

## ارزیابی روش‌های خوشه‌بندی فازی و شبکه عصبی مصنوعی در پهنه‌بندی فضایی بارش سالانه ایران

علی شاهبایی کوتنایی\* و حسین عساکره<sup>۱</sup>

(تاریخ دریافت: ۱۴۰۰/۱۱/۲۰؛ تاریخ پذیرش: ۱۴۰۱/۲/۲۴)

### چکیده

بارش یکی از عناصر اصلی چرخه هیدرواقليمی کره زمین است که چگونگی تغییرپذیری آن تابع روابط پیچیده و غیرخطی بین سامانه اقليم و عوامل محیطی است. شناخت این روابط و انجام برنامه‌ریزی‌های محیطی بر مبنای آنها در حالت عادی امری دشوار بوده و در نتیجه طبقه‌بندی داده‌ها و تقسیم اطلاعات به دسته‌های متجانس و کوچک می‌تواند در این زمینه کمک‌کننده باشد. در پژوهش حاضر تلاش شد تا داده‌های بارش، ارتفاع، شیب، جهت دامنه‌ها و تراکم ایستگاهی برای ۳۴۲۳ ایستگاه همدید، اقليم‌شناسی و باران‌سنجی ایران در دوره ۱۹۶۱-۲۰۱۵ به مدل‌های فازی (FCM) و شبکه عصبی خودسازمانده (SOM-ANN) وارد و ضمن انجام پهنه‌بندی بارشی- محیطی، خروجی‌های دو مدل از لحاظ دقت و کارایی مورد مقایسه قرار گیرد. نتایج حاصل از خروجی مدل‌ها، شرایط بارشی ایران را در ارتباط با عوامل محیطی به چهار پهنه تقسیم‌بندی کرده است. همچنین، بررسی شاخص‌های ارزیابی نشان داد هر دو مدل از دقت بالایی در امر طبقه‌بندی فراسنج بارش برخوردار هستند؛ با این حال مدل فازی از لحاظ دقت نتایج برتری نسبی به مدل شبکه عصبی دارد.

واژه‌های کلیدی: شبکه عصبی خودسازمانده، مدل فازی، پهنه‌بندی، بارش، ایران

۱. گروه جغرافیا، دانشکده علوم انسانی، دانشگاه زنجان

\*: مسئول مکاتبات: پست الکترونیکی: Shahbai2008@yahoo.com

## مقدمه

اقلیمی و غیراقلیمی صورت پذیرفته است، از جمله می‌توان به پژوهش چتینی و بوگاپاتی (۷) اشاره کرد که با ترکیب روش‌های فازی C-Means و شبکه عصبی و اعمال آن بر روی داده‌های آزمایشی Iris و Thyroid دقت این روش‌ها در ارائه طبقه‌بندی از داده‌های مذکور را مورد بررسی قرار دادند. یان و همکاران (۲۸) سامانه ANFIS را برای طبقه‌بندی کیفیت آب رودخانه به‌کار برده و به این نتیجه رسیدند که طبقه‌بندی با این سامانه دارای دقت ۹۰ درصدی بوده و از سایر شاخص‌های شبکه عصبی کیفیت بهتری دارد. لی و همکاران (۱۶) خصوصیات ۵۳ حوضه آبریز در آلمان را با الگوریتم‌های شبکه عصبی طبقه‌بندی کرده و آن را با طبقات به دست آمده از سایر روش‌های آماری مقایسه کردند. نتیجه این پژوهش یک همپوشانی ۶۷ درصدی بین خروجی‌های دو روش طبقه‌بندی را آشکار ساخت. مکاریان و همکاران (۱۹) نیز با همین روش مطالعه‌ای در زمینه طبقه‌بندی میزان تبخیر سالانه ایستگاه‌های تبخیرسنجی انجام دادند. گلیان و همکاران (۱۰) به مقایسه کارایی روش‌های خوشه‌بندی در طبقه‌بندی الگوهای فضایی بارش پرداختند. شیرانی و همکاران (۲۴) دقت روش‌های شبکه عصبی و رگرسیون چند متغیره در پهنه‌بندی خطر زمین لغزش را بررسی کردند. ساهینا و سگیزاوغلو (۲۱) روش‌های سنتی و نروفازی را به‌منظور طبقه‌بندی اقلیمی و رژیم‌های بارشی ترکیه مورد بررسی قرار دادند و دریافتند که روش‌های مبتنی بر شبکه عصبی مصنوعی کارایی ۹۵/۳۳ درصدی داشته و از دقت بالاتری نسبت به روش‌های سنتی خوشه‌بندی برخوردار هستند. هیبو و همکاران (۱۲) کاربرد و دقت روش‌های خوشه‌بندی شبکه عصبی و منطق فازی را در برآورد و طبقه‌بندی رواناب‌ها و سیلاب‌ها مورد تجزیه و تحلیل قرار دادند. بررسی پژوهش‌های بالا نشان داد که روش‌های نوین طبقه‌بندی، از کاربرد گسترده‌ای در شناسایی روابط پیچیده و پنهان عوامل و عناصر مختلف برخوردار بوده و شناسایی و استفاده از بهینه‌ترین روش‌ها می‌تواند در افزایش دقت و کاهش عدم

یکی از متغیرترین عناصر کلیدی آب و هوا که الگوی مکانی و زمانی دسترسی به منابع آب را تحت تأثیر قرار می‌دهد بارش است (۲۷). تغییرات این فراسنج حاصل کنش‌ها و واکنش‌های بی‌وقفه و پیوسته اجزای مختلف سامانه آب و هوا بوده (۳) و ارتباطی تنگاتنگ با عوامل محلی - مکانی دارد (۲۹). با وجود پیچیدگی‌های پیدا و پنهان در این روابط و با توجه به اینکه انجام هر نوع برنامه‌ریزی محیطی به شناخت نسبی از این کنش‌ها و واکنش‌ها بستگی دارد (۲۳)، انجام فرایندهای طبقه‌بندی در این زمینه، ضروری به نظر می‌رسد. طبقه‌بندی به مجموعه روش‌هایی اطلاق می‌شود که در آنها هر نمونه در یک طبقه از پیش تعیین شده قرار گرفته و با استفاده از یک سری اطلاعات اولیه، نمونه‌ها به دسته‌های خاصی نسبت داده می‌شوند (۹). در این روش‌ها تلاش می‌شود تا داده‌ها به خوشه‌هایی تقسیم شوند که شباهت بین داده‌های درون هر خوشه حداکثر و شباهت بین داده‌های درون خوشه‌های متفاوت حداقل شود (۱۹). گروهی از روش‌های طبقه‌بندی که در سال‌های اخیر مورد توجه پژوهشگران قرار گرفته و در زمینه‌های مختلف از آنها استفاده شده است، روش‌های ریاضی مبتنی بر الگوریتم‌های رایانه‌ای هستند (۲۴). از پرکاربردترین این روش‌ها می‌توان به روش خوشه‌بندی فازی (FCM) و شبکه عصبی مصنوعی (ANN) اشاره کرد (۲۰). در خوشه‌بندی فازی، درجه عضویت یک نمونه در یک خوشه، مقداری بین صفر و یک است. در این شرایط هرچه نمونه به سایر اعضای خوشه شباهت بیشتری داشته باشد، درجه عضویت آن به یک نزدیک‌تر خواهد بود و برعکس (۶). در روش شبکه عصبی مصنوعی با تقلید از شبکه عصبی مغز انسان، الگوهای برای پردازش اطلاعات ساخته شده و با بررسی ساختارهای موجود در این الگوها امکان مدل‌سازی، پیش‌بینی و خوشه‌بندی مجموعه داده‌ها فراهم می‌شود (۵). در زمینه طبقه‌بندی خوشه‌بندی داده‌ها با منطق فازی و شبکه عصبی و همچنین مقایسه کارایی این دو روش، پژوهش‌های مختلفی در زمینه‌های

بارش سالانه ۳۴۲۳ ایستگاه همدید، اقلیم‌شناسی و باران‌سنجی در دوره ۱۹۶۱ تا ۲۰۱۵ (ایستگاه‌هایی با حداقل ۱۵ سال طول دوره آماری) بوده که از سازمان هواشناسی کشور دریافت شده است. گروه دوم داده‌ها، مربوط به مشخصات مکانی است. در این قسمت از داده‌های ارتفاع، شیب، جهت دامنه‌ها و تراکم ایستگاهی برای هر یک از ایستگاه‌ها استفاده شده است. این داده‌ها از نقشه مدل رقومی ارتفاع (DEM) با توان تفکیک ۳۰ متر و در محیط نرم افزار GIS استخراج شده است. در ادامه با توجه به وسیع بودن محدوده مورد مطالعه تلاش شده نوعی پوشش مشترک برای یاخته‌های مکانی مورد استفاده در پژوهش ایجاد شود. بر این اساس و با توجه به تعداد و پراکنش ایستگاه‌های هواشناسی موجود در سطح کشور و محاسبه نسبت این پراکنش با مساحت محدوده مورد مطالعه، اندازه مناسب هر یک از یاخته‌ها جهت استفاده در مراحل بعدی پژوهش، ۸ کیلومتر مربع تعیین و در نهایت ماتریس داده‌های اصلی در سطح کشور با ابعاد ۲۵۳۸۱×۵ از طریق میان‌یابی با روش کریجینگ تشکیل شد (شکل ۲).

در ادامه با هدف یکسان‌سازی بعد متغیرها، فرایند استانداردسازی بر روی ماتریس داده‌ها اعمال شد. رابطه اصلی این فرایند به شرح زیر است (۲):

$$Z_i = \frac{x_i - \bar{x}}{S} \quad (1)$$

در رابطه ۱،  $x_i$  مقدار هریک از مشاهدات،  $\bar{x}$  میانگین کل مشاهدات و  $S$  انحراف معیار مشاهدات است.

در بخش بعد، تعداد مناسب خوشه‌ها مشخص شده است. در این قسمت تعداد بهینه خوشه‌ها به نحوی انتخاب شده که خوشه‌های مشخص شده دارای بیشینه پراش بیرونی و کمینه پراش درونی باشند (۸). بدین منظور از روش دیویس-بولدین استفاده شده است. در این شاخص،  $\delta(X_i, X_j)$  مقدار فاصله درون خوشه‌ای برای خوشه‌های  $X_i, X_j$ ،  $\Delta(X_i)$  قطر خوشه و  $C$  تعداد خوشه‌های قسمت‌های  $U$  است (۹):

$$DB(U) = \frac{1}{C} \sum_{i=1}^c \max \left[ \frac{\Delta(X_i) + \Delta(X_j)}{\delta(X_i, X_j)} \right] \quad i \neq j \quad (2)$$

قطعیت در نتایج پژوهش‌ها مؤثر باشد. بر این اساس، در پژوهش حاضر تلاش شده تا میزان کارایی و دقت دو روش خوشه‌بندی فازی (FCM) و شبکه عصبی مصنوعی (ANN) در پهنه‌بندی بارش سالانه ایران مورد ارزیابی قرار گیرد. با توجه به شرایط مورفولوژیک ایران و نقش شرایط محیطی در تغییرات فراسنج بارش، علاوه بر داده‌های بارش سالانه، از داده‌های ارتفاع، شیب، جهت شیب و تراکم ایستگاهی نیز به‌عنوان داده‌های کمکی استفاده شده و این جامع‌نگری در طبقه‌بندی، نقطه تفاوت این پژوهش با سایر پژوهش‌های مشابه است.

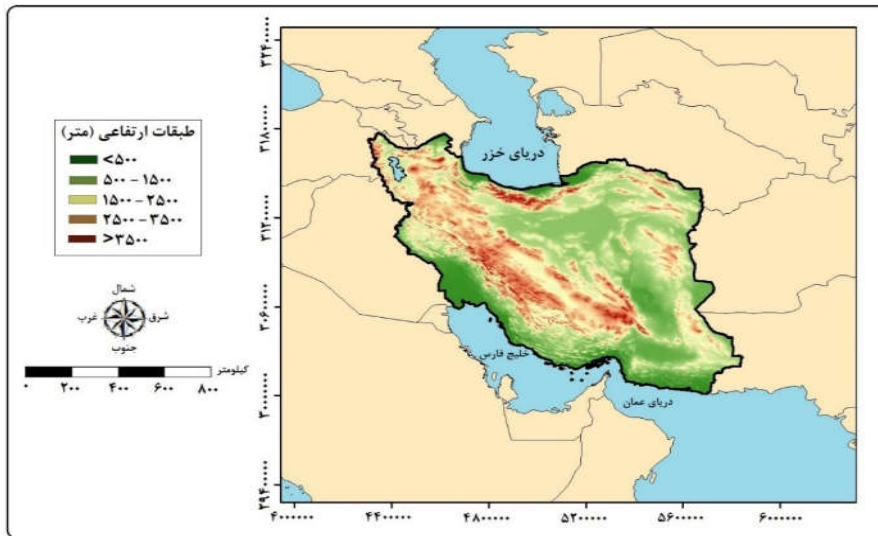
## مواد و روش‌ها

### محدوده مورد مطالعه

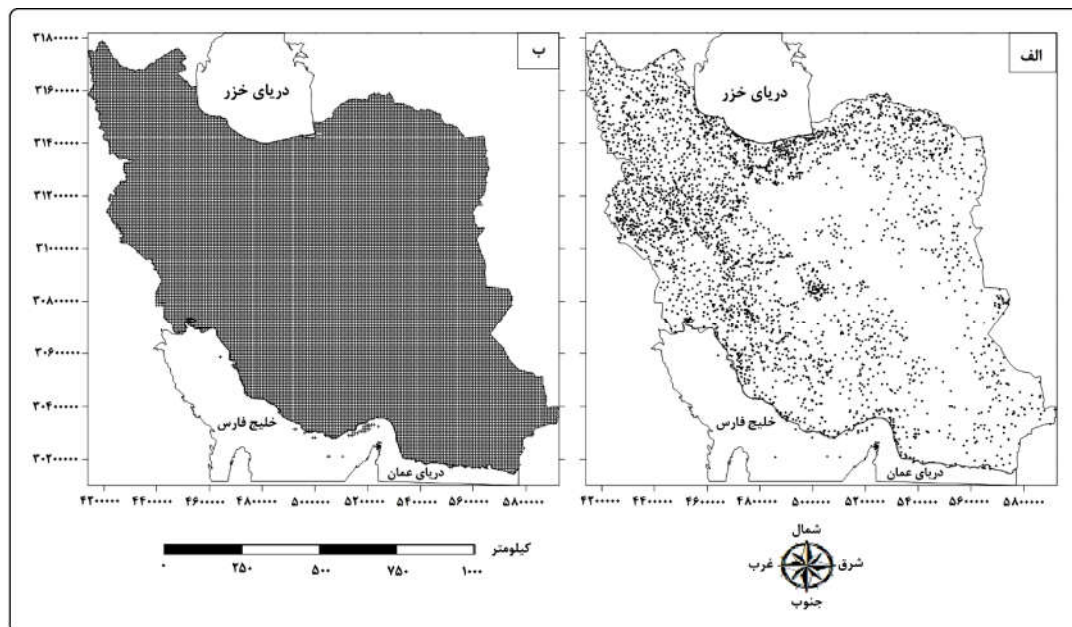
محدوده مورد مطالعه در این پژوهش شامل تمامی پهنه‌ی کشور ایران است (شکل ۱). سرزمین ایران با وسعت ۱۶۴۸۱۹۵ کیلومتر مربع، مابین ۲۵ تا ۴۰ درجه عرض شمالی و ۴۴ تا ۶۳ درجه طول شرقی واقع شده است (۴). ایران سرزمینی بسیار متنوع است و این گوناگونی در تمام ویژگی‌های جغرافیایی آن از مسائل انسانی گرفته تا خصوصیات طبیعی به چشم می‌خورد و شاید بتوان گفت که بهترین جلوه‌گاه این همه تنوع و گوناگونی، آب و هوای کشور است (۱۷). ترکیب عوامل محیطی مختلف از قبیل وسعت، ارتفاع، شکل و جهت ناهمواری‌ها و موقعیت طبیعی سبب شده این سرزمین از نظر آب و هوایی تنوع زیادی داشته باشد و این موضوع از نحوه تغییرات دو عامل عمده اقلیمی یعنی دما و بارش در آن قابل مشاهده بوده (۱۵) که از تغییرپذیری زیادی در سطح کشور برخوردار هستند.

### روش تحقیق

به‌منظور انجام مقایسه دو مدل فازی و عصبی و انجام پهنه‌بندی بر مبنای آنها، در پژوهش حاضر از دو گروه داده اقلیمی و محیطی استفاده شده است. داده‌های اقلیمی شامل داده‌های



شکل ۱. موقعیت محدوده‌ی مورد مطالعه و چگونگی پراکنش ناهمواری‌ها در سطح آن



شکل ۲. الف) پراکنش ایستگاه‌های مورد استفاده در پژوهش ب) پوشش یاخته‌های ۸×۸ کیلومتری در پهنه‌ی ایران

است. الگوریتم فازی C-Means (FCM) برخلاف روش‌های سنتی خوشه‌بندی که هر داده را به صورت قطعی در یک دسته خاص جای می‌دهند، می‌تواند نوعی خوشه‌بندی چندگانه را بر داده‌ها اعمال کند. در این الگوریتم هر یک از داده‌ها دارای یک درجه عضویت از صفر تا صد درصد هستند و خوشه‌بندی در این روش بر اساس بالاترین درجه

این شاخص بر اساس هر تعداد خوشه محاسبه می‌شود و تعداد خوشه متناظر با حداقل مقدار شاخص دیویس-بولدین به عنوان تعداد خوشه بهینه لحاظ می‌شود (۲۰).

با مشخص شدن تعداد بهینه خوشه‌ها، در بخش بعد از طریق هر یک از مدل‌های فازی و شبکه عصبی اقدام به خوشه‌بندی و در نهایت پهنه‌بندی داده‌های بارشی شده

به‌طور معمول از لایه ورودی، لایه‌های میانی یا مخفی و لایه خروجی تشکیل شده است. لایه ورودی یک لایه انتقال‌دهنده و وسیله‌ای برای تهیه کردن داده‌هاست. آخرین لایه یا لایه خروجی شامل مقادیر پیش‌بینی شده به وسیله شبکه است و بنابراین، خروجی مدل را معرفی می‌کند. لایه میانی یا مخفی که از نرون‌های پردازشگر تشکیل شده‌اند، محل پردازش داده‌هاست. شمار لایه‌ها و شمار نرون‌ها در هر لایه مخفی به‌طور معمول به وسیله روش آزمون و خطا مشخص می‌شود (۵). در پژوهش حاضر از شبکه عصبی خودسازمانده (SOM-ANN) استفاده شده است. هدف اصلی این الگوریتم، کاهش ابعاد و خوشه‌بندی مجموعه داده‌ها است. هر شبکه خودسازمانده از تعدادی نرون تشکیل شده است. هر نرون دارای برداری از وزن‌ها است. ابعاد این بردار با ابعاد فضای ورودی برابر است. پس از آموزش شبکه هر ناحیه، نرون‌های شبکه به الگوهای خاصی از داده‌های ورودی واکنش نشان می‌دهند (۹). روش آموزش شبکه، یادگیری رقابتی است. وقتی نمونه آموزشی جدید به شبکه اعمال می‌شود فاصله اقلیدسی آن از بردار وزن تمام نرون‌های شبکه حساب می‌شود. نرونی که بردار وزنش بیشترین شباهت را با بردار ورودی داشته باشد برنده است. این نرون بهترین واحد تطبیق (BMU) نامیده می‌شود. بردار وزن نرون برنده و همسایه‌هایش در شبکه عصبی در راستای نزدیک شدن به بردار ورودی تغییر می‌کنند. میزان تغییرات بر اساس زمان و فاصله از بهترین واحد تطبیق کاهش می‌یابد. وزن‌های هر نرون مطابق رابطه زیر بروزسانی می‌شود:

$$W_i(t+1) = W_i(t) + \theta(v,t)\alpha(t)(X(t) - W_i) \quad (6)$$

که در آن  $\alpha(t)$  نرخ یادگیری کاهشی و  $X(t)$  بردار ورودی است. تابع همسایگی  $\theta(v,t)$  وابسته به فاصله بین نرون  $i$  و بهترین واحد تطبیق در شبکه است. همچنین شعاع همسایگی و نرخ یادگیری در طول زمان توسط روابط زیر کاهش می‌یابد:

$$\alpha(t) = \sigma_0 \exp\left(-\frac{t}{\lambda}\right) \quad t = 1, 2, 3, \dots \quad (7)$$

که در آن  $t$  شماره ی تکرار و  $\lambda$  مقداری ثابت و  $\sigma_0$  نرخ

عضویت هر یک از داده‌ها در هر خوشه صورت می‌پذیرد. هدف نهایی الگوریتم مذکور آن است که خروجی تابع زیر کمینه شود:

$$J(\mu, v) = \sum_{i=1}^n \sum_{j=1}^c (\mu_{ij})^m \|x_i - v_j\|^2 \quad (3)$$

در روابط مربوط به این الگوریتم  $n$  تعداد نقاط داده،  $v_j$  نشان دهنده  $j$  امین مرکز خوشه‌ای،  $m$  بیانگر معیار Fuzziness ( $m \in [1, \infty)$ )،  $C$  تعداد مراکز خوشه‌ای،  $\mu_{ij}$  عضویت  $i$  امین داده در  $j$  امین مرکز خوشه‌ای،  $d_{ij}$  نشان دهنده فاصله اقلیدسی بین  $i$  امین داده و  $j$  امین مرکز خوشه‌ای بوده و  $x = x_1, x_2, x_3, \dots, x_n$  به‌عنوان مجموعه‌ای از داده‌ها و  $v = v_1, v_2, v_3, \dots, v_n$  به‌صورت مجموعه‌ای از مراکز دسته‌ها محسوب می‌شوند.

مراحل پیاده‌سازی این الگوریتم شامل چهار مرحله به شرح زیر است (۲۶):

۱- انتخاب تصادفی مراکز خوشه‌ای

۲- محاسبه توابع عضویت فازی  $\mu_{ij}$  با استفاده از رابطه زیر:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{ik}}\right)^{\frac{2}{m-1}}} \quad (4)$$

۳- محاسبه مراکز فازی  $v_j$  بر اساس رابطه (۵):

$$v_j = \frac{\sum_{i=1}^n (\mu_{ij})^m x_i}{\sum_{i=1}^n (\mu_{ij})^m} \quad \forall j = 1, 2, 3, \dots, c. \quad (5)$$

۴- تکرار مراحل ۲ و ۳ تا زمانی که حداقل مقدار  $J$  به دست آید یا  $\|U^{(k+1)} - U^k\| < \beta$  شود. در این بخش  $k$  تعداد مراحل تکرار،  $J$  تابع هدف و  $\beta$  معیار سنجش خاتمه فرایند است که مقدار آن بین ۰ و ۱ است؛ همچنین  $U$  ماتریس عضویت فازی بوده که از طریق رابطه  $U = (\mu_{i,j})_{n,c}$  به دست می‌آید.

شبکه عصبی مصنوعی (ANN)، مجموعه‌ای از الگوریتم‌های رایانه‌ای هستند که می‌توانند روابط مهم بین تعداد زیادی فراسنج را از یک بانک داده استخراج کرده و به پیش‌بینی و طبقه‌بندی همان بانک داده و یا یک بانک جدید داده به کار ببرند (۱۴). ساختار عادی یک شبکه عصبی مصنوعی،

یادگیری اولیه است.

$$\theta(t) = \exp\left(-\frac{\text{dist}^2}{2\sigma_0^2}\right) \quad (8)$$

که  $\text{dist}$  فاصله‌ی هر گره از BMU و  $\sigma(t)$  با استفاده از رابطه‌ی زیر محاسبه می‌شود:

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\lambda}\right) \quad t = 1, 2, 3, \dots \quad (9)$$

در رابطه‌ی بالا  $\sigma_0$  عرض شبکه در زمان  $t_0$  و  $\lambda$  مقدار ثابت و  $t$  شماره‌ی تکرار جاری است (۲۵).

با توجه به موارد اشاره شده، در این بخش بر اساس تنظیمات هر یک از مدل‌های مذکور داده‌ها مرتب‌سازی شده و با معرفی تعداد بهینه‌ی خوشه‌ها (که توسط متوسط شاخص دیویس-بولدین مشخص شده است)، تقسیم داده‌ها به دو بخش یادگیری (۸۰ درصد) و ارزیابی (۲۰ درصد)، معرفی قوانین فازی حاکم، توابع عضویت مدل فازی، تعداد تکرارها (۱۰۰۰ تکرار)، نرخ یادگیری (۰/۰۱ تا ۰/۹) و شرایط نرونها، داده‌های مربوط به فراسنج‌های مورد مطالعه در محیط نرم‌افزار Matlab 2021 به مدل‌های فازی و شبکه عصبی مصنوعی وارد شد و مورد پردازش قرار گرفت تا در نهایت پس از ارزیابی خروجی مدل‌ها و انجام پهنه‌بندی، نتایج دو مدل مذکور با یکدیگر مورد مقایسه قرار گیرد.

در زمینه‌ی ارزیابی خروجی مدل‌ها از دو روش ضریب تعیین (رابطه‌ی ۱۰) و میانگین خطای مطلق (رابطه‌ی ۱۱) استفاده شده است (۲):

$$R^2 = \left(\frac{\frac{1}{n} \sum_m^n (p_i - \bar{p})(x_i - \bar{x})}{s_p \cdot s_o}\right)^2 \quad (10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x_p| \quad (11)$$

در این روابط،  $x_i$  مقدار مشاهده شده متغیر مورد نظر؛  $p_i$  و  $x_p$  مقادیر شبیه‌سازی شده متغیر مورد نظر،  $\bar{x}$  میانگین مقادیر متغیرها،  $n$  تعداد داده‌ها و  $s_p$  و  $s_o$  انحراف معیار داده‌های مشاهداتی و برآورد شده است. شاخص‌های  $R^2$  و MAE بیانگر دقت و کیفیت طبقه‌بندی در خروجی مدل‌ها بوده و میزان اشرف

مدل به روابط درونی بین فراسنج‌ها را نشان می‌دهند. به صورت کلی میزان شاخص  $R^2$  هرچه به یک و مقادیر شاخص MAE هر چه به صفر نزدیک‌تر باشند نشانه‌ی کارایی مدل‌ها در پیش‌بینی یا طبقه‌بندی بهینه‌ی داده‌های ورودی است.

در بخش نهایی این پژوهش بر اساس داده‌های طبقه‌بندی‌شده هر یک از مدل‌ها، نقشه‌های پهنه‌بندی فضایی بارش در سطح ایران ترسیم و نتایج شاخص‌های ارزیابی مدل‌ها (مربوط به داده‌های یادگیری هر یک از مدل‌ها) در قالب جدولی نمایش داده شده و تلاش شده است ضمن ارزیابی عددی نتایج، خروجی مدل‌ها از لحاظ پهنه‌بندی شرایط بارشی کشور و ارتباط آن با متغیرهای محیطی مورد تحلیل قرار گیرد.

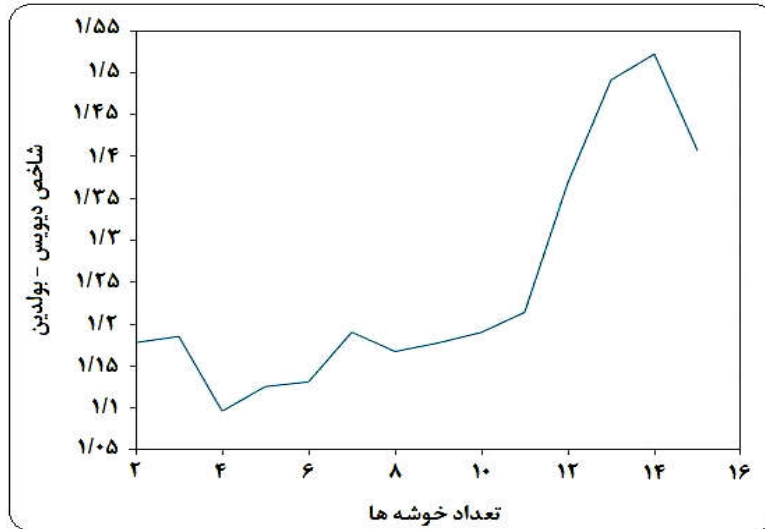
## نتایج و بحث

### تعیین تعداد بهینه‌ی خوشه‌ها

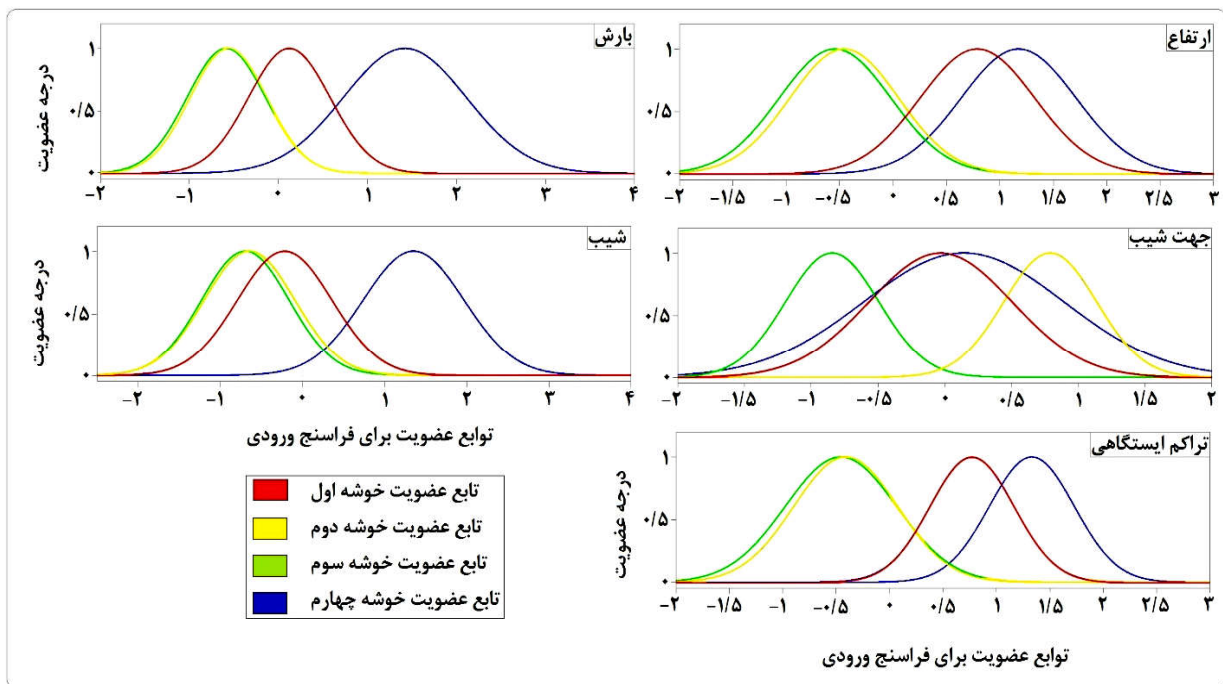
در این قسمت با هدف تشخیص تعداد بهینه‌ی خوشه‌ها، شاخص دیویس-بولدین بر داده‌های مورد استفاده اعمال شده و نتایج آن به تفکیک تعداد خوشه‌های مختلف، در شکل ۳ آمده است. در این شکل مقادیر شاخص دیویس-بولدین برای خوشه‌های ۲ تا ۱۶ محاسبه و نمایش داده شده است. همان‌طور که از تغییرات نمودار مشخص است، کمترین مقدار شاخص ۱/۱ بوده که نشان می‌دهد تعداد بهینه‌ی خوشه‌ها برای معرفی به مدل‌های فازی و شبکه عصبی چهار خوشه است. بر این اساس می‌توان استنباط کرد که طبقه‌بندی داده‌ها به چهار خوشه موجب افزایش دقت خروجی مدل‌ها خواهد شد.

### ارزیابی و تحلیل عددی خروجی مدل FCM

نتایج عددی حاصل از طبقه‌بندی داده‌های پنج گانه ورودی به مدل فازی (FCM) در قالب شکل ۴ و جدول ۱ آمده است. در پژوهش حاضر بر اساس نوع داده‌های محیطی-بارشی مورد استفاده و مقایسه چندین نوع تابع فازی، بر اساس روش آزمون و خطا، تابع گوسی به‌عنوان بهینه‌ترین تابع عضویت در



شکل ۳. مقادیر شاخص دیویس - بولدین برای خوشه‌های مختلف



شکل ۴. توابع عضویت در مدل FCM برای هر یک از پارامترهای ورودی (رنگی در نسخه الکترونیکی)

درجه عضویت داده‌های هر یک از فراسنج‌ها در هر خوشه را نشان می‌دهد. با بررسی این نمودارها می‌توان به تصویری کلی از چگونگی پراکنش داده‌ها و میزان تعلق‌پذیری آنها به هر یک از خوشه‌های چهارگانه طبقه‌بندی‌شده بر اساس منطق فازی دست یافت. در این راستا اگر درجه عضویت یک عنصر از مجموعه برابر با صفر باشد، آن عضو کاملاً از مجموعه

مدل فازی انتخاب شده است. نمودارهای توابع عضویت این تابع بر اساس هر یک از قوانین فازی مورد استفاده در مدل و برای هر یک از فراسنج‌های مورد استفاده در شکل بالا نمایش داده شده است. در این نمودارها، محور افقی مقادیر استانداردشده مربوط به هر یک از متغیرهای ورودی به مدل و محور عمودی نمودار

جدول ۱. مقادیر عددی درجه عضویت یاخته ها در خوشه های مختلف

شماره یاخته	خوشه اول	خوشه دوم	خوشه سوم	خوشه چهارم	واریانس	خوشه اصلی
۱	۰/۰۶۹	۰/۳۵۱۴	۰/۴۶۷۷	۰/۱۱۱۷	۰/۰۳۶	۳
۲	۰/۰۶۹۷	۰/۴۱۳۱	۰/۱۱۱۶	۰/۴۰۵۵	۰/۰۳۴	۲
۳	۰/۰۸۱۱	۰/۲۵۴۵	۰/۵۳۸۵	۰/۱۲۵۸	۰/۰۴۲	۳
۴	۰/۰۶۹۱	۰/۳۵۳۱	۰/۴۶۷۳	۰/۱۱۰۳	۰/۰۳۶	۳
۵	۰/۰۷۳۳	۰/۳۶۰۱	۰/۴۵۲۶	۰/۱۱۳۹	۰/۰۳۴	۳
⋮	⋮	⋮	⋮	⋮	⋮	⋮
۲۵۳۷۷	۰/۰۳۷۳	۰/۵۰۰۲	۰/۰۹۷۷	۰/۳۶۴۶	۰/۰۴۸	۲
۲۵۳۷۸	۰/۱۲۴۸	۰/۴۵۴۹	۰/۲۰۱۴	۰/۲۱۸۷	۰/۰۲	۲
۲۵۳۷۹	۰/۰۴۱۴	۰/۱۲۳۹	۰/۷۳۷۶	۰/۰۹۶۹	۰/۱۰۶	۳
۲۵۳۸۰	۰/۰۵۸۵	۰/۱۳۳	۰/۶۹۹۵	۰/۱۰۸۸	۰/۰۹	۳
۲۵۳۸۱	۰/۰۶۰۸	۰/۱۴۵۹	۰/۶۸۵۴	۰/۱۰۷۷	۰/۰۸۵	۳

گرفته‌اند. این میزان برای خوشه های ۲، ۳ و ۴ به ترتیب ۲۳۰۱ (بازه ۰/۵- تا ۰/۶-)، ۲۱۵۰ (بازه ۰/۶- تا ۰/۷-) و ۶۵۹ (بازه ۰/۷- تا ۰/۸-) است.

نکته مهم در بررسی این نمودارها و تفسیر چگونگی کارایی مدل در پردازش داده‌ها، توجه به فشردگی داده‌ها در محدوده خوشه‌های ۲ و ۳ بوده که موجب شده در بازه‌های عددی کوچک (با اختلاف ۰/۱) تعداد زیادی از داده‌ها از نظر مدل و بر اساس منطق فازی در یک ناحیه قرار بگیرند. این شرایط به صورت تقریباً مشابه (با اندکی اختلاف) برای دیگر داده‌های ورودی به مدل نیز تکرار شده است. سایر داده‌های مربوط به هر فراسنج که وابستگی کامل و صد درصدی به یک خوشه‌ی خاص نداشته‌اند، با مقادیر توابع عضویت در بازه کمتر از ۱ تا صفر، بر اساس میزان شباهت خود به یکی از خوشه‌ها، طبقه‌بندی شده و در خوشه‌های ۱ تا ۴ قرار گرفته‌اند.

یکی از ویژگی‌های روش‌های خوشه‌بندی مبتنی بر منطق فازی آن است که هر یک از داده‌ها دارای مقادیری از عضویت جزئی در سایر خوشه‌ها هستند و این امر در جدول ۱ به خوبی مشخص شده است. این جدول مقادیر عددی درجه عضویت برخی از یاخته‌ها در خوشه‌های مختلف به همراه

خارج است و در صورتی که درجه عضویت یک عضو برابر با یک شود، آن عضو کاملاً در مجموعه قرار دارد. حال اگر درجه عضویت یک عضو در بازه مابین صفر و یک قرار بگیرد، این عدد بیانگر درجه عضویت تدریجی خواهد بود.

بر اساس توضیحات بیان شده، در این قسمت نمودارهای توابع عضویت داده‌های ورودی به مدل شامل بارش، ارتفاع، شیب، جهت شیب و تراکم ایستگاهی درج شده‌اند. با توجه به نمودارها مشخص است که برای هر فراسنج چهار تابع عضویت بر اساس قوانین فازی درونی مدل و همچنین تعداد خوشه‌ها، تشکیل شده است. در نمودارهای مربوط به این توابع، برای هر یک از خوشه‌ها نقاط اوجی وجود دارند که نشانه عضویت کامل اعضا در آن خوشه خاص هستند.

همچنین، بین نمودارهای توابع عضویت، مناطق همپوشانی نیز مشاهده می‌شوند که نشانه عدم قطعیت در اختصاص داده‌ها به خوشه خاصی هستند. در این شرایط، اندازه‌های مربوط به توابع عضویت، مشخص‌کننده تخصیص یا عدم تخصیص هر یک از یاخته‌ها به خوشه‌های مختلف هستند؛ به‌عنوان مثال بر اساس نمودار مربوط به فراسنج بارش، ۱۶۴۵ عدد از یاخته‌ها (بازه ۰/۰۵- تا ۰/۲) به صورت قطعی در خوشه ۱ قرار



خوشه‌ها و انحراف معیار کلی خوشه‌ها در نرخ‌های یادگیری مختلف قابل مشاهده است. بر اساس اطلاعات این جدول مشخص می‌شود که انحراف معیار خوشه‌ها در هر یک از نرخ‌های یادگیری، فاقد روند مشخص و یا اختلافات معنی‌دار است. این شرایط نشان‌دهنده آن است که مدل در شرایط مختلف یادگیری، توانایی مناسبی در درک شرایط محیطی و یافتن ارتباط آن با تغییرات بارش در سطح کشور داشته است. با این حال به منظور انتخاب بهینه‌ترین خروجی‌های حاصل از مدل، پایین‌ترین میزان انحراف معیار به عنوان مبنای انتخاب نرخ یادگیری بهینه انتخاب شد. بر این اساس، نرخ یادگیری ۰/۴، دارای انحراف معیار ۳۶۰۶ بوده و از نتایج خروجی مدل در این نرخ به عنوان بهینه‌ترین خروجی در بخش بعدی این پژوهش استفاده شده است.

#### مقایسه و تحلیل خروجی مدل‌ها

با هدف مقایسه خروجی مدل‌های فازی و شبکه عصبی، در این بخش اطلاعات مربوط به خوشه‌های چهارگانه خروجی مدل‌ها در قالب جدول ۳ و شکل ۶ نمایش داده شده و مورد تحلیل قرار گرفته است؛ همچنین در بخش پایانی پژوهش، اطلاعات آماری کارایی مدل‌ها در قالب دو روش  $R^2$  و MAE محاسبه و در جدول ۴ درج شده است.

بر اساس داده‌های جدول ۳ و شکل ۶ مشخص می‌شود که خروجی مدل فازی (FCM)، ۵۵۷۷ یاخته (۲۲ درصد کل یاخته‌ها) را در خوشه اول قرار داده است. این مقدار برای مدل شبکه عصبی (SOM-ANN) ۴۸۶۳ یاخته (۱۹/۲ درصد کل یاخته‌ها) است. در این خوشه درصد یاخته‌های مشترک ۸۷/۲ درصد است. با توجه به نقشه‌های شکل ۶، یاخته‌های خوشه اول از لحاظ پراکنش جغرافیایی مناطقی از شمال و غرب کشور در امتداد رشته کوه‌های البرز و زاگرس، همچنین بخش‌هایی از شمال غرب، سواحل شمالی و شمال شرق کشور را دربرگرفته‌اند. این نواحی منطبق ارتفاعات بلند و مناطق پربارش کشور بوده و متوسط بارش این مناطق ۳۱۸ میلی‌متر و متوسط ارتفاع ۱۶۷۷ متر

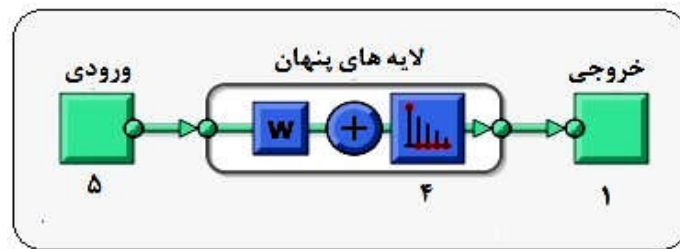
مقادیر واریانس و انحراف معیار مربوط به آنها را نمایش می‌دهد؛ به عنوان مثال در یاخته شماره ۱، بالاترین درجه عضویت مربوط به خوشه سوم است که نشان می‌دهد بیشترین میزان اثرگذاری عوامل محیطی مانند ارتفاع و شیب بر فراسنج بارش، در این خوشه وجود داشته و با کاهش اثرگذاری پارامترهای ورودی، مقادیر درجات عضویت در سایر خوشه‌ها کاهش یافته است. با این حال، درجات عضویت یاخته مورد اشاره در برخی خوشه‌ها (مانند خوشه دوم) همچنان قابل توجه است که دلیل آن نسبی بودن طبقه بندی‌ها در مدل‌های فازی و وجود تأثیر متغیرهای محیطی بر فراسنج بارش (هرچند به میزان کمتر) است.

#### ارزیابی و تحلیل عددی خروجی مدل SOM-ANN

همانطور که پیشتر اشاره شد، در پژوهش حاضر از یکی از انواع مدل شبکه عصبی به نام شبکه عصبی خودسازمانده (SOM-ANN) استفاده شده است. در این مدل داده‌های ورودی شامل بارش سالانه، ارتفاع، شیب، جهت ناهمواری‌ها و تراکم ایستگاهی بوده و خروجی مدل، شامل داده‌های خوشه‌بندی شده است. تعداد بهینه خوشه‌ها نیز قبلاً توسط شاخص دیویس - بولدین مشخص و به مدل اعلام شده است. ساختار شبکه متشکل از پنج لایه ورودی، چهار لایه میانی و یک لایه خروجی است (شکل ۵).

پس از عرضه نمونه‌های آموزشی به شبکه، وزن‌های شبکه (پارامترهای آزاد) به گونه‌ای اصلاح می‌شوند که تفاوت پاسخ مطلوب و خروجی واقعی شبکه کمینه شود. این فرایند با نمونه‌های ورودی تکرار می‌شود تا اینکه شبکه به حالت پایدار برسد و تغییر قابل توجهی در وزن‌ها اتفاق نیفتد. مدل بهینه زمانی حاصل می‌شود که شبکه‌ی عصبی بهترین پردازش را با کمترین خطا انجام دهد (۲۴). بر این اساس مدل مذکور در نرخ‌های یادگیری مختلف و با ۱۰۰۰ تکرار اجرا شده و نتایج آن در جدول ۲ درج شده است.

در این جدول اطلاعاتی مانند تعداد یاخته‌ها برای هر یک از



شکل ۵. ساختار مدل شبکه عصبی خودسازمانده استفاده شده در پژوهش (رنگی در نسخه الکترونیکی)

جدول ۲. نتایج عددی خوشه بندی مدل SOM-ANN

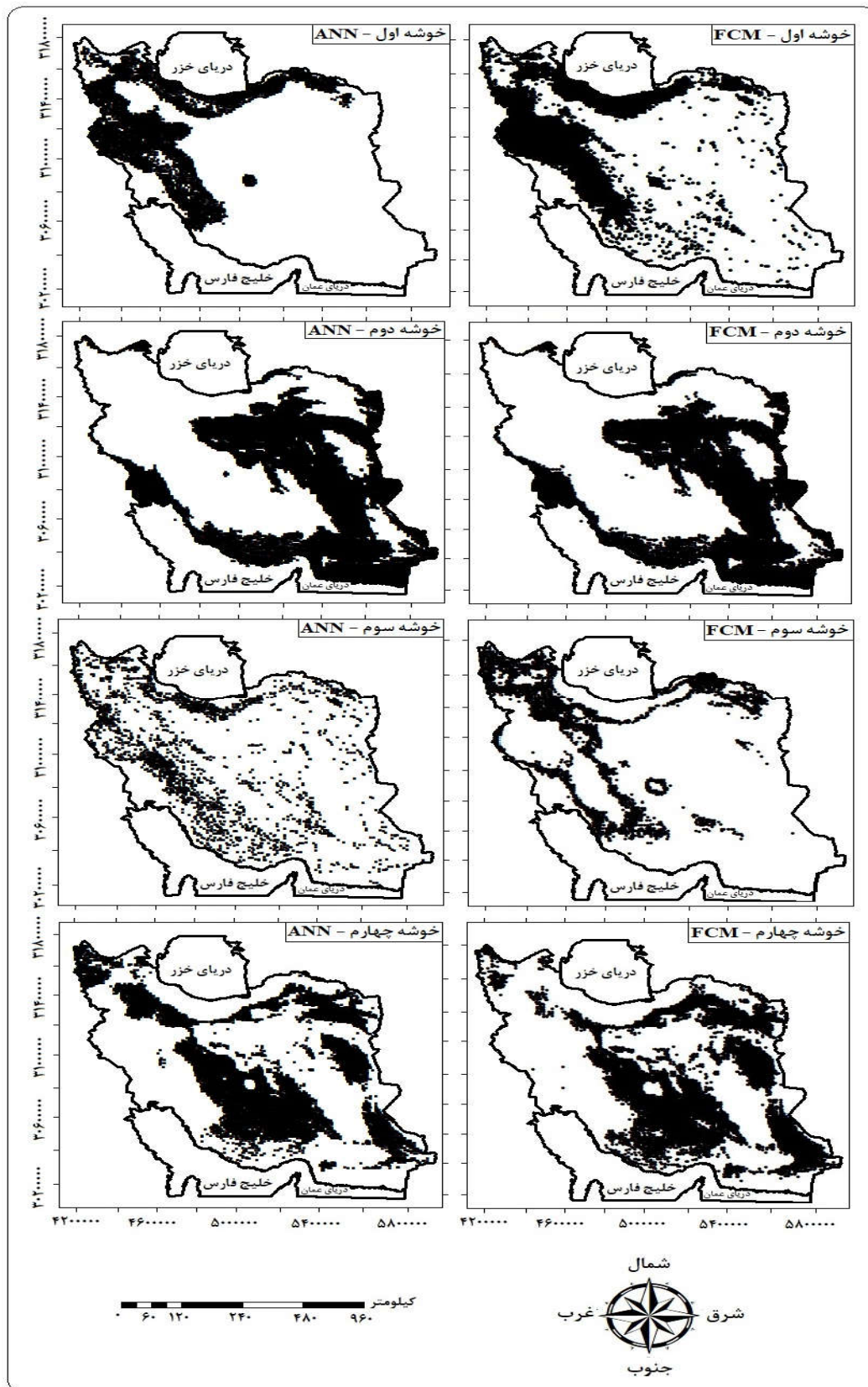
ردیف	نرخ یادگیری	انحراف معیار	خوشه اول	خوشه دوم	خوشه سوم	خوشه چهارم	کل یاخته ها
۱	۰/۰۱	۳۷۲۱	۴۸۴۶	۱۰۸۰۰	۲۱۲۱	۷۶۱۴	۲۵۳۸۱
۲	۰/۱	۳۶۲۷	۴۸۱۹	۱۰۷۵۲	۲۳۰۰	۷۵۱۰	۲۵۳۸۱
۳	۰/۲	۳۶۳۰	۴۸۳۰	۱۰۷۷۱	۲۲۹۹	۷۴۸۱	۲۵۳۸۱
۴	۰/۳	۳۶۴۶	۴۷۹۸	۱۰۷۸۲	۲۲۹۰	۷۵۱۱	۲۵۳۸۱
۵	۰/۴	۳۶۰۶	۴۸۶۳	۱۰۷۱۴	۲۲۹۷	۷۵۰۷	۲۵۳۸۱
۶	۰/۵	۳۷۲۹	۴۸۴۲	۱۰۸۳۰	۲۱۲۵	۷۵۸۴	۲۵۳۸۱
۷	۰/۶	۳۶۱۶	۴۸۴۸	۱۰۷۳۰	۲۲۹۴	۷۵۰۹	۲۵۳۸۱
۸	۰/۷	۳۶۳۴	۴۸۱۵	۱۰۷۶۳	۲۲۹۴	۷۵۰۹	۲۵۳۸۱
۹	۰/۸	۳۶۶۳	۴۷۶۹	۱۰۸۱۵	۲۲۹۱	۷۵۰۶	۲۵۳۸۱
۱۰	۰/۹	۳۶۵۶	۴۸۰۲	۱۰۷۷۴	۲۲۶۳	۷۵۴۲	۲۵۳۸۱

جدول ۳. تعداد یاخته‌ها در خوشه‌های مختلف

مدل - خوشه	خوشه اول	خوشه دوم	خوشه سوم	خوشه چهارم
FCM	۵۵۷۷	۹۳۷۱	۳۱۳۴	۷۲۹۹
SOM-ANN	۴۸۶۳	۱۰۷۱۴	۲۲۹۷	۷۵۰۷

شبکه عصبی نسبت به یاخته‌های مدل فازی است. همانطور که در نقشه‌ها مشخص است، با وجود اینکه بخش اعظم یاخته‌های مدل فازی در نواحی ذکر شده تمرکز یافته‌اند، اما برخی یاخته‌ها که تعداد به نسبت قابل توجهی دارند در سایر نواحی مرکزی، شرقی و جنوبی کشور نیز پراکنده شده‌اند که این شرایط در مدل شبکه عصبی به‌طور محسوسی متفاوت بوده و هیچ یاخته پراکنده‌ای در خروجی این مدل وجود

است. همچنین این نواحی، در مسیر سیستم‌های بارشی عبوری از غرب و شمال غرب ایران و همچنین سامانه‌های نفوذی از مناطق شمالی به سواحل دریای خزر قرار دارند که عبور این سامانه‌ها و برخورد آنها به ارتفاعات (۱۸) در کنار نقش مثبت دریای خزر موجب افزایش بارش در سرتاسر این نواحی و قرارگیری آنها در یک گروه شده است. نکته مهم در بررسی نقشه‌های این خوشه، تمرکز بیشتر یاخته‌های خروجی مدل



شکل ۶. پهنه‌بندی بارشی - محیطی ایران با روش‌های ANN و FCM

جدول ۴. مقایسه ی شاخص های ارزیابی مدل

مدل	R <sup>2</sup>	MAE
FCM	۰/۸۶	۰/۱۴
SOM-ANN	۰/۷۸	۰/۲

یکنواخت تر این نواحی در نظر گرفت.

تعداد یاخته های خوشه سوم در مدل FCM ۳۱۳۴ یاخته (۱۲/۳ درصد کل یاخته ها) و در مدل SOM-ANN ۲۲۹۷ یاخته (۹ درصد کل یاخته ها) است که از نظر تعداد و وسعت جغرافیایی، کمترین میزان در بین سایر خوشه ها است. تعداد یاخته های مشترک بین دو مدل ۵۳/۱ درصد بوده و از نظر پراکنش نیز با یکدیگر متفاوت و ناهمگون هستند. مدل FCM یاخته های مربوط به دامنه ها و مناطق کوهپایه ای را در این خوشه طبقه بندی کرده است. از سوی دیگر یاخته های طبقه بندی شده توسط مدل SOM-ANN با وجود تمرکز نسبی در بلندی های غرب و شمال کشور، فاقد نظم خاصی بوده و در بیشتر مناطق کشور پراکنده شده اند. این مورد می تواند نشان دهنده افزایش خطای داخلی و ناتوانی مدل شبکه عصبی برای تشخیص روابط بین ناهمواری های دامنه ای با تغییرات فراسنج بارش باشد.

نواحی طبقه بندی شده در خوشه سوم در دامنه های غربی و شرقی رشته کوه زاگرس، دامنه های شمالی و جنوبی رشته کوه البرز و مناطق شمال غربی و بخش هایی از شمال شرق کشور قرار گرفته اند. ویژگی مهم آب و هوایی این مناطق که در ارتباط مستقیم با توپوگرافی سطح زمین بوده و در تغییرات میزان بارش نقش چشمگیری دارد، پدیده همرفت دامنه ای است. همرفت دامنه ای فرایندی است که طی آن هوای مناطق شیب دار دامنه کوه ها، در حضور رطوبت کافی هوا و ریزموج های تراز میانی جو، با دریافت تابش خورشید نسبت به اطراف خود گرم تر شده و شروع به صعود می نماید (۱). وقوع این پدیده موجب وقوع بارش های رگباری در این نواحی شده و موجب شده تا میانگین میزان بارش دریافتی در این نواحی ۱۸۱

ندارد. به نظر می رسد این شرایط به دلیل حساسیت بیشتر مدل FCM در تشخیص ارتفاعات منفرد نسبت به مدل شبکه عصبی است که موجب شده این مدل نواحی ارتفاعی کوچک و نقش آنها در اقلیم بارشی نواحی اطراف خود را نیز در محاسبات و طبقه بندی های خود لحاظ کند. مورد دیگر عدم قرارگیری بخش های بسیار مرتفع رشته کوه زاگرس و بخش غربی رشته کوه البرز در طبقه بندی مدل شبکه عصبی برای خوشه اول است که می تواند به دلیل خطای مدل ایجاد شده باشد.

خروجی مدل ها برای خوشه دوم شامل ۹۳۷۱ یاخته (۳۶/۹ درصد کل یاخته ها) برای مدل FCM و ۱۰۷۱۴ یاخته (۴۲/۲ درصد کل یاخته ها) برای مدل SOM-ANN است و ۸۷/۵ درصد یاخته ها بین این دو مدل مشترک است. این خوشه بزرگ ترین پوشش جغرافیایی را در بین سایر خوشه ها به خود اختصاص داده و شامل بخش های جلگه ای جنوب غرب (استان خوزستان) و سرتاسر نوار ساحلی جنوب کشور، نواحی کم ارتفاع مرکزی در استان های سمنان، یزد و کرمان) و همچنین مناطق کوچک تری از نواحی شرقی، شمال شرقی و شمال غرب کشور را در بر گرفته است. ویژگی مشترک این مناطق، عدم وجود ناهمواری های مرتفع، ارتفاع کم در مناطق جلگه ای و ساحلی و فاصله داشتن از مسیر معمول سامانه های عبوری است که موجب شده تنها در دوره سرد سال یا از طریق سامانه های موسمی فصل گرم (در مناطق جنوب شرق و جنوب) بارش دریافت کنند (۲۲). متوسط میزان بارندگی در این نواحی ۱۰۱ میلی متر و متوسط ارتفاع آن ۹۴۵ متر است. خروجی دو مدل مورد بررسی برای این خوشه تا حدود زیادی هماهنگ بوده و غیر از تفاوت اندک در نوار جنوبی کشور، اختلاف چندانی مشاهده نشده که دلیل آن را می توان شرایط اقلیمی - محیطی

متغیره آب و هوایی است. با اینکه مدل FCM به میزان اندکی از کارایی بهتری در طبقه‌بندی برخوردار بوده، می‌تواند به‌عنوان انتخاب اول در مباحث مربوط به طبقه‌بندی مورد استفاده قرار گیرد؛ با این حال مدل شبکه عصبی نیز دارای دقت قابل قبولی است؛ ضمن اینکه با بهینه‌سازی‌های بیشتر در ساختار داخلی هر دو مدل فازی و شبکه عصبی، می‌توان دقت خروجی‌های آنها را بهبود بخشید.

### نتیجه‌گیری

در این مطالعه کارایی دو مدل فازی (FCM) و شبکه عصبی خودسازمانده (SOM-ANN) در زمینه پهنه‌بندی داده‌های بارش در ارتباط با عوامل محیطی مورد مقایسه و ارزیابی قرار گرفت. بدین منظور داده‌های شبکه‌بندی شده پنج فراسنج بارش سالانه، ارتفاع، شیب، جهت دامنه‌ها و تراکم ایستگاهی به مدل‌ها وارد و با تعیین تعداد مناسب خوشه‌ها، فرایند پهنه‌بندی انجام شد و نتایج مورد ارزیابی قرار گرفت. نتایج حاصل از خروجی مدل‌ها نشان داد هر دو مدل، از دقت خوبی برای طبقه‌بندی داده‌های آب و هوایی برخوردار هستند؛ هرچند کارایی و دقت مدل فازی به میزان اندکی بهتر از مدل شبکه عصبی بوده است. بر این اساس، مقادیر ضریب تعیین در مدل فازی (FCM)  $0/86$  و در مدل شبکه عصبی (SOM-ANN)  $0/78$  به دست آمده است. در بررسی نقشه‌های مربوط به پهنه‌های چهارگانه حاصل از خروجی مدل‌ها مشخص شد در خوشه‌های اول، دوم و چهارم بین دو مدل هماهنگی بالایی وجود دارد. در خوشه اول، نقش مثبت ارتفاعات و موقعیت جغرافیایی مناسب در مناطق غربی و شمالی کشور در برابر سامانه‌های بارش‌زای عبوری عامل افزایش مقادیر بارش در این نواحی بوده است. کاهش سطوح ارتفاعی و قرارگیری در پشت دیواره‌های کوهستانی از جمله عواملی بوده که موجب شده است تا مناطق مرکزی، نوار جنوبی و مناطق جنوب غربی کشور در خوشه دوم قرار گیرند. در خوشه چهارم، ناهمواری‌های کوچک‌تر و منفرد و

میلی‌متر باشد که نسبت به مناطق کم ارتفاع داخل فلات ایران بیشتر است.

چهارمین خوشه طبقه‌بندی شده حاوی  $7299$  یاخته  $(28/7)$  درصد کل یاخته‌ها) در مدل FCM و  $7507$  یاخته  $(29/6)$  درصد کل یاخته‌ها) در مدل SOM-ANN است. بالاترین میزان اشتراک یاخته‌ها به مقدار  $92/7$  درصد در این خوشه وجود دارد. نواحی تحت پوشش این خوشه در مناطق مرکزی، بخش‌هایی از دامنه‌های جنوبی البرز و قسمت‌هایی از شرق کشور، بین دو مدل مورد استفاده مشترک بوده است. در بخش‌های شمال غربی کشور میزان تفاوت بین دو مدل افزایش یافته و مدل SOM-ANN وزن بیشتری برای تمرکز یاخته‌ها در این نواحی قائل شده است. از لحاظ شرایط اقلیمی، ناهمواری‌ها در این نواحی دارای ارتفاعی کمتر از ناهمواری‌های رشته کوه‌های البرز و زاگرس بوده و از نظر موقعیت جغرافیایی در نواحی داخلی فلات ایران قرار گرفته‌اند. این شرایط ارتفاعی و موقعیت جغرافیایی آنها موجب شده است تا سامانه‌های بارشی در شرایط ضعیف‌تر و همراه با رطوبت کمتر از آنها عبور کند  $(13)$ ، در نتیجه میزان بارش‌های دریافتی کاهش یافته و میانگین مقادیر بارش سالانه در این نواحی  $102$  میلی‌متر است که مقداری تقریباً مشابه با بارش‌های دریافتی مناطق خوشه دوم دارد.

در قسمت پایانی این پژوهش نتایج مربوط به شاخص‌های ارزیابی کارایی مدل‌های FCM و SOM-ANN در جدول ۴ آمده است:

با توجه به نتایج درج شده در جدول بالا مشخص می‌شود که مقادیر پارامترهای سنجش کارایی هر دو مدل در انجام پهنه‌بندی به یکدیگر نزدیک بوده و در سطح بالایی قرار دارد. میزان شاخص  $R^2$  در مدل فازی (FCM)  $0/86$  بوده که در مقایسه با مقدار این شاخص در مدل شبکه عصبی  $(0/78)$  اندکی بالاتر است. همچنین مقدار شاخص میانگین مطلق خطا (MAE) برای مدل فازی  $0/14$  و در مدل شبکه عصبی  $0/2$  محاسبه شده است. مقادیر شاخص‌های ارزیابی، حاکی از دقت بالا و کارایی بهینه هر دو مدل در انجام پهنه‌بندی‌های چند

پژوهش در مقایسه میزان دقت و کارایی روش‌های نوین طبقه‌بندی با نتایج تحقیقات قربانی و همکاران (۹)، مکاریان و همکاران (۱۹)، گلیان و همکاران (۱۰) و گومز و همکاران (۱۱) مطابقت داشته و توانایی مناسب این مدل‌ها در طبقه‌بندی انواع پدیده‌های پیچیده‌ی محیطی را تأیید می‌کند. پدیده‌های پیچیده‌ای که شناخت روابط و چگونگی تغییرات آنها می‌تواند موجب بهینه‌سازی روش‌های مدیریت و برنامه‌ریزی محیطی و در نهایت کاهش هزینه‌ها و افزایش بهره‌وری طرح‌ها خواهد شد.

تشکیل اقلیم‌های کوچک در اطراف آنها موجب قرارگیری برخی نواحی مرکزی و شرقی کشور در این خوشه شده است. در خوشه سوم نتایج خروجی مدل‌ها متفاوت بوده است. خروجی مدل شبکه عصبی در این خوشه پراکنده و فاقد نظم مشخصی بوده است. مدل فازی نواحی دامنه‌ای اطراف رشته کوه‌های البرز و زاگرس و مناطق شمال غربی کشور را در این خوشه قرار داده که علت آن نقش ارتفاعات دامنه‌ای در وقوع پدیده همرفت دامنه‌ای در این نواحی و نقش آن در تغییرات مقادیر بارش بوده است. نتایج این

### منابع مورد استفاده

1. Alijani, B. 2016. Climate of Iran. Payam-e-Noor University Press, Tehran, Iran (In Farsi).
2. Asakereh, H. 2011. Fundamentals of Statistical Climatology. Zanjan University Press, Zanjan, Iran (In Farsi).
3. Asakereh, H. and H. Shadman. 2015. Identification of spatial relationships of comprehensive warm Days in Iran. *Geographical Researches* 30(1): 53-70 (In Farsi).
4. Asakereh, H. and A. Shahbaee Kotenaee. 2018. Synoptic analysis of the weather pattern of the most pervasive cold day in Iran during the years 1339 to 1388. *Geography and Planning* 22(64): 211-228 (In Farsi).
5. Azizi, H. R. and M. Montazeri. 2015. Anticipated monthly temperatures for selected stations in Isfahan province using artificial neural network multi-layer perceptron. *Geographical Researches* 30(3): 241-258 (In Farsi).
6. Chaghari, A. and M. Feizi-Derakhshi. 2017. Automatic clustering using improved imperialist competitive algorithm. *Journal of Signal and Data Processing* 14(2): 159-169 (In Farsi).
7. Chittineni, S. and R. B. Bhogapathi. 2015. Neural network based fuzzy C-Means clustering algorithm. *International Journal of Education and Social Sciences* 1(1): 100-104.
8. Farshadfar, E. 2010. Multivariate Principles and Procedures of Statistics. Razi University Press, Kermanshah, Iran (In Farsi).
9. Ghoorbani, M. A., H. Jabari, S. Poorbabak, E. Asadi and M. Fazelifard. 2016. Classification of annual temperature of Iran meteorological stations using fuzzy cluster analysis and Kohonen artificial neural networks. *Journal of Geography and Planning* 20(55): 45-63 (In Farsi).
10. Golian, S., B. Saghafian, S. Sheshangosht and H. Ghalkani. 2010. Comparison of classification and clustering methods in spatial rainfall pattern recognition at Northern Iran. *Theoretical and Applied Climatology* 102(1): 319-329.
11. Gomes, E. P., C. J. C. Blanco and F. C. L. Pessoa. 2019. Identification of homogeneous precipitation regions via fuzzy C-means in the hydrographic region of Tocantins-Araguaia of Brazilian Amazonia. *Applied Water Science* 9(6): 1-12.
12. Haibo, C., W. Jiahua and W. Wenyan. 2020. Streamflow prediction using LASSO-FCM-DBN approach based on hydro-meteorological condition classification. *Journal of Hydrology* 580(1): 124-130.
13. Halabian, A. H., F. Hosseinalipour and F. Jafari Shendi. 2012. The analysis of spatial-temporal distribution of daily precipitation on the central and south western Iran. *Geography and Environmental Sustainability* 1(2): 31-41 (In Farsi).
14. Hosseini, A., A. Deljavi and M. M. Sadeghi. 2016. Evaluation of different methods of landslide risk zonation in forest ecosystems. *Journal of Extension and Development of Watershed Management* 4(13): 7-14 (In Farsi).
15. Jedari Eyvazi, J. 2008. Geomorphology of Iran. Payam-e-Noor University Press, Tehran, Iran (In Farsi).
16. Ley, R., M. C. Casper, H. Hellebrand and R. Merz. 2011. Catchment classification by runoff behavior with self-organizing maps (SOM). *Hydrology and Earth Systems Science* 15(1): 2947-2962.
17. Massodian, S. A. and M. Kaviani. 2008. Climatology of Iran. Isfahan University Press, Isfahan, Iran (In Farsi).
18. Massodian, S. A., M. Darand and S. Karsaz. 2011. Precipitation regions of west and northwestern of Iran using cluster analysis. *Journal of Physical Geography*. 4(11): 35-44 (In Farsi).
19. Mokarian, R., H. Sedghi, S. Nemati and H. Babazadeh. 2018. Classification of evaporation stations using fuzzy

- cluster analysis and Kohonen artificial neural networks. *Geography and Planning* 22(63): 283-304 (In Farsi).
20. Raju, K. S. and D. N. Kumar. 2007. Classification of Indian meteorological stations using cluster and fuzzy cluster analysis and kohonen artificial neural networks. *Nordic Hydrology* 38(3): 303-314.
21. Sahina, S. and H. Kerem Cigizoglu. 2012. The sub-climate regions and the sub-precipitation regime regions in Turkey. *Journal of Hydrology* 450(1): 180-189.
22. Saligheh, M. 2006. Rainfall mechanism in southeast of country. *Geographical Research Quarterly* 38(55): 1-13 (In Farsi).
23. Shahbaee Kotenae, A. and H. Asakereh. 2019. Spatial analysis features of autumn rainfall in north west Iran. *Geographic Space* 19(65): 247-267 (In Farsi).
24. Shirani, K., F. Heydari and A. Arabameri. 2018. Comparison of artificial neural network and multivariate regression methods in landslide hazard zonation, case study: Vanak Basin, Isfahan province. *Watershed Engineering and Management* 9(4): 451-464 (In Persian).
25. Soltani, A., H. Sadoghi Yazdi, S. Ashkzari Toosi and M. Rouhani. 2011. Improving the Kohonen self-organizing network with the aim of fuzzy data clustering. In: 10th Iranian Fuzzy Systems Conference, Shahid Beheshti University, Tehran, Iran (In Persian).
26. Sumit, S. H. and S. Akhter. 2019. C-means clustering and deep neuro-fuzzy classification for road weight measurement in traffic management systems. *Soft Computing* 23(1): 4329-4340.
27. Taxak A. K., A. R. Murumkar and D. S. Arya. 2014. Long term spatial and temporal and homogeneity analysis in Wainganga basin (Central India). *Weather and Climate Extremes* 4(1): 50-61.
28. Yan, H., Z. Zou and H. Wang. 2010. Adaptive neuro fuzzy inference system for classification of water quality status. *Journal of Environmental Sciences* 22(12): 1891-1896.
29. Yarahmadi, D., M. Halimi and Z. Zarei Chaghbalki. 2015. Analysis of spatial patterns of monthly precipitation in west and northwest Iran using spatial autocorrelation. *Physical Geography Research Quarterly* 47(3): 451-464 (In Farsi).

# Evaluation of Fuzzy Clustering and Artificial Neural Network Methods in Spatial Zoning of Annual Precipitation in Iran

A. Shahbaee Kotenae\* and H. Asakereh<sup>1</sup>

(Received: February 9-2022 ; Accepted: May 14-2022)

## Abstract

Precipitation is one of the main elements of the Earth's hydro-climatic cycle and its variability depends on the complex and non-linear relationships between the climate system and environmental factors. Understanding these relationships and doing environmental planning based on them is difficult. Therefore, classifying data and dividing information into homogeneous and small categories can be helpful in this regard. In the present study, an attempt was made to prepare precipitation, altitude, slope, slope direction, and station density data for 3423 synoptic, climatological, and gauge stations in Iran in the 1961-2015 years' period. These data were entered into fuzzy (FCM), self-organizing map neural network (SOM-ANN) models and precipitation-spatial zoning. The outputs of the two models were compared in terms of accuracy and efficiency. The results obtained from the output of the models have divided the rainfall conditions of Iran into four zones concerning environmental factors. Evaluations also showed that both models had high accuracy in classifying precipitation parameters; However, the fuzzy model has a relative advantage over the neural network model in the accuracy of results.

**Keywords:** Self-organizing map neural network, Fuzzy model, Zoning, Precipitation, Iran

---

1. Department of Geography, Faculty of Humanities, Zanjan University, Zanjan, Iran.

\*: Corresponding author, Email: Asakereh1@yahoo.com