Statistics and Forecasting of Aftershocks During the 2019 Ridgecrest, California, Earthquake Sequence

Robert Shcherbakov

1Department of Earth Sciences, University of Western Ontario, London, ON, Canada, 2Department of Physics and Astronomy, University of Western Ontario, London, ON, Canada

Abstract  The 2019 Ridgecrest, California, earthquake sequence represents a complex pattern of seismicity that is characterized by the occurrence of a well-defined foreshock sequence followed by a mainshock and subsequent aftershocks. In this study, a detailed statistical analysis of the sequence is performed. Particularly, the parametric modeling of the frequency-magnitude statistics and the earthquake occurrence rate is carried out. It is shown that the clustering of earthquakes plays an important role during the evolution of this sequence. In addition, the problem of constraining the magnitude of the largest expected aftershocks to occur during the evolution of the sequence is addressed. In order to do this, two approaches are considered. The first one is based on the extreme value theory, whereas the second one uses the Bayesian predictive framework. The latter approach has allowed to incorporate the complex earthquake clustering through the Epidemic Type Aftershock Sequence (ETAS) process and the uncertainties associated with the model parameters into the computation of the corresponding probabilities. The results indicate that the inclusion of the foreshock sequence into the analysis produces higher probabilities for the occurrence of the largest expected aftershocks after the M7.1 mainshock compared to the approach based on the extreme value distribution combined with the Omori-Utsu formula for the earthquake rate. Several statistical tests are applied to verify the forecast.

1. Introduction

The occurrence of a significant mainshock presents an opportunity to test different existing or novel statistical approaches to model the evolution of the corresponding sequences of earthquakes that precede and follow the mainshock. Among several statistical measures, the computation of the probability to have the magnitude of the largest expected earthquake to be above a certain value during a predefined future time interval is of critical importance. In this respect, the 2019 Ridgecrest, California, earthquake sequence represents the latest highly productive and nonstandard sequence to be analyzed in detail.

The problem of constraining the magnitudes of the largest expected aftershocks is important as these aftershocks can inflict further damage to already weakened by a mainshock structures or the evolution of the sequence can trigger even larger subsequent events (Gerstenberger et al., 2005; Omi et al., 2013; Page et al., 2016; Shebalin et al., 2011). The standard approach is to use the past seismicity to compute the probabilities of having subsequent strong earthquakes during a finite future time interval. The most recognized model was formulated by Reasenberg and Jones (1989) for California based on the analysis of the past
aftersehock sequences. In that model, the probabilities are computed from the extreme value distribution by assuming that the occurrence of earthquakes follows a nonhomogeneous Poisson process, the earthquake rate is approximated by the Omori-Utsu formula and the frequency-magnitude statistics is described by the left-truncated exponential distribution. Reasenberg and Jones (1989) estimated the average values of the model parameters to be used in California. However, a recent work by Hardebeck et al. (2019) introduced improvements to the original model by analyzing more recent sequences, introducing the ability to control the early incompleteness of aftershock sequences, and using the Bayesian updating of the model parameters. These developments contributed to the introduction of the operational aftershock forecasting in the U.S. by the U.S. Geological Survey (Michael et al., 2019). A similar approach has been undertaken in Japan to create a real-time system for automatic aftershock forecasting (Omi et al., 2016, 2019). Earthquake forecasting centers also operate in New Zealand (Rhoades et al., 2018) and Italy (Taroni et al., 2018), where the evaluation of earthquake probabilities and assessment of earthquake hazard are routinely performed.

A critical aspect of any earthquake forecasting methods is their prospective/retrospective testing and validation (Kagan & Jackson, 1995). This is consistently implemented by the Collaboratory for the Study of Earthquake Predictability (CSEP) (Gerstenberger et al., 2020; Schorlemmer et al., 2007, 2018; Zechar et al., 2010). Within the CSEP framework several statistical methods were developed to test the short/long term earthquake forecasts. Those methods test the consistency of a given forecasting scheme to reproduce the observed number of earthquakes, their spatial and magnitude distributions during the forecasting time interval (Zechar et al., 2010). They also incorporate likelihood based approaches to compare various forecasting schemes. For example, this framework was used to test the performance of aftershock forecasts during the 2011 Tohoku, Japan, earthquake (Nanjo et al., 2012), the 2010 Canterbury, New Zealand, earthquake sequence (Cattania et al., 2018; Rhoades et al., 2016), and the 2016 Kaikoura, New Zealand, earthquake sequence (Rhoades et al., 2018).

An early systematic empirical study of aftershocks concluded that the largest occurred aftershock on average was \( \sim 1.2 \) magnitude less than the magnitude of a mainshock (Båth, 1965). This is referred to as Båth's law. Subsequently, it was proposed that the difference was independent of the number of events and its mean value was proportional to the inverse of the \( b \)-value (Vere-Jones, 1969, 1975). More recent studies have provided further details on this difference (Console et al., 2003; Shcherbakov et al., 2013; Shcherbakov & Turcotte, 2004; Shearer, 2012; Tahir et al., 2012). The aftershock sequences also exhibit scaling with respect to the lower magnitude cutoff (Shcherbakov et al., 2004; Shcherbakov et al., 2005a, 2006, 2015).

An important limitation of all earthquake catalogs is the early aftershock incompleteness (Hainzl, 2016a, 2016b; Kagan, 2004; Peng et al., 2006). This incompleteness can affect the estimation of the model parameters if the magnitude of completeness is underestimated. As a result, this can significantly influence the calculation of the probabilities for the occurrence of extreme earthquakes. To recover partially the true rate a variable magnitude of completeness can be considered (Helmstetter et al., 2006; Omi et al., 2014; Page et al., 2016). Several approaches were suggested to recover the aftershock rate by using the information of early aftershocks in order to estimate the probability of larger subsequent events during future evolution of the sequences (Ebrahimian et al., 2014; Omi et al., 2013, 2016).

The occurrence of strong earthquakes typically produces spatial and temporal clusters. This clustering is a result of triggering by preceding earthquakes that can lead to a cascade of events with a complicated branching structure (Felzer et al., 2004). To describe such a clustering, the ETAS model was introduced that offers a realistic and quantifiable approximation to the earthquake occurrence rate (Ogata, 1988, 1999, 2017). Particularly, it can model the rate of earthquakes punctuated by the occurrence of strong earthquakes. This also allows to quantify the increased earthquake hazard after a mainshock by incorporating the triggering ability of foreshocks, a mainshock, and subsequent aftershocks. It also can be used for short-term forecasting of large earthquakes by studying past seismicity (Ebrahimian & Jalayer, 2017; Harte, 2017; Helmstetter et al., 2006; Ogata, 2017; Omi et al., 2019; Werner et al., 2011).

After the occurrence of the 2019 Ridgecrest earthquakes, several approaches have been used to study the statistical and triggering aspects of this sequence. The operational earthquake forecasting was documented based on the UCERF3-ETAS model (Milner et al., 2020; Savran et al., 2020). Retrospective analysis of the historic seismicity in California and its relation to the initiation of the 2019 Ridgecrest sequence was
performed (Ogata & Omi, 2020). Predictive skills of the models based on the Coulomb stress transfer were analyzed (Mancini et al., 2020; Toda & Stein, 2020). The triggering of aftershocks during the evolution of the sequence was studied using the stress-similarity model (Hardebeck, 2020). The question of changes in the stress field inferred from past seismicity and its relation to the initiation of the Ridgecrest sequence and subsequent relaxation was addressed in Nanjo (2020).

In this study, a detailed statistical analysis of the 2019 Ridgecrest earthquake sequence was performed to study its temporal evolution and frequency-magnitude statistics. In addition, several methods were considered to estimate the probabilities to have the largest expected aftershock to be above a certain magnitude during several stages of the evolution of the sequence. The computation of probabilities was performed using two approaches, that is, the one based on the extreme value theory and the second one using the Bayesian predictive distribution. These approaches assume parametric models for the earthquake occurrence rate and the frequency-magnitude statistics. Specifically, the Omori-Utsu (OU) law (Omori, 1894; Utsu, 1961; Utsu et al., 1995), the compound Omori-Utsu law (Ogata, 1983), and the Epidemic Type Aftershock Sequence (ETAS) process (Ogata, 1988, 1999, 2017) were used to approximate the earthquake rate. The frequency-magnitude statistics of earthquakes was modeled by the left-truncated exponential distribution (Vere-Jones, 2010). The obtained results, which are reported below, suggest that the clustering of earthquakes plays an important role in approximating the earthquake rate and as a consequence can significantly affect the computation of the probabilities for the occurrence of the largest expected aftershocks.

The paper is organized as follows. In Section 2, the statistical methods used in the study are summarized and explained. In Section 3, a detailed analysis of the sequence is presented. The retrospective validation of the forecasting results is given in Section 4. In Section 5, the obtained results are summarized and evaluated. The last section presents concluding remarks.

2. Data and Methods

2.1. The 2019 Ridgecrest Earthquake Sequence

The 2019 Ridgecrest earthquake sequence started on July 4th when several small events of low magnitude occurred not far away from the town of Ridgecrest in California. Then, two strong foreshocks of magnitudes M3.98 and M6.4 struck on July 04, 2019 at 17:02:55 UTC and 17:33:49 UTC, respectively (Figure 1). These events were followed by a well-developed aftershock sequence that culminated in the occurrence of M7.1 mainshock on July 06, 2019 (03:19:53 UTC), which in turn generated a more prolific aftershock sequence. The M6.4 foreshock ruptured several predominantly strike-slip, left-lateral fault segments, whereas the M7.1 mainshock occurred on a system of several right-lateral fault segments conjugate to the rupture of the M6.4 foreshock (Barnhart et al., 2019; Ross et al., 2019). Many of the foreshocks and subsequent aftershocks of the M7.1 mainshock occurred on numerous secondary faults adjacent to the main rupture faults. It was suggested that this earthquake sequence occurred in an immature fault zone with a complex fault structure (Liu et al., 2019; Ross et al., 2019).

To analyze the 2019 Ridgecrest earthquake sequence, the earthquake catalog provided by the Southern California Seismic Network (SCSN, 2020) was used. The spatial distribution of seismicity during 14 days starting from 2019/07/04 (17:02:55 UTC) is shown in Figure 1. This includes the occurrence of the M6.4 foreshock on 2019/07/04 (17:33:49 UTC) and the occurrence of the M7.1 mainshock on 2019/07/06 (03:19:53 UTC). Their focal mechanisms are also shown and were obtained from the SCSN Moment Tensor catalog (SCSN, 2020). The foreshock-aftershock zone for the sequence is defined as an elliptical region outlining the majority of earthquakes that occurred near the ruptures of both the M6.4 foreshock and M7.1 mainshock. Figure 1 also shows the quaternary faults for this region extracted from the U.S.G.S. Quaternary fault and fold database (USGS, 2006).

When analyzing seismicity, several time intervals, during which the parameters of statistical models can be estimated or future evolution of the seismicity can be quantified, are defined. Specifically, the past seismicity is extracted during the training time interval $[T_0, T_t]$. To minimize the effect of earlier earthquakes in the sequence, the training time interval is typically subdivided into a preparatory time interval $[T_0, T_s]$ and a target time interval $[T_s, T_t]$ during which the parameters of the earthquake models are estimated. One also
considers a forecasting time interval \([T_e, T_e + \Delta T]\) during which specific measures of seismicity can be computed or evolution of seismicity can be forecasted. For properly estimating the parameters of earthquake models, it is also important to consider the seismicity above the magnitude of completeness \(m_c\) as typical earthquake catalogs have missing events below this magnitude.

For the statistical modeling of seismicity, the occurrence of earthquakes can be considered as a realization of a stochastic marked point process in time (Daley & Vere-Jones, 2003; Vere-Jones, 2010). In this representation, the earthquakes are characterized by their occurrence times \(t_i\) and magnitudes \(m_i\) represent corresponding marks. The occurrence of earthquakes during a specified time interval can be arranged in an ordered set \(S = \{(t_i, m_i)\}: i = 1, \ldots, n\). In one simplified assumption, the occurrence of earthquakes in the sequence can be described by a nonhomogeneous Poisson marked point process (Shcherbakov et al., 2005b; Utsu et al., 1995), where magnitudes and the time intervals between successive events are not correlated.

### 2.2. Exponential Distribution and the Gutenberg-Richter Scaling Relation

The frequency-magnitude statistics of earthquake magnitudes is typically modeled by the left-truncated exponential distribution (Vere-Jones, 2010):

\[
f_\beta(m) = \beta \exp\left[-\beta(m - m_\beta)\right],
\]

\[
F_\beta(m) = 1 - \exp\left[-\beta(m - m_\beta)\right] \quad \text{for} \quad m \geq m_\beta,
\]

\[\text{(1)}\]

\[\text{(2)}\]
where $f_\theta(m)$ is the probability density, $F_\theta(m)$ is the cumulative distribution function, and $\theta = [\beta]$ is the model parameter. $m_0$ is a given lower magnitude cutoff set above the catalog completeness level $m_c$. All earthquakes above $m_0$ during the target time interval $[T_s, T_e]$ are used to estimate the model parameter $\beta$.

The parameter $\beta$ is related to the $b$-value of the Gutenberg-Richter (GR) scaling relation, $\beta = \ln(10)b$ (Gutenberg & Richter, 1944):

$$\log_{10} N(\geq m) = a - bm,$$

(3)

where $N(\geq m)$ is the cumulative number of earthquakes above magnitude $m$. The GR relation combines two aspects of the occurrence of earthquakes, i.e. the frequency-magnitude statistics of earthquake magnitudes and the average rate of the occurrence of earthquakes, which is quantified through the parameter $a$. $N(\geq 0) = 10^a$ gives the total number of earthquakes above magnitude zero that occurred during the corresponding time interval.

The standard method to estimate the parameter $\beta$ (or $b$-value) is to use the maximum likelihood approach, which has an analytic solution for the point estimator of the parameter of the exponential distribution. However, in typical earthquake catalogs the magnitudes are binned and not continuous variables. Therefore, one needs to apply a corrected estimator, which explicitly assumes the binning of the magnitudes (Bender, 1983). For the estimation of the parameter uncertainties at a given confidence level in case of binned magnitudes one can use the method suggested in Tinti and Mulargia (1987).

### 2.3. Omori-Utsu Law

The occurrence of moderate to large earthquakes, in most cases, triggers subsequent aftershock sequences and results in the rise of seismic activity. The most accepted model that reproduces the rate of the occurrence of aftershocks is known as the Omori-Utsu (OU) law (Omori, 1894; Utsu, 1961; Utsu et al., 1995):

$$\lambda_\omega(t) = \frac{K_0}{(t + c_0)^p},$$

(4)

where $\lambda_\omega$ is the rate of aftershocks per unit time for events above a certain magnitude $m_\omega$. $\omega = [K_0, c_0, p_0]$ are the OU model parameters. The time $t$ is elapsed since $T_0 = 0$, which corresponds to the time of the occurrence of the mainshock. The parameter $K_0$ describes the productivity of the sequence, $c_0$ is a characteristic time, and $p_0$ specifies how fast or slow the sequence decays in time. The parameters can be estimated using the maximum likelihood method and parameter uncertainties are computed using the inverse of the Fisher information matrix, which is derived from the likelihood function (Ogata, 1983, 1999). In this model, it is assumed that the occurrence of earthquakes can be approximated by a nonhomogeneous Poisson process, where earthquake magnitudes are independent and identically distributed (i.i.d.) random numbers and do not influence the future earthquake rate. The Bayesian approach to estimate the parameters and their uncertainties of the OU law was also implemented (Holschneider et al., 2012).

The Omori-Utsu law, Equation 4, is applicable to “standard” aftershock sequences with a single mainshock and a consistently decaying rate. However, in some cases the earthquake sequence can be punctuated by several strong shocks each one of them producing their own aftershocks. In that case, a compound Omori-Utsu model can be considered (Ogata, 1983; Shcherbakov et al., 2012). In a case of two strong earthquakes, it is written as:

$$\lambda_\omega(t) = \frac{K_1}{(t + c_1)^p} + \frac{K_2}{(t - \tau_m + c_2)^p},$$

(5)

where $\omega = [K_1, c_1, p_1, K_2, c_2, p_2]$, time $t$ is elapsed since the occurrence of the first event at $T_0 = 0$ and $\tau_m$ is the time of the occurrence of the second strong event. $H(x)$ is a Heaviside step function and is equal to one for positive $x \geq 0$ and is zero otherwise. For the times past the occurrence of the second strong earthquake ($t \geq \tau_m$), Equation 5 defines the earthquake rate as a superposition of two aftershock sequences triggered by the both strong earthquakes.
2.4. Epidemic Type Aftershock Sequence (ETAS) Model

The occurrence of earthquakes is characterized by the clustering of seismicity. This clustering is a direct manifestation of the ability of earthquakes to trigger subsequent events. The ETAS model was introduced to reflect this essential aspect of the occurrence of earthquakes (Ogata, 1988, 1999, 2017). In the temporal version of the model, the conditional earthquake rate \( \lambda_{w}(t | \mathcal{H}_t) \) at a given time \( t \) is given as (Harte, 2010; Ogata, 1988):

\[
\lambda_{w}(t | \mathcal{H}_t) = \mu + K \sum_{i \in \mathcal{I}} e^{\alpha(m_i - m_0)} \left( \frac{t - t_i}{\epsilon} + 1 \right)^c,
\]

(6)

where \( \omega = [\mu, K, c, p, \alpha] \) is a set of parameters and \( m_0 \) is a reference magnitude. The summation is performed over the history, \( \mathcal{H}_t \), of past events up to time \( t \) during the time interval \([T_0, t]\). \( N_t \) is the number of earthquakes in the interval \([T_0, t]\) above the lower magnitude cutoff \( m_0 \). In the ETAS process, a certain fraction of earthquakes occurs randomly with a constant rate \( \mu \). These earthquakes are associated with background seismicity driven by tectonic loading and can be modeled as a homogeneous Poisson process. It is also postulated that each earthquake is capable of triggering its own offsprings. As a result, the total earthquake rate at a given time, is a superposition of the background rate given by \( \mu \) and the contribution from each already occurred earthquake.

As the ETAS rate, Equation 6, is conditioned on past seismicity \( \mathcal{H}_t \), one has to minimize the effect of lack of earthquakes at the start of the sequence when estimating the ETAS parameters. For this, one can consider a short time interval \([T_0, T_s]\) before the target time interval \([T_s, T_e]\). The earthquakes in the interval \([T_0, T_s]\) can be used to properly estimate the conditional earthquake rate during the target time interval \([T_s, T_e]\). The explicit forms of the log-likelihood function and the productivity of the ETAS model were given in Shcherbakov et al. (2019). The ETAS parameters \( \omega = [\mu, K, c, p, \alpha] \) are estimated in the target time interval \([T_s, T_e]\) by maximizing the likelihood function and the uncertainties are computed using the inverse of the Fisher information matrix.

2.5. Extreme Value Distribution

For the sequence of earthquake that can be described as a nonhomogeneous Poisson process, the probability that the magnitude of the largest expected event will exceed \( m \) for all possible number of events during a future time interval \([T_e, T_e + \Delta T]\) can be computed from the extreme value distribution (EVD) (Campbell, 1982; Coles, 2001; Daley & Vere-Jones, 2003):

\[
P_{Ev}(m_{ex} > m | \theta, \omega, \Delta T) = 1 - \exp \left[ -\Lambda_{w}(\Delta T) \left[ 1 - F_{\theta}(m) \right] \right],
\]

(7)

where the productivity is \( \Lambda_{w}(\Delta T) = \int_{T_0}^{T_e + \Delta T} \lambda_{w}(t) dt \). Using the exponential model for the magnitude distribution, Equation 2, this results in the Gumbel distribution for the magnitudes of extreme earthquakes:

\[
P_{Ev}(m_{ex} > m | \theta, \omega, \Delta T) = 1 - \exp \left[ -\Lambda_{w}(\Delta T) \exp \left[ -\beta(m - m_0) \right] \right].
\]

(8)

Assuming that the earthquake rate is described by the OU law (4), the productivity \( \Lambda_{w}(\Delta T) \) can be computed explicitly and takes the following form for \( p_o \neq 1 \):

\[
\Lambda_{w}(\Delta T) = K_o \frac{(T_e + c_0)^{1-p_0} - (T_e + \Delta T + c_0)^{1-p_0}}{p_o - 1}.
\]

(9)

Given a set of parameters \( \theta, \omega \), which can be estimated from past seismicity during the training time interval \([T_0, T_s]\), Equations 8 and 9 allow to compute the probability to have the extreme earthquake above magnitude \( m \) during a future time interval \( \Delta T \). It is equivalent to the computation of the probabilities given in Reasenberg and Jones (1989).
For the compound OU model (5) the productivity \( \Lambda_\omega(\Delta T) \) can be expressed as follows for \( p_1 \neq 1 \) and \( p_2 \neq 1 \):

\[
\Lambda_\omega(\Delta T) = K_1 \left( T_c + c_1 \right)^{1-p_1} - \left( T_c + \Delta T + c_1 \right)^{1-p_1} + K_2 \left( T_c - \tau_m + c_2 \right)^{1-p_2} - \left( T_c + \Delta T - \tau_m + c_2 \right)^{1-p_2},
\]

where \( \tau_m \) is the time of the occurrence of the second strong earthquake during the training time interval \([T_x, T_f]\).

### 2.6. Bayesian Predictive Distribution

The computation of the EVD (7) using specific parametric models for the earthquake rate and frequency-magnitude statistics, requires the knowledge of the model parameters. However, the true values of the model parameters are not known for specific earthquake sequences. As a result, the parameter estimates are used, which are computed with a given range of uncertainties. Those uncertainties can significantly affect the computation of the corresponding probabilities. The incorporation of the model uncertainties into the computation of probabilities can be achieved through the Bayesian predictive distribution (BPD) (Shcherbakov et al., 2018, 2019; Zöller et al., 2013). The BPD for the largest expected event \( m_{\text{ex}} \) to be greater than a certain value \( m \) and during the forecasting time interval \( \Delta T \) is:

\[
P_B(m_{\text{ex}} > m \mid S, \Delta T) = \int_\Theta \int_\Omega \int P_{\text{EV}}(m_{\text{ex}} > m \mid \theta, \omega, \Delta T) p(\theta, \omega \mid S) d\theta d\omega d\Omega.
\]

where \( \Theta \) and \( \Omega \) define the multidimensional domains of the frequency-magnitude distribution and earthquake rate parameters, respectively. When computing the predictive distribution, Equation 11, the model parameter uncertainties are fully integrated into the BPD (Renard et al., 2013; Shcherbakov et al., 2019). This is done through the use of the posterior distribution function \( p(\theta, \omega \mid S) \), which characterizes the distribution of the model parameter uncertainties.

For the ETAS model, the extreme value distribution for the extreme events does not follow Equation 7, due to stochastic nature of the process, which deviates from a nonhomogeneous Poisson process. In this case, one can compute the extreme value distribution by stochastic simulation of the ETAS model and extracting the maximum magnitude from each simulated sequence (Shcherbakov et al., 2019).

To compute the BPD, Equation 11, for a given training time interval, first, the Markov Chain Monte Carlo (MCMC) sampling of the posterior distribution is performed to generate a chain of the ETAS parameters using the Metropolis-within-Gibbs algorithm. The generated chains of length \( N_{\text{sim}} \) are used to simulate the ensemble of the ETAS processes forward in time during the forecasting time interval \( \Delta T \). From each simulated sequence the maximum event is extracted and the distribution of these maxima approximates the BPD (Shcherbakov et al., 2019).

### 3. Results

#### 3.1. Frequency-Magnitude Statistics

The earthquakes within an elliptical region, given in Figure 1, were extracted during predefined target time intervals. The frequency-magnitude statistics of earthquake magnitudes were computed for the foreshock sequence starting from July 04, 2019 (17:02:55 UTC) which corresponds to \( T_0 = 0 \) and during the target time interval \([T_x, T_f]\) = [10\(^{-4}\), 1.428] days. It was also computed for the aftershocks of the M7.1 mainshock starting from July 06, 2019 (03:19:53 UTC) during 7 days after the mainshock. The frequency-magnitude statistics was also computed for the whole sequence including both foreshocks and aftershocks during 31 days. The results are given in Figure 2 as open symbols for events larger than \( m \geq 2.0 \). The maximum likelihood fits of the exponential distribution, Equation 2, to the frequency-magnitude data above \( m \geq 3.2 \) are shown as...
GR plots with estimated $b$-values using the method of Bender (1983) and their 95% confidence intervals according to Tinti and Mulargia (1987).

The sequence exhibited a change in the slope of the frequency-magnitude statistics around a magnitude 3.2. This was the reason to use only the events above this value in the analysis. This change in the behavior can be the result of the early aftershock incompleteness observed right after the M6.4 foreshock and the M7.1 mainshock or it can be related to the fact that the aftershocks occurred on a distributed fault network and the geometrical distribution of fault sizes affected the statistics of earthquake magnitudes. The fit of the exponential distribution, Equation 2, (or the corresponding Gutenberg-Richter relation) to the frequency-magnitude statistics of the foreshock and aftershock sequences produced the $b$-values which were typical for tectonic earthquakes as illustrated in Figure 2. The largest aftershock of the M7.1 mainshock had a magnitude 5.5 and occurred less than half an hour after the mainshock. Two more strong aftershocks of magnitude 4.7 and 5.0 occurred later in the sequence on 20th and 48th days after the mainshock. The value of the largest occurred aftershock is lower than what would be expected from Båth’s law (Båth, 1965). It is possible that the M6.4 foreshock partially released the accumulated strain energy in the region and this resulted in a lower magnitude of the largest occurred aftershock.

3.2. Earthquake Rate Evolution and Modeling

First, the earthquake rate was modeled separately for the foreshock and aftershock sequences using the OU law (4). The results are given in Figure 3 for all earthquakes above magnitude $m \geq 3.2$. For the foreshock sequence, the following target time interval was used $[T_s, T_e] = [10^{-3}, 1.407]$ days with $T_0 = 0$ corresponding to July 04, 2019 (17:33:49 UTC). For the aftershock sequence, $T_0 = 0$ was set to the occurrence of the M7.1 mainshock on July 06, 2019 (03:19:53 UTC) with the target time interval $[T_s, T_e] = [10^{-3}, 30]$ days. The OU law parameters for the foreshock and aftershock sequences are given in the legend with the corresponding
The earthquake decay rates after the M6.4 foreshock and M7.1 mainshock exhibited a consistent pattern observed in other prominent aftershock sequences. The fit of the OU law (4) produced $p = 0.99 \pm 0.18$ for the foreshock sequence and $p = 1.28 \pm 0.07$ for the aftershock sequence (Figure 3). The smaller $p$-value for the foreshock sequence can be the result of a strong M5.36 foreshock that occurred 16.2 h before the M7.1 mainshock and triggered its own sequence of events.

Next, the compound OU model (5) was used to fit the sequence starting from the occurrence of the M6.4 foreshock on July 04, 2019 (17:33:49 UTC) corresponding to $T_0 = 0$ and during the following target time interval $[T_s, T_e] = [10^{-3}, 8.407]$ days. This is illustrated in Figure 4 and Figure S1. The maximum likelihood fitting of the compound OU model yielded the following parameters $\{K_1, c_1, p_1, K_2, c_2, p_2\} = \{23.22, 0.0026, 0.93, 40.3, 0.034, 1.59\}$.

The ETAS model, Equation 6, was fitted to the 2019 Ridgecrest sequence using a number of target time intervals. In one particular example, the training time interval $[T_s, T_e] = [0.03, 8.428]$ days was used with $T_0 = 0$ corresponding to the start date July 04, 2019 (17:02:55 UTC). The estimated conditional rate, Equation 6, and the corresponding earthquake magnitudes above $m \geq 3.2$ are plotted in Figure 5 and Figure S2. For comparison, the separate fits of the Omori-Utsu law to the foreshocks and aftershocks of the M7.1 mainshock are also plotted with the parameters given in Figure 3.

Finally, the point estimates of the model parameters and their 95% confidence intervals were computed at predefined times during the evolution of the sequence (Figure 6). The reported $b$-value at time 1.428 days corresponds to the foreshock sequence starting from the occurrence of the M3.98 event on July 04, 2019 (17:02:55 UTC). The frequency-magnitude statistics and the fitting of the GR relation to the foreshock sequence is also illustrated in Figure 2. The subsequent estimates of $b$-values at 1 day, 2 days, etc., correspond to the time duration of the aftershock sequence since the M7.1 mainshock (Figure 6a). Similarly, the parameters of the OU law (4) were estimated during the same time intervals (Figure 6b). In addition, the point estimates of the ETAS model parameters were also computed (Figure 6c). The parameter $\mu$ was held constant at $\mu = 0.05$ to improve the stability of the parameter estimation. It was assumed that the background
seismicity rate for earthquakes above magnitude $m = 3.2$ was relatively low in this region prior to the start of the sequence.

### 3.3. Forecasting the Magnitude of the Largest Expected Earthquake

The EVD (7) and the BPD (11) were used to compute retrospectively the probabilities of having the largest expected earthquakes to occur during predefined times of the evolution of the 2019 Ridgecrest earthquake sequence. This was done both before and after the occurrence of the M7.1 mainshock using the OU (4), compound OU (5), or ETAS (6), parametric models for the earthquake rate and the exponential distribution, Equation 2, for the distribution of earthquake magnitudes. When computing the probabilities for the aftershock sequence generated by the M7.1 mainshock two cases were analyzed. In the first consideration, only the aftershocks were used. However, when using the ETAS model and the compound OU model the foreshock sequence was also incorporated into the analysis.

First, the only aftershocks of the M7.1 mainshock were used to compute the probabilities of having the strongest aftershocks above a specified magnitude during a future time interval of $\Delta T = 7$ days. The occurrence of the M7.1 mainshock on July 04, 2019 (03:19:53 UTC) corresponded to $T_0 = 0$ with the target time interval $[T_s, T_e] = [10^{-4}, 1]$ days. One particular example is given in Figure 7, where the EVD (8) was computed after 1 day and plotted as a short dashed violet curve. The following model parameter estimates were used: $[\beta, K_0, c_0, p_0] = [2.28, 39.85, 0.038, 1.65]$. The corresponding probabilities to have strong aftershocks above $m_{ex} \geq 5.0, 6.1, 7.1$ are also given.

Next, the BPD (11) was computed using the aftershocks of the M7.1 mainshock during different training time intervals to forecast the magnitudes of the largest expected earthquakes to occur during the evolution of the sequence. The OU law (4) was used to approximate the earthquake rate. The exponential distribution, Equation 2, was used to model the frequency-magnitude statistics. The forecasting time interval was fixed at $\Delta T = 7$ days. The computed BPD to estimate probabilities for the largest expected aftershocks above magnitude $m = 3.2$ during one day after the mainshock is plotted in Figure 7 as a dash-dotted cyan curve. This was

![Figure 4](image-url)
done by employing the MCMC sampling of the posterior distribution and the Gamma distribution for the priors of the model parameters (Shcherbakov et al., 2019). The total number of 200,000 MCMC sampling steps were performed for each model. The first 100,000 steps were discarded as “burn in” and the remaining \( N_{\text{sim}} = 100,000 \) sampling steps were used for the synthetic model simulations or analysis. For the OU model, this is given in Figure S3. The distribution of the OU model parameters computed from the MCMC chain is illustrated in Figure S4. The matrix plot of the pairs of the OU model parameters is given in Figure S5. The values for the mean and variance of the prior distribution (Gamma) of the OU model parameters are provided in Table S1.

To investigate the influence of the foreshocks on the computation of the probabilities for the largest expected aftershocks, the EVD (8) using the compound OU law (5) and the BPD using the ETAS model (6) were computed for the earthquake sequence starting from the occurrence of the first M3.98 foreshock on July 04, 2019 (17:02:55 UTC). The estimated conditional earthquake rate (solid curve) is plotted using the following ETAS parameters: \( \mu = 0.05, K = 2.64, c = 0.015, p = 1.41, \) and \( \alpha = 2.10 \). For comparison, the Omori-Utsu law fit, Equation 4, is plotted as a short-dashed curve.

Figure 5. The occurrence of earthquakes during the evolution of the 2019 Ridgecrest sequence and the fitting of the ETAS model, Equation 6. The start of the sequence \( T_0 = 0 \) corresponds to the time of the occurrence of the M3.98 foreshock on July 04, 2019 (17:02:55 UTC). The ETAS model is fitted to the sequence during the target time interval \([T_s, T_e] = [0.03, 8.428]\) days. The estimated conditional earthquake rate (solid curve) is plotted using the following ETAS parameters: \( \mu = 0.05, K = 2.64, c = 0.015, p = 1.41, \) and \( \alpha = 2.10 \). For comparison, the Omori-Utsu law fit, Equation 4, is plotted as a short-dashed curve.
Figure 1: Time evolution of (a) $k_1$, (b) $K_0$, (c) $C_0$, (d) $P_0$, and (e) $\alpha$ over 30 days.
The probabilities to have the largest expected earthquake above a certain magnitude can be computed at specified times during the evolution of the sequence. This can be done by increasing progressively the upper limit $T_e$ of the target time interval $[T_s, T_e]$ for a fixed forecasting interval $ΔT$. Figure 8 illustrates the computed probabilities from the BPD (11) with the ETAS model (6) for the earthquake rate, and the exponential distribution, Equation 2, for the frequency-magnitude statistics. $T_0 = 0$ corresponded to the date July 04, 2019 (17:02:55 UTC) and $T_s = 0.03$ days. The MCMC sampling steps, the distribution of the ETAS model parameters, and the matrix plot of the pairs of the ETAS parameters are given in Figures S9–S11 for the following target time interval $[T_s, T_e] = [0.03, 2.4284]$ days. The probabilities were estimated for the largest expected earthquakes to be above $m_{ex} ≥ 4.5, 5.0, 6.1, 6.4, \text{and} 7.1$. First, the probabilities were computed using only the foreshock sequence right before the occurrence of the M7.1 mainshock with $T_e = 1.4284$ days. After that, the probabilities were recomputed for each subsequent day after the M7.1 mainshock by incorporating the information from the newly occurred aftershocks. For reference, the fit of the ETAS model is also shown as a red curve using the following estimated model parameters $[β, μ, K, c, p, α] = [2.39, 0.05, 3.47, 0.01, 1.35, 2.05]$ during the training time interval $[T_s, T_e] = [0.03, 15.4284]$ days. The forecast evolution during 330 days after the occurrence of the M7.1 mainshock is given in Figure S12. It also illustrates the computed probabilities before the occurrence of the M5.5 event, which occurred on June 4, 2020.

Finally, Figure 9 provides a comparison of the results for the computation of the probabilities to have the expected largest aftershock to be greater than $m_{ex} ≥ 6.1$ after progressively increasing times $T_e$ during the evolution of the sequence by using several methods examined in this work. The forecasting time interval was

**Figure 7.** The extreme value and the Bayesian predictive distributions for the 2019 Ridgecrest sequence. The BPD is shown as a solid red curve using the ETAS model and MCMC sampling with the Gamma prior for the foreshocks and 1 day of aftershocks after the M7.1 mainshock. For the same sequence of events, the EVD using the compound OU law is shown as a dashed blue curve. For the rest of the distributions, 1 day of aftershocks after the M7.1 mainshock was used: the OU rate using the MCMC sampling with the Gamma prior (dash-dotted cyan curve); the Gumbel distribution with OU rate (short dashed violet curve).

The probabilities to have the largest expected earthquake above a certain magnitude can be computed at specified times during the evolution of the sequence. This can be done by increasing progressively the upper limit $T_e$ of the target time interval $[T_s, T_e]$ for a fixed forecasting interval $ΔT$. Figure 8 illustrates the computed probabilities from the BPD (11) with the ETAS model (6) for the earthquake rate, and the exponential distribution, Equation 2, for the frequency-magnitude statistics. $T_0 = 0$ corresponded to the date July 04, 2019 (17:02:55 UTC) and $T_s = 0.03$ days. The MCMC sampling steps, the distribution of the ETAS model parameters, and the matrix plot of the pairs of the ETAS parameters are given in Figures S9–S11 for the following target time interval $[T_s, T_e] = [0.03, 2.4284]$ days. The probabilities were estimated for the largest expected earthquakes to be above $m_{ex} ≥ 4.5, 5.0, 6.1, 6.4, \text{and} 7.1$. First, the probabilities were computed using only the foreshock sequence right before the occurrence of the M7.1 mainshock with $T_e = 1.4284$ days. After that, the probabilities were recomputed for each subsequent day after the M7.1 mainshock by incorporating the information from the newly occurred aftershocks. For reference, the fit of the ETAS model is also shown as a red curve using the following estimated model parameters $[β, μ, K, c, p, α] = [2.39, 0.05, 3.47, 0.01, 1.35, 2.05]$ during the training time interval $[T_s, T_e] = [0.03, 15.4284]$ days. The forecast evolution during 330 days after the occurrence of the M7.1 mainshock is given in Figure S12. It also illustrates the computed probabilities before the occurrence of the M5.5 event, which occurred on June 4, 2020.

Finally, Figure 9 provides a comparison of the results for the computation of the probabilities to have the expected largest aftershock to be greater than $m_{ex} ≥ 6.1$ after progressively increasing times $T_e$ during the evolution of the sequence by using several methods examined in this work. The forecasting time interval was
Figure 8. The probabilities for the largest expected earthquake to be above the magnitudes $m_{ex} \geq 4.5, 5.0, 6.1, 6.4,$ 7.1 and during the progressively increasing time intervals since July 04, 2019 (17:02:55 UTC). The probabilities are estimated using the BPD combined with the ETAS model for the earthquake rate during the forecasting time interval $\Delta T = 7$ days and plotted in a logarithmic scale. The earthquake magnitudes of the 2019 Ridgecrest sequence are plotted as open diamonds for all events above magnitude $m \geq 3.2$. The fit of the ETAS model is shown as a solid curve.

Figure 9. The comparison of the computed probabilities for the largest expected aftershock to be above magnitude $m_{ex} \geq 6.1$ during the progressively increasing time intervals since July 04, 2019 (17:02:55 UTC) for the fixed forecasting time interval $\Delta T = 7$ days. The four models were considered: the EVD with the OU law (solid squares), the EVD with the compound OU formula (solid circles), the BPD with the ETAS model (solid triangles), and the BPD with the OU law (solid diamonds).
set to $\Delta T = 7$ days. Specifically, the EVD with the OU law, Equations 8 and 9, was used and the estimated probabilities are plotted as solid squares. Next, the compound OU law (10) was used in the EVD computation and the results are plotted as solid circles. The computed probabilities from the BPD (11) with the ETAS model (6) as the earthquake rate are plotted as solid triangles. Finally, the probabilities were computed from the BPD with the earthquake rate modeled using the standard OU law (4) and are plotted as solid diamonds.

4. Forecast Validation

The extreme value distribution, Equation 8, and the Bayesian predictive distribution, Equation 11, allow to compute the probability of having the expected largest event during the forecasting time interval $\Delta T$. This computation critically depends on the proper simulation of the earthquake rate and the frequency-magnitude distribution of earthquakes during $\Delta T$. Therefore, it is important to perform specific statistical tests to validate retrospectively as to how the models, that are used to describe those aspects of seismicity, accurately reproduce the observed earthquakes during the forecasting time intervals. One such test has been developed for the CSEP testing framework and is known as the N-test (Kagan & Jackson, 1995; Schorlemmer et al., 2007; Zechar et al., 2010). This test is used to quantify as to how accurately a given stochastic process reproduces the observed number of earthquakes above a certain magnitude during the forecasting time interval.

The following implementation of the N-test is considered in this work. It is assumed that $N_{\text{obs}}$ earthquakes above magnitude $m_0$ occurred during a given forecasting time interval $[T_t, T_t + \Delta T]$. The posterior distribution of the parameters of a given stochastic point process model is sampled by the MCMC method $N_{\text{sim}}$ times using the information of the earthquakes that occurred during the training time interval $[T_t, T_t]$. The MCMC sets of the model parameters are used to model forward in time a given point process during the forecasting time interval $\Delta T$. The synthetic simulations produce the distribution of the number of the forecasted events at the end of the interval $\Delta T$ corresponding to each MCMC set of model parameters. The N-test statistically assesses whether the observed number of earthquakes $N_{\text{obs}}$ is consistent with the forecast.

The two quantile scores are computed (Zechar et al., 2010):

$$\delta_1 = 1 - P(N_{\text{obs}} - 1 \mid N_{\text{fore}}),$$

$$\delta_2 = P(N_{\text{obs}} \mid N_{\text{fore}}),$$

where $N_{\text{sim}}$ is the average number of forecasted events above magnitude $m_0$ at the end of the forecasting time interval $T_t + \Delta T$. $P(x \mid \lambda)$ is the cumulative Poisson distribution with the expectation $\lambda$. As a result, $\delta_1$ gives the probability of observing at least $N_{\text{sim}}$ events and $\delta_2$ gives the probability of observing at most $N_{\text{obs}}$ events. The forecast underpredicts the observations if $\delta_1$ is very small and the forecast overpredicts the observation if $\delta_2$ is very small. Therefore, one can consider a one-sided test with an effective significance level $\alpha_{\text{eff}}$. If the computed probabilities $\delta_1$ and $\delta_2$ are smaller than $\alpha_{\text{eff}}$ then the forecast can be rejected.

The second test, which is known as M-test, has been suggested to check whether the distribution of the forecasted magnitudes is consistent with the observed magnitudes (Schorlemmer et al., 2007; Zechar et al., 2010). The M-test is performed by computing a quantile score $\kappa$. The values of $\kappa$ below a significance level $\alpha_{\text{eff}}$ signify that the distribution of forecasted earthquake magnitudes is inconsistent with observations. The details of computing the $\kappa$ score can be found in Zechar et al. (2010).

Two more tests have been introduced to compare the performance of different forecasting models. These are known as R- and T-test (Rhoades et al., 2011; Schorlemmer et al., 2007). The R-test is performed by computing the log-likelihood ratio for two models under consideration. The joint log-likelihood for given earthquake observations during the forecasting time interval can be written as follows:

$$L(M \mid A) = \log[\Pr(M \mid A)] = \sum_{i \in B} \left\{ -\lambda(i) + m(i) \log[\lambda(i)] - \log[m(i)!] \right\},$$

where $M = \{m(i) \mid i \in B\}$ is the set of the number of earthquakes $m(i)$ in each magnitude bin above a certain magnitude threshold. $A = \{\lambda(i) \mid i \in B\}$ is the earthquake forecast produced by a given point process in each magnitude bin, where $\lambda(i)$ is the number of earthquakes forecasted in bin $i$ and the magnitude
binning coincides with the binning of the earthquake catalog. In the definition of the joint log-likelihood, Equation 14, it is assumed that the number of earthquakes in a forecast bin follows a Poisson distribution:

\[ \Pr(m \mid \lambda) = \frac{\lambda^m e^{-\lambda}}{m!} \]

To compare two models, \( \Lambda^1 \) and \( \Lambda^2 \), that forecast the same sequence of events one can compute the log-likelihood ratio:

\[ \Lambda\delta = \log \left( \frac{L(M \mid \Lambda^1)}{L(M \mid \Lambda^2)} \right) \]

In applying the R-test, one of the two models is assumed to be correct and is used to simulate the ensemble of synthetic earthquake events and compute the log-likelihood ratios for each synthetic record by using both models. These ratios are compared with the log-likelihood ratio computed for the observed earthquake sequence during the forecasting interval. The properly normalized fraction of the simulated ratios that are less than the observed ratio gives the quantile score \( \alpha \) (Schorlemmer et al., 2007). The values of \( \alpha \), that are larger than a certain significance level, support the model that was assumed to be correct. This test is symmetric with respect to both models and can result in the situations when both models reject each other (Rhoades et al., 2011). To overcome this difficulty, a so called T-test was introduced along with the sample information gain per earthquake (Rhoades et al., 2011). The sample information gain per earthquake of the model \( \Lambda^2 \) over the model \( \Lambda^1 \) is defined as

\[ I_\delta(\Lambda^2, \Lambda^1) = \frac{\Lambda_\delta}{N_{\text{obs}}} \]

where \( N_{\text{obs}} \) is the number of observed earthquakes during the forecasting time interval \( \Delta T \). The T-test checks whether the sample information gain is statistically different from zero that indicates a significant difference between the two models (Rhoades et al., 2011).

One important difference in performing the above tests is implemented in this work. To account for the stochastic variability of the model parameters and the uncertainty associated with the prior information on the model parameters, the MCMC sampling of the posterior distribution of the model parameters is performed to produce a chain of model parameters that are used when simulating the models forward in time during the forecasting time interval.

The N-, M-, R-, and T-tests check the consistency of the underlying earthquake rate and frequency-magnitude distribution models. To test the consistency of the Bayesian predictive distribution, Equation 11, with the observed largest earthquakes during the forecasting time interval \([T_e, T_e + \Delta T]\), one can evaluate the posterior predictive \( p_{\delta} \) value (Gelman et al., 2013). The Bayesian \( p_{\delta} \) value gives the probability that the largest simulated earthquakes can be more extreme than the observed largest earthquake during the forecasting time interval. It is defined as follows:

\[ p_{\delta} = \Pr\left[ T(\hat{y}, \theta, \omega) \geq T(y, \theta, \omega) \mid y \right], \tag{15} \]

where \( T(y, \theta, \omega) \) is a test quantity computed for an observed variable \( y \) and simulated variable \( \hat{y} \). The test quantity \( T(y, \theta, \omega) \) characterizes data \( y \) with given model parameters \( \theta \) and \( \omega \). It is used for model checking in Bayesian analysis similar to a test statistic in classical testing. One possible choice for the test quantity is:

\[ T(y, \hat{\theta}, \hat{\omega}) = \max(y) \]

In practice, the Bayesian \( p_{\delta} \) value can be computed from the MCMC chain of the model parameters \( \theta \) and \( \omega \). For each set of the model parameters, the stochastic forecasting model is simulated forward in time and the largest event is extracted. This will allow to compute \( T(y, \hat{\theta}, \hat{\omega}) = \max(\hat{y}) \). The realized test quantity \( T(y, \hat{\theta}, \hat{\omega}) = \max(y) \) is simply the value of the largest observed earthquake during the forecasting time interval. Therefore, the estimated \( p_{\delta} \) value is the proportion of the test quantities for the simulated maximum events that are larger than the observed largest event:

\[ p_{\delta} = \frac{\left\lfloor T(\hat{y}, \theta, \omega) \geq T(y) \mid i = 1, \ldots, N_{\text{sim}} \right\rfloor}{N_{\text{sim}}} \tag{16} \]

where \( N_{\text{sim}} \) is the total number of simulated sequences from the MCMC chain and \( \left\lfloor x \right\rfloor \) gives the size of the set \( x \).

### 4.1. Application to the 2019 Ridgecrest Sequence

The three-point process models (OU, compound OU, and ETAS) were examined to see whether they were consistent with the observed seismicity during the forecasting time intervals \([T_e, T_e + \Delta T]\). For this, N- and
M-tests were performed. Figure 10a shows the observed number of earthquakes above magnitude $m \geq 3.2$ (as solid black diamonds) during a fixed forecasting time interval $\Delta T = 7$ days and varying training time interval $[T_s, T_e]$. The numbers are plotted at the end of the forecasting time interval with the training interval ending after 1, 2, 3, 4, 5, 6, 7, 10, 14, 30 days after the M7.1 mainshock (the corresponding $T_e = 2.4284, 3.4284, \ldots, 22.4284, 31.4284$). For example, the first symbol at $T_e + \Delta T = 9.4284$ days gives 89 earthquakes above magnitude $m \geq 3.2$. The symbols indicate the number of the observed (black solid diamonds) and the mean number of forecasted earthquakes during $\Delta T = 7$ days computed at times $T_e + \Delta T$. (b) The end of the training time interval is fixed at $T_e = 3.428$ days while the forecasting time interval is increasing as $\Delta T = 1, 2, 5, 7, 10, 14$. The shaded bands correspond to 95% confidence intervals.

Figure 10. The observed and forecasted numbers of earthquakes starting after one day of aftershocks post M7.1 mainshock and during specified forecasting and training time intervals by using the three rate models: Omori-Utsu (OU), compound OU, and ETAS. (a) The forecasting time interval $\Delta T = 7$ days is fixed while the end of the training time interval $T_e$ is progressively increasing as $T_e = 2.428, 3.428, \ldots, 22.428, 31.428$ days. The symbols indicate the number of the observed (black solid diamonds) and the mean number of forecasted earthquakes during $\Delta T = 7$ days computed at times $T_e + \Delta T$. (b) The end of the training time interval is fixed at $T_e = 3.428$ days while the forecasting time interval is increasing as $\Delta T = 1, 2, 5, 7, 10, 14$. The shaded bands correspond to 95% confidence intervals.
magnitude 3.2 that occurred during 7 days starting after 1 day ($T_e = 2.4284$) after the M7.1 mainshock. It also shows the average forecasted numbers of earthquakes with the corresponding 95% bands (plotted as shaded regions) simulated by the three models. Each model was simulated $N_{sim} = 100,000$ times forward in time during $\Delta T = 7$ days and for the varying ends of the training time interval $T_e$. For each model simulation, the parameters were chosen from the MCMC chain obtained by sampling the posterior distribution of the model parameters. This allowed to incorporate the variability of the model parameters into the forecasted numbers. Similarly, Figure 10b illustrates the observed and forecasted number of earthquakes when the end of the training time interval was held fixed at $T_e = 3.4284$ days (2 days after the M7.1 mainshock) and the forecasting time interval varied $\Delta T = 1, 2, 5, 7, 10, 14$ days. For the compound OU and ETAS models the preceding foreshock sequence was used. For the OU model only the aftershocks of the M7.1 mainshock were used.

To analyze to what extent the considered models underpredicted or overpredicted the observed sequence of earthquakes, the N-test was performed. The quantile scores computed during the N-test corresponding to the forecasting of the number of earthquakes are illustrated in Figures 11a and 11b. Two threshold quantiles are plotted at 0.025 and 0.05 levels, $\delta_1$ and $\delta_2$ scores, Equations 12 and 13, were computed and plotted for the three models for the same forecasting time intervals of duration $\Delta T = 7$ days as used in Figure 10a. In addition, the results of the M-test for the three models and for the same forecasting time intervals are plotted in Figure 11c, where the quantile score $\kappa$ characterizes the consistency of the forecasted earthquake magnitudes compared to the observed ones in each forecasting time interval. The quantile scores in a case of the varying forecasting time interval $\Delta T = 1, 2, 5, 7, 10, 14$ days and fixed training time interval $T_e = 3.4284$ days are given in Figure S13.

The models were also compared among each other by applying the R- and T-tests. Two pairs of the models were considered, that is the forecasts produced by the ETAS model versus the model with the OU law and the ETAS model versus the model with the compound OU law. The results of the quantile score $\alpha$ for the R-test are plotted in Figure 12. The scores $\alpha$ were computed at the end of each forecasting time interval of duration $\Delta T$ as in Figure 10a. The corresponding sample information gain $I_{n}(A^1, A^2)$ for each pair of the models is given in Figure 13. The quantile score $\alpha$ and the information gain per earthquake in a case of the varying forecasting time interval $\Delta T = 1, 2, 5, 7, 10, 14$ days and fixed training time interval $T_e = 3.4284$ days are given in Figure S14 and S15. In both pairs of models, it was assumed that the ETAS model (with the forecast $A^2$) is the correct model to simulate the synthetic sequences of events during the forecasting time intervals.

Finally, the Bayesian $p_n$-value, Equation 16, was computed for the three models. This is plotted in Figure 14 for the varying training time intervals. Figure S16 illustrates the dependency of the $p_n$-value on the varying forecasting time interval as in Figure S14.

5. Discussion

The 2019 Ridgecrest earthquake sequence occurred in a complex network of fault structures. It generated a prominent foreshock sequence that culminated in the occurrence of the M7.1 mainshock, which was followed by a productive aftershock sequence. This complexity of the sequence was partially reflected in the frequency-magnitude statistics of foreshocks and aftershocks. It also manifested in the clustering of earthquakes in time and in space. The complex pattern of multi-segmented ruptures of the two strongest events in the sequence contributed to the assumed stress transfer pattern, which affected the distribution of subsequent triggered aftershocks.

One of the main objectives of this work was to provide a framework to compute the probabilities for the occurrence of the largest expected aftershocks during different stages of the evolution of this sequence by incorporating the preceding seismicity. This was accomplished through two main approaches. The first one was based on the assumption that the occurrence of earthquakes could be modeled as a nonhomogenous Poisson process with a specified parametric model for the earthquake rate and the frequency-magnitude distribution. Specifically, one can use the OU law (4) or the compound OU law (5) and the exponential distribution for the earthquake magnitudes, Equation 2. Then, the probabilities can be estimated from the...
EVD (8) for a specific forecasting time interval $\Delta T$ by using the point estimates of the model parameters. The second approach employed the computation of the BPD (11), which allowed to incorporate the uncertainties of the model parameters into the computation of the corresponding probabilities. This approach also requires to provide certain \textit{a priori} knowledge on the model parameters specified through the prior distributions.

**Figure 12.** Plot of the quantile score $\alpha$ (R-test) for the comparative test of the ETAS model versus the forecast based on the OU model and on the compound OU model. The scores are computed at the end of each forecasting time interval as in Figure 11.

**Figure 13.** The sample information gain for the pairs of the models. The solid squares correspond to the comparison of the forecasts based on the ETAS model versus the forecasts based on the OU model. The solid diamonds correspond to the comparison of the forecasts based on the ETAS model versus the forecast based on the compound OU model. The 95% confidence intervals are given.
The comparison of these two approaches with the combination of the three models for the earthquake rate and either including or excluding the foreshocks is illustrated in Figure 7. The results clearly illustrate that the inclusion of the foreshocks along with the earthquake rate models that favor earthquake clustering produces higher probabilities for the occurrence of the largest expected earthquakes during the specified forecasting period of $\Delta T = 7$ days.

It is interesting to note, the 2019 Ridgecrest earthquake sequence bears a striking similarity to the 2016 Kumamoto, Japan, earthquake sequence. Both sequences had a pronounced foreshock sequence which was triggered by the strong foreshocks of similar magnitudes (M6.4 vs. M6.5) and duration. They occurred on the different fault segments than the mainshock fault rupture. The $b$-values of the GR relation and $p$-values of the OU law were also smaller than the values for the aftershocks generated by the mainshocks. The mainshock magnitudes were also similar (M7.1 vs. M7.3) and had the strike-slip mechanisms. It is difficult to pin point the common stress conditions and state of faults that lead to the occurrence of both sequences but some clues may be inferred from the seismicity patterns that preceded and followed both events and can be related to the changes in the stress field (Nanjo, 2020; Nanjo et al., 2019).

To validate the three stochastic models, several statistical tests (N-, M-, R-, and T-tests) were applied retrospectively for several combinations of the training and forecasting time intervals. The results of the N-test indicate that the OU model underestimated the observed number of earthquakes for most of the forecasting time intervals. The compound OU model performed better especially in the early stages of the evolution of the sequence. The ETAS model approximated the observed number of earthquakes during the all considered forecasting time intervals, however, the ETAS model also had wider 95% spread in the number of forecasted earthquakes (Figure 10). This is the consequence of the branching nature of the ETAS process and the deviation of the distribution of the number of events from the Poisson distribution. The ETAS model was also consistent in reproducing the distribution of the magnitudes in each bin that is illustrated in Figure 11c through the $\kappa$ quantile score of the M-test.

The comparative analysis of the ETAS model versus the OU and the compound OU models also confirmed that the forecast based on the ETAS model outperformed the forecasts based on the other two models. This is illustrated in Figure 12, where the quantile score $\alpha$ from the R-test is plotted at the end of each forecasting time interval. The values of the score above the threshold level 0.025 indicate that the ETAS model outperformed the other two models. The similar conclusion is drawn from the plot (Figure 13) of the sample infor-
The results of the T-test confirmed that the ETAS model provided a statistically significant information gain with respect to the models based on the OU or compound OU rates except for the last forecasting interval ending at 38.4284 days, where the ETAS model and the model based on the compound OU rate performed similarly. For the last forecasting time interval ending at $T_e + \Delta T = 38.4284$ days, there were only three events above magnitude $m \geq 3.2$. The compound OU model produced relatively close results when computing the probabilities for the occurrence of the largest expected earthquakes (Figure 9).

One limitation of the above tests (M-, R-, T-) based on the computing of the joint log-likelihoods, Equation 14, is that they assume that the distribution of the number of earthquakes in the forecasting time interval is Poisson. This is true for the both point process models based on the OU law. However, the ETAS model deviates from the Poisson assumption. This was already demonstrated in Shcherbakov et al. (2019) when computing the Bayesian predictive distribution. Therefore, the application of these tests to the ETAS based models has to be considered approximate.

The above tests implemented in this work used the MCMC sampling of the posterior distribution of the model parameters. This allowed to incorporate the stochastic variability of the model parameters and the uncertainties associated with the prior information on the model parameters into the computation of the resulting probabilities and when performing the statistical tests. The consistency of the Bayesian predictive distribution was evaluated by estimating the Bayesian $p_B$-value, Equation 16. The values of $p_B$ within a reasonable range (say $[0.05, 0.95]$) indicate that a model is expected to reproduce a specific aspect of the data given by the test quantity $T(y)$. Whereas, the values close to 0 or 1 signify that this aspect of the data is not captured by the model. All the three models were consistent in reproducing the observed largest earthquakes in each forecasting time interval.

The analysis of the 2019, Ridgecrest earthquake sequence showed that the Bayesian predictive framework combined with the ETAS model outperformed more traditional approaches based on the Omori-Utsu type models when using the extreme value distribution to compute the probabilities for the occurrence of the largest events. The latter approach uses point estimates of the model parameters to compute the corresponding probabilities. However, large uncertainties associated with these model parameters can result in significant underestimation/overestimation of the probabilities for the largest expected events or the numbers of earthquakes above a certain magnitude during the forecasting time intervals. This is particularly evident for the Omori-Utsu law, where the productivity of the process is controlled by the $K_o$ parameter, which is typically estimated with large uncertainties (Marsan & Helmstetter, 2017; Shebalin et al., 2020). On the other hand, the Bayesian framework fully incorporates these model uncertainties into the computation of the probabilities. It also allows to account for the correlations among the model parameters. In addition, the Bayesian approach provides a flexible way of separating those uncertainties into epistemic and aleatory types (Gerstenberger et al., 2020; Kiureghian & Ditlevsen, 2009). It allows to control the epistemic uncertainties through the prior information of the model parameters and incorporates the aleatory variability of the stochastic process through the earthquake rate models and the frequency-magnitude distributions.

6. Conclusions

The 2019 Ridgecrest earthquake sequence was characterized by the complex clustering of seismicity with earthquakes occurring on a distributed fault network. It also presented a good opportunity to analyze the sequence retrospectively in order to test several statistical approaches to study the sequence in temporal and magnitude domains and to forecast the occurrence of the largest expected aftershocks during the evolution of the sequence.

Two approaches were used to compute the probabilities of having the largest expected earthquakes to be above certain magnitudes after specified time intervals and during the fixed forecasting time interval $\Delta T = 7$ days. For the first approach, the EVD (8) with the OU law (4) or the compound OU formula (5) was used. In the second approach, the Bayesian predictive distribution, Equation 11, combined with the OU law or the ETAS model (6) was used. The comparison of these approaches are illustrated in Figure 9.

Applying these two approaches to the 2019 Ridgecrest earthquake sequence revealed that the incorporation of the foreshock sequence for the subsequent computation of the probabilities to have the largest expected
after shocks above a certain magnitude was important. This was also relevant to the choice of the model to approximate the earthquake rate. Specifically, the compound OU law (5) and the ETAS model (6) provided a better approximation for the earthquake rate than the OU law (4) applied separately to the foreshock and aftershock sequences during the forecasting time intervals. These conclusions have been verified by several statistical tests. In addition, a new test based on the Bayesian $p_2$-value was implemented and applied to check the consistency of the Bayesian predictive distribution. Overall, the ETAS model passed the tests most of the time and was successful in reproducing the observed number of earthquakes and the distribution of magnitudes. As a result, the computed probabilities using the Bayesian predictive distribution (Figure 8) for the largest expected earthquake during the evolution of the 2019 Ridgecrest sequence can be considered accurate.

**Data Availability Statement**

The Southern California Seismic Network database, SCSN (2020), [https://service.scegd.caltech.edu/eq-catalogs/date_mag_loc.php](https://service.scegd.caltech.edu/eq-catalogs/date_mag_loc.php), was used to download the seismic catalog (last accessed on December 1, 2020). U.S. Geological Survey and California Geological Survey quaternary fault and fold database for the United States, USGS (2006), was downloaded from the USGS web site: [https://earthquake.usgs.gov/hazards/qfaults/](https://earthquake.usgs.gov/hazards/qfaults/)(last accessed on June 1, 2020). The data analysis was performed using computer scripts written in Matlab and can be requested from the author. The supporting information for this article includes Tables S1–S3 with the parameters of the Gamma distribution, which was used as a prior distribution for the parameters of the three models considered in the work. It also includes plots illustrating the fit of the compound OU (Figure S1) and the ETAS (Figure S2) models. The MCMC sampling of the model parameters for the OU (Figure S4 and S5), the compound OU (Figures S6–S8), the ETAS (Figures S9–S11) models are provided for one specific training and forecasting time intervals. The forecast evolution during 330 days after the occurrence of the M7.1 mainshock is given in Figure S12. The additional quantile scores of the plots are given in Figures S13–S16.

**Acknowledgments**

The constructive comments by two anonymous reviewers helped to improve the manuscript. This work has been supported by the NSERC Discovery grant.

**References**

Barnhart, W. D., Hayes, G. P., & Gold, R. D. (2019). The July 2019 Ridgecrest, California, earthquake sequence: Kinematics of slip and stressing in cross-fault ruptures. *Geophysical Research Letters*, 46, 11859–11867. [https://doi.org/10.1029/2019GL084741](https://doi.org/10.1029/2019GL084741)

Båth, M. (1965). Lateral in homogeneities of the upper mantle. *Tectonophysics*, 2, 483–514.

Bender, B. (1983). Maximum likelihood estimation of $b$ values for magnitude grouped data. *Bulletin of the Seismological Society of America*, 73(3), 831–851.

Campbell, K. W. (1982). Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model. *Bulletin of the Seismological Society of America*, 72(5), 1649–1705.

Cattania, C., Werner, M. J., Marzocchi, W., Hainzl, S., Rhoades, D., Gerstenberger, M., et al. (2018). The forecasting skill of physics-based seismicity models during the 2010–2012 Canterbury, New Zealand earthquake sequence. *Seismological Research Letters*, 89(4), 1238–1250. [https://doi.org/10.1785/0220180033](https://doi.org/10.1785/0220180033)

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London: Springer.

Console, R., Lombardi, A. M., Murru, M., & Rhoades, D. (2003). Bhil’s law and the self-similarity of earthquakes. *Journal of Geophysical Research*, 108(B2), 2128. [https://doi.org/10.1029/2003JB001651](https://doi.org/10.1029/2003JB001651)

Daley, D. J., & Vere-Jones, D. (2003). An introduction to the theory of point processes (1, 2nd ed.). New York, NY: Springer.

Ebrahimian, H., & Jalayer, F. (2017). Robust seismicity forecasting based on Bayesian parameter estimation for epidemiological spatio-temporal aftershock clustering models. *Scientific Reports*, 7, 8903. [https://doi.org/10.1038/s41598-017-09962-z](https://doi.org/10.1038/s41598-017-09962-z)

Ebrahimian, H., Jalayer, F., Asproune, D., Lombardi, A. M., Marzocchi, W., Prota, A., & Manfredi, G. (2014). Adaptive daily forecasting of seismic aftershock hazard. *Bulletin of the Seismological Society of America*, 104(1), 145–161. [https://doi.org/10.1785/0120130040](https://doi.org/10.1785/0120130040)

Felzer, K. R., Abercrombie, R. E., & Ekström, G. (2004). A common origin for aftershocks, foreshocks, and multiplets. *Bulletin of the Seismological Society of America*, 94(1), 88–98. [https://doi.org/10.1785/0120030069](https://doi.org/10.1785/0120030069)

Gelman, A., Carlin, J. Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: CRC Press.

Gerstenberger, M. C., Marzocchi, W., Allen, T., Pagani, M., Adams, J., Danciu, L., et al. (2020). Probabilistic seismic hazard analysis at regional and national scale: State of the art and future challenges. *Reviews of Geophysics*, 58(2), e2019RG000653. [https://doi.org/10.1029/2019RG000653](https://doi.org/10.1029/2019RG000653)

Gerstenberger, M. C., Wiemer, S., Jones, L. M., & Reasenberg, P. A. (2005). Real-time forecasts of tomorrow’s earthquakes in California. *Nature*, 435(7040), 328–331. [https://doi.org/10.1038/Nature03622](https://doi.org/10.1038/Nature03622)

Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34, 185–188.

Hainzl, S. (2016a). Apparent triggering function of aftershocks resulting from rate-dependent incompleteness of earthquake catalogs. *Journal of Geophysical Research*, 121(9), 6499–6509. [https://doi.org/10.1002/2016JB013319](https://doi.org/10.1002/2016JB013319)

Hainzl, S. (2016b). Rate-dependent incompleteness of earthquake catalogs. *Seismological Research Letters*, 87(2), 337–344. [https://doi.org/10.1785/0220150211](https://doi.org/10.1785/0220150211)

Hardebeck, J. L. (2020). A stress-similarity triggering model for aftershocks of the M6.4 and 7.1 Ridgecrest earthquakes. *Bulletin of the Seismological Society of America*, 110(4), 1716–1727. [https://doi.org/10.1785/0120200015](https://doi.org/10.1785/0120200015)
Rhoades, D. A., Schorlemmer, D., Gerstenberger, M. C., Christophersen, A., Zechar, J. D., & Imoto, M. (2011). Efficient testing of earthquake forecasting models. *Acta Geophysica*, 59(4), 728–747. https://doi.org/10.2478/s11600-011-0013-5
Ross, Z. E., Ildini, B., Jia, Z., Stephenson, O. L., Zhong, M. Y., Wang, X., et al. (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence. *Science*, 366(6463), 346–351. https://doi.org/10.1126/science.aaz3109
Savran, W. H., Werner, M. J., Marzocchi, W., Rhoades, D. A., Jackson, D. D., Milner, K., et al. (2020). Pseudosynoptic evaluation of UCERF3-ETAS forecasts during the 2019 Ridgecrest sequence. *Bulletin of the Seismological Society of America*, 110(4), 1799–1817. https://doi.org/10.1785/0120200026
Schorlemmer, D., Gerstenberger, M. C., Wiemer, S., Jackson, D. D., & Rhoades, D. A. (2007). Earthquake likelihood model testing. *Seismological Research Letters*, 78(1), 17–29. https://doi.org/10.1785/1100605078.1.17
Schorlemmer, D., Werner, M. J., Marzocchi, W., Jordan, T. H., Ogata, Y., Jackson, D. D., et al. (2018). The collaborative study for the earthquake predictability: Achievements and priorities. *Seismological Research Letters*, 89(4), 1305–1313. https://doi.org/10.1785/0220180053
SCSN (2020). *Southern California seismic network*. https://doi.org/10.7909/C3WD3xH1. https://service.scedc.caltech.edu/eq-catalogs/date_mag_loc.php
Shcherbakov, R., Goda, K., Ivianan, A., & Atkinson, G. M. (2013). Aftershock statistics of major subduction earthquakes. *Bulletin of the Seismological Society of America*, 103(6), 3222–3234. https://doi.org/10.1785/0120120337
Shcherbakov, R., Nguyen, M., & Quigley, M. (2012). Statistical analysis of the 2010 Mw 7.1 Darfield earthquake aftershock sequence. *New Zealand Journal of Geology and Geophysics*, 55(3), 305–311. https://doi.org/10.1080/00288306.2012.676556
Shcherbakov, R., & Turcotte, D. L. (2004). A modified form of Båth's law. *Bulletin of the Seismological Society of America*, 94(5), 1968–1975. https://doi.org/10.1785/012003162
Shcherbakov, R., Turcotte, D. L., & Rundle, J. B. (2004). A generalized Omori's law for earthquake aftershock decay. *Geophysical Research Letters*, 31(11), L11613. https://doi.org/10.1029/2004GL019808
Shcherbakov, R., Turcotte, D. L., & Rundle, J. B. (2005). Aftershock statistics. *Pure and Applied Geophysics*, 162(6–7), 1051–1076. https://doi.org/10.1007/s00024-004-2661-8
Shcherbakov, R., Turcotte, D. L., & Rundle, J. B. (2006). Scaling properties of the Parkfield aftershock sequence. *Bulletin of the Seismological Society of America*, 96(4B), S376–S384. https://doi.org/10.1785/1100605015
Shcherbakov, R., Turcotte, D. L., & Rundle, J. B. (2015). Complexity and earthquakes. In H. Kanamori (Ed.), Earthquake seismology (4, 2nd ed., pp. 627–653) Elsevier.
Shcherbakov, R., Yakovlev, G., Turcotte, D. L., & Rundle, J. B. (2005). Model for the distribution of aftershock interoccurrence times. *Physical Review Letters*, 95, 218501. https://doi.org/10.1103/PhysRevLett.95.218501
Shcherbakov, R., Zhuang, J., & Ogata, Y. (2018). Constraining the magnitude of the largest event in a foreshock-mainshock-aftershock sequence. *Geophysical Journal International*, 212(1), 1–13. https://doi.org/10.1093/gji/ggy407
Shcherbakov, R., Zhuang, J., Zöller, G., & Ogata, Y. (2019). Forecasting the magnitude of the largest expected earthquake. *Nature Communications*, 10, 4051. https://doi.org/10.1038/s41467-019-11958-4, Art.
Shearer, P. M. (2012). Self-similar earthquake triggering, Båth's law, and foreshock/aftershock magnitudes: Simulations, theory, and results for southern California. *Journal of Geophysical Research*, 117, B06S10. https://doi.org/10.1029/2011JB008957
Shebalin, P., Narteau, C., & Baranov, S. V. (2020). Earthquake productivity law. *Geophysical Journal International*, 222(2), 1264–1269. https://doi.org/10.1093/gji/ggaa252
Shebalin, P., Narteau, C., Holcschnieder, M., & Schorlemmer, D. (2011). Short-term earthquake forecasting using early aftershock statistics. *Bulletin of the Seismological Society of America*, 101(11), 297–312. https://doi.org/10.1785/0220100119.
Tahiri, M., Grasso, J. R., & Amorese, D. (2012). The largest aftershock: How strong, how far away, how delayed?. *Geophysical Research Letters*, 39, L04301. https://doi.org/10.1029/2011gl050604
Taroni, M., Marzocchi, W., Schorlemmer, D., Werner, M. J., Wiemer, S., Zechar, J. D., & Euchner, F. (2018). Prospective CSEP evaluation of 1-day, 3-month, and 5-year earthquake forecasts for Italy. *Seismological Research Letters*, 89(4), 1251–1261. https://doi.org/10.1785/0220180031
Tinti, S., & Mulargia, F. (1987). Confidence intervals of b-values for grouped magnitudes. *Bulletin of the Seismological Society of America*, 77(6), 2125–2134.
Toda, S., & Stein, R. S. (2020). Long- and short-term stress interaction of the 2019 Ridgecrest sequence and Coulomb-based earthquake forecasts. *Bulletin of the Seismological Society of America*, 110(4), 1765–1780. https://doi.org/10.1785/1100605016
USGS (2006). *U.S. Geological Survey and California Geological Survey quaternary fault and fold database* (last accessed: 2020-07-01). https://earthquake.usgs.gov/hazards/igfaults/
Utsu, T. (1961). A statistical study on the occurrence of aftershocks. *Geographical Magazine*, 30, 521–605.
Utsu, T., Ogata, Y., & Matsusura, R. S. (1995). The centenary of the Omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth*, 43(1), 1–33.
Vere-Jones, D. (1975). Stochastic models for earthquake sequences. *Geophysical Journal International*, 42(2), 811–826. https://doi.org/10.1111/j.1365-246X.1975.tb05893.x
Vere-Jones, D. (2010). Foundations of statistical seismology. *Pure and Applied Geophysics*, 167(6–7), 645–653. https://doi.org/10.1007/s00024-010-0079-z
Werner, M. J., Helmstetter, A., Jackson, D. D., & Kagan, Y. Y. (2011). High-resolution long-term and short-term earthquake forecasts for California. *Bulletin of the Seismological Society of America*, 101(4), 1630–1648. https://doi.org/10.1785/0220090340
Zechar, J. D., Gerstenberger, M. C., & Rhoades, D. A. (2010). Likelihood-based tests for evaluating space-rate-magnitude earthquake forecasts. *Bulletin of the Seismological Society of America*, 100(3), 1184–1195. https://doi.org/10.1785/0220090192
Zöller, G., Holcschnieder, M., & Hainzl, S. (2013). The maximum earthquake magnitude in a time horizon: Theory and case studies. *Bulletin of the Seismological Society of America*, 103(2A), 860–875. https://doi.org/10.1785/10.1785/0220190113