

خوشه بندی زمانی - مکانی، از جمله نوآوری های این تحقیق محسوب می شود. نتایج این تحقیق نشان دهنده پتانسیل استفاده از نتایج خوشه بندی در پیش بینی رفتار فصلی بارش در منطقه جنوب شرق ایران می باشد.

واژه های کلیدی : خوشه بندی، پیش بینی بارش، الگوریتم ژنتیک، دمای سطح دریا (SST)

مقدمه

روش های خوشه بندی، مجموعه ای اطلاعات را در یک فضای چند بعدی به گروه هایی تقسیم می کنند که دارای دو ویژگی اساسی باشند. ویژگی اول این است که بیشترین شباهت بین داده های قرار گرفته در خوشه ها برقرار باشد و از طرف دیگر کمترین شباهت بین خوشه ها به صورت دو به دو وجود داشته باشد. به طور کلی روش های مختلف خوشه بندی به دو بخش سلسله مراتبی^۱ و غیر سلسله مراتبی تقسیم می شوند. همچنین در برخی روش ها تمامی اطلاعات و در برخی دیگر، تنها برخی از اطلاعات خوشه بندی می شوند. از طرفی خوشه بندی می تواند در مقیاس زمانی و مکانی و یا هر دو صورت گیرد. در علوم هیدرولوژی و سایر زمینه های مرتبط با منابع آب، همواره لزوم توجه به مطالعات اقلیمی و بررسی تغییرات آب و هوا و شناخت سیگنال های هواشناسی مورد توجه بوده است. در حقیقت بررسی این تغییرات کمک شایانی به پیش بینی بارش و برنامه ریزی دراز مدت منابع آب هر منطقه خواهد نمود. در بین بحث های مربوط به اقلیم و پیش بینی متغیرهای هیدرولوژیکی از جمله بارش، بحث خوشه بندی سیگنال های هواشناسی و اطلاعات دما و فشار سطح دریا و ارتباط آن با این متغیرها بخصوص با مسئله بارندگی اهمیت بسیار زیادی دارد. تا به امروز خوشه بندی کاربردهای متعددی در زمینه های مختلف از جمله هواشناسی و هیدرولوژی داشته است.

امروزه روش های متعددی جهت خوشه بندی انواع متغیرها مثل روش های EM^A، Fuzzy C-means Map to map، K-means

3- SST

4- GA-Clustering

5- Crossover

6- Mutation

7- Hierarchy

8- Expectation Maximization

خوشه بندی سیگنال های هواشناسی

با توجه به تغییرات بارش

(مطالعه موردی: پیش بینی بارندگی استان سیستان و بلوچستان)

بنفشه زهرایی^۱ و عباس روزبهانی^۲

چکیده

امروزه لزوم مطالعه و بررسی تغییرات آب و هوایی و شناخت رفتار متغیرهای مختلف هواشناسی مثل بارش، تبخیر، دما و فشار هوا در نقاط مختلف جهان بخصوص در کشورهایی که با تنوع آب و هوایی گوناگون و وقوع دوره های خشک و تر شدید مواجه هستند، هم از جهت برنامه ریزی منابع آبی و هم از جهت مدیریت شرایط بحران اهمیت زیادی دارد. از جمله این کشورها، کشور ایران می باشد که در منطقه جنوب شرق آن و بخصوص استان سیستان و بلوچستان، ستادهای بحران سیل و خشکسالی بطور همزمان فعال هستند و این منطقه از جمله بخش هایی از کشور است که دارای تنوع آب و هوایی شدید و تغییرات بسیار قابل ملاحظه در سالهای مختلف است. تا کنون روشهای متنوعی جهت خوشه بندی اطلاعات، بخصوص داده های هواشناسی به کار گرفته شده است. در این مقاله، روشی نوین برای خوشه بندی اطلاعات ماهواره ای دمای سطح دریا^۳ در مناطق موثر بر آب و هوای استان سیستان و بلوچستان شامل دریای عمان، دریای عرب و اقیانوس هند، با استفاده از روش الگوریتم ژنتیک^۴ ارائه شده است. مراحل مختلف این تحقیق شامل انتخاب ساختار مدل، تابع هدف، طول کروموزومها و عملگرهای تزویج^۵ و جهش^۶ می باشد. در تدوین مدل خوشه بندی، علاوه بر رفتار زمانی و مکانی تغییرات SST، ارتباط این متغیر با وقوع بارندگی های کمتر و بیشتر از نرمال مورد بررسی قرار گرفت. در مطالعه موردی ارائه شده، اطلاعات بارش در ۲۰ ایستگاه مختلف استان سیستان و بلوچستان در نظر گرفته شد. استفاده از روش الگوریتم ژنتیک جهت خوشه بندی SST و ارائه الگوریتمی پویا برای بیان ارتباط بارش منطقه مورد مطالعه در زمان های مختلف با الگوهای

۱- استادیار دانشکده مهندسی عمران و عضو قطب علمی مهندسی و مدیریت

زیرساختارها، دانشگاه تهران، bzahraie@ut.ac.ir

۲- دانشجوی دکتری مهندسی آب، دانشکده مهندسی عمران، دانشگاه تهران،

roozbahany@ut.ac.ir

غیره وجود دارد. از جمله این روش‌ها، روش K-means است که روشی ساده بوده و به وفور در خوشه‌بندی داده‌ها مورد استفاده قرار گرفته است. پاسخ نهایی این روش بستگی زیادی به تعداد خوشه‌هایی دارد که از ابتدا تعیین می‌شود. با این حال، گاهی این روش در شرایط خاص، در همگرا شدن به جواب نهایی موفق عمل نمی‌کند. روش شاخه و مرز^۱ نیز توسط کونتز و همکاران [۴] گسترش داده شد که نیازمند زمان زیادی برای محاسبات بود. لورت و همکاران [۵] روش خوشه‌بندی فازی را با نگرش همپوشانی خوشه‌های مختلف تعمیم داد.

به طور کلی مهمترین مسئله در خوشه‌بندی، مشکل تقسیم‌بندی مجموعه‌ای از داده‌های همگن به یک سری خوشه‌های طبیعی بدون دانش قبلی در مورد تعداد و مراکز خوشه‌ها می‌باشد که از بین این روش‌ها و روش‌های دیگر موجود، روش خوشه‌بندی با استفاده از الگوریتم ژنتیک روشی نوین و کارآمدی است و بوسیله آن می‌توان با یک فرآیند جستجوی غیر قطعی، ساختار خوشه‌بندی مناسب را تعیین نمود. یکی از ویژگی‌های اساسی الگوریتم ژنتیک انعطاف پذیر بودن آن در جستجوی فضای جواب‌ها می‌باشد که منجر به یافتن بهترین ساختار خوشه‌بندی می‌شود. هیچ یک از روش‌های دیگر قادر نیستند به طور موثری گروه‌های همگن از داده‌ها را با توجه به ماهیت تصادفی آنها تشکیل دهند بخصوص زمانی که تعداد خوشه‌ها زیاد باشند. خوشه‌بندی سیگنال‌های هواشناسی مثل SST و SLP^۲ (فشار سطح دریا) توسط روش‌های ذکر شده در بالا صورت گرفته است اما تا کنون از الگوریتم ژنتیک برای خوشه‌بندی این متغیرهای هواشناسی استفاده نشده است و بیشتر کاربرد آن در مسائل اقتصادی و اجتماعی بوده است.

توسعه روش خوشه‌بندی با استفاده از الگوریتم ژنتیک به اوایل دهه ۹۰ باز می‌گردد. آلی پی و کوکچیارا، باندیوپادها و مولیک تلاش‌های متعددی را در این مسیر صرف نمودند و روش مناسبی را برای خوشه‌بندی با الگوریتم ژنتیک با استفاده از ساختار اعداد واقعی ژن‌ها و با فرض نامشخص بودن تعداد خوشه‌ها ارائه دادند [۱ و ۲].

روش نوین دیگری برای خوشه‌بندی خودکار توسط چاد هاری ارائه شد [۳].

لین و همکاران [۶] یک روش موثرتر بر مبنای کدگذاری دو دویی برای خوشه‌بندی با تعداد نامشخص خوشه‌ها ارائه دادند که در آن با ارائه ساختار متفاوتی از عملگرهای جهش و تزیج، سرعت و دقت رسیدن به جواب بهینه را افزایش دادند.

به طور کلی مشابهت زیادی بین کارهای صورت گرفته در مقالات اخیر دیده می‌شود. در این روش‌ها تمامی داده‌ها به صورت ژن و در موقعیت‌های تصادفی در کروموزوم‌ها چیده می‌شوند و عمل خوشه‌بندی با محاسبه فاصله داده‌ها از مراکز خوشه‌ها که ممکن

است واقعی یا تصادفی باشند، صورت می‌گیرد. خروجی این مدل‌ها به این دلیل که کلیه متغیرهای مورد خوشه‌بندی در تعیین مختصات خوشه‌ها دخالت داده می‌شوند، قابل استفاده برای پیش‌بینی نیستند. در ادامه این بخش، سابقه چند تحقیق مرتبط در مورد بررسی ارتباط SST با بارش ارائه شده است.

ناظم السادات [۷] نشان داده است که تغییرات دمای سطح آب در خلیج فارس تأثیر معنی‌داری بر تغییرات بارندگی در مناطق جنوب و جنوب غرب ایران دارد. مطالعات وی نشان می‌دهد که بارش زمستانه (ژانویه تا مارچ) در مناطق مزبور با SST خلیج فارس رابطه معکوس دارد. روکونک و همکاران [۸] همبستگی همراه با تأخیر بین SST و بارندگی مانسون تابستانه شرق چین را بررسی نموده‌اند. ایشان برای این کار از دمای سطحی مناطقی از اقیانوس هند، آرام و اطلس استفاده کردند. آنها بدون اینکه به دنبال توجیه فیزیکی ارتباط بین این دو پارامتر باشند، مشاهده نمودند که دمای سطح غرب اقیانوس آرام و جنوب شرق اقیانوس هند، بیشترین همبستگی همراه با تأخیر را با بارندگی‌های مزبور دارد.

اسملر و ژاکوب [۹] به بررسی زمانی- مکانی اثر SST بر روی تداوم و شدت بارش‌های شدید در منطقه وسیعی از اروپا پرداختند و به نتایج جالبی دست یافتند. از جمله این نتایج، تأثیر چشمگیر بالا رفتن مقدار SST در دریای بالتیک بر افزایش بارش‌های سنگین در مناطق اطراف آن دریا بوده است.

مسئله پیش‌بینی بارش از مهمترین مسائل با اهمیت در بحث مدیریت کلان‌آبی هر کشور است و با توجه به تحقیقات گذشته که ارتباط نزدیکی بین SST و بارش در برخی مناطق را نشان داده است، انتظار می‌رود در صورت دخالت دادن ارتباط SST و بارش در فرآیند خوشه‌بندی SST، بتوان خوشه‌ها را به نوعی شناسایی کرد که از اطلاعات آن بتوان در مورد تغییرات بارش اظهار نظر نمود.

در این مقاله، روش جدیدی ارائه شده است که علاوه بر نوین بودن آن در استفاده از الگوریتم ژنتیک برای خوشه‌بندی SST، ساختار مدل به نحوی تدوین شده است که خوشه‌بندی متغیر SST با توجه به متغیر دیگری مانند بارندگی (P) صورت گیرد به نوعی که ارتباط بین این دو متغیر بدون وارد کردن مستقیم بارندگی در فرآیند خوشه‌بندی، درون مدل گنجانده شده است. این قابلیت کمک شایانی به پیش‌بینی و تقسیم‌بندی وضعیت بارش یک منطقه به سال‌های کمتر از نرمال، نرمال و بیش از نرمال با در نظر گرفتن تغییرات سیگنال‌های هواشناسی در مناطق موثر بر بارش منطقه مورد نظر نموده است.

همچنین از آنجا که یکی از مشکلات الگوریتم ژنتیک در خوشه‌بندی، محدودیت تعداد داده‌های ورودی می‌باشد، در روش پیشنهادی با توجه به استفاده بهینه از ژن‌ها در تعریف مراکز خوشه‌ها، امکان بهره‌گیری از تعداد داده‌های بیشتر فراهم شده است که جزئیات آن در بخش‌های بعدی ارائه شده است.

1- Branch & Bound

2- Sea Level Pressure

مواد و روش‌ها

الگوریتم ژنتیک یک روش نوین بهینه‌سازی است که ساختار آن شامل تابع هدف، کروموزوم‌ها، عملگرهای جهش و تزیوج است. ساختار متداول خوشه‌بندی با استفاده از این الگوریتم، به طور خلاصه شامل مراحل زیر است:

- جایگزینی تصادفی داده‌ها در ژن‌ها و تولید جمعیت اولیه
- انتخاب تصادفی مراکز خوشه‌ها از داده‌های موجود
- محاسبه فاصله داده‌ها از مراکز خوشه‌ها و محاسبه مقدار تابع برازش برای هر کروموزوم در جمعیت کنونی
- به کار بردن عملگرهای اصلی جهش و تزیوج برای تولید نسل جدید
- جایگزینی نسل حاضر با نسل قبلی
- تکرار این روند تا همگرا شدن جواب‌ها و تعیین ساختار خوشه‌بندی

در مدل پیشنهادی ارائه شده در این مقاله، طول کروموزوم‌ها بر حسب تعداد نواحی انتخاب شده برای بررسی SST و همچنین تعداد خوشه‌ها، متفاوت است. در این مدل، تعداد نواحی مورد بررسی ۳ و تعداد حداکثر خوشه‌ها نیز ۳ در نظر گرفته شده و بنابراین طول کروموزوم معادل ۹ ژن خواهد بود که نمونه آن در شکل (۱) دیده می‌شود. در روش پیشنهادی، برای ساده‌تر کردن محاسبات، تنها از مراکز خوشه‌ها به عنوان اعداد واقعی داخل ژن‌ها استفاده می‌شود. همانطور که در شکل (۱) دیده می‌شود، $X_{i,j}$ معرف مقدار متغیر مورد نظر برای خوشه‌بندی در خوشه i ام از ناحیه جغرافیایی j ام است. در این پژوهش، متغیر مورد خوشه‌بندی دمای سطح دریا بصورت فصلی در سه ناحیه جغرافیایی مختلف و منتخب می‌باشد که در بخش مطالعه موردی مقاله شرح داده شده‌اند. مقیاس و بازه زمانی مورد استفاده در قسمت‌های بعدی مقاله تشریح شده است. نکته قابل توجه در این روش، ساختار خوشه‌بندی زمانی-مکانی آن است.

روش کار بدین صورت است که در ابتدا با تعیین محدوده تغییرات دمای سطح آب دریا (SST) در هر ناحیه، به تعداد اعداد ژن‌های هر کروموزوم، عدد تصادفی در محدوده مقادیر تاریخی مشاهده شده تولید می‌شود. به عنوان مثال برای شکل (۱)

$X_{1,1}$	$X_{1,2}$	$X_{1,3}$	$X_{2,1}$	$X_{2,2}$	$X_{2,3}$	$X_{3,1}$	$X_{3,2}$	$X_{3,3}$
-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

شکل ۱ - ساختار پیشنهادی برای کروموزوم نمونه در مدل GA برای خوشه‌بندی

$$SST_{\min,j} \leq X_{i,j} \leq SST_{\max,j} \quad (1)$$

در رابطه فوق، $SST_{\min,j}$ و $SST_{\max,j}$ به ترتیب مقادیر مشاهده

شده حداکثر و حداقل دمای سطح آب دریا در محدوده جغرافیایی است. در گام بعدی با استفاده از جدول داده‌ها که شامل سری زمانی SST و بارش می‌باشد، به خوشه‌بندی سال‌ها لازم به توضیح است که مقصود از خوشه‌بندی سال‌ها در واقع خوشه‌بندی میانگین مقادیر SST در ماه‌های مورد نظر در سال‌های متوالی می‌باشد. در حقیقت هر سطر اطلاعات این جدول که مربوط به یک سال خاص می‌باشد، به صورت مقابل است: $(SST_{1,t}, SST_{2,t}, SST_{3,t}, P_t)$ $SST_{j,t}$ در عبارت فوق نشان دهنده مقدار SST مشاهده شده در منطقه جغرافیایی j ام در فصل مورد نظر در سال t ام می‌باشد. P_t مقدار متغیر کمکی خوشه‌بندی می‌باشد که در این تحقیق میزان بارندگی در فصل مورد نظر در سال t ام می‌باشد و بنابراین همانطور که مشاهده می‌شود چهار سری زمانی به عنوان ورودی مدل انتخاب شده‌اند. برای خوشه‌بندی، به دلیل تغییرات اندک SST از رابطه نسبتاً ساده فاصله اقلیدسی استفاده شده است به طوری که فاصله اقلیدسی SST هر سال از مختصات مراکز خوشه‌ها که قبلاً به صورت تصادفی انتخاب شده بودند، از رابطه (۲) بدست می‌آید:

$$d_{i,t} = \sqrt{\sum_{j=1}^3 (SST_{t,j} - SST_{i,j}^c)^2} \quad (2)$$

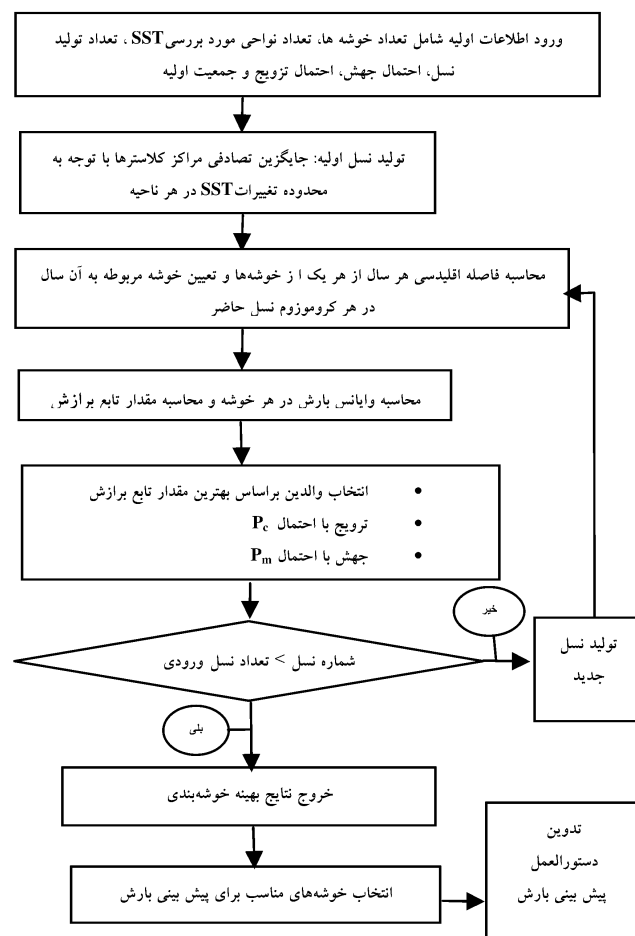
که در آن $d_{i,t}$ نشان دهنده فاصله اقلیدسی SST مشاهده شده در فصل مورد نظر در سال t ام از مرکز انتخاب شده برای خوشه i ام، $SST_{i,j}^c$ دمای سطح آب مشاهده شده در فصل مورد نظر در سال t ام در ناحیه جغرافیایی j ام و دمای سطح آب مرکز خوشه i ام در ناحیه j ام در فصل مورد نظر می‌باشد که به صورت تصادفی در الگوریتم ژنتیک (GA) تولید شده است.

این مقدار برای هر سال نسبت به هر یک از خوشه‌ها محاسبه می‌شود و کوچکترین مقدار این فاصله نشان دهنده تعلق آن سال به خوشه مورد نظر است. نکته مهم در این روش آن است که بارش منطقه مورد پیش‌بینی تأثیری در نحوه خوشه‌بندی ندارد بلکه تأثیر اصلی آن در محاسبه مقدار تابع برازش و انتخاب بهترین کروموزوم‌ها برای تولید نسل بعدی خواهد بود. پس از خوشه‌بندی سال‌ها، روند محاسبه مقدار تابع برازش به صورت رابطه (۳) است:

$$\text{Min} [\text{Max} (\text{Var} (\underline{P}))] \quad (3)$$

در این رابطه، Var نشان دهنده واریانس و \underline{P} نشان دهنده مجموعه مقادیر بارندگی سال‌هایی است که در هر خوشه قرار گرفته‌اند. $[\text{Max} (\text{Var} (\underline{P}))]$ تابع برازش مدل می‌باشد. حداقل این مقدار در هر نسل به عنوان تابع هدف مسئله شناخته می‌شود. رابطه (۳) نشان دهنده انتخاب کروموزوم‌هایی به عنوان کروموزوم برتر است که بارندگی متناظر با سال‌های قرار گرفته در هر خوشه، دارای حداقل واریانس مطابق تعریف فوق باشند. در نهایت در هر نسل، کروموزومی به عنوان کروموزوم برتر شناخته می‌شود که تابع هدف کوچکتری داشته باشد. برای ساده‌تر شدن درک نحوه کار

مدل، الگوریتم آن در شکل (۲) نشان داده شده است.



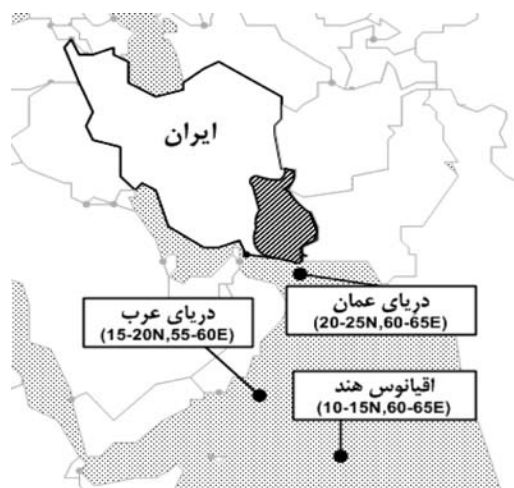
شکل ۲- الگوریتم مدل پیشنهادی برای خوشه بندی SST با استفاده از اطلاعات بارش

منطقه مورد مطالعه:

برای بررسی کارایی مدل خوشه بندی SST با استفاده از بارش، از اطلاعات بارندگی استان سیستان و بلوچستان استفاده شده است. استان سیستان و بلوچستان با مساحتی در حدود ۱۸۱۰۰۰ کیلومتر مربع و جمعیت ۱/۷ میلیون نفر، دارای میانگین سالانه ریزش جوی حدود ۱۱۰ میلیمتر می باشد. بر اساس تقسیم بندی وزارت نیرو تمامی یا بخشی از زیر حوضه های آبریز هیرمند، کویر لوت، جازموریان، ماشکید و دریای عمان در این استان قرار دارند. این استان دارای آب و هوای خشک با بارندگی ناچیز است و رژیم بارندگی آن اغلب به صورت رگبارهای سیل آسا است. کمبود نزولات و پراکندگی و نامشخص بودن زمان بارش از یکسو و بالا بودن سطح میزان تبخیر از سوی دیگر باعث شده تا منابع آب در دسترس با کمبود شدیدی مواجه شود. به طور کلی مشکلات متعدد منابع آبی استان مانند نوسانات شدید منابع آبی، محدودیت منابع آب زیرزمینی، بارندگی اندک، بادهای موسمی مخرب و غیره،

اهمیت پیش بینی بارش به منظور برنامه ریزی و مدیریت صحیح منابع آبی در آن را بیشتر نمایان می سازد. بالغ بر ۷۰ ایستگاه باران سنجی توسط وزارت نیرو در این استان نصب شده است که در این تحقیق پس از بررسی ایستگاه های گوناگون، ۲۰ ایستگاه باران سنجی برای بررسی کارایی مدل انتخاب گردیده اند. ۱۰ ایستگاه از این تعداد، ایستگاه های شاخص باران سنجی استان سیستان و بلوچستان، معرفی شده توسط شرکت سهامی آب منطقه ای استان سیستان و بلوچستان می باشند که عبارتند از ایستگاه های چابهار، ایرانشهر، سراوان، زاهدان، زابل، لادیز، نصرت آباد، خاش، قصر قند و پیشین که برخی از این ایستگاه ها در برخی سالها مشکل کمبود اطلاعات دارند.

همچنین تعداد ۱۰ ایستگاه باران سنجی دیگر نیز که دارای طول دوره آماری بیشتری بوده اند مورد تحلیل قرار گرفتند. طول دوره آماری مورد استفاده نیز ۲۱ سال می باشد. علاوه بر این برای بررسی تاثیر و ارتباط بین SST و بارش، نقشه های تغییرات SST در دریاهای اطراف منطقه در ماه های مختلف سال مورد بررسی قرار گرفت و در نهایت سه منطقه منتخب در دریای عمان، دریای عرب و اقیانوس هند جهت بررسی و خوشه بندی SST انتخاب گردید. موقعیت این مناطق در شکل (۳) و همچنین مشخصات جغرافیایی این ۳ منطقه در جدول (۱) آورده شده است.



شکل ۳- مناطق تحت مطالعه برای در نظر گرفتن اطلاعات SST

جدول ۱- مشخصات جغرافیایی مناطق مورد بررسی برای در نظر گرفتن اطلاعات SST

ردیف	نواحی جغرافیایی	عرض جغرافیایی	طول جغرافیایی
۱	دریای عمان	۲۵N-۲۰N	۶۵E-۶۰E
۲	دریای عرب	۲۰N-۱۵N	۶۰E-۵۵E
۳	اقیانوس هند	۱۵N-۱۰N	۶۵E-۶۰E

نتایج مدل

در مدل خوشه بندی، به منظور بدست آوردن بهترین مقادیر پارامترهای مختلف الگوریتم ژنتیک، تحلیل حساسیت انجام شد و در نهایت تعداد نسل اولیه الگوریتم ژنتیک مورد نظر ۵۰ و بهترین مقادیر احتمال جهش و توزیع به ترتیب ۰/۰۶ و ۰/۶، انتخاب شدند. لازم به ذکر است که در این تحلیل، حساسیت سرعت همگرایی به مقادیر مختلف این پارامترها مورد بررسی قرار گرفته است.

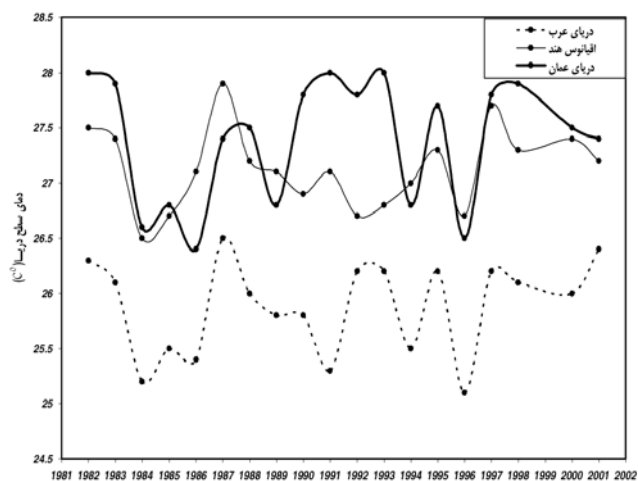
بررسی نتایج نشان داد که از بین ایستگاه‌های انتخاب شده، بارندگی ثبت شده در ایستگاه‌های سراوان، کجدار، لادیز و چابهار تاثیر بیشتری از تغییرات SST گرفته اند که نتایج آن در شکل (۸) دیده می شود. در این شکل، نقاط مشخص شده با علائم دایره و مربع، نشان دهنده نقاط واقع شده در خوشه‌های منتخب هستند. زمانی که تاثیر SST درای عمان بر ایستگاه سراوان بررسی می شود، مطابق شکل (۸-الف)، مدل خوشه بندی، دو خوشه را انتخاب می کند که دسته اول شامل ۹ سال با احتمال ۷۸٪ دارای بارندگی بیش از میانگین و مابقی سال‌ها در دسته دوم و با احتمال ۸۰٪ دارای بارش کمتر از میانگین دراز مدت می باشند. همچنین خوشه بندی بر اساس SST دریای عرب نیز تحت بررسی قرار گرفت که همانطور که در شکل (۸-ب) مشاهده می شود، ۵ سال در یک خوشه قرار گرفته اند که نشان دهنده بارش بیش از میانگین با احتمال ۸۰٪ می باشند.

در مورد ایستگاه کجدار (شکل ۸-ج) نیز تنها تاثیر منطقه دریای عمان قابل توجه بوده و در آن ۴ سال در یک خوشه واقع شده و همگی دارای بارش کمتر از میانگین هستند. در شکل (۸-د) وقتی تاثیر اقیانوس هند بر بارش ایستگاه لادیز بررسی می شود ۴ سال با بارندگی بیش از نرمال در یک خوشه و ۱۱ سال در یک خوشه قرار دارند که نمایان کننده بارش زیر نرمال با احتمال ۸۰٪ می باشد.

در مورد ایستگاه چابهار و با توجه به شکل (۸-و) تاثیر دریای عمان حاکی از این است که ۴ سال با بارش بالای نرمال در یک خوشه و ۹ سال با احتمال ۸۰٪ در یک خوشه با بارش زیر نرمال قرار می گیرند. تاثیر هر سه منطقه جغرافیایی در شکل (۸-ه) نشان می دهد که ۱۴ سال با احتمال ۸۵٪ با بارش زیر نرمال در یک خوشه قرار می گیرند.

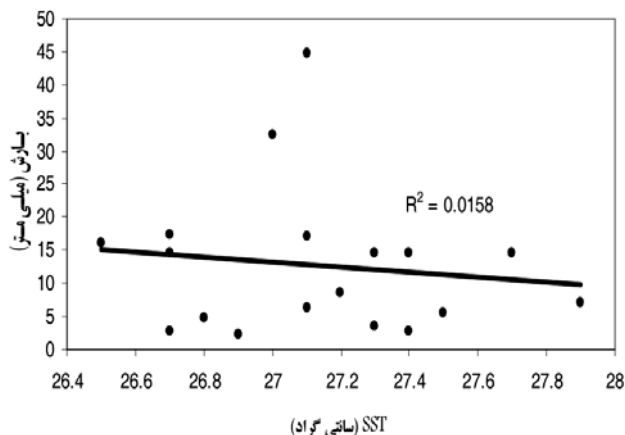
همچنین جهت مقایسه نتایج روش پیشنهادی خوشه بندی بر اساس GA با روش‌های مرسوم که در تحقیقات گذشته در جهت بیان ارتباط و همبستگی بین بارش و SST صورت گرفته است، از روش رگرسیون خطی مرتبه اول استفاده شده است. چند نمونه از نتایج این رگرسیون‌ها در اشکال (۵) الی (۷) نشان داده شده است که همانطور که مشخص است روش رگرسیون خطی قادر به بیان همبستگی این دو متغیر نمی باشد و از دقت کافی برای پیش بینی برخوردار نمی باشد.

لازم به ذکر است که اطلاعات SST از مرکز تحقیقات ملی اتمسفری کشور آمریکا (NCAR^۱) بدست آمده است. پس از بررسی نقشه‌های SST در منطقه و بررسی ارتباط تغییرات آنها با وقوع فصول کم بارش و پر بارش در منطقه مورد مطالعه، برای ورودی‌های مدل از اطلاعات بارندگی در ماه‌های دی، بهمن و اسفند که ماه‌های پر بارش در سال‌های ۱۳۶۱ تا ۱۳۸۱ و همچنین از اطلاعات SST در ماه‌های آگوست، سپتامبر و اکتبر سال‌های ۱۹۸۲ تا ۲۰۰۱ که مصادف با ماه‌های مرداد، شهریور و مهر می باشند، استفاده شد. لازم به ذکر است که سعی شده انتخاب ماه‌ها به نحوی انجام شود که اطلاعات SST برای پیش بینی ماه‌های پر بارش منطقه مورد نظر قابل استفاده باشد که بتوان در مورد وضعیت آبی منطقه از لحاظ کم بارش یا پر بارش بودن با تاخیر زمانی مشخص تصمیم‌گیری کرد. به عبارت دیگر با انتخاب ماه‌های فوق‌الذکر، در انتهای ماه اکتبر (۹ آبان ماه) می توان در مورد بارش محتمل در زمستان بعد پیش بینی انجام داد. بررسی اولیه تغییرات اطلاعات SST و بارش در فصول ذکر شده، نشان دهنده ارتباط عددی بین اطلاعات بوده که مبنای طراحی مدل خوشه بندی بوده است. شکل (۴) سری زمانی دمای سطح دریا در این ۳ منطقه را نشان می دهد.



شکل ۴- متوسط دمای سطح دریا در ماه‌های آگوست، سپتامبر و اکتبر

لازم به ذکر است که مدل مورد نظر به راحتی قابلیت تغییر دادن تعداد نواحی مورد بررسی از ۱ به ۳ ناحیه را داراست. لذا تاثیر SST هر یک از نواحی به تنهایی، دو به دو و یا هر سه ناحیه بر روی بارش ایستگاه‌های مورد نظر بررسی شده است و در نهایت نتایج قابل توجهی در مورد برخی ایستگاه‌ها بدست آمده است. در مورد سایر ایستگاه‌ها، نتایج نشان گر عدم وجود همبستگی مناسب بین بارش این ایستگاه‌ها و SST مناطق جغرافیایی منتخب می باشد. همچنین در مورد ایستگاه‌هایی که نتایج قابل توجه بوده است نیز تعداد نواحی SST آنها بر بارش موثر است، از ۱ تا ۳ منطقه جغرافیایی متغیر است.

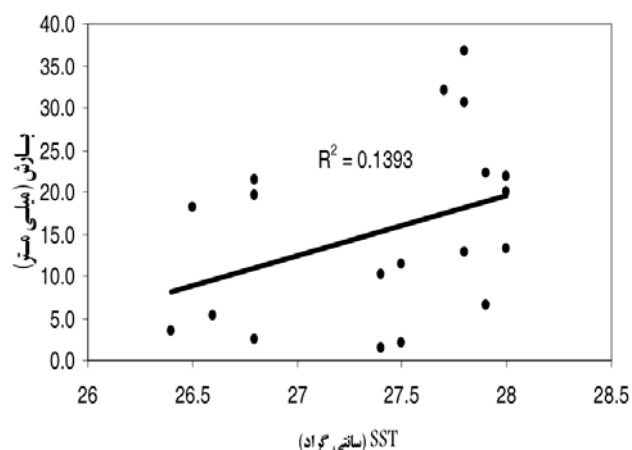


شکل (۷): معادله همبستگی خطی مرتبه اول بین بارش ایستگاه لادیز و SST در اقیانوس هند

جدول (۲) دستور العمل تدوین شده را برای پیش بینی بارش در ایستگاه‌های منتخب در ماه‌های دی و بهمن و اسفند، بر اساس مقدار SST در ماه‌های اگوست، سپتامبر و اکتبر بر اساس نتایج خوشه بندی نشان می‌دهد. روش کار بدین صورت است که بخشی از اطلاعات موجود برای مدلسازی و بخشی دیگر جهت آزمون مدل مورد استفاده قرار می‌گیرد. در واقع با دانستن اینکه به عنوان مثال بارش حوزه ایستگاه سراوان تحت تاثیر SST دریای عمان، دارای چه ترکیب خوشه بندی خواهد بود، با در دست داشتن متوسط مقدار SST دریای عمان در ماه‌های نام برده شده و با مراجعه به خوشه‌ای که SST مورد نظر در آن قرار دارد، می‌توان به وضعیت کلی متوسط بارش به صورت زیر نرمال یا بالای نرمال در ماه‌های دی و بهمن و اسفند همان سال پی برد و وضعیت بدست آمده را با وضعیت واقعی مقایسه کرد. با تکرار این آزمایش برای چند سال مختلف و مقایسه با وضعیت حقیقی بارش آن سال می‌توان به صحت عملکرد مدل تدوین شده پی برد و به پیش بینی وضعیت دراز مدت بارش در سال‌های آینده و در حوزه‌های مختلف استان پرداخت. در مثال یاد شده و با توجه به جدول (۲)، اگر متوسط اندازه گیری شده SST دریای عمان در ماه‌های مزبور (Aug.-Oct.) بین ۲۷/۷ و ۲۸ درجه سانتی‌گراد باشد، وضعیت بارش ایستگاه سراوان در زمستان همان سال بیشتر از میانگین ($P > 15/3\text{mm}$) و چنانچه SST بین ۲۶/۴ و ۲۷ درجه سانتی‌گراد باشد، بارش کمتر از میانگین ($P < 15/3\text{mm}$) مورد انتظار خواهد بود.

بحث و نتیجه‌گیری

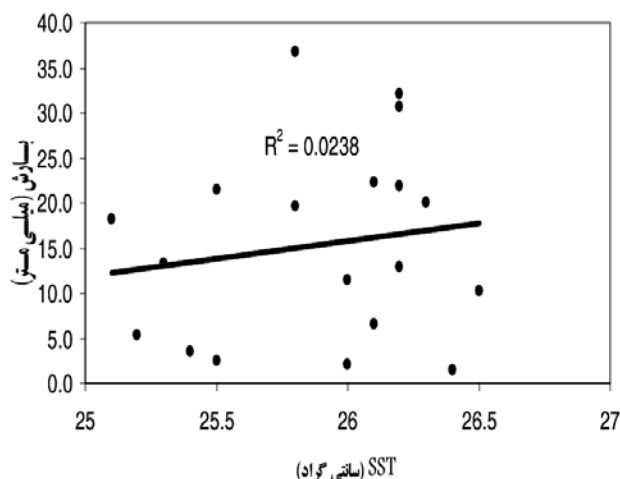
از گذشته تا کنون روش‌های گوناگون توصیفی و عددی برای تعیین وضعیت آب و هوایی مناطق مختلف وجود داشته است اما بهره‌گیری از مدل‌های هوشمند و سریع مانند الگوریتم ژنتیک تحول قابل توجهی را در این زمینه ایجاد کرده است. همانطور که ملاحظه شد روش ارائه شده نه تنها قابلیت خوشه بندی و جستجوی کامل فضای



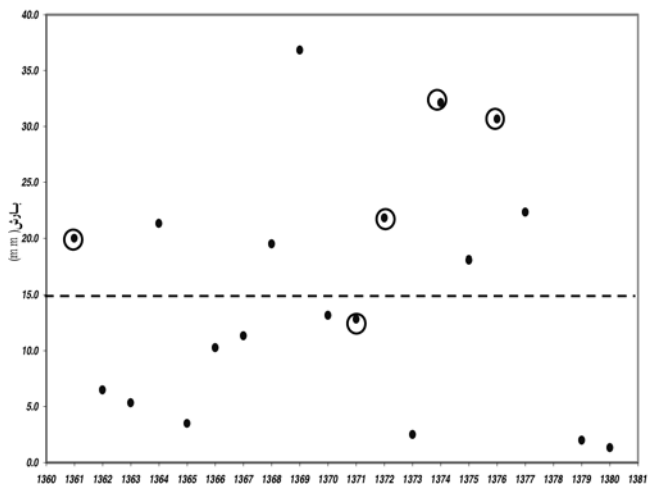
شکل ۵- معادله همبستگی خطی مرتبه اول بین بارش ایستگاه سراوان و SST در دریای عمان

پیش بینی بارش با استفاده از نتایج خوشه بندی

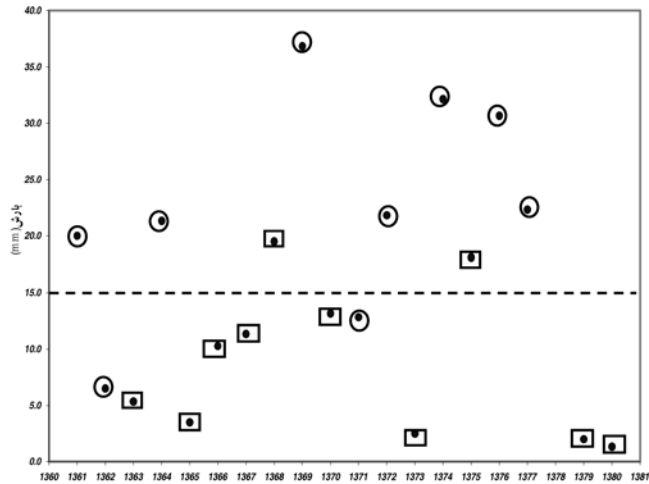
یکی از مهمترین کاربردهای خوشه بندی SST علاوه بر تعیین رفتار درازمدت این متغیر، بررسی تاثیر آن بر بارندگی و استفاده از آن در بحث پیش بینی بارش است. در تحقیق صورت گرفته در این مقاله، با استفاده از نتایج بدست آمده در هر ایستگاه و تعیین سال‌های با بارندگی بیش از نرمال و کمتر از نرمال می‌توان با توجه به میزان تغییرات SST در مناطق جغرافیایی انتخاب شده، وقوع بارش‌های بیشتر یا کمتر از نرمال را پیش بینی نمود. البته این امر مستلزم داشتن سری زمانی طولانی تری از بارش و SST می‌باشد. زیرا از طرفی می‌توان رفتار دراز مدت تری را از خوشه بندی صورت گرفته بررسی کرد و از طرفی امکان آزمون مدل تدوین شده برای پیش بینی بارش وجود خواهد داشت.



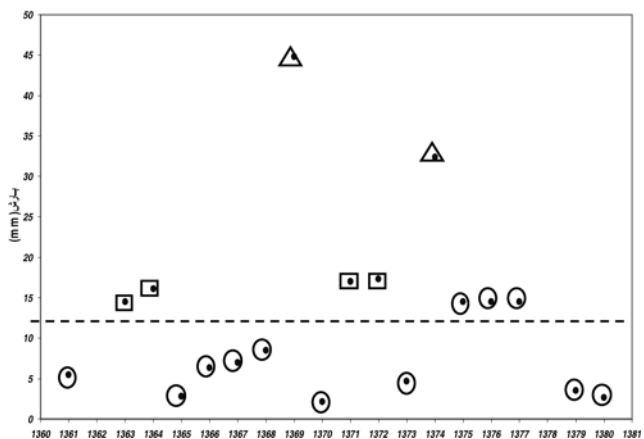
شکل ۶- معادله همبستگی خطی مرتبه اول بین بارش ایستگاه سراوان و SST در دریای عرب



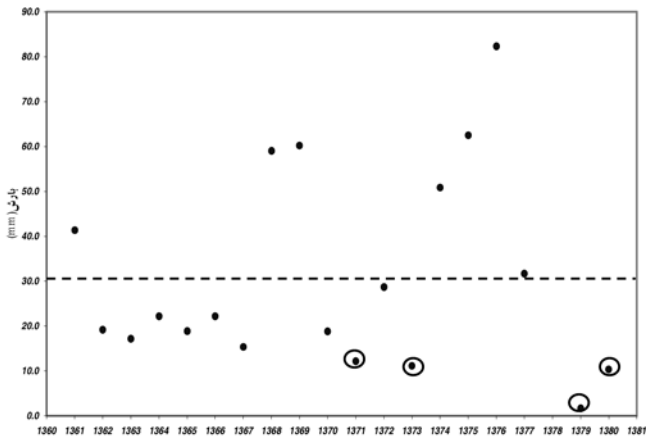
۸- ب : دریای عرب- ایستگاه سراوان



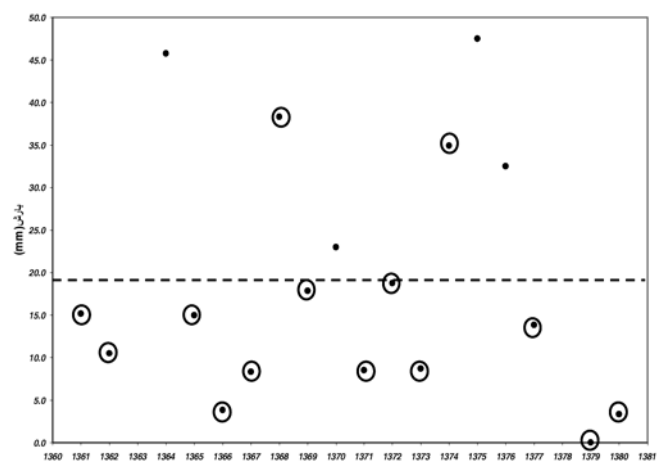
۸- الف : دریای عمان- ایستگاه سراوان



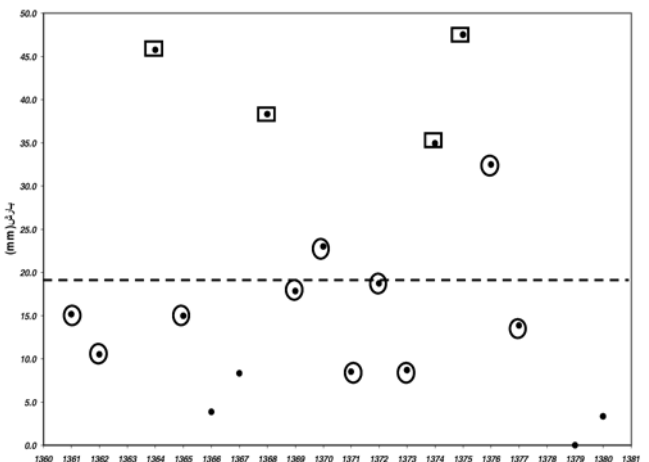
۸- د : اقیانوس هند- ایستگاه لادیز



۸- ج : دریای عمان- ایستگاه کجدار



۸- ه : دریای عمان ، دریای عرب و اقیانوس هند- ایستگاه چابهار



۸- و : دریای عمان- ایستگاه چابهار

شکل (۸): نتایج خوشه بندی SST با توجه به تغییرات بارندگی ایستگاه های منتخب

جدول ۲ - نتایج استفاده از خوشه بندی با الگوریتم ژنتیک در پیش بینی بارش ایستگاه های مختلف

نام ایستگاه	ناحیه جغرافیایی پیش بینی کننده (SST)	محدوده تغییرات SST(C°)	وضعیت پیش بینی بارش
سراوان	دریای عمان	$27/7 \leq SST_O \leq 28$	$P > 15/3 \text{ mm}$
		$26/4 \leq SST_O \leq 27/7$	$P < 15/3 \text{ mm}$
سراوان	دریای عرب	$26/2 \leq SST_A \leq 26/3$	$P > 15/3 \text{ mm}$
کجدار	دریای عمان	$27/4 \leq SST_O \leq 27/7$	$P < 30/08 \text{ mm}$
لادیز	اقیانوس هند	$26/5 \leq SST_I \leq 26/8$	$P > 12/7 \text{ mm}$
		$26/9 \leq SST_I \leq 27/9$	$P < 12/7 \text{ mm}$
چابهار	دریای عمان	$26/5 \leq SST_O \leq 26/8$	$P > 19/2 \text{ mm}$
		$27/7 \leq SST_O \leq 28$	$P < 19/2 \text{ mm}$
چابهار	دریای عمان ، دریای عرب و اقیانوس هند	$26/2 \leq SST_O \leq 26/3$ $25/4 \leq SST_A \leq 26/5$ $26/8 \leq SST_I \leq 28$	$P < 19/2 \text{ mm}$

P : مقدار بارندگی فصلی در ایستگاه

SST_A: دمای سطح دریا در دریای عرب
SST_I: دمای سطح دریا در اقیانوس هند
SST_O: دمای سطح دریا در دریای عمان

دریا)، مدل پیشنهادی را توسعه داد. همچنین با توجه به قابلیت متغیر بودن طول کروموزوم ها در مدل می توان به بررسی نواحی بیشتری در سطح اقیانوس هند و دریاهای اطراف منطقه مورد مطالعه پرداخت.

نتایج این تحقیق کمک شایانی در پیش بینی فصلی بارش و در نهایت برنامه ریزی سودمند منابع آبی کشور خواهد نمود و می تواند مورد استفاده سازمان ها، مشاوران و محققان مرتبط با علوم هیدرولوژی، هیدروکلیماتولوژی، آبخیزداری و غیره قرار گیرد.

قدردانی و تشکر

بخشی از نتایج ارائه شده در این مقاله در قالب پروژه تحقیقاتی "خوشه بندی سیگنال های هواشناسی با توجه به تغییرات بارش در استان سیستان و بلوچستان" با حمایت مالی معاونت فنی و پژوهش های شرکت سهامی مدیریت منابع آب ایران بدست آمده است. بدین وسیله از توجه این شرکت تشکر و قدردانی می شود.

جواب ها را ممکن می سازد، بلکه قادر خواهد بود. طوری عمل خوشه بندی را صورت دهد که ارتباط SST سال هایی که در یک خوشه قرار می گیرند با بارش همان سال ها در منطقه مورد مطالعه حفظ گردد و بتوان از این طریق به پیش بینی در مورد وقوع سال های با بارش کمتر از نرمال و بیشتر از نرمال دست یافت.

این نتیجه مهم فواید زیادی را در پی خواهد داشت از جمله این که با اندازه گیری و محاسبه دمای سطح آب می توان به پیش بینی فصلی بارش در مناطق تحت تاثیر قرار گرفته پرداخت و به طور عملی در مورد تاثیر منطقه ای و محلی سیگنال های هواشناسی و SST نیز، دستاوردهای ارزشمندی را به دست آورد. بوسیله این روش و با تدوین جدول پیش بینی می توان به مقایسه شباهت رفتاری ایستگاه های مختلف پرداخت. در مورد مطالعه موردی صورت گرفته امکان پیش بینی پر بارشی و کم بارشی دیده می شود در حالی که در صورت استفاده از مدل در مناطق دیگر ممکن است تنها یکی از این دو حالت قابل پیش بینی باشد. در ادامه این تحقیق، می توان با دخالت دادن دیگر متغیرهای مهم هواشناسی نظیر SLP (فشار سطح

using fuzzy clustering and local membership variations detection IEEE International Conference on Fuzzy Systems Volume , Issue , 8-12, pp. 221 - 228.

6- Lin H. F. Yang. and Y. Ta Kao. 2005. An Efficient GA-based Clustering Technique", Journal of Science and Engineering., 8(2), pp. 113-122.

7- Nazemosadat, M.J. 1998. Persian Gulf sea surface temperature as a drought diagnostic for southern parts of Iran, J. of Drought News Network, Vol. 10, pp. 12-14.

8- Rucong, Y. Minghua, Z. Yongqiang, Y. and Yimin, L. 2001. Summer monsoon rainfalls over Mid-Eastern China lagged correlated with global SSTs, J. of Advances in Atmospheric Sciences, Vol. 18, No. 2, pp.179-196.

9- Semmler, S., and Jacob, J. 2004. Modeling extreme precipitation events-a climate change simulation for Europe. J. of Global and Planetary Change, Vol. 44, pp.119-127.

1- Alippi, C. and Cucchiara, R. 1992. Cluster partitioning in image analysis and classification: A genetic algorithm approach. Proc.CompEuro 1992 on Comput. System Software Eng., IEEE Computer Society Press, Silver Spring, MD, pp. 139-144.

2- Bandyopadhyay, S. and Maulik, U. 2002. Genetic algorithm-based clustering technique, Journal of Pattern Recognition Society., Vol. 33, pp. 1455-1465.

3- Garai, G., and Chaudhuri, B.B. 2004. A novel genetic algorithm for automatic clustering, Journal of Pattern Recognition Letters, Vol. 25, pp. 173-187.

4- Koontz, W.L. Narendra, P.M. and Fukunaga, K. 1975. A branch and bound clustering algorithm, IEEE Trans. Comput. Vol. 24, pp. 908-915.

5- Levrat, E. Bombardier, V. Lamotte, M. and Bremont, J. 1992. Multi-level image segmentation

Abstract**Clustering Climate Signals based on Precipitation Variation (Case Study of Sistan and Baluchestan Province Precipitation Forecasting)**B. Zahraie¹ and A. Roozbahany²

Today, studying climatic variations with respect to different climate variables such as precipitation, evaporation, temperature, and air pressure in various parts of the world is important. This topic has received more attention especially in areas with intense wet and dry periods for water resources and crisis management. In Southeast of Iran, especially in Sistan & Baluchestan Province, very intense climatic variations have been observed in recent years. Various methods for clustering climate signals have been developed, which are suitable for modeling of climate variations in such areas. In this paper, a new method for clustering satellite information of Sea Surface Temperature (SST) using Genetic Algorithm Clustering is presented. Different parts of this research are based on selection of best model framework, objective function, length of chromosomes and mutation and crossover operators. This model tries to find the relation between SST data and above and below normal precipitation. 20 rain gauges in Sistan & Baluchestan Province have been considered in the case study. Using the Genetic Algorithm for mapping, the relation between precipitation in various seasons and temporal-spatial clustering are innovative aspects of this paper. The results show the potential of using the clustering method for forecasting the seasonal behavior of precipitation in Southeast of Iran.

Keywords: *Clustering, Climate Signals, Genetic Algorithm, Sea Surface Temperature (SST).*

1- Assistant Professor and Member of Center of Excellence for Infrastructure Engineering and Management, bzahraie@ut.ac.ir

2- PhD Student, School of Civil Engineering, University of Tehran, Iran, roozbahany@ut.ac.ir