

## Research Paper

# Integration of SVR and GPR Algorithms with Wavelet in Modeling Monthly Drought Forecasting

Jahanbakhsh Mohammadi<sup>1</sup>, Alireza Vafaeinejad<sup>2\*</sup>, Saeed Behzadi<sup>3</sup>, Hossein Aghamohammadi<sup>4</sup>, Amirhooman Hemmasi<sup>5</sup>

1. Ph.D. Student, Department of Remote Sensing and GIS, Faculty of Natural Resources and Environment, Science and Research Branch, Islamic Azad University, Tehran, Iran

2. Associate Professor, Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran

3. Assistant Professor, Faculty of Civil Engineering Shahid Rajaei Teacher Training University, Tehran, Iran

4. Assistant Professor, Department of Remote Sensing and GIS, Faculty of Natural Resources and Environment, Science and Research Branch, Islamic Azad University, Tehran, Iran

5. Professor, Natural Resources Engineering, Faculty of Natural Resources and Environment, Tehran science and Research Branch, Islamic Azad University, Tehran, Iran.

Received: 2021/11/21

Revised: 2021/12/12

Accepted: 2021/12/31

Use your device to scan and read the article online



DOI:

10.30495/wej.2023.29419.2344

### Keywords:

Drought, Wavelet, SPI, SVR, GPR

### Abstract

**Introduction:** Drought is one of the natural hazards that have random and nonlinear behavior due to its various climatic parameters. SPI index is the most common index extracted from rainfall that has been used in modeling drought by various researchers.

**Methods:** The use of computational intelligence methods to model drought in recent years has been much considered by researchers in the field of water resources. In this research, SVR and GPR algorithms individually and also the combination of these algorithms with wavelet algorithms have been modeled and predicted by SPI index, and the purpose was to evaluate the improvement of computational intelligence algorithms in combination with wavelet. In this research, the time series data of 10 synoptic stations in Iran in the period 1961 to 2017 have been used on a monthly basis for modeling the drought as the input of the studied algorithms.

**Findings:** The results of this study showed that the use of the wavelet method in combination with SVR and GPR computational intelligence algorithms improved the results in all time scales. Also, the modeling improvement is due to the use of wavelet in combination with the SVR model with an average RMSE difference of -0.1540 and  $R^2$  difference of 0.1491 and the GPR model with an average RMSE difference of -0.1554 and  $R^2$  difference of 0.1530. Compared to the single SVR and GPR models showed that the GPR model in general (all time scales and all stations) had a better improvement in the hybrid model than the single model.

**Citation:** Mohammadi J, Vafaeinejad A, Behzadi S, Aghamohammadi H, Hemmasi A. Integration of SVR and GPR algorithms with wavelet in modeling and monthly drought forecasting. Water Resources Engineering Journal. 2023; 16(56): 95-108

**\*Corresponding author:** Alireza Vafaeinejad

**Address:** Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Ira

**Tell:** +989122875986

**Email:** a\_vafaei@sbu.ac.ir

## Extended Abstract

### Introduction

Drought is one of the natural disasters that can cause great damage to various sectors of agriculture and other economic and social activities of the human system. Drought occurs mainly due to extreme climate change, usually triggered by human influences. Drought occurs naturally in almost any climate as a result of reduced rainfall from normal for long periods. Drought indices are an effective tool for monitoring and evaluating drought. Many indicators have been developed to identify and quantify drought events based on different types of variables. The standardized precipitation index (SPI) is one of the most useful indicators for meteorological drought analysis since most drought monitoring indices require a number of parameters other than rainfall. So, the SPI index can be as one of the most widely used drought indices extracted from rainfall data. Today, the most common methods used to assess and predict drought are data-based (data-driven) methods and the use of computational intelligence algorithms. The drought index used in this study is the standardized precipitation index (SPI). In this study, using two GPR and SVR algorithms in combination with wavelet in 10 synoptic stations (Abadan, Babolsar, Bandar Abbas, Kerman, Mashhad, Rasht, Saqez, Tehran, Tabriz, and Zahedan) in different climates in Iran. The model Drought construction and forecasting has been done using SPI drought index in time scales of 3, 6, 12, 18, 24 and 48 months.

### Materials and Methods

In this study, using monthly precipitation data between 1961 and 2017, the SPI drought index was implemented in 3, 6, 12, 18, 24, and 48-month time scales through programming in MATLAB software environment. The results of this step were validated using the available scientific software MDM and Drinc. Then, prediction models were designed using Markov chain. In this research, a total of four computational intelligence models, including two single support vector regression (SVR) models,

Gaussian process regression (GPR), and two hybrids wavelet models with these two models (WSVR-WGPR) for modeling and predicting SPI index. It has been used in 10 stations of this research. In the implementation of all four models, the MATLAB software programming environment has been used. In this study, four types of discrete wavelets were used, including Daubechies, Symlets, Coiflets, and Biorthogonal. Due to the better performance of the Daubechies wavelet, this type of wavelet was used as the final option in the research. In the Daubechies wavelet used between levels 1 to 45, level 3 showed the best performance among different SPI time scales; therefore, Daubechies level 3 wavelet was used in all hybrid models of this study. After training all four algorithms used, the results with the evaluation criteria of coefficient of determination ( $R^2$ ) and the square root of error (RMSE) were used to measure the difference between actual and estimated values.

### Findings and Discussion

The results of this study showed that computational intelligence methods have high accuracy in modeling and predicting SPI drought index. In the first stage, the results showed that the individual SVR and GPR models, if properly trained, have close results in modeling and predicting the SPI drought index. In the next step, it was observed that the wavelet technique would improve the modeling results. In using the wavelet technique in combination with two single models SVR and GPR, the choice of wavelet type is also effective in better modeling, so that in this research, the first four types of Daubechies, Simlet, Qoiflet and Biorthogonal wavelets are used in combination with two single models of this research. The results of these four types of wavelets showed the relative superiority of the Daubechies wavelet over the other three wavelets. In using the Daubechies wavelet, since this wavelet has 45 times and the order of choice was also effective in modeling, it was observed by testing the wavelet 45 times that the 3rd wavelet, in general, has higher accuracy in all time scales of SPI index, i.e., 3,6,12, 18, 24 and 48 months and

also in both SVR and GPR algorithms. Therefore, in this research, the third-order Daubechies wavelet was used in both algorithms of this research and also in all time scales. The results showed that the combination of wavelet techniques with both SVR and GPR models would improve the results. Distribution diagrams of SPI data related to Babolsar station showed that the modeling results for single and hybrid mode in 3 and 6-month time scales are less accurate than other time scales, and the fitting line separation, as well as its uncertainty, is higher. However, in all the models used in the research and in all time scales, the hybrid method has shown more accuracy. The numerical results of the study indicate that in all SPIs and stations under study, the differential values of  $R^2$  are positive, which indicates higher values of  $R^2$  in the hybrid model than in the individual modeling, which in some way indicates the improvement of hybrid modeling compared to the individual models. Also, the differential values of RMSE are negative in all studied models and stations, which indicates that the amount of RMSE in predicting hybrid models is lower than single models. In the research graphs, it can be seen that the differences in RMSE and  $R^2$  indicates a greater difference in time scales 3 and 6 than the time scales 12, 18, 24, and 48, which somehow goes back to the nature of the data of these time scales. The greatest improvement in  $R^2$  and RMSE is from the 3-month low to the 48-month high, respectively.

## Conclusion

From the findings of this study, it can be concluded that artificial neural network algorithms are efficient methods in modeling and predicting the SPI drought index. The use of wavelets in both models of computational intelligence algorithm (SVR and GPR) will also improve the results. It can also be concluded that for better modeling of the SPI drought index, it is necessary to select the optimal wavelet type and order. From the results of this study, it can be concluded that the wavelet technique has a

greater effect on the lower time scales, i.e., 3 and 6 months, than the higher scales, i.e., 24 and 48 months.

## Ethical Considerations compliance with ethical guidelines

The cooperation of the participants in the present study was voluntary and accompanied by their consent.

## Funding

No funding.

## Authors' contributions

Each author contributed to all sections of this article.

## Conflicts of interest

The authors declared no conflict of interest.

## مقاله پژوهشی

## بهره‌گیری از الگوریتم‌های تلفیقی SVR و GPR با موجک در مدل‌سازی و پیش‌بینی ماهانه خشک‌سالی

جهانبخش محمدی<sup>۱</sup>، علیرضا وفایی نژاد<sup>۲\*</sup>، سعید بهزادی<sup>۳</sup>، حسین آقامحمدی زنجیر آباد<sup>۴</sup>، امیر هومن حمصی<sup>۵</sup>

۱. دانشجوی گروه تخصصی سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده منابع طبیعی و محیط زیست، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

۲. دانشیار دانشکده مهندسی عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران

۳. استادیار دانشکده مهندسی عمران، دانشگاه تربیت دبیر شهید رجایی، تهران، ایران

۴. استادیار گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده منابع طبیعی و محیط زیست، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

۵. استاد، مهندسی منابع طبیعی، دانشکده منابع طبیعی و محیط‌زیست، واحد علوم و تحقیقات تهران، دانشگاه آزاد اسلامی، تهران، ایران

## چکیده

**مقدمه:** خشک‌سالی یکی از مخاطرات طبیعی است که به دلیل متأثر بودن از پارامترهای مختلف اقلیمی دارای رفتاری تصادفی و غیرخطی است. شاخص SPI به‌عنوان مرسوم‌ترین شاخص مستخرج شده از بارندگی می‌باشد که در مدل‌سازی خشک‌سالی مورد استفاده محققین مختلف قرار گرفته است. استفاده از روش‌های هوش محاسباتی جهت مدل‌سازی خشک‌سالی در سال‌های اخیر بسیار مورد توجه محققین حوزه منابع آب بوده است.

**روش:** در این تحقیق از الگوریتم‌های SVR و GPR به صورت منفرد و همچنین تلفیق این الگوریتم‌ها با فن موجک به مدل‌سازی و پیش‌بینی شاخص SPI پرداخته شده است و هدف بررسی میزان بهبود نتایج الگوریتم‌های هوش محاسباتی در تلفیق با موجک بوده است. در این تحقیق از اطلاعات سری زمانی ۱۰ ایستگاه سینوپتیک ایران در بازه زمانی ۱۹۶۱ تا ۲۰۱۷ به صورت ماهانه جهت مدل‌سازی خشک‌سالی به عنوان ورودی الگوریتم‌های مورد مطالعه استفاده شده است.

**یافته‌ها:** نتایج این پژوهش نشان داد که استفاده از روش موجک در تلفیق با الگوریتم‌های هوش محاسباتی SVR و GPR سبب بهبود نتایج در تمامی مقیاس‌های زمانی گردید. همچنین میزان بهبود مدل‌سازی ناشی از استفاده از موجک در تلفیق با مدل SVR با میانگین تفاضل RMSE برابر با  $-0/1540$  و تفاضل  $R^2$  برابر  $0/1491$  و در مدل GPR با میانگین تفاضل RMSE برابر با  $-0/1554$  و تفاضل  $R^2$  برابر  $0/1530$  نسبت به مدل‌های منفرد SVR و GPR نشان داد که مدل GPR در حالت کلی (همه مقیاس‌های زمانی و تمامی ایستگاه‌ها) بهبود بهتری در مدل هیبریدی نسبت به مدل منفرد داشته است.

تاریخ دریافت: ۱۴۰۰/۰۸/۳۰

تاریخ داوری: ۱۴۰۰/۰۹/۲۱

تاریخ پذیرش: ۱۴۰۰/۱۰/۱۰

از دستگاه خود برای اسکن و خواندن مقاله به صورت آنلاین استفاده کنید



DOI: 10.30495/wej.2023.29419.2344

## واژه‌های کلیدی:

خشک‌سالی، موجک، SPI، SVR، GPR

\* نویسنده مسئول: علیرضا وفایی نژاد

نشانی: دانشکده مهندسی عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران

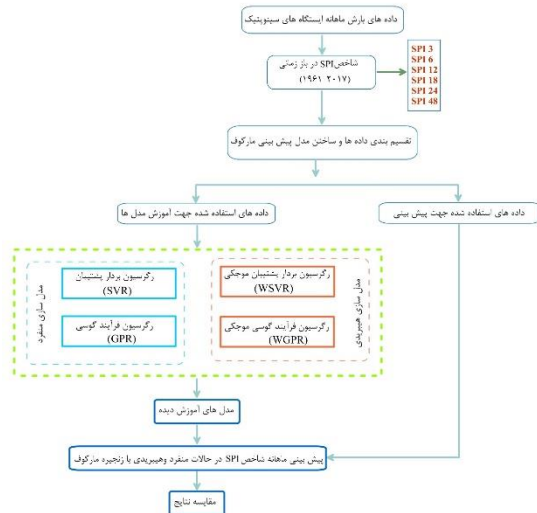
تلفن: ۰۹۱۲۲۸۷۵۹۸۶

پست الکترونیکی: a\_vafaei@sbu.ac.ir

## مقدمه

خشک‌سالی از جمله بلاای طبیعی است که می‌تواند خسارات هنگفتی را به بخش‌های مختلف کشاورزی و سایر فعالیت‌های اقتصادی و اجتماعی سیستم بشری وارد کند. خشک‌سالی اساساً در نتیجه تغییرات شدید آب و هوایی که معمولاً با تأثیرات انسانی تحریک می‌شود رخ می‌دهد. خشک‌سالی به‌طور طبیعی تقریباً در هر اقلیمی در نتیجه کاهش بارندگی از مقادیر نرمال برای مدت‌زمان طولانی رخ می‌دهد. شاخص‌های خشک‌سالی ابزاری کارآمد برای پایش و ارزیابی خشک‌سالی هستند. شاخص‌های زیادی برای شناسایی و کمی‌سازی وقایع خشک‌سالی بر اساس انواع مختلف متغیرها ایجاد شده است. شاخص بارش استاندارد شده (SPI) از جمله کاربردی‌ترین شاخص برای تجزیه و تحلیل خشک‌سالی هواشناسی است، از آنجایی که اکثر شاخص‌های پایش خشک‌سالی به تعدادی پارامتر دیگر به‌غیر از بارندگی نیاز دارند، لذا شاخص SPI را می‌توان به‌عنوان یکی از پرکاربردترین شاخص خشک‌سالی مستخرج از داده‌های فقط بارش دانست که در مطالعات مختلفی مورد بررسی قرار گرفته است. (۲۱، ۳۱) با توجه به هدف از مطالعه معمولاً شاخص SPI در مقیاس‌های زمانی مختلف محاسبه می‌گردد. به دلیل رفتار پیچیده بارش، مدل‌سازی و پیش‌بینی پدیده خشک‌سالی نیز که متأثر از بارش است همواره با مشکلاتی همراه بوده است و لذا این بسیار مورد توجه محققین بوده است. لذا تقویت تحقیقات در زمینه پیش‌بینی و مدل‌سازی خشک‌سالی برای جلوگیری از بلاای خشک‌سالی و کاهش خسارات ناشی از بلاای خشک‌سالی مهم است. کشور ایران به دلیل قرار گرفتن در اقلیمی عمدتاً خشک و نیمه‌خشک، همواره در طول تاریخ با تنش‌های آبی فراوانی همراه بوده است و همیشه با بحران خشک‌سالی مواجه بوده است. لذا پرداختن به موضوع خشک‌سالی در کشور ایران از اهمیت دوچندان برخوردار است. امروزه، متداول‌ترین روش‌های مورداستفاده برای ارزیابی و پیش‌بینی خشک‌سالی، روش‌های مبتنی بر داده‌ها (داده محور) و استفاده از الگوریتم‌های هوش محاسباتی می‌باشند. از جمله تحقیقات صورت گرفته در زمینه مدل‌سازی و پیش‌بینی خشک‌سالی با بهره‌گیری از الگوریتم‌های هوش محاسباتی می‌توان به کار از جمله این تحقیق‌ها می‌توان به مطالعه ده ژو و همکاران (۲۵) جهت پیش‌بینی خشک‌سالی در استان هنان در کشور چین اشاره کرد. در این تحقیق از مدل ARIMA و مدل ترکیبی ARIMA-SVR برای پیش‌بینی شاخص خشک‌سالی در مقیاس‌های زمانی ۱، ۳، ۶ و ۱۲ ماهه SPI استفاده شده است. نتایج این تحقیق نشان داد که مدل ترکیبی ARIMA-SVR دارای دقت بالاتری در تمامی مقیاس‌های زمانی نسبت به مدل منفرد ARIMA است و نیز نتایج پیش‌بینی شده ۱ تا ۲ ماه آینده تطابق نسبتاً خوبی با داده‌های واقعی دارد. تایلان و همکاران (۲۲) برای مدل‌سازی شاخص خشک‌سالی SPI در چاناک قلعه ترکیه از روش‌های هوش محاسباتی استفاده کردند. در این تحقیق ابتدا شاخص SPI در مقیاس‌های زمانی ۳، ۶، ۹ و ۱۲ ماهه با استفاده از سه مدل ANFIS، SVM و شبکه عصبی مصنوعی مدل‌سازی شد. در ادامه با استفاده از تکنیک موجک سیگنال‌های SPI تجزیه و دوباره از سه الگوریتم تحقیق جهت مدل‌سازی دوباره شاخص SPI استفاده شد. نتایج

این تحقیق نشان داد که مدل‌های ترکیبی توسعه‌یافته با استفاده از تکنیک پیش‌پردازش (موجک) بهتر عمل می‌کنند. در بین این مدل‌ها مشاهده شد که مدل W-ANFIS (ترکیب موجک با ANFIS) بهترین نتایج را برای دوره ۶ ماهه ارائه می‌دهد. کماسی و شرقی (۲۷) به پیش‌بینی خشک‌سالی با استفاده از موجک - ماشین بردار پشتیبان و شاخص استاندارد بارش در دریاچه ارومیه-ایران پرداختند. در این تحقیق سری زمانی SPI حوضه آبخیز دریاچه ارومیه با تبدیل موجک به سری‌های زمانی مکرر متعدد تجزیه شد. سپس این سری‌های زمانی به‌عنوان داده‌های ورودی به مدل ماشین بردار پشتیبان (SVM) برای پیش‌بینی خشک‌سالی اعمال شدند. نتایج این تحقیق نشان داد که مدل ترکیبی پیشنهادی (موجک با ماشین بردار پشتیبان) نسبتاً مناسب‌تر از مدل‌های خود رگرسیون کلاسیک مانند ANN است. در تحقیقی دیگر درسیا و همکاران (۶) به ارزیابی خشک‌سالی هیدرولوژیکی از طریق پیش‌بینی جریان با استفاده از شبکه‌های عصبی مصنوعی در تلفیق با تکنیک موجک در حوزه آبخیز بین دو ایالت تامیل نادو و کرالا در کشور هند پرداختند. در این تحقیق از یک شبکه عصبی پیش‌خور معمولی و شبکه عصبی در تلفیق با موجک جهت پیش‌بینی خشک‌سالی استفاده شد. نتایج این تحقیق نشان داد که ترکیب موجک با شبکه عصبی پیش‌خور معمولی تأثیر قابل‌توجهی بر نتایج پیش‌بینی خشک‌سالی دارد و مدل WANN به‌مراتب بهتر از مدل منفرد شبکه عصبی در منطقه مورد مطالعه تحقیق بود. در تحقیقی دیگر جربوایی و سواگ‌گامانه (۵) به پیش‌بینی خشک‌سالی با استفاده از شبکه‌های عصبی، شبکه‌های عصبی موجک و مدل‌های تصادفی در حوضه التجروا در شمال الجزایر پرداختند. در این تحقیق جهت پیش‌بینی شاخص خشک‌سالی SPI از مدل‌های شبکه عصبی، ARIMA، SARIMA و یک شبکه عصبی موجکی (WANN) در مقیاس‌های زمانی ۳، ۶ و ۱۲ ماهه استفاده شد. نتایج این تحقیق نشان داد که شبکه عصبی موجکی (WANN) دارای دقت بالاتری نسبت به سه مدل دیگر استفاده‌شده در تحقیق در پیش‌بینی شاخص خشک‌سالی در تمامی مقیاس‌های زمانی بود. تکنیک موجک به‌عنوان ابزاری توانمند جهت تجزیه سری زمانی به مؤلفه‌های فرکانسی، بسیار مورد توجه محققین حوزه منابع آب و خشک‌سالی در سال‌های اخیر بوده است. (۲۲، ۵) این تکنیک در ترکیب با الگوریتم‌های هوش محاسباتی و ساختن مدل‌های هیبریدی باعث افزایش توانایی این الگوریتم‌ها در مدل‌سازی پدیده‌های مختلف از جمله خشک‌سالی شده است. تمامی پژوهش‌های قبلی، بر کارایی روش‌های هوش محاسباتی از جمله انواع شبکه‌های عصبی مصنوعی، مدل رگرسیون بردار پشتیبان، الگوریتم فرآیند گوسی و نیز تبدیل موجک جهت مدل‌سازی و پیش‌بینی پدیده‌های مختلف از جمله خشک‌سالی دارد. تاکنون تحقیقی در مدل‌سازی و پیش‌بینی خشک‌سالی با استفاده از دو الگوریتم GPR و SVR به‌صورت منفرد و در ترکیب با موجک در ایستگاه‌های با اقلیم‌های مختلف صورت نگرفته است. در این تحقیق با استفاده از دو الگوریتم GPR و SVR در ترکیب با موجک در ۱۰ ایستگاه سینوپتیک در اقلیم‌های مختلف در کشور ایران به مدل‌سازی و پیش‌بینی خشک‌سالی با استفاده از شاخص خشک‌سالی SPI در مقیاس‌های زمانی ۳، ۶، ۹، ۱۲، ۱۸، ۲۴ و ۴۸ ماهه پرداخته شده است. در این



شکل ۲- فلوجارت روند انجام مراحل پژوهش

### شاخص بارش استاندارد شده

مکی و همکاران (۱۹۹۳) شاخص بارش استاندارد شده را برای اولین بار در مقیاس های زمانی مختلف جهت مدل سازی احتمالاتی بارش ارائه دادند (۲۵). این شاخص به عنوان یکی از شاخص های اصلی در مدل سازی خشک سالی مورد توجه محققین مختلفی قرار گرفته است. در این مطالعه مقدار شاخص SPI بر اساس روابط بیان شده در (۱۶) برای سری زمانی بارش ۱۰ ایستگاه نام برده شده در بخش های قبل محاسبه و به عنوان ورودی مدل های مورد مطالعه در این پژوهش در نظر گرفته شده است.

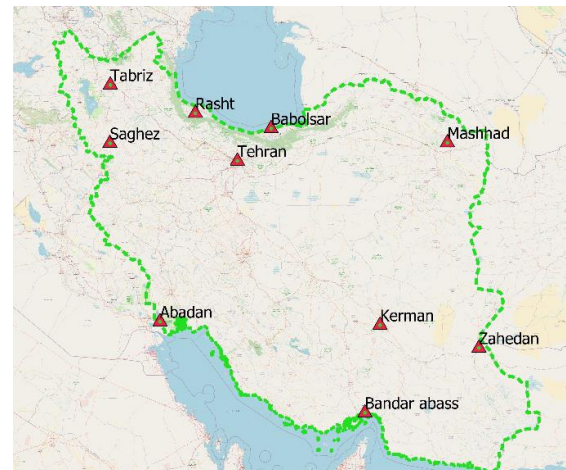
### رگرسیون بردار پشتیبان

رگرسیون بردار پشتیبان که در واقع تعمیمی است از ماشین بردار پشتیبان در سال ۱۹۹۵ توسط وپنیک معرفی شد. رگرسیون بردار پشتیبان برخلاف انواع شبکه های عصبی از جمله MLP, RBF و حتی ANFIS به جای کمینه کردن خطای مدل سازی، ریسک عملیاتی را به عنوان تابع هدف در نظر گرفته و مقدار بهینه آن را محاسبه می کند. منظور از رگرسیون، به دست آوردن یک ابر صفحه است که بر داده های مورد نظر برازش داده شود. فاصله هر نقطه از این ابر صفحه نشان دهنده خطای آن نقطه خاص است. بهترین روشی که تاکنون برای رگرسیون خطی پیشنهاد شده است، روش کمترین مربعات است. با این وجود، برای مسائل رگرسیون، در حضور داده های پرت عملکرد کمترین مربعات مطلوب نیست؛ بنابراین می بایست یک رگرسور نیرومند که نسبت به تغییرات کوچک در مدل حساس نباشد را توسعه داد. این کار با استفاده از رگرسیون بردار پشتیبان امکان پذیر است. یک مسئله رگرسیون ساده را با مجموعه داده  $X = \{u_i, v_i; i = 1, \dots, n\}$  در نظر بگیرید که در آن  $u_i$  ورودی و  $v_i$  خروجی باشند. تابع  $g(u)$  که یک تابع

تحقیق ابتدا به بررسی منطقه مورد مطالعه و ایستگاه های سینوپتیک مورد استفاده پرداخته می شود. سپس توضیحی در خصوص شاخص SPI و نحوه استخراج آن از داده های بارندگی پرداخته می شود. در ادامه به بررسی مبانی ریاضی روش موجک و روش های SVR و GPR جهت مدل سازی شاخص SPI پرداخته می شود.

### منطقه مورد مطالعه و داده های مورد استفاده

ایران از نظر اقلیمی عمدتاً کشوری خشک و نیمه خشک است که عمده اقلیم ایران متأثر از رشته کوه البرز، زاگرس و همچنین دشت های کویر در مرکز ایران هست. در این تحقیق از داده های بارش ۱۰ ایستگاه سینوپتیک آبادان، بابلرس، بندرعباس، کرمان، مشهد، رشت، سقز، تهران و تبریز جهت استخراج شاخص SPI مورد استفاده قرار گرفته است. دلیل استفاده از این ۱۰ ایستگاه داشتن ۵۷ سال اطلاعات بارش در زمان انجام تحقیق و نیز توزیع مکانی مناسب در کل پهنه کشور ایران و همچنین قرار گرفتن این ۱۰ ایستگاه در اقلیم های مختلف است. در شکل (۱) موقعیت ایستگاه ها به همراه نقشه ایران نمایش داده شده است.



شکل ۱- موقعیت ایستگاه های سینوپتیک مورد مطالعه در این تحقیق

### روش شناسی تحقیق

در این تحقیق از تلفیق روش موجک و دو الگوریتم SVR و GPR جهت مدل سازی سری زمانی SPI مورد استفاده قرار گرفته است. همان طور که در فلوجارت شکل (۲) دیده می شود ابتدا شاخص SPI از داده های بارندگی ایستگاه های هواشناسی استخراج می گردد و در مرحله بعد مقادیر سری زمانی SPI به دسته آموزشی و ارزیابی تقسیم بندی می گردند. از مقادیر سری زمانی آموزشی SPI جهت برآورد پارامترهای مدل های مورد استفاده بهره گیری شد؛ و با استفاده از مقادیر ارزیابی در زمان های مورد نظر مقادیر SPI به شیوه زنجیره مارکوف و ماهانه پیش بینی می گردند.

مختلف یا انواع مختلف ساختارهای پیوسته می‌شود و امکان انتخاب صحیح را برای محققین فراهم می‌کند. (۷) فرآیندهای گاوسی در مدل‌سازی آماری بسیار مهم هستند زیرا دارای ویژگی‌های نرمال هستند. (۱۲) مدل‌های فرایند گاوسی به دلیل دانش قبلی درباره وابستگی‌های تابعی و داده‌ها، برای تعمیم به هیچ فرایند اعتبارسنجی نیاز ندارند و مدل‌های رگرسیون فرایند گاوسی قادر به درک توزیع پیش‌بینی متناظر با ورودی آزمون هستند. (۲۷) از جمله روش‌های متداول برای حل مسائل غیرخطی، استفاده از توابع کرنل است. از مهم‌ترین توابع کرنل می‌توان به کرنل خطی، چندجمله‌ای، کرنل گاوسی و کرنل مربع گاوسی اشاره کرد. در این پژوهش از تابع کرنل مربع گاوسی که رابطه آن در معادله (۵) آمده است، استفاده شد. (۲۷، ۲۶، ۱۹، ۱۸، ۷)

$$K(x_i, x_j) = \exp\left(-\frac{1}{2\sigma^2} |x_i - x_j|^2\right) \quad (5)$$

که در آن  $\sigma$  پارامتر کرنل است که در این پژوهش با روش سعی و خطا محاسبه شده است.

### تبدیل موجک

تبدیل موجک (ویولت) یک تابع ریاضی است که توانایی بالایی در استخراج ویژگی سری‌های زمانی دارد. این ابزار می‌تواند سری‌های زمانی را به مؤلفه‌های فرکانسی آن‌ها تجزیه کند تا بر ضعف روش‌های داده محور در مواجهه با سری‌های زمانی غیر ایستا غلبه کند تبدیل ویولت به دو صورت تبدیل ویولت پیوسته و تبدیل ویولت گسسته انجام می‌شود. (۱۰، ۸) ضرایب تبدیل ویولت پیوسته از رابطه (۶) محاسبه می‌شود. در این رابطه موجک مادر از رابطه (۷) محاسبه می‌گردد.

$$W_f(a, b) = \int_{-\infty}^{+\infty} f(t) \Psi_{a,b}^*(t) dt \quad (6)$$

$$\Psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) \quad a, b \in R, a \neq 0 \quad (7)$$

در رابطه (۶) علامت \* نماد مزدوج مختلط  $a, b, t$  به ترتیب پارامترهای مقیاس، انتقال و زمان هستند. (۱۰، ۸، ۱۷) محاسبه ضرایب CWT باعث افزایش زمان محاسبات و تولید حجم زیاد ضرایب می‌شود. با توجه به گسسته بودن داده‌های ورودی، در این تحقیق از روش تبدیل ویولت گسسته جهت تجزیه سری‌های زمانی SPI استفاده شده است. DWT به صورت رابطه (۸) تعریف می‌شود

$$W_f(j, k) = \int_{-\infty}^{+\infty} f(t) \Psi_{j,k}^*(t) dt \quad (8)$$

$$\Psi_{j,k}(t) = a_0^{-1/2} \Psi(a_0^{-1} t - b_0 k) \quad (9)$$

خطی است به عنوان تابع رگرسیون در نظر گرفته می‌شود. رگرسیون استاندارد SVR از یک تابع جریمه  $L_\varepsilon(v, g(u))$  استفاده می‌کند که انحراف تابع تخمین زده شده از تابع اصلی را توصیف می‌کند. انواع مختلفی از توابع جریمه را می‌توان استفاده کرد، مانند خطی، درجه دوم، نمایی، تابع جریمه هوبر و غیره. (۴) در این پژوهش از تابع استاندارد وینیک استفاده شده است که در رابطه (۱) تعریف شده است.

$$L_\varepsilon(v, g(u)) = \begin{cases} 0 & \text{for } |v - g(u)| \leq \varepsilon \\ |v - g(u)| - \varepsilon & \text{otherwise} \end{cases} \quad (1)$$

خروجی مدل رگرسیون بردار پشتیبان به صورت رابطه (۲) است.

$$g(u) = w \cdot u + b \quad (2)$$

که در رابطه فوق  $w \in X$  فضای ورودی است و  $b \in R$  ترم بایاس هست و نیز  $w \cdot u$  ضرب داخلی  $w$  و  $u$  است. بنابراین صورت مسئله رگرسیون را می‌توان به عنوان مسئله بهینه‌سازی مطابق رابطه (۳) بیان کرد.

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (3)$$

که در آن قیود به صورت روابط (۴) بیان می‌شود.

$$\begin{aligned} v_i - (w \cdot u_i + b) &\leq \varepsilon + \xi_i \\ (w \cdot u_i + b) - v_i &\leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* &\geq 0, \quad i = 1, 2, \dots, n \end{aligned} \quad (4)$$

که در روابط فوق  $\xi_i, \xi_i^*$  متغیرهای جریمه می‌باشند و درواقع میزان انحراف نمونه‌های آموزشی خارج از منطقه  $\varepsilon$  معرفی شده‌اند و  $C$  عدد ثابت مثبتی است. یکی از مسائل مهم مدل رگرسیون بردار پشتیبان تنظیم کردن (میزان بهینه) مقادیر  $C$  و  $\varepsilon$  است.

### رگرسیون فرایند گاوسی

الگوریتم رگرسیون فرایند گاوسی از جمله روش‌های تقریب توابع یادگیری نظارت شده در حوزه یادگیری ماشین است. (۹) مدل‌های رگرسیون فرایند گاوسی بر مبنای این فرض هستند که مشاهدات تنظیم باید حامل اطلاعاتی درباره همدیگر باشد. فرایندهای گاوسی راهی برای مشخص کردن اولویتی به صورت مستقیم روی فضای تابع هستند. (۱۹) این کار تعمیم طبیعی توزیع گاوسی است که میانگین و کوواریانس آن به ترتیب بردار و ماتریس هستند به عبارت دیگر توزیع گاوسی روی بردارهاست درحالی که فرایند گاوسی روی توابع است. روش رگرسیون فرایند گاوسی به دلیل مزایای نان پارامتریک و احتمالاتی بودن در بسیاری از زمینه‌ها استفاده شده است. برخلاف روش ماشین بردار پشتیبان، رگرسیون فرایند گاوسی از روش احتمالی برای توصیف تابع رگرسیون استفاده می‌کند. (۲۶) یکی از مهم‌ترین ویژگی‌های فرایند گاوسی تنوع توابع کوواریانس است که منجر به ایجاد توابع با درجات

که در روابط فوق  $a_0$  و  $b_0$  ثابت،  $z$  سطح تجزیه و  $k$  فاکتور زمان انتقال است.

### زنجیره مارکوف

مدل زنجیره مارکوف رویکردی ریاضی جهت ساخت مدل‌های سری‌های زمانی پدیده‌های تصادفی است. (۲۰) وابستگی این زنجیره به زمان یا از طریق ضرایب همبستگی و یا با استفاده از ماتریس‌های احتمالاتی انتقال بیان می‌شود. (۱۱) مدل زنجیره مارکوف مرتبه اول اصلی‌ترین شکل مدل زنجیره مارکوف هست که در واقع با داشتن چند مرحله و گام عقب‌تر یک پدیده می‌توان یک مرحله جلوتر را پیش‌بینی کرد که در این پژوهش همانند برخی پژوهش‌های انجام‌شده جهت پیش‌بینی خشک‌سالی (۲۰، ۱۱) از سه مرحله متوالی SPI جهت پیش‌بینی یک مرحله جلوتر استفاده شده است.

### معیارهای ارزیابی نتایج مدل‌سازی

برای ارزیابی مدل‌ها و مقایسه بین نتایج چهار روش مدل‌سازی ذکر شده، از دو معیار ارزیابی ضریب تعیین ( $R$ ) و جذر میانگین مربعات خطا (RMSE) جهت اندازه‌گیری اختلاف بین مقادیر مشاهداتی و برآورد شده استفاده شده است. این معیارها به صورت روابط ۱۰ و ۱۱ به ترتیب تعریف می‌شوند. (۱۵)

$$R = \sqrt{1 - \frac{\sum_{i=0}^n (y_0 - y_p)^2}{\sum_{i=0}^n (y_0 - \frac{\sum y_0}{n})^2}} \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^n (y_0 - y_p)^2} \quad (11)$$

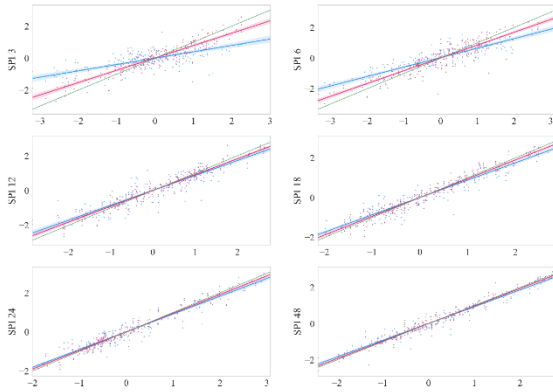
در این روابط،  $y_0$  مقادیر مشاهداتی (واقعی) پارامتر SPI هست و  $y_p$  مقادیر برآورد شده SPI به روش مدل‌سازی با استفاده از چهار مدل مورد استفاده در این مطالعه هست. پارامتر  $n$  بیانگر تعداد داده‌های تست مورد استفاده هست.

### نتایج

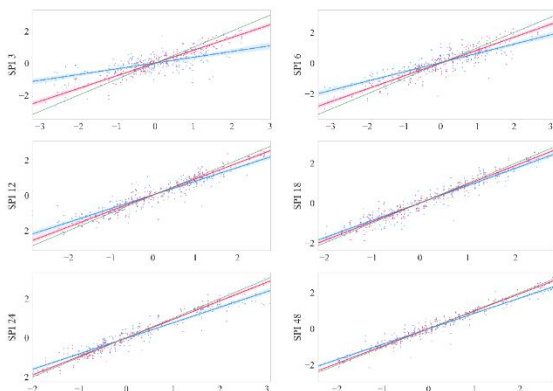
همان‌طور که در بخش روش تحقیق بیان شد ابتدا مقادیر سری زمانی SPI برای ایستگاه‌های منطقه مورد مطالعه استخراج گردید و سپس داده‌های SPI به دو بخش آموزشی مدل (جهت استخراج پارامترهای مدل‌سازی) و بخش پیش‌بینی ماهانه تقسیم‌بندی گردید. جهت انتخاب موجک مادر بهینه برای تمامی مدل‌های هیبریدی مورد مطالعه در این پژوهش و ایستگاه‌های مورد نظر، مقادیر RMSE و  $R^2$  را برای موجک‌های مادر دابیشز از مرتبه ۱ تا ۴۵ مورد بررسی قرار گرفت که نتایج بیانگر بهینه بودن مقدار موجک مادر دابیشز مرتبه سوم نسبت به دیگر مرتبه‌ها در کلیه مدل‌سازی‌ها بود و در این پژوهش از موجک مادر دابیشز مرتبه سوم استفاده گردید. با توجه به تعداد زیاد نمودار خروجی،

در این تحقیق صرفاً نمودارهای سری زمانی و پراکنش داده‌های مربوط به پیش‌بینی ماهانه ایستگاه بابلسر نمایش داده شده است. شکل (۳) و (۴) به ترتیب نتایج مربوط به پیش‌بینی ماهانه مقادیر SPI مربوط به ایستگاه بابلسر را برای مدل‌سازی به روش SVR و GPR را به صورت منفرد و همچنین به صورت هیبریدی و تلفیق با موجک نشان می‌دهد. این نمودارها شامل سری زمانی پیش‌بینی ماهانه SPI در مقیاس‌های زمانی ۳، ۶، ۱۲، ۱۸، ۲۴ و ۴۸ ماهه هست. همان‌طور که در این اشکال دیده می‌شود سری زمانی داده‌های SPI مشاهداتی (محاسبه شده از داده‌های بارش ایستگاه سینوپتیک) با ستاره سیاه‌رنگ، مقادیر پیش‌بینی ماهانه به روش ترکیبی موجک و روش‌های رگرسیون فرآیند گوسی و رگرسیون بردار پشتیبان با خط قرمز رنگ همچنین مقادیر مدل‌سازی شده به روش الگوریتم‌های رگرسیون فرآیند گوسی و رگرسیون بردار پشتیبان منفرد با خط آبی‌رنگ نمایش داده شده است. در حالت کلی می‌توان رفتار هموارتری از مدل‌های هیبریدی در بیش‌بینی نسبت به مدل‌های منفرد SVR و GPR در رفتار سری زمانی مشاهده کرد. در مدل GPR مدل‌سازی منفرد در بین سال‌های ۲۰۰۴ تا ۲۰۰۶ برای مقیاس‌های زمانی ۲۴ و ۴۸ ماهه نتوانسته به خوبی به مقدار بیشینه برسد و با اختلاف بیشتری نسبت به دیگر مقیاس‌های زمانی و همچنین ماه‌های دیگر مقادیر SPI را کمتر برآورد کرده است در حالی که در مدل هیبریدی به خوبی به داده مشاهداتی SPI نزدیک بوده است و این ضعف را ندارد. شکل‌های (۵ و ۶) نشان‌دهنده پراکنش داده‌های مربوط به SPI مشاهداتی ایستگاه بابلسر مربوط به مدل‌های SVR و GPR را در حالت منفرد و هیبریدی (محور افقی SPI مشاهداتی) و مقادیر پیش‌بینی شده داده‌های SPI با استفاده از مدل منفرد و مدل هیبریدی (محور قائم)، نقاط قرمز و خط برازش داده شده قرمز رنگ مربوط به مدل هیبریدی و نقاط آبی و خط برازش داده شده آبی‌رنگ مربوط به مدل منفرد هست. خط سبز بیانگر محور ۴۵ درجه جهت مقایسه بصری نتایج مدل‌ها نسبت به حالت ایدئال است که در واقع هرچه میزان برآورد شده به خط ۴۵ درجه نزدیکتر باشد به معنی نزدیکی به مقدار واقعی و دقت بالاتر است. همان‌طور که در اشکال (۵ و ۶) دیده می‌شود نتایج مدل‌سازی برای حالت منفرد و هیبریدی در مقیاس‌های زمانی ۳ و ۶ ماهه به نسبت دقت کمتری نسبت به دیگر مقیاس‌های زمانی دارد و جدایی خط برازش و همچنین میزان عدم قطعیت آن (پهنای خط برازش بیانگر میزان عدم قطعیت هست) نسبت به خط ۴۵ درجه (حالت ایدئال) بیشتر است. همان‌طور که در نمودارهای سری زمانی نیز کاملاً مشهود بود در مقیاس زمانی ۲۴ و ۴۸ ماهه مدل GPR خط برازش داده شده مربوط به داده‌های منفرد جدایی بیشتری نسبت به مدل هیبریدی به خط برازش ایدئال دارد. این در حالی است که در مدل SVR چنین اختلافی مشاهده نمی‌شود که می‌تواند مربوط به ناتوانی



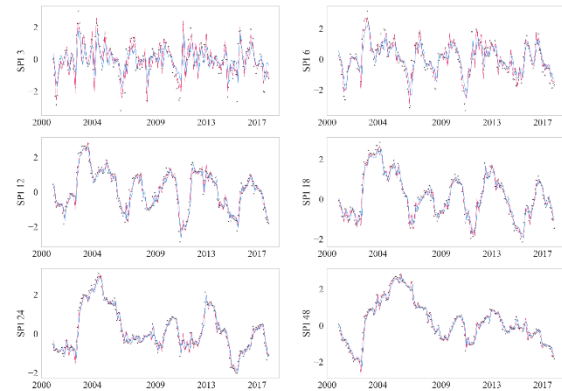


شکل ۵- نمودار پراکنش داده‌های SPI مشاهداتی ایستگاه بابلسر مربوط به مدل منفرد و هیبریدی (محور افقی SPI مشاهداتی) و مقادیر پیش‌بینی شده داده‌های SPI با استفاده از مدل منفرد و مدل هیبریدی (محور قائم). نقاط قرمز و خط برازش داده‌شده قرمز رنگ مربوط به مدل هیبریدی و نقاط آبی و خط سبز داده‌شده آبی رنگ مربوط به مدل منفرد هست. خط سبز بیانگر محور ۴۵ درجه جهت مقایسه بصری نتایج مدل‌ها نسبت به حالت ایدئال است.



شکل ۶- نمودار پراکنش داده‌های SPI مشاهداتی ایستگاه بابلسر مربوط به مدل GPR منفرد و هیبریدی (محور افقی SPI مشاهداتی) و مقادیر پیش‌بینی شده داده‌های SPI با استفاده از مدل منفرد و مدل هیبریدی (محور قائم). نقاط قرمز و خط برازش داده‌شده قرمز رنگ مربوط به مدل هیبریدی و نقاط آبی و خط سبز داده‌شده آبی رنگ مربوط به مدل منفرد هست. خط سبز بیانگر محور ۴۵ درجه جهت مقایسه بصری نتایج مدل‌ها نسبت به حالت ایدئال است.

مدل منفرد GPR در برآورد بیشینه SPI در بازه‌های زمانی ۲۰۰۴ تا ۲۰۰۶ باشد.



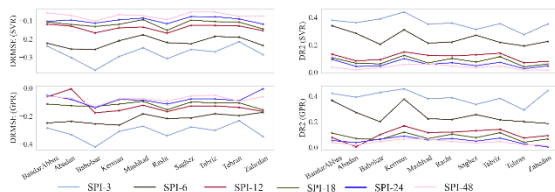
شکل ۳- نمایش سری زمانی SPI ایستگاه بابلسر مربوط به داده‌های مشاهداتی (ستاره سیاه رنگ) - مدل‌سازی با مدل ترکیبی موجک و مدل‌های هوش محاسباتی (نمودار قرمز رنگ) - مدل‌سازی با الگوریتم‌های هوش محاسباتی SVR (نمودار آبی رنگ)



شکل ۴- نمایش سری زمانی SPI ایستگاه بابلسر مربوط به داده‌های مشاهداتی (ستاره سیاه رنگ) - مدل‌سازی با مدل ترکیبی موجک و مدل‌های هوش محاسباتی (نمودار قرمز رنگ) - مدل‌سازی با الگوریتم‌های هوش محاسباتی GPR (نمودار آبی رنگ)

## مقادیر تفاضل $R^2$ مدل SVR با WSVR (D) مقادیر تفاضل RMSE مربوط به مدل SVR با WSVR.

نمودارهای مربوط به شکل (۸) بیانگر میزان اختلاف بین مدل‌های هیبریدی و منفرد برای تمامی ایستگاه‌ها به صورت گراف‌های رنگی هست. همان‌طور که در این گراف‌ها دیده می‌شود مقدار اختلاف‌ها در RMSE و  $R^2$  بیانگر میزان اختلاف بیشتر در مقیاس‌های زمانی ۳ و ۶ نسبت به مقیاس‌های زمانی ۱۲، ۱۸، ۲۴ و ۴۸ هست که به نحوی به ماهیت داده‌های این مقیاس‌های زمانی برمی‌گردد. همان‌طور که در این شکل دیده می‌شود به ترتیب بیشترین بهبود در مقدار  $R^2$  و RMSE از مقیاس پایین ۳ ماهه به سمت مقیاس‌های زمانی بالاتر ۴۸ ماهه هست. در کل در مقیاس‌های زمانی ۱۲، ۱۸، ۲۴ و ۴۸ (یک سال به بالا) اختلاف کمی باهم دارند که می‌تواند مربوط به ماهیت هموارتر بودن رفتار سری زمانی داده‌های SPI در این مقیاس‌های زمانی باشد.



شکل ۸- نمودار مقادیر تفاضل RMSE و تفاضل  $R^2$  برای ده ایستگاه مورد مطالعه در SPI های ۳ تا ۴۸ ماهه

## نتیجه‌گیری و بحث

خشک‌سالی دوره‌های طولانی مدت کاهش بارندگی است که منجر به کمبود آب در دسترس و پیامدهای نامطلوب ناشی از آن بر محصولات کشاورزی، حیوانات و جوامع بشری می‌شود. بر این اساس مدل‌سازی و پیش‌بینی خشک‌سالی با کاهش تبعات ناشی از خشک‌سالی، نقش مهمی در برنامه‌ریزی و مدیریت منابع آب ایفا می‌کند. مدل‌های زیادی تاکنون به وسیله محققین مختلف جهت مدل‌سازی و پیش‌بینی خشک‌سالی استفاده شده است از جمله جدیدترین و کارآمدترین روش‌ها، مدل‌های داده محور و استفاده از الگوریتم‌های هوش محاسباتی است. در این تحقیق با بهره‌گیری از دو الگوریتم هوش محاسباتی شامل SVR و GPR و استفاده از تکنیک مویج در تلفیق با آن‌ها به مدل‌سازی و پیش‌بینی خشک‌سالی پرداخته شده است. در این پژوهش در ابتدا مقدار شاخص SPI در مقیاس‌های زمانی ۳، ۶، ۱۲، ۱۸، ۲۴ و ۴۸ در ۱۰ ایستگاه سینوپتیک در بازه زمانی ۱۹۶۱ تا ۲۰۱۷ محاسبه گردید. در ادامه با تقسیم داده‌های SPI به دو بخش داده‌های مورد استفاده جهت آموزش مدل‌های مورد مطالعه در این پژوهش و همچنین بخش داده‌های مورد استفاده جهت پیش‌بینی ماهانه تقسیم گردید. از زنجیره مارکوف با سه گام عقب‌تر جهت پیش‌بینی یک ماه آینده مورد استفاده قرار گرفت. در این پژوهش از دو مدل SVR و GPR به صورت منفرد و نیز ترکیب این دو مدل با مویج استفاده شد. در تمامی مقیاس‌های زمانی مورد بررسی مدل‌های هیبریدی دقت بالاتری نسبت به مدل‌های

جداول آورده شده در شکل (۷) مربوط به تفاضل RMSE و  $R^2$  مدل‌های هیبرید با مدل‌سازی SVR و GPR به صورت منفرد جهت پیش‌بینی مقدار SPI ماهانه مربوط به ۱۰ ایستگاه مورد مطالعه در مقیاس‌های زمانی مورد بررسی هست. مقادیر عددی به دست آمده در این جداول نشان‌دهنده این است که در تمامی SPI ها و ایستگاه‌های مورد مطالعه، مقادیر تفاضلی  $R^2$  مثبت است که بیانگر مقادیر بیشتر  $R^2$  مدل هیبریدی نسبت به مدل‌سازی‌های منفرد هست که به نحوی نشان‌دهنده بهبود مدل‌سازی هیبریدی نسبت به مدل‌های منفرد است. همچنین مقادیر تفاضلی RMSE نیز در کلیه مدل‌های مورد بررسی و ایستگاه‌های مورد مطالعه منفی هست که نشان‌دهنده کمتر بودن مقدار RMSE در پیش‌بینی مدل‌های هیبریدی نسبت به مدل‌های منفرد هست. در رنگ‌بندی جداول هر مقدار سلولی رنگ سبز پررنگ‌تر باشد یعنی مدل‌سازی هیبریدی نسبت به منفرد اختلاف بیشتری دارد و بهبود مدل‌سازی در این موارد بیشتر بوده است. همان‌طور که در این جداول دیده می‌شود اختلاف بین مدل هیبریدی و مدل منفرد در همه ایستگاه‌ها برای مقیاس‌های پایین‌تر بیشتر است و بیشترین اختلاف بین مدل‌سازی هیبریدی و منفرد مربوط به مقیاس زمانی ۳ ماهه هست. بیشترین اختلاف مربوط به  $R^2$  بین مدل‌سازی هیبریدی و منفرد مربوط به ایستگاه کرمان و مقیاس زمانی ۳ ماهه و مربوط به مدل GPR هست که مقدار آن برابر ۰/۴۵۷۳ بوده است که در آن میزان اختلاف RMSE برابر با ۰/۳۰۹۴- است. بیشترین اختلاف RMSE بین مدل‌سازی هیبریدی و منفرد مربوط به ایستگاه بابل‌سر و مقیاس زمانی ۳ ماهه و مربوط به مدل GPR هست که مقدار آن برابر ۰/۴۲۱۳- بوده است که در آن میزان اختلاف  $R^2$  برابر با ۰/۴۲۸۷- بوده است.

A)						B)						
SPI_3	SPI_6	SPI_12	SPI_18	SPI_24	SPI_48	Station	SPI_3	SPI_6	SPI_12	SPI_18	SPI_24	SPI_48
0.4212	0.3649	0.6763	0.1109	0.0531	0.0380	BandarAbbas	-0.2848	-0.2489	-0.0022	-0.1137	-0.0514	-0.0578
0.3915	0.2726	0.0046	0.0687	0.0370	0.0181	Abadan	-0.3321	-0.2380	-0.0059	-0.1260	-0.0805	-0.0722
0.4287	0.2000	0.0996	0.0601	0.0639	0.0461	Babolour	-0.4213	-0.2545	-0.3766	-0.1305	-0.1411	-0.1326
0.4373	0.3378	0.1674	0.1190	0.0883	0.0650	Kerman	-0.3094	-0.2668	-0.1611	-0.1145	-0.0818	-0.0771
0.3793	0.2230	0.1150	0.0684	0.0576	0.0633	Mashhad	-0.2723	-0.1835	-0.1223	-0.0937	-0.0852	-0.0755
0.3878	0.2179	0.1185	0.1025	0.0688	0.0458	Razi	-0.3416	-0.2181	-0.1679	-0.1548	-0.1146	-0.0941
0.3352	0.2559	0.1300	0.0757	0.0480	0.0345	Saghez	-0.2776	-0.2127	-0.1271	-0.0993	-0.0773	-0.0558
0.3530	0.2148	0.1406	0.1136	0.0724	0.0446	Tabriz	-0.3026	-0.1835	-0.1275	-0.1052	-0.0773	-0.0517
0.2912	0.2005	0.0729	0.0396	0.0269	0.0202	Tehran	-0.2325	-0.1972	-0.1375	-0.1050	-0.0899	-0.0633
0.4443	0.1858	0.0922	0.0658	0.0025	0.0144	Zahedan	-0.3464	-0.1732	-0.1667	-0.1552	-0.0649	-0.0615

C)						D)						
SPI_3	SPI_6	SPI_12	SPI_18	SPI_24	SPI_48	Station	SPI_3	SPI_6	SPI_12	SPI_18	SPI_24	SPI_48
0.3797	0.3390	0.1314	0.1061	0.0940	0.0382	BandarAbbas	-0.2393	-0.2233	-0.1186	-0.1070	-0.1034	-0.0571
0.3620	0.2839	0.0839	0.0661	0.0430	0.0178	Abadan	-0.3027	-0.2558	-0.1295	-0.1204	-0.0962	-0.0703
0.3895	0.2039	0.0917	0.0610	0.0460	0.0344	Babolour	-0.3716	-0.2574	-0.1667	-0.1316	-0.1131	-0.1025
0.4398	0.3101	0.1489	0.1235	0.0994	0.0556	Kerman	-0.2967	-0.2699	-0.1391	-0.1199	-0.0954	-0.0667
0.3513	0.2125	0.1239	0.0680	0.0574	0.0606	Mashhad	-0.2488	-0.1772	-0.1349	-0.0939	-0.0854	-0.0725
0.3595	0.2207	0.1193	0.1013	0.0701	0.0457	Razi	-0.3089	-0.2205	-0.1688	-0.1531	-0.1168	-0.0939
0.3129	0.2303	0.1274	0.0744	0.0478	0.0331	Saghez	-0.2575	-0.2262	-0.1247	-0.0966	-0.0769	-0.0529
0.3551	0.2181	0.1390	0.1135	0.0729	0.0447	Tabriz	-0.2708	-0.1869	-0.1202	-0.1050	-0.0779	-0.0520
0.2743	0.1920	0.0684	0.0404	0.0267	0.0153	Tehran	-0.2151	-0.1895	-0.1287	-0.1074	-0.0903	-0.0700
0.3536	0.2235	0.0799	0.0604	0.0467	0.0168	Zahedan	-0.2865	-0.2357	-0.1546	-0.1451	-0.1184	-0.0737

شکل ۷- مقایسه نتایج مدل‌سازی مدل SVR و GPR در SPI های مختلف. A) مقادیر تفاضل  $R^2$  مدل GPR با WGPR (B) مقادیر تفاضل RMSE مربوط به مدل GPR با WGPR (C)

تلفیق با الگوریتم‌های هوش محاسباتی مانند SVR و GPR دارد. در پژوهش‌های بعدی می‌توان با استفاده از تصاویر ماهواره‌ای به مدل‌سازی و پیش‌بینی خشک‌سالی در مقایسه با روش‌های هوش محاسباتی پرداخت. علاوه بر دو الگوریتم استفاده‌شده در این تحقیق، در کارهای آتی می‌توان به تلفیق الگوریتم‌های دیگری از خانواده هوش محاسباتی از جمله ANFIS، درخت تصمیم‌گیری، شبکه عصبی GMDH و شبکه‌های عصبی عمیق با موجک پرداخت. پیشنهاد می‌شود در کارهای آتی از تعدادی بیشتری ایستگاه سینوپتیک جهت ساخت سامانه پیش‌بینی خشک‌سالی در کشور ایران استفاده کرد. از آنجایی که همه مدل‌های هوش محاسباتی دارای پارامترهایی هستند که یافتن مقدار بهینه آن‌ها تأثیر بالایی در نتایج مدل‌سازی دارد، پیشنهاد می‌شود در پژوهش‌های بعدی به یافتن مقادیر بهینه این پارامترها با استفاده از الگوریتم‌های بهینه‌سازی پرداخت. از جمله الگوریتم‌های بهینه‌سازی می‌توان به الگوریتم ژنتیک، کلونی مورچگان، کلونی زنبورعسل، PSO و بقیه روش‌های ابتکاری و فرا ابتکاری اشاره کرد.

### ملاحظات اخلاقی پیروی از اصول اخلاق پژوهش

همکاری مشارکت‌کنندگان در تحقیق حاضر به صورت داوطلبانه و با رضایت آنان بوده است.

### حامی مالی

هزینه تحقیق حاضر توسط نویسندگان مقاله تامین شده است.

### مشارکت نویسندگان

تمامی نویسندگان در بخش‌های مختلف مقاله مشارکت داشتند.

### تعارض منافع

بنابر اظهار نویسندگان، مقاله حاضر فاقد هرگونه تعارض منافع بوده است.

### References

1. Abeyasingha NS, Rajapaksha URLN. 2020. SPI-Based Spatiotemporal Drought over Sri Lanka. *Advances in Meteorology*.
2. Azimi S, Azhdary Moghaddam M. 2020. Modeling Short Term Rainfall Forecast Using Neural Networks, and Gaussian Process Classification Based on the SPI Drought Index. *Water Resources Management*. 34(4):1369-405.
3. Bhunia P, Das P, Maiti R. 2020. Meteorological drought study through SPI in three drought prone districts of

منفرد از خود نشان دادند. این موضوع با نتیجه حاصل از پژوهش خان و همکاران (۱۰) که به پیش‌بینی خشک‌سالی با استفاده از دو مدل MLP و ARIMA به‌صورت منفرد و در ترکیب با موجک در منطقه حوضه رودخانه لانگات مالزی استفاده کردند و نشان دادند که در تمامی مقیاس‌های زمانی مدل ترکیبی دارای دقت بالاتری نسبت به مدل‌های منفرد است، مطابقت دارد. دقت بالاتر مدل‌های هیبریدی نسبت به مدل‌های منفرد با کار کماسی و شرقی (۱۲) نیز مطابقت دارد. اختلاف بین مدل‌های هیبریدی و منفرد در مقیاس زمانی سه و شش‌ماهه بیشتر از دیگر مقیاس‌های دیگر زمانی بود. این موضوع نیز با نتایج تحقیق تایلان و همکاران (۲۳) مطابقت دارد. میانگین تمامی تفاضل بین RMSE مدل‌های هیبریدی نسبت به منفرد در هر ده ایستگاه مورد مطالعه و در همه مقیاس‌های زمانی نشان داد که به میزان بهبود ناشی از استفاده از موجک در تلفیق با مدل SVR و GPR به ترتیب در مدل SVR با میانگین تفاضل RMSE برابر با  $-0/1540$  و تفاضل  $R^2$  برابر  $0/1491$  و در مدل GPR با میانگین تفاضل RMSE برابر با  $-0/1554$  و تفاضل  $R^2$  برابر  $0/1530$  نشان داد که مدل GPR در حالت کلی (همه مقیاس‌های زمانی و در تمامی ایستگاه‌های مورد مطالعه) بهبود بهتری در مدل هیبریدی نسبت به مدل منفرد داشته است و این بدین مفهوم است که استفاده از موجک تأثیر بیشتری بر الگوریتم GPR نسبت به SVR دارد. در این تحقیق در مجموع از ۴ روش جهت مدل‌سازی و پیش‌بینی خشک‌سالی در ۱۰ ایستگاه در اقلیم‌های مختلف استفاده شد که در تحقیقات قبلی به این صورت به آن پرداخته نشده است. نتیجه‌گیری کلی که می‌توان از پژوهش حاضر داشت این است که الگوریتم‌های هوش محاسباتی قابلیت بالایی در مدل‌سازی و پیش‌بینی پدیده خشک‌سالی دارند که با ترکیب این الگوریتم‌ها با موجک توانایی آن‌ها در مدل‌سازی و پیش‌بینی خشک‌سالی به میزان قابل‌توجهی افزایش می‌یابد. علاوه بر این می‌توان نتیجه گرفت که در مقیاس‌های زمانی پایین‌تر شاخص خشک‌سالی SPI، مانند مقیاس‌های ۳ و ۶ ماهه که سری زمانی پیچیده‌تری نسبت به مقیاس‌های زمانی بالاتر دارند، میزان بهبود دقت مدل‌های هیبریدی نسبت به مدل‌های منفرد بیشتر است که نشان‌دهنده اهمیت بالای استفاده از موجک در

West Bengal, India. *Earth Systems and Environment*. 4(1):43-55.

4. Raghavendra S, Deka PC. 2014. Support vector machine applications in the field of hydrology: A review. *Applied soft computing*. 19:372-86.
5. Djerbouai S, Souag-Gamane D. 2016. Drought forecasting using neural networks, wavelet neural networks, and stochastic models: case of the Algerois Basin in North Algeria. *Water Resources Management*. 30(7):2445-64.
6. Drisya J, Kumar DS, Roshni T. 2021. Hydrological drought assessment through streamflow forecasting using

- wavelet enabled artificial neural networks. *Environment, Development and Sustainability*. 23(3):3653–72.
7. Ghasemi P, Karbasi M, Nouri AZ, Tabrizi MS, Azamathulla HM. 2021. Application of Gaussian process regression to forecast multi-step ahead SPEI drought index. *Alexandria Engineering Journal*. 60(6):5375–92.
  8. Tarebari H, Javid AH, Mirbagheri SA, Fahmi H. 2018. Multi-objective surface water resource management considering conflict resolution and utility function optimization. *Water Resources Management*. 32:4487–509.
  9. Hong X, Ding Y, Ren L, Chen L, Huang B. 2018. A weighted heteroscedastic Gaussian Process Modelling via particle swarm optimization. *Chemometrics and Intelligent Laboratory Systems*. 172:129–38.
  10. Khan MMH, Muhammad NS, El-Shafie A. 2020. Wavelet based hybrid ANN-ARIMA models for meteorological drought forecasting. *Journal of Hydrology*. 590:125380.
  11. KIŞI Ö. 2006. Generalized regression neural networks for evapotranspiration modelling. *Hydrological Sciences Journal*. 51(6):1092–105.
  12. Komasi M, Sharghi S. 2020. Drought Forecasting Using Wavelet-Support Vector Machine and Standardized Precipitation Index (Case Study: Urmia Lake-Iran). *Journal of Environmental Science and Technology*. 22(7):83–101.
  13. Lazri M, Ameer S, Brucker JM, Lahdir M, Sehad M. 2015. Analysis of drought areas in northern Algeria using Markov chains. *Journal of Earth System Science*. 124:61–70.
  14. Li G, Liu Z, Li J, Fang Y, Liu T, Mei Y, et al. 2018. Application of general regression neural network to model a novel integrated fluidized bed gasifier. *International Journal of Hydrogen Energy*. 43(11):5512–21.
  15. Li L, She D, Zheng H, Lin P, Yang Z-L. 2020. Elucidating diverse drought characteristics from two meteorological drought indices (SPI and SPEI) in China. *Journal of Hydrometeorology*. 21(7):1513–30.
  16. McKee TB, Doesken NJ, Kleist J. 1993. The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference on Applied Climatology*. p. 179–83.
  17. Mehdizadeh S, Ahmadi F, Mehr AD, Safari MJS. 2020. Drought modeling using classic time series and hybrid wavelet-gene expression programming models. *Journal of Hydrology*. 587:125017.
  18. Neal RM. 1997. Monte Carlo implementation of Gaussian process models for Bayesian regression and classification. *arXiv preprint physics /9701026*.
  19. Pal M, Deswal S. 2010. Modelling pile capacity using Gaussian process regression. *Computers and Geotechnics*. 37(7–8):942–7.
  20. Paulo AA, Pereira LS. 2007. Prediction of SPI drought class transitions using Markov chains. *Water resources management*. 21:1813–27.
  21. Pei Z, Fang S, Wang L, Yang W. 2020. Comparative analysis of drought indicated by the SPI and SPEI at various timescales in inner Mongolia, China. *Water*. 12(7):1925.
  22. Tay FEH, Cao L. 2001. Application of support vector machines in financial time series forecasting. *Omega*. 29(4):309–17.
  23. Taylan ED, Terzi Ö, Baykal T. 2021. Hybrid wavelet-artificial intelligence models in meteorological drought estimation. *Journal of Earth System Science*. 130:1–13.
  24. Won J, Choi J, Lee O, Kim S. 2020. Copula-based Joint Drought Index using SPI and

- EDDI and its application to climate change. *Science of the Total Environment*. 744:140701.
25. Xu D, Zhang Q, Ding Y, Huang H. 2020. Application of a hybrid ARIMA-SVR model based on the SPI for the forecast of drought—a case study in Henan Province, China. *Journal of Applied Meteorology and Climatology*. 59(7):1239-59.
26. Yang D, Zhang X, Pan R, Wang Y, Chen Z. 2018. A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve. *Journal of Power Sources*. 384:387-95.
27. Zhang C, Wei H, Zhao X, Liu T, Zhang K. 2016. A Gaussian process regression based hybrid approach for short-term wind speed prediction. *Energy conversion and management*. 126:1084-92.

