

## پهنه‌بندی شرق دریاچه ارومیه بر مبنای شاخص توزیع زمانی بارش با روش وارد و K-Means

پروا محمدی<sup>۱\*</sup> و احمد فاخری فرد<sup>۲</sup>

### چکیده

با توجه به اهمیت بارش و کمبود منابع آب، تحلیل مکانی بارش‌های روزانه توأم با فواصل زمانی مربوطه، یکی از ضروریات می‌باشد. هدف این مطالعه، معرفی یک شاخص برای شناسایی مناطق مستعد از نظر توان دیم‌کاری، با استفاده از داده‌های توأم بارش و فواصل زمانی آن‌ها در شرق حوضه دریاچه ارومیه می‌باشد. با استفاده از مقدار بارش و فواصل زمانی آن، در دوره آماری سال ۱۳۷۰ تا ۱۳۹۲ شاخص توزیع زمانی بارش برای ۲۳ ایستگاه محاسبه شد. برای خوشه‌بندی شاخص توزیع زمانی بارش، روش‌های K-Means و وارد به کار گرفته شد. آزمون همگنی خوشه‌های به دست آمده از طریق روش آماره H انجام گرفت. مقایسه نواحی همگن حاصل از خوشه‌بندی با دو روش ذکر شده با توزیع مکانی خطوط هم‌شاخص نشان داد که روش K-Means نواحی همگن را بهتر از روش وارد تفکیک کرد. بخش‌های جنوبی، مرکزی، شمال شرقی و جنوب غربی ناحیه مورد مطالعه توانایی کشت دیم بیشتری را نسبت به سایر مناطق در شرق دریاچه ارومیه دارند.

**واژه‌های کلیدی:** آزمون همگنی، توزیع مکانی، خوشه‌بندی، دیم.

ارجاع: محمدی پ. و فاخری فرد ا. ۱۳۹۸. پهنه‌بندی شرق دریاچه ارومیه بر مبنای شاخص توزیع زمانی بارش با روش وارد و K-Means. مجله پژوهش آب ایران. ۳۳: ۴۴-۴۴.

۱- دانشجوی دکتری مهندسی منابع آب، گروه مهندسی آبیاری و آبادانی، دانشکده مهندسی و فناوری کشاورزی، دانشگاه تهران  
۲- استاد گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز.

\* نویسنده مسئول: [mohammadi.parva@yahoo.com](mailto:mohammadi.parva@yahoo.com)

تاریخ پذیرش: ۱۳۹۶/۰۱/۱۵

تاریخ دریافت: ۱۳۹۵/۰۸/۱۰

## مقدمه

برای کاهش تهدید نابودی منابع آب و در عین حال رفع نیازهای غذایی آینده مردم، اهمیت کشاورزی دیم اجتناب‌ناپذیر خواهد بود. در این راستا، یکی از اقدامات مؤثری که می‌توان انجام داد یافتن مناطق پرباران با امکان کشت دیم می‌باشد. با توجه به کمبود منابع آب، دیم‌کاری از اهمیت ویژه‌ای برخوردار است که در کشاورزی دیم، عامل بارش تنها منبع تأمین آب در دیم‌کاری محسوب می‌شود. میزان تولید محصولات کشاورزی همبستگی بالایی با نزولات جوی و مناسب بودن شرایط آب و هوایی دارد (دینار و همکاران، ۱۹۹۸). رومرو و همکاران (۱۹۹۹) برای پهنه‌بندی مناطق مدیترانه‌ای اسپانیا به نواحی همگن بارشی از آمار ۴۱۰ ایستگاه باران‌سنجی در طول دوره آماری ۱۹۹۳-۱۹۶۴ استفاده کردند. نتایج پهنه‌بندی منطقه مورد مطالعه، ۱۲ منطقه بارشی را معرفی کرد. سدراس و همکاران (۲۰۰۳) به بررسی شرایط کشت متغیر در زراعت دیم پرداختند. نتایج نشان داد با بالابودن تغییرات بارش در منطقه می‌توان برنامه کشت پویا و متغیر را برنامه‌ریزی کرد. یونال و همکاران (۲۰۰۳) پهنه‌بندی جدیدی برای اقلیم کشور ترکیه با استفاده از روش تحلیل خوشه‌ای سلسله‌مراتبی<sup>۱</sup> انجام دادند. از داده‌های ۱۱۳ ایستگاه هواشناسی، شامل دمای حداقل، حداکثر و متوسط و کل بارش ماهانه استفاده شد. نتایج با روش وارد<sup>۲</sup> کشور ترکیه را به هفت منطقه اقلیمی تقسیم کرد. کایسلی و همکاران (۲۰۰۷) مناطق همگن بارشی را براساس ویژگی‌های آماری حداکثر بارش در جمهوری چک بررسی کردند؛ بارش‌های روزانه در طول دوره آماری ۲۰۰۰-۱۹۶۱ به‌کار گرفته شد. برای پهنه‌بندی از روش‌های تجزیه خوشه‌ای وارد و AL<sup>۳</sup> استفاده کردند. در نهایت، جمهوری چک به چهار منطقه همگن بارشی تفکیک شد. ویمالاسوریا و همکاران (۲۰۰۸) تغییرات بارش بر کشت دیم را در ویکتوریا بررسی کردند. نتایج نشان داد که میانگین بارش سالانه از سالی به سال دیگر متغیر می‌باشد. جاگاس (۲۰۰۹) برای پهنه‌بندی بارش در حوضه دریای بالتیک با استفاده از روش PCA، از آمار بارش ماهانه و فصلی در طول دوره آماری ۱۹۹۶-۱۹۰۰

استفاده کرد. نتایج نشان داد که چهار ناحیه بارشی شمالی، جنوبی، شرقی و غربی برای حوضه مذکور وجود دارد. براو کابرا و همکاران (۲۰۱۲) از داده‌های بارش ۳۴۹ ایستگاه طی دوره ۳۵ تا ۴۰ ساله در مکزیک به‌منظور خوشه‌بندی از روش تحلیل سلسله‌مراتبی K-Means استفاده کردند. نتایج نشان داد دو گروه بارشی در مکزیک وجود دارد. سانتوس و لوسی (۲۰۱۵) برای پهنه‌بندی بارش آمازون برزیل از روش تحلیل خوشه‌ای سلسله‌مراتبی در ۳۰۵ ایستگاه باران‌سنجی استفاده کردند. نتایج نشان داد که آمازون برزیل دارای شش منطقه همگن بارشی می‌باشد. پاراچو و همکاران (۲۰۱۶) برای پهنه‌بندی بارش در شبه‌جزیره ایبری، داده‌های روزانه بارش در طول دوره آماری ۱۹۵۱ تا ۲۰۰۳ را به کار بردند. از روش تحلیل غیرسلسله‌مراتبی K-Means برای خوشه‌بندی استفاده شد. نتایج نشان داد در چهار منطقه میزان بارش سالانه متمایز می‌باشد. جهانبخش و ذوالفقاری (۱۳۸۱) برای بررسی الگوهای زمانی و مکانی بارش‌های روزانه، داده‌های ۲۲ ایستگاه همدید و اقلیم‌شناسی منطقه غرب ایران را طی یک دوره آماری ۲۰ ساله (۱۹۷۱ تا ۱۹۹۰) مورد استفاده قرار دادند. ابتدا با استفاده از روش آماری تحلیل عاملی<sup>۴</sup>، ماتریس‌های بارش‌های روزانه ایستگاه‌های مذکور، خوشه‌بندی گردید. براساس نتایج تحلیل عاملی، پنج ناحیه بارش روزانه مشخص شد. رضایی و عزیز (۱۳۸۸) برای شناخت مناطق همگن بارشی غرب ایران با استفاده از ۹ متغیر وابسته به بارش، ۱۴۰ ایستگاه هواشناسی در دوره آماری ۲۰۰۰-۱۹۶۵ را مورد استفاده قرار دادند و با بهره‌گیری از روش سلسله‌مراتبی وارد ایستگاه‌ها خوشه‌بندی شدند. براساس نتایج به‌دست آمده، غرب ایران به چهار زیرمنطقه همگن بارشی تقسیم شد. تیموری و عظیم‌زاده (۱۳۹۰) با استفاده از شاخصی به‌نام شاخص مقدار بارش مؤثر به پهنه‌بندی ایران و بررسی نوسانات آن به‌صورت فصلی و سالانه پرداختند. در این مطالعه، از داده‌های بارش و درجه حرارت ماهانه ۳۹ ایستگاه هواشناسی کشور در دوره آماری ۱۳۸۴-۱۳۵۵ استفاده شد. شاخص بارش مؤثر ایستگاه‌های مختلف به‌صورت ماهانه، فصلی و سالانه محاسبه و نقشه‌ها و خطوط هم‌تراز با فواصل مناسب توسط روش خطی جامع کریجینگ تهیه شد که براساس

1- Hierarchical Cluster Analysis  
2- Ward Method  
3- Average Linkage

4- Factor Analysis

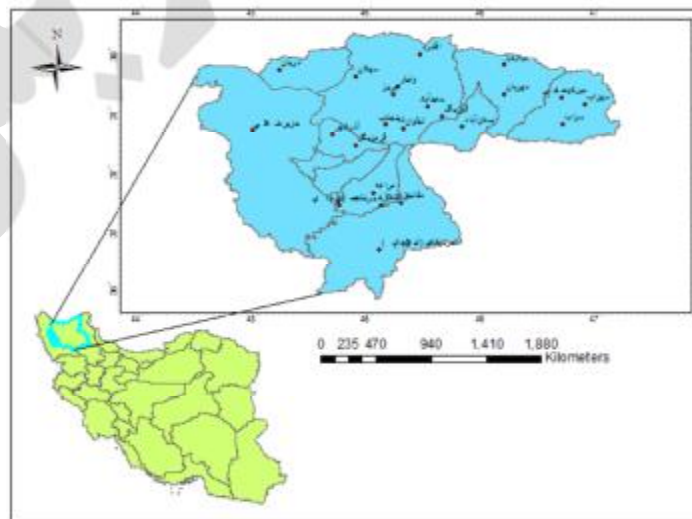
(۱۳۹۶) بررسی تأثیر تغییرات فصلی بارش بر میزان عملکرد سالانه دیم را مطالعه کردند. نتایج نشان داد اثر بارش در سه فصل بهار، پاییز و زمستان در میزان عملکرد دیم مثبت می‌باشد و بارش در این فصل‌ها بر روی افزایش عملکرد تأثیر قابل ملاحظه‌ای دارد.

با توجه به محدودیت منابع آب و اینکه بیشترین میزان مصرف آب کشور، مربوط به بخش کشاورزی می‌باشد، هدف این پژوهش تعیین شاخصی با استفاده از بارش و فواصل زمانی آن می‌باشد که با پهنه‌بندی این شاخص، میزان بارش و توزیع زمانی آن در منطقه شرق دریاچه ارومیه مشخص می‌شود که ارزیابی قابلیت کشت دیم بیشتر را در منطقه مورد مطالعه تعیین می‌کند.

#### مواد و روش‌ها

منطقه مورد مطالعه، بخش شرقی حوضه دریاچه ارومیه در شمال غربی ایران می‌باشد که در موقعیت جغرافیایی ۳۳' شمالی تا ۴۴' طول شرقی و ۴۲' ۳۵' تا ۲۵' ۳۸' عرض شمالی قرار گرفته است. آمار روزانه بارش ۲۳ ایستگاه باران‌سنجی شرق دریاچه ارومیه با توجه به دوره مشترک آماری ۱۳۹۲-۱۳۷۰ از سازمان آب منطقه‌ای استان آذربایجان شرقی به دست آمد. برای بازسازی داده‌های گمشده از برنامه‌ای در محیط فرترن استفاده شد. در شکل ۱، موقعیت جغرافیایی منطقه مورد مطالعه نشان داده شده است.

نتایج این تحقیق، ایران از نظر شاخص مذکور به پنج ناحیه اقلیمی تقسیم‌بندی شد. سهرابی ملایوسف و همکاران (۱۳۹۱) اثرها و تغییرات بارش بر میزان عملکرد سالانه گندم دیم را با استفاده از شاخص زمان- بارش بررسی کردند. نتایج نشان داد استفاده از شاخص زمان- بارش، اثر و تفاوت بین بارش‌های نوبتی را بهتر نشان می‌دهد. خورشیددوست و شیرزاد (۱۳۹۳) بارش‌های ناحیه شمال ایران را با استفاده از تحلیل خوشه‌ای و تجزیه تابع تشخیص بررسی کردند. از داده‌های بارش ماهانه دوره آماری ۱۹۹۵ تا ۲۰۰۷ سه استان گلستان، گیلان و مازندران استفاده شد. نتایج حاصل از روش وارد نشان داد بر حسب مقدار و زمان بارش، منطقه مورد مطالعه به چهار خوشه بارشی تقسیم می‌شود. عساکره و قلندری (۱۳۹۳) با استفاده از آمار بارندگی ماهانه ۷۳ ایستگاه طی دوره آماری ۲۰۰۵-۱۹۶۶ رژیم بارندگی شمال غرب کشور را تحلیل کردند. نتایج نشان داد که کمترین ضریب تغییرات و بیشترین یکنواختی بارش در شمال شرق پهنه مورد بررسی رخ می‌دهد. در نقطه مقابل آن، بخش جنوب غربی پهنه، دارای رژیم بارش کاملاً متمرکز است. جهانپخش اصل و همکاران (۱۳۹۴) برای بررسی توزیع زمانی و مکانی بارش در سطح شهرستان تبریز از داده‌های روزانه ۲۲ ایستگاه استفاده کردند. تحلیل خوشه‌ای سلسله‌مراتبی با روش وارد انجام گرفت. نتایج نشان داد که شهرستان تبریز از لحاظ بارش به سه گروه مجزای، بارش کم، متوسط و زیاد قابل تفکیک است. محمدی و همکاران



شکل ۱- موقعیت جغرافیایی منطقه مورد مطالعه

می‌یابد. این کار آن‌قدر تکرار می‌شود تا تابع خطا حداقل شود و یا اعضای خوشه‌ها تغییر نیابند.

اگر  $D$ ، مجموعه داده‌ها با  $n$  بردار باشد و  $C_1, C_2, \dots, C_K$ ، بیانگر  $K$  خوشه مختلف حاوی کلیه عناصر  $D$  باشند؛ در این صورت، تابع خطا (EF) مجموع فواصل هر بردار از مرکز خوشه خودش است که به صورت معادله (۱) تعریف می‌شود:

$$EF = \sum_{i=1}^K \sum_{X=C_i} d(X, \mu(C_i)) \quad (1)$$

که در آن  $\mu$ ، نشان‌دهنده مرکز (میانگین) خوشه؛  $d(X, \mu(C_i))$ ، فاصله هر بردار از مرکز خوشه‌ای است که خود درون آن است. فاصله هر بردار از خوشه خود بر پایه فاصله اقلیدسی محاسبه می‌شود.

برای خوشه‌بندی K-Means، گام‌هایی به‌قرار زیر طی می‌شود:

گام آغازین: تفکیک کلیه بردارها به  $K$  خوشه به صورت دلخواه؛

گام تکراری: الف) محاسبه فاصله هر بردار از مرکز خود، ب) محاسبه تابع خطا؛

گام بهبود: جابه‌جایی عضوی که بیشترین فاصله را با مرکز خوشه خودش دارد به خوشه‌ای که کمترین فاصله را با آن دارد؛ و

دستور توقف: تغییر نیافتن اعضای خوشه‌ها یا کاهش نیافتن مقدار تابع خطا (مؤمنی، ۱۳۹۳).

### روش وارد

وارد (۱۹۶۳)، روشی برای خوشه‌بندی داده‌ها ارائه کرد که این روش، یکی از روش‌های خوشه‌بندی سلسله‌مراتبی<sup>۱</sup> می‌باشد. در این روش، ابتدا هر یک از بردارها به صورت خوشه جداگانه‌ای در نظر گرفته می‌شوند. سپس، در هر گام دو بردار با هم ادغام می‌شوند. این کار آن‌قدر تکرار می‌شود تا در پایان کار، خوشه یکتایی شکل گیرد. در این روش، در هر گام، همه ترکیب‌های دوتایی ممکن از ادغام دیده می‌شوند و شاخصی به نام مجموع مربعات خطا (SSE)<sup>۲</sup> محاسبه می‌شود. هر ترکیبی که خطای کمتری داشته باشد، برگزیده می‌شود و بر پایه آن ادغام انجام می‌پذیرد. روشن است که هر بار مقدار حداقل خطا افزایش می‌یابد. این کار آن‌قدر تکرار می‌شود تا همه بردارها با هم

خوشه‌بندی، یکی از مهم‌ترین روش‌های طبقه‌بندی محسوب می‌شود. تحلیل خوشه‌ای، یکی از روش‌های پرکاربرد در بسیاری از شاخه‌های علمی است. در تحلیل خوشه‌ای تلاش می‌شود تا مشاهدات واقع در هر خوشه (گروه) بیشترین تشابه را از نظر متغیرهای مورد نظر با هم داشته باشند و مشاهدات هر خوشه از مشاهدات خوشه‌های دیگر بیشترین فاصله (عدم تشابه) را داشته باشند (مؤمنی، ۱۳۹۳). در این پژوهش، از روش K-Means و وارد برای خوشه‌بندی استفاده شد. روش وارد به دلیل ماهیت حداقل کردن خطاهای ناشی از الحاق خوشه‌ها به همدیگر، در زمره کاراترین روش‌های خوشه‌بندی در بین روش‌های سلسله‌مراتبی می‌باشد. روش K-Means، جزء روش‌های غیرسلسله‌مراتبی محسوب می‌شود. این روش، داده‌ها را به صورت انبوه در خوشه‌ها وارد می‌کند؛ ولی روش سلسله‌مراتبی وارد، خوشه‌ها را یک به یک به هم متصل می‌کند.

### روش K-Means

روش K-Means، یکی از کاربردی‌ترین روش‌های خوشه‌بندی داده‌هاست. این روش، اولین بار توسط مک کوپین (۱۹۶۷) ارائه شد. تعداد خوشه‌ها در این روش، ثابت و از پیش تعیین شده است. این روش براساس کمینه‌کردن مربعات خطا یا واریانس درون گروهی که معادل با بیشینه‌کردن واریانس بین خوشه‌هاست، بنا نهاده شده است. بنابراین، هدف کلی این الگوریتم، به دست آوردن قسمت‌هایی است که با تعداد ثابت خوشه‌ها به‌طور کلی مربع خطاها را کمینه کند. با توجه به  $K$  خوشه اولیه، از طریق تخصیص داده‌های باقی‌مانده به نزدیک‌ترین مرکز و پس از آن بارها تغییر عضویت در خوشه‌ها براساس تابع خطا، صورت می‌گیرد؛ تا زمانی که تابع خطا به‌طور معنی‌داری تغییر نکند یا اعضای خوشه‌ها تغییرات زیادی نداشته باشند. این روش برای خوشه‌بندی داده‌هایی طراحی شد که به صورت عددی (کمی) باشند و خوشه دارای مرکزی به‌نام میانگین باشد. در این روش، ابتدا بردارها به صورت تصادفی به  $K$  خوشه تقسیم می‌شوند. در گام بعد، فاصله هر یک از بردارها از مرکز خوشه خود محاسبه می‌شود. در صورتی که فاصله بردار مورد نظر از میانگین خوشه خود زیاد و به خوشه دیگری نزدیک‌تر باشد، این بردار به خوشه‌ای که نزدیک‌تر است، اختصاص

$$[F(X)]^{1-h} \quad \alpha > 0$$

و تابع چندک توزیع کاپا از معادله (۶) محاسبه می‌شود:

$$X(F) = \xi + \frac{\alpha}{k} \left[ 1 - \left( \frac{1-F}{h} \right)^k \right] \quad (۶)$$

۵- تعداد بسیار زیادی خوشه یا ناحیه ( $N_{Sm}$ ) که هر کدام دارای  $N$  ایستگاه با طول داده‌های مساوی با طول داده‌های ایستگاه‌های خوشه واقعی می‌باشد، با استفاده از توزیع چهارپارامتری کاپا ایجاد شد. می‌توان  $H_1$  را از معادله (۷) به‌دست آورد:

$$H_1 = \frac{V_1 - \mu_{V_{1s}}}{\sigma_{V_{1s}}} \quad (۷)$$

۶-  $V_{1s}$ ، برای هر ناحیه یا خوشه شبیه‌سازی شده محاسبه شد؛ در نتیجه، تعداد  $N_{Sm}$  عدد برای  $V_{1s}$  حاصل شد که میانگین و انحراف معیار آن‌ها، یعنی  $\mu_{V_{1s}}$  و  $\sigma_{V_{1s}}$  به‌دست آمد.

۷- عدد همگنی (معیار همگنی)  $H$  تحت عنوان  $H$ -Statistics از طریق معادله زیر محاسبه می‌شود که در آن  $V_1$  برای خوشه واقعی می‌باشد. می‌توان معیار همگنی  $H$  را بر مبنای گشتاورهای خطی مراتب بالاتر، یعنی  $L$ -Skew و  $L$ -Kurtosis نیز محاسبه کرد که در این صورت، به جای  $V_1$  دیگر معیارهای تغییرپذیری بین ایستگاهی گشتاورهای خطی، مثل  $L$ -Skew و  $L$ -CV و میانگین شبیه‌سازی شده آن‌ها به‌کار برده می‌شوند. به ازای  $L$ -CV و  $L$ -Skew می‌توان  $V_2$  را از معادله (۸) به‌دست آورد:

$$V_2 = \frac{\sum_{i=1}^N n_i [(t^{(i)} - \bar{t})^2 + (t_3^{(i)} - \bar{t}_3)^2]^{1/2}}{\sum_{i=1}^N n_i} \quad (۸)$$

به ازای  $L$ -Skew و  $L$ -Kurt می‌توان  $V_3$  را از معادله (۹) تخمین زد:

$$V_3 = \frac{\sum_{i=1}^N n_i [(t_3^{(i)} - \bar{t}_3)^2 + (t_4^{(i)} - \bar{t}_4)^2]^{1/2}}{\sum_{i=1}^N n_i} \quad (۹)$$

$V_2$ ، متوسط وزنی (توزین شده) فاصله گشتاورهای  $L$ -CV،  $L$ -CV،  $L$ -CV،  $L$ -CV از گشتاورهای میانگین ناحیه‌ای  $L$ -CV،  $L$ -Skew؛ و  $V_3$ ، میانگین وزنی فاصله گشتاورهای  $L$ -Skew،  $L$ -Skew،  $L$ -Skew،  $L$ -Skew از گشتاورهای میانگین ناحیه‌ای  $L$ -Skew،  $L$ -Skew،  $L$ -Skew،  $L$ -Skew می‌باشد؛ بنابراین، مقادیر را می‌توان از معادلات (۱۰) تخمین زد:

$$H_2 = \frac{V_2 - \mu_{V_{2s}}}{\sigma_{V_{2s}}}, \quad H_3 = \frac{V_3 - \mu_{V_{3s}}}{\sigma_{V_{3s}}} \quad (۱۰)$$

که در آن  $\mu_{V_{2s}}$ ، میانگین خوشه‌های شبیه‌سازی از طریق  $V_2$ ،  $\sigma_{V_{2s}}$ ، انحراف معیار خوشه‌های شبیه‌سازی از طریق

ادغام شوند و خوشه یکتایی به‌وجود آید. گاهی از این روش با نام روش «کمترین واریانس» یاد می‌شود.

### آزمون همگنی آماره H

در این تحقیق، برای آزمون همگنی خوشه‌ها از آماره  $H$  توسط هاسکینگ (۱۹۹۳)، که براساس گشتاورهای خطی مشاهداتی استاندارد شده بر مبنای گشتاورهای داده‌های مصنوعی حاصل از بهترین توزیع برازش داده شده بر داده‌های مشاهداتی ایجاد می‌شود، استفاده شده است.

آزمون  $H$ -Statistics طی مراحل زیر انجام شد:

۱- مقادیر  $L$ -CV،  $L$ -Skew،  $L$ -Kurtosis ( $t, t_3, t_4$ ) برای هر ایستگاه واقع در یک خوشه محاسبه شدند.  $L$ -CV، نظیر ضریب تغییرات معمول؛  $L$ -Skew، ضریب چولگی؛ و  $L$ -Kurtosis، ضریب کشیدگی می‌باشند.

۲- مقادیر  $L$ -CV،  $L$ -Skew،  $L$ -Kurtosis ( $\bar{t}, \bar{t}_3, \bar{t}_4$ ) ناحیه‌ای برای خوشه یا ناحیه محاسبه گردید.

۳- میانگین وزنی انحراف معیار  $L$ -CV‌های نسبی ایستگاهی به‌صورت معادله (۲) محاسبه شد:

$$V_1 = \frac{\sum_{i=1}^N n_i (t_r^{(i)} - \bar{t})^2}{\sum_{i=1}^N n_i} \quad (۲)$$

برای یک خوشه یا ناحیه حاوی  $N$  ایستگاه، گشتاورهای خطی میانگین ناحیه‌ای یا گروهی به‌صورت معادله (۳) تعریف می‌شوند:

$$\bar{t} = \frac{\sum_{i=1}^N n_i t^{(i)}}{\sum_{i=1}^N n_i}, \quad \bar{t}_r = \frac{\sum_{i=1}^N n_i t_r^{(i)}}{\sum_{i=1}^N n_i} \quad (۳)$$

که در آن  $n_i$  و  $t_r^{(i)}$  به‌ترتیب طول داده‌های ثبت شده و گشتاورهای خطی نسبی نسبت در ایستگاه نام خوشه می‌باشند و  $L$ -CV،  $t^{(i)}$ ، ایستگاه نام می‌باشد.

۴- با استفاده از گشتاورهای خطی نسبی ناحیه‌ای  $\bar{t}_2=1$ ،  $\bar{t}_3$  و  $\bar{t}_4$  پارامترهای توزیع کاپا یعنی  $\xi$ ،  $\alpha$ ،  $k$  و  $h$  محاسبه شدند. توزیع مذکور دارای تابع احتمالاتی تراکمی به‌صورت معادله (۴) است:

$$F(X) = \left\{ 1 - h \left[ \frac{1 - k(X - \xi)}{\alpha} \right]^{1/k} \right\}^{1/h} \quad (۴)$$

و تابع چگالی احتمال توزیع کاپا به شرح معادله (۵) است:

$$f(x) = \frac{1}{\alpha} \left[ \frac{1 - k(X - \xi)}{\alpha} \right]^{(1/k)-1} \quad (۵)$$

پهنه‌بندی خوشه‌های همگن به‌دست آمده با دو روش ذکرشده رسم گردید.

### نتایج و بحث

با استفاده از آمار بارش روزانه ایستگاه‌های مورد مطالعه، فاصله زمانی وقوع بارش‌ها تعیین شد. شاخصی بر مبنای این دو متغیر (بارش و فواصل زمانی آن) معرفی شد و پهنه‌بندی منطقه براساس شاخص توزیع زمانی بارش با هدف ارزیابی دیم و توزیع مکانی بارش منطقه مورد مطالعه صورت گرفت. در زیر، نتایج مربوط به دو روش خوشه‌بندی آورده شده است.

#### نتایج پهنه‌بندی روش K-Means

در این روش، کل منطقه مورد مطالعه به‌صورت دو خوشه در نظر گرفته شد و نتایج آزمون همگنی آماره H نشان داد که دو خوشه همگن می‌باشند (جدول ۱). خوشه یک دارای همگنی ملایم و خوشه دوم کاملاً همگن می‌باشد. در شکل ۲، مناطق همگن به‌دست آمده در منطقه شرق دریاچه ارومیه بر مبنای شاخص توزیع زمانی بارش با استفاده از روش K-Means نشان داده شده‌اند که ناحیه اول، قسمت‌های جنوبی، بخش‌هایی از مرکز، بخش‌هایی از شمال شرقی و ایستگاه دربان را در بر گرفته است؛ ناحیه دوم، قسمت‌هایی از غرب، مرکز، شمال و ایستگاه سهزاب را شامل شده است.

#### نتایج پهنه‌بندی روش وارد

در خوشه‌بندی شاخص با روش وارد، منطقه مورد مطالعه به دو خوشه تقسیم شد. بررسی همگنی خوشه‌ها با آزمون همگنی آماره H نشان داد دو خوشه همگن می‌باشند (جدول ۱). هر دو خوشه دارای همگنی ملایم می‌باشند. در پهنه‌بندی ناحیه یک، قسمت اعظم مساحت را در بر گرفته است. ناحیه دوم، بخش‌هایی از جنوب و ایستگاه‌های دربان، میرکوه حاجی، برازین، سعیدآباد و ليقوان را شامل می‌شود. در شکل ۳، مناطق همگن شرق دریاچه ارومیه بر مبنای شاخص توزیع زمانی بارش با روش وارد نشان داده شده است. نتایج مربوط به مقدار آماره آزمون همگنی شاخص توزیع زمانی بارش در جدول ۱ ارائه شده است.

$V_2$ ؛  $\mu_{V_{3S}}$  میانگین خوشه‌های شبیه‌سازی از طریق  $V_3$ ؛ و  $\sigma_{V_{3S}}$  انحراف معیار خوشه‌های شبیه‌سازی از طریق  $V_3$  می‌باشند. آزمون H-Statistics به‌صورت زیر ارزیابی همگنی را انجام می‌دهد:

اگر  $H \leq 1$  باشد، خوشه یا ناحیه کاملاً همگن خواهد بود؛  
اگر  $1 < H \leq 2$  باشد، خوشه یا ناحیه دارای همگنی ملایم خواهد بود؛ و  
اگر  $H > 2$  باشد، خوشه یا ناحیه کاملاً ناهمگن خواهد بود (هاسکینگ، ۱۹۹۳).

برای تعیین شاخص توزیع زمانی بارش، ابتدا برنامه‌ای در محیط فرترن نوشته شد که با استفاده از داده‌های بارش، ابتدا نسبت بارش به فواصل زمانی آن محاسبه شود. سپس، با استفاده از این دو متغیر، برای بررسی مناطق مستعدتر دیم و توزیع زمانی بارش منطقه مورد مطالعه شاخصی معرفی گردید. ایده معرفی شاخص بر مبنای نسبت بارش به فواصل زمانی آن می‌باشد که اگر مقدار این نسبت بزرگ حاصل شود، بیانگر تأمین آب (منطقه پرباران بوده) و یا به‌عبارت دیگر، بیانگر قابلیت دیم بالا خواهد بود که می‌تواند به‌عنوان شاخص ارزیابی عمل نماید. در این شاخص، عمق بارندگی نقش مثبت و فواصل بارندگی نقش منفی را در ارزیابی دارند. هرچه مقدار شاخص بزرگ‌تر شود، قابلیت دیم افزایش می‌یابد؛ زیرا مقدار بارش بایستی بزرگ‌تر شود که حاکی از مقدار زیاد بارش می‌باشد یا فواصل زمانی آن کوچک‌تر شود که افزایش تعدد بارش را به دنبال خواهد داشت و یا اگر به‌صورت ترکیبی موجب افزایش  $D_i$ ، فاصله زمانی بارش در زمان  $t_i$ ؛  $P_i$  مقدار بارش در زمان  $t_i$  و  $n_p$  تعداد بارش در طول سال می‌باشند. بعد از محاسبه شاخص، خوشه‌بندی داده‌های سالانه شاخص با دو روش K-Means و وارد انجام گردید. در خوشه‌بندی، ایستگاه‌های باران‌سنجی به‌عنوان متغیر در نظر گرفته شد؛ یعنی هر ستون از داده‌های شاخص که به‌صورت ماتریس به‌دست آمده بود، به کل داده‌های یک ایستگاه اختصاص یافت؛ یعنی در هر ستون، عددهای شاخص از سال ۱۳۷۰ تا ۱۳۹۲ به‌طور متوالی درج شدند. ماتریس همبستگی  $23 \times 23$  (۲۳ مربوط به تعداد سال‌ها و تعداد ایستگاه‌ها) به‌عنوان ورودی روش K-Means و وارد در نرم‌افزار SPSS برای خوشه‌بندی به‌کار گرفته شد؛ سپس، بررسی همگنی خوشه‌های حاصل‌شده با استفاده از آزمون آماره H صورت گرفت. در نهایت، با استفاده از نرم‌افزار ArcGIS نقشه‌های

جدول ۱- مقدار آماره آزمون همگنی H شاخص توزیع زمانی

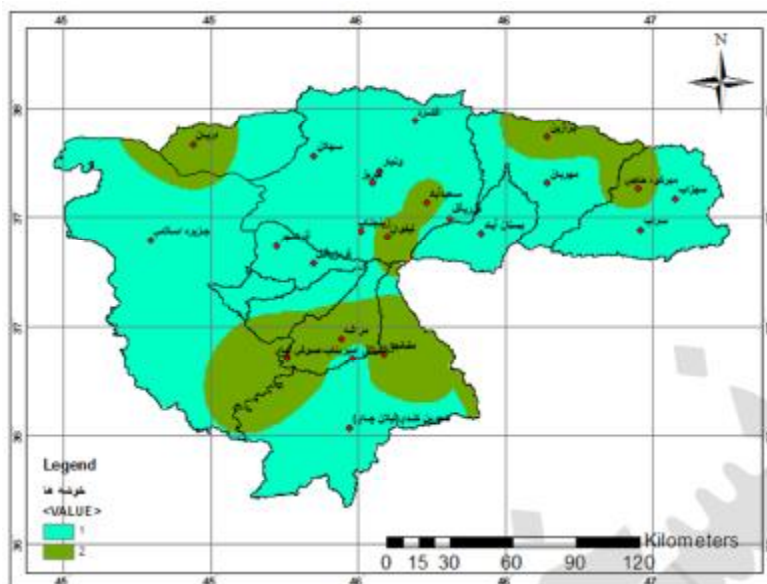
بارش منطقه مورد مطالعه		
شماره خوشه	K-Means	وارد
۱	-۱/۵۹	۱/۴۵
۲	۰/۷۲	-۱/۱۰

ایستگاه‌های واقع در این ناحیه نسبت به ایستگاه‌های ناحیه دوم می‌باشد. ایستگاه‌های موجود در ناحیه یک از نظر مقدار شاخص توزیع زمانی بارش از مقادیر بالایی برخوردار هستند که با توجه به تعریف شاخص توزیع زمانی بارش، بالا بودن میزان شاخص، نشانگر مناطق پرباران و قابلیت کشت دیم بیشتری در منطقه است. میزان خطوط هم‌شاخص رسم‌شده در این ناحیه نیز قابل توجه بوده که این ناحیه، مناطق مستعدتر را برای کشت دیم شامل می‌شود. میزان خطوط هم‌شاخص توزیع زمانی بارش با تعداد ایستگاه‌های بسیار کمتری در دو ناحیه در روش وارد، از افزایشی همسو برخوردارند. همچنین، ناحیه دوم مناطق همگن به‌دست آمده با روش وارد در ایستگاه‌های محدودی با خطوط هم‌شاخص همخوانی دارند. با توجه به توانایی خوشه‌بندی روش K-Means، این روش به‌عنوان روش برتر در خوشه‌بندی شاخص توزیع زمانی بارش منطقه شرق دریاچه ارومیه عمل کرد. از نظر تغییرات مکانی با استفاده از شاخصی که با توزیع زمانی محاسبه شد، مناطق جنوبی، بخش‌هایی از مرکز، شمال شرقی و شمال غرب، مناطق پرباران منطقه مورد مطالعه می‌باشند و از نظر توان دیم‌کاری مناطق مستعدتر دیم را دربر می‌گیرند. این مناطق نیز از نظر پتانسیل کشاورزی و توزیع زمانی بارش مناسب می‌باشند.

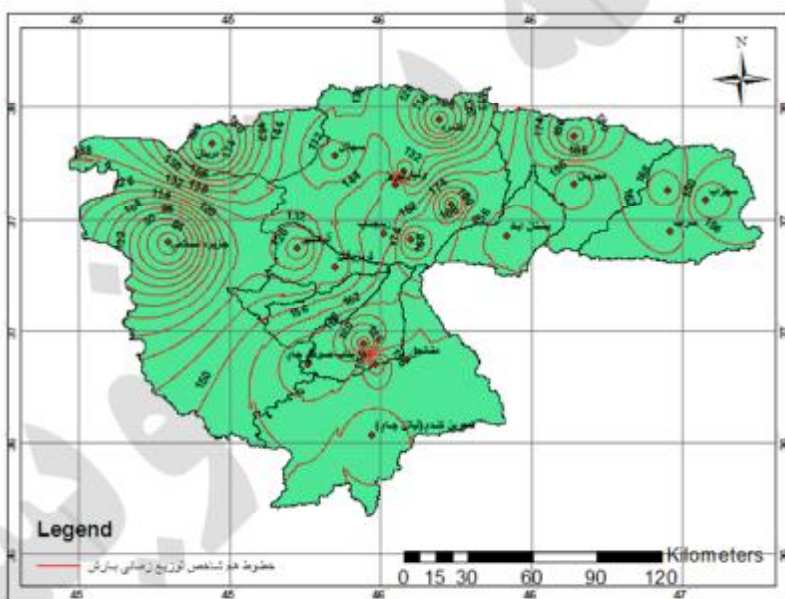
با توجه به نتایج خوشه‌های همگن به‌دست آمده با دو روش K-Means و وارد، منطقه مورد مطالعه به دو ناحیه تفکیک شد. ناحیه یک روش وارد قسمت اعظم منطقه را در بر گرفت و این روش، ایستگاه‌های مشترک محدودی را با روش K-Means شامل شدند. برای بررسی توانایی روش‌های ذکرشده در خوشه‌بندی و شناسایی مناطق مستعدتر دیم، خطوط هم‌شاخص توزیع زمانی بارش منطقه مورد مطالعه رسم شد. شکل ۴، خطوط هم‌شاخص توزیع زمانی بارش منطقه شرق دریاچه ارومیه طی دوره ۱۳۷۰-۱۳۹۲ را نشان می‌دهد. شناسایی مناطق پرباران منطقه با بررسی این خطوط (بیشترین نسبت موجود به دست آمده) انجام گرفت. مناطق پرباران برای کشاورزی دیم مناسب‌اند و از نظر اقلیمی، توزیع زمانی بارش در این مناطق بهتر می‌باشد. مقایسه دو نقشه حاصل از پهنه‌بندی با روش K-Means و وارد و خطوط هم‌شاخص توزیع زمانی بارش نشان داد که در پهنه‌بندی با روش K-Means، ناحیه یک این پهنه‌بندی با خطوط هم‌شاخص افزایشی هم‌سو دارند که حاکی از پرباران‌تر بودن



شکل ۲- مناطق همگن شرق دریاچه ارومیه بر مبنای شاخص توزیع زمانی بارش با روش K-Means



شکل ۳- مناطق همگن شرق دریاچه ارومیه بر مبنای شاخص توزیع زمانی بارش با روش وارد



شکل ۴- خطوط هم شاخص شرق دریاچه ارومیه طی دوره ۱۳۷۰-۱۳۹۲

### نتیجه‌گیری

با توجه به مسئله بحران آب در کشور، پایین‌رفتن سطح آب زیرزمینی و اهمیت بخش کشاورزی در تأمین مواد غذایی و ایجاد اشتغال، مدیریت منابع آب امری ضروری می‌باشد. بنابراین، تعیین اراضی مستعد دیم در شرایط فعلی با وجود کاهش بارندگی، ارزش مدیریتی بالایی خواهد داشت. در این تحقیق، برای شناسایی مناطق مستعدتر دیم، شاخص ارزیابی مناسبی بر مبنای بارش و

فواصل زمانی آن تعیین شد. با استفاده از این شاخص مناطقی را که برای کشاورزی دیم از نظر تأمین آب مناسب می‌باشند، شناسایی شد که این شاخص می‌تواند در مدیریت منابع آب عمل کند. با شناسایی مناطق مستعدتر دیم در مناطق بحرانی می‌توان از آثار مخرب در سطح منطقه ممانعت کرد. همچنین، در روش K-Means با توجه به افزایشی هم‌سو در ایستگاه‌های واقع در هر خوشه و خطوط هم‌شاخص توزیع زمانی بارش، این روش



تبخیر- تعرق با روش‌های وارد و K-Means. مجله اکوهیدرولوژی. ۴(۲): ۴۸۹-۴۹۸.

۹. مؤمنی م. ۱۳۹۳. خوشه‌بندی داده‌ها (تحلیل خوشه‌ای). تهران. چاپ دوم. ۲۹۶ص.

10. Bravo Cabrera J. L. Azpra Romero E. Zarraluqui Such V. Gay Garcia C. and Estrada Porum F. 2012. Cluster analysis for validated climatology stations using precipitation in Mexico. *Atmosfera*. 25(4): 339-354.
11. Dinar A. Mendelsohn R. Evenson R. E. Parikh J. Sanghi A. Kumar K. Mc Kinsey J. and Lonergan S. 1998. Measuring the Impact of Climate Change on Indian Agriculture. World Bank Technical, Washington. 402 p.
12. Hosking J. R. M. and Wallis J. R. 1993. Some statistics useful in regional frequency analysis. *Water Resource Research*. 29(2): 271-281.
13. Jaagus J. 2009. Regionalization of precipitation pattern in the Baltic Sea drainage basin and its dependence on large-scale atmospheric circulation. *Boreal Environment Research Journal*. 14: 31-44.
14. Kysely J. Piecek J. and Huth R. 2007. Formation of homogeneous regions for regional frequency analysis of extreme precipitation events in the Czech Republic. *Studio Geophysical et Geodetic*. 51: 327-344.
15. Macqueen J. 1967. Some methods for classification and analysis of multivariate observation. In *Proceeding of the 5th Berkeley symposium on mathematical statistics and probability*. Berkeley, CA: University of California. 1: 281:297.
16. Parracho A. C. Melo-Goncalves P. and Rocha A. 2016. Regionalisation of precipitation for the Iberian Peninsula and climate change. *Physics and Chemistry of the Earth, Parts A/B/C*.
17. Romero R. Ramis C. Guijarr J. A and Summer G. 1999. Daily rainfall affinity areas in Mediterranean Spain. *International Climatology Journal*. 19: 557-578.
18. Sadras V. Rogeta D. and Krause K. 2003. Dynamic cropping strategies for risk management in dry-land farming systems. *Agricultural System Journal*. 76(3): 929-948.
19. Santos E. B. and Lucio P. S. 2015. Precipitation regionalization of the Brazilian Amazon. *Atmospheric Science Letters*. 16(3): 185-192.
20. Unal Y. Kindap. T. and Karacab M. 2003. Redefining the climate zones of Turkey

عملکرد بهتری را برای خوشه‌بندی شاخص، نسبت به روش وارد ارائه کرد. روش K-Means، به‌عنوان روش برتر برای خوشه‌بندی در این مطالعه انتخاب شد. در نهایت، مناطق جنوبی، بخش‌هایی از مرکز، شمال شرقی و شمال غرب منطقه مورد مطالعه، به‌عنوان مناطق مستعدتر دیم شناسایی شدند.

## منابع

۱. تیموری م. و عظیم‌زاده س. ج. ۱۳۹۰. ناحیه‌بندی اقلیمی ایران با استفاده از شاخص بارش مؤثر. اولین کنفرانس ملی خشکسالی و تغییر اقلیم، کرج، مرکز تحقیقات کم‌آبی و خشکسالی در کشاورزی و منابع طبیعی. ۲۸ اردیبهشت ۱۳۹۰، ۸ صفحه.
۲. جهانبخش اصل س. ابطحی و. قربانی م. ع. تدینی م. و والایی ا. ۱۳۹۴. بررسی توزیع زمانی و مکانی بارش شهرستان تبریز با روش تحلیل خوشه‌ای. فصلنامه علمی- پژوهشی فضای جغرافیایی. ۱۵(۵۰): ۵۹-۸۱.
۳. جهانبخش س. و ذوالفقاری ح. ۱۳۸۱. بررسی الگوهای بارش‌های روزانه در غرب ایران. فصلنامه تحقیقات جغرافیایی. ۶۴: ۲۳۴-۲۵۸.
۴. خورشیددوست ع. م. و شیرزاد ع. ا. ۱۳۹۳. بررسی و تحلیل بارش‌های ناحیه شمال ایران با استفاده از تحلیل خوشه‌ای و تجزیه تابع تشخیص. نشریه جغرافیا و برنامه‌ریزی. ۱۸(۴۹): ۱۰۱-۱۱۸.
۵. رضایی ط. و عزیزی ف. ۱۳۸۸. شناخت مناطق همگن بارشی در غرب ایران. مجله جغرافیا و برنامه‌ریزی محیطی. ۳۴: ۶۵-۸۶.
۶. سهرابی ملا یوسف س. فاحری‌فرد ا. و بزرگ حداد ا. ۱۳۹۱. بررسی نوبتی بارش‌های پاییز و زمستان بر عملکرد سالانه دیم با استفاده از شاخص زمان- بارش (RTI). نشریه آب و خاک (علوم و صنایع غذایی). ۲۶(۱): ۷۵-۸۴.
۷. عساکره ح. و رزمی قلندری ر. ۱۳۹۳. توزیع زمانی و رژیم بارش در شمال غرب ایران. فصلنامه تحقیقات جغرافیایی. ۲۹(۱۱۲): ۱۴۵-۱۶۰.
۸. محمدی پ. فاحری‌فرد ا. دین‌پژوه ی. و اسدی س. ۱۳۹۶. پهنه‌بندی دریاچه ارومیه بر مبنای بارش بر

- using cluster analysis. International Climatology Journal. 23: 1045-1055.
21. Ward Jr. 1963. Hierarchical grouping to optimize an objective function. American Statistical Association Journal. 58(301): 236-244.
  22. Wimalasuriya R. Ha A. Tsafack E. and Larson K. 2008. Rainfall Variability and its Impact on Dryland Cropping in Victoria. In Proceeding of the 52 Annual Conference of the Australian Agricultural and Resource Economics Society, 5-8 Feb. Canberra. Australia.

مجله علمی پژوهشی  
پژوهش‌های جغرافیایی

## Regionalization of the East Part of Urmia Lake Basin based on temporal distribution of Precipitation Index using the K- Means and Ward methods

P. Mohammadi<sup>1\*</sup>, A.Fakheri Fard<sup>2</sup>

### Abstract

Considering the importance of precipitation and scarcity of water resources, spatial analysis of daily rainfall together with corresponding time intervals is one of the requirements. Among atmospheric factors affecting the rainfed farming, precipitation is the most important factor in numerous studies to rainfed cultivation. For reducing the destruction threat of water resources and resolve the future food needs of the people, the importance of rainfed agriculture would be inevitable. In this regard, one of the effective measures that can be done to find rainy areas with the possibility of rainfed cultivation. The highest water consumption is corresponding to the agricultural sector, according to the water crisis in the country, decline in the groundwater levels and the importance of agriculture in job creation, the water resource management is essential. Therefore it will have a value of upper management the determination of fertile lands in the current situation despite the reduction in rainfall, also it is necessary to analytical study of rainfed and to determine the fertile lands in the country especially in Urmia lake basin. This study aimed to introduce an index for identification of the eligible areas from the viewpoint of rainfed ability, by using daily rainfall in corresponding with the time intervals for each year in the east part of the Urmia lake basin.

Firs a program is written in Fortran, by using rainfall data, the proportion of rainfall intervals is determined, then Fortran program is developed for this propose, based on the concept of the temporal intensity precipitation and their related mechanistic, an index is defined which is the main outcome of Fortran programming. The index was a criterion for analysis and disintegration of the quantity of rainfed. The idea of introduce an index is based on the ratio of rainfall to the interval corresponding. If the value of this ratio is high, it represents the water supply, in other words that rainfed capability will be high, and it can play the role of an index rating. In this index the depth of rainfall have positive role and intervals have negative role in the assessment. When the index amount is high, rainfed capability will increase. Because the amount of precipitation must be bigger which is the sign of the high amount of rainfall or the smaller time interval consequently the increase of the number of rainfall will be followed, or in compound cause the increase of the proportion which shows the increase of rainfed ability. Statistical period was 1992-2014 for 23 rainfall stations. Index values were calculated for each year of stations. Clustering is one of the most numerous classification methods. In cluster analysis one attempts to actual observations of each cluster which have the most similarity in terms of variables together. In this study the clustering methods of K-

1- Ph.D. student of Water Resources Engineering, Department of Irrigation and Reclamation Engineering, Faculty of Agricultural Engineering and Technology, University of Tehran

2- Professor, professor of Department Water Engineering, University of Tabriz, Tabriz, Iran.

\*- Corresponding author: [mohammadi.parva@yahoo.com](mailto:mohammadi.parva@yahoo.com)

Means and Ward were carried out for delineation of homogeneous regions based on developed index values. Clustering in rainfall stations is intended as a variable, the correlation matrix  $23 \times 23$ , 23 related to the number of the years and the number of stations is used. The homogeneity of cluster were checked through H-Statistics method and homogeneous clusters were shown in GIS environment.

For regionalization with K- Means method in the whole area of the study was considered as two clusters and the results of homogeneity test H-Statistics showed that the two cluster are homogeneous. The first cluster have a mild homogeneous and the second one is completely homogeneous. In the clustering index with Ward method in the study area was divided in two clusters. The observation of the of homogeneous clusters with homogeneous test of H-Statistics showed both clusters are mild homogeneous. On the other hand, from the view point of spatial variation of Iso-Index lines were drawn over the study area. Identify rainy areas was conducted by examining the lines of these areas that are suitable for rainfed agriculture and also the climate temporal distribution of rainfall in these areas are better. The comparison of two maps regionalization with two methods and the Iso-Index lines show that the regionalization by K- means method, the first part of this regionalization with Iso-Index lines have the same incremental direction, that suggests more rain stations in this area in comparison with the second area. The second cluster of homogeneous region obtained by Ward method having only in the limited stations have similarity with Iso-Index lines. The comparison regions resulted from clustering methods with the spatial distribution of the Iso-Index lines over the study area implied that K-Means method isolated the regions better than ward method. The results show that South, the northeast, the northwest and the center of the study area were more eligible for rainfed agriculture than the other parts, this areas regarding to agriculture, temporal distribution of precipitation and richness of groundwater are better.

**Keywords:** Clustering, Homogeneity Test, Rainfed, Spatial Distribution,

**Citation:**