \quad k = 1,2,3,\dots,n-1 \qquad \text{رابطهٔ ۴)}$$
$$(S_0^* = 0, S_n^* = 0)$$

$$S_k^{**} = \frac{S_k^*}{\sigma} \quad 0 \le k \le n \qquad \text{رابطهٔ ۵)}$$

$$Q = \max(|S_k^{**}|) \qquad \text{رابطهٔ ۶)}$$

که در آن S^* انحرافات تجمعی از میانگین، S^{**} انحرافات تجمعی مقیاس‌شده و Q مقدار آزمون است.

آزمون همگنی احتمال ورسلی (WLR)

این آزمون توسط ورسلی و با وزن‌دار کردن انحرافات تجمعی S_k^* مورد استفاده در آزمون همگنی انحرافات تجمعی ارایه گردید (Worsley, 1979). با توجه به این امر، مقادیر ابتدا و انتهای سری زمانی نسبت به داده‌های میانی از وزن بیشتری برخوردار است. در نتیجه حساسیت این آزمون به ابتدا و انتهای سری زمانی بیشتر خواهد بود که به صورت زیر محاسبه می‌شود.

$$Z_k^* = \frac{1}{\sqrt{k(n-k)}}S_k^* \quad k = 1,2,3,\dots,n-1 \qquad \text{رابطهٔ ۷)}$$

$$Z_k^{**} = \frac{Z_k^*}{\sigma} \quad 0 \le k \le n \qquad \text{رابطهٔ ۸)}$$

$$V = \max(|Z_k^{**}|) \qquad \text{رابطهٔ ۹)}$$

$$W = \frac{V\sqrt{n-2}}{\sqrt{1-V^2}}$$ رابطهٔ ۱۰)

که در آن Z^* انحرافات تجمعی از میانگین وزن‌دار، K^{**} انحرافات تجمعی مقیاس‌شده وزن‌دار، V مقدار بیشینه قدر مطلق انحرافات تجمعی مقیاس‌شده وزن‌دار و W مقدار آزمون است.

آزمون همگنی نرمال استاندارد (SNHT)

آزمون همگنی نرمال استاندارد توانایی تشخیص زمان ناپیوستگی یا وقوع ناهمگنی با حساسیت بیشتر به مقادیر ابتدایی و انتهایی سری زمانی را دارد. این آزمون با این فرض که داده‌ها از توزیع نرمال پیروی می‌کنند، محتمل‌ترین زمان برای یک تغییر ناگهانی در میانگین (a) فرض صفر و یک را به صورت زیر تعریف می‌کند:

$$H_0: Z_i \approx N(0,1) \qquad for\ i = 1,2,3,\dots,n$$
$$H_1: \begin{cases} Z_i \approx N(\mu_1, 1) & for\ i = 1,2,3,\dots,a \\ Z_i \approx N(\mu_2, 1) & for\ i = a+1, a+2, a+3,\dots,n \end{cases}$$ رابطهٔ ۱۱)

که در آن Z نمره معیار و N توزیع نرمال استاندارد است. سپس، به منظور بررسی فرض‌های مورد نظر بر اساس بیشینه مقدار دنباله T_0 از رابطه زیر استفاده می‌شود.

$$T_0 = max\left\{a\overline{Z_1}^2 + (n+a)\overline{Z_2}^2\right\} \qquad for\ 1 \le a \le n-1$$ رابطهٔ ۱۲)

که در آن $\overline{Z_1}$ و $\overline{Z_2}$ مقادیر میانگین Z_i قبل و بعد از نقطه تغییر هستند. مقدار α برابر با محتمل‌ترین زمان بروز تغییر ناگهانی در داده‌ها می‌باشد.

ب) رویکرد آماری-اقلیمی

رویکرد آماری-اقلیمی مبتنی بر یک روش همگنی مطلق و بر این اصل استوار می‌باشد که انتظار می‌رود نقاط دارای تغییرات مشابه را در یک منطقه مشاهده نمود. در نگاه دیگر، رخدادهای اقلیمی بصورت نقطه‌ای تاثیرگذار نبوده و می‌تواند اثرات منطقه‌ای داشته باشند. با این حال نباید از اثرات خرد-اقلیمی حاکم بر یک منطقه غافل شد. در این مطالعه از آزمون همگنی ناپارامتری پتیت و مفهوم ایستگاه(های) مجاور (نزدیکترین ایستگاه یا ایستگاه‌ها) به منظور بررسی همگنی (همگنی مشروط) استفاده شد. ایستگاه(های) مجاور باید با ایستگاه مورد نظر در یک فاصله معقول قرار داشته و یا خصوصیات اقلیمی -بوم‌سازگانی یا آبشناسی مشابه با ایستگاه مورد مطالعه داشته باشد. از این جهت سه مجموعه داده شامل پنج ایستگاهی که در نزدیکترین فاصله، ایستگاه‌های واقع در یک حوضه آبریز و ایستگاه‌هایی که دارای خصوصیات اقلیمی-بوم‌سازگانی مشابه هستند جمع آوری می‌شود. سپس حداکثر ۵ ایستگاهی که بیشترین همبستگی معنی‌دار را با ایستگاه مورد نظر دارند به عنوان ایستگاه(های) مجاور مورد بررسی قرار می‌گیرند. در ادامه و با توجه به جدول ۲ بعد از بررسی همگنی ایستگاه‌های مد نظر در مورد همگنی آنها تصمیم‌گیری می‌شود.

جدول ۲. شرط همگنی (همگنی مشروط) داده‌های مورد مطالعه مبتنی بر رویکرد آماری-اقلیمی

کلاس	شرط مورد بررسی
همگن	آزمون مورد بررسی همگنی را تایید کند
همگن مشروط	حداقل یک نقطه شکست متناظر با ایستگاه مورد مطالعه دیده شود
ناهمگن	در ایستگاه‌های مجاور نقطه شکست مشاهده نشود.

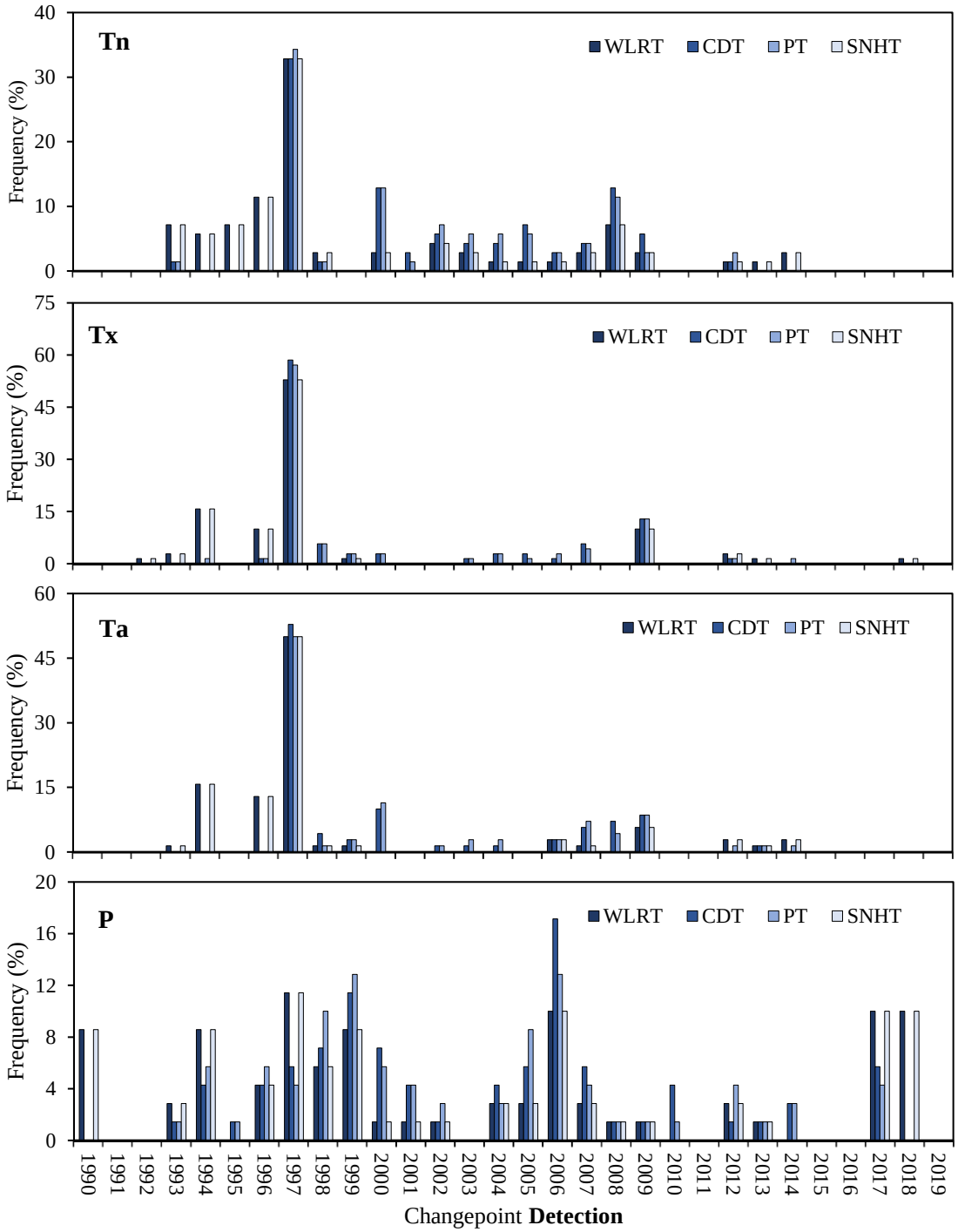
نتایج و بحث

نتایج

بعد از جمع‌آوری داده‌های دما شامل دمای کمینه، بیشینه و متوسط و بارش به بررسی همگنی داده‌ها با دو رویکرد آماری و آماری-اقلیمی در مقیاس زمانی سالانه پرداخته شد. نتایج این بخش به تفکیک دو رویکرد به شرح زیر است.

رویکرد آماری

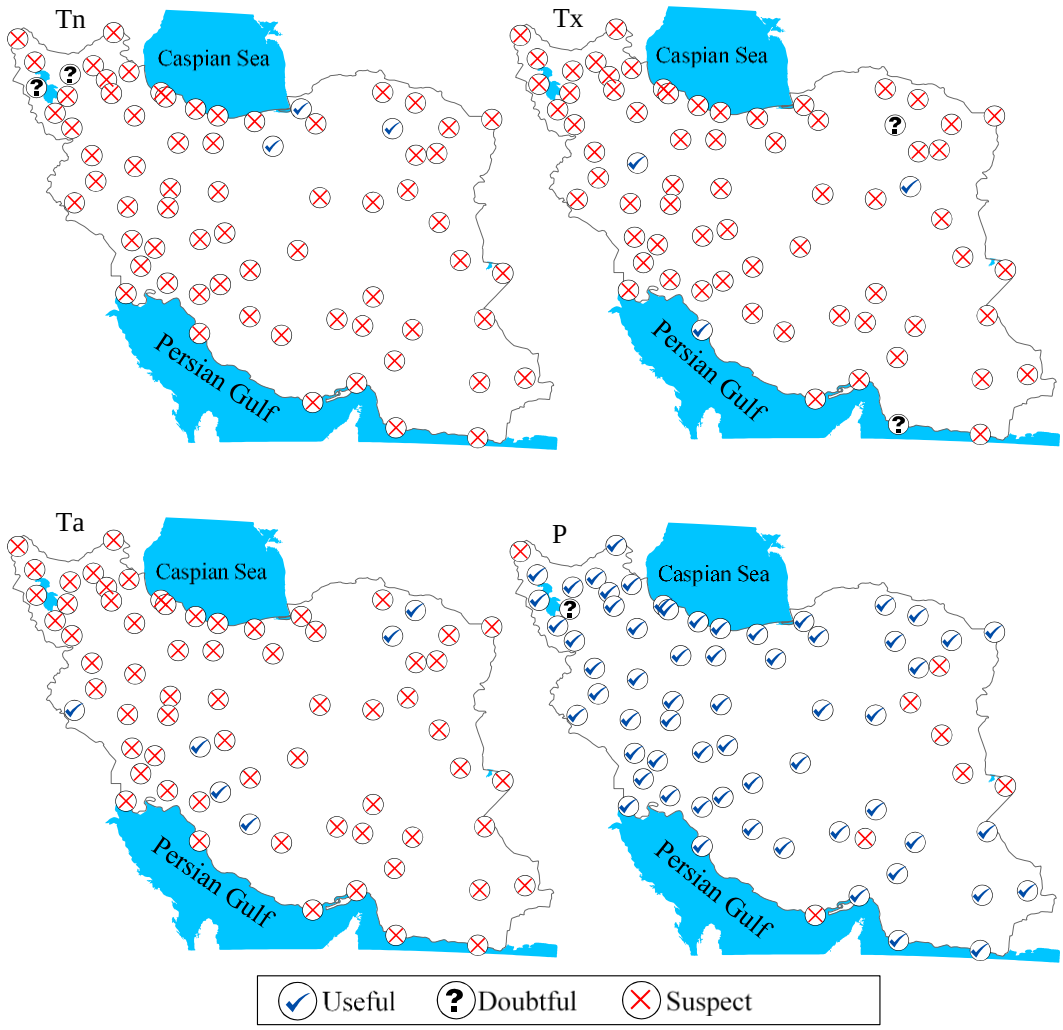
این رویکرد مبتنی بر روش Wijngaard et al. (2003)می‌باشد. شکل ۲ فراوانی وقوع سال شکست مبتنی بر چهار روش آماری را نشان می‌دهد. بیشترین سال شکست مشاهده شده برای داده‌های دما سال ۱۹۹۷ و برای بارش سال ۲۰۰۶ می‌باشد.



شکل ۲. فراوانی سال شکست متغیرهای دمایی (کمینه: Tn، بیشینه: Tx و متوسط: Ta) و بارش (P) با روش‌های آماری مختلف
(Worsley's Likelihood Ratio Test (WLRT), Cumulative Deviation Test (CDT), Pettitt Test (PT), and the Standard Normal Homogeneity Test)

برای داده‌های دما (کمینه، بیشینه و متوسط) بیش از ۹۰ درصد سال‌های شکست معنی‌دار و بیشترین فراوانی وقوع برابر با سال ۱۹۹۷ می‌باشد. بررسی نتایج مربوط به سال شکست ۱۹۹۷ متغیرهای دمای کمینه، متوسط و بیشینه نشان دادکه آزمون‌های WLRT و SNHT فراوانی وقوع (درصد معنی‌داری) برابر با ۳۲/۳ (۱۰۰)، ۵۰ (۹۷/۱) و ۵۲/۹ (۹۱/۹) درصد و برای CDT برابر با ۳۲/۳ (۹۵/۶)، ۵۲/۹ (۹۱/۹) و ۵۸/۶ (۹۰/۲) درصد می‌باشد. بررسی نتایج PT برای سال شکست ۱۹۹۹ نشان داد که فراوانی وقوع (درصد معنی‌داری)

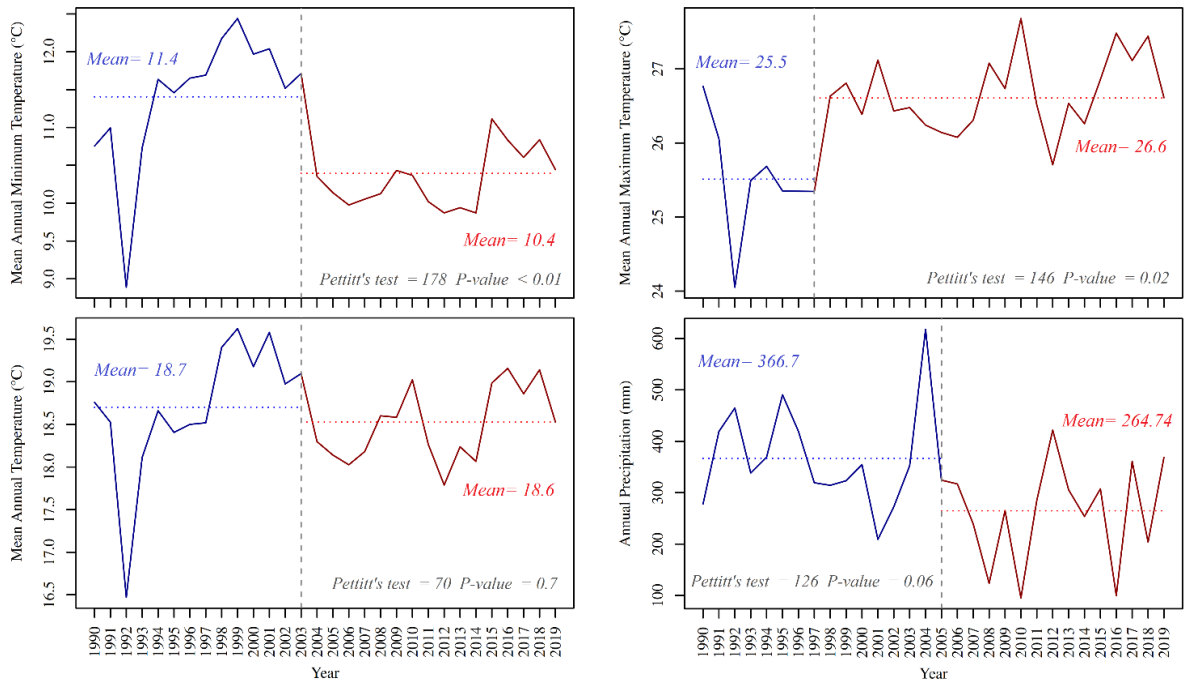
دمای بیشینه برابر با ۵۷/۱ (۷۷/۵) درصد، دمای متوسط برابر با آزمون WLRT و برای دمای کمینه برابر با ۵۷/۱ (۷۷/۵) درصد می‌باشد. به طور کلی بیشترین فراوانی وقوع سال شکست مشاهده شده برای دمای بیشینه CDT برابر با ۵۸/۶ درصد و کمترین فراوانی سال شسکت برای دمای کمینه توسط WLRT، CDT، PT و SNHT برابر با ۳۲/۳ درصد مشاهده شد. . کمتر از ۱۵ درصد سال‌های شکست متغیر بارش معنی‌دار بوده و بیشترین فراوانی شکست در سال‌های ۲۰۰۶، ۱۹۹۹ و ۱۹۹۷ مشاهده می‌شود، با این حال سال‌های شکست مختلفی در طول سری زمانی وجود دارد. برای سال شکست ۲۰۰۶ بیشترین فراوانی وقوع در حدود ۱۷/۱ درصد توسط CDT مشاهده شد که از این مقدار ۲ مورد معنی‌دار می‌باشد. فراوانی وقوع PT برابر با ۱۲/۸ درصد و برای WLRT و CDT برابر با ۱۰ درصد مشاهده شد که سال شکست معنی‌داری توسط این آزمون‌ها گزارش نشد. در بررسی سال شکست ۱۹۹۹ مشاهده شد که درصد فراوانی (درصد معنی داری) برای WLRT، CDT، PT و SNHT به ترتیب برابر با ۸/۶ (۶۶/۷)، ۱۱/۶ (۶۲/۵)، ۱۲/۹ (۵۵/۵) و ۸/۶ (۶۶/۷) می‌باشد. برخلاف سال ۱۹۹۹ که اکثرا نقاط شکست معنی‌دار می‌باشند برای سال ۱۹۹۷، ۲۵ درصد هر یک از آزمون‌های WLRT با فراوانی وقوع ۱۱/۶ درصد، CDT با فراوانی وقوع ۵/۷ درصد و SNHT با فراوانی وقوع ۱۱/۶ معنی‌دار بوده و سال شکست معنی‌داری برای PT با فراوانی وقوع ۳/۴ درصد مشاهده نشد. بعد از آشکارسازی سال‌های شکست، طبقه‌بندی داده‌های مربوطه در سه کلاس مفید (همگن)، مشکوک و رد (ناهمگن) انجام شد. با توجه به شکل ۳ نتایج این بخش نشان می‌دهد که داده‌های دمای کمینه همانند دمای بیشینه به ترتیب در حدود ۴/۳ و ۹۲/۹ درصد همگن و ناهمگن بوده و مابقی ایستگاه‌ها در کلاس مشکوک طبقه‌بندی می‌شوند. ۱۰ درصد داده‌های دمای متوسط همگن بوده و بقیه ایستگاه‌ها در طبقه ناهمگن دسته‌بندی می‌شوند. بررسی داده‌های بارش در گستره ایران نشان داد که در حدود ۸۷/۱، ۱۱/۴ و ۱/۵ درصد ایستگاه‌ها به ترتیب همگن، ناهمگن و مشکوک می‌باشند.



شکل ۳. پراکنش ایستگاه‌های مفید، مشکوک و رد متغیرهای دمایی (کمینه: Tn، بیشینه: Tx و متوسط: Ta) و بارش (P)

رویکرد آماری-اقلیمی

این رویکرد با این فرض که هر یک از آزمون‌های همگنی می‌تواند در هر اقلیمی مورد استفاده قرار گیرد و نوع اقلیم در کاربرد نوع آزمون همگنی تاثیری ندارد مورد بررسی قرار گرفت (بازگیر و همکاران، ۱۳۹۸). بعد از مشخص نمودن سال شکست هر متغیر و تعیین ایستگاه(های) مجاور به بررسی همگنی داده‌ها پرداخته شد. به عنوان مثال، بررسی همگنی متغیرهای دما (کمینه، بیشینه و متوسط) و بارش سالانه ایستگاه شیراز در شکل ۴ نشان داده شده است. بر اساس آزمون همگنی پتیت، بررسی همگنی متغیرهای مورد مطالعه در ایستگاه شیراز نشان داد که متغیر دمای کمینه دارای نقطه شکست معنی‌داری در سال ۲۰۰۳ است که این رخداد در ایستگاه‌های مجاور گزارش نشده است. در حالی که نقطه شکست معنی‌دار متغیر دمای بیشینه در ایستگاه‌های مجاور فسا و آباده با متوسط همبستگی معنی دار ۰/۸۶ گزارش شده است. در نتیجه متغیرهای دمای کمینه و بیشینه به ترتیب ناهمگن و همگن مشروط فرض می‌شوند. نقاط شکست در سال‌های ۲۰۰۳ و ۲۰۰۵ به ترتیب در سری زمانی دمای متوسط و بارش مشاهده شد که به دلیل عدم معنی‌داری آماره پتیت، این دو متغیر همگن در نظر گرفته می‌شوند.



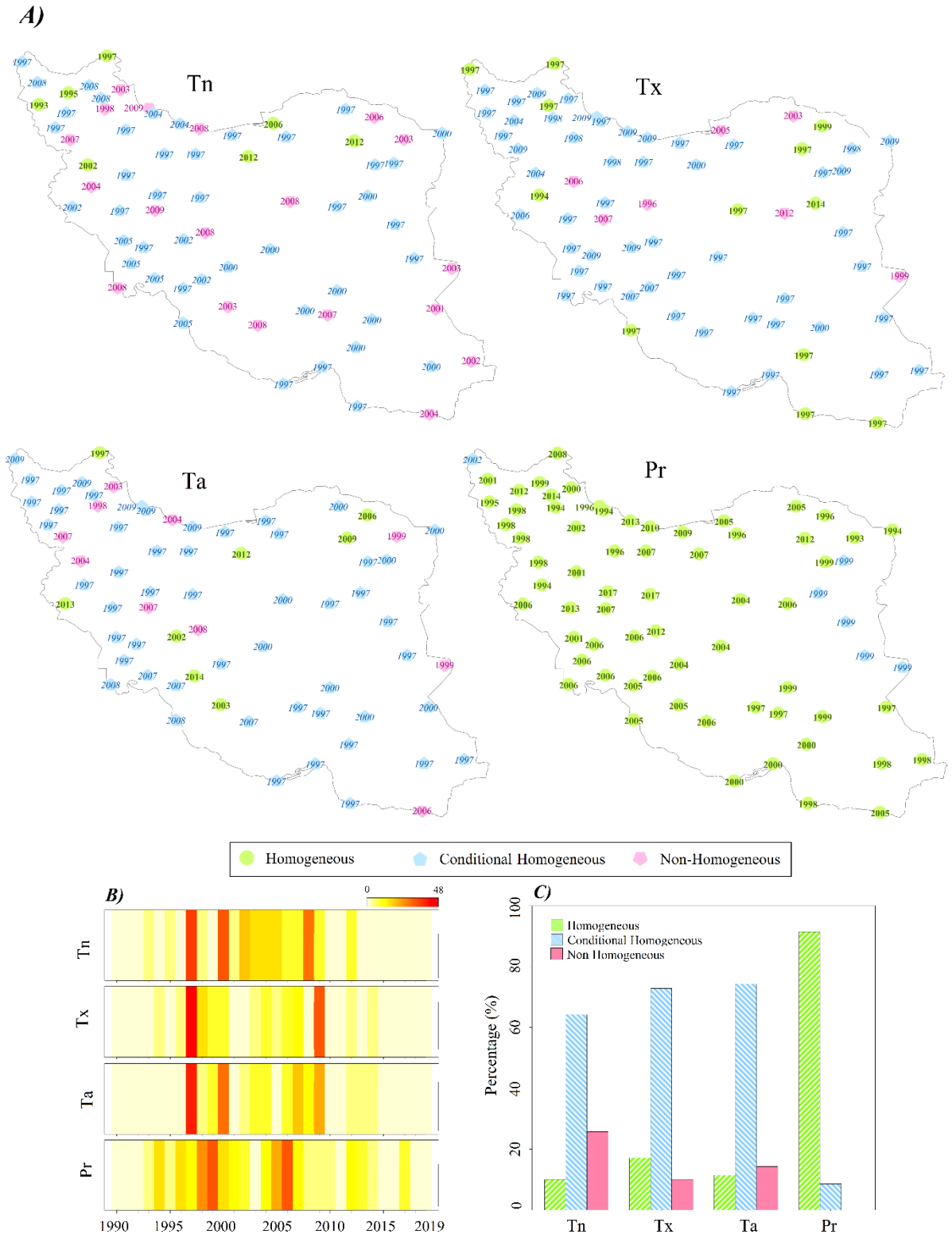
شکل ٤. سری زمانی سالانه و سال شکست متغیرهای دمایی (کمینه، بیشینه و متوسط) و بارش سالانه شیراز

نتایج همگنی داده‌ها برای کلیه ایستگاه‌ها مبتنی بر آزمون پتیت و رویکرد آماری-اقلیمی در سطح اطمینان ۹۵ درصد در شکل ۵ نشان داده شده است. به طور کلی، بررسی نتایج همگنی برای هر متغیر به شرح زیر می باشد.

دمای کمینه: بیشترین نقاط شکست (درصد معنی‌داری) در سری زمانی ایستگاه‌های مورد مطالعه مربوط به سال‌های ۱۹۹۷، ۲۰۰۰ و ۲۰۰۸ به ترتیب برابر با ۳۴/۳ (۹۱/۷)، ۱۲/۹ (۱۰۰) و ۱۱/۴ (۱۰۰) درصد است. بررسی مکانی توزیع سال‌های شکست دمای کمینه نشان دهنده تراکم توزیع مکانی در برخی از مناطق است. به عنوان مثال، سال شکست ۱۹۹۷ در ایستگاه‌های جاسک، بندرعباس و بندرلنگه؛ سال شکست ۲۰۰۰ در ایستگاه‌های کرمان، سیرجان و بم و برای سال شکست ۲۰۰۸ ایستگاه‌های خوی، اهر و سراب نقاط شکست مشابهی داشته که همه موارد با رویکرد آماری ناهمگن و با رویکرد آماری-اقلیمی همگن مشروط فرض می‌شوند. نتایج این بخش با توجه به رویکرد آماری-اقلیمی نشان داد که ۱۰، ۶۴/۳ و ۲۵/۷ درصد ایستگاه‌ها به ترتیب همگن، همگن مشروط و ناهمگن می‌باشند.

دمای بیشینه: بیشترین نقاط شکست (درصد معنی‌داری) در سری زمانی ایستگاه‌های مورد مطالعه مربوط به سال‌های ۱۹۹۷ و ۲۰۰۹ به ترتیب برابر با ۵۷/۱ (۷۷/۵) و ۱۲/۹ (۱۰۰) درصد است. بررسی توزیع مکانی داده‌های بیشینه برای سال شکست ۱۹۹۷ در گستره ایران نشان داد که پراکندگی سال شکست در بیشتر مناطق جنوبی و شرق از تراکم بیشتری برخوردار است. به عنوان نمونه تمامی ایستگاه‌های نوار ساحلی جنوب کشور سال شکست ۱۹۹۷ را نشان می‌دهند که با توجه به رویکرد آماری معنی‌دار و ناهمگن می‌باشند در

صورتی که با رویکرد آماری–اقلیمی همگن مشروط در نظر گرفته می‌شوند. برای سال شکست ۲۰۰۹ نیز می‌توان ایستگاه‌های بندرانزلی، رامسر و نوشهر را بیان نمود که همگی سال شکست معنی‌داری را داشته اما با رویکرد آماری–اقلیمی همگن مشروط در نظر گرفته شدند. نتایج این بخش با توجه به رویکرد آماری–اقلیمی نشان داد که ۱۷/۱، ۷۲/۹ و ۱۰ درصد ایستگاه‌ها به ترتیب همگن، همگن مشروط و ناهمگن می‌باشند.



شکل ۵. نقشه توزیع جغرافیایی سال‌های شکست (A) و نمودارهای فراوانی سال‌های شکست (B) و درصد ایستگاه‌های همگن، همگن مشروط و ناهمگن (C) متغیرهای دمایی (کمینه: Tn، بیشینه: Tx و متوسط: Ta) و بارش (P)

دمای متوسط: بیشترین نقاط شکست (درصد معنی‌داری) در سری زمانی ایستگاه‌های مورد مطالعه مربوط به سال‌های ۱۹۹۷ و ۲۰۰۰ به ترتیب برابر با ۵۰ (۹۷/۱)و ۱۱/۴ (۱۰۰) درصد است. توزیع مکانی دادهای دمای متوسط برای سال شکست ۱۹۹۷ در گستره ایران با تراکم بیشتر در محدوده‌های از شمال غرب، غرب، جنوب کرمان و سواحل نزدیک به تنگه هرمز مشهودتر است. با توجه به رویکرد آماری تمامی سال‌های شکست ۱۹۹۷ به جز یک مورد ناهمگن و با رویکرد آماری–اقلیمی تمامی موارد همگن مشروط در نظر گرفته می‌شود. نتایج این بخش با توجه به رویکرد آماری–اقلیمی نشان داد که ۱۱/۴، ۷۴/۳ و ۱۴/۳ درصد ایستگاه‌ها به ترتیب همگن، همگن مشروط و ناهمگن می‌باشند.

بارش: بررسی سری زمانی بارش ایستگاه‌های مورد مطالعه نشان داد که در حدود ۱۱/۴ درصد سال‌های شکست معنی‌دار و فراوانی سال‌های شکست در ۷۰ درصد سال‌های دوره مطالعاتی مشاهده شد که تنوع در سال شروع تغییرات بارش در ایستگاه‌های مختلف در مطالعه بازگیر و همکاران (۱۳۹۸) نیز مشاهده شد. بیشترین نقاط شکست (درصد معنی‌داری) در سری زمانی ایستگاه‌های مورد مطالعه مربوط به سال‌های ۱۹۹۹ و ۲۰۰۶ به ترتیب برابر با ۱۲/۹ (۵۵/۶) و ۱۲/۹ (۰) درصد است. متمرکزین توزیع مکانی سال شکست ۱۹۹۹ در شرق ایران و برای سال شکست ۲۰۰۶ در استان خوزستان می‌باشد. برای سال شکست ۱۹۹۹ در شرق کشور ایستگاه‌های تربت حیدریه، فردوس، بیرجند، نهبندان معنی‌دار و برای سال شکست ۲۰۰۶ ایستگاه‌های اهواز، آبادان، امیدیه و مسجدسلیمان غیرمعنی‌دار می‌باشند که با رویکرد آماری–اقلیمی سال ۱۹۹۹ همگن مشروط و سال ۲۰۰۶ همگن در نظر گرفته شد. نتایج این بخش با توجه به رویکرد آماری–اقلیمی نشان داد که ۹۱/۴، ۸/۶ درصد ایستگاه‌ها به ترتیب همگن و همگن مشروط می‌باشند.

نتیجه‌گیری

این مطالعه بدون پیش آگاهی نسبت به فراداده یا رخدادهای اقلیمی به بررسی همگنی داده‌های دمایی و بارش در مقیاس سالانه با دو رویکرد آماری و آماری–اقلیمی پرداخت. به منظور بررسی همگنی مبتنی بر رویکرد آماری بیشتر (کمتر) از ۹۰ درصد از داده‌های دمایی (بارش) را باید ناهمگن فرض کرد که این امر می‌تواند در مطالعات مبتنی بر داده‌ها تاثیرگذار باشد. از این رو، باید به بررسی عوامل اقلیمی و غیراقلیمی در دوره مورد مطالعه برای ایستگاه مورد نظر پرداخت. در برخی از موارد آگاهی از عوامل اقلیمی و غیراقلیمی میسر نمی‌باشد و پژوهشگر باید از روش‌های غیرمستقیم نظیر ایستگاه مجاور کمک بگیرد تا عدم قطعیت‌های مطالعه خود را کاهش دهد. این پژوهش نشان داد که با در نظر گرفتن مفهوم ایستگاه مجاور می‌توان ۷۵ (۱۰۰) درصد متغیرهای دمایی (بارش) ایستگاه های ناهمگن را به همگن مشروط تبدیل نمود. در نتیجه، از محدودیت‌های استفاده از داده های مورد مطالعه در پژوهش کاسته می‌شود. از این جهت که هدف این پژوهش بررسی تاثیر سینگال‌های اقلیمی بر سری زمانی داده‌های هواشناسی نمی‌باشد، همگنی داده‌ها بصورت مشروط مورد بررسی قرار داده شد.

به طور کلی نتایج بدست آمده از این پژوهش به شرح زیر است:

به منظور بررسی همگنی مبتنی بر نقاط شکست ابتدا باید محدوده نقاط شکست مبتنی بر آزمون‌های آماری با حساسیت به ابتدا/انتها و میانه سری زمانی مورد بررسی قرار گیرد تا آشکارسازی سال شکست بصورت صحیح انجام شود. نتایج این پژوهش نشان داد که هر چهار آزمون همگنی شامل پتیت، انحرافات تجمعی، درستنمایی ورسلی و نرمال استاندارد با حساسیت‌های مختلف توانایی تشخیص محدوده نقاط شکست را دارا بودند. به طور کلی هر چهار آزمون بیشینه نقاط شکست متغیرهای دمایی (کمینه، متوسط و بیشینه) را در اواخر دهه ۱۹۹۰ میلادی و برای متغیر بارش اواخر دهه ۱۹۹۰ و میانه دهه ۲۰۰۰ میلادی تشخیص دادند. نتایج مطالعه Ghasemi (2015) در بررسی نقاط شکست متغیرهای دمایی (کمینه، متوسط و بیشینه) ۳۸ ایستگاه هواشناسی با استفاده از آزمون همگنی پتیت برای دوره مطالعاتی ۲۰۱۰–۱۹۶۱ در ایران مشابه با پژوهش حاضر می‌باشد. با توجه به تشابه نتایج آزمون‌های همگنی و همچنین توانایی آزمون ناپارامتری همگنی پتیت در آشکارسازی نقاط شکست مشابه در میانه (مطالعه حاضر) و انتهای (Ghasemi, 2015) سری‌های زمانی مختلف با طول دوره‌های متفاوت از یک جامعه آماری توصیه می‌شود از آزمون همگنی پتیت به منظور آشکارسازی نقاط شکست استفاده شود.

بررسی همگنی بدون دسترسی به فراداده یا اطلاعات اقلیمی تاثیرگذار بر نقاط شکست می‌تواند تصمیم‌گیری در مورد همگنی/ناهمگنی داده‌ها را تحت شعاع قرار دهد یکی از نقاط ضعف استفاده از رویکردهای آماری عدم تشخیص عوامل اقلیمی به عنوان عامل همگنی در سری زمانی داده‌های هواشناسی است. این مطالعه نشان داد که صرف استفاده از رویکردهای آماری می‌تواند محدودیت

استفاده از داده‌های هواشناسی به ویژه متغیرهای دما را به وجود آورد. اما با این فرض که رخدادهای اقلیمی بصورت نقطه‌ای تاثیرگذار نبوده و می‌تواند اثرات منطقه‌ای داشته باشند، فارغ از تطابق صحیح سال شکست و رخداد اقلیمی به دلیل فرایندهای حاکم بر خرداقلیم یک منطقه می‌توان سال‌های شکست مشابه در ایستگاه‌های مجاور را همگن مشروط در نظر گرفت. به عنوان مثال، گزارش هیئت بین‌دولتی تغییر اقلیم[1] (IPCC) نقطه تغییر در میانگین دمای جهانی را در سال ۱۹۹۸ گزارش نمود (IPCC, 2014). با در نظر گرفتن این پیش آگاهی انتظار می‌رود که نقاط شکست معنی‌داری در مجاورت سال ۱۹۹۸ مشاهده شود. نتایج این پژوهش نشان داد که حدود ۵۰ درصد نقاط شکست معنی‌دار در سال ۱۹۹۷ در مجاورت خود سال شکستی مشابهی را مشاهده کرده‌اند و در نتیجه انتساب همگن مشروط به سال ۱۹۹۷ می‌تواند مفید واقع گردد. در اکثر مقالات داخلی و خارجی ارتباط قوی بین سیگنال‌ها و پارامترهای اقلیمی به خصوص بارش در مناطق مختلف ایران با و بدون گام زمانی مشاهده شد (محجوبی و همکاران، ۱۴۰۰). بررسی داده‌های بارش در پژوهش حاضر نشان داد که تمامی ایستگاه‌های ناهمگن در شرق منطقه مطالعاتی سال شکست ۱۹۹۹ را تجربه کرده که در مجاورت زمانی رخداد حدی ال نینو در سال‌های ۱۹۹۸-۱۹۹۷ و مطابقت زمانی با رخداد حدی لانینا سال‌های ۱۹۹۸-۱۹۹۹ است (Cai et al., 2014; McPhaden, 1999; Paek et al., 2017).

مشاهده نقاط شکست مشابه در ایستگاه‌های مجاور نشان‌دهنده ناهمگنی مبتنی بر آزمون‌های آماری و همگنی مشروط احتمالاً مبتنی بر سیگنال‌های اقلیم منطقه‌ای است که مورد اخیر نیازمند بررسی‌های بیشتر در مطالعات آینده است.

"هیچ‌گونه تعارض منافع بین نویسندگان وجود ندارد"

منابع

بازگیر، سعید، عباسی، فائزه، اسعدی اسکویی، ابراهیم، حقیقت، مسعود، و رضازاده، پرویز. (۱۳۹۸). تحلیل همگنی داده‌های دما و بارش در ایران با رویکرد اقلیمی. تحلیل فضایی مخاطرات محیطی، ۶(۱)، ۵۱-۷۰. SID. https://sid.ir/paper/379112/fa

علی‌پور، حسن، ملکیان، آرش. (۱۳۹۸). تحلیل همگنی و روند بارش‌های آستانه با رویکرد آماری ناپارامتری در شمال غرب ایران. مهندسی و مدیریت آبخیز، ۱۱(۴)، ۹۱۷-۹۲۸. doi: 10.22092/ijwmse.2018.109070.1256

بهاروندی، کبرا، خورشیددوست، علی محمد، نساجی زواره، مجتبی. (۱۴۰۰). آشکارسازی نوسانات اقلیمی با استفاده از روش آزمون همگنی نرمال استاندارد (مطالعه موردی: ایستگاه خرم‌آباد). نشریه علمی جغرافیا و برنامه ریزی، ۲۵(۷۵)، ۵۱-۶۳. doi: 10.22034/gp.2021.10829

محجوبی، عماد، بخشش رباط، سلمان، حسین‌پور، مسعود. (۱۴۰۰). مروری بر برخی مطالعات پیرامون تأثیر دورپیوندها بر بارش ایران در بازه سال‌های ۱۳۸۳ تا ۱۳۹۷. نیوار، ۴۵(۱۱۲-۱۱۳)، ۲۸-۴۳. doi: 10.30467/nivar.2021.246857.1167

مرتضی پور، سامان، اسعدی اسکویی، ابراهیم، و عباسی، فائزه. (۱۳۹۹). ارزیابی کارآیی چند آزمون همگنی در شناسایی جهش در داده‌های اقلیمی دما، بارش و فشار؛ مطالعه موردی: ایستگاه فرودگاهی رشت. نیوار، ۴۴(۱۰۸-۱۰۹)، ۱۱-۳۰. doi: 10.30467/nivar.2020.188070.1132

REFERENCES

Aguilar, E., Auer, I., Brunet, M., Peterson, T. C., & Wieringa, J. (2003). Guidelines on climate metadata and homogenization. World Meteorol. *Organ*, *1186*, 1–52.

Ahmed, K., Shahid, S., Ismail, T., Nawaz, N., & Wang, X.-J. (2018). Absolute homogeneity assessment of precipitation time series in an arid region of Pakistan. *Atmósfera*, *31*(3), 301–316.

Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *Journal of Climatology*, *6*(6), 661–675.

Alipoor, H., & Malekian, A. (2019). Aassessment of homogeneity and trend of precipitation thresholds based on nonparametric approaches in north-west Iran. *Watershed Engineering and Management*, *11*(4), 917-928. doi: 10.22092/ijwmse.2018.109070.1256 .(In Persian).

Azizi Jalilian, M., Shayesteh, K., Danehkar, A., & Salmanmahiny, A. (2020). A new ecosystem-based land classification of Iran for conservation goals. *Environmental Monitoring and Assessment*, *192*(3). https://doi.org/10.1007/s10661-020-8145-1

Baharvandi, K., Khorshiddoust, A. M., & Nassaji Zavareh, M. (2021). The Clarification of Climactic Fluctuations through SNHT Method (Case Study: Khorramabad Station). *Geography and Planning*, *25*(75), 51-63. doi: 10.22034/gp.2021.10829 .(In Persian).

Bazgeer, S., Abbasi, F., Asaadi Oskouei, E., Haghighat, M., Rezazadeh, P. (2019). Assessing the Homogeneity

1 Intergovernmental Panel on Climate Change

of Temperature and Precipitation Data in Iran with Climatic Approach, *Journal of Spatial Analysis Environmental Hazarts*, 6(1), 51-70.(In Persian).

Bazrafshan, J., & Cheraghalizadeh, M. (2021). Verification of abrupt and gradual shifts in Iranian precipitation and temperature data with statistical methods and stations metadata. *Environmental Monitoring and Assessment*, *193*(3), 1–19.

Buishand, T. A. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology*, *58*(1–2), 11–27.

Cai, W., Borlace, S., Lengaigne, M., Van Rensch, P., Collins, M., Vecchi, G., Timmermann, A., Santoso, A., McPhaden, M. J., & Wu, L. (2014). Increasing frequency of extreme El Niño events due to greenhouse warming. *Nature Climate Change*, *4*(2), 111–116.

Conrad, V., & Pollak, L. W. (1950). *Methods in Climatology*. Harvard University Press. https://doi.org/doi:10.4159/harvard.9780674187856

Elzeiny, R., Khadr, M., Zahran, S., & Rashwan, E. (2019). Homogeneity analysis of rainfall series in the upper blue nile river basin, Ethiopia. *Journal of Engineering Research*, *3*(September), 46–53.

Ghasemi, A. R. (2015). Changes and trends in maximum, minimum and mean temperature series in Iran. *Atmospheric Science Letters*, *16*(3), 366–372

Hänsel, S., Medeiros, D. M., Matschullat, J., Petta, R. A., & de Mendonça Silva, I. (2016). Assessing Homogeneity and Climate Variability of Temperature and Precipitation Series in the Capitals of North-Eastern Brazil. *Frontiers in Earth Science*, *4*. https://www.frontiersin.org/article/10.3389/feart.2016.00029

Intergovernmental Panel on Climate Change (IPCC). (2014). Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9781107415324

Khalil, A. (2021). Inhomogeneity detection in the rainfall series for the Mae Klong River Basin, Thailand. *Applied Water Science*, *11*(9), 147. https://doi.org/10.1007/s13201-021-01474-6

Kolendowicz, L., Czernecki, B., Półrolniczak, M., Taszarek, M., Tomczyk, A. M., & Szyga-Pluga, K. (2019). Homogenization of air temperature and its long-term trends in Poznań (Poland) for the period 1848–2016. *Theoretical and Applied Climatology*, *136*(3), 1357–1370. https://doi.org/10.1007/s00704-018-2560-z

Mahjoobi, E., Bakhshesh Robat, S., & Hosseinpour, M. (2021). A Review of Studies on the Effect of Teleconnections on Precipitation in Iran from 2004 to 2018. *Nivar*, *45*(112–113), 28–43. https://doi.org/10.30467/nivar.2021.246857.1167. (In Persian)

McPhaden, M. J. (1999). The child prodigy of 1997-98. *Nature*, *398*(6728), 559–561.

Paek, H., Yu, J., & Qian, C. (2017). Why were the 2015/2016 and 1997/1998 extreme El Niños different? *Geophysical Research Letters*, *44*(4), 1848–1856.

Mortezapour, S., Asaadi Oskouei, E., & Abbasi, F. (2020). Evaluation of some Homogeneity Tests on mutation detection in climatic data, case study: Rash station. *NIVAR*, 44(108-109), 12-31 .(In Persian).

Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *28*(2), 126–135.

Rafati, S., & Karimi, M. (2018). Assessment of homogenization of climate data and trend of temperature. *Journal of the Earth and Space Physics*, *44*(1), 199–214. https://doi.org/10.22059/jesphys.2017.61674

Reid, P. C., Hari, R. E., Beaugrand, G., Livingstone, D. M., Marty, C., Straile, D., Barichivich, J., Goberville, E., Adrian, R., & Aono, Y. (2016). Global impacts of the 1980s regime shift. *Global Change Biology*, *22*(2), 682–703.

Ribeiro, S., Caineta, J., & Costa, A. C. (2016). Review and discussion of homogenisation methods for climate data. *Physics and Chemistry of the Earth, Parts A/B/C*, *94*, 167–179.

Salehi, S., Dehghani, M., Mortazavi, S. M., & Singh, V. P. (2020). Trend analysis and change point detection of seasonal and annual precipitation in Iran. *International Journal of Climatology*, *40*(1), 308–323. https://doi.org/https://doi.org/10.1002/joc.6211

Suhaila, J., & Yusop, Z. (2018). Trend analysis and change point detection of annual and seasonal temperature series in Peninsular Malaysia. *Meteorology and Atmospheric Physics*, *130*(5), 565–581. https://doi.org/10.1007/s00703-017-0537-6

Wijngaard, J. B., Klein Tank, A. M. G., & Können, G. P. (2003). Homogeneity of 20th century European daily temperature and precipitation series. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *23*(6), 679–692.

Worsley, K. J. (1979). On the likelihood ratio test for a shift in location of normal populations. *Journal of the American Statistical Association*, *74*(366a), 365–367.

Investigating the homogeneity of temperature and precipitation data using statistical and statistical-climatic approaches

EXTENDED ABSTRACT

Introduction

Managing and policymaking in studies pertaining to natural hazards, such as drought and floods, are significantly influenced by the quality of the data under examination. Consequently, the paramount consideration in addressing management issues lies in ensuring the reliability and quality of the requisite data. The initial step in ensuring the quality control of the data under scrutiny involves verifying data homogeneity. Discontinuities within homogenous time series can be attributed to climatic factors. In contrast, heterogeneous time series exhibit discontinuities due to non-climatic factors, thereby amplifying the level of uncertainty in the findings. The accurate interpretation and reduction of uncertainty in such studies necessitate a meticulous assessment of data quality.

Methodology

This study aimed to examine the homogeneity of temperature variables (average, minimum, and maximum) as well as precipitation within the timeframe spanning from 1990 to 2019. This investigation was carried out using both statistical and statistical-climatic approaches across a dataset comprising 70 synoptic stations in Iran. The statistical approach employed a battery of homogeneity tests, which encompassed Petit homogeneity (PT), homogeneity of cumulative deviations (CDT), Worsley likelihood homogeneity (WLR), standard normal homogeneity (SNHT), and a statistical-climatic approach based on PT in conjunction with data from adjacent stations.

Results and Discussion

The results of the homogeneity tests conducted through the statistical approach revealed that the most frequent breaking year for temperature and precipitation variables was 1997 and 2006, respectively. Significantly, over 90% of the breaking years for temperature data (including minimum, maximum, and average) were deemed statistically significant. Specifically, the highest frequency of breaking years for minimum and maximum temperature occurred in 1997, 2008, and 2000, while for average temperature, it was in 1997, 1994, and 2009. In contrast, less than 15% of the breaking years for precipitation were considered significant, with the highest frequency of breaking years occurring in 2006, 1999, and 1997. Regarding temperature, it was found that 4.3% of stations exhibited homogeneity in minimum temperature, while 92.9% displayed heterogeneity. Other stations were categorized as doubtful. For average temperature, 10% were homogeneous, and 90% were heterogeneous. The analysis of precipitation indicated that 1.87% of stations were homogeneous, 11.4% were heterogeneous, and 1.5% fell into the doubtful category. In the homogeneity investigation based on the statistical-climatic approach utilizing adjacent stations, it was determined that 10% of minimum temperature stations (17.1% of maximum temperature stations) were homogeneous, 64.3% of minimum temperature stations (72.9% of maximum temperature stations) were conditionally homogeneous, and 25.7% of minimum temperature stations (10.0% of maximum temperature stations) were heterogeneous. For average temperature, 11.4% were homogeneous, 91.4% were conditionally homogeneous, and 14.3% were heterogeneous. Similarly, for precipitation, 74.3% were conditionally homogeneous, and 8.6% were heterogeneous.

Conclusion

This study aimed to assess the homogeneity of annual temperature and precipitation data using both statistical and statistical-climatic approaches, with no prior knowledge of metadata or climatic factors. In the homogeneity assessment based on the statistical approach, it was observed that over 90% of the temperature data and precipitation data were likely to be heterogeneous, which could potentially impact data-driven research. Therefore, it is imperative to investigate both climatic and non-climatic factors during the studied period to better understand the data's characteristics. Conversely, recognizing that climatic signals can influence broad geographical areas, it becomes feasible to attribute similar breaking years in a region to climatic factors, irrespective of their specific causes. In such instances, heterogeneity can be regarded as conditional homogeneity, with the heterogeneity factor considered as part of the climatic norm. The results obtained from the statistical-climatic approach indicated that 75% of the heterogeneous temperature data and 100% of the precipitation data could potentially be treated as conditionally homogeneous. However, it is advisable to conduct further in-depth investigations into the statistical-climatic approach to ensure its robustness and reliability.

Keywords: Adjacent station, Iran, Statistical-climate, Climate signal, Homogeneity.