Environmental Earth Sciences (2023) 82:387
https://doi.org/10.1007/s12665-023-11072-1

ORIGINAL ARTICLE

Earthquake magnitude prediction in Turkey: a comparative study of deep learning methods, ARIMA and singular spectrum analysis

Hatice Öncel Çekim[1] · Hatice Nur Karakavak[1] · Gamze Özel[1] · Senem Tekin[2]

Received: 2 February 2023 / Accepted: 21 July 2023 / Published online: 2 August 2023

Abstract

The Aegean region is geologically situated at the western end of the Gediz Graben system, influenced by the Western Anatolian Regime. In addition, the region is characterized by various active fault lines that can generate earthquake activity. Numerous earthquakes have been recorded in the region, causing significant material and moral damage from the past to the present. In this study, earthquake data from three different catalogs are examined. The non-clustered catalog is compiled for the years 1970 to 2020, including earthquakes with a moment magnitude (Mw) greater than 3.0. The monthly average magnitudes of earthquakes in the region are obtained and analyzed using ARIMA, singular spectrum analysis (SSA), and deep learning methods including convolutional neural network (CNN) and long short-term memory (LSTM), as these methods have not been compared for the region previously. Each method has a different benefit. ARIMA analyzes time series trends and seasonal patterns, while SSA focuses on decomposition and feature extraction. LSTM attempts to capture complex relationships using memory mechanisms, while CNN is powerful at pattern recognition and extracting important features. Thanks to this diversity, our study allows for more comprehensive and reliable forecasts of average earthquake magnitudes for the next 36 periods. The estimation capabilities and error rates of each method were analyzed based on earthquake magnitude data, and it was determined that the LSTM method provided the most effective and accurate predictions.

Keywords ARIMA · Deep learning · Earthquake · Long short-term memory · Magnitude · Neural network · Singular spectrum analysis

Introduction

Earthquakes are the leading natural disasters causing loss of life and property both in our country and worldwide, particularly in residential areas. While the exact timing of earthquakes remains uncertain, they are more likely to occur in regions with active faults and high seismic activity (Gioncu and Mazzolani 2011). Among the globally recognized earthquake-prone areas, the Pacific and Alpine–Himalaya seismic belts stand out as the most intense ones. Turkey, located within the Alpine–Himalayan belt, faces a high risk of earthquakes (Cekim et al. 2021). Moreover, a significant portion, around 43%, of the country's surface lies in a highly seismically hazardous zone, contributing to its long history of devastating earthquakes (Akkar et al. 2018). As the most active region in the Mediterranean, Turkey experiences frequent earthquake activity due to its location between three major tectonic plates: the African, Arabian, and Eurasian plates. The dynamic interaction among these plates accounts for most of the tectonic activity in the region (Cohen et al. 1995). As a result, the Anatolian Plate carries a substantial seismic hazard. Particularly, Western Anatolia, which encompasses the Aegean region of Turkey, is characterized by intense deformation and a series of active faults aligned from north to south with an east–west extension. Consequently, the Aegean region is highly susceptible to destructive earthquakes (Seyitoğlu and Scott 1992; Bayrak et al. 2017; Coban and Sayıl 2019, 2022).

The Aegean Region holds significant importance both in terms of geology and tectonics, making it a compelling area for further investigation. With its high values of maximum ground acceleration in relation to earthquake hazards, the

✉ Hatice Öncel Çekim
oncelhatice@hacettepe.edu.tr

[1] Department of Statistics, Hacettepe University, 06800 Ankara, Turkey

[2] Department of Mining and Mineral Extraction, School of Technical Science, Adıyaman University, 02040 Adıyaman, Turkey

region is prone to seismic activity. The westward movement of Anatolia leads to east–west compression and north–south expansion, activating the faults in the Aegean Region. Upon examining its seismicity, it is evident that the region houses numerous faults capable of generating earthquakes with magnitudes of six or higher. Tectonically, there are several noteworthy faults, each with independent characteristics and distinct focal solutions. The earthquake hazard map of Turkey indicates that the region's hazard values pose a significant risk. It is believed that the Graben basins in the Aegean region formed due to thermal collapse during the Early–Middle Miocene period. While there is limited information on regional tectonics during the Middle–Late Miocene, gradual stretching is thought to have been initiated during this period. The primary movement of east–west grabens is believed to have occurred in the Pliocene (Erdik et al. 2004). The active tectonics of the Aegean Graben system are influenced by two crucial geological events: the subduction system and the compression of Anatolia by the Arabian plate in the north. Situated in the western part of Turkey, the Aegean region is characterized by numerous graben systems. These grabens are delineated by faults on both sides and appear as elongated, narrow flat regions. Among these, the Gediz graben system holds significant importance in Western Anatolia and is one of the vital graben systems in Turkey (Ketin 1968; Polat et al. 2008). In recent years, east–west main grabens and active graben systems in the NE–SW and NW–SE directions have developed in Western Anatolia due to the influence of the NNE–SSW extension. Prominent neotectonic structures that have formed in Western Anatolia under this tectonic regime include the Gökova, Büyük Menderes, Gediz, Bakırçay, and Simav grabens. The Büyük Menderes graben, a significant structure in the east–west direction, extends from the Denizli basin in the east to the Aegean Sea in the west (Şengör 1982; Emre et al. 2018). The northern and southern boundaries of the approximately 150 km-long Büyük Menderes graben are marked by faults (Paton 1992). Previous studies have revealed that the faults on the northern side of the Graben exhibit greater activity compared to those on the southern side, and most of the activity during the Holocene period occurred in the northern part of the Graben (Altunel 1999; Bozkurt 2000; Sözbilir 2001).

Apart from the Gediz graben, the region is characterized by several active faults that can serve as sources of earthquake activity. These include the Gülbahçe, Güzelhisar, Menemen, Seferihisar, Tuzla, Gümüdür, Mordagan, İzmir, Kiraz, and Yeni Foça faults (Bayrak and Bayrak 2012). According to the Disaster and Emergency Management Authority (AFAD) in Turkey, the earthquake with a moment magnitude (Mw) of 6.9 occurred off the coast of Samos on October 30, 2020, at 14:51 Turkish time. The earthquake was recorded as one of the largest to occur in the region, with a magnitude equivalent to its scale. It lasted approximately 16 s, and within the following 2 months, 5799 aftershocks were recorded, ranging in magnitude from 0.7 to 5.1. The town of Seferihisar, located 27 km away from the epicenter, was the closest area in Turkey affected by the earthquake (Karadaş and Öner 2021). This recent seismic event resulted in significant damage and underscored the importance of studying earthquakes in the region. Consequently, numerous research studies have been conducted to understand the characteristics of seismic activities and their relationship with faults in the highly active Aegean Region (Mouslopoulou and Hristopulos 2011; Tağil and Alevkayali 2013; Akın and Yağmurlu 2020). In addition, there have been investigations specifically focused on earthquakes with magnitudes of four and above, which possess significant destructive potential in the Aegean Region (Stein et al. 1997; Durak 2008; Danese et al. 2009; Tagil 2004).

Predicting earthquake magnitudes and their occurrence times is a crucial task in seismic analysis, as it provides valuable information for probabilistic seismic hazard assessment studies (Abdalzaher et al. 2020; Elhadidy et al. 2021). These studies utilize ground motion estimation equations to investigate seismicity and assess the potential hazards associated with earthquakes. Furthermore, time series analysis plays a significant role in earthquake prediction by examining statistical data collected at regular intervals over time. By analyzing past observations, reliable predictions can be made, aiding in the understanding and management of seismic risks. Among the widely used models in this field, the autoregressive integrated moving average (ARIMA) model stands out (Shishegaran et al. 2019; Keskin and Külahcı, 2022). For example, Amei et al. (2012) applied the ARIMA model to predict earthquake sizes worldwide between 1896 and 2009, with a specific focus on earthquakes of magnitude 8.0 or higher on the Richter scale. Similarly, Todelo et al. (2019) utilized an ARIMA (1,0,6) model to predict earthquake occurrences based on earthquake count datasets from 2000 to 2017. ARIMA models have also been employed in studies focusing on smaller regions, aiming to predict earthquake magnitudes. Shishegaran et al. (2019) utilized ARIMA, GARCH, and a combination of both to predict potential earthquake magnitudes in Iran. Yuan et al. (2022) used the ARIMA model to analyze earthquake catalog data from the Longmen Mountain fault zone in China, aiming to estimate earthquake origin times. However, the application of the ARIMA model to analyze seismicity and predict magnitudes in the Aegean graben region of Turkey remains unexplored.

Singular spectrum analysis (SSA) is a powerful non-parametric method widely employed for spectral estimation in time series analysis, offering a broad range of applications. It integrates components from dynamical systems, multivariate statistics, classical time series analysis, multivariate

geometry, and signal processing. SSA proves to be highly valuable in various tasks, including trend detection at different resolutions, smoothing, extraction of seasonal components, simultaneous extraction of cycles with small and large periods, identification of periodicities with varying amplitudes, extraction of complex trends and periodicities, analysis of short time series for structural patterns, and detection of change points (Hassani 2007). Unlike ARIMA, SSA does not rely on assumptions about the underlying time series and only requires adjusting a single parameter. Consequently, it has gained significant popularity across different fields in recent years (Lin et al. 2022; Coussin 2022; Putriasari et al. 2022). While there have been limited applications of SSA in seismological studies, Yadav et al. (2015) employed SSA to investigate the presence of annual cyclicity in seismicity data, facilitating the establishment of statistical correlations between seismic activity and reservoir operations in India. SSA is capable of analyzing and reconstructing a time series with or without specific constituents based on the desired objectives. In this study, we employ the SSA model for the following purposes: (i) constructing a smoothed version of the time series by focusing on a selected subset of its constituents, (ii) analyzing the periodic components of the time series to gain insights into the fundamental processes that gave rise to it, (iii) eliminating all trends and periodic components from the series, isolating the noise, and (iv) reconstructing the original time series while excluding its cyclic constituents.

Convolutional neural network (CNN) is a widely used neural network architecture in the field of deep learning. While initially designed for image processing tasks, CNNs can also be effectively utilized for analyzing different types of data, including time series data. Earthquake magnitude prediction refers to the process of measuring the intensity of an earthquake based on given seismic data. CNN can assist in automatically learning data-specific features that are used to estimate the magnitude of earthquakes.

The rapid determination of the magnitude and location of the earthquake is very important in reducing the earthquake risk and loss of life. In order to quickly determine these parameters, studies are carried out on earthquake early warning systems with various CNN-based deep learning models (Saad et al. 2020; Abdalzaher et al. 2021; Zhang et al. 2022). These studies investigate how a CNN-based model can be utilized for estimating the magnitudes of earthquakes. As far as we know, CNN has not been used in Turkey for earthquake magnitude prediction before. Another deep learning method, namely long short-term memory (LSTM) can improve the accuracy of these predictions and help in taking more effective measures related to earthquakes. LSTM is a type of recurrent neural network that is used to model long-term dependencies in time series data. Earthquake data are typically represented as time series, and it is crucial to accurately analyze the patterns and relationships in this data. Earthquake magnitude prediction involves understanding the characteristics of earthquake data, detecting patterns, and making predictions about future magnitudes. LSTM, with its ability to understand the complex dynamics of time series data over time, can be an effective tool for earthquake magnitude prediction. The memory cells in LSTM can capture long-term dependencies and learn hidden relationships, which can be used to predict future earthquake magnitudes. The prediction of earthquake with high accuracy is tried to achieve by Wang et al. (2017) using LSTM. Berhich et al. (2021) used the LSTM to predict earthquake magnitudes using LSTM networks. The model was trained using data collected from different earthquake regions and utilized for predicting future earthquake magnitudes. The study evaluates the effectiveness and performance of LSTM in earthquake magnitude prediction. Al Banna et al. (2021) performed the LSTM to predict earthquakes in the Bangladesh region. On the other hand, the utilization of this method for analyzing earthquake magnitudes in Turkey remains unexplored.

Let us point out that CNN or LSTM-based deep learning methods alone may not be sufficient for earthquake magnitude prediction. It is often preferred to compare different methods and combine various data analysis techniques to obtain more accurate predictions. Therefore, it is important to evaluate the performance of CNN and LSTM in comparison with other methods such as SSA and ARIMA. Comparative studies of this nature can help assess which method is more effective in earthquake magnitude prediction. Therefore, in the study, the average magnitudes of earthquakes that occurred within the specified area in the Aegean Region were compiled on a monthly basis, covering the period from January 1970 to December 2020. The aim of this study is to identify the most suitable models, including ARIMA, SSA, CNN-based, and LSTM-based deep learning models, using average series data for accurate prediction of potential average earthquake magnitudes in this region. These models will be utilized to generate predictions based on the collected data.

The paper is structured as follows: the next section provides an overview of the study area and details the dataset preparation. Subsequently, the methods used in this study are introduced. Moving on to the third section, the selection of the optimal model for describing the dataset is discussed, with a comparative analysis of the results obtained from the ARIMA, SSA, CNN and LSTM-based deep learning models. Diagnostic tests are also performed in this section. Finally, the last section presents the concluding remarks.

Materials and methods

Materials

The study focuses on the Aegean Region and its surrounding 200 km buffer zone, as depicted in Fig. 1. This region, characterized by a complex tectonic structure, exhibits numerous independent fault lines. In his work, Şengör (1979a, b) proposed the existence of four distinct neotectonic provinces across Turkey: (1) the North Anatolian province, (2) the Eastern Anatolian contractional province, (3) the Central Anatolian 'Ova' province, and (4) the Western Anatolian extensional province. Each province displays unique tectonic characteristics.

The earthquake magnitude data used in this study were sourced from the "Artificial Intelligence and Probabilistic Model Based Earthquake Hazard Map" project, which is part of the TUBITAK 1001 initiative. To ensure the accuracy of the study's results, the earthquake data were obtained from reputable sources, including the Republic of Turkey Prime Ministry of Disaster and Emergency Management, the Presidential Authority of Earthquake Department (AFAD), and the Bogazici University Kandilli Observatory and Earthquake Research Institute Regional Earthquake Tsunami Monitoring Center (KOERI), as well as the Tan (2021) catalogs. Figure 2 illustrates the earthquakes that occurred in the region between the years 1970 and 2020.

To ensure consistency in the earthquake catalog used in this study, magnitude conversion relations from Kadiroğlu and Kartal (2016), KOERI, and Tan (2021) were applied. These conversions were necessary to homogenize the catalog and express all magnitudes on a single scale, specifically Moment Magnitude (Mw). Various magnitude types, such as Time-dependent Magnitude (Md), Local Magnitude (Ml), Surface Wave Magnitude (Ms), Body Wave Magnitude (Mb), and Moment Magnitude (Mw), are used to measure earthquake magnitudes. Among these, Moment Magnitude (Mw) is considered the most reliable and comprehensive measurement. Calculating the moment magnitude involves developing a mathematical model of the earthquake occurrence, making it a complex process. Typically, if the moment magnitude is known, other magnitude types are not calculated separately. The earthquake catalog was declustered to remove foreshocks and aftershocks using distance (r) and time (t) relations based on the work by Gardner and Knopoff (1974) and modified by Urhammer (1986). Subsequently, the

Fig. 1 Location map of the study area

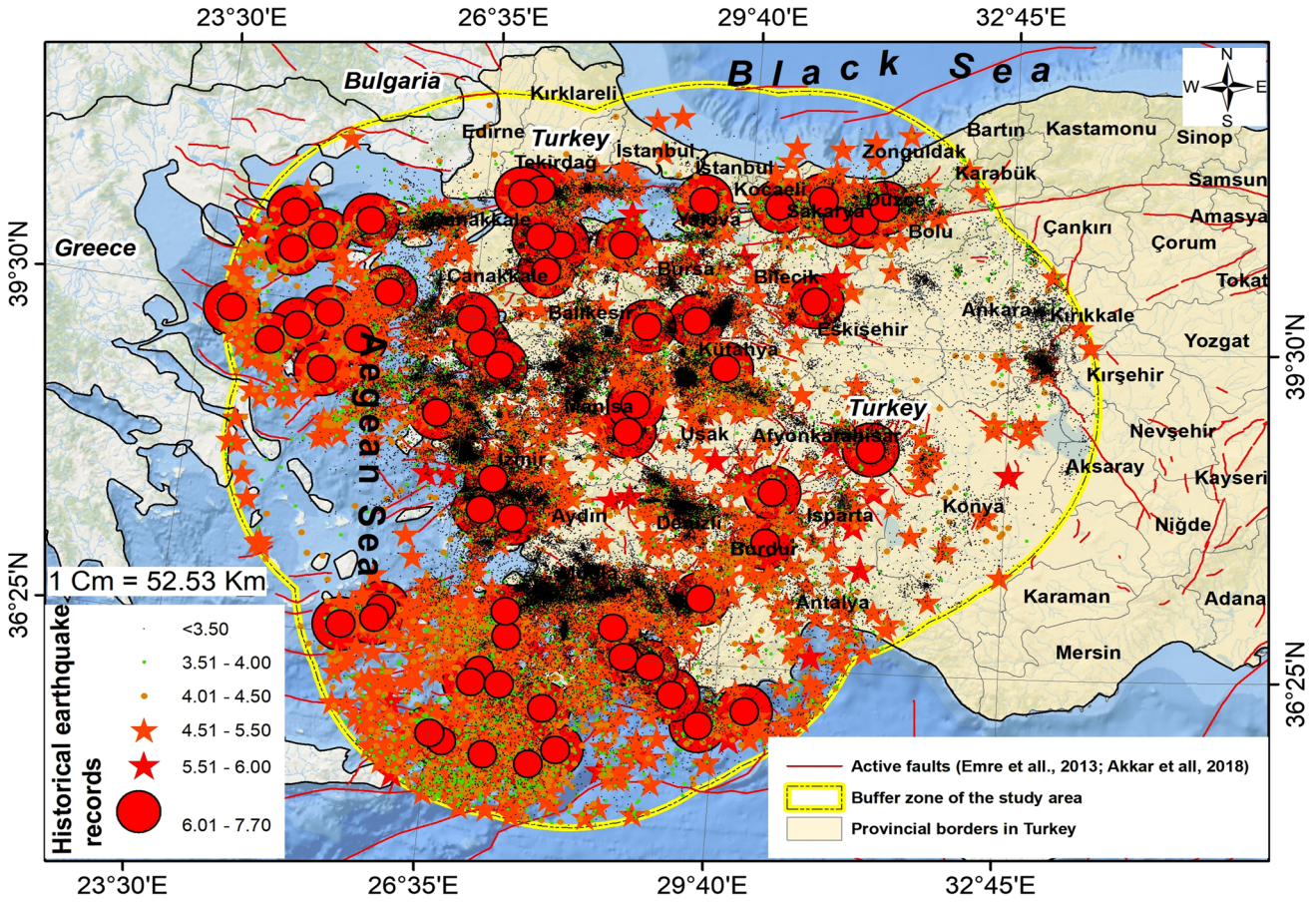


Fig. 2 Earthquakes that occurred between 1970 and 2020 in the region

earthquake magnitudes were compiled to cover the period between January 1970 and December 2020, and a series of monthly average earthquake magnitudes was obtained.

Table 1 presents the basic statistics derived from the earthquake magnitude data. The values Q1, Q2, and Q3 correspond to the 1st, 2nd, and 3rd quartiles, respectively. Analyzing the quantiles of the monthly average magnitudes of earthquakes in the region, it is observed that approximately 25% of the magnitudes are below 3.4. Furthermore, 50% of the magnitudes are less than 3.5, and 75% of the magnitudes are less than 4.2.

Figure 3 shows the histogram and scatter graphs of the monthly average earthquake magnitude data.

Based on Table 1 and Fig. 3, it is evident that the monthly average magnitudes of earthquakes in the Aegean Graben System range from a minimum of 3.1 to a maximum of 5.0, with an average magnitude value of 3.7. The distribution of the dataset is right-skewed, as indicated by the positive skewness value. Furthermore, the kurtosis value of 2.080, as shown in Table 1, suggests that the distribution of monthly average earthquake magnitudes is platykurtic, meaning it has lighter tails compared to a normal distribution. This indicates a higher risk of experiencing extremely large magnitudes.

The scatterplot in Fig. 3 shows that the observed pattern of no clear trend in the monthly average earthquake magnitudes in the Aegean region between 1970 and 1980, followed by a decrease from 1980 to 1990, can be explained by the following factors: the Aegean region is characterized by active tectonic movements. Fault lines, tectonic plate boundaries, and other geological features contribute to variations in earthquake magnitudes. The absence of a distinct trend during 1970–1980 may indicate the complex nature of geological features in the region that influence

Table 1 Basic statistics of monthly average earthquake magnitudes (Mw)

Min	Q_1	Q_2	Q_3	Max	Mean	Skewness	Kurtosis	St. Dev	Variance
3.144	3.377	3.548	4.155	4.967	3.745	0.679	2.080	0.452	0.205

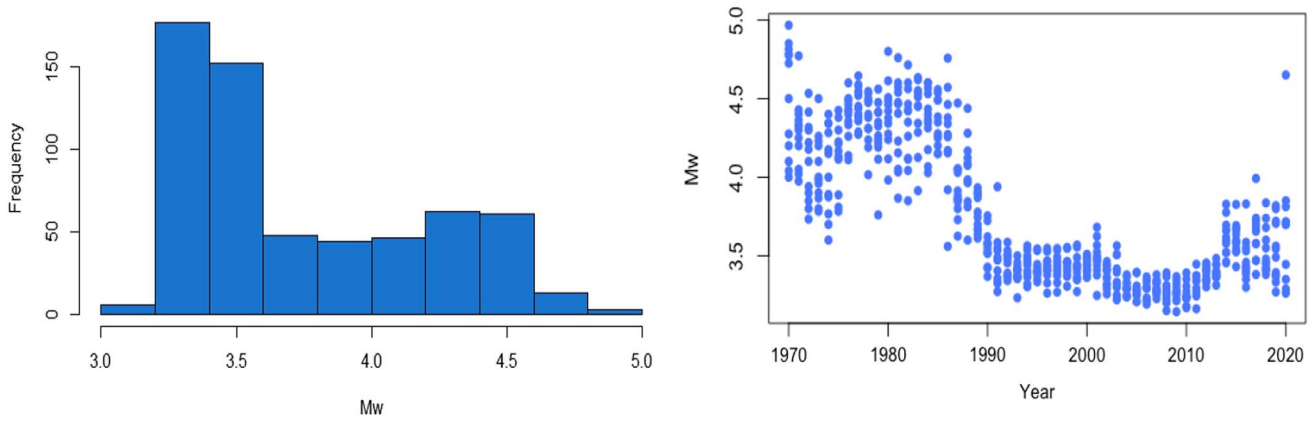


Fig. 3 Histogram (left side) and scatterplot (right side) for monthly average earthquake magnitudes

earthquake magnitudes. Furthermore, earthquakes occur when accumulated stresses in the Earth's crust are released. Stress distribution can vary based on the rate of tectonic movements, friction along plate boundaries, and the activity of faults. The decrease observed from 1980 to 1990 could suggest a change or reduction in stress distribution during that period. On the other hand, accurate and comprehensive collection of earthquake data is crucial. The lack of a clear trend during 1970–1980 might be attributed to limitations in data sources or monitoring systems during that period. In later years, advancements in technology and more extensive monitoring networks may have led to more precise and reliable data.

Methods

ARIMA (Box–Jenkins) method

In the study, we conduct time series analysis to obtain the prediction of earthquake magnitudes. ARIMA is a classical time series forecasting model that has been widely used for magnitude estimation. While there are newer methodologies such as machine learning and deep learning that have gained popularity, ARIMA still offers several incentives for magnitude estimation in certain contexts. ARIMA models are based on statistical assumptions and principles, which can provide a solid foundation for magnitude estimation. These models assume stationarity, which means that the statistical properties of the time series data do not change over time. This assumption can be useful in certain applications where stationarity is a reasonable assumption.

ARIMA models can also be effective when dealing with small or limited datasets. They can handle situations where the data have a short history or lacks the vast amounts of training data required by some machine learning or deep learning approaches. ARIMA models can capture patterns and dependencies in the data even with limited information. Furthermore, ARIMA models are specifically designed for time series data, which have inherent temporal dependencies. They take into account the autocorrelation and moving average components present in the data. This makes ARIMA models particularly useful when working with data that exhibit trends, seasonality, or other time-related patterns.

Figure 4 shows a flowchart for the steps of ARIMA method in general.

The Box–Jenkins method is a statistical forecasting technique commonly used for forward prediction and control of univariate time series. It assumes that the time series being analyzed consists of equally spaced observations and is stationary. However, in reality, time series often exhibit time-dependent changes in mean and variance due to trends, regular fluctuations, irregular fluctuations, and random fluctuations. To apply the Box–Jenkins methodology to these non-stationary time series, it is necessary to transform them into stationary series using various methods. Models applied to non-stationary series that are converted into stationary by taking differences are known as non-stationary linear stochastic models. These models combine elements of Moving Average (MA) and Autoregressive (AR) models applied to series with differencing of order 'd' (Kaynar and Taştan 2009). The representation of such models is denoted as ARIMA (p, d, q), where 'p' and 'q' represent the degrees of the AR and MA models, respectively, and 'd' represents the degree of differencing. ARIMA (p, d, q) model is defined as

$$\begin{aligned} Z_t =& \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \cdots + \Phi_p Z_{t-p} + \delta \\ &+ \varepsilon_t - \Theta_1 \varepsilon_{t-1} - \Theta_2 \varepsilon_{t-2} - \cdots - \Theta_q \varepsilon_{t-q}. \end{aligned} \tag{1}$$

The Box–Jenkins model follows a five-stage process to determine the appropriate model for a time series: model identification, parameter estimation, model diagnostic checking, model prediction, and forecasting. It is important that the time series being analyzed consists of observation values with equal time intervals. To apply the Box–Jenkins

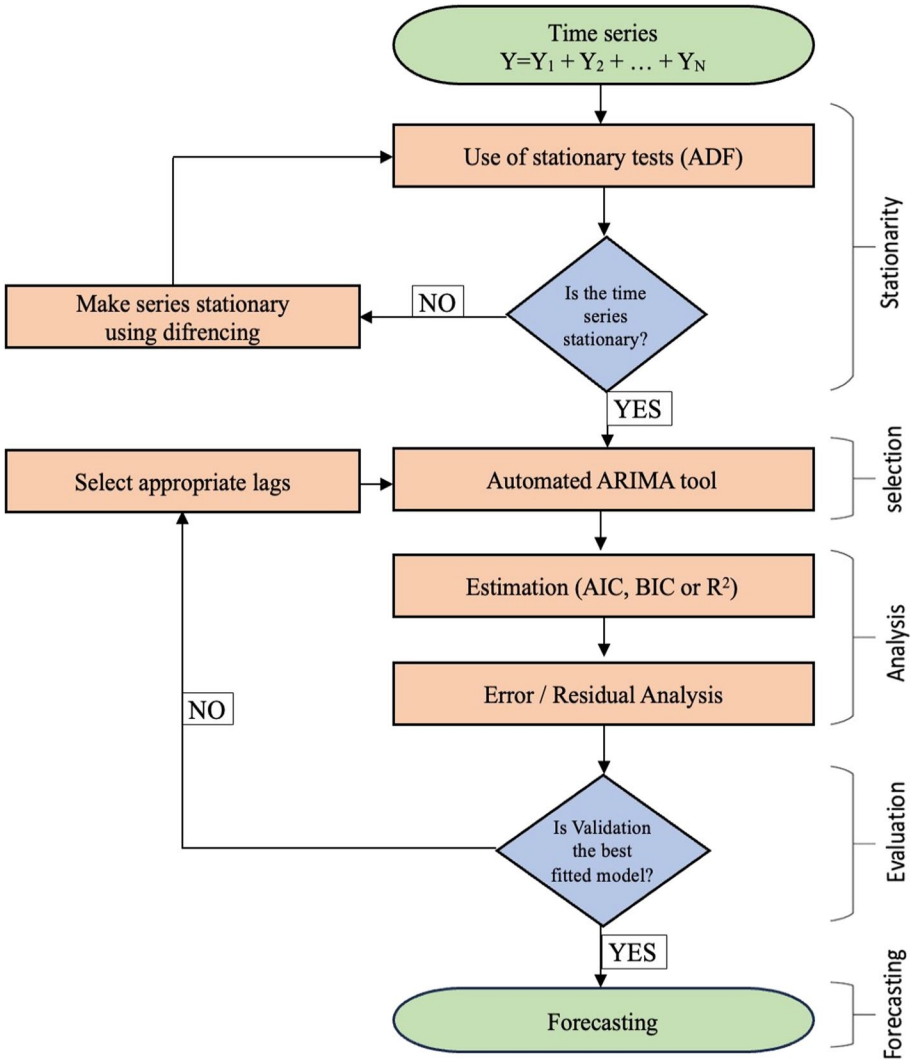


Fig. 4 Flow chart showing the time series analysis process (Zhai et al. 2020)

methodology, the series should be transformed into a stationary series by removing trends and seasonal fluctuations. The determination of the appropriate model for the series is based on the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the stationary series. If the series is stationary, the suitable models are represented as AR(p), MA(q), or the combination of both, known as the autoregressive moving average model (ARMA)(p, q). However, if the series exhibits a trend and is made stationary by differencing, the term I(d) is added to the model, resulting in the notation ARIMA(p, d, q) (Kadilar and Öncel Çekim 2020). By following this process, the Box–Jenkins model enables the identification and estimation of parameters, diagnostic checking of the model, and prediction and forecasting of the time series.

In the AR(p) model, the parameter "p" represents the order or degree of the autoregressive model. When determining the suitable model for a series, the ACF and PACF plots are analyzed to observe the pattern of correlation. In the PACF plot, if only the first lag shows a statistically significant correlation, while the correlations in the ACF plot gradually decrease as the number of lags increases, it indicates that the model is an AR(1) model. This means that in the AR(1) model, the relations in the ACF plot decrease gradually, while the relations of the first lag in the PACF plot are significant. By comparing the patterns of correlation decay in the ACF and PACF plots, the appropriate value of "p" can be determined for the AR model. AR(p) model is given by

$$Z_t = \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \cdots + \Phi_p Z_{t-p} + \varepsilon_t. \quad (2)$$

In the MA(q) model, the parameter "q" represents the order or degree of the moving average pattern. When analyzing the ACF and PACF plots to determine the suitable model, specific patterns can be observed for the moving average models. In the PACF plot, as the number of lags increases, the correlations gradually decrease in the moving

average models. However, in the ACF plot, this decrease is much faster. This pattern indicates that the relations of the first q lags in the ACF plot are significant for the MA(q) model. By examining the decay of correlations in the ACF and PACF plots, the appropriate value of "q" can be determined for the moving average model. The model is defined as follows:

$$Z_t = \varepsilon_t - \Theta_1\varepsilon_{t-1} - \Theta_2\varepsilon_{t-2} - \cdots - \Theta_q\varepsilon_{t-q}. \quad (3)$$

In the ARMA(p, q) model, both the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots are analyzed to determine the appropriate values for the parameters "p" and "q". For the AR component (p), the important relationships are identified by observing the first lags in the PACF plot. If there are significant correlations at the first p lags in the PACF plot, it suggests an autoregressive pattern of order p. Similarly, for the MA component (q), the significant relationships are determined by examining the first lags in the ACF plot. If there are significant correlations at the first q lags in the ACF plot, it indicates a moving average pattern of order q. By considering the decay of correlations in both ACF and PACF plots, the appropriate values for p and q can be determined for the ARMA(p, q) model. ARMA(p, q) model is given by

$$\begin{aligned} Z_t =& \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \cdots + \Phi_p Z_{t-p} \\ &+ \varepsilon_t - \Theta_1\varepsilon_{t-1} - \Theta_2\varepsilon_{t-2} - \cdots - \Theta_q\varepsilon_{t-q}. \end{aligned} \quad (4)$$

SSA (singular spectrum analysis) method

SSA is a powerful non-parametric method used for analyzing nonlinear time series data. It offers a flexible approach to decompose the original time series into distinct components, including trends, periodic patterns, and noise. This technique is applicable to a wide range of time series, such as annual, monthly, and hourly data, allowing researchers to recognize and predict critical information within these series. The process of SSA involves several steps. First, a trajectory matrix is generated based on the one-dimensional time series. Next, this matrix is decomposed and reconstructed, enabling the calculation of various components such as trend, period, and noise. By carefully analyzing the structure of the different constituent signals obtained through SSA, researchers can gain valuable insights into the underlying patterns and dynamics of the time series, facilitating prediction and forecasting tasks. Overall, SSA provides a robust framework for exploring and understanding nonlinear dynamics in time series data. Its non-parametric nature allows for flexibility and adaptability, making it a valuable tool in a wide range of applications.

The SSA flowchart in Fig. 5 shows two main stages: time series decomposition and reconstruction of the desired

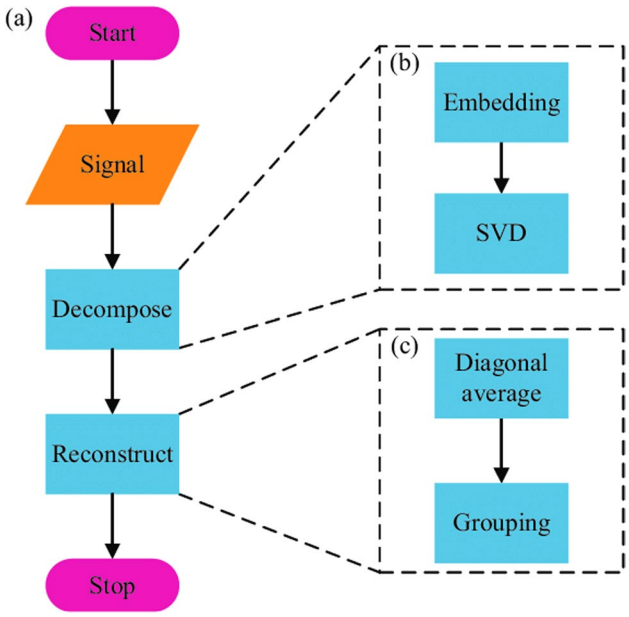


Fig. 5 Flowchart of the basic SSA architecture (Source: Cao et al. 2019)

additive constituents. In the decomposition stage, the algorithm first performs embedding, which maps the original time series to a sequence of multidimensional lagged vectors using a chosen window length. This forms a trajectory matrix. Then, the trajectory matrix undergoes singular value decomposition (SVD), which decomposes it into a sum of rank-one biorthogonal matrices. In the reconstruction stage, the algorithm involves two steps. The first step is grouping, where the matrices obtained from SVD are split into several groups and then summed up. This representation represents the trajectory matrix as a sum of resultant matrices. The second step is diagonal averaging, which transforms the resultant matrix into a time series that represents an additive constituent of the initial series. By following this procedure, the SSA algorithm decomposes the initial time series into several additive constituents. For a more in-depth understanding of the mathematical details of the SSA technique, researchers can refer to works such as Ghil and Vautard (1991), Golyandina et al. (2001), Schoellhamer (2001), Ghil et al. (2002), and Hassani (2007). These references provide comprehensive explanations and discussions on the mathematical aspects of the SSA technique.

The original series, with the length N of the series considered in the study assuming $\boldsymbol{Z} = \left[Z_1 Z_2 \cdots Z_N\right]^T$, the detailed steps of SSA are given as the following subsections:

a. Generating the trajectory matrix

The window length in SSA, denoted as L ($1 < L < N/2$), is typically determined experimentally and can vary depending on the specific problem and prior knowledge

of the time series. There is no universal rule for selecting the window length, as it depends on the characteristics of the data and the objectives of the analysis. However, it is often recommended to choose a window length that is a common multiple of 12. The choice of a common multiple of 12 is often motivated by the desire to capture seasonal patterns or periodicities in the data. By selecting a window length that aligns with the seasonal cycle (e.g., 12 months for monthly data), the SSA algorithm can effectively identify and extract seasonal components from the time series. It is important to note that the selection of the window length is a subjective decision and may require some experimentation and domain knowledge to find the most appropriate value for a given analysis.

The trajectory matrix Z is represented as follows (Gao et al. 2020; Yang et al. 2022):

$$Z = \begin{bmatrix} z_1 & z_2 & \cdots & z_{N-L+1} \\ z_2 & z_3 & \cdots & z_{N-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ z_L & z_{L+1} & \cdots & z_N \end{bmatrix} \quad (5)$$

b. Singular value decomposition (SVD)

Let identify $\lambda_i (i = 1, 2, \ldots, L)$ as a singular value. Separation of the matrix $\boldsymbol{Z}$ is built to achieve ZZ^T by decomposing the eigenvalue $\left(\lambda_1 \geq \ldots \geq \lambda_L \geq 0\right)$. $Z = Z_1 + Z_2 + \cdots + Z_L$ is expressed from this stage, describe as $Z_i = \sqrt{\lambda_i}\Lambda_i A_i^T$, $A_i = Z^T \Lambda_i \Big/ \sqrt{\lambda_i}$. Further, $\Lambda = diag\left(\lambda_1, \lambda_2, \ldots, \lambda_L\right)$ is the diagonal matrix and $\Lambda_1, \Lambda_2, \ldots, \Lambda_L$ is the *j*th eigenvector (Golyandina and Zhigljavsky 2013).

c. Restructuring

At this stage, using the group of v (number of eigenvalues) group eigenvectors obtained from TDA, a v-dimensional hyperplane of Z_j vectors is created in an L-dimensional space. This process involves determining the optimal window length, L, and the number of principal constituents, v. It is recommended to choose a window length that is a multiple of 12. In this process, cyclic sorting and series separation techniques are applied to each dimension of the window length. Furthermore, the W-correlation matrix is used to identify strong relationships among the principal constituents (v), which are then grouped together. The hyperplane is then decomposed into constituents such as trend, error, and cyclical movements. By utilizing these constituents, estimation and prediction values can be obtained. It is important to note that the choice of window length and the determination of principal constituents require careful consideration and may involve trial and error, as well as domain knowledge.

CNN (convolutional neural network) method

CNN is a deep learning method widely used for analyzing various types of data, including time series data. While CNNs are commonly associated with image processing tasks, they have also proven to be effective for analyzing text, audio, and time series data. When applied to time series analysis, CNNs aim to discover intricate relationships, identify patterns, and extract meaningful features that represent these patterns. By utilizing CNNs in time series analysis, more accurate predictions can be made, and significant changes within the time series can be detected. Several studies, such as those conducted by Zhao et al. (2017), Liu et al. (2018), and Li et al. (2021), have demonstrated the successful application of CNNs in time series analysis. The architecture of a CNN involves sequential layers that process the input data. These layers typically include convolutional layers, activation functions, pooling layers, and fully connected layers. Convolutional layers play a crucial role in feature extraction by applying filters to the data, enabling the network to recognize patterns and capture important features within the time series. Overall, CNNs provide a powerful approach for analyzing time series data by leveraging their ability to discover complex relationships and extract meaningful features, ultimately leading to improved analysis and prediction capabilities.

In earthquake magnitude prediction, a CNN model is utilized to estimate the magnitude of earthquakes. The CNN model takes preprocessed or transformed versions of earthquake data as input and performs feature extraction operations on this data. These operations involve applying convolutional filters to the data, enabling the network to identify spatial and temporal patterns within the earthquake signals. The preprocessed earthquake data can include various types of information, such as seismic waveforms, frequency spectra, or spectrograms. By representing the earthquake data in these different forms, the CNN model can capture different aspects of the seismic signals and extract relevant features.

The multi-layered structure of the CNN allows for the extraction of hierarchical representations of the earthquake data. The initial layers of the CNN learn lower level features, such as edges or corners, while deeper layers learn higher level features that capture more complex patterns and correlations. This hierarchical representation learning enables the CNN to effectively model the complexity of earthquake data. Moreover, the parameters of the CNN, such as the size of convolutional filters, the number of filters, and the architecture of the network, can be customized based on the specific characteristics of the earthquake data and the desired prediction task. By appropriately tuning these parameters,

the CNN model can optimize its performance in earthquake magnitude estimation. In summary, CNNs provide a powerful framework for earthquake magnitude prediction by leveraging their ability to extract features from earthquake data and model the underlying patterns. The customizable nature of CNNs allows for flexibility in capturing the complexity of earthquake signals, leading to improved accuracy in magnitude estimation.

The foundations of the method based on the visual perception ability of living beings were established by LeCun et al. (1989), and later developed by LeCun et al. (1998). Although various CNN architectures exist in the literature, they share similar basic structures, typically comprised of three types of layers: convolutional, pooling, and fully connected (Ramanjaneyulu et al. 2018; Gu et al. 2018). The convolutional and pooling layers are responsible for extracting features from the input data, while the fully connected layer aids in classifying these extracted features.

As seen in Fig. 6, a typical CNN structure includes multiple iterations of convolutional layers, followed by a pooling layer, and finally one or more fully connected layers (Yamashita et al. 2018). This layered approach allows the network to progressively extract and learn complex features from the input data. The convolutional layers apply filters to capture patterns and local dependencies in the data, while the pooling layers down sample the output, reducing the spatial dimensions and extracting the most salient information. The fully connected layers connect all the neurons of the previous layer to the next layer, enabling classification or regression based on the learned features. Overall, CNNs have become a widely adopted method due to their ability to effectively extract features and learn hierarchical representations from data. The combination of convolutional, pooling, and fully connected layers allows CNNs to capture complex patterns and relationships, making them particularly effective in tasks such as image recognition, natural language processing, and, as in the case mentioned, earthquake magnitude prediction.

The input layer of a CNN receives the input information in the form of $N \times k$, where k represents the number of variables in the input time series, and N represents the length of each univariate series (Zhao et al. 2017). Following the input layer, the convolutional and pooling layers come into play as data preprocessing layers. These layers filter the input data and extract relevant information to be fed into the fully connected layer (Livieris et al. 2020). The convolutional layer plays a crucial role in defining the input's properties. It typically involves a combination of linear and nonlinear operations, along with an activation function (Yamashita et al. 2018). Through various calculations, the convolutional layer determines the output based on the input. In addition, rectified linear units (ReLUs) are often applied as activation functions, such as sigmoid, to the output of the previous layer's activation. ReLUs help accelerate learning by introducing non-linearity (O'Shea and Nash 2015). In short, the input layer receives the input data, and subsequent convolutional and pooling layers preprocess the data to extract important features. These features are then passed on to the fully connected layers for further analysis and prediction in the CNN architecture.

The pooling layer plays a crucial role in downsampling and reducing the complexity of the CNN architecture (Albawi et al. 2017; Zhao et al. 2017). It gradually decreases the model's size, effectively reducing the number of parameters and computational complexity (O'Shea and Nash 2015). Among the various pooling operations available, Max-pooling is the most widely used method for nonlinear downsampling. Typically, CNNs employ a 2×2 filtered pooling layer implemented in two steps. Positioned between sequential convolutional layers, the pooling layer helps capture important features while reducing spatial dimensions (Ramanjaneyulu et al. 2018). Importantly, pooling layers do not contain any learnable parameters (Yamashita et al. 2018). Following the pooling layers, the fully connected layer comes into play. This layer establishes direct connections with all activations from the previous layer, similar

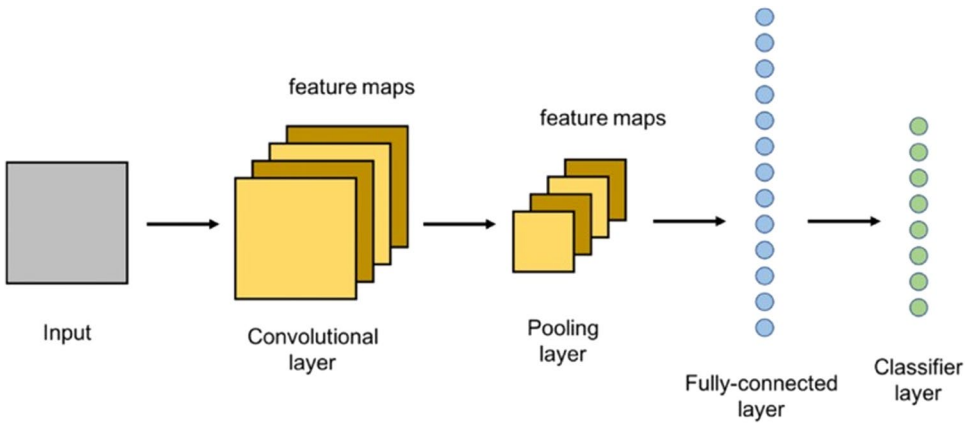


Fig. 6 Flowchart of the basic CNN architecture (O'Shea and Nash 2015)

to traditional neural networks (Albawi et al. 2017; Ramanjaneyulu et al. 2018). Acting as a classification layer, it processes the features learned from the convolutional and pooling layers, transforming them into a format suitable for prediction or classification (Yang and Li 2017). To enhance the performance of the CNN architecture, ReLU activation functions can be employed between these layers. However, one drawback of the fully connected layer is that it often contains a significant number of parameters, requiring computational resources for processing (Albawi et al. 2017).

LSTM (long short-term memory) method

The LSTM neural network, a variant of recurrent neural networks (RNN), has gained significant attention since its introduction by Hochreiter and Schmidhuber (1997). Over the years, researchers have proposed various approaches to enhance the LSTM model, leading to its widespread application in time series prediction problems (Cho et al. 2014; Gers and Schmidhuber 2000; Graves and Schmidhuber 2005; Schmidhuber et al. 2007; Zaremba et al. 2014; Yu et al. 2019). Its importance in earthquake magnitude prediction is evident through its utilization in fields such as earthquake risk assessment, disaster management, and structural design. In time series forecasting, it is crucial to consider not only the latest data but also the previous data to capture dependencies and patterns (Cao et al. 2019). While traditional RNN models, such as Elman (1990), can handle long-term dependencies, they often face challenges due to the issue of vanishing gradients, which can hinder or even halt the learning process. LSTM addresses this problem by incorporating memory cells that store relevant information and gates that control the flow of information, making it more effective than classic RNNs (Livieris et al. 2020).

The core idea behind LSTM lies in its cell state and gate structure (Zha et al. 2022). Each LSTM unit consists of three primary gates: an input gate that determines whether new input should be allowed, a forget gate that discards irrelevant information, and an output gate that regulates the information to be passed forward (Livieris et al. 2020; Siami-Namini et al. 2019). The cell state, represented by a horizontal line passing through the LSTM unit, carries information throughout the network (Yadav et al. 2020). The overall structure of LSTM is illustrated in Fig. 7. To summarize, LSTM's ability to capture long-term dependencies, store relevant information in memory cells, and regulate the flow of information through gates makes it a powerful tool for modeling time series data, including earthquake magnitude prediction.

(i) The Input Gate: The input gate in LSTM plays a crucial role in deciding whether to incorporate new information into the memory. It consists of two key components: a sigmoid layer and a tanh layer. The sigmoid layer computes the values between 0 and 1, representing the significance of each element to update. On the other hand, the tanh layer generates a vector of potential new values that could be added to the memory. The outputs of sigmoid and *tanh* layers are, respectively, calculated as follows:

$$i_t = \sigma\left(W_{i_h}\left[h_{t-1}\right] + W_{i_x}\left[X_t\right] + b_i\right), \tag{6}$$

$$\widetilde{c_t} = \tanh\left(W_{c_h}\left[h_{t-1}\right] + W_{c_x}\left[X_t\right] + b_c\right). \tag{7}$$

Here, i_t determines whether the value should be updated and $\widetilde{c_t}$ specifies a vector of new candidate values to be added

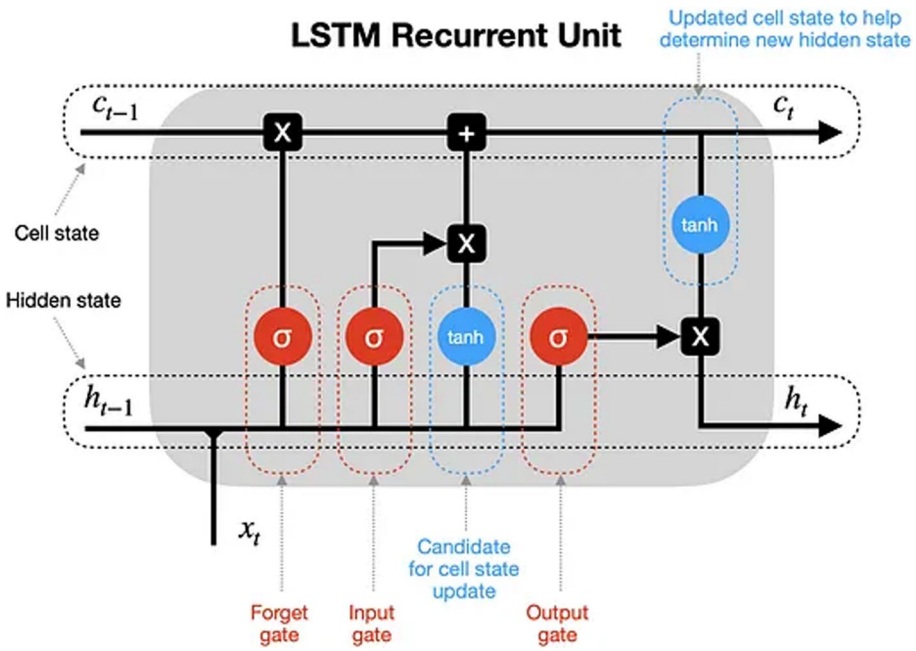


Fig. 7 Flowchart of a basic LSTM architecture (Dobilas 2022)

to LSTM memory. The sigmoid function σ squashes the values between 0 and 1, while the tanh function (tan$h$) ensures that the values lie between − 1 and 1. The outputs of the sigmoid layer and the *tanh* layer are then combined to compute the new information that will be added to the memory cell:

$$c_t = f_t \odot c_{t-1} + i_t * \widetilde{c}_t. \tag{8}$$

Here, ⊙ denotes element-wise multiplication, and f_t represents the forget gate output. The forget gate determines how much of the previous memory cell state should be retained.

(ii) The Forget Gate: The forget gate in LSTM is responsible for determining the relevant information to retain from the LSTM memory. It makes this decision by considering the values of h_{t-1} and X_t. The output of the gate is denoted as f_t and lies between 0 and 1. A value of 0 signifies that the corresponding information should be completely discarded, while a value of 1 indicates that the information should be retained. The calculation for the output of the forget gate is as follows:

$$f_t = \sigma\left(W_{f_h}\left[h_{t-1}\right] + W_{f_x}\left[X_t\right] + b_f\right). \tag{9}$$

Here, W_{f_x} and W_{f_h} are weight matrices, and b_f is the bias term associated with the forget gate. The sigmoid function σ squashes the input values between 0 and 1, determining the extent to which the information should be forgotten.

The output of the forget gate f_t is then used to modulate the previous memory cell state c_{t-1}. It performs element-wise multiplication (⊙) with c_{t-1} to determine which information to retain and which to forget:

$$c_t = f_t \odot c_{t-1}. \tag{10}$$

By applying this forget gate mechanism, the LSTM can selectively retain important information from previous time steps while discarding irrelevant or outdated information. This adaptive memory management enables the LSTM to effectively capture long-term dependencies in time series data and make accurate predictions.

(iii) The Output Gate: This gate plays a crucial role in determining which portion of the LSTM memory contributes to the final output. It starts by utilizing a sigmoid layer to determine the extent of contribution. Subsequently, a non-linear tanh function is employed to map values within the range of -1 to 1. Finally, the result is multiplied by the output of a sigmoid layer. The operations involved in this process can be expressed as

$$o_t = \sigma\left(W_{o_h}\left[h_{t-1}\right] + W_{o_x}\left[X_t\right] + b_o\right), \tag{11}$$

$$h_t = o_t \odot \tanh\left(c_t\right). \tag{12}$$

Here, i_t is the vector of the input gate, f_t is the vector of the forget gate, ot is the output gate, $\widetilde{c}_t$ is the vector of the candidate values added to the new cell state, and c_t is the vector of the memory cell. W_i, W_f, W_o and W_c are weight matrices and b_i, b_f, b_o and b_c are bias vectors. σ denotes the sigmoid function and finally, h_t is the output matrix. The final output h_t is the result of element-wise multiplication between the output gate o_t and the hyperbolic tangent of the current memory cell state (c_t). This output represents the information that the LSTM model decides to pass on to the subsequent time step or use for prediction (Siami-Namini et al. 2019).

Results

We created a homogenized earthquake catalog specifically for the region of Turkey, focusing on earthquakes with magnitudes equal to or greater than Mw ≥ 3. The selected region spans from 35 to 42° N latitude and 23 to 32° E longitude. To predict earthquake magnitudes, we utilized the monthly average magnitudes from this catalog. To accomplish this, we constructed a time series dataset and performed various tasks such as modeling the time series, conducting exploratory analysis, and carrying out prediction processes. These tasks were executed using the "Rssa", "forecast", "TensorFlow", and "Keras" packages, which are popular and widely used open-source software environments for statistical computing and graphics in both R and Python. By leveraging these tools and methodologies, we aimed to gain insights into the earthquake magnitudes and develop predictive models to forecast future magnitudes. This approach enabled us to analyze and interpret the patterns and dynamics of earthquake magnitudes in the designated region of Turkey, ultimately contributing to a better understanding and assessment of seismic activity in the area.

In this study, we use the monthly average of earthquake magnitudes since it is a commonly used tool in time series analysis, providing a general overview of seismic activity and helping assess earthquake risk. Taking the monthly average in a time series helps provide a summary of seismic activity during a specific period and enables the observation of longer-term trends or changes. Earthquake data in a time series can often be noisy and subject to instantaneous fluctuations. The monthly average can be used to smooth out these fluctuations, allowing us to see a more stable underlying trend. Furthermore, in certain regions, seismic activity may vary seasonally. The monthly average can be used to identify seasonal variations. For example, some areas may experience more earthquakes during the winter months. This information is important for assessing earthquake risk and preparedness strategies. The monthly average can also help observe long-term trends. Over a large time, span (years

or decades), changes in monthly average magnitudes can assist in determining a long-term increase or decrease in earthquake activity in a specific region. Monthly average magnitudes may also be more suitable for other statistical analysis. They can be used to assess trends, variances, or other statistical properties and to develop models.

The time series plot in Fig. 8 displays the monthly average earthquake magnitudes. From the graph, we observe that the series does not exhibit a distinct trend between 1970 and 1980. However, it demonstrates a declining pattern from 1980 to 1990.

ARIMA model

The ACF and PACF graphs, displayed in Fig. 9, offer a more comprehensive analysis of the series. These graphs are useful in identifying the suitable model for the ARIMA method. In this case, Fig. 9 indicates the presence of a trend in the series, which can provide valuable insights for further modeling and forecasting.

The graph displayed in Fig. 10 represents the first difference of the series, along with the corresponding ACF and PACF. This graph reveals that the series becomes trend-free and stationary after taking the first difference. Consequently, we determine that the parameter d should be set to 1, as it achieves stationarity. Upon examining the graphs in Fig. 10 to determine the appropriate values for the p and q parameters, we observe that the correlation decreases significantly between the first and second lags in the ACF graph, whereas the decrease in the PACF graph is less pronounced. This observation suggests that the model is a Moving Average (MA) model with $p=0$. To determine the value of the q parameter, we need to assess the significance of the relationships in the initial lags depicted in the ACF graph. Since only the relationship at the first lag appears to be significant in the ACF graph, we set $q=1$. Hence, in the initial stage, the most suitable model for the series is identified as ARIMA (0,1,1). The results for the ARIMA (0,1,1) model are presented in Table 2.

Table 2 presents the estimation results for the ARIMA (0,1,1) model. The estimated coefficient for the MA1 term

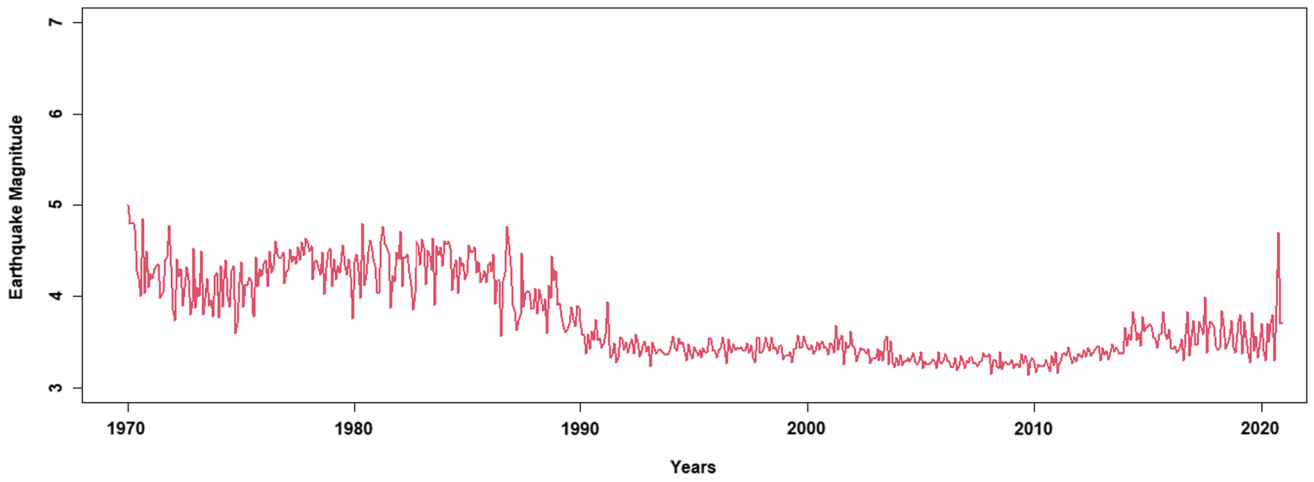


Fig. 8 Time series graph of monthly average earthquake magnitude series

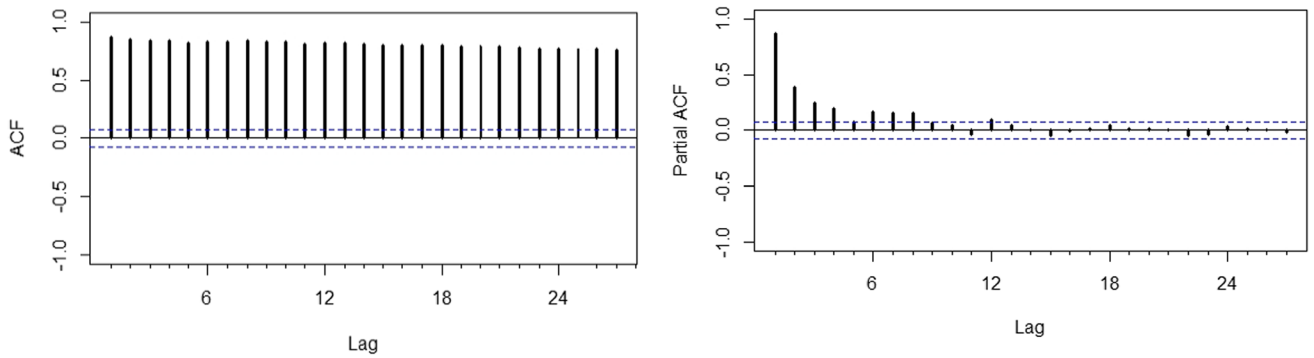


Fig. 9 The ACF (left side) and PACF (right side) graphs of the monthly average earthquake magnitude series

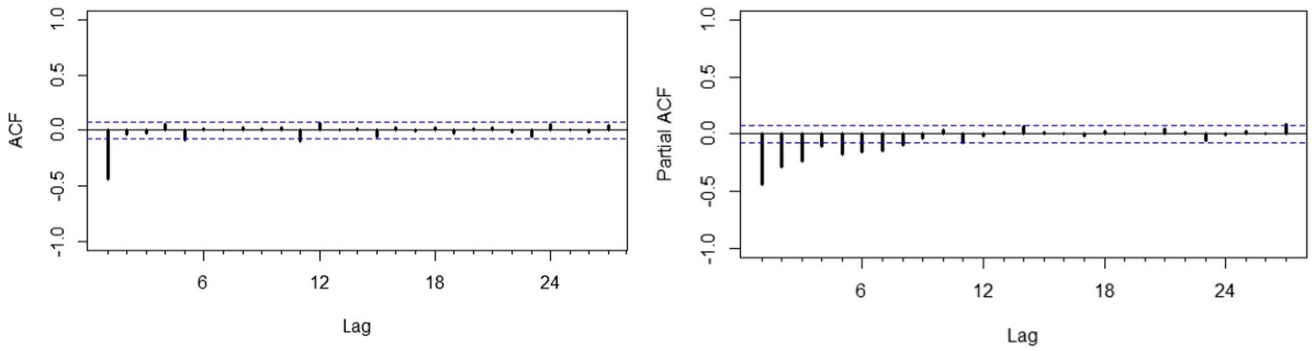


Fig. 10 First-order difference ACF (left side) and PACF (right side) graphs of the monthly average earthquake magnitude series

Table 2 The results for the ARIMA (0,1,1) model

	Estimation	Standard Error	*Z*-value	*P*-value
MA1	−0.8218497	0.0234982	−34.9751	<0.00
Constant	−0.0014418	0.0013076	−1.1026	0.2702

is −0.8218, with a standard deviation of 0.0235. The corresponding *z*-value is −34.9751, and the *P*-value is approximately 0, indicating that the MA1 term is statistically significant. Hence, the selected model is considered suitable for the series. In addition, the Akaike Information Criterion (AIC) value is −354.77, and the Bayesian Information Criterion (BIC) value is −341.53. By testing various alternative models, it was determined that the ARIMA (0,1,1) model exhibits the most favorable AIC and BIC values for the series. Consequently, based on the results presented in Table 2, the model for the monthly average earthquake magnitude series is derived as follows:

$$Z_t = Z_{t-1} - 0.0014 + e_t + 0.8218e_{t-1}. \quad (13)$$

SSA model

The first step in determining the optimal model in the SSA method involves setting a window length. The window length (L) should be chosen in proportion to the size (T) of the cyclic component. In the case of monthly data, where $T=12$, the window length should be a multiple of 12. In long-term series, the maximum value for L is typically $N/2$. In our study, the dataset size is $N=612$, so a value of $L=12$ M is determined, where M is an integer. Figure 11 displays the plots for the W-correlation matrix.

The W-correlation matrix plays a crucial role in determining the window length and identifying the principal constituents, namely the essential rebuild constituents matrix. When two reconstructed core constituents (v) exhibit a high correlation, it indicates a strong relationship between them. Figure 11a to d display the correlation matrices, illustrating that the relationships become stronger as the L value increases. Consequently, a window length of 240 is chosen. Figure 11e to h indicates that the most significant values of the principal constituents are observed at 12. Hence, the optimal model is determined with the parameters $L=240$ and $v=12$. The resulting SSA (240.12) model estimates the three constituents through the separation process, as depicted in Fig. 12a to c. These constituents represent the trend, seasonal, and error components, respectively.

As anticipated, the error constituent is characterized by values close to zero. Subsequently, the predicted values of all constituents are combined through an integrated rebuilding process, as described in the diagnostic test conducted in "The diagnostic test of models".

CNN model

The primary objective of using CNN in time series analysis is to uncover intricate relationships within the data, identify patterns, and extract relevant features that represent those patterns. By applying the CNN method to our dataset, which comprises the temporal changes of earthquake magnitudes, we can emphasize features such as trends, seasonality, and fluctuations in the series. The analysis of our dataset using the CNN method involves the following steps:

- Step 1 (data preparation): The dataset is divided into 70% training data and 30% test data, and Min–Max normalization is applied.
- Step 2 (model architecture): The activation function ReLU is utilized, and a max pooling method with a stride of 2 is applied in the pooling layer.
- Step 3 (training process): The training data are processed with 1000 epochs, a batch size of 32, and an optimization algorithm using mean squared error as the loss function.
- Step 4 (results): Predictions, forecasts, and confidence intervals are computed.

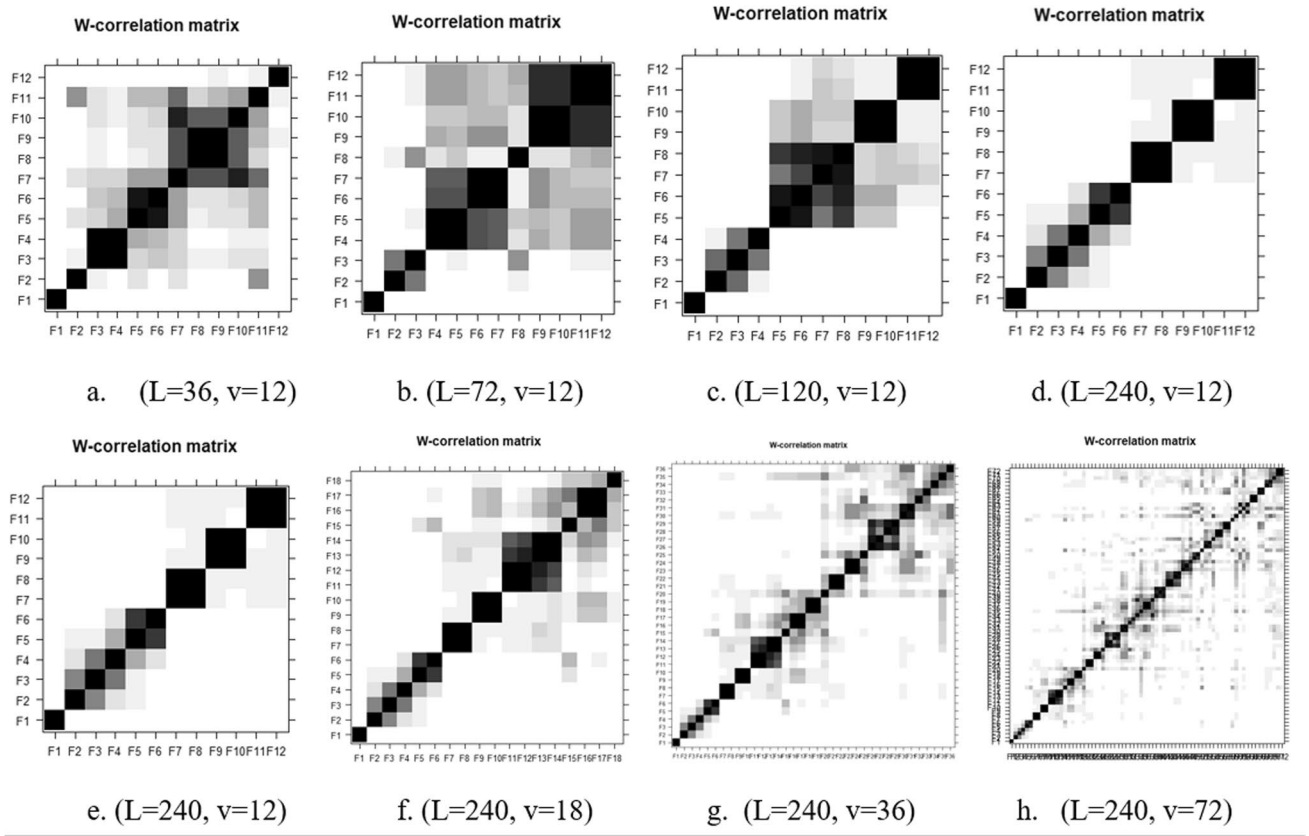


Fig. 11 Plots for W-correlation matrix

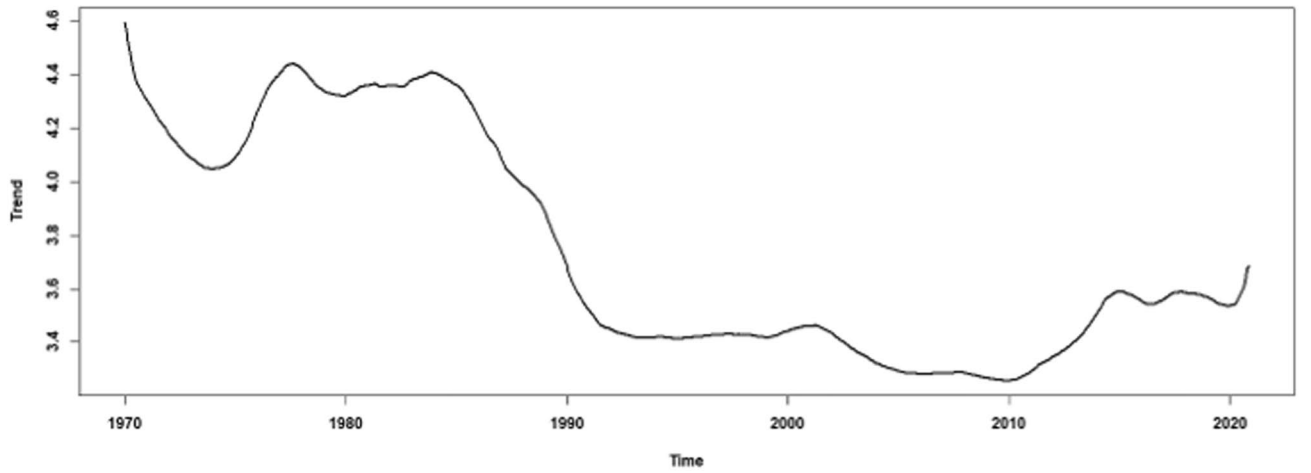
It is important to highlight that CNN does not depend on the traditional assumptions of statistical model validity. Instead, these methods are guided by data-centric learning principles and aim to investigate complex relationships within the data. Their focus is on uncovering patterns and extracting meaningful features from the data rather than conforming to specific statistical assumptions.

LSTM model

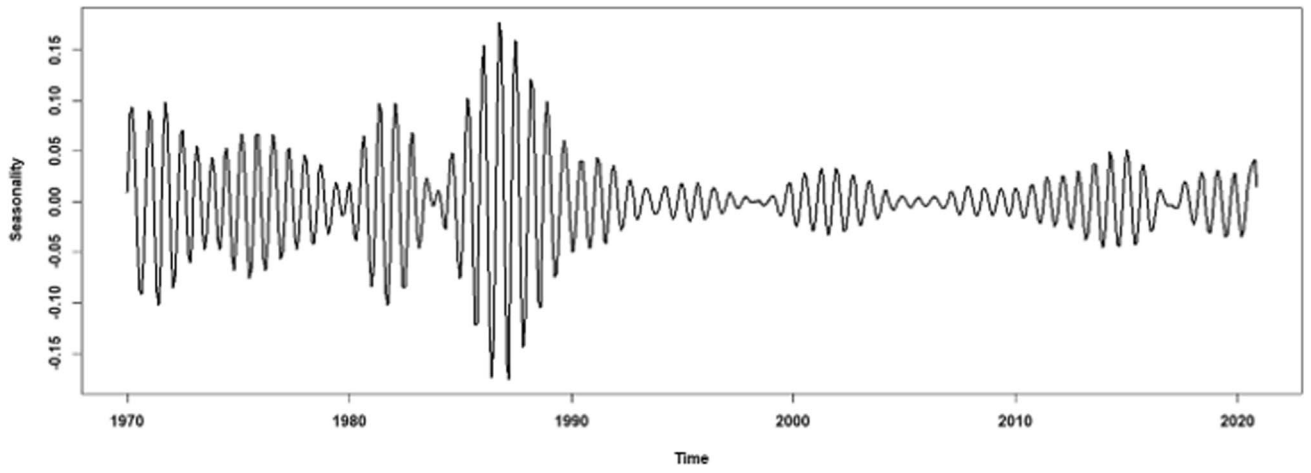
LSTM, as a deep learning method, is well suited for capturing time-dependent structures by analyzing past observations in a time series. In the case of earthquake data, which exhibit repeating patterns due to tectonic features, LSTM can effectively capture the relationships between past and present earthquake occurrences. Its ability to learn hidden patterns and relationships within the data enables it to make predictions for future time steps (Lakshmi and Tiwari 2009). The analysis steps for applying the LSTM method to our earthquake dataset are as follows:

- Step 1 (data preparation): The dataset is divided into 70% training data and 30% test data. Then, Min–Max normalization is applied to ensure that the data falls within a specific range.
- Step 2 (model architecture): The LSTM model employs the activation function *Tanh* and the parameter "return_sequences = True", allowing it to preserve the sequence of outputs.
- Step 3 (training process): The training data are processed using an optimization algorithm, such as stochastic gradient descent, with an epoch number of 1000. The batch size is set to 128, and the mean squared error is chosen as the loss function.
- Step 4 (results): The trained LSTM model can then be used to generate predictions, forecasts, and confidence intervals for future time steps.

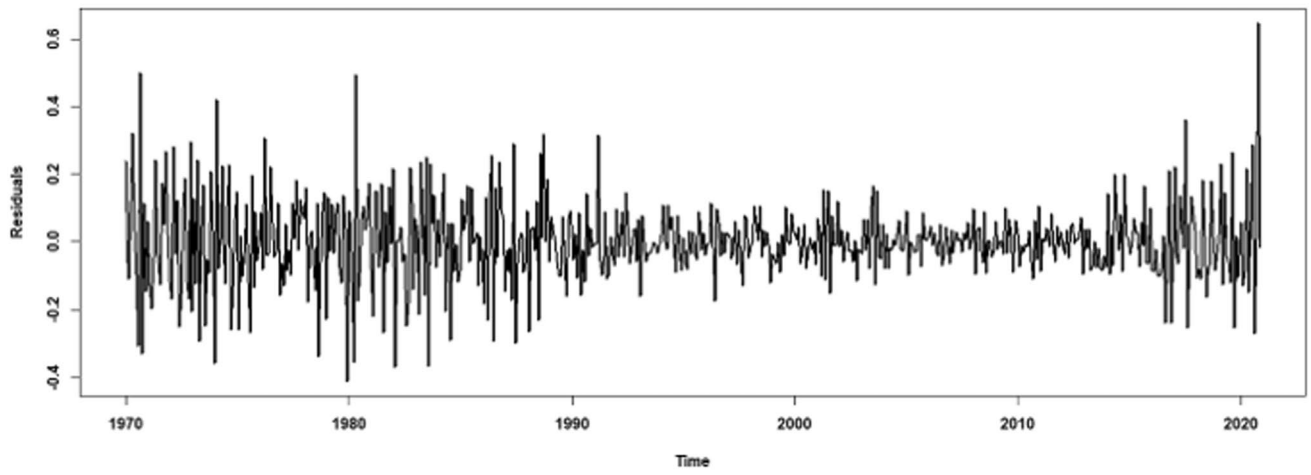
Deep learning methods, including LSTM, operate on the principle of automatically extracting high-level features from the data and making predictions based on these features. Unlike the CNN method, the LSTM method does not require any specific assumptions for its validity, as it relies on the data-driven learning approach to uncover patterns and relationships in the time series data.



a. Trend constituent of the series



b. Seasonality constituent of the series



c. Error constituent of the series

Fig. 12 Graphs for the trend, seasonal and error constituents of the series

The diagnostic test of models

Performing diagnostic tests is a crucial step in time series modeling. Therefore, in this subsection, we present a graphical comparison of the forecast values generated by the ARIMA, SSA, CNN, and LSTM models, along with the original series. In addition, we analyze the ACF and PACF of the error series.

Figure 13 displays the graph of the original series and the forecast series generated by the ARIMA model. This graph allows for a visual assessment of the performance of the ARIMA model in capturing the patterns and trends present in the original series.

Figure 14 depicts the graph of the original series and the forecast series generated by the SSA model. From Fig. 14, it can be observed that the forecasted components align well with the trend of the earthquake magnitude series. The SSA model successfully captures the fluctuations and variations within the series, particularly during the period from 1970 to 1990.

Figures 13 and 14 demonstrate a notable alignment in the trend between the earthquake magnitude series and the prediction series generated by the ARIMA and SSA models, respectively. However, further analysis is required to assess the accuracy of these predictions. To this end, the ACF and PACF graphs of the errors from the ARIMA model, as depicted in Fig. 15, are provided. These graphs allow us to examine the autocorrelation and partial autocorrelation of the residuals, providing insights into any remaining patterns or correlations in the model's errors. The ACF and PACF graphs displayed in Fig. 15 indicate that the error series exhibits characteristics of a white noise process. This implies that the ARIMA (0, 1, 1) model is statistically valid and reliable for modeling the monthly average earthquake magnitude series. These findings validate the suitability of the chosen model for accurately capturing the underlying patterns and dynamics of the data.

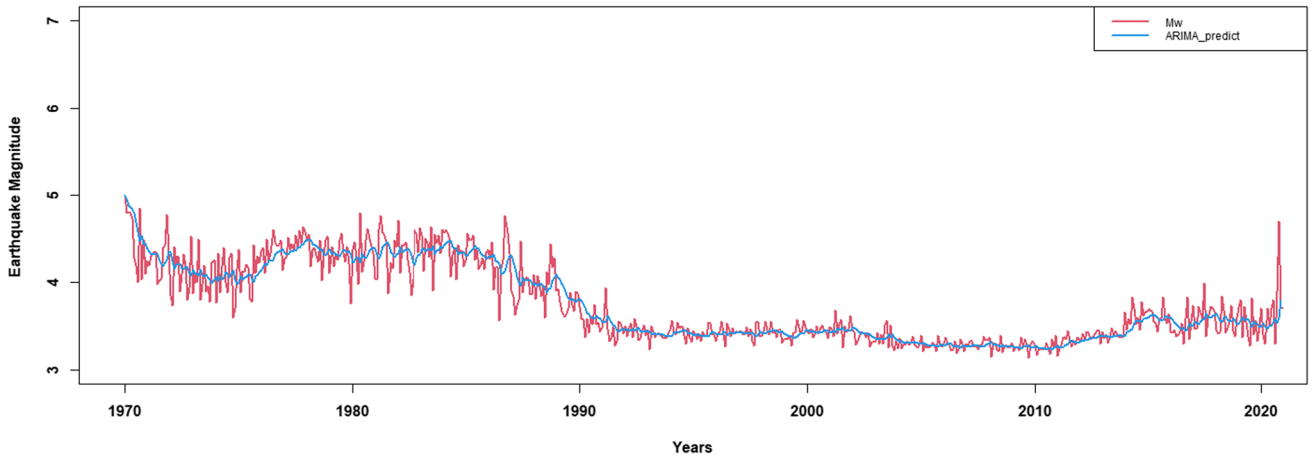


Fig. 13 The graph of original series and ARIMA forecast series

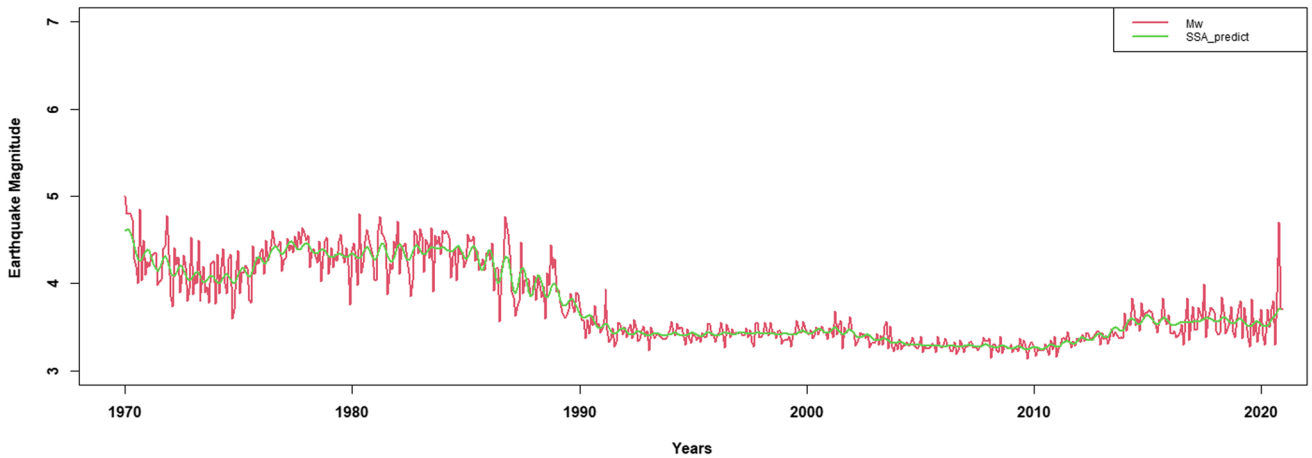


Fig. 14 The graph of original series and SSA forecast series

Figures 16 and 17 depict the graphs of the original series alongside the prediction series generated by the CNN and LSTM models. It is evident that the CNN and LSTM predictions exhibit a stronger agreement with the magnitude series compared to the ARIMA and SSA predictions. The CNN and LSTM models successfully

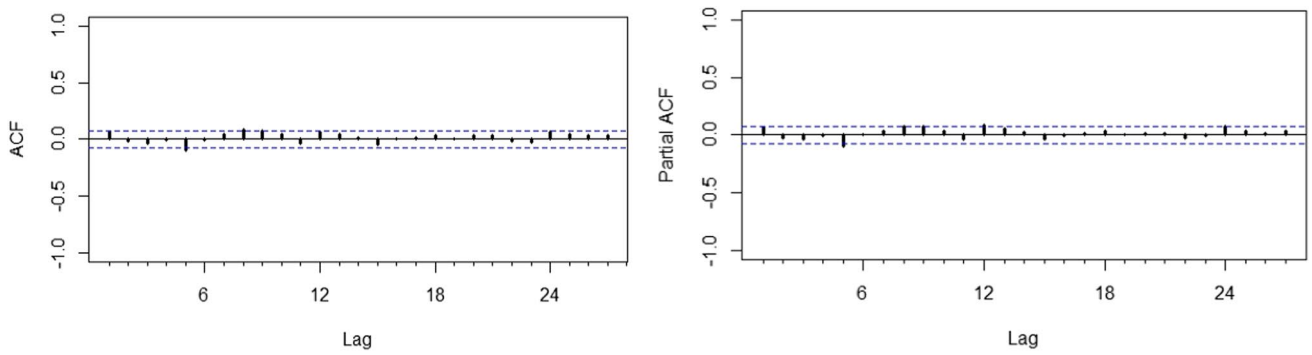


Fig. 15 The ACF (left side) and PACF (right side) graphs for the errors of the generated ARIMA model

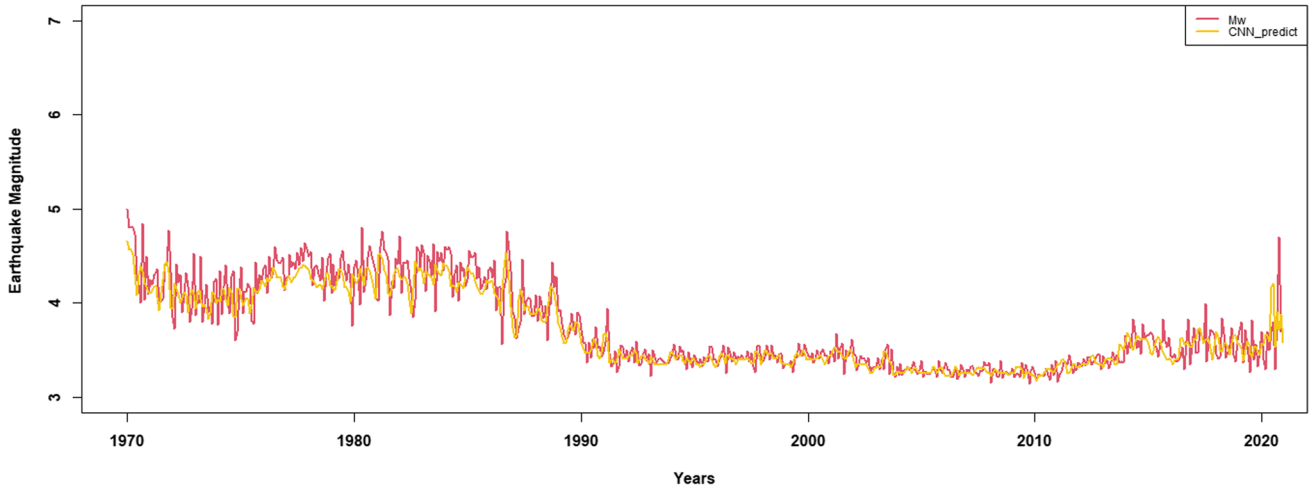


Fig. 16 The graph of original series and CNN forecast series

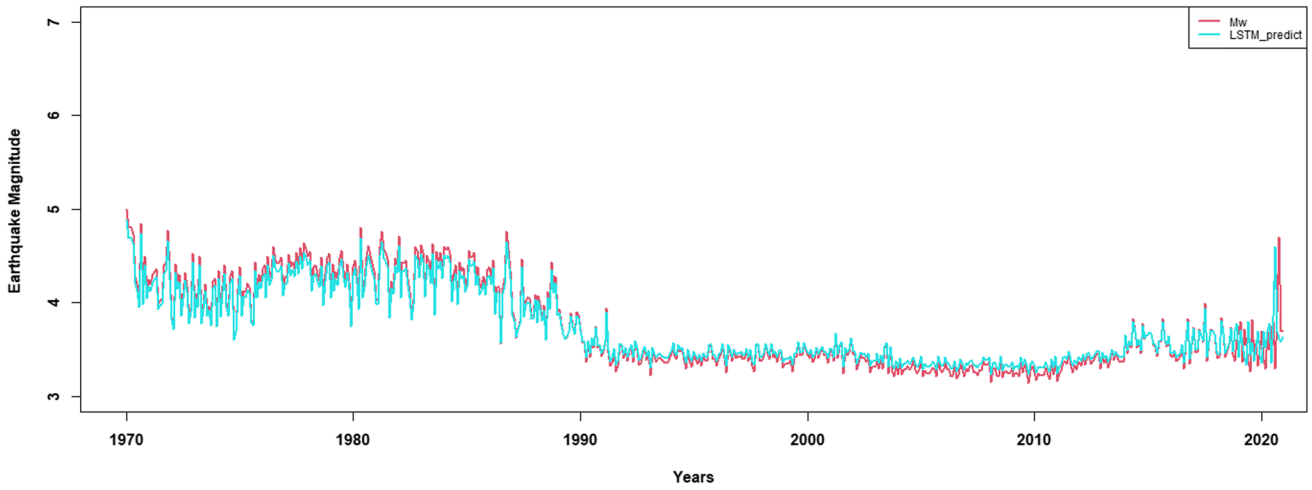


Fig. 17 The graph of original series and LSTM forecast series

capture the fluctuations and variations in the magnitude series, with LSTM showing a remarkable alignment with the original series. These results highlight the superior performance of CNN and LSTM in accurately capturing the underlying patterns and dynamics of the earthquake magnitude series.

Figure 17 displays the mean squared error (MSE) values obtained in each iteration using the Loss function, which serves as a measure of the performance of the CNN and LSTM models during the training process. The MSE values provide insight into the accuracy and effectiveness of the models in minimizing the differences between the predicted and actual values of the earthquake magnitude series.

It can be observed in Fig. 18 that the CNN method begins with a higher initial MSE value compared to the LSTM method. However, as the training progresses, both models exhibit a convergence behavior, with the LSTM method quickly reaching a lower MSE value. This suggests that the LSTM model is able to learn and capture the patterns in the earthquake magnitude series more effectively, resulting in better predictive performance.

Table 3 presents the calculated values of mean absolute error (MAE), mean squared error (MSE) and coefficient of determination (R^2) to compare the performance of the four methods used in the study. A smaller MAE indicates better predictive performance, while MSE provides a more precise measure that emphasizes large errors. R^2, on the other hand, represents how well the prediction model describes the dataset, with a value closer to 1 indicating a better fit.

Based on these values in Table 3, it is evident that the LSTM method achieves the most accurate estimations of the original series.

Table 3 The model selection criteria values

Model	Criteria		
	MAE	MSE	R^2
ARIMA	0.1264	0.0324	0.844
SSA	0.1074	0.0229	0.888
CNN	0.1283	0.0300	0.864
LSTM	0.0618	0.0100	0.963

Forecasts of earthquake magnitude

Once the suitable models were determined using the ARIMA, SSA, CNN, and LSTM methods, the forecast values obtained from these models are presented in Table 4.

Based on the forecast values obtained from the ARIMA model in Table 4, it is projected that the monthly average earthquake magnitudes will be 3.8 in 2021, 3.7 in 2022, and 3.7 in 2023. The SSA model predicts a gradual increase in the monthly average magnitudes, ranging from 3.7 to 4.1 over the 3-year period. In contrast, both the CNN and LSTM models anticipate consistent average magnitudes for the 3 years, with values of 3.8 and 3.6, respectively. These forecasts provide insights into the expected trends and patterns in earthquake magnitudes, highlighting the varying predictions obtained from the different models employed in the analysis.

The forecast values for 36 periods are presented in Table 4, and Figs. 19, 20, 21, and 22 display the corresponding graphs comparing the forecasted values with the original series. These visual representations allow for a clearer understanding of the predicted trends and patterns in the data. The forecast values provide valuable insights into the future behavior of the time series and facilitate

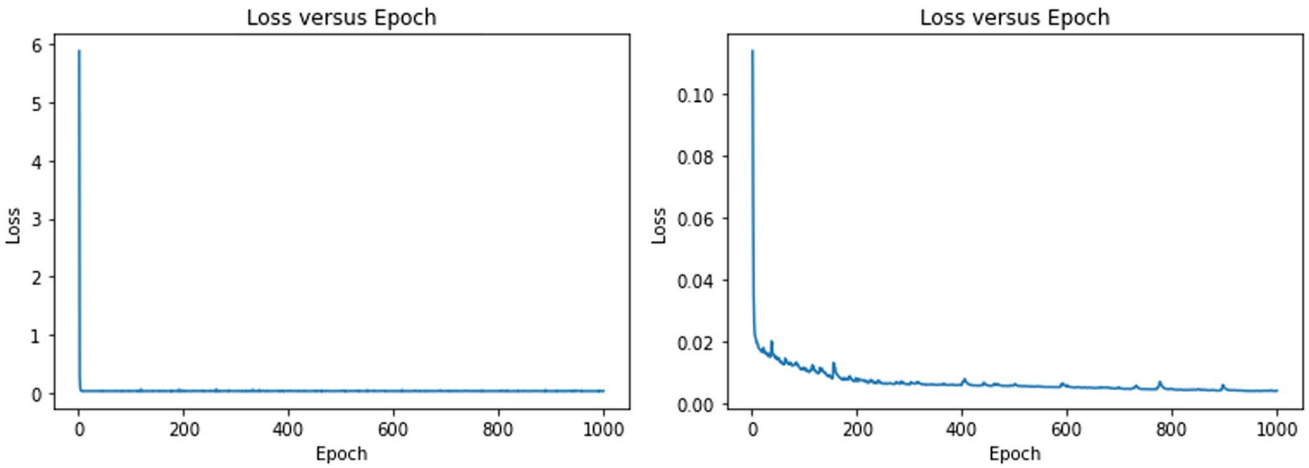


Fig. 18 Loss versus iterations (epoch) graph for the CNN (left side) and LSTM (right side)

Table 4 The forecast values of earthquake magnitudes from the models

Year	Month	1	2	3	4	5	6	7	8	9	10	11	12
2021	ARIMA	3.778	3.776	3.775	3.773	3.772	3.771	3.769	3.768	3.766	3.765	3.764	3.762
	SSA	3.735	3.827	3.842	3.750	3.760	3.910	3.982	3.890	3.831	3.916	3.981	3.908
	CNN	3.963	3.652	3.842	3.860	3.950	3.686	4.038	3.511	3.707	3.712	3.858	3.889
	LSTM	3.526	3.568	3.532	3.542	3.596	3.609	3.600	3.589	3.661	3.713	3.704	3.622
2022	Month	1	2	3	4	5	6	7	8	9	10	11	12
	ARIMA	3.761	3.759	3.758	3.756	3.755	3.754	3.752	3.751	3.749	3.748	3.747	3.745
	SSA	3.851	3.934	4.015	3.951	3.870	3.934	4.041	4.023	3.956	4.001	4.094	4.067
	CNN	3.728	3.837	3.863	3.574	4.024	3.634	3.826	3.868	3.892	3.668	4.068	3.620
	LSTM	3.663	3.683	3.699	3.746	3.679	3.682	3.654	3.593	3.606	3.556	3.585	3.522
2023	Month	1	2	3	4	5	6	7	8	9	10	11	12
	ARIMA	3.743	3.742	3.740	3.739	3.738	3.737	3.735	3.734	3.732	3.731	3.730	3.728
	SSA	3.969	3.976	4.066	4.072	4.001	4.017	4.125	4.156	4.078	4.046	4.107	4.121
	CNN	3.917	3.485	3.696	3.843	3.873	3.582	4.038	3.664	4.176	3.775	4.264	3.829
	LSTM	3.494	3.551	3.508	3.616	3.623	3.598	3.550	3.557	3.588	3.600	3.692	3.667

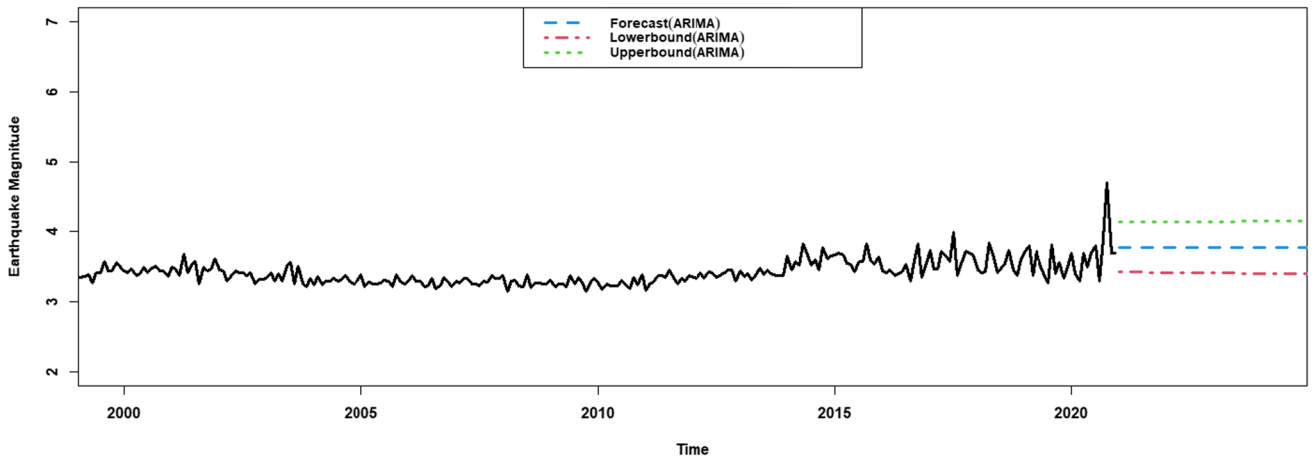


Fig. 19 The graph for the original and ARIMA forecast series

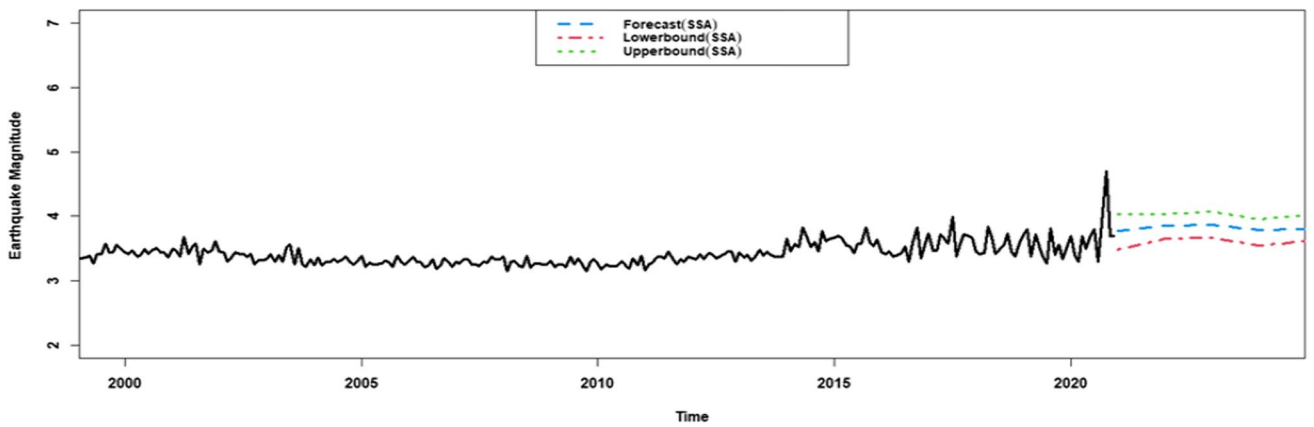


Fig. 20 The graph for the original and SSA forecast series

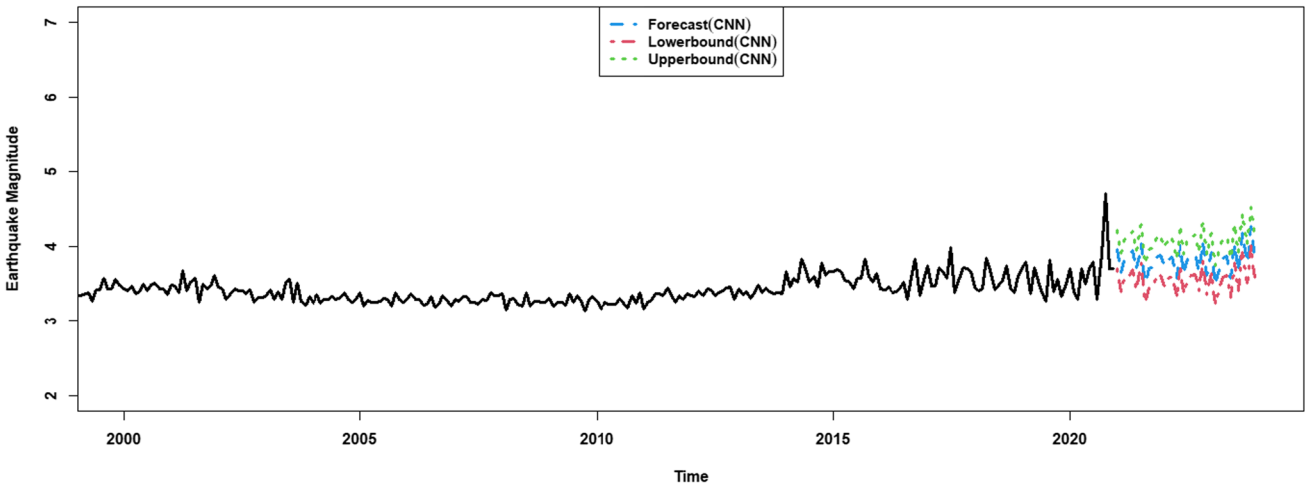


Fig. 21 The graph for the original and CNN forecast series

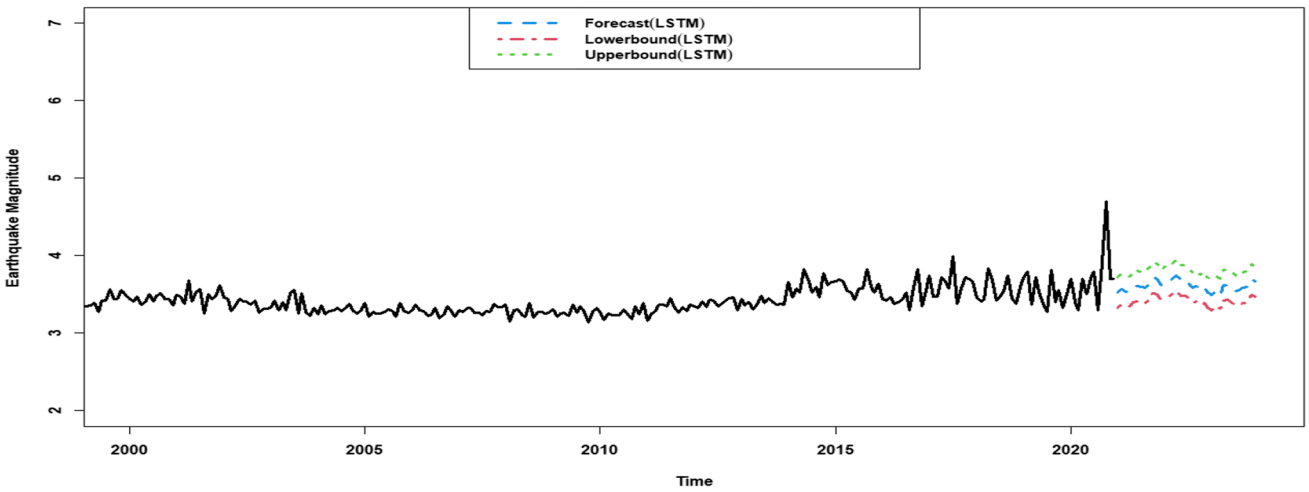


Fig. 22 The graph for the original and LSTM forecast series

decision-making processes based on the anticipated magnitudes of future earthquakes.

Based on the obtained results, it is predicted that there will be no significant increase or decrease in the monthly average earthquake magnitudes over the course of 3 years, according to all the methods employed. Regarding the prediction intervals, the ARIMA method provides forecasts with a wider range, whereas the SSA method yields forecasts with values that are relatively closer. In addition, the CNN method produces monthly average forecasts characterized by sudden fluctuations, whereas the LSTM method generates forecasts with larger fluctuations spanning several years. These findings shed light on the different characteristics and strengths of each method in predicting earthquake magnitudes.

ARIMA, SSA, CNN and LSTM models have not been used for the Aegean seismicity studies. Therefore, it is not possible to directly compare the results of this study with previous time series models employed in Aegean seismicity studies. However, it is worth mentioning that there are other studies in the literature that have explored the application of these models in predicting earthquake magnitudes. These studies may provide valuable insights and serve as references for future research in the field. Indeed, several studies in the literature have demonstrated the effectiveness of ARIMA, CNN, and LSTM models in predicting earthquake magnitudes and occurrences.

Shishegaran et al. (2019) conducted a study using earthquake events along the Zagrous fault from 2009 to 2018, specifically focusing on magnitudes greater than 2.5. Their results demonstrated that the ARIMA model is an effective tool for predicting earthquake magnitudes. Todelo et al. (2019) utilized an ARIMA (1, 0, 6) model to forecast the future occurrences of earthquakes within different

magnitude ranges for the years 2018 to 2022. Their findings revealed that the highest count of earthquake occurrences, specifically in the magnitude level of 5.0–5.9, is expected in the year 2022, with an estimated number of 1580 occurrences. This highlights the effectiveness of the ARIMA (1, 0, 6) model in predicting earthquake occurrences. Yuan et al. (2022) obtained optimal prediction models for various magnitude thresholds in the Longmen Mountain fault zone, situated on the eastern edge of the Qinghai Tibet Plateau. Specifically, they found the ARIMA $(10, 2, 1)\times(0, 1, 1)_{20}$ model suitable for series with magnitudes greater than or equal to 2.5, the ARIMA $(8, 2, 1)\times(0, 1, 1)_{40}$ model for magnitudes greater than or equal to 3.0, and the ARIMA $(1, 2, 3)\times(0, 1, 1)_{3}$ model for magnitudes greater than or equal to 4.5. These results further support the applicability of the ARIMA model in earthquake prediction for specific fault zones. Fuentes et al. (2022) proposed a parallel neural network model based on CNN and LSTM in their study. They utilized accumulated shell velocity and density data as inputs to estimate the daily mean and predict the average number of seismic events in Chile. By employing data from the last 20 days, they obtained forecasts for the following day. The proposed method achieved favorable results, with MSE, MAE, and R2 values of 0.03, 0.03, and 0.81, respectively, indicating the success of their CNN and LSTM-based approach. In another study by Nicolis et al. (2021), LSTM, CNN, and Algae Type Aftershock Sequences (ETAS) methods were employed to estimate seismic velocity and predict the number and location of future seismic events in Chile. They categorized the size of the Chilean catalogue to develop a model based on LSTM and CNN, and compared the performance using MAE, MSE, and R^2 values.

Deep learning methods, LSTM and CNN, have proven to be successful in earthquake prediction, considering factors such as spatial–temporal relationships, image data, earthquake signals, and seismic indicators. LSTM and CNN deep learning methods are heavily involved in earthquake prediction, taking into account different situations such as the spatial–temporal relationship of earthquakes in different regions, image data, earthquake signals and seismic indicators (Wang et al. 2017; Huang et al. 2018; Al Banna et al. 2021; Bhandarkar et al. 2019; Mousavi and Beroza 2020; Nicolis et al. 2021). Overall, the literature demonstrates the effectiveness of ARIMA, CNN, and LSTM models in earthquake magnitude and occurrence prediction, highlighting their success in various scenarios and regions.

Conclusion

Earthquakes are natural disasters that can result in significant damage and loss of life worldwide. The existing measures to mitigate the consequences of earthquakes are often inadequate, as the exact timing and magnitude of these events are unpredictable. Therefore, accurate earthquake forecasts are crucial for effective preparedness and response efforts. Time series analysis is a practical and commonly employed method for predicting future events, making it particularly relevant in the context of earthquake magnitude estimation.

The study focused on analyzing earthquake magnitude data obtained through the Scientific and Technological Research Council of Turkey (TUBITAK) ARDEB 1001 program [Project number: 121F208]. In this study, we focused on analyzing earthquake magnitudes in the Aegean region of Turkey, along with a 200 km buffer zone. This specific area has not been extensively investigated using ARIMA before, emphasizing the novelty and importance of our research. Then, we employed the singular spectrum analysis (SSA) technique, which has not been commonly utilized for earthquake magnitude prediction in Turkey's earthquake catalogs. By applying these methodologies, we aimed to contribute to the understanding and prediction of earthquake magnitudes in the region. In time series analysis, calculating the monthly average allows us to gain insights into seismic activity over specific time periods and facilitates the identification of long-term patterns or variations. The raw earthquake data in a time series often exhibit random fluctuations and noise, which can make it challenging to observe underlying trends. By calculating the monthly average, we can reduce the impact of these instantaneous fluctuations, resulting in a smoother representation of the overall seismic activity and enabling a clearer understanding of the underlying trend over time. Therefore, the monthly average earthquake magnitudes have been used for the ARIMA and SSA models.

The monthly average earthquake magnitudes are also utilized to compare deep learning methods, namely long short-term memory (LSTM) and conventional neural network (CNN), with ARIMA and SSA methods. In this study, we decide to use ARIMA and SSA methods as baseline models that can help evaluate the performance of more advanced techniques such as machine learning or deep learning models and we conduct deep learning model in the revised paper. In this way, ARIMA and SSA models provide reference points to assess the effectiveness of newer methods. Among the various models examined, the LSTM method demonstrated the best estimations for the series. In the analysis of the data, the classical time series method ARIMA was employed, and it was observed that there would be no significant change in the monthly average earthquake magnitudes over the next 36 periods. Moving forward, the study incorporated the use of the non-parametric method SSA, known for its effectiveness in analyzing magnitude data. The most suitable model identified for the series was SSA (240,12). According to the estimated values obtained from this model, there is an

expectation of a slight increase in the monthly average earthquake magnitudes between 2021 and 2023.

Lastly, the study delved into the application of deep learning methods, specifically CNN and LSTM, which have proven to be effective in earthquake research. Through experimentation, the most accurate results were obtained using specific configurations such as batch size = 32, epoch number = 1000, and specific activation functions for both CNN and LSTM models. The findings suggest that there will be no significant changes in the monthly average earthquake magnitudes from 2021 to 2023, as indicated by both the ARIMA and SSA methods. However, the CNN method predicts periodic ups and downs, while the LSTM method indicates larger fluctuations in earthquake magnitudes during the same period.

The significance of this study lies in its potential to enhance our ability to forecast earthquake magnitudes accurately. This, in turn, can facilitate the implementation of proactive measures to mitigate the impact of earthquakes and reduce potential damages. By utilizing ARIMA, SSA and exploring new techniques such as LSTM and CNN, we hope to provide valuable insights into earthquake prediction in the Aegean Region of Turkey.

Acknowledgements We thank the support of The Scientific and Technological Research Council of Turkey (TUBITAK) ARDEB 1001 [Project number: 121F208] program. The authors also thank Assistant Prof. Dr. Tuba Eroğlu Azak for preparing dataset, Prof. Dr. Tolga Çan for his valuable advice and reviewers for their constructive comments.

Author contributions H. O. Cekim and H. N. Karakavak composed methods and results, G. Ozel wrote the main manuscript text, and S. Tekin prepared the data and Figures 1 and 2. All authors reviewed the manuscript.

Data Availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors report that they have no conflict of interest.

References

Abdalzaher MS, El-Hadidy M, Gaber H, Badawy A (2020) Seismic hazard maps of Egypt based on spatially smoothed seismicity model and recent seismotectonic models. J Afr Earth Sc 170:103894

Abdalzaher MS, Soliman MS, El-Hady SM, Benslimane A, Elwekeil M (2021) A deep learning model for earthquake parameters observation in IoT system-based earthquake early warning. IEEE Internet Things J 9(11):8412–8424

Akın C & Yağmurlu F (2020) Seismicity of Akşehir Graben. J Adv Eng Stud Technol 1(2), 162–170. Retrieved from https://dergipark.org.tr/en/pub/imctd/issue/59372/830343 (**in Turkish**)

Akkar S, Azak T, Çan T, Çeken U, Demircioğlu Tümsa MB, Duman TY et al (2018) Evolution of seismic hazard maps in Turkey. Bull Earthquake Eng 16:3197–3228

Al Banna MH, Ghosh T, Nahian MJ, Taher KA, Kaiser MS, Mahmud M, Hossain MS, Andersson K (2021) Attention-based bi-directional long-short term memory network for earthquake prediction. IEEE Access 9:56589–56603

Albawi S, Mohammed TA & Al-Zawi S (2017) Understanding of a convolutional neural network. In: 2017 international conference on engineering and technology (ICET), 1–6, IEEE

Altunel E (1999) Geological and geomorphological observations in relation to the 20 September 1899 Menderes earthquake, Western Turkey. J Geol Soc Lond 156:241–246

Amei A, Fu W & Ho CH (2012) Time series analysis for predicting the occurrences of large scale earthquakes. Int J Appl Sci Technol 2(7):64–75

Bayrak Y, Bayrak E (2012) An evaluation of earthquake hazard potential for different regions in Western Anatolia using the historical and instrumental earthquake data. Pure Appl Geophys 169(10):1859–1873

Bayrak E, Yılmaz Ş, Bayrak Y (2017) Temporal and spatial variations of Gutenberg—Richter parameter and fractal dimension in Western Anatolia, Turkey. J Asian Earth Sci 138:1–11

Berhich A, Belouadha F & Kabbaj MI (2021) LSTM-based Models for Earthquake Prediction, Conference: NISS2020: The 3rd International Conference on Networking, Information Systems & Security

Bhandarkar T, Satish N, Sridhar S, Sivakumar R, Ghosh S (2019) Earthquake trend prediction using long short-term memory RNN. Int J Elect Comput Eng 9(2):1304–1312

Bozkurt E (2000) Timing of extension on the Büyük Menderes Graben, Western Turkey and its tectonic implications. Geol Soc Lond 173(1):385–403 (**Special Publications**)

Cao J, Li Z, Li J (2019) Financial time series forecasting model based on CEEMDAN and LSTM. Physica A 519:127–139

Cekim HO, Tekin S, Ozel G (2021) Prediction of the earthquake magnitude by time series methods along the East Anatolian Fault, Turkey. Earth Sci Inform 14(3):1339–1348

Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H & Bengio Y (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078

Coban KH, Sayil N (2019) Evaluation of earthquake recurrences with different distribution models in western Anatolia. J Seismolog 23:1405–1422

Coban KH, Sayil N (2022) Magnitude type conversion models for earthquakes in turkey and its vicinity with machine learning algorithms. J Earthquake Eng 27(9):2533–2554

Cohen HA, Dart CJ, Akyüz HS, Barka A (1995) Syn-rift sedimentation and structural development of the Gediz and Büyük Menderes graben, Western Turkey. J Earth Soc Lond 152:629–638

Coussin M (2022) Singular spectrum analysis for real-time financial cycles measurement. J Int Money Financ 120:102532

Danese M, Lazzari M & Murgante B (2009) Integrated Geological, Geomorphological and Geostatistical analysis to study macroseismic effects of 1980 Irpinian earthquake in urban areas (southern Italy). In International conference on computational science and ıts applications (pp 50–65). Springer, Berlin, Heidelberg

Dobilas S (2022) LSTM Recurrent Neural Networks—How to Teach a Network to Remember the Past. Medium. https://towardsdatascience.com/lstm-recurrent-neural-networks-how-to-teach-a-network-to-remember-the-past-55e54c2ff22e. Accessed 28 Jun 2022

Durak S (2008) The masonry structures commonly used in the Aegean Region and the earthquake safety of these structures (Master's

thesis. Pamukkale University Institute of Science and Technology) (**in Turkish**)
Elhadidy M, Abdalzaher MS, Gaber H (2021) Up-to-date PSHA along the Gulf of Aqaba-Dead Sea transform fault. Soil Dyn Earthq Eng 148:106835
Elman JL (1990) Finding structure in time. Cogn Sci 14(2):179–211
Emre Ö, Duman TY, Özalp S, Şaroğlu F, Olgun Ş, Elmacı H, Çan T (2018) Active fault database of Turkey. Bull Earthq Eng 16(8):3229–3275
Erdik M, Demircioglu M, Sesetyan K, Durukal E, Siyahi B (2004) Earthquake hazard in Marmara region, Turkey. Soil Dyn Earthq Eng 24(8):605–631
Fuentes AG, Nicolis O, Peralta B, Chiodi M (2022) Spatio-temporal seismicity prediction in Chile using a multi-column ConvLSTM. IEEE Access 10:107402–107415
Gao W, Guo J, Zhou M, Yu H, Chen X & Ji B (2020) Gravity tides extracted from SSA-denoised superconducting gravity data with the harmonic analysis: a case study at Wuhan station. China. Acta Geodaetica et Geophysica 1–17
Gardner JK, Knopoff L (1974) Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? Bull Seismol Soc Am 64(5):1363–1367
Gers FA & Schmidhuber J (2000) Recurrent nets that time and count. In IJCNN, Vol. 3 (pp 189–194). IEEE
Ghil M, Vautard R (1991) Interdecadal oscillations and the warming trend in global temperature time series. Nature 350(6316):324–327
Ghil M, Allen MR, Dettinger MD, Ide K, Kondrashov D, Mann ME et al (2002) Advanced spectral methods for climatic time series. Rev Geophys 40(1):1–3
Gioncu V, Mazzolani F (2011) Earthquake engineering for structural design. CRC Press, London
Golyandina N, Zhigljavsky A (2013) Singular Spectrum Analysis for time series. Springer Science & Business Media
Golyandina N, Nekrutkin V, Zhigljavsky AA (2001) Analysis of time series structure: SSA and related techniques. CRC Press
Graves A, Schmidhuber J (2005) Framewise phoneme classification with bidirectional lstm and other neural network architectures. Neural Netw 18(5–6):602–610
Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B et al (2018) Recent advances in convolutional neural networks. Pattern Recogn 77:354–377
Hasan Al Banna M, Ghosh T, Taher KA, Kaiser MS & Mahmud M (2021) An earthquake prediction system for bangladesh using deep long short-term memory architecture. In: Intelligent Systems: Proceedings of ICMIB 2020 (pp 465–476). Springer Singapore
Hassani H (2007) Singular spectrum analysis: methodology and comparison, MPRA Paper No: 4991
Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
Huang JP, Wang XA, Zhao Y, Xin C, Xiang H (2018) Large earthquake magnitude prediction in Taiwan based on deep learning neural network. Neural Network World 28(2):149–160
Kadılar C, Öncel Çekim H (2020) Introduction to SPSS and R applied time series analysis. Seçkin Publication (**in Turkish**)
Kadirioğlu FT, Kartal RF (2016) The new empirical magnitude conversion relations using an improved earthquake catalogue for Turkey and its near vicinity (1900–2012). Turkish J Earth Sci 25(4):300–310
Karadaş A, Öner E (2021) The effects of alluvial geomorphology of the Bornova plain on the damage caused by the 30 October 2020 Samos earthquake in İzmir-Bayraklı. J Geogr 42:139–153 (**in Turkish**)
Kaynar O, Taştan S (2009) Comparison of Mlp artificial neural networks and purification model in time series analysis. J Erciyes Univ Faculty Econ Admin Sci 33:161–172 (**in Turkish**)
Keskin S & Külahcı F (2023) ARIMA model simulation for total electron content, earthquake and radon relationship identification. Nat Hazards 115(3):1955–1976
Ketin İ (1968) Relationships between general tectonic situation of Turkey and major earthquake zones. J Min Res Explor 71(71):129–134 ((**in Turkish**))
Lakshmi SS, Tiwari RK (2009) Model dissection from earthquake time series: a comparative analysis using modern non-linear forecasting and artificial neural network approaches. Comput Geosci 35(2):191–204
LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W & Jackel L (1989) Handwritten digit recognition with a back-propagation network. Adv Neural İnf Process Syst 2
Li T, Zhang Y, Wang T (2021) SRPM–CNN: a combined model based on slide relative position matrix and CNN for time series classification. Complex Intell Syst 7:1619–1631
Lin Y, Ling BWK, Xu N, Zhou X (2022) Two dimensional quaternion valued singular spectrum analysis with application to image denoising. J Franklin Inst 359(8):3808–3830
Liu CL, Hsaio WH, Tu YC (2018) Time series classification with multivariate convolutional neural network. IEEE Trans Industr Electron 66(6):4788–4797
Livieris IE, Pintelas E, Pintelas P (2020) A CNN–LSTM model for gold price time-series forecasting. Neural Comput Appl 32:17351–17360
Mousavi SM, Beroza GC (2020) A machine-learning approach for earthquake magnitude estimation. Geophys Res Lett 47(1):e2019GL085976
Mouslopoulou V & Hristopulos DT (2011) Patterns of tectonic fault interactions captured through geostatistical analysis of microearthquakes. J Geophys Res 116(B7):1–18
Nicolis O, Plaza F, Salas R (2021) Prediction of intensity and location of seismic events using deep learning. Spatial Stat 42:100442
O'Shea K & Nash R (2015) An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.
Paton S (1992) Active normal faulting. Drainage patterns ve sedimentation in Southwestern Turkey. J Earth Soc Lond. 149:1031–1044
Polat O, Gok E, Yilmaz D (2008) Earthquake hazard of the Aegean extension region (West Turkey). Turkish J Earth Sci 17(3):593–614
Putriasari N, Nugroho S, Rachmawati R, Agwil W, Sitohang YO (2022) Forecasting a weekly red Chilli Price in Bengkulu City using autoregressive integrated moving average (ARIMA) and Singular Spectrum Analysis (SSA) methods. J Stat Data Sci 1(1):1–6
Ramanjaneyulu K, Swamy KV & Rao CS (2018) Novel CBIR system using CNN architecture. In: 2018 3rd International conference on inventive computation technologies (ICICT) (pp 379–383). IEEE
Saad OM, Hafez AG, Soliman MS (2020) Deep learning approach for earthquake parameters classification in earthquake early warning system. IEEE Geosci Remote Sens Lett 18(7):1293–1297
Schmidhuber J, Wierstra D, Gagliolo M, Gomez F (2007) Training recurrent networks by evolino. Neural Comput 19(3):757–779
Schoellhamer DH (2001) Singular spectrum analysis for time series with missing data. Geophys Res Lett 28(16):3187–3190
Şengör AMC (1979a) Mid-Mesozoic closure of Permo-Triassic Tethys and its implications. Nature 279:590–593
Şengör AMC (1979b) The north Anatolian transform fault: its age, offset and tectonic significance. J Geol Soc Lond 136:269–282

Şengör AMC, Yilmaz Y (1981) Tethyan evolution of Turkey: a plate tectonic approach. Tectonophysics 75(3–4):181–241
Şengör AMC (1982) Factors governing the neotectonic evolution of the Aegean. Young Tectonics and Volcanism of Western Anatolia Panel. 59–71. (**in Turkish**)
Seyitoğlu G, Scott B (1992) The age of the Büyük Menderes Graben (West Turkey) and its tectonic implications. Geol Mag 129:239–242
Shishegaran A, Taghavizade H, Bigdeli A, Shishegaran A (2019) Predicting the earthquake magnitude along Zagros fault using time series and ensemble model. J Soft Comput Civ Eng 3(4):67–77
Siami-Namini S, Tavakoli N & Namin AS (2019) The performance of LSTM and BiLSTM in forecasting time series. In 2019 IEEE International Conference on Big Data, 3285–3292
Sözbilir H (2001) Young-tectonics of Nazilli and its surroundings (Büyük Menderes Graben). Great Menderes Earthquakes Geophysical Meeting. 54–61 (**in Turkish**)
Stein RS, Barka AA, Dieterich JH (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. Geophys J Int 128(3):594–604
Tağil Ş (2004) Neotectonic characteristics and seismicity of Balikesir plain and its neighborhood. J Geog Sci 2(1):73–92 ((**in Turkish**))
Tağil Ş, Alevkayali C (2013) Spatial Distribution of Earthquake in the Aegean Region: A Geostatistical Approach, Social. Int J Stud 6(28):69–379 (**in Turkish**)
Tan O (2021) Turkish Homogenized Earthquake Catalogue (TURHEC). Natural Hazards and Earth System Sciences (NHESS). Zenodo. https://doi.org/10.5281/zenodo.5056801
Todelo TL & Chris Jordan GA (2019) Predictability of earthquake occurrence using auto regressive integrated moving average (ARIMA) model. In: Lecture notes in engineering and computer science: proceedings of the ınternational multiconference of engineers and computer scientists, pp 13–15
Uhrhammer RA (1986) Characteristics of northern and central California seismicity. Earthquake Notes 57(1):21
Wang Q, Guo Y, Yu L, Li P (2017) Earthquake prediction based on spatio-temporal data mining: an LSTM network approach. IEEE Trans Emerg Top Comput 8(1):148–158
Yadav A, Gahalaut K, Mallika K, Purnachandra Rao N (2015) Annual periodicity in the seismicity and water levels of the Koyna and Warna reservoirs, western India: a singular spectrum analysis. Bull Seismol Soc Am 105(1):464–472
Yadav A, Jha CK, Sharan A (2020) Optimizing LSTM for time series prediction in Indian stock market. Procedia Comput Sci 167:2091–2100
Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9:611–629
Yang Q, Deng C, Chang X (2022) Ultra-short-term/short-term wind speed prediction based on improved singular spectrum analysis. Renew Energy 184:36–44
Yang J & Li J (2017) Application of deep convolution neural network. In: 2017 14th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) (pp 229–232). IEEE
Yu R, Gao J, Yu M, Lu W, Xu T, Zhao M et al (2019) LSTM-EFG for wind power forecasting based on sequential correlation features. Fut Gener Comput Syst 93:33–42
Yuan X, Dan H, Qiuyin Y, Wenjun Z, Min R & Jing Y (2022) ARIMA model analysis of the regularities of earthquake origin times in the Longmen mountain fault zone. (Preprint in Research Square)
Zaremba W, Sutskever I & Vinyals O (2014) Recurrent neural network regularization. arXiv preprint arXiv:1409.2329.
Zha W, Liu Y, Wan Y, Luo R, Li D, Yang S, Xu Y (2022) Forecasting monthly gas field production based on the CNN-LSTM model. Energy 260:124889
Zhai D, Zhang X, Xiong P (2020) Detecting thermal anomalies of earthquake process within outgoing longwave radiation using time series forecasting models. Ann Geophys 63(5):PA48–PA548
Zhang H, Melgar D, Sahakian V, Searcy J, Lin JT (2022) Learning source, path and site effects: CNN-based on-site intensity prediction for earthquake early warning. Geophys J Int 231(3):2186–2204
Zhao B, Lu H, Chen S, Liu J, Wu D (2017) Convolutional neural networks for time series classification. J Syst Eng Electron 28(1):162–169