

هیدروژئومورفولوژی، شماره ۸، پاییز ۱۳۹۵، صص ۹۲-۷۳

وصول مقاله: ۱۳۹۴/۰۵/۲۶ تأیید نهایی مقاله: ۱۳۹۵/۰۷/۲۸

کاربرد مدل درختی M5 در پیش‌بینی خشکسالی (مطالعه‌ی موردی: مراغه، ایران)

محمدتقی ستاری^{*۱}

رسول میرعباسی نجف آبادی^۲

مسعود علیمحمدی^۳

چکیده

در کشورهای خشک و نیمه‌خشکی مانند ایران، پیش‌بینی دقیق خشکسالی‌ها، نقش بسیار مهمی در مقابله با بحران ناشی از خشکسالی و مدیریت سیستم‌های منابع آب ایفا می‌کند. با توجه به اینکه شاخص بارندگی استاندارد شده (SPI) به عنوان یکی از مناسب‌ترین شاخص برای تحلیل خشکسالی شناخته شده است، در این تحقیق، جهت پیش‌بینی SPI از مدل درختی M5 استفاده گردید. بدین منظور از داده‌های بارش ماهانه ایستگاه هم‌مدیدی مراغه در یک دوره‌ی ۲۵ ساله (۱۳۶۵-۸۹) استفاده و شاخص SPI در مقیاس ۶ ماهه استخراج گردید. نتایج نشان داد که شهرستان مراغه در دو دهه‌ی اخیر با خشکسالی‌های پی در پی و شدیدی مواجه بوده است. سپس با استفاده از مدل درختی M5 اقدام به پیش‌بینی مقادیر شاخص SPI در مقیاس زمانی ۶ ماهه برای ۱ تا ۱۲ ماه آینده گردید. نتایج به دست آمده نشان داد که میزان شاخص SPI در مقیاس‌های زمانی گذشته بیشترین تأثیر را نسبت به پارامترهای دیگر در پیش‌بینی شاخص بارش استاندارد

۱ - استادیار گروه مهندسی آب، دانشکده‌ی کشاورزی، دانشگاه تبریز (نویسنده‌ی مسئول). Email: mtsattar@gmail.com

۲ - استادیار گروه مهندسی آب، دانشکده‌ی کشاورزی، دانشگاه شهرکرد.

۳ - کارشناس ارشد مهندسی عمران - آب.

شده دارد و با افزایش طول دوره‌ی پیش‌بینی از دقت مدت کاسته می‌شود. به طوری که در محاسبه SPI6 برای یک ماه آینده مقدار ضریب همبستگی حدود ۰/۹۴ به دست آمد که این مقدار برای ۱۲ ماه آینده به حدود ۰/۴۰ کاهش پیدا کرد. با این وجود نتایج نشان داد که مدل درختی M5 با ارایه‌ی روابط خطی کاربردی و قابل فهم از دقت و توانایی نسبتاً بالایی در پیش‌بینی خشکسالی برخوردار است.

کلمات کلیدی: خشکسالی، داده کاوی، مدل درختی M5، مراغه.

مقدمه

تأثیر خشکسالی بر بخش‌های مختلف، یکسان نمی‌باشد. در شرایطی که منطقه‌ای با کاهش محسوس بارندگی مواجه شده باشد، این امکان وجود دارد که تأثیر آن بر منابع آب هنوز ملموس نشده و یا کاهش منابع آب‌های سطحی تأثیری بر تولیدات کشاورزی نداشته باشد. از این رو وجود نیازها و زمینه‌های کاری مختلف موجب شده است که محققان با تمرکز بر انواع مختلف کمبودها، خشکسالی‌ها را طبقه‌بندی کنند. برای این اساس خشکسالی‌ها به چهار طبقه‌ی عمده‌ی خشکسالی شامل هواشناختی، کشاورزی، هیدرولوژی و اقتصادی-اجتماعی تقسیم شده‌اند (دراکوپ و همکاران^۱، ۱۹۸۰). برای مثال، خشکسالی‌های هواشناسی بر اساس کمبود در بارش، خشکسالی‌های هیدرولوژیکی بر اساس کمبود در جریان رودخانه و خشکسالی‌های کشاورزی بر اساس کمبود در رطوبت خاک تعریف می‌شوند. خشکسالی اقتصادی-اجتماعی هنگامی بروز می‌یابد که میزان تقاضای یک کالای اقتصادی بیشتر از عرضه‌ی آن باشد که این عدم تعادل (میان عرضه و تقاضا) می‌تواند ناشی از کمبود منابع آب در ارتباط با وضع هوا می‌باشد. در این نوع خشکسالی تنها مقدار منابع

1- Dracup et al.,

آب مورد توجه قرار نمی‌گیرد، بلکه تقاضا برای آب نیز مدنظر قرار خواهد گرفت. محققان اغلب برای ارزیابی خشکسالی‌ها از شاخص‌های خشکسالی مختلف که از متغیرهای هیدروکلیماتولوژی مشتق شده‌اند، استفاده می‌کنند و تصمیم‌گیری در مورد مدیریت خشکسالی‌ها بر اساس این شاخص‌ها صورت می‌پذیرد (میشرا و سینگ، ۲۰۱۰). تاکنون شاخص‌های زیادی توسط محققان برای پیش‌بینی خشکسالی بسط و گسترش یافته است. یکی از شاخص‌های معروف که به طور گسترده در سراسر دنیا استفاده می‌شود، شاخص بارش استاندارد شده^۱ (SPI)، می‌باشد که در این مطالعه نیز مورد بررسی قرار گرفته است. این شاخص دارای مزایای زیادی از جمله سادگی محاسبات می‌باشد. چرا که SPI بر پایه‌ی بارش قرار دارد، در حالی که شاخص‌های دیگر پیچیده‌تر بوده و محاسبه‌ی آن‌ها مستلزم فرآیندهای طولانی‌تری است. مزیت دیگر این شاخص انعطاف‌پذیری زمانی آن است، چرا که در هر مقیاس زمانی قابل محاسبه می‌باشد. این ویژگی باعث شده است تا SPI قابلیت پیش‌توأم اقلیمی، هیدرولوژیکی و کشاورزی را داشته باشد.

تاکنون روش‌های متعددی توسط محققان برای پیش‌بینی شاخص‌های خشکسالی ارائه شده است که از جمله می‌توان به زنجیره مارکوف (پائولو همکاران^۲، ۲۰۰۵) و مدل های لوگ خطی (موریرا و همکاران^۳، ۲۰۰۸) اشاره نمود. در سال‌های اخیر، استفاده از روش‌های داده‌کاوی بسیار مورد توجه محققان علوم مختلف از جمله مهندسان آب قرار گرفته است. در حال حاضر، داده‌کاوی مهم‌ترین فناوری جهت بهره‌برداری مؤثر از داده‌های حجیم است و اهمیت آن رو به فزونی است (ستاری و همکاران، ۲۰۱۲a).

1- Standardized Precipitation Index (SPI)

2- Paulo et al.,

3- Moreira et al.,

باتاچاریا و سولماتین^۱ (۲۰۰۵) در مطالعه‌ای کاربرد شبکه‌ی عصبی مصنوعی و مدل درختی M5 در پیش‌بینی منحنی دبی-اشل (منحنی سنج‌ی آب) در رودخانه‌ی Bhagirathi واقع در کشور هند را مورد بررسی قرار دادند. در این تحقیق، مقایسه‌ی مدل‌های معمولی آماری دبی-اشل با روش شبکه‌ی عصبی مصنوعی و مدل درختی M5 نشان داد که روش شبکه‌ی عصبی مصنوعی و مدل درختی M5 قابلیت بهتری در پیش‌بینی منحنی دبی-اشل دارا می‌باشند. یوره‌کلی و همکاران^۲ (۲۰۱۲) با استفاده از روش تصمیم‌گیری درختی، خشکسالی منطقه‌ای فصلی و سالانه را با استفاده از داده‌های سری بارندگی تجمعی برای ۵ دوره‌ی مرجع (۴ دوره‌ی فصلی و ۱ دوره‌ی سالیانه) از ۱۷ ایستگاه در حوضه‌ی Cekerek ترکیه بر اساس شاخص SPI مورد بررسی قرار دادند. نتایج نشان داد که تفاوت قابل ملاحظه‌ای میان دسته‌بندی خشکسالی محاسبه شده با الگوریتم متعارف SPI و روش تصمیم‌گیری درختی وجود ندارد. ستاری و همکاران (۲۰۱۲) با استفاده از روش تصمیم‌گیری درختی نسبت به تعیین دوره‌های خشکسالی ممکن در آنکارا اقدام نمودند. نتایج به دست آمده حاکی از آن است که استان آنکارا دارای آب و هوایی معتدل و خشک نزدیک به معتدل است و مقادیر بارش در تمام فصل‌ها و به خصوص ماه ژانویه جهت تعیین خشکی باید در نظر گرفته شود. زهیری و همکاران (۲۰۱۴) عملکرد روش برنامه‌ریزی ژنتیک خطی و مدل درختی M5 را در پیش‌بینی تخلیه‌ی جریان در کانال‌های مرکب را مورد مقایسه قرار دادند. آن‌ها از دو مدل جدید LGP و M5 برای برآورد ظرفیت انتقال جریان استفاده کردند که مقایسه‌ی تخلیه‌ی جریان محاسبه شده و مشاهده شده نشان داد گرچه هر دو مدل دقت بالایی در پیش‌بینی ظرفیت جریان دارند، ولی مدل LGP با $R=0/989$ عملکرد بسیار خوبی داشته است.

1- Bhattacharya and Solomatine

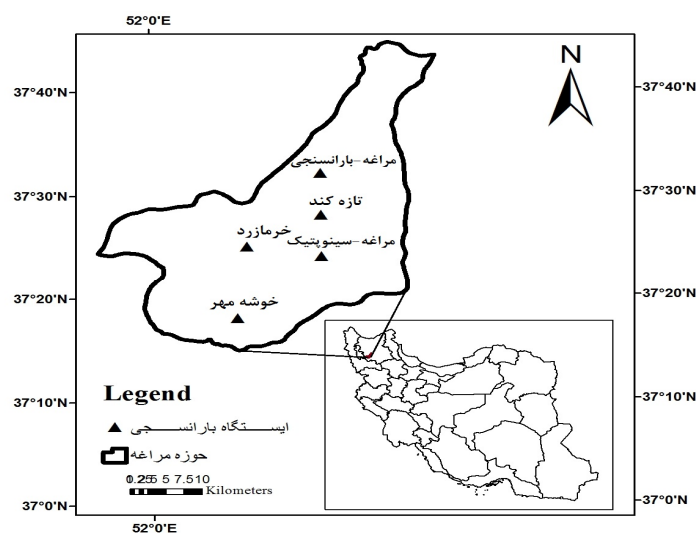
2- Yurekli et al.,

بررسی منابع صورت گرفته نشان می‌دهد تاکنون در ایران، مطالعه‌ای در زمینه کاربرد مدل درختی M5 در پیش‌بینی خشکسالی صورت نگرفته است. لذا هدف اصلی این مطالعه، پیش‌بینی خشکسالی‌های شهرستان مراغه بر اساس شاخص SPI در مقیاس زمانی ۶ ماهه با استفاده از مدل‌های درختی M5 می‌باشد.

مواد و روش‌ها

معرفی منطقه‌ی مورد مطالعه و داده‌های مورد استفاده

شهرستان مراغه در جنوب غربی استان آذربایجان شرقی قرار گرفته است. این شهرستان با وسعتی معادل ۲۱۸۵/۶۵ کیلومتر مربع در گستره جغرافیایی ۳۷ درجه و ۱ دقیقه الی ۳۷ درجه و ۴۵ دقیقه عرض شمالی و ۴۶ درجه و ۹ دقیقه الی ۴۶ درجه و ۴۴ دقیقه طول شرقی واقع شده است. در شکل (۱) موقعیت شهرستان مراغه در استان آذربایجان شرقی نشان داده شده است.



شکل (۱) موقعیت جغرافیایی منطقه‌ی مورد مطالعه و ایستگاه‌های بارانسنجی

در این مطالعه، به منظور محاسبه شاخص SPI از داده‌های ۴ ایستگاه بارانسنجی و ۱ ایستگاه سینوپتیک که دارای حداقل ۲۵ سال (دوره‌ی ۸۹-۱۳۶۵) داده ماهانه می‌باشند، استفاده شد. از ایستگاه سینوپتیک اداره هواشناسی شهرستان مراغه که دارای داده‌های هواشناسی با طول دوره‌ی آماری مناسب می‌باشد، به عنوان ایستگاه منتخب در این تحقیق استفاده گردید.

جدول (۱) ویژگی‌های آماری داده‌های بارش ماهانه‌ی ایستگاه سینوپتیک مراغه در دوره‌ی ۸۹۰-

۱۳۶۵ بر حسب میلی متر			
میانگین	کمینه	بیشینه	انحراف معیار
۲۵/۵۶	۰	۱۳۷/۵	۲۷/۴

شاخص بارش استاندارد شده (SPI)

اولین مرحله‌ی محاسبه‌ی شاخص SPI، برازش تابع توزیع احتمالی گاما بر مقادیر بارندگی ایستگاه مورد نظر می‌باشد. پس از آن پارامترهای مربوط به این تابع برای هر مقیاس زمانی دلخواه برآورد می‌گردند. در نهایت، تابع توزیع تجمعی مربوطه محاسبه و به یک تابع توزیع تجمعی نرمال جهت محاسبه SPI تبدیل می‌شود. تابع چگالی توزیع گامای دو پارامتری به صورت زیر می‌باشد:

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad (۱) \text{ رابطه‌ی}$$

که در رابطه‌ی (۱) پارامتر شکل، β پارامتر مقیاس، x مقدار بارش در مقیاس زمانی مشخص و $\Gamma(\alpha)$ تابع گاما می‌باشد که به صورت زیر تعریف می‌شود:

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy \quad (۲) \text{ رابطه‌ی}$$

پس از برازش تابع گاما بر داده‌های ایستگاه مورد مطالعه، احتمال تجمعی برای هر رویداد بارش در مقیاس‌های زمانی مختلف به صورت زیر محاسبه می‌شود:

$$F(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^x x^{\alpha-1} e^{-\frac{x}{\beta}} dx \quad (۳) \text{ رابطه‌ی}$$

از آنجائی که تابع توزیع گاما در نقطه صفر تعریف نشده و مقادیر بارش ممکن است شامل داده صفر نیز باشند، لذا در این حالت احتمال تجمعی از رابطه‌ی زیر محاسبه می‌گردد (اصغری تبریزی و همکاران، ۲۰۱۰):

$$H(x) = q + (1-q).F(x) \quad (۴) \text{ رابطه‌ی}$$

که در آن q احتمال وقوع بارش ماهانه صفر است. در مرحله بعد احتمال تجمعی $H(x)$ به متغیر نرمال استاندارد Z با میانگین صفر و واریانس ۱ تبدیل می‌گردد که برابر با شاخص SPI می‌باشد. در این مطالعه، برای تبدیل احتمال تجمعی گاما به متغیر نرمال استاندارد از روش پیشنهادی آبراموویتز و استگان^۱ (۱۹۶۵) استفاده گردید. در این روش مقادیر متغیر نرمال استاندارد (Z) به ترتیب زیر محاسبه می‌شود:

$$Z = SPI = - \left(t - \frac{C_0 + C_1.t + C_2.t^2}{1 + d_1.t + d_2.t^2 + d_3.t^3} \right) \quad \begin{array}{l} \text{رابطه‌ی (۵) برای} \\ 0 < H(x) \leq 0.5 \end{array}$$

$$Z = SPI = + \left(t - \frac{C_0 + C_1.t + C_2.t^2}{1 + d_1.t + d_2.t^2 + d_3.t^3} \right) \quad \begin{array}{l} \text{رابطه‌ی (۶) برای} \\ 0.5 < H(x) < 1 \end{array}$$

که در رابطه‌های فوق:

$$t = \sqrt{\ln\left(\frac{1}{(H(x))^2}\right)} \quad \circ < H(x) \leq 0.5 \quad (۷) \text{ رابطه‌ی}$$

$$t = \sqrt{\ln\left(\frac{1}{(1-H(x))^2}\right)} \quad 0.5 < H(x) < 1 \quad (۸) \text{ رابطه‌ی}$$

که در روابط فوق C_0, C_1, C_2, d_1, d_2 و d_3 به ترتیب برابر $۲/۵۱۵۵۱۷$ ، $۰/۸۰۲۸۵۳$ ، $۰/۱۰۳۲۸$ ، $۱/۴۳۲۷۸۸$ ، $۰/۱۱۹۲۶۹$ و $۰/۰۰۱۳۰۸$ می‌باشند.

طبقه‌بندی شدت خشکسالی بر اساس مقادیر SPI در جدول (۲) آورده شده است. در این مطالعه، بر طبق توصیه‌ی لوکاس و واسیلیادز^۱ (۲۰۰۴) رویداد خشکسالی به صورت دوره‌ای که در آن مقادیر SPI کمتر از صفر هستند، تعریف شد.

جدول (۲) طبقه‌بندی خشکسالی بر اساس شاخص بارش استاندارد شده (مک کی و همکاران،

(۱۹۹۳)

توصیف وضعیت	طبقات شاخص SPI
خشکسالی ملایم	۰ تا ۰/۹۹-
خشکسالی متوسط	۱- تا ۱/۴۹-
خشکسالی شدید	۱/۵- تا ۱/۹۹-
خشکسالی بسیار شدید	کمتر از ۲-

مدل درخت M5

اساس مدل‌های درختی مبتنی بر روش تصمیم و غلبه کردن هستند. جایگزینی معادله‌ی رگرسیون خط به جای برچسب در گره‌ها، شیوه‌ای است که در مدل M5

1- Loukas, A., Vasiliades

اجرا می‌شود که می‌تواند متغیرهای عددی پیوسته را پیش‌بینی یا برآورد کند. ساختار درخت تصمیم‌گیری شبیه یک درخت بوده که از ریشه^۱، شاخه‌ها^۲، گره‌ها^۳ و برگ‌ها^۴ تشکیل یافته است. ساخت مدل درخت در دو مرحله انجام می‌گیرد. در مرحله اول، درخت تصمیم با انشعاب‌سازی داده‌ها تشکیل می‌شود. معیار انشعاب در مدل M5 بیشینه‌سازی کاهش انحراف معیار داده‌ها در گره فرزند است. زمانی که امکان کاهش انحراف معیار داده‌های گره فرزند میسر نبود، گره والد آن منشعب نشده و به گره پایانی و یا برگ رسیده است. فرمول مورد استفاده برای محاسبه کاهش انحراف معیار استاندارد عبارت است از (آلبرگ و همکاران^۵، ۲۰۱۲):

$$SDR = Sd(T) - \sum_{i=1}^N \frac{|T_i|}{|T|} Sd(T_i) \quad \text{رابطه‌ی (۹)}$$

که در رابطه‌ی (۹)، T نشان‌دهنده‌ی مجموعه‌ای از نمونه‌هاست که بر هر گره وارد شده‌اند، T_i نشان‌دهنده‌ی زیر مجموعه‌ای از نمونه‌هاست که i امین نتیجه تست پتانسیلی را دارند و Sd انحراف معیار استاندارد داده‌های ورودی است و به صورت زیر محاسبه می‌شود (آلبرگ و همکاران، ۲۰۱۲).

$$Sd(T) = \sqrt{\frac{1}{N} \left(\sum_{i=1}^N y_i^2 - \frac{1}{N} \left(\sum_{i=1}^N y_i \right)^2 \right)} \quad \text{رابطه‌ی (۱۰)}$$

در رابطه‌ی (۱۰)، Y_i مقدار عددی ویژگی هدف نمونه i و N تعداد داده‌ها را نشان می‌دهد. از آنجایی که فرآیند انشعاب (دسته‌بندی) در گره‌های فرزند دارای

-
- 1- Root
 - 2- Branches
 - 3- Nodes
 - 4- Leaves
 - 5- Alberg

انحراف معیار کمتری نسبت به گره‌های والد است، از این رو دارای نتیجه دقیق‌تری هستند و از همگنی بیشتری برخوردارند. پس از بررسی همه دسته‌بندی‌های ممکن مدل M5 یکی را انتخاب می‌کند که دارای حداقل خطای مورد انتظار باشد. مرحله‌ی دوم طراحی مدل درختی، شامل کوچک کردن درخت بیش از حد بزرگ شده از طریق هرس کردن شاخه‌ها و جایگزین شدن با توابع رگرسیون خطی است. در این مطالعه از نرم‌افزار WEKA برای ایجاد مدل درختی استفاده شد (ویتن^۱، ۲۰۱۱). این نرم‌افزار مجموعه‌ای از بروزترین الگوریتم‌های یادگیری ماشینی و ابزارهایی برای پیش‌پردازش داده‌ها می‌باشد.

معیارهای ارزیابی

هدف از کاربرد معیارهای ارزیابی، برآورد میزان خطا در محاسبات و تعیین الگو و ساختاری است که کمترین میزان خطا را در پیش‌بینی شاخص SPI داشته باشد. در این مطالعه، از دو معیار ضریب همبستگی R و ریشه‌ی میانگین مربعات خطا (RMSE) استفاده شده است که به صورت زیر محاسبه می‌شوند (میرعباسی و همکاران، ۲۰۱۲):

$$R = 1 - \frac{\sum_{i=1}^n (SPI - \widehat{SPI}_m)}{\sum_{i=1}^n (SPI - \overline{SPI}_0)} \quad \text{رابطه‌ی (۱۴)}$$

$$RMSE = \left[\frac{\sum (SPI - \overline{SPI}_0)^2}{N} \right]^{1/2} \quad \text{رابطه‌ی (۱۵)}$$

که در رابطه‌ی فوق SPI: مقدار SPI مشاهداتی، \overline{SPI}_0 مقدار SPI محاسباتی، \widehat{SPI}_m میانگین مقدار SPI مشاهداتی و N تعداد سری داده‌ها می‌باشند. در مورد

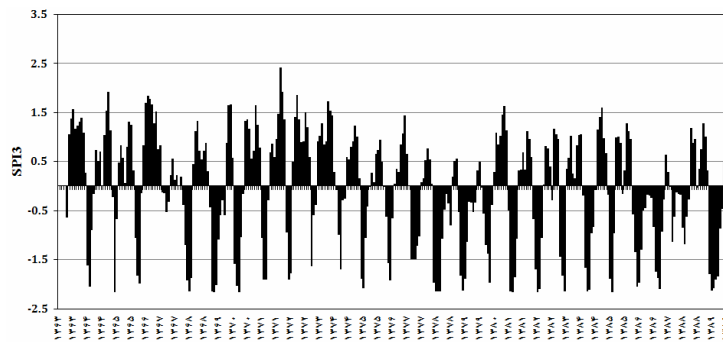
ضریب همبستگی R، هر چقدر به یک نزدیک‌تر باشد، مدل دارای دقت بیشتری است و هرچه مقادیر خطای ریشه‌ی میانگین مربعات RMSE کمتر و به صفر نزدیک‌تر باشد، میزان دقت مدل بیشتر خواهد بود.

بحث و نتایج

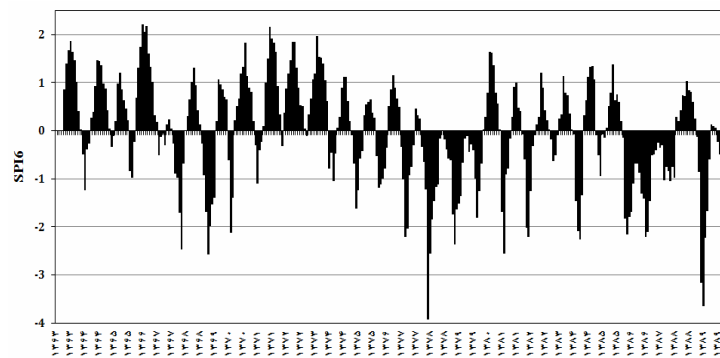
در این مطالعه، از بین مقیاس‌های مختلف شاخص SPI، مقیاس ۶ ماهه مورد بررسی قرار گرفت و با مدل درختی پیش‌بینی شد. دلیل انتخاب شاخص SPI در مقیاس ۶ ماهه این است که این مقیاس میان مدت می‌باشد و می‌تواند به طور متوسط شرایط رطوبتی منطقه را توصیف کند. شاخص SPI در مقیاس‌های ۳ و ۶ ماهه برای ایستگاه مراغه در شکل (۲) آورده شده است. با دقت در این نمودارها مشخص می‌شود که با افزایش مقیاس زمانی در شاخص SPI از نوسانات آن کاسته شده و در مقابل در خشکسالی‌های با دوره‌ی طولانی‌تر، شدت خشکسالی ممکن است بیشتر نشان داده شود. به عنوان مثال، در یک دوره‌ی شش ماهه کمبود بارندگی، مقدار این شاخص در انتهای دوره‌ی کم‌بارش بیش از مقدار آن در زمان مشابه و در مقیاس سه ماهه خواهد بود. چرا که در مقیاس شش ماهه اثر تمام دوره‌ی کم‌بارش لحاظ ولی در مقیاس سه ماهه تنها اثر دو ماه ماقبل آخر در نظر گرفته می‌شود (شکل ۲).

– تهیه الگوی ورودی مدل‌ها

برای بدست آوردن پارامترهای مناسب برای ورود به مدل و حذف پارامترهای ناکارآمد میزان همبستگی پارامترها با شاخص SPI مورد نظر را به دست آورده و از پارامترهایی که همبستگی قابل قبولی با شاخص SPI داشتند، استفاده گردید. همچنین از مقدار همبستگی بین ورودی‌ها با تأخیر در گام‌های زمانی مختلف استفاده گردید (جدول ۳).



(الف)



(ب)

شکل (۲) شاخص SPI در مقیاس‌های ۳ و ۶ ماهه برای ایستگاه مراغه (الف و ب).

– نتایج پیش‌بینی شاخص SPI6

در این مقاله به بررسی نتایج پیش‌بینی شاخص بارش استاندارد ۶ ماهه برای زمان‌های ۱، ۳، ۶ و ۱۲ ماهه پرداخته شده است. برای رعایت اصل خلاصه‌نویسی فقط نتایج مربوط به پیش‌بینی شاخص SPI6 برای ۱ و ۱۲ ماه آینده ارائه خواهد شد. بر اساس میزان همبستگی پارامترهای گوناگون با شاخص بارش استاندارد ۶ ماهه، ۷ سناریو به عنوان ورودی مدل درختی M5 تعریف گردید.

جدول (۳) ضرایب همبستگی بین شاخص SPI6 و کل متغیرها

ضریب همبستگی	پارامتر	ضریب همبستگی	پارامتر	ضریب همبستگی	پارامتر	ضریب همبستگی	پارامتر	ضریب همبستگی	پارامتر
۰/۱۷۲۶	R(t-4)	۰/۱۹۵۴	R(t-3)	۰/۰۳۰۱	tmean(t-2)	۰/۰۱۷	R(t-1)	۰/۳۶۲۱	R(t)
۰/۳۴۳۹	Tmean(t-4)	۰/۱۹۹۳	tmean(t-3)	۰/۱۴۴۶	tmean(t-2)	۰/۰۲۱۹	tmean(t-1)	۰/۱۹۸۸	Tmean(t)
۰/۳۶۵۷	Rhmean(t-4)	۰/۳۸۸۹	Rhmean(t-3)	۰/۱۵۲۹	Rhmean(t-2)	۰/۱۴	Rhmean(t-1)	۰/۴۶۰۴	rhmean(t)
۰/۲۹۲۱	SPI1(t-4)	۰/۲۸۹۵	SPI1(t-3)	۰/۲۱۶۶	SPI1(t-2)	۰/۰۴۴۷	SPI1(t-1)	۰/۳۷۹۲	SPI1(t)
۰/۳۱۵۱	SPI3(t-4)	۰/۲۰۴۴	SPI3(t-3)	۰/۰۶۷۱	SPI3(t-2)	۰/۳۹۲۲	SPI3(t-1)	۰/۷۰۰۸	SPI3(t)
۰/۰۸۱۸	SPI6(t-4)	۰/۳۳۳	SPI6(t-3)	۰/۵۹۷۳	SPI6(t-2)	۰/۸۴۹	SPI6(t-1)	۱	SPI6(t)

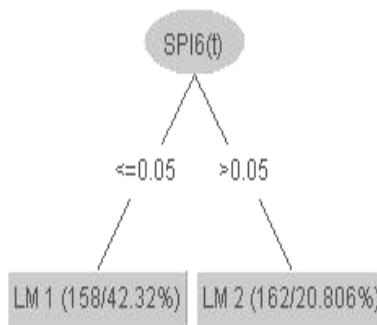
در جدول (۳) نمادهای R نشان‌دهنده‌ی بارش ماهانه، tmean نشان‌دهنده‌ی دمای متوسط ماهانه، rhmean نشانگر رطوبت متوسط ماهانه و t گام زمانی است.

```

LM num: 1
SPI6(t+1) =
0.0007 * t_mean(t-4)
- 0.0165 * rain(t)
- 0.012 * rh_mean(t)
+ 0.8788 * SPI6(t)
+ 0.7981

LM num: 2
SPI6(t+1) =
0.0136 * t_mean(t-4)
- 0.0072 * rain(t)
- 0.0158 * rh_mean(t)
+ 1.0934 * SPI6(t)
+ 0.6964

```



شکل (۳) مدل درختی و روابط به دست آمده از مدل درختی M5 جهت پیش‌بینی شاخص SPI6 برای یک ماه آینده

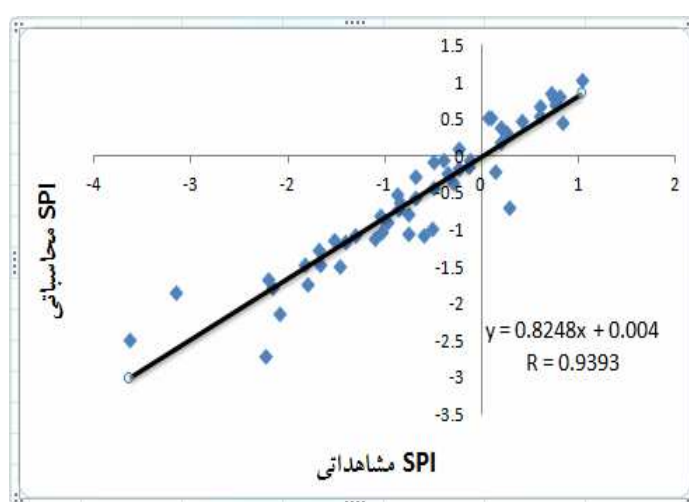
پیش‌بینی شاخص SPI6 برای ۱ ماه آینده با استفاده از مدل درختی M5

برای پیش‌بینی ۱ ماه آینده شاخص SPI6، سناریوی شماره‌ی ۴ با ۸۰ درصد داده‌ها به عنوان داده‌های آموزشی و تولید ۱ رابطه‌ی خطی بهترین نتیجه را ارائه نمود. در شکل (۳) نمودار درختی و روابط خطی ایجاد شده آورده شده است.

جدول (۴) نتایج سناریوهای مختلف جهت پیش‌بینی شاخص SPI6 برای ۱ ماه آینده با استفاده از مدل درختی M5

شماره سناریو	ورودی	میزان آموزش (%)	تعداد قوانین	میزان خطا	
				R	RMSE
S 1	R(t)	۶۶	۱	۰/۰	۱/۱۳۶
		۷۰		۰/۰	۱/۱۴۶۴
		۷۵		۰/۰	۱/۱۷۰۸
		۸۰		۰/۰	۱/۲۷۲
S 2	(t), R(t) و SPI	۶۶	۲	۰/۹۱۱۲	۰/۴۴۳۵
		۷۰		۰/۹۰۹۹	۰/۴۵
		۷۵		۰/۹۱۴۲	۰/۴۵۱۳
		۸۰		۰/۹۲۸۸	۰/۴۲۹۸
S 3	SPI6(t), R(t), tmean(t-4),	۶۶	۲	۰/۹۱۸۵	۰/۴۲۲۳
		۷۰		۰/۹۱۷۳	۰/۴۲۸۱
		۷۵		۰/۹۲۴۸	۰/۴۱۸۱
		۸۰		۰/۹۲۱۸	۰/۴۳۵۱
S 4	SPI6(t), R(t), tmean(t-4), rhmean(t)	۶۶	۲	۰/۹۲۳۸	۰/۴۰۸۷
		۷۰		۰/۹۲۸۲	۰/۴۰۱۳
		۷۵		۰/۹۳۱۲	۰/۴۰۱۴
		۸۰		۰/۹۳۹۳	۰/۳۷۹۷
S 5	SPI6(t), R(t), tmean(t-4), rhmean(t), SPI6(t-1)	۶۶	۲	۰/۹۲۶۱	۰/۴۰۱۴
		۷۰		۰/۹۲۵۳	۰/۴۰۷۸
		۷۵		۰/۹۲۸۱	۰/۴۰۸۸
		۸۰		۰/۹۳۴۷	۰/۳۹۰۸
S 6	SPI6(t), R(t), tmean(t-4), rhmean(t), SPI6(t-1), SPI3(t)	۶۶	۲	۰/۹۲۶۲	۰/۴۰۱
		۷۰		۰/۹۲۷۸	۰/۴۰۳۳
		۷۵		۰/۹۲۸۴	۰/۴۰۹۱
		۸۰		۰/۹۳۴۲	۰/۴۰۳۳
S 7	SPI6(t), R(t), tmean(t-4), rhmean(t), SPI6(t-1), SPI3(t), SPI1(t)	۶۶	۲	۰/۹۲۶۲	۰/۴۰۱
		۷۰		۰/۹۲۷۸	۰/۴۰۳۳
		۷۵		۰/۹۲۸۴	۰/۴۰۹۱
		۸۰		۰/۹۳۴۲	۰/۴۰۳۳

در شکل (۴) نمودار پراکنش مقادیر مشاهداتی و محاسباتی SPI6 برای ۱ ماه آینده آورده شده است. همانگونه که مشاهده می‌شود، توافق خوبی بین مقادیر محاسباتی و مشاهداتی وجود دارد.



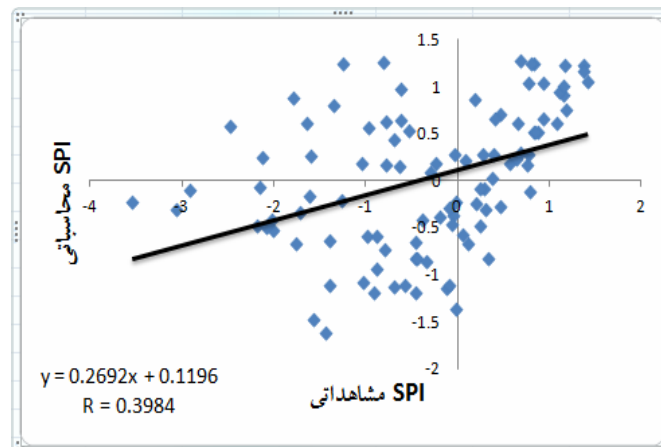
شکل (۴) نمودار پراکنش SPI6 مشاهداتی و محاسباتی برای یک ماه آینده

پیش‌بینی شاخص SPI6 برای ۱۲ ماه آینده با استفاده از مدل درختی M5 برای پیش‌بینی ۱۲ ماه آینده شاخص SPI6، سناریوی ۴ با ۶۶ درصد داده‌ها به عنوان داده‌های آموزشی و تولید یک رابطه‌ی خطی بهترین نتیجه را ارائه نمود. در شکل (۵) مدل درختی و روابط به دست آمده از مدل درختی M5 جهت پیش‌بینی شاخص SPI6 برای ۱۲ ماه ارائه شده است. در معادله‌ی فوق که برای همه شرایط صادق می‌باشد، SPI6(t) دارای ضریب بالاتری نسبت به دیگر پارامترها می‌باشد که نشان‌دهنده‌ی تأثیر زیاد این پارامتر در پیش‌بینی شاخص SPI6 برای ۱۲ ماه آینده دارد.

$$\begin{aligned}
 \text{LM num: 1} \\
 \text{SPI6}(t+12) = & \\
 & 0.0351 * t_meant(t-4) \\
 & + 0.0093 * rh_mean(t) \\
 & + 0.3064 * \text{SPI6}(t) \\
 & - 0.9342
 \end{aligned}$$

شکل (۵) مدل درختی و روابط به دست آمده از مدل درختی M5 جهت پیش‌بینی شاخص SPI6 برای ۱۲ ماه آینده

در شکل (۶) نمودار پراکنش مقادیر مشاهداتی و محاسباتی SPI6 برای ۱۲ ماه آینده آورده شده است. همانگونه که مشاهده می‌شود، بهترین سناریو در این حالت نتوانسته است نتایج دقیقی را پیش‌بینی نماید و توافق قابل قبولی بین مقادیر محاسباتی و مشاهداتی وجود ندارد. به نظر می‌رسد دلیل آن این است که در پیش‌بینی‌های بلندمدت باید از شاخص بلندمدت استفاده نمود، مثل SPI12 یا SPI24 و نمی‌توان شاخص میان مدت SPI6 را برای دوره‌ی بلندمدت به کار برد.



شکل (۶) نمودار پراکنش SPI6 مشاهداتی و محاسباتی برای ۱۲ ماه آینده

جدول (۵) نتایج سناریوهای مختلف جهت پیش‌بینی شاخص SPI6 برای ۱۲ ماه آینده با استفاده از مدل درختی M5

شماره سناریو	ورودی	میزان آموزش (%)	تعداد قوانین	میزان خطا	
				R	RMSE
S 1	R(t)	۶۶	۲	-۰/۱۱۳۷	۱/۱۷۶۲
		۷۰		-۰/۰۸۸۵	۱/۲۵۷۲
		۷۵		-۰/۰۱۲۶	۱/۲۹۱۲
		۸۰		-۰/۰۰۵	۱/۳۵۰۹
S 2	(t), R(t) & SPI	۶۶	۱	-۰/۲۶۶۵	۱/۱۲۹۸
		۷۰		-۰/۲۲۷۴	۱/۲۰۳۸
		۷۵		-۰/۱۵۴۴	۱/۲۵۷۴
		۸۰		-۰/۰۵۳۷	۱/۲۹۴
S 3	SPI6(t), R(t), tmean(t-4)	۶۶	۱	-۰/۳۸۴۳	۱/۱۰۶۱
		۷۰		-۰/۳۵۸۶	۱/۱۸۱۷
		۷۵		-۰/۲۹۳۲	۱/۲۵۳۶
		۸۰		-۰/۲۰۵۴	۱/۳۰۵۶
S 4	SPI6(t), R(t), tmean(t-4), rhmean(t)	۶۶	۱	-۰/۳۹۸۴	۱/۱۱۵۱
		۷۰		-۰/۳۷۰۴	۱/۱۹۱۳
		۷۵		-۰/۲۸۸۷	۱/۲۷
		۸۰		-۰/۱۹۸۷	۱/۳۵۱۹
S 5	SPI6(t), R(t), tmean(t-4), rhmean(t), SPI6(t-1)	۶۶	۱	-۰/۳۹۸۴	۱/۱۱۵۱
		۷۰		-۰/۳۷۰۴	۱/۱۹۱۳
		۷۵		-۰/۲۸۸۷	۱/۲۷
		۸۰		-۰/۱۹۸۷	۱/۳۵۱۹
S 6	SPI6(t), R(t), tmean(t-4), rhmean(t), SPI6(t-1), SPI3(t)	۶۶	۱	-۰/۳۹۸۴	۱/۱۱۵۱
		۷۰		-۰/۳۷۴	۱/۱۹۱۳
		۷۵		-۰/۲۸۸۷	۱/۲۷
		۸۰		-۰/۱۹۸۷	۱/۳۵۱۹
S 7	SPI6(t), R(t), tmean(t-4), rhmean(t), SPI6(t-1), SPI3(t), SPI1(t)	۶۶	۱	-۰/۳۹۸۴	۱/۱۱۵۱
		۷۰		-۰/۳۷۰۴	۱/۱۹۱۳
		۷۵		-۰/۲۸۸۷	۱/۲۷
		۸۰		-۰/۱۹۸۷	۱/۳۵۱۹

نتیجه‌گیری

از آنجا که خشکسالی، بخش‌های مختلف جامعه مانند منابع آب، کشاورزی، صنعت، اقتصاد، بهداشت و... را تحت تأثیر قرار می‌دهد، لذا پایش و ارزیابی این بلیه‌ی طبیعی در آینده به منظور ارایه‌ی برنامه‌ریزی صحیح در بخش‌های مختلف جامعه، امری لازم و ضروری است. در این تحقیق، میزان کارایی مدل درختی M5 در پیش‌بینی خشکسالی‌های شهرستان مراغه مورد ارزیابی قرار گرفت. بدین منظور از داده‌های بارش ماهانه‌ی ایستگاه مراغه در یک دوره‌ی ۲۵ ساله استفاده شد و شاخص SPI در مقیاس ۶ ماهه استخراج گردید. سپس با استفاده از مدل درختی M5 اقدام به پیش‌بینی مقادیر شاخص SPI در مقیاس زمانی ۶ ماهه برای ۱ تا ۱۲ ماه آینده گردید. نتایج نشان داد که میزان شاخص SPI در مقیاس‌های زمانی گذشته بیشترین تأثیر را نسبت به پارامترهای دیگر در پیش‌بینی شاخص بارش استاندارد شده دارد، ولی در بهترین سناریو، با افزایش طول دوره‌ی پیش‌بینی از دقت نتایج کاسته می‌شود. براساس بررسی منابع انجام گرفته تاکنون مشخص گردید، مطالعه‌ی مستقیم در زمینه‌ی کاربرد مدل درختی M5 جهت پیش‌بینی خشکسالی انجام نیافته است و لذا مقاله‌ای جهت مقایسه یافت نشد. ولیکن نتایج نشان داد این مطالعه از نظر دقت قابل مقایسه با مقالات ستاری و همکاران (۲۰۱۳a) در زمینه‌ی پیش‌بینی تبخیر- تعرق مرجع در آنکارا و همچنین مقاله ستاری و همکاران (۲۰۱۳b) در زمینه پیش‌بینی دبی جریان در رودخانه‌ی سوهو واقع در کشور ترکیه، می‌باشد. نهایتاً براساس نتایج به دست آمده مشخص گردید که مدل درختی M5 با ارائه‌ی روابط خطی کاربردی و قابل فهم از توانایی نسبتاً بالایی در پیش‌بینی خشکسالی برخوردار است.

منابع

- Bhattacharya, B. and Solomatine, D.P. (2005), **Neural Networks and M5 Model Trees in Modeling Water Level Discharge Relationship for an Indian river**, Publisher Elsevier Science Publishers B.V. Amsterdam, The Netherlands, The Netherlands, PP. 381-396.
- Dracup JA, Lee KS, and Paulson EGR, (1980), **On the Definition of Droughts**, Water Resources Research, 16: PP. 297-302.
- Huang, J., Van den Dool, H.M., Georgarakos, K.P., (1996), **Analysis of Model-calculated Soil Moisture over the United States (1931-1993) and Applications to Long-range Temperature Forecasts**, Journal of Climate, 9(6): PP. 1350-1362.
- Kogan, F.N., 1995, Droughts of the late (1980s), **In the United States as Derived from NOAA Polar-Orbiting Satellite Data**, Bulletin of the American Meteorological Society, 76(5): PP. 655-668.
- Londhe, S.N. Dixit, R.R., (2011), **Forecasting Stream Flow Using model Trees**, International Journal of Earth Sciences and Engineering, 4(6): 282-285.
- Loukas, A., Vasiliades, L., (2004), **Probabilistic Analysis of Drought Spatiotemporal Characteristics in Thessaly Region, Greece**, Natural Hazards and Earth System Sciences, 4: PP. 719-731
- Mishra AK and Singh VP, (2010), **A Review of Drought Concepts**, Journal of Hydrology 391: PP. 202-216.
- Moreira, E.E., Coelho, C.A., Paulo, A.A., Pereira, L.S., and Mexia, J.T., (2008), **SPI-based Drought Category Prediction Using Log Linear Models**, Journal of Hydrology, 354: PP. 116-130.
- Nalbantis, I., (2008), **Evaluation of a Hydrological Drought Index**, European Water, 23/24: PP. 67-77.
- Paulo, A.A., Ferreira, E., Coelho, C., Pereira, L.S, (2005), **Drought Class Transition Analysis through Markov and Loglinear**

- Models, an Approach to Early Warning**, Agricultural Water Management, 77: PP. 59-81.
- Quinlan JR, (1992), **Learning with Continuous Classes**, Proceedings of Australian Joint Conference on Artificial Intelligence, Singapore: World Scientific Press, PP. 343-348.
- Sattari, MT., Pal, M., Apaydin, H., (2013a), **M5 Model Tree Application in Daily River Flow Forecasting in Sohu Stream**, Turkey, Water Resources, 4(3), PP. 233-242.
- Sattari, MT., Pal, M., Yurekli, K., Unlukara, A., (2013b), **M5 Model Trees and Neural Networks Based Modelling of ET0 in Ankara**, Turkey, Turkish Journal of Engineering and Environmental Sciences, 37: PP. 211-219.
- Shiau, J.T., (2006), **Fitting Drought Duration and Severity with Two-dimensional Copulas**, Water Resources Management, 20: PP. 795-815
- Tsakiris, G., Pangalou, D., Vangelis, H., (2007), **Regional Drought Assessment Based on the Reconnaissance Drought Index (RDI)**, Water Resources Management, 21: PP. 821-833.
- Wilhite, D.A., (2000), **Drought as a Natural Hazard: Concepts and Definitions, Drought: A Global Assessment**, London, Routledge, UK. PP. 3-18.
- Wiltshire, S.E., (1986), **Identification of Homogeneous Regions for Flood Frequency Analysis**, Journal of Hydrology, 84: PP. 287-302.
- Yurekly, K. Sattari, MT. Anli AS. and Hinis, MA, (2012), **Seasonal and Annual Regional Drought Prediction by Using Data-mining Approach**, Atmósfera, 25(1), PP. 85-105.