



## مدیریت آب و آبیاری

دوره ۱۲ ■ شماره ۱ ■ بهار ۱۴۰۱

صفحه‌های ۱۳-۱

DOI: 10.22059/jwim.2022.336530.956

مقاله پژوهشی:

### ارزیابی رویکردهای ساختاری فضای حالت نسبت به کلاسیک در پیش‌بینی سری‌های زمانی بارش (حوضه آبریز دز)

محمد رضا شریفی<sup>۱\*</sup>، امین محمدزاده شوبادegar<sup>۲</sup>، فریدون رادمنش<sup>۳</sup>، بهزاد منصوری<sup>۴</sup>

۱. استادیار، گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران.
  ۲. دانشجوی دکتری منابع آب، گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران.
  ۳. دانشیار، گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران.
  ۴. دانشیار، گروه آمار، دانشکده علوم ریاضی و کامپیوتر، دانشگاه شهید چمران اهواز، اهواز، ایران.
- تاریخ دریافت مقاله: ۱۴۰۰/۱۰/۰۹ تاریخ پذیرش مقاله: ۱۴۰۱/۰۱/۰۷

#### چکیده

در این مقاله، مطالعه‌ای در مورد استفاده از تکنیک‌های پیش‌بینی بارش با داده‌های سری زمانی ارائه شده است. سری‌های زمانی ابزاری کارآمد برای شناخت ماهیت پدیده‌های هیدرولوژیکی هستند که با داشتن شناخت کافی از آن‌ها می‌توان تغییرات آینده را مدل‌سازی و پیش‌بینی کرد. مدل‌های مختلف آماری با هدف کاهش خطا و بالا بردن دقت پیش‌بینی، در نظر گرفته شده است. فضای حالت به واسطه ساختاری بودن و انعطاف‌پذیری آن، امکان مدل‌بندی هر یک از مؤلفه‌های تشکیل‌دهنده متغیر شامل سطح، فصلی و تصادفی را به‌طور مجزا دارد. از این‌رو با شناسایی سیستم در نحوه مدل‌سازی متغیر مورد مطالعه، امکان کنترل و حداقل نمودن خطای برآورد، به‌طور هوشمندانه‌تری در مقایسه با مدل‌های کلاسیک را دارد. در پژوهش حاضر به منظور ارزیابی قابلیت مدل‌سازی فضای حالت و مقایسه با مدل‌های کلاسیک، اقدام به مدل‌سازی بارش ماهانه در سه ایستگاه باران‌سنجی، در حوضه آبریز دز، با چهار مدل ساختاری فضای حالت شامل فیلتر کالمن، مدل هموارسازی نمایی ETS و مدل‌های هموارسازی نمایی اصلاح شده BATS و TBATS و مدل کلاسیک ARIMA شد. نتایج نشان داد در ایستگاه سپیددشت سزار براساس معیار RMSE و MAE مدل TBATS و در ایستگاه تنگ پنج بختیاری براساس معیار RMSE و MAE مدل فیلتر کالمن و در ایستگاه تله زنگ براساس معیار RMSE و MAE مدل TBATS بهترین مدل‌ها انتخاب شدند.

**کلیدواژه‌ها:** باکس-جنکینز، فضای حالت، فیلتر کالمن، هموارسازی نمایی.

### Evaluation of structural approaches of state space compared to classical in predicting precipitation time series (Dez catchment)

Mohammadreza Sharifi<sup>1\*</sup>, Amin Mohammadzadeh Shobadegar<sup>2</sup>, Fereydoon Radmanesh<sup>3</sup>, Behzad Mansouri<sup>4</sup>

1. Assistant Professor, Department of Hydrology and Water Resources, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
2. Ph.D. Student in Water Resources, Department of Hydrology and Water Resources, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
3. Associate Professor, Department of Hydrology and Water Resources, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
4. Associate Professor, Department of Statistics, Faculty of Mathematics and Computer Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

Received: December 30, 2021

Accepted: March 07, 2022

#### Abstract

In this paper, a study on the use of precipitation prediction techniques with time series data was presented. Time series are an effective tool for understanding the nature of hydrological phenomena that with sufficient knowledge of them, future changes can be modeled and predicted. Various statistical models have been considered with the aim of reducing error and increasing forecast accuracy. Due to its structural and flexibility, state space makes it possible to model each of the components of a variable, including surface, seasonal and random separately. Therefore, by identifying the system in the way of modeling the studied variable, it is possible to control and minimize the estimation error, more intelligently compared to classical models. In the present study, in order to evaluate the modeling capability of state space and compare it with classical models, monthly precipitation modeling was performed in three rain gauge stations in Dez catchment, with four structural models of state space including Kalman filter, ETS exponential smoothing model and Modified exponential smoothing models were BATS and TBATS and the classic model was ARIMA. The results showed that at Sepiddasht Sesar station based on RMSE and MAE criteria of TBATS model and in Tangpanj Bakhtiyari station based on RMSE and MAE criterion of Kalman filter model and in Telezang station according to RMSE and MAE criterion of TBATS model the best models were chosen.

**Keywords:** Box-Jenkins, Exponential smoothing, Kalman Filter, State Space.

## مقدمه

مدل سازی سری های زمانی، با اهدافی از جمله پیش بینی مقادیر متغیر مورد مطالعه در آینده، در تحلیل و بازنگری سیاست ها و انتخاب گزینه های مناسب مدیریتی در بهره برداری از سامانه های آبی، نقش مهمی را ایفا می نماید.

Harting (2010) به اهمیت پیش بینی میزان بارندگی در حوزه انرژی پرداخت و بیان کرد مناطق پربارش در انتخاب محل ساخت نیروگاه ها به منظور تولید انرژی برق آبی در نظر گرفته می شود. رویکردهای مختلف آماری با هدف کاهش خطای مدل سازی و بالابردن دقت پیش بینی، پیشنهاد شده است. مطالعات متعددی در مورد مدل سازی داده های بارش وجود دارد. به عنوان مثال، Brath et al. (2000) و al. (2002) از ARIMA و شبکه عصبی مصنوعی (ANN) برای پیش بینی بارش و سیل در ایتالیا استفاده کردند و نشان دادند که تحلیل سری های زمانی، بهبودی در دقت پیش بینی سیل ایجاد می کند. Soltani et al. (2007) مدل  $ARIMA(0, 1, 1)$  فصلی را برای ایستگاه اصفهان پیشنهاد کردند و از مدل های مختلف ARIMA فصلی برای مناطق مختلف استفاده نمودند. Sun et al. (2019) از سری بارش تک متغیره و شبکه های عصبی استفاده کردند؛ بهترین میانگین مطلق خطا (MAE) به دست آمده  $31/25$  mm بود. در برخی موارد برای دستیابی به مقادیر بارش، از الگوریتم های بارش ماهواره ای استفاده می شود. برای نمونه، Sadeghi et al. (2019) الگوریتم PERSIANN-CNN را برای پیش بینی داده های ماهواره ای با استفاده از شبکه های عصبی کانولوشن توسعه دادند. کوچک ترین MAE و RMSE که آن ها به دست آوردند به ترتیب  $0/12$  mm و  $0/88$  mm/hr بود. Asemota et al. (2016) بارندگی در شمال نیجریه انجام دادند. آن ها از داده های بارندگی ماهانه جمع آوری شده از سال 1981 تا 2013 استفاده کردند تا راه را برای برنامه ریزی کشاورزی جدید در

شمال نیجریه هموار کنند. آن ها این کار را با استفاده از مدل فضای حالت<sup>1</sup> از طریق فیلتر کالمن انجام دادند. Zulfi et al. (2018) مطالعه ای در مورد توسعه یک مدل پیش بینی بارندگی با استفاده از فضای حالت و فیلتر کالمن منتشر کردند. در مطالعه آن ها، روش های فضای حالت و ARIMA از نظر عملکرد در رابطه با پیش بینی بارندگی مقایسه شدند. داده های اعتبارسنجی جمع آوری شده از سال 2005 تا 2015 با استفاده از الگوریتم k-means به خوشه ها تقسیم شدند که از الگوریتم فیلتر کالمن برای مدل سازی و پیش بینی در هر خوشه استفاده شد. در پایان مطالعه به این نتیجه رسیدند که عملکرد مدل فضای حالت با فیلتر کالمن بهتر از مدل ARIMA برای پیش بینی بارندگی است.

Masazade et al. (2019) بر میزان بارندگی تخمین زده شده توسط فضای حالت با فیلتر کالمن با اندازه گیری های بازتاب راداری متمرکز شدند. هدف از اعمال فیلتر کالمن در آن مطالعه مدل سازی مقادیر واقعی بارندگی بود.

Neslihanoglu et al. (2021) با استفاده از مدل فضای حالت از طریق فیلتر کالمن، به پیش بینی بارش پرداختند. میزان انطباق نتایج حاصل از روش فضای حالت و مقادیر مشاهده ای متناظر، حاکی از انطباق کامل نتایج شبیه سازی بارش ماهانه با مشاهدات اندازه گیری در ایستگاه مورد مطالعه آن ها بود. Soumik Ray et al. (2021) به مدل سازی سری زمانی SARIMA و پیش بینی بارندگی و دما ماهانه در کشورهای جنوب آسیا پرداختند و نتایج آن ها نشان داد که تمامی سری داده ها هم برای بارندگی و هم دما شامل سه مؤلفه سری زمانی مانند روند تصادفی، فصلی و جزء نامنظم است. Gabriel (2021) بارش ماهانه منطقه انیوگو را با مقایسه مدل های SARIMA و ASARIMA در یک سری زمانی فصلی غیرایستا مورد مطالعه قرار داد؛ نتایج نشان داد که زیر کلاس  $ASARIMA(2,1,1)_2$  بهتر از همه زیر کلاس های SARIMA توسط معیار آکائیک است.

ارزیابی رویکردهای ساختاری فضای حالت نسبت به کلاسیک در پیش‌بینی سری‌زمانی بارش (حوضه آبریز دز)

پیش‌بینی بارش ماهانه منطقه حوضه آبریز دز در سه ایستگاه (سپیددشت سزار، تنگ پنج بختیاری، تله زنگ) پردازد و با مقایسه آن‌ها با یکدیگر نشان دهد که کدام یک از این مدل‌ها بهتر بوده و می‌تواند به مدل‌های آمار در هیدرولوژی اضافه شود.

## مواد و روش‌ها

### محدوده مورد مطالعه

حوضه آبریز رودخانه دز به‌عنوان یک حوضه درجه سه، زیر مجموعه‌ای از حوضه کارون بزرگ محسوب می‌شود و در تقسیم‌بندی بزرگ‌تر در زیرمجموعه حوضه خلیج فارس و دریای عمان قرار می‌گیرد. مساحت کل حوضه بالغ بر ۲۱۷۲۰ کیلومتر مربع و متوسط ارتفاع حوضه حدود ۱۶۰۰ متر می‌باشد. محدوده مورد مطالعه در این پژوهش حوضه آبریز دز در بالادست سد دز می‌باشد که مساحت آن حدود ۱۷۳۶۵ کیلومتر مربع است. شکل (۱) موقعیت جغرافیایی سه ایستگاه باران‌سنجی در حوضه آبریز دز را نشان می‌دهد. همان‌طور که در نقشه مشاهده می‌شود هر سه ایستگاه از ارتفاع حدوداً یکسانی برخوردار هستند.

Zeng et al. (2016) به تجزیه و تحلیل روندهای سری

زمانی برای پیش‌بینی کوتاه‌مدت بیماری اپیدمی سیاه‌سرفه در کشور چین به کمک مدل‌های ETS (M, N, M) و مدل ARIMA پرداختند. نتایج نشان داد مدل ETS (M, N, M) با حداقل  $RSME = 0.202$  در واسنجی و صحت‌سنجی بالاترین کارایی را دارد. Livera et al. (2011) به بررسی پیش‌بینی سری‌های زمانی با الگوهای پیچیده فصلی با استفاده از هموارسازی نمایی پرداختند و نشان دادند که مدل BATS واضح‌ترین تعمیم مدل‌های جدید فصلی است که برای دوره‌های فصلی متعدد امکان‌پذیر است و می‌تواند تعداد بسیار زیادی حالت داشته باشد. جزء فصلی اولیه حاوی مقادیر غیر صفر است، که این مقدار برای الگوهای فصلی با دوره‌های بالا به کار می‌رود. مدل TBATS به تخمین مقادیر فصلی اولیه نیاز دارد، که احتمالاً بسیار کوچک‌تر از تعداد پارامترهای فصلی در مدل‌های BATS است. لذا براساس مطالعات انجام‌شده می‌توان گفت رویکرد فضای حالت با توجه به صحت نتایج به‌دست‌آمده، بهترین روش برای پیش‌بینی بارش است. در نتیجه این مقاله در نظر دارد با چهار مدل فضای حالت و یک مدل کلاسیک به

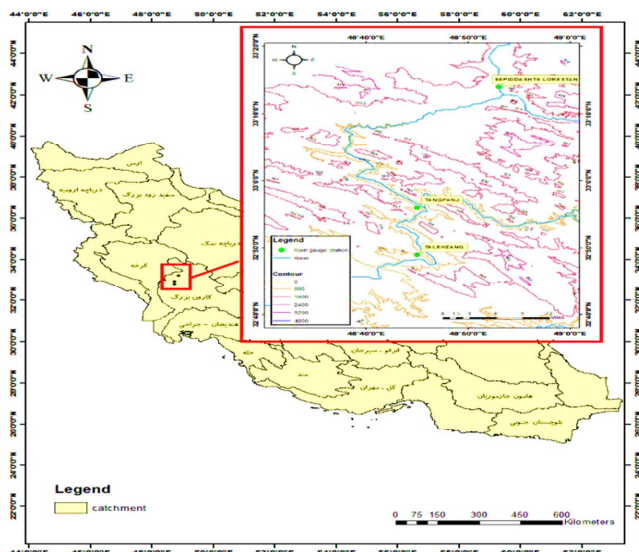


Figure 1. Geographical location of the studied rain gauge stations in Dez catchment

سیستم را در زمان‌های مختلف نشان می‌دهند. هم‌چنین فرض می‌کنیم که مشاهدات  $y_1, y_2, \dots$  از این پدیده یا سیستم در اختیار باشد و مشخص باشد که  $y_t$  ها و  $\alpha_t$  ها مرتبط هستند. مدل فضای حالت رابطه بین  $y_t$  ها و  $\alpha_t$  را مشخص می‌کند. در واقع هدف مدل‌بندی فضای حالت کسب دانش درباره  $\alpha_t$  ها از طریق اطلاعات به‌دست‌آمده از مشاهدات  $y_t$  است (Durbin & Koopman, 2012).

### مدل‌های فضای حالت در فرم کلی

نمایش فضای حالت شامل تمام موارد موردعلاقه در بحث سری‌های زمانی است که به‌عنوان یک مدل خطی پویا شناخته می‌شود (Shumway et al., 2016). فرم کلی یک مدل فضای حالت خطی و گوسین به‌صورت زیر تعریف می‌شود:

$$y_t = Z_t \alpha_t + \varepsilon_t \quad \varepsilon_t \sim NID(0, H_t) \quad (1)$$

$$\xi_t \sim NID(0, Q_t), \quad \alpha_1 \sim NID(\alpha_1, P_1) \quad (2)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \xi_t$$

در معادلات فوق،  $t = 1, \dots, n$  و بردار  $P \times 1$

مشاهدات و  $\alpha_t$  بردار  $m \times 1$  غیرقابل مشاهده حالت است. معادله (۱) را معادله اندازه<sup>۳</sup> و معادله (۲) را معادله حالت<sup>۴</sup> می‌گویند. شایان ذکر است که مشاهدات  $y_t$  تابعی خطی از متغیرهای پنهان (حالت)  $\alpha_t$  هستند. هم‌چنین  $\varepsilon_t$  اغتشاش مشاهده در زمان  $t$ ،  $\xi_t$  اغتشاش حالت در زمان  $t$  است که از یکدیگر مستقل هستند. به همین دلیل این مدل را مدل فضای حالت با چندین منبع خطا<sup>۵</sup> (MSOE) نیز می‌گویند. در معادلات (۱) و (۲) ماتریس‌های  $H_t$  و  $Q_t$  ماتریس واریانس اغتشاش مشاهده و حالت،  $\alpha_1$  و  $P_1$  مقادیر اولیه برای میانگین و واریانس سطح  $T_t$  ماتریس انتقال و  $R_t$  ماتریس انتخاب است. کالمن (۱۹۶۹) با نوشتن معادلات به‌فرم (۱) و (۲) یک روش بازگشتی مؤثر را برای برآورد بردار پنهان  $\alpha_t$  با استفاده از بردار مشاهدات  $y_t$  پیشنهاد کرد که به فیلتر

در حوضه آبریز دز، میانگین سالانه بارش طی دوره ۵۴ ساله (۱۹۶۷-۲۰۲۱) در هر یک از سه ایستگاه سپیددشت سزار، تنگ‌پنج بختیاری و تله‌زنگ، به‌ترتیب ۶۰، ۱۳۷ و ۷۱ میلی‌متر می‌باشد. مقادیر مزبور در جدول (۱)، به‌همراه مختصات ایستگاه‌های باران‌سنجی آورده شده است. بازسازی مقادیر ماهانه گم‌شده، با استفاده از الگوریتم EM<sup>۶</sup> توسط Little & Rubin (1987) و Yozgatligil et al. (2013) انجام شده است. این الگوریتم یک روش تکرار است که در هر تکرار دو گام انجام می‌گیرد؛ گام اول امید ریاضی داده‌های گم‌شده به شرط داده‌های مشاهده‌شده محاسبه و جایگزین داده‌های گم‌شده قرار می‌گیرد و پارامترهای موردنظر برآورد می‌شوند و در گام دوم، لگاریتم تابع درست‌نمایی حداکثر می‌شود. در ضمن داده‌های این سه ایستگاه از سازمان مدیریت منابع آب کشور و سازمان آب و برق خوزستان اخذ شده است.

Table 1. Details of the studied rain gauge stations in the study area

Rows	River	Station	Longitude (DD)	Latitude (DD)	Average annual precipitation (mm)
1	Sezar	Sepiddasht	48.89	33.22	60
2	Bakhtiyari	Tangpanj	48.76	32.93	137
3	Dez	Telezang	48.77	32.82	71

### روش‌شناسی

تفاوت اساسی مدل‌های فضای حالت نسبت به مدل‌بندی کلاسیک (باکس-جنکینز)، ساختاری بودن آن‌ها است. به این معنی که در مدل‌بندی فضای حالت، عملاً هر یک از مؤلفه‌های تشکیل‌دهنده متغیر، شامل سطح، روند، اثرات فصلی و اغتشاش، به‌صورت مجزا مدل‌بندی می‌شوند. فرض کنید که  $\alpha_1, \alpha_2, \dots$  مجموعه‌ای از بردارهای حالت غیرقابل مشاهده هستند که وضعیت یک پدیده فیزیکی یا

ارزیابی رویکردهای ساختاری فضای حالت نسبت به کلاسیک در پیش‌بینی سری‌زمانی بارش (حوضه آبریز دز)

$\tilde{w}_{j,t+1}$  و  $\tilde{w}_{j,t+1}^*$  متغیرهای مستقل هم‌توزیع با  $N(0, \sigma_w^2)$  هستند (Durbin et al., 2012).

$$\lambda_j = \frac{2\pi j}{s}, j = 1, \dots, \left\lfloor \frac{s}{2} \right\rfloor \quad (5)$$

$$\gamma_t = \sum_{j=1}^{\left\lfloor \frac{s}{2} \right\rfloor} (\tilde{\gamma}_j \cos \lambda_j t + \tilde{\gamma}_j^* \sin \lambda_j t)$$

$$j = 1, \dots, \left\lfloor \frac{s}{2} \right\rfloor, t = 1, \dots, n \quad (6)$$

$$\tilde{w}_{j,t+1}^* \tilde{\gamma}_{j,t} + \tilde{\gamma}_{j,t+1}^* = \tilde{\gamma}_{j,t+1} = 1, \dots, n$$

$$\tilde{\gamma}_{j,t} + \tilde{w}_{j,t+1}$$

#### مدل‌های فضای حالت تک منبع خطا (SSOE)

در یک مدل فضای حالت با چند منبع خطا  $\varepsilon_t$  و  $\xi_t$  دو منبع خطای مستقل هستند. فرم کلی مدل فضای حالت با یک منبع خطا به صورت زیر است که در آن،  $c$  یک ثابت است. بنابراین خطای حالت و خطای اندازه‌گیری همبسته هستند. هر مدل MSOE خطی را می‌توان به فرم یک مدل SSOE<sup>v</sup> نوشت و بر عکس (Hyndman et al., 2008).

$$y_t = Z_t \alpha_t + \varepsilon_t \quad \varepsilon_t \sim NID(0, H_t) \quad (7)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t c \varepsilon_t \quad (8)$$

از مدل‌های SSOE به عنوان مدل، برای روش‌های هموارسازی نمایی<sup>^</sup> استفاده می‌شود. روش‌های هموارسازی نمایی به خاطر دقت و ثبات بالای پیش‌بینی‌هایی که ارائه می‌کنند، معروف هستند. در این روش‌ها فرض می‌کنند که یک سری زمانی از اجزای مجزایی تشکیل شده که این اجزای رفتار سری را در طول زمان تعیین می‌کنند. در حالت معمول این اجزا را روند (T)، نمود فصلی (S) و جزء تصادفی (E) در نظر می‌گیرند. در مدل، هر یک از این مؤلفه‌ها می‌تواند به صورت جمعی (A) یا به صورت ضربی (M) باشند یا این‌که در مدل نباشند (N). هم‌چنین مؤلفه روند ممکن

کالمن مشهور شد. الگوریتم فیلتر کالمن پردازش‌گری بازگشتی با حالت بهینه است (Ribeiro, 2000). واژه "بهینه" بهترین تخمین را براساس ایده به حداقل رساندن خطای حالت در حین برآورد آن نشان می‌دهد، Ribeiro (2004). واژه "بازگشتی" به این معنی است که داده‌های گذشته نباید توسط فیلتر کالمن ذخیره شوند و با اندازه‌گیری جدید، هر بار توسط فیلتر دوباره ارزیابی می‌شوند که این کار سرعت پردازش اطلاعات را بالا می‌برد. مدل‌بندی فضای حالت یک کلاس بسیار گسترده و انعطاف‌پذیر را تشکیل می‌دهد، به گونه‌ای که تقریباً تمام مدل‌های مختلف در سری‌های زمانی و رگرسیون را می‌توان در قالب فضای حالت نوشت. برخی از زیر کلاس‌های مدل شامل مدل‌های سطح موضعی با اثرات فصلی، ETS، BATS و TBATS در شبیه‌سازی بارش ماهانه و هم‌چنین مدل‌های باکس-جنکینز، به شرح ذیل می‌باشند.

#### مدل سطح موضعی با اثر فصلی (فیلتر کالمن)

چنانچه نمودار داده‌های بارش فاقد روند بوده و فقط دارای اثر فصلی باشند، مدل سطح موضعی با اثر فصلی به فرم ذیل برای برازش به داده‌های بارش در نظر گرفته می‌شود. معادلات (۳) و (۴) حالات خاص معادلات کلی (۱) و (۲) هستند که در آن  $\mu_t$  اثر سطح و  $\gamma_t$  اثر فصلی را نشان می‌دهد و  $\varepsilon_t$  و  $\xi_t$  به ترتیب اغتشاش مشاهده و سطح را نشان می‌دهند.

$$y_t = \mu_t + \gamma_t + \varepsilon_t \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \quad (3)$$

$$\mu_{t+1} = \mu_t + \xi_t \quad \xi_t \sim NID(0, \sigma_\xi^2) \quad (4)$$

مؤلفه فصلی  $\gamma_t$  را می‌توان به فرم یک متغیر ظاهری<sup>6</sup> یا بفرم مثلثاتی به صورت معادله (۵) نوشت که در آن  $s$  دوره فصلی،  $[x]$  جزء صحیح  $x$  است و مقادیر  $\tilde{\gamma}_j^*$  و  $\tilde{\gamma}_j$  ثابت‌هایی معلوم هستند؛ برای یک مدل فصلی زمان-متغیر این عبارت‌ها را می‌توان به فرم معادله (۶) نوشت که

برای  $i = 1, \dots, T$  نمایش داده شده اند و  $\varphi$  پارامتر میرایی است که روند را در بلندمدت کنترل می کند. با وجود عمومیت مدل BATS، این مدل دارای ضعف هایی مانند تعداد بسیار زیاد حالت اولیه است که باید برآورد شوند. *Livera et al.* (2011) به منظور دست یافتن به یک مدل با تعداد کم تر پارامتر و حالت های اولیه از نمایش مثلثی معادله (۱۰) برای مؤلفه های فصلی استفاده کردند:

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)} \quad (10)$$

$$s_{j,t}^{*(i)} = s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + \gamma_1^{(i)} d_t$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + \gamma_2^{(i)} d_t$$

که در آن،  $\gamma_1^{(i)}$  و  $\gamma_2^{(i)}$  پارامترهای هموارسازی و

$\lambda_j^{(i)} = \frac{2\pi j}{m_i}$  هستند. این مدل را مدل TBATS می گویند.

### مدل کلاسیک ARIMA

مدل های ARIMA توسط باکس و جنکینز برای پیش بینی و کنترل سری های زمانی توسعه پیدا کردند. به همین خاطر این مدل ها را مدل های کلاس باکس- جنکینز نیز می گویند. یک مدل ARIMA از تلفیق یک مدل میانگین متحرک و خودبازگشت برای داده هایی که با تفاضل گیری ایستا شده اند، تشکیل شده است. با وجود گستردگی استفاده از این مدل ها در سری های زمانی، محدودیت آن ها در لزوم ایستابودن سری تحت بررسی، باعث می شود که گاهی کاربرد مدل های ARIMA در داده های واقعی مشکل پیدا کند. برای پیش بینی سری های زمانی که دارای دوره فصلی هستند، مدل های ARIMA گسترش یافته و از مدل SARIMA استفاده می شود. مدل SARIMA در معادله (۱۱) تعریف می شود:

$$\varphi(B_s)\varphi(B)(1-B)^d(1-B^s)^D y_t = \theta(B^s)\theta(B)\varepsilon_t \quad (11)$$

در اینجا، آن را با  $(P, D, Q)_s \times (p, d, q)$  ARIMA نشان می دهیم که در آن s دوره فصلی است (Box &

است میرا (d) باشد. از ترکیب های مختلف این اجزا ۳۰ مدل مختلف به وجود می آید که به آن ها به اختصار مدل های کلاس ETS گفته می شود. برای مثال مدل ETS (A,N,N) مدل هموارسازی نمایی ساده با خطای جمعی و مدل ETS (M,N,N) مدل هموارسازی نمایی ساده با خطای ضربی است. برای جزئیات کم تر در خصوص مدل های ETS می توان به *Hyndman et al.* (2008) مراجعه کرد.

### مدل های BATS و TBATS

برخی از سری های زمانی دارای الگوهای فصلی پیچیده هستند. *Livera et al.* (2011) برای چنین سری هایی قبل از مدل های اصلاح شده، هموارسازی نمایی را در قالب فضای حالت به منظور امکان مدل بندی طیف گسترده تری از الگوهای فصلی را معرفی کردند. فرض کنید که  $y_t^{(\omega)}$  مشاهده تبدیل شده در زمان t با روش باکس-کاکس با پارامتر  $\omega$  باشد. آن ها مدل BATS را به شکل معادله (۹) تعریف کردند:

$$y_t^{(\omega)} = \begin{cases} \frac{y_t^{(\omega)} - 1}{\omega}, \omega \neq 0 \\ \log y_t, \omega = 0 \end{cases} \quad (9)$$

$$y_t^{(\omega)} = l_{t-1} + \phi b_{t-1} + \sum_{i=1}^T s_{t-m_i}^{(i)} + d_t$$

$$l_t = l_{t-1} + \phi b_{t-1} + \alpha d_t$$

$$b_t = (1 - \phi)b + \phi b_{t-1} + \beta d_t$$

$$s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t$$

$$d_t = \sum_{i=1}^p \varphi_i d_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

در روابط فوق،  $m_1, \dots, m_T$  دوره های فصلی،  $l_t$

سطح موضعی<sup>۱۱</sup> در دوره t، b روند طولانی مدت<sup>۱۱</sup>،  $b_t$

روند کوتاه مدت<sup>۱۲</sup> در دوره t،  $s_t^{(i)}$  بیانگر i-امین مؤلفه

فصلی در زمان t،  $d_t$  بیانگر یک فرایند ARMA(p,q) و  $\varepsilon_t$

یک فرایند نوفه سفید گاوسی با میانگین صفر و واریانس

ثابت  $\sigma^2$  هستند. پارامترهای هموارساز به وسیله  $\alpha$ ،  $\beta$  و  $\gamma_i$

ساخته‌شده در پایان یک مدل معنادار است (Willmott *et al.*, 2005). معادلات معیارهای MAE، MSE و RMSE در روابط (۱۲)، (۱۳) و (۱۴) نشان داده شده است.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_t| \quad (12)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_t)^2 \quad (13)$$

$$RMSE = \sqrt{MSE} \quad (14)$$

### نتایج و بحث

شکل (۲) تجزیه سری زمانی میزان بارش در سه ایستگاه سپیددشت سزار، تنگ پنج بختیاری و تله‌زنگ به همراه اجزای تشکیل‌دهنده آن را نشان می‌دهد. این تجزیه به کمک دستور decompose در برنامه R انجام شده است. همان‌طور که مشاهده می‌شود پس از محاسبه روند و تغییرات فصلی، تنها چیزی که در سری زمانی به جا می‌ماند نوسانات نامنظم است. بسیاری از نوسانات نامنظم ناشی از وقایع غیرمعمول هستند که قابل پیش‌بینی نمی‌باشند. با توجه به این‌که اجزای سری زمانی، همیشه به‌تنهایی و به‌صورت منفرد رخ نمی‌دهند و می‌توانند با ترکیبات دیگر و یا هم‌زمان رخ دهند، نمی‌توان تنها یک مدل را به‌عنوان بهترین مدل پیش‌بینی در نظر گرفت، زیرا مدلی که تنها مشخصه روند را با خود به‌همراه دارد، نمی‌تواند برای پیش‌بینی سری زمانی که هم روند و هم تغییرات فصلی را به‌طور هم‌زمان در بر دارد، مناسب باشد.

و همان‌طور که در شکل (۲) مشاهده می‌شود، هیچ الگو یا روند قطعی در بارش وجود ندارد. این سری زمانی به سه جزء تفکیک می‌شود؛ جزء روندی، جزء فصلی و جزء تصادفی، که جزء فصلی به‌وضوح مشاهده می‌شود. تنوع قله‌ها و فرورفتگی‌ها از داده‌های ارائه‌شده ناشی می‌شود. داده‌های بارندگی سری زمانی در طول ۵۴ سال مورد مطالعه، افزایش و کاهش می‌یابد.

(Jenkins, 1970). با توجه به فصلی مشاهده‌شده در داده‌های بارش نمودارهای خودهمبستگی و خودهمبستگی جزئی سری، مدل SARIMA با استفاده از تابع auto.arima در نرم‌افزار R و به‌صورت SARIMA  $(P, D, Q)_{12} \times (p, d, q)$  برازش شده و پس از برآورد پارامترها، مدل مناسب براساس معیارهای اطلاعات بیزین آکائیک و براساس الگوریتم خودکار انتخاب شد. این رویکرد در سراسر این پژوهش مدل SARIMA نامیده می‌شود (Hyndman *et al.*, 2008).

### معیارهای ارزیابی مدل

به‌منظور مقایسه عملکرد پیش‌بینی مدل‌ها، در این مطالعه از سه ملاک MAE، MSE و RMSE استفاده شد. به‌دلیل این‌که میزان بارش در برخی از ماه‌ها صفر است، نسبتی خطا مانند MAPE وجود ندارد. برای این منظور سری زمانی به دو بخش آموزش مدل (واسنجی) و مجموعه آزمایشی (صحت‌سنجی) تقسیم شد. دو سال آخر سری به‌عنوان مجموعه آزمایشی در نظر گرفته شد. سپس معیارهای عملکرد ارزیابی هم برای آموزش مدل و هم برای آزمایش مدل محاسبه شدند. قابل ذکر است که به‌منظور تثبیت واریانس سری‌ها از یک تبدیل رادیکالی استفاده کرده و سپس از تبدیل معکوس برای پیش‌بینی مدل‌ها استفاده شده است.

MAE قدر مطلق تفاوت بین مقادیر پیش‌بینی‌شده و مشاهداتی را اندازه‌گیری می‌کند. MSE به‌عنوان مقدار میانگین مربعات خطاها عمل می‌کند و وقتی تفاوت بین MAE و MSE افزایش می‌یابد، به این معنی است که اندازه خطا با یکی از این دو معیار سازگارتر است. مقادیر کوچک معیارهای ارزیابی پیش‌بینی به این معنی است که پیش‌بینی‌های صحت‌سنجی به‌درستی تخمین زده شده و مدل

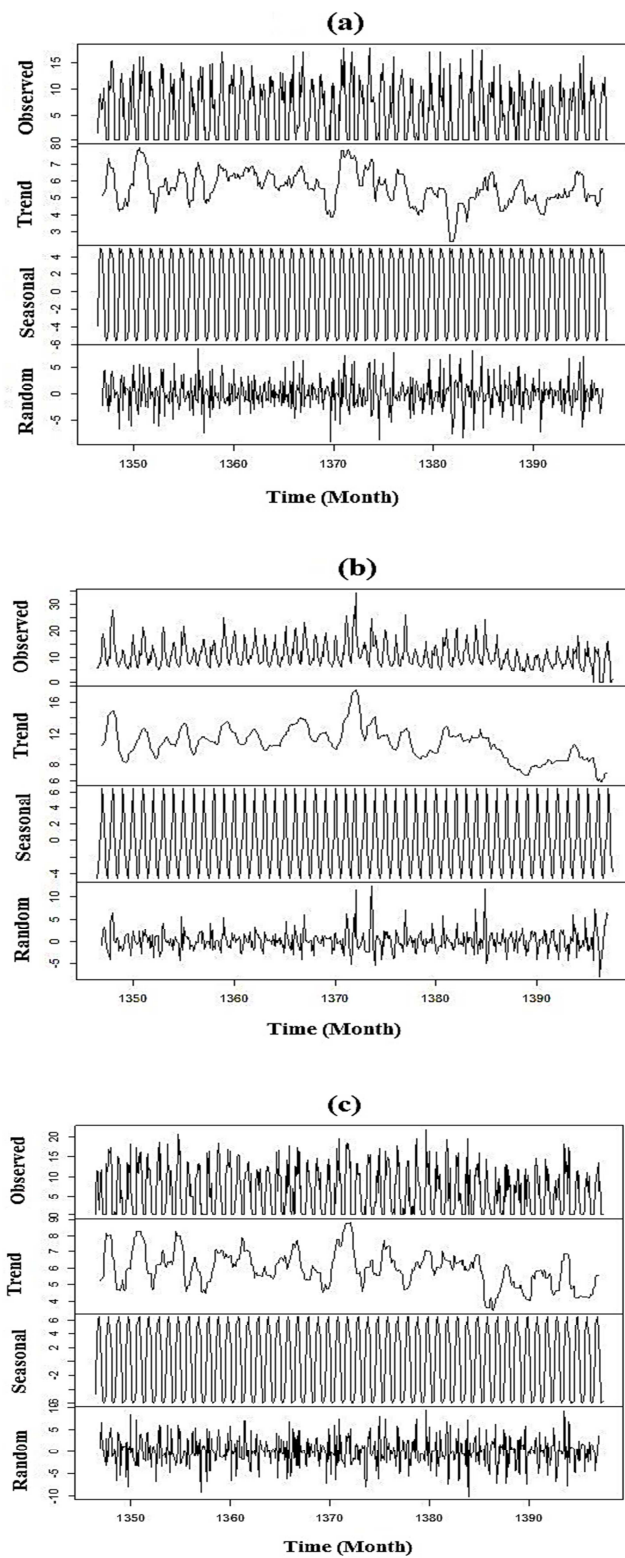


Figure 2. Decomposition of precipitation time series (a) Sepiddasht Sezar Station (b) Tangpanj Bakhtiyari Station (c) Telezang Station with its components



ارزیابی رویکردهای ساختاری فضای حالت نسبت به کلاسیک در پیش‌بینی سری‌زمانی بارش (حوضه آبریز دز)

دقت مدل به‌حد مورد انتظار رسید و مدل کالیبره و هم‌خوانی لازم ایجاد شد. بعد از انجام واسنجی برای دوسال دیگر مدل اعتبارسنجی شد. برای مقایسه مقادیر شبیه‌سازی با مقادیر مشاهداتی ثبت‌شده در ایستگاه‌ها از شاخص‌های معیار ارزیابی مدل به‌ترتیب MAE (رابطه ۱۲)، MSE (رابطه ۱۳) و RMSE (رابطه ۱۴) استفاده شده است. این نتایج در ایستگاه‌های مختلف بارش حوضه آبریز دز در شکل‌های (۳) و (۴) و نیز نتایج به‌صورت کمی در جدول (۲) آورده شده است.

اوج و فرودهای منظم که هر ۱۲ ماه یک‌بار اتفاق می‌افتد به‌دلیل فصلی‌بودن است. این بخش مقایسه‌ای از عملکرد واسنجی و صحت‌سنجی برای مدل‌های Kalman Filter، SARIMA، ETS، TBATS، BATS ارائه می‌کند، که برای واسنجی مدل ۷۰ درصد داده‌ها و برای صحت‌سنجی ۳۰ درصد داده‌ها در نظر گرفته شد. هدف از واسنجی، مطلوب‌سازی درصد خطای محاسباتی نسبت به مقادیر مشاهداتی در طول دوره مورد مطالعه می‌باشد. با انجام سعی و خطا و تغییر پارامترهای واسنجی در نهایت

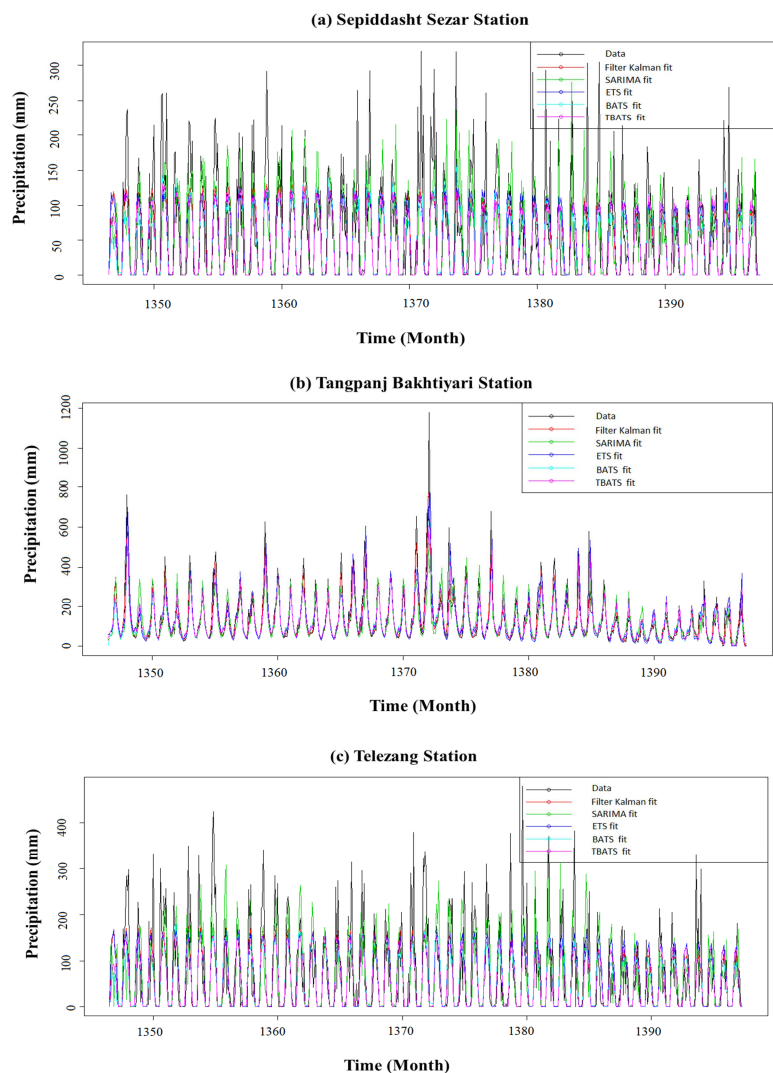


Figure 3. Calibration chart of monthly precipitation time series (1967-2019) (a) Sepiddasht Sezar Station (b) Tangpanj Bakhtiyari Station (c) Telezang Station

مدیریت آب و آبیاری

دوره ۱۲ ■ شماره ۱ ■ بهار ۱۴۰۱

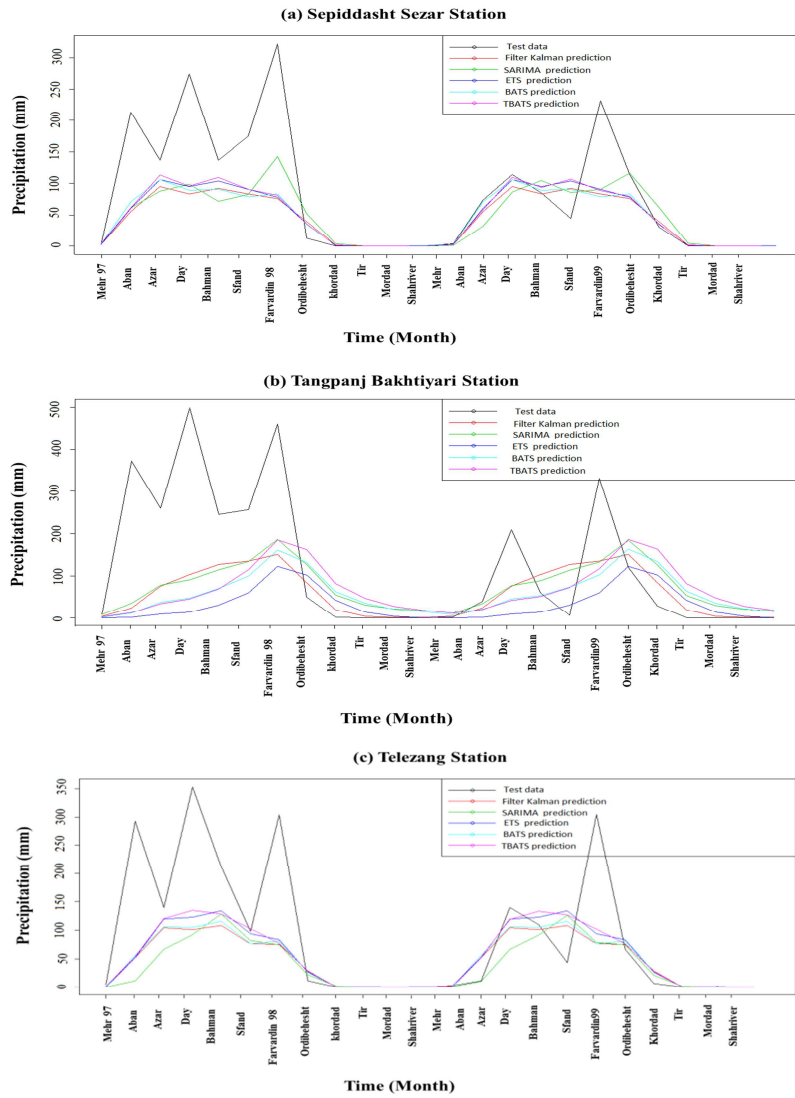


Figure 4. Monthly precipitation validation chart (2019-2021) (a) Sepiddasht Sezar Station (b) Tangpanj Bakhtiyari Station (c) Telezang Station

Table 2. Comparison of calibration errors and validation of different methods with two criteria of MAE, MSE and RMSE in stations (Sepiddasht Sezar, Tangpanj Bakhtiyari and Telezang)

Station	Time series models	Monthly precipitation calibration (1967-2021)			24-month precipitation validation (1967-2021)		
		MAE (mm/month)	MSE (mm/month)	RMSE (mm/month)	MAE (mm/month)	MSE (mm/month)	RMSE (mm/month)
Sepiddasht Sezar	Kalman Filter	29.64	2524.48	50.24	45.23	6709.55	81.91
	ETS	30.80	2644.83	51.42	42.67	6244.85	79.02
	SARIMA	34.76	3447.65	58.71	44.63	5362.09	73.22
	BATS	30.97	2708.83	52.04	42.37	6293.81	79.33
	TBATS	30.38	2628.90	51.27	41.91	6146.21	78.39
Tangpanj Bakhtiyari	Kalman Filter	20.61	1591.91	39.89	89.96	21487.75	146.58
	ETS	41.14	5309.53	72.86	112.04	31876.64	178.54
	SARIMA	41.31	5993.08	77.41	100.84	21636.94	147.09
	BATS	37.42	4918.66	70.13	110.89	26793.60	163.68
	TBATS	37.07	4877.34	69.83	116.12	26968.64	164.22
Telezang	Kalman Filter	36.82	4035.48	63.52	55.18	10296.97	101.47
	ETS	38.12	4201.84	64.82	51.58	9271.09	96.28
	SARIMA	43.46	5440.09	73.75	58.25	11550.25	107.47
	TBATS	37.72	4274.08	65.37	51.21	8999.84	94.86

می‌کند و به ظرفیت محاسباتی کم‌تری نیاز دارد و نتایج نسبتاً دقیقی نسبت به رویکرد کلاسیک از خود نشان می‌دهد. ساختار تصادفی بارش باعث می‌شود که هرکدام از یک سری الگو پیروی کند. شناسایی کردن این الگو به مهندسان می‌تواند کمک کند، به‌همین منظور الگوریتم‌ها روزبه‌روز توسعه کم‌تری پیدا می‌کنند. انتخاب یک الگوریتم نسبت به دیگر الگوریتم‌ها تحت تأثیر عوامل مختلفی از جمله نرخ هم‌گرایی تعیین می‌شود. مدل‌های فضای حالت به دلیل این‌که جزء به جزء در مدل‌سازی عمل می‌کنند به‌عنوان مدل‌های جعبه سفید شناخته می‌شوند. با توجه به این‌که مدل‌های فضای حالت نیازی به ایستاکردن ندارند، مؤلفه روند از سری زمانی حذف نمی‌شود، زیرا حذف روند باعث تغییر شدید در میانگین‌های بعد و قبل از روندزایی می‌شود که توجه و تأیید آن عملکرد روش را با ابهام مواجه می‌سازد. در این پژوهش برای بررسی و تحلیل روند سری زمانی از روش‌های آماری پارامتری مدل‌های فضای حالت ETS، Kalman Filter، TBATS، BATS استفاده شده است. نتایج نشان داد که این مدل‌ها نسبت به مدل کلاسیک توانایی کم‌تری در شناسایی سیستم، در مرحله واسنجی و هم‌صحت‌سنجی از خود نشان می‌دهد. با توجه به نتایج به‌دست‌آمده می‌توان پیشنهاد کرد چون عناصر اقلیمی، رفتاری غیرخطی داشته و از تابع توزیع آماری خاصی پیروی نمی‌کنند، بهتر است جهت بررسی روند از مدل‌های غیرخطی فضای حالت مانند EKF و UKF استفاده شود.

### پی‌نوشت‌ها

1. State Space Model
2. Expectation Maximization
3. Measurement equation
4. Equation state
5. Multiple Sources of Error
6. Dummy Variable
7. Single Sources of Error
8. Exponential Smoothing Methods
9. Modified Models

همان‌طورکه در جدول (۲) مشاهده می‌شود، مدل‌های ساختاری فضای حالت به‌خوبی توانسته‌اند نسبت به رویکرد کلاسیک در هیدرولوژی (باکس - جنکینز) در هر دو مرحله واسنجی و هم‌صحت‌سنجی و براساس سه معیار ارزیابی (MAE)، (MSE) و (RMSE) دارای کم‌ترین مقادیر خطا باشند. از نظر مقایسه بین انواع مدل‌های ساختاری فضای حالت در مرحله واسنجی مشاهده شد که مدل فیلتر کالمن در سه ایستگاه نسبت به دیگر مدل‌ها بهتر عمل کرده است، اما در مرحله هم‌صحت‌سنجی تفاوت‌هایی مشاهده شد، که در مرحله هم‌صحت‌سنجی در ایستگاه‌های تله‌زنگ و سپیددشت سزار براساس معیار (MAE) و (RMSE)، مدل TBATS و در ایستگاه تنگ پنج بختیاری فیلتر کالمن مدل برتر شناخته شد. در این راستا، (2021) Yu *et al.*، (2018) Tamatta، (2018) Naim *et al.* و (2021) Neslihanoglu *et al.* به نتایج مشابهی دست یافتند.

### نتیجه‌گیری

پیش‌بینی سری زمانی بارش برای اهدافی نظیر مدیریت حوضه آبریز اعم از سیلاب و خشکسالی دارای اهمیت زیادی است. استفاده از مدل‌های سری‌زمانی با توجه به ماهیت تصادفی داده‌ها، به‌عنوان یکی از روش‌های توصیه‌شده در پیش‌بینی متغیرهای اقلیمی از جمله بارش می‌باشد. تغییر اقلیم باعث شده است که میانگین بارش یا دبی در سال‌های مختلف مقادیر ثابتی نداشته باشد. این در حالی است که هیدرولوژیست‌ها اغلب از رویکرد باکس - جنکینز جهت مدل‌سازی متغیرهای هیدرولوژی استفاده می‌کنند. این رویکرد با فرض ایستایی شروع به مدل‌سازی می‌کند. در یک سری زمانی ایستا، قوانین حاکم بر تغییرات مقادیر، وابسته به زمان نیست. در مقابل چنین رویکردی، رویکرد فضای حالت قرار دارد که به فرضیات فوق نیاز ندارد. فضای حالت از ساختارهای زمانی ذاتی سری استفاده

- Electrical Engineering and Computer Sciences*, 27, 1198-1212.
12. Neslihanoglu, S., Ünal, E., Yozgatlıgil, C. (2021). Performance comparison of filtering methods on modelling and forecasting the total precipitation amount: a case study for Muğla in Turkey. *Journal of Water and Climate Change*, 12.4, 1071-1085.
  13. Naim, I., Mahara, T., Idrisi, A.R. (2018). Effective Short-Term Forecasting for Daily Time Series with Complex Seasonal Patterns, *Procedia Computer Science* 132, 1832-1841.
  14. Ribeiro, M. I. (2000). Introduction to Kalman Filtering: A Set of Two Lectures.
  15. Ribeiro, M. I. (2004). Kalman and Extended Kalman Filters: Concept, Derivation and Properties. Lisboa: Institute for Systems and Robotics.
  16. Soltani, S., Modarres, R. & Eslamian, S. S. (2007). The use of time series modelling for the determination of rainfall climates of Iran. *International Journal of Climatology*, 27, 819-829.
  17. Sun, M., Li, X. & Kim, G. (2019). Precipitation analysis and forecasting using singular spectrum analysis with artificial neural networks. *Cluster Computing*, 22, 12633-12640.
  18. Sadeghi, M., Asanjan, A. A., Faridzad, M., Nguyen, P., Hsu, K., Sorooshian, S. & Braithwaite, D. (2019). PERSIANN-CNN: Precipitation estimation from remotely sensed information using artificial neural networks—convolutional neural networks. *Journal of Hydrometeorology*, 20 (12), 2273-2289.
  19. Shumway, R. H., & Stoffer, D. S. (2016). State Space Models. In *Time Series Analysis and Its Applications with R Examples*. 287-295. New York: Springer.
  20. Soumik, R., Soumitra, S. D., Pradeep, M., & Khatib, A.M.G.A. (2021). Time Series SARIMA Modelling and Forecasting of Monthly Rainfall and Temperature in the South Asian Countries. *Earth Systems and Environment*, 5, 531-546.
  21. Toth, E., Brath, A. & Montanari, A. (2000). Comparison of shortterm rainfall prediction models for real-time flood forecasting. *Journal of Hydrology*, 239, 132-147.
  22. Tamatta, R. K., (2018). *Time series forecasting of hospital Inpatients and Day case waiting list using ARIMA, TBATS and Neural Network Models*.
  23. Willmott, C.J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *CLIMATE RESEARCH Clim Res*, 30, 79-82.
  10. Local level
  11. Long-run trend
  12. Short-run trend

## تعارض منافع

هیچ گونه تعارض منافع توسط نویسندگان وجود ندارد.

## منابع

1. Asemota, O. J., Bamanga, M. A. & Alaribe, O. J. (2016). Modelling seasonal behavior of rainfall in northeast Nigeria. A state space approach. *International Journal of Statistics and Applications*, 6 (4), 203-222.
2. Brath, A., Montanari, A. & Toth, E. (2002). Neural networks and nonparametric methods for improving real-time flood forecasting through conceptual hydrological models. *Hydrology and Earth System Sciences Discussions*, 6(4), 627-639.
3. Box, G. & Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
4. Durbin, J. & Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University 343.
5. Gabriel, A.C. (2021). A SARIMA and Adjusted SARIMA Models in a Seasonal Nonstationary Time Series; Evidence of Enugu Monthly Rainfall. *European Journal of Mathematics and Statistics*, 2 (1), 13-18.
6. Harting, C. (2010). Rainfall as an Energy Source. Available from: <http://large.stanford.edu/courses/2010/ph240/harting2/>.
7. Hyndman RJ, Khandakar Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 27 (3), 1-22.
8. Hyndman, R. J., Koehler, A. B., Ord, J. K. & Snyder, R.D. (2008). *Forecasting with Exponential Smoothing: The State Space Approach*. Springer-Verlag, Berlin.
9. Little, R. J. A. & Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. John Wiley & Sons, New York, NY.
10. Livera, Alysha M., Rob J. Hyndman, and Ralph D. Snyder. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496), 1513-1527.
11. Masazade, E., Bakır, A. K. & Kırcı, P. (2019). A Kalman filter application for rainfall estimation using radar reflectivity. *Turkish Journal of*

24. Yozgatligil, C., Aslan, S., Iyigun, C. & Batmaz, I. (2013). Comparison of missing value imputation methods for Turkish meteorological time series data. *Theoretical and Applied Climatology*, 112, 143-167.
25. Yu, C., Xu, C., Li, Y., Yao, S., Bai, Y., Li, J., Wang, L., Wu, W., & Wang, Y., (2021). Time Series Analysis and Forecasting of the Hand-Foot-Mouth Disease Morbidity in China Using an Advanced Exponential Smoothing State Space TBATS Model. *Infection and Drug Resistance*, 14, 2809-2821.
26. Zulfi, M., Hasan, M. & Purnomo, K. D. (2018). The development rainfall forecasting using Kalman filter. *Journal of Physics: Conference Series*, 1008 (1), 012006.
27. Zeng, Q., Li, D., Huang, G., Xia, J., Wang, X., Zhang, Y., Tang, W., & Zhou, H. (2016). Time series analysis of temporal trends in the pertussis incidence in Mainland China from 2005 to 2016. *Scientific Reports*, 6, 32367.