



Monthly precipitation prediction improving using the integrated model based on kernel-wavelet and complementary ensemble empirical mode decomposition

Kiyoumars Roushangar^{1,*}, Roghayeh Ghasempour²

¹ Professor, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

² PhD Candidate, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

ABSTRACT: Estimates of monthly rainfall are important for various purposes such as flood estimation, drought, irrigation planning, and river basin management. In the present study, the monthly rainfall of Tabriz station was investigated using the intelligent Gaussian Process Regression (GPR) method based on Complementary Ensemble Empirical Mode Decomposition (CEEMD) and Wavelet Transform (WT). Different models were defined based on teleconnection patterns and climatic elements, and the impact of different input parameters was assessed. The obtained results proved high capability and efficiency of the applied method in predicting the monthly precipitation. The results showed that time series decomposition based on wavelet transformation led to more accurate outcomes compared to the complementary ensemble empirical mode decomposition. The best evaluation of test series using wavelet transform decomposition was obtained for the state of modeling based on teleconnection patterns and climatic elements with the values of $DC=0.889$, $R=0.961$ and $RMSE=0.036$. Also, based on the sensitivity analysis, Pt-3 was found to be the most effective parameter in modeling.

Review History:

Received: 2019-03-23

Revised: 2019-04-29

Accepted: 2019-09-17

Available Online: 2019-09-29

Keywords:

Climatic elements

Empirical mode

GPR, Precipitation

Wavelet Transform

1. INTRODUCTION

Precipitation plays an important role in determining the climate of a region. Precise estimation of precipitation is required to manage and plan water resources, as well as other related applications such as hydrology, climatology, meteorology and agriculture. So far, numerous precipitation prediction methods have been proposed in literature, including time series models, regression models, probabilistic models, machine learning models, physical models, and a host of hybrid models such as Tofani et al [1] and Soltani et al., [2]. In recent years, the Meta model approaches such as Artificial Neural Networks (ANNs), Neuro-Fuzzy models (NF), Genetic Programming (GP), Support Vector Machine (SVM), and Gaussian Process Regression (GPR) have been applied in investigating the hydraulic and hydrologic complex phenomena [3]. Hybrid models involving signal decomposition have also been shown to be effective in improving prediction accuracy of time series prediction methods, as indicated in [4]. Complementary Ensemble Empirical Mode Decomposition analysis is one of the widely used signal decomposition methods for hydrological time series prediction. Decomposition of time series reduces the difficulty of forecasting, thereby improving forecasting accuracy. Due to the complexity of the precipitation phenomenon and the effect of various parameters on its prediction, in this study, the capability of GPR as a kernel based approach and also integrated CEEMD-GPR and DWT-

GPR models were assessed for precipitation modeling in Tabriz station during the period time of 1978-2017. Different input combinations were considered using climatic data such as precipitation, monthly temperature and relative humidity with one to four time lag and telecommunication patterns. Considering two different types of modeling based on original time series data (without data decomposition) and time series decomposition data, the capability of used methods was investigated. Also, sensitivity analysis was performed to determine the effective parameters in monthly precipitation prediction process.

2. METHODOLOGY

In the current study, monthly rainfall data of the Tabriz station was used during the period of 1978-2017. GPR model are based on the assumption that adjacent observations should convey information about each other. Gaussian processes are a way of specifying a prior directly over function space. This is a natural generalization of the Gaussian distribution, whose mean and covariance are a vector and matrix, respectively. The Gaussian distribution is over vectors, whereas the Gaussian process is over functions. Thus, due to prior knowledge about the data and functional dependencies, no validation process is required for generalization, and GP regression models are able to understand the predictive distribution corresponding to the test input. A GP is defined as a collection of random variables, any finite number of which has a joint multivariate Gaussian distribution. The WT is a popular method and a very

*Corresponding author's email: kroschangar@yahoo.com



Table 1. Statistical parameters results for Test series

| Model | Performance criteria | | | | | | | | | | |
|---|----------------------|-------|------------|--|-------|-------|----------|--|-------|-------|-------|
| | R | DC | RMSE | | R | DC | RMSE | | R | DC | RMSE |
| Based on precipitation data | | | | | | | | | | | |
| GPR | | | CEEMD- GPR | | | | DWT- GPR | | | | |
| P(I) | 0.378 | 0.219 | 0.236 | | 0.623 | 0.484 | 0.119 | | 0.725 | 0.525 | 0.099 |
| P(II) | 0.536 | 0.333 | 0.222 | | 0.766 | 0.548 | 0.090 | | 0.906 | 0.816 | 0.057 |
| P(III) | 0.605 | 0.508 | 0.193 | | 0.822 | 0.721 | 0.081 | | 0.914 | 0.835 | 0.054 |
| P(IV) | 0.587 | 0.420 | 0.195 | | 0.805 | 0.699 | 0.086 | | 0.913 | 0.828 | 0.062 |
| Based on precipitation and teleconnection patterns data | | | | | | | | | | | |
| CEEMD- GPR | | | DWT- GPR | | | | | | | | |
| I(I) | 0.638 | 0.508 | 0.106 | | 0.708 | 0.588 | 0.091 | | | | |
| I(II) | 0.834 | 0.735 | 0.069 | | 0.926 | 0.851 | 0.046 | | | | |
| I((III) | 0.845 | 0.742 | 0.064 | | 0.938 | 0.859 | 0.042 | | | | |
| I(IV) | 0.822 | 0.726 | 0.075 | | 0.912 | 0.841 | 0.050 | | | | |
| I(V) | 0.818 | 0.724 | 0.080 | | 0.908 | 0.838 | 0.053 | | | | |
| I(VI) | 0.827 | 0.732 | 0.072 | | 0.918 | 0.848 | 0.048 | | | | |
| I(VII) | 0.866 | 0.768 | 0.054 | | 0.961 | 0.889 | 0.036 | | | | |
| I(VIII) | 0.471 | 0.351 | 0.250 | | 0.523 | 0.407 | 0.166 | | | | |

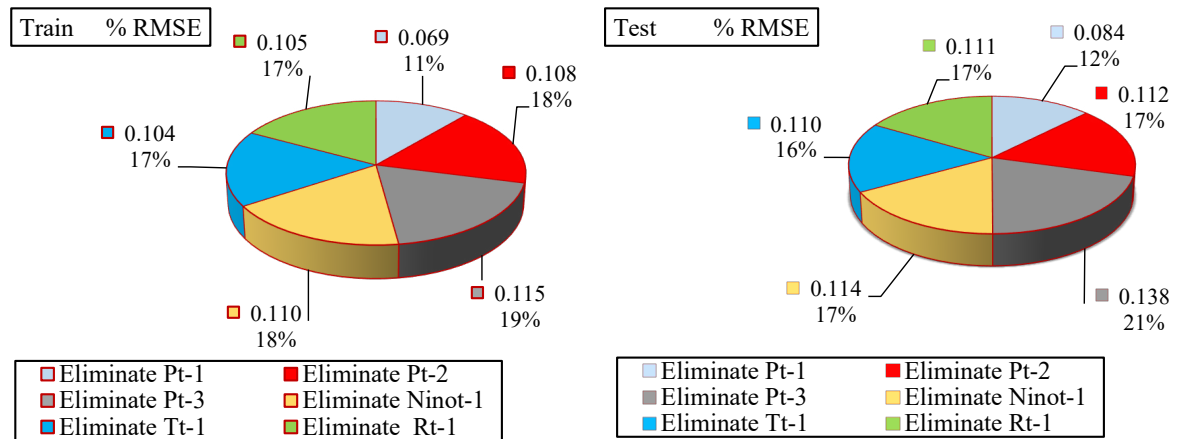


Fig. 1. Relative significance of each of input parameters of the best model

precise method for time series processing. While the general theory behind WT is quite analogous to that of the short time Fourier transform, WT allows for a completely flexible window function (called the mother wavelet), which can be changed over time based on the shape and compactness of the signal. Given this property, WT can be used to analyze the time-frequency characteristics of any kind of time series. CEEMD was proposed to solve the mode mixing issue of empirical mode decomposition (EMD) which specifies the true IMF as

the mean of an ensemble of trials [5]. Each trial consists of the decomposition results of the signal plus a white noise of finite amplitude. EMD can be used to decompose any complex signal into finite intrinsic mode functions and a residue, resulting in subtasks with simpler frequency components and stronger correlations that are easier to analyze and forecast. Another important feature of empirical mode decomposition is that it can be used for noise reduction of noisy time series, which can be effective in improving the accuracy of model predictions.

3. RESULTS AND DISCUSSION

In order to evaluate and review the performance of the developed models and determine the accuracy of the selected models, three performance criteria named Correlation Coefficient (R), Determination Coefficient (DC), and Root Mean Square Errors (RSME) were used according to Table 1. According to the results, it could be seen that the accuracy of single GPR model did not yield to accurate prediction. However, both integrated WT-GPR and EEMD-GPR models led to higher accuracy than the GPR model. The use of these two methods decreased the error criteria approximately between 25 to 35%. It was observed that in prediction of precipitation, climatic elements including monthly average temperature, relative humidity, and previous months precipitation as well as teleconnection patterns were effective in prediction process.

According to the Fig. 1, sensitivity analysis was performed to determine the most significant parameters in modeling process. It was observed that P_{t-3} is the most effective parameter in precipitation modeling.

4. CONCLUSIONS

The comparison of the developed models' accuracy revealed that both integrated CEEMD-GPR and DWT-GPR

models had better performance compared with the GPR model in predicting the monthly precipitation. The use of these two methods decreased the error criteria 25 to 35%. Also, based on the sensitivity analysis, P_{t-3} was found to be the most effective parameter in modeling process.

REFERENCES

- [1] P. Tofani, E. Mosaedi, A. Fakheri Fard, Precipitation forecast using wavelet theory, *Water and Soil Journal (Agriculture Sciences and Technology)*, 25(5) (2011) 1217-1226. [in Persian]
- [2] A. S. Soltani, A. Saberi, M. Gheisouri, Determination of the best time series model for forecasting annual rainfall of selected stations of Western Azerbaijan province, *Researches in Geographical Sciences*, 17(44) (2017) 87-105.
- [3] K. Roushangar, R. Ghasempour, The study of the performance of classical and artificial intelligence methods in the estimation of roughness coefficients in pontoons, *Irrigation and Drainage Journal of Iran*, 12(4) (2019) 811-822. [in Persian]
- [4] K.M. Lau, H.Y. Weng, Climate signal detection using wavelet transform, *How to make time series sing*, *Bull Am Meteorol Soc*, 76 (1995) 2391-2402.
- [5] Y. Amirat, M. Benbouzidb, T. Wang, K. Bacha, G. Feld, EEMD-based notch filter for induction machine bearing faults detection, *Applied Acoustics*, 133 (2018) 202-209.

HOW TO CITE THIS ARTICLE

K. Roushangar, R. Ghasempour, *Monthly precipitation prediction improving using the integrated model based on kernel-wavelet and complementary ensemble empirical mode decomposition*, *Amirkabir J. Civil Eng.*, 52(10) (2021) 649-652.

DOI: [10.22060/ceej.2019.16043.6109](https://doi.org/10.22060/ceej.2019.16043.6109)





بهبود پیش بینی بارش ماهانه با استفاده از مدل تلفیقی بر پایه روش کرنل - تبدیل موجک و تجزیه ی یکپارچه مد تجربی کامل

کیومرث روشنگر^{۱*}، رقیه قاسم پور^۲

^۱ دانشیار گروه آب دانشکده عمران دانشگاه تبریز

^۲ دانشجوی دکتری عمران مهندسی و مدیریت منابع آب دانشگاه تبریز

تاریخچه داوری:

دریافت: ۱۳۹۸-۰۱-۰۳

بازنگری: ۱۳۹۸-۰۲-۰۹

پذیرش: ۱۳۹۸-۰۶-۲۶

ارائه آنلاین: ۱۳۹۸-۰۷-۰۷

کلمات کلیدی:

بارش

تبدیل موجک

عناصر اقلیمی

مد تجربی

GPR

خلاصه: بارش یکی از مهم ترین اجزای چرخه آب بوده و در سنجش خصوصیات اقلیمی هر منطقه، نقش بسیار مهمی ایفا می کند. تخمین مقادیر بارش ماهانه برای اهداف مختلفی چون برآورد سیلاب، خشکسالی، برنامه ریزی آبیاری و مدیریت حوضه های آبریز اهمیت زیادی دارد. در تحقیق حاضر، پیش بینی بارش ماهانه ایستگاه تبریز با استفاده از روش هوشمند رگرسیون فرآیند گاوسی (GPR) بر پایه روش تجزیه ی یکپارچه مد تجربی کامل (CEEMD) و تبدیل موجک (WT) مورد بررسی قرار گرفته است. در این راستا، مدل های متفاوتی بر اساس شاخص های پیوند از دور و عناصر اقلیمی تعریف شد و نرخ تأثیر پارامترهای ورودی مختلف مورد بررسی قرار گرفت. نتایج حاصل از تحلیل مدل ها قابلیت و کارایی بالای روش به کار رفته را در تخمین میزان بارش ماهانه به خوبی نشان داد. ملاحظه گردید که در پیش بینی بارش ماهانه، شاخص های پیوند از دور NAO, Nino3, MEI و عناصر اقلیمی شامل میانگین دمای ماهانه و رطوبت نسبی و همچنین بارش مربوط به ماه های گذشته در پیش بینی مقادیر بارش تأثیرگذار بوده و موجب بهبود نتایج مدل ها می گردد. در بررسی روش های تجزیه ی یکپارچه مد تجربی کامل و تبدیل موجک گسسته مشاهده گردید که تجزیه بر اساس تبدیل موجک منجر به نتایج دقیق تری می گردد. بهترین حالت ارزیابی برای داده های آزمون با استفاده از تجزیه ی تبدیل موجک در حالت مدلسازی بر اساس داده های اقلیمی و عناصر پیوند از دور با مقادیر $DC=0/889$ ، $R=0/961$ و $RMSE=0/036$ بدست آمد. همچنین بر اساس نتایج آنالیز حساسیت مشخص گردید که P_{t3} تأثیرگذارترین پارامتر در مدل سازی است.

۱- مقدمه

مشخص، اقلیم و بارش پدیده های غیر خطی در طبیعت هستند. امروزه محققین، با ابداع و پیشرفت علمی چون روش های هوشمند که ابزاری توانمند، انعطاف پذیر و مستقل از مدل های دینامیکی سیستم هستند در جستجوی راه هایی برای پیشرفت و پیش بینی پارامترهای مهم هواشناسی هستند. در این بین سری زمانی به مجموعه ای از دیده بانها با مقادیر ثبت شده از یک متغیر گفته می شود که بر حسب زمان مرتب شده است. هدف از سری زمانی، تعیین قانونمندی و شناسایی رفتار آن جهت پیش بینی آینده است. پیش بینی بارش به ویژه در نواحی مختلف سرزمین کم آبی نظیر ایران که منابع آب آن متکی به بارش و با تقاضای روزافزون جمعیت روبه رشد همراه

بررسی پیش بینی و روند بارش یکی از اقدامات مفید و مؤثر جهت برنامه ریزی های کلان و خرد اقتصادی- اجتماعی است. در صورتی که برآوردهای قابل اعتمادی از وضعیت اقلیمی چند ماه یا چند سال آینده در دسترس باشد، برنامه ریزان، مدیران و اقشار مختلف اجتماعی می توانند ضمن آمادگی جهت مقابله با حوادث نامطلوب، نسبت به بهره برداری بهینه از منابع آب و خاک و نیروی انسانی اقدام نمایند [۱]. در این راستا پیش بینی بارش برای کشورهای در حال توسعه که اقتصاد آن ها مبتنی بر کشاورزی است، خیلی مهم است. به طور

* نویسنده عهده دار مکاتبات: kroshangar@yahoo.com

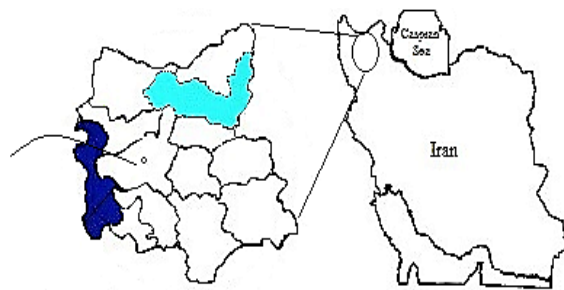
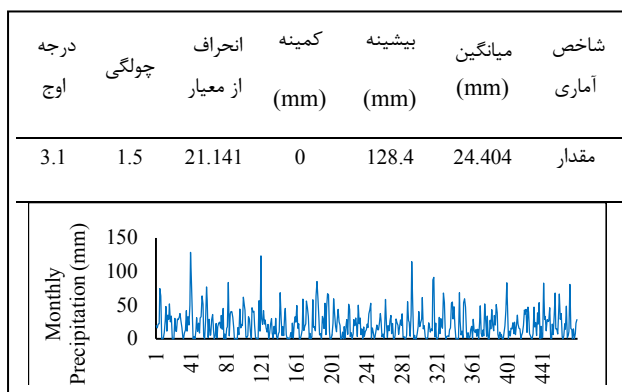


است، از اهمیت ویژه‌های برخوردار است. در این راستا، برنامه‌ریزی‌های صحیح و کسب اطلاعات از طریق انجام مطالعات علمی و پژوهش‌های دقیق ضروری است. طوفانی و همکاران [۲] با استفاده از تبدیل موجک (WT) به پیش بینی بارندگی ماهانه در ایستگاه زرينگل استان گلستان پرداختند و نتایج حاصل از آن را با مدل سری زمانی خطی فصلی (SARIMA) مقایسه کردند. نتایج نشان داد که تجزیه سیگنال با موجک، منجر به ساده تر شدن سیگنال بارندگی می شود. به طوریکه همبستگی میان داده های مشاهداتی و محاسباتی ۸۴ درصد به دست آمد و پیش بینی سیگنال بارندگی با دقت بیشتری صورت گرفت. شریفان و قهرمان [۳] به ارزیابی پیش بینی بارش با مدل SARIMA در استان گلستان پرداختند. هدف از انجام این پژوهش بررسی مناسب ترین مدل برآورد بارش بود. با مقایسه مقادیر برآوردی باران ۱۰ روزه، ماهانه و سالانه با مقادیر واقعی متناظر مشخص شد برای پیش بینی بارش فقط در مقیاس ۱۰ روزه استفاده از باران های ۱۰ روزه از دقت بیشتری برخوردار هستند. گلابی و همکاران [۴] بارندگی سه ماهانه سه ایستگاه اهواز، آبادان و دزفول را در دوره آماری ۴۸ ساله (۱۳۴۰-۱۳۸۷) با استفاده از مدل های خطی سری زمانی مورد بررسی و مدل سازی قرار دادند. نتایج تحقیق نشان داد که مدل های آریما از دقت قابل قبولی در پیش بینی بارش برخوردار هستند. سلطانی و همکاران [۵] بهترین مدل سری زمانی در پیش بینی بارندگی سالانه ایستگاه های منتخب استان آذربایجان غربی را بررسی کردند و به این نتیجه رسیدند که مدل (۱,۰,۰) ARIMA، ARIMA(۰,۱,۱) و مدل (۰,۱,۱) ARIMA به ترتیب در ایستگاه های ارومیه، ماکو و مهاباد به عنوان مدلی مناسب جهت پیش بینی بارندگی سالانه هستند.

با وجود اینکه مدل های هیدرولوژیکی مختلفی جهت مدل سازی بارندگی پیشنهاد شده اند، اما با توجه به ماهیت غیرخطی و ویژگی های متغیر زمانی و مکانی در سیستم گردش آبی، هیچ یک از مدل های آماری و مفهومی پیشنهاد شده به منظور الگوسازی دقیق بارش نتوانسته اند به عنوان یک مدل برتر و توانا شناخته شوند و عدم قطعیت در پیش بینی ها از اعتبار برآوردها از طریق مدل های مختلف می کاهد. امروزه شبکه های غیرخطی به عنوان یکی از سیستم های هوشمند در پیش بینی چنین پدیده های پیچیده بسیار مورد استفاده

قرار می گیرند. این روش ها که اکثراً الهام گرفته از پدیده های موجود در طبیعت هستند و می توانند برای پیش بینی هرگونه متغیری که در آن رابطه بین متغیرهای مربوطه به خوبی درک نمی شود، پیدا کردن اندازه و شکل راه حل نهایی دشوار است، و روش های تحلیلی ریاضی رایج قادر به حل آن ها نبوده و یا بسیار زمان بر هستند، استفاده شوند (Banzhaf et al., ۱۹۹۸)، به عنوان ابزاری توانمند در حل مسائل پیچیده به شمار می آیند و معمولاً دارای جواب بهتری نسبت به روش های کلاسیک هستند [۶]. به عنوان نمونه سیواپراگاسم و لیونگ [۷] از مدل SVM برای پیش بینی بارش - رواناب حوضه تریجوآلد^۲ واقع در کشور هلند استفاده کردند. روشنگر و قاسم پور [۸] با استفاده از روش های هوشمند رگرسیون بردار پشتیبان و سیستم استنتاج عصبی- فازی انطباقی مقاومت جریان را در رودخانه های ماریچی بررسی کردند. کیشی و سیمن [۹] با استفاده از روش ماشین بردار پشتیبان به پیش بینی بارش روزانه پرداختند. مارزانو و همکاران [۱۰] به بررسی کارایی روش شبکه عصبی در مقایسه با روش های رگرسیون در پیش بینی شدت بارش پرداختند. رزاق زاده و همکاران [۱۱] به شبیه سازی بارش با استفاده از ریزمقیاس نمایی آماری مبتنی بر ترکیب مدل های شبکه عصبی مصنوعی و نروفازی پرداختند. کومار و همکاران [۱۲] در مقایسه روش های سری زمانی و شبکه عصبی مصنوعی، روش شبکه عصبی مصنوعی را با توجه به سری زمانی پیچیده بارش های هند مناسبتر دانستند. نایاک و همکاران [۱۳] با مروری بر روش های پیش بینی بارش در هند، شبکه عصبی مصنوعی را در مقایسه با روش های پیش بینی عددی مناسب تر معرفی کردند.

بررسی منابع صورت گرفته، نشان از کاربرد بالای روش های هوشمند مانند شبکه عصبی مصنوعی در زمینه پیش بینی بارش است. این مدل ها به طور معناداری، خطاهای کوچکتری را در پیش بینی نسبت به مدل های آماری نشان می دهند و دارای سرعت پیش بینی بالاتری هستند. با این حال، مطالعات چندانی در زمینه کاربرد روش های مبتنی بر کرنل مانند رگرسیون فرآیند گاوسی در تخمین میزان بارش ماهانه انجام نشده است. از طرفی، در سال های اخیر استفاده از روش های تجزیه سیگنال به عنوان روشی نوین و بسیار موثر در زمینه تحلیل سیگنال ها و سری های زمانی مورد توجه



شکل ۱. موقعیت منطقه مطالعاتی و متوسط بارندگی ماهانه در ایستگاه سینوپتیک تبریز (۱۹۷۸-۲۰۱۷)

Fig. 1. Location of the study area and monthly average precipitation in Tabriz synoptic station (1978-2017)

۴۶ درجه و ۲۵ دقیقه طول شرقی و ۳۸ درجه و ۲ دقیقه عرض شمالی از نصف‌النهار گرینویچ واقع شده است و وسعتی معادل ۱۷۸۱ کیلومترمربع دارد. ارتفاع تقریبی آن از سطح دریا ۱۳۰۰ تا ۲۱۰۰ متر متغیر است. شهرستان تبریز از دو ناحیه کوهستانی و دشت تشکیل شده است. ارتفاع از سطح دریا در دشت ۱۳۱۰ و در مناطق کوهستانی تا ۲۱۰۰ متر می‌رسد. در شکل ۱ منطقه مطالعاتی و اطلاعات آماری سری زمانی بارندگی ماهانه ایستگاه سنوپتیک تبریز در بازه زمانی ۱۹۷۸-۲۰۱۷ نمایش داده شده است.

۲-۲- تبدیل موجک

موجک (Wavelet) دسته‌ای از توابع ریاضی است که برای تجزیه سیگنال پیوسته به مؤلفه‌های فرکانسی آن بکار می‌رود که سطح تفکیک هر مؤلفه برابر با مقیاس آن است. تبدیل موجک تجزیه یک تابع بر مبنای توابع موجک است. موجک‌ها (که به عنوان موجک‌های مادر شناخته می‌شوند) نمونه‌های انتقال یافته و مقیاس شده یک تابع (موجک مادر) با طول متناهی و نوسانی شدیداً میرا هستند. تابع موجک، تابعی است که دو ویژگی مهم نوسانی بودن و کوتاه مدت بودن را دارد. $\Psi(x)$ تابعی موجک است اگر و فقط اگر تبدیل فوریه آن $\Psi(\omega)$ شرط زیر را ارضا کند [۱۴].

$$\int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|}{|\omega|^2} d\omega < +\infty \quad (1)$$

این شرط با عنوان شرط پذیرفتگی برای موجک $\Psi(x)$ شناخته می‌شود $\Psi(x)$ تابع موجک مادر است (رابطه ۲) که ضرایب a و b در

قرار گرفته است. در این روش سیگنال اصلی به چندین زیرسری تجزیه شده و داده‌های حاصل از تجزیه با معادلات مناسب برازش می‌یابد و مدل بدست‌آمده برای پیش‌بینی استفاده می‌شود. از جمله این روش‌ها روش جدید تجزیه مد تجزیه یکپارچه کامل است که در روند تجزیه‌سازی آن، حجم محدودی از نویز سفید وارد سیگنال اصلی شده و با استفاده از جنبه‌های مثبت آماری نویز سفید که توزیع متوازن در حوزه فرکانس دارد، اثر نویز متناوب از روند تجزیه حذف می‌شود. بنابراین، با توجه به اهمیت دقت پیش‌بینی بارش در هیدرولوژی و پیش‌بینی سیلاب، در تحقیق کنونی کارایی روش رگرسیون فرآیند گاوسی (GRP) به عنوان یکی از روش‌های نوین داده‌کاوی در تخمین بارش ماهانه ایستگاه سنوپتیک تبریز مورد بررسی قرار گرفت. مدل‌های ورودی متفاوتی با استفاده از داده‌های مربوط به عناصر اقلیمی مانند بارش، میانگین دما و رطوبت نسبی ماهانه با تاخیر زمانی یک تا چهار ماهه و الگوهای پیوند از دور در بازه زمانی (۱۹۷۸-۲۰۱۷) تعریف گردیده و با در نظر گرفتن دو نوع مدل‌سازی متفاوت براساس داده‌های سری زمانی اصلی (بدون تجزیه داده‌ها) و داده‌های سری زمانی اصلی تجزیه شده با استفاده از تبدیل موجک گسسته و روش جدید تجزیه یکپارچه مد تجزیه کامل مورد بررسی قرار گرفت. همچنین با استفاده از آنالیز حساسیت موثرین پارامترها در بارش ماهانه تعیین شد.

۲- مواد و روش‌ها

۲-۱- موقعیت و ویژگی‌های منطقه مورد مطالعه

منطقه‌ی مورد مطالعه شهر تبریز است که با مختصات جغرافیای

آن نقش انتقال و مقیاس سیگنال مورد نظر را دارند [۱۴].

$$\varphi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (2)$$

توابع موجک دارای انواع بسیاری هستند که مهم ترین و پرکاربردترین آن ها شامل تابع موجک Mexican, Meyer, Haar, Symlet, Daubechies, Coiflet است. تبدیل موجک پیوسته به عنوان روش جایگزین تبدیل فوریه زمان-کوتاه ارائه گردید و هدف آن، فائق آمدن بر مشکلات مربوط به تفکیک پذیری در تبدیل فوریه زمان-کوتاه است. در عمل اکثر سری های زمانی هیدرولوژیکی به صورت گسسته در زمان هایی مشخص اندازه گیری می شوند. بنابراین کاربرد موجک گسسته می تواند تناسب بهتری با این نوع سری زمانی داشته باشد. تبدیل موجک گسسته در عمل همان تبدیل موجک پیوسته است (نسخه گسسته شده آن) که در آن یک سری موجک از تبدیل موجک پیوسته نمونه گرفته است و سیستم های کامپیوتری قابلیت محاسبه آن را با هزینه زمانی کمتر دارا است. اصول تبدیل موجک گسسته به روشی تحت عنوان کدینگ زیرباند بر می گردد که در سال ۱۹۷۶ سنگ بنای اولیه آن گذارده شد و در آن نوعی توصیف زمان-مقیاس از سیگنال گسسته با استفاده از فیلترهای دیجیتال ارائه می گردد. تبدیل موجک، حاصل شباهت سنجی (کورولیشن) بین محتوای فرکانسی (مقیاسی) سیگنال و تابع موجک در مقیاس های مختلف است [۱۵]. تجزیه کردن سری های زمانی با استفاده از فیلترهای مشخص (موجک و توابع مقیاس گذاری شده) دو نوع زیرسری را به وجود می آورند؛ بردارهای تقریبی یا باقی مانده و بردارهای جزییات [۱۶]. این ضرایب حاصل از در هم پیچیدن سری زمانی اصلی با فیلتر پایین-گذر و فیلتر بالا-گذر می باشد. فیلتر پایین-گذر همان تابع مقیاس گذاری است و فیلتر بالاگذر همان تابع موجک است. تابع تبدیل موجک قابلیت تجزیه سری زمانی به چندین زیرسری زمانی با مقیاس های مختلف را دارد و با بررسی زیرسری های زمانی منتهی از سری زمانی کلی، رفتار کوچک مقیاس و بزرگ مقیاس یک فرآیند هیدرولوژیکی را آنالیز می کند. روابطی که بیشتر برای تخمین سطح تجزیه مورد استفاده قرار می گیرد به صورت زیر است که در آن L سطح تجزیه، N تعداد نقاط داده در سری زمانی و p برابر با زمان محو شونده برای موجک مادر انتخاب شده است [۱۶].

$$L = \text{int}[\log(N)] \quad , \quad L = \frac{\log\left(\frac{N}{2^{p-1}}\right)}{\log 2} \quad (3)$$

۳-۲- روش تجزیه مد تجربی EMD^۱

روش تجزیه مد تجربی روشی برای تجزیه سیگنال های گوناگون است که این عمل در فرآیندی تحت عنوان غربال کردن صورت می گیرد. طی این فرایند سیگنال اصلی به تعدادی مولفه با محتوای بسامدی متفاوت تجزیه می شود. روش EMD طبق رابطه (۴) سیگنال اصلی $x(n)$ را به تعدادی مد ذاتی (IMF^۲) تجزیه می کند [۱۷].

$$x(n) = \sum_{i=1}^n c_i(x) + r_n(x) \quad (4)$$

که $r_n(x)$ همان مولفه باقی مانده، بعد از n تعداد IMFs و $c_i(x)$ است. یک داده ممکن است در یک زمان دارای چندین مد ذاتی باشد. این مدهای نوسانی، تابع های مد ذاتی (IMF) نام دارند و دارای دو شرایط هستند: (۱) در کل داده ها، تعداد نقاط فرین (اکسترمم) و نقاط صفر با هم برابر و یا حداکثر دارای یک واحد اختلاف هستند، (۲) در هر نقطه میانگین پوش برازش داده شده بر نقاط بیشینه محلی و پوش برازش داده شده بر نقاط کمینه محلی باید صفر باشد. به علت وجود تناوب و نویز در سیگنال ها، در برخی موارد به علت اختلاط مدها؛ توزیع حوزه زمان-فرکانس دچار انقطاع می شود و عملکرد EMD دچار نقض می گردد (چون میانگین IMF ها حالتی نامشخص به خود می گیرد). جهت رفع این ایراد، وو و هوانگ [۱۸] روشی متفاوت با عنوان روش تجزیه مد تجربی یکپارچه (EEMD^۳) را پیشنهاد کردند. در روند تجزیه سازی آن، حجم محدودی از نویز سفید وارد سیگنال اصلی می شود. با استفاده از جنبه های مثبت آماری نویز سفید که توزیع متوازن در حوزه فرکانس دارد، اثر نویز متناوب از روند تجزیه حذف می شود. در روش تجزیه مد تجربی یکپارچه کامل (CEEMD^۴) نوفه سفید بصورت دوتایی (یکی مثبت و دیگری منفی) به داده اصلی اضافه می شود تا دو سری IMF مجموع ساخته شود. بنابراین ترکیبی متشکل از داده اصلی و نوفه ای اضافی داریم که مجموع IMF ها برابر سیگنال اصلی می شود

1 Empirical Mode Decomposition

2 Intrinsic Mode Functions

3 Ensemble Empirical Mode Decomposition

4 Complementary Ensemble Empirical Mode Decomposition

$$f(x) \approx GP(m(x), k(x, x')) \quad (7)$$

که معمولا جهت ساده سازی، مقدار تابع میانگین برابر با صفر در نظر گرفته می شود. در فرآیند گاوسی، رابطه بین بردار ورودی و هدف به فرم زیر است:

$$y_i = f(x_i) + \varepsilon \quad (8)$$

که در آن $f(x)$ بیانگر تابع رگرسیون دلخواه و ε نیز مقدار نویز توزیع گاوسی با میانگین صفر و واریانس σ^2 می باشد، یعنی $\varepsilon \sim N(0, \sigma^2)$. علاوه بر این، چنین فرض می گردد که داشته به نحوی که $p(f|X) = N(0, K)$ که در آن K ماتریس کواریانس با درایه های $k_{i,j} = k(x_i, x_j)$ است.

$$K(x, x) = \begin{pmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \dots & k(x_n, x_n) \end{pmatrix} \quad (9)$$

$k_{i,j}$ کواریانس بین مقادیر توابع نهان $f(x_i)$ و $f(x_j)$ می باشد. رگرسیون فرآیند گاوسی به منظور محاسبه توزیع پیش بینی شده برای مقادیر تابع f^* در نقاط تست $X^* = [x_1^*, x_2^*, \dots, x_m^*]$ به کار می رود.

۵-۲- توابع کرنل

یکی از روش های متداول برای حل مسائل غیرخطی استفاده از توابع کرنل است؛ این توابع بر اساس ضرب داخلی داده های مفروض تعریف می شود. طراحی روش های رگرسیون مبتنی بر فرآیند گاوسی

و نوفه ای اضافی دامنه ای موجی از مرتبه ای pm خواهد بود.

۴-۲- رگرسیون فرآیند گاوسی (GPR)

مجموعه داده S با n مشاهده را در نظر بگیرید $S = \{(x_i, y_i) | i=1, \dots, n\}$ که در آن x_i بردار ورودی با D بعد و y_i خروجی اسکالر یا هدف می باشد. این مجموعه متشکل از دو جز ورودی و خروجی به عنوان نقاط نمونه یا تجربی معرفی خواهند شد. به منظور سهولت کار، ورودی های مجموعه در ماتریس $X = [x_1, x_2, \dots, x_n]$ و خروجی ها نیز در ماتریس $Y = [y_1, y_2, \dots, y_n]$ تجمع می گردند. وظیفه رگرسیون، ایجاد یک ورودی جدید x^* به منظور دستیابی به توزیع پیش بینی شده برای مقادیر متناظر داده های مشاهداتی y^* و بر مبنای مجموعه داده S است. فرآیند گاوسی مجموعه ای از متغیرهای تصادفی است که تعداد محدودی از آنها با توزیع های گاوسی ادغام شده اند. فرآیند گاوسی تعمیمی از توزیع گاوسی است. توزیع گاوسی در واقع توزیع بین متغیرهای تصادفی بوده در حالی که فرآیند گاوسی بیانگر توزیع بین توابع است. فرآیند گاوسی $f(x)$ توسط توابع میانگین $m(x)$ و کواریانس به شکل زیر تعریف می شود [۱۹]:

$$m(x) = E(f(x)) \quad (5)$$

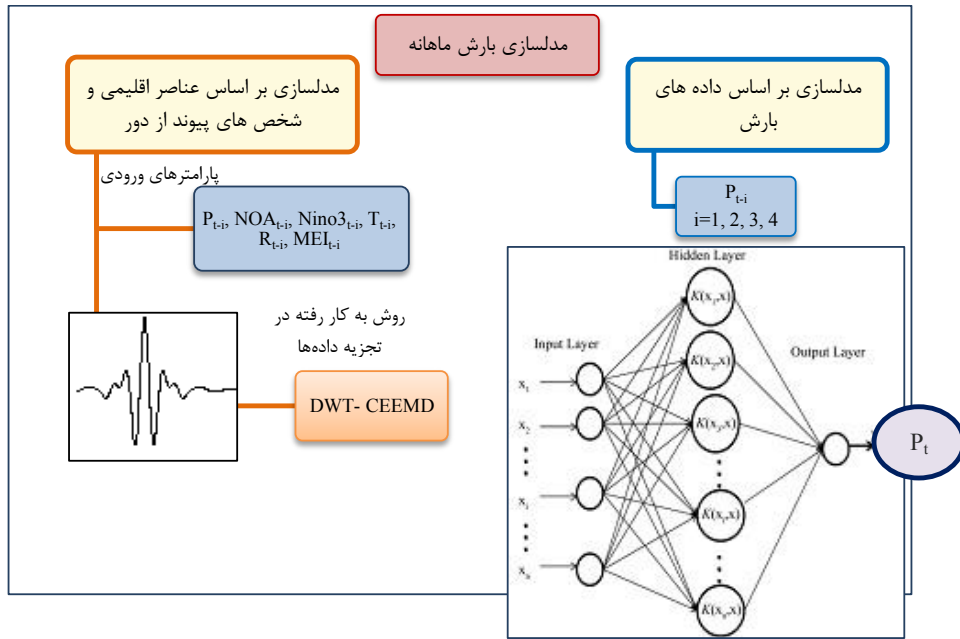
$$k(x, x') = E((f(x) - m(x))(f(x') - m(x'))) \quad (6)$$

که در روابط فوق، $k(x, x')$ تابع کواریانس (یا کرنل) بوده که در نقاط x و x' محاسبه می شود. فرآیند گاوسی $f(x)$ می تواند به صورت زیر بیان گردد:

جدول ۱. انواع تابع کرنل

Table 1. Types of kernel functions

| نوع کرنل | تابع کرنل | پارامتر کرنل |
|-----------------------|---|--------------|
| چند جمله ای ساده | $K(x_i, x_j) = (x_i^T x_j + 1)^d$ | d |
| چند جمله ای نرمال شده | $K(x_i, x_j) = (x_i^T x_j + 1)^d / \sqrt{(x_i^T x_i)(x_j^T x_j)}$ | d |
| تابع شعاع محور | $K(x_i, x_j) = \exp(-\frac{\ x_i - x_j\ ^2}{2\sigma^2})$ | γ |
| تابع کرنل پیرسون | $K(x_i, x_j) = \frac{1}{1 + 2\sqrt{\ x_i - x_j\ ^2} \sqrt{2^{1/w} - 1/\sigma}}^w$ | w, σ |



شکل ۲. شماتیک مدلسازی صورت گرفته در تحقیق

Fig. 2. The schematic view of considered modeling in the study

یک مدل به یک نزدیک تر و مقدار $RMSE$ کوچکتر باشد به معنی مطلوب بودن آن مدل است. روابط این پارامترهای آماری به صورت زیر است [۲۰]:

$$DC = 1 - \frac{\sum_{i=1}^N (I_o - I_p)^2}{\sum_{i=1}^N (I_o - \bar{I}_o)^2}, \quad R = \frac{\sum_{i=1}^N (I_o - \bar{I}_o) \times (I_p - \bar{I}_p)}{\sqrt{\sum_{i=1}^N (I_o - \bar{I}_o)^2 \times \sum_{i=1}^N (I_p - \bar{I}_p)^2}} \quad (10)$$

$$, RMSE = \sqrt{\sum_{i=1}^N \frac{(I_o - I_p)^2}{N}}$$

در روابط بالا L_m : مقدار اندازه گیری شده، \bar{L}_m : متوسط مقدار اندازه گیری شده، L_p : مقدار پیش‌بینی شده، \bar{L}_p : متوسط مقدار پیش‌بینی شده و N تعداد داده‌هاست. از آنجا که وارد کردن داده‌ها به صورت خام باعث کاهش سرعت و دقت شبکه می‌شود لذا نرمالیزه کردن داده‌ها کمک شایانی به آموزش بهتر و سریع‌تر مدل می‌کند. در تحقیق کنونی داده‌ها مطابق رابطه (۱۱) نرمالیزه شدند که در آن x_{max} ، x_{min} به ترتیب حداقل و حداکثر مقدار داده‌های مشاهده شده و x_n داده نرمال شده است [۲۰].

$$x_n = 0.05 + 0.95 \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right) \quad (11)$$

نیز شامل استفاده از مفهوم تابع کرنل است. در واقع، با یک تبدیل غیرخطی از فضای ورودی به فضای خنثی با ابعاد بیشتر (حتی نامتناهی) می‌توان مسائل را به صورت خطی تفکیک پذیر ساخت. با تبدیل نمونه‌ها از فضای ورودی به فضای ویژگی، تفکیک کننده غیرخطی به حالت خطی تبدیل خواهد شد. از مهمترین توابع کرنل می‌توان به کرنل چند جمله‌ای ساده، چند جمله‌ای نرمال شده، تابع شعاع محور و تابع کرنل پیرسون اشاره کرد. در ادامه روابط مربوط به این کرنل‌ها در جدول ۱ آورده شده است. مطابق با جدول ۱ توابع کرنل دارای پارامترهای مختلفی (γ, d, w, σ) است که نحوه تنظیم آن‌ها در حصول جواب هرچه بهتر مسئله بسیار موثر است.

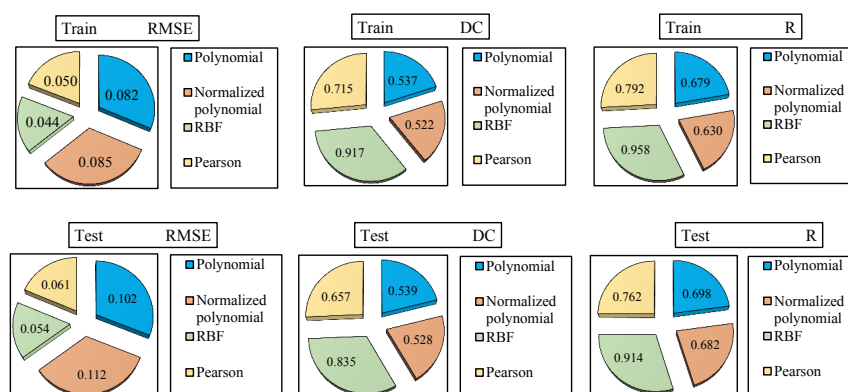
۲-۶- معیارهای ارزیابی

به منظور ارزیابی کارایی روش استفاده شده در تخمین میزان بارش ماهانه، از سه پارامتر آماری استفاده گردید که عبارتند از: ضریب همبستگی بین مقادیر مشاهداتی و محاسباتی (R)، ضریب تبیین (DC) که همبستگی خطی بین مقادیر پیش‌بینی شده و واقعی را نشان می‌دهد و یکی از معیارهای مهم ارزیابی قابلیت یک مدل است و ریشه میانگین مربعات خطاها ($RMSE$) که بیانگر میزان انحراف بین مقادیر پیش‌بینی شده است. هر چه مقدار R و DC برای

جدول ۲. مدل‌های تعریف شده در تحقیق

Table 2. Developed models in the study

| مدلسازی بر اساس بارش | | مدلسازی بر اساس عناصر اقلیمی و عناصر پیوند از دور | |
|----------------------|---|---|---|
| مدل | پارامترهای ورودی | مدل | پارامترهای ورودی |
| P(I) | $P_t=f(P_{t-1})$ | I(I) | $P_t=f(P_{t-1}, Nino3_{t-1}, T_{t-1})$ |
| P(II) | $P_t=f(P_{t-1}, P_{t-2})$ | I(II) | $P_t=f(P_{t-1}, P_{t-2}, P_{t-3}, T_{t-1})$ |
| P(III) | $P_t=f(P_{t-1}, P_{t-2}, P_{t-3})$ | I(III) | $P_t=f(P_{t-1}, P_{t-2}, P_{t-3}, T_{t-1}, R_{t-1})$ |
| P(IV) | $P_t=f(P_{t-1}, P_{t-2}, P_{t-3}, P_{t-4})$ | I(IV) | $P_t=f(P_{t-1}, P_{t-2}, P_{t-3}, NAO_{t-1})$ |
| | | I(V) | $P_t=f(P_{t-1}, P_{t-2}, P_{t-3}, MEI_{t-1})$ |
| | | I(VI) | $P_t=f(P_{t-1}, P_{t-2}, P_{t-3}, Nino3_{t-1})$ |
| | | I(VII) | $P_t=f(P_{t-1}, P_{t-2}, P_{t-3}, Nino3_t, T_{t-1}, R_{t-1})$ |
| | | I(VIII) | $P_t=f(NOA_{t-1}, Nino3_{t-1}, T_{t-1}, MEI_{t-1})$ |



شکل ۳. پارامترهای آماری مدل GPR با توابع کرنل متفاوت

Fig. 3. Statistical parameters of GPR model with different kernels function

اساسی را جهت بررسی سهم و درک نقش آن‌ها در سیستم اقلیم سیاره‌ای شناسایی نموده و در قالب شاخص‌هایی ارائه داده است [۲۱]. از انواع این شاخص‌ها می‌توان شاخص مد جنوبی اقیانوس اطلس (AMM)، آتلانتیک شرقی (EA)، نوسانات اطلس شمالی (NAO)، شاخص چند متغیره انسو (MEI)، شاخص‌های دمای سطح اقیانوس آرام (Nino) و شاخص بارش موسمی ناحیه جنوب غربی (SW monsoon) را نام برد. شاخص‌های پیوند از دور و عناصر اقلیمی مانند بارش (P)، دما (T) و رطوبت نسبی (R) در فرآیند پیش‌بینی می‌توانند به عنوان متغیرهای مستقل به کار روند؛ بنابراین به عنوان پیش‌بینی در مدل ترکیبی WT-GRP و CEEMD-GRP مورد استفاده قرار گرفتند. در تحقیق کنونی دو نوع مدل‌سازی در نظر گرفته شده است: مدل‌سازی بر اساس فقط

۳- نتایج و بحث

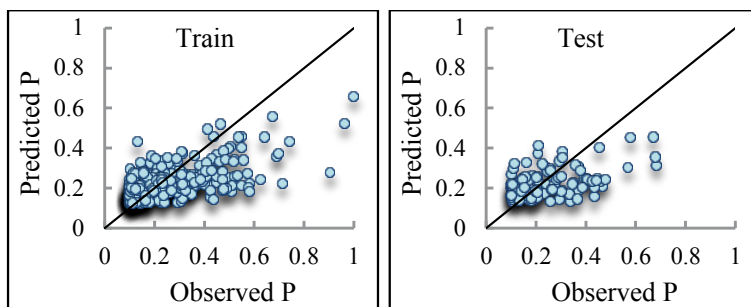
۳-۱- تعریف مدل‌های ورودی

با توجه به اینکه انتخاب مدل‌های ورودی در سامانه‌های هوشمند می‌تواند بر روی دقت جواب‌های حاصل از تحلیل تأثیر گذارد، لذا سعی گردید در مدل‌سازی پارامترهای مناسب و مؤثر در تعیین میزان بارش ماهانه انتخاب گردد. بررسی منابع متعدد نشان می‌دهد که ارتباط بین الگوهای پیوند از دور با بارش و دما در مناطق وسیعی از کره زمین به اثبات رسیده است، الگوهای پیوند از دور منعکس‌کننده‌ی تغییرات در مقیاس وسیع در امواج اتمسفری و موقعیت رودبادهای بوده و بر درجه حرارت، بارش، مسیر توفان‌ها و موقعیت و شدت رودباد بر روی مناطق وسیع تأثیر می‌گذارند. اداره ملی اقیانوس و اتمسفر ایالات متحده آمریکا الگوهای ارتباط از دور

جدول ۳. نتایج ارزیابی مدل‌های GPR بدون تجزیه داده‌ها

Table 3. Results of the GPR models evaluation without data decomposition

| مدل | آزمون | | | تست | | |
|--------|-------|-------|-------|-------|-------|-------|
| | R | DC | RMSE | R | DC | RMSE |
| P(I) | 0.422 | 0.322 | 0.232 | 0.378 | 0.219 | 0.236 |
| P(II) | 0.583 | 0.374 | 0.218 | 0.536 | 0.333 | 0.222 |
| P(III) | 0.612 | 0.522 | 0.177 | 0.605 | 0.508 | 0.193 |
| P(IV) | 0.609 | 0.444 | 0.183 | 0.587 | 0.420 | 0.195 |



شکل ۴. نتایج مرحله آزمون و تست مدل برتر بدون تجزیه داده‌ها

Fig. 4. Results of the train and test steps of superior model without data decomposition

۳-۳- نتایج به دست آمده برای مدل GPR بدون تجزیه داده‌ها جهت تخمین میزان بارش ماهانه بر اساس سری زمانی اصلی و بدون تجزیه آن به چندین زیرسری، چهار مدل بر اساس داده‌های ماهانه بارش مربوط به ماه‌های قبل تعریف گردید و کارایی روش رگرسیون فرایند گاوسی مورد بررسی قرار گرفت. بنابراین، این مدل‌ها با روش GPR آنالیز گردید و نتایج حاصله از تحلیل مدل‌ها در جدول ۳ و شکل ۴ نشان داده شد. با توجه به نتایج به دست آمده از سه معیار ارزیابی مدل برای داده‌های آموزش و آزمون، مشاهده می‌شود که مدل P(III) با پارامترهای ورودی P_{t-1} ، P_{t-2} ، P_{t-3} که مربوط به بارش ماهانه یک، دو و سه ماه قبل است، بیشترین کارایی را داراست. مطابق با نتایج ملاحظه می‌گردد که تنها استفاده از مقادیر بارش ماه قبل نمی‌تواند منجر به نتایج دقیقی گردد و همچنین مقادیر بارش چهار ماه قبل نیز در مدل سازی تاثیرگذار بوده و موجب بهبود نتایج شده است. نمودار مدل برتر برای سری داده‌های آموزش و آزمون در شکل ۴ نشان داده شده است.

۳-۴- نتایج به دست آمده برای مدل GPR بر پایه تجزیه ی یکپارچه مد تجرپی کامل و تبدیل موجک

در این قسمت با استفاده از دو روش تجزیه ی یکپارچه مد

داده‌های بارش و مدل سازی بر اساس عناصر اقلیمی و عناصر پیوند از دور. بنابراین، پس از نرمالیزه کردن داده‌ها ابتدا سری‌های زمانی توسط تبدیل موجک و تجزیه یکپارچه مد تجرپی کامل به زیر سری‌هایی تجزیه گردید و سپس این زیر سری‌ها به عنوان ورودی مدل GPR استفاده شد. در شکل ۲ شماتیک مدل سازی صورت گرفته در تحقیق نشان داده شده است. همچنین جدول ۲ بیان گر مدل‌های تعریف شده است. لازم به توضیح است که داده‌های سری زمانی بارش از سال ۱۹۷۸-۲۰۰۷ به عنوان داده‌های آموزش و بقیه داده‌ها (۲۰۰۸-۲۰۱۷) به عنوان داده‌های آزمون مورد استفاده قرار گرفتند.

۳-۲- انتخاب تابع کرنل مناسب GPR

جهت انتخاب تابع کرنل مناسب رگرسیون فرایند گاوسی مدل P(III) پس از تجزیه داده‌ها توسط تبدیل موجک با کرنل‌های مختلف مورد ارزیابی قرار گرفت. از سه معیار ارزیابی R ، $RMSE$ و DC جهت تعیین کارایی هر یک از کرنل‌ها استفاده گردید. مطابق شکل ۳، نتایج نشان داد که تابع کرنل شعاع مداری در تخمین میزان بارش ماهانه نسبت به توابع کرنل دیگر دارای ضریب همبستگی و تبیین بیشتر و خطای کمتری است. بنابراین در تجزیه و تحلیل تمامی مدل‌ها از کرنل شعاع مداری استفاده شد.

جدول ۴. نتایج ارزیابی مدل‌ها بر اساس تجزیه سریهای زمانی

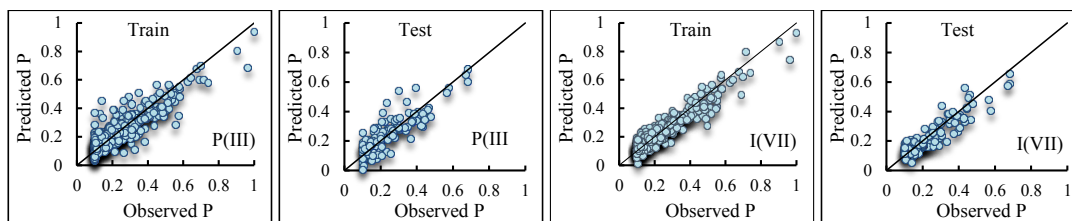
Table 4. Results of the models evaluation with time series decomposition

| نتایج ارزیابی مدل‌ها بر اساس داده‌های بارش | | | | | | |
|--|--------------|--------------|--------------|--------------|--------------|--------------|
| مدل | آزمون | | | تست | | |
| | R | DC | RMSE | R | DC | RMSE |
| CEEMD- GPR | | | | | | |
| P(I) | 0.734 | 0.531 | 0.108 | 0.623 | 0.484 | 0.119 |
| P(II) | 0.877 | 0.762 | 0.073 | 0.766 | 0.548 | 0.090 |
| P(III) | 0.875 | 0.802 | 0.066 | 0.822 | 0.721 | 0.081 |
| P(IV) | 0.901 | 0.801 | 0.066 | 0.805 | 0.699 | 0.086 |
| DWT- GPR | | | | | | |
| P(I) | 0.763 | 0.582 | 0.101 | 0.725 | 0.525 | 0.099 |
| P(II) | 0.943 | 0.882 | 0.050 | 0.906 | 0.816 | 0.057 |
| P(III) | 0.958 | 0.917 | 0.044 | 0.914 | 0.835 | 0.054 |
| P(IV) | 0.933 | 0.862 | 0.057 | 0.913 | 0.828 | 0.062 |
| نتایج ارزیابی مدل‌ها بر اساس عناصر اقلیمی و الگوهای پیوند از دور | | | | | | |
| CEEMD- GPR | | | | | | |
| I(I) | 0.652 | 0.535 | 0.101 | 0.638 | 0.508 | 0.106 |
| I(II) | 0.894 | 0.820 | 0.056 | 0.834 | 0.735 | 0.069 |
| I(III) | 0.895 | 0.823 | 0.055 | 0.845 | 0.742 | 0.064 |
| I(IV) | 0.883 | 0.805 | 0.059 | 0.822 | 0.726 | 0.075 |
| I(V) | 0.880 | 0.795 | 0.065 | 0.818 | 0.724 | 0.080 |
| I(VI) | 0.883 | 0.812 | 0.058 | 0.827 | 0.732 | 0.072 |
| I(VII) | 0.903 | 0.834 | 0.048 | 0.866 | 0.768 | 0.054 |
| I(VIII) | 0.570 | 0.374 | 0.214 | 0.471 | 0.351 | 0.250 |
| DWT- GPR | | | | | | |
| I(I) | 0.711 | 0.601 | 0.089 | 0.708 | 0.588 | 0.091 |
| I(II) | 0.974 | 0.938 | 0.037 | 0.926 | 0.851 | 0.046 |
| I(III) | 0.976 | 0.941 | 0.037 | 0.938 | 0.859 | 0.042 |
| I(IV) | 0.962 | 0.921 | 0.040 | 0.912 | 0.841 | 0.050 |
| I(V) | 0.959 | 0.909 | 0.044 | 0.908 | 0.838 | 0.053 |
| I(VI) | 0.963 | 0.929 | 0.039 | 0.918 | 0.848 | 0.048 |
| I(VII) | 0.984 | 0.954 | 0.032 | 0.961 | 0.889 | 0.036 |
| I(VIII) | 0.621 | 0.428 | 0.143 | 0.523 | 0.407 | 0.166 |

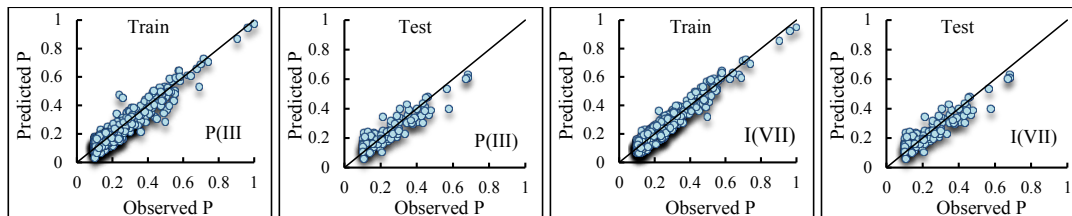
وقتی ادامه می یابد که تقریباً سیگنال باقیمانده ثابت شود. در نهایت زیر سری های به دست آمده از هر دو روش، به عنوان ورودی به مدل GPR داده شد. نتایج حاصل از مدل سازی های در نظر گرفته شده در جدول ۴ و شکل ۵ ارائه شده است. با مقایسه نتایج جدول ۳ و ۴ ملاحظه می شود که تجزیه سری های زمانی تا حدود زیادی باعث بهبود نتایج گردیده و تاثیر تجزیه با استفاده از تبدیل موجک گسسته بیشتر از روش تجزیه ی یکپارچه مد تجزیه کامل است. مطابق با جدول ۴، در حالت مدل سازی بر اساس تنها داده های بارش، مدل P(III) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} با داشتن بیشتر R و DC و کمترین مقادیر خطای RMSE مدل برتر است. برای حالتی که از عناصر اقلیمی و الگوهای پیوند از دور استفاده شده

تجزیه کامل و تبدیل موجک گسسته ابتدا سری های زمانی تجزیه گردید و سپس زیر سری های به دست آمده به عنوان ورودی به مدل GPR داده شد. مدل سازی نیز در دو حالت بر اساس تنها داده های بارش و بر اساس عناصر اقلیمی شامل میانگین بارش، دما و رطوبت نسبی ماهانه و الگوهای پیوند از دور شامل NOA, Nino3, MEI انجام شد. اولین قدم برای تجزیه سیگنال با تبدیل موجک، انتخاب موجک مادر است. در این تحقیق از موجک مادر db4 با دو سطح تجزیه استفاده شد. اساس روش CEEMD نیز تجزیه سیگنال به IMF های مختلف و یک سیگنال باقیمانده است که در نهایت مجموعه ی این ها، همان سیگنال اصلی خواهد بود. روش تشکیل IMF ها بر اساس تفریق تابع پایه از سیگنال اصلی است. این پروسه تا

(الف)

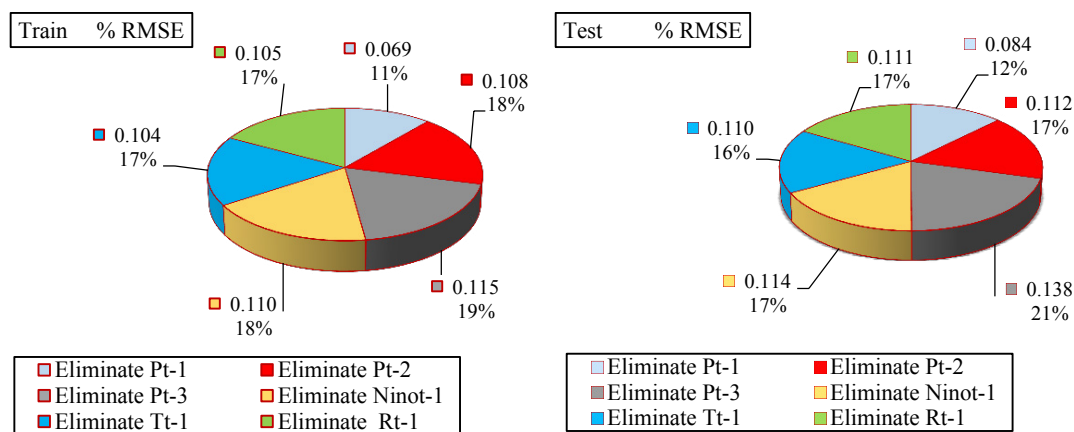


(ب)



شکل ۵. نتایج مرحله آزمون و تست مدل برتر بر اساس تجزیه دادهها با استفاده از (الف): CEEMD و (ب): DWT

Fig. 5. Results of the train and test steps of superior model based on data decomposition via (a): CEEMD and (b): DWT



شکل ۶. تاثیر نسبی هر یک از پارامترهای ورودی مدل برتر

Fig. 6. Relative significance of each of input parameters of the superior model.

تنها از الگوهای پیوند از دور مربوط به بازه زمانی قبل نتایج دقیقی به دست نیامده است (نتایج مدل I(VIII)). لازم به توضیح است که در مورد مدل انتخاب شده علاوه بر اینکه این مدل از نظر معیارهای ارزیابی جهت تخمین پارامتر وابسته باید دارای دقت مطلوبی باشد، تعداد پارامترهای به کار رفته در مدل و ساده بودن آن نیز مهم است. استفاده از مدل با تعداد پارامترهای کمتر هم راحت تر بوده و هم زمان و حجم محاسبات را کاهش می دهد. با در نظر گرفتن این نکته و با توجه به نتایج جدول ۴ و مدل I(II) مشاهده می گردد که در صورت استفاده از پارامترهای بارش و دما، پس از تجزیه سری داده ها روش استفاده شده در تحقیق کارایی خوبی را در فرایند پیش بینی نشان داده است. همچنین مدل P(III) نیز با استفاده

است نیز ملاحظه می شود که مدل I(VII) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} , $Nino3_{t-1}$, T_{t-1} , R_{t-1} با مقایسه مدل های تعریف شده مشاهده می گردد که استفاده از پارامترهای میانگین دما و رطوبت نسبی ماهانه باعث افزایش دقت مدل ها شده اند. همچنین تاثیر الگوی پیوند از دور Nino3 از NOA و MEI بیشتر است. در حالت کلی مطابق با جدول ۴، هر دو عناصر اقلیمی و الگوهای پیوند از دور در بهبود دقت نتایج تاثیر گذار هستند. باین حال، داده های اقلیمی در پیش بینی بارش موفق تر بوده و استفاده از داده های مربوط به الگوهای پیوند از دور در کنار داده های بارش ماه های گذشته دقت مدل را اندکی افزایش داده اند و در حالت استفاده

در حالت مدل سازی بر اساس عناصر اقلیمی و الگوهای پیوند از دور نیز مدل (VII) با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} , Nino3, T_{t-1} , R_{t-1} منجر به نتایج دقیق تری گردید. مشاهده شد که استفاده از پارامترهای میانگین دما و رطوبت نسبی ماهانه باعث افزایش دقت مدل ها می گردد و همچنین در بین الگوهای پیوند از دور نیز Nino3 نسبت به NOA و MEI تاثیرگذارتر است. در حالت کلی مشخص گردید که تنها استفاده از مقادیر بارش ماه قبل نمی تواند منجر به نتایج دقیقی گردد و هر دو عناصر اقلیمی و الگوهای پیوند از دور سبب بهبود دقت مدل ها می گردند. نتایج حاصل از آنالیز حساسیت نشان داد که پارامتر P_{t-3} بیشترین تاثیر را در تخمین میزان بارش ماهانه داراست.

۵- فهرست علائم

علائم انگلیسی

| | |
|----------|------------------------------|
| P | بارش |
| T | دما |
| R | رطوبت نسبی |
| Nino | شاخص دمای سطح اقیانوس آرام |
| NOA | نوسانات اطللس شمالی |
| MEI | شاخص چند متغیره انسو |
| N | تعداد نقاط داده در سری زمانی |
| L | سطح تجزیه |
| $r_n(x)$ | مولفه باقی مانده |

علائم یونانی

| | |
|---------------|------------------------|
| ε | مقدار نویز توزیع گاوسی |
|---------------|------------------------|

منابع

- [1] M.j. Nazemsadat, A.A. Gamgar Haghighi, M. Sharifzadeh, M. Ahmadvand, Adoption of long-term rainfall forecasts (studied by wheat farmers in Fars Province), Journal of Iranian Agricultural Science and Education, 22 (2006) 1-15. [in Persian]
- [2] P. Tofani, E. Mosaedi, A. Fakheri Fard, Precipitation forecast using wavelet theory, Water and Soil Journal (Agriculture Sciences and Technology), 25(5) (2011) 1217-1226. [in Persian]
- [3] H. Sharifan, B. Ghahreman, Estimation of Rain Forecast

از تنها پارامتر بارش نتایج قابل قبولی را ارائه داده است و می توان در صورت عدم دسترسی به سایر داده های اقلیمی یا پیوند از دور از این مدل جهت تخمین بارش ماهانه است کرد. نمودار مدل های برتر برای سری داده های آموزش و آزمون در شکل ۵ نشان داده شده است.

۵-۳- آنالیز حساسیت

جهت بررسی تاثیر پارامترهای به کار رفته در مدل برتر (مدل I(VII) در حالت تجزیه با تبدیل موجک گسسته) آنالیز حساست انجام گرفت. برای این منظور با حذف تک تک پارامترهای مدل برتر و اجرای دوباره مدل و تعیین معیار ارزیابی RMSE، میزان تاثیر پارامتر حذف شده در کاهش دقت مدل مورد بررسی قرار گرفت. نتایج حاصل از آنالیز حساسیت به صورت شکل ۶ ارائه گردیده است. همانطور که از شکل مشخص است با حذف پارامتر P_{t-3} دقت مدل تا حدود زیادی کاهش یافته است. بنابراین، می توان نتیجه گرفت که پارامتر P_{t-3} بیشترین تاثیر را در تخمین میزان بارش ماهانه داراست.

۴- نتیجه گیری کلی

پیش بینی بارش به دلیل ماهیت تصادفی آن در مکان و زمان همواره با مشکلات بسیاری مواجه بوده است و این عدم قطعیت از اعتبار بسیاری از مدل های پیش بینی می کاهد. در این تحقیق به پیش بینی میزان بارش ماهانه با استفاده از روش هوشمند GPR بر پایه تجزیه ی یکپارچه مد تجزیه کامل و تبدیل موجک پرداخته شد. مدل های ورودی متفاوتی با استفاده از داده های مربوط به عناصر اقلیمی مانند بارش، میانگین دما و رطوبت نسبی ماهانه ایستگاه تبریز در بازه زمانی (۱۹۷۸-۲۰۱۷) تعریف گردید. دو نوع مدل سازی با استفاده از داده های سری های زمانی اصلی و همچنین با استفاده از داده های سری های زمانی تجزیه شده با دو روش تجزیه ی یکپارچه مد تجزیه کامل و تبدیل موجک گسسته انجام شد. نتایج حاصله نشان داد که تجزیه سری های زمانی با هر دو روش CEEMD و DWT تا حدود زیادی باعث بهبود نتایج می گردد و تجزیه داده ها با استفاده از تبدیل موجک گسسته منجر به نتایج دقیق تری می شود. مطابق با نتایج حاصله، برای مدل های تعریف شده بر اساس تنها داده های بارش، مدل (III) P با پارامترهای ورودی P_{t-1} , P_{t-2} , P_{t-3} با داشتن بیشتر R و DC و کمترین مقادیر خطا مدل برتر گردید.

- statistical downscaling models, *Amirkabir Journal of Civil Engineering*, DOI: 10.22060/ceej.2018.14986.5806, (2018). [in Persian]
- [12] S. Kumar, D. Tripathy, S. Nayak, S. Mohaparta, Prediction of rainfall in India using artificial neural network models, *International Journal of intelligent system and applications*, 12 (2013) 1-22.
- [13] D. Nayak, A. Mahapatra, P. Mishra, A survey on rainfall prediction using artificial neural network, *International journal of computer applications*, 72(16) (2013) 32-40.
- [14] K.M. Lau, H.Y. Weng, Climate signal detection using wavelet transform, *How to make time series sing*, *Bull Am Meteorol Soc*, 76 (1995) 2391-2402.
- [15] K. Adamowski, A. Prokoph, J. Adamowski, Development of a new method of wavelet aided trend detection and estimation, *Hydrology Process*, 23(18) (2009) 2686-2696.
- [16] C.M. Chou, Complexity analysis of rainfall and runoff time series based on sample entropy in different temporal scales, *Stochastic Environmental Research and Risk Assessment*, 6 (2011) 1401-1408.
- [17] Y. Amirat, M. Benbouzidb, T. Wang, K. Bacha, G. Feld, EEMD-based notch filter for induction machine bearing faults detection, *Applied Acoustics*, 133 (2018) 202-209.
- [18] Z. Wu, N.F. Huang, A study of the characteristics of white noise using the empirical mode decomposition method, *Proc RS Lond 460A*: 1597-1611, (2004).
- [19] C.E. Rasmussen, C.K.I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, MA, (2006).
- [20] W.C. Dawson, R. Wilby, An artificial neural network approach to rainfall-runoff modelling. *Hydrological Sciences Journal*, 43(1) (1998) 47-66.
- [21] NOAA Earth System Research laboratory, <https://www.esrl.noaa.gov/psd/data/climateindices/list/>, (2009).
- Using ARIMA Technique in Golestan Province, *Journal of Agricultural Sciences and Natural Resources*, 14 (2008) 13-14. [in Persian]
- [4] M. Gholabi, A. Akhund, Ali, F. Radmanesh, M. Kashifipour, Comparison of Predicting the Jenkins Box Models in Seasonal Modeling (Case Study: Selected Stations in Khuzestan Province), *Quarterly Journal of Geographic Research*, 29(3) (2012) 61-72. [in Persian]
- [5] A. S. Soltani, A. Saberi, M. Gheisouri, Determination of the best time series model for forecasting annual rainfall of selected stations of Western Azerbaijan province, *Researches in Geographical Sciences*, 17(44) (2017) 87-105.
- [6] ASCE, Task Committee on Application of Artificial Neural Networks in Hydrology, *Artificial Neural Networks in hydrology. I: Preliminary concepts*, *Hydrological Engineering*, ASCE. 5(2) (2000) 115-123.
- [7] C. Siviapragasam, S. Liong, Rainfall and runoff forecasting with SSA-SVM approach, *Hydroinformatics*, 3(2001) 141-152.
- [8] K. Roushangar, R. Ghasempour, The study of the performance of classical and artificial intelligence methods in the estimation of roughness coefficients in pontoons, *Irrigation and Drainage Journal of Iran*, 12(4) (2019) 811-822. [in Persian]
- [9] O. Kisi, M. Cimen, Precipitation forecasting by using wavelet-support vector machine conjunction model, *Engineering Application Artificial Intelligence*, 25 (2012) 783-792.
- [10] F.S. Marzano, E. Fionda, P. Ciotti, Neural-network approach to ground-based passive microwave estimation of precipitation intensity and extinction, *Hydrology*, 328 (2006) 121-131.
- [11] Z. Razzaghzadeh, V. Nourani, N. Behfar, The conjunction of feature extraction method with AI-based ensemble

چگونه به این مقاله ارجاع دهیم
K. Roushangar, R. Ghasempour, Monthly precipitation prediction improving using the integrated model based on kernel-wavelet and complementary ensemble empirical mode decomposition, Amirkabir J. Civil Eng., 52(10) (2021) 2649-2660.

DOI: 10.22060/ceej.2019.16043.6109

