Sari Agriculture Sciences and Natural Resources University

Journal of Watershed Management Research, Vol. 16, Issue 1, 2025 p: 121-132

OPEN ACCESS

Research Paper

Spectral Clustering of Precipitation Time Series in Golestan Province

Nader Jandaghi[1], Mojtaba Ghare Mahmoodlu[2], Majid Azimmohseni[3] and Mahnaz Khalafi[4]

1- Assistant Professor, Department of Rangeland and Watershed Management, Faculty of Agriculture and Natural Resources, University of Gonbad Kavous, Gonbad Kavous, Iran,
(Corresponding author: nader.jandaghi@gmail.com)
2- Associate Professor, Department of Rangeland and Watershed Management, Faculty of Agriculture and Natural Resources, University of Gonbad Kavous, Gonbad Kavous, Iran
3- Associate Professor, Faculty of Sciences, Department of Statistics, Golestan University, Gorgan, Iran
4- Assistant Professor, Faculty of Sciences, Department of Statistics, Golestan University, Gorgan, Iran

Received: 16 April 2024 Revised: 05 August 2024 Accepted: 31 August 2024

Extended Abstract

Background: Clustering time series of precipitation and other hydrological elements by direct use of classic methods, such as K-means, can be misleading because there is a time-lagged correlation in time series observations that is ignored in classic methods. A periodically correlated time series is a type of non-stationary time series with a periodic covariance function. Since there is periodic behavior in time series observed in hydrology, meteorology, climatology, etc., the use of periodically correlated time series has attracted the attention of experts in recent years. Time series data can be studied with two different approaches, time domain and frequency domain. Frequency domain is usually used to identify the structure of a time series. In this approach, time series are studied using Fourier transforms, which are functions of frequency. The clustering of meteorological stations in terms of precipitation time series provides important information about a geographical area and plays an important role in water resources management in that area. The main purpose of this research is to calculate the distance between the monthly precipitation of meteorological stations based on the observations of periodically correlated time series in the frequency domain and then use a clustering method to group the meteorological stations. For clustering, the fuzzy clustering method is compared with the usual k-means clustering method.

Methods: In this research, the monthly precipitation time series of 34 meteorological stations in Golestan Province were collected over a common period of 15 years from the Regional Water Company of Golestan. Considering that monthly precipitation values are time-dependent, these data were first arranged as a time series. Because of the dispersion of precipitation values and the presence of zero values in a number of meteorological stations, the one-to-one transformation was used to stabilize the variance. Since monthly precipitation data has a period of 12, they can be studied as periodically correlated time series. In this research, the periodically correlated time series of monthly precipitation was clustered in the spectral domain. First, the distance between the monthly precipitation times series of meteorological stations in the frequency domain was measured based on the periodic multiple correlation coefficient index. Then, clustering was performed by a fuzzy method based on the calculated distance matrix. After calculating the distance matrix, clustering was also done using the k-means method. The mean square error index of monthly precipitation was used and the results were compared to compare the accuracy of the clustering methods. All calculations of this research were done in MINITAB 17 and R 4.3.1 software.

Results: Investigating the seasonal trend for the monthly precipitation time series revealed a period of 12 in all the selected meteorological stations of Golestan Province; therefore, these data were analyzed based on periodically correlated time series. The distance between meteorological stations was determined based on the periodic multiple correlation index. Then, using two methods of fuzzy and k-means clustering, eight groups were identified for 34 selected meteorological stations in Golestan Province. The largest and smallest groups in fuzzy clustering included eight and one stations, respectively, and seven and one meteorological stations in the k-means method. Comparing the accuracy of the methods based on the mean square error index showed that the value of this index for the fuzzy clustering method was 13.67, while this index was calculated at 165.11 in the k-means method. In this research, it was also found that the accuracy of the spectral domain method along with fuzzy clustering was almost 12 times higher

Copyright ©2025 Jandaghi et al. Published by Sari Agricultural Sciences and Natural Resources University.
This work is licensed under a Creative Commons Attribution-Non-Commercial 4.0 Unported License, which allows users to read, copy, distribute, and make derivative works for non-commercial purposes from the material, as long as the author of the original work is cited properly.

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]

[DOI: 10.61882/jwmr.2024.1282]

than the spectral domain method based on the k-means method. Moreover, the investigation of the groups in the fuzzy clustering method shows the similarity of precipitation changes in the meteorological stations assigned to each group, which indicates the accuracy of the spectral domain method along with the fuzzy clustering.

Conclusion: In this research, the distance between the monthly percipitation time series of 34 meteorological stations in Golestan Province was measured in the frequency domain based on the alternating multiple correlation index. Then, the selected meteorological stations were clustered in the spectral domain based on fuzzy and k-means clustering methods, and eight groups were identified in each method. The mean square error index for fuzzy and k-means clustering methods was calculated at 13.67 and 165.11, respectively. These values showed that the accuracy of the fuzzy method for clustering the monthly percipitation time series of the selected meteorological stations was almost 12 times higher than the usual k-means method. The similar trend of changes in the monthly precipitation of meteorological stations assigned to each of the clusters shows the considerable efficiency of using the frequency domain method with fuzzy clustering for grouping periodically correlated time series such as monthly precipitation. In addition to the simplicity and accuracy of the clustering method in the frequency domain, the other advantage is to consider the periodic structure of time series in clustering. Furthermore, spectral clustering can be used for time series with unequal lengths recorded in hydrological studies.

Keywords: Fuzzy Clustering, Golestan Province, K-means Method, Monthly Precipitation, Seasonal Trend

How to Cite This Article: Jandaghi, N., Ghare Mahoodlu, M., Azimmohseni, M. M., & Khalafi, M. (2025). Spectral Clustering of Percipitation Time Series in Golestan Province. *J Watershed Manage Res*, 16(1), 121-132. DOI: 10.61882/jwmr.2024.1282

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]

[DOI: 10.61882/jwmr.2024.1282]

مقاله پژوهشی

خوشه‌بندی طیفی سری‌های زمانی بارش در استان گلستان

نادر جندقی[۱]، مجتبی قره محمودلو[۲]، مجید عظیم محسنی[۳] و مهناز خلقی[۴]

۱- استادیار، گروه مرتع و آبخیزداری، دانشکده کشاورزی و منابع طبیعی، دانشگاه گنبد کاووس، گنبد کاووس، ایران،
(نویسنده مسوول: nader.jandaghi@gmail.com)
۲- دانشیار، گروه مرتع و آبخیزداری، دانشکده کشاورزی و منابع طبیعی، دانشگاه گنبد کاووس، گنبد کاووس، ایران
۳- دانشیار، گروه آمار، دانشکده علوم پایه، دانشگاه گلستان، گرگان، ایران
۴- استادیار، گروه آمار، دانشکده علوم پایه، دانشگاه گلستان، گرگان، ایران

تاریخ دریافت: ۱۴۰۳/۰۱/۲۸ تاریخ ویرایش: ۱۴۰۳/۰۵/۱۵ تاریخ پذیرش: ۱۴۰۳/۰۶/۱۰
صفحه ۱۲۱ تا ۱۳۲

چکیده مبسوط

مقدمه و هدف: خوشه‌بندی سری‌های زمانی بارش و سایر عناصر هیدرولوژیکی با استفاده مستقیم از روش‌های کلاسیک مانند روش K میانگین می‌تواند گمراه‌کننده باشد زیرا در مشاهدات سری زمانی یک همبستگی تاخیری وجود دارد که در روش‌های کلاسیک نادیده گرفته می‌شود. سری‌های زمانی همبسته متناوب نوعی از سری‌های زمانی ناایستا هستند که در ساختار کوواریانس آن‌ها دوره تناوب مشاهده می‌شود. از آن‌جا که در سری‌های زمانی مربوط به داده‌های هیدرولوژی، هواشناسی، اقلیم‌شناسی و غیره رفتار تناوبی مشاهده می‌شود، لذا استفاده از سری‌های زمانی همبسته متناوب در سالیان گذشته مورد توجه متخصصین قرار گرفته‌است. سری‌های زمانی را می‌توان با دو رویکرد مختلف دامنه زمان و دامنه فرکانس مورد مطالعه قرارداد. از دامنه فرکانس معمولاً برای ساختارشناسی سری‌های زمانی استفاده می‌شود. در این رویکرد، سری‌های‌زمانی با استفاده از تبدیلات فوریه که توابعی از فرکانس هستند مورد مطالعه قرار می‌گیرند. خوشه‌بندی ایستگاه‌های هواشناسی از لحاظ سری‌های زمانی بارش اطلاعات مهمی را در مورد یک ناحیه جغرافیایی در اختیار کارشناسان قرار می‌دهد و نقش مهمی در مدیریت منابع آب آن ناحیه دارد. هدف اصلی این پژوهش محاسبه فاصله بین ایستگاه‌های هواشناسی استان گلستان بر اساس مشاهدات سری‌های زمانی همبسته متناوب بارش ماهانه ایستگاه‌های هواشناسی در دامنه فرکانس و سپس استفاده از یک روش خوشه‌بندی برای گروه‌بندی ایستگاه‌ها است. برای خوشه‌بندی نیز دو روش خوشه‌بندی فازی و روش معمول خوشه‌بندی k میانگین مقایسه می‌شوند.

مواد و روش‌ها: در این پژوهش، سری زمانی بارش ماهانه ۳۴ ایستگاه هواشناسی استان گلستان در یک بازه زمانی مشترک ۱۵ ساله از شرکت آب منطقه‌ای گلستان جمع‌آوری شد. با توجه به این‌که مقادیر بارش ماهانه وابسته به زمان هستند، ابتدا این داده‌ها به‌صورت سری زمانی مرتب شدند. با توجه به نوع پراکنش مقادیر بارش و وجود مقادیر صفر در تعدادی از ایستگاه‌های هواشناسی، برای تثبیت واریانس از تبدیل یک به یک در مدل‌سازی استفاده شد. از آن‌جا که سری‌های زمانی بارش ماهانه دارای دوره تناوب ۱۲ هستند لذا می‌توان آن‌ها را براساس سری‌های زمانی همبسته متناوب مورد مطالعه قرارداد. در این پژوهش، خوشه‌بندی سری‌های زمانی همبسته متناوب در دامنه طیفی مورد استفاده قرارگرفت. در ابتدا فاصله بین سری‌های زمانی بارش ماهانه ایستگاه‌های هواشناسی در دامنه فرکانس بر مبنای شاخص ضریب همبستگی چندگانه متناوب سنجیده شد. سپس بر اساس ماتریس فاصله، خوشه‌بندی به روش فازی انجام گرفت. پس از محاسبه ماتریس فاصله، خوشه‌بندی به‌روش k میانگین نیز انجام گرفت. برای مقایسه دقت روش‌های خوشه‌بندی مورد استفاده از شاخص میانگین مربعات خطا بارش ماهانه خوشه‌ها استفاده شد و نتایج آن‌ها باهم مقایسه شد. کلیه محاسبات این پژوهش در محیط نرم‌افزار مینی‌تب ۱۷ و آر ۴/۳/۱ انجام شدند.

یافته‌ها: نتایج بررسی روند فصلی برای سری‌های زمانی بارش ماهانه نشان دادند که در همه ایستگاه‌های هواشناسی منتخب در استان گلستان دوره تناوب ۱۲ وجود داشت. لذا می‌توان آن‌ها را بر اساس سری‌های زمانی همبسته متناوب مورد مطالعه قرار داد. بر اساس شاخص ضریب همبستگی متناوب فاصله بین ایستگاه‌های هواشناسی تعیین و با استفاده از دو روش خوشه‌بندی فازی و k میانگین، هشت گروه برای ۳۴ ایستگاه هواشناسی مورد بررسی در استان گلستان شناسایی شدند. بزرگ‌ترین و کوچک‌ترین گروه در خوشه‌بندی فازی به‌ترتیب شامل ۸ و ۱ ایستگاه و در روش k میانگین ۷ و ۱ ایستگاه هواشناسی بودند. نتایج بررسی دقت دو روش مورد استفاده بر اساس شاخص میانگین مربعات خطا نشان دادند که مقدار این شاخص برای خوشه‌بندی بر اساس روش فازی ۱۳/۶۷ بود در صورتی‌که این شاخص در روش k میانگین ۱۶۵/۱۱ محاسبه شد. هم‌چنین در این تحقیق مشخص شد که دقت روش دامنه طیفی به‌همراه خوشه‌بندی فازی تقریبا ۱۲ برابر بیشتر از روش دامنه طیفی براساس روش k میانگین بود. هم‌چنین، بررسی خوشه‌ها در روش فازی نشان‌دهنده شباهت روند تغییرات بارش در ایستگاه‌های هواشناسی واقع در هر خوشه بود که بر دقت روش دامنه طیفی به‌همراه خوشه‌بندی فازی دلالت دارد.

نتیجه‌گیری: در این پژوهش، فاصله بین سری‌های زمانی بارش ماهانه ۳۴ ایستگاه هواشناسی در استان گلستان در دامنه فرکانس بر مبنای شاخص همبستگی چندگانه متناوب سنجیده شد. سپس در دامنه طیفی و بر اساس دو روش خوشه‌بندی فازی و k میانگین ایستگاه‌های هواشناسی منتخب گروه‌بندی شدند و در هر یک از روش‌ها هشت گروه شناسایی شدند. بر این اساس، میانگین مربعات خطا برای روش خوشه‌بندی فازی و k میانگین به‌ترتیب ۱۳/۶۷ و ۱۶۵/۱۱ محاسبه شد. این موضوع نشان داد دقت روش خوشه‌بندی فازی برای گروه‌بندی مقادیر بارش ماهانه ایستگاه‌های هواشناسی منتخب تقریبا ۱۲ برابر بیشتر از روش معمول k میانگین بود. شباهت بسیار زیاد روند تغییرات بارش ماهانه ایستگاه‌های هواشناسی واقع در هر یک از گروه‌ها، بیانگر کارایی قابل ملاحظه استفاده از روش دامنه فرکانس و بر اساس خوشه‌بندی فازی برای گروه‌بندی سری‌های زمانی همبسته متناوب همانند بارش‌ماهانه است. علاوه بر سادگی و دقت روش خوشه‌بندی در دامنه فرکانس، از محاسن دیگر آن لحاظ کردن ساختار متناوب سری‌های زمانی در خوشه‌بندی است. اهمیت دیگر استفاده از این روش این است که در برخی مطالعات هیدرولوژی که طول سری‌های زمانی باهم برابر نباشند، در دامنه فرکانس امکان خوشه‌بندی آن‌ها وجود دارد.

واژه‌های کلیدی: استان گلستان، بارش ماهانه، خوشه‌بندی فازی، روش k میانگین، روند فصلی

مقدمه

سری‌های زمانی همبسته متناوب نوعی از سری‌های زمانی ناایستا هستند که در ساختار کوواریانس آن‌ها دوره تناوب مشاهده می‌شود (Najafiamiri *et al.*, 2024). از آن‌جا که در ساختار کواریانس سری‌های زمانی مربوط به داده‌های هیدرولوژی، هواشناسی، اقلیم‌شناسی و غیره رفتار تناوبی مشاهده می‌شود، لذا در سالیان اخیر استفاده از سری‌های زمانی همبسته متناوب مورد توجه متخصصین قرار گرفته‌است.

سری‌های زمانی را می‌توان با دو رویکرد مختلف مورد مطالعه قرارداد. رویکرد اول مطالعه سری زمانی در دامنه زمان

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]
[DOI: 10.61882/jwmr.2024.1282]

است که در آن، هدف اصلی پیش‌بینی سری زمانی براساس مقادیر گذشته آن است. اما رویکرد دیگر که معمولا در رشته‌های مهندسی مورد بررسی قرار می‌گیرد، مطالعه یک‌سری زمانی در دامنه فرکانس است (Ghezelsefla, 2022). از دامنه فرکانس معمولا برای ساختارشناسی سری‌های زمانی استفاده می‌شود. به‌عنوان مثال، می‌توان دوره تناوب سری زمانی را مشخص کرد (Sang *et al.*, 2009). در سال‌های اخیر، استنباط‌های آماری برای سری‌های زمانی در دامنه فرکانس گسترش زیادی پیدا کرده است (Percival & Walden, 2020). ابزاری که در دامنه فرکانس مورد استفاده قرار می‌گیرد تابع چگالی طیفی $f(\lambda)$ است که اطلاعاتی شبیه تابع خودهمبستگی در دامنه زمان دارد اما تابعی از فرکانس است و می‌توان آن‌را با تابع دوره‌نگار[1] برآورد کرد. سری‌های زمانی دارای وابستگی هستند و نمی‌توان از روش‌های معمول برای استنباط آماری در مورد آن‌ها استفاده کرد؛ اما در دامنه فرکانس به‌ازای فرکانس‌های مختلف دوره‌نگارها به‌طور مجانبی از هم مستقل هستند و می‌توان تحلیل‌های کلاسیک را در این دامنه انجام داد (Shumway & Stoffer, 2000).

در زمینه کاربرد دامنه فرکانس بر روی سری‌های زمانی هیدرولوژیک تاکنون مطالعات چندانی در ایران انجام نشده‌است. به‌خصوص تحلیل‌های آماری مبتنی بر نمایش سری زمانی در دامنه فرکانس بسیار محدود هستند. در این زمینه می‌توان به مطالعات زیر اشاره داشت.

میرزایی و همکاران (Mirzaee *et al.*, 2005) جهت تعیین زمان تاخیر رواناب کوتاه‌مدت و بلندمدت در حوضه جهان‌بین از تحلیل چگالی طیفی سری زمانی استفاده نمودند. نتایج نشان دادند که در این حوضه برای رژیم‌های سالیانه تأخیری معادل ۱/۲ ماه و در نوسانات کوتاه‌مدت بارش با دوره پریودهای ۳، ۴ و ۲ ماه، تأخیر حوضه ۰/۱۸، ۰/۵ و ۰/۰۸۳ بودند. نورانی و همکاران (Nourani *et al.*, 2016) اثر پارامترهای هیدروکلیماتولوژی را بر روی تراز آب دریاچه ارومیه با استفاده از آزمون من‌کندال مطالعه کردند. در این پژوهش، دوره تناوب ۸ ماهه به‌عنوان دوره تناوب موثر در روند تراز آب زیرزمینی شناخته شد. داداشی و همکاران (Dadashi Roudbar *et al.*, 2016) در بررسی تغییرات بارش حوضه آبریز هراز از روش‌های آماری و تحلیل طیفی جهت تعیین چرخه‌های بارش استفاده نمودند. نتایج مطالعات آن‌ها نشان دادند که علاوه بر چرخه‌های دوساله، چرخه‌های بلندمدت نیز بر بارش منطقه مؤثر بودند. دانشور وثوقی و شاکر (Daneshvar Vousoughi & Shaker, 2018) دوره تناوب تأثیرگذار را برای پارامتر تراز آب زیرزمینی در دشت اردبیل بررسی کردند. نتایج مطالعه آن‌ها نشان دادند که تناوب ۸ ماهه، به‌عبارت دیگر زیرسری D3، در روند این پارامتر در دشت مؤثر بود. دانشور وثوقی (Daneshvar Vousoughi, 2021) به بررسی تعیین تناوب تاثیرگذار در سری‌های زمانی بارش و رواناب در مقیاس ماهانه در ۳ ایستگاه بارش و یک ایستگاه هیدرومتری در حوزه دشت اردبیل در طول دوره آماری ۱۳۵۰ تا ۱۳۹۴ پرداخت. وی برای تعیین تناوب تأثیرگذار از ترکیب آزمون من‌کندال و تبدیل موجک استفاده نمود.

سلطانی و عظیم محسنی (Soltani & Azimmohseni, 2007) تبدیل فوریه و دوره‌نگار یک سری زمانی همبسته متناوب را به‌صورت یک بردار و ماتریس تعریف کردند که بر اساس آن‌ها می‌توان تحلیل‌های چندمتغیره مرتبط با سری‌های زمانی‌های همبسته متناوب را در دامنه فرکانس انجام داد. همچنین Hurd & Miamee (2007) با رویکردی متفاوت به تحلیل سری‌های زمانی همبسته متناوب در دامنه فرکانس پرداختند.

یکی از کاربردهای تبدیل فوریه در دامنه فرکانس، مقایسه سری‌های زمانی همبسته متناوب است. کایادو و همکاران (Caiado *et al.*, 2009) به مقایسه سری‌های زمانی با طول‌های نابرابر در دامنه فرکانس پرداختند. اهمیت این روش در این است که در برخی مطالعات هیدرولوژی که طول سری‌های زمانی باهم برابر نیست می‌توان آن‌ها را در دامنه فرکانس مقایسه نمود. عظیم‌محسنی و همکاران (Azimmohseni *et al.*, 2019) مقایسه دبی رسوب معلق را با متغیر کمکی دبی رودخانه براساس آنالیز کوواریانس طرح یک عامله در دامنه فرکانس انجام دادند. اما از آن‌جا که داده‌های هیدرولوژیکی دارای دوره تناوب هستند، در تحلیل‌های مربوط به این سری‌های زمانی باید دوره تناوب در نظر گرفته شود. محمودی و همکاران (Mahmoudi *et al.*, 2020) به مقایسه سری‌های زمانی همبسته متناوب در دامنه فرکانس پرداختند. آن‌ها براساس مشاهدات خشک‌سالی به روش RDI در ۱۶ ایستگاه هواشناسی سینوپتیک در کشور ایران به مقایسه ایستگاه‌ها در دامنه فرکانس پرداختند. پیشینه و اطلاعات مشاهداتی ایستگاه‌های هواشناسی منتخب با نتایج این تحقیق مطابقت داشت. قزل‌سفلی (Ghezelsefla, 2022) به مقایسه مقادیر دبی‌های ماهانه در دامنه زمان و فرکانس پرداخت. وی با انتخاب آمار دبی‌های متوسط ماهانه ۴ ایستگاه هیدرومتری در حوضه آبریز قره‌سو واقع در استان گلستان برای یک دوره ۳۶ ساله، به مقایسه سری‌های زمانی در دامنه فرکانس در قالب طرح بلوکی ساده پرداخت. نتایج نشان دادند که بین سری‌های زمانی مقادیر دبی متوسط ماهانه ۲ ایستگاه هیدرومتری غازمحله و پل تفاوت معنی‌داری در دامنه طیفی وجود نداشت. همچنین مشخص شد که بین سری‌های زمانی ایستگاه‌های هیدرومتری نهارخوران و سیاه‌آب با سایر ایستگاه‌ها در دامنه طیفی تفاوت معنی‌داری وجود داشت.

یکی از نتایج مهم مقایسه سری‌های زمانی، خوشه‌بندی آن‌ها است. به‌خصوص، در این پژوهش به خوشه‌بندی سری‌های زمانی بارش که همبسته متناوب است در دامنه فرکانس پرداخته می‌شود. با توجه به اینکه در سری‌های زمانی همبستگی وجود دارد، خوشه‌بندی با استفاده مستقیم از روش‌های کلاسیک مانند K میانگین می‌تواند گمراه‌کننده باشد (Gorgin, 2010). از طرفی، باید روشی ارائه شود که ساختار متناوب سری‌های زمانی را نیز در نظر بگیرد. از آن‌جا که در دامنه فرکانس بدون از دست دادن اطلاعات می‌توان سری‌های

1. Periodogram

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]
[DOI: 10.61882/jwmr.2024.1282]

زمانی را در فرکانس‌های مختلف به مشاهدات مستقل تبدیل کرد، از این‌رو می‌توان از روش‌های کلاسیک در این دامنه استفاده کرد (Chen *et al*., 2021). به‌خصوص با در نظر گرفتن دوره تناوب و تعریف بردارهای تبدیل فوریه، می‌توان فاصله بین ایستگاه‌های هواشناسی را براساس سری‌های زمانی بارش محاسبه نمود.

خوشه‌بندی سری‌های زمانی نقش مهمی را در تعیین گروه‌های همگن از لحاظ پدیده‌های دینامیکی دارند. یولاح و همکاران (Ullah *et al*., 2020) جهت تعیین مناطق همگن اقلیمی در کشور پاکستان از تحلیل خوشه‌ای و رویکرد L-moment براساس شاخص خشکسالی RDI استفاده نمودند. بر اساس نتایج این بررسی، پنج منطقه همگن اقلیمی در محدوده مورد مطالعه شناسایی شدند و مشخص شد که ارتفاع نقش کلیدی در الگوی بارندگی و خشکی مناطق داشت. در پژوهشی دیگر، نجفی و همکاران (Najafiamiri *et al*., 2024) چندین شاخص شباهت را در دامنه طیفی سری‌های زمانی همبسته متناوب معرفی کردند و به‌همراه روش فازی، خوشه‌بندی سری‌های زمانی همبسته متناوب بیان ژن در دامنه طیفی را ارائه دادند.

آب منبع حیاتی برای هر پدیده زیستی و انسانی و یکی از منابع مهم پایه و اساسی برای توسعه کشورها است (Yousefi & Mahdian, 2015). خوشه‌بندی صحیح سری‌های زمانی بارش می‌تواند نقش مهمی در مدیریت منابع آب به‌خصوص در منطقه‌ای مانند استان گلستان با تنوع آب و هوایی مختلف داشته باشد. لذا هدف اصلی این پژوهش محاسبه فاصله بین ایستگاه‌های هواشناسی بر اساس مشاهدات سری‌های زمانی بارش ماهانه ایستگاه‌های هواشناسی استان گلستان در دامنه فرکانس و سپس استفاده از یک روش خوشه‌بندی برای گروه‌بندی ایستگاه‌ها است. برای خوشه‌بندی نیز دو روش خوشه‌بندی فازی و روش معمول خوشه‌بندی k-میانگین مقایسه می‌شوند.

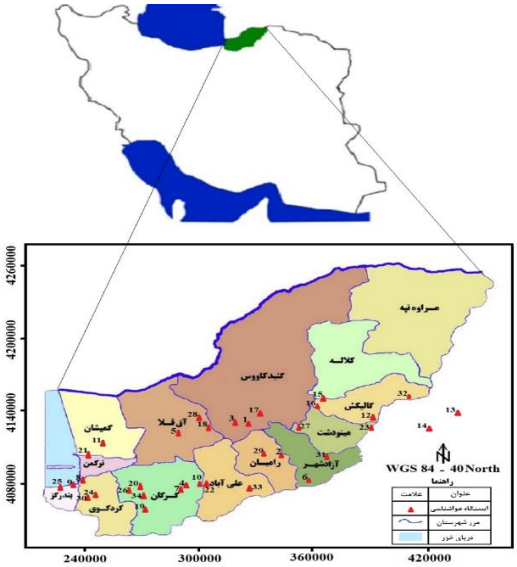
روش انجام تحقیق

محدوده مورد مطالعه

استان گلستان با مساحت ۲۰۴۳۸ کیلومتر مربع در شمال ایران و در جنوب شرق دریای خزر واقع شده است. این استان بین عرض جغرافیایی ´۳۰ °۳۶ تا ´۰۸ °۳۸ شمالی و طول جغرافیایی ´۵۷ °۵۳ تا´۲۲ °۵۶ شرقی واقع شده است. میزان بارندگی متوسط سالانه در جنوب استان گلستان ۷۰۰ میلی‌متر، در شمال استان ۲۰۰ میلی‌متر و میانگین آن ۴۵۰ میلی‌متر است (Jandaghi *et al*., 2019). در این پژوهش، در ابتدا آمار بارش ماهانه تمام ایستگاه‌های هواشناسی و تبخیرسنجی در محدوده مورد مطالعه از بدو تأسیس تا سال آبی ۱۳۹۶-۱۳۹۵ از شرکت آب منطقه‌ای استان گلستان جمع‌آوری شد. در نهایت، پس از بررسی کیفیت داده‌ها، تعداد ۳۴ ایستگاه هواشناسی برای این پژوهش انتخاب شدند که اطلاعات آن‌ها در جدول (۱) نمایش داده شده است. موقعیت ایستگاه‌های هواشناسی منتخب نیز بر روی شکل (۱) نمایش داده شده است. لازم به توضیح است که در این پژوهش، دوره آماری مشترک برای ایستگاه‌های هواشناسی منتخب ۱۵ ساله و از سال آبی۱۳۸۲-۱۳۸۱ تا ۱۳۹۵-۱۳۹۶ بود و مقادیر بارش به‌صورت ماهانه جمع‌آوری شد.

آماده‌سازی و مرتب کردن داده‌ها

با توجه به این که مقادیر بارش ماهانه وابسته به زمان هستند، در ابتدا این داده‌ها به‌صورت سری زمانی مرتب شدند. در این تحقیق، مقادیر بارش ۳۴ ایستگاه هواشناسی منتخب واقع در استان گلستان در یک دوره زمانی ۱۵ ساله به‌صورت سری زمانی مرتب شدند. لازم به توضیح است که با توجه به نوع پراکنش مقادیر بارش و وجود مقادیر صفر در تعدادی از ایستگاه‌های هواشناسی، برای تثبیت واریانس از تبدیل یک به یک $log(X_t + 1)$ در مدل‌سازی استفاده گردید (Jandaghi *et al*., 2021).



شکل ۱- نقشه موقعیت جغرافیایی ۳۴ ایستگاه هواشناسی منتخب در استان گلستان
Figure 1. The geographical location map of 34 selected meteorological stations in Golestan province

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]
[DOI: 10.61882/jwmr.2024.1282]

جدول ۱- مشخصات ایستگاه‌های هواشناسی منتخب در استان گلستان
Table 1. Characteristics of the selected meteorological stations in Golestan province

شماره ایستگاه Station Number	نام ایستگاه هواشناسی Meteorological Station Name	ارتفاع ایستگاه (متر) Station Height (m)	بارش متوسط ماهانه (میلی‌متر) Average Monthly Precipitation (mm)	مختصات UTM UTM Coordinates	
				طول Longitude	عرض Latitude
1	ارازکوسه	41	40.5	336111	4121413
2	نوده‌خرمالو	255	73.3	345784	4104520
3	قزاقلی	38	30.0	325009	4122323
4	تقی‌آباد گرگان	148	49.6	289060	4083488
5	آق‌قلا	-12	34.6	278102	4100003
6	تیل‌آباد	1000	21.7	363417	4086826
7	پل اردوگاه	444	60.7	284355	4073407
8	سیاه‌آب	-12	49.9	237187	4079985
9	باغوجنگل‌بانی	26	54.2	234440	4070700
10	فاضل‌آباد	202	59.0	299107	4085455
11	قلعه‌جیق	-20	32.1	250562	4114471
12	تنگراه	438	67.6	390288	4139471
13	رباط قره‌بیل	1295	16.8	439940	4133425
14	چشمه‌خان	1174	19.7	421774	4128174
15	تمر گرگان	183	50.2	367584	4150505
16	گالیکش	213	72.7	361086	4125346
17	گنبد	64	41.4	339743	4123727
18	بهلکه داشلی	16	33.3	304117	4103202
19	زیارت	885	42.9	274976	4065750
20	محوطه اداره	44	52.4	270830	4082460
21	غفارحاجی	34	38.8	245052	4098462
22	سرمو	500	66.0	305257	4076759
23	دشت شاد	1596	32.0	390361	4125971
24	کردکوی بالابلوک	76	58.3	243359	4072714
25	وطنا	90	60.3	235023	4067557
26	شصت‌کلاته	276	68.6	264810	4074080
27	مینودشت	210	70.7	358269	4120983
28	سد وشمگیر	4	28.1	299510	4119958
29	رامیان	208	81.3	334246	4098755
30	پل جاده غازمحله	6	48.3	241821	4075295
31	نرآب	1484	30.3	374176	4097435
32	پارک ملی گلستان	482	82.2	393568	4139903
33	شیرین‌آباد	882	48.1	323213	4073644
34	ناهارخوران	330	63.9	273886	4071982

روش‌های آماری مورد استفاده

در این پژوهش، پس از مشخص کردن این که سری‌های زمانی بارش همبسته متناوب هستند، فاصله بین تبدیلات فوریه سری‌های زمانی بارش ایستگاه‌های منتخب در دامنه فرکانس محاسبه و به‌عنوان فاصله بین میزان بارش دو ایستگاه در نظر گرفته شد. فاصله بین تبدیلات فوریه سری‌های زمانی همبسته متناوب توسط شاخص همبستگی چندگانه متناوب انجام گرفت و سپس بر اساس ماتریس فاصله محاسبه‌شده بین ایستگاه‌ها، خوشه‌بندی فازی و k میانگین انجام شد. در ادامه به جزئیات بیشتر از روش‌های ذکر شده پرداخته می‌شود.

در این پژوهش، پس از انتخاب ۳۴ ایستگاه هواشناسی، بررسی کیفیت و آماده‌سازی داده‌ها، برای تعیین دوره تناوب بارش‌های ماهانه ایستگاه‌های منتخب، از نمودارهای خودهمبستگی نگار[1] استفاده شد. این نمودارها در محیط نرم‌افزار MINITAB 17 ترسیم و به کمک آن‌ها وجود یا عدم وجود دوره تناوب فصلی بارش ایستگاه‌ها تعیین شد. وجود این الگوهای فصلی در مقایسه سری‌های زمانی در دامنه فرکانس مورد استفاده قرار می‌گیرد.

بر اساس یک سری زمانی $X_1,\ldots,X_n$ تابع دوره‌نگار به‌صورت رابطه (۱) تعریف می‌شود:

$$\boldsymbol{I(\lambda)=\frac{1}{n}\left|\sum_{i=1}^{n}X_t e^{-it\lambda}\right|^2=|d_X(\lambda)|^2} \qquad (۱)$$

که در آن $d_X(\lambda)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i$ تبدیل فوریه سری زمانی در دامنه فرکانس است. با رسم نمودار $I(\lambda)$ بر حسب فرکانس‌ها می‌توان دوره تناوب موثر سری زمانی را مشخص کرد.

سری‌های زمانی همبسته متناوب

سری زمانی همبسته متناوب به سری زمانی اطلاق می‌شود که در ساختار آن دوره تناوب وجود داشته باشد، مانند داده‌های بارش ماهانه که دارای دوره تناوب ۱۲ است. سری‌های زمانی همبسته متناوب در مطالعات هیدرولوژیکی کاربرد فراوانی دارند.

1- Autocorrelation plot

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]
[DOI: 10.61882/jwmr.2024.1282]

یک سری زمانی $X_t\,;\ t \in Z$ همبسته متناوب با دوره تناوب T نامیده می‌شود که با رابطه (۲) معرفی می‌شود:

$$\boldsymbol{Cov(X_t, X_s) = Cov(X_{t+T}, X_{s+T})} \qquad (۲)$$

برای یک سری زمانی همبسته متناوب $X_1, \dots, X_n$ با دوره تناوب T در دامنه طیفی یا دامنه فرکانس، بردار تبدیل فوریه به‌صورت رابطه (۳) ارائه شده است (Soltani & Azimmohseni, 2007):

$$\boldsymbol{d_X(\lambda) = (d_X(\lambda), d_X\left(\lambda + \frac{2\pi}{T}\right), \dots, d_X\left(\lambda + \frac{2\pi(T-1)}{T}\right))',\ \lambda \in [0, \frac{2\pi}{T})} \qquad (۳)$$

همچنین، ماتریس دوره‌نگار برای برآورد ماتریس تابع چگالی طیفی سری‌های زمانی همبسته متناوب به‌صورت رابطه (۴) ارائه شده است:

$$\boldsymbol{I_X(\lambda) = d_X(\lambda) d_X^*(\lambda),\ \lambda \in [0, \frac{2\pi}{T})} \qquad (۴)$$

که در آن * نشانگر ترانهاد مزدوج است.

معیار شباهت سری‌های زمانی همبسته متناوب در دامنه طیفی (فرکانس)

نجفی و همکاران (Najafiamiri *et al.*, 2024) روش‌هایی را برای خوشه‌بندی سری‌های زمانی همبسته متناوب در دامنه طیفی ارائه داده‌اند. شاخص همبستگی چندگانه متناوب (PRV) بر اساس ضریب همبستگی چندگانه ساخته شده است و شاخصی برای ارزیابی شباهت بین دو فرآیند همبسته متناوب در دامنه طیفی است. چنانچه فرض شود دو فرآیند همبسته متناوب X_t و Y_t با دوره تناوب T و با تبدیلات فوریه $d_X(\lambda)$ و $d_Y(\lambda)$ باشند، به‌صورت رابطه (۵) تا (۷) تعریف می‌شود (Szekely & Rizzo. 2014):

$$PRV = \frac{\sum_{i,j=1}^{k} I_X(\lambda_i, \lambda_j)\bar{I}_Y(\lambda_i, \lambda_j)}{\sqrt{\sum_{i,j=1}^{k}\left|I_X(\lambda_i, \lambda_j)\right|^2}\sqrt{\sum_{i,j=1}^{k}\left|I_Y(\lambda_i, \lambda_j\right.}} \qquad (۵)$$

$$I_X(\lambda_i, \lambda_j) = \sum_{j=0}^{T-1} d_X\left(\lambda_i + \frac{2\pi j}{T}\right)\overline{d_X\left(\lambda_j + \frac{2\pi J}{T}\right)} \qquad (۶)$$

$$I_Y(\lambda_i, \lambda_j) = \sum_{j=0}^{T-1} d_Y\left(\lambda_i + \frac{2\pi j}{T}\right)\overline{d_Y\left(\lambda_j + \frac{2\pi J}{T}\right)} \qquad (۷)$$

که در آن $\lambda_i, i = 1, \dots, k$ فرکانس‌هایی در فاصله $[0, \frac{2\pi}{T})$ هستند. همچنین توجه شود که $0 \le PRV \le 1$ است و هرچه PRV به ۱ نزدیک‌تر باشد دو سری زمانی به‌هم شباهت بیشتری دارند. در نتیجه، معیار فاصله بین دو سری زمانی همبسته متناوب از رابطه (۸) محاسبه می‌شود.

$$D = 1 - PRV \qquad (۸)$$

خوشه‌بندی فازی (C-means)

در این روش تخصیص یک مشاهده به یک خوشه با یک تابع عضویت بیان می‌شود. برای مشاهده iام، تابع عضویت به c خوشه با مقادیر $u_{1i}, \dots u_{ci}$ به‌صورت رابطه (۹) بیان می‌شود.

$$\sum_{j=1}^{c} u_{ji} = 1 \qquad (۹)$$

اگر n مشاهده چندمتغیره $X_1, \dots, X_n$ در اختیار باشد و قرار باشد آن‌ها به c خوشه تخصیص داده شوند، باید مقادیر $u_{ji}, j = 1, \dots, c\ \ i = 1, \dots, n$ به‌گونه‌ای انتخاب شود که تابع زیان را به حداقل مقدار خود برساند. در این حالت تابع زیان از رابطه (۱۰) محاسبه می‌شود:

$$L(u, v) = \sum_{i=1}^{n}\sum_{j=1}^{c} u_{ji}^m \left\|X_i - v_j\right\|^2 \qquad (۱۰)$$

که در آن u مجموعه مقادیر برای u_{ji}، v مجموعه بردارها برای مراکز خوشه، m پارامتر توان فازی که در بازه $[1, \infty]$ تغییر می‌نماید و درجه فازی بودن تخصیص مشاهدات را تنظیم می‌کند، و $\|\,.\,\|$ فاصله بین دو بردار را محاسبه می‌کند.

الگوریتم خوشه‌بندی فازی به‌صورت یک رابطه بازگشتی بردار مراکز خوشه و مقادیر عضویت را برای رسیدن به بهترین پاسخ به‌روزرسانی می‌کند (Miyamoto *et al.*, 2008).

لازم به ذکر است که کلیه الگوریتم‌های خوشه‌بندی از جمله خوشه‌بندی فازی را می‌توان تنها با در اختیار داشتن فاصله بین مشاهدات بازنویسی کرد و نیازی به دانستن خود مشاهدات نیست. در این تحقیق، پس از محاسبه فواصل بین سری‌های زمانی همبسته متناوب، از این روش خوشه‌بندی برای تکمیل خوشه‌بندی سری‌های زمانی همبسته متناوب استفاده شد.

در این پژوهش پس از محاسبه ماتریس فاصله و خوشه‌بندی به‌روش فازی، خوشه‌بندی به‌روش k–میانگین نیز انجام گرفت و با یکدیگر مقایسه شد. مقایسه دقت روش خوشه‌بندی بر اساس شاخص میانگین مربعات خطای بارش ماهانه (MSE) خوشه‌ها انجام و نتایج مقایسه گردیدند (Moriasi *et al.*, 2007).

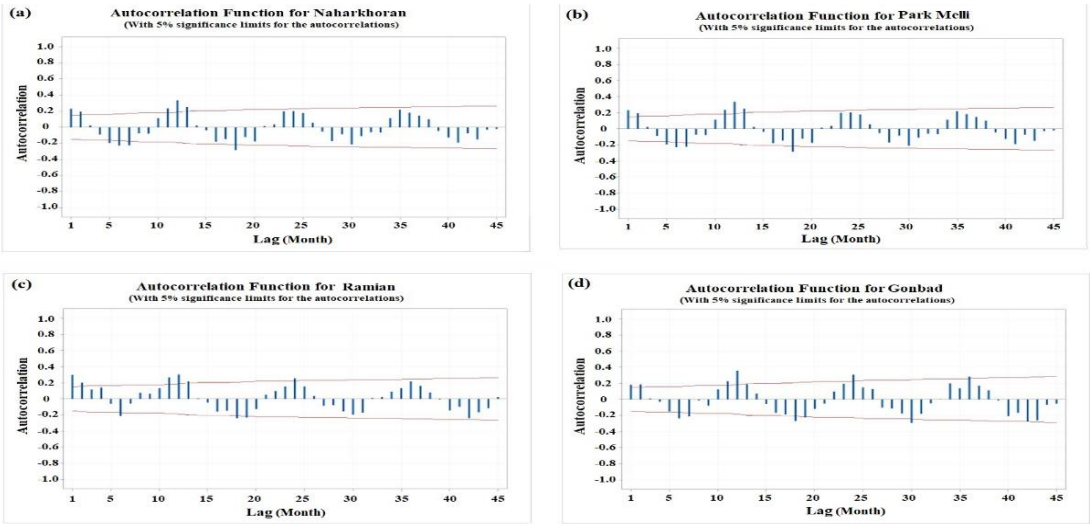
نتایج و بحث

بررسی روند فصلی سری‌های زمانی بارش ماهانه نشان داد که در همه ایستگاه‌های هواشناسی منتخب دوره تناوب ۱۲ وجود داشت. شکل (۲)، نمودار خودهمبستگی نگار مربوط به بارش ماهانه را در چهار ایستگاه هواشناسی نهارخوران، پارک ملی گلستان، رامیان و گنبد که در قسمت‌های مختلف استان واقع شده‌اند، به تفکیک نشان می‌دهد. با توجه به شکل (۲) می‌توان متناوب بودن سری‌های زمانی را مشاهده نمود. وجود این الگوهای فصلی نه تنها در مدل‌های باکس و جنکینز بلکه در مقایسه سری‌های زمانی در دامنه فرکانس نیز مورد استفاده قرار گرفت.

خوشه‌بندی ۳۴ ایستگاه هواشناسی از لحاظ الگوی سری زمانی مقادیر بارش‌های ماهانه در دامنه فرکانس و بر اساس بردار تبدیل فوریه انجام شد. فرکانس‌ها در فاصله $[0, \frac{2\pi}{12}]$ به‌صورت فرکانس‌های فوریه در نظر گرفته شدند ($\lambda_k = \frac{2\pi k}{m}$, $k = 0, \dots, m-1$) که در آن $m = \frac{n}{12}$ است. سپس شاخص

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]
[DOI: 10.61882/jwmr.2024.1282]

همبستگی چندمتغیره متناوب PRV برای هر زوج سری زمانی بارش ماهانه در فرکانس‌های فوریه محاسبه و بر اساس آن ماتریس فاصله بین کلیه ایستگاه‌ها تعیین شد. بر اساس روش خوشه‌بندی فازی، هشت خوشه برای ۳۴ ایستگاه هواشناسی موردبررسی در استان گلستان حاصل شد. جدول (۲) گروه‌بندی ایستگاه‌های هواشناسی مورد بررسی در استان گلستان را با استفاده از روش خوشه‌بندی فازی (C-means) نشان می‌دهد.



شکل ۲- نمودارهای خودهمبستگی نگار بارش در ایستگاه‌های نهارخوران (a)، پارک ملی گلستان(b)، رامیان (c) و گنبد (d)
Figure 2. Autocorrelation plots of precipitation in Naharkhoran (a), Golestan National Park (b), Ramiaan (c), and Gonbad (d) stations

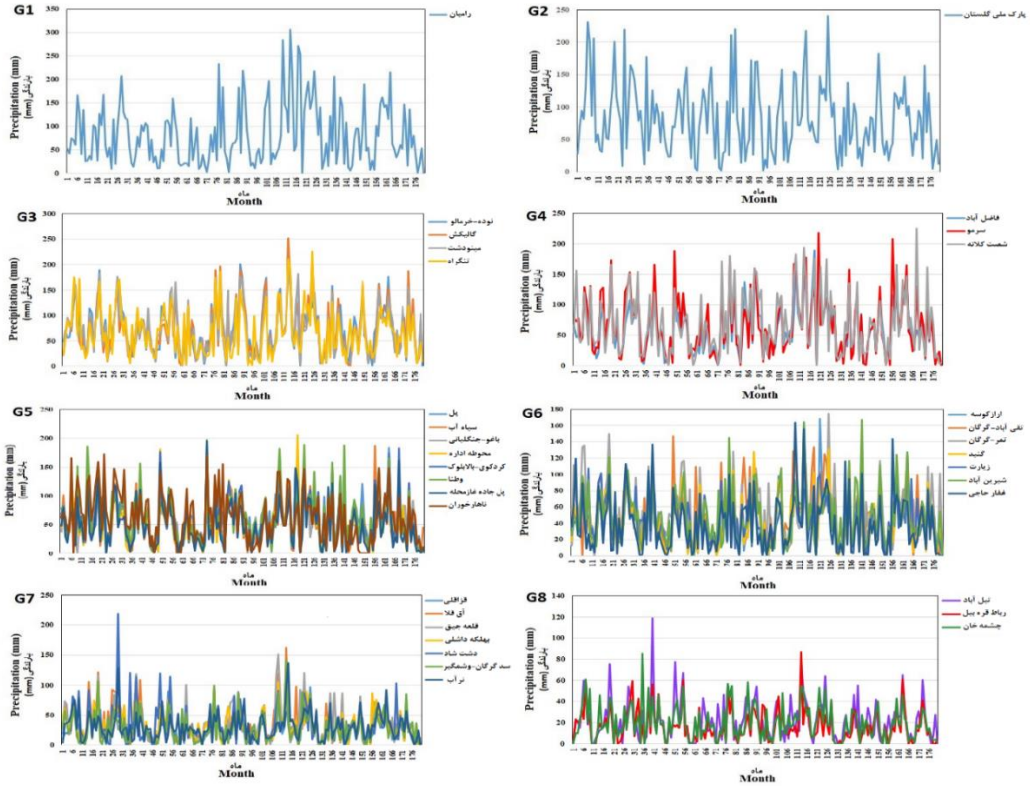
جدول ۲- گروه‌بندی ایستگاه‌های هواشناسی در استان گلستان با استفاده از روش خوشه‌بندی فازی (C-means)
Table 2. Grouping of meteorological stations in Golestan Province using the fuzzy clustering method (C-means)

شماره گروه Group Nnmber	نام ایستگاه‌های هواشناسی Meteorological Station Name	تعداد ایستگاه هواشناسی Number of Meteorological Stations	میانگین بارش ماهانه گروه Average Monthly Precipitation of the Group (mm)
1	رامیان	1	81.3
2	پارک ملی گلستان	1	82.2
3	تنگراه، مینودشت، گالیکش و نوده-خرمالو	4	71.1
4	شصت‌کلاته، سرمو و فاضل‌آباد	3	64.5
5	محوطه اداره، باغو-جنگلبانی، سیاهآب، پل اردوگاه، نهارخوران، پل جاده غازمحله، وطنا و کردکوی-بالابلوک	8	56.0
6	غفارحاجی، شیرین‌آباد، زیارت، گنبد، تمر-گرگان، تقی‌آباد-گرگان و ارازکوشه	7	44.5
7	نرآب، سدگرگان وشمگیر، دشت‌شاد، بهلکه داشلی، قلعه‌جیق، آق‌قلا و قزاقلی	7	31.5
8	چشمه‌خان، رباط‌قره‌بیل و تیل‌آباد	3	19.4

بر اساس جدول (۲)، بیشترین میانگین بارش ماهانه مربوط به گروه ۲ با ۸۲/۲ میلی‌متر و مربوط به ایستگاه هواشناسی پارک ملی گلستان است و کمترین میانگین بارش ماهانه ۱۹/۴ میلی‌متر به گروه ۸ مربوط به ایستگاه‌های هواشناسی چشمه‌خان، تیل‌آباد و رباط‌قره‌بیل تعلق دارد. همچنین، با مقایسه سری‌های زمانی بارش‌های ماهانه مشخص شد که گروه ۵ با هشت ایستگاه هواشناسی به‌عنوان بزرگترین گروه و گروه‌های ۱ و ۲ هرکدام تنها با یک ایستگاه به‌عنوان کوچکترین گروه در دامنه فرکانس شناسایی شدند.

شکل (۳) نمودارهای سری زمانی مقادیر بارش ماهانه ایستگاه‌های هواشناسی مورد بررسی را در هریک از هشت گروه که با استفاده از روش خوشه‌بندی فازی در دامنه طیفی شکل گرفته است، نشان می‌دهد. با توجه به شکل (۳)، می‌توان شباهت تغییرات روند بارش ماهانه ایستگاه‌های واقع‌شده در هر یک از گروه‌ها را به‌وضوح مشاهده نمود.

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]
[DOI: 10.61882/jwmr.2024.1282]



شکل ۳- نمودارهای گروه‌بندی شده سری‌های زمانی بارش ماهانه ایستگاه‌های هواشناسی با روش خوشه‌بندی فازی (C-means)
Figure 3. Grouped graphs of monthly precipitation time series of meteorological stations with the fuzzy clustering method (C-means)

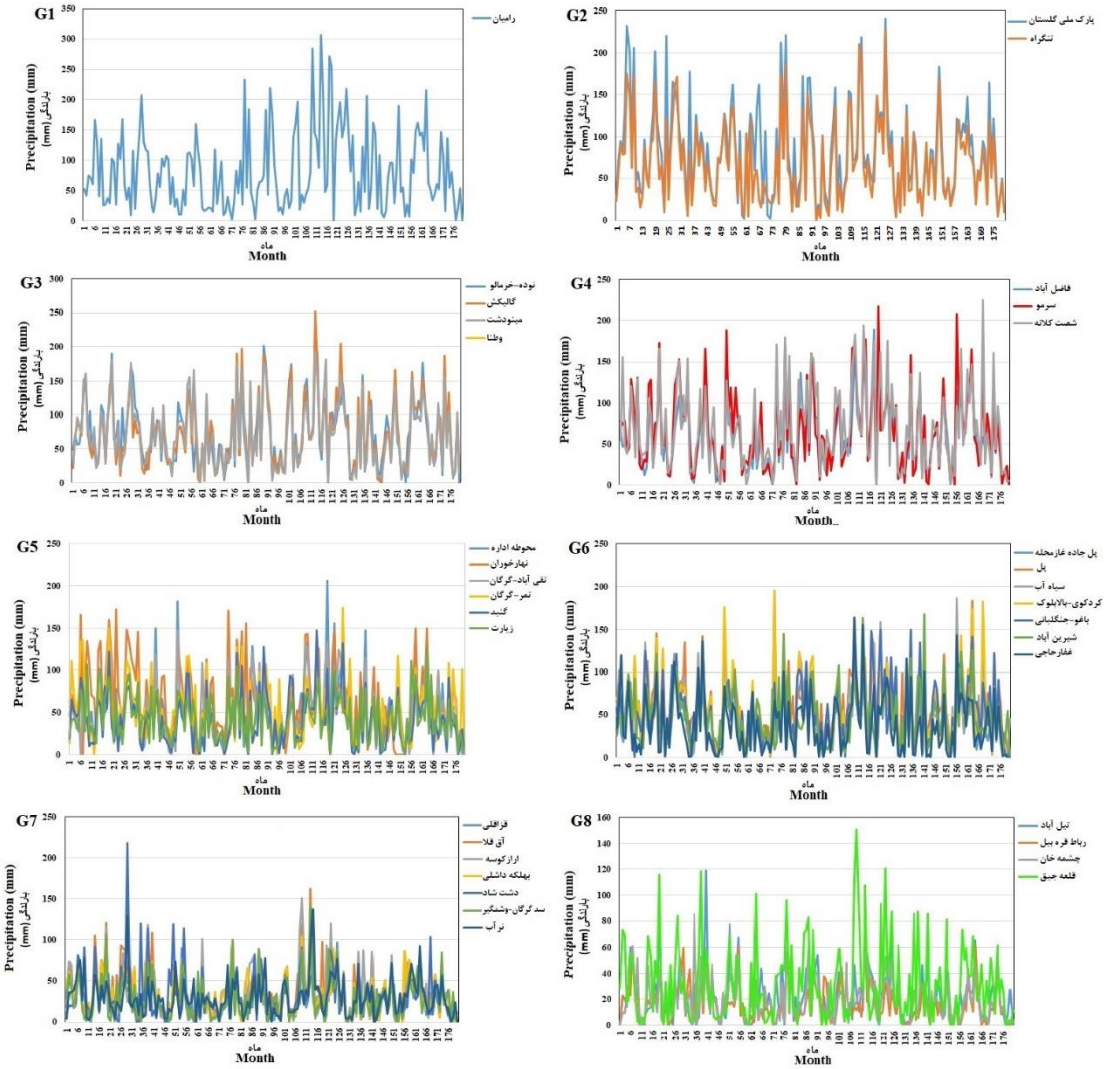
شباهت بسیار زیاد روند تغییرات بارش ماهانه ایستگاه‌های هواشناسی واقع در هریک از گروه‌ها بیانگر کارایی قابل ملاحظه استفاده از روش دامنه فرکانس و بر اساس بردار تبدیل فوریه به همراه خوشه‌بندی فازی (C-means) را برای گروه‌بندی سری‌های زمانی همبسته متناوب همانند بارش‌های ماهانه است. لذا، از این روش می‌توان برای گروه‌بندی سایر سری‌های زمانی همبسته متناوب مانند دبی‌های متوسط ماهانه، دمای متوسط ماهانه و غیره که دارای الگوی فصلی هستند نیز استفاده نمود.

پس از محاسبه ماتریس فاصله در دامنه طیفی، خوشه‌بندی به‌روش معمول k-میانگین نیز انجام گرفت. جدول (۳) نتایج گروه‌بندی ایستگاه‌های هواشناسی منتخب را به روش k-میانگین نشان می‌دهد. براساس روش k-میانگین، ایستگاه‌های هواشناسی در هشت گروه تقسیم شده‌اند که بزرگترین و کوچکترین گروه به‌ترتیب شامل ۷ و ۱ ایستگاه هواشناسی هستند. همچنین بیشترین و کمترین بارش ماهانه ۸۱/۳ و ۲۲/۶ میلی‌متر به‌ترتیب به گروه های ۱ و ۸ تعلق دارند. همان طور که ملاحظه می‌شود، ایستگاه‌های هواشناسی گروه‌های اول و سوم در هر دو روش مشابه هم هستند ولی در سایر گروه‌ها تفاوت‌هایی وجود دارند. شکل (۴) نمودارهای گروه‌بندی شده سری زمانی مقادیر بارش ماهانه ایستگاه‌های هواشناسی را با روش معمول k-میانگین نشان می‌دهد.

جدول ۳- گروه‌بندی ایستگاه‌های هواشناسی در استان گلستان با استفاده از روش K-means
Table 3. Grouping of meteorological stations in Golestan Province using the K-means method

شماره گروه Group Nnmber	نام ایستگاه‌های هواشناسی Meteorological Station Name	تعداد ایستگاه هواشناسی Number of Meteorological Stations	میانگین بارش ماهانه گروه Average Monthly Precipitation of the Group (mm)
1	رامیان	1	81.3
2	پارک ملی گلستان، تنگراه	2	74.9
3	مینودشت، گالیکش، نوده-خرمالو و وطنا	4	69.2
4	شصت‌کلاته، سرمو و فاضل‌آباد	3	64.5
5	محوطه اداره، نهارخوران، زیارت، گنبد، تمر-گرگان، تقی‌آباد-گرگان	6	50.1
6	غفارحاجی، شیرین‌آباد، باغو-جنگلبانی، سیاه‌آب، پل اردوگاه، پل جاده غازمحله و کردکوی-بالابلوک	7	51.2
7	نرآب، سدگرگان-وشمگیر، دشت‌شاد، بهلکه داشلی، ارازکوشه، آق‌قلا و قزاقلی	7	32.7
8	چشمه‌خان، رباط‌قره‌بیل، تیل‌آباد و قلعه‌جیق	4	22.6

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]
[DOI: 10.61882/jwmr.2024.1282]



شکل ۴- نمودارهای گروه‌بندی شده سری‌های زمانی بارش ماهانه ایستگاه‌های هواشناسی با روش K-means
Figure 4. Grouped graphs of monthly precipitation time series of meteorological stations with the K-means method

مطابق شکل (۴)، ناهماهنگی در تغییرات روند بارش در برخی از ایستگاه‌های هواشناسی واقع در برخی خوشه‌ها قابل مشاهده است. برای بررسی و مقایسه دقیق‌تر دو روش مورد استفاده، میانگین مربعات خطای (MSE) مربوط به خوشه‌ها بر اساس میانگین بارش ماهانه ایستگاه‌های هواشناسی محاسبه شد. بر این اساس، میانگین مربعات خطا برای خوشه‌بندی بر اساس روش فازی (C-means) ۱۳/۶۷ محاسبه شد در صورتی‌که این کمیت در روش معمول k-میانگین ۱۶۵/۱۱ محاسبه شد. این نتایج بیانگر شباهت بیشتر ایستگاه‌های هواشناسی در خوشه‌های روش فازی بر اساس میانگین بارش ماهانه نسبت به روش k-میانگین هستند. در این تحقیق، همچنین مشخص شد که دقت روش دامنه طیفی به‌همراه روش فازی تقریباً ۱۲ برابر بیشتر از روش دامنه طیفی بر اساس روش k-میانگین بود.

نتیجه‌گیری کلی

در این پژوهش، در ابتدا با بررسی نمودارهای خودهمبستگی نگار مشخص شد که برای ۳۴ ایستگاه هواشناسی منتخب روند فصلی با دوره تناوب ۱۲ ماهه وجود داشت. سپس بردارهای تبدیلات فوریه سری‌های زمانی همبسته متناوب بارش محاسبه و فاصله بین سری‌های زمانی بارش ماهانه ایستگاه‌ها بر اساس شاخص PRV در دامنه فرکانس محاسبه گردید. بر اساس ماتریس فاصله حاصل شده، این ایستگاه‌ها از لحاظ بارش به‌روش فازی خوشه‌بندی شدند و با روش معمول k-میانگین با استفاده از شاخص میانگین مربعات خطا (MSE) مورد مقایسه قرار گرفتند.

بر اساس ماتریس فاصله محاسبه‌شده و متعاقب آن استفاده از خوشه‌بندی فازی، هشت خوشه برای ۳۴ ایستگاه هواشناسی مورد بررسی در استان گلستان حاصل شد. شباهت بسیار زیاد روند تغییرات بارش ماهانه ایستگاه‌های هواشناسی واقع در هریک از گروه‌ها بیانگر کارایی قابل ملاحظه استفاده از روش

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]
[DOI: 10.61882/jwmr.2024.1282]

دامنه فرکانس و براساس بردار تبدیل فوریه برای گروه‌بندی سری‌های زمانی همبسته متناوب همانند بارش‌های ماهانه است. سپس گروه‌بندی ایستگاه‌های هواشناسی منتخب به‌روش k–میانگین انجام و هشت گروه شناسایی شدند. بر این اساس، میانگین مربعات خطا (MSE) برای خوشه‌بندی بر اساس روش فازی (C-means) ۱۳/۶۷ محاسبه شد در صورتی‌که این کمیت در روش k–میانگین ۱۶۵/۱۱ محاسبه شد. در این تحقیق همچنین مشخص شد که دقت روش دامنه طیفی به‌همراه روش فازی تقریباً ۱۲ برابر بیشتر از روش دامنه طیفی بر اساس روش k–میانگین بود. این نتایج بیانگر شباهت بیشتر ایستگاه‌های هواشناسی در خوشه‌های روش فازی (C-means) بر اساس میانگین بارش ماهانه نسبت به روش k–میانگین هستند.

با توجه به نتایج این پژوهش، مشخص شد که استفاده از روش دامنه فرکانس انعکاس‌دهنده بسیار خوبی از رفتار سری‌های زمانی بود، به این معنی که دو ایستگاه با تبدیلات فوریه تقریباً یکسان (مثلا داده‌های بارش) دارای متوسط بارش تقریبا یکسانی هستند. به‌عنوان مثال، می‌توان به دو ایستگاه هواشناسی زیارت با موقعیت کوهستانی در جنوب غرب استان با ارتفاع ۸۸۵ متر و ایستگاه هواشناسی گنبد با موقعیت دشت در شمال شرق استان با ارتفاع ۶۴ متر اشاره داشت. این دو ایستگاه علی‌رغم داشتن موقعیت‌های بسیار متفاوت از نظر ارتفاعی و جغرافیایی که در مقدار بارش بسیار تأثیرگذار هستند، در گروه‌بندی با روش دامنه فرکانس در یک گروه واقع شدند و بررسی روند و الگوی بارش ماهانه آن‌ها، صحت این موضوع را تأیید می‌نماید. نتایج این تحقیق با مطالعات قزل‌سفلی (Ghezelsefla, 2022) که به مقایسه مقادیر دبی‌های ماهانه در دامنه فرکانس پرداختند و محمودی و همکاران (Mahmoudi *et al.,* 2020) که بر اساس مشاهدات خشکسالی به‌روش RDI در ۱۶ ایستگاه هواشناسی سینوپتیک در کشور ایران به مقایسه ایستگاه‌ها در دامنه فرکانس پرداختند، هم‌راستایی دارند.

از محاسن دیگر استفاده از روش خوشه‌بندی در دامنه فرکانس این است که می‌تواند ساختار متناوب سری زمانی را در خوشه‌بندی لحاظ کند. اهمیت دیگر استفاده از این روش این است که در برخی مطالعات هیدرولوژی که طول سری‌های زمانی با هم برابر نباشند، در دامنه فرکانس امکان خوشه‌بندی آن‌ها وجود دارد.

لذا با توجه به سادگی، دقت و دیگر محاسن ذکر شده در استفاده از روش دامنه فرکانس و بر اساس بردار تبدیل فوریه، پیشنهاد می‌شود برای خوشه‌بندی سری‌های زمانی همبسته متناوب مانند دبی متوسط ماهانه، دمای متوسط ماهانه و غیره که دارای الگوی فصلی هستند، استفاده شود. همچنین با توجه به این که شاخص PRV شاخصی مبتنی بر دوره‌نگارها است، پیشنهاد می‌شود که در مطالعات آینده از فرم وزنی آن استفاده شود که باعث ثبات بیشتر در نتایج خوشه‌بندی شود.

References

Azimmohseni, M., Khalafi, M., & Kordkatuli, M. (2019). Time series analysis of covariance based on linear transfer function models. *Statistical Inference for Stochastic Processes,* 22(1), 1-16. DOI: 10.1007/s11203-018-9182-z

Caiado, J., Crato, N., & Peña, D. (2009). Comparison of times series with unequal length in the frequency domain. *Communications in Statistics—Simulation and Computation,* 38(3), 527-540. DOI: 10.1080/03610910802562716

Chen, X., Wang, X., & Lian, J. (2021). Applicability Study of Hydrological Period Identification Methods: Application to Huayuankou and Lijin in the Yellow River Basin, China. *Water,* 13(9), 1265. https://doi.org/10.3390/w13091265

Dadashi Roudbar, A., Fallah Ghalhari, Gh., Karami, M., & Baaghide, M. (2016). Analysis of Precipitation Variations of Haraz Watershed Using by Statistical Methods and Spectrum Analysis Technique. *Hydrogeomorphology*, 3(7), 59-86. DOI: 20.1001.1.23833254.1395.3.7.4.0 [In Persian].

Daneshvar Vousoughi, F. (2021). Determination of Dominant Periods in Rainfall and Runoff Time Series (Case Study: Ardabil Plain). *Water Resources Engineering*, 13(4), 27-42. [In Persian]. DOI: 20.1001.1.20086377.1399.13.47.3.0

Daneshvar Vousoughi, F., & Shaker, R. (2018). Assessment of Trend in Groundwater Level using Hybrid Mann-Kendall and Wavelet Transform Method (Case Study: Ardabil Plain). *Environment and Water Engineering*, 4(3), 243-253. https://doi.org/10.22034/jewe.2018.112575.1221 [In Persian]

Ghezelsefla, H. (2022). Comparison of monthly discharges in time and frequency domains (Case study: Gharsu basin). Master's thesis in watershed management, Gonbad Kavous University, 54 p. [In Persian]

Gorgin, V. (2010). Least Squares Estimation and Analysis of Variance for Intermittently Correlated Time Series. Master's thesis of Statistics, Payam Noor University, Fars province.

Holan, S. H., & Ravishanker, N. (2018). Time series clustering and classification via frequency domain methods. *Wiley Interdisciplinary Reviews: Computational Statistics*, 10(6), e1444.

Hurd, H. L., & Miamee, A. (2007). Periodically correlated random sequences: spectral theory and practice (Vol. 355). John Wiley and Sons.

Jandaghi, N., Azimmohseni, M., & Ghareh Mahmoodlu, M. (2021). Rainfall-runoff process modeling using time series transfer function. *Environmental Erosion Research,* 11(2), 111-128. DOI: 20.1001.1.22517812.1400.11.2.5.5. [In Persian].

Jandaghi, N., Seyyedian, M., Mohamadi Ostadkelayeh, A., Fathabadi, A. Ghareh Mahmoodlu, M., & Mohammad esmaeili, M. (2019). Analysis and documentation of March 1997 flood in Gonbad Kavos city, Gonbad Kavous University, 126 p. [In Persian]

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]

[DOI: 10.61882/jwmr.2024.1282]

Mahmoudi, M. R., Maleki, M., Borodin, K., Pho, K. H. & Baleanu, D. (2020). On comparing and clustering the spectral densities of several almost cyclostationary processes, *Alexandria Engineering Journal,* 59(4), 2555-2565. https://doi.org/10.1016/j.aej.2020.03.043

Mirzaee, S. Y., Chitsazan, M., & Chinipardaz, R. (2005). Application of Time Series Analysis in Determination of Lag Time in Jahanbin Basin. *Jounal of Water and Wastewater,* 15(3), 53-59, [In Persian].

Miyamoto, S., Ichihashi, H., Honda, K., & Ichihashi, H. (2008). Algorithms for fuzzy clustering. *Heidelberg: Springer*, 10.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Transactions of the American Society of Agricultural and Biological Engineers,* 50(3), 885-900. DOI: 10.13031/2013.23153

Najafiamiri, F., Khalafi, M., Golalipour, M., & Azimmohseni, M. (2024). On clustering of periodically correlated processes based on Hilbert-Schmidt inner product of Fourier transforms. *Communications in Statistics-Simulation and Computation*, 53(10), 5028-5046.

Nourani, V., Azad, N., Ghasemzade, M., & Sharghi, E. (2016). Assessment of Ajichay Hydro-Climatologic Parameters Impacts on Urmia Lake Level Using Hybrid Wavelet-Mann Kendall. *Hydrogeomorphology,* 3(7), 141-159. DOI: 20.1001.1.23833254.1395.3.7.8.4. [In Persian]

Percival, D. B., & Walden, A. T. (2020). Spectral analysis for univariate time series. *Cambridge University Press*, 51.

Sang, Y. F., Wang, D., Wu, J. C., Zhu, Q. P., & Wang, L. (2009). The relation between periods' identification and noises in hydrologic series data. *Journal of Hydrology*, 368(1-4), 165-177. https://doi.org/10.1016/j.jhydrol.2009.01.042

Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. (2000). Time series analysis and its applications. *New York: springer*, 3, p 4.

Soltani, A. R. & Azimmohseni, M., (2007). Periodograms asymptotic distributions in periodically correlated processes and multivariate stationary processes: An alternative approach. *Journal of Statistical Planning and Inference*, 137(4), 1236-1242. https://doi.org/10.1016/j.jspi.2006.01.015

Szekely, G. J., & Rizzo, M. L. (2014). Partial distance correlation with methods for dissimilarities. *The Annals of Statistics*, 42(6), 2382–2412. doi :10.1214/14-AOS1255.

Ullah, H., Akbar, M., & Khan, F. (2020). Construction of homogeneous climatic regions by combining cluster analysis and L-moment approach on the basis of Reconnaissance Drought Index for Pakistan. *International Journal of Climatology*, 40(1), 324-341. https://doi.org/10.1002/joc.6214

Yousefi, A., & Mahdian, Sh. (2015). The Economic and Social Necessity of Water Reuse in Iran. *Journal of Water Reuse,* 1(1), 1-7. [In Persian]

[Downloaded from jwmr.sanru.ac.ir on 2025-09-01]

[DOI: 10.61882/jwmr.2024.1282]