



مدیریت آب و آبیاری

دوره ۶ ■ شماره ۲ ■ پاییز و زمستان ۱۳۹۵

صفحه‌های ۳۳۱-۳۴۶

روش ترکیبی درخت تصمیم و قوانین انجمنی در پیش‌بینی بلندمدت بارش

وحید نورانی^{۱*}، محمدتقی ستاری^۲ و امیر مولاجو^۳

۱. استاد، گروه مهندسی منابع آب، دانشکده عمران، دانشگاه تبریز، ایران

۲. استادیار، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، ایران

۳. کارشناسی ارشد مهندسی و مدیریت منابع آب، دانشکده عمران، دانشگاه تبریز، ایران

تاریخ پذیرش مقاله: ۱۳۹۵/۱۲/۲۱

تاریخ وصول مقاله: ۱۳۹۵/۰۹/۱۰

چکیده

پیش‌بینی‌های بلندمدت متغیرهای هیدروکلیماتولوژیکی مهم نظری بیشینه بارش ماهانه ابزاری مهم در مدیریت منابع آب است. تحقیقات گذشته نشان داده است که یافتن روابط بین پدیده‌های بزرگ مقیاس اقیانوسی- اتمسفری مانند دمای سطح دریاها و متغیرهای هیدروکلیماتولوژیکی، نظری بارش، به پیش‌بینی بلندمدت این متغیرها کمک می‌کند. در این مطالعه ترکیب دو تکنیک داده‌کاوی (درخت تصمیم و قوانین انجمنی) در استخراج وابستگی‌های مهم بین بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز و دمای سطح دریاهای سیاه، مدیترانه و سرخ معروفی شده است. دو گام اصلی در این مطالعه طبقه‌بندی دمای تفاضلی سطح آب دریاهای مذکور و انتخاب مؤثرترین گروه‌ها، همچنین استخراج الگوهای پنهان‌نهرفته در میان داده‌ها بود و از تکنیک‌ها و الگوریتم‌های درخت تصمیم در طبقه‌بندی و انتخاب مؤثرترین گروه‌ها و از قوانین انجمنی برای کشف الگوهای اطلاعات پنهان‌نهرفته در میان داده‌های مشاهداتی استفاده شد. نتایج نشان داد، بین هر یک از دمای سطوح آب دریاهای سیاه، مدیترانه و سرخ با بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز همیستگی نسبی وجود دارد، بهطوری که شاخص اعتماد بین هر یک از دمای سطح پنهان‌های آبی فوق، با بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز، بالاتر از ۶۰ درصد برآورد شد.

کلیدواژه‌ها: ارومیه، بیشینه بارش ماهانه، تبریز، داده‌کاوی، دمای سطح دریا، طبقه‌بندی.

مقدمه

سامانه و منبع رطوبتی به طور پیوسته برقرار باشد تا کاهش رطوبتی سامانه به علت بارش از سوی منبع رطوبتی جبران شود و بارش ادامه یابد. این عمل باعث تقویت دوباره سامانه می‌شود و این چرخه با مساعدت شرایط جغرافیایی، وقوع بارش چشمگیری را در پی خواهد داشت. اقیانوس‌ها و دریاها منشأ اصلی و تأمین‌کننده آب مورد نیاز جوّ برای بارش است. هرچه دمای سطح آب دریاها و اقیانوس‌ها بیشتر باشد، تبخیر از سطح آن‌ها بیشتر و آب بیشتری وارد جوّ زمین می‌شود. از طرفی، افزایش دمای سطح آب اقیانوس‌ها و دریاها موجب ناپایدار شدن توده‌های هوایی عبوری از روی مناطق دریایی با دمای گرم‌تر می‌شود؛ لذا، بارش بیشتر در این مناطق را به دنبال دارد. الگوی کلی انتقال رطوبت وابستگی شدیدی به الگوی جريان و گردش جوّ منطقه دارد و همین گردش عمومی هوا باعث می‌شود که در ایام مختلف سال، سطوح آبی مقادیر مختلفی از بخار آب را به سمت خشکی‌ها انتقال دهند.

از همین‌رو، محققان با توجه به وسعت دریاها و اقیانوس‌ها و با توجه به اینکه این منابع آبی تأمین‌کننده رطوبت اتمسفر است، با در نظر گرفتن اندکش اقیانوسی - اتمسفری در مقیاس‌های منطقه‌ای و جهانی به تحلیل تغییرات فیزیکی دمای سطح آب اقیانوس‌ها و دریاها و تغییرات بارش پرداخته و می‌پردازنند.

در پیش‌بینی‌های بلندمدت بارش، روش‌های گوناگونی برای ایجاد ارتباط بین بارش و الگوهای بزرگ مقیاس اقیانوسی - اتمسفری ارائه شده است. این روش‌ها را می‌توان به‌طور کلی در دو دسته‌بندی قرارداد: (الف) دینامیکی، و (ب) تجربی. مدل‌های دینامیکی نظیر مدل گردش عمومی جوّ^۷ بر اساس قوانین فیزیکی پایه‌ریزی

بارش یکی از مهم‌ترین داده‌های ورودی به سیستم‌های هیدرولوژیکی و هیدرولیماتولوژیکی است. مطالعه و اندازه‌گیری آن در اکثر موارد برای مطالعات رواناب، آب‌های زیرزمینی، سیلان، رسوب و جزان ضروری است. به دلیل رشد سریع جمعیت و افزایش مضلاع دسترسی به آب آشامیدنی در مناطق خشک و نیمه‌خشکی همچون ایران، اهمیت مدیریت صحیح و کارآمد منابع آب و ارائه مدل‌های قابل اعتماد از بارش روزبه‌روز بیشتر می‌شود. در این بین پیش‌بینی‌های قابل اعتماد و بلندمدت بارش در مقیاس زمانی ماهانه و فصلی اهمیت ویژه‌ای دارد.

تحقیقات گذشته نشان می‌دهد که درک روابط بین بارش و تغییرات الگوهای بزرگ مقیاس اقیانوسی - اتمسفری و استفاده از آن در پیش‌بینی بارش منجر به بهبود پیش‌بینی‌های بلندمدت می‌شود. پدیده‌های اقیانوسی - اتمسفری نظیر نوسانات دهه‌ای اقیانوس آرام^۱، نوسانات اطلس شمالی^۲، نوسانات النیوی جنوبی^۳، شاخص نوسانات جنوبی^۴، دمای سطح دریا^۵ و فشار سطح دریا^۶ اطلاعات پیش‌بینی‌کننده مهمی در مورد تغییرات هیدرولوژیکی در سراسر مناطق جهان فراهم می‌کند (۲۶، ۲۱، ۱۵، ۱۴، ۱۱، ۷، ۵).

برای وقوع بارش علاوه بر سازوکارهای دینامیکی و ترمودینامیکی در ایجاد سرعت قائم مناسب، وجود رطوبت به اندازه کافی نیز ضروری است. وجود بخار آب (رطوبت) به اندازه کافی و مناسب، نه تنها در زمان بارش لازم است، بلکه باید در طول مدت معینی نیز ارتباط بین

-
1. Pacific Decadal Oscillation (PDO)
 2. North Atlantic Oscillation (NAO)
 3. El Nino-Southern Oscillation (ENSO)
 4. Southern Oscillation Index (SOI)
 5. Sea Surface Temperature (SST)
 6. Sea Level Pressure (SLP)

مدیریت آب و آبیاری

دوره ۶ ■ شماره ۲ ■ پاییز و زمستان ۱۳۹۵

شده است برای شناسایی روابط و الگوهای پنهان بین چندین متغیر استفاده می‌شود (۳، ۸، ۲۵).

هدف از این پژوهش بررسی واپستگی بین بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز و دمای سطح دریاهای سیاه، مدیترانه و سرخ با درنظرگرفتن تأخیرهای چند ماهه است. در این تحقیق، روشی ترکیبی از دو روش داده‌کاوی (درخت تصمیم و قوانین انجمنی) در بررسی تأثیرات دمای تفاضلی سطح دریاهای مذکور با بیشینه بارش ماهانه ایستگاه‌های ارومیه و تبریز پیشنهاد داده شده است. دلیل این انتخاب آن بود که هنگام مواجهه و بررسی حوادث شدید در مقیاس‌های بزرگ که تحت تأثیر فرایندهای دورمحور است (مانند بیشینه بارش ماهانه) تکنیک‌های داده‌کاوی جایگزین قابل اعتمادی در مقابل روش‌ها و ابزارهای کلاسیک پیش‌بینی سری‌های زمانی، مانند رگرسیون و شبکه‌های عصبی مصنوعی است که عمولاً از اتورگرسیو داده‌های میدانی محصور در پیش‌بینی‌های کوتاه‌مدت استفاده می‌کند (۶، ۱۶، ۱۷، ۱۸).

مواد و روش‌ها

منطقه مورد مطالعه و داده‌ها

در این تحقیق برای ارزیابی کارایی روش پیشنهادی، داده‌های بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز که هر دو در شمال‌غربی ایران واقع شده‌اند، و نیز داده‌های ماهانه دمای سطح دریاهای سیاه، مدیترانه و سرخ مطالعه شده است (شکل ۱). داده‌های بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز، در دوره زمانی ۱۳۹۳ تا ۱۳۹۳، از سایت سازمان هواشناسی کشور^۱ و داده‌های ماهانه دمای سطح دریاهای سیاه، مدیترانه و سرخ در همان دوره زمانی از پایگاه اینترنتی NOAA^۲ دریافت

شده است و در پیش‌بینی اقلیم استفاده می‌شود. با توجه به پیچیدگی‌های محاسباتی مدل‌های گردش عمومی جو، روش‌های تجربی هنوز هم به طور وسیعی در پیش‌بینی‌ها استفاده می‌شود. مدل‌های تجربی بر اساس روابط مشاهده شده بین متغیر پیش‌بینی‌شونده و متغیرهای پیش‌بینی کننده پایه‌ریزی شده است. منطق و اساس این مدل‌ها، پیداکردن ویژگی‌ها و روابط مکانی و زمانی داده‌های تاریخی بارش و متغیرهای پیش‌بینی کننده است تا از آن در پیش‌بینی بارش آینده استفاده شود (۱۲، ۱۳).

در سال‌های اخیر، در کنار روش‌های تجربی و آماری، روش‌های داده‌کاوی به طور گسترش‌های توسعه یافته و برای مدل‌سازی پارامترهای هیدرولیکی استفاده شده است. هدف داده‌کاوی، توسعه و بسط روش‌هایی برای کشف روابط معنادار و پیش‌بینی نشده از میان انبوهی از داده‌های است که به صورت دستی و با ابزار محاسباتی سنتی تشخیص پذیر نیست. با استفاده از این روش استخراج روابط علت و معلولی، تعیین اینکه کدام متغیرها ارتباط قوی با هم دارد و پیش‌بینی برای آینده امکان‌پذیر می‌شود. داده‌کاوی را می‌توان در چهار روش یا زمینه مطالعاتی تقسیم‌بندی کرد (۲۵): الف) تشخیص ناهنجاری، ب) طبقه‌بندی، ج) خوش‌بندی، د) تجزیه و تحلیل روابط انجمنی.

روش تشخیص ناهنجاری در تشخیص و شناسایی داده‌های غیرعادی و ناهنجاری‌ها استفاده می‌شود. روش‌های طبقه‌بندی و خوش‌بندی نیز در تقسیم و گروه‌بندی داده‌ها در چندین طبقه و خوش‌بندی به کار می‌رود. یکی از معروف‌ترین و کارآمدترین ابزارهای داده‌کاوی در طبقه‌بندی و خوش‌بندی، درخت تصمیم است. قوانینی که با درخت تصمیم ارائه می‌شود به آسانی بیان می‌شود و به همین علت، تجزیه و تحلیل روابط ارائه شده به آسانی انجام می‌شود (۲۲، ۲۳). تجزیه و تحلیل روابط انجمنی نیز که به صورت خیلی محدودی در مطالعات محیط‌زیستی استفاده

1. www.weather.ir

2. <https://www.esrl.noaa.gov/psd/cgi-bin/data/timeseries/timeseries1.pl>

و شمال اروپا به این منطقه وارد می‌شود و این امر در کاهش دمای این شهر تأثیر بسیاری دارد. تبریز نیز دارای آب‌وهوایی اقلیمی و کم رطوبت است و در اثر ارتفاعی که از سطح دریا دارد، دارای زمستان‌های طولانی و سرد و تابستان‌های معتمد است. اطلاعات آماری مربوط به بارش و دمای ماهانه این دو ایستگاه در جدول ۱ ارائه شده است.

شد. با توجه به اینکه خلیج فارس تأثیر چندانی در تزریق رطوبت به مناطق غربی و شمال‌غربی ایران ندارد (۲)، در این مطالعه برای کاهش دیمانسیون ماتریس ورودی‌ها، دمای سطح آب خلیج فارس در نظر گرفته نشد. ارومیه دارای اقلیم کوهستانی است که تحت تأثیر جبهه‌های مدیترانه‌ای و توده‌های غربی قرار می‌گیرد، ولی در فصول مختلف سال، توده‌های سرد قاره‌ای نیز از سبیری



شکل ۱. موقعیت جغرافیایی ایستگاه‌های ارومیه و تبریز

جدول ۱. مشخصات آماری داده‌های ماهانه مشاهداتی در دو ایستگاه تبریز و ارومیه در دوره زمانی ۱۳۹۳ تا ۱۳۳۳

ایستگاه‌ها	دماه ماهانه (°C)					
	انحراف معیار	میانگین	انحراف معیار	میانگین	کمینه	بیشینه
	بارش ماهانه (mm)	میانگین	کمینه	بیشینه	کمینه	بیشینه
تبریز	۹/۹۴	-۱۳/۰۰	۲۸/۸۰	۱۲/۲۷	۰	۱۲۸/۴۰
ارومیه	۹/۲۷	-۱۱/۸۰	۲۶/۵۰	۱۱/۲۴	۰	۱۵۴/۶۰
۲۳/۴۵	۲۲/۳۸	۲۲/۳۸	۱۲/۲۷	۱۲/۲۷	۱۲/۴۵	۱۲۸/۴۰
۲۷/۷۷	۳۰/۱۷	۳۰/۱۷	۱۱/۲۴	۱۱/۲۴	۱۵۴/۶۰	۱۵۴/۶۰

سردتر از تبریز است. دریای سیاه، مدیترانه و سرخ در موقعیت‌های جغرافیایی متفاوتی قرار گرفته‌اند. به همین دلیل، شرایط آب‌وهوایی متفاوتی دارند، به‌طوری که میانگین ماهانه دمای سطح دریای سرخ که به خط استوا نزدیک‌تر است در حدود ۱۰ درجه سانتی‌گراد گرم‌تر از دمای سطح دریای مدیترانه و آن نیز در حدود ۷ درجه

فاصله بین شهرهای ارومیه و تبریز در حدود ۱۴۶ کیلومتر است و همان‌طور که در جدول ۱ نیز مشاهده می‌شود، متوسط بارش ماهانه ایستگاه سینوپتیکی ارومیه در حدود ۷ میلی‌متر بیش از متوسط بارش ماهانه ایستگاه سینوپتیکی تبریز است. همچنین، متوسط دمای ماهانه ایستگاه سینوپتیکی ارومیه در حدود ۱ درجه سانتی‌گراد

مدیریت آب و آبیاری

در این مطالعه به منظور بررسی آثار نوسان‌های دمای تفاضلی سطح دریاهای سیاه، مدیترانه و سرخ بر بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز، روشی ترکیبی از دو روش داده‌کاوی پیشنهاد شده است، یعنی درخت تصمیم و قوانین انجمنی. مطابق شکل ۲ چارچوب کلی روش ترکیبی پیشنهادی شامل چهار گام مجاز است. در گام نخست، داده‌های ماهانه دمای تفاضلی شده سطح آب دریاهای سیاه، مدیترانه و سرخ با استفاده از الگوریتم درخت تصمیم به گروه‌های خیلی زیاد (VH)، زیاد (H)، متوسط (M)، کم (L) و خیلی کم (VL)، همچنین داده‌های بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز با تعریف حد آستانه ۲۵ درصد بارش به عنوان بیشینه بارش ماهانه به گروه‌های زیاد (H) و کم (L) تقسیم شد. دلیل این گروه‌بندی این بود که در روش قوانین انجمنی (که در نهایت استفاده خواهد شد)، داده‌ها باید به صورت گستته باشد.

در گام دوم، با استفاده از تکنیک‌های درخت تصمیم، برای زمان‌های تأخیر مختلف، مؤثرترین گروه دمای تفاضلی شده سطح آب دریاهای مذکور شناسایی شد.

در گام سوم، برای کشف قوانین، سه سناریو تعریف شد که در سناریوهای نخست و دوم به ترتیب تأثیر مستقل و آثار دوبعدی دمای تفاضلی شده سطح دریاهای مذکور بر بیشینه بارش ماهانه ایستگاه‌های ارومیه و تبریز بررسی شد. در سناریوی سوم تأثیر توأم‌ان دمای تفاضلی سطح هر سه پهنه آبی دریاهای سیاه، مدیترانه و سرخ بر بیشینه بارش ماهانه ایستگاه‌های ارومیه و تبریز بررسی شد.

سرانجام در گام چهارم، به منظور ارزیابی قوانین استخراج شده، شاخص‌های اعتماد قوانین محاسبه شد.

در ادامه، به اجزای اصلی این روش (درخت تصمیم و قوانین انجمنی) و دلایل استفاده از هر کدام به طور مختصراً اشاره می‌کنیم.

سانسی‌گراد گرم‌تر از دمای سطح دریای سیاه است. لازم به ذکر است که در این مطالعه برای ایستاکردن سری زمانی دمای سطح آب دریاهای سیاه، مدیترانه و سرخ، از تفاضلی کردن مرتبه اول استفاده شد ($\Delta SST = SST(t) - SST(t-1)$).

داده‌کاوی و روش ترکیبی پیشنهادی

در دهه‌های اخیر، پیشرفت‌های بسیاری در فناوری‌های جمع‌آوری و ذخیره داده‌ها باعث شده است که محققان در بسیاری از علوم با حجم انبوهی از داده‌ها مواجه شوند. به همین دلیل، امروزه استفاده از روش‌های سنتی آماری در تجزیه و تحلیل داده‌ها هزینه‌های، همچنین وقت‌گیر بودن آن، مناسب نیست. بنابراین، استفاده از روش‌های نوین در تجزیه و تحلیل داده‌ها مانند داده‌کاوی لازم و ضروری است. داده‌کاوی در مقایسه با سایر روش‌ها، به ویژه روش‌های سنتی آماری، چهار ویژگی و مزیت اصلی دارد:

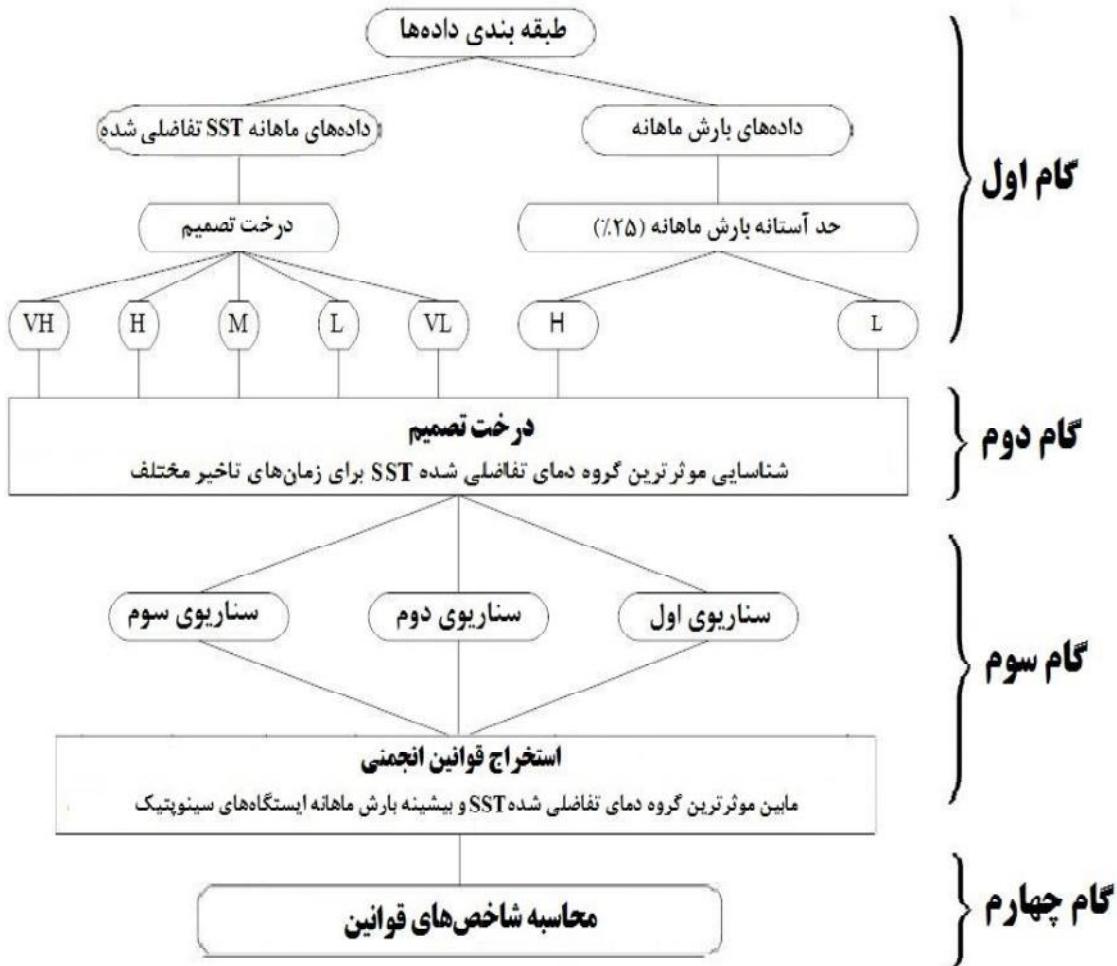
الف) داده‌کاوی انعطاف‌پذیری بسیار بالایی در آنالیز سری‌های زمانی در کشف روابط بین چندین پارامتر ایجاد می‌کند که با هم اتفاق می‌افتد یا با تأخیرهایی مشخص بر یکدیگر اثرگذارند. به عبارتی، داده‌کاوی ابزاری برای سنجش همبستگی بین چندین متغیر یا پارامتر است.

ب) قوانین و مدل‌هایی که الگوریتم‌ها و تکنیک‌های داده‌کاوی ارائه می‌کند به راحتی تفسیرپذیر است. به عبارتی، قوانین و مدل‌های آن ساده، قابل فهم و در عین حال کاربردی است.

ج) استفاده از تکنیک‌های داده‌کاوی محققان را قادر می‌سازد تا دید گستره‌ای از روابط بین پارامترهای مختلف داشته باشند تا بتوانند فیزیک مسئله را راحت‌تر درک کنند.

د) از تکنیک‌های داده‌کاوی می‌توان در مسائلی استفاده کرد که حجم بالایی از داده‌ها با چندین متغیر وجود دارد؛ همچنین، برای نظارت بر حوادث و رویدادهای شدید با حد آستانه بالا.

مدیریت آب و آبیاری



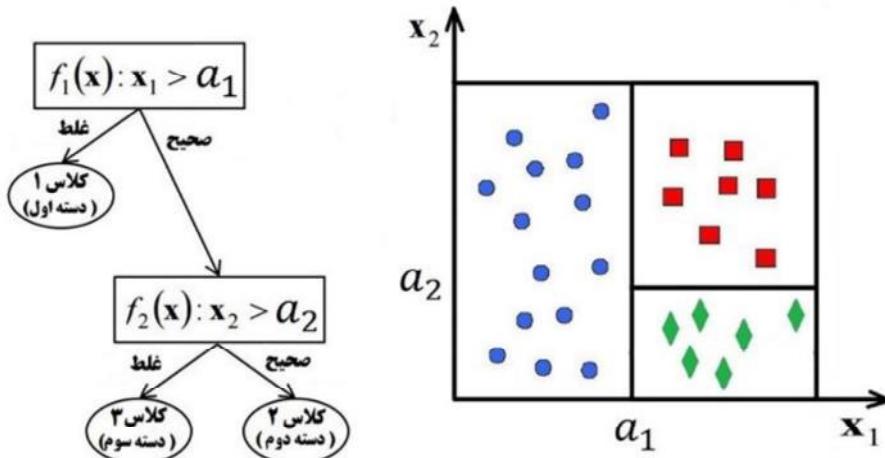
شکل ۲. شمایی از روش ترکیبی پیشنهادی

مهم که تفکیک دسته‌ها و طبقه‌ها بر اساس آن انجام می‌شود، تصمیم‌گیری صورت می‌گیرد و هر شاخه با توجه به نتیجه شرط (صحیح یا غلط بودن شرط) به گره بعدی اشاره می‌کند. درخت تصمیم به دلیل سادگی و قابل فهم بودن، تکنیک محبوبی در داده‌کاوی محسوب می‌شود. به عبارت دیگر، درخت تصمیم خود به تنها یعنی همه مطالب را توصیف می‌کند و نیاز به فرد خبره‌ای نیست تا خروجی اتفاقیس کند (۱).

الف) درخت تصمیم

روش درخت تصمیم مدلی تخمینی و ارزیابی و یکی از پرکاربردترین ابزارها و روش‌ها در طبقه‌بندی و خوشبندی داده‌ها در کاربردهای مختلف داده‌کاوی است. همان‌طور که در شکل ۳ مشاهده می‌شود، ساختار درخت تصمیم ساختاری درختی، شبیه فلوچارت است که از تعدادی گره و شاخه تشکیل شده است. در درخت تصمیمی که عمل طبقه‌بندی انجام می‌شود، برگ‌ها بیانگر کلاس‌هاست. در هر یک از گره‌های دیگر با توجه به یک یا چند مشخصه

مدرسہ آب و آبادی



شکل ۳. ساختار کلی درخت تصمیم

داده‌ها متعلق به کلاس آم باشد. با استفاده از معیار آنتروپی، معیار IG برای صفت خاصه A از رابطه (۲) به دست می‌آید (۱۰).

$$(2)$$

$IG(A) = Entropy(D) - \sum_{j=1}^v \frac{|D_j|}{|D|} Entropy(D_j)$

در این رابطه، v تعداد اعضای دامنه صفت خاصه A و D_j قسمتی از داده‌های اولیه‌ای است که مقدار صفت خاصه آن v_j است. همچنین، $|D_j|$ دلالت بر اندازه داده‌های D دارد. یکی دیگر از معیارهایی که درخت تصمیم برای انتخاب صفت خاصه از آن استفاده می‌کند، معیار Gain (GR) است. در واقع، نرمال شده معیار IG است که طبق رابطه (۳) به دست می‌آید (۱۰).

$$Gain Ratio(A) = \frac{IG(A)}{\sum_{j=1}^v \frac{|D_j|}{|D|} Entropy(D_j)}$$

در این رابطه، v تعداد اعضای دامنه صفت خاصه A و D_j قسمتی از داده‌های اولیه است که مقدار صفت خاصه آن v_j است. چنانچه مخرج کسر مقدار صفر داشته باشد، این معیار قابل تعریف نیست (۱).

ب) قوانین انجمنی

یکی از الگوریتم‌های معروف و بسیار کاربردی درخت تصمیم، الگوریتم C4.5 است که در سال ۱۹۹۲ کوئینلان معرفی کرد (۲۰). این الگوریتم برای طبقه‌بندی در داده‌کاوی استفاده می‌شود و از معیار Information Gain(IG) در انتخاب ویژگی مهم یا همان صفت خاصه برای ایجاد گره استفاده می‌کند. این معیار خود معیاری مبتنی بر آنتروپی و یکی از معیارهای رایج در گسسته‌سازی است. گسسته‌سازی بر اساس آنتروپی تکنیک جداسازی بالا به پایین است که در آن با کمک اطلاعات توزیع کلاس، نقاط جداسازی برای داده‌ها محاسبه می‌شود. برای گسسته‌سازی صفت خاصه‌ای مانند A در این روش، مقداری از A نقطه جداسازی انتخاب می‌شود که مقدار آنtronپی را حداقل کند. این کار به صورت بازگشتی، در انتخاب نقاط جداسازی دیگر ادامه می‌یابد. از رابطه (۱) برای محاسبه آنتروپی مجموعه داده‌ای با D نمونه استفاده می‌شود (۱۹).

$$Entropy(D) = \sum_{i=1}^c -p_i \log_2 p_i \quad (1)$$

در این رابطه، c تعداد برچسب کلاس‌های موجود در داده‌های آموزشی و p_i احتمال این است که نمونه‌ای از

دیریت آب و آبیاری

که مجموعه اقلام X را در خود دارد، به دنبال مجموعه اقلام Y می‌گردد (۱۰).

$$Support(X \rightarrow Y) = P(X \cup Y) \quad (4)$$

$$Confidence(X \rightarrow Y) = \frac{P(X \cap Y)}{P(X)} \quad (5)$$

در این روابط، پارامترهای X و Y به ترتیب نشان‌دهنده رویدادهای پیشین و پی‌آیند و نماد P بیانگر احتمال است.

نتایج و بحث

برای استخراج قوانین انجمنی، نخست باید داده‌ها را در چند گروه طبقه‌بندی کرد. به عبارت بهتر، چون قوانین انجمنی با داده‌های کیفی یا اسمی سروکار دارد، داده‌ها باید گسترش‌سازی شود. در مطالعات قبلی (۳، ۸، ۹، ۲۵) همه شاخص‌ها و داده‌های اقلیمی در تمام ورودی‌ها، به صورت صلب و غیرقابل تغییر، در تعداد گروه‌های یکسان و برابر طبقه‌بندی شده است. این مطالعات با فرض توزیع احتمالاتی نرمال برای ورودی‌ها و شاخص‌های آب‌وهوایی، مرزهای صلب و یکسانی را برای سری زمانی همه متغیرهای ورودی تعریف کرده‌اند.

در این مطالعه، برخلاف مطالعات قبلی، با استفاده از تکنیک‌های درخت تصمیم که بر مبنای معیار آنتروپی به طبقه‌بندی داده‌ها می‌پردازد، مرزهای انعطاف‌پذیر و غیریکسانی برای سری‌های زمانی داده‌های دمای تفاضلی شده سطح آب دریاهای سیاه، مدیترانه و سرخ تعریف شد. نتایج گروه‌بندی داده‌های دمای تفاضلی شده سطح آب دریاهای مذکور در جدول ۲ ارائه شده است. همان‌طور که مشاهده می‌شود، دمای تفاضلی سطح آب دریاهای مدیترانه و سرخ، در پنج گروه و دمای تفاضلی سطح آب دریای سیاه در چهار گروه طبقه‌بندی شده است.

هدف اصلی از داده‌کاوی، کشف دانش در محیط مورد بررسی است. این دانش شکل‌های گوناگونی دارد. دانش استخراج شده ممکن است به فرم الگوهای موجود در داده‌ها باشد که کشف این الگوها منجر به شناخت بهتر سیستم نیز می‌شود. تکنیک‌های متنوعی در داده‌کاوی وجود دارد که الگوهای مختلفی را تولید می‌کند. طی سال‌های گذشته، در میان تکنیک‌های داده‌کاوی توجه خاصی به الگوریتم‌های کشف الگوهای مکرر وجود داشته است. همان‌طور که از نام این الگوریتم‌ها مشخص است، به دنبال الگوهایی هستیم که به دفعات در مجموعه داده‌ها دیده می‌شود. در این میان به الگوریتم‌های کشف مجموعه اقلام مکرر بیشتر پرداخته شده است که در نهایت به تولید قوانین انجمنی منجر می‌شود. در قوانین انجمنی واپسگی‌های مهم میان اقلام موجود در پایگاه داده تراکنشی مشخص می‌شود، به‌نحوی که حضور بعضی اقلام در تراکنش‌ها بر حضور برخی اقلام دیگر در همان تراکنش‌ها دلالت دارد (۱).

قانونی انجمنی که در این بررسی از آن به عنوان روش تحلیل بهره برده شده است، با عبارت $Y \rightarrow X$ بیان می‌شود که در آن X و Y مجموعه اقلام غیرنهایی است که هیچ گونه اشتراکی ندارند ($X \cap Y = \emptyset$). پارامتر X نشان‌دهنده رویداد پیشین و پارامتر Y رویداد پی‌آیند است. هدف از قانون انجمنی یافتن تعداد فراوانی‌ها در مجموعه سری‌ها یا پایگاه داده است که رویدادهای پیشین و پی‌آیند با هم اتفاق افتد. لازم به ذکر است که در این پژوهش برای استخراج قوانین انجمنی از نرم‌افزار Weka استفاده شده است.

مفید بودن هر قانون انجمنی با دو معیار پشتیبان و اعتماد اندازه‌گیری می‌شود. مقدار پشتیبان طبق رابطه (۴) نشان می‌دهد که در چند درصد از تراکنش‌های پایگاه داده‌ها می‌توان مجموعه اقلام X و Y را همراه یکدیگر پیدا کرد و مقدار اطمینان طبق رابطه (۵)، در میان تراکنش‌هایی

مدیریت آب و آبیاری

جدول ۲. گروههای دمای تفاضلی سطح آب دریای مدیترانه، سرخ و سیاه

VH	H	M	L	VL	
> ۳/۰۴	(۱/۱۸) – (۳/۰۴)	(-۱/۲۶) – (۱/۱۸)	(-۲/۶۵) – (-۱/۲۶)	<-۲/۶۵	دریای مدیترانه (°C)
> ۱/۶۳	(۰/۶۱) – (۱/۶۳)	(-۰/۶۴) – (۰/۶۱)	(-۲/۱۴) – (-۰/۶۴)	<-۲/۱۴	دریای سرخ (°C)
-	> ۲/۹۲	(-۱/۰۸) – (۲/۹۲)	(-۳/۲۲) – (-۱/۰۸)	<-۳/۲۲	دریای سیاه (°C)

گستته‌سازی و طبقه‌بندی داده‌ها را تعریف و تفاوت‌های موجود در زیرگروه‌ها را شناسایی می‌کند. به عبارتی، خود درخت تصمیم برای ورودی‌های مختلف بهترین تعداد گروه را مشخص می‌کند. همچنین، این روش قادر به سازگار کردن داده‌های فاقد مقدار، استفاده از ویژگی‌های متفاوت برای نمونه‌های مختلف و نیز کار با داده‌های پیوسته و گستته است که در این مطالعه اهمیت بسیار دارد.

مطابق جدول ۳، برای مقایسه طبقه‌بندی صورت گرفته با درخت تصمیم از روش K-Means استفاده شد که روش متداولی برای خوشبندی است و ساختار نسبتاً ساده و مؤثری دارد. در این روش، امکان خوشبندی با دانستن تعداد اولیه خوش‌ها فراهم می‌شود و همین ملزم بودن برای مشخص بودن تعداد خوش‌ها از ابتدای حل، از جمله مشکلات روش K-Means است (۲۶). در مقابل، درخت تصمیم مرزهای انعطاف‌پذیر و غیریکسان برای

جدول ۳. مرکز خوش‌های دمای تفاضلی سطح آب دریای مدیترانه، سرخ و سیاه در روش K-Means

خوشة ۵	خوشة ۴	خوشة ۳	خوشة ۲	خوشة ۱	دمای تفاضلی سطح آب دریای مدیترانه (°C)
-۳/۲۸	-۱/۶۷	۰/۱۹	۱/۶۷	۳/۲۴	دمای تفاضلی سطح آب دریای سرخ (°C)
-۲/۸۹	-۱/۴۷	۰/۰۱	۱/۴۹	۲/۹۱	دمای تفاضلی سطح آب دریای سیاه (°C)
-۵/۱۲	-۲/۶۵	-۰/۱۹	۲/۳۹	۴/۹۵	

تصمیم و به روش K-Means مشاهده می‌شود که هر دو روش گروه‌بندی‌های تقریباً یکسانی دارد، به نحوی که میانگین بازه در درخت تصمیم برای داده‌های دمای تفاضلی سطح آب دریاهای مذکور، تقریباً مرکز خوش‌های است که K-Means ارائه داده است. برای مثال، میانگین گروه L در درخت تصمیم برای داده‌های دمای تفاضلی شده سطح آب دریای مدیترانه (بازه -۲/۶۵ تا -۱/۲۶) تقریباً با مرکز خوشة چهارم در روش K-Means. یعنی -۱/۶۷، هماهنگ است، با این تفاوت که درخت تصمیم، برخلاف

همان‌طور که قبل اشاره کردیم، از جمله نقاط ضعف روش K-Means حساس بودن به پارامترهای اولیه نظیر تعداد اولیه خوش‌های است. از این‌رو، مطابق جدول ۳، برای خوشبندی داده‌های دمای تفاضلی سطح آب دریای مدیترانه، سرخ و سیاه با استفاده از سعی و خطای مقدار ۵ برای K انتخاب شد و مرکز خوش‌ها که در واقع به عبارتی میانگین آن خوشه است در جدول آمد. با مقایسه گروه‌بندی‌های جدول ۲ و ۳ برای داده‌های دمای تفاضلی شده سطح آب دریاهای مذکور به روش درخت

میریت آب و آبیاری

و تبریز است، پارامتر بارش شامل دو گروه می‌شود. برای تعیین این دو گروه ۷۲۰ داده بارش ماهانه هر دو ایستگاه به ترتیب صعودی به نزولی مرتب شد و ۲۵ درصد بیشترین آن‌ها بیشینه بارش ماهانه ایستگاه‌ها در نظر گرفته شد. به این ترتیب، داده‌های بارش ماهانه در دو گروه H و L قرار گرفت. در گام بعدی، برای زمان‌های تأخیر مختلف، مؤثرترین گروه‌های دمای تفاضلی سطح آب دریای سیاه، مدیترانه و سرخ شناسایی شد (جدول ۲). الگوریتم C4.5 درخت تصمیم برای این منظور استفاده شد (جدول ۴).

روش K-Means، برای ورودی‌های مختلف بهترین تعداد گروه را مشخص می‌کند و نیازی به سعی و خطا نیست. تفاوتی که در گروه‌بندی‌های این دو روش مشاهده می‌شود، تعداد گروه‌های ارائه شده برای دمای تفاضلی سطح آب دریای سیاه است که به نظر می‌رسد دلیل آن پراکندگی زیاد داده‌های دمای تفاضلی شده سطح آب دریای سیاه نسبت به داده‌های دمای تفاضلی سطح آب دریای مدیترانه و سرخ باشد.

در ادامه، با توجه به اینکه هدف از این مطالعه بررسی پیش‌بینی بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه

جدول ۴. مؤثرترین گروه دمای تفاضلی سطح آب دریای سیاه، مدیترانه و سرخ بر بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز در زمان‌های تأخیری مختلف

ایستگاه سینوپتیکی ارومیه			ایستگاه سینوپتیکی تبریز			
دریای سیاه	دریای مدیترانه	دریای سرخ	دریای سیاه	دریای مدیترانه	دریای سرخ	بدون تأخیر
H	H	VH	H	H	VH	یک ماه تأخیر
H	M	VH	M	M	H	دو ماه تأخیر
M	L	H	M	M	H	

بسته به افق پیش‌بینی یا همان زمان تأخیر بین فرایندها و رویدادهای پیشین و پی‌آیند متفاوت است. در سناریوی نخست، بالاترین شاخص اعتماد قوانین انجمنی محاسبه شده ۶۳ درصد است که بین دمای تفاضلی سطح آب دریای مدیترانه و بیشینه بارش ماهانه ایستگاه سینوپتیکی تبریز برقرار است. این بدان معناست که اگر دمای تفاضلی سطح آب دریای سیاه در گروه H قرار گیرد، با شاخص اعتماد برابر با ۶۳ درصد، بارش متناظری بیش از ۳۴ میلی متر (حد استانداری ۲۵٪ بارش ماهانه تبریز) در همان ماه (بدون تأخیر) در شهر تبریز رخ خواهد داد. همچنین، بالاترین شاخص اعتماد محاسبه شده بین مؤثرترین گروه دمای تفاضلی سطح آب دریای مدیترانه و

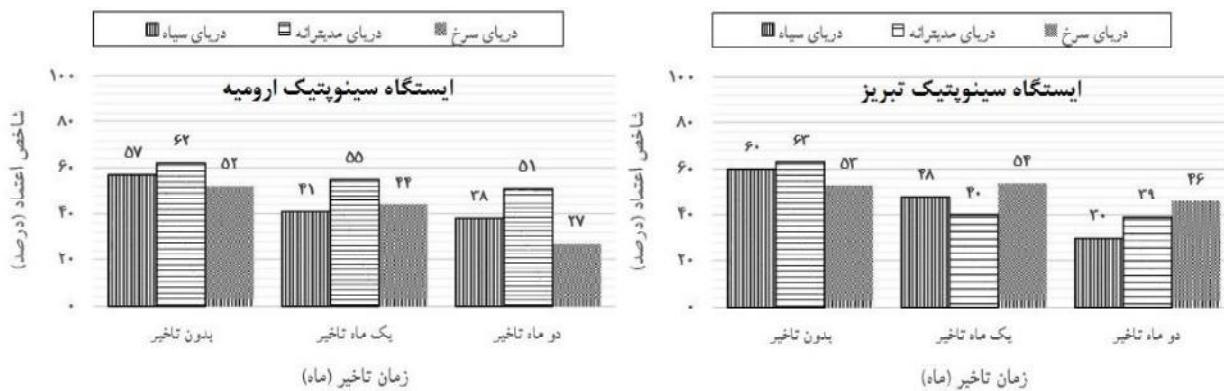
در ادامه، با تعریف سه سناریوی مختلف، قوانین انجمنی بین مؤثرترین گروه‌های دمای تفاضلی شده سطح آب دریای سیاه، مدیترانه و سرخ و بیشینه بارش ماهانه هر دو ایستگاه سینوپتیکی ارومیه و تبریز استخراج شد. برای سناریوی نخست، مطابق شکل ۴، تأثیر مستقل دمای سطح آب هر یک از پهنه‌های آبی فوق بر بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز با درنظر گرفتن زمان‌های تأخیر مختلف بررسی شد. در شکل ۴ شاخص اعتماد قوانین استخراج شده ارائه شده است. مطابق نتایج در شکل ۴، دمای تفاضلی شده سطح آب دریای سیاه، مدیترانه و سرخ بر بیشینه بارش ایستگاه‌های سینوپتیکی ارومیه و تبریز مؤثر است، اما این اثرگذاری‌ها

مدیریت آب و آبیاری

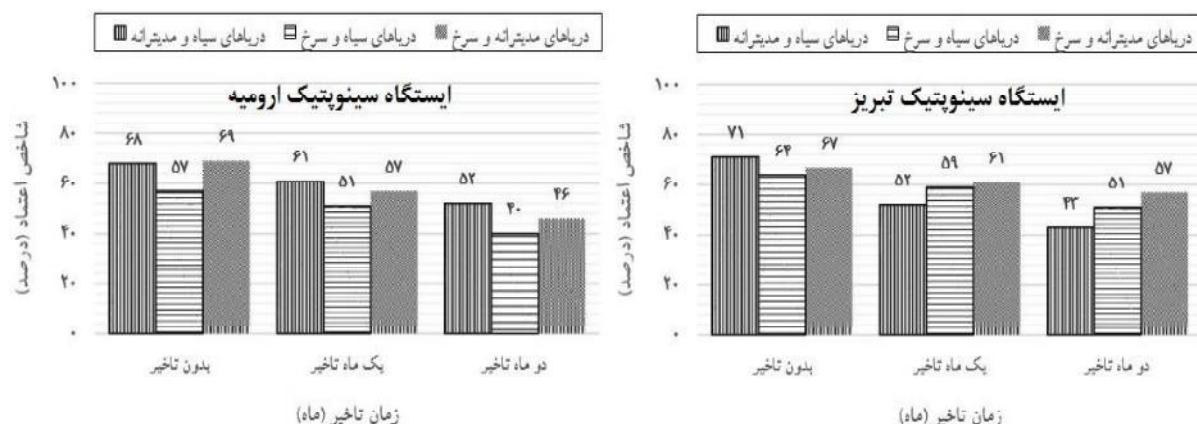
روش ترکیبی درخت تصمیم و قوانین انجمنی در پیش‌بینی بلندمدت بارش

در ادامه بررسی‌ها، مطابق شکل ۵، در سناریوی دوم، تأثیر توأم‌ان دمای تفاضلی سطح آب پهنه‌های آبی مذکور به صورت دوبه‌دو، با بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز بررسی شد. نتایج این سناریو با سناریوی نخست مقایسه شد. نتایج حاکی از افزایش شاخص‌های اعتماد نسبت به سناریوی نخست بود.

بیشینه بارش ایستگاه سینوپتیکی ارومیه نیز برابر با ۶۲ درصد است؛ به این معنا که اگر دمای تفاضلی سطح آب دریای سیاه در گروه دمایی H قرارگیرد، با شاخص اعتماد برابر با ۶۲ درصد، بارش چشمگیری بیش از ۴۲ میلی‌متر (حد آستانه ۲۵٪ بارش ماهانه ارومیه) در همان ماه در شهر ارومیه اتفاق خواهد افتاد.



شکل ۴. شاخص اعتماد قوانین انجمنی بین مؤثرترین گروه دمای تفاضلی سطح آب دریای سیاه، مدیترانه و سرخ و بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز در زمان‌های تأخیری مختلف



شکل ۵. شاخص اعتماد قوانین انجمنی بین مؤثرترین گروه‌های دمای تفاضلی شده سطح آب دریای سیاه، مدیترانه و سرخ به صورت دوبه‌دو و بیشینه بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز در زمان‌های تأخیری مختلف

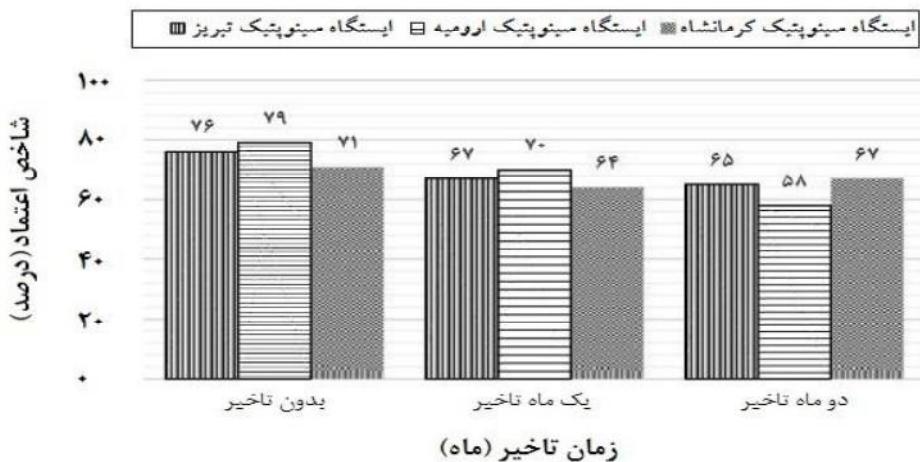
مدیریت آب و آبیاری

دوره ۶ ■ شماره ۲ ■ پاییز و زمستان ۱۳۹۵

سناریوی سوم، تأثیر توأم ان دمای تفاضلی سطح آب هر سه پهنه آبی دریای سیاه، مدیترانه و سرخ با بیشینه بارش ماهانه ایستگاههای سینوپتیکی ارومیه و تبریز با درنظر گرفتن زمانهای تأخیر مختلف بررسی می‌شود. همانند سناریوهای قبلی، شاخصهای اعتماد قوانین انجمانی محاسبه و نتایج با سناریوی نخست و دوم مقایسه شد. نتایج حاکی از این بود که در نظر گرفتن دمای تفاضلی سطح آب هر سه دریا به صورت توأم و همزمان به عنوان قسمت رویدادهای پیشین قوانین انجمانی یا همان پیش‌بینی‌کننده‌ها، باعث افزایش نسبی مقادیر شاخصهای اعتماد نسبت به دو سناریوی قبلی می‌شود. در این سناریو، برای ارزیابی روش‌شناسی پیشنهادی، ایستگاه سینوپتیکی کرمانشاه نیز بررسی شد تا به دلیل فاصله‌ای که بین این ایستگاه با ایستگاههای سینوپتیکی ارومیه و تبریز وجود دارد، بتوان آثار مکانی، همچنین روابط همبستگی بین ایستگاهی را نیز بررسی کرد.

با مقایسه نتایج سناریوی نخست با دوم (شکل ۴ و ۵) مشاهده می‌شود که درنظر گرفتن دمای تفاضلی سطح آب دو دریا به عنوان قسمت رویدادهای پیشین قوانین انجمانی (عامل‌های پیش‌بینی‌کننده) در سناریوی دوم باعث افزایش مقادیر شاخصهای اعتماد نسبت به سناریوی نخست می‌شود. در سناریوی دوم، بالاترین شاخص اعتماد قوانین انجمانی در ایستگاه سینوپتیکی تبریز برابر با ۷۱ درصد است. بین مؤثرترین گروه‌های دمای تفاضلی سطح آب دریای سیاه و مدیترانه بیشینه بارش ماهانه برقرار است. این بدان معنا است که اگر دمای تفاضلی سطح آب دریای سیاه در گروه دمایی H قرار گیرد و همزمان با آن دمای تفاضلی سطح آب دریای مدیترانه نیز در گروه دمایی H باشد، در آن صورت با شاخص اعتماد ۷۱ درصد، بارش متناظری بیش از ۳۴ میلی‌متر در همان ماه (بدون تأخیر) در شهر تبریز رخ خواهد داد.

در ادامه، مطابق با شکل ۶ و جدول ۵، به عنوان



شکل ۶. شاخص اعتماد قوانین انجمانی بین مؤثرترین گروه‌های دمای تفاضلی سطح آب دریای سیاه، مدیترانه و سرخ به صورت توأم و بیشینه بارش ماهانه ایستگاههای سینوپتیکی تبریز، ارومیه و کرمانشاه در زمانهای تأخیر مختلف

جدول ۵. قوانین انجمنی منتخب برای سناریوی سوم

شناخت اعتماد	گروه دمایی دریای سرخ	گروه دمایی دریای مدیترانه	گروه دمایی دریای سیاه	
۷۶	VH	H	H	بدون تأخیر
۶۷	H	M	M	تأخر یکماهه
۶۵	H	M	M	تأخر دوماهه
۷۹	VH	H	H	بدون تأخیر
۷۰	VH	M	H	تأخر یکماهه
۵۸	VH	L	M	تأخر دوماهه
۷۱	VH	M	H	بدون تأخیر
۶۴	H	M	M	تأخر یکماهه
۶۷	M	L	VL	تأخر دوماهه

سیاه و شهر کرمانشاه، آثار تغییرات دمای سطح آب دریای سیاه با زمان تأخیر بیشتری نسبت به دو ایستگاه سینوپتیکی ارومیه و تبریز، در این شهر مشاهده می‌شود. همچنین، نتایج بررسی همبستگی بین ایستگاه‌های نیز حاکی از آن بود که رژیم آب و هوایی ایستگاه‌ها، همچنین موقعیت جغرافیایی ایستگاه‌ها (فاصله از منابع رطوبتی دریاها) دو عامل اصلی و تعیین‌کننده در تعیین روابط بین بیشینه بارش ماهانه ایستگاه‌ها و دمای تفاضلی شده سطح آب دریاهاست، چنانچه موقعیت جغرافیایی ایستگاه‌ها در تعیین مؤثرترین گروه‌های دمای تفاضلی سطح آب دریاها نقش بسیار پررنگی ایفا می‌کند.

در ادامه، مطابق جدول ۶، برای داشتن ارزیابی کلی از روش ترکیبی پیشنهادی در این مطالعه و مقایسه آن با روش‌های سنتی آماری، همبستگی خطی^۱ بین دمای تفاضلی سطح آب دریای سیاه، مدیترانه و سرخ و بارش ماهانه ایستگاه‌های سینوپتیکی ارومیه و تبریز با درنظرگرفتن اثر هم‌زمان یا بدون تأخیر ($Lag = 0$ ، تأخیر یکماهه ($Lag = 1$) و تأخیر دوماهه ($Lag = 2$) محاسبه شد.

1. linear correlation

مطابق شکل ۶ و جدول ۵، در سناریوی سوم بالاترین شناخت اعتماد قوانین انجمنی در ایستگاه سینوپتیکی ارومیه برابر با ۷۹ درصد است و بین مؤثرترین گروه‌های دمای تفاضلی سطح آب دریای سیاه و مدیترانه و سرخ و بیشینه بارش ماهانه این ایستگاه برقرار است. این بدان معناست که اگر دمای تفاضلی سطح آب دریای سیاه و مدیترانه به صورت هم‌زمان در گروه دمایی H قرار گیرد و هم‌زمان با آن دمای تفاضلی سطح آب دریای سرخ در گروه دمایی VH قرار گیرد، در آن صورت با شناخت اعتماد ۷۹ درصد، بارش متناظری بیش از ۴۲ میلی‌متر در همان ماه (بدون تأخیر) در شهر ارومیه رخ خواهد داد. همچنین، بالاترین شناخت اعتماد قوانین انجمنی در ایستگاه سینوپتیکی تبریز نیز برابر با ۷۶ درصد است. به این معناست که اگر دمای تفاضلی سطح آب دریای سیاه و مدیترانه به صورت هم‌زمان در گروه دمایی H و هم‌زمان با آن دمای تفاضلی سطح آب دریای سرخ در گروه دمایی VH قرار گیرد، در آن صورت با شناخت اعتماد ۷۶ درصد، بارش متناظری بیش از ۳۴ میلی‌متر در همان ماه (بدون تأخیر) در شهر تبریز رخ خواهد داد.

در مورد ایستگاه سینوپتیکی کرمانشاه نیز به نظر می‌رسد که به دلیل وجود رشته‌کوه‌ها و ارتفاعات بلند بین دریای

میریت آب و آبیاری

جدول ۶. ضریب همبستگی خطی بین دمای تفاضلی سطح آب دریای سیاه، مدیترانه و سرخ و بارش ماهانه ایستگاه‌های ارومیه و تبریز

ایستگاه سینوپتیکی تبریز		ایستگاه سینوپتیکی ارومیه			
بدون تأخیر	تأخیر یکماهه	بدون تأخیر	تأخیر یکماهه	بدون تأخیر	تأخیر یکماهه
۰/۱۲	۰/۱۴	۰/۲۹	-۰/۰۹	۰/۱۹	۰/۳۲
۰/۲۲	۰/۲۷	۰/۳۴	۰/۱۲	۰/۱۴	۰/۳۸
-۰/۰۷	۰/۱۶	۰/۲۵	۰/۱۷	۰/۲۷	۰/۲۶
دمای تفاضلی سطح آب دریای سیاه		دمای تفاضلی سطح آب دریای مدیترانه		دمای تفاضلی سطح آب دریای سرخ	

شهرها مؤثر است، اما این اثرگذاری‌ها بسته به مدت زمان تأخیر ممکن است متفاوت باشد. در این تحقیق نیز همانند مطالعات قبلی، نتایج نشان داد که استفاده توأم‌ان از دمای سطوح پهنه‌های آبی در پیش‌بینی، کارایی بهتری دارد تا اینکه به صورت مستقل از دمای سطح پهنه‌های آبی در پیش‌بینی استفاده شود. برخلاف تحقیقات قبلی، در این مطالعه نشان داده شد که استفاده از دمای تفاضلی سطح آب دریاها، همچنین استفاده از الگوریتم C4.5 درخت تصمیم برای شناسایی مؤثرترین گروه‌های دمای تفاضلی سطح دریاها در زمان‌های تأخیر مختلف نتایج پیش‌بینی‌ها را بهبود می‌بخشد. با توجه به اینکه هر چه متغیرهای پیشین به عنوان عوامل پیشگوکننده در مدل داده‌کاوی بیشتر شود، امکان وقوع با هم چندین رویداد کمتر می‌شود، از این‌رو برای معتبربودن شاخص‌های ارزیابی کننده الگوها لازم است تعداد داده‌ها زیاد باشد. بهمین منظور، در این تحقیق با توجه به محدود بودن داده‌ها، فقط سه متغیر پیشین دمای سطح آب دریای سیاه، مدیترانه و سرخ استفاده شد، در صورتی که بارش به غیر از دمای سطوح آبی به پارامترهای دیگری نیز ارتباط دارد. توصیه می‌شود در مطالعات آینده، در صورتی که امکان استفاده از داده‌های بیشتری فراهم شد، امکان استفاده از سایر سیگنال‌های بزرگ‌مقیاس اقلیمی دیگر نظری فشار سطح دریاها، شاخص نوسانات جنوبی و شاخص نوسانات اطلس شمالی، همراه

البته، باید توجه داشت که ضریب همبستگی^۱ ابزار سنجش همبستگی بین تنها دو متغیر (یک ورودی و یک خروجی) است، در حالی که در این مطالعه نشان داده شد که با استفاده از قوانین انجمنی و روش ترکیبی پیشنهادی، می‌توان چندین رویداد پیشین را پیش‌بینی کننده در نظر گرفت (همانند سناریوهای نخست و سوم).

نتیجه‌گیری

پیش‌بینی‌های بلندمدت بارش برای برنامه‌ریزی در حوزه‌های مختلف همچون صنعت، کشاورزی و بهداشت لازم و ضروری است. تحقیقات گذشته نشان داده است که استفاده از پدیده‌های اقیانوسی - اتمسفری نظیر دمای سطح دریاها در پیش‌بینی‌های بلندمدت باعث بهبود مدل‌ها می‌شود. همچنین، مطالعات قبلی نشان داده است که در مناطق نزدیک به دریاها، اولویت ممکن است استفاده از دمای سطح دریاهای مجاور و نزدیک باشد.

از آنجا که ایران در شمال، جنوب و غرب با دریا احاطه شده است، در این تحقیق از تکنیک‌های داده‌کاوی برای یافتن روابط بین بیشینه بارش ماهانه دو شهر در شمال‌غرب ایران استفاده و نشان داده شد که دمای سطح دریای سیاه، مدیترانه و سرخ بر بیشینه بارش ماهانه این

1. correlation coefficient

میریت آب و آبیاری

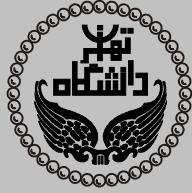
- hydrological predictors. *Water Resources Research*, 42(3): W03431.
8. Dadaser F. Celik M. and Dokuz A.S. (2012) Associations between stream flow and climatic variables at Kizilirmak river basin in Turkey. *Global NEST Journal*, 14(3): 354-361.
9. Dhanya C.T. and Kumar D. (2009) Data mining for evolution of association rules for droughts and floods in India using climate inputs. *Geophysical Research*, 114: 1-15.
10. Han J. Kamber M. and Pei J. (2012) Data mining: concepts and techniques. 3rd Ed. Morgan Kaufmann Publishers, San Francisco, California, United States of America, 744 p.
11. Harshburger B. Ye H. and Dzialoski J. (2002) Observational evidence of the influence of Pacific sea surface temperatures on winter precipitation and spring stream discharge in Idaho. *Hydrology*, 264(1-4): 157-169.
12. He X. Guan H. and Qin J. (2015) A hybrid wavelet neural network model with mutual information and particle swarm optimization for forecasting monthly rainfall. *Hydrology*, 527: 88-100.
13. Lim E.P. Hendon H.H. Hudson D. Wang G. and Alves O. (2009) Dynamical forecast of inter-El Nino variations of tropical sea surface temperature and Australian spring rainfall. *Monthly Weather Review*, 137(11): 3796-3810.
14. Meidani E. and Araghinejad S. (2014) Long-lead streamflow forecasting in the Southwest of Iran by sea surface temperature of the Mediterranean Sea. *Hydrologic Engineering*, 19(8): 1-10.
15. Nicholls N. (1989) Sea surface temperatures and Australian winter rainfall. *Climate*, 2(9): 965-973.

با دمای تفاضلی سطح آب دریاهای مذکور در بررسی امکان افزایش شاخص اعتماد پیش‌بینی بررسی شود.

منابع

۱. اسماعیلی م. (۱۳۹۱) مفاهیم و تکنیک‌های داده‌کاوی. انتشارات نیاز دانش، تهران، ۳۱۵ ص.
۲. خسروی غ.ر. خورانی ا. و خسروی م. (۱۳۹۲) نقش خلیج فارس در وضعیت اقلیمی کشور. رشد آموزش جغرافیا، ۲۸(۲): ۴۳-۴۸.
۳. رحیمی خوب ع. (۱۳۹۰) پیش‌بینی حداکثر بارندگی ماهانه ایستگاه ایلام از روی دمای سطح آب خلیج فارس و دریای سرخ با استفاده از روش داده‌کاوی. *تحقیقات آب و خاک ایران*, ۴۲(۱): ۷-۱۷.
۴. رضایی بنفشه م. جهانبخش س. بیاتی خطیبی م. و زینالی ب. (۱۳۸۹) پیش‌بینی بارش پاییزه و زمستانه نیمه غربی ایران، با استفاده از دمای سطح آب دریای مدیترانه در فصول تابستان و پاییز. *پژوهش‌های جغرافیای طبیعی*, ۴۷-۶۲: ۷۴.
۵. نظام‌السادات س.م.ج. و قاسمی ا.ر. (۱۳۸۳) تأثیر نوسان‌های دمای سطح آب دریای خزر بر بارش فصول زمستان و بهار نواحی شمال و جنوب‌غربی ایران. *علوم و فنون کشاورزی و منابع طبیعی*, ۸(۴): ۱-۱۴.
۶. Akrami S.A. Nourani V. and Hakim S.J.S. (2014) Development of nonlinear model based on Wavelet-ANFIS for rainfall forecasting at Klang Gates Dam. *Water Resources Management*, 28(10): 2999-3018.
۷. Araghinejad S. Burn D.H. and Karamouz M. (2006) Long-lead probabilistic forecasting of streamflow using ocean-atmospheric and

16. Nourani V. Alami M.T. and Aminfar M.H. (2009) A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence*, 22(3): 466-472.
17. Nourani V. Hosseini Baghanam A. Adamowski J. and Kisi O. (2014) Applications of hybrid Wavelet-Artificial Intelligence models in hydrology: A review. *Hydrology*, 514: 358-377.
18. Nourani V. Alizadeh F. and Roushangar K. (2016) Evaluation of a two-stage SVM and spatial statistics methods for 460 modeling monthly river suspended sediment load. *Water Resources Management*, 30(1): 393-407.
19. Otero F.E.B. Freitas A.A. and Johnson C.G. (2012) Inducing decision trees with an ant colony optimization algorithm. *Applied Soft Computing*, 12: 3615-3626.
20. Quinlan J.R. (1992) Learning with continuous classes. *Proceedings 5th Australian Joint Conference on Artificial Intelligence*, Australian.
21. Rowell D.P. (2003) The impact of Mediterranean Sea surface temperatures on the Sahelian rainfall season. *Climate*, 16(5): 849-862.
22. Saghebian S.M. Sattari M.T. Mirabbasi R. and Pal M. (2014) Ground water quality classification by decision tree method in Ardeabil region, Iran. *Arabian Geosciences*, 7(11): 4767-4777.
23. Sattari M.T. Anil A.S. Apaydin H. and Kodal S. (2012) Decision trees to determine the possible drought periods in Ankara. *Atmosfera*, 25(1): 65-83.
24. Soukup T.L. Aziz O.A. Tootle G.A. Piechota T.C. and Wulff S.S. (2009) Long lead-time streamflow forecasting of the North Platte river incorporating ocean-atmospheric climate variability. *Hydrology*, 368(1-4): 131-142.
25. Tadesse T. Wilhite D.A. Harms S.K. Hayes M.J. and Goddard S. (2004) Drought monitoring using data mining techniques: A case study for Nebraska, United States of America. *Natural Hazards*, 33(1): 137-159.
26. Zahraie B. and Roozbahani A. (2011) SST clustering for winter precipitation prediction in southeast of Iran: comparison between modified k-means and genetic algorithm-based clustering methods. *Expert System with Applications*, 38(5): 5919-5929.



مدیریت آب و آبیاری

دوره ۶ ■ شماره ۲ ■ پاییز و زمستان ۱۳۹۵

صفحه‌های ۳۴۷-۳۶۵

مطالعه تجمع فلزات سنگین در محصولات مختلف تحت تأثیر آبیاری با فاضلاب تصفیه‌شده شهری کرمانشاه

میلاد فرمائی‌فرد^۱، هوشنگ قمرنیا^{۲*}، مقداد پیرصاحب^۳ و نظری فتاحی^۴

۱. دانشجوی دکتری آبیاری و زهکشی، گروه مهندسی آب، پردیس کشاورزی و منابع طبیعی، دانشگاه رازی، کرمانشاه، ایران
۲. استاد، گروه مهندسی آب، پردیس کشاورزی و منابع طبیعی، دانشگاه رازی، کرمانشاه، ایران
۳. استاد، گروه مهندسی بهداشت محیط، دانشگاه علوم پزشکی کرمانشاه، کرمانشاه، ایران
۴. استادیار، گروه شیمی، دانشکده علوم، دانشگاه رازی، کرمانشاه، ایران

تاریخ پذیرش مقاله: ۱۳۹۵/۱۲/۲۲

تاریخ وصول مقاله: ۱۳۹۵/۱۰/۰۲

چکیده

در این تحقیق به بررسی تجمع فلزات سنگین شامل مس، آهن، روی، کادمیم و منگنز در اندام‌های مختلف (ریشه، اندام هوایی، میوه یا دانه) ذرت، جو، بامیه و جعفری در شرایط واقعی تحت آبیاری با دو تیمار فاضلاب تصفیه‌شده شهری خروجی تصفیه‌خانه کرمانشاه (TWW) و آب چاه موجود در اراضی همان منطقه (WEW) در سه تکرار و قالب طرح کاملاً تصادفی کرت‌های خردشده برای هر محصول (در سال زراعی ۱۳۹۳) پرداخته شده است. نتایج تحلیل‌های آماری نشان داد که اثر آبیاری با پساب بر تجمع فلزات سنگین در تمامی محصولات در سطح ۵ درصد معنادار بود. همچنین، تجمع آهن، روی و کادمیم در اندام‌های محصولات مورد بررسی تحت تیمار پساب بیش از تیمارهای آب چاه و از نظر آماری نیز معنادار بود و در بیشتر موارد بالاتر از حدود استانداردها نیز قرارداشت، ولی از نظر تجمع مس و منگنز، علی‌رغم مشاهده اختلاف معنادار آماری و مقداری بین تیمارهای مختلف، میزان تجمع این دو فلز در اندام‌های مختلف محصولات در اکثر موارد پایین‌تر از حدود مجاز قرارداشت. دامنه تغییرات فلزات سنگین در محصولات مختلف تحت تیمارهای آبیاری با پساب، ۱۴/۱-۳۷/۷، ۷۱/۱-۱۱۶۵، ۲۴/۳-۹۸/۶، ۰/۱۰-۲/۳۰ و ۱۵/۸-۲۶۹ میلی‌گرم در کیلوگرم به ترتیب برای مس، آهن، روی، کادمیم و منگنز اندازه‌گیری شد. علاوه‌بر این، تجمع فلزات مختلف در ریشه بیش از اندام‌های دیگر به دست آمد.

کلیدواژه‌ها: بامیه، پساب شهری تصفیه‌شده، جعفری، جو، ذرت، فلزات سنگین.