Time Series Analysis for Predicting the Occurrences of Large Scale Earthquakes

Amei Amei*

Wandong Fu**

Chih-Hsiang Ho***

Abstract

Earthquakes that have occurred worldwide during the period of 1896 to 2009 with magnitude of 8.0 or greater on the Richter scale are assumed to follow a Poisson process. Several autoregressive integrated moving average (ARIMA) models with different time steps are proposed to predict the occurrences of large scale earthquakes by fitting the model with a sequence of empirical recurrence rates (ERRs) time series. The last five or ten data points are used as prediction sets to check the predictive ability of the candidate models developed by the time series modeling techniques. For a full scale forecast, the best fitted model predicts a total number of 12 large scale earthquakes in the next 6 years worldwide. The application of ERR based ARIMA models to long-term earthquake prediction not only serves as a linking bridge between point processes and the classical time series but also extends the usage of statistical methods to a wide area of natural disaster predictions.

Introduction

On January 12, 2010, a 7.0 magnitude earthquake hit Port-au-Prince, Haiti. The earthquake lasted one minute, just enough time to kill thousands of people and destroy numerous buildings. According to the government's estimate, 200,000 people were killed, 250,000 were injured, and consequently, 1.5 million people became homeless. On February 27, 2010, a magnitude of 8.8 earthquake occurred off the coast of the Maule Region of Chile, which lasted 90 seconds. Six Chilean cities experienced intense vibrations. The earthquake triggered a tsunami which devastated several coastal towns in south-central Chile and damaged the port at Talcahuano. Large earthquake, as is pointed out in Mogi (1985), occurs unexpectedly and sometimes inflicts enormous damage, and earthquake prediction is not only an extremely fascinating topic in seismology but also its ultimate goal. In recent years, tremendous progress has been made toward this goal in a wide range of research area of earthquake prediction and hazard assessment (e.g., Bakun et al. 2005, Felzer et al. 2003, Helmstetter et al. 2006, Hong and Guo 1995, Jackson and Kagan 2006, Kagan 1993, Savage and Cockerham 1987, and references therein).

In this study, we use the earthquake data worldwide from 1896 to 2009 with magnitude greater than or equal to 8.0 on the Richter scale and assume that they follow a Poisson process. We then construct a discrete time series based on the empirical recurrent rates (ERRs) of the assumed Poisson process, computed sequentially at equidistant time intervals during the observation period. The time-plot of the ERRs, referred to as the "fingerprint" or the ERR plot, offers the possibility of further insight into the data and provides a technical basis for model developments for the earthquake data. The three main objectives of this study are: (i) convert point process to ERR time series, (ii) fit the time series data into the ARIMA model (to be defined later), and (iii) develop methods to retrieve the counterparts of the predicted ERRs.

Theory and Method

Empirical Recurrence Rate (ERR)

Let $t_1,...,t_N$ be the time of the *N*-ordered earthquakes during an observation period of $(t_0,0)$. If t_0 is adopted as the time origin and h as time step then $\{z_\ell\}$, a discrete time series generated sequentially at equidistant time intervals $t_o+h,\ t_0+2h,\ \ldots,\ t_0+Nh$ (= 0 = present time), can be regarded as observations at times $t_l = t_0 + lh$, ℓ=1, 2, …, N for the earthquakes to be modeled.

A key parameter in the modeling is the recurrence rate of the targeted earthquake data. The time series based on the ERRs (Ho 2008) is generated as follows:

$$z_l = \frac{n_l}{lh} = \frac{\text{total number of earthquakes in } (t_0, t_0 + lh)}{lh}, \quad l = 1, 2, ..., N.$$

Note that z_ℓ evolves over time and is simply the maximum likelihood estimator (MLE) of the mean recurrence rate, if the underlying process observed in $(t_0,\ t_l)$ is a homogeneous Poisson process. The time-plot of the ERRs, referred to as the "ERR plot", offers the possibility of further insights into the data. Specifically, if we start at time T, the value z_{T+k}, $k \geq 1$ can be predicted based on the sample observations $(z_1, \ldots, z_T)$ of an ERR time series. In the traditional regression modeling, the observations are assumed to be independent and this is not a reasonable assumption for a process that evolves over time. Therefore the autoregressive integrated moving average (ARIMA) models are introduced.

Arima Model

Autoregressive integrated moving average (ARIMA) models, introduced by Box and Jenkins (1976), are mathematical models of persistence, or autocorrelation, in a time series. They can be expressed by a series of equations. One subset of ARIMA models is called autoregressive, or AR models. The name autoregressive refers to the regression on self. An AR model describes a time series as a linear function of its past values plus a noise term ε_t. The order of the AR model shows the number of past values involved. The simplest AR model is the first-order autoregressive, or AR (1) model. The equation for this model is given by:

$$X_t = \phi X_{t-1} + \varepsilon_t, \quad t = 1, 2, ..., N,$$

where X_t is a stationary mean zero time series and ϕ is the first-order autoregressive coefficient. We can see that the AR (1) model has the form of a regression model in which X_t is regressed on its previous value, and the error term ε_t is analogous to the regression residuals and represents a "white noise" (uncorrelated with mean 0 and variance σ^2) process.

The moving average (MA) model is another form of ARIMA model in which the time series is described as a linear function of its prior errors plus a current error ε_t. The first-order moving average, or MA (1), model is given by:

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1}, \quad t = 1, 2, ..., N,$$

where X_t is a stationary mean zero time series, ε_t, ε_{t-1} are the error terms at time t and $t-1$, and θ is the first-order moving average coefficient.

A general autoregressive moving average (ARMA) model, ARMA (p, q), is given by:

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = \varepsilon_t + \theta_1\varepsilon_{t-1} + ... + \theta_q\varepsilon_{t-q}.$$

The integrated ARMA (ARIMA) model is a broadening of the class of ARMA that includes differencing.

Differencing is an important technique in data transformation, which attempts to de-trend to control autocorrelation and achieve stationary time series. The first order differencing is defined by:

$$\nabla X_t = X_t - X_{t-1} = (1-B)X_t,$$

where B is the backshift operator. Consequently, the differencing of order d is denoted by:

$$\nabla^d X_t = (1-B)^d X_t.$$

Usually, single differencing is used to remove linear trend and double differencing is used to remove quadratic trend. We can eliminate seasonality and trend of period d by introducing the lag-d differencing operator ∇_d:

$$\nabla_d X_t = X_t - X_{t-d} = (1-B^d)X_t.$$

In general, ARIMA modeling involves three stages. The first stage is to identify a model by specifying the type of the model (AR, MA, ARMA, or ARIMA) and its order. Sometimes identification is done by looking at plots of the sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF). Sometimes it is done by an auto fit procedure – fitting many different possible model structures and orders and using a goodness-of-fit statistic to select the best model. The second stage is to estimate the coefficients of the model by minimizing the sum of squared residuals. The final stage is model diagnostics. At this stage, it is crucial to check that the residuals of the candidate model are random and normally distributed and the estimated parameters are statistically significant. The fitting process is usually guided by the principle of parsimony, by which the best model is the one which has fewest parameters among all models that fit the data.

Transformation and Data Splitting

If the data set is large enough, it can be split into two sets: training sample and prediction set. Training sample is used to develop a model for prediction. Prediction set is used to evaluate the reasonableness and predictive ability of the selected model. This validation procedure named cross-validation is the statistical practice of splitting a sample of data into two subsets so that the analysis is initially performed on one subset and the other subset is retained for subsequent use in confirming and validating the initial analysis.

In proving a fitted ARMA model meaningful, it must be at least plausible that the data are in fact a realization of an ARMA process and in particular a realization of a stationary process. A stationary time series is the one whose statistical properties such as mean, variance, autocorrelation, etc. are all constant over time. In order to obtain such stationary time series, we use a sequence of mathematical transformations, such as Box-Cox transformation, mean subtraction, and the differencing.

For a sequence of observations $Y_1,\ Y_2,\ldots,Y_n$, the Box-Cox transformation f_λ is given by

$$f_\lambda(y) = \begin{cases} \dfrac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log(y), & \lambda = 0 \end{cases}$$

This transformation is useful when the variability of the data increases or decreases with the level. By suitable choice of λ, the variability can be made nearly constant. For instance, the variability of a set of positive data whose standard deviation increases linearly can be stabilized by choosing $\lambda = 0$ (Brockwell et al. 2002).

Normally, the correct amount of differencing is the lowest order of differencing that yields a time series which fluctuates around a well-defined mean value and whose autocorrelation function (ACF) plot decays rapidly to zero, either from above or below. Thus, at every stage of differencing, we check the plots of sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) to see where the ACF/PACF "cuts off" the bounds $\pm 1.96/\sqrt{n}$. It is desirable to find a sample ACF that decays fairly rapidly. We say that a series is stationary if the sample ACF has very few significant spikes at very small lags and then cuts off drastically or dies down very quickly. If the sample ACF dies slowly, the series still has some trend. If ACF has periodicity, the series has seasonality and we should do some more differencing for the data.

Model Diagnostics

We check the residual ACF/PACF of the models and the randomness of the residuals. For large n, the sample autocorrelations of an independent and identically distributed (*iid*) sequence $Y_1,\ldots,Y_n$ with zero mean and finite variance are approximately *iid* with normal distribution $N(0,1/n)$. We test whether or not the observed residuals are consistent with *iid* noise by examining the sample correlations of the residuals and rejecting the *iid* noise hypothesis if more than two or three out of 40 fall outside the bounds $\pm 1.96/\sqrt{n}$ or if one falls far outside the bounds (Brockwell et al. 2002).

Ljung-Box test, proposed by Ljung and Box (1978), is commonly used to check whether the residuals of a fitted model are *iid* in ARIMA modeling. It is based on the autocorrelation plot and it tests the overall independence based on a few of the time lags. Formally, the definition of Ljung-Box test is as follows.

H_0 : The sequence data are *iid*

H_a : The sequence data are not *iid*

The test statistic is $\hat{Q}(\hat{r}) = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\hat{r}_k^2$, where $\hat{r}_k = \sum_{l=k+1}^{n}\hat{a}_l\hat{a}_{l-k} \Big/ \sum_{l=1}^{n}\hat{a}_l^2$, the estimated autocorrelation at lag-k, n = sample size, m = number of lags being tested and $\hat{a}_1,...,\hat{a}_n$ are the residuals after a model has been fitted to a series $z_1,...,z_n$. If no model is being fitted, $\hat{a}_1,...,\hat{a}_n$ are the "mean corrected" series of $z_1,...,z_n$.

For large n, the distribution of $\hat{Q}(\hat{r})$ is approximately χ^2_{m-p-q} under the null hypothesis, where $p+q$ is the number of parameters of the fitted model. The hypothesis of *iid* is rejected if $\hat{Q} > \chi^2_{1-\alpha;m-p-q}$ at level α and we say that the sequence data do have autocorrelations significantly different from zero and a new search for a fitted ARMA model for a mean-corrected data set will follow.

Model Comparison

We use *AICC* statistic (Akaike 1974), the bias-corrected version of *AIC* statistic, as an information criterion to select candidate models using the ITSM2000 package (Brockwell et al. 2002). Small value of *AICC* is indication of a good model, but it should be used only as rough guide. Final decisions between models are based on maximum likelihood estimation. Some other model-selection statistics, such as *BIC* statistic, are also available in ITSM2000. The *BIC* statistic (Schwarz 1978) is a Bayesian modification of the *AIC* statistic. It is evaluated at the same time and used in the same way as the *AICC*. Each information statistic is defined as:

$$AIC_{p,q} = N\log\hat{\sigma}_\varepsilon^2 + 2r$$

$$AICC_{p,q} = N\log\hat{\sigma}_\varepsilon^2 + 2rN/(N-r-1)$$

$$BIC_{p,q} = N\log\hat{\sigma}_\varepsilon^2 + r\log N$$

where $\hat{\sigma}_\varepsilon^2$ is the maximum likelihood estimator of σ_ε^2, and $r = p+q+1$ is the number of parameters estimated in the model, including a constant term. The second term in all three equations is a penalty for increasing *r*. Hence, the best model is the model adequately describes data and has fewest parameters.

The candidate ARIMA models are used to predict future values of the time series from the past values. The predicting function $z_t = f(z_{t-1},...,z_1) + \varepsilon_t$ has minimum mean square error. The first part of the equation $f(z_{t-1,},...,z_1)$ is a function of the past values of the series and the second part ε_t, called noise part, is a sequence of *iid* variables. Predictions of the original series are achieved by forecasting the residuals and then feeding them back into the inverted transformations. The best fitting model is selected based on the predictive ability of both the training sample and the prediction set. Finally, we combine the training sample and the prediction set as a full data set to forecast earthquakes for future occurrences. Note that the cumulated mean numbers inverted from the forecasted ERRs should be non-decreasing and sometimes need to be adjusted accordingly (e.g., Ho 2010).

Application

Data

Earthquakes that have occurred worldwide during the period of 1896 to 2009 with Magnitude $\mathrm{M} \geq 8.0$ on the Richter scale are obtained from the *U.S. Geological Survey* (http://www.usgs.gov/). The data set contains 55 earthquakes from the time origin $t_0 = 1896$ to the present time 0 (year 2009). (Table A1).

A dot plot is first constructed to observe any possible trends presented by the data (Figure 1). We then count the number of earthquakes at a pre-chosen time step and calculate the z_l values to do further analysis.

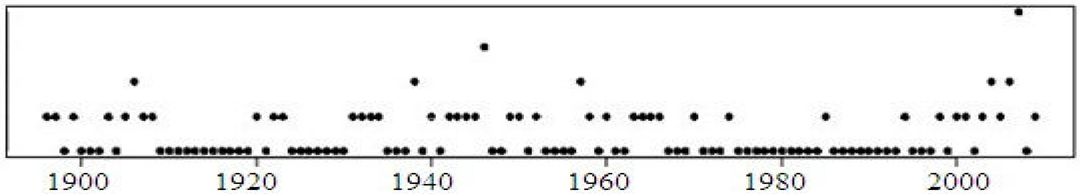

Figure 1. Dot plot of large earthquakes worldwide between 1896 and 2009.

Three ERR plots with three different time-steps are shown in Figure 2. Since there are 55 large earthquakes in 114 years, approximately one in every two years, we first choose $h = 2$ years as the time-step.

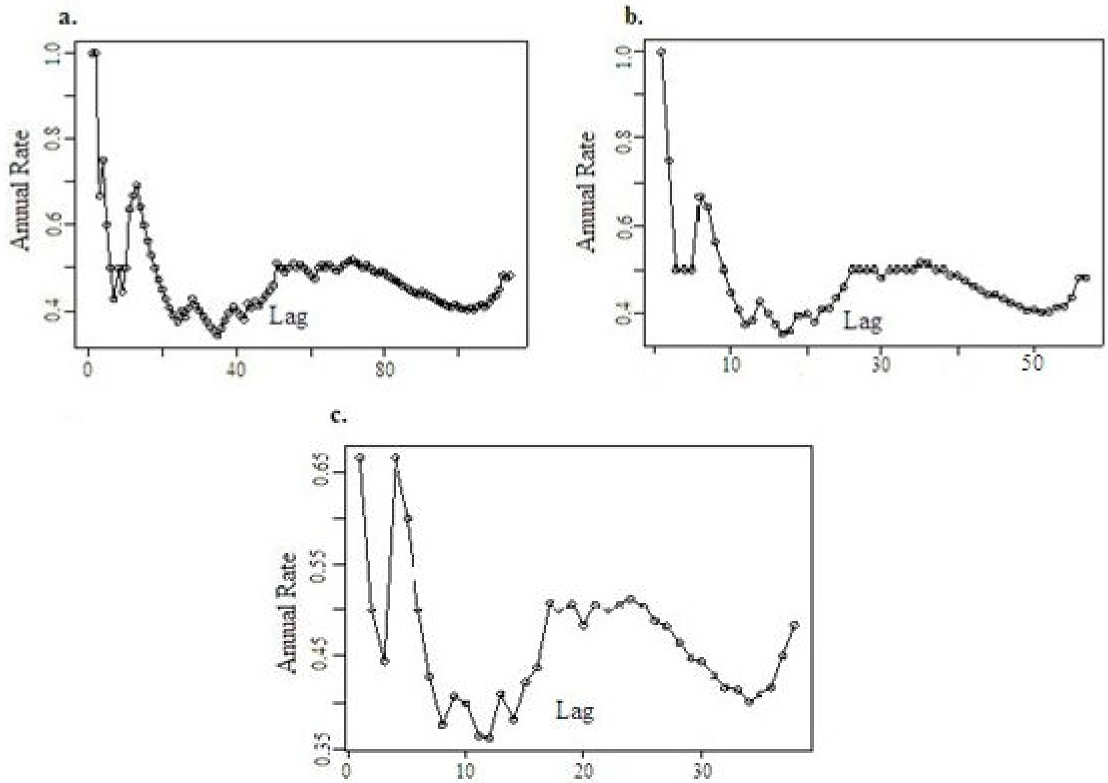

Figure 2. ERR plots with different time-steps h: (**a**) $h = 1$ year, (**b**) $h = 2$ years, (**c**) $h = 3$ years.

Arima Modeling with Time Step Set at Two Years

We use the ITSM2000 software to model the ERR data with time step $h = 2$ years. The data set, containing 57 lags in total, is split into two sets: training sample and prediction set. Specifically, the training sample is the original data set excluding the last 5 ERRs, which is the prediction set (Figure 3). The size of a prediction set is quite flexible as long as it fits a common goal of model selection. We focus on the training sample and plot the sample ACF and PACF to capture possible trend and seasonality of the data (Figure 4). From the plot of sample ACF, we found that the spikes die slowly and have periodicity. This indicates that the underlying process described by the data is not stationary and differencing is necessary.

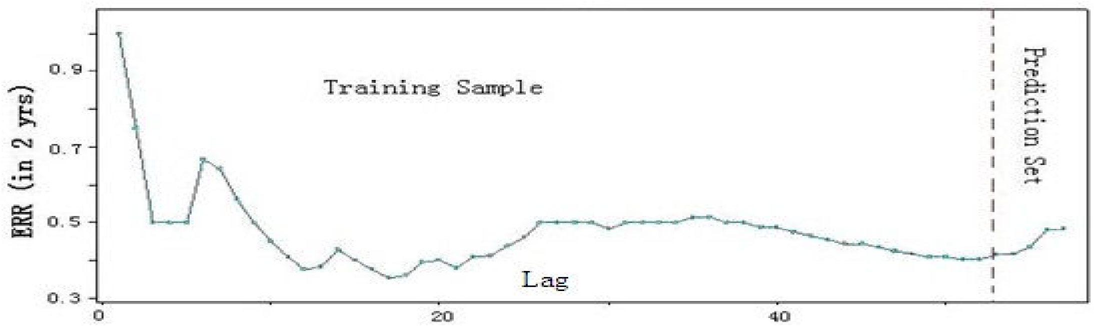

Figure 3. Training sample and prediction set with $h = 2$ years. Each lag corresponds to 2 years.

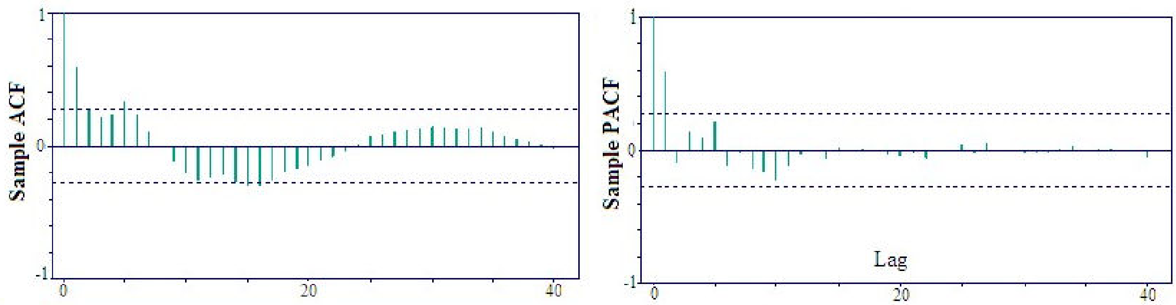

Figure 4. Sample ACF and sample PACF of the original training set with $h = 2$ years.

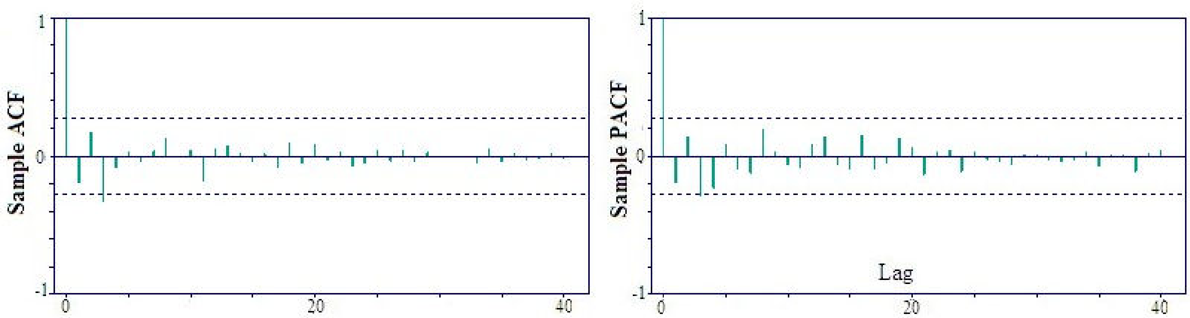

Figure 5. Sample ACF and sample PACF of the twice-differenced training sample with $h = 2$ years.

We find that it is sufficient to produce a series with no apparent trend by applying the differencing operator ∇, twice at lag-1, to the training sample (Figure 5). We then subtract the sample mean from each observation of the twice-differenced series to generate a "mean-corrected" series. The sample ACF and PACF plots as well as the *AICC* statistics for several ARMA(p,q) with $0 \le p, q \le 5$ suggest that model MA(3) is appropriate for the mean-corrected and twice-differenced data X_t (the ACF is cutting off at lag 3 and the PACF is tailing off). The estimated (MLE) model is:

$$X_t = \varepsilon_t - 0.2475\varepsilon_{t-1} + 0.1471\varepsilon_{t-2} - 0.4985\varepsilon_{t-3}$$

with $\hat{\sigma}^2 = 0.00224$. Recall that $\hat{\sigma}^2$ is the MLE of σ^2, the variance of the white noise process ε_t, and the unit of the time series is presented as an annual rate. A set of diagnostic plots (Figure 6) for the MA(3) model is produced by the ITSM2000 package, consisting of the plot of the residuals and its ACF and PACF. The *AICC* statistic is -153.367 and the Ljung-Box test is not significant ($\text{p-value} = 0.96067$), indicating that the residuals are approximately white noise. The numerical values of the actual ERRs and mean numbers of occurrences in the prediction set are compared with their predicted counterparts and are listed in Table 1.

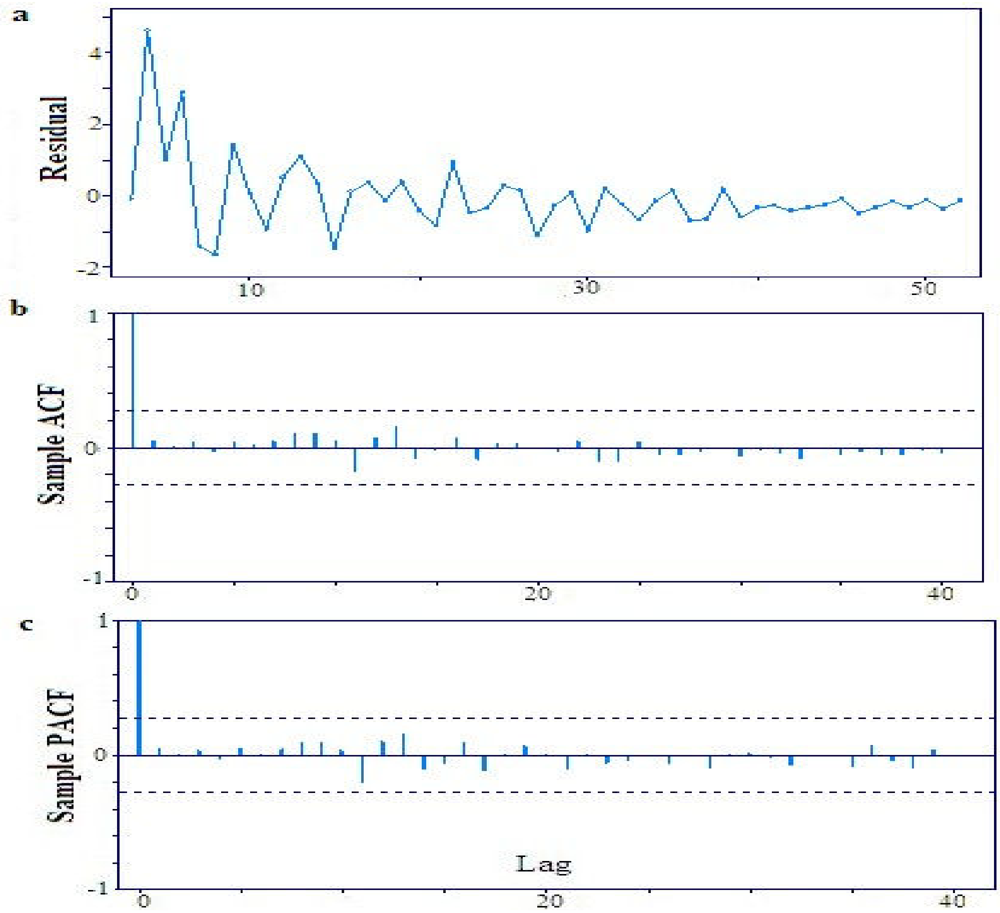

Figure 6. (**a**) time-plot, (**b**) sample ACF, and (**c**) sample PACF for the residuals of the MA (3) model fitted to the mean-corrected and twice-differenced training sample.

Table 1. The numerical values of the actual ERRs and mean numbers versus the predicted ERRs and mean numbers based on the proposed MA (3) model.

Year	Annual ERR		Mean number	
	Actual	Prediction	Actual	Prediction
2000-2001	0.415094	0.41238	2	1.71228
2002-2003	0.416667	0.43368	1	2.83744
2004-2005	0.436364	0.46318	3	5.9498
2006-2007	0.482143	0.49771	6	7.74352
2008-2009	0.482456	0.53728	1	7.24992

For a causal model (Brockwell et al. 2002), the ITSM2000 package calculates the ratios of $(\text{estimated coefficients})/(1.96 \times \text{standard error})$. If the ratio is greater than 1 in absolute value, we conclude (at level 0.05) that the corresponding coefficient in the model may not be zero (Brockwell et al. 2002) and a subset model comes up after dropping the non-significant coefficients. In our case, the ratios for the proposed MA(3) model are $-0.892166,\ 0.444319,$ and -2.217303 respectively and the subset MA(3) model is

$$X_t = \varepsilon_t - 0.5136\varepsilon_{t-3}$$

where $AICC = -155$ and $\text{p-value} = 0.828$ for the Ljung-Box test. The values predicted by the subset model are very similar to those by the original MA(3) model and we choose the original MA(3) model for full data forecasting.

Full-Data Forcasting

We re-estimate the coefficients of the MA(3) model by using the full ERR time series to forecast the number of earthquakes in the future. This yields the best-fitted MA (3) model for the mean-corrected and twice-differenced (at lag 1) data (same as before). The estimated (MLE) model is:

$$X_t = \varepsilon_t - 0.2708\varepsilon_{t-1} + 0.1450\varepsilon_{t-2} - 0.5025\varepsilon_{t-3}$$

with the estimated white noise variance $\hat{\sigma}^2 = 0.002100$. The ratios, defined in last section, for the fitted model are -1.030746, 0.448350 and -2.236312, which leads to a subset MA (3) model:

$$X_t = \varepsilon_t - 0.2213\varepsilon_{t-1} - 0.4597\varepsilon_{t-3}, \quad \text{with} \quad \hat{\sigma}^2 = 0.002137$$

Even though the *AICC* statistic and the p-value of the Ljung-Box test of the subset MA(3) model ($AICC = -174.738$, p-value $= 0.97676$) are a little bit better than MA(3) model

($AICC =$ -173.294, p-value $=$ 0.96568), there is no big difference and we will keep both of them. The predictions of the next ten years, from 2010 to 2019, are shown in Table 2.

Table 2. The predicted ERRs and mean numbers using the MA(3) and the subset MA(3).

Year	Full model ERR		Mean number	
	MA (3)	Subset MA (3)	MA (3)	Subset MA (3)
2010-2011	0.50365	0.49733	3.4234	2.69028
2012-2013	0.50785	0.50620	1.5029	2.04132
2014-2015	0.54365	0.54526	5.3117	5.69960
2016-2017	0.58401	0.58886	6.01122	6.40972
2018-2019	0.62891	0.63702	6.73562	7.14956

The predicted mean numbers of occurrences in Table 1 are roughly close to the actual numbers except for the last two year period for which the model predicted 7 large scale earthquakes while there was only one happened during that time. One possible reason could be due to the use of the two-year time step, which reduced the sample size from 114 to 57 and, in general, models generated on larger data sets have better performance. This suggests us to find a model by adjusting the time step to one year.

Arima Modeling with Time Step Set at One Year

The data set, after choosing the time step $h = 1$ year, contains 114 lags in total and is split into the training sample with 104 lags and the prediction set with 10 lags (Figure 7). The plots of sample ACF and PACF on the training sample (Figure 8) indicate that the series is not stationary and differencing is needed. Upon applying the same approach as we did for the case of two-year time step to the mean-corrected and twice-differenced (at lag 4 and 1 respectively) training data (contains 104 ERRs) we find the best fitted model ARMA(5, 5). Specifically, the estimated (MLE) model is:

$$\begin{aligned} X_t = &-0.1717X_{t-1} + 0.1597X_{t-2} + 0.5383X_{t-3} - 0.08758X_{t-4} - 0.4073X_{t-5} + \varepsilon_t \\ &+0.2547\varepsilon_{t-1} + 0.4622\varepsilon_{t-2} - 0.2237\varepsilon_{t-3} - 0.2444\varepsilon_{t-4} - 0.6749\varepsilon_{t-5} \end{aligned}$$

with $\hat{\sigma}^2 = 0.00108$, $AICC =$ 362.268, and p-value $= 0.86584$.

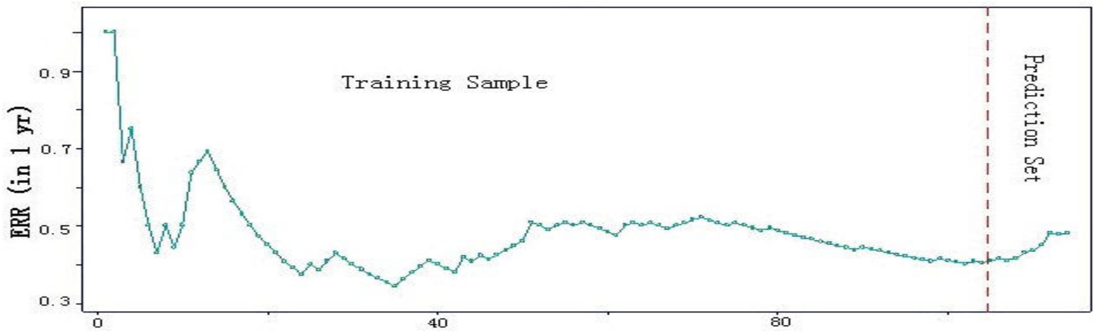

Figure 7. Training sample and prediction set of data set with $h=1$, each lag corresponds to 1 year.

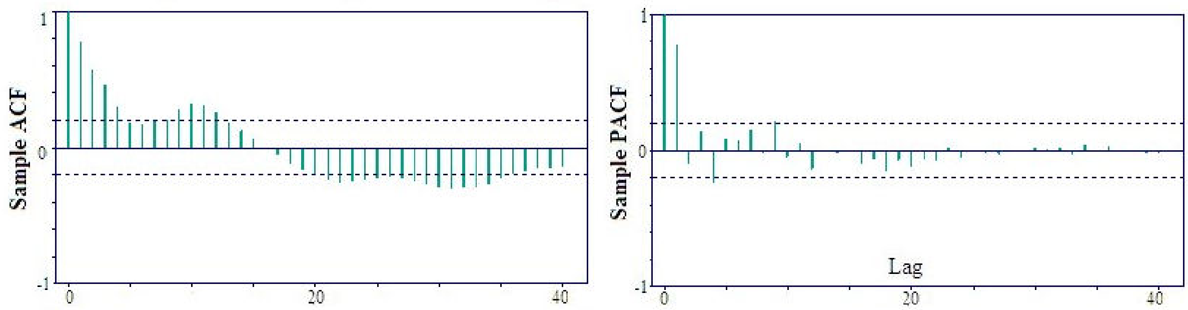

Figure 8. Sample ACF and sample PACF of the training sample with $h=1$ year.

Moreover, the best fitted model by using the complete ERR time series (contains 114 ERRs) for the mean-corrected and twice-differenced data is also an ARMA(5,5) and the estimated (MLE) model is:

$$X_t = -0.1766X_{t-1} + 0.1507X_{t-2} + 0.5407X_{t-3} - 0.08186X_{t-4} - 0.4143X_{t-5} + \varepsilon_t$$
$$+0.2635\varepsilon_{t-1} + 0.4700\varepsilon_{t-2} - 0.2179\varepsilon_{t-3} - 0.2439\varepsilon_{t-4} - 0.6681\varepsilon_{t-5}$$

with $\hat{\sigma}^2 = 0.00099$. Unfortunately, all of the subset ARMA(5, 5) models neither pass model diagnostic tests nor outperform the ARMA(5, 5). The forecasting results for the model based on the training sample are summarized in Table 3. It turns out that, by comparing the results in Table 3 with those in Table 1, the predictive ability of the model improved significantly when the time step h equals 1 year. Finally, the predictions of large scale earthquakes for the next ten years are shown in Table 4.

Table 3. The numerical values of the actual ERRs and mean numbers versus the predicted ERRs and mean numbers based on the proposed ARMA (5, 5).

Year	Annual ERR		Mean number	
	Actual	Prediction	Actual	Prediction
2000	0.40952381	0.41519	1	1.59495
2001	0.41509434	0.4218	1	1.11585
2002	0.41121495	0.43678	0	2.02466
2003	0.41666667	0.44336	1	1.14742
2004	0.43119266	0.46441	2	2.73781
2005	0.43636364	0.47262	1	1.36751
2006	0.45045045	0.49327	2	2.76477
2007	0.48214286	0.50355	4	1.64463
2008	0.47787611	0.52445	0	2.86525
2009	0.48245614	0.53609	1	1.85141

Table 4. Predictions of large earthquakes using the full data forecasting model ARMA(5,5).

Year	Full model ERR	Mean number
2010	0.4976	2.224
2011	0.50603	1.47548
2012	0.50149	0 (adjusted)
2013	0.53084	3.96479
2014	0.54151	1.80057
2015	0.5598	2.73631
2016	0.57991	2.99311
2017	0.60608	3.77265
2018	0.62047	2.37605
2019	0.65601	5.02743

Conclusions

A method of predicting future occurrences of large scale earthquakes by using time series modeling technique is presented in this study. A total of 114 documented earthquakes with magnitude greater than or equal to 8.0 on the Richter scale occurred during 1896 to 2009. We assume these occurrences follow a Poisson process and fingerprint them with a sequence of ERR time series. The data set is split into a training sample and a prediction set for which the prediction set is used as a holdout sample to check the predictive ability of the candidate model developed by the training sample. Model validation results are successful for the proposed method. The best fitted model for time step $h = 2$ years is a subset MA(3) model and for $h = 1$ year is an ARMA(5,5). For a full scaled prediction, the MA(3) model predicts a total of about 10 occurrences for the next six years and 12 for the ARMA(5,5) model. Our two candidate models, MA(3) and ARMA(5,5), forecast a mean number of about 24 and 19 occurrences for the prediction set while the actual number of events is 13. Therefore, we conclude that the ARMA(5,5) model is the best fitted model overall. The application of ARIMA models for long-term earthquake prediction is a natural extension of the methodologies developed for the volcanic risk assessment studies (Ho 2008, 2010). An important observation is that the modeling technique presented in this study will not only further facilitate the research in the areas of natural disaster predictions such as prediction of occurrences of dust storms and hurricanes but can also be extended to monitor the occurrence rates of cancer, genetic mutation, teen pregnancy, etc.

Although our model can predict future occurrences of large scale earthquakes, further studies should be carried out to improve model accuracy by quantifying the error bounds/confidence limits on the forecast.

Appendix

Table A1. Large earthquakes worldwide since 1896 ($M \geq 8.0$)

Date	Location	Magnitude
06/15/1896	Sanriku, Japan	8.5
06/12/1897	Assam, India	8.3
09/10/1899	Yakutat Bay, Alaska	8
08/11/1903	Southern Greece	8.3
07/09/1905	Mongolia	8.4
01/31/1906	Off the Coast of Esmeraldas, Ecuador	8.8
08/17/1906	Valparaiso, Chile	8.2
10/21/1907	Qaratog, Tajikistan	8
12/12/1908	Off the Coast of Central Peru	8.2
06/05/1920	Taiwan region	8
11/11/1922	Chile-Argentina Border	8.5
02/03/1923	Kamchatka	8.5
08/10/1931	Xinjiang, China	8
06/03/1932	Jalisco, Mexico	8.1
03/02/1933	Sanriku, Japan	8.4
01/15/1934	Bihar, India - Nepal	8.1
02/01/1938	Banda Sea, Indonesia	8.5
11/10/1938	Shumagin Islands, Alaska	8.2
05/24/1940	Callao, Peru	8.2
08/24/1942	Off the coast of central Peru	8.2
04/06/1943	Illapel - Salamanca, Chile	8.2
12/07/1944	Tonankai, Japan	8.1
11/27/1945	Makran Coast, Pakistan	8
04/01/1946	Unimak Island, Alaska	8.1
08/04/1946	Samana, Dominican Republic	8
12/20/1946	Nankaido, Japan	8.1
08/22/1949	Queen Charlotte Island, British Columbia,Canada	8.1
08/15/1950	Assam - Tibet	8.6
11/04/1952	Kamchatka	9
03/09/1957	Andreanof Islands, Alaska	8.6
12/04/1957	Gobi-Altay, Mongolia	8.1
11/06/1958	Kuril Islands	8.3
05/22/1960	Chile	9.5
10/13/1963	Kuril Islands	8.5
03/28/1964	Prince William Sound, Alaska	9.2
02/04/1965	Rat Island, Alaska	8.7
10/17/1966	Near the Coast of Peru	8.1
07/31/1970	Colombia	8
10/03/1974	Near the Coast of Central Peru	8.1
09/19/1985	Michoacan, Mexico	8
06/09/1994	Bolivia	8.2
03/25/1998	Balleny Islands Region	8.1
11/16/2000	New Ireland Region, Papua New Guinea	8
06/23/2001	Near the Coast of Peru	8.4
09/25/2003	Hokkaido, Japan Region	8.3
12/23/2004	North of Macquarie Island	8.1
12/26/2004	Sumatra-Andaman Islands	9.1
03/28/2005	Northern Sumatra, Indonesia	8.6
05/03/2006	Tonga	8
11/15/2006	Kuril Islands	8.3
01/13/2007	East of the Kuril Islands	8.1
04/01/2007	Solomon Islands	8.1
08/15/2007	Near the Coast of Central Peru	8
09/12/2007	Southern Sumatra, Indonesia	8.5
09/29/2009	Samoa Islands region	8.1

References

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans: Aut. Contr. AC* **19**, 203-217.

Bakun, W. H., Aagaard, B., and Dost, B., 2005. Implication for Prediction and Hazard Assessment From the 2004 Parkfield Earthquake. *Nature* **437**, 969-974.

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

Brockwell, P. J., and Davis, R. A., 2002. *Introduction to Time Series and Forecasting*. 2nd Edition. Springer-Verlag, New York.

Felzer, K. R., Abercrombie, R. E., and Ekstrom, G., 2003. Secondary Aftershocks and Their Importance for Aftershock Forecasting. *Bulletin of the Seismological Society of America* **93**, 1433-1448.

Helmstetter, A., Kagan, Y. Y., and Jackson D. D., 2006. Comparison of short-term and long-term earthquake forecast models for southern California. *Bulletin of the Seismological Society of America* **96**, 90-106.

Ho, C.-H., 2008. Empirical recurrent rate time series for volcanism: application to Avachinsky volcano, Russia. *Volcanol Geotherm Res*. **173**, 15-25.

Ho, C.-H., 2010. Hazard area and recurrence rate time series for determining the probability of volcanic disruption of the proposed high-level radioactive waste repository at Yucca Mountain, Nevada, USA. *Bulletin of Volcanology* **72**, 205-219.

Hong, L.-L., and Guo, S.-W., 1995. Nonstationary Poisson Model for Earthquake Occurrences. *Bulletin of the Seismological of America* **85**, 814-824.

Jackson, D. D., and Kagan, Y. Y., 2006. The 2004 Parkfield Earthquake, the 1985 Prediction, and Characteristic Earthquakes: Lessons for the Future. *Bulletin of the Seismological Society of America* **96**, 397-409.

Kagan, Y. Y., 1993. Statistics of Characteristic Earthquakes. *Bulletin of the Seismological Society of America* **83**, 7-24.

Ljung, G. M., and Box, G. E. P., 1978. On A Measure of Lack of Fit in Time Series Models. *Biometrika* **65**, 297-303.

Mogi, K., 1985. *Earthquake Prediction*. Academic Press Japan, Inc., Tokyo.

Savage, J. C., and Cockerham, R. S., 1987. Quasi-Periodic Occurrence of Earthquakes in the 1978-1986 Bishop-Mammoth Lakes Sequence, Eastern California. *Bulletin of the Seismological Society of America* **77**, 1347-1358.

Schwarz, G., 1978. Estimating the dimensions of a model. *Annals of Statistics* **6(2)**, 461-464.