



مدیریت آب و آبیاری

دوره ۱۲ ■ شماره ۲ ■ تابستان ۱۴۰۱

صفحه‌های ۳۷۳-۳۵۹

DOI: 10.22059/jwim.2022.339537.972

مقاله پژوهشی:

بررسی عملکرد شبکه‌های عصبی گازی در خوشه‌بندی هیدرولوژیک

محمدرضا محمودی^۱، سید سعید اسلامیان^{۲*}، سید علیرضا گوهری^۳، معین طحانیان^۴

۱. کارشناسی ارشد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران.

۲. استاد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران.

۳. استادیار، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران.

۴. کارشناسی ارشد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران.

تاریخ پذیرش مقاله: ۱۴۰۱/۰۴/۱۰

تاریخ دریافت مقاله: ۱۴۰۰/۱۲/۳

چکیده

طراحی بسیاری از زیرساخت‌ها و پروژه‌های عمرانی نیازمند مطالعات گسترده‌ای در زمینه شرایط جغرافیایی منطقه و ویژگی‌های اقلیمی آن ناحیه می‌باشد. کارایی این پژوهش‌ها خود وابسته به اطلاعات و داده‌های موردنیاز است. در بسیاری از مواقع منطقه طرح در موقعیتی قرار دارد که هیچ‌گونه اطلاعات اقلیمی مانند بارش موجود نیست. از این‌رو، تحلیل فراوانی منطقه‌ای بسیار موردتوجه قرار گرفته است. در این شیوه با شرایط و ابزار خاصی اطلاعات موجود در نواحی دیگر قابل بسط و انتقال به سایر نواحی می‌شود. در این مسیر خوشه‌بندی یکی از تأثیرگذارترین مراحل می‌باشد که منطقه و ایستگاه‌های موجود را به مناطق همگن هیدرولوژیک تقسیم می‌نماید. از این‌رو، در این پژوهش علاوه بر روش‌های رایج در خوشه‌بندی از دو مدل جدید شبکه عصبی گازی و شبکه عصبی گازی رشدیابنده به منظور تعیین مناطق همگن در سطح استان خوزستان استفاده شد. یکی از ویژگی‌های منحصربه‌فرد این الگوریتم‌ها یادگیری توپولوژی یا شکل توزیع حاکم بر فضای داده‌ها می‌باشد. با استفاده از متغیرهای طول جغرافیایی، عرض جغرافیایی، ارتفاع، متوسط بارش سالانه و حداکثر بارش ۲۴ ساعته سالانه ایستگاه، منطقه طرح به دو ناحیه همگن تقسیم و فرایند خوشه‌بندی انجام پذیرفت. نتایج نشان‌دهنده کارایی و دقت بالای شبکه‌های عصبی گازی در مبحث خوشه‌بندی می‌باشد. متوسط میزان خطا و هم‌چنین ضریب تغییرات خطا در این مدل به ترتیب ۱۵/۵۶ و ۲۴/۳۹ درصد برآورد شد که نسبت به روش‌های معمول برتری قابل توجهی از خود نشان داد.

کلیدواژه‌ها: تحلیل فراوانی منطقه‌ای، خوزستان، همگنی هیدرولوژیک، یادگیری توپولوژی.

Investigation of the performance of neural gas networks in hydrological clustering

Mohammadreza Mahmoudi¹, Saeid Eslamian^{2*}, Alireza Gohari³, moein tahananian

1. M.Sc., Department of Water Engineering, College of Agriculture, Isfahan University of Technology, Isfahan, Iran.

2. Professor, Department of Water Engineering, College of Agriculture, Isfahan University of Technology, Isfahan, Iran.

3. Assistant Professor, Department of Water Engineering, College of Agriculture, Isfahan University of Technology, Isfahan, Iran.

4. M.Sc., Department of Water Engineering, College of Agriculture, Isfahan University of Technology, Isfahan, Iran.

Received: February 22, 2022

Accepted: July 01, 2022

Abstract

The design of many infrastructures and construction projects requires the extensive studies in the area's geographical conditions and climatic characteristics. The effectiveness of this research itself depends on the information and data required. In many cases, the project area is in a situation where no climatic information such as rainfall is available. Hence, regional frequency analysis has been received much attention. In this way, having specific conditions and mechanisms, the available information in other sites can be expanded and transferred to the other areas. In this research, clustering is one of the most effective steps that divide the existing stations into the hydrologically homogeneous areas. Therefore, in this study, in addition to the common methods in clustering, two new models, neural gas network and growing neural gas network, were used to determine the homogeneous regions in Khuzestan province, Iran. One of the unique features of these algorithms is learning the topology or shape of the distributions that governs the data space. Using the variables of longitude, latitude, altitude, mean annual rainfall, and maximum 24-hour rainfall of the station, the design area was divided into two hydrologically homogeneous areas, and the clustering process was performed. The results show that the neural gas networks has a high efficiency and accuracy in view of clustering. The Mean Percentage Difference and Coefficient of Variation of Root Mean Square Error in neural gas were estimated to be 15.56 and 24.39 percent, respectively, which showed a considerable advantages over the conventional methods.

Keywords: Clustering, Hydrological homogeneity, Khuzestan, Regional Frequency Analysis, Topology learning.

مقدمه

و همچنین عدم سهولت دسترسی ممکن است تراکم ایستگاه‌های باران‌سنجی کم‌تر از میزان استاندارد باشد. در این وضعیت شبیه‌سازی بارش در نقاط فاقد ایستگاه از اهمیت ویژه‌ای در مطالعات مربوط به منابع آب برخوردار است. به‌منظور برطرف‌نمودن این مشکل تحلیل منطقه‌ای پارامترهای هیدرولوژیک پیشنهاد شده است.

تحلیل منطقه‌ای به‌جای استفاده از داده‌های یک ایستگاه از آمار و ویژگی‌های گروهی از ایستگاه‌های هم‌رفتار استفاده می‌کند. پژوهش‌های مختلف با استفاده از روش‌های منطقه‌ای بر روی بارش‌های حدی حاکی از آن است که این تکنیک‌ها تردیدها را در زمینه برآوردهای ناشی از دیدگاه ایستگاهی به‌طور فزاینده‌ای کاهش می‌دهند (Lee & Maeng, 2003).

در پژوهشی Durrans & Kirby (2004) تابع توزیع GEV را به‌منظور الگوبندی بارش‌های حدی روزانه و ساعتی انتخاب کردند. بدین منظور با تحلیل خوشه‌ای و تخمین‌های بارش حدی اقدام به منطقه‌ای‌کردن این بارش‌ها برای اطلس باران آلاباما کردند و روش منطقه‌ای‌کردن را برای گسترش نتایج در مناطق فاقد ایستگاه مناسب ارزیابی کردند.

در پژوهشی دیگر Soltani *et al.* (2017) با استفاده از خصوصیات مقیاس زمانی بارش و با استفاده از سه متغیر میانگین شدت بارش روزانه، انحراف معیار شدت بارش روزانه و نمایه مقیاس، منحنی‌های IDF منطقه‌ای را برای استان خوزستان ترسیم و میزان خطای مطلق این روش را به‌طور عمده زیر ۲۵ درصد و نتایج را قابل قبول اعلام کردند.

هم‌چنین Ariff *et al.* (2016) با استفاده از گشتاورهای خطی و منطقه‌ای‌نمودن توزیع مقدار حدی، IDFهای منطقه‌ای را برای ناحیه پنین‌سولار مالزی استخراج و کاربرد این روش را با توجه به‌سادگی و کارایی آن در مناطق فاقد ایستگاه مناسب ارزیابی نمودند.

مطالعات هیدرولوژی دارای اهمیت فراوانی در بسیاری از طرح‌ها هم‌چون سدسازی و راه‌سازی می‌باشد. در اکثر موارد این مطالعات به‌وسیله مدل‌های هیدرولوژیک انجام می‌شود که به داده‌های اقلیمی با دقت مکانی و زمانی مناسب نیاز دارند. از مهم‌ترین این داده‌ها می‌توان به بارندگی اشاره کرد که در صورت وجود محدودیت‌هایی مثل پراکندگی نامناسب ایستگاه‌ها و طول دوره آماری کوتاه و هم‌چنین نداشتن دقت مکانی لازم، موجب ایجاد خطای تخمین در محاسبات و عملکرد مدل‌ها می‌شود (Chaubey *et al.*, 1999; Goovaerts, 1999). لذا لازم است که داده‌های بارش با دقت مکانی و زمانی مناسب جهت استفاده در این مدل‌ها تهیه شود.

بارندگی یکی از مهم‌ترین پارامترهای چرخه هیدرولوژی است که در اثر آن رخدادهایی هم‌چون سیلاب، آلودگی آب‌های سطحی و زیرزمینی، فرسایش و رسوب‌گذاری به وقوع می‌پیوندد. بارش از بیش‌ترین تغییرات زمانی و مکانی برخوردار بوده و کمی‌کردن آن در بُعد مکان و زمان موردتوجه پژوهش‌گران مختلف بوده است. در بسیاری از موارد داده‌ها و هم‌چنین پراکندگی مکانی ایستگاه‌ها ناکافی بوده و برای تحلیل‌های مختلف نامناسب است. به‌علاوه، تخمین درست توزیع منطقه‌ای بارش به شبکه بسیار متراکمی از تجهیزات نیازمند است که هزینه بسیار زیادی در پی خواهد داشت. در مواردی که ایستگاه موردنظر خارج از منطقه مطالعه طرح قرار دارد مشکلات زیادی برای تخمین بارندگی و در نتیجه دبی ایجاد می‌کند (Lee & Maeng, 2003).

در سطح جهانی برای اندازه‌گیری داده‌های بارندگی از ایستگاه‌های هواشناسی به‌طور عمده شامل ایستگاه‌های سینوپتیک، کلیماتولوژی و باران‌سنجی استفاده می‌شود. در برخی کشورها به دلایل مختلفی از جمله کمبود منابع مالی

تشخیص مناطق همگن معمولاً در بین تمامی مراحل تحلیل فراوانی منطقه‌ای مهم‌ترین و سخت‌ترین گام و پیش‌نیاز فرضیه تحلیل فراوانی می‌باشد. در این مرحله ایستگاه‌هایی که به‌جز عامل مقیاس در نوع تابع توزیع مشترک باشند گروه‌بندی می‌شوند.

در این پژوهش ما روشی مبتنی بر شبکه‌های عصبی گازی^۲ و گازی رشدیابنده^۳ را به‌منظور خوشه‌بندی داده‌های هیدرولوژیک و تعیین مناطق همگن ارائه می‌دهیم. شبکه عصبی گازی یکی از انواع شبکه‌های عصبی رقابتی است و از شیوه آموزش غیرنظارت‌شده استفاده می‌کند. این شبکه نخستین مرتبه توسط Martinetz & Schulten (1991) معرفی شد. یکی از ویژگی‌های منحصر به فرد این الگوریتم یادگیری توپولوژی یا شکل توزیع حاکم بر فضای داده‌ها می‌باشد. یکی از مشکلات این الگوریتم این است که در ابتدا با تعداد عناصر زیادی شروع به کار می‌کند که باعث می‌شود الگوریتم در ابتدا کند عمل کند. این نقص چهار سال بعد با ارائه شبکه عصبی گازی رشدیابنده توسط Fritzsche (1995) برطرف شد. در شبکه عصبی گازی رشدیابنده تعداد نوروها در طی فرایند یادگیری بدون در نظر گرفتن دانش قبلی و ساختار حاکم بر ورودی‌ها افزایش پیدا می‌کند (Fink & Weidmann, 2015).

در مطالعه‌ای Abdi et al. (2017) توانایی شبکه عصبی گازی رشدیابنده در منطقه‌ای نمودن شاخص خشک‌سالی برای ۴۰ ایستگاه سینوپتیک در ایران را بررسی و نتایج حاصل از ارزیابی ناهمگونی براساس گشتاورهای خطی نشان‌دهنده موفقیت این الگوریتم نسبت به روش‌های دیگر در تعیین زیر مناطق همگن بود. استفاده از شبکه‌های عصبی گازی در خوشه‌بندی در رشته‌های دیگر علمی مانند ریاضیات (Carlevarino et al., 2000; Ferrer, 2014)، پزشکی (Angelopoulou et al., 2009) و (Cselényi, 2005; Oliveira Martins, 2015) و

نتایج پژوهش‌های Alemaw & Chaoka (2016) که با استفاده از ویژگی‌های توپوگرافیک و بارندگی، بوتسوانا را با مدل خوشه‌بندی K-Means به سه منطقه همگن تقسیم و توابع توزیع احتمال گاما و لوگ نرمال را به‌منظور تخمین عمق بارندگی در دوره‌های بازگشت تا ۱۰۰ سال برای مناطق ذکر شده استفاده نمودند، مناسب ارزیابی شد.

پژوهش‌گران از مدل‌های مختلف خوشه‌بندی استفاده می‌کنند و برای هر یک از این مدل‌ها، الگوریتم‌های مختلفی را می‌توان ارائه داد. مفهوم یک خوشه همان‌طور که توسط الگوریتم‌های مختلف یافت می‌شود به‌طور خاصی در ویژگی‌ها تفاوت دارند. هیچ توافقی میان پژوهش‌گران در مورد برتر بودن یک روش خاص وجود ندارد.

خوشه‌بندی را می‌توان براساس سختی تمایز به دو دسته طبقه‌بندی کرد. اول خوشه‌بندی سخت که هر شیء یا متعلق به یک خوشه می‌باشد یا خیر. دوم خوشه‌بندی نرم (خوشه فازی) که هر شیء یا داده به میزان خاصی متعلق به یک خوشه باشد.

با توجه به شکل‌های متنوع مناطق و اثرات ویژگی‌های مربوط به حوزه‌های مختلف که در منطقه‌بندی هیدرولوژیک اجتناب‌ناپذیر است، انتخاب بهترین روش بسیار مهم است (Alemaw & Chaoka, 2016).

یکی از مشکلات عمده بسط نتایج تحلیل فراوانی از یک یا چند ایستگاه به یک منطقه، عدم همگنی منطقه از نظر هیدرولوژیک می‌باشد. با وجود مناسب بودن تحلیل‌های خوشه‌ای برای گروه‌بندی ویژگی‌های هیدرولوژیک، همگن بودن مناطق کاملاً حاصل نمی‌شود. به این ترتیب پیشنهاد می‌شود نتایج تحلیل خوشه‌ای با روش‌های دیگر مورد بررسی و آزمون قرار گیرد (Rousseuw, 1987).

در این پژوهش با استفاده از دو مدل جدید شبکه عصبی گازی و گازی رشدیابنده، استان خوزستان از نظر تعداد مناطق همگن بررسی و سپس نتایج حاصل با استفاده از آزمون‌های مبتنی بر گشتاورهای خطی با روش‌های مرسوم خوشه‌بندی از جمله مدل‌های وارد^۵، کامیانگین^۶، شبکع عصبی نقشه خودسازمانده^۶ و مدل فازی سی- میانگین^۷ مقایسه شدند. سپس به منظور بررسی بهتر کارایی این دو مدل در تحلیل فراوانی منطقه‌ای، منحنی‌های شدت- مدت- فراوانی منطقه‌ای بارش در این نواحی استخراج شد.

مواد و روش‌ها

استان خوزستان با مساحتی حدود ۴ درصد از مساحت کل کشور، وسیع‌ترین استان در نیمه غربی کشور است. این استان بین ۴۷ درجه و ۴۱ دقیقه تا ۵۰ درجه و ۳۹ دقیقه طول شرقی و ۲۹ درجه و ۵۸ دقیقه تا ۳۳ درجه و ۴ دقیقه عرض شمالی از خط استوا قرار دارد. خوزستان با حجم ورودی بیش از ۳۳ میلیارد مترمکعب آب در سال و دارا بودن ۳ میلیارد مترمکعب آب ذخیره‌شده در سفره‌های زیرزمینی، از موقعیت ممتازی در کشور برخوردار است. این استان بیش از ۳۰ درصد از آب‌های سطحی کشور را در خود جای داده است. میانگین بارندگی استان طی دوره آماری بلندمدت ۲۸۴/۳ میلی‌متر است. بیش‌ترین میزان بارش سالانه مربوط به ایستگاه ایذه با ۶۱۴/۸ میلی‌متر و کم‌ترین میزان بارش سالانه در آبادان با ۱۴۹/۷ میلی‌متر ثبت شده است.

تعداد ۲۱ ایستگاه هواشناسی شامل ۱۷ ایستگاه تبخیرسنجی و چهار ایستگاه باران‌سنجی انتخاب شد. حداقل طول دوره آماری مربوط به ایستگاه پای‌پل با ۱۶ سال و حداکثر طول دوره آماری مربوط به ایستگاه‌های اهواز و سد شهدا با ۴۲ سال می‌باشد. میانگین دوره آماری ایستگاه‌های انتخاب شده بیش از ۲۹ سال می‌باشد. موقعیت جغرافیایی ایستگاه‌ها در شکل (۱) نشان داده شده است.

اقتصاد (Decker, 2005; Lisboa, 2000) هم مورد توجه قرار گرفته است، از این‌رو از این الگوریتم می‌توان برای خوشه‌بندی استفاده نمود.

Modarres (2010) پژوهش‌هایی را در مورد حداکثر طول دوره خشک‌سالی در استان اصفهان انجام داد. در این پژوهش برای انتخاب تعداد خوشه مناسب از معیار R^2 استفاده و استان اصفهان به دو خوشه تقسیم شد. ایستگاه‌های گروه اول در منطقه غرب استان اصفهان و ایستگاه‌های گروه دوم در مناطق شرقی و جنوب‌شرقی استان قرار گرفتند (Reza Modarres, 2010).

Adib et al. (2021) شاخص خشک‌سالی جریان رودخانه (SDI) را با استفاده از پارامتر شدت خشک‌سالی هیدرولوژیکی در فواصل زمانی سه، شش، نه و ۱۲ ماهه منطقه‌ای کردند. براساس پارامتر مقیاس زمانی سه ماهه، استان خوزستان در ایران به سه منطقه تقسیم شد، اما برای مقیاس زمانی دیگر، این استان به دو منطقه تقسیم شد (Adib et al., 2021).

Ghadami et al. (2020) مطالعاتی را با هدف منطقه‌ای کردن احتمال شاخص شدت خشک‌سالی شدید (SD) در سراسر ایران انجام دادند. در این پژوهش سری زمانی ۱۲ ماهه شاخص بارش استاندارد شده (SPI) برای ایستگاه‌های مورد مطالعه محاسبه و با اعمال تحلیل خوشه‌ای (CA) بر روی ایستگاه‌ها، منطقه مورد مطالعه به چهار ناحیه طبقه‌بندی شد (Ghadami et al., 2020).

بعد از کاربرد شبکه‌های عصبی گازی و دیگر مدل‌ها برای خوشه‌بندی لازم است مناطق تشکیل شده و ایستگاه‌های موجود در هر ناحیه از نظر همگنی و هم‌نوایی بررسی شوند. بدین منظور از آزمون‌های هاسکینگ و والیس که بر مبنای گشتاورهای خطی هستند و به‌عنوان بهترین روش در تحلیل‌های منطقه‌ای شناخته شده‌اند استفاده شده است (Masselot et al., 2017).

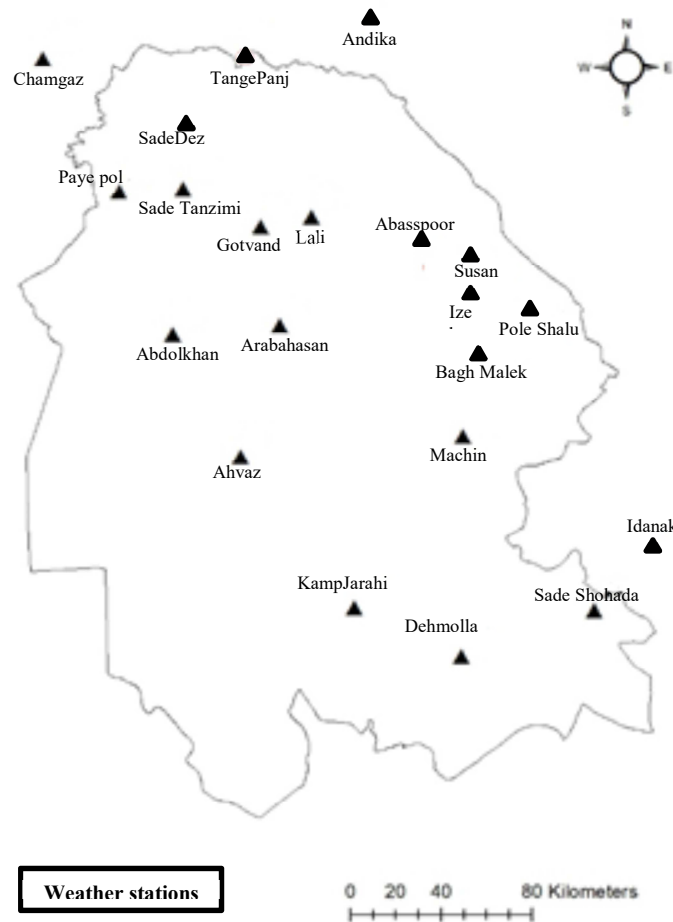


Figure 1. Study area and location of the stations

تعیین نواحی همگن

تشخیص مناطق همگن معمولاً در بین تمامی مراحل تحلیل فراوانی منطقه‌ای مهم‌ترین و سخت‌ترین گام و پیش‌نیاز فرضیه تحلیل فراوانی می‌باشد.

به دلیل متفاوت بودن جنس داده‌ها و واحدهای متفاوت آن‌ها لازم است قبل از وارد کردن آن‌ها به مدل‌های خوشه‌بندی و هم‌چنین مدل‌های مربوط به تعیین تعداد خوشه بهینه، داده‌ها نرمال شده و از نظر واحد یکسان شوند. به این منظور از رابطه سیگموئید^۱ به شرح زیر استفاده شد.

$$Q_n = 0.05 + 0.95 \frac{Q_r - Q_{min}}{Q_{max} - Q_{min}} \quad (\text{رابطه ۱})$$

که در آن، Q_n مقادیر نرمال‌شده، Q_r مقادیر اصلی و

اولیه و Q_{min} و Q_{max} به ترتیب مقادیر حداقل و حداکثر داده‌ها می‌باشند.

پیش از به‌کارگیری مدل‌های خوشه‌بندی لازم است تعداد خوشه بهینه موجود در منطقه تخمین زده شود. بدین منظور از شاخص‌های چو و سو (CS) (Chou et al., 2004)، نیم‌رخ (Rousseeuw, 1987) و کالینسکی هاراباسز (Caliński & Harabasz, 1974) استفاده شد.

سپس مناطق همگن احتمالی با استفاده از مدل‌های خوشه‌بندی شامل مدل‌های الگوریتم فازی سی- میانگین، نقشه خودسازنده، حداقل واریانس وارد، کا- میانگین، شبکه عصبی گازی و شبکه عصبی گازی رشدیابنده می‌باشد تعیین شدند.

فاصله زمانی (سن) از آخرین دیدار یا تجدید یال بوده است که اگر از یک اندازه بیش تر شود همسایگی شکسته می شود. این رویکرد کمک می کند تا شبکه عصبی توپولوژی را یاد بگیرد.

الگوریتم NG را می توان در مراحل زیر خلاصه کرد: مرحله صفر- موقعیت تصادفی w_i ها در فضای داده ها ایجاد می شود.

مرحله ۱- یک ورودی به نام x از داده های مدنظر انتخاب می شود.

مرحله ۲- رتبه بندی^{۱۲}، شامل محاسبه فاصله بین x و مراکز w_i ها و رتبه بندی k_i برای هر مرکز.

مرحله ۳- تطبیق یا یادگیری^{۱۳}

$$w_i^{new} = w_i^{old} + \epsilon e^{-\frac{k_i}{\lambda}} (x - w_i^{old}) \quad (\text{رابطه ۴})$$

نکته مهم این است که در طول دوره آموزش با جلوتر رفتن الگوریتم باید از سرعت یادگیری کاسته شود، در غیر این صورت شبکه عصبی دچار تکرار شده و چرخه نادرستی ایجاد خواهد شد. بدین منظور باید مقدار λ و ϵ با پیش رفتن یادگیری کاهش یابد. بدین منظور از تابع زیر استفاده خواهد شد.

$$G(t) = G_i \left(\frac{G_f}{G_i} \right)^{\frac{t}{t_{max}}} \quad \lambda_i > \lambda_f, \quad \epsilon_i > \epsilon_f, \quad T_i < T_f \quad (\text{رابطه ۵})$$

که در آن اندیس i نشان دهنده مقدار پارامتر در ابتدای یادگیری و اندیس f نمایشگر مقدار پارامتر در انتهای فرایند یادگیری می باشد. به عنوان مثال اگر $t=0$ باشد در نتیجه $\lambda_i(t) = \lambda_i$ و چنانچه $t = t_{max}$ باشد در نتیجه $\lambda_i(t) = \lambda_f$.

مرحله ۴- (ایجاد همسایگی) یالی بین دو رتبه اول از نظر نزدیکی ایجاد می شود و سن این یال برابر صفر در نظر گرفته می شود.

مرحله ۵- سن همه یال ها افزایش می یابد $(t_{i,j} \rightarrow t_{i,j} + 1)$.

گام بعدی تأیید خوشه است که در این مرحله همگنی^۹ و ناجوری^{۱۰} ایستگاه های هر خوشه با آماره های مناسب گشتاور خطی بررسی می شود.

شبکه عصبی گازی

شبکه عصبی گازی یکی از انواع شبکه های عصبی رقابتی است و از شیوه آموزش غیر نظارت شده استفاده می کند. از این رو، از این الگوریتم می توان برای خوشه بندی هم چنین بخش بندی تصویر^{۱۱} استفاده نمود.

قاعده یادگیری در شبکه عصبی گازی به شکل زیر است:

$$w_i^{new} = w_i^{old} + \alpha_i (x - w_i^{old}) \quad (\text{رابطه ۲})$$

$$\alpha_i = \epsilon e^{-\frac{k_i}{\lambda}} \quad (\text{رابطه ۳})$$

که در آن w_i مولکول گاز ایجاد شده روی فضای داده می باشد. تعداد این مولکول ها در ابتدا یک مقدار فرض می شود و در انتها به منظور کارکرد منطقی تر و بهینه تر الگوریتم اصلاح می شوند. هم چنین این عناصر به صورت اتفاقی در دامنه داده های اصلی انتخاب می شوند. α_i پارامتری است که نرخ یادگیری را مشخص می کند و وابسته به k_i و λ می باشد. به این صورت که اگر λ به سمت بی نهایت میل کند یادگیری همه نورون ها با هم برابر و اگر به سمت صفر میل کند نزدیک ترین نورون شروع به یادگیری می کند. حالت های حدی λ هیچکدام به تنهایی مناسب نمی باشد و معمولاً حالتی بین آنها انتخاب می شود. k_i نشان دهنده تعداد نورون برتر نسبت به نورون i ام می باشد. ϵ نیز عددی ثابت است که نرخ یادگیری را کنترل می کند.

به منظور ایجاد همسایگی بین اولین و دومین نورون از نظر نزدیکی یک یال ایجاد می شود. برای هر نورون داریم $c_{i,j} \in \{0,1\}$ که نشان دهنده وجود یا عدم وجود یال یا همسایگی و هم چنین $t_{i,j} \in \{0,1,2, \dots\}$ که نشان دهنده

مرحله ۳- نزدیک‌ترین نورون S_1 و نزدیک‌ترین نورون بعد از آن S_2 پیدا می‌شود.

مرحله ۴- افزایش سن تمام لبه‌هایی که از S_1 به همسایگان خود می‌رسند.

مرحله ۵- خطای محلی s_1 با استفاده از فاصله اقلیدسی بین دو بردار افزایش داده می‌شود:

$$\Delta E_{s_1} = \|w_{s_1} - x\|^2 \quad (\text{رابطه ۶})$$

مرحله ۶- S_1 و همسایگی‌های توپولوژیک مستقیم آن به سمت x با نرخ یادگیری ϵ_b و ϵ_n حرکت داده می‌شود به ترتیب از فاصله کل:

$$\Delta w_{s_1} = \epsilon_b(x - w_{s_1}) \quad (\text{رابطه ۷})$$

$$\Delta w_n = \epsilon_n(x - w_n) \quad (\text{رابطه ۸})$$

که در این روابط، n نشان‌دهنده همسایگان مستقیم s_1 است.

مرحله ۷- اگر s_1 و s_2 با یک لبه متصل شوند، سن این لبه صفر تنظیم می‌شود. اگر چنین لبه‌ای وجود نداشته باشد، آن را باید ایجاد نمود.

مرحله ۸- لبه‌های با سن بزرگ‌تر از amax حذف می‌شود. اگر این نتایج در نقاط بدون لبه ظاهر شوند، آن‌ها نیز حذف می‌شوند.

مرحله ۹- اگر تعداد بردارهای ورودی ارائه شده تاکنون یک عدد صحیح از یک پارامتر λ است، یک نورون جدید به صورت زیر وارد می‌شود:

• نورون q با بزرگ‌ترین متغیر خطا تعیین می‌شود.

• نورون f با بزرگ‌ترین متغیر خطا در میان همسایگان نورون q پیدا می‌شود.

• یک نورون جدید به نیمه‌راه بین q و f به عنوان $w_f = 0.5(w_q + w_f)$ قرار داده می‌شود.

• ایجاد لبه‌های اتصال نورون r با نورون q و f و حذف لبه اصلی بین q و f

مرحله ۶- فرض می‌شود $k_i = 0$ و به ازای هر j که $t_{i,j} > T$ باشد، $c_{i,j} = 0$ در نظر گرفته می‌شود. در این مرحله نیز باید به دلیل ذکر شده در مرحله سه، T در طول دوره یادگیری افزایش یابد تا از میزان سخت‌گیری کاسته شود یعنی اجازه داده می‌شود پال‌ها بیش‌تر دوام داشته باشند. مرحله ۷- در صورت برآورده نشدن شرایط خاتمه (به عنوان مثال، حداکثر تعداد نورون‌ها و یا هر اندازه عملکرد)، از مرحله یک تکرار می‌شود، در غیر این صورت مراحل الگوریتم پایان می‌یابد.

شبکه عصبی گازی رشدیابنده

GNG یک الگوریتم خوشه‌بندی است که به صورت قدم‌به‌قدم کار می‌کند، یعنی تعداد نورون‌ها در طی فرایند آموزش بدون استفاده از دانش پیشین در مورد ساختار الگوهای ورودی، افزایش می‌یابد (Fink et al., 2015; Angelopoulou et al., 2015). برخلاف الگوریتم‌های خوشه‌بندی کلاسیک، الگوریتم GNG دارای یک ساختار شبکه سازگار است که آن را برای یادگیری توپولوژی مجموعه داده‌های با ابعاد بزرگ مناسب می‌سازد (Linda & Manic, 2009; Zaki & Yin, 2008). ایده اصلی شبکه GNG این است که به طور پیوسته گره‌های جدید (نورون‌ها) را به یک شبکه ابتدا کوچک در یک ساختار در حال رشد اضافه کند. در GNG، نورون‌های شبکه برای تعیین آن‌هایی که دارای بیش‌ترین شباهت به مجموعه داده ورودی هستند، رقابت می‌کنند (Moreli et al., 2014). الگوریتم GNG را می‌توان در مراحل زیر خلاصه کرد:

مرحله ۱- با دو نورون a و b در موقعیت‌های تصادفی w_b و w_a در فضای ورودی شروع می‌شود.

مرحله ۲- یک بردار ورودی (x) از مجموعه داده‌های مدنظر ارائه می‌شود.

در مرحله بعد برای محاسبه پراکنش نسبت گشتاورهای خطی ناحیه‌ای اقدام به شبیه‌سازی داده‌های منطقه‌ای معادل می‌شود. برای این منظور باید تابع توزیع مناسبی در نظر گرفته و با استفاده از آمار واقعی ناحیه‌ای، پارامترهای توزیع انتخابی در هر ناحیه تعیین شود. برای هر ناحیه شبیه‌سازی شده V_1 محاسبه می‌شود. همچنین از داده‌های شبیه‌سازی شده، میانگین μ_v و انحراف استاندارد σ_v تعیین و معیار غیر همگنی از طریق رابطه (۱۱) مشخص می‌شود.

$$H_i = \frac{(V_i - \mu_v)}{\sigma_v} \quad (\text{رابطه ۱۱})$$

در پژوهشی Hosking & Wallis (1993) پیشنهاد کردند که یک ناحیه وقتی می‌تواند ناحیه همگن قابل قبول باشد که H_i کم‌تر از یک باشد و می‌تواند ناحیه نسبتاً غیرهمگن باشد اگر H_i بین یک و دو باشد و به‌عنوان ناحیه به‌طور مشخص غیرهمگن خواهد بود اگر H_i بزرگ‌تر از دو باشد. در عمل عنوان می‌شود که معیار H_1 مناسب‌تر است.

علاوه بر آماره‌های گشتاورهای خطی که برای مقایسه مدل‌های خوشه‌بندی به‌کار گرفته شد اختلاف میان منحنی‌های IDF منطقه‌ای حاصل از مدل‌های خوشه‌بندی که براساس توابع توزیع منطقه‌ای بنا شده‌اند با منحنی‌های IDF ایستگاهی که بر مبنای توابع توزیع در تداوم‌های مختلف در هر ایستگاه می‌باشند، از سه رابطه، ضریب تغییرات جذر میانگین مربع خطا^{۱۴} (CV_{RMSE})، میانگین درصد اختلاف^{۱۵} (Δ) و انحراف نتایج^{۱۶} (MBE) استفاده شد. هر چقدر مقادیر CV_{RMSE} و Δ کم‌تر باشند، نشان‌دهنده دقت بیشتر مدل به‌کاررفته در خوشه‌بندی می‌باشد. همچنین مقادیر منفی MBE نشان‌دهنده تخمین بیش‌تر و مقادیر مثبت نشان‌دهنده تخمین کم‌تر مقادیر منطقه‌ای نسبت به مقادیر ایستگاهی بارش‌ها می‌باشد.

• متغیرهای خطای q و f با ضرب آن‌ها در یک کسر (a)، کاهش داده می‌شوند. متغیر خطای r با مقدار جدیدی از متغیر خطای q تنظیم می‌شود.

مرحله ۱۰- تمام متغیرهای خطا با ضرب آن‌ها در کسر β کاهش داده می‌شود.

مرحله ۱۱- اگر معیار توقف (به‌عنوان مثال، حداکثر تعداد نوروها و یا هر اندازه عملکرد) هنوز انجام نشده است، به مرحله دو برگشت داده می‌شود.

به‌طور خلاصه، پارامتر سن در نظر گرفته‌شده در مرحله چهارم نشان می‌دهد که ارتباط بین نوروها چقدر قوی است. در مرحله پنجم، خطای محلی عصبی برای شناسایی مناطقی که نوروها به اندازه کافی برای بردارهای ورودی سازگار نیستند، استفاده می‌شود. سازگاری شبکه با فضای ورودی در مرحله ششم اتفاق می‌افتد. قرارداد لبه‌ها (مرحله هفتم) بین دو نزدیک‌ترین نورو برای الگوهای ورودی، بخشی از ساختار ساختمان توپولوژیک است. حذف لبه‌ها (مرحله هشتم) برای از بین بردن لبه‌های بین نوروهایی که دیگر فعال نیستند ضروری است. در مرحله نهم نورو جدید در مناطق فضای ورودی با استفاده از خطای انباشته (مرحله پنجم) وارد می‌شود. در نهایت، شبکه تا زمانی که شرایط پایان را به‌دست آورد ادامه می‌یابد (Fišer et al., 2013; Quintana-Pacheco et al., 2014).

آزمون همگنی خوشه‌ها

برای محاسبه میزان ناهمگنی ابتدا V_1 با استفاده از رابطه (۹) برای داده‌های مشاهداتی به‌دست می‌آید.

$$V_1 = \sum_{i=1}^N n_i (t^{(i)} - \bar{t})^2 / \sum_{i=1}^N n_i \quad (\text{رابطه ۹})$$

$$\bar{t} = (\sum_{i=1}^N N_i t^{(i)}) / (\sum_{i=1}^N n_i) \quad (\text{رابطه ۱۰})$$

که در آن n_i اندازه نمونه در ایستگاه i ، $t^{(i)}$ گشتاور خطی نمونه (L-CV)، \bar{t} میانگین نقطه‌ای گشتاور نمونه (L-CV) می‌باشند.

نتایج و بحث تعداد خوشه بهینه

با استفاده از داده‌های میانگین بارش سالانه، ارتفاع، حداکثر بارش ۲۴ ساعته، طول جغرافیایی و عرض جغرافیایی اقدام به تعیین تعداد خوشه بهینه شد. بدین منظور از نرم‌افزار Matlab (2016 a) کمک گرفته شد که نتایج آن در شکل (۲) آورده شده است. در شاخص CS هرچه مقدار شاخص عدد کم‌تر و در شاخص‌های کالینسکی-هاراباسز و نیم‌رخ هرچه مقدار شاخص بیش‌تر باشد تعداد خوشه‌های در نظر گرفته شده مناسب‌تر است. با توجه به نتیجه این سه روش تعداد خوشه بهینه، دو خوشه در نظر گرفته شد.

$$CV_{RMSE} = \frac{\sqrt{\frac{1}{N_d N_T} \sum_{d=1}^{N_d} \sum_{T=1}^{N_T} (x_{d,T} - z_{d,T})^2}}{\frac{1}{N_d N_T} \sum_{d=1}^{N_d} \sum_{T=1}^{N_T} x_{d,T}} \times 100 \quad (\text{رابطه } 12)$$

$$\Delta = \frac{1}{N_d N_T} \sum_{d=1}^{N_d} \sum_{T=1}^{N_T} \frac{|x_{d,T} - z_{d,T}|}{x_{d,T}} \times 100 \quad (\text{رابطه } 13)$$

$$MBE = \frac{\sum_{d=1}^{N_d} \sum_{T=1}^{N_T} (x_{d,T} - z_{d,T})}{N_d N_T} \quad (\text{رابطه } 14)$$

در روابط بالا، $x_{d,T}$ و $z_{d,T}$ به ترتیب شدت حداکثر بارندگی در تداوم d ام و دوره بازگشت T ام در ایستگاه هواشناسی موردنظر و منطقه همگنی که آن ایستگاه در آن قرار گرفته می‌باشد. همچنین N_d و N_T به ترتیب، تعداد تداوم‌ها و دوره‌های بازگشت مورد استفاده است.

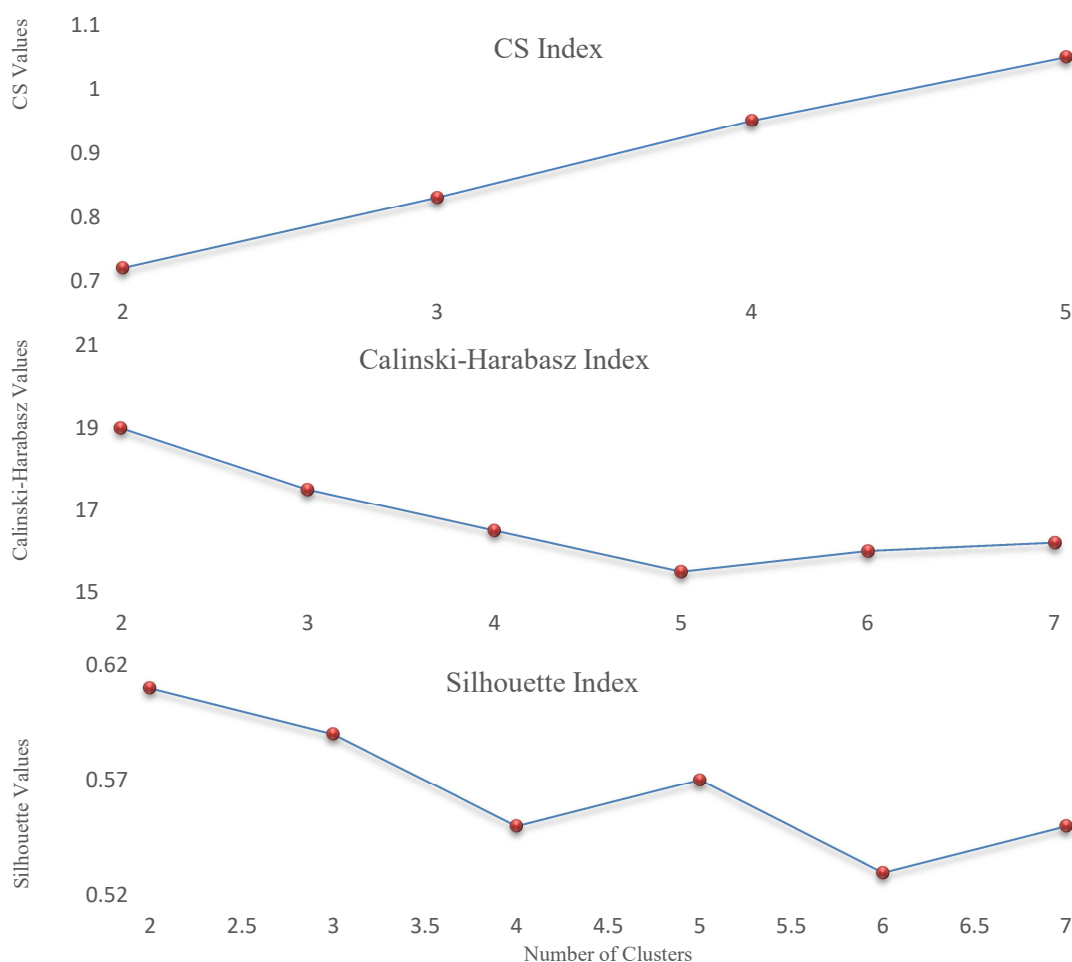


Figure 2. Calinsky-Harabasz values versus the number of clusters. Silhouette values versus the number of clusters CS values versus the number of clusters.

نقشه‌های استخراج شده از خوشه‌بندی به روش‌های مختلف

پس از تعیین تعداد مناطق بهینه روش‌های خوشه‌بندی شبکه عصبی گازی (NG)، شبکه عصبی گازی رشدیابنده (GNG)، شبکه عصبی خودسازمانده (SOM) و مدل فازی

سی میانگین (FCM) در نرم‌افزار متلب و روش‌های کا- میانگین (K-Means) و وارد (Ward) در نرم‌افزار R (3.5.0) پیاده‌سازی شدند. در نتیجه تعلق هر ایستگاه به مناطق دوگانه بررسی و در شکل‌های (۳) و (۴) ارائه شده است.

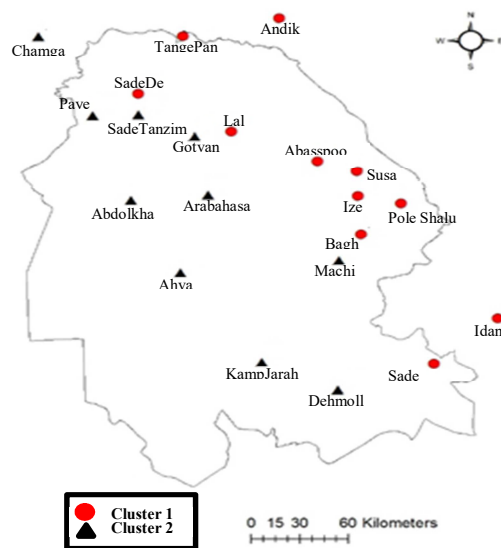


Figure 3. Location of stations in the cluster identified by NG

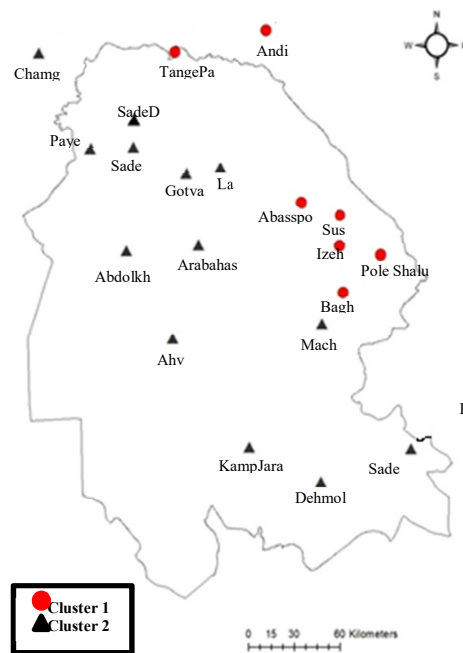


Figure 4. Location of stations in the cluster identified by GNG

می‌باشد، که نشان‌دهنده برتری این روش نسبت به دیگر روش‌ها می‌باشد. هم‌چنین منفی بودن شاخص MBE برای این مدل نمایانگر این موضوع است که مقادیر عددی منحنی‌های IDF منطقه‌ای حاصل از این روش نسبت به مقادیر ایستگاهی تا حدودی دارای مقادیر بزرگ‌تری هستند. هم‌چنین با در نظر گرفتن هر سه شاخص می‌توان شبکه عصبی گازی رشدیابنده را به‌عنوان دومین مدل مناسب در نظر گرفت. با در نظر گرفتن مقادیر خطای بین IDF منطقه‌ای با IDF ایستگاهی که در جدول (۳) آورده شده است، به این نتیجه می‌رسیم که در ایستگاه‌های دارای طول آماری کوتاه میزان خطا افزایش پیدا کرده که در صورت حذف این ایستگاه‌ها می‌توان عملکرد بهتری را از مدل‌های خوشه‌بندی استفاده‌شده انتظار داشت. مقایسه بین دو نوع منحنی منطقه‌ای و ایستگاهی در چهار ایستگاه مختلف در شکل (۵) نمایش داده شده است.

نتایج بررسی همگنی خوشه‌ها و ناجوری ایستگاه‌ها

به‌منظور بررسی همگنی منطقه‌ای و ناجوری ایستگاه‌های موجود در هر منطقه به‌ترتیب از آماره‌های H و D_i که آزمون‌هایی براساس گشتاورهای خطی می‌باشند استفاده شد. این مقادیر برای حداکثر شدت بارش‌های سالانه در تداوم‌های مختلف و هم‌چنین مدل‌های متنوع به‌کاررفته برای خوشه‌بندی تعیین شدند که در جدول (۱) ارائه شده است.

نتایج بررسی بهترین مدل خوشه‌بندی

به‌منظور تعیین بهترین مدل خوشه‌بندی، مقادیر عددی منحنی‌های منطقه‌ای با مقادیر نظیر آن‌ها در منحنی‌های ایستگاهی مقایسه شدند و نتایج در جدول (۲) ارائه شده است. همان‌طور که از جدول مشخص است با توجه به مقادیر خطاهای محاسبه‌شده، مدل خوشه‌بندی شبکه عصبی گازی کم‌ترین میزان خطا را در هر دو شاخص دارا

Table 1. Results of heterogeneity and discordancy measures for 24hr rainfall duration

Cluster	Clustering models	Number of stations	Discordant stations	Heterogeneity measure			Heterogeneity situation
				H_1	H_2	H_3	
1	NG	11	-	-1.12	-1.13	-1.04	Definitely homogeneous
	GNG	8	-	-0.91	-1.91	-1.63	Definitely homogeneous
	SOM	9	-	-1.11	-2.01	-1.81	Definitely homogeneous
	FCM	9	-	-0.64	-1.79	-1.59	Definitely homogeneous
	Kmeans	9	-	-1.11	-2.01	-1.81	Definitely homogeneous
	Ward	9	-	-1.11	-2.01	-1.81	Definitely homogeneous
	2	NG	10	-	0.17	-1.28	-2.33
GNG		13	-	0.23	-0.12	-1.3	Definitely homogeneous
SOM		12	-	0.39	-0.04	-1.21	Definitely homogeneous
FCM		12	-	0.12	-0.01	-1.09	Definitely homogeneous
Kmeans		12	-	0.39	-0.04	-1.21	Definitely homogeneous
Ward		12	-	0.39	-0.04	-1.21	Definitely homogeneous

Table 2. Average values of goodness-of-fit indices of the difference between IDF curves based on regional and at-site probability distributions for used clustering models

Model	Values of goodness-of-fit indices		
	CV_{RMSE}	Δ	MBE
NG	24.39	15.56	-0.04
GNG	24.41	16.03	0.23
FCM	25.19	16.35	0.32
SOM, K-Means, Ward	24.67	15.91	0.31

Table 3. Values for goodness-of-fit indices of the difference between IDF curves based on regional and at-site probability distributions at 21 rainfall stations for NG clustering model

Stations	Values of goodness-of-fit indices		
	CV_{RMSE}	MBE	Δ
1	29.46	-2.11	10.75
2	42.46	-6.61	24.84
3	26.38	-2.89	11.67
4	14.68	-1.42	10.04
5	15.61	-1.77	9.14
6	23.07	-2.05	8.51
7	17.49	0.28	14.29
8	23.38	-0.11	10.32
9	13.59	1.43	5.91
10	21.3	6.85	28.14
11	18.67	1.58	8.33
12	16.27	0.80	13.06
13	23.36	-1.38	20.99
14	46.52	9.48	28.56
15	27.05	-2.34	14.41
16	26.02	-3.71	18.87
17	29.19	0.27	19.62
18	15.55	2.13	22.00
19	35.81	-3.70	21.42
20	30.17	5.12	14.33
21	16.11	-0.69	11.72
Mean	24.39	-0.04	15.56

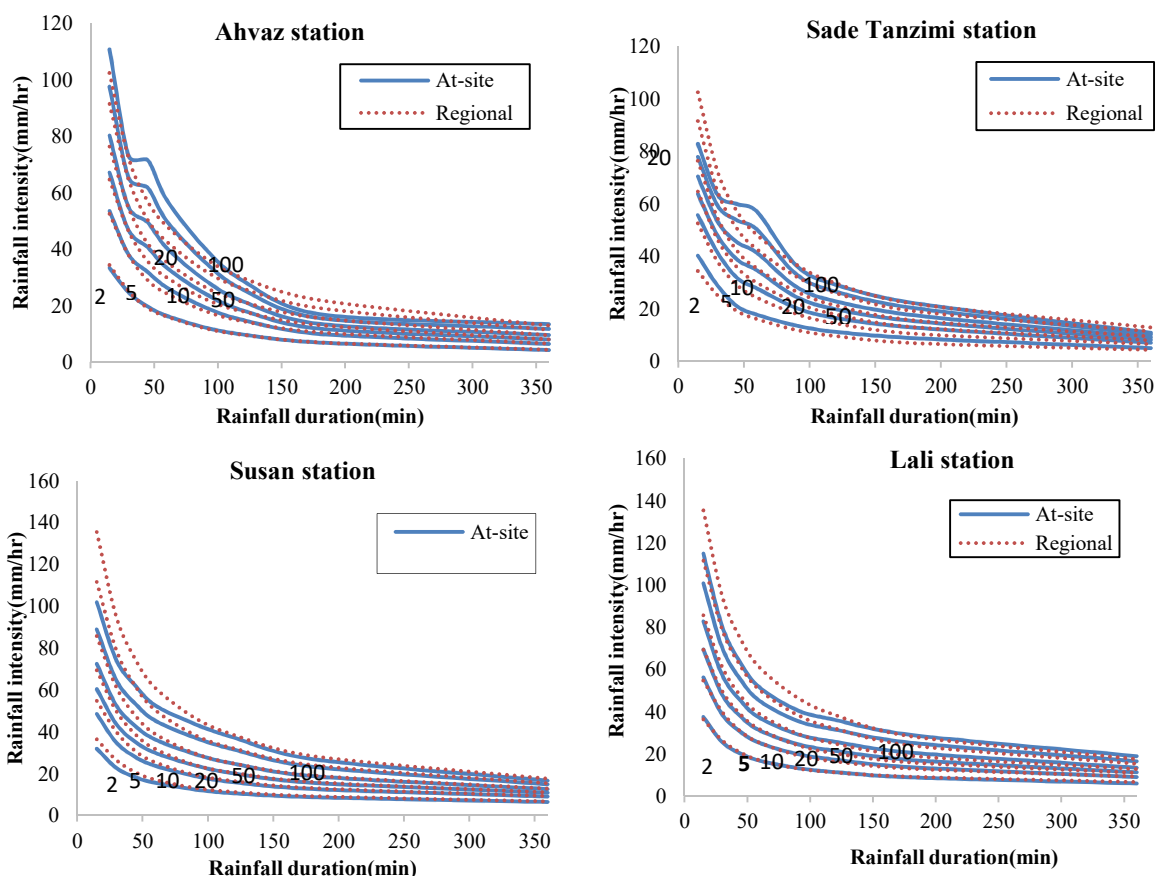


Figure 5. IDF curves for four rainfall stations in Khuzestan province (Note: The values on each curve is the return period T for the curve)

نتیجه‌گیری

در این پژوهش دو مدل جدید شبکه‌های عصبی گازی و گازی رشدیابنده به منظور منطقه‌ای نمودن منحنی‌های IDF ارائه شد. به این منظور با در نظر گرفتن مشخصه‌های طول جغرافیایی، عرض جغرافیایی، میانگین بارندگی سالانه، ارتفاع از سطح دریا و حداکثر بارش ۲۴ ساعته سالانه مربوط به هر ایستگاه و با استفاده از سه شاخص چو و سو (CS)، نیم‌رخ (Silhouette) و کالینسکی-هاراباسز (CH) مشخص شد که استان خوزستان در بهینه‌ترین حالت دارای دو منطقه مجزا و احتمالاً همگن می‌باشد. سپس با استفاده از مدل‌های مختلف خوشه‌بندی نسبت به منطقه‌ای نمودن ایستگاه‌های موجود اقدام شد. خوشه‌بندی به دلیل حساسیت و تأثیرگذاری فراوان در نتیجه پایانی، یکی از مهم‌ترین و اصلی‌ترین مراحل انجام این پژوهش بود. از این‌رو، با استفاده از شش مدل متفاوت عملیات خوشه‌بندی انجام شد. روش‌های خوشه‌بندی، وارد (Ward)، کا میانگین (K-Means)، فازی سی میانگین (FCM) و شبکه عصبی خودسازمانده (SOM) از جمله روش‌های رایجی هستند که به‌طور گسترده استفاده می‌شوند. علاوه بر چهار روش ذکر شده، از دو مدل جدید شبکه عصبی گازی (NG) و شبکه عصبی گازی رشدیابنده (GNG) به منظور خوشه‌بندی استفاده شد. به منظور بررسی همگنی دو منطقه ایجاد شده و همچنین ناجور بودن یا نبودن ایستگاه‌های موجود در هر منطقه، در هر شش مدل در تداوم‌های یازده‌گانه، از آزمون‌های همگنی ناحیه‌ای و آزمون ناجوری که بر اساس گشتاورهای خطی می‌باشند استفاده شد. در اکثر مدل‌ها و تداوم‌های مختلف، مناطق ایجاد شده در اثر خوشه‌بندی دارای همگنی مناسبی بودند. پس از مشخص شدن موقعیت هر کدام از ایستگاه‌ها در مناطق دوگانه و همگن تشخیص داده شدن هر دو ناحیه، نسبت به تعیین تابع توزیع منطقه‌ای و پس از آن استخراج

IDFهای منطقه‌ای با استفاده از روش گشتاورهای خطی اقدام شد. منحنی IDF منطقه‌ای به دست آمده برای هر ناحیه با منحنی‌های IDF ایستگاهی به دست آمده در همان مناطق مورد مقایسه قرار گرفت. نتایج نشان داد که در تمام ایستگاه‌ها، منحنی‌های IDF منطقه‌ای و ایستگاهی تطابق بالایی داشته و روند یکسانی را به نمایش می‌گذارند. این پژوهش برای اولین بار کارایی شبکه‌های عصبی گازی را در مبحث منطقه‌ای نمودن منحنی‌های IDF مورد ارزیابی قرار داده است. در بین IDFهای منطقه‌ای استخراج شده، منحنی‌های حاصل از مناطق استخراج شده از مدل‌های شبکه عصبی گازی و شبکه عصبی گازی رشدیابنده، دارای بیش‌ترین دقت و بیش‌ترین تطابق با منحنی‌های ایستگاهی بودند که نشان‌دهنده کارایی این مدل‌ها در مبحث تحلیل فراوانی منطقه‌ای می‌باشد. از این‌رو، کیفیت عملکرد شبکه‌های عصبی گازی می‌تواند موجب بهبود مسائل و مشکلات مختلف در مباحث مربوط به آب و مدیریت و برنامه‌ریزی منابع آب شود.

پی‌نوشت‌ها

1. Generalized Extreme Value Distribution
2. Neural Gas Network
3. Growing Neural Gas Network
4. Ward
5. K-Means
6. Self-Organizing Map (SOM)
7. Fuzzy C-Means
8. Sigmoid
9. Homogeneity
10. Discordancy
11. Image Segmentation
2. Aging
3. Adaptation
4. Coefficient of Variation of Root Mean Square Error
5. Mean Percentage Difference
6. Mean Bias Error

تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد.

منابع

1. Abdi, A., Hassanzadeh, Y., & Ouarda, T. B. (2017). Regional frequency analysis using Growing Neural Gas network. *Journal of Hydrology*, 550, 92-102.
2. Adib, A., Kashani, A., & Ashrafi, S. M. (2021). Merge L-Moment Method, Regional Frequency Analysis and SDI for Monitoring and Zoning Map of Short-Term and Long-Term Hydrologic Droughts in the Khuzestan Province of Iran. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 45(4), 2681-2694.
3. Alemaw, B. F., & Chaoka, R. T. (2016). Regionalization of Rainfall Intensity-Duration-Frequency (IDF) Curves in Botswana. *Journal of Water Resource and Protection*, 8(12), 1128.
4. Angelopoulou, A., Psarrou, A., Garcia-Rodriguez, J., Orts-Escolano, S., Azorin-Lopez, J., & Revett, K. (2015). 3D reconstruction of medical images from slices automatically landmarked with growing neural models. *Neurocomputing*, 150, 16-25.
5. Ariff, N. M., Jemain, A. A., & Bakar, M. A. A. (2016). Regionalization of IDF curves with L-moments for storm events. *International Journal of Mathematical and Computational Sciences*, 10(5), 217-223.
6. Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3(1), 1-27.
7. Carlevarino, A., Martinotti, R., Metta, G., & Sandini, G. (2000, July). An incremental growing neural network and its application to robot control. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, 5, 323-328. IEEE.
8. Chaubey, I., Haan, C. T., Grunwald, S., & Salisbury, J. M. (1999). Uncertainty in the model parameters due to spatial variability of rainfall. *Journal of Hydrology*, 220(1-2), 48-61.
9. Chou, C. H., Su, M. C., & Lai, E. (2004). A new cluster validity measure and its application to image compression. *Pattern Analysis and Applications*, 7(2), 205-220.
10. Cselényi, Z. (2005). Mapping the dimensionality, density and topology of data: The growing adaptive neural gas. *computer methods and programs in biomedicine*, 78(2), 141-156.
11. de Oliveira Martins, L., Silva, A. C., De Paiva, A. C., & Gattass, M. (2009). Detection of breast masses in mammogram images using growing neural gas algorithm and Ripley's K function. *Journal of Signal Processing Systems*, 55(1), 77-90.
12. Decker, R. (2005). Market basket analysis by means of a growing neural network. *The International Review of Retail, Distribution and Consumer Research*, 15(2), 151-169.
13. Durrans, S. R., & Kirby, J. T. (2004). Regionalization of extreme precipitation estimates for the Alabama rainfall atlas. *Journal of Hydrology*, 295(1-4), 101-107.
14. Ferrer, G. J. (2014). Creating Visual Reactive Robot Behaviors Using Growing Neural Gas. In *MAICS* (pp. 39-44).
15. Fink, O., Zio, E., & Weidmann, U. (2015). Novelty detection by multivariate kernel density estimation and growing neural gas algorithm. *Mechanical Systems and Signal Processing*, 50, 427-436.
16. Fišer, D., Faigl, J., & Kulich, M. (2013). Growing neural gas efficiently. *Neurocomputing*, 104, 72-82.
17. Fritzsche, B. (1995). A growing neural gas network learns topologies. *Advances in neural information processing systems*, 7, 625-632.
18. Ghadami, M., Raziqi, T., Amini, M., & Modarres, R. (2020). Regionalization of drought severity-duration index across Iran. *Natural Hazards*, 103(3), 2813-2827.
19. Goovaerts, P. (1999). Using elevation to aid the geostatistical mapping of rainfall erosivity. *Catena*, 34(3-4), 227-242.
20. Hosking, J.R.M., & Wallis, J.R. (1993). Some statistics useful in regional frequency analysis. *Water resources research*, 29(2), 271-281.
21. Lee, S. H., & Maeng, S. J. (2003). Frequency analysis of extreme rainfall using L-moment. *Irrigation and Drainage: The journal of the International Commission on Irrigation and Drainage*, 52(3), 219-230.
22. Linda, O., & Manic, M. (2009, June). GNG-SVM framework-classifying large datasets with Support Vector Machines using Growing Neural Gas. In *2009 International Joint Conference on Neural Networks*. (pp. 1820-1826). IEEE.
23. Lisboa, P. J., Edisbury, B., & Vellido, A. (2000). *Business applications of neural networks: the state-of-the-art of real-world applications* (Vol. 13). World scientific.
24. Martinetz, T., & Schulten, K. (1991). *A "neural-gas" network learns topologies*.
25. Masselot, P., Chebana, F., & Ouarda, T. B. (2017). Fast and direct nonparametric procedures in the L-moment homogeneity test. *Stochastic Environmental Research and Risk Assessment*, 31(2), 509-522.

26. Modarres, R. (2010). Regional dry spells frequency analysis by L-moment and multivariate analysis. *Water resources management*, 24(10), 2365-2380.
27. Moreli, V., Cazorla, M., Orts-Escolano, S., & Garcia-Rodriguez, J. (2014, July). 3d maps representation using gng. In 2014 International Joint Conference on Neural Networks (IJCNN) (pp. 1482-1487). IEEE.
28. Quintana-Pacheco, Y., Ruiz-Fernández, D., & Magrans-Rico, A. (2014). Growing Neural Gas approach for obtaining homogeneous maps by restricting the insertion of new nodes. *Neural networks*, 54, 95-102.
29. Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20, 53-65.
30. Soltani, S., Helfi, R., Almasi, P., & Modarres, R. (2017). Regionalization of rainfall intensity-duration-frequency using a simple scaling model. *Water Resources Management*, 31(13), 4253-4273.
31. Zaki, S. M., & Yin, H. (2008). A semi-supervised learning algorithm for growing neural gas in face recognition. *Journal of Mathematical Modelling and Algorithms*, 7(4), 425-435.