

## پیش‌بینی الگوی خشکسالی هواشناسی با استفاده از دسته‌بندی متغیرهای هواشناسی

علیرضا نیکبخت شهبازی<sup>۱\*</sup>، بنفشه زهرایی<sup>۲</sup>

### چکیده

پیش‌بینی خشکسالی، شدت و زمان رخداد آن از دغدغه‌های بزرگ پژوهشگران هواشناسی و آب‌شناسی است. در این میان، اهمیت متغیرهای هواشناسی و جوی در تخمین مناسب شدت خشکسالی از موضوع‌های جالب برای محققین است. در این پژوهش با به‌کارگیری متغیرهای دمای هوا و ارتفاع زمین توانش در لایه‌های گوناگون جوی به بررسی پیوند این متغیرها در ۳۱ سال گذشته (۵۴-۱۳۸۶) با استفاده از نمایه‌ی بارش معیار شده (SPI) به‌عنوان شاخص خشکسالی اقلیمی در حوزه آبخیز سدهای تأمین‌کننده‌ی آب تهران پرداخته شده است. پس از برآورد میانگین بارش با استفاده از شاخص آماری اطلاعات متقابل نقطه-متغیرهای دما و ارتفاع زمین توانش در سطوح مختلف جو که پیوند نیرومندتری را با نمایه SPI در گستره‌های مورد مطالعه شده داشته‌اند، به‌عنوان متغیرهای مناسب برگزیده شدند. سپس در راستای پیش‌بینی طبقه خشکسالی اقلیمی SPI از شبیه‌ماشین بردار پشتیبان (SVM) به‌عنوان یکی از روش‌های یادگیری آماری استفاده شده است. نتایج گویای پاسخ مناسب رویکرد استفاده شده در پیش‌بینی وقوع خشکسالی است. همچنین، با استفاده از جستجوی مکانی ارتباط متغیرهای هواشناسی و مقادیر هدف، بهترین ترکیب این متغیرها نیز برگزیده شده است و در اکثر موقعیت‌ها پیش‌بینی با دقت مناسب به منظور کاربردهای واقعی دارای قابلیت به‌کارگیری است. دقت پیش‌بینی در فصول پاییز و زمستان بیش از سایر فصل‌ها بوده است.

**کلمات کلیدی:** خشکسالی هواشناسی، دسته‌بندی نمایه بارش معیار شده، ماشین بردار پشتیبان، حوضه‌ی آبخیز تهران.

<sup>۱</sup> گروه علوم و مهندسی آب، واحد اهواز، دانشگاه آزاد اسلامی، اهواز، ایران

Email: nikbakhta@gmail.com

<sup>۲</sup> دانشیار، دانشکده مهندسی عمران، پردیس دانشکده‌های فنی، دانشگاه تهران

## مقدمه

خشکسالی هواشناسی معرف دوره‌هایی از خشکسالی با بارش کمتر از یک حد آستانه است. نمایه‌های زیادی جهت پیش‌بینی خشکسالی وجود دارد که از آن جمله می‌توان به شاخص درصد از بهنجار<sup>۳</sup> که مفهوم اساسی آن تقسیم بارش واقعی بر بارش بهنجار است، شاخص دهک‌ها<sup>۴</sup>، شاخص رطوبتی گیاه<sup>۵</sup>، شاخص خشکسالی احیایی<sup>۶</sup>، شاخص پالمر<sup>۷</sup> و شاخص معیار شده-ی بارش<sup>۸</sup> اشاره نمود. در میان شاخص‌های خشکسالی، شاخص پالمر و SPI کاربرد وسیعی را در تعیین و درجه‌بندی خشکسالی زراعی و هواشناسی در اکثر کشورها داشته‌اند (موریرا و همکاران ۲۰۰۸). شاخص معیار شده‌ی بارش تنها به دلیل نیاز به آمار بارندگی ماهانه که در اکثر ایستگاه‌های کشور در دسترس است، قابلیت درجه‌بندی و ارزیابی خشکسالی در مقیاس‌های زمانی مختلف کوتاه مدت، میان مدت و دراز مدت، قدرت پایش نزدیک به زمان واقعی و بالاخره با توجه به کاستی‌هایی آماری موجود در کشور می‌تواند به عنوان شاخص مناسبی جهت ارزیابی خشکسالی در ایران به کار رود (ابراهیمی و همکاران، ۲۰۱۱).

نویسندگان متعددی نقش متغیرهای هواشناسی نظیر دما و فشار سطح دریا را برای خشکسالی و بارندگی فصلی بررسی کرده، و جهت پیش‌بینی شاخص SPI از شبیه‌های مختلفی نظیر زنجیره‌ی مارکوف و یا شبکه‌های عصبی مصنوعی استفاده نموده‌اند (زهراپی و روزبهانی، ۲۰۰۷؛ ابراهیمی و همکاران، ۲۰۱۱؛ موریرا و همکاران، ۲۰۰۶؛ باروس و بودن، ۲۰۰۸؛ کارآموز و همکاران، ۲۰۰۹؛ بلاینیش و همکاران، ۲۰۱۴؛ بنارکوسو و همکاران، ۲۰۱۵؛ راوینش و همکاران، ۲۰۱۷). از روشهای گوناگون و نوین آماری و محاسباتی از جمله شبکه‌های عصبی مصنوعی، خوشه‌بندی و زنجیره‌ی مارکوف به منظور تخمین خشکسالی بهره‌وری شده است (لوکاس و واسیلادوس، ۲۰۰۴؛ پائولو و همکاران، ۲۰۰۵؛ پای و هانگ، ۲۰۰۵؛ زهرایی و روزبهانی، ۲۰۰۷). ماشین‌های

بردار پشتیبان<sup>۹</sup> (SVM) از جمله سامانه‌های یادگیری آماری هستند که از یک فضای فرضی به نام فضای مشخصه<sup>۱۰</sup>، شامل توابع خطی با بعد زیاد، استفاده می‌کنند و می‌توانند به منظور پیش‌بینی (دسته‌بندی و همبستگی) استفاده شوند. این سامانه‌ها به وسیله‌ی الگوریتم یادگیری، که بر اساس تئوری بهینه‌سازی استوار است، آموزش داده می‌شوند. این تئوری که به وسیله‌ی وپنیک و همکاران (۱۹۹۵) به عنوان روشی قدرتمند و مهم معرفی گردید از پیش‌قدر یادگیری<sup>۱۱</sup>، که از تئوری یادگیری آماری استخراج شده، استفاده می‌کند. توانایی این روش به گونه‌ای است که از آغاز معرفی از آن در کاربردهای مختلفی استفاده شده است و تاکنون در زمینه‌های گوناگونی مانند آب شناسی و هواشناسی نیز از آن بهره‌گرفته شده است. دیبایک و همکاران (۲۰۰۱) قابلیت‌های این روش را، به خصوص در پیش‌بینی‌های آب شناسی، نشان دادند. آنها با کاربرد از این روش در زمینه‌ی دسته‌بندی داده‌های سنجش از دور و شبیه‌سازی بارندگی و رواناب، و مقایسه آن با روش‌های شبکه‌ی عصبی مصنوعی به نتایج خوبی برای پیش‌بینی رسیده‌اند. تریپاتی و همکاران (۲۰۰۶) از روش SVM برای برآورد آماری بارندگی ماهانه در هند استفاده کردند. ایشان از روش SVM برای دسته‌بندی فصل‌های خشک و نمناک با کاربرد داده‌های ماهواره‌ای استفاده کردند. نتایج نشان داده‌اند که شبیه SVM می‌تواند به عنوان گزینه برتر نسبت به سایر روش‌های معمول برای پیش‌بینی بارش بروش آماری به کار گرفته شود. وانگ و همکاران (۲۰۰۸) از این شبیه بمنظور پیش‌بینی میزان آلاینده‌های هوا در یک گروه زمانی بر اساس پایش آلاینده‌های هوای موجود در هنگ کنگ استفاده نمودند. چن و همکاران (۲۰۰۹) به منظور دسته‌بندی روزهای خشک و روزهای بارانی از دسته‌بندی‌کننده‌ی ماشین بردار پشتیبان استفاده نموده، سپس با استفاده از همبستگی میزان بارش را در روزهای بارانی پیش‌بینی کردند. بهزاد و همکاران (۲۰۰۹) نیز در پیش‌بینی رواناب با زمان پایش دید یک روزه در رودخانه بختیاری ایران از SVM استفاده نمودند. لین و همکاران (۲۰۰۶) از SVM برای پیش‌بینی دبی دراز مدت استفاده

<sup>3</sup> percent of normal

<sup>4</sup> deciles

<sup>5</sup> crop moisture index

<sup>6</sup> reclamation drought index

<sup>7</sup> palmer drought severity index

<sup>8</sup> standardized precipitation index

<sup>9</sup> support vector machine

<sup>10</sup> feature space

<sup>11</sup> learning bias

افزایش شاخص NAO کاهش می یابد. موریرا (۲۰۱۵) یک شبیه خطی-لگاریتمی را برای پیش بینی طبقه شاخص خشکسالی SPI ماهانه در پرتغال برای فصل های خشک و مرطوب در پرتغال ارائه نمود. نتایج پیش بینی خشکسالی ۶ ماهه و ۱۲ ماهه با ۷۰ و ۸۰ درصد دقت پیش بینی طبقه خشکسالی انجام گرفت. راوینش و همکاران (۲۰۱۷) برای پیش بینی خشکسالی با شاخص SPI از سه روش همبستگی چند متغیره ی سازگار، ماشین بردار پشتیبان و شبیه درخت تصمیم در شرق استرالیا استفاده نمودند. داده های بارندگی و دمای سطح دریا به عنوان پیش بینی کننده به کار رفتند. نتایج پیش بینی ها نشان دادند که شبیه همبستگی چند متغیره سازگار و درخت تصمیم نتایج بهتری را با توجه به معیار RMSE, MAE و ضریب همبستگی از خود نشان می دهند. همچنین نتایج نشان دادند که دوره ای بودن خشکسالی در پیش بینی های فصلی اهمیت بالایی دارد و می بایست بررسی شود. رضیئی (۲۰۱۶) با استفاده از شبیه های گروه زمانی و زنجیره ی مارکف مرتبه ی اول امکان پیش بینی پدیده ی خشکسالی در بخش خشک و نیمه خشک مرکز و شرق ایران بررسی کرد. برای این منظور، از داده های بارش ۶۹ ایستگاه سینوپتیک و اقلیم شناسی منطقه مطالعه شده در دوره ی ۱۹۷۵-۲۰۰۵ نمایه SPI در مقیاس زمانی ۳، ۶ و ۱۲ ماهه بهره وری شد. نتایج نشان دادند که شبیه ARMA با گروه های منطقه ای SPI سه و شش ماهه و SARIMA با گروه های منطقه ای SPI دوازده ماهه بهترین برازش را دارد. نتایج نشان دادند که شبیه های انتخاب شده حداکثر می توانند تا سه ماه آینده را با دقت مناسب پیش بینی کنند اما از دقت پیش بینی بیش از سه ماه به سرعت کاسته می شود.

شبیه‌هایی که تاکنون برای پیش‌بینی خشکسالی توسعه داده شده‌اند عمدتاً برای گروه های زمانی با ساختار خودهمبستگی و خودهمبستگی جزئی هستند، و نیز در تعداد بازه های زمانی قابل پیش بینی در آینده نیز دارای محدود نیستند. با توجه به اینکه خشکسالی عمدتاً یک فرآیند میان مدت بوده و از چند ماه تا بیش از یک سال می تواند تداوم داشته باشد، کاربرد شبیه‌های گروه زمانی در پیش بینی خشکسالی عملاً در بسیاری از مناطق امکان پذیر نبوده و منجر به پیش بینی های قابل

نموده و با روش های شبیه شبکه های عصبی مصنوعی و آریمای مقایسه کردند. نتایج حاصل از پیش بینی کننده SVM دقت بسیار بالاتری را در پیش بینی دبی نسبت به سایر روش های استفاده شده داشته است. بلائینیش و همکاران (۲۰۱۴) از چهار روش داده کاوی جهت پیش بینی بلند (۶ تا ۱۲ ماهه) مدت شاخص خشکسالی SPI (۱۲ ماهه و ۲۴ ماهه) در حوضه ی رود آواش در اتیوپی استفاده کردند. روشهای استفاده شده شامل شبیه ARIMA و شبیه شبکه های عصبی مصنوعی و شبیه ماشین بردار پشتیبان بودند که درباره ی این دو، تبدیل موجی برای ورودی های شبکه عصبی و ماشین بردار به کار رفت. نتایج نشان دادند شبیه ترکیبی موجی با شبکه های عصبی بهترین برآورد را برای SPI در حوضه مطالعه شده دارد. موسا خدر (۲۰۱۵) شبیه های مختلف زنجیره ی مارکف را پنهان برای پیش بینی کوتاه مدت شاخص خشکسالی SPI با توجه به بارش در ۲۲ ایستگاه در حوضه ی آبخیز نیل آبی واقع در اتیوپی بررسی کرد. نتایج نشان دادند که با توجه به معیارهای مختلف تفاوت بین مقادیر پیش بینی شده و مشاهده شده بسیار کم بوده و می توان در حوضه ی مزبور از این شبیه برای پیش هشدار زمان خشکسالی استفاده کرد. ریبیرو و پایرز (۲۰۱۶) قابلیت پیش بینی خشکسالی را در پرتغال با استفاده از روش های آماری-پویایی بررسی نمودند. هدف آنها توسعه ی شبیه پیش بینی شاخص خشکسالی SPI سه ماهه با زمان پیش دید ۶ ماهه بر اساس شبیه پیش بینی بارش UKMO بود. قابلیت پیش بینی بر اساس ضریب همبستگی به دست آمد. نتایج برای ۴ فصل بدست آمدند که نشان دادند شبیه مزبور برای فصل زمستان قابلیت پیش بینی بالاتری را دارد. نتایج منجر به تهیه راهنمایی برای پیش بینی خشکسالی در پرتغال شد که قابل استفاده برای کشاورزان نیز هست. بنارکوسو و همکاران (۲۰۱۵) پیش بینی احتمالاتی کوتاه مدت تغییر طبقه شاخص خشکسالی SPI را با استفاده از اطلاعات شاخص NAO در ایتالیا بررسی نمودند. با توجه به همبستگی معکوس بین شاخص SPI و NAO در منطقه-ی سیسیل، ایتالیا، در دهه های اخیر این دو شاخص در ۴ دسته تقسیم بندی شدند. نتایج نشان دادند که احتمال تغییر دسته خشکسالی بستگی به شاخص NAO دارد. نتایج نشان دادند که احتمال انتقال به طبقه مرطوب با

مطالعه شده انتخاب شده اند. بر اساس تحقیقات گذشته، سامانه‌های مؤثر بر اقلیم مناطق مرکزی و غرب ایران را می‌توان به چهار گروه عمده تقسیم کرد: مرکز پر فشار سیبری (Siberia, 50° N, 90° E)، مرکز پر فشار آزور (Azores, 40° N, 25° W)، مرکز کم‌فشار مدیترانه یا چرخند زای مدیترانه‌ای (East of Mediterranean, 35° N, 50° E)، مرکز کم‌فشار سودان (Sudan, 10° N, 30° E) (کارآموز و همکاران ۲۰۰۹). در شکل (۳) کلیه‌ی مناطقی که بر بارندگی‌های ایران تأثیر می‌گذارند نشان داده شده اند. شکل (۴) محدوده جغرافیایی در نظر گرفته شده برای بررسی ارتباط اطلاعات هواشناسی با بارش فصلی در حوضه‌های آبخیز مطالعه شده را نشان می‌دهد. کلیه‌ی مناطقی که به نحوی در شکل‌گیری سامانه‌های بارش مؤثر هستند در محدوده مورد بررسی در نظر گرفته شده است. با توجه به وسیع بودن محدوده مطالعه گردیده، منطقه نمایش داده شده در شکل (۳) به مناطق جغرافیایی با ابعاد پنج درجه عرض جغرافیایی در پنج درجه طول جغرافیایی دسته‌بندی شده‌اند. تعداد منطقه‌های بررسی شده در این تحقیق ۱۸۰ است.

داده‌های هواشناسی استفاده شده در این تحقیق، که برای هر یک از منطقه‌های جغرافیایی یادشده طول دوره‌ی تاریخی ۳۱ ساله استفاده شده است، شامل آمار ذیل است: دمای هوا در سطح زمین، دمای هوا در سطوح ۳۰۰، ۵۰۰، ۷۰۰ و ۸۵۰ میلی بار و ارتفاع معادل فشار<sup>۱۲</sup> در سطوح ۳۰۰، ۵۰۰، ۷۰۰ و ۸۵۰ میلی بار. در جدول (۱) متغیرهای مورد استفاده و علامت مشخصه نظیر هر یک نشان داده شده است.

به منظور بررسی ارتباط بارش متوسط و شاخص SPI با متغیرهای هواشناسی فوق‌الذکر، این متغیرها در ماه‌های اکتبر تا آوریل که ۶ ماهه‌ی اول سال آبی را پوشش می‌دهند در نظر گرفته شده‌اند. بدین ترتیب در صورت تشخیص ارتباط آماری مناسب، این متغیرها قابل استفاده به عنوان پیش‌بینی‌کننده<sup>۱۳</sup> با زمان پیش‌دید<sup>۱۴</sup> مناسب برای پیش‌بینی دوره‌های کم بارش خواهند بود. با استفاده از ۱۱۳۴۰ گروه زمانی تحلیل زمانی- مکانی

کاربرد نخواهد شد. یکی از عمده‌ترین متغیرهای مهم که تا کنون کمتر به آن توجه شده است در نظر نگرفتن الگوهای تغییرات متغیرهای مختلف هواشناسی در سطوح مختلف جو در مسیر حرکت سامانه‌های باران زا بوده است.

هدف از این پژوهش پیش‌بینی زمان وقوع و تداوم خشکسالی بر اساس شاخص SPI و اطلاعات هواشناسی موجود و ثبت شده در ایستگاه‌های زمینی و ماهواره‌ای با استفاده از ماشین بردار پشتیبان است. در این پژوهش الگوهای تغییرات متغیرهای دما و ارتفاع معادل فشار در سطوح مختلف جوی در پهنه جغرافیایی مؤثر بر بارش منطقه‌ی تحقیق بررسی و نقش این متغیرها در پیش‌بینی دراز مدت بارندگی (۱ تا ۱۲ ماهه) در محدوده‌ی پژوهش بررسی شده است. در انتها بازه‌ی زمانی مناسب به منظور برقراری ارتباط آماری بین متغیرهای هواشناسی و متغیر پیش‌بینی شونده (SPI) تعیین گردید. ساختار به کار گرفته شده در این پژوهش، حجم قابل توجه داده‌های مکانی در لایه‌های متعدد جو را پردازش کرده، و از این نتایج در رسیدن به بهترین ساختار ورودی‌های شبیه‌بهره‌جسته است.

## مواد و روش‌ها

### الف: منطقه مطالعه شده و داده‌های به کار رفته

در ابتدا محدوده مطالعاتی معرفی شده و در ادامه روش به کار گرفته شده شرح داده خواهد شد. محدوده مورد مطالعه، حوضه‌های آبخیز سدهای کرج، ماملو و طالقان است آب تهران را تأمین کننده می‌کند. اطلاعات ۳۱ ایستگاه باران سنجی در این محدوده برای تحلیل بارش و خشکسالی هواشناسی مورد استفاده قرار گرفت. بارندگی در تهران ۳۴۰ میلی متر و حجم آب تجدید شونده ۴/۶ میلیارد متر مکعب و حجم ریزش‌های جوی ۶/۵۶ میلیارد متر مکعب در سال است. شکل (۱) حوضه‌های مربوطه و موقعیت ایستگاه‌های استفاده شده را نشان می‌دهد. طول آماری ایستگاه‌های مزبور (۳۲ ایستگاه بارانسنجی) از سال ۱۳۵۴-۵۵ الی ۱۳۸۴-۸۵ است. این داده‌ها از مؤسسه تحقیقات آب وزارت نیرو، بخش آمار و اطلاعات دریافت شده است.

متغیرهای هواشناسی استفاده شده در این تحقیق از محدوده‌ی مؤثر بر سامانه‌های بارش زا در محدوده مورد

<sup>12</sup> geopotential Height

<sup>13</sup> predictor

<sup>14</sup> lead-time

جدول ۱- شماره‌ی مشخص کننده متغیرهای هواشناسی.

شماره	متغیرهای هواشناسی	علامت
۱	دما در سطح زمین	AT1000
۲	دما در سطح ۳۰۰ میلی بار	AT300
۳	دما در سطح ۵۰۰ میلی بار	AT500
۴	دما در سطح ۷۰۰ میلی بار	AT700
۵	دما در سطح ۸۵۰ میلی بار	AT850
۶	ارتفاع معادل فشار در سطح ۳۰۰ میلی بار	GH300
۷	ارتفاع معادل فشار در سطح ۵۰۰ میلی بار	GH500
۸	ارتفاع معادل فشار در سطح ۷۰۰ میلی بار	GH700
۹	ارتفاع معادل فشار در سطح ۸۵۰ میلی بار	GH850

جدول ۲- مقادیر SPI و وضعیت اقلیمی متناظر با آن

(مک کی و همکاران، ۱۹۹۳).

نمایه	حدود SPI	اقلیم
VW	+۲	فوق العاده مرطوب
W	۱/۹۹ تا ۱/۵	خیلی مرطوب
NW	۱ تا ۱/۴۹	مرطوب متوسط
N	تا ۰/۹۹	نزدیک به بهنجار
ND	-۱/۴۹ تا -۱	خشک ملایم
D	-۱/۹۹ تا	خیلی خشک
VD	-۲ و کمتر	فوق العاده خشک

## محاسبه متوسط بارش حوضه ای و نمایه بارش

### معیار شده

بارش ماهانه‌ی حوضه ای با استفاده از داده های بارش ایستگاهی کل سه حوضه محاسبه شده است. به منظور تخمین مکانی، شبیه معکوس فاصله‌ی وزندار<sup>۱۵</sup> به دلیل روانی و سهولت استفاده و نیز تناسب آن با کوهستانی بودن بخش قابل توجهی از محدوده‌ی مطالعه شده با توجه به ارتفاع متوسط ایستگاه های هواشناسی، انتخاب شده است و اطلاعات بدست آمده به مقادیر SPI تبدیل شده اند.

جهت بررسی تغییرات بارندگی و توسعه‌ی شبیه پیش‌بینی برای محدوده‌ی مطالعه شده که به وسیله ی سیستم برنامه‌ریزی و مدیریت منابع آب قابل بهره‌برداری باشد، شش نمایشنامه به شرح زیر تعریف شده اند:

(۱) مجموع بارش فصل بهار ((SPI(S)، (۲) مجموع بارش فصل پائیز ((SPI(A)، (۳) مجموع بارش فصل زمستان ((SPI(W)، (۴) مجموع بارش فصل‌های پائیز + زمستان ((SPI(AW)، (۵) مجموع بارش فصل‌های زمستان

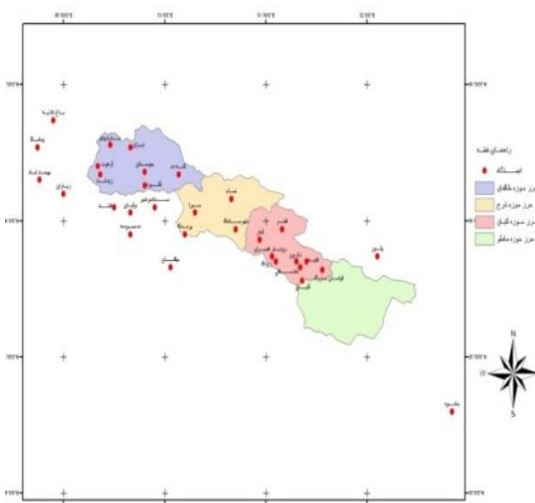
متغیرهای هواشناسی انجام می‌گیرد: ۹ متغیر هواشناسی  $\times 180$  منطقه‌ی جغرافیایی  $\times 7$  ماه از سال = ۱۱۳۴۰.

### ب: روش تحقیق

#### ۱- بارش معیار شده (SPI)

یکی از متغیرهای معتبر در خصوص تعیین رژیم بارش، نمایه‌ی بارش معیار شده است. مک کی و همکاران (۱۹۹۳) این متغیر را به منظور پایش هواشناختی بارش ارائه نمودند. این شاخص معرف شدت بارش (و یا خشکسالی ناشی از آن) با استناد به توزیع زمانی و البته بلند مدت آن است، همچنین می‌تواند معرف میزان و کیفیت وقوع یک حادثه خشکسالی یا یک ترسالی در مقیاس زمانی مشخص و در منطقه‌ای معین باشد. این شاخص معرف مقدار عددی انحراف معیاری است که یک پدیده‌ی بارشی از میانگین دارد. محاسبه‌ی شاخص مطلوب SPI منوط به گذراندن یک تابع چگالی احتمال گاما، بر توزیع تناوب بارش کل برای یک ایستگاه مشخص است.

با توجه به اینکه در مقالات نحوه محاسبه شاخص SPI ارائه شده اند، خوانندگان برای مطالعه بیشتر می‌توانند به مراجع مک کی و همکاران (۱۹۹۳) و موریرا (۲۰۰۸، ۲۰۰۶) مراجعه کنند. در جدول (۲) اقلیم‌های متناظر با حدود شاخص‌های SPI با توجه به پیشنهاد مک کی (۱۹۹۳) آمده اند.



شکل ۱- موقعیت ایستگاه‌های استفاده شده نسبت

به سه حوضه‌ی بررسی گردیده

<sup>15</sup>inverse weighted distance (IWD)

$I(x, y)$  شاخص مشخص مطلب و

به ترتیب توابع چگالی احتمال متغیرهای  $x$  و  $y$  و تابع توزیع احتمالاتی متغیرهای  $x$  و  $y$  است. عبارت فوق بیانگر این است که دانستن هر یک از متغیرهای  $x$  و  $y$  چه میزان از عدم قطعیت دانستن دیگری می‌کاهد (وینیک و چروونیک، ۱۹۹۱).

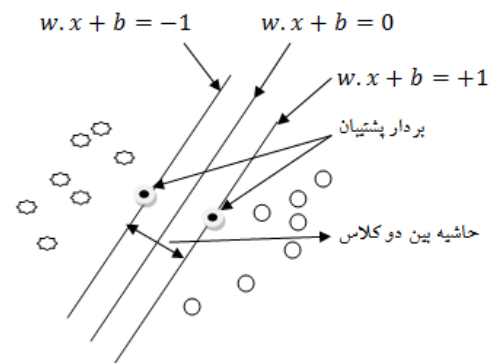
شاخص اطلاعات متقابل برای متغیرهای هواشناسی و نمایه بارش معیار شده برای هر نمایشنامه محاسبه شد. شاخص اطلاعات متقابل به دست آمده از کل متغیرهای هواشناسی به ترتیب نزولی مرتب شده و شماره‌ی منطقه‌هایی که دارای بیشترین شاخص اطلاعات متقابل است بدست آمد. شاخص‌های اطلاعات متقابل بزرگتر جهت تعیین نقطه متغیرهای پیش‌بینی‌کننده انتخاب شدند. با توجه به اینکه هر نقطه‌ی متغیر با یک نمایشنامه بیشترین ارتباط را داشته است، آن نقطه متغیرها به عنوان پیش‌بینی‌کننده همان نمایشنامه انتخاب می‌شوند.

### ۳- افزایش اطلاعات به وسیله ماشین‌های بردار

#### پشتیبان<sup>۱۸</sup> و درخت تصمیم

ماشین‌های بردار پشتیبان از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی را نسبت به روش‌های قدیمی‌تر برای دسته‌بندی و افزایش اطلاعات نشان داده است. ماشین بردار پشتیبان یا SVM یک افزازکننده دودویی است که دو (یا چند) کلاس را با استفاده از یک مرز خطی در فضای متغیرهای ورودی از یکدیگر جدا می‌کند. در این روش با استفاده از تمامی گستره اطلاعات و یک الگوریتم بهینه‌ساز، نمونه‌هایی که مرزهای کلاس مطلوب را تشکیل می‌دهند، به دست می‌آید که به این نمونه‌ها بردارهای پشتیبان می‌گویند. تعدادی از نقاط آموزشی که کمترین فاصله را تا مرز تصمیم‌گیری دارند می‌توانند به عنوان زیرمجموعه‌ای برای تعریف مرزهای تصمیم‌گیری و به عنوان بردار پشتیبان در نظر گرفته شوند. در شکل (۲) دو کلاس و بردارهای پشتیبان مربوط به آنها نشان داده شده اند. فرض

+ بهار (SPI(WS))، ۶ مجموع بارش فصل‌های پائیز + زمستان + بهار (SPI(AWS)).



شکل ۲- نحوه‌ی دسته‌بندی در ماشین بردار پشتیبان.

### ۲- انتخاب متغیرهای پیش‌بینی‌کننده

انتخاب متغیرهای مؤثر ورودی شبیه در فرایند شبیه‌سازی حائز اهمیت بالایی است. ابزارهای مختلفی در تحلیل‌های آماری برای این منظور وجود دارند. یکی از معیارهای مهم که در ادبیات انتخاب متغیر اخیراً کاربرد گسترده‌ای یافته، شاخص اطلاعات متقابل<sup>۱۶</sup> است. به کاربرد این معیار برای انتخاب تأخیرهای مناسب از یک پدیده زمانمند به منظور شبیه‌سازی در متون مورد تأکید گردیده، و عملکرد آن تا حدود قابل توجهی در ادبیات فنی موضوع تبیین شده است (پنگ و همکاران، ۲۰۰۵). این الگوریتم براساس عدم قطعیت در سری‌های گسسته پایه‌ریزی شده است. این شاخص آن گونه که در ضریب همبستگی تشابه مورد ارزیابی قرار می‌گیرد، کاهش عدم قطعیت در شناسایی یک متغیر  $x$  را با داشتن متغیر دیگر  $y$  بیان می‌کند. این نمایه شاخص اطلاعات متقابل نام دارد. روابط زیر، شاخص اطلاعات متقابل را در صورتی که دنباله‌های  $x$  و  $y$  گسسته و یا پیوسته باشند، بیان می‌دارد:

$$I(x, y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left( \frac{p(x, y)}{p_1(x)p_2(y)} \right) \quad (1)$$

$$I(x, y) = \iint_{YX} p(x, y) \log \left( \frac{p(x, y)}{p_1(x)p_2(y)} \right) dx dy \quad (2)$$

<sup>17</sup> joint probability distribution function

<sup>18</sup> support vector classification

<sup>16</sup> mutual information index

حل کردن مسئله بهینه‌سازی بالا کار مشکلی است. برای ساده‌تر کردن آن با استفاده از روش ضرایب نامعین لاگرانژ این مسئله بهینه‌سازی را می‌توان به فرم زیر تبدیل کرد که  $\alpha_i$  ها ضرایب لاگرانژ هستند:

$$(7) \quad \begin{cases} \max(-\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i y_i (x_i x_j) \alpha_j y_j + \sum_{i=1}^l \alpha_i) \\ \alpha_i \geq 0, i = 1, 2, \dots, L \\ \sum_{i=1}^l \alpha_i y_i = 0 \end{cases}$$

پس از حل مسئله بهینه‌سازی بالا و یافتن ضرایب لاگرانژ،  $w$  با استفاده از رابطه‌ی زیر محاسبه می‌شود:

$$w = \sum_{i=1}^l \alpha_i x_i y_i \quad (8)$$

$\alpha_i$  بردارهای پشتیبان بزرگتر از صفر، و  $\alpha_i$  نقاط دیگر صفر خواهند بود. بنابراین با توجه به معادله‌ی (۸) و صفر بودن  $\alpha_i$  مربوط به  $x_i$  هایی که بردار پشتیبان نیستند، برای به دست آوردن مرز تصمیم‌گیری فقط نیاز به تعداد معدودی از نقاط آموزشی است که همان بردارهای پشتیبان بوده است و همه‌ی آنها لازم نیستند. در نتیجه، دسته‌بندی وضعیت خشکسالی با استفاده از ماشین بردار پشتیبان به تعداد معدودی نقطه آموزشی نیاز خواهد داشت. پس از یافتن  $w$  با استفاده از رابطه‌ی زیر مقدار  $b$  به ازاء بردارهای پشتیبان مختلف محاسبه شده و  $b$  نهایی با میانگین‌گیری از  $b$  های حاصل به دست می‌آید:

$$\alpha_i [y_i (w \cdot x_i + b) - 1] = 0, \quad i = 1, \dots, L \quad (9)$$

طبقه‌بندی‌کننده‌ی نهایی نیز از رابطه‌ی (۱۰) به دست می‌آید:

$$f(x, w, b) = \text{sgn}(w \cdot x + b) \quad (10)$$

الگوریتم بالا مرز خطی دو کلاس کاملاً جدا از هم را به دست می‌دهد، اما در حالتی که کلاس‌ها با هم همپوشانی داشته باشند جدا کردن کلاس‌ها به وسیله مرز تصمیم‌گیری خطی همواره با خطا همراه خواهد بود. برای حل این مشکل می‌توان ابتدا داده‌ها را از فضای اولیه با استفاده از یک تبدیل غیرخطی  $\phi$  به فضایی با ابعاد بیشتر منتقل کرد. سپس در فضای جدید با استفاده از

کنید داده‌ها از نقطه آموزشی  $x_i, i=1, \dots, L$  نقطه‌ی آموزشی باشند که  $x_i$  یک بردار است. این دو کلاس با  $y_i = \pm 1$  برچسب زده می‌شوند. برای محاسبه مرز تصمیم‌گیری دو کلاس کاملاً جدا از هم، از روش حاشیه‌ی بهینه استفاده می‌شود. در این روش مرز خطی بین دو کلاس به گونه‌ای محاسبه می‌گردد که تمام نمونه‌های کلاس +۱ در یک طرف مرز و تمام نمونه‌های کلاس -۱ در طرف دیگر مرز واقع شوند؛ همچنین مرز تصمیم‌گیری به گونه‌ای باشد که فاصله نزدیکترین نمونه‌های آموزشی هر دو کلاس از یکدیگر در راستای عمود بر مرز تصمیم‌گیری تا جایی که ممکن است حداکثر شود. یک مرز تصمیم‌گیری خطی را در حالت کلی می‌توان به صورت زیر نوشت (ویتن و همکاران، ۲۰۰۵):

$$f(x) = \langle w, x \rangle + b, \quad (3)$$

به طوری که  $w, x \in X, b \in R$ . به  $x$  یک نقطه روی مرز تصمیم‌گیری و  $w$  یک بردار  $n$  بعدی عمود بر مرز تصمیم‌گیری است.  $\langle w, x \rangle$  بیانگر ضرب داخلی دو بردار و  $b/w$  فاصله‌ی مبدأ تا مرز تصمیم‌گیری است. از آنجا که با ضرب یک ثابت در دو طرف (۱) باز هم تساوی برقرار خواهد بود، برای تعریف یکتای مقدار  $b$  و  $w$  شرایط زیر برای آنها اعمال می‌شود:

$$\begin{cases} y_i (w \cdot x_i + b) = 1 & \Rightarrow \text{اگر } x_i \text{ یک بردار پشتیبان باشد} \\ y_i (w \cdot x_i + b) > 1 & \Rightarrow \text{اگر } x_i \text{ یک بردار پشتیبان نباشد} \end{cases} \quad (4)$$

اولین مرحله برای محاسبه مرز تصمیم‌گیری بهینه، پیدا کردن نزدیکترین نمونه‌های آموزشی دو کلاس است. در مرحله بعد فاصله آن نقاط از هم در راستای عمود بر مرزهایی که دو کلاس را به طور کامل جدا می‌کنند محاسبه می‌شود. مرز تصمیم‌گیری بهینه مرزی است که حداکثر حاشیه را داشته باشد. مرز تصمیم‌گیری بهینه با حل مسئله‌ی بهینه‌سازی زیر محاسبه می‌شود:

$$\min y_i \frac{(w \cdot x_i + b)}{|w|}, \quad i = 1, 2, \dots, L \quad (5)$$

با توجه به رابطه‌ی (۵) و انجام یک سری عملیات ریاضی، رابطه‌ی بالا به رابطه زیر تبدیل می‌شود:

$$\min_{w, b} \frac{1}{2} |w|^2, y_i (w \cdot x_i + b) - 1 \geq 0, \quad i = 1, \dots, L \quad (6)$$

با توجه به شاخص SPI محاسبه شده برای هر فصل - حوضه دسته‌بندی بر اساس جدول مک‌کی و همکاران صورت می‌گیرد. دوره‌هایی که نمایه بارش معیار شده در حالت بهنجار (بین -۱ و +۱) قرار دارد را صفر، و دوره‌های کم بارش (کمتر از -۱) را -۱ و دوره‌های پر بارش (بیشتر از +۱) را +۱ برای دسته‌بندی و پیش‌بینی به وسیله ی شبیه SVC در نظر گرفته شده اند. پس از آماده سازی داده‌ها برای شبیه دسته‌بندی، جهت بهینه‌سازی، متغیرهای شبیه داده‌های آموزشی (۷۰ درصد داده‌ها) و واسنجی و صحت سنجی (۳۰ درصد داده‌ها) به ساختار افزاز کننده ارائه می‌شوند. در این تحقیق از نرم افزار LIBSVM برای دسته‌بندی شبیه ماشین بردار پشتیبان استفاده شده است. در این نرم افزار از توابع مختلف کرنل اعم از خطی، چند جمله‌ای، سیگموئید و شعاعی استفاده شد و با تکرار اجرای برنامه و انتخاب بازه‌های مناسب برای متغیر شبیه و متغیرهای تابع هسته، نتایج شامل نوع کرنل، متغیرها به دست آمدند و مناسب ترین تابع با توجه به معیار ارزیابی عملکرد شبیه یعنی میانگین مربعات خطا<sup>۱۹</sup> تعیین شد:

$$MSE = \frac{\sum_1^n (X_p - X_o)^2}{n} \quad (13)$$

$X_p$  مقادیر پیش بینی شده و  $X_o$  مقادیر مشاهده شده و  $n$  تعداد داده‌ها است.

### درخت تصمیم

درخت تصمیم یکی از ابزارهای نوین داده کاوی است. به کارگیری این روش به منظور تخمین و دسته بندی در آب شناسی دارای کاربردهای متعددی در ادبیات فنی است (شنسی و همکاران ۲۰۰۸). دسته‌بندی با کاربرد روش درخت تصمیم عبارت است از یادگیری درخت تصمیم از نمونه‌های آموزش دهنده‌ی دسته‌بندی شده. درخت تصمیم یک ساختار درختی شبیه نمودار گردش کار است که هر گره آن معرف یک آزمون برای یک مقدار صفت است، هر شاخه نمایانگر یک برآمد از آزمون، و برگ‌های درخت نشان دهنده دسته‌ها هستند. از مزایای اصلی درخت تصمیم می توان سادگی ساختار و بازدهی بالای آن را در دسته بندی ذکر کرد. برای دسته بندی با کاربرد روش درخت تصمیم در این تحقیق از نرم

معادلات قبلی و جایگزینی  $x_i$  با  $\phi(x_i)$  و در نظر گرفتن مقداری خطا مرز تصمیم‌گیری بهینه محاسبه می‌گردد. با توجه به این امر و رابطه‌ی (۷) در حالت مزبور یافتن مرز تصمیم‌گیری بهینه به حل مسئله بهینه‌سازی (۱۱) تبدیل می‌شود:

$$(11)$$

$$\begin{cases} \text{Max} \left[ -\frac{1}{2} \sum_{i=1}^L \sum_{j=1}^L \alpha_i y_i (\phi(x_i) \phi(x_j)) \alpha_j y_j + \sum_{i=1}^L \alpha_i \right] \\ C \geq \alpha_i \geq 0, i = 1, 2, \dots, L \\ \sum_{i=1}^L \alpha_i y_i = 0 \end{cases}$$

در این مسئله بهینه‌سازی  $C$  متغیر شبیه و یک عدد ثابت است. اگر  $C$  به بینهایت میل کند، بهینه‌سازی به سمت یافتن یک مرز بهینه برای کلاس‌هایی با تداخل بسیار زیاد پیش می‌رود. از طرفی، اگر  $C$  به صفر میل کند، مسئله بهینه‌سازی به سمت یافتن مرز بهینه‌ی جداکننده دو طبقه، که تداخل بسیار کمی دارند، هدایت خواهد شد. در رابطه‌ی (۱۱) معمولاً به جای استفاده از  $\phi$  از یک تابع هسته که به صورت زیر تعریف می‌گردد، استفاده می‌شود:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) T \phi(\mathbf{x}_j) \quad (12)$$

پس از تعیین یک  $k(x_i, x_j)$  مناسب، در رابطه‌ی

(۱۲) و به جای  $\phi(x_i) \phi(x_j)$ ، تابع  $k(x_i, x_j)$  قرار داده شده و مسئله بهینه‌سازی حل می‌شود.  $k(x_i, x_j)$  در واقع یک تابع در فضای اولیه است که برابر با ضرب داخلی دو بردار در فضای ویژگی است. برای معادل بودن تابع  $k(x_i, x_j)$  با ضرب داخلی دو بردار در فضای ویژگی، باید هسته یک تابع معین مثبت متقارن بوده و شرط مرکز صدق کند. برخی از مهمترین توابع هسته، که در این شرط صدق می‌کنند، عبارتند از توابع هسته خطی، شعاعی و چند جمله‌ای. در این تحقیق از توابع هسته خطی (رابطه‌ی ۱۳) و شعاعی (رابطه‌ی ۱۴) استفاده شده است:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle \quad (13)$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\gamma}\right) \quad (14)$$

<sup>19</sup> mean square error



### جدول ۳- متغیرهای هواشناسی انتخاب شده برای نمایشنامه های مختلف.

SPI	حوضه ی آبخیز	متغیرهای پیش بینی کننده مورد استفاده با توجه به شاخص MI
بهار	ماملو	AT1000,AT300,AT500, AT850,GH300,GH500
	طالقان	GH500,GH700
پائیز	ماملو	AT300,GH500,GH850
زمستان	کرج	AT500,GH500,GH850
پائیز+زمستان	طالقان	AT1000,AT500,AT700, AT850,GH300,GH700,GH850
	طالقان	GH500,GH700,GH850
پائیز+زمستان+ بهار	کرج	AT300,AT500,AT850, GH700,GH850

### جدول ۴- نتیجه نهایی متغیرهای شبیه SVC.

SPI	MSE	۱۰۰۰× C	تابع هسته	متغیر تابع هسته
بهار	۰/۸۹	۶۵۰	خطی	-
	۰/۹	۳۴۵	خطی	-
پائیز	۱/۸۵	۶۸	خطی	-
زمستان	۱/۳	۱۲۰	خطی	-
پائیز+زمستان	۱/۲	۱۵۴	خطی	-
زمستان+بهار	۰/۷	۶۶	شعاعی	۱/۱۶
پائیز+زمستان+ بهار	۱/۴	۲۹	خطی	-

### جدول ۵- خطای دسته بندی داده ها(درصد) در دو شبیه

#### ماشین بردار پشتیبان و درخت تصمیم.

SPI	حوض ه	روش ماشین بردار پشتیبان		روش درخت تصمیم	
		آزمون	آموزش	آزمون	آموزش
بهار	ماملو	۲۷	۵	۳۶	۵
	طالقان	۲۷	۱۰	۵۴	۱۰
پائیز	ماملو	۳۶	۵	۵۴	۱۰
زمستان	کرج	۲۷	۱۵	۴۵	۵
پائیز+زمستان	طالقان	۳۶	۵	۵۴	۱۰
	طالقان	۳۶	۱۰	۵۴	۱۰
پائیز+زمستان+ بهار	کرج	۳۶	۱۵	۳۶	۱۰

### جدول ۶- زمان پیش دید پیش بینی نمایه ی بارش معیار شده برای هر نمایشنامه.

حوضه	زمان پیش دید مربوط به هر نمایشنامه (ماه)					
	SPI- 1	SPI- 2	SPI- 3	SPI- 4	SPI- 5	SPI- 6
کرج	-	-	۲	-	-	۲
ماملو	۵	۲	-	-	-	-
طالقان	۵	-	-	۲	-	-

افزار C-TREE استفاده شده است. جهت مقایسه پیش بینی های انجام گرفته به وسیله شبیه دسته بندی ماشین بردار پشتیبان با استفاده از همان ورودی های انتخاب شده در مرحله قبل دسته بندی SPI بر اساس متغیرهای هواشناسی با استفاده از روش درخت تصمیم برای نمایشنامه های مطلوب انجام گرفت. به طوری که از ۷۰ درصد داده ها برای آموزش ساختار درخت و از ۳۰ درصد داده ها برای آزمون شبیه درخت تصمیم استفاده شد. ساختار مناسب درخت تصمیم با تغییر در تعداد شاخه ها و با توجه به معیار میانگین مربعات خطا تعیین شد.

### نتایج

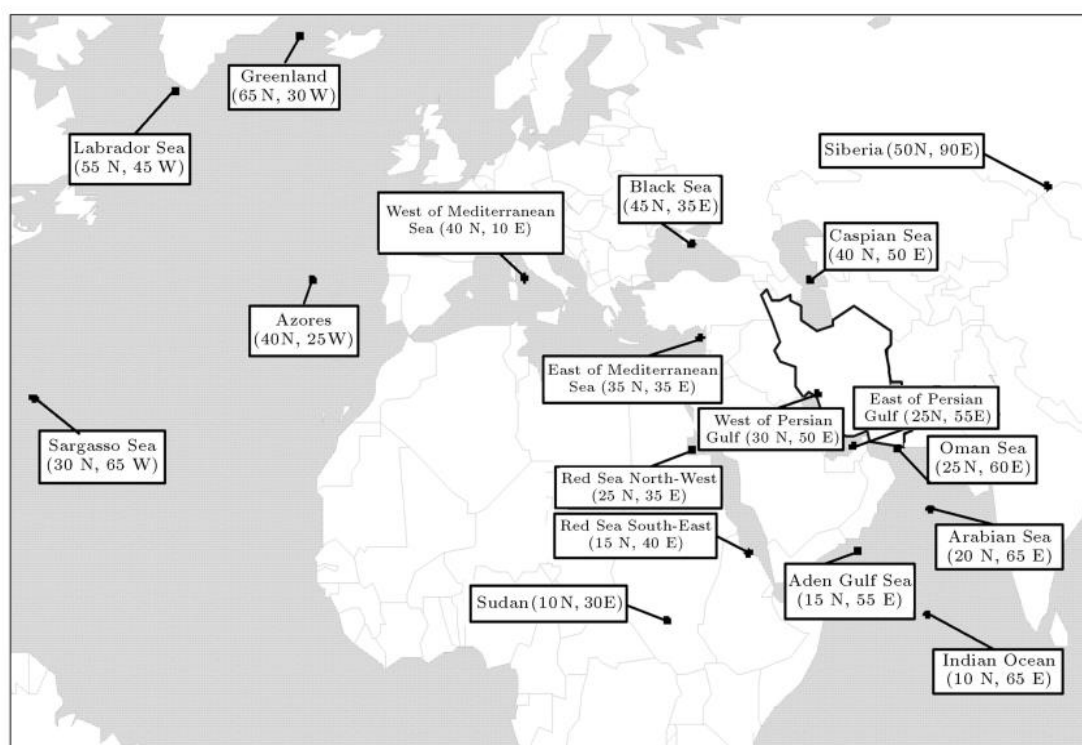
متغیرهای پیش بینی کننده ی انتخاب شده برای نمایشنامه های منتخب به منظور ایجاد شبیه پیش بینی کننده در جدول (۳) ارائه شده اند. مقدار مبنای شاخص اطلاعات متقابل برای انتخاب پیش بینی کننده ها پس از بررسی نتایج برای هر نمایشنامه فصل-حوضه شبیهی که کمترین MSE را دارا بود انتخاب گردید، سپس پیش بینی وضعیت خشکسالی و دسته SPI استفاده شد (جدول ۴). آزمون شبیه برای هر نمایشنامه حوضه به صورت جداگانه انجام گرفت. نتایج آزمون شبیه و مقایسه ی آن با مقادیر مشاهده شده در شکل های (۵) نشان داده شده اند. نتایج حاصل از پیش بینی دسته ی SPI در مقایسه با نتایج ماشین بردار پشتیبان برای هر نمایشنامه در جدول (۵) آورده شده است. همان طور که مشاهده می شود، دستاوردهای بدست آمده از ماشین بردار پشتیبان دارای خطای (شمار نقاطی که در گام آزمون شبیه به اشتباه دسته بندی شده است) کمتری در مقایسه با روش دسته بندی درخت تصمیم است.

شکل (۵) وضعیت پیش بینی ها را برای داده های آزمون شبیه نشان می دهد. به طوری که در این تخمین یک سال تر و دو سال خشک در طبقه میانه پیش بینی شده است. در جدول (۶) حداقل زمان پیش دید برای هر نمایشنامه و حوضه ی آبخیز با توجه به متغیرهای هواشناسی انتخاب شده مشخص شده است. همان طور که ملاحظه می شود زمان پیش دید پیش بینی نمایه ی

نمایشنامه های فصل پائیز از متغیرهای هواشناسی کمتری استفاده کرد.

نتایج این تحقیق نشان می‌دهد که با توجه به نقطه متغیرهای انتخاب شده، محدوده‌های دارای ارتباط مؤثر با بارش محدوده مطالعه شده در فصل بهار بیشتر در جنوب، جنوب غربی و شمال غربی کشور و در فصل پائیز، در شمال، شمال غربی و جنوب و در زمستان، در شمال غربی و غرب کشور قرار دارند. این موضوع تقریباً منطبق بر مطالعات صورت گرفته پیشین است (کارآموز و همکاران، ۲۰۰۶). همچنین متغیرهای پیش‌بینی کننده استفاده شده به ترتیب بیشتر در سطوح ۸۵۰، ۷۰۰، ۵۰۰ و ۱۰۰۰ میلی بار سطح جو قرار گرفته‌اند.

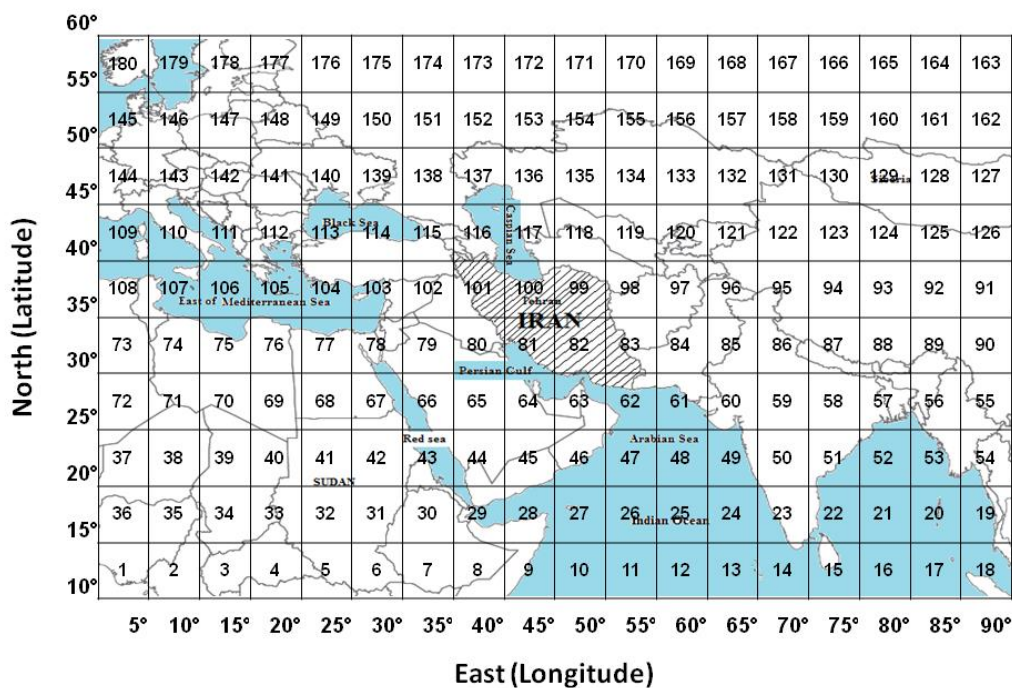
بارش معیار شده برای حوضه کرج، ماملو و طالقان برای هر فصل متفاوت است. زمان پیش‌دید نمایشنامه فصل بهار برای کلیه حوضه‌های تحقیق شده برابر با ۵ ماه بوده، در حالی که زمان پیش‌دید در نمایشنامه فصل زمستان بین ۲ تا ۴ ماه است. این زمان در نمایشنامه هایی که با فصل پائیز آغاز می‌شود، ۲ تا ۳ ماه است. بدیهی است که هر چه زمان پیش‌دید افزایش یابد متناسب با آن میزان خطای شبیه نیز افزایش می‌یابد، اما در شبیه پیش‌بینی بردار پشتیبان نتایج حاصل برای استفاده در بازه زمانی مطلوب با کمترین خطا صورت گرفته است. همچنین می‌توان به منظور افزایش زمان پیش‌دید به خصوص برای



شکل ۳- منابع مؤثر بر رطوبت بارندگی‌های ایران (کارآموز و همکاران، ۲۰۰۹)

بیشترین تعداد پیش‌بینی کننده را شامل می‌شود. در مقایسه با روش‌های پیشنهادی دیگر محققین، رابطه‌ی شاخص MI ارتباط آماری متغیرهای هواشناسی با شاخص SPI را با توجه به خطای پایین محاسبه شده نشان می‌دهد. ضمناً، روش اطلاعات متقابل یک روش جدید برای بررسی ارتباط آماری و انتخاب پیش‌بینی کننده ها بوده است که از آن در هیچ یک از تحقیقات گذشته مورد استفاده قرار نگرفته است.

۶۰٪ درصد داده‌های انتخاب شده برای کلیه نمایشنامه ها متغیر ارتفاع معادل فشار است که نشان دهنده‌ی تأثیر بیشتر این متغیر بر نمایه‌ی بارش معیار شده است. از متغیرهای هواشناسی پیش‌بینی کننده ارتفاع معادل فشار در سطح ۸۵۰ میلی بار و ۷۰۰ میلی بار در کلیه نمایشنامه ها استفاده شده و بیشترین تعداد متغیری است که در پیش‌بینی‌ها به کار رفته است. پس از این دو متغیر، متغیرهای ارتفاع معادل فشار در سطح ۵۰۰ میلی بار و دما در سطح ۵۰۰ میلی بار و دما در سطح زمین

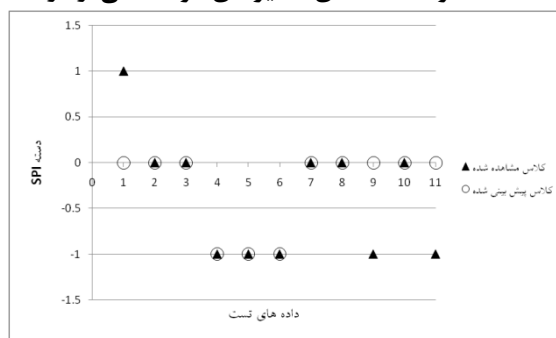


شکل ۴- گستره ی محدوده ۱۸۰ گانه ی جغرافیایی

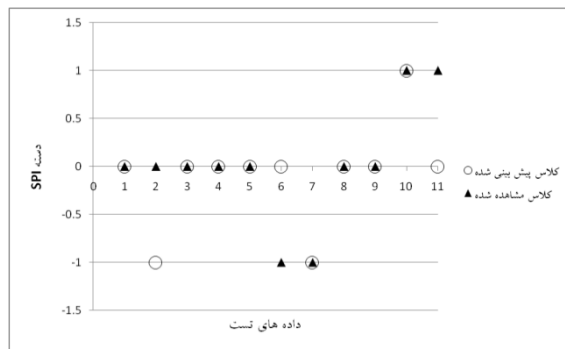
(متوسط ماهانه ی متغیرهای هواشناسی از درگاه NCAR در هر محدوده ی جغرافیایی اخذ شده است)



(ب)



(الف)



(پ)

شکل ۵- مقایسه طبقه های پیش بینی شده و مشاهده شده نمایه بارش معیار شده در (الف) فصل بهار حوضه ماملو، (ب) فصل بهار حوضه طالقان، (پ) فصل زمستان حوضه کرج

استفاده شده است، پیش‌بینی‌ها فقط در روش شبکه‌های عصبی مصنوعی قابل مقایسه و نزدیک به روش ارائه شده در این تحقیق است.

### جمع‌بندی و پیشنهادها

استفاده از اطلاعات جوی و نحوه‌ی تغییرات آنها می‌تواند یکی از جدی‌ترین ریشه‌های رفتار سنجی اقلیمی باشد. این جستجو چنانچه در پهنه‌ی مناسب و همراه با تعیین فاصله‌ی زمانی مناسب همراه باشد، ارتباط ارگانیک در ساختار جوی را نمایان می‌سازد، که شاید تنها در بحث‌هایی کیفی و توصیفی ردپایی از آن را بتوان جست. این کاوش هنگامی که به منظور تخمین شدت و یا طبقه‌بندی خشکسالی باشد، و زمان پیش‌دید مناسبی نیز داشته باشد، کارساز است، به ویژه اینکه از روش‌هایی آماری اما با ساختاری نوین همچون SVM بهره‌گرفته شود. ساختار به کار گرفته شده در این پژوهش، حجم قابل توجه داده‌های مکانی در لایه‌های متعدد جو را پردازش شده و از این نتایج در رسیدن به بهترین ساختار ورودی‌های شبیه‌بهره‌جسته است. متغیرهای هواشناسی پیش‌بینی‌کننده‌ی ارتفاع معادل فشار در سطح ۸۵۰ میلی‌بار و ۷۰۰ میلی‌بار در کلیه‌ی نمایاننامه‌ها به کار رفته و بیشترین تعداد متغیری است که در پیش‌بینی‌ها به کار رفته است. پیش‌بینی‌دسته‌ی SPI با دو روش ماشین بردار پشتیبان و درخت تصمیم نشان از دقت بیشتر شبیه‌دسته‌بندی‌کننده‌ی SVM نسبت به درخت تصمیم دارد. نتایج نشان دادند که می‌توان با زمان پیش‌دید ۲ تا ۵ ماهه در هر یک از حوضه‌های مطالعه شده پیش‌بینی‌دسته‌ی SPI با دقت قابل قبول صورت گیرد. پیشنهاد می‌گردد تا در ادامه این تحقیق با انتخاب خودکار داده‌های هواشناسی همراه با فرآیند شبیه‌سازی اقدام گردد تا بتوان نتایج انتخاب گسسته متغیرهای شبیه، و روش‌های آماری همچون شاخص اطلاعات متقابل ارزیابی گردند. دیگر اینکه، عملکرد شبیه SVM با روش‌های هوشمند و عصبی تنظیم گردد و علامت‌های کوتاه جوی نیز در این فرآیند ارزیابی گردد.

### منابع

1) Barros, P. and Bowden, J. 2008. Forward long-lead operational forecasts of drought: An experimental study in the Murray-Darling

برای نمایاننامه‌هایی که با فصل پاییز شروع می‌شوند متغیرهای هواشناسی سطوح ۸۵۰ و ۷۰۰ میلی‌بار بر نمایه‌ی بارش معیار شده با استفاده از شاخص اطلاعات متقابل مؤثر شناخته شده‌اند. در حالی که برای نمایاننامه‌هایی که با فصل زمستان آغاز می‌شوند تقریباً کلیه‌ی سطوح بر نمایه‌ی بارش معیار شده مؤثر شناخته شده‌اند. می‌توان نتیجه گرفت که سامانه‌های مؤثر بر SPI برای فصول زمستان پیچیده‌تر و با دخالت کلیه سطوح در نظر گرفته شده در تحقیق حاضر است. درصد قابل توجهی از پیش‌بینی‌کننده‌های انتخاب شده در ماه مه و ژوئن قرار گرفته‌اند که امکان پیش‌بینی با زمان پیش‌دید حداقل ۳ ماهه و حداکثر ۵ ماهه را برای نمایاننامه‌های مختلف فراهم می‌کند. زمان پیش‌دید در تحقیقات گذشته بین ۱ تا ۳ ماهه بوده و دقت کمتری حاصل شده است.

نتایج پیش‌بینی‌دسته‌ی مطلوب SPI در مناطق بررسی شده نشان داده که در پیش‌بینی کلیه‌ی نمایاننامه‌ها، بیش از ۶۰ درصد مواقع دسته‌ی پیش‌بینی شده مشابه دسته‌ی مشاهداتی است. نتایج دسته‌بندی و دقت تخمین داده‌ها در هر دسته با بکارگیری توابع مختلف هسته‌ی برای هر بار دسته‌بندی نشان می‌دهد که تغییر قابل توجهی در دقت تعیین نوع دسته حاصل نمی‌شود. همچنین، در اکثر وضعیت‌ها امکان دسته‌بندی با دقت بالای ۷۰ درصد امکان پذیر است. این دقت در نمایاننامه‌های مربوط به پاییز و زمستان بیشتر از سایر نمایاننامه‌ها می‌باشد.

نتایج دسته‌بندی داده‌ها با استفاده از روش درخت تصمیم برای نمایاننامه‌های مطلوب نشان از دقت بالاتر ماشین بردار پشتیبان دارد. با بررسی نقطه‌ی متغیرهای استفاده شده در هر نمایاننامه، بررسی اینکه چند دوره در حالت خشکسالی قرار گرفته است، در حالتی که پیش‌بینی‌دسته‌ی مورد نظر نمایه بارش معیار شده به درستی انجام نگرفته است، مشاهده شد که در نمایاننامه فصل بهار ۷۰ درصد داده‌های استفاده شده آموزش شبیه و در نمایاننامه فصل زمستان ۵۶ درصد داده‌های آموزش شبیه با SPI بالای یک قرار دارند، که می‌تواند علت خطای برآورد زیاد SPI در این دو نمایاننامه باشد. در تحقیقات پیشین که از روش‌های شبکه‌های عصبی مصنوعی، زنجیره‌ی مارکوف و لگاریتم چند متغیره و ...

- and duration to time scales. In: Proceedings of the Eighth Conference on Applied Climatology. Am. Meteor. Soc., Boston, 179–184.
- 13) Moreira, E.E., Coelho, C.A., Paulo, A.A., Pereira, L.S. and Mexia J.T. 2008. SPI-based drought category prediction using loglinear models. *Journal of Hydrology*. 354: 116–130.
- 14) Moreira E.E., Paulo A.A., Pereira L.S., Mexia J.T., 2006. Analysis of SPI drought class transitions using loglinear models. *J. Hydrol.* 331: 349–359.
- 15) Mosaa, Kh. 2015. Forecasting of meteorological drought using Hidden Markov Model (case study: The upper Blue Nile river basin, Ethiopia). *Ain Shams University*. 7(1): 47–56.
- 16) Pai, P-F. and Hong, W-C. 2007. A recurrent support vector regression model in rainfall forecasting, *Hydrological Processes*. 21: 819-827.
- 17) Paulo, A., Ferreira, E., Coelho, C., and Pereira, L.S. 2005. Drought Class Transition Analysis through Markov and Loglinear Models, an Approach to early Warning. *Agricultural water Management*. 77: 59-81.
- 18) Peng, H.C., Long, F. and Ding, C. 2005. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 27:1226-1232.
- 19) Ravinesh, C., Ozgur Kisi, D., Vijay, P. and Singh. P. 2017. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmospheric Research*. 184: 149–175.
- 20) Razi, Tayeb. 2016. Drought prediction in dry and semi-arid region of Iran using time-series and Markov chain simulations. *Scientific and Research Journal of Watershed Engineering and Management*. (4): 454-477 (In Persian).
- 21) Ribeiro, A. and Pires, C.A.L. 2016. Seasonal drought predictability in Portugal using statistical–dynamical techniques. *Physics and Chemistry of the Earth*. 94: 155–166.
- 22) Sesnie, S.E., Finegan, G.B. and Thessler, S. 2008. Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. *Remote Sensing of Environment*, 112 (5): 2145-2159.
- River Basin. *Journal of Hydrology*, 357: 349-367.
- 2) Behzad, M., Asghari, K., Eazi, M. and Palhang, M. 2009. Generalization performance of support vector machines and neural networks runoff modeling. *Expert System with Applications*, 36: 7624-7629.
- 3) Belayneh, J. Adamowski, B. Khalil, B. and Ozga-Zielinski, B. 2014. Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression models. *Journal of Hydrology*. 508: 418–429.
- 4) Bonaccorso Brunell, A., Cancelliere, Giuseppe Rossi. 2015. Probabilistic forecasting of drought class transitions in Sicily (Italy) using Standardized Precipitation Index and North Atlantic Oscillation Index. *Journal of Hydrology*. 526: 136–150.
- 5) Chen Shien-Tsung, Pao-Shan Yu, Yi-Hsuan Tang., 2009. Statistical downscaling of daily precipitation using support vector machines and multivariate analysis. *Journal of Hydrology*, doi:10.1016/j.jhydrol.2010.01.021.
- 6) Dibike, Y.B., Velickov, S., Solomatine, D. and Abbott, M.B. 2001. Model induction with support vector machines: Introduction and applications. *Journal of Computing in Civil Engineering* 15 (3): 208–216.
- 7) Ebrahimi, R., Zahraie, B., and Naseri. M. 2011. Meteorological Drought Forecasting Using Fuzzy Inference Method. *Journal of Water and Sewage Engineering*, 78: 112-125 (In Persian).
- 8) Elsa, E. and Moreira. M. 2016. SPI drought class prediction using log-linear models applied to wet and dry seasons. *Journal of Hydrology*, 94: 136–145.
- 9) Karamouz, M., and Fallahi, M., Nazif, S and M Rahimi Farahani. 2009. Long Lead Rainfall Prediction Using Statistical Downscaling and Artificial Neural Network Modeling. *Scintia Iranica*. 16: 165-172
- 10) Lin J.-Y., Cheng, C.-T., and Chau K.W. 2006. Using support vector machines for longterm discharge prediction. *Hydrological Sciences Journal*. 51(4): 599–612.
- 11) Loukas, A., and Vasiliades, L. 2004. Probabilistic Analysis of Drought Spatiotemporal Characteristics in Thessaly Region, Greece. *Natural Hazards and Earth System Sciences*. 4: 719-731.
- 12) McKee, T.B., Doesken, N.J. and Kleist J. 1993. The relationship of drought frequency

- 23) Tripathi, Sh., Srinivas, R. and Nanjundiah, S. 2006. Downscaling of precipitation for climate change scenarios: A support vector machine approach. *Journal of Hydrology*. 330: 621– 640.
- 24) Vapnik, V. and Chervonenkis, A. 1991. The necessary and sufficient conditions for consistency in the empirical risk minimization method, *Pattern Recognition and Image Analysis*. 1(3): 283-305.
- 25) Vapnik, V. N. and Cortes, C. 1995. Support vector networks. *Machine Learning*. 20: 273–297.
- 26) Wang, W. C. and Men, W.Lu. 2008. Online prediction model based on support vector machine. *Neurocomputing*. 71: 550-558.
- 27) Witten, Ian H., and Frank, Eibe. 2005. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, Amsterdam.
- 28) Zahraie B., Roozbahani A., 2007. Climate Signal Clustering Using Genetic Algorithm for Precipitation Forecasting: A Case Study of Southeast of Iran Proceedings of the World Environmental & Water Resources Congress (ASCE), Tampa, Florida, USA
- 29) Zahraie, B., and Roozbahani, A. 2007. Clustering of Meteorological Signs Due to Precipitation Changes (Case Study: Rainfall Forecast of Sistan and Baluchestan Province), *Journal of Watershed Management Sciences and Engineering*, 1 (2): 21-29 (In Persian).