

## مطالعه تطبیقی نمایه‌های هواشناسی خشکسالی SPI, SIAP به روش داده‌کاوی (مطالعه موردی استان کرمانشاه)

خلیل قربانی<sup>۱\*</sup> - علی خلیلی<sup>۲</sup> - سیدکاظم علوی پناه<sup>۳</sup> - غلامرضا نخعی‌زاده<sup>۴</sup>

تاریخ دریافت: ۸۸/۶/۱۰

تاریخ پذیرش: ۸۸/۱۱/۱۱

### چکیده

بارندگی یکی از مهم‌ترین متغیرهایی است که از آن در تعریف خشکسالی استفاده می‌شود. نمایه‌های مختلفی بر پایه بارش برای ارزیابی خشکسالی ارائه شده است که می‌توان به نمایه‌های بارش استاندارد شده (SPI) و معیار بارندگی سالانه (SIAP) اشاره کرد. هر یک از این نمایه‌ها برای توصیف خشکسالی، در طبقاتی تقسیم‌بندی می‌شوند که هر طبقه بیانگر وضعیتی از شدت خشکسالی می‌باشد. بررسی حالات وقوع همزمان طبقات هر نمایه معین خشکسالی می‌تواند معیار مناسبی برای ارزیابی تطابق نمایه‌ها باشد. قواعد انجمنی در روش‌های داده‌کاوی برای کشف قواعد و الگوها در مجموعه‌ای از داده‌ها بکار می‌رود. برای بررسی تطابق نمایه‌های SPI و SIAP بعد از محاسبه این نمایه‌ها برای یازده ایستگاه هواشناسی متعلق به وزارت نیرو واقع در استان کرمانشاه، با بکار بردن قواعد انجمنی، وقوع همزمانی هر یک از طبقات خشکسالی از این دو نمایه در مقیاس‌های زمانی فصلی، نیم‌سالانه و سالانه بررسی شد. نتایج نشان داد که در هیچ یک از مقیاس‌های زمانی، بین طبقات خشکسالی از هر یک از این دو نمایه در بیشتر موارد تطابق خوبی مشاهده نشد (کمتر از ۵۰ درصد) و رفتار متفاوتی از خشکسالی را نشان می‌دهند.

**واژه‌های کلیدی:** داده‌کاوی، خشکسالی، قواعد انجمنی، کرمانشاه

### مقدمه

مختلف، لازم است خشکسالی پایش شود. یکی از مشکلات این شناخت، عدم وجود یک تعریف دقیق و جامعی از خشکسالی است. به عقیده دراکوپ خشکسالی را می‌توان به سه نوع هواشناسی، هیدرولوژی و کشاورزی دسته‌بندی کرد (۱۰). خشکسالی هواشناسی زمانی که مقدار بارش طی دوره زمانی معین کمتر از مقدار متوسط آن باشد اتفاق می‌افتد بسته به هدف، این دوره می‌تواند ماه، و یا سال در نظر گرفته شود. در اثر خشکسالی هیدرولوژیکی جریان آب‌های سطحی و سطح آب دریاچه‌ها و منابع آبی (آب پشت سدها) کاهش می‌یابد و سطح سفره آب‌های زیرزمینی پایین می‌افتد. خشکسالی کشاورزی در اثر کمبود آب طی یک دوره خاص از رشد محصول یا کل دوره رشد آن اتفاق می‌افتد که باعث ایجاد تنش در گیاه و کاهش عملکرد محصول می‌شود.

بنابراین خشکسالی پدیده‌ای نیست که بطور مستقیم اندازه‌گیری شود. برای شناخت و اندازه‌گیری خشکسالی باید اثرات آن را کمی کرد و از روی آن کمیت‌ها خشکسالی را ارزیابی نمود. این کمیت‌ها معمولاً براساس اندازه‌گیری‌های زمینی بدست می‌آیند. بارندگی یکی از مهم‌ترین و اصلی‌ترین متغیرهایی است که از آن در تعریف

خشکسالی یکی از پدیده‌های زیان‌بخش جوی است که وقوع آن تقریباً در تمام نقاط جهان محتمل است ولی ویژگی‌های آن از منطقه‌ای به منطقه دیگر بطور محسوسی فرق می‌کند. برخلاف تصور اکثر افراد، خشکسالی رخدادی نادر و تصادفی نیست بلکه حالتی نرمال و مستمر از اقلیم است ولی از آنجایی که بصورت تدریجی ظاهر می‌شود روند آهسته‌تر و نامحسوس‌تری نسبت به سایر بلایای طبیعی دارد.

لازمه شناخت و مقایسه رویدادها بیان کمی و یا اندازه‌گیری آنها است. بنابراین برای مطالعه خشکسالی و مقایسه آن در زمان‌های

۱ و ۲- به ترتیب دانشجوی دکتری هواشناسی کشاورزی و استاد گروه آبیاری، دانشکده مهندسی و فناوری کشاورزی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج

(\*- نویسنده مسئول: Email:ghorbani.khalil@yahoo.com)

۳- استاد گروه کارتوگرافی، دانشکده جغرافیا، دانشگاه تهران

۴- استاد دانشکده اقتصاد، دانشگاه کالسروهه آلمان - کالسروهه آلمان

غربی آمریکا انجام دادند و نتیجه گرفتند که نمایه بارش استاندارد شده دست کم یک ماه زودتر از نمایه پالمر وقوع خشکسالی ۱۹۹۶ را پایش می‌نماید. و پیشنهاد کردند نمایه بارش استاندارد شده عملاً به‌عنوان ابزاری سودمند برای پایش خشکسالی، جایگزین نمایه پالمر در ایالات متحده شود (۱۳).

در زمینه استفاده از قواعد انجمنی در مطالعات خشکسالی نیز می‌توان به مطالعات زیر اشاره کرد.

تادسه<sup>۱۲</sup> پارامترهای اقیانوسی، اقلیمی و هواشناسی را برای کشف قواعدی که بتوانند رخداد خشکسالی را تشخیص دهند بکار گرفت. و نتیجه گرفت که بیشتر رخدادهای خشکسالی با مقادیر مثبت نمایه نوسانات جنوبی<sup>۱۳</sup>، مقادیر منفی نمایه ENSO چند متغیره<sup>۱۴</sup> و مقادیر منفی نوسانات دهکی آرام<sup>۱۵</sup> مرتبط هستند (۲۲).

هارمز و دئوگان<sup>۱۶</sup> از داده‌کاوای برای مدیریت خشکسالی استفاده کردند بطوریکه در کاوش قواعد انجمنی متوالی با قیود تأخیر زمانی ارتباطاتی بین حوادث ضمنی اقلیمی در مقیاس جهانی و شرایط محلی خشکسالی پیدا کردند (۱۲).

دهانیا و ناگش کومار<sup>۱۷</sup> برای ارتباط خشکسالی‌ها و سیل‌ها در هندوستان با داده‌های اقلیمی، از قواعد انجمنی داده‌کاوای استفاده کردند و روابط مستحکمی بین نمایه‌های اقلیمی منتخب از قبیل فشار سطح دریا داروین، نوسانات آتلانتیک شمالی، ال‌نینو و مقادیر درجه حرارت سطح دریا با رویدادهای بارش حدی پیدا کردند (۹).

## مواد و روش‌ها

### منطقه مطالعاتی

منطقه مطالعاتی استان کرمانشاه با مساحتی معادل ۴۳۴۲۴ کیلومتر مربع با میانگین بارش سالانه بین ۳۰۰ تا ۸۰۰ میلی‌متر می‌باشد. این تحقیق بر روی یازده ایستگاه هواشناسی متعلق به وزارت نیرو (شکل ۱) طی دوره آماری ۵۴-۱۳۵۳ تا ۸۴-۱۳۸۳ صورت گرفته است.

### استخراج نمایه‌های خشکسالی

#### نمایه بارش استاندارد SPI

نمایه SPI توسط مک‌کی و همکاران در سال ۱۹۹۳ بر اساس اختلاف بین مقادیر بارش و میانگین بارش برای یک بازه زمانی مشخص و تقسیم این مقدار بر انحراف معیار بارش پیشنهاد شد.

خشکسالی استفاده می‌شود. بر پایه کمیت بارش، نمایه‌های متعددی برای ارزیابی خشکسالی ارائه شده است که می‌توان به نمایه شدت خشکسالی پالمر<sup>۱</sup> (۱۹)، نمایه ناهنجاری بارندگی<sup>۲</sup> (۲۱)، نمایه دهک‌ها<sup>۳</sup> (۱۱)، نمایه رطوبت محصول<sup>۴</sup> (۲۰)، نمایه خشکسالی بالم و مولی<sup>۵</sup> (۶)، نمایه استاندارد شده بارش سالانه (۳)، نمایه بارش استاندارد (۱۵ و ۱۶)، نمایه درصد از نرمال<sup>۶</sup> (۱۴)، روش نیچه<sup>۷</sup> (۱۸) اشاره کرد.

مطالعات متعددی در ارزیابی خشکسالی توسط محققین در مناطق مختلف انجام شده است و به ارزیابی نمایه‌ها پرداخته‌اند. در زیر به چند تا از این تحقیقات اشاره می‌شود.

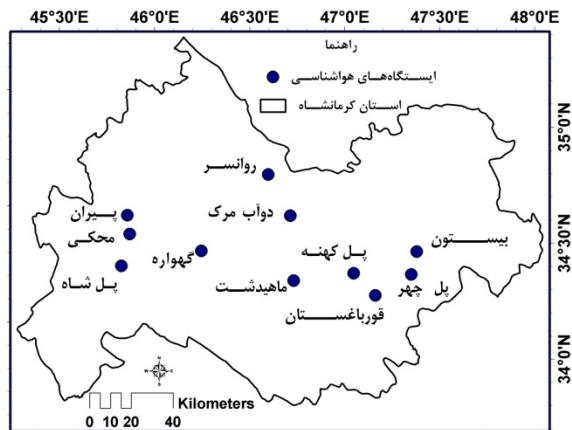
خلیلی و بذرافشان مطالعه‌ی تطبیقی برخی نمایه‌های خشکسالی هواشناسی در چند نمونه‌ی اقلیمی ایران را به انجام رساندند و نهایتاً به این جمع‌بندی رسیدند که نمایه‌های SIAP و DPI از نظر توالی، احتمالات شرطی و شدت خشکسالی، کاملاً مشابه بوده، در ارزیابی‌های خشکسالی نتایج تقریباً مشابهی عاید می‌سازند و نیز نمایه‌های SPI، BMDI، EPI<sup>۸</sup> و PNPI مشابه بوده و ارزیابی‌های تقریباً یکسانی از نظر توالی، شدت و احتمال وقوع خشکسالی ارائه می‌دهند. نمایه‌ی RAI<sup>۹</sup> از نظر توالی سال‌های خشک، نرمال و تر و نیز از نظر احتمال شرطی وقوع سال‌های خشک، نرمال و تر، تفاوت زیادی با سایر نمایه‌های خشکسالی داشته و در گروه مستقلی قرار می‌گیرد ولی از نظر شدت خشکسالی، مشابه نمایه‌های SPI، BMDI، EPI، و PNPI رفتار می‌نماید (۱).

مرید و همکاران برای پایش خشکسالی هیدرولوژیکی در استان تهران نمایه‌های خشکسالی شامل CZI، SPI، DPI، PNPI، ZSI<sup>۱۱</sup>، MCZI<sup>۱۰</sup> و EDI را در طول یک دوره‌ی آماری ۳۰ ساله مورد مطالعه قرار دادند و به این نتیجه رسیدند که: (۱) نمایه‌های SPI، CZI و ZSI رفتار نسبتاً مشابهی دارند. (۲) نمایه MCZI که با جایگزینی میانه بجای میانگین در نمایه CZI محاسبه می‌شود رفتار خشکسالی را در منطقه‌ی مطالعاتی توجیه نمی‌نماید. (۳) نمایه PNPI نیز برای منطقه‌ی مطالعاتی، مناسب تشخیص داده نشد (۴).

هیز و همکاران مقایسه‌ای بین دو نمایه بارش استاندارد شده و نمایه پالمر جهت بررسی خشکسالی‌های دشت‌های جنوبی و جنوب

- 1 - PDSI : Palmer Drought Severity Index
- 2 - RAI : Rainfall Anomaly Index
- 3 - DPI : Deciles of Precipitation Index
- 4 - CMI : Crop Moisture Index
- 5- BMDI : Bahlme and Mooley Drought Index
- 6 -PNPI : Percent of Normal Precipitation Index
- 7 - Nitzche
- 8 - EPI : Effective Precipitation Index
- 9 - RAI: Rainfall Anomaly Index
- 10-MCZI : Modified CZI
- 11-ZSI : Z-Score Index

- 12 - Tadesse
- 13 - SOI: Southern Oscillation Index
- 14 - MEI: Multivariate ENSO Index
- 15 - PDO: Pacific Decadal Oscillation
- 16- Harms and Deogun
- 17 - Dhanya and Nagesh Kumar



شکل ۱- شبکه ایستگاه‌های هواشناسی مطالعاتی

برای داده‌ها می‌باشد و این مسأله برای دوره‌های زمانی کوتاه مدت چندان مورد قبول نیست و باعث بروز اشکالاتی می‌شود (۲). در این تحقیق نمایه SPI در مقیاس‌های زمانی سه و شش ماهه و نیز سالانه محاسبه شد. و بر اساس جدول ۱ در پنج دسته طبقه‌بندی می‌شوند (۱).

### نمایه سیپا SPAP

خلیلی (۳) به منظور بررسی روند خشکسالی‌ها و ترسالی‌ها در گستره ایران نمایه معیار بارندگی سالانه را به صورت زیر ارائه کرد.

$$SIAP = \frac{P_i - \bar{P}}{SD}$$

جدول ۲- طبقات نمایه‌ی معیار بارندگی سالانه SIAP

نماد	مقادیر نمایه	ویژگی بارشی
SA <sub>1</sub>	بیشتر از ۰/۸۴	بسیار تر
SA <sub>2</sub>	۰/۵۲ تا ۰/۸۴	تر
SA <sub>3</sub>	۰/۵۲ تا -۰/۵۲	نرمال
SA <sub>4</sub>	-۰/۵۲ تا -۰/۸۴	خشک
SA <sub>5</sub>	کمتر از -۰/۸۴	بسیار خشک

P<sub>i</sub>: بارندگی در سال هیدرولوژی،  $\bar{P}$ : میانگین بارش‌های سالانه و SD: انحراف استاندارد بارش‌های سالانه می‌باشد. توضیح اینکه می‌توان از بارندگی‌های ماهانه نیز در فرمول فوق استفاده نمود. بنابراین P<sub>i</sub>: مقدار بارندگی در ماه هیدرولوژی و  $\bar{P}$ : میانگین بارش‌های ماهانه و SD: انحراف استاندارد بارش‌های ماهانه می‌باشد. با این وصف ویژگی‌های بارشی، با شرایط اخیر در مقیاس ماهانه ارائه می‌شود.

اساس نمایه بارش استاندارد بر محاسبات احتمالات وقوع بارندگی برای هر مقیاس زمانی استوار است. این نمایه صرفاً از داده‌های بارندگی ماهانه استفاده می‌کند و برای تشخیص کمبود میزان بارندگی در مقیاس‌های زمانی چندگانه (۳، ۶، ۱۲، ۲۴ و ۴۸ ماه) طراحی شده است.

محاسبه SPI شامل برازش تابع چگالی احتمال گاما بر توزیع فراوانی بارندگی کل برای یک ایستگاه معین می‌باشد. سپس، پارامترهای  $\alpha$  و  $\beta$  مربوط به تابع چگالی احتمال گاما برای هر ایستگاه، هر مقیاس زمانی (۳، ۶، ۱۲، ۲۴، ۴۸ ماه و غیره) و برای هر ماه سال برآورد می‌شوند. در مرحله بعد، از پارامترهای  $\hat{\beta}, \hat{\alpha}$  برای بدست آوردن تابع احتمال تجمعی بارندگی برای ماه مورد نظر و مقیاس زمانی معین در ایستگاه تحت بررسی استفاده می‌شود. پس از محاسبه تابع تجمعی کل، تغییر شکل هم احتمالی تابع تجمعی گاما به متغیر تصادفی نرمال استاندارد Z (یا SPI) با میانگین صفر و واریانس ۱ صورت می‌گیرد (۱).

جدول ۱- طبقات نمایه خشکسالی SPI

نماد	مقادیر نمایه	ویژگی بارشی
SP <sub>1</sub>	بزرگتر از ۱/۵	بسیار تر
SP <sub>2</sub>	۱ تا ۱/۵	تر
SP <sub>3</sub>	۱ تا -۱	نرمال
SP <sub>4</sub>	-۱ تا -۱/۵	خشک
SP <sub>5</sub>	کمتر از -۱/۵	بسیار خشک

کاربرد نمایه SPI درجهان بدلیل داشتن مزایایی چون سادگی در محاسبات، نیاز نداشتن به داده‌های زیاد و مستقل بودن از میانگین بارش رو به افزایش بوده و برای مقایسه‌ی دامنه وسیعی از اقلیم‌ها قابلیت استفاده دارد. از معایب این نمایه در نظر گرفتن توزیع نرمال

تفاوت اصولی که بین داده‌کاوی و آمار وجود دارد این است که داده‌کاوی یک رهیافت بدون پیش‌فرض است در حالی که بیشتر تکنیک‌های آماری معمول نیاز به پیش‌فرض دارند و آماردان‌ها در جستجوی معادلاتی برای مطابقت دادن با پیش‌فرض‌ها هستند. در مقابل الگوریتم‌های داده‌کاوی می‌توانند این معادلات را بطور اتوماتیک از اطلاعات موجود در مجموعه داده‌ها توسعه دهند (۸). داده‌کاوی تکنیکی است که آزمونهای فرض را ترکیب می‌کند و داده‌های مشتقی را کشف می‌کند. در آزمونهای فرض، محقق ایده را در برابر داده آزمون می‌کند تا اعتبار آن را تأیید یا رد کند. واندربرگ و همکاران توضیح می‌دهند که در کشف، محقق نتیجه را از داده‌ها ترسیم می‌کند و به داده‌ها اجازه می‌دهد تا نتیجه را بپذیرد. اغلب مسائل داده‌کاوی با استفاده از یک ترکیب از هر دو روش حل می‌شود. بطور مثال از نتیجه ممکن است فرضیه جدیدی برخیزد که می‌تواند آزمون شود و این آزمون تأیید یا رد شود (۲۳).

### قواعد انجمنی<sup>۲</sup>

قوانین انجمنی یکی از تکنیک‌های اصلی داده‌کاوی است و البته احتمالاً مهمترین شکل از کشف و استخراج الگوهای محلی در سیستم‌های یادگیری غیر هدایت شده می‌باشد. این روش از داده‌کاوی شباهت خیلی نزدیک به فرآیندهایی دارد که اغلب مردم وقتی سعی می‌کنند فرآیندهای داده‌کاوی را بفهمند، درباره آن فکر می‌کنند.

قواعد انجمنی در داده‌کاوی از زمان توسعه الگوریتم آپریوری<sup>۳</sup> بطور گسترده مورد استفاده قرار گرفته است و برای سنجش نکویی قاعده، معیارهای سنجش متعددی پیشنهاد شده است به طوری که حمایت<sup>۴</sup> و اطمینان<sup>۵</sup> معمولی‌ترین آنها می‌باشند و به این صورت تعریف می‌شوند:

قاعده انجمنی که بصورت  $x \rightarrow y$  تعریف می‌شود را در یک مجموعه تراکنش  $\Pi$  تایی در نظر می‌گیریم که  $x$  پیش‌آیند<sup>۶</sup> قاعده و  $y$  پس‌آیند<sup>۷</sup> آن می‌باشد. اگر تعداد تراکنش‌هایی که شامل آیت  $x$  یا  $y$  می‌باشند را به ترتیب با  $n_x$  و  $n_y$  همچنین تعداد تراکنش‌هایی که شامل آیت  $x$  هستند و آیت  $y$  را نیز شامل می‌شوند با  $n_{xy}$  نمایش دهیم سنجش‌های حمایت (supp) و اطمینان (conf) بصورت زیر تعریف می‌شوند:

الف) حمایت از پیش‌آیند ( $supp(x)$ ) که عبارتست از احتمال

2 - Association Rules

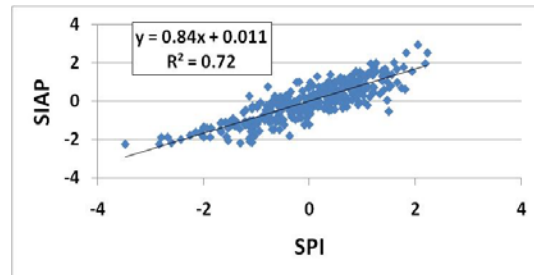
3 - Apriori

4 - Support

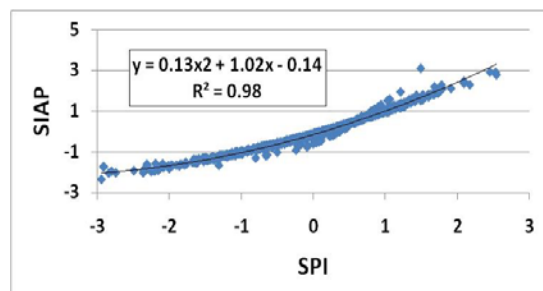
5- Confidence

6 - Antecedent

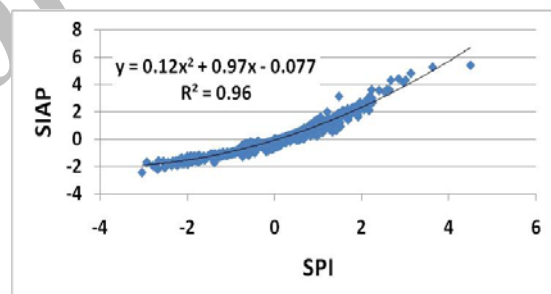
7 - Consequence



شکل ۲- همبستگی بین دو نمایه SPI و SIAP در مقیاس زمانی سالانه



شکل ۳- همبستگی بین دو نمایه SPI و SIAP در مقیاس زمانی نیم سالانه



شکل ۴- همبستگی بین دو نمایه SPI و SIAP در مقیاس زمانی فصلی

### داده‌کاوی

داده‌کاوی طریقه کشف عملی الگویی معنی‌دار، شکل و روند به روش سیفتینگ<sup>۱</sup> (روشی که به‌وسیله آن رکوردها حرکت داده می‌شوند تا ورود رکوردهای دیگر را مجاز سازند) از میان داده‌ها با استفاده از تکنیک تشخیص الگو مانند شبکه عصبی، آموزش ماشین و الگوریتم ژنتیک می‌باشد که توانایی تشخیص خودکار دانش موجود را دارد و از دیگر روش‌های تجزیه و تحلیل داده‌ها به دلیل کشف ساختار پنهان، نسبت‌ها، الگوها و علائم متفاوت است. این روش منحصرأ دینامیکی است چون که ناگزیر به تمهید برای سوالات نیست و می‌تواند روابط را بازیابی کند (۱۷).

1 - Sifting

در نظر گرفتن ۵ طبقه برای SIAP و ۵ طبقه برای SPI، تعداد ۲۵ قاعده مورد ارزیابی قرار گرفت.

با توجه به تعریف ارائه شده برای سنجش اطمینان، این سنجش مناسب تشخیص داده شد. در تفسیر این سنجش می‌توان گفت که: هریک از طبقات نمایه خشکسالی SPI را با  $SP_i$  و طبقات نمایه خشکسالی SIAP را با  $SA_j$  نمایش می‌دهیم به طوری که  $i, j$  بین یک تا پنج متغیر باشد. اگر رابطه  $SA_j \rightarrow SP_i$  درصد اطمینان بالایی داشته باشد شرط تطابق بین هریک از طبقات این دو نمایه این است که رابطه  $SA_j \rightarrow SP_i$  نیز از درصد اطمینان بالایی برخوردار باشد.

### نتایج و بحث

بر اساس نمایه‌های محاسبه شده SIAP و SPI برای ایستگاه‌های مطالعاتی، در سه مقیاس زمانی فصلی، نیم‌سالانه و سالانه، بانک اطلاعاتی داده‌ها تشکیل شد و الگوریتم آپروری بر آنها برآزش داده شد. قواعدی که درصد اطمینان بیش از ۵۰ درصد داشتند انتخاب شدند و نتایج آن به قرار زیر است:

#### قواعد انجمنی بین نمایه‌های SIAP و SPI در مقیاس زمانی فصلی

با جستجوی قواعد انجمنی بین نمایه‌های فصلی SIAP و SPI، ۹ قاعده با درصد اطمینان بالای ۵۰ درصد بدست آمد ولی با توجه به شرط تطابق ذکر شده، بین هیچ یک از طبقات خشکسالی مطابقتی مشاهده نشد (جدول ۳).

#### جستجوی قواعد انجمنی بین نمایه‌های SIAP و SPI در مقیاس زمانی نیم‌سالانه

برای نمایه‌ها در مقیاس زمانی نیم‌سالانه نیز، ۹ قاعده با درصد اطمینان بالای ۵۰ درصد بدست آمد ولی هیچ تطابقی بین آنها نیز مشاهده نشد (جدول ۴).

#### قواعد انجمنی بین نمایه‌های SIAP و SPI در مقیاس زمانی سالانه

در این مقیاس زمانی نیز بین ۹ قاعده بدست آمده هیچ تطابق خوبی بین نمایه‌ها دیده نشد (جدول ۵).

با افزودن آیتم زمان، روابط موجود بین نمایه‌های SIAP و SPI در دوره‌های زمانی فصلی در چهار فصل مجزا (بهار، تابستان، پاییز و زمستان) و دوره زمانی نیم‌سالانه در دوره شش ماه اول (از مهر تا فروردین) و شش ماه دوم (از فروردین تا مهر) به تفکیک جستجو شد که نتایج آن به قرار زیر است.

وقوع پیش‌آیند در یک مجموعه تراکنش:

$$supp(x) = p(x) = \frac{n_x}{n}$$

(ب) حمایت از پس‌آیند ( $supp(y)$ ) عبارتست از احتمال

وقوع پس‌آیند در یک مجموعه تراکنش:

$$supp(y) = p(y) = \frac{n_y}{n}$$

(ج) حمایت از قاعده ( $supp(x \rightarrow y)$ ) عبارتست از احتمال

وقوع قاعده انجمنی  $x \rightarrow y$  در یک مجموعه تراکنش:

$$supp(x \rightarrow y) = p(x \cap y) = \frac{n_{xy}}{n}$$

سنجش حمایت از آنجائی که تعداد تراکنش‌ها را به حساب می‌آورد

اغلب به‌عنوان قید تکرار نیز نامیده می‌شود. مجموعه آیتم‌هایی که حمایتی بیشتر از مقدار آستانه داشته باشند به‌عنوان مجموعه آیتم‌های بزرگ یا مکرر نامیده می‌شوند. مشکل عمده این سنجش عدم در نظر گرفتن تراکنش‌های نادر است. این سنجش توسط آگروال و همکاران (۵) معرفی شده است.

(د) اطمینان قاعده عبارتست از احتمال وقوع همزمانی پیش‌آیند و

پس‌آیند در مجموعه تراکنش‌هایی که شامل آیتم  $x$  می‌باشند:

$$conf(x \rightarrow y) = \frac{n_{xy}}{n_x}$$

این سنجش نیز توسط آگروال و همکاران معرفی شد. مسئله‌ای

که در مورد این سنجش وجود دارد این است که به تکرار پس‌آیند در یک مجموعه داده حساس می‌باشد به طوری که پس‌آیندهای با مقادیر حمایت بالا بطور خودکار، اطمینان با مقادیر بالا تولید می‌کنند حتی اگر قاعده انجمنی بین آیتم‌ها موجود نباشد مقدار این سنجش نیز بین صفر تا ۱۰۰ درصد تغییر می‌کند.

#### لیفت<sup>۱</sup>

لیفت یکی دیگر از سنجش‌های نکویی قاعده است که توسط

براین و همکاران معرفی شد (۷). سنجش لیفت بیانگر این است که چند بار انتظار داریم پیش‌آیند و پس‌آیند باهم اتفاق بیافتد در جایی که آنها از لحاظ آماری مستقل باشند. این سنجش مشکلی با وقایع نادر ندارد و بین صفر تا بینهایت تغییر می‌کند. مقادیر کمتر از یک به این معنی است که شرایط رضایت‌بخش از پیش‌آیند، احتمال پس‌آیند را در یک احتمال غیر شرطی کاهش می‌دهد. مقادیر بزرگتر از یک به این معنی است که شرایط رضایت‌بخش از پس‌آیند، احتمال پیش‌آیند را در یک احتمال غیر شرطی افزایش می‌دهد. اگر پیش‌آیند و پس‌آیند مستقل باشند مقدار لیفت برابر یک می‌باشد.

$$lift = \frac{p(x/y)}{p(y)} = \frac{p(x|y)}{p(x)p(y)} = \frac{n_{xy}n}{n_x n_y}$$

در این تحقیق هر یک از طبقات خشکسالی SIAP و SPI بطور

متناوب به‌عنوان پیش‌آیند و پس‌آیند در نظر گرفته شدند بنابراین با

جدول ۳- قواعد انجمنی بدست آمده بین نمایه‌های فصلی SIAP و SPI

لیفت	درصد اطمینان	درصد حمایت	آیند پس	آیند پیش
۶/۵	۱۰۰	۵/۴	SA5	SP4
۶/۵	۱۰۰	۶/۱	SA5	SP5
۵/۷	۱۰۰	۴/۷	SA1	SP1
۵/۴	۹۵	۹/۸	SA1	SP2
۵/۴	۵۳	۱۷/۵	SP2	SA1
۱/۴	۱۰۰	۹/۵	SP3	SA4
۱/۳	۱۰۰	۵۰/۴	SP3	SA3
۱/۳	۶۸	۷۴	SA3	SP3
۱/۳	۹۵	۷/۲	SP3	SA2

جدول ۴- قواعد انجمنی بدست آمده بین نمایه‌های نیم‌سالانه SIAP و SPI

لیفت	درصد اطمینان	درصد حمایت	آیند پس	آیند پیش
۴/۸	۱۰۰	۸/۱	SA5	SP5
۴/۸	۱۰۰	۸/۷	SA5	SP4
۴/۷	۱۰۰	۴/۳	SA1	SP1
۴/۷	۱۰۰	۱۲/۲	SA1	SP2
۴/۷	۵۸	۲۱/۱	SP2	SA1
۱/۵	۱۰۰	۱۰	SP3	SA2
۱/۵	۱۰۰	۱۲	SP3	SA4
۱/۵	۱۰۰	۳۶	SP3	SA3
۱/۵	۵۴	۶۶/۸	SA3	SP3

جدول ۵- قواعد انجمنی بدست آمده بین نمایه‌های سالانه SIAP و SPI

لیفت	درصد اطمینان	درصد حمایت	آیند پس	آیند پیش
۴/۴	۱۰۰	۸/۲	SA5	SP5
۳/۹	۸۹	۸/۵	SA5	SP4
۳/۵	۷۱	۴/۳	SA1	SP1
۲/۸	۵۸	۱۱/۶	SA1	SP2
۱/۵	۱۰۰	۴/۹	SP3	SA4
۱/۴	۹۱	۳۸/۹	SP3	SA3
۱/۴	۵۳	۶۷/۵	SA3	SP3
۱/۱	۷۲	۱۳/۱	SP3	SA2
۰/۸	۵۲	۲۰/۴	SP3	SA1

#### قواعد انجمنی بین نمایه‌های SIAP و SPI در مقیاس زمانی فصلی بتفکیک فصول

با تفکیک نمایه‌ها در ۴ فصل بهار، تابستان، پاییز و زمستان، قواعد انجمنی بین نمایه‌های SIAP و SPI جستجو شد و ۳۲ قاعده با درصد اطمینان بالای ۵۰ درصد بدست آمد ولی با این تفکیک نیز هیچ تطابقی بین این دو نمایه مشخص نشد (جدول ۷).

#### قواعد انجمنی بین نمایه‌های SIAP و SPI در مقیاس زمانی نیم سالانه بتفکیک

با تفکیک نمایه‌ها در دو مقیاس زمانی شش ماه اول و دوم، ۱۸ قاعده با درصد اطمینان بالای ۵۰ درصد بدست آمد. با این تفکیک نیز هیچ تطابقی بین این دو نمایه مشخص نشد (جدول ۶).

### نتیجه گیری

سنجه اطمینان به عنوان معیار سنجش، مشخص شد که در طبقات خشکسالی تعریف شده برای هر یک از این دو نمایه، تطابق خوبی مشاهده نمی‌شود. از این رو می‌توان نتیجه گرفت با وجود اینکه دو نمایه SIAP و SPI از بارندگی استخراج می‌گردند و همبستگی نسبی بالایی با هم دارند ولی در بیشتر از پنجاه درصد موارد، رفتار متفاوتی در تحلیل خشکسالی‌ها بروز می‌دهند.

با محاسبه نمایه‌های هواشناسی خشکسالی SIAP و SPI در مقیاس‌های زمانی فصلی، نیمسالانه و سالانه و طبقه‌بندی آنها در پنج گروه که بیانگر وضعیت خشکسالی شدید، خشکسالی، نرمال، ترسالی و ترسالی شدید می‌باشند هم‌رخدادی هر یک از این گروه‌های خشکسالی بطور جداگانه در هر یک از مقیاس‌های زمانی مذکور با استفاده از قواعد انجمنی از روش داده‌کاوی بررسی شد و با انتخاب

جدول ۶- قواعد انجمنی بدست آمده بین نمایه‌های نیمسالانه SIAP و SPI

لیفت	درصد اطمینان	درصد حمایت	نیمسال آبی	پس آیند	پیش آیند
۴/۷	۱۰۰	۱/۵	دوم	SA1	SP1
۴/۷	۱۰۰	۲/۸	اول	SA1	SP1
۴/۸	۱۰۰	۴	دوم	SA5	SP5
۴/۸	۱۰۰	۴/۱	اول	SA5	SP5
۴/۸	۱۰۰	۳/۸	دوم	SA5	SP4
۴/۸	۱۰۰	۴/۸	اول	SA5	SP4
۱/۵	۱۰۰	۴/۳	دوم	SP3	SA2
۱/۵	۱۰۰	۵/۷	اول	SP3	SA2
۱/۵	۱۰۰	۶/۶	دوم	SP3	SA4
۱/۵	۱۰۰	۵/۴	اول	SP3	SA4
۴/۷	۱۰۰	۶/۹	دوم	SA1	SP2
۴/۷	۱۰۰	۵/۳	اول	SA1	SP2
۱/۵	۱۰۰	۱۷/۳	دوم	SP3	SA3
۱/۵	۱۰۰	۱۸/۶	اول	SP3	SA3
۴/۹	۶۰/۳	۱۱/۵	دوم	SP2	SA1
۱/۶	۵۶/۴	۳۳	اول	SA3	SP3
۴/۵	۵۴/۵	۹/۷	اول	SP2	SA1
۱/۴	۵۱/۳	۳۳/۸	دوم	SA3	SP3

جدول ۷- قواعد انجمنی بدست آمده بین نمایه‌های فصلی SIAP و SPI

لیفت	درصد اطمینان	درصد حمایت	فصل	پس آیند	پیش آیند
۵/۷	۱۰۰	۱/۸	تابستان	SA1	SP1
۵/۷	۱۰۰	۱	پائیز	SA1	SP1
۵/۷	۱۰۰	۱/۲	زمستان	SA1	SP1
۵/۷	۱۰۰	۰/۷	بهار	SA1	SP1
۶/۵	۱۰۰	۱/۹	پائیز	SA5	SP4
۶/۵	۱۰۰	۱/۵	زمستان	SA5	SP4
۶/۵	۱۰۰	۲	بهار	SA5	SP4
۶/۵	۱۰۰	۲/۳	پائیز	SA5	SP5
۶/۵	۱۰۰	۲	زمستان	SA5	SP5
۶/۵	۱۰۰	۱/۸	بهار	SA5	SP5
۱۰/۲	۱۰۰	۰/۴	تابستان	SP2	SA2
۱/۴	۱۰۰	۲/۵	پائیز	SP3	SA2

۱/۴	۱۰۰	۲/۲	زمستان	SP3	SA2
۱/۴	۱۰۰	۲/۱	بهار	SP3	SA2
۱/۴	۱۰۰	۲/۹	پائیز	SP3	SA4
۱/۴	۱۰۰	۳/۱	زمستان	SP3	SA4
۱/۴	۱۰۰	۳/۴	بهار	SP3	SA4
۵/۷	۱۰۰	۲/۸	پائیز	SA1	SP2
۵/۷	۱۰۰	۲/۹	زمستان	SA1	SP2
۵/۷	۱۰۰	۳/۲	بهار	SA1	SP2
۲	۱۰۰	۲۲/۴	تابستان	SA3	SP3
۱/۴	۱۰۰	۹/۶	پائیز	SP3	SA3
۱/۴	۱۰۰	۹/۷	زمستان	SP3	SA3
۱/۴	۱۰۰	۸/۷	بهار	SP3	SA3
۱/۳	۹۹/۷	۲۲/۵	تابستان	SP3	SA3
۱۷/۶	۸۲/۸	۲/۱	تابستان	SP1	SA1
۶/۲	۶۰/۳	۴/۶	پائیز	SP2	SA1
۶	۵۸/۸	۵	زمستان	SP2	SA1
۱/۱	۵۶/۲	۱۷/۱	پائیز	SA3	SP3
۱/۱	۵۵/۹	۱۷/۳	زمستان	SA3	SP3
۵/۷	۵۵/۷	۵/۸	بهار	SP2	SA1
۱	۵۰/۲	۱۷/۲	بهار	SA3	SP3

## منابع

- ۱- بذرافشان ج. ۱۳۸۱. مطالعه‌ی تطبیقی برخی نمایه‌های خشکسالی هواشناسی در چند نمونه‌ی اقلیمی ایران، پایان‌نامه‌ی کارشناسی ارشد، دانشگاه تهران، دانشکده کشاورزی کرج.
- ۲- چناری م. ۱۳۸۵. بررسی تغییرات چند نمایه‌ی مختلف خشکسالی، با استفاده از زنجیره‌ی مارکف (در نمونه‌های اقلیمی البرز جنوبی)، پایان‌نامه‌ی کارشناسی ارشد، دانشگاه تهران، دانشکده کشاورزی کرج.
- ۳- خلیلی ع. ۱۳۷۰. گزارش‌های حوضه‌ای هواشناسی طرح جامع آب کشور، وزارت نیرو، جاماب، تهران.
- ۴- مرید س. مقدسی م.، پایمرد ش. و قائمی ه. ۱۳۸۴. طراحی سیستم پایش خشکسالی استان تهران، پژوهشکده‌ی مهندسی آب دانشگاه تربیت مدرس.
- 5- Agrawal R., Imielinski T., and Swami A.N. 1993. Mining Association Rules between Sets of Items in Large Databases, in Peter Buneman and Sushil Jajodia (eds.), Proc. ACM SIGMOD International Conference on Management of Data, 207{216, 1993}.
- 6- Bhalme, H. N., and Mooley D. A. 1980. Large-scale drought/floods and monsoon circulation. Mon. Wea. Rev., 108, 1197-1211.
- 7- Brin S., Motwani R., and Silverstein C. 1997. Beyond market baskets: Generalizing association rules to correlations. In J. M. Peckman (ed.), Proc. ACM SIGMOD Conference on Management of Data (SIGMOD'97), pages 265 { 276, ACM, May 1997}.
- 8- Cabena, P. H., Stadler R., Verhees J., and Zanasi A. 1998. Discovering Data Mining: From Concept to Implementation, IBM, New Jersey, 195 pp.
- 9- Dhanya C. T., and Nagesh Kumar D. 2009. Data mining for evolution of association rules for droughts and floods in India using climate inputs, J. Geophys. Res., 114, D02102, doi:10.1029/2008JD010485.
- 10- Dracup J. A., Lee K. S and Paulson E. G. 1980. On the definition of droughts. Water Resource Research, 16: 297-302.
- 11- Gibbs W. J., and Maher J. V. 1967. Rainfall deciles as drought indicators. Bureau of Meteorology Bull. 48, Commonwealth of Australia, Melbourne, Australia.
- 12- Harms S.K., and Deogun J. S. 2004. Sequential association rule mining time lags. Journal of Intelligent Information Systems, Volume 22, Number 1, January 2004 , pp. 7-22(16).
- 13- Hayes M. 1996. Drought indexes. National Drought Mitigation Center, University of Nebraska-Lincoln, 7 pp.



- [Available from University of Nebraska Lincoln, 239LW Chase Hall, Lincoln, NE 68583.]
- 14- Hayes M. J., Svoboda M. D., Wilhite D. A., and Vanyarkho O. V. 1999. Monitoring the 1996 drought using the standardized precipitation index. *Bulletin of the American Meteorological Society*, 80 (3): 429-438.
  - 15- McKee T. B., Doesken N. J., and Kleist J. 1993. The relationship of drought frequency and duration to time scale. Preprints, Eighth Conf. on Applied Climatology, Anaheim, CA, Amer. Meteor. Soc., 179-184.
  - 16- McKee T. B., Doesken N. J., and Kleist J. 1995. Drought monitoring with multiple time scales. Preprints, Ninth Conf. on Applied Climatology, Dallas, TX, Amer. Meteor. Soc., 233-236.
  - 17- Mena, J. 1999. *Data Mining Your Website*. Woburn, Digital Press.
  - 18- Nitzche M. H. 2002. Drought Quantification and Preparedness in Brazil, the Example of Sao Paulo State, Working Paper, No 7, Londrina-PR, Brazil.
  - 19- Palmer W. C. 1965. Meteorological drought. U.S. Weather Bureau Teach Paper 45, 1-58.
  - 20- Palmer W. C. 1968. Keeping track of crop moisture conditions, nationwide: The new crop moisture index. *Weatherwise*, 21, 156-161.
  - 21- Rooy M., and van P. 1965. A rainfall anomaly index independent of time and space. *Notos*, 14, 43.
  - 22- Tadesse T. 2002. Identifying Drought and its association with climatic and Oceanic Parameters Using Data Mining Techniques. Nebraska, Graduate college University of Nebraska.
  - 23- Vanderberg H., Sogard P., and Motoroni S. 1999. *MineSetTM 3.0 Enterprise Edition Tutorial for Windows*, Doc. No. 007-4006-001, Silicon Graphics Inc.

Archive of SID

## Comparative Study of the Meteorological Drought Indices (Spi and Siap) Using Data Mining Method (Case Study of Kermanshah Province)

Kh. Ghorbani<sup>1\*</sup>- A. Khalili<sup>2</sup>- S.K. Alavipanah<sup>3</sup>- Gh. Nakhaezadeh<sup>4</sup>

### Abstract

Precipitation is an important variable which is used in definition of drought. Based on precipitation amounts, some indices are devised for drought monitoring including, the Standardized Precipitation Index (SPI) and the Standard Index of Annual Precipitation (SIAP). Each of these Drought indices are classified into some classes so that each class describes a given severity of drought. Investigation of simultaneously occurrence situation for two drought indices can be an appropriate measure to evaluate the agreement of indices. Association Rules in DATA MINING is used to find rules and patterns in database. In this paper, two drought indices, SPI and SIAP, were computed at 11 meteorological stations belong to Ministry of Energy in Kermanshah province. Then, based on Association Rules, simultaneously occurrence situation of drought severity classes for both indices were determined in seasonally, half yearly and yearly time scales. Results showed that there is not any good agreement between most of drought category from these indices (less than 50 percent) and Shows different behavior of drought.

**Keywords:** Data Mining, Drought, Association Rules, Kermanshah

Archive of SID

1,2 - PhD. Student of Agrometeorology and Professor, respectively, Irrigation & Reclamation Eng. Dept., Faculty of Agricultural Engineering & Technology., College of Agriculture and Natural Resources, University of Tehran, Karaj (\*- Corresponding author Email:ghorbani.khalil@yahoo.com)

3- Professor of Remote Sensing- Cartography Dept., Geography Faculty, University of Tehran

4- Professor of Economic, Institute of Statistic & Economic – Karlsruhe-Germany