



ISSN 2251-7480

پیش‌بینی نمایه خشکسالی SPI به‌روش‌های رگرسیون بردار پشتیبان و خطی چندگانه

سعید صمدیان فرد^{۱*} و اسماعیل اسدی^۲

^۱ استادیار گروه مهندسی آب؛ دانشکده کشاورزی؛ دانشگاه تبریز؛ تبریز؛ ایران

* نویسنده مسئول مکاتبات: s.samadian@tabrizu.ac.ir

^۲ استادیار گروه مهندسی آب؛ دانشکده کشاورزی؛ دانشگاه تبریز؛ تبریز؛ ایران

تاریخ پذیرش: ۱۳۹۵/۱۰/۲۶

تاریخ دریافت: ۱۳۹۵/۰۴/۳۰

چکیده

خشکسالی پدیده‌ای طبیعی است که به‌علت برهم‌کنش عوامل مختلف هواشناسی، دارای فرآیند پیچیده‌ای بوده و در همه شرایط اقلیمی و در تمام مناطق کره زمین به وقوع می‌پیوندد. لذا پیش‌بینی نمایه‌های خشکسالی و ارزیابی زمانی آن‌ها، یکی از راه‌های مؤثر در مدیریت بحران خشکسالی و تدوین طرح‌هایی به‌منظور کاهش اثرات آن به نظر می‌رسد. در پژوهش حاضر، برای پیش‌بینی نمایه خشکسالی بارش استاندارد شده (SPI) در ایستگاه سینوپتیک تبریز در بازه‌زمانی سال‌های ۱۳۵۸ تا ۱۳۹۱، از روش‌های رگرسیون بردار پشتیبان و رگرسیون خطی چندگانه استفاده گردید. بدین منظور، برای پیش‌بینی نمایه SPI در دوره‌های ۳، ۶، ۹، ۱۲، ۲۴ و ۴۸ ماهه از شش ترکیب متفاوت ورودی مشکل از مقادیر متناظر قبلی همان نمایه استفاده گردید. نتایج حاصل از تحلیل پارامترهای آماری نشان داد که هر دو روش مورد مطالعه دقت مناسبی در پیش‌بینی نمایه‌های خشکسالی داشته‌اند ولی با این وجود، روش رگرسیون بردار پشتیبان در پیش‌بینی نمایه‌های SPI-6، SPI-9 و SPI-24 به‌ترتیب با داشتن جذر میانگین مربعات خطای ۰/۴۹۸۵، ۰/۴۳۴۰ و ۰/۲۴۲۷ عملکرد بهتری در مقایسه با رگرسیون خطی چندگانه داشته است. ولی در پیش‌بینی نمایه‌های SPI-3، SPI-12 و SPI-48، روش رگرسیون خطی چندگانه خطای نسبی کم‌تری را نشان داد. با این وجود، می‌توان چنین نتیجه‌گیری نمود که هر دو روش مورد مطالعه شامل رگرسیون بردار پشتیبان و خطی چندگانه پیش‌بینی‌های مناسبی از نمایه‌های خشکسالی داشته‌اند و می‌توانند برای مدیریت عواقب ناشی از پدیده مذکور، با اطمینان قابل قبولی مورد استفاده قرار گیرند.

کلید واژه‌ها: پیش‌بینی؛ تحلیل آماری؛ خشکسالی؛ نمایه بارش استاندارد شده

مقدمه

مستقیم آن‌ها به رطوبت ناشی از بارندگی و منابع آب‌های سطحی و زیرزمینی، وارد می‌سازد (قبایی سوق و مساعدی، ۱۳۹۲). خشکسالی در چهار حالت خشکسالی هواشناسی، خشکسالی آب‌شناسی، خشکسالی کشاورزی و خشکسالی اقتصادی-اجتماعی قابل تصور است. خشکسالی هواشناسی ناشی از کمبود بارندگی است که در نهایت می‌تواند منجر به خشکسالی آب‌شناسی و کشاورزی گردد. نیاز به پیش‌بینی و پایش شدت وقوع پدیده‌های خشکسالی انگیزه‌ای برای پژوهشگران شد که که نمایه‌های کمی را برای مطالعه خصوصیات آن معرفی کنند. در بین نمایه‌هایی

خشکسالی پدیده‌ای طبیعی ناشی از کاهش تدریجی بارش و افزایش دماست که احتمال وقوع آن در هر وضعیت اقلیمی وجود دارد (علیزاده، ۱۳۹۴)، اما اثرات زیست محیطی آن در مناطق خشک و نیمه‌خشک، مثل ایران، بیش‌تر از مناطق مرطوب احساس می‌گردد. از طرفی دیگر، پدیده مذکور یکی از بزرگترین بلایای طبیعی است که خسارت‌های فراوانی به بخش‌های مختلف اقتصادی و به خصوص کشاورزی و منابع طبیعی، به دلیل وابستگی

هواشناسی با طول دوره آماری کوتاه مورد استفاده قرار گیرد. اژدری مقدم و همکاران (۱۳۹۱) خشکسالی‌های پاییزه زاهدان را به وسیله متغیرهای بارش، مقادیر پیشین نمایه SPI و ۱۹ عدد از نمایه‌های اقلیمی را با استفاده از سیستم استنتاج فازی - عصبی^۶ (ANFIS) پیش‌بینی نمودند. آن‌ها به این نتیجه رسیدند که مقادیر پیشین نمایه خشکسالی SPI تنها در تاخیر زمانی یک ماهه نتایج قابل قبولی را برای پیش‌بینی خشکسالی پاییزه ارائه داده‌اند. حسین پور نیکنام و همکاران (۱۳۹۱) برای پیش‌بینی خشکسالی یک سال بعد در شهر زاهدان از مقادیر پیشین نمایه خشکسالی بارش استاندارد شده استفاده نموده و قابلیت‌های سیستم استنتاج فازی - عصبی برای ساخت مدل‌های پیش‌بینی نمایه خشکسالی را مورد آزمون قرار دادند. در مطالعه مذکور، ترکیبات مختلفی از متغیرهای ورودی در مدل‌های پیش‌بینی فازی - عصبی وارد گردیدند. نتایج نشان داد که تنها استفاده از سری‌های زمانی مشابه سال قبل SPI در پیش‌بینی خشکسالی‌های ۱۲ ماهه موثر است. کماسی و همکاران (۱۳۹۲) از مبانی شبکه‌های عصبی به همراه تحلیل خوشه‌بندی برای پیش‌بینی خشکسالی با نمایه SPI استفاده نموده و اذعان داشتند که نمایه مذکور قابلیت بیش‌تری در پیش‌بینی نسبت به نمایه‌هایی چون پالمر^۷ (PDSI)، پالفی^۸ (PAI) و دیگر نمایه‌ها داشته و از سویی راهبرد شبکه عصبی - فازی بر مبنای روش خوشه‌بندی C-mean در امر مدل‌سازی برای پیش‌بینی خشکسالی از کارایی بالایی برخوردار است. ققمقامی و بذرافشان (۱۳۹۱) در پایش خشکسالی ایران با استفاده از نمایه SPI و مدل زنجیره مارکف نشان دادند که شدت خشکسالی‌های شدید در کشور بر مبنای سری بارش متوسط منطقه‌ای رو به افزایش است. میرعباسی و همکاران (۱۳۹۲) پایش بلندمدت خشکسالی‌های ارومیه را با استفاده از نمایه کمبود توام مورد مطالعه قرار دادند. آن‌ها

که ارائه شده است، نمایه بارش استاندارد شده^۱ (SPI)، یکی از نمایه‌هایی است که برای پایش دوره‌های خشکسالی مورد استفاده بسیاری از محققین قرار گرفته و همچنین کاربرد گسترده‌ای در تعیین و درجه‌بندی خشکسالی هواشناسی در اکثر مناطق جهان دارد (صادقی نیا و همکاران، ۱۳۹۲).

از سویی دیگر، آن چه که علاوه بر استفاده از نمایه مناسب برای تحلیل و پیش‌بینی خشکسالی اهمیت دارد پیش‌بینی دقیق آن با استفاده از داده‌های متناظر رویدادهای به وقوع پیوسته قبلی است. بدین منظور، پژوهشگران برای ساختن مدل‌های پیش‌بینی کننده معتبر، به استفاده از روش‌های داده‌کاوی و هوش مصنوعی روی آورده‌اند. عیوضی و همکاران (۱۳۸۸) به منظور پیش‌بینی نمایه خشکسالی SPI در دوره‌های زمانی ۱، ۳، ۶ و ۹ ماهه از مدل‌های شبکه عصبی مصنوعی^۲ MLP و تابع پایه شعاعی^۳ RBF در حوضه آبخیز گرگان رود استفاده نموده و نشان دادند که مدل شبکه عصبی مصنوعی MLP در مقایسه با مدل RBE پیش‌بینی‌های دقیق‌تری را ارائه می‌کنند. حسن زاده و همکاران (۱۳۹۱) در پژوهشی به منظور پیش‌بینی خشکسالی‌های آتی، از تلفیق الگوریتم ژنتیک^۴ و مدل ترکیبی شبکه عصبی - موجکی در تحلیل نمایه بارش استاندارد شده استفاده نموده و نشان دادند که بکارگیری روش تلفیقی مذکور در مقایسه با تلفیق الگوریتم ژنتیک و شبکه عصبی، نتایج مطلوب‌تری را ارائه می‌دهد. نیکبخت شهبازی و همکاران (۱۳۹۱) از ماشین‌های بردار پشتیبان^۵ برای پیش‌بینی فصلی خشکسالی هواشناسی با استفاده از نمایه بارش استاندارد شده در ۳۱ ایستگاه باران سنجی محدوده حوضه آبریز سدهای ماملو و طالقان استفاده نموده و نتیجه گرفتند که روش مورد مطالعه با دقت مناسبی می‌تواند در پیش‌بینی رفتارهای غیرخطی داده‌های

6. Adaptive neuro-fuzzy inference system
7. Palmer drought severity index
8. Palfai aridity index

1. Standardized precipitation index
2. Artificial neural networks
3. Radial based function
4. Genetic algorithm
5. Support vector machines

نمایه در محدوده ۰/۶۶ تا ۰/۷۹ متغیر بود. همچنین پیش‌بینی‌های نمایه خشکی موثر دقیق‌تر از نمایه بارش استاندارد شده گزارش گردید. Abrishamchi و همکاران (۲۰۰۷) از مدل تلفیقی شبکه عصبی مصنوعی و تبدیلات موجک برای پیش‌بینی ۱، ۳ و ۶ ماهه خشکسالی بهره گرفتند. نتایج به‌دست آمده از مدل تلفیقی برای چهار ایستگاه واقع در حوضه آبریز سد زاینده رود، نشان دهنده کارایی مدل فوق در بهبود دقت پیش‌بینی خشکسالی نسبت به مدل شبکه عصبی بود. Adamowski و Belayneh (2012) در مطالعه‌ای به بررسی تاثیرگذاری روش‌های شبکه عصبی مصنوعی، رگرسیون بردار پشتیبان^۲ و نرو ویولت^۳ در پیش‌بینی نمایه‌های SPI-3 و SPI-12 در حوضه رودخانه آواش^۴ در اتیوپی پرداختند. ارزیابی عملکرد روش‌های مورد مطالعه نشان داد که روش نرو ویولت پیش‌بینی‌های دقیق‌تری از نمایه‌های خشکسالی ارائه کرده است. Danandeh Mehr و همکاران (۲۰۱۴) روش جدیدی بر مبنای ترکیب روش‌های ویولت و برنامه‌ریزی ژنتیک خطی را به‌منظور پیش‌بینی طولانی مدت خشکسالی ارائه کردند. آن‌ها از نمایه‌های NINO 3.4^۵ و پالمر به‌عنوان پارامترهای ورودی برای پیش‌بینی نمایه پالمر در بازه‌های زمانی ۳، ۶ و ۱۲ ماه آینده در ایالت تگزاس آمریکا استفاده نموده و دقت روش ارائه شده را با روش‌های برنامه‌ریزی ژنتیک خطی کلاسیک^۶، نرو ویولت و فازی ویولت^۷ مقایسه کردند. نتایج حاصل از تحقیق مذکور نشان داد که روش برنامه‌ریزی ژنتیک خطی کلاسیک از توانایی کافی در پیش‌بینی نمایه‌های خشکسالی در بازه‌های زمانی بیش‌تر از ۳ ماه برخوردار نبوده، در حالی که روش ترکیبی ویولت برنامه‌ریزی ژنتیک خطی قادر به پیش‌بینی دقیق نمایه‌های مذکور در تمام بازه‌های زمانی می‌باشد. Khadr (۲۰۱۶) چندین مدل مختلف مارکوف همگن، برای پیش‌بینی

در تحقیق خود از نمایه SPI اصلاح شده Kao and Govindaraju (۲۰۱۰) استفاده کردند. نتایج تحقیق آن‌ها نشان داد نمایه SPI اصلاح شده تا حدود زیادی معایب SPI متداول برطرف و تغییرات فصلی بارش را در محاسبه این نمایه لحاظ نموده است. عابدینی و همکاران (۱۳۹۳) در تحقیقی به‌منظور پیش‌بینی پدیده خشکسالی در ایستگاه سینوپتیک مشهد در فاصله زمانی ۱۹۵۱ تا ۲۰۰۷ از مدل‌های پرسپترون و گرادیان نزولی با الگوریتم‌های آموزش مونتوم و گرادیان نزولی استفاده و گزارش نمودند که به‌علت دامنه زیاد تغییرات بارندگی در مقیاس‌های ۱، ۲ و ۳ ماهه، دقت نتایج پیش‌بینی شده توسط این سری‌های زمانی بسیار کم بوده و با افزایش مقیاس زمانی نمایه، به‌علت کاهش دامنه تغییرات بارندگی، بر دقت نتایج افزوده می‌شود. معروفی و همکاران (۱۳۹۳) از سری زمانی SARIMA در پیش‌بینی بارش و از نمایه SPI برای پیش‌بینی خشکسالی در ایستگاه‌های سینوپتیک بهار، فرودگاه همدان، کبودر آهنگ و نوژه استان همدان در بازه زمانی ۱۳۵۶ تا ۱۳۹۱ استفاده نمودند. نتایج به‌دست آمده حاکی از وقوع خشکسالی ضعیف در همه ایستگاه‌های مورد مطالعه بود. هاشمی نسب و همکاران (۱۳۹۴) در پایش خشکسالی ایران، آمار بارندگی ماهانه ۴۲ ایستگاه سینوپتیک را در طول دوره آماری ۲۰۱۰-۱۹۶۶ با استفاده از نمایه SPI اصلاح شده مورد ارزیابی قرار دادند. نتایج تحقیق آن‌ها نشان داد که نمایه کمبود نظری جهت پایش خشکسالی تحت شرایط اقلیمی ایران مناسب می‌باشد. در پژوهشی دیگر، Morid و همکاران (۲۰۰۷) از مدل‌های شبکه عصبی مصنوعی برای پیش‌بینی نمایه‌های خشکسالی بارش استاندارد شده و خشکی موثر^۱ (EDI) در چندین ایستگاه باران‌سنجی در استان تهران استفاده کردند. بهترین مدل‌ها در هر دو حالت شامل نمایه متناظر در همان ماه در سال قبلی بود. نتایج نشان دادند که ضریب همبستگی برای بهترین مدل‌های شبکه عصبی مصنوعی در پیش‌بینی هر دو

2. Support vector regression
3. Neuro-wavelet
4. Awash river basin
5. El Niño-southern oscillation indicator
6. Classic linear genetic programing
7. Fuzzy-wavelet

1. Effective drought index

۱۲، ۲۴ و ۴۸ ماهه در شهر تبریز و به دست آوردن بیش کلی نسبت به توانایی هر کدام از روش‌های مذکور برای پیش‌بینی وقوع خشکسالی در آینده می‌باشد.

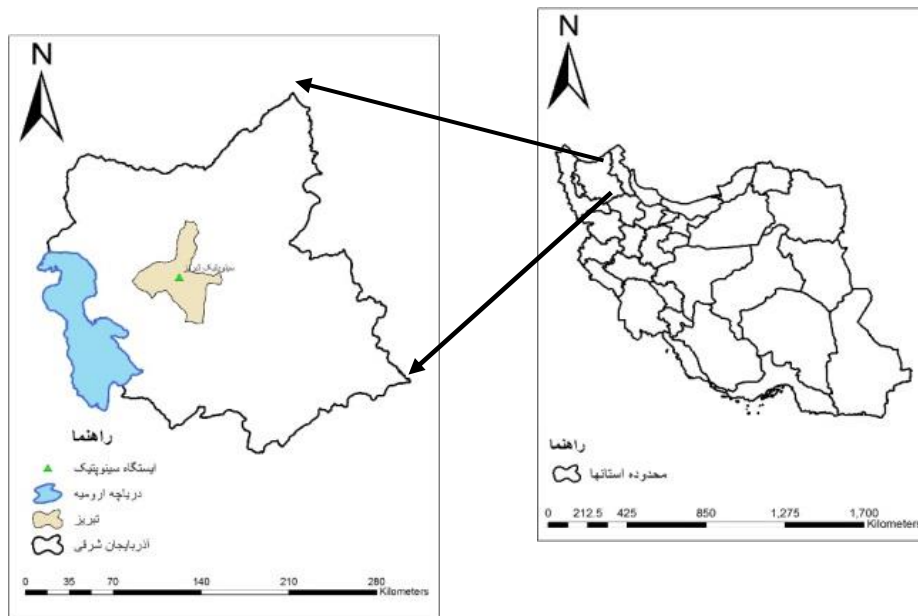
مواد و روش‌ها

منطقه مورد مطالعه

ایستگاه سینوپتیک تبریز یکی از قدیمی‌ترین ایستگاه‌های همدیدی وابسته به سازمان هواشناسی کشور است که در استان آذربایجان شرقی در عرض جغرافیایی ۰۵' ۳۸ شمالی و طول ۱۷' ۴۶ شرقی و ارتفاع ۱۳۶۴ متر از سطح دریا واقع شده است. در این تحقیق، مقادیر بارش ماهانه طی دوره آماری ۱۳۸۹-۱۳۳۱ (۲۰۱۲-۱۹۵۹) به عنوان دوره پایه مورد استفاده قرار گرفته است. میانگین بارندگی سالانه در منطقه مورد مطالعه معادل ۲۸۵ میلی‌متر، پرباران‌ترین ماه اردیبهشت و کم بارش‌ترین ماه مرداد می‌باشد. با این حال، در حدود ۳۸/۵ درصد بارش‌ها در فصل بهار اتفاق می‌افتد. موقعیت جغرافیایی ایستگاه مورد مطالعه در شکل ۱ نشان داده شده است.

دوره‌های کوتاه مدت خشکسالی با استفاده از نمایه SPI در ۲۲ ایستگاه حوضه آبریز نیل آبی در اتیوپی توسعه داد. نتایج حاصله نشان از دقت نسبتاً مناسب مدل‌های توسعه داده شده در پیش‌بینی خشکسالی داشت.

با در نظر گرفتن این که نمایه بارش استاندارد شده ابزاری قوی برای تحلیل داده‌های بارش و بررسی رویدادهای خشکسالی و ترسالی می‌باشد، شناسایی و پیش‌بینی نمایه مذکور در دوره‌های زمانی مختلف از اهمیت پژوهشی بالایی برخوردار است. با بررسی منابع انجام یافته مشخص گردید که تاکنون مطالعه‌ای در مورد پیش‌بینی نمایه‌های خشکسالی با روش رگرسیون بردار پشتیبان در سطح ملی و منطقه‌ای صورت نگرفته است. در این راستا هدف پژوهش حاضر، بررسی عملکرد روش‌های رگرسیون بردار پشتیبان و رگرسیون خطی چندگانه در برآورد نمایه خشکسالی SPI در دوره‌های زمانی ۳، ۶، ۹،



شکل ۱. موقعیت جغرافیایی ایستگاه تبریز

بر پایه تئوری یادگیری آماری معرفی شد. آن‌ها طی سال‌های بعد تئوری ابرصفحه^۱ بهینه را به عنوان طبقه‌بندی

رگرسیون بردار پشتیبان (SVR)

ماشین بردار پشتیبان یکی از روش‌های یادگیری است که در سال ۱۹۹۲ میلادی توسط Boser و همکاران (۱۹۹۲)

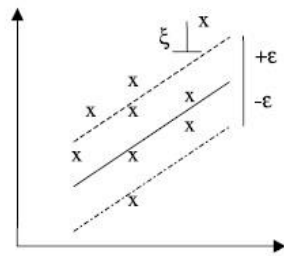
ضعیفی را از خود به نمایش بگذارد. بنابراین می‌بایست یک برآوردکننده نیرومند که نسبت به تغییرات کوچک در مدل حساس نباشد را توسعه داد. در واقع، یک تابع جریمه به‌صورت زیر تعریف می‌شود (Kisi and Cimen, 2011):

$$L^\varepsilon(x, y, f) = \begin{cases} 0 & \text{if } |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon & \text{otherwise} \end{cases} \quad (1)$$

مجموعه داده‌های آموزشی به‌صورت $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ و کلاس تابع به‌صورت $f(x) = \{w^T x + b, w \in R^m, b \in R\}$ می‌باشد. در صورتی که داده‌ها از مقدار ε تخطی نمایند، می‌بایست متناسب با مقدار تخطی، متغیر کمبود تعریف شود که در شکل ۲ نمونه‌ای از این تخطی نمایش داده شده است. مطابق با تابع جریمه اشاره شده، کمینه‌سازی به‌صورت زیر تعریف می‌گردد (اسدپور و ستاری، ۱۳۹۴).

$$\begin{aligned} & \text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (2) \\ & \text{Subject to } \begin{cases} f(x) - y_i \leq \varepsilon + \xi_i \\ f(x) - y_i \leq \varepsilon + \xi_i^* & i = 1, 2, 3, \dots, n \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned}$$

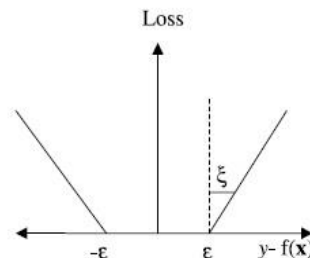
در رابطه ۲، $\|w\|^2$ نرم بردار وزن و ξ_i و ξ_i^* متغیرهای کمبود کمکی هستند، که در شکل ۲ نشان داده شده‌اند و پارامتر C ضریب تعادل پیچیدگی میان ماشین و تعداد نقاط تفکیک ناپذیر است که با سعی و خطا بدست می‌آید.



شکل ۲. میزان تخطی داده‌ها از مقدار مجاز ε .

کننده خطی ارائه داده و طبقه‌بندی کننده‌های غیرخطی را با کمک توابع کرنل معرفی نمودند. اصول اساسی آن چه که امروز به‌عنوان ماشین بردار پشتیبان شناخته می‌شود نتایج کارهای Boser و همکاران است و در نهایت گسترش ماشین بردار پشتیبان بر اساس رگرسیون نیز در سال ۱۹۹۵ توسط Vapnik به نتیجه رسید (Vapnik, 1995). ماشین بردار پشتیبان مبتنی بر کمینه کردن ساختاری ریسک می‌باشد که از نظریه آموزش آماری گرفته شده است (Vapnik, 1998). مدل‌های ماشین‌های بردار پشتیبان به دو گروه عمده الف) مدل طبقه‌بندی ماشین بردار پشتیبان و ب) مدل رگرسیون بردار پشتیبان تقسیم‌بندی می‌شوند. از مدل طبقه‌بندی ماشین بردار پشتیبان جهت حل مسائل طبقه‌بندی داده‌هایی که در کلاس‌های مختلف قرار می‌گیرند، استفاده می‌گردد و مدل رگرسیون بردار پشتیبان در حل مسائل پیش‌بینی کاربرد دارد.

منظور از رگرسیون، به‌دست آوردن یک ابرصفحه است که بر داده‌های مورد نظر برازش داده می‌شود. فاصله هر نقطه از این ابرصفحه نشان‌دهنده خطای آن نقطه خاص است. بهترین روشی که تاکنون برای رگرسیون خطی پیشنهاد شده است، روش حداقل مربعات می‌باشد. با این وجود، برای مسائل رگرسیون، این امکان وجود دارد که استفاده از برآوردکننده کم‌ترین مربعات در حضور داده‌های پرت، بطور کامل شدنی نباشد و در نتیجه رگرسیون عملکرد



می‌توان مسائل را به‌صورت خطی، تفکیک پذیر ساخت. در تحقیق حاضر از سه تابع کرنل، که عموماً در کاربردهای مهندسی متداول هستند (جدول ۱)، استفاده شده است.

یکی از روش‌های متداول برای حل مسائل غیرخطی، استفاده از توابع کرنل می‌باشد که بر اساس ضرب داخلی داده‌ها تعریف می‌شود. در واقع، با یک تبدیل غیرخطی از فضای ورودی به فضایی با ابعاد بیش‌تر (حتی نامتناهی)

جدول ۱. توابع کرنل مورد استفاده در تحقیق حاضر

رابطه ریاضی	توابع کرنل
$k(x_i, x_j) = (x_i^T x_j + 1)^p$	تابع کرنل چند جمله‌ای
$k(x_i, x_j) = e^{-\gamma x_i-x_j ^2}$	تابع کرنل با پایه شعاعی
$k(x_i, x_j) = \frac{1}{\left[1 + \left(2\sqrt{\ x_i - x_j\ ^2} \cdot \sqrt{2^{1/\omega} - \frac{1}{\sigma}}\right)^2\right]^\omega}$	تابع کرنل پیرسون

بارندگی در هر پنجره متحرک تعیین شود. برای مثال، بارندگی تجمعی با پنجره متحرک سه ماهه از مجموع بارندگی همان ماه و دو ماه قبل محاسبه می‌شود. با برازش یک توزیع آماری مناسب بر سری بارش در پنجره زمانی معین و انتقال این توزیع به توزیع نرمال استاندارد، مقادیر نمایه بارش استاندارد به دست می‌آید (قمقامی و بذرافشان، ۱۳۹۱). این شاخص برای اولین بار توسط McKee و همکاران (۱۹۹۳) معرفی و تاکنون در پژوهش‌های بسیاری مورد استفاده قرار گرفته است (Szalai et al., 2000, Sims et al., 2002, Jones and Moberg, 2003, Vicente-Serrano and Lopez-Moreno, 2005, Khan et al., 2008). به تجربه ثابت شده است که توزیع احتمال بارش در اکثر مواقع از توزیع گاما پیروی می‌کند (Shukla and Wood, 2008, Nalbantis and Tsakiris, 2009). توزیع گاما به صورت رابطه ۵ می‌باشد (Mckee, et al., 1993).

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad x > 0 \quad (5)$$

در این رابطه، x مقدار متغیر، α پارامتر شکل، β پارامتر مقیاس توزیع و $\Gamma(\alpha)$ تابع گاما (مطابق رابطه ۶) می‌باشند.

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy \quad (6)$$

پارامترهای α و β بهینه به صورت روابط ۷، ۸ و ۹ برای هر یک از ماه‌های ۱۲ گانه از پنجره متحرک زمانی، با استفاده از روش حداکثر درستنمایی^۲ محاسبه می‌گردند.

$$\hat{\alpha} = \frac{1}{4A} \left[1 + \sqrt{1 + \frac{4A}{3}} \right] \quad (7)$$

رگرسیون خطی چندگانه^۱ (MLR)

رگرسیون خطی چندگانه، روشی است که با استفاده از آن می‌توان به تحلیل و بررسی توام چندین متغیر مختلف پرداخت. این روش نیازمند تعداد زیاد نمونه‌ها می‌باشد که در عین حال باید از دقت و صحت بالایی برخوردار باشند (جلالی لیچایی و نبی بیدهدنی، ۱۳۸۵). رگرسیون خطی چندگانه در حقیقت، ارتباط یک سری از متغیرهای مستقل را با یک متغیر وابسته بیان می‌کند. در صورت وجود متغیرهای مستقل x_1, x_2, \dots, x_n اگر ارتباط خطی بین آن‌ها و متغیر y که وابسته به آن‌ها است ایجاد گردد، رابطه ۳ حاصل خواهد شد.

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n + e \quad (3)$$

که در آن، مقادیر a_1, a_2, \dots, a_n ضرایب رگرسیون می‌باشند. در صورتی که از طرفین رابطه ۳، امید ریاضی گرفته شود، به دلیل این که در امید ریاضی مقدار خطا برابر با صفر می‌باشد، می‌توان نوشت:

$$E(y) = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n \quad (4)$$

که $E(y)$ مقدار مورد انتظار تابع تحت تاثیر ورود مقادیر متغیرهای x_1, x_2, \dots, x_n می‌باشد (Balan et al., 1995).

روش محاسبه نمایه خشکسالی SPI

در این تحقیق به منظور تشخیص کمبود میزان بارندگی و کمی‌سازی خشکسالی هواشناختی، از نمایه خشکسالی SPI اصلاح شده استفاده گردید. برای محاسبه شاخص SPI ابتدا لازم است یک پنجره متحرک زمانی در مقیاس‌های ۱، ۳، ۶، ۸، ۱۲ ماهه تعریف گردد. سپس سری زمانی مجموع

$$Z = \text{SPI} = \left[t - \frac{C_0 + C_1 t + C_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right] \quad (13)$$

$$0.5 < H(x) < 1$$

$$t = \sqrt{\ln\left(\frac{1}{H(x)^2}\right)} \quad 0 < H(x) \leq 0.5 \quad (14)$$

$$t = \sqrt{\ln\left(\frac{1}{(1-H(x))^2}\right)} \quad 0.5 < H(x) \leq 1 \quad (15)$$

در روابط ۱۲ الی ۱۵، $C_0, C_1, C_2, d_1, d_2, d_3$ ضرایب ثابت به‌ترتیب معادل $C_0=2.5165, C_1=0.8029, C_2=0.0103, d_1=1.4328, d_2=0.1893, d_3=0.0013$ می‌باشند.

بدین ترتیب، SPI نرمال شده تبدیل به نمره Z نرمال می‌گردد که منعکس‌کننده مقدار انحرافات بالاتر و یا پایین‌تر از میانگین می‌باشد. بر طبق تعریف McKee و همکاران (۱۹۹۳) اگر این نمایه منفی باشد، دوره خشکی آغاز شده و هر جا به صفر برسد، خشکی پایان یافته است.

ارزیابی دقت روش‌های مورد استفاده

یکی از روش‌های آماده‌سازی مجموعه داده‌ها، تقسیم کل نمونه‌ها به دو دسته واسنجی و صحت‌سنجی به‌صورت تصادفی است که در این صورت، مدل پیش‌بینی با مجموعه داده‌های واسنجی ایجاد و دقت آن با مجموعه صحت‌سنجی ارزیابی می‌شود. اما تقسیم تصادفی داده‌ها به دو قسمت واسنجی و صحت‌سنجی باعث ایجاد تفاوت در نتایج آماری می‌گردد. برای کاهش این اختلاف، منابع متعددی استفاده از روش K باره^۱ را توصیه می‌کنند (Cios et al., 2007, Olson and Delen, 2008, Witten et al., 2005). در این روش، مجموعه داده اصلی به‌صورت تصادفی در K زیرگروه داده‌ای جدا قرار می‌گیرند. مدل پیش‌بینی در هر بار با K-1 زیرگروه ایجاد و با یک زیرگروه باقیمانده صحت‌سنجی می‌شود. این فرایند واسنجی و صحت‌سنجی، K بار تکرار می‌گردد. مزیت این روش این است که سرانجام همه داده‌های موجود در هر دو

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \quad (8)$$

$$A = \ln(\bar{x}) - \frac{\sum \ln(x)}{N} \quad (9)$$

در روابط ۷ الی ۹، N تعداد مشاهدات و \bar{x} میانگین سری داده‌ها می‌باشند. با استفاده از تابع توزیع تراکمی گاما (رابطه ۱۰) مقادیر احتمال تجمعی توزیع گاما محاسبه می‌گردد:

$$F(x) = \int_0^x f(x) dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^x x^{\alpha-1} e^{-x/\beta} dx \quad (10)$$

پس از محاسبه مقادیر احتمال تجمعی گاما، بایستی این مقادیر به مقادیر توزیع نرمال استاندارد با میانگین صفر و انحراف معیار یک تبدیل گردد. بهترین حالت انتقال یک متغیر از توزیع آماری به توزیع آماری دیگر، انتقال هم‌احتمال می‌باشد. در انتقال هم‌احتمال، احتمال‌های کم‌تر یا مساوی دو متغیر از دو توزیع، برابر فرض می‌شوند. از آنجایی که تابع گاما برای $X=0$ تعریف نشده و داده‌های بارندگی همیشه در برگیرنده تعداد زیادی مشاهده با بارندگی صفر می‌باشند، احتمال تجمعی بارندگی به‌صورت رابطه ۱۱ محاسبه می‌شود.

$$H(x) = q + (1-q) \times F(x) \quad (11)$$

در رابطه ۱۱، q احتمال بارندگی صفر موجود در سری داده‌ها می‌باشد که از تقسیم تعداد صفر بر تعداد کل داده‌ها به‌دست می‌آید ($q = \frac{m}{N}$). با محاسبه رابطه ۱۱ و با استفاده از روابط ۱۲ و ۱۳، توزیع نرمال استاندارد (Z) با میانگین صفر و انحراف معیار یک بدست خواهد آمد که برابر مقدار نمایه SPI در نظر گرفته می‌شود (McKee, 1993). در تبدیل مقادیر احتمال تجمعی گاما به Z نرمال استاندارد مطابق با مطالعه (Edwards and McKee (1997) از روش عددی (Abramowitz, M. and Stegun (1965) به شرح زیر استفاده می‌شود:

$$Z = \text{SPI} = \left[t - \frac{C_0 + C_1 t + C_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right] \quad (12)$$

$$0 < H(x) \leq 0.5$$

روش تحقیق

در مطالعه حاضر داده‌های اندازه‌گیری شده بارش ایستگاه تبریز طی سال‌های ۱۳۵۸ تا ۱۳۹۱ جمع‌آوری شد. با استفاده از داده‌های مذکور، مقادیر نمایه خشکسالی بارش استاندارد شده در بازه‌های زمانی ۳، ۶، ۹، ۱۲، ۲۴ و ۴۸ ماهه تعیین گردیدند. سپس با تعریف شش سناریو ترکیبی از داده‌های پیشین متناظر شامل $SPI-A_t$, $SPI-A_{t-1}$, $SPI-A_{t-2}$, $SPI-A_{t-3}$, $SPI-A_{t-4}$, $SPI-A_{t-5}$ که $SPI-A_{t-1}$ مقدار نمایه بارش استاندارد شده A ماهه مربوط به یک دوره قبل می‌باشد، مقدار SPI در بازه زمانی آینده توسط روش‌های مورد مطالعه پیش‌بینی گردید. در نهایت دقت روش‌های SVR و MLR با استفاده از روش K باره و بهره‌گیری از پارامترهای آماری جذر میانگین مربعات خطا، میانگین خطای مطلق و ضریب همبستگی تعیین گردید.

نتایج و بحث

به‌منظور پیش‌بینی خشکسالی در ایستگاه سینوپتیک تبریز، ابتدا مقادیر نمایه SPI در دوره‌های زمانی ۳، ۶، ۹، ۱۲، ۲۴ و ۴۸ ماهه محاسبه گردید. روند تغییرات نمایه خشکسالی SPI و مقادیر متناظر بارش در دوره‌های زمانی اشاره شده در شکل‌های ۳ تا ۸ نشان داده شده است.

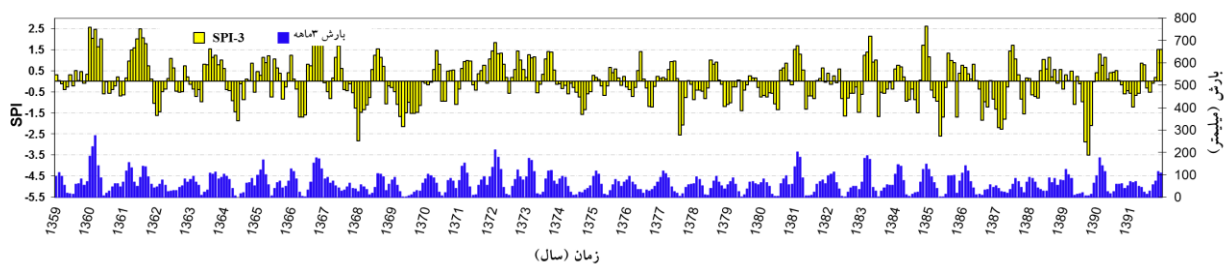
فرایند واسنجی و صحت‌سنجی شرکت می‌کنند. در پژوهش حاضر، خطای بین مقادیر ضرایب نمایه خشکسالی SPI مشاهداتی و محاسباتی با استفاده از معیارهای جذر میانگین مربعات خطا (RMSE)، میانگین خطای مطلق (MAE) و ضریب همبستگی (R) مورد بررسی قرار گرفت.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (16)$$

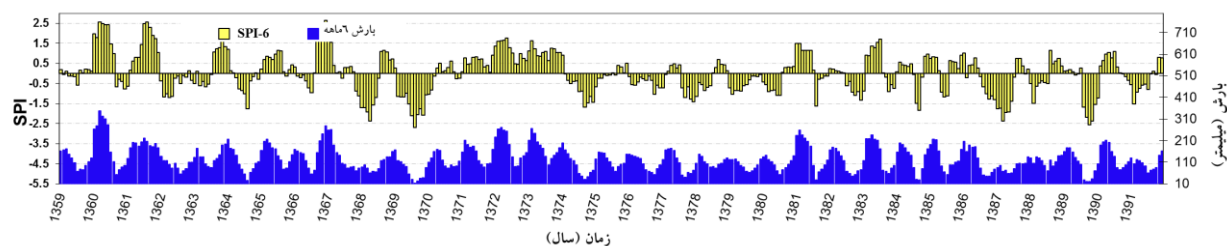
$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (17)$$

$$R = \frac{\left(\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i \right)}{\left(\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right) \left(\sum_{i=1}^n y_i^2 - \frac{1}{n} \left(\sum_{i=1}^n y_i \right)^2 \right)} \quad (18)$$

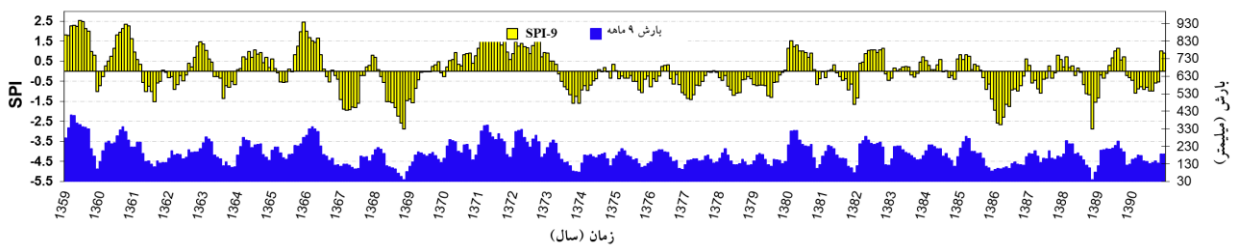
در روابط ۱۶ الی ۱۸، x_i و y_i به ترتیب مقادیر ضریب نمایه خشکسالی SPI و محاسبه شده توسط روش‌های مورد مطالعه و n تعداد داده‌ها می‌باشند. لازم به ذکر است که در روش K باره، مقادیر پارامترهای آماری به‌صورت میانگین K تکرار محاسبه می‌گردد.



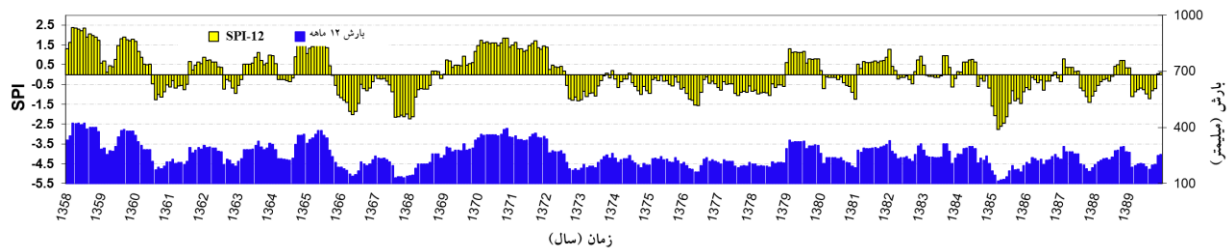
شکل ۳. روند تغییرات نمایه خشکسالی SPI-3 ایستگاه تبریز در دوره زمانی مورد مطالعه



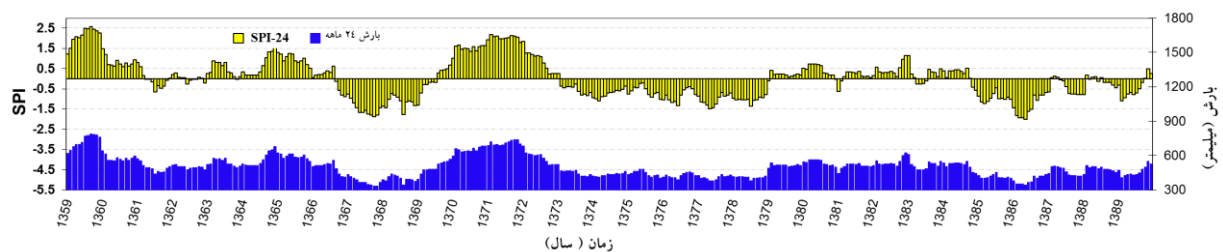
شکل ۴. روند تغییرات نمایه خشکسالی SPI-6 ایستگاه تبریز در دوره زمانی مورد مطالعه



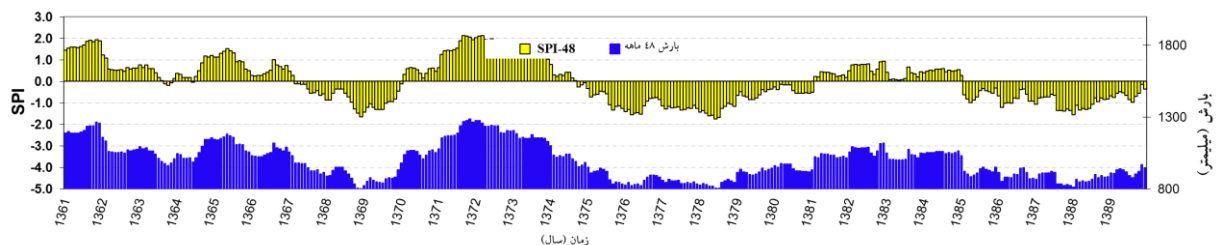
شکل ۵. روند تغییرات نمایه خشکسالی SPI-9 ایستگاه تبریز در دوره زمانی مورد مطالعه



شکل ۶. روند تغییرات نمایه خشکسالی SPI-12 ایستگاه تبریز در دوره زمانی مورد مطالعه



شکل ۷. روند تغییرات نمایه خشکسالی SPI-24 ایستگاه تبریز در دوره زمانی مورد مطالعه



شکل ۸. روند تغییرات نمایه خشکسالی SPI-48 ایستگاه تبریز در دوره زمانی مورد مطالعه

جدول ۲ نشان داده شده‌اند. لازم به ذکر است که محاسبات مربوط به پیش‌بینی SPI در بازه زمانی آینده با روش‌های WEKA و SVR با استفاده از نرم‌افزار داده کاوی انجام پذیرفته و دقت روش‌های مذکور در مرحله صحت‌سنجی، بر اساس معیارهای آماری و اجرای روش K باره (روابط ۱۶ الی ۱۸) محاسبه گردیده و در جدول‌های ۳ تا ۸ ارائه شده‌اند.

به‌منظور بررسی امکان استفاده از داده‌های پیشین SPI در پیش‌بینی نمایه مشابه در دوره زمانی آینده، شش سناریو ترکیبی از داده‌های پیشین متناظر شامل $SPI-A_t$, $SPI-A_{t-1}$, $SPI-A_{t-2}$, $SPI-A_{t-3}$, $SPI-A_{t-4}$, $SPI-A_{t-5}$ که مقدار نمایه بارش استاندارد شده A ماهه مربوط به یک دوره قبل می‌باشد، مورد بررسی قرار گرفت. ترکیب‌های ورودی اشاره شده برای پیش‌بینی نمایه خشکسالی در

جدول ۲. ترکیب های مختلف ورودی در پیش بینی نمایه SPI

شماره سناریو	پارامترهای ورودی
I	SPI-A _t
II	SPI-A _t , SPI-A _{t-1} ,
III	SPI-A _t , SPI-A _{t-1} , SPI-A _{t-2} ,
IV	SPI-A _t , SPI-A _{t-1} , SPI-A _{t-2} , SPI-A _{t-3}
V	SPI-A _t , SPI-A _{t-1} , SPI-A _{t-2} , SPI-A _{t-3} , SPI-A _{t-4}
VI	SPI-A _t , SPI-A _{t-1} , SPI-A _{t-2} , SPI-A _{t-3} , SPI-A _{t-4} , SPI-A _{t-5}

جدول ۳. ارزیابی عملکرد روش های رگرسیون بردار پشتیبان و رگرسیون خطی چندگانه در پیش بینی نمایه SPI-3

MLR			SVR			شماره سناریو	نمایه خشکسالی
RMSE	MAE	R	RMSE	MAE	R		
۰/۷۵۶۱	۰/۵۸۶۳	۰/۶۵۹۳	۰/۷۵۷۹	۰/۵۸۷۶	۰/۶۵۹۲	I	SPI-3
۰/۷۴۷۷	۰/۵۸۱۰	۰/۶۶۸۸	۰/۷۴۷۳	۰/۵۷۹۲	۰/۶۷۱۸	II	
۰/۷۲۳۴	۰/۵۶۲۹	۰/۶۹۴۸	۰/۷۲۷۸	۰/۵۶۴۷	۰/۶۹۰۹	III	
۰/۷۱۴۲	۰/۵۵۹۵	۰/۷۰۴۱	۰/۷۲۲۲	۰/۵۶۱۱	۰/۶۹۷۰	IV	
۰/۷۱۲۰	۰/۵۵۸۷	۰/۷۰۶۴	۰/۷۲۳۱	۰/۵۶۳۷	۰/۶۵۹۷	V	
۰/۷۱۳۳	۰/۵۶۰۴	۰/۷۰۵۱	۰/۷۲۳۴	۰/۵۶۶۱	۰/۶۹۵۶	VI	

جدول ۴. ارزیابی عملکرد روش های رگرسیون بردار پشتیبان و رگرسیون خطی چندگانه در پیش بینی نمایه SPI-6

MLR			SVR			شماره سناریو	نمایه خشکسالی
RMSE	MAE	R	RMSE	MAE	R		
۰/۵۲۱۷	۰/۳۷۴۰	۰/۸۴۸۲	۰/۵۲۳۶	۰/۳۷۲۰	۰/۸۴۸۵	I	SPI-6
۰/۵۰۹۵	۰/۳۶۹۴	۰/۸۵۵۹	۰/۵۱۳۵	۰/۳۶۳۰	۰/۸۵۵۴	II	
۰/۵۰۷۸	۰/۳۶۷۲	۰/۸۵۶۹	۰/۵۰۹۷	۰/۳۶۱۹	۰/۸۵۷۲	III	
۰/۵۰۲۵	۰/۳۶۳۵	۰/۸۶۰۱	۰/۵۰۴۲	۰/۳۶۰۵	۰/۸۵۹۷	IV	
۰/۵۰۴۹	۰/۳۶۵۹	۰/۸۵۸۷	۰/۵۰۴۴	۰/۳۵۹۸	۰/۸۵۹۹	V	
۰/۴۹۶۹	۰/۳۵۷۵	۰/۸۶۳۵	۰/۴۹۸۵	۰/۳۵۶۲	۰/۸۶۳۱	VI	

جدول ۵. ارزیابی عملکرد روش های رگرسیون بردار پشتیبان و رگرسیون خطی چندگانه در پیش بینی نمایه SPI-9

MLR			SVR			شماره سناریو	نمایه خشکسالی
RMSE	MAE	R	RMSE	MAE	R		
۰/۴۴۱۹	۰/۳۳۴۳	۰/۸۹۴۳	۰/۴۴۲۸	۰/۳۳۳۸	۰/۸۹۴۵	I	SPI-9
۰/۴۴۱۶	۰/۳۳۶۸	۰/۸۹۴۴	۰/۴۴۳۱	۰/۳۳۴۰	۰/۸۹۴۶	II	
۰/۴۳۶۵	۰/۳۳۱۹	۰/۸۹۷۰	۰/۴۳۶۵	۰/۳۲۶۲	۰/۸۹۸۱	III	
۰/۴۳۶۴	۰/۳۳۰۷	۰/۸۹۷۱	۰/۴۳۴۲	۰/۳۲۴۴	۰/۸۹۹۳	IV	
۰/۴۳۷۹	۰/۳۳۳۳	۰/۸۹۶۳	۰/۴۳۴۰	۰/۳۲۵۸	۰/۸۹۹۱	V	
۰/۴۳۶۱	۰/۳۳۲۴	۰/۸۹۷۲	۰/۴۳۵۴	۰/۳۲۷۸	۰/۸۹۸۵	VI	

جدول ۶. ارزیابی عملکرد روش‌های رگرسیون بردار پشتیبان و رگرسیون خطی چندگانه در پیش‌بینی نمایه SPI-12

MLR			SVR			شماره سناریو	نمایه خشکسالی
RMSE	MAE	R	RMSE	MAE	R		
۰/۳۸۴۱	۰/۲۷۲۷	۰/۹۲۰۸	۰/۳۸۵۲	۰/۲۶۷۷	۰/۹۲۱۹	I	SPI-12
۰/۳۸۱۱	۰/۲۷۴۳	۰/۹۲۲۰	۰/۳۸۵۰	۰/۲۶۷۸	۰/۹۲۱۳	II	
۰/۳۸۰۴	۰/۲۷۴۷	۰/۹۲۲۳	۰/۳۸۴۴	۰/۲۶۸۲	۰/۹۲۱۶	III	
۰/۳۸۰۴	۰/۲۷۴۷	۰/۹۲۲۳	۰/۳۸۴۴	۰/۲۶۷۹	۰/۹۲۱۷	IV	
۰/۳۸۱۰	۰/۲۷۵۷	۰/۹۲۲۱	۰/۳۸۵۰	۰/۲۶۸۸	۰/۹۲۱۳	V	
۰/۳۸۰۹	۰/۲۷۶۸	۰/۹۲۲۱	۰/۳۸۴۳	۰/۲۷۰۷	۰/۹۲۱۶	VI	

جدول ۷. ارزیابی عملکرد روش‌های رگرسیون بردار پشتیبان و رگرسیون خطی چندگانه در پیش‌بینی نمایه SPI-24

MLR			SVR			شماره سناریو	نمایه خشکسالی
RMSE	MAE	R	RMSE	MAE	R		
۰/۲۴۵۲	۰/۱۷۷۳	۰/۹۶۵۲	۰/۲۴۴۸	۰/۱۷۶۱	۰/۹۶۵۵	I	SPI-24
۰/۲۴۴۴	۰/۱۷۶۶	۰/۹۶۵۴	۰/۲۴۴۰	۰/۱۷۵۲	۰/۹۶۵۷	II	
۰/۲۴۲۹	۰/۱۷۵۸	۰/۹۶۵۹	۰/۲۴۲۷	۰/۱۷۴۳	۰/۹۶۶۱	III	
۰/۲۴۲۹	۰/۱۷۵۸	۰/۹۶۵۹	۰/۲۴۲۷	۰/۱۷۴۸	۰/۹۶۶۱	IV	
۰/۲۴۲۹	۰/۱۷۵۸	۰/۹۶۵۹	۰/۲۴۳۰	۰/۱۷۵۴	۰/۹۶۶۰	V	
۰/۲۴۲۹	۰/۱۷۵۸	۰/۹۶۵۹	۰/۲۴۳۲	۰/۱۷۵۴	۰/۹۶۵۹	VI	

جدول ۸. ارزیابی عملکرد روش‌های رگرسیون بردار پشتیبان و رگرسیون خطی چندگانه در پیش‌بینی نمایه SPI-48

MLR			SVR			شماره سناریو	نمایه خشکسالی
RMSE	MAE	R	RMSE	MAE	R		
۰/۱۹۹۲	۰/۱۴۳۸	۰/۹۷۸۲	۰/۱۹۸۱	۰/۱۴۲۸	۰/۹۷۸۵	I	SPI-48
۰/۱۹۵۴	۰/۱۴۴۲	۰/۹۷۹۰	۰/۱۹۶۱	۰/۱۴۱۷	۰/۹۷۸۹	II	
۰/۱۹۵۴	۰/۱۴۴۲	۰/۹۷۹۰	۰/۱۹۶۱	۰/۱۴۱۸	۰/۹۷۸۹	III	
۰/۱۹۵۹	۰/۱۴۴۶	۰/۹۷۸۹	۰/۱۹۵۸	۰/۱۴۱۵	۰/۹۷۸۹	IV	
۰/۱۹۵۹	۰/۱۴۴۶	۰/۹۷۸۹	۰/۱۹۵۹	۰/۱۴۲۱	۰/۹۷۸۹	V	
۰/۱۹۶۵	۰/۱۴۵۴	۰/۹۷۸۸	۰/۱۹۶۳	۰/۱۴۲۷	۰/۹۷۸۸	VI	

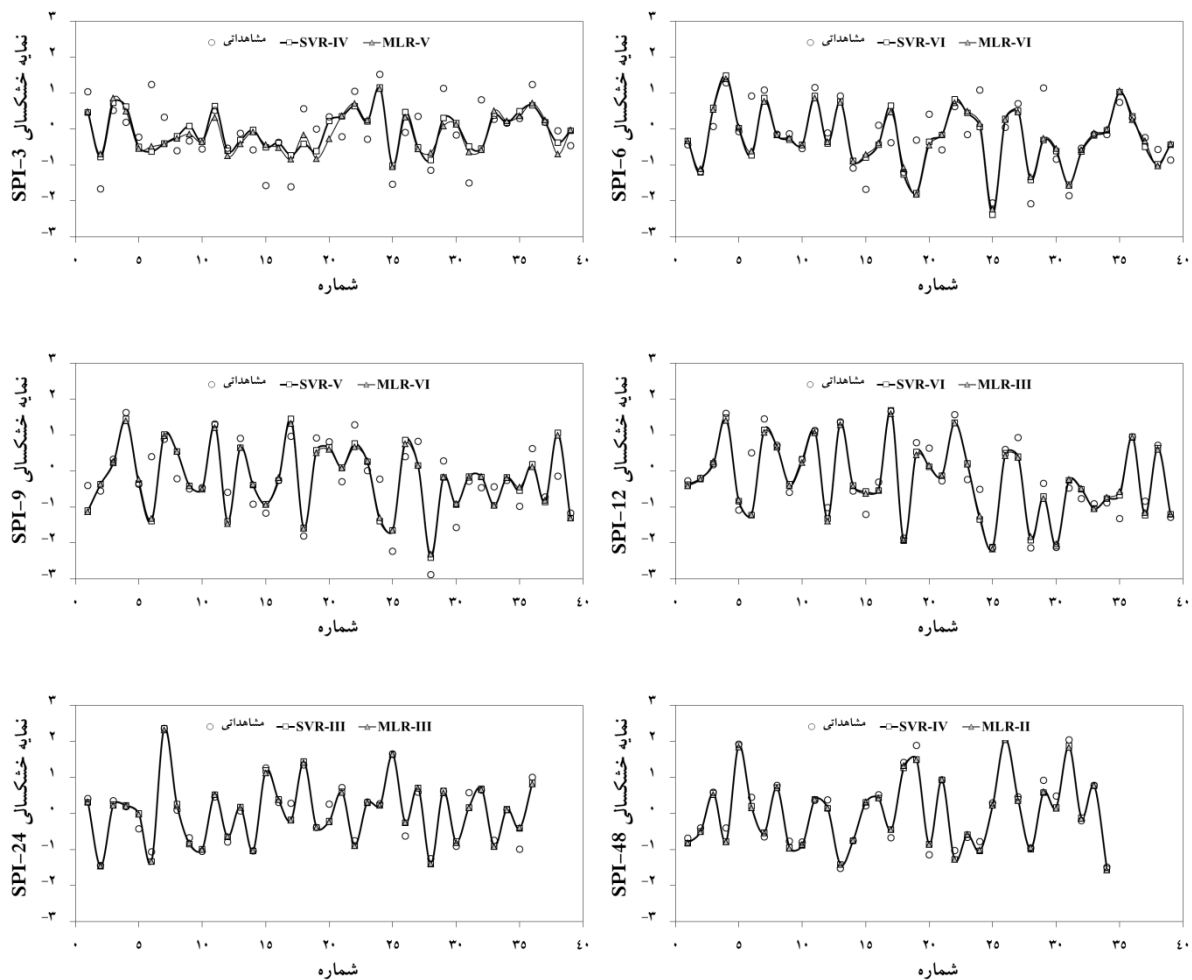
برده‌اند. با این وجود، باز هم دقت مدل MLR-VI بیش‌تر از دقت متناظر SVR-VI در پیش‌بینی نمایه SPI-6 می‌باشد. در مدل‌سازی نمایه SPI-9 و برخلاف دو نمایه قبلی، از جدول ۵ مشاهده می‌گردد که دقت مدل SVR-V با مقدار RMSE برابر با ۰/۴۳۴۰ بیش‌تر از مدل MLR-VI با مقدار RMSE برابر با ۰/۴۳۶۱ می‌باشد. در پیش‌بینی نمایه SPI-12 نیز مشاهده می‌گردد که مدل MLR-III نسبت به مدل SVR-VI به‌صورت جزئی، عملکرد مناسب‌تری از خود نشان داده است. ولی با توجه به این که مدل MLR-III از پارامترهای ورودی کم‌تری نسبت به مدل SVR-VI استفاده

همانطوری که از جدول ۳ مشاهده می‌گردد، در محاسبه SPI-3 روش SVR با سناریو IV و روش MLR با سناریو V در مقایسه با دیگر ترکیب‌های ورودی بهترین پیش‌بینی‌ها را داشته‌اند. اگر چه مدل MLR-V دقت بیش‌تری در مقایسه با SVR-IV در پیش‌بینی SPI-3 داشته است. همچنین از جدول ۳ چنین استنباط می‌گردد که منظور کردن داده SPI-3₅ در محاسبات، تاثیر مثبتی در کاهش خطای پیش‌بینی نداشته است. برخلاف روند مشاهده شده قبلی، در پیش‌بینی SPI-6 هر دو مدل SVR و MLR از سناریو VI برای افزایش دقت پیش‌بینی بهره

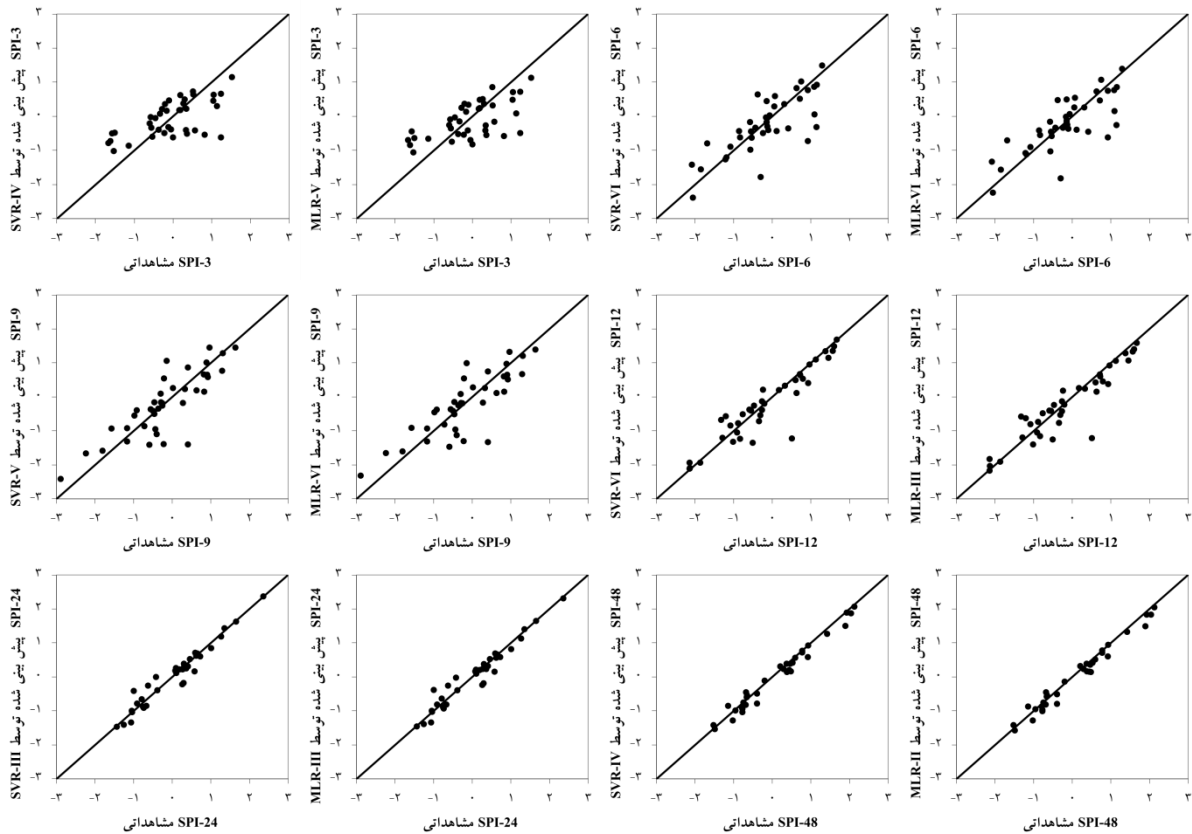
از شکل ۹ مشاهده می‌گردد که با افزایش دوره زمانی نمایه خشکسالی از SPI-3 به SPI-48، دقت پیش‌بینی‌ها افزایش پیدا نموده و منحنی‌های رسم‌شده در دوره‌های مختلف به نقاط مشاهداتی نزدیک می‌شوند.

روند اشاره شده قبلی در مورد کاهش خطای مدل‌سازی با افزایش دوره زمانی نمایه خشکسالی از شکل ۱۰ نیز قابل نتیجه‌گیری است. در شکل ۱۰ نیز به وضوح مشاهده می‌شود که پراکنش نقاط رسم شده حول خط یک به یک (خط نیم ساز) با افزایش دوره زمانی نمایه SPI کم‌تر شده و دقت پیش‌بینی‌ها بالاتر رفته است.

می‌کند، ارزش کاربردی بالایی در مقایسه با آن دارد. با توجه به نتایج حاصل شده در پیش‌بینی نمایه‌های SPI-24 و SPI-48 (جدول‌های ۷ و ۸)، هر دو مدل SVR و MLR عملکردهای تقریباً یکسانی در پیش‌بینی خشکسالی در دوره‌های دو و چهارساله از خود نشان داده‌اند، ولی استفاده از روابط خطی رگرسیونی و استفاده از مدل MLR-II در پیش‌بینی نمایه SPI-48 ساده‌تر از روش SVR می‌باشد. در نهایت تغییرات نمایه‌های خشکسالی مشاهداتی و پیش‌بینی شده توسط روش‌های بهینه مورد مطالعه در شکل ۹ و پراکنش مقادیر مشاهداتی و مقادیر متناظر محاسبه شده در شکل ۱۰ ارائه گردیده‌اند.



شکل ۹. تغییرات نمایه‌های خشکسالی مشاهداتی و محاسباتی در مرحله صحت‌سنجی



شکل ۱۰. پراکنش مقادیر مشاهداتی نمایه‌های متفاوت خشکسالی (محور X ها) و مقادیر متناظر پیش‌بینی شده آن توسط روش‌های مورد مطالعه بهینه (محور Y ها)

نتیجه‌گیری

نتایج حاصل از پژوهش نشان دادند که روش رگرسیون بردار پشتیبان در پیش‌بینی نمایه‌های SPI-6، SPI-9 و SPI-24 به‌ترتیب با داشتن جذر میانگین مربعات خطای ۰/۴۹۸۵، ۰/۴۳۴۰ و ۰/۲۴۲۷ عملکرد بهتری در مقایسه با رگرسیون خطی چندگانه داشته، ولی در پیش‌بینی نمایه‌های SPI-3، SPI-12 و SPI-48، روش رگرسیون خطی چندگانه خطای نسبی کم‌تری را نشان داد. همچنین مشاهده گردید که با افزایش دوره زمانی نمایه خشکسالی از SPI-3 به SPI-48، دقت پیش‌بینی‌ها افزایش پیدا نموده و اختلاف بین مقادیر مشاهداتی و محاسباتی نمایه SPI کم‌تر می‌شود.

در مطالعه حاضر، برای مدل‌سازی پدیده خشکسالی از روش‌های رگرسیون بردار پشتیبان و خطی چندگانه استفاده گردید. بدین منظور، نمایه بارش استاندارد شده، به‌عنوان یکی از مهمترین و معتبرترین نمایه‌های بررسی خشکسالی هواشناسی، در بازه زمانی سال‌های ۱۳۵۸ تا ۱۳۹۱ در ایستگاه سینوپتیک تبریز محاسبه و سپس با استفاده از داده‌های پیشین SPI و تعریف شش ترکیب متفاوت ورودی متشکل از مقادیر متناظر قبلی همان نمایه در دوره‌های زمانی ۳، ۶، ۹، ۱۲، ۲۴ و ۴۸ ماهه دقت روش‌های مورد مطالعه مورد بررسی و آزمون قرار گرفت.

فهرست منابع

اسدپور اصل، ر. و ستاری، م. ۱۳۹۴. پیش‌بینی عمق حفره آب شستگی اطراف پایه پل با استفاده از روش رگرسیون ماشین بردار پشتیبان. دهمین کنگره بین‌المللی مهندسی عمران، تبریز، ایران.

ازدري مقدم، م.، خسروي، م.، حسين پور نيکنام، ح. و جعفري ندوشن، ا. ۱۳۹۱. پيش بينی خشکسالی با استفاده از مدل فازی - عصبی، شاخص های اقليمي، بارندگی و شاخص خشکسالی (مطالعه موردی: زاهدان). جغرافيا و توسعه. 26: 61-72.

جلالی لیچایی، م. و نبی بیدهندي، م. ۱۳۸۵. مقایسه روش های رگرسیون خطی چندگانه و شبکه های عصبی مصنوعی برای برآورد تخلخل و نفوذپذیری. علوم زمین. ۶۱: ۱۴۹-۱۴۰.

حسن زاده، ی.، عبدی کردانی، ا. و فاخری فرد، ا. ۱۳۹۱. پيش بينی خشکسالی با استفاده از الگوریتم ژنتیک و مدل ترکیبی شبکه عصبی - موجکی. آب و فاضلاب. ۳: ۵۹-۴۸.

حسین پور نيکنام، ح.، ازدري مقدم، م. و خسروي، م. ۱۳۹۱. پيش بينی خشکسالی یک ساله با استفاده از مدل فازی - عصبی، سری های زمانی خشکسالی و شاخص های اقليمي (مطالعه موردی: زاهدان). آب و فاضلاب. ۲: ۵۱-۴۲.

صادقی نیا، ع.، حجازی زاده، ز.، حمیدیان پور، م. و پورسربندان، ر. ۱۳۹۲. برآورد احتمالات خشکسالی و ترسالی با استفاده از نمایه SPI و مدل زنجیره مارکف مطالعه موردی: تهران. فصلنامه ی فضای جغرافیایی. ۴۳: ۸۱-۶۵.

عابدینی، ا.، موسوی بایگی، م. و انصاری، ح. ۱۳۹۳. کاربرد مدل های شبکه عصبی مصنوعی در تخمین نمایه های خشکسالی SPI و PDSI در ایستگاه سینوپتیک شهر مشهد. نشریه پژوهش آب در کشاورزی. ۲۸(۱): ۲۳۹-۲۲۷.

علیزاده، ا. ۱۳۹۴. اصول هیدرولوژی کاربردی. چاپ بیست و هشتم. دانشگاه امام رضا (ع). مشهد. ایران. ۹۴۶ صفحه.

عیوضی، م.، مساعدی، ا. و دهقانی، ا. ۱۳۸۸. مقایسه روش های مختلف پيش بينی شاخص خشکسالی SPI. مجله پژوهش های حفاظت آب و خاک. ۱۶(۲): ۱۶۷-۱۴۵.

قبایی سوق، م. و مساعدی، ا. ۱۳۹۲. پایش خشکسالی گندم با استفاده از مدل اسکالوگرام تعمیم یافته در محدوده ایستگاه های سینوپتیک مشهد و شیراز. نشریه آبیاری و زهکشی ایران. ۱(۷): ۳۵-۲۳.

قمقماقی، م. و بذرافشان، ج. ۱۳۹۱. پيش آگاهی وضعیت خشکسالی هواشناسی در گستره ایران با استفاده از مدل زنجیره مارکف. نشریه حفاظت منابع آب و خاک. ۱(۳): ۱۱-۱.

کماسی، م.، اعلمی، م.ت. و نورانی، و. ۱۳۹۲. پيش بينی خشکسالی با نمایه SPI به روش مدل سازی ANFIS بر مبنای خوشه بندی C-mean فازی. آب و فاضلاب. ۴: ۱۰۲-۹۰.

معروفی، ص.، ختار، ب.، صادقی فر، م.، پارسا فر، ن. و ایلدورمی، ع. ۱۳۹۳. پيش بينی خشکسالی با استفاده از سری زمانی SARIMA و شاخص SPI در ناحیه مرکزی استان همدان. نشریه پژوهش آب در کشاورزی. ۲۸(۱): ۲۲۵-۲۱۳.

میرعباسی نجف آبادی، ر.، فاخری فرد، ا.، دین پژوه، ی. و اسلامیان س. ۱۳۹۲. پایش بلند مدت خشکسالی های ارومیه با استفاده از شاخص کمبود توأم. نشریه دانش کشاورزی (آب و خاک). ۲۳(۴): ۱۰۳-۸۷.

نیک بخت شهبازی، ع.، زهرایی، ب. و ناصری، م. ۱۳۹۱. پيش بينی فصلی خشکسالی هواشناسی با استفاده از ماشین های بردار پشتیبان. آب و فاضلاب. ۲: ۸۴-۷۲.

هاشمی نسب، آ.، بذرافشان، ج. و نازی قمشلو، آ. ۱۳۹۴. ارزیابی شاخص خشکسالی کمبود توأم تحت شرایط اقليمي ایران. نشریه حفاظت منابع آب و خاک. ۴(۳): ۶۴-۵۳.

- Abramowitz, M. and Stegun I.A. 1965. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover Publications, Inc., New York. 1046 pp.
- Abrishamchi, A., Mehdikhani, H., Tajrishy, M., and Marino, M.A. 2007. Application of conjunctive nonlinear model based on wavelet transforms and artificial neural networks to drought forecasting. American Geophysical Union, Fall Meeting, Abstract Number: PA33A-1021.
- Balan, B., Mohaghegh, S., Ameri, S. 1995. State - of -Art in permeability determination from well log data : Part - 1 -A comparative study , Model development. SPE 30978 , 17-25.

- Belayneh, A. and Adamowski, J. 2012. Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression. *Applied Computational Intelligence and Soft Computing*, 2012, 1-13.
- Boser, B.E., Guyon, I.M. and Vapnik, V.N. 1992. A training algorithm for optimal margin classifiers. In D.Haussler, editor, 5th Annual ACM Workshop on COLT , pages 144-152, Pittsburgh, PA.
- Cios KJ, Pedrycz W, Swiniarski RW, Kurgan LA. 2007. *Data Mining: A Knowledge Discovery Approach*. USA: Springer; 473-4.
- Danandeh Mehr, A., Kahya, E. and Özger, M. 2014. A gene-wavelet model for long lead time drought forecasting. *Journal of Hydrology*, 517: 691-699.
- Edwards, D.C. and McKee, T.B. 1997. Characteristics of 20th century drought in the United States at multiple timescales. Colorado State University: Fort Collins. *Climatology Report No. 97-2*.
- Jones, P.D. and Moberg, A. 2003. Hemispheric and large-scale surface air temperature variations: an extensive revision and an update to 2001. *Journal of Climate*. 16: 206-223.
- Kao, S.C., and Govindaraju, R.S. 2010. A copula-based joint deficit index for droughts. *Journal of Hydrology*. 380: 121-134.
- Khadr, M. 2016. Forecasting of meteorological drought using Hidden Markov Model (case study: The upper Blue Nile river basin, Ethiopia). *Ain Shams Engineering Journal*, 7: 47-56.
- Khan, S., Gabriel, H. F. and Rana, T. 2008. Standardized precipitation index to track drought and assess impact of rainfall on water tables in irrigation areas. *Irrigation and Drainage Systems*, 22: 159-177.
- Kisi, O. and Cimen, M. 2011. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *Journal of Hydrology*, 399: 132-140.
- McKee, T. B., Doesken, N. J. and Kleist, J. 1993. The relationship of drought frequency and duration to time scales, Preprints, Eighth Conference on Applied Climatology, Anaheim, California, American Meteorology Society, 179-184.
- Morid, S., Smakhtin, V., and Bagherzadeh, K. 2007. Drought forecasting using artificial neural networks and time series of drought indices. *International Journal of Climatology*, 27(15): 2103-2111.
- Nalbantis, I., and Tsakiris, G. 2009. Assessment of hydrological drought revisited. *Water Resource Management*, 23(5), 881-897.
- Olson D.L., Delen D. 2008. *Advanced Data Mining Techniques*. USA: Springer; 141:3.
- Shukla, S.H., and Wood, A.W. 2008. Use of a Standardized Runoff Index index for Characterizing characterizing Hydrologic hydrologic Droughtdrought. *Geophysical Research Letters*, 35: 1-7.
- Sims, A. P., Nigoyi, D. S., and Raman, S. 2002. Adopting indices for estimating soil moisture: A North Carolina case study. *Geophysical Research Letters*, 29: 1-4.
- Szalai, S., Szinell, C. S., and Zoboki, J. 2000. Drought monitoring in Hungary, in *Early warning systems for drought preparedness and drought management*. World Meteorological Organization. Lisboa, 182-199.
- Vapnik, V. 1995. *The Nature of Statistical Learning Theory*. Springer, New York.
- Vapnik, V.N. 1998. *Statistical Learning Theory*. Wiley, New York. p. 736.
- Vicente-Serrano, S. M. and Lopez-Moreno, J. I. 2005. Hydrological response to different time scales of climatological drought: an evaluation of the standardized precipitation index in a mountainous Mediterranean basin, *Hydrology and Earth System Sciences*, 9: 523-533.
- Witten, I.H., Frank, E. and Hall, M.A. 2005. *Data Mining: Practical Machine Learning Tools and Techniques*. USA: EsvierElsevier. p. 664.



ISSN 2251-7480

Prediction of SPI drought index using support vector and multiple linear regressions

Saeed Samadianfard^{1*} and Esmail Asadi²

1) Assistant Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz.

* Corresponding author email: s.samadian@tabrizu.ac.ir

2) Assistant Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz.

Received: 20-07-2016

Accepted: 15-01-2017

Abstract

Drought is a natural phenomenon, which has a complex mechanism as a result of interactions of meteorological parameters and usually occurs in all climates. So, predicting drought indices and their chronological evaluation is an effective way for the drought management and adaptation with its consequences. In the current research, prediction of drought indices are carried out for Tabriz synoptic station, using support vector regression, multiple linear regression and standard precipitation index (SPI) for the time period of 1979 to 2012. In this regard, for predicting SPI indices in the periods of 3, 6, 9, 12, 24 and 48 months, six different input combinations including the antecedent correspondent values of the mentioned index have been utilized. The results of statistical analysis showed that both methods had significant accuracy. Nonetheless, the support vector regressions for predicting SPI-6, SPI-9 and SPI-24 had better performances, regarding the root mean squared errors of 0.4985, 0.4340 and 0.2427, respectively. However, the multiple linear regressions showed lower relative errors, for predicting SPI-3, SPI-12 and SPI-48. Meanwhile, it can be concluded that both examined methods including support vector and multiple linear regressions had acceptable predictions of drought index and can be used with an admissible confidentiality for the management of drought consequences.

Keywords: drought, prediction, standardized precipitation index, statistical analysis