

Original Article

Iranian Remote Sensing GIS
Journal of

Evaluation of the Performance of the Singular Spectrum Analysis (SSA) Algorithm in Reconstructing Missing Data with Different Intensities in the Hourly Land Surface Temperature Time Series

Hadi Zare Khormizi¹, Mohammad Jafari^{1*}, Hamidreza Ghafarian Malamiri², Ali Tavili¹, Hamidreza Keshtkar¹

Affiliation

1. Dep. of Arid and Mountainous Regions Reclamation, Faculty of Natural Resources, University of Tehran, Karaj, Iran
2. Dep. of Geography, Dep. of Environmental Planning, Yazd University, Yazd, Iran

ABSTRACT

Introduction and Purpose: Generating Land Surface Temperature (LST) data with temporal and spatial continuity is in great demand for hydrology, meteorology, ecology, environment, and, etc. studies. Approximately, 60 to 75 percent of the Earth is covered by clouds at any given moment. Therefore, clouds, by creating an obstacle, absorb part of the thermal energy emitted from the earth by affecting thermal infrared energy, creating gaps and outliers in LST time series data. Removing the effect of cloud cover is always a common problem in the field of using satellite images. The purpose of this research is to evaluate the performance of Multi-channel Singular Spectrum Analysis (M-SSA) in order to reconstruct gaps and remove outlier data due to the cloud coverage in the hourly LST time series of the Meteosat-9 satellite.

Materials and Methods: The study area in the present research was whole Iran. Also, the hourly LST time series of the SEVIRI sensor from the Meteosat-9 geostationary satellite in 2022 was used. At first, using SSA software and the Monte Carlo test, the window size and the number of significant components of an hourly LST time series were determined. Then, using the identified significant components, LST time series were reconstructed using M-SSA algorithm. Reconstruction error in clear sky conditions with available time series data and reconstruction error in cloudy sky conditions by creating artificial missing data (artificial cloud) with intensities of 10, 20, 30, ..., 90% in time series were evaluated using root mean square error (RMSE) and coefficient of determination (R^2) statistics.

Results: On average, in Iran, 25.5% of the hourly LST time series in 2022 was lost due to cloud cover, and the highest percentage of lost data was observed at the edge of the Caspian Sea. The results of analyzing the annual hourly LST time series in a window size of 96 hours with the Monte Carlo test showed that components 1 to 5 are significant components of this time series. These components control 97.5% of the LST time series variance. The frequency of the first, second-third, and fourth-fifth components are respectively 0, 0.042 and 0.083 cycles per image. The first component indicates annual periodic changes, the second and third components indicate 24-hour or daily temperature changes, and the fourth and fifth components indicate 12-hour periodic temperature changes. Based on the results, the RMSE and the R^2 between the original and the reconstructed data in clear sky conditions were 1.38 and 0.99 Kelvin, respectively. Also, in cloudy sky conditions, the RMSE error up to the level of 80% of randomly lost data (artificial cloud) was always less than 2.1 Kelvin.

Discussion and Conclusion: The main key to reconstructing time series with periodic behavior is to identify significant periodic components and trends. In hourly LST time series, annual, 24-, 12- and 8-hour periods are the most important components of the time series. These components are formed due to the rotation of the earth around itself and the sun and the deviation of its axis. Therefore, these components are generally the same for the reconstruction of hourly LST time series in the major part of the globe. Based on the findings, M-SSA algorithm can be effective in reconstructing lost data with large distance in LST time series due to consideration of periodic components and trends as well as using temporal and spatial correlation. One of the significant cases in reconstructing the effect of cloud cover in the present study and many other studies is the reconstruction of LST with the clear sky condition. Therefore, reconstruction of LST under cloud cover can be a challenge and suggestion for further studies in the future.

Keywords: Cloud cover, Empirical Orthogonal Functions, Singular Spectrum Analysis, Temporal-spatial interpolation, Time series.

Citation:

Zare Khormizi, H., Jafari, M., Ghafarian Malamiri, H., Tavili, A., Keshtkar, H., Evaluation of the Performance of the Singular Spectrum Analysis (SSA) Algorithm in Reconstructing Missing Data with Different Intensities in the Hourly Land Surface Temperature Time Series, *Iran J Remote Sens GIS*. 17(4): 39-58.

* Corresponding Author: jafari@ut.ac.ir
DOI: <https://doi.org/10.48308/gisj.2024.235694.1219>

Received: 2024.05.15
Accepted: 2024.09.07





ارزیابی عملکرد الگوریتم تحلیل طیفی منفرد در بازسازی داده‌های از دست رفته با شدت‌های متفاوت در سری زمانی ساعتی دمای سطح زمین

هادی زارع خورمیزی^۱، محمد جعفری^{۱*}، حمیدرضا غفاریان مالگیری^۲، علی طویلی^۱، حمیدرضا کشت کار^۱

چکیده

مقدمه و هدف: تولید داده‌های دمای سطح زمین (LST) با پیوستگی زمانی و مکانی، برای مطالعات هیدرولوژی، هواشناسی، اکولوژی، زیست‌محیطی و دیگر موارد، تقاضای چشمگیری دارد. تقریباً ۶۰ تا ۷۵٪ کره زمین در هر لحظه با ابرها پوشیده می‌شود؛ بنابراین ابرها، با ایجاد مانعی، باعث جذب بخشی از انرژی حرارتی بازتاب یافته از زمین می‌شوند و با تأثیر در انرژی فروسرخ حرارتی، سبب ایجاد داده‌های از دست رفته و نامعتبر در سری‌های زمانی LST می‌شوند. رفع اثر پوشش ابر همواره مسئله‌ای رایج در زمینه استفاده از تصاویر ماهواره‌ای است. هدف از پژوهش حاضر ارزیابی عملکرد الگوریتم تحلیل طیفی منفرد چندکاناله (M-SSA)، به منظور بازسازی داده‌های از دست رفته و دورافتاده با استفاده از پوشش ابر در سری زمانی LST ساعتی ماهواره Meteosat-9 است.

مواد و روش‌ها: منطقه مطالعاتی، در پژوهش حاضر، کشور ایران در نظر گرفته شد و نیز از سری زمانی LST ساعتی سنجنده SEVIRI از ماهواره زمین ثابت Meteosat-9 در سال ۲۰۲۲ استفاده شد. در ابتدا، با استفاده از نرم‌افزار SSA و آزمون مونت کارلو، اندازه پنجره و تعداد مؤلفه‌های معنی‌دار یک سری زمانی LST ساعتی تعیین شد. سپس با استفاده از مؤلفه‌های معنی‌دار شناسایی شده و نیز الگوریتم M-SSA، سری‌های زمانی LST بازسازی شد. خطای بازسازی در شرایط آسمان صاف، با داده‌های موجود سری زمانی و خطای بازسازی در شرایط آسمان ابری، از طریق ایجاد داده‌های از دست رفته مصنوعی (ابر مصنوعی) با شدت‌های ۱۰، ۲۰، ۳۰، ...، ۹۰٪ در سری زمانی، با استفاده از آماره‌های خطای جذر میانگین مربعات (RMSE) و ضریب تبیین (R^2) اندازه‌گیری شد.

نتایج: به‌طور میانگین، در سطح ایران، ۲۵/۵٪ از سری زمانی LST ساعتی در سال ۲۰۲۲ به‌دلیل پوشش ابر از دست رفته است که بیشترین میزان داده‌های از دست رفته در حاشیه دریای خزر مشاهده شد. نتایج تحلیل سری زمانی LST ساعتی سالیانه در اندازه پنجره ۹۶ ساعته با آزمون مونت کارلو نشان داد مؤلفه‌های ۱ تا ۵ جزء مؤلفه‌های معنی‌دار این سری زمانی‌اند. این مؤلفه‌ها ۹۷/۵٪ از تغییرات سری زمانی LST را در کنترل دارند. فرکانس مربوط به مؤلفه‌های اول، دوم با سوم، چهارم با پنجم به‌ترتیب ۰/۰۴۲، ۰/۰۸۳ و سیکل بر تصویر است. مؤلفه اول تغییرات دوره‌ای سالیانه، مؤلفه دوم با سوم تغییرات دوره‌ای ۲۴ ساعته یا روزانه دما، و مؤلفه چهارم با پنجم تغییرات دوره‌ای ۱۲ ساعته دما را نشان می‌دهد. براساس نتایج، RMSE و R^2 بین داده‌های موجود و داده‌های بازسازی شده سری زمانی LST ساعتی با استفاده از الگوریتم M-SSA، در شرایط آسمان صاف، به‌ترتیب ۱/۳۸ کلون و ۰/۹۹ بود. همچنین در شرایط آسمان ابری، خطای RMSE تا سطح ۸۰٪ داده از دست رفته به‌صورت تصادفی (ابر مصنوعی) همواره کمتر از ۲/۱ کلون بود.

بحث و نتیجه‌گیری: کلید اصلی بازسازی سری‌های زمانی، با رفتار دوره‌ای، شناسایی مؤلفه‌های معنی‌دار دوره‌ای و روندهاست. در سری‌های زمانی LST ساعتی، دوره‌های سالیانه، ۲۴، ۱۲ و ۸ ساعته از مهم‌ترین مؤلفه‌های سری زمانی‌اند. این مؤلفه‌ها از چرخش زمین به دور خود و خورشید و انحراف محور آن شکل می‌گیرند. بنابراین، این مؤلفه‌ها اغلب برای بازسازی سری‌های زمانی LST ساعتی در بیشتر بخش‌های کره زمین، یکسان‌اند. براساس یافته‌ها، الگوریتم M-SSA با توجه به در نظر گرفتن مؤلفه‌های دوره‌ای و روندها و همچنین استفاده از همبستگی زمانی و مکانی، می‌تواند در بازسازی داده‌های از دست رفته دارای فاصله بزرگ در سری‌های زمانی LST مؤثر باشد. یکی از موارد شایان توجه در بازسازی اثر پوشش ابر، در پژوهش حاضر و بسیاری از پژوهش‌های دیگر، بازسازی دمای سطح زمین با شرط آسمان صاف است. بنابراین بازسازی دمای سطح زمین زیر پوشش ابر می‌تواند چالش و پیشنهادی برای مطالعات بیشتر در آینده باشد.

واژه‌های کلیدی: پوشش ابر، تحلیل طیفی منفرد، توابع متعامد تجربی، درون‌یابی زمانی-مکانی، سری زمانی.

سمت

۱. گروه احیای مناطق خشک و کوهستانی، دانشکده منابع طبیعی، دانشگاه تهران، کرج، ایران
۲. گروه جغرافیا، بخش برنامه‌ریزی محیطی، دانشگاه یزد، یزد، ایران

استناد:

زارع خورمیزی، ه.، جعفری، م.، غفاریان مالگیری، ح.ر.، طویلی، ع.، کشت کار، ح.ر.، ارزیابی عملکرد الگوریتم تحلیل طیفی منفرد در بازسازی داده‌های از دست رفته با شدت‌های متفاوت در سری زمانی ساعتی دمای سطح زمین، نشریه سنجش از دور و GIS ایران، سال ۱۷، شماره ۴، زمستان ۱۴۰۴، ۵۸-۳۹.



۱- مقدمه

دمای سطح زمین (LST)^۱ نقش مهمی در فرایندهای فیزیکی و بیولوژیکی سطح زمین، در مقیاس جهانی و منطقه‌ای، ایفا می‌کند (Yu et al., 2022). این پارامتر ارتباط نزدیکی با تغییرات در متغیرهای کلیدی اکوسیستم‌های زمین، مانند محتوای بخار آب، رطوبت خاک، وضعیت تبخیر و تعرق، و فرایندهای فنولوژی دارد (Yu et al., 2022). بنابراین دمای سطح زمین در مطالعات اکولوژی، زیست‌محیطی، هیدرولوژی، هواشناسی، تولیدات کشاورزی و دیگر موارد، کاربردی گسترده دارد (Helman et al., 2017; Zhou et al., 2017; Zare khormizi et al., 2021; Wang et al., 2021; Yu et al., 2022). یکی از روش‌های اندازه‌گیری دمای سطح زمین اندازه‌گیری‌های زمینی و میدانی است. این روش داده‌های پیوسته و مورد اعتمادی را در مکان اندازه‌گیری تولید می‌کند. با این حال، در مورد مناطق وسیع و ناهمگن دارای توزیع نامناسب ایستگاه‌های اندازه‌گیری، نمی‌توان به این داده‌ها اعتماد کرد (Shiff et al., 2021). از طرفی، تهیه و تولید این داده‌ها هزینه‌های مالی زیادی دربردارد. یکی از روش‌های نوین، برای دستیابی به دمای سطح زمین، استفاده از فناوری سنجش از دور ماهواره‌ای است. این فناوری دمای سطح زمین را با کمک سنجنده‌های حرارتی، با وضوح زمانی و مکانی مورد قبول در سراسر جهان، به دست می‌دهد. در زمینه استفاده از فناوری سنجش از دور برای محاسبه دمای سطح زمین، همواره ممکن است عواملی از قبیل ذرات معلق در هوا، گردوغبارهای جو و حضور ابرها در انرژی بازتابیده شده از سطح تأثیری شایان توجه بگذارند و به دنبال آن، قرائت سنجنده‌های حرارتی را با خطا مواجه کنند (Julien & Sobrino, 2010). تقریباً ۶۰ تا ۷۵٪ کره زمین، در هر لحظه، زیر پوشش ابرها قرار می‌گیرد (Kleipool et al., 2008). ابرها، با ایجاد مانع، باعث جذب بخشی از انرژی حرارتی بازتاب یافته از زمین می‌شوند و با تأثیر در انرژی فروسرخ حرارتی،

سبب ایجاد دمای پایین در تصاویر دمای سطح زمین می‌شوند (Ghafarian Malamiri et al., 2018). در مراحل بعد ممکن است تأثیرات پوشش ابر در تصاویر دمای سطح زمین، با استفاده از الگوریتم‌های ماسک ابر^۲، شناسایی و حذف شود؛ در نتیجه، در تصاویر سری‌های زمانی دمای سطح زمین، داده‌های از دست‌رفته^۳ و داده‌های با مقادیر نامرتب^۴ حضور خواهند داشت. این مشکل باعث می‌شود روش‌های سنجش از دور، با توجه به اثر پوشش ابر، همواره در هر زمان و مکان قادر به ارائه دمای سطح زمین نباشند. بیش از بیست روش، به منظور رفع اثر پوشش ابر و بازیابی داده‌های از دست‌رفته و دورافتاده در سری‌های زمانی تصاویر ماهواره‌ای (شاخص‌های گیاهی یا دمای سطح زمین)، وجود دارد (Geng et al., 2014). به‌طور کلی، این روش‌ها را می‌توان در سه دسته زمانی، مکانی و در نهایت، زمانی-مکانی قرار داد (Chen et al., 2022). در روش‌های زمانی، مقادیر داده‌های از دست‌رفته و نامعتبر براساس حرکت در زمان تخمین زده می‌شود و مقادیر پیکسل‌های مجاور جغرافیایی در نظر گرفته نمی‌شود. برای نمونه، الگوریتم تجزیه و تحلیل هارمونیک سری‌های زمانی (HANTS)^۵ براساس رویکرد زمانی، برای بازسازی سری‌های زمانی LST و شاخص NDVI^۶، به کار رفته است (Verhoef et al., 1996; Xu & Shen, 2013; Zhou et al., 2015). روش‌های مکانی فقط مقادیر پیکسل‌های مجاور را در نظر می‌گیرند و مقادیر در دوره‌های زمانی متفاوت را شامل نمی‌شوند (Pede & Mountrakis, 2018). از بین روش‌های مکانی، می‌توان روش‌های زمین آماری مانند کوکریجینگ (Zhang et al., 2009; Ke et al., 2013) یا رگرسیون‌های خطی یا چندجمله‌ای در مکان (Mukherjee et al., 2014) را نام

1. Land Surface Temperature

2. Cloud Mask Algorithm

3. Gap

4. Outlier

5. Harmonic Analysis of Time Series

6. Normalized Difference Vegetation Index

برد. روش‌های مکانی به راحتی تحقق پذیرند اما کاستی‌های خاصی، مانند نداشتن وضوح و دقت تصاویر بازسازی شده، دارند (Cui et al., 2022). در روش‌های زمانی-مکانی، هم تغییرات زمانی و هم تغییرات مکانی در بازسازی داده‌های ابری و از دست رفته استفاده می‌شوند. از روش‌های درون‌یابی زمانی-مکانی، برای بازسازی مقادیر داده‌های از دست رفته و نامعتبر در سری‌های زمانی، می‌توان به روش تحلیل طیفی منفرد چندکاناله (M-SSA)^۱ اشاره کرد (Ghafarian Malamiri et al., 2018).

یکی از استراتژی‌های غالب در پر کردن فاصله داده‌های از دست رفته در روش‌های زمانی و زمانی-مکانی، شناسایی رفتار دوره‌ای و روندها و پارامترهای تغییر در سری‌های زمانی و سپس به کارگیری این پارامترها برای بازسازی سری زمانی است (Cui et al., 2022). سری‌های زمانی LST دارای اجزای منظم (چرخه‌ای)^۲ و اجزای نامنظم (نویز) هستند. اجزا یا مؤلفه‌های منظم از حرکت زمین به دور خود و خورشید ناشی می‌شود و تغییرات فصلی از انحراف محور کره زمین و اجزای نامنظم یا نویز نیز از فعل‌وانفعال تعامل میان انرژی تابشی و انرژی آشفته محیط ایجاد می‌شود (Ghafarian Malamiri et al., 2018). الگوریتم تحلیل طیفی منفرد (SSA) از همبستگی زمانی (تک‌کاناله: SSA) و همبستگی زمانی-مکانی (چندکاناله) برای بازسازی سری‌های زمانی استفاده می‌کند. این الگوریتم یک سری زمانی را به مؤلفه‌های ساده‌تر و تفسیرپذیر، مانند روندها^۳، نوسانات دوره‌ای و نیمه‌دوره‌ای، و نویزها تجزیه می‌کند (Golyandina et al., 2001).

در پژوهش‌هایی، دقت و کارایی روش‌های گوناگون برای رفع داده‌های از دست رفته و نامعتبر از طریق پوشش ابر در سری‌های زمانی LST بررسی شد و پژوهشگرانی هم دقت روش‌های زمانی، مکانی و زمانی-مکانی در بازسازی اثر پوشش ابر در سری‌های زمانی LST سنجنده مادیس (MODIS)^۴ در حوضه رودخانه هیله^۵، چین، را بررسی کردند. براساس نتایج این

مطالعه، روش درون‌یابی زمانی خطی بیشترین دقت را با خطای جذر میانگین مربعات (RMSE)^۶ ۷/۱ کلوین داشت (Chen et al., 2022). در مطالعه‌ای دیگر، از الگوریتم HANTS برای بازسازی و پر کردن فاصله داده‌های از دست رفته و نامعتبر در سری زمانی هشت‌روزه LST سنجنده مادیس استفاده شد. نتایج این مطالعه نشان داد الگوریتم HANTS می‌تواند کیفیت داده‌های دمای سطح زمین سنجنده مادیس را بهبود مؤثری ببخشد (Xu & Shen, 2013). در مطالعه‌ای دیگر، داده‌های از دست رفته و نامعتبر از طریق پوشش ابر، در محصول LST سنجنده مادیس و با استفاده از داده‌های آسمان صاف پیکسل‌های مجاور در زمان، بازسازی شد (Sun et al., 2017). طی پژوهشی، روش درون‌یابی داده‌ها از طریق توابع متعامد تجربی (DINEOF)^۷ برای بازسازی و پر کردن فاصله داده‌های از دست رفته و نامعتبر در سری‌های زمانی LST سنجنده مادیس، در فلات تبت، به کار رفت. در این پژوهش، خطای بازسازی داده‌های از دست رفته در تصاویر LST، طی روز و شب، به ترتیب ۴/۵ و ۲/۳ کلوین بود (Zhou et al., 2017). در پژوهشی دیگر، با استفاده از الگوریتم M-SSA، سری زمانی داده‌های ساعتی دمای سطح زمین ماهواره زمین ثابت Fengyun-2C در فلات تبت چین بازسازی شد. نتایج این مطالعه نشان داد، در یک سری زمانی ساعتی با میانگین ۶۳٪ داده از دست رفته با پوشش ابر، خطای میانگین مطلق (MAE)^۸ ۲/۲۵ درجه کلوین است (Ghafarian Malamiri et al., 2012). در بازسازی سری‌های زمانی روزانه LST سنجنده مادیس با استفاده از الگوریتم HANTS، در منطقه مطالعاتی ایران،

1. Multi-Singular Spectrum Analysis
2. Cycles
3. Trend
4. Moderate Resolution Imaging Spectroradiometer
5. Heihe River
6. Root Mean Square Error
7. Data Interpolating Empirical Orthogonal Functions (DINEOF)
8. Mean Absolute Error

۲- مواد و روش‌ها

۲-۱- معرفی منطقه مورد مطالعه

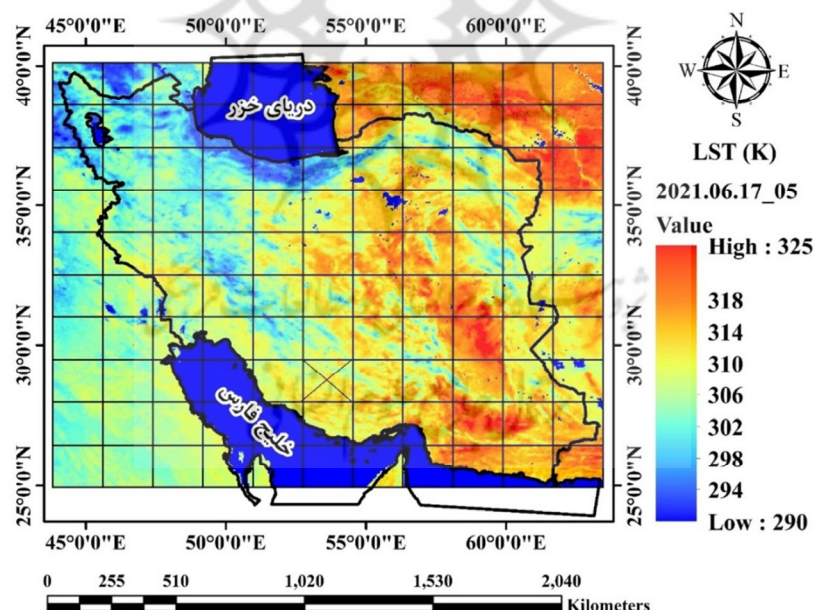
منطقه مطالعاتی، در پژوهش حاضر، کل سطح کشور ایران در نظر گرفته شد. علت این انتخاب داشتن منطقه‌ای وسیع، با میزان داده‌های ازدست‌رفته دارای شدت‌های متفاوت در سری‌های زمانی LST ساعتی بوده است. شکل ۱ نمونه‌ای از یک تصویر LST در تاریخ ۲۷ خرداد ۱۴۰۱ (۲۰۲۲/۰۶/۱۷)، در ساعت ۵ صبح را نشان می‌دهد.

۲-۲- تصاویر ماهواره‌ای

در پژوهش حاضر، از محصول دمای سطح زمین ماهواره زمین ثابت^۱ Meteosat-9 استفاده شد^۲. این محصول دارای قدرت تفکیک مکانی پنج کیلومتر و قدرت

مشخص شد خطای جذر میانگین مربعات بین داده‌های غیرابری و بازسازی‌شده در سری‌های زمانی LST، طی روز و شب، به‌ترتیب ۳/۸۷ و ۲/۶۸ کلوین است (Ghafarian Malamiri & Zare khormizi, 2017).

هدف از پژوهش حاضر ارزیابی الگوریتم M-SSA، به‌منظور بازسازی داده‌های ازدست‌رفته و دورافتاده برآثر پوشش ابر در سری زمانی LST ساعتی ماهواره زمین ثابت Meteosat-9 است. با توجه به اینکه تولید داده‌های دمای سطح زمین با پیوستگی زمانی و مکانی برای مطالعات هیدرولوژی، هواشناسی، اکولوژی، مطالعات زیست‌محیطی و دیگر موارد تقاضای چشمگیری دارد، نتایج این پژوهش می‌تواند در اغلب این مطالعات، با توجه به داده‌های پیوسته و خالص تولیدشده، کاربرد داشته باشد.



شکل ۱. منطقه مورد مطالعه به‌همراه تصویری از دمای سطح زمین در ساعت ۵ صبح، به‌تاریخ ۱۴۰۱/۰۳/۲۷، به‌علاوه شبکه ایجادشده برای پردازش جزءبه‌جزء سری زمانی

1. Indian Ocean Data Coverage (IODC)
2. <https://land.copernicus.eu/global/products/lst>

تفکیک زمانی ساعتی است. مجموعه ماهواره‌های Meteosat، از جمله ماهواره‌های هواشناسی، با هدف پیش‌بینی و تشخیص سریع آب‌وهوا طی چند ساعت آینده‌اند و ماهواره 9-Meteosat از ماهواره‌های نسل دوم Meteosat است.^۱ سنجنده SEVIRI^۲ ماهواره 9-Meteosat در دوازده باند طیفی و هر پانزده دقیقه (در حالت دید کامل^۳) تصویربرداری می‌کند. همچنین تصاویر دمای سطح زمین دریافتی از این سنجنده با استفاده از الگوریتم پنجره مجزای دوکاناله^۴ به دست می‌آید (Freitas et al., 2013; Martins et al., 2023).

۳-۲- اجرای الگوریتم تجزیه و تحلیل طیفی منفرد

به منظور بازسازی سری زمانی یک‌ساله LST ساعتی با ۸۷۶۰ تصویر (۳۶۵ روز × ۲۴ ساعت) در طول سال ۲۰۲۲، از نرم‌افزار SSA-MTM استفاده شد. دسترسی به این نرم‌افزار، از لینک پیوست^۵، به رایگان امکان‌پذیر است. برای بازسازی سری‌های زمانی در نرم‌افزار SSA، ابتدا باید اندازه پنجره^۶ و تعداد مؤلفه‌های معنی‌دار^۷ تعیین شود. به منظور تعیین این پارامترها، نخست، مجموعه‌ای زمانی تکی یا برداری بدون داده از دست‌رفته، به منزله نماینده‌ای برای کل سری زمانی انتخاب شد. سپس اندازه پنجره ۹۶ ساعت در نظر گرفته شد. در مرحله بعد با استفاده از آزمون‌های گوناگون در نرم‌افزار SSA-MTM، مانند آزمون مونت کارلو^۸، تعداد مؤلفه‌های معنی‌دار در پنجره زمانی ۹۶ ساعته مشخص شد؛ در بخش نتایج، مفصلاً به این مورد پرداخته خواهد شد. همچنین باید اشاره کنیم که با توجه به طولانی بودن بحث آزمون‌های نرم‌افزار SSA، نظیر آزمون مونت کارلو و آزمون فرض صفر EOFs^۹ و بحث نظریه درباره روندها و دوره‌ها، برای کسب اطلاعات بیشتر می‌توان به پژوهش غفاریان ملامیری^{۱۰} (۲۰۱۵) مراجعه کرد. اندازه پنجره، در پژوهش حاضر، ۹۶ ساعت (چهار روز) در نظر گرفته شد زیرا کوچک در نظر گرفتن اندازه پنجره، در مناطق با داده‌های از دست‌رفته طی فاصله طولانی، به ایجاد سری زمانی

نامعتبر منجر خواهد شد و از سویی، هنگامی که اندازه پنجره بزرگ در نظر گرفته شود، منابع حافظه موقت و پردازنده رایانه به شدت درگیر خواهد شد و مدت زمان بازسازی افزایش خواهد یافت.

در نرم‌افزار SSA-MTM، سری‌های زمانی را می‌توان به صورت برداری و ماتریسی پردازش کرد. هنگام پردازش یک سری زمانی تکی، از الگوریتم SSA استفاده می‌شود و اگر تعداد n سری زمانی در کنار هم، به صورت ماتریس قرار داشته باشند، در حالت M-SSA پردازش می‌شوند.

پس از تعیین اندازه پنجره و مؤلفه‌های معنی‌دار، سری‌های زمانی LST ساعتی سالیانه با استفاده از M-SSA بازسازی شد. شایان ذکر است که با توجه به وسعت منطقه مورد مطالعه و اندازه پیکسل‌سایز تصاویر LST (پنج کیلومتر)، این منطقه ۳۴۰ پیکسل در ردیف و ۴۴۰ پیکسل در ستون دارد و طبق داده‌های LST ساعتی، سالیانه ۸۷۶۰ تصویر دمای سطح زمین دریافت خواهد شد. پردازش یک‌جای این سری زمانی، با استفاده از الگوریتم M-SSA، در رایانه‌های معمولی امکان‌پذیر نیست. بنابراین در این مطالعه، با توجه به طرح شبکه‌بندی نشان داده‌شده در شکل ۱، این سری زمانی به قطعات کوچک‌تر تقسیم و هر بخش جداگانه پردازش شد. در ادامه، مبانی نظری الگوریتم SSA توضیح داده می‌شود.

۴-۲- مبانی نظری الگوریتم تحلیل طیفی منفرد

تحلیل طیفی منفرد را ابتدا برومهد و کینگ^{۱۱} (۱۹۸۶) توسعه دادند و سپس ووتارد و گیل^{۱۲} (۱۹۹۲) و

1. Meteosat Second Generation (MSG)
2. Spinning Enhanced Visible and InfraRed Imager (SEVIRI)
3. Full View
4. Generalized Split-Window Algorithm from Two Adjacent Channels
5. <http://research.atmos.ucla.edu/tcd/ssa/form.html>
6. Window Size
7. Significant Components
8. Monte Carlo Test
9. Empirical Orthogonal Functions
10. Ghafarian Malamiri
11. Broomhead & King
12. Vautard & Ghil

هر بردار ویژه مورد نظر، توصیف‌های مثبتی از بهترین نمایش ماتریس X به‌منزله مجموعه ماتریس‌های $X_i, i=1, 2, \dots, d$ است (Golyandina et al., 2001).

۳. در مرحله سوم، d به تعداد p زیرمجموعه متمایز تقسیم می‌شود؛ در نتیجه، جمع همه مؤلفه‌ها در داخل هر زیرمجموعه مطابق رابطه (۳) است.

$$X = \sum_{n'=1}^p X_{I_{n'}} \rightarrow X_{I_{n'}} = \sum_{i \in I_{n'}} X_i \quad \text{رابطه (۳)}$$

در حالتی ایدئال، ماتریس‌های $X_{I_{n'}}$ به‌صورت یک ماتریس هانکل^۷ هستند و در نتیجه، بر ماتریس‌های تراژکتوری منطبق می‌شوند.

۴. از آنجاکه حالت ایدئال توصیف‌شده در مرحله ۳ معمولاً رخ نمی‌دهد، در این مرحله، ماتریس‌های $X_{I_{n'}}$ باید به‌صورت ماتریس هانکل، برای انطباق بر ماتریس‌های تراژکتوری تبدیل شوند. این مرحله با عنوان میانگین‌گیری قطری شناخته می‌شود؛ بدین مفهوم که ماتریس اصلی می‌تواند به‌صورت مجموع این ماتریس‌ها بازسازی شود (رابطه (۴)).

$$X_t = \sum_{n'=1}^p X_t^{n'}, \quad t = 0, 1, \dots, n-1 \quad \text{رابطه (۴)}$$

در این معادله، برای هر p ، سری $X_t^{n'}$ نتیجه میانگین قطری ماتریس $X_{I_{n'}}$ است.

SSA چندکاناله زمانی و مکانی استفاده می‌شود که بیش از یک سری زمانی وجود داشته باشد. M-SSA تعداد L پروفایل مکانی زمان را (که $L \leq m$ ؛ اندازه پنجره انتخاب‌شده در SSA است) به کار می‌گیرد و از اطلاعات مکانی (برخی مؤلفه‌های معنی‌دار مکانی PCA)، همراه با اطلاعات زمانی (T-PCA) برای بازسازی دقیق‌تر سری‌های زمانی استفاده می‌کند.

کاندراشو^۱ و گیل (۲۰۰۶) این الگوریتم را برای پر کردن فاصله داده‌های از دست‌رفته یا گم‌شده در سری‌های زمانی، پیشنهاد کردند. عملکرد الگوریتم تحلیل طیفی منفرد در بازسازی سری‌های زمانی را می‌توان در چهار مرحله توصیف کرد (Ghafarian et al., 2018; Musial et al., 2011).

۱. در مرحله اول، یک سری زمانی اسکالر $F(t)$ ؛ $t=1, 2, 3, \dots, n$ به یک ماتریس تراژکتوری^۲ چندبعدی از بردارهای $X=[f_1, \dots, f_k]$ تبدیل می‌شود که در این ماتریس، $k=n-m+1$ (اندازه پنجره m). هر بردار این ماتریس به این صورت تعریف می‌شود:

$$X_j = (f_j, \dots, f_{j+m-1}); j=1, \dots, k$$

این ماتریس، درواقع، حاوی رکوردهای کامل الگوهای موجود در اندازه پنجره زمانی ($m < n$) است (رابطه (۱)).

$$X = \begin{pmatrix} f_1 & f_2 & f_3 & \dots & f_k \\ f_2 & f_3 & f_4 & \dots & f_{k+1} \\ f_3 & f_4 & f_5 & \dots & f_{k+2} \\ \dots & \dots & \dots & \dots & \dots \\ f_m & f_{m+1} & f_{m+2} & \dots & f_n \end{pmatrix} \quad \text{رابطه (۱)}$$

۲. در مرحله دوم، ماتریس $m \times k$ تراژکتوری X با استفاده از روش SVD^۳ تجزیه می‌شود. محصول این تجزیه به‌صورت رابطه (۲) خواهد بود.

$$X = DLE^T \quad \text{رابطه (۲)}$$

عناصر ماتریس L ، به‌منزله مقادیر منحصربه‌فرد، ریشه‌های مربع مقادیر ویژه^۴ ماتریس $S=XX^T$ با اندازه $m \times m$ هستند. ستون‌های ماتریس D درواقع بردارهای ویژه^۵ ماتریس S هستند که به‌منزله توابع متعامد تجربی (EOFs) شناخته می‌شوند. ردیف‌های ماتریس E^T نیز بردارهای ویژه^۶ ماتریس XX^T محسوب می‌شوند.

اگر مقادیر اعداد تکین^۶ را به‌ترتیب صعودی به نزولی ترسیم کنیم، اغلب نموداری با شیب تند اولیه، نشان‌دهنده یک سیگنال و یک سطح صاف با شیب کم، به دست می‌آید که بیانگر نویزهاست (Vautard et al., 1992). هر زیرمجموعه از بردارهای ویژه^۷ d (EOFs)، $1 \leq d \leq m$ ، برای

1. Kondrashov
2. Trajectory
3. Singular Value Decomposition
4. Eigenvalues
5. Eigenvectors
6. Singular Values
7. Henkel Matrix

شود، RMSE بین داده‌های بازسازی‌شده جدید و مقادیر مشخص داده‌های از دست‌رفته طبق تکرار قبلی محاسبه می‌شود (Kondrashov et al., 2010). سپس فرایند تکرار، برای دومین EOF پیشرو، آغاز می‌شود و تا زمانی که همگرایی برای EOF دوم به دست آید، ادامه می‌یابد. به‌منظور به دست آوردن مقدار بهینه برای اندازه پنجره و تعداد حالت‌های غالب SSA، اعتبارسنجی متقابل اعمال می‌شود تا داده‌های از دست‌رفته پر شوند. با هدف انتخاب بهترین اندازه پنجره و تعداد EOFها، بخشی از داده‌های موجود (انتخاب‌شده به صورت تصادفی) به داده از دست‌رفته تبدیل و سپس خطای RMSE بازسازی محاسبه می‌شود. در الگوریتم SSA هیچ راه مستقیمی برای حذف داده‌های نامعتبر وجود ندارد؛ باین حال با انتخاب تعداد اندک مؤلفه‌های معنی‌دار، سیگنال بازسازی‌شده صاف می‌شود و داده‌های نامعتبر به‌طور غیرمستقیم هموار می‌شوند. اما باید توجه داشت که داده‌های نامعتبر در سیگنال بازسازی تأثیر می‌گذارند.

۵-۲- ارزیابی عملکرد الگوریتم M-SSA

به‌منظور ارزیابی عملکرد الگوریتم M-SSA در بازسازی تصاویر LST، از خطای جذر میانگین مربعات و ضریب تبیین (R^2) استفاده شد. RMSE با استفاده از رابطه (۶) و ضریب تبیین طبق روابط (۷)، (۸) و (۹) به دست آمد. در این روابط، x_i ، y_i ، SS_{err} و SS_{tot} ، به ترتیب داده‌های واقعی، داده‌های مدل‌شده، مجموع مربعات خطا و مجموع مربعات کل هستند.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad \text{رابطه (۶)}$$

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}} \quad \text{رابطه (۷)}$$

$$SS_{err} = \sum_{i=1}^n (x_i - y_i)^2 \quad \text{رابطه (۸)}$$

این قابلیت، به‌ویژه در مواردی که سری زمانی دارای داده‌های از دست‌رفته طولانی باشد، کاربرد دارد. اگر بیش از یک سری زمانی از مشاهدات داشته باشیم، در این صورت X_{li} ؛ که $i = 1, 2, \dots, N$ و $l = 1, 2, \dots, L$ و تعمیم SSA برای ساخت ماتریس کوواریانس چندمتغیره (T) به صورت رابطه (۵) خواهد بود.

$$T = \begin{bmatrix} T_{1,1} & T_{1,2} & \dots & T_{1,L} \\ T_{2,1} & T_{2,2} & \dots & \\ \vdots & \vdots & \dots & \\ T_{L,1} & T_{L,2} & \dots & T_{L,L} \end{bmatrix} \quad \text{رابطه (۵)}$$

در این معادله، T_{li} ماتریس کوواریانس چندمتغیره بین کانال ۱ و l' است. T_{li} ، مانند S ، متعلق به SSA تک‌کاناله است ($l = 1$). با قطری کردن ماتریس T ، تعداد $L \times m$ بردار ویژه و مقدار ویژه از ماتریس T با استفاده از M-SSA ایجاد می‌شود. هر بردار ویژه (E^k) مؤلفه اصلی مکان-زمان (ST-PCs)^۱ نامیده می‌شود که مانند SSA، با انتخاب p از ST-PCs، سری زمانی اصلی را می‌توان بازسازی کرد.

عملکرد SSA برای پر کردن شکاف داده‌های از دست‌رفته در سری‌های زمانی بدین صورت است (Ghafarian Malamiri et al., 2018):

در یک سری زمانی با پنجره‌ای به اندازه (m) ، سری زمانی اصلی از طریق محاسبه مقادیر میانگین بدون اریب، در مرکز قرار می‌گیرد و داده‌های از دست‌رفته به صورت صفر تنظیم می‌شوند. اولین EOF طی روندی تکراری یافت می‌شود که الگوریتم SSA روی داده‌های صفر و متمرکز شده اعمال می‌کند. مقادیر از دست‌رفته براساس مؤلفه بازسازی‌شده EOF فعلی به‌روزرسانی می‌شوند. سپس الگوریتم SSA دوباره روی این مجموعه به‌روزر شده اعمال و EOF بار دیگر محاسبه می‌شود و مقادیر از دست‌رفته بازسازی می‌شوند. این روند، تا زمانی که نتایج آزمون همگرایی رضایت‌بخش باشد، تکرار می‌شود. تا وقتی تفاوت بین دو مقدار متوالی خطای جذر میانگین مربعات نرمال‌شده کمتر از ۰.۲/۵٪

۳- یافته‌های پژوهش

۳-۱- بررسی میزان داده‌های ازدست‌رفته در زمان و مکان

به‌منظور بررسی عملکرد الگوریتم M-SSA، ابتدا میزان و نحوه پراکنش داده‌های ازدست‌رفته در سری زمانی LST ساعتی بررسی شد. میزان داده‌های ازدست‌رفته، در هر تصویر سری زمانی LST سالیانه، در شکل ۲ نشان داده شده است. با توجه به این شکل، تقریباً در تمامی سال، داده‌های ازدست‌رفته وجود دارد (پراکنش زمانی). نقشه مکانی درصد داده‌های ازدست‌رفته در سری زمانی ساعتی دمای سطح زمین، طی یک سال، در شکل ۳ ارائه شده است. با توجه به این شکل، با حرکت از مناطق شمالی ایران به سمت جنوب، میزان داده‌های ازدست‌رفته کاهش می‌یابد. به عبارت دیگر، در مناطق شمالی ایران، شدت ابرناکی یا روزهای ابری بیشتر از مناطق جنوبی ایران است. حداکثر میزان داده‌های ازدست‌رفته، در سری زمانی ساعتی LST، در حاشیه دریای خزر مشاهده شد. به‌طور میانگین، در سطح ایران ۲۵/۵٪ از سری زمانی LST ساعتی طی سال ۲۰۲۲، با پوشش ابر، از بین رفته است.

۳-۲- بررسی روند و دوره‌ها در سری زمانی ساعتی دمای سطح زمین

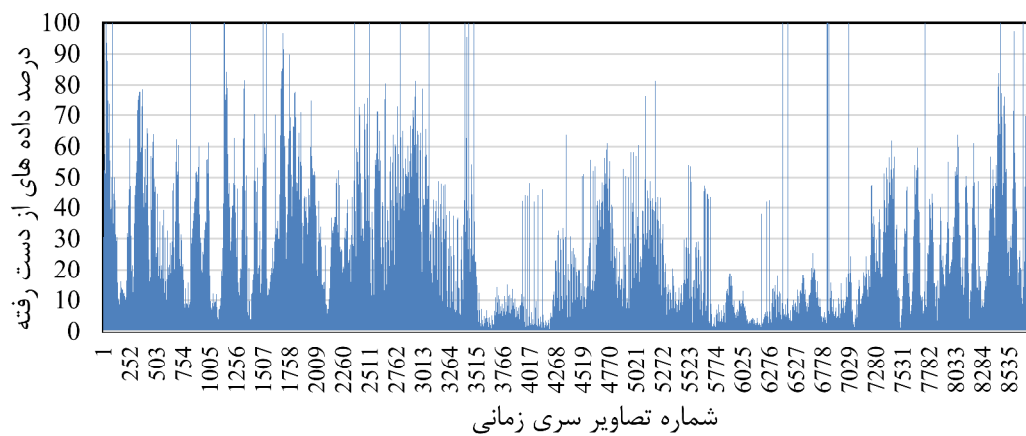
همان‌طور که در بخش مواد و روش‌ها توضیح داده شد، به‌منظور پر کردن فاصله داده‌های ازدست‌رفته یا گم‌شده در سری‌های زمانی با استفاده از الگوریتم SSA، ابتدا لازم است اندازه پنجره و سپس تعداد مؤلفه‌های معنی‌دار سری‌های زمانی در آن اندازه پنجره تعیین شود. منحنی اعداد ویژه، با پنجره زمانی ۹۶ ساعته در طول یک سری زمانی LST ساعتی سالیانه (۸۷۶۰ داده)، در شکل ۴-الف نشان داده شده است. با توجه به انتخاب اندازه پنجره ۹۶ ساعته، در محور عمودی شکل ۴-الف و ۴-ب، ۹۶ حالت یا مؤلفه داریم.

$$SS_{tot} = \sum_{i=1}^n (x_i - \bar{x})^2 \quad (9) \text{ رابطه}$$

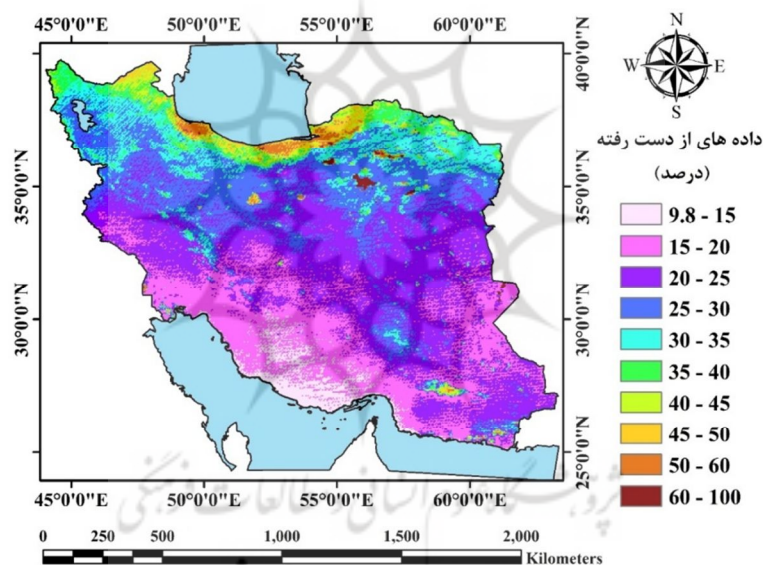
عملکرد الگوریتم M-SSA، در بازسازی سری زمانی ساعتی LST، به دو صورت بررسی شد. در حالت اول، بین داده‌های واقعی موجود در سری زمانی LST (آسمان صاف) و داده‌های بازسازی‌شده از طریق الگوریتم M-SSA، خطای RMSE و ضریب تبیین اندازه‌گیری شد. شایان ذکر است که در این حالت، عملکرد الگوریتم M-SSA در پر کردن نقاط بدون داده (آسمان ابری) همچنان مبهم خواهد بود. درواقع، اگر بخواهیم عملکرد الگوریتم M-SSA در مناطق دارای داده‌های ازدست‌رفته را بررسی کنیم (خطای پر کردن)^۱، باید به داده ازدست‌رفته یا داده‌های زمینی دسترسی داشته باشیم. برای رفع این مشکل و بررسی عملکرد الگوریتم M-SSA در پر کردن نقاط بدون داده، ماتریسی از سری زمانی ساعتی LST سالیانه، با اندازه ۸۷۶۰×۱۳۶۰ انتخاب شد. به عبارت دیگر، ۱۳۶۰ سری زمانی LST ساعتی سالیانه، بدون داده ازدست‌رفته (در مناطق جنوبی ایران، با کمترین شدت ابرناکی) انتخاب شد (شکل ۱؛ قطع مشخص‌شده با ضربدر). سپس داده‌های ازدست‌رفته مصنوعی (ابر مصنوعی) با شدت‌های متفاوت ۱۰، ۲۰، ۳۰، ۴۰، ...، ۹۰٪ به‌صورت تصادفی ایجاد شد. همان‌طور که اشاره شد، این سری زمانی ۸۷۶۰ داده در سال دارد. نقطه شروع ایجاد داده ازدست‌رفته (گپ)، به‌صورت تصادفی، از عدد ۱ تا ۸۷۶۰ انتخاب شد و این روند، تا رسیدن به سطح داده ازدست‌رفته دلخواه (۱۰، ۲۰، ...، ۹۰٪) در نرم‌افزار متلب^۲ ادامه یافت. سپس این سری‌های زمانی با داده‌های ازدست‌رفته مصنوعی، با استفاده از الگوریتم M-SSA، بازسازی شد و در نهایت، RMSE و ضریب تبیین، بین داده‌های ازدست‌رفته به‌صورت مصنوعی و داده‌های بازسازی‌شده از طریق الگوریتم M-SSA، محاسبه شد. در روش دوم، عملکرد الگوریتم M-SSA در بازسازی داده‌های ازدست‌رفته با شدت‌های متفاوت بررسی شد.

1. Gap Filling Error

2. MATLAB



شکل ۲. میزان داده‌های از دست رفته هر تصویر (پراکنش زمانی) در سری زمانی ساعتی دمای سطح زمین، طی یک سال (۸۷۶۰ تصویر)



شکل ۳. نقشه مکانی درصد داده‌های از دست رفته در سری زمانی ساعتی دمای سطح زمین، در طول یک سال

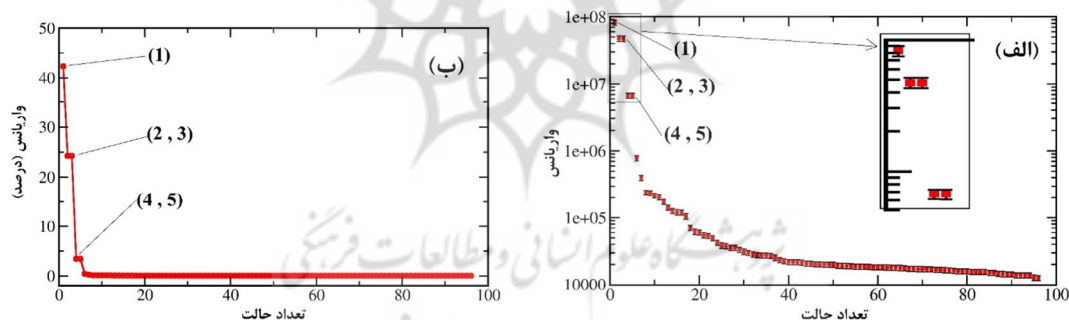
نشان‌دهنده نويز در سيگنال است. به عبارت ديگر، سمت راست منحنی تعداد مؤلفه‌های معنی‌دار سری زمانی را نمایش می‌دهد. محور عمودی شکل ۴-الف گویای اثر واریانس هریک از مؤلفه‌هاست. با توجه به شکل ۴-ب، مؤلفه‌های اول تا نهم به گونه‌ای معنی‌دار، بیشترین میزان واریانس را در کنترل دارند (مؤلفه‌های ۱ تا ۵، در شکل ۴-الف، با بزرگ‌نمایی مشخص شده است). مؤلفه‌های ۲ و ۳ یا ۴ و ۵ یا ۸ و ۹، با توجه به اینکه به صورت جفت در کنار هم قرار گرفته‌اند، توابع متعامد

هنگامی که منحنی اعداد تکین یک سری زمانی به ترتیب صعودی به نزولی مرتب شود، نموداری از سمت چپ به راست، با شیب زیاد به مناطق دارای شیب کم به دست می‌آید. هنگامی که منحنی اعداد تکین یک سری زمانی به ترتیب صعودی به نزولی مرتب شود، نموداری از سمت چپ به راست، با شیب زیاد به مناطق کم‌شیب به دست می‌آید. سمت راست نمودار منحنی (یعنی جایی که شیب تندتر است) نشان‌دهنده مؤلفه‌های اصلی یک سیگنال، و مناطق دارای شیب کم

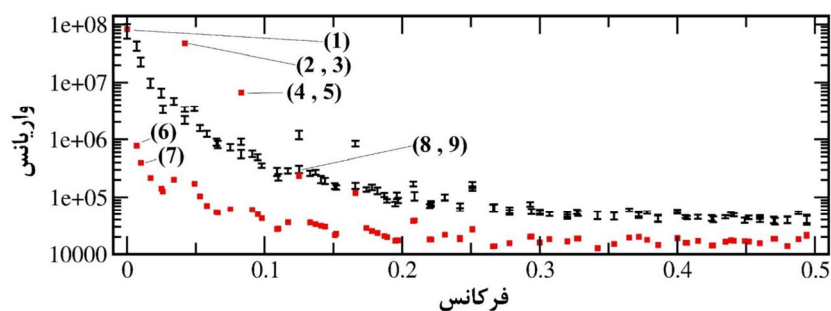
سری زمانی‌اند (این مؤلفه‌ها، در شکل، با شماره مشخص شده‌اند). در شکل ۵، ترتیب نام‌گذاری مؤلفه‌ها براساس میزان واریانس بزرگ‌تر است. مؤلفه‌های ۲ با ۳، ۴ با ۵ و ۸ با ۹، در این شکل، روی یکدیگر قرار گرفته‌اند. فرکانس مربوط به مؤلفه‌های اول، دوم با سوم، چهارم با پنجم و هشتم با نهم، به ترتیب ۰،۰۴۲، ۰/۰۸۳ و ۰/۱۲۵ سیکل بر تصویر است. با توجه به اینکه معکوس فرکانس بیانگر دوره زمانی است، مؤلفه اول به روند سالیانه تغییرات دما (شکل ۶-الف، خط آبی) اشاره دارد. مؤلفه‌های دوم و سوم به تغییرات دوره‌ای ۲۴ ساعته (روزانه) دما (شکل ۶-الف) اشاره دارند. مؤلفه‌های چهارم و پنجم گویای تغییرات دوره‌ای ۱۲ ساعته دما (شکل ۶-ب) و مؤلفه‌های هشتم و نهم نیز نشان‌دهنده تغییرات دوره‌ای ۸ ساعته دما (شکل ۶-ج) هستند. در مجموع، براساس نتایج این تحلیل، از نه مؤلفه در پنجره زمانی ۹۶ ساعته برای بازسازی سری‌های زمانی ساعتی ماهواره Meteosat-9 استفاده شد.

تجربی آن‌ها در ربع فاز انتقال می‌یابد (اختلاف فاز)؛ این مؤلفه‌ها به گونه‌ای مؤثر نشان‌دهنده مؤلفه‌های دوره‌ای‌اند (Vautard & Ghil, 1989). منحنی اعداد تکین نرمال شده در اندازه پنجره ۹۶ در شکل ۴-ب نشان داده شده است. با توجه به این شکل، مؤلفه ۱ در این سری زمانی تقریباً ۴۲/۵٪، مؤلفه ۲ و ۳ هریک ۲۴٪ و هریک از مؤلفه‌های ۴ و ۵ تقریباً ۳/۵٪ تغییرات را در کنترل دارند. مؤلفه‌های ۱ تا ۵ جمعاً ۹۷/۵٪ تغییرات این سری زمانی را در کنترل دارند. به عبارتی، انتخاب نکردن هریک از این مؤلفه‌ها، در بازسازی این سری زمانی، باعث ایجاد سیگنالی غیرقابل قبول یا ناقص خواهد شد.

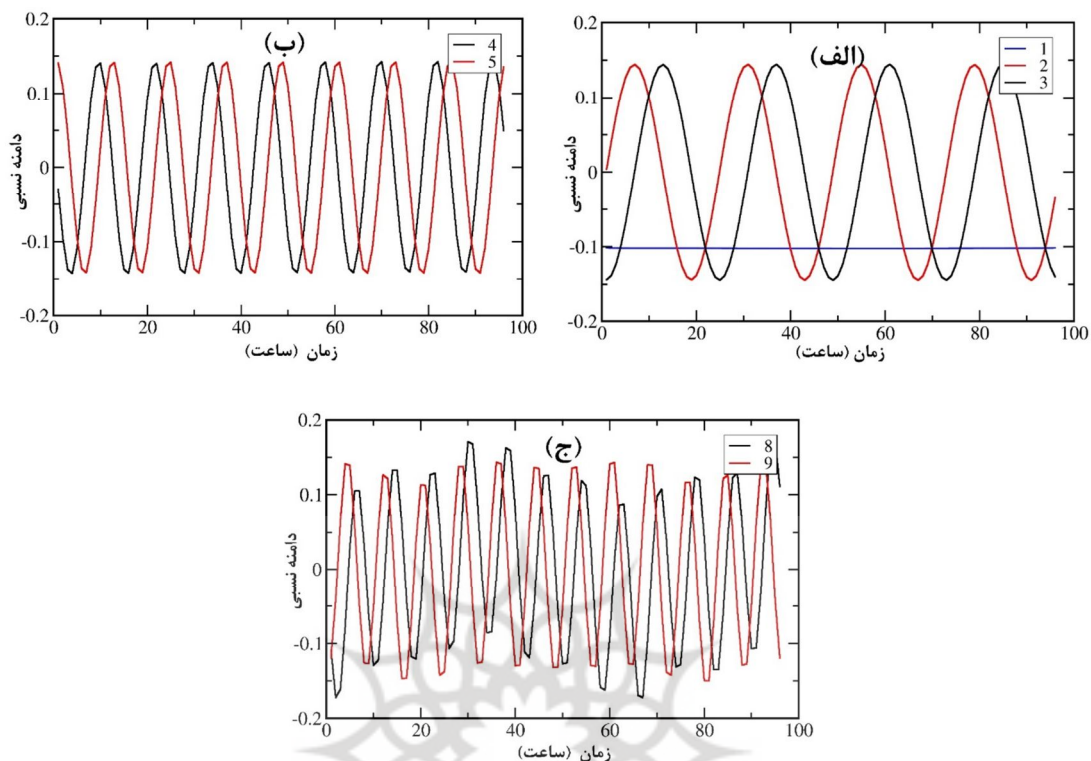
خروجی آزمون مونت کارلو SSA با داده‌های ساعتی LST، در پنجره زمانی ۹۶ ساعته، در شکل ۵ نشان داده شده است. شکل ۵، درواقع، مقادیر اعداد ویژه را در مقابل فرکانس مورد نظر نشان می‌دهد. با توجه به این شکل، مؤلفه‌های ۱ تا ۵ مؤلفه‌های اصلی و معنی‌دار این



شکل ۴. منحنی اعداد تکین در پنجره ۹۶ ساعته (راست) و منحنی اعداد تکین نرمال شده (تغییرات واریانس نسبت به هر مؤلفه) در پنجره ۹۶ ساعته (چپ)



شکل ۵. آزمون مونت کارلو SSA با داده سالیانه سری زمانی دمایی سطح زمین (پنجره ۹۶ ساعته)

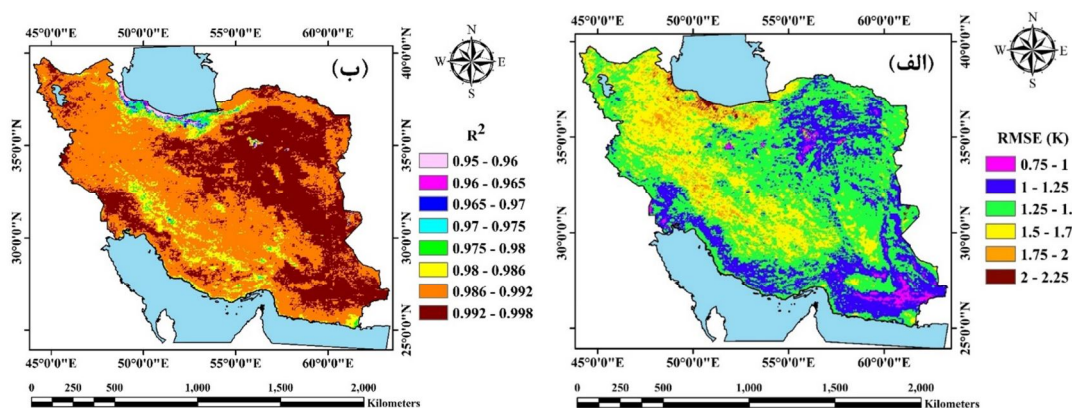


شکل ۶. توابع متعامد تجربی زمانی برای مؤلفه‌های ۱، ۲ و ۳ (الف)؛ مؤلفه‌های ۴ و ۵ (ب)؛ مؤلفه‌های ۸ و ۹ (ج)

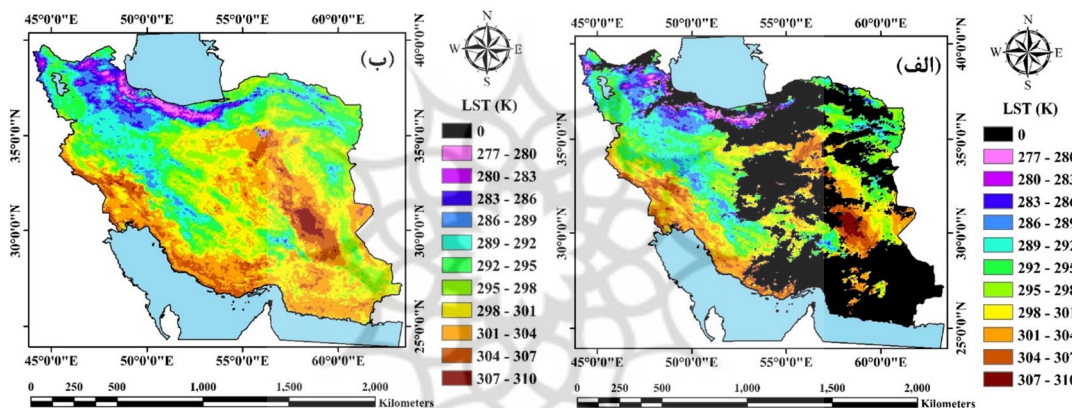
نمونه‌ای از تصاویر LST قبل و پس از بازسازی (متعلق به تاریخ ۱۴۰۱/۰۴/۲۲، در ساعت ۲۱) در شکل ۸ نشان داده شده است. با توجه شکل، نواحی شامل داده‌های ازدست‌رفته برآثر پوشش ابر (نواحی با رنگ سیاه) به خوبی با الگوریتم M-SSA بازسازی شده است زیرا، با توجه به شکل ۸، الگوها و طبقات تغییرات دما ثابت مانده و نواحی بدون داده پر شده است (نمایش مکانی). نمونه‌ای از سری زمانی LST ساعتی سالیانه (۸۷۶۰ داده)، با سری زمانی بازسازی شده از طریق الگوریتم M-SSA، در شکل ۹-الف ارائه شده است. با توجه به تراکم داده‌های زیاد، در شکل ۹-ب بخشی از این سری زمانی دیده می‌شود. مطابق شکل ۹، الگوریتم M-SSA به خوبی توانسته است نواحی شامل داده‌های ازدست‌رفته را بازسازی کند (نمایش زمانی).

۳-۳- ارزیابی عملکرد M-SSA در شرایط آسمان صاف

نقشه خطای RMSE و ضریب تبیین، بین داده‌های موجود سری زمانی LST ساعتی سالیانه (داده‌های آسمان صاف) و داده‌های بازسازی شده با الگوریتم M-SSA، در شکل ۷ نشان داده شده است. با توجه به شکل ۷-الف، حداقل خطای RMSE ۰/۷۵ و حداکثر آن، در نواحی غربی حاشیه دریای خزر، ۲/۲۵ کلوین مشاهده شد. همچنین با توجه به نقشه ضریب تبیین، کمترین میزان همبستگی در سواحل دریای خزر (۰/۹۵) وجود داشت. به طور میانگین در سطح کل ایران، خطای RMSE و ضریب تبیین بین داده‌های اصلی موجود و داده‌های بازسازی شده با الگوریتم M-SSA، به ترتیب ۱/۳۸ کلوین و ۰/۹۹ بود.



شکل ۷. نقشه خطای RMSE (الف) و ضریب تبیین (ب)، بین داده‌های موجود سری زمانی LST ساعتی سالیانه و داده‌های بازسازی شده با الگوریتم M-SSA

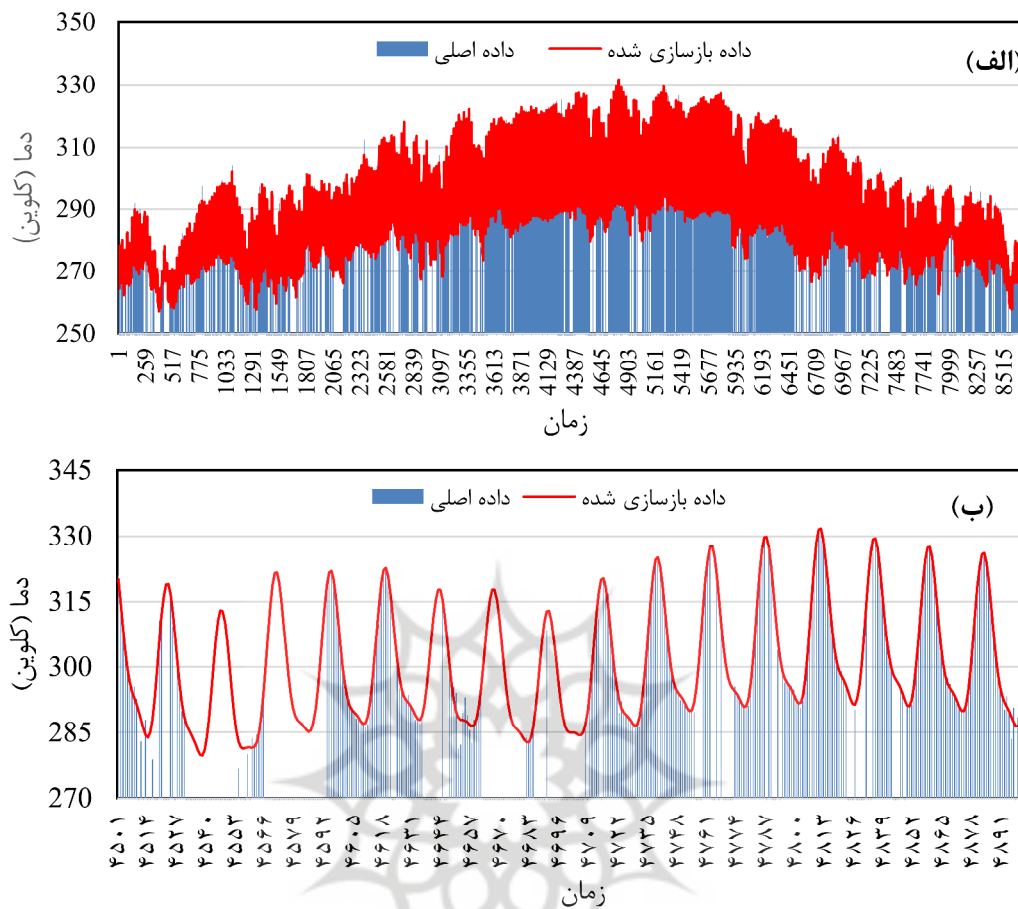


شکل ۸. نمونه‌ای از تصاویر با داده‌های از دست‌رفته (الف) و تصویر نظیر بازسازی از طریق الگوریتم M-SSA (ب)

۳-۴- ارزیابی عملکرد M-SSA در شرایط آسمان ابری

همان‌طور که در بخش مواد و روش‌ها توضیح داده شد، برای ارزیابی خطای پر کردن نقاط بدون داده در سری زمانی ماتریسی دارای ۱۳۶۰ سری زمانی LST ساعتی سالیانه با طول ۸۷۶۰ داده، ابر مصنوعی یا گپ ایجاد شد. نقشه داده‌های از دست‌رفته مصنوعی در سری‌های زمانی ماتریسی، با ۱۳۶۰ سری زمانی LST ساعتی طی یک سال (۸۷۶۰ داده)، در شکل ۱۰ نشان داده شده است. خطای RMSE و ضریب تبیین، در سری‌های زمانی LST ساعتی سالیانه با شدت‌های متفاوت داده‌های از دست‌رفته، در شکل ۱۱ ارائه شده است. مطابق شکل ۱۱، با افزایش شدت داده‌های از دست‌رفته به‌صورت تصادفی، خطای RMSE افزایش و ضریب

تبیین کاهش می‌یابد. اما این افزایش خطای RMSE تا سطح ۸۰٪ داده از دست‌رفته در سری زمانی، اندک بود. در سطح ۹۰٪ داده از دست‌رفته در سری زمانی، خطای RMSE و ضریب تبیین بین داده‌های حذف‌شده به‌صورت مصنوعی و بازسازی‌شده با الگوریتم M-SSA، به‌ترتیب، ۲/۷ کلون و ۰/۹۶ بود که در قیاس با سطح ۸۰٪ داده از دست‌رفته، افزایشی معنی‌دار یافت. باین‌حال این نتایج بسیار مطلوب و امیدوارکننده است زیرا الگوریتم M-SSA، در سری زمانی با ۹۰٪ داده از دست‌رفته، عملکرد خوبی داشت و خطای RMSE، در مقایسه با سری‌های زمانی دارای سطح ۱۰٪ داده از دست‌رفته، فقط ۱/۱ کلون افزایش یافت.



شکل ۹. نمونه سری زمانی LST ساعتی سالیانه (میللهای آبی)، با سری زمانی بازسازی شده (خط قرمز) از طریق الگوریتم M-SSA (الف) و بزرگنمایی بخشی از سری زمانی (ب)

زمین، سه مؤلفه دوره‌ای معنی‌دار ۲۴، ۱۲ و ۸ ساعته در سری‌های زمانی شناسایی شد. در بازسازی LST ساعتی ماهواره زمین ثابت Fengyun-2C نیز، مؤلفه‌های دوره‌ای معنی‌دار ۸، ۱۲ و ۲۴ ساعته در اندازه پنجره ۷۲ ساعته شناسایی شد (Ghafarian Malamiri et al., 2012). در فرایند بازسازی اثر پوشش ابر در تصاویر دمای سطح زمین با استفاده از داده‌های توالی شب و روز ماهواره مادیس، سه مؤلفه معنی‌دار شناسایی شد. مؤلفه اول تغییرات دمایی ناشی از چرخه شبانه‌روزی را تبیین می‌کند. مؤلفه دوم نشان‌دهنده الگوهای تغییرات دمایی سالیانه است و مؤلفه سوم تغییرات دمایی مرتبط با فصل‌های گوناگون را بازنمایی می‌کند (Ghafarian

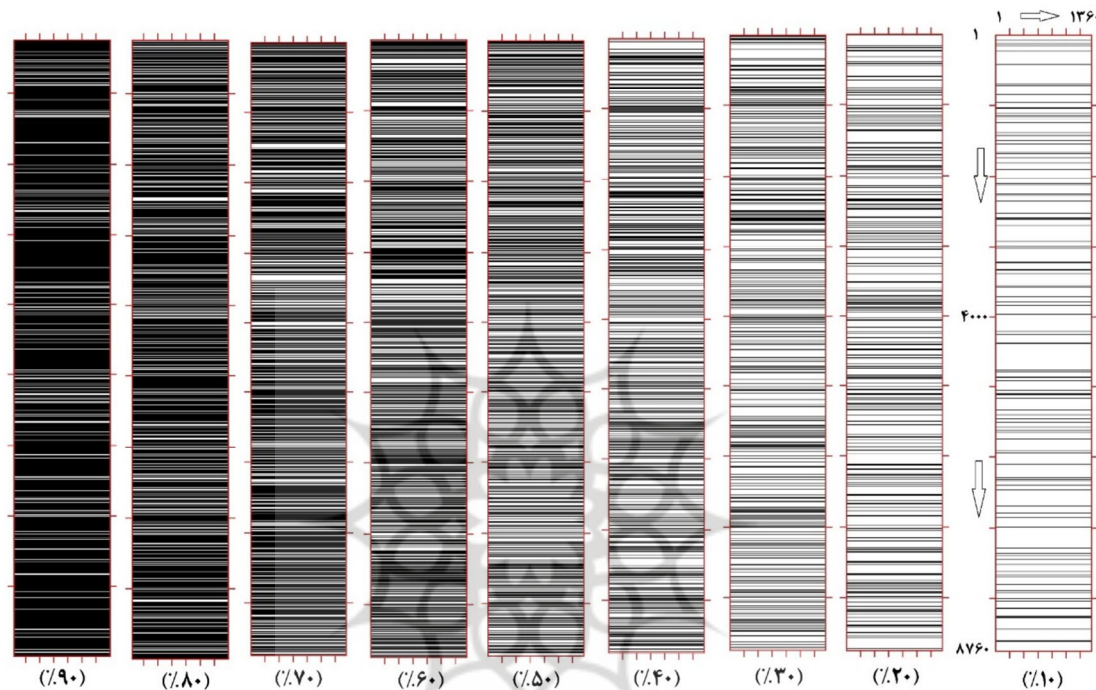
یک نمونه سری زمانی LST ساعتی سالیانه با داده‌های اصلی و داده‌های از دست‌رفته به صورت مصنوعی، همراه با داده‌های بازسازی شده، در شکل ۱۲-الف نشان داده شده است. در شکل ۱۲-ب، بخش کوچکی از این سری زمانی برای وضوح ارائه شده است. با توجه به این شکل، داده‌های بازسازی شده در سری زمانی با ۹۰٪ داده از دست‌رفته، به داده‌های اصلی و اولیه بسیار نزدیک است.

۴- بحث و نتیجه‌گیری

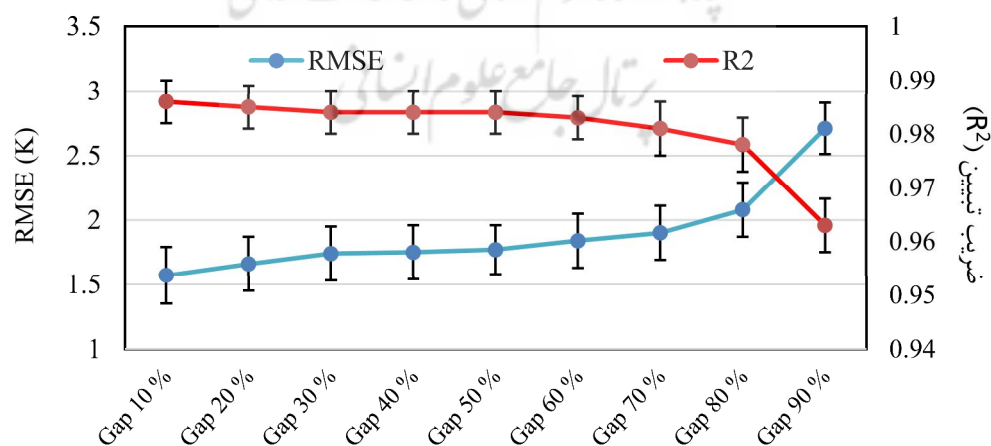
کلید اصلی بازسازی سری‌های زمانی، با رفتار دوره‌ای، شناسایی مؤلفه‌های معنی‌دار دوره‌ای است. در پژوهش حاضر، با توجه به فرکانس ساعتی داده‌های دمای سطح

سطح زمین، از چرخش زمین به دور خود و خورشید و انحراف محور آن پدید می‌آیند. بنابراین مؤلفه‌های یادشده اغلب برای بازسازی دمای سطح زمین، در هرکجای کره زمین، یکسان‌اند.

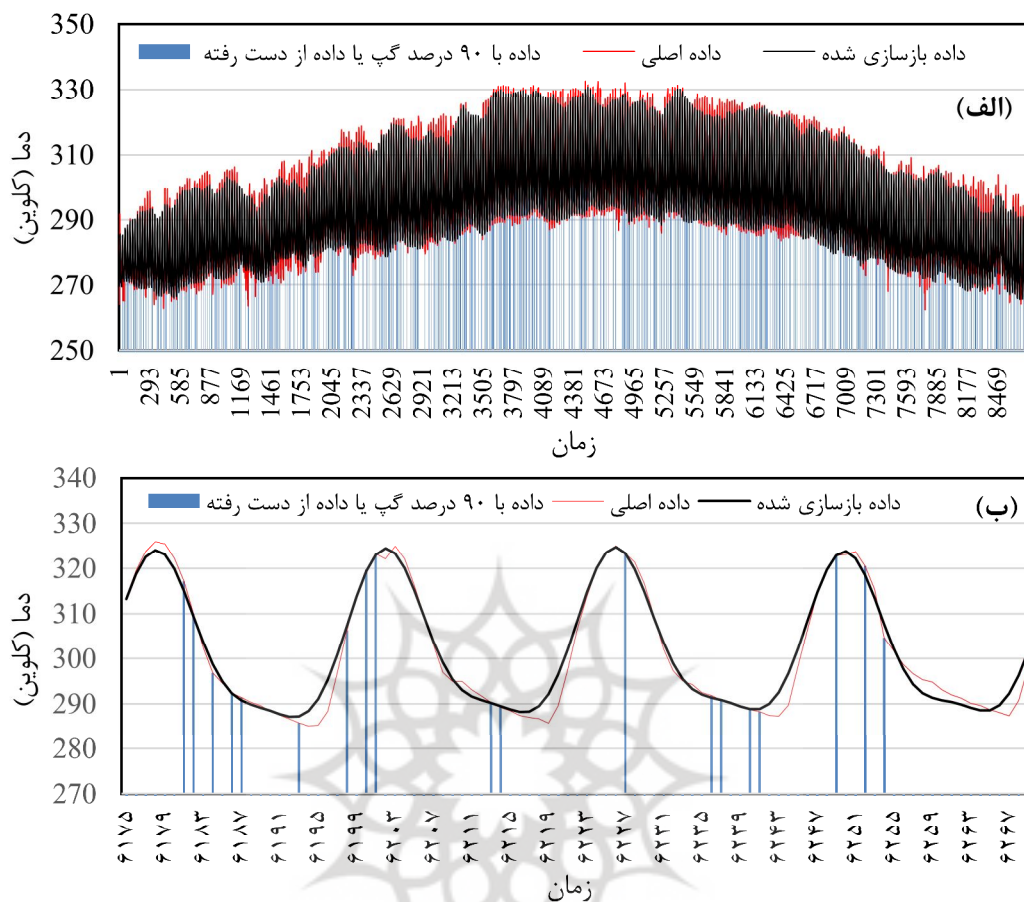
(Malamiri et al., 2018). در بازسازی سری زمانی سالیانه دمای سطح زمین ماهواره لندست-۸، با توالی شانزده‌روزه، فقط دوره‌های معنی‌دار سالیانه شناسایی شد (Zare Khormizi et al., 2023). به‌طور کلی تغییرات و مؤلفه‌های دوره‌ای، در سری‌های زمانی دمای



شکل ۱۰. نقشه داده‌های از دست‌رفته مصنوعی (گپ) در سری‌های زمانی ماتریسی، با ۱۳۶۰ سری زمانی LST ساعتی طی یک سال (۸۷۶۰ داده). از سمت چپ به راست، داده‌های از دست‌رفته افزایش می‌یابد (خطوط سیاه: داده‌های از دست‌رفته به صورت تصادفی)



شکل ۱۱. خطای RMSE و ضریب تبیین در سری‌های زمانی LST ساعتی سالیانه، با شدت‌های متفاوت داده‌های از دست‌رفته



شکل ۱۲. یک نمونه سری زمانی LST ساعتی سالیانه (خطوط قرمز)، به همراه داده‌های از دست‌رفته مصنوعی با شدت ۹۰٪ (خطوط آبی) و سری زمانی بازسازی‌شده از طریق الگوریتم M-SSA (الف) و بزرگنمایی بخشی از سری زمانی (ب)

براساس نتایج، خطای RMSE و ضریب تبیین بین داده‌های اولیه موجود و داده‌های بازسازی‌شده از طریق الگوریتم M-SSA، در شرایط آسمان صاف، به ترتیب ۱/۳۸ کلوین و ۰/۹۹ بود. همچنین براساس نتایج بررسی عملکرد M-SSA در شرایط آسمان ابری، خطای RMSE تا سطح ۸۰٪ داده از دست‌رفته به صورت تصادفی، همواره کمتر از ۲/۱ کلوین بود. در مجموع، خطاهای بازسازی، در شرایط آسمان صاف و ابری، در محدوده مورد قبولی قرار گرفتند زیرا با خطاهای تخمین دمای سطح زمین با استفاده از سنجش از دور (Li et al., 2013, ۱ تا ۳ کلوین) هم‌سو هستند (Chakraborty et al., 2015). نتایج سایر پژوهش‌ها با

نتایج پژوهش حاضر هم‌راستاست. در پژوهشی، خطای RMSE بین داده‌های غیرابری و بازسازی‌شده، در سری زمانی روزانه سنجنده مادیس با ۷۳۰ داده در سال، ۲/۹۵ کلوین به دست آمد (Ghafarian Malamiri et al., 2018). در مطالعه‌ای دیگر، مشخص شد خطای میانگین مطلق در بازسازی سری زمانی ساعتی با میانگین ۶۳٪ داده از دست‌رفته برآثر پوشش ابر، ۲/۲۵ کلوین است (Ghafarian Malamiri et al., 2012). خطای بازسازی داده از طریق روش درون‌یابی داده‌ها با توابع متعامد تجربی، در تصاویر دمای سطح زمین سنجنده مادیس طی روز و شب، به ترتیب ۴/۵ و ۲/۳ کلوین بود (Zhou et al., 2017). تغییرات سری‌های زمانی دمای

از جمله ساعت ابرناکی (روز یا شب)، مدت زمان ابرناکی، فصل، توپوگرافی، رطوبت خاک، نوع پوشش زمین، ویژگی‌های حرارتی خاک و اینرسی حرارتی خاک. بنابراین بازسازی دمای سطح زمین، زیر پوشش ابر، می‌تواند چالش و پیشنهادی برای مطالعات بیشتر در آینده باشد.

۵- منابع

- Broomhead, D.S. & King, G.P., 1986, **Extracting Qualitative Dynamics from Experimental Data**, Physica D: Nonlinear Phenomena, 20, PP. 217-236, [https://doi.org/10.1016/0167-2789\(86\)90031-X](https://doi.org/10.1016/0167-2789(86)90031-X).
- Chakraborty, S.D., Kant, Y. & Mitra, D., 2015, **Assessment of Land Surface Temperature and Heat Fluxes over Delhi Using Remote Sensing Data**, Journal of Environmental Management, 148, PP. 143-152, <https://doi.org/10.1016/j.jenvman.2013.11.034>.
- Chen, D., Zhuang, Q., Zhu, L. & Zhang, W., 2022, **Comparison of Methods for Reconstructing MODIS Land Surface Temperature under Cloudy Conditions**, Applied Sciences, 12(12), P. 6068, <https://doi.org/10.3390/app12126068>.
- Cui, J., Zhang, M., Song, D., Shan, X. & Wang, B., 2022, **MODIS Land Surface Temperature Product Reconstruction Based on the SSA-BiLSTM Model**, Remote Sensing, 14(4), P. 958, <https://doi.org/10.3390/rs14040958>.
- Freitas, S.C., Trigo, I.F., Macedo, J., Barroso, C., Silva, R. & Perdigão, R., 2013, **Land Surface Temperature from Multiple Geostationary Satellites**, International Journal of Remote Sensing, 34(9-10), PP. 3051-3068, <https://doi.org/10.1080/01431161.2012.716925>.
- Geng, L., Ma, M., Wang, X., Yu, W., Jia, S. & Wang, H., 2014, **Comparison of Eight Techniques for Reconstructing Multi-Satellite Sensor Time-Series NDVI Data Sets in the Heihe River Basin, China**, Remote Sensing, 6(3), PP. 2024-2049, <https://doi.org/10.3390/rs6032024>.
- Ghafarian Malamiri, H.R., 2015, **Reconstruction of Gap-Free Time Series Satellite Observations of Land Surface Temperature to Model Spectral Soil Thermal Admittance**, Doctoral Dissertation, Technische Universiteit Delft, The Netherlands, <https://doi.org/10.4233/uuid:63dc3402-9fd6-4594-a00e-7aa5ae2501aa>.

سطح زمین تأثیر معتنابهی از مؤلفه‌های دوره‌ای می‌پذیرند. این ویژگی امکان بازسازی دقیق و با خطای کم را برای تصاویر دمای سطح زمین، فراهم می‌کند. باین حال دقت بازسازی سری‌های زمانی از سه عامل ناشی می‌شود: اندازه بزرگ‌ترین فاصله داده‌های از دست‌رفته (بزرگ‌ترین شکاف)، میزان داده‌های از دست‌رفته و تعداد مکان‌های داده از دست‌رفته (تعداد شکاف‌ها) (Zhou et al., 2012). اوج‌ها و پایه‌ها یا نقاط عطف تغییرات سری زمانی نیز در دقت بازسازی تأثیر بسیار دارند (Zhou et al., 2012). باین‌همه الگوریتم M-SSA، با توجه به در نظر گرفتن مؤلفه‌های دوره‌ای و روندها و همچنین استفاده از همبستگی زمانی و مکانی، می‌تواند در بازسازی داده‌های از دست‌رفته با فاصله بزرگ مؤثر باشد.

یکی از موارد شایان توجه در بازسازی اثر پوشش ابر، در پژوهش حاضر و بسیاری از پژوهش‌های دیگر، بازسازی دمای سطح زمین با شرط آسمان صاف است. به عبارت دیگر، در بسیاری از روش‌های درون‌یابی زمانی، LST زیر پوشش ابر به دست نمی‌آید. باین حال با توجه به فرکانس ساعتی داده‌های سطح زمین در پژوهش حاضر، الگوریتم SSA مؤلفه‌های دوره‌ای و روندهای تغییرات را با دقت استخراج می‌کند و فرکانس بالای داده‌ها به برازش دقیق سیگنال بر داده‌های موجود منجر می‌شود. این مسئله باعث می‌شود دمای بازسازی شده به LST زیر پوشش ابر نزدیک باشد. به طور کلی، بازسازی LST زیر پوشش ابر دارای پیچیدگی بیشتری است (Sun et al., 2017). در برخی پژوهش‌ها، از معادلات رگرسیون چندمتغیره و بیلان انرژی سطح برای تخمین LST زیر پوشش ابر استفاده شده است (Lu et al., 2011; Sun et al., 2017; Zhao & Duan, 2020).

ابرها یکی از مهم‌ترین عوامل ایجاد داده‌های گم‌شده (شکاف) و نامعتبر در سری‌های زمانی دمای سطح زمین‌اند. به طور کلی، در پژوهش پیش رو، شکاف‌های ایجادشده بر اثر پوشش ابر به گونه‌ای مؤثر و با خطای اندک، از طریق الگوریتم M-SSA پر شدند. تخمین دقیق LST زیر پوشش ابرها به عوامل بسیاری بازمی‌گردد؛

- Ghafarian Malamiri, H.R. & Zare Khormizie, H., 2017, **Reconstruction of Cloud-Free Time Series Satellite Observations of Land Surface Temperature (LST) Using Harmonic Analysis of Time Series Algorithm (HANTS)**, Journal of RS and GIS for Natural Resources, 8(3), PP. 37-55, <http://dorl.net/dor/20.1001.1.26767082.1396.8.3.3.0>.
- Ghafarian, H.R., Menenti, M., Jia, L.H. & den Ouden, R., 2012, **Reconstruction of Cloud-Free Time Series Satellite Observations of Land Surface Temperature**, EARSeL eProceedings (PP. 121-131).
- Ghafarian Malamiri, H.R., Roustaei, I., Olafsson, H., Zare, H. & Zhang, H., 2018, **Gap-Filling of MODIS Time Series Land Surface Temperature (LST) Products Using Singular Spectrum Analysis (SSA)**, Atmosphere, 9(9), P. 334, <https://doi.org/10.3390/atmos9090334>.
- Golyandina, N., Nekrutkin, V. & Zhigljavsky, A., 2001, **Analysis of Time Series Structure: SSA and Related Techniques**, Washington DC, USA: Chapman & Hall/CRC.
- Helman, D., Lensky, I.M., Yakir, D. & Osem, Y., 2017, **Forests Growing under Dry Conditions Have Higher Hydrological Resilience to Drought than Do More Humid Forests**, Global Change Biology, 23(7), PP. 2801-2817, <https://doi.org/10.1111/gcb.13551>.
- Julien, Y. & Sobrino, J.A., 2010, **Comparison of Cloud-Reconstruction Methods for Time Series of Composite NDVI Data**, Remote Sensing of Environment, 114(3), PP. 618-625, <https://doi.org/10.1016/j.rse.2009.11.001>.
- Ke, L., Ding, X. & Song, C., 2013, **Reconstruction of Time-Series MODIS LST in Central Qinghai-Tibet Plateau Using Geostatistical Approach**, IEEE Geoscience and Remote Sensing Letters, 10(6), PP. 1602-1606, <https://doi.org/10.1109/LGRS.2013.2263553>.
- Kleipool, Q.L., Dobber, M.R., de Haan, J. & Levelt, P.F., 2008, **Earth Surface Reflectance Climatology from 3 Years of OMI Data**, Journal of Geophysical Research: Atmospheres, 113(D18), <https://doi.org/10.1029/2008JD010290>.
- Kondrashov, D. & Ghil, M., 2006, **Spatio-Temporal Filling of Missing Points in Geophysical Data Sets**, Nonlinear Processes in Geophysics, 13(2), PP. 151-159, <https://doi.org/10.5194/npg-13-151-2006>.
- Kondrashov, D., Shprits, Y. & Ghil, M., 2010, **Gap Filling of Solar Wind Data by Singular Spectrum Analysis**, Geophysical Research Letters, 37(15), <https://doi.org/10.1029/2010GL044138>.
- Li, J., Carlson, B.E. & Lacis, A.A., 2013, **Application of Spectral Analysis Techniques in the Intercomparison of Aerosol Data: 1. An EOF Approach to Analyze the Spatial-Temporal Variability of Aerosol Optical Depth Using Multiple Remote Sensing Data Sets**, Journal of Geophysical Research: Atmospheres, 118(15), PP. 8640-8648, <https://doi.org/10.1002/jgrd.50686>.
- Lu, L., Venus, V., Skidmore, A., Wang, T. & Luo, G., 2011, **Estimating Land-Surface Temperature under Clouds Using MSG/SEVIRI Observations**, International Journal of Applied Earth Observation and Geoinformation, 13(2), PP. 265-276, <https://doi.org/10.1016/j.jag.2010.12.007>.
- Martins, J.P., Coelho e Freitas, S., Trigo, I.F., Barroso, C. & Macedo, J., 2023, **Copernicus Global Land Operations-Lot I "Vegetation and Energy" Algorithm Theoretical Basis Document**, Land Surface Temperature—LST, 1. V2.1. ISSUE 12.23.
- Mukherjee, S., Joshi, P.K. & Garg, R.D., 2014, **A Comparison of Different Regression Models for Downscaling Landsat and MODIS Land Surface Temperature Images over Heterogeneous Landscape**, Advances in Space Research, 54(4), PP. 655-669, <https://doi.org/10.1016/j.asr.2014.04.013>.
- Musial, J.P., Verstraete, M.M. & Gobron, N., 2011, **Comparing the Effectiveness of Recent Algorithms to Fill and Smooth Incomplete and Noisy Time Series**, Atmospheric Chemistry and Physics, 11(15), PP. 7905-7923, <https://doi.org/10.5194/acp-11-7905-2011>.
- Pede, T. & Mountrakis, G., 2018, **An Empirical Comparison of Interpolation Methods for MODIS 8-Day Land Surface Temperature Composites across the Conterminous United States**, ISPRS Journal of Photogrammetry and Remote Sensing, 142, PP. 137-150, <https://doi.org/10.1016/j.isprsjprs.2018.06.003>.
- Shiff, S., Helman, D. & Lensky, I.M., 2021, **Worldwide Continuous Gap-Filled MODIS**

- Land Surface Temperature Dataset**, Scientific Data, 8(1), PP. 74, <https://doi.org/10.1038/s41597-021-00861-7>.
- Sun, L., Chen, Z., Gao, F., Anderson, M., Song, L., Wang, L., ... & Yang, Y., 2017, **Reconstructing Daily Clear-Sky Land Surface Temperature for Cloudy Regions from MODIS Data**, Computers & Geosciences, 105, PP. 10-20, <https://doi.org/10.1016/j.cageo.2017.04.007>.
- Vautard, R. & Ghil, M., 1989, **Singular Spectrum Analysis in Nonlinear Dynamics, with Applications to Paleoclimatic Time Series**, Physica D: Nonlinear Phenomena, 35, PP. 395-424, [https://doi.org/10.1016/0167-2789\(89\)90077-8](https://doi.org/10.1016/0167-2789(89)90077-8).
- Vautard, R., Yiou, P. & Ghil, M., 1992, **Singular-Spectrum Analysis: A Toolkit for Short, Noisy Chaotic Signals**, Physica D: Nonlinear Phenomena, 58(1-4), PP. 95-126, [https://doi.org/10.1016/0167-2789\(92\)90103-T](https://doi.org/10.1016/0167-2789(92)90103-T).
- Verhoef, W., Menenti, M. & Azzali, S., 1996, **Cover A colour Composite of NOAA-AVHRR-NDVI Based on Time Series Analysis (1981-1992)**, International Journal of Remote Sensing, 17(2), PP. 231-235, <https://doi.org/10.1080/01431169608949001>.
- Wang, H., Mao, K., Yuan, Z., Shi, J., Cao, M., Qin, Z., ... & Tang, B., 2021, **A Method for Land Surface Temperature Retrieval Based on Model-Data-Knowledge-Driven and Deep Learning**, Remote Sensing of Environment, 265, P. 112665, <https://doi.org/10.1016/j.rse.2021.112665>.
- Xu, Y. & Shen, Y., 2013, **Reconstruction of the Land Surface Temperature Time Series Using Harmonic Analysis**, Computers & Geosciences, 61, PP. 126-132, <https://doi.org/10.1016/j.cageo.2013.08.009>.
- Yu, P., Zhao, T., Shi, J., Ran, Y., Jia, L., Ji, D. & Xue, H., 2022, **Global Spatiotemporally Continuous MODIS Land Surface Temperature Dataset**, Scientific Data, 9(1), P. 143, <https://doi.org/10.1038/s41597-022-01214-8>.
- Zare Khormizi, H., Tavili, A. & Ghafarian Malamiri, H.R., 2021, **Estimation of Actual Evapotranspiration Using SEBAL Algorithm and Comparison with Improved FAO 56 Standard Evapotranspiration with KC-NDVI Relationship**, Iranian Journal of Remote Sensing & GIS, 13(3), PP. 73-92, <https://doi.org/10.52547/gisj.13.3.73>.
- Zare Khormizi, H., Ghafarian Malamiri, H.R. & Alian, S., 2023, **Calculation of Land Surface Temperature Using a Generalized Split-Window Algorithm and the Reconstruction of Its Lost Data by Cloud Cover Through a Singular Spectral Analysis (SAA)-Algorithm**, Desert, 28(1), PP. 27-48, <https://doi.org/10.22059/jdesert.2023.93537>.
- Zhang, C., Li, W. & Travis, D.J., 2009, **Restoration of Clouded Pixels in Multispectral Remotely Sensed Imagery with Cokriging**, International Journal of Remote Sensing, 30(9), PP. 2173-2195, <https://doi.org/10.1080/01431160802549294>.
- Zhao, W. & Duan, S.B., 2020, **Reconstruction of Daytime Land Surface Temperatures under Cloud-Covered Conditions Using Integrated MODIS/Terra Land Products and MSG Geostationary Satellite Data**, Remote Sensing of Environment, 247, P. 111931, <https://doi.org/10.1016/j.rse.2020.111931>.
- Zhou, J., Jia, L. & Menenti, M., 2015, **Reconstruction of Global MODIS NDVI Time Series: Performance of Harmonic Analysis of Time Series (HANTS)**, Remote Sensing of Environment, 163, PP. 217-228, <https://doi.org/10.1016/j.rse.2015.03.018>.
- Zhou, J., Jia, L., Hu, G. & Menenti, M., 2012, **Evaluation of Harmonic Analysis of Time Series (HANTS): Impact of Gaps on Time Series Reconstruction**, In 2012 Second International Workshop on Earth Observation and Remote Sensing Applications (PP. 31-35), IEEE, <https://doi.org/10.1109/EORSA.2012.6261129>.
- Zhou, W., Peng, B. & Shi, J., 2017, **Reconstructing Spatial-Temporal Continuous MODIS Land Surface Temperature Using the DINEOF Method**, Journal of Applied Remote Sensing, 11(4), PP. 046016-046016, <https://doi.org/10.1117/1.JRS.11.046016>.

This Page is Intentionally Left Blank



پروہشگاہ علوم انسانی و مطالعات فرہنگی
پرتال جامع علوم انسانی