Article

Earthquake Magnitude and Frequency Forecasting in Northeastern Algeria using Time Series Analysis

Mouna Merdasse [1], Mohamed Hamdache [2], José A. Peláez [3,*], Jesús Henares [4] and Tarek Medkour [5]

1   Department of Probability and Statistics, Faculty of Mathematics, University of Sciences USTHB, Algiers 16111, Algeria
2   Seismological Survey Department, CRAAG, Algiers 16032, Algeria
3   Department of Physics, University of Jaén, 23071 Jaén, Spain
4   International University of La Rioja, 26006 La Rioja, Spain
5   Department of Intelligent Systems Engineering, National School of Artificial Intelligence, Algiers 16309, Algeria
*   Correspondence: japelaez@ujaen.es

Abstract: This study uses two different time series forecasting approaches (parametric and non-parametric) to assess a frequency and magnitude forecasting of earthquakes above Mw 4.0 in Northeastern Algeria. The Autoregressive Integrated Moving Average (ARIMA) model encompasses the parametric approach, while the non-parametric method employs the Singular Spectrum Analysis (SSA) approach. The ARIMA and SSA models were then used to train and forecast the annual number of earthquakes and annual maximum magnitude events occurring in Northeastern Algeria between 1910 and 2019, including 287 main events larger than Mw 4.0. The SSA method is used as a forecasting algorithm in this case, and the results are compared to those obtained by the ARIMA model. Based on the root mean square error (RMSE) criterion, the SSA forecasting model appears to be more appropriate than the ARIMA model. The consistency between the observation and the forecast is analyzed using a statistical test in terms of the total number of events, denoted as N-test. As a result, the findings indicate that the annual maximum magnitude in Northeastern Algeria between 2020 and 2030 will range from Mw 4.8 to Mw 5.1, while between four and six events with a magnitude of at least Mw 4.0 will occur annually.

Keywords: earthquake magnitude forecasting; time series analysis; singular spectrum analysis (SSA); autoregressive integrated moving average (ARIMA) model

1. Introduction

Earthquakes are one of the most challenging natural disasters for populations to manage. Even though earthquakes are unpredictable and typically occur without warning, a detailed analysis of the seismic hazard and risk-reduction measures can help to reduce subsequent economic and social losses after the occurrence of an earthquake.

The north of Algeria is located on the border of the Nubian Plate, which is in a compressional movement with the Eurasian Plate [1–3] (Figure 1a). The origin of the seismicity in this region is the compressional movement between these two plates. Known seismic activity (Figure 1b) includes several damaging earthquakes, especially in the last 50 years, where severe earthquakes have been recorded in the El Asnam region (now Chleff), including the earthquakes of 9 September 1954 (Mw 6.8) and 10 October 1980 (Ms 7.3) (see [4]). The most recent significant event was the Zemmouri/Algiers earthquake 21 May 2003 (Mw 6.9), occurring roughly 50 km offshore from the northeast of Algiers (e.g., [5,6]). In addition, in this period, several earthquakes affected regions close to important cities and caused some damages, such as the Constantine (northeast Algeria) earthquake (Ms 5.9) of 27 October 1985.
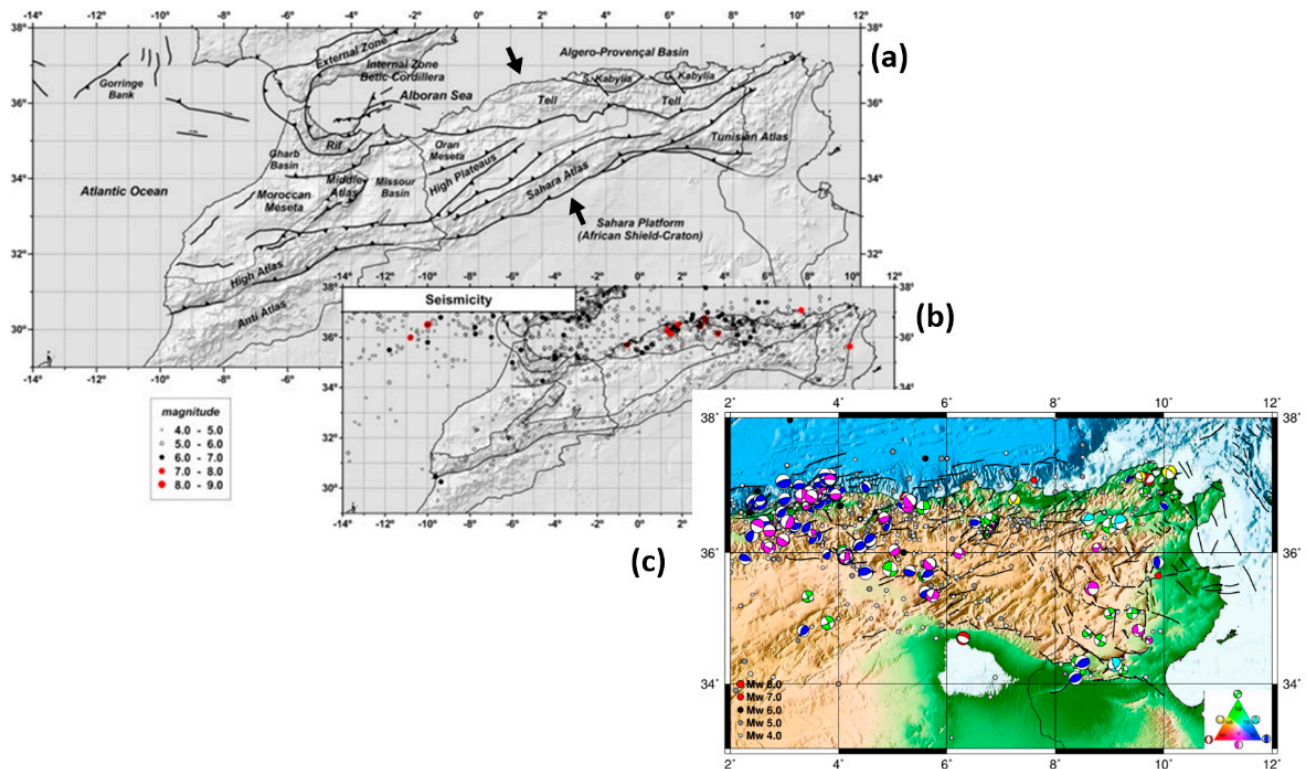
Figure 1. Tectonic and seismicity frame. (**a**) Tectonic sketch for the studied region. (**b**) Sha-low seismicity with depth less than 30 km and magnitude above magnitude Mw 4.0. (**c**) Distribution of seismicity with magnitude above Mw 4.0 and distribution of focal mechanism solutions.

The interest of the scientific community in approaches to reduce the risk of damaging earthquakes and related seismic risk assessments of urban areas in Northern Algeria is steadily growing in response to this seismic activity. There is always a pressing need for studies on earthquake forecasting, the implementation of building regulations, and safe constructions, especially in developing countries like Algeria, due to the severity of large earthquakes and the damage they produce. The resulting improvement in earthquake risk assessment and hazard management leads to significant savings in human life and properties.

The effects of the earthquakes previously mentioned in Northern Algeria [7] indicate the importance of improving estimates of the forecasting from different source zones. Due to its complex tectonics, the Northeastern region of Algeria was chosen as the studied area.

On both short- and long-term time periods, earthquake forecasting has significant social and economic consequences. In addition, it plays a main role in earthquake preparedness. On the basis of the quantification of patterns in seismicity data, a wide range of forecasting methods have been proposed with varied degrees of effectiveness [8–17]. For instance, Mignan [17] develops an analysis of the steady increase in seismic activity around a potential earthquake epicenter, known as the accelerating moment release (AMR) approach, while Keilis–Borok [8] presents and analyses step-by-step the topic of the earthquake prediction. A significant advance in this research field is the earthquake forecasting method based on Regional Earthquake Likelihood Models (RELM) [14,18–21]. Nowadays, a new algorithm known as pattern informatics (PI) is also emerging as an advanced method [14]. In addition, new methods based on deep learning are being developed to find a static-stress-based criterion predicting the location of aftershocks [22], as well as the interest in real-time post-seismic forecasting based on ground velocity recorded within the first hour after the mainshock by means of the perceived magnitude [23].

Several studies have been conducted from the perspective of seismic hazard assessment and forecasting. The Gutenberg Richter (GR) model [24] is widely used and states that the logarithm of the cumulative number of events is linearly proportional to the magnitude.

It is worth noting that several authors [25–27] suggested that the extreme value distribution is a more suitable analytical model and should be employed for the distribution of the magnitude set rather than the GR model. Stochastic processes, particularly Poisson processes, are also used to forecast earthquake frequencies by taking the earthquake occurrence time into account. The Poisson model is characterized by a constant hazard function and an exponential recurrence time distribution. This assumption leads to the unsatisfying result of only time-independent seismic hazard or seismic forecasting estimates, suggesting that the probability of an earthquake at any given time is independent of its magnitude and the time since the last one, respectively. Additionally, compared to the Poisson law, the distribution of earthquakes is over-dispersed, since the events are grouped in time and location. Various other models based on stochastic processes have also been investigated to predict the number of events, such as compound Poisson [28,29], branching [15,30] and stochastic point processes [31,32]. These stochastic procedures made predicting the magnitude of the earthquake challenging. In addition, these approaches are ineffective for determining earthquake seasonality and patterns.

Recently, a few studies attempted to use forecasting models based on time series, such as the Auto-Regressive Integrated Moving Average (ARIMA) and the generalized autoregressive conditional heteroscedasticity (GARCH) models [33,34], to forecast earthquake magnitudes by considering the seasonality and trends of earthquake series. Another model, the singular spectrum analysis model, or SSA [35–37], is becoming more attractive in order to forecast earthquake magnitudes in a specific seismic zone. It is a reliable and advanced non-parametric time series analysis method combining dynamical systems, signal processing, multivariate statistics and traditional forecasting analysis. According to several authors [35–37], this approach is useful for determining the magnitude of earthquakes in a given area. The SSA model can identify significant seismic time series components exhibiting typical irregular behavior and provide accurate forecasts for them. Despite the fact that a few studies on earthquake magnitude forecasting using the SSA approach have been conducted [37], the primary goal of our study is to apply this method in Northeastern Algeria to predict earthquake magnitudes with time effects, which are frequently missed in estimates, while also capturing the dynamics of earthquake occurrences. Subsequently, the ARIMA and SSA models provided the annual maximum earthquake magnitudes for the first time in the studied area.

The root mean square error approach, usually denoted as RMSE, is a common tool for comparing different models and/or methodologies in time series [38,39]. In this study, the estimated RMSE values show that the SSA is the best model for describing the number of earthquakes and the annual maximum magnitude in this region. The seismicity in this area is continuous, and it is characterized by low-to-moderate seismic activity. Despite the interest of the scientific community in regional seismology and seismic hazards, the studied area remains of great interest due mainly to the tectonic complexity, which resulted in the identification of several active faults, the characteristics of which are still under investigation [40]. Then, this study is the first one in this region to address the issue of forecasting using time series, and it is part of the scientific community's efforts to manage and reduce seismic risk. It is worth noting that, previously, a model has been developed for Northern Morocco and Algeria to assess the probability of exceeding magnitudes Mw 5.0 and 6.0 in 10 years [41], but assuming a Poissonian process on a spatially smoothed seismicity model. This study enabled for the spatial variation of such results.

In the current study, the consistency between the observation and the forecast is analyzed using the CSEP (Collaborative for the Study of Earthquake Predicability) test [42] in terms of the total number of events, also denoted as N-test. The results are presented as quantile scores, δ_1 and δ_2 indices. According to Nanjo et al. [43], the forecast rate is too high (an overestimation) if δ_2 is very small, and too low (an underestimation) if δ_1 is very small for the N-test. Furthermore, according to the original CSEP testing framework, a model fails the test if its score is below a significance level of 2.5%, indicating an inconsistency between the forecast and the observation.

2. Seismicity and Earthquake Database

The study area is situated in the Tell Atlas in Northeastern Algeria (Figure 1a). This important geological formation resulted from a collision caused by the Algero–Provençal Basin opening within the Nubia Plate in the Early Miocene. Furthermore, it is most likely related to the Nubian Plate subduction, which is dipping to the north [44,45].

Onshore and offshore folds and thrust faults extending from NE–SW and from E–W represent the majority of the current tectonic features (e.g., [46,47]). The compressional movement between the Nubian and Eurasian Plates causes considerable seismicity, which is mostly represented in a moderately diffused seismic area in Northern Africa in its western domain (Figure 1b). This collision zone includes Northern Algeria, where recent geodetic studies reveal an actual plate convergence rate of about 5 ± 1 mm/yr in a N60°W direction, as illustrated in Figure 1a [2,3,48–50]. Earthquakes with a magnitude above Mw 5.0 frequently occur in the area [7,51], sometimes causing significant damage and casualties [52,53].

Most well-studied thrusting earthquakes occur on land, including the largest-recorded Ms 7.3 10 October 1980, El Asnam earthquake [54–56]. Examples of destructive earthquakes in the past include the 3 January 1365, and 5 May 1716 earthquakes, both felt with intensity X (European Macroseismic Scale, EMS-98) [57]. The earthquake that struck on 2 March 1825, with a felt intensity of X–XI on the Modified Mercalli Intensity Scale (MMI), was another historical occurrence. The Mw 6.9 earthquake that struck the area under consideration on 21 May 2003 was the most recent destructive earthquake [7]. The seismicity of the studied area displayed in Figure 1b has been the subject of various studies, both on historical and instrumental seismicity. For instance, Harbi et al. [51] performed a clear and comprehensive analysis of the historical seismicity, proposing the re-appraisal of several historical events located in the region. On the other hand, recent instrumental seismicity recorded by the Algerian seismic network has also been the subject of several works [58–62]. Figure 1c depicts the shallow depth (h $\leq$ 30 km) focal mechanism data compiled in the region [44,45]; the Frohlich diagram [63] is displayed, and the Zoback classification [64] is used. In previous studies [44,45], a detailed analysis of the inversion of the focal mechanism solutions and the inferred stress pattern was performed. The tectonic framework of the Ibero–Maghrebian region is detailed from these data, and the results are found to be in agreement with more recent studies (e.g., [50]). Similar horizontal maximum principal stress directions can be seen throughout the area, and they almost exactly coincide with the previously identified horizontal P axis in the NW–SE direction, as previously established by Henares et al. [65]. Additional tectonic stress regimes have been identified, including extensional, compressional, transpressional and strike-slip schemes.

The earthquake data file considered in this analysis was obtained from a seismic catalog previously compiled [7] for Northern Algeria, which has been updated until December 2019. It is well recognized that having an up-to-date Poissonian earthquake data file is a prerequisite for any seismic hazard assessment [66–69]. A combination of available published studies, bulletins and original data was used to produce the early earthquake data file. This led to the development of a unified earthquake data file, including information on magnitude, regional extent (between 32° and 38° latitudes and between 3°W and 10°E longitude), and date range (between AD 856 and June 2008) [7].

Specifically, for the current study, this initial earthquake data file has been updated to December 2019, with data coming from the Spanish Instituto Geográfico Nacional (IGN) and the Algerian Centre de Recherche en Astronomie, Astrophysique et Géophysique (CRAAG). It is important to note that in order to maintain the magnitude homogenization, this process was performed using the same relationships between reported magnitudes and moment magnitude that were used to generate the initial catalog [7]. The initial recorded events were described using several scales (surface-wave Ms, body wave m_b, body wave from Lg phase amplitude m_{bLg} and local duration M_L magnitudes). The next stage was to identify and remove any dependent event or non-Poissonian earthquake (foreshocks, aftershocks and swarms). In the present study, the method of Gardner and Knopoff [70]

was applied using the particular temporal and spatial window introduced in the initial catalog [7]. Figure 1b shows the spatial distribution of shallow seismicity for earthquakes with magnitudes greater than Mw 4.0 using the declustered dataset. Figure 1c, on the other hand, shows the distribution of focal mechanism solutions and seismicity with magnitudes greater than Mw 4.0.

Finally, the completeness of this catalog is performed at the base of a visual approach [71] in order to evaluate the threshold magnitude of the earthquake data file. Following previous studies [7,72], this approach is applied to our data above different magnitude values: if the cumulative annual number of earthquakes over this magnitude is approximately linear, then the seismic catalog is complete and Poissonian for a given threshold magnitude during a specific period of time. The cumulative number of events with a magnitude above Mw 4.0, 4.5, 5.0, 5.5, 6.0 and 6.5 is depicted in Figure 2. It shows that magnitudes above Mw 4.0 and 4.5 appear roughly complete and Poissonian since 1920 and 1910, with rates of 7.7 and 4.2 events/year, respectively. However, approximately since 1870 and 1885, with rates of 2.1 and 0.81 events/year, magnitudes exceeding Mw 5.5 and 5.0 can be considered complete and Poissonian, respectively. For magnitudes greater than Mw 6.0 and 6.5, they are likely complete and Poissonian since 1860 and 1700, respectively, with rates of 0.21 and 0.08 events/year.
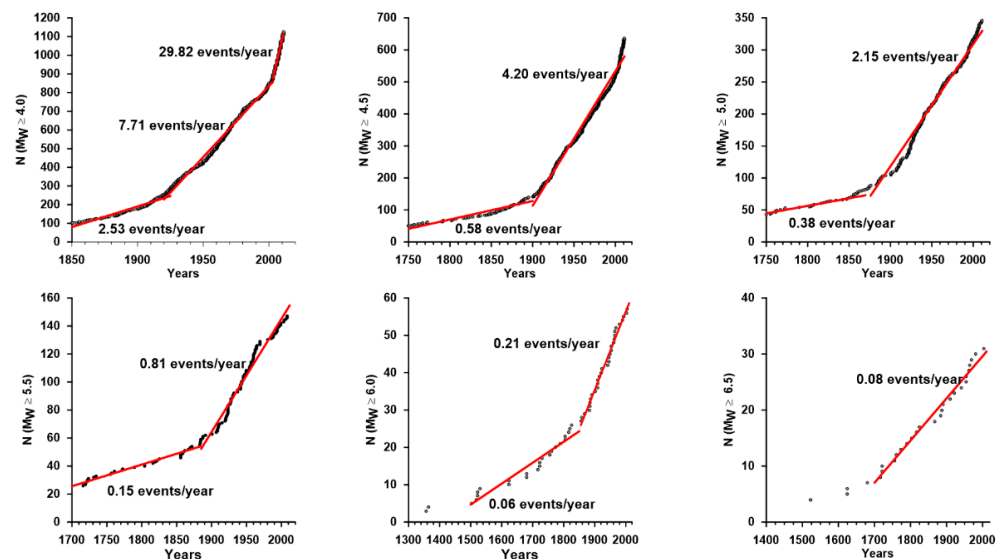

Figure 2. Cumulative number of earthquakes above magnitudes Mw 4.0, 4.5, 5.0, 5.5, 6.0 and 6.5 vs. time (black circles). The straight lines (in red) show linear fits.

3. Methodology Outline

The annual earthquake number series, denoted as $\{N_k; k \geq 0\}$, and the annual maximum magnitude in Northeastern Algeria that occurred during the period from 1910 to 2019, named $\{m_k, k \geq 0\}$, have been modelled using two different approaches. Initially, a parametric approach, abbreviated as ARIMA, was based on the Autoregressive Integrated Moving Average model, and a non-parametric approach was based on the Singular Spectrum Analysis model, abbreviated as SSA. ARIMA models [73], also known as Box–Jenkins models, are powerful tools in time series analysis aiming to describe the autocorrelations in the data and forecast values in the univariate time series that are non-stationaries, which are the time series with a trend component [74]. Usually, the notation *ARIMA* (*p*,*d*,*q*) is used, where *q* is the parameter of the moving average (MA) model, *p* is the parameter of the autoregression (AR) model, and d is the number of the differentiation procedure necessary to ensure the stationarity in the series. The extensions to ARIMA models are the Seasonal Box–Jenkins models [75], which support the direct modelling of the seasonal component of the non-stationaries time series exhibiting both the trend and seasonal fluctuations [74]. For seasonal series of a given period *S*, a Seasonal Autoregressive Integrated Moving Average (SARIMA, or

Seasonal ARIMA) is introduced with the notation $SARIMA(p, d, q) \times (P, D, Q)_S$, with P and Q being the orders of the seasonal autoregressive and seasonal moving average polynomials. Let $(X_t)_{t \geq 0}$ be a time series. The $ARIMA(p, d, q)$ equation model takes the form

$$\Phi_p(B) \triangledown^d X_t = \Theta_p(B) \varepsilon_t \tag{1}$$

where ε_t is the error series representing the white noise of mean 0 and variance σ_ε^2, and

$$\triangledown^d = (1 - B)^d \tag{2}$$

represents the difference operator of order d, being d, the order of integration required to achieve the stationarity in the data. The polynomial

$$\Phi_p(B) = 1 - \varphi_1 B^1 - \varphi_2 B^2 - \ldots - \varphi_p B^p \tag{3}$$

corresponds to the AR term at the pth order and

$$\Phi_p(B) = 1 - \varphi_1 B^1 - \varphi_2 B^2 - \ldots - \varphi_p B^p \tag{4}$$

is the MA polynomial at the qth order [76].

The first stage in the Box–Jenkins analysis procedure is to ensure that the analyzed series is stationary; that is, free of trend and seasonal terms. The plot of the autocorrelation (ACF) and partial autocorrelation functions (PACF) are used to determine the parameters p and q, which control substantially the model. According to Cowpertwait and Metcalfe [77], for a second-order stationary time series, the autocovariance function of the lag k is given by

$$\gamma_k = E[(x_t - \mu)(x_{t+k} - \mu)] \tag{5}$$

noticing that the number of time steps between the variables is known as the lag.

The lag k autocorrelation function (ACF) ρ_k, is defined as

$$\rho_k = \frac{\gamma_k}{\sigma^2} \tag{6}$$

being μ and σ^2 the mean and the variance of the time series, respectively, where $\rho_0 = 1$. In general, the partial autocorrelation function (PACF) of stationary time series at lag k is the k-th coefficient of a fitted AR(k) model; if the underlying process is AR(p), then the coefficients will be zero for all $k > p$. It measures the correlation between observations that are separated by k time units (e.g., x_t and x_{t-k}) after removing the effect of any correlation resulting by the terms at shorter lags (e.g., $x_{t-1}, \ldots \ldots \ldots \ldots, x_{t-k-1}$). The Akaike information criterion [78] and the Bayesian information criterion [79], denoted as AIC and BIC, respectively, are used to select the best model, the one with minimum information criterion values and white noise error series. According to Fabozzi et al. [80], it is worth noting that the AIC criterion represents the relationship between the Kullback–Leibler measure [81] and the maximum likelihood estimation method. The Kullback–Leibler measure is developed to capture the lost information in the estimation, which means that this measure selects the good model minimizing the loss of information. Usually, the AIC criterion is given by the equation

$$AIC = -2logL\left(\widetilde{\theta}\right) + 2K \tag{7}$$

where θ is the set of model parameters, $L(\widetilde{\theta})$ is the likelihood of the candidate model given the data when evaluated at the maximum likelihood estimate of θ, and K is the number of the estimated parameters in the candidate model.

For small samples, the corrected Akaike information criterion, denoted as AIC_c, should be used instead of the *AIC* criterion described previously. The relation giving the AIC_c is

$$AIC_c = -2 \, logL(\widetilde{\theta}) + 2K + \frac{2K+1}{N-K-1} \qquad (8)$$

where *N* is the number of observations. Hurvich and Tsai [82] define a small sample size as one that is less than 40. It is noteworthy that, as N increases, the third term in AIC_c approaches zero, producing the same result as the *AIC* criterion.

The Bayesian information criterion [79], denoted as *BIC*, is another model selection criterion based on the information theory but set within a Bayesian context. The difference between the *BIC* and the *AIC* criteria is that the former puts a greater penalty for the number of parameters than the latter. It is computed using the following relation

$$BIC = -2 \, logL(\widetilde{\theta}) + KlogN \qquad (9)$$

where the terms are the same as described in the definition of the *AIC* criterion. As in the previous criterion, the best model is the one providing the minimum *BIC* value.

The Single Spectrum Analysis (SSA) model is a time series forecasting method commonly used to analyze time series with periodic oscillations. The application of the SSA method is advantageous because it encompasses decomposing a time series into simpler components, such as a gradually changing trend, oscillations and noise. Various disciplines have implemented the SSA method, including signal processing [83], nonlinear dynamics [84,85], climate data [86–88], medical science [89] and mathematical statistics [90]. Furthermore, when combined with neural networks or other comparable techniques, it is a powerful pre-processing tool for time series forecasting [91,92]. Decomposition and reconstruction are the two complementary processes that make up the SSA implementation algorithm.

According to Golyandina and Zhigljavsky [93], we consider a real-valued time series $X = (x_1, \dots, x_N)$ with length *N* (*N* > 2), satisfying $x_j \neq 0$ for at least one *j*. The first step in the decomposition stage, named the embedding step, consists in transforming the original time series into a sequence of *L*-dimensional vectors, $X_i = (x_i, \dots, x_J)^T$, where $J = N - L + 1$, and *L* is the window length. The *J*-formed vectors are called *L*-lagged vectors and present the columns of the trajectory matrix (or *L*-trajectory matrix) of the series X ($L \times J$) given by

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_J \\ x_2 & x_3 & \cdots & x_{J+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{pmatrix} \qquad (10)$$

The window length *L* is obtained through experimentation, and the appropriate *L* parameter is determined according to the problem being considered and some preliminary information from the time series [94]. It is worth noting that there are no general rules for the determination of this parameter. However, it is advised that for time series exhibiting seasonality, the window length can be chosen as common multiples of 12 [95].

The second step in this stage, called the singular value decomposition, consists of performing a singular value decomposition method to the trajectory matrix X. Here, we will define the matrix $S = X.X^T$, and denote by $\lambda_1, \lambda_2, \dots, \lambda_L$ the eigenvalues of S that are taken in the decreasing order ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$), and by $U_1, U_2, \dots, U_L$ the orthonormal system corresponding to the eigenvalues of *S*. Let *d = rank (X)*, which is equal to *max{i, such that $\lambda_i > 0$}* and

$$V_i = X^T U_i / \sqrt{\lambda_i} \text{ for } i = 1, \dots, d \qquad (11)$$

Using this notation, the singular value decomposition of the trajectory matrix *X* can be written as

$$X = X_1 + X_2 + \dots + X_d \qquad (12)$$

with

$$X_i = \sqrt{\lambda_i} U_i V_i^T \tag{13}$$

The matrices X_i, called elementary matrices, with the rank equal to one, and the set of triplets $(\sqrt{\lambda_i}, U_i, V_i)$ are called the ith eigen-triplet of the singular values decomposition.

The second stage of the algorithm implementation corresponding to the reconstruction stage includes two other steps: the eigentriple grouping and the diagonal averaging steps. The eigentriple grouping step consists of dividing the elementary matrices X_i ($i = 1, \ldots, L$) into r groups, $1 \leq r \leq d$ [94]. The resulting matrices are then produced by adding the r eigen-triples in each group.

Whereas, the diagonal averaging step is based on the reconstruction of the one-dimensional series of length N, that can be considered as an approximation of the original series, by applying the diagonal averaging method on the grouped matrices that are resulted in the previous step [93,96].

In order to compare between the different approaches applied, an important convention is widely used and based on the root mean square error (RMSE) criterion [38,39], which is calculated by

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (x_k - \hat{x}_k)^2} \tag{14}$$

where x_k indicates the actual value and $\hat{x}_k$ represents the kth forecasted value based on the previous data. Therefore, the optimum model is the one presenting the smallest RMSE value.

4. Earthquake Magnitude Forecasting

The magnitude forecasting analysis for earthquakes with magnitudes equal to or greater than Mw 4.0 is implemented in this section using the previously described models. The data file used in this section includes the main shocks that occurred between 1910 and 2019 in Northeastern Algeria.

Figure 2 shows that the occurrence process during this time period can be approximated by a Poisson process. In this section, we will focus on two time series that represent the annual number of earthquakes and the annual maximum magnitudes in this region.

Here, we denote for $j = 1, \ldots, N$ by $m_j^{(k)}$ the magnitudes above Mw 4.0 of events occurred during the k-th year and define m_k as the annual maximum magnitude $\max_{0 \leq j \leq N} m_j^{(k)}$ during the k-th year, whereas the number of earthquakes during the considered year, the k-th year, is denoted as N_k.

Figure 3a displays the $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ plots covering the considered time period using a threshold magnitude equal to Mw 4.0. The ARIMA and SSA time series models are then employed to forecast both the frequency and annual maximum magnitude of earthquakes using the R-packages [97] RSSA and FORECAST [98,99]. By comparing the observed and predicted values, it is critical to identify the fitting models [74]. To increase the accuracy and reduce the rate of uncertainty, we have divided the dataset into training and testing data. We have identified that between 1910 and 2000, nearly 80% of the total events in the series $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ are included, corresponding to 70 observations. As a result, we estimate that 80% of the data are composed of training data, while the remaining 20% are events that were recorded between 2001 and 2019.
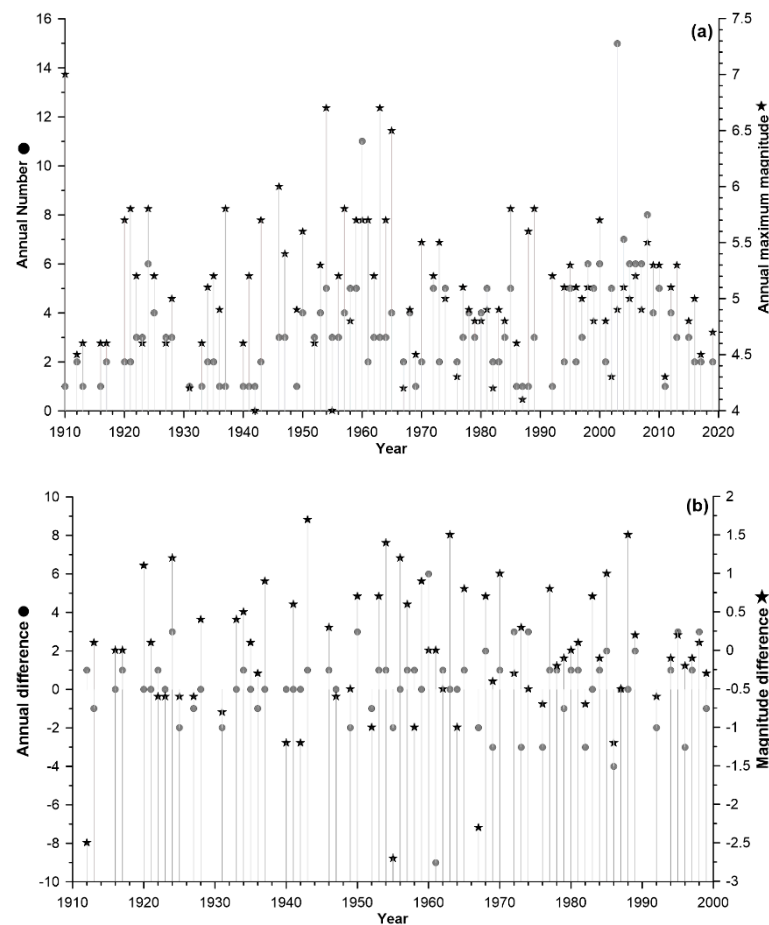
Figure 3. (**a**) Annual number of earthquakes (circles) and maximum annual magnitude (stars) in the period 1910–2019. (**b**) Differential time series for annual earthquake number (circles) and for maximum annual magnitude (stars) in the period 1910–2000.

Figure 3a shows two significate peaks in the $\{N_k; k \geq 0\}$ plot that refer to the high number of earthquakes that occurred between 1960 and 2003, and an important peak in the $\{N_k; k \geq 0\}$ plot corresponding to the biggest earthquakes that occurred in the studied region during this period.

The stationarity and the normality of the studied series are tested using the Augmented Dickey Fuller (ADF) [100] and the Shapiro–Wilk (SW) tests [101]. The obtained results are shown in Table 1. The ADF test indicates the non-stationarity in $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ time series with a *p*-value greater than 5%, and according to SW test (Table 1), the normality hypothesis is rejected for both series (*p*-value less than 5%). Thus, the $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ series are not normally distributed variables with non-stationary behaviour, which means that these non-parametric methods can be more appropriate for our study.

The stationarity of the data is an important hypothesis in ARIMA modelling. As a result, the first order differentiated series of $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ are computed. Figure 3b shows the differentiated resultant time series, with the two plots indicating a clear stationarity in the data. Thus, the integration parameter d is fixed in this study to 1. The *p* and *q* parameters of the ARIMA model are determined from the ACF and PACF plots of the differentiated series of $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$, respectively, which are shown in Figure 4. In the set of candidate ARIMA models, the optimum models for $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ are the ones that minimize the AIC and BIC criteria and have a normally uncorrelated residual. Hence, according to Table 1, the ARIMA (2,1,1) and ARIMA (2,1,2) models are the selected ones for $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ time series, respectively. Figure 5 illustrates the residuals analysis of the two models, where Figure 5a shows the residuals of the selected models. Figure 5b shows the ACF residual plots, indicating that

the residuals of the two models are uncorrelated. Finally, Figure 5c shows that the residuals are normally distributed according to the residuals Quantile-Quantile plots (Q-Q plot) statistical test [102,103].

Table 1. Annual earthquake number and annual maximum magnitude time series stationarity (ADF) and normality (SW) tests results. ARIMA model selection using the AIC and BIC criterions are also showed.

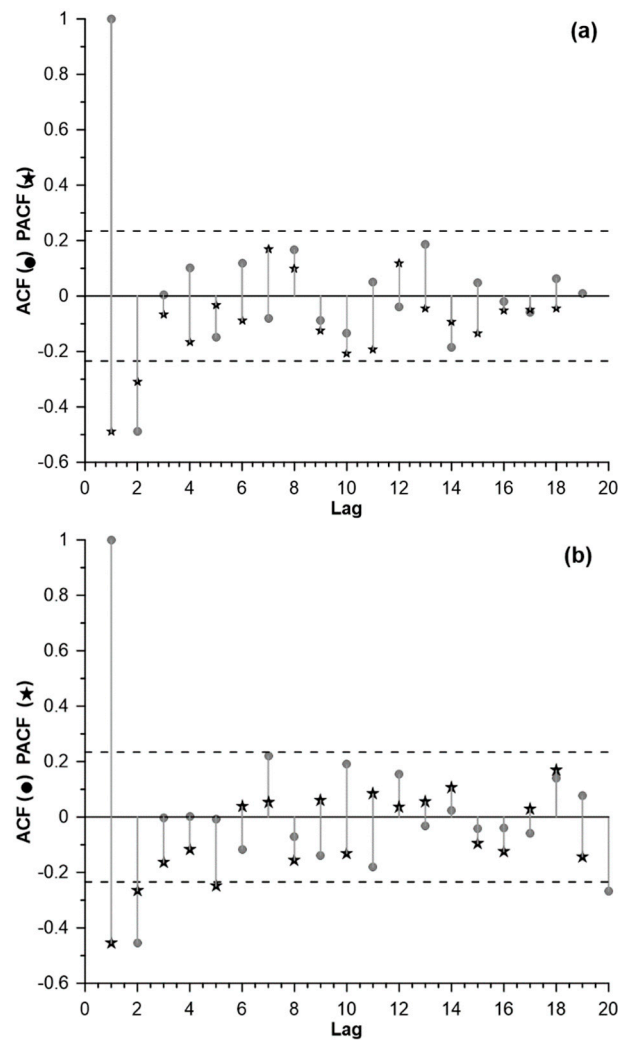| Series	ADF Test	SW Test	Model	AIC	BIC
Eartq_Num	0.345	7.21×10^{-7}	ARIMA (2,1,2)	4.061	4.255
			ARIMA (2,1,1)	3.995	4.175
			SARIMA (2,1,2) $(1,0,1)_{(S=12)}$	4.068	4.294
Max_Mag	0.063	0.029	ARIMA (2,1,2)	2.169	2.364
			ARIMA (2,1,1)	2.191	2.425
			SARIMA (2,1,2) $(1,0,1)_{(S=12)}$	2.198	2.425

Figure 4. ACF (circles) and PACF (stars) values (**a**) for the differential annual earthquake number and (**b**) for the differential annual maximum magnitude.
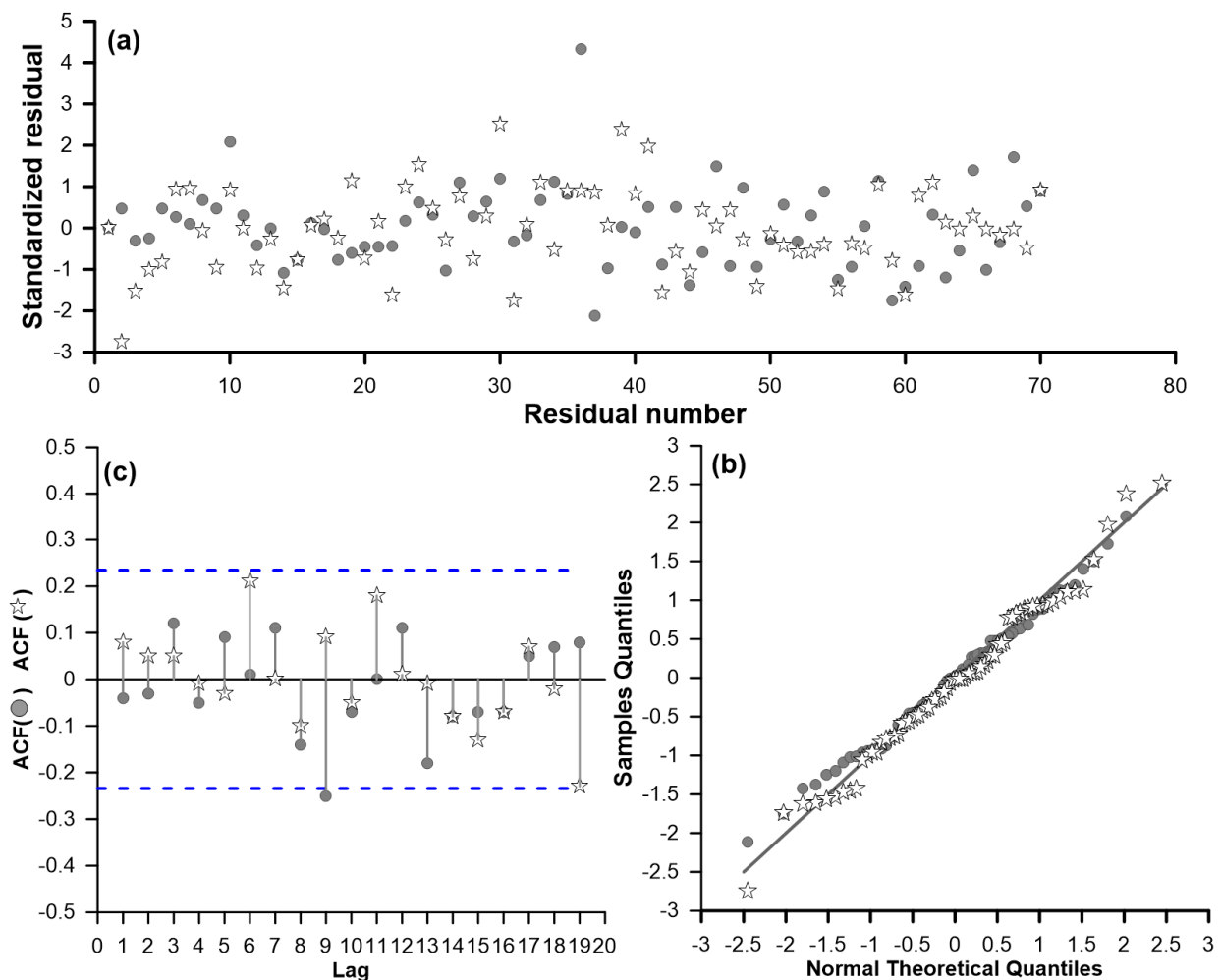
Figure 5. Residual analysis of the selected ARIMA model for earthquakes number time series (circles) and for annual maximum magnitudes time series (stars). (**a**) residuals of the selected models; (**b**) ACF residual plots, indicating that the residuals of the two models are uncorrelated; (**c**) residuals are normally distributed according to the residuals Quantile-Quantile plots (Q-Q plot).

The second approach considered in our study is based on the SSA method. The first stage of the SSA method consists of decomposing the $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ time series into principal independent components, which are the trend, the seasonality and the noise. The second stage consists of reconstructing the original series using only the trend and the seasonality. Figure 6 indicates that the seasonality and trend components used together (blue line) describe the observed earthquake number and the annual maximum magnitude in the studied region better than the trend component alone (red line).

The parameters of the SSA model were empirically chosen based on a visual presentation and clear separability of the main independent components. According to Hassani [104], a window length less than half of the sample size is considered adequate. According to Golyandina et al. [94], for seasonal time series, this parameter must be a common multiple of the seasonality period. Several window lengths (12, 24 and 36 samples) were thus evaluated to select the best one, which has been chosen using the RMSE criterion. The reconstruction stage parameter (see Section 3), denoted r, is determined using the eigenvalues plot of the $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ series presented in Figure 7, where the slow decrease in these plots suggests the beginning of the noise component. Then, according to Figure 7, two components are used both in the reconstruction process of the $\{N_k; k \geq 0\}$ series and in the reconstruction of the $\{m_k, k \geq 0\}$ series. These results can be confirmed using the w-correlation matrices shown in Figure 8. The w-correlation matrices

indicate that the first three components are uncorrelated in the $\{N_k; k \geq 0\}$ series (Figure 8a), and two components in the $\{m_k, k \geq 0\}$ series (Figure 8b). Therefore, from the w-correlation matrices, the reconstruction parameter is fixed to 3 and 2 for the $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ time series, respectively.
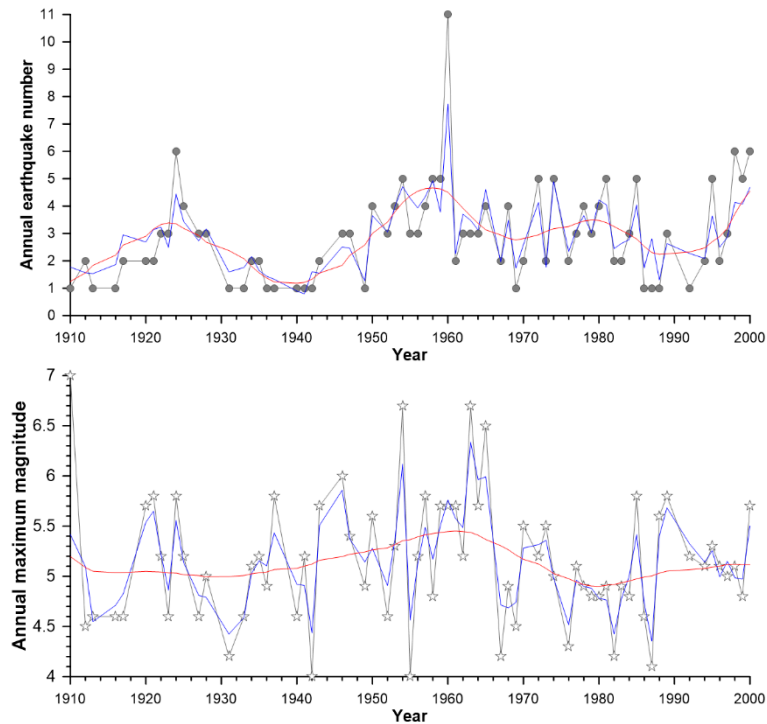
Figure 6. Fit of the recorded annual number and the annual maximum earthquake magnitude (grey line) by the seasonality and trend components combined (blue line) and by the trend component alone (red line).
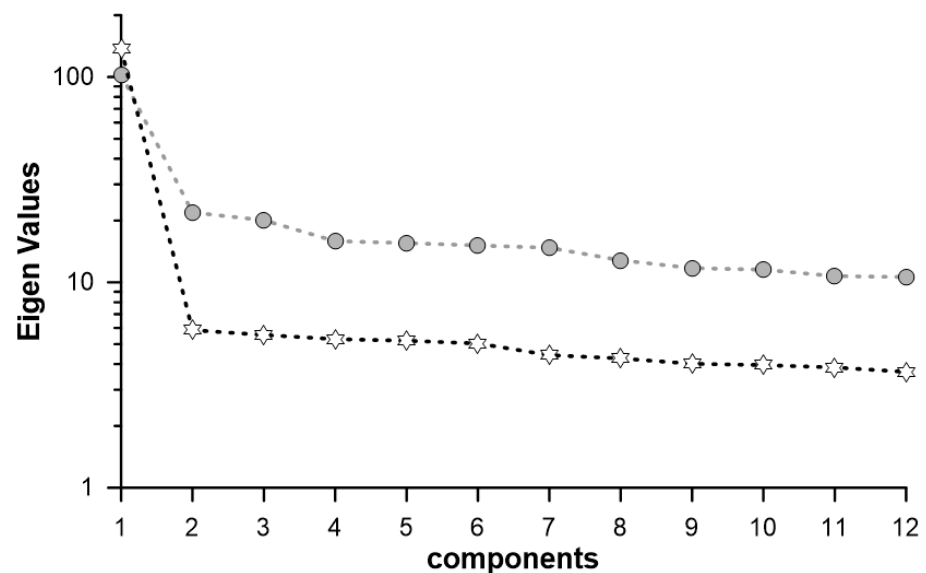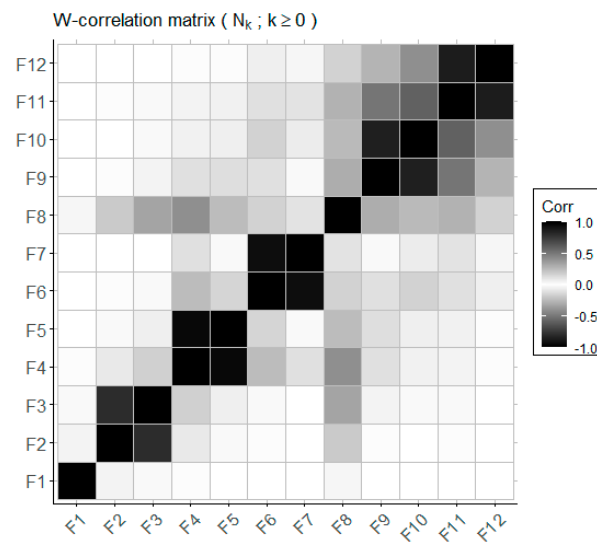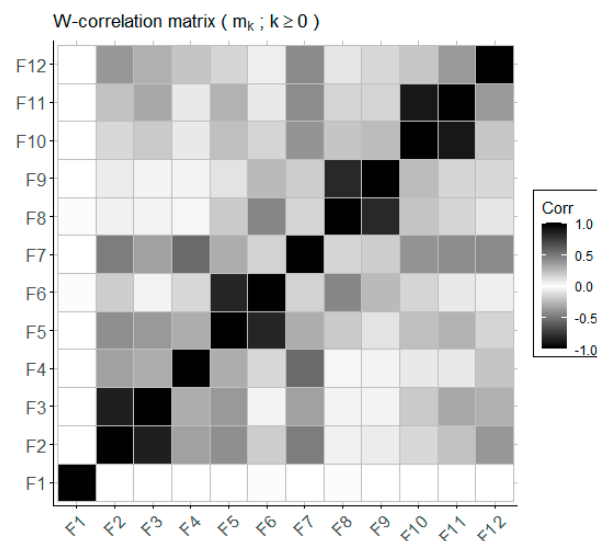
Figure 7. Eigenvalues plots of the first 12 components for the annual earthquake number series (circles) and for the annual maximum magnitudes series (stars).

(a)

(b)

Figure 8. W-correlation matrix plots for the first 12 components of the earthquake number series (**a**) and the annual maximum magnitude series (**b**).

The first group is composed of solely the first component, which is uncorrelated to the other components. Because the second and third components are highly correlated, the pair (2,3) forms the second group. After linking the components from 4 to 8, it is shown that the fourth component has a correlation with the fifth, sixth and eighth components. In addition, the sixth component has a high correlation with the seventh component. Thus, the pairs (4,5) and (6,7), as well as the eighth component, configure the last group.

Therefore, the rest of components in $\{N_k; k \geq 0\}$ series correspond to the noise. For $\{m_k, k \geq 0\}$ series, according to the w-correlation matrix in Figure 8b, two groups can be used in the reconstruction of the original series. The first group consists of mainly the first component, whereas the second group consists of Components 2–8 and Component 12, which are all correlated with one another. As a result, all of the other components form the noise in the $\{m_k, k \geq 0\}$ series. It is worth noting that the first component in both matrices that is uncorrelated with the other component corresponds to the trend component in both

series. Therefore, the reconstruction parameters in this case are fixed to 3 and 2 for the $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ time series, respectively.

Table 2 presents the RMSE value of the candidate SSA models for the studied series. According to the results given in Table 2, the best SSA models for $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ are the SSA (24,3) and SSA (12,2) models, respectively. Figure 9 depicts the independent main component extracted from the $\{N_k; k \geq 0\}$ (a) and $\{m_k, k \geq 0\}$ (b) series, representing the trend, seasonality and noise components.

Table 2. The SSA chosen model based on the RMSE of the analyzed time series.

Series	Model	*RMSE*
Earthq_Num	SSA (12,2)	3.847
	SSA (24,3)	3.048
	SSA (36,2)	3.426
Max_Mag	SSA (12,2)	0.314
	SSA (24,2)	0.365
	SSA (36,2)	0.395

Finally, the RMSE value is used to compare the different approaches. Table 3 displays the RMSE values of the ARIMA and SSA models for the $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ time series, with the smaller RMSE value indicating the better model.

Table 3. Best model selected based on the RMSE value for the two considered time series.

Series	ARIMA (*p*,*d*,*q*) RMSE	SSA (*L*,*r*) RMSE
Earthq_Num	(2,1,1) 3.317	(24,3) 3.048
Max_Mag	(2,1,2) 0.374	(12,2) 0.314

As a result, the SSA (24,3) and SSA (12,2) models are the best ones for describing the number of earthquakes and the maximum magnitude series, respectively. Table 4 also shows the forecasted values from 2020 to 2030 using the selected SSA models. Finally, Figure 10 depicts the original (grey dashed line) and forecasted (red dashed line) series, where a stationary behavior can be observed in both time series from 2020 to 2030.

Table 4. Forecasted annual earthquake number and annual maximum magnitude time series from 2020 to 2030 from the selected SSA model.

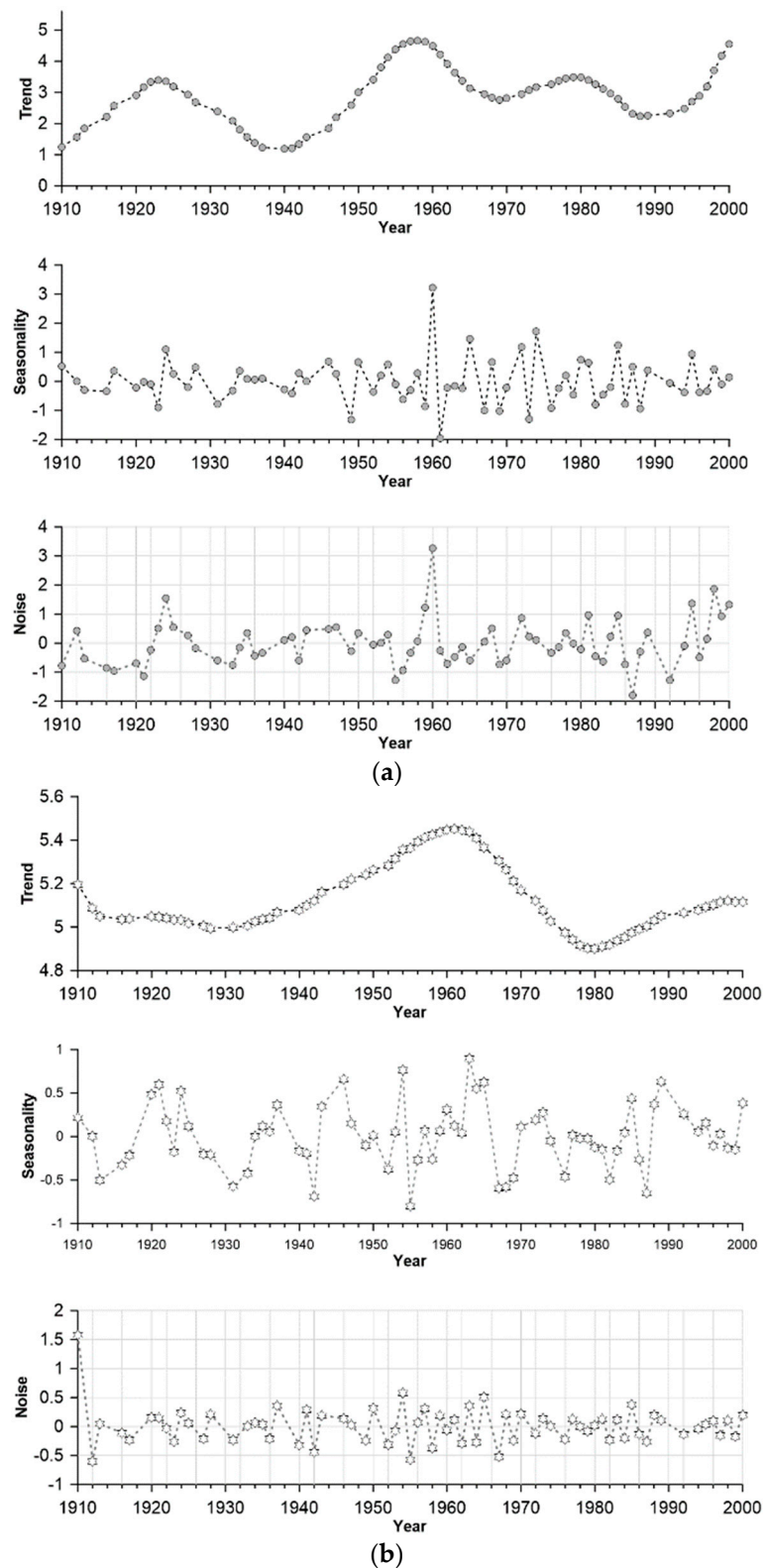| Years	Series	
	Earthq_Num	**Max_Mag**
2020	5	4.9
2021	6	5.1
2022	5	4.8
2023	5	4.9
2024	5	5.1
2025	4	5.0
2026	5	4.9
2027	6	5.0
2028	4	5.1
2029	4	5.0
2030	5	5.1

(**a**)

(**b**)

Figure 9. Trend, seasonality and noise components for the annual earthquake number (**a**) and maximum annual magnitude (**b**) series.
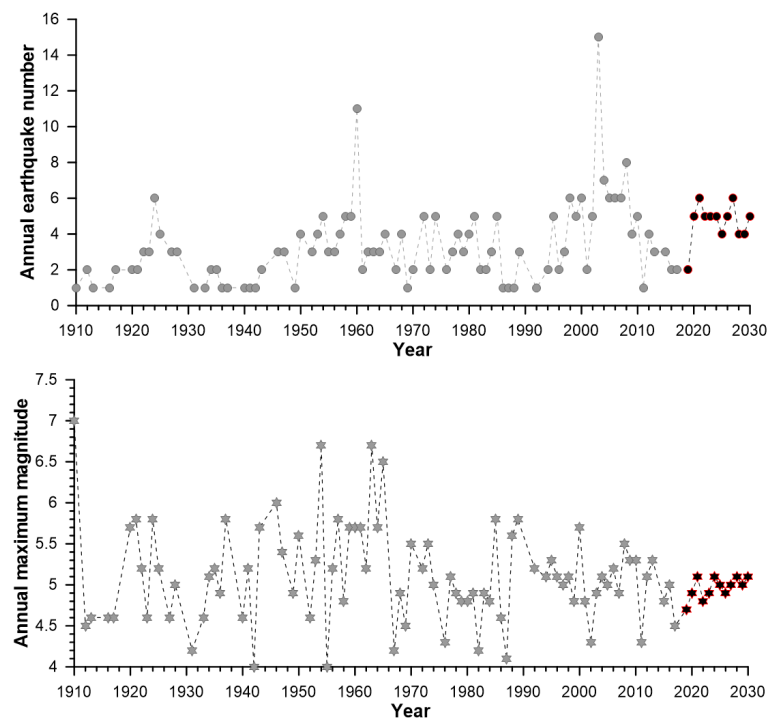
Figure 10. Observed and forecasted (in red) values for the number of earthquakes (top) and the annual maximum magnitude (bottom) using the SSA model.

5. Validation of the Procedure

After a certain earthquake forecast method has been developed using training data and set up by best fit on real data, and before any practical implementation, it must be validated [105].

The SSA method is used to perform forecasting analysis of the earthquake data file from 1910 to 2015, with the purpose of generating forecast values and comparing them to observed values from 2016 to 2019. As previously stated, the time series are divided into two sets, the first of which includes 80% of the data used as training data [74] and encompasses the years 1910 to 1997, while the remaining 20% covers the years 1998 to 2015, considered as testing data.

The stationarity and the normality of the studied series in the considered training period are tested again using the ADF and the SW tests. The derived ADF values are 0.34 and 0.077 for the annual earthquake number and maximum magnitude, respectively, whereas the obtained SW values are 2.0×10^{-7} and 0.035 for the two investigated series, respectively. The ADF test results for the two series are higher than 5%, showing that the two series are non-stationary during the training period. Moreover, the two series are not normally distributed and exhibit a non-stationary behaviour as a result of the computed SW values. Therefore, as proceeded initially, and in order to achieve the stationarity in the data, the differentiated series of $\{N_k; k \geq 0\}$ and $\{m_k, k \geq 0\}$ are computed.

Considering the d parameter of the ARIMA model equal to 1, the ACF and PACF functions of the resultant series are used to select the *p* and *q* parameters. Thus, ARIMA (2,1,2) and ARIMA (2,1,3) are the model-candidates for the time series $\{N_k; k \geq 0\}$, and ARIMA (2,1,2) and ARIMA (3,1,2) for the time series $\{m_k, k \geq 0\}$, derived during the training period 1910–1997. Analyzing the annual earthquake number series, AIC values equal to 3.99 and 4.00 are obtained for the ARIMA (2,1,2) and ARIMA (2,1,3) models, respectively, whereas BIC values equal to 4.20 and 4.24 are obtained for the same model candidates, respectively. In addition, analyzing the annual maximum magnitude time series, the obtained AIC values are equal to 2.21 and 2.23 for the ARIMA (2,1,2) and ARIMA (3,1,2) models, respectively. For the same model-candidates, the BIC values are equal to 2.41 and 2.16. Thus, on the base of AIC and BIC criteria, the ARIMA (2,1,2) model is considered the

best one to describe the annual earthquake number and the annual maximum magnitude during the training period.

The parameters of the SSA model are then obtained, as previously proceeded, by testing several window lengths (12, 24 and 36 samples). The best model is also derived using the RMSE criterion. Analyzing the annual earthquake number time series, the lower RMSE value equal to 3.62 is derived for the model SSA (24,2), whereas for the models SSA (12, 2) and SSA (36,2), the values are equal to 4.40 and 3.89, respectively. Concerning the annual maximum magnitude time series, the lower RMSE is derived for the model SSA (36,2), whereas for the models SSA (12,2) and SSA (36,2), the values are equal to 0.379 and 0.375, respectively. Finally, for the two considered time series, SSA (24,2) is considered the best model.

Then, using the previously described procedure, the chosen models ARIMA (2,1,2) and SSA (24,2) are used to forecast the annual number of earthquakes and the annual maximum magnitude from 2016 to 2019. Uncommonly, no events with magnitudes greater than Mw 4.0 occurred in the studied region in 2018, whereas two events with magnitudes greater than Mw 4.0 occurred in 2016, 2017 and 2019. Then, the period from 2016–2019 appears as a low seismicity epoch. That being the case, the forecasted values using the model ARIMA (2,1,2) are greater than real ones. In 2016 and 2019, about four events were predicted, and three events were predicted in 2017.

On the other hand, values found using the SSA (24,2) model are of the order of 3 for each year, and hence closer to the observed values. Regarding the annual maximum magnitude, similar values are obtained using the ARIMA (2,1,2) and SSA (24,2) models, of the order of Mw 5.0 for 2016, 2017 and 2019. These forecasted values are in agreement with the observed ones, which are equal to Mw 5.0, 4.5 and 4.7 for 2016, 2017 and 2019, respectively.

Thus, clearly the SSA (24,2) model makes it possible to find realistic results corroborating the observed data.

A statistical test is used to assess the consistency of the results obtained using the SSA algorithm in the period from 2016 to 2019. According to Schorlemmer et al. [19], to take into account the uncertainty, the likelihood test, named L-test, can be conducted by simulating the observed events. In the current study, we focus on the statistical N-test [42], which consists to test the rate forecast. The N-test is intended to measure, in a probabilistic manner, how the forecasted number of events will match the observed number of earthquakes.

Then, the N-test [19,42,106] is implemented. The N-test takes into account two linked quantile scores, δ_1 and δ_2, reflecting whether the produced sequences produced forecasted event numbers N_{fore} that were higher or lower than the observed values N_{obs}, as given in the equations

$$\delta_1 = 1 - P\Big((N_{obs} - 1)\Big|N_{fore}\Big) \tag{15}$$

$$\delta_2 = P\Big(N_{obs}\Big|N_{fore}\Big) \tag{16}$$

bearing in mind that $P(\omega|\lambda) = \frac{\lambda^\omega}{\omega!}exp(-\lambda)$.

The quantity δ_1 is the probability of observing at least N_{obs}, and δ_2 is the probability of observing at most N_{obs}. Both the overall forecast rate and observed number of events are assumed to be Poissonian and described by N_{fore} and N_{obs}, respectively. The quantity δ_2 describes the right-continuous Poisson cumulative distribution with the expectation N_{fore} at corresponding N_{obs} at the times evaluated. This score describes the fraction of forecast expectations smaller than the observed events. The probability that more than N_{obs} events are forecasted is given by $(1 - \delta_2)$. The problem with this approach is addressed in [42]. Instead, the δ_1 probability was added in addition to the original N-test to describe at least N_{obs}, in which the user only needs to be concerned about low probability values [19,42].

The quantile scores in this statistical N-test assess if the number of forecasted occurrences is inconsistent with N_{obs}. A small δ_1 indicates that the forecast underestimates the observed sequence, whereas a small δ_2 suggests that the forecast overestimates the number of occurrences. The forecast can thus be rejected if the probabilities of δ_1 and δ_2 are less than

the effective significance level [42]. The effective significance threshold for the one-sided N-tests was set at 2.5% to coincide with a single quantile score, which corresponds to a 5% error rate for the test [42,107].

Table 5 presents the obtained results, the observed N_{obs} and the forecasted number of earthquakes N_{fore}, and the quantile scores δ_1 and δ_2. The N-test performance can be interpreted by observing whether the quantile scores δ_1 and δ_2 fall the region between 0.025 and 1. In Figure 11, we can observe that high δ_1 scores tend to correspond to low δ_2 scores.

Table 5. Observed and forecasted earthquake number using the selected SSA model quantile score δ_1 and δ_2.

Year	N Observed	N Forecasted	N-Test δ_1	N-Test δ_2
2016	2	3.0	0.8506	0.2240
2017	2	3.2	0.8705	0.2079
2018	0	3.5	1	0.0302
2019	2	3.4	0.8865	0.1929
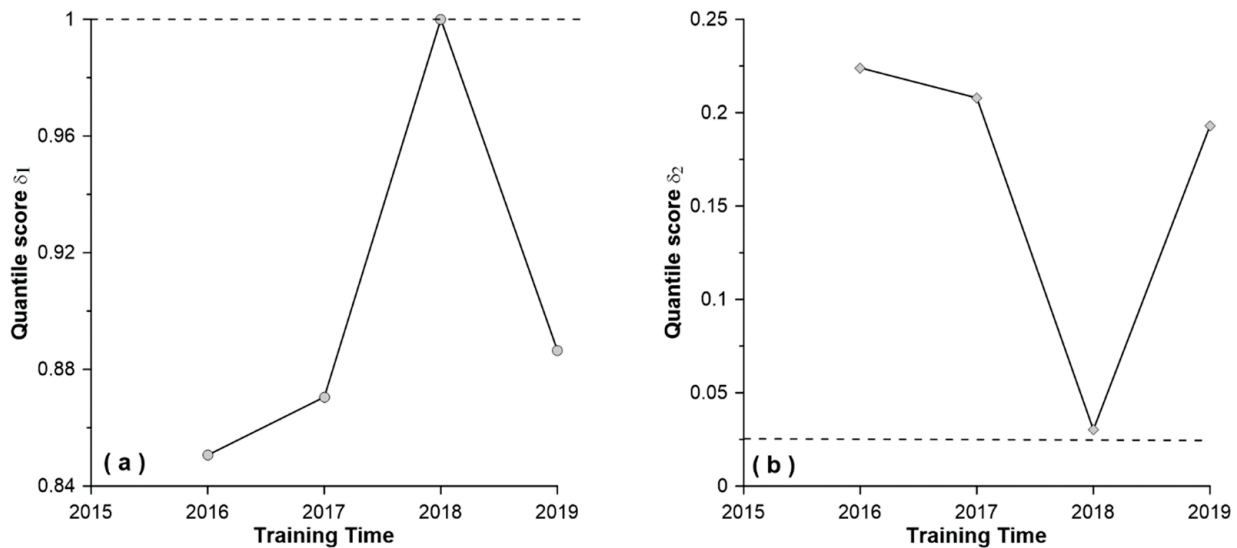
Figure 11. N-test values for (**a**) δ_1 index and (**b**) δ_2 index.

6. Discussion and Conclusions

The SSA technique is a well-known and effective time series analysis tool. On the studied earthquake magnitude time series, the forecasting capabilities of the SSA method were evaluated and compared with those of the standard ARIMA one, which is recognized to have a more suitable structure for forecasting.

The results of this study show that the SSA technique could be effectively applied as an algorithm for forecasting earthquake number and sizes.

According to Abacha et al. [61], the seismicity in the studied area is continuous and moderate. Several swarm sequences occurred in the recent past, showing intense activity with low magnitudes and limited time. An analysis of the series covering the years 1910 to 2015 is performed in order to validate the method employed in the current study by forecasting for the years 2016 to 2019. The forecasting values agree very well with the observed events both for the frequency and the annual maximum magnitude. Moreover, according to the recent recorded seismicity in the studied area, the forecasted magnitudes computed in the current study agree very well with the recorded seismicity during the years 2020 and 2021. The recorded magnitude events during these two years are lower than the maximum forecasted magnitudes, equal to 4.9 and 5.1, respectively. These results should to be considered as the maximum magnitude that is not exceeded.

The SSA model successfully modelled the considered data file and provided an accurate forecast of future earthquake magnitudes. Furthermore, the SSA model's trend component is a clear deterministic component for estimating earthquake magnitude.

The N-test was employed in this study to assess the consistency between the observation and forecasting in terms of the total number of events. The results were given in terms of quantile scores δ_1 and δ_2. According to the results, the test is passed, which validates the suggested model.

The results also show that, while significant parts of the variability in those seismological time series do not appear to have an adequate time structure to be forecasted, the key components may still be identified and forecasted using the SSA model.

It is worth noting that the data file used in this study, which spans the years 1910 to the end of 2019, including events with magnitudes greater than Mw 4.0, could be modelled by an inhomogeneous Poisson process with a variable activity rate (Figure 2). A future study could forecast the number of earthquakes using such an inhomogeneous Poisson process, while the annual maximum magnitude could also be forecasted using, for instance, the extreme probability statistics [25]. An issue that appears to be a weakness of these procedures, and deserves special attention in the future, is the assessment of uncertainties of the forecasted values.

References

1. DeMets, C.; Gordon, R.G.; Argus, D.F.; Stein, S. Current plate motions. *Geophys. J. Int.* **1990**, *101*, 425–478. [CrossRef]
2. Calais, E.; Demets, C.; Nocquet, J.M. Evidence for a post-3.16 Ma change in Nubia-Eurasia plate motion. *Earth Planet. Sci. Lett.* **2003**, *216*, 81–92. [CrossRef]
3. Serpelloni, E.; Vannucci, G.; Pondrelli, S.; Argnani, A.; Casula, G.; Anzidei, M.; Baldi, P.; Gasperini, P. Kinematics of the western Africa-Eurasia plate boundary from focal mechanisms and GPS data. *Geophys. J. Int.* **2007**, *169*, 1180–1200. [CrossRef]
4. EERI-Earthquake Engineering Research Institute. El-Asnam. In *Algeria Earthquake of October 10, 1980: A Reconnaissance and Engineering Report*; National Academy of Sciences: Washington, DC, USA, 1983.
5. Hamdache, M.; Peláez, J.A.; Yelles-Chauche, A.K. The Algiers, Algeria earthquake (Mw 6.8) of 21 May 2003: Preliminary report. *Seismol. Res. Lett.* **2004**, *75*, 360–367. [CrossRef]
6. ASCE-American Society of Civil Engineers. *Zemmouri, Algiers, Mw 6.8 Earthquake of May 21, 2003*; Report to the Technical Council on Lifeline Earthquake Engineering Monograph No. 27; American Society of Civil Engineers: Reston, VA, USA, 2004.
7. Hamdache, M.; Peláez, J.A.; Talbi, A.; López Casado, C. A Unified Catalog of Main Earthquakes for Northern Algeria from A.D. 856 to 2008. *Seismol. Res. Lett.* **2010**, *81*, 732–739. [CrossRef]
8. Keilis-Borok, V.I. What comes next in the dynamics of lithosphere and earthquake prediction? *Phys. Earth Planet. Inter.* **1999**, *111*, 179–185. [CrossRef]
9. Orihara, Y.; Noda, Y.; Nagao, T.; Uyeda, S. A possible case of SES selectivity at Kozushima Island, Japan. *J. Geodyn.* **2002**, *33*, 425–432. [CrossRef]
10. Tiampo, K.F.; Rundle, J.B.; McGinnis, S.; Gross, S.J.; Klein, W. Eigenpatterns in southern California seismicity. *J. Geophys. Res.* **2002**, *107*, ESE 8-1–ESE 8-17. [CrossRef]
11. Tiampo, K.F.; Rundle, J.B.; Mcginnis, S.; Klein, W. Pattern dynamics and forecast methods in seismically active regions. *Pure Appl. Geophys.* **2002**, *159*, 2429–2467. [CrossRef]

12. Kanamori, H. Earthquake prediction: An overview. In *International Handbook of Earthquake and Engineering Seismology*; Lee, W.H.K., Kanamori, H., Jennings, P.C., Kisslinger, C., Eds.; Academic Press: Amsterdam, The Netherlands, 2003; pp. 1205–1216.

13. Evison, F.F.; Rhoades, D.A. Demarkation and scaling of long-term seismogenesis. *Pure Appl. Geophys.* **2004**, *161*, 21–45. [CrossRef]

14. Holliday, J.; Chen, C.; Tiampo, K.F.; Rundle, J.B.; Turcotte, D.L. A RELM earthquake forecast based on pattern informatics. *Seismol. Res. Lett.* **2007**, *78*, 87–93. [CrossRef]

15. Kagan, Y. On Earthquake Predictability Measurement: Information Score and Error Diagram. *Pure Appl. Geophys.* **2007**, *164*, 1947–1962. [CrossRef]

16. Papazachos, B.C.; Karakaisis, G.F.; Papazachos, C.B.; Scordilis, E.M. Evaluation of the results for an intermediate-term prediction of the 8 January 2006 Mw $\frac{1}{4}$ 6.9 Cythera earthquake in the Southwestern Aegean. *Bull. Seismol. Soc. Am.* **2007**, *97*, 347–352. [CrossRef]

17. Mignan, A. The stress accumulation model: Accelerating moment release and seismic hazard. *Adv. Geophys.* **2008**, *49*, 67–191. [CrossRef]

18. Field, E.H. Overview of the working group for the development of regional earthquake likelihood models (RELM). *Seismol. Res. Lett.* **2007**, *78*, 7–16. [CrossRef]

19. Schorlemmer, D.; Gerstenberger, M.C.; Wiemer, S.; Jackson, D.D.; Rhoades, D.A. Earthquake likelihood model testing. *Seismol. Res. Lett.* **2007**, *78*, 17–29. [CrossRef]

20. Schorlemmer, D.; Zechar, J.D.; Werner, M.J.; Field, E.H.; Jackson, D.D.; Jordan, T.H.; Relm Working Group. First Results of the regional earthquake likelihood models experiment. *Pure Appl. Geophys.* **2010**, *167*, 859–876. [CrossRef]

21. Zechar, J.D.; Schorlemmer, D.; Werner, M.J.; Gerstenberger, M.C.; Rhoades, D.A.; Jordan, T.H. Regional earthquake likelihood models I: First-order results. *Bull. Seismol. Soc. Am.* **2013**, *103*, 787–798. [CrossRef]

22. DeVries, P.M.R.; Viégas, F.; Wattenberg, M.; Meade, B.J. Deep learning of aftershock patterns following large earthquakes. *Nature* **2018**, *560*, 632–634. [CrossRef]

23. Lippiello, E.; Petrillo, G.; Godano, C.; Tramelli, A.; Papadimitriou, E.; Karakostas, V. Forecasting of the first hour aftershocks by means of the perceived magnitude. *Nat. Commun.* **2019**, *10*, 2953. [CrossRef]

24. Gutenberg, B.; Richter, C.F. Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.* **1944**, *34*, 185–188. [CrossRef]

25. Dargahi-Noubary, G.R. A method for predicting future large earthquakes using extreme order statistics. *Phys. Earth Planet. Inter.* **1986**, *42*, 241–245. [CrossRef]

26. Kagan, Y. Statistics of characteristic earthquakes. *Bull. Seismol. Soc. Am.* **1993**, *83*, 7–24.

27. Firuzan, E. Statistical Earthquake Frequency Analysis for Western Anatolia. *Turk. J. Earth Sci.* **2008**, *17*, 741–762.

28. Özel, G.; Inal, C. The probability function of the compound Poisson process and an application to aftershock sequence in Turkey. *Environmetrics* **2008**, *19*, 79–85. [CrossRef]

29. Özel, G. A bivariate compound Poisson model for the occurrence of foreshock and aftershock sequences in Turkey. *Environmetrics* **2011**, *22*, 847–856. [CrossRef]

30. Jackson, D.D.; Kagan, Y.Y. The 2004 Parkfield earthquake, the 1985 prediction, and characteristic earthquakes: Lessons for the future. *Bull. Seismol. Soc. Am.* **2006**, *96*, S397–S409. [CrossRef]

31. Helmstetter, A.; Sornette, D. Importance of direct and indirect triggered seismicity in the ETAS model of seismicity. *Geophys. Res. Lett.* **2003**, *30*, 1576. [CrossRef]

32. Ogata, Y. Statistics of Earthquake Activity: Models and Methods for Earthquake Predictability Studies. *Annu. Rev. Earth Planet. Sci.* **2017**, *45*, 497–527. [CrossRef]

33. Zhu, F. Modelling over dispersed or under dispersed count data with generalized Poisson integer-valued GARCH models. *J. Math. Anal. Appl.* **2012**, *389*, 58–71. [CrossRef]

34. Shishegaran, A.; Taghavizade, H.; Bigdeli, A.; Shishegaran, A. Predicting the earthquake magnitude along zagros fault using time series and ensemble model. *J. Soft Comput. Civ. Eng.* **2019**, *3*, 67–77. [CrossRef]

35. Serita, A.; Hattor, K.; Yoshino, C.; Hayakawa, M.; Isezaki, N. Principal component analysis and singular spectrum analysis of ULF geomagnetic data associated with earthquakes. *Nat. Hazards Earth Syst. Sci.* **2005**, *5*, 685–689. [CrossRef]

36. Guo, J.; Shi, K.; Liu, X.; Sun, Y.; Li, W.; Kong, Q. Singular spectrum analysis of ionospheric anomalies preceding great earthquakes: Case studies of Kaikoura and Fukushima earthquakes. *J. Geodyn.* **2019**, *124*, 1–13. [CrossRef]

37. Cekim, H.O.; Tekin, S.; Özel, G. Prediction of the earthquake magnitude by time series methods along the East Anatolian Fault, Turkey. *Earth Sci. Inform.* **2021**, *14*, 1339–1348. [CrossRef]

38. Silva, E.S.; Hassani, H.; Heravi, S. Modelling European industrial production with multivariate singular spectrum analysis: A cross industry analysis. *J. Forecast.* **2018**, *37*, 371–384. [CrossRef]

39. Hassani, H.; Yeganegi, M.R.; Khan, A.; Silva, E.S. The Effect of Data Transformation on Singular Spectrum Analysis for Forecasting. *Signals* **2020**, *1*, 4-25. [CrossRef]

40. Rabaute, A.; Chamot-Rooke, N. Active Inversion Tectonics from Algiers to Sicily. In *On Significant Applications of Geophysical Methods, Proceedings of the 1st Springer Conference of the Arabian Journal of Geosciences (CAJG 2018), Hammamet, Tunisia, 12–15 November 2018*; Sundararajan, N., Eshagh, M., Saibi, H., Meghraoui, M., Al-Garni, M., Giroux, B., Eds.; Springer: Cham, Switzerland, 2019; pp. 249–251. [CrossRef]

41. Peláez, J.A.; Hamdache, M.; Sanz de Galdeano, C.; Sawires, R.; García Hernández, M.T. Forecasting moderate earthquakes in Northern Algeria and Morocco. In *Earthquakes and Their Impact on Society*; Springer Natural Hazards Series; D'Amico, S., Ed.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 81–95.

42. Zechar, J.D.; Gerstenberger, M.C.; Rhoades, D.A. Likelihood-based tests for evaluating space-rate-magnitude earthquake forecasts. *Bull. Seismol. Soc. Am.* **2010**, *100*, 1184–1195.

43. Nanjo, K.Z.; Tsuruoka, H.; Yokoi, S.; Ogata, Y.; Falcone, G.; Hirata, N.; Ishigaki, Y.; Jordan, T.H.; Kasahara, K.; Obara, K.; et al. Predictability study on the aftershock sequence following the 2011 Tohoku-Oki, Japan, earthquake: First results. *Geophys. J. Int.* **2012**, *191*, 653–658. [CrossRef]

44. Peláez, J.A.; Henares, J.; Hamdache, M.; Sanz de Galdeano, C. A seismogenic zone model for seismic hazard studies. In *Northwestern Africa in Moment Tensor Solutions. A Useful Tool for Seismotectonics*; Springer Natural Hazards Series; D'Amico, S., Ed.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 643–680.

45. Peláez, J.A.; Henares, J.; Hamdache, M.; Sanz de Galdeano, C. An updated seismic model for the northwestern Africa. In Proceedings of the 16th European Conference (EECE), Thessaloniki, Greece, 18–21 June 2018.

46. Morel, J.L.; Meghraoui, M. The Goringe Alboran-Tell (Galtel) tectonic zone: A transpression system along the Africa-Eurasia plate boundary. *Geology* **1996**, *24*, 755–758.

47. Déverchère, J.; Yelles, K.; Domzig, A.; Mercier de Lépinay, B.; Bouillin, J.P.; Gaullier, V.; Bracène, R.; Calais, E.; Savoye, B.; Kherroubi, A.; et al. Active thrust faulting offshore Boumerdès Algeria, and its relations to the 2003 Mw 6.9 earthquake. *Geophys. Res. Lett.* **2005**, *32*, L04311. [CrossRef]

48. Nocquet, J.M.; Calais, E. Crustal velocity field of western Europe from permanent GPS array solutions, 1996–2001. *Geophys. J. Int.* **2003**, *154*, 72–88. [CrossRef]

49. Nocquet, J.M.; Calais, E. Geodetic measurements of crustal deformation in the western Mediterranean and Europe. *Pure Appl. Geophys.* **2004**, *161*, 661–681. [CrossRef]

50. Sparacino, F.; Palano, M.; Peláez, J.A.; Fernández, J. Geodetic deformation versus seismic crustal moment-rates: Insights from the Ibero-Maghrebian region. *Remote Sens.* **2020**, *12*, 952. [CrossRef]

51. Harbi, A.; Benouar, D.; Benhallou, H. Re-appraisal of seismicity and seismotectonics in the north-eastern Algeria. Part I: Review of historical seismicity. *J. Seismol.* **2003**, *7*, 115–136. [CrossRef]

52. Adams, R.D.; Barazangi, M. Seismotectonics and seismology in the Arab region: A brief summary and future plans. *Bull. Seismol. Soc. Am.* **1984**, *74*, 1011–1030.

53. Yelles-Chaouche, A.; Kherroubi, A.; Beldjoudi, H. The large Algerian earthquakes (267 A.D.-2017). *Fís. Tierra* **2017**, *29*, 159–182. [CrossRef]

54. Deschamps, A.; Gaudemer, Y.; Cisternas, A. The El Asnam, Algeria, earthquake of 10 October 1980: Multiple-source mechanism determined from long-period records. *Bull. Seismol. Soc. Am.* **1982**, *72*, 1111–1128. [CrossRef]

55. Meghraoui, M.; Philip, H.; Albarede, F.; Cisternas, A. Trench investigations through the trace of the 1980 El Asnam thrust fault: Evidence for paleoseismicity. *Bull. Seismol. Soc. Am.* **1988**, *78*, 979–999. [CrossRef]

56. Bezzeghoud, M.; Dimitrov, D.; Ruegg, J.C.; Lammali, K. Faulting mechanism of the El Asnamn (Algeria) 1954 and 1980 earthquakes from modelling of vertical movements. *Tectonophysics* **1995**, *249*, 249–266. [CrossRef]

57. Harbi, A.; Maouche, S.; Vaccari, F.; Aoudia, A.; Oussadou, F.; Panza, G.F.; Benouar, D. Seismicity, seismic input and site effects in the Sahel-Algiers region (North Algeria). *Soil Dyn. Earthq. Eng.* **2007**, *27*, 427–447. [CrossRef]

58. Boulahia, O.; Abacha, I.; Yelles-Chaouche, A.; Bendjama, H.; Kherroubi, A.; Mohammedi, Y.; Aidi, C.; Chami, A. Recent Seismic Activity in the Bejaia-Babors Region (Northeastern Algeria): The case of the 2012–2013 Bejaia earthquake sequences. *Pure Appl. Geophys.* **2021**, *178*, 1253–1279. [CrossRef]

59. Boulahia, O.; Abacha, I.; Yelles-Chaouche, A.; Bendjama, H.; Kherroubi, A. Analysis of the Bejaia Seismic Sequence of 2012–2013, Northeastern, Algeria. In *Advances in Geophysics, Tectonics and Petroleum Geosciences, Proceedings of the 2nd Springer Conference of the Arabian Journal of Geosciences (CAJG 2019), Sousse, Tunisia, 25–28 November 2019*; Meghraoui, M., Ed.; Springer: Cham, Switzerland, 2022; pp. 135–139. [CrossRef]

60. Bendjama, H.; Yelles-Chaouche, A.; Boulahia, O.; Abacha, I.; Mohammedi, Y.; Beldjoudi, H.; Rahmani, S.T.E.; Belheouane, O. The March 2017 earthquake sequence along the E-W trending Mcid Aïcha-Debbagh Fault, northeast Algeria. *Geosci. J.* **2021**, *25*, 697–713. [CrossRef]

61. Abacha, I.; Yelles-Chaouche, A.; Boulahia, O. Statistical Study of Earthquake Swarms in North-eastern Algeria with Special Reference to the Ain Azel Swarm; Hodna Chain, 2015. In *Advances in Geophysics, Tectonics and Petroleum Geosciences, Proceedings of the 2nd Springer Conference of the Arabian Journal of Geosciences (CAJG 2019), Sousse, Tunisia, 25–28 November 2019*; Meghraoui, M., Ed.; Springer: Cham, Switzerland, 2022; pp. 145–148. [CrossRef]

62. Yelles-Chaouche, A.; Abacha, I.; Boulahia, O.; Aidi, C.; Chami, A.; Belheouane, A.; Rahmani, S.T.E.; Roubeche, K. The 13 July 2019 Mw 5.0 Jijel Earthquake, northern Algeria: An indicator of active deformation along the eastern Algerian margin. *J. Afr. Earth Sci.* **2021**, *177*, 104149. [CrossRef]

63. Frohlich, C. Triangle diagrams: Ternary graphs to display similarity and diversity of earthquake focal mechanisms. *Phys. Earth Planet. Inter.* **1992**, *75*, 193–198. [CrossRef]

64. Zoback, M.L. First and second order patterns of stress in the lithosphere: The world stress map project. *J. Geophys. Res.* **1992**, *97*, 11703–11728. [CrossRef]

65. Henares, J.; López Casado, C.; Sanz de Galdeano, C.; Delgado, J.; Peláez, J.A. Stress fields in the Ibero-Maghrebian region. *J. Seismol.* **2003**, *7*, 65–78. [CrossRef]

66. Cornell, A. Engineering seismic risk analysis. *Bull. Seismol. Soc. Am.* **1968**, *58*, 1583–1606. [CrossRef]

67. Esteva, L. Seismicity prediction: A Bayesian approach. In Proceedings of the Fourth World Conference on Earthquake Engineering, Santiago, Chile, 13–18 January 1969; Volume 1, pp. 172–184.

68. Esteva, L. Seismic risk and seismic design decisions. In *Seismic Design for Nuclear Power Plants*; Hansen, R.J., Ed.; MIT Press: Cambridge, MA, USA, 1970; pp. 142–182.

69. McGuire, R.K. *FORTRAN Computer Program for Seismic Risk Analysis*; Open-File Report 76-67; U.S. Geological Survey: Reston, VA, USA, 1976.

70. Gardner, J.K.; Knopoff, L. Is the sequence of earthquakes in Southern California, with afershocks removed, Poissonian? *Bull. Seismol. Soc. Am.* **1974**, *64*, 1363–1367. [CrossRef]

71. EPRI-Electric Power Research Institute. *Seismic Hazard Methodology for the Central and Eastern United States*; EPRI Report NP-4726; U.S. Department of Energy Office of Scientific and Technical Information: Oak Ridge, TN, USA, 1986.

72. Peláez, J.A.; Hamdache, M.; López Casado, C. Seismic Hazard in Northern Algeria Using Spatially-smoothed Seismicity. Results for Peak Ground Acceleration. *Tectonophysics* **2003**, *372*, 105–119. [CrossRef]

73. Box, G.E.P.; Pierce, D.A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *J. Am. Stat. Assoc.* **1970**, *65*, 1509–1526. [CrossRef]

74. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*, 2nd ed.; OTexts: Melbourne, Australia, 2018.

75. Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*; Holden-Day: San Francisco, CA, USA, 1976.

76. Lai, Y.; Dzombak, D.A. Use of the Autoregressive Integrated Moving Average (ARIMA) Model to Forecast Near-Term Regional Temperature and Precipitation. *Weather Forecast.* **2020**, *35*, 959–976. [CrossRef]

77. Cowpertwait, P.S.P.; Metcalfe, A.V. *Introductory Time Series with R*; Springer: New York, NY, USA, 2009. [CrossRef]

78. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]

79. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]

80. Fabozzi, F.J.; Focardi, S.M.; Rachev, S.T.; Arshanapalli, B.G.; Markus, H. *The Basics of Financial Econometrics. Tools, Concepts, and Asset Management Applications*; John Wiley and Sons, Inc.: Hoboken, NJ, USA, 2014.

81. Kullback, S.; Leibler, R.A. On Information and Sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]

82. Hurvich, C.M.; Tsai, C.L. Regression and Time Series Model Selection in Small Samples. *Biometrika* **1989**, *76*, 297–307. [CrossRef]

83. Kumaresan, R.; Tufts, D.W. Data-adaptive principal component signal processing. In Proceedings of the 1980 19th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes, Albuquerque, NM, USA, 10–12 December 1980; pp. 949–954.

84. Broomhead, D.S.; King, G.P. Extracting qualitative dynamics from experimental data. *Phys. D Nonlinear Phenom.* **1986**, *20*, 217–236. [CrossRef]

85. Fraedrich, K. Estimating the dimensions of weather and climate attractors. *J. Atmos. Sci.* **1986**, *43*, 419–432. [CrossRef]

86. Vautard, R.; Ghil, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Phys. D Nonlinear Phenom.* **1989**, *35*, 395–424. [CrossRef]

87. Ghil, M.; Vautard, R. Interdecadal oscillations and the warming trend in global temperature time series. *Nature* **1991**, *350*, 324–327. [CrossRef]

88. Yiou, P.; Baert, E.; Loutre, M.F. Spectral analysis of climate data. *Surv. Geophys.* **1996**, *17*, 619–663. [CrossRef]

89. Pereira de Albuquerque, W.C.; Maciel, C.D. Performance of ultrasound echo decomposition using singular spectrum analysis. *Ultrasound Med. Biol.* **2001**, *27*, 1231–1238. [CrossRef] [PubMed]

90. Moskvina, V.; Zhigljavsky, A. An algorithm based on singular spectrum analysis for change-point detection. *Commun. Stat.-Simul. Comput.* **2003**, *32*, 319–352. [CrossRef]

91. Sivapragasam, C.; Liong, S.Y.; Pasha, M.F.K. Rainfall and runoff forecasting with SSA-SVM approach. *J. Hydroinform.* **2001**, *3*, 141–152. [CrossRef]

92. Deng, C. Time Series Decomposition Using Singular Spectrum Analysis. Master's Thesis, East Tennessee State University, Johnson City, TN, USA, 2014.

93. Golyandina, N.; Zhigljavsky, A. *Singular Spectrum Analysis for Time Series*; Springer: Berlin/Heidelberg, Germany, 2013.

94. Golyandina, N.; Nekrutkin, V.; Zhigljavsky, A. *Analysis of Time Series Structure: SSA and Related Techniques*, 1st ed.; Chapman & Hall/CRC: New York, NY, USA, 2001.

95. Gao, W.; Guo, J.; Zhou, M.Y.H.; Chen, X.; Ji, B. Gravity tides extracted from SSA-denoised superconducting gravity data with the harmonic analysis: A case study at Wuhan station, China. *Acta Geod. Geophys.* **2020**, *55*, 609–625. [CrossRef]

96. Golyandina, N.; Korobeynikov, A.; Zhigljavsky, A. *Singular Spectrum Analysis with R*; Springer: Berlin/Heidelberg, Germany, 2018. [CrossRef]

97. Goulet, V. Introduction à la Programmation en R, 5th ed.École D'actuariat, Université Laval: Québec, QC, Canada, 2016.

98. Korobeynikov, A. Computation- and space-efficient implementation of SSA. *Stat. Its Interface* **2010**, *3*, 357–368. [CrossRef]

99. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications*; Springer: Cham, Switzerland, 2017. [CrossRef]

100. Elliott, G.; Stock, J.H.; Rothenberg, T.J. Efficient Tests for an Autoregressive Unit Root. *Econometrica* **1996**, *64*, 813–836. [CrossRef]

101. Shapiro, S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]

102. Baddeley, A.; Turner, R.; Moller, J.; Hazelton, M. Residual analysis for spatial point processes. *J. R. Stat. Soc. Ser. B* **2005**, *67*, 617–666. [CrossRef]
103. Andersen, A.J.; Dennison, J. An Introduction to Quantile-Quantile Plots for the Experimental Physicist. *Am. J. Phys.* **2019**, *87*, 1–4.
104. Hassani, H. Singular spectrum analysis: Methodology and comparison. *J. Data Sci.* **2007**, *5*, 239–257. [CrossRef]
105. Console, R. Testing earthquake forecast hypotheses. *Tectonophysics* **2001**, *338*, 261–268. [CrossRef]
106. Shcherbakov, R. Statistics and forecasting of aftershocks during the 2019 Ridgecrest, California, earthquake sequence. *J. Geophys. Res. Solid Earth* **2021**, *126*, e2020JB020887. [CrossRef]
107. Taroni, M.; Marzocchi, W.; Schorlemmer, D.; Werner, M.J.; Wiemer, S.; Zechar, J.D.; Heiniger, L.; Euchner, F. Prospective CSEP Evaluation of 1-Day, 3-Month, and 5-Yr Earthquake Forecasts for Italy. *Seismol. Res. Lett.* **2018**, *89*, 1251–1261. [CrossRef]