

Combining a Locally Restricted Anisotropic Spatial Kernel with an ETAS Incomplete Model for Better Forecasts of the 2019 Ridgecrest Sequence

Christian Grimm (✉ Christian.Grimm@stat.uni-muenchen.de)

Ludwig Maximilians University Munich: Ludwig-Maximilians-Universität München

<https://orcid.org/0000-0002-2190-2981>

Sebastian Hainzl

GFZ: Deutsches Geoforschungszentrum Potsdam

Martin Käser

University of Munich: Ludwig-Maximilians-Universität München

Helmut Küchenhoff

Ludwig Maximilians University Munich: Ludwig-Maximilians-Universität München

Research Article

Keywords: ETAS, short-term incompleteness, anisotropic spatial kernel, Ridgecrest

Posted Date: December 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1128731/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 Combining a Locally Restricted Anisotropic Spatial Kernel with an 2 ETAS Incomplete Model for Better Forecasts of the 2019 Ridgecrest 3 Sequence

4 **Christian Grimm^a, Sebastian Hainzl^b, Martin Käser^{c,d}, Helmut Küchenhoff^a**

5 ^aLudwig-Maximilians-University Munich, Department of Statistics, Ludwigstraße 33, 80539 Munich, Germany,
6 christian.grimm@stat.uni-muenchen.de, <https://orcid.org/0000-0002-2190-2981> (CG);

7 kuechenhoff@stat.uni-muenchen.de (HK)

8 ^bGFZ German Research Centre for Geoscience, Physics of Earthquakes and Volcanoes, Helmholtzstraße 6/7, 14467
9 Potsdam, Germany; hainzl@gfz-potsdam.de, <https://orcid.org/0000-0002-2875-0933> (SH)

10 ^cLudwig-Maximilians-University Munich, Department of Earth and Environmental Sciences, Geophysics,
11 Theresienstraße 41, 80333 Munich, Germany; martin.kaeser@geophysik.uni-muenchen.de (MK)

12 ^dalso at Munich Re, Section GeoRisks, Königinstr. 107, 80802 Munich, Germany

13 **Abstract.** Strong earthquakes cause aftershock sequences that are clustered in time according to a power decay
14 law, and in space along their extended rupture, shaping a typically elongate pattern of aftershock locations. A widely
15 used approach to model seismic clustering is the *Epidemic Type Aftershock Sequence* (ETAS) model, that shows three
16 major biases: First, the conventional ETAS approach assumes isotropic spatial triggering, which stands in conflict
17 with observations and geophysical arguments for strong earthquakes. Second, the spatial kernel has unlimited extent,
18 allowing smaller events to exert disproportionate trigger potential over an unrealistically large area. Third, the ETAS
19 model assumes complete event records and neglects inevitable short-term aftershock incompleteness as a consequence
20 of overlapping coda waves. These three effects can substantially bias the parameter estimation and particularly lead
21 to underestimated cluster sizes. In this article, we combine the approach of [Grimm et al. \(2021\)](#), which introduced a
22 generalized anisotropic and locally restricted spatial kernel, with the *ETAS-Incomplete* (ETASI) time model of [Hainzl](#)
23 [\(2021\)](#), to define an ETASI space-time model with flexible spatial kernel that solves the abovementioned shortcomings.
24 We apply different model versions to a triad of forecasting experiments of the 2019 Ridgecrest sequence, and evaluate
25 the prediction quality with respect to cluster size, largest aftershock magnitude and spatial distribution. The new model
26 provides the potential of more realistic simulations of on-going aftershock activity, e.g. allowing better predictions of
27 the probability and location of a strong, damaging aftershock, which might be beneficial for short term risk assessment
28 and disaster response.

29 **Keywords:** ETAS, short-term incompleteness, anisotropic spatial kernel, Ridgecrest.

30 **Main author contact information:** Christian.Grimm@stat.uni-muenchen.de

31 **1 Introduction**

32 Strong earthquakes are usually observed to cause a pronounced spatio-temporal pattern of after-
33 shocks. More precisely, according to the Omori-Utsu Law (Utsu et al., 1995), the temporal after-
34 shock rate is subject to a power law decrease with time $t - t_{main}$ after the main triggering event,
35 that is,

$$g(t - t_{main}) = (t - t_{main} + c)^{-p} \quad (1)$$

36 with the delay parameter $c > 0$ (usually a few minutes to hours) and the exponent p (usually in the
37 range between 0.8 – 1.2). It means that the temporal pattern of aftershocks is dominated by events
38 occurring within short time after the mainshock. Fig. 1(a) demonstrates this temporal behavior
39 for the Ridgecrest sequence in California, which produced an M6.4 foreshock on July 4, 2019,
40 followed by an M7.1 mainshock within 34 hours on July 6, 2019.

41

42 The observed spatial patterns of aftershock sequences stem from their tendency to occur on or
43 close to the mainshock rupture plane (Marsan and Lengliné, 2008). The larger the length-to-width
44 ratio of this plane gets, the more elongate the typical aftershock region becomes. In addition, a
45 higher dip angle reduces the width of the 3D-to-2D projection of the rupture plain to the earth's
46 surface and therefore results in a scatter of two-dimensional aftershock epicenters that can be in-
47 creasingly well approximated by a line segment.

48

49 The prevailing continental tectonic regime in southern California with typically steep, strike-
50 slip faulting favors such elongated aftershock patterns in this region. With the exception of the
51 M6.7 1994 Northridge earthquake, all of the most prominent mainshock-aftershock sequences of
52 the last 40 years (M6.6 1987 Superstition Hill, M7.3 1992 Landers, M7.1 1999 Hector Mine, M7.2
53 2010 Baja California, M7.1 2019 Ridgecrest) demonstrate distinct linearly elongate scattering of
54 aftershock locations (Hainzl, 2021).

55

56 In this context, the Ridgecrest sequence is a special case as the M6.4 foreshock simultaneously
57 ruptured two almost orthogonal faults, leading to a double pattern of separate linearly elongate af-
58 tershock clouds (Marsan and Ross, 2021). Fig. 1(b) shows that the triggering M6.4 event (yellow
59 pentagram) is located close to the intersection of the two ruptured faults. In contrast, the M7.1
60 mainshock (yellow hexagram) ruptured only one fault which appears to be the extension of one of
61 the faults activated by the foreshock.

62

63 Analyzing and forecasting clustered seismicity is an established discipline in seismological re-
64 search. Its goal is to understand the evolution of large aftershock sequences and to predict their size,
65 largest aftershock magnitude, spatial distribution etc. A prominent approach to model clustered
66 seismicity is the so-called *Epidemic Type Aftershock Sequence (ETAS)* model, which describes
67 earthquake records as a superposition of independent background seismicity and triggered earth-
68 quake sequences (Ogata, 1988, 1998). The earthquake triggering component is designed in terms
69 of a branching process and characterized by the triad of (1) trigger-magnitude dependent aftershock
70 productivity, (2) a temporal distribution of aftershock times typically derived from the Omori Law
71 (see Equation 1), and (3) an usually isotropic spatial distribution of aftershock locations (e.g. Jalil-

72 [ian, 2019; Zhuang et al., 2002](#)). Particularly, the aftershock productivity (i.e. expected number of
73 offsprings) for a trigger event with magnitude m is

$$k_{A,\alpha}(m) = A \exp(\alpha (m - M_c)), \quad (2)$$

74 where parameters $A > 0$ and $\alpha > 0$ control the exponential growth of the trigger potential and M_c
75 is the cut-off magnitude of the analyzed earthquake catalog.

76

77 Despite generally producing successful and insightful estimation and forecast results, ETAS
78 models are known to be limited by a number of potential biases. In this article, we present an
79 approach that combines solutions for three main short-comings of the conventional ETAS model,
80 (1) the isotropic spatial aftershock distribution, (2) the infinite extent of the spatial kernel and (3)
81 the short-term incompleteness of earthquake records after strong triggering events.

82 *Bias 1: Isotropic spatial distribution*

83 The common assumption in ETAS models is that spatial aftershock locations are distributed isotrop-
84 ically around the triggering event. It is named as a shortcoming in many publications because it
85 stands in conflict with the abovementioned observation that aftershocks tend to occur close to the
86 (elongate) rupture plane of the triggering event ([Hainzl et al., 2008, 2013](#); [Ogata, 1998, 2011](#);
87 [Ogata and Zhuang, 2006](#); [Seif et al., 2017](#); [Zhang et al., 2018](#)). The assumption of isotropy is
88 reasonably valid for weak earthquakes with small rupture extensions, but becomes problematic for
89 larger magnitudes, e.g. see the spatial pattern of the Ridgecrest sequence in Fig. 1(b). It has been
90 shown that inadequate spatial models can lead to an underestimation of the productivity parame-

91 ter α (Equation 2) because the numerous small events take over the role of mimicking the "true"
92 anisotropic distribution (Grimm et al., 2021; Hainzl et al., 2008, 2013).

93 *Bias 2: Infinite Spatial Extent*

94 Under the common definition of an infinite spatial kernel, aftershock triggering is disproportion-
95 ately associated with the more numerous small events, that can more flexibly mimic anisotropic
96 event alignments than the few strong mainshocks. This can lead to unrealistically far trigger impact
97 of small magnitudes and to a substantial underestimation of the direct aftershock productivity of
98 strong events, resulting in a smoothing of temporal event distributions (Grimm et al., 2021).

99 *Bias 3: Short-term aftershock incompleteness (STAI)*

100 Strong earthquakes typically cause incomplete aftershock records immediately after their occur-
101 rence, mainly due to an overlap of coda waves (de Arcangelis et al., 2018; Hainzl, 2021). Fig. 1(c)
102 and (d) confirms this phenomenon for the aftershock sequences of the M6.4 and M7.1 Ridgecrest
103 events, respectively. Apparently, records of smaller sized aftershocks are missing in the first min-
104 utes to hours, somewhat foiling the power law decay of event rates expected from the Omori-Utsu
105 law (Equation 1). The short-term incomplete event records can therefore hide to a large extent
106 both the "true" Omori Law decay (Equation 1) and the "true" aftershock productivity of the trigger
107 event (Equation 2) and lead to an overestimation of Omori parameter c and an underestimation of
108 productivity parameter α (Hainzl, 2016a,b; Page et al., 2016; Seif et al., 2017).

109

110 Data-driven uncertainties of event locations and cut-off magnitude as well as the assumption of
111 a *time-invariant seismic background* may lead to further inaccuracies in the parameter estimation

112 ([Harte, 2013, 2016](#); [Seif et al., 2017](#)), However, they can be neglected in our study because they
113 are either expected to be small in southern California datasets (e.g. location and magnitude uncer-
114 tainty) or do not apply in an isolated sequence analysis (background miss-specification).

115

116 *Scope of this article*

117 In this article, we combine an ETAS approach accounting for short-term incomplete event records
118 with the application of a generalized, anisotropic spatial model that restricts the spatial kernel to
119 the local surrounding of the trigger source. We demonstrate the functionality and superiority of
120 our approaches over the conventional, isotropic ETAS model by means of forecasting experiments
121 for the Ridgecrest sequence.

122

123 We utilize the generalized anisotropic and locally restricted spatial kernel suggested by [Grimm](#)
124 [et al. \(2021\)](#), which assumes uniform trigger density along an estimated rupture line segment, with
125 power-law decay to the sides and at the end points of the rupture. [Zhang et al. \(2018\)](#) pursued an
126 even more detailed approach, which assumed constant trigger rate in the entire rupture plane, with
127 power-law decay outside of it. Different versions of elliptic Gaussian distributions were introduced
128 and discussed by [Ogata \(1998, 2011\)](#) and [Ogata and Zhuang \(2006\)](#). The latter approaches suc-
129 cessfully modeled spatial aftershock patterns, however, they require a new set of parameters and
130 are therefore not flexibly combinable with the conventional, isotropic functionality. In contrast, the
131 kernel of [Grimm et al. \(2021\)](#) represents a generalization of the isotropic function and therefore
132 allows simultaneous anisotropic modeling of some events (e.g. above a certain magnitude thresh-
133 old) and isotropic modeling of the rest. In order to address the abovementioned particularity of the

134 M6.4 Ridgecrest foreshock, rupturing two almost orthogonal faults, we further generalize the ap-
135 proach by allowing a spatial kernel composed by a weighted superposition of two distinct rupture
136 line segments.

137

138 Additionally, we account for STAI by applying an ETAS model version that incorporates rate-
139 dependent incompleteness of event records. Recognizing alternative approaches that will be briefly
140 described in the *Methods* section, we choose for the *ETAS-Incomplete (ETASI)* model as recently
141 suggested by [Hainzl \(2021\)](#). For simplicity and to sharpen its focus on the incompleteness de-
142 tection, [Hainzl \(2021\)](#) neglected the space dimension in his model. As this article combines the
143 ETASI time model of [Hainzl \(2021\)](#) with an adequate, anisotropic spatial kernel it can be seen as
144 the space-including extension of the latter. The focus of this study, however, is on the benefit of
145 modeling the spatial aftershock distribution by a generalized anisotropic spatial kernel, rather than
146 the benefit of the ETASI model.

147

148 This article is structured as follows. In the *Methods* section, we introduce the conventional
149 ETAS model and its ETASI extension and define the anisotropic, locally restricted spatial kernel.
150 This section includes a description of the estimation procedures for strikes and rupture positions
151 and the spatial integral over anisotropic kernels. Next, the *Application* section explains the three
152 forecasting experiments, introducing the data and time-space windows for the parameter estimation
153 and forward simulations. Finally, we interpret and discuss our forecasting results and draw our
154 conclusions.

155 2 Methods

156 The ETAS model, first introduced by [Ogata \(1988, 1998\)](#), is a branching-tree type model which de-
157 scribes clustered earthquake occurrences by consecutive triggering evolving over multiple parent-
158 child generations (i.e. allowing secondary aftershocks). The triggered seismicity is overlaying a
159 time-invariant background process.

160

161 In this section, we will first introduce the conventional, isotropic ETAS model approach. Next,
162 we will extend the model to obtain a time-space version of the ETASI model suggested by [Hainzl](#)
163 [\(2021\)](#), which involves STAI into parameter estimation. Mostly, notations are consistent with
164 [Hainzl \(2021\)](#). We will then define the anisotropic generalization of the spatial kernel, which is
165 compatible with both the ETAS and ETASI model, and introduce the local restriction of the kernel.
166 Finally, we explain the fitting algorithm for the strike angle and rupture position of anisotropic
167 trigger events and the methods for spatial integral estimation.

168 2.1 ETAS-Model

169 In the conventional ETAS model approach, the occurrence rate of an earthquake with *magnitude*
170 *m*, occurring at *time t* and at *location (x, y)* is modeled by an inhomogeneous Poisson process with
171 discontinuous intensity function

$$\lambda(t, x, y, m) = f_0(m) R_0(t, x, y) \quad (3)$$

172 where

$$f_0(m) = \beta e^{-\beta(m-M_c)} \quad (4)$$

173 is the "true" probability density function (pdf) of the frequency-magnitude distribution (FMD) with
 174 Gutenberg-Richter parameter $b = \beta/\ln(10)$ (Gutenberg and Richter, 1944), and

$$R_0(t, x, y) = \mu u(x, y) + \sum_{i:t_i < t} k_{A,\alpha}(m_i) g_{c,p}(t - t_i) h_{D,\gamma,q}(r_i(x, y), m_i, l_i) \quad (5)$$

175 is the "true" occurrence rate of events with magnitude $m \geq M_c$, at time t and at location (x, y) .
 176 The "true" event rate is modeled by a superposition of the time-invariant *seismic background*
 177 *rate* $\mu u(x, y)$ with parameter $\mu > 0$ and a sum of the trigger rate contributions of all events i
 178 that occurred prior to current time t . $k_{A,\alpha}(m_i)$ and $g_{c,p}(t - t_i)$ denote the aftershock produc-
 179 tivity and Omori-Utsu Law decay functions as defined in Equations (1) and (2), respectively.
 180 $h_{D,\gamma,q}(r_i(x, y), m_i, l_i)$ models distribution of aftershock locations triggered by event i , with pa-
 181 rameters D, γ and q . The precise inputs and shape of the spatial kernel are discussed later.

182

183 The term "true" means that the (physical) relationships are expected to be observed with per-
 184 fect earthquake records. The main assumption of the conventional ETAS model is that STAI does
 185 not significantly distort the "true" magnitude distribution and the "true" event rates.

186

187 2.2 ETASI Model

188 *Rate-Dependent Incompleteness*

189 The concept of rate-dependent earthquake record incompleteness assumes that the "true" rela-
 190 tionships underlying $f_0(m)$ and $R_0(t, x, y)$ are not accurately identifiable in available earthquake
 191 catalogs because especially events with small magnitudes are detected with lower probability in

192 periods of high seismic activity. In these periods, the detection ability is limited typically due to
193 overlapping seismic waves (Hainzl, 2016a, 2021).

194 Fitting the "true" relationships to incomplete data records may therefore lead to significantly bi-
195 ased parameter estimates (Hainzl, 2016a,b, 2021; Page et al., 2016; Seif et al., 2017).

196

197 In recent years, there has been growing research interest in how to account for short-term
198 incomplete datasets. For instance, Zhuang et al. (2017) developed a replenishment algorithm to
199 fill up likely incomplete time intervals by simulated events, in order to obtain artificially complete
200 pseudo-records. Other authors, particularly mentionable Omi et al. (2013, 2014), Lippiello et al.
201 (2016), de Arcangelis et al. (2018), Mizrahi et al. (2021) and Hainzl (2021), tried to incorporate
202 STAI directly into the ETAS model fit. A rather simple workaround approach is to remove likely
203 incomplete time periods from the fitted time interval using empirical completeness functions, such
204 as performed in Hainzl et al. (2013) and Grimm et al. (2021). A comprehensive discussion and
205 comparison of various ETASI models is not in the scope of this article. The choice for the ETASI
206 model proposed by Hainzl (2021) was made for rather practical reasons, mainly because of its
207 compatibility with existing code.

208 2.2.1 Model Formulation

209 The working assumption of the ETASI model described here is that an earthquake with magnitude
210 m and occurring at time t can only be detected by the operating seismic network if no event of
211 equal or larger magnitude occurred within the blind time $[t - T_b, t]$, where T_b is typically in the
212 range of some seconds to few minutes (Hainzl, 2021). Similar assumptions have formerly been
213 formulated by Lippiello et al. (2016), de Arcangelis et al. (2018) and Hainzl (2016a).

214

215 Let $N_0(t)$ be the expected number of events occurring within the entire spatial window S during
 216 blind time $[t - T_b, t]$,

$$N_0(t) = \int_{t-T_b}^t \iint_S R_0(t, x, y) dx dy dt \approx T_b \iint_S R_0(t, x, y) dx dy,$$

217 where the approximation holds under the assumption that event rates are approximately constant
 218 during the blind time (Hainzl, 2021). According to the "true" FMD (Equation 4), each of the $N_0(t)$
 219 events has a probability of $e^{-\beta(m-M_c)}$ to exceed magnitude m . We define the detection probability
 220 $p_d(m, t)$ of an earthquake at time t with magnitude m as the probability that no equal or larger
 221 event occurred during blind time T_b , i.e.

$$p_d(m, t) = e^{-N_0(t) e^{-\beta(m-M_c)}}.$$

222 Following the derivations in Hainzl (2016b, 2021), we obtain the "apparent", incompleteness-
 223 biased FMD

$$f(m, t) := f_0(m) N_0(t) \frac{p_d(m, t)}{1 - e^{-N_0(t)}}$$

224 and the "apparent" event rate

$$R(t, x, y) := \frac{R_0(t, x, y)}{N_0(t)} (1 - e^{-N_0(t)}).$$

225 The term "apparent" signals that the functions f and R do not represent the "true", but
 226 the observable relationships that are possibly distorted by rate-dependent record incompleteness.

227 In periods of high seismic activity, the "apparent" FMD exhibits a larger relative frequency of
 228 strong events (because they are more likely to be detected) and an event rate lowered by detection
 229 capacity. We obtain the ETASI intensity function

$$\lambda(t, x, y, m) = f(m, t) R(t, x, y) = f_0(m) R_0(t, x, y) p_d(m, t)$$

230 The two underlying, simplifying assumptions in the ETASI model are that (1) the blind time T_b
 231 is magnitude-independent, which [Hainzl \(2021\)](#) justifies by typically shorter source durations than
 232 travel times of coda waves, and (2) that the seismic network is equally occupied for blind time T_b
 233 by any event in the entire investigated spatial window. The second assumption is reasonable for a
 234 small spatial window, e.g. when analyzing an isolated sequence. When fitting the ETASI model
 235 over a larger region, it needs to be assured that not two relevant clusters evolve at the same time but
 236 at distinct locations as they would be assumed to simultaneously occupy the entire seismic network.
 237 A reasonable approach to prevent undesired biases is to choose a larger cut-off magnitude.

238 *Log-Likelihood Optimization*

239 The parameter vector $\theta = \{\mu, A, \alpha, c, p, D, \gamma, q, \beta, T_b\}$ of the ETASI model is estimated by maxi-
 240 mizing its log-likelihood function $LL = LL_1 - LL_2$ with

$$\begin{aligned} LL_1 &= \sum_{\text{events } j} \ln (f_0(m_j) R_0(t_j, x_j, y_j) p_d(m_j, t)), \\ LL_2 &= \int_{M_c}^{\infty} \int_{T_1}^{T_2} \iint_S \lambda(t, x, y, m) dx dy dt dm \\ &\approx \frac{T_2 - T_1}{T_b} - \frac{1}{T_b} \int_{T_1}^{T_2} e^{-T_b} \iint_S R_0(t, x, y) dx dy dt \end{aligned} \quad (6)$$

241 where the sum in LL_1 goes over all *target events* in the magnitude-time-space window $[M_c, \infty) \times$
 242 $[T_1, T_2] \times S$ and LL_2 integrates over this model space. In our study we optimized the parameter vec-
 243 tor θ using the gradient-based Davidson-Fletcher-Powell algorithm (Jalilian, 2019; Ogata, 1998;
 244 Zhuang et al., 2002).

245 2.3 Generalized Anisotropic Spatial Kernel

246 Conventional Isotropic Kernel

The *spatial kernel* $h_{D,\gamma,q}(r_i, m_i, l_i)$ in Equation (5) models the 2D-distribution of aftershocks lo-
 cations. In conventional ETAS model approaches, the triggering event is assumed to be a point
 source, distributing its offsprings isotropically around its epicenter. A classical definition of an
 isotropic kernel (see Grimm et al., 2021; Jalilian, 2019; Ogata, 1998) is

$$h_{D,\gamma,q}^{iso}(r_i(x, y), m_i) := \frac{q - 1}{D \exp(\gamma(m_i - M_c))} \left(1 + \frac{\pi r_i(x, y)^2}{D \exp(\gamma(m_i - M_c))} \right)^{-q} \quad (7)$$

247 where $r_i(x, y)$ denotes the point-to-point distance between a potential aftershock location (x, y)
 248 and the coordinates (x_i, y_i) of the triggering event i , and m_i is the magnitude of the event i . The
 249 kernel is constrained by the parameters D and γ that control the magnitude-dependent width of the
 250 kernel, and parameter q that describes the exponential decay of the function with growing spatial
 251 distance.

Here we use the anisotropic generalization of the spatial kernel that was first introduced by [Grimm et al. \(2021\)](#),

$$h_{D,\gamma,q}(r_i(x,y), m_i, l_i) := \frac{q-1}{D \exp(\gamma(m_i - M_c))} \left(1 + \frac{2l_i r_i(x,y) + \pi r_i(x,y)^2}{D \exp(\gamma(m_i - M_c))} \right)^{-q} \quad (8)$$

253 where l_i denotes the rupture length estimate for the triggering event i . In this spatial model, the
 254 distance term $r_i(x,y)$ denotes the point-to-line distance between the potential aftershock location
 255 (x,y) and the estimated rupture segment of event i . Note that

$$h_{D,\gamma,q}(r_i(x,y), m_i, 0) = h_{D,\gamma,q}^{iso}(r_i(x,y), m_i),$$

256 i.e. the anisotropic kernel is a generalization and collapses to the isotropic model if the triggering
 257 location is assumed to be a point source with rupture extension $l_i = 0$. Therefore, the generalized
 258 spatial model can be used for mixing approaches of both kernels, e.g. applying anisotropy to events
 259 i with magnitudes $m_i \geq M_{aniso}$:

$$l_i = \begin{cases} 0, & \text{for } m_i < M_{aniso}, \quad (\text{isotropic trigger}) \\ 10^{-2.57+0.62m_i}, & \text{for } m_i \geq M_{aniso}, \quad (\text{anisotropic trigger}) \end{cases} \quad (9)$$

260 The scaling relationship for anisotropic events is taken from the estimate of subsurface rupture
 261 lengths for strike-slip faulting events, provided in [Wells and Coppersmith \(1994\)](#).

262 *Local Spatial Restriction*

263 Both the conventional isotropic and the generalized anisotropic kernels are defined in terms of a
 264 probability density function (pdf) over infinite space. Realistically, however, small earthquakes
 265 should exert only a locally restricted trigger influence. [Grimm et al. \(2021\)](#) showed that an infinite
 266 spatial extent may lead to an underestimation of the aftershock productivity parameter α because
 267 it overestimates the triggering power of smaller events at the cost of the larger events. A manual
 268 analysis of the spatial aftershock patterns of the six Californian mainshocks named in the intro-
 269 duction shows that the cloud of potential aftershocks typically lies within one estimated rupture
 270 length (by [Wells and Coppersmith, 1994](#)) from the epicenter. In case of an anisotropic shape of
 271 the kernel, the area of half a rupture length around the extended rupture segment seems sufficient.
 272 According to this observation, we choose a spatial restriction R_i for event i according to

$$R_i := \begin{cases} 10^{-2.57+0.62m_i}, & \text{for } m_i < M_{aniso}, \quad (\text{isotropic trigger}) \\ 0.5 \cdot 10^{-2.57+0.62m_i}, & \text{for } m_i \geq M_{aniso}, \quad (\text{anisotropic trigger}) \end{cases} \quad (10)$$

273 where again we use the strike-slip faulting subsurface rupture length scaling from [Wells and Cop-](#)
 274 [persmith \(1994\)](#).

275 In other words, the spatial kernel for event i is only defined in the restricted area

$$S_i(R_i) := \{(x, y) \in \mathbb{R}^2 | r_i(x, y) \leq R_i\}$$

276 and set to 0 outside of it. Note that the restricted area $S_i(R_i)$ can assume isotropic and anisotropic
 277 shapes, depending on the point-to-point or point-to-line definition of the distance term $r_i(x, y)$. In

278 order to retain the property of a pdf, we need to rescale the kernel within the restricted area by its
 279 analytical integral

$$\begin{aligned}
 H_{D,\gamma,q}(R_i, m_i, l_i) &:= \iint_{S_i(R_i)} h_{D,\gamma,q}(r_i(x, y), m_i, l_i) dx dy \\
 &= 1 - \left(1 + \frac{2 l_i R_i + \pi R_i^2}{D \exp(\gamma(m_i - M_c))} \right)^{1-q}.
 \end{aligned} \tag{11}$$

280 The integral term holds true for both isotropic ($l_i = 0$) and anisotropic triggers ($l_i > 0$). We obtain
 281 the generalized, restricted and anisotropic spatial kernel

$$h_{D,\gamma,q}^{restr}(r_i(x, y), m_i, l_i) = \begin{cases} \frac{h_{D,\gamma,q}^{restr}(r_i(x,y),m_i,l_i)}{H_{D,\gamma,q}(R_i,m_i,l_i)}, & \text{if } r_i(x, y) \leq R_i, \\ 0, & \text{if } r_i(x, y) > R_i. \end{cases} \tag{12}$$

282 2.4 Estimation of Strike and Epicenter Location

283 The restricted, anisotropic spatial kernel in Equation (12) requires a strike angle to define the ori-
 284 entation of the extended rupture for anisotropic triggers with $l_i > 0$. In retrospect, the strike angle
 285 could be taken from one of the numerous publications about the Ridgecrest sequence or from focal
 286 mechanism datasets such as the Global Moment Tensor Catalog, the ISC-GEM Global Instrumen-
 287 tal Earthquake Catalog or from local datasets of the Southern California Earthquake Data Center
 288 (SCEDC). In order to perform a realistic forecasting test case, however, we should build upon in-
 289 stantaneous strike estimates such as from local agencies (e.g. the United States Geological Survey),
 290 which are typically available within several minutes to hours.

291

292 Here, we used the strike estimation algorithm proposed by [Grimm et al. \(2021\)](#), that optimally fits

293 the rupture segment through the cloud of early aftershock locations by maximizing the contributed
 294 spatial rate under initial spatial parameter guesses. To be more precise, we ran through possible
 295 strikes in 1° steps (i.e. $\{1^\circ, \dots, 180^\circ\}$ where we can neglect all strikes larger than 180° because we
 296 do not account for dip direction in our model). We also moved the rupture along each strike angle
 297 in order to test different positions of the rupture segment relative to the fix epicenter. Here, we
 298 go through possible relative positions in 0.01-steps (i.e. $\{0, 0.01, 0.02, \dots, 0.99, 1\}$), where 0 and 1
 299 means that one of the ends of the rupture segment coincides with the epicenter, and 0.5 denotes the
 300 situation where the rupture embeds the epicenter directly in its center. Among all combinations,
 301 we searched the orientation and rupture position that maximizes the forward trigger contribution
 302 of the anisotropic event i to subsequent events j within a fixed duration $\Delta t = 1 \text{ hour}$, i.e. with
 303 $t_i < t_j < t_i + \Delta t$. The forward trigger contribution of event i is computed as

$$\sum_{t:t_i < t_j < t_i + \Delta t} h_{D,\gamma,q}^{rest}(r_i(x_j, y_j), m_i, l_i). \quad (13)$$

304 In order to avoid that the rupture orientation and position is dominated by single events that oc-
 305 curred very close to the segment candidate, we applied a minimum distance of 0.2 kilometers.

306

307 Here, we use the initial spatial parameters $D = 0.0025$, $\gamma = 1.78$ and $q = 1.71$ derived from the
 308 results of a locally restricted, isotropic ETAS model for California by [Grimm et al. \(2021\)](#). Tests
 309 have shown that modified initial parameters changed the level of the sum of forward rate contribu-
 310 tions, but led to similar strike and epicenter location estimates. We also tested multiple durations
 311 Δt up to 30 hours and found that the estimation procedure provided very similar estimates for
 312 strike and rupture position. It shows that the rupture orientation and position can be appropriately

313 identified soon after the trigger event occurred.

314

315 In the *Application* section we present the strike and rupture position estimation for the example of
316 the M6.4 and M7.1 Ridgecrest events.

317 2.5 Estimation of spatial integral

318 The computation of the log-likelihood function in Equation (6) requires the estimation of the spa-
319 tial integral of R_0 and therefore $h_{D,\gamma,q}^{restr}$. In the isotropic case, the integral can be estimated semi-
320 analytically by the radial partitioning method suggested by Ogata (1998) and applied in the R
321 package *ETAS* (Jalilian, 2019). It uses the property, that the isotropic spatial kernel can be inte-
322 grated analytically over circular areas $S_i(R)$, according to Equation (11).

323 As Figure 2(a) illustrates, the polygon defining the spatial window S is partitioned into a fine grid,
324 with two neighboring boundary grid points having approximately equal distances \tilde{R} to the point
325 source coordinate of event i . The integral of $h_{D,\gamma,q}^{restr}$ over each of these thin segments of a circle can
326 then be approximated by the analytical full integral, weighted by the ratio of the circle segment
327 $\phi/360^\circ$, where ϕ is the angle enclosed by the circle segment (Jalilian, 2019; Ogata, 1998).

328

329 Analogously, an anisotropic spatial kernels can be integrated analytically over an anisotropic
330 region $S_i(\tilde{R})$ with maximum distance \tilde{R} to the extended rupture.

331

332 Due to the non-circular shape of region $S_i(R)$ for anisotropic triggers, radial partitioning can
333 be only performed at both ends of the rupture segment. As Figure 2(b) illustrates, in a similar way
334 we use "bin partitioning" in the space orthogonal to the rupture. Unfortunately, in the anisotropic

335 case, the weights $\phi/360^\circ$ of the circle segments at both ends of the rupture only relate to the part of
336 the integral that is estimated by radial partitioning. Similarly, the weight of a bin of size Δl is $\frac{\Delta l}{2l_i}$
337 relative to only the orthogonal space on both sides of the rupture segment. In each iteration of the
338 parameter estimation, the shares of the radial and the orthogonal integral parts change and need to
339 be determined numerically before each iteration.

340 **3 Application**

341 We carry out three forecasting experiments to check the quality of the previously defined models
342 in predicting the observed Ridgecrest M6.4 and M7.1 sequences. Each forecasting experiment
343 consists of three main steps, represented as blue boxes in Fig. 3:

- 344 • **Parameter Estimation:** Estimate model parameters for a specified event sub-set of southern
345 Californian earthquake data
- 346 • **Forward Simulation:** Use the fitted model parameters to conduct 10,000 forward simula-
347 tions of the Ridgecrest M6.4 or M7.1 sequence, respectively.
- 348 • **Evaluation:** Analyze simulated sequences and compare to the observation.

349 In the following, we first introduce the basic earthquake event set for California underlying
350 this study, and define the time-space windows used to obtain the event sub-sets applied for pa-
351 rameter estimation. Next, we describe the three forecasting experiments, rigorously defining the
352 magnitude-time-space windows applied for parameter estimation and forward simulations. Each
353 forecasting experiment is repeated for five versions of the models introduced in the *Methods* sec-
354 tion, which are defined in detail. Finally, we specify the forward simulation process and attributes

355 and measures to assess the quality of the forecasts. Here, we also give an example of the estimation
356 of strike angles and rupture positions for the Ridgecrest M6.4 and M7.1 events.

357 3.1 Data

358 As our basic event set, we downloaded the earthquake catalog from the Southern California Earth-
359 quake Data Center (SCEDC, [Hauksson et al., 2012](#)). The entire dataset comprises 699,175 event
360 occurrences between January 1, 1981, and December 31, 2019. Because magnitudes seem to be
361 clustered at values with one decimal, we round all magnitudes to one decimal and use the cut-off
362 magnitude $M_c = 2.05$ ([Hainzl, 2021](#); [Hutton et al., 2010](#)). We remove events at depths larger than
363 40 km to avoid completeness issues.

364 3.2 Forecasting Experiments

365 Here, we describe in detail the design of the forecasting experiments, summarized in Fig. 3.

366 3.2.1 Experiment 1

367 We estimate generic, long-term California model parameters for the sub-set of events occurring
368 between January 1, 1987, and December 31, 2018, and within the hexagonal polygon of southern
369 California defined in [Hutton et al. \(2010\)](#). In particular, it includes the five prominent earthquake
370 sequences named in the *Introduction* section that occurred before the Ridgecrest sequence. For
371 numerical runtime reasons and in order to avoid potentially biased estimates of the blind time
372 parameter T_b in large spatial areas, we choose the larger cut-off magnitude $M_c = 2.95$. The
373 magnitude-time-space window contains 7,215 fitted *target* events. We account for external trig-
374 gering impact by including *complementary* events that occurred after January 1, 1986, and in the
375 surrounding of 0.5° geographical degrees of the fitted area.

376 The estimated models are then applied to forecast the Ridgecrest M6.4 foreshock sequence above
377 cut-off magnitude $M_c = 2.05$, within a circular polygon with radius 40 km centered in the M6.4
378 event location. The simulated time window starts in the moment of the M6.4 event (July 4, 2019)
379 and ends at the M7.1 mainshock event time (July 6, 2019), thus it has a duration of approximately
380 34 hours. We initialize triggering seismicity by the event history from June 1, 2019.

381 3.2.2 Experiment 2

382 In the second experiment, we use the same set of generic, long-term California parameters, but
383 apply it in a forecast of the Ridgecrest M7.1 mainshock sequence above cut-off magnitude $M_c =$
384 2.95, starting at the M7.1 event time for a duration of ten days. The spatial simulation window is
385 defined by a disk with radius of 75 km, centered in the M7.1 event location. Again, we initialize
386 triggering seismicity by the event history from June 1, 2019, here until the M7.1 event time.

387 3.2.3 Experiment 3

388 In the third experiment, we simulate Ridgecrest M7.1 sequences with the same settings as for *Ex-*
389 *periment 2*, but based on parameter estimates fitted over the immediately preceding M6.4 foreshock
390 sequence. For the parameter estimation, we use the same magnitude-time-space target window as
391 for the M6.4 sequence simulations in *Experiment 1*. We account for external triggering by includ-
392 ing complementary events that occurred after June 1, 2019, and within a disk with increased radius
393 of 50 km.

394 3.3 Fitted Models

395 Each forecasting experiment is carried out for five different versions of the model introduced in the
396 *Methods* section, summarized in table 1. The "ETAS conventional" model serves as our benchmark

397 and uses the most standard set-up of an ETAS model (e.g. [Jalilian, 2019](#); [Ogata, 1998](#); [Zhuang](#)
398 [et al., 2002](#)). It applies an isotropic spatial kernel with infinite spatial extent to all triggers. The
399 "ETAS iso-r" model applies an isotropic kernel, but restricts the spatial extent to one rupture length
400 for all events, according to Equation (10). In the "ETAS aniso-r" model, all events with magnitudes
401 $m_i \geq M_{aniso} = 6.0$ are modeled as an anisotropic trigger source with a spatial restriction to half a
402 rupture length (Equations 9 and 10). The other events are modeled as isotropic triggers, restricted
403 to one rupture length. The "ETASI iso-r" and "ETASI aniso-r" models combine the spatial kernel
404 settings of the latter models with the ETASI approach accounting for STAI.

405 3.4 Simulation Process

406 For each forecasting experiment and model version, we carry out 10,000 realizations of synthetic
407 sequences. At the beginning of each simulation, we distribute the Poisson-sampled number of
408 background events, scaled by the size of the spatial area, uniformly over space and time. The
409 assumption of an uniform spatio-temporal background event distribution appears plausible within
410 the short and small space-time simulation windows.

411 Next, we sample the numbers of offsprings for the initiating event history and the background
412 events. The number of offsprings, depending on trigger magnitude m , is drawn from a Poisson
413 distribution with expected value

$$N(m) = k(m) \frac{1}{1-p} \left((T+c)^{1-p} - c^{1-p} \right). \quad (14)$$

414 where $k(m)$ is the aftershock productivity function in Equation (2) and the latter term is the in-
415 tegral from $t = 0$ to a maximum trigger duration $t = T$ (in days) over the Omori-Utsu function

416 in Equation (1). We need to rescale the aftershock productivity to obtain the expected number of
 417 offsprings within T days, because the Omori-Utsu law is no pdf and therefore typically integrates
 418 to values larger than 1. Thus, it interacts with the scaling parameter A of the productivity function.
 419 Each triggered event is then assigned an event time and location according to inversion sampling
 420 from the (rescaled) Omori-Utsu law and the spatial kernel. The magnitude is sampled by the in-
 421 version method from the estimated FMD in Equation (4), applying a maximum magnitude of 7.5
 422 for California. The aftershock sampling routine is repeated for every newly triggered event in the
 423 simulated time-space window until no more events are sampled.

424

425 In order to make fair comparisons of simulated sequences with the observed ones, we need to
 426 consider the implications of STAI in the forecasts. The ETASI models account for incomplete
 427 records in the parameter estimation and therefore forecast the "true", i.e. complete, aftershock
 428 sequence. According to its definition of event detectability, we would need to delete all events that
 429 occurred within the blind time T_b of an earlier event with larger magnitude.

430 For the sake of transparency and consistency with the observations, we used an alternative approach
 431 and manually fitted empirical magnitude completeness functions

$$M_c(t) = \begin{cases} -1.4 \log_{10}(t) + M - M_c - 4.75, & \text{(Ridgecrest M6.4),} \\ -0.99 \log_{10}(t) + M - M_c - 3.75, & \text{(Ridgecrest M7.1).} \end{cases} \quad (15)$$

432 to the logarithmic event time-magnitude scatter data of the observed Ridgecrest M6.4 and Ridge-
 433 crest M7.1 sequences in Fig. 1(c) and (d).

434 In forecasts generated by the ETASI iso-r and aniso-r models, we delete all events that fall in the

435 supposedly undetected range below the line of the simulated sequence. In contrast, ETAS models
436 estimate STAI-biased aftershock productivities and therefore lead to predictions of the incomplete,
437 rather than the "true" size of the sequence. Therefore, in forecasts generated by these models we
438 do not delete events.

439 3.4.1 Attributes and Measures

440 For each model version and experiment, we want to assess the quality of the forecasts with respect
441 to three attributes, in comparison with the observed sequence evaluated over the same magnitude-
442 time-space window.

443 We compute the predicted cdf of the number of aftershocks and the predicted pdf of the largest
444 aftershock magnitude out of the 10,000 forecasted sequences. As a quantitative measure of the fit,
445 we determine the exceedance probability that the predicted distribution would forecast a larger or
446 the observed value. Extreme exceedance probabilities, either close to 0 or 1, indicate an inadequate
447 prediction of the attribute.

448

449 To test the spatial distribution of aftershock locations, we define a $1\text{km} \times 1\text{km}$ spatial grid over
450 the spatial simulation window of the forecasting experiment and count the number of aftershocks
451 in each simulation run, that occurred closest to the respective grid points. We determine the spatial
452 distribution D_{ij} of the i -th simulation run by dividing the number of events occurred at each grid
453 point j , N_{ij} , by the number of events in the i -th simulation run, N_i , i.e.

$$D_{ij} = N_{ij}/N_i.$$

454 By repeating the same procedure for each simulation run, we obtain 10,000 simulated spatial dis-
 455 tributions D_{ij} for each model version. Finally, we average the individual simulated distributions to
 456 obtain the predicted probability P_j that an event occurs at grid point j .

457 The more events of the observed sequence have occurred at grid points with high predicted proba-
 458 bilities P_j , the better is the forecast. Therefore, we quantify the goodness of the spatial fit with the
 459 likelihood $L_{space} = \prod_{grid\ points\ j} P_j^{N_j^{obs}}$ where N_j^{obs} is the number of observed events at grid point
 460 j with corresponding log-likelihood

$$LL_{space} = \sum_{grid\ points\ j} N_j^{obs} \ln(P_j).$$

461 We compute the information gain of the models' spatial predictions relative to the ETAS conven-
 462 tional model by

$$IG = \frac{LL_{space} - LL_{space}^0}{N_{obs}}$$

463 where LL_{space}^0 is the benchmark result for the ETAS conventional model (Hainzl, 2021; Rhoades
 464 et al., 2014).

465 3.4.2 Strike and Rupture Position Estimates

466 For anisotropic models, both the parameter estimation and the forward simulations of the Ridge-
 467 crest M6.4 and M7.1 sequences require estimates of the strike angle and rupture position of all
 468 events with magnitude $M > 6.0$.

469 Figure 4(a) demonstrates the methodology, described in the *Methods* section, for the Ridgecrest
 470 M6.4 foreshock. The forward trigger rate contribution (y axis) from Equation (13) is plotted against
 471 the strike sample (x axis) and the sample of relative rupture positions (red lines). The curves clearly

472 show a bi-modal shape, with peaks at strikes 34° and 132° and relative rupture positions 0.76 and
473 0.77. Fig. 4(c) visualizes the optimized rupture orientation and position as a fit through the cloud
474 of potential aftershocks within the first (dark red) or within 30 hours (light red). It confirms the
475 earlier mentioned particularity of two almost orthogonally ruptured faults. The strike 34° rupture
476 segment does not perfectly fit the aftershock alignment, as segment must go through the fixed M6.4
477 epicenter location which seems to be slightly off the ruptured fault. Apparently, later aftershocks
478 have a very similar spatial distribution as the events occurred within the first hour. For larger Δt ,
479 the M6.4 strike estimates would vary by only 1° or 2° , respectively.

480 Figure 4(b) shows the analogous analysis for the M7.1 Ridgecrest mainshock. Here, the maxi-
481 mizing properties are strike 142° and a relatively central rupture position 0.55. The M7.1 event
482 ruptured a single fault, resulting in an uni-modal shape of the forward trigger contribution curves.

483 **4 Results and Discussion**

484 In this section, we analyze and discuss the results of the three forecasting experiments, summarized
485 in Fig. 3. We use the attributes and measures presented in the *Application* section to evaluate the
486 quality of the forecasts, compare to the observed sequences. The model parameter estimation
487 results of both the generic California and the Ridgecrest M6.4 sequence parameter fits are listed
488 in table 2 and will help us to understand and explain features in the forecasts. After a rigorous
489 discussion of the forecasting results, we will mention some sensitivity tests that we applied to
490 check the robustness of our findings.

491 4.1 Forecasting Experiment 1

492 In the first forecasting experiment, we simulated Ridgecrest M6.4 sequences, starting in the mo-
493 ment of the M6.4 trigger event, based on generic parameters, fitted on a long-term and spacious
494 Californian event set. The simulation period covers the 34 hours until (but non-including) the oc-
495 currence of the M7.1 mainshock.

496

497 Fig. 5(a) shows the predicted cdfs of the number of aftershocks for each model, compared to the
498 observed M6.4 sequence, which produced 633 events in the same time-space window.

499 Evidently, the ETAS conventional model (with isotropic, unlimited spatial kernel) provides a very
500 inappropriate estimate, as it does not reach the observed number in any of the 10,000 simulations.

501 According to the ETAS iso-r and ETAS aniso-r models, the observed number of events would be
502 a possible, but rather unlikely scenario, with approximately 2.4% and 3.7% probability to exceed
503 the observed value. The ETASI models tend to only moderately (ETASI iso-r) or slightly (ETASI
504 aniso-r) underestimate the observed number.

505 To explain this observation, we consider that the size of this relatively short sequence is predomi-
506 nantly influenced by the amount of direct aftershocks of the initial M6.4 trigger event. According
507 to the model parameter estimates in table 2 and Equation (14), the M6.4 trigger event would only
508 produce approximately 17 direct aftershocks in the ETAS conventional model, compared to 46
509 (ETAS iso-r), 49 (ETAS aniso-r), 66 (ETASI iso-r) and 74 (ETASI aniso-r) in the other models.

510 The larger the parameter α , the more direct aftershocks are expected for the M6.4 event.

511 As argued in the *Methods* section, the local restriction of the spatial kernels prevents a dispropor-
512 tionate triggering power of small events and in return increases the direct aftershock productivity

513 of the stronger events, characterized by a considerable increase of the parameter α in the non-
514 conventional models (Grimm et al., 2021). Besides, the application of the ETASI model accounts
515 for missing aftershock records after strong trigger events and corrects for the biased, underesti-
516 mated aftershock productivity, leading to an additional increase of α in the ETASI models (Hainzl,
517 2021). Finally, we note that the majority of the $M > 6$ mainshocks included in the estimation time
518 window from 1987 until 2018, are characterized by anisotropic aftershock patterns. Consequently,
519 more events are associated as direct aftershocks of the strong mainshocks when we estimate the
520 parameters with the ETAS aniso-r or the ETASI aniso-r model.

521

522 Fig. 5(b) shows the predicted pdfs of the largest aftershock magnitude in the synthetic sequences.
523 For each of the five models, a kernel density function was computed for the 10,000 largest mag-
524 nitude samples. In all models, the observed M7.1 event would have been an extremely rare case,
525 with exceedance probabilities $P(\text{largest magnitude} \geq 7.1) \leq 0.43\%$. Even the second largest,
526 observed aftershock magnitude ($M = 5.4$) was not reached in approximately 75% of the simula-
527 tions of the best model (ETASI aniso-r).

528 To interpret this result, think of the largest aftershock magnitude as the largest order statistic of the
529 sample of simulated events in a simulation run. Then, the expected value of the sample maximum
530 (i.e. the largest aftershock) increases if (1) the sequence size becomes larger or (2) if the magni-
531 tudes in the sample are distributed in a way that they favor high values.

532 The underestimations of the observed sequence size, shown in Fig. 5(a) and discussed earlier, can-
533 not sufficiently explain the miss-match of the predicted largest aftershock magnitudes. Even the
534 observed sample size (633 events) would produce a M7.1 event with a probability of less than 1%,
535 given the generic California estimates for the FMD with $b = 0.98$ (ETAS models) or $b = 1.01$

536 (ETASI models; see table 2). If $b = 1$, then each magnitude increment by 1 leads to a 10 times
537 smaller probability of occurrence. Therefore, we would need a sample of the size 100,000 events
538 to obtain on average one M7.1 event.

539 According to the estimation results in table 2, all models estimated significantly smaller values
540 $b < 0.8$ for the observed Ridgecrest M6.4 sequence, which favors the occurrence of strong events.

541 Note that the b estimates of the three ETAS models are biased, because they are fitted for the "true"
542 FMD using the evidently short-term incomplete M6.4 sequence event record (see Fig. 1(c)). The
543 ETASI models account for the missing smaller magnitudes and therefore lead to corrected, larger
544 b values.

545 If we would simulate the Ridgecrest M6.4 sequence using its own estimation results (note that
546 this is not a valid forecasting experiment, but used for illustration purposes), we would obtain an
547 $M \geq 7.1$ event with 10.0% (ETAS conventional), 25.9% (ETAS iso-r), 53.7% (ETAS aniso-r),
548 15.6% (ETASI iso-r) and 25.3% (ETASI aniso-r) chance.

549

550 The branching ratios ν_{branch} , i.e. the average number of direct aftershocks per event, clearly
551 exceed 1 in each model (see table 2), indicating that the M6.4 sequence was in an unstable state.
552 According to these models, an earthquake would trigger more than one direct aftershock on aver-
553 age, which would cause the sequence to propagate exponentially. The large branching ratios are
554 mainly driven by the small b values, which substantially increase the occurrence probability of the
555 more productive, strong earthquakes.

556 The instability of the M6.4 sequence could be interpreted as an indication of an imminent, strong
557 mainshock. On the other hand, it is unlikely that the instability is based on a model error, e.g. due
558 to a substantial misfit of the b-value due to few magnitude outliers. First, the FMD is estimated

559 accounting for all earthquakes at equal weights, regardless of their magnitude. Therefore, the b
560 value estimate is primarily controlled by the more numerous, small magnitudes. Secondly, the
561 M7.1 event magnitude was *not* included in the b value estimation.

562 In summary, the generic California parameters are fitted to a long-term catalog mainly consisting of
563 stable earthquake sequences and seismically quiet periods. Therefore, it cannot adequately predict
564 the magnitude distribution of the M6.4 foreshock sequence, which is characterized by instability
565 due to a particularly flat FMD.

566

567 Fig. 6(a) and (b) show the predicted spatial event distributions, averaged over the 10,000 simulation
568 runs and evaluated on the $1 \text{ km} \times 1 \text{ km}$ grid described in the *Application* section, for the ETASI
569 iso-r model (in (a)) and the ETASI aniso-r model (in (b)). The observed event locations are plotted
570 overlaying the logarithmic heat map of grid cell probabilities. Intuitively, the anisotropic forecast
571 in (b) leads to a better spatial prediction than the isotropic counterpart.

572 In the isotropic model, all direct aftershocks are distributed point-symmetrically around the M6.4
573 trigger event. Subsequent secondary triggering then takes place around the new initiators. In the
574 anisotropic model, the direct aftershocks are distributed around the fitted rupture segments of the
575 two orthogonal faults (see Fig. 4). Subsequent trigger generations then spread isotropically (if
576 $M < M_{aniso}$) or anisotropically (if $M \geq M_{aniso}$) around their new initiators. In both plots, we
577 can see a pronounced boundary from green to blue color, indicating the spatial restriction to one
578 rupture length (isotropic model) and half a rupture length (anisotropic model) around the trigger
579 source, according to Equation (10). Spatial grid cells outside of this boundary can only be activated
580 by secondary triggering or background seismicity.

581 To quantify the quality of the spatial forecasts, we computed the information gains relative to the

582 ETAS conventional model as described in the *Application* section. Fig. 7(c) shows the results for
583 *Experiment 1* in the left group of bars. Both anisotropic models lead to a pronounced improvement,
584 which confirms the visual impression in Fig. 6(a) and (b). The ETAS and ETASI iso-r models,
585 which differ from the conventional approach in terms of the local spatial restriction, show a small
586 information gain. As we can see in Fig. 6(a), none of the observed events occurred outside of the
587 spatial restriction. Therefore, the restriction leads to a slightly more pronounced accumulation of
588 simulated event locations closer to the M6.4 trigger, which coincides better with the observation.

589 4.2 Forecasting Experiment 2

590 In the second forecasting experiment, we simulated the Ridgecrest M7.1 sequence for a duration
591 of 10 days based on the same generic California parameters as used for *Experiment 1*.

592
593 Fig. 5(c) compares the number of aftershocks, predicted by the five models, to the observed num-
594 ber of 3,273 events. The forecasts show a very similar setup of curves as in *Experiment 1* (see
595 Fig. 5(a)). The ETAS conventional model clearly underestimates the observed number of events.
596 The ETAS iso-r and aniso-r models reach the observation in about 6.5% and 14.1% of the simula-
597 tion runs. Again, the ETASI models provide the best approximations.

598 According to Equation (14), the M7.1 trigger event would on average trigger only roughly 53 di-
599 rect aftershocks in the ETAS conventional model, compared to 219 in the ETAS iso-r, 242 in the
600 ETAS aniso-r, 328 in the ETASI iso-r and 387 in the ETASI aniso-r model. As explained in detail
601 for *Experiment 1*, the reason is found in the parameter estimate for α .

602
603 Fig. 5(d) shows the predicted pdfs for the largest aftershock magnitude of the Ridgecrest M7.1

604 sequence. In contrast to *Experiment 1*, all but the conventional model provide very good forecasts,
605 indicating that the generic, long-term California estimates of the FMD with $b \approx 1$ coincide well
606 with the FMD of the Ridgecrest M7.1 sequence and the instability of the sequence ended with
607 the occurrence of the M7.1 mainshock. The moderate underestimation by the ETAS conventional
608 model can be explained by the underestimated sequence size, which substantially reduces the sam-
609 ple size of event magnitudes.

610

611 Fig. 6(c) and (d) show the predicted spatial distributions of aftershock locations, again for the
612 ETASI iso-r and aniso-r model. The visual impression, that the anisotropic model provides the
613 substantially better forecast, is confirmed by the bar plot in Fig. 7(c). The information gain by the
614 anisotropic models is more pronounced for the Ridgecrest M7.1 sequence, because it has a longer
615 rupture extension ($\sim 68km$ by Wells and Coppersmith, 1994) than the M6.4 event and it did not
616 rupture two orthogonal faults, which can be approximated more easily by an isotropic kernel.

617 4.3 Forecasting Experiment 3

618 In the third forecasting experiment, we simulated the 10-days Ridgecrest M7.1 sequence based on
619 the parameters fitted to the local Ridgecrest M6.4 foreshock sequence. Since the instability of the
620 sequence would lead to exploding forecasts, we assumed the long-term estimated FMD with $b = 1$.

621

622 Fig. 5(e) shows that the number of aftershocks is predicted much more similarly by the five models
623 than in *Experiments 1* and 2. It suggests that the particular features of the model versions play a
624 smaller role in the estimation over a closed, local sequence than in the generic fit over a long-term
625 catalog with several sequences and seismically quiet periods in between. The ETAS conventional

626 model reaches the observation in 4.4% of the simulation runs, the ETASI aniso-r even overesti-
627 mates the size of the sequence in 94.1% of the simulations. The other models show very good
628 predictions.

629

630 According to Fig. 5(f), our manual choice of $b = 1$ led to very realistic predictions of the largest
631 aftershock magnitude. Together with the results for the number of aftershocks, it shows that the
632 Ridgecrest left the unstable state after the M7.1 event by returning to the generic FMD, while re-
633 taining a similar structure of aftershock productivity.

634

635 Finally, Fig. 6(e) and (f) suggests that, compared to *Experiment 2*, the spatial kernels fitted over
636 the Ridgecrest M6.4 sequence are much narrower than those coming from the generic, long-term
637 model fit. This is confirmed by the larger estimates of q and the smaller estimates of γ in table 2.
638 Fig. 7(c) shows that the narrower spatial distribution leads to a more pronounced information gain
639 by the local restriction and the anisotropy, relative to the ETAS conventional model.

640 Note that, to some extent, the predicted spatial distributions show traces of late or secondary af-
641 tershocks triggered along the orthogonal M6.4 Ridgecrest fault, in contrast to very few observed
642 events in that area. This might be an indication of an underestimated Omori parameter p or an
643 overestimated c , favoring pronounced triggering over a longer time period.

644 4.4 Summary of Forecast Quality

645 Fig. 7 shows a summary of the quality measures for the three experiments, with respect to the
646 predicted number of aftershocks in Fig. 7(a), largest aftershock magnitude in Fig. 7(b) and spatial
647 aftershock distribution in Fig. 7(c). The conventional model scores worst in each category. It

648 confirms the results in [Grimm et al. \(2021\)](#), which argued that the isotropic and unlimited spatial
649 kernel assumes an implausibly far trigger reach and leads to underestimated cluster sizes.

650 According to Fig. 7(a), the ETASI models seem to estimate more realistic aftershock productivities
651 than the ETAS models when fitted over the long-term Californian catalog (see *Experiments 1* and
652 2). When fitted over the specific Ridgecrest M6.4 sequence, the bias of an underestimated aftershock
653 productivity seems to be balanced out by not cutting out undetected events. Anisotropic models
654 always lead to larger predicted sequence sizes, in the case of *Experiment 3* even to a substantial
655 overestimation.

656 The predictions of the largest aftershock magnitude, shown in Fig. 7(b), are reasonable for all but
657 *Experiment 1*. Apparently, the short-term incompleteness bias in the ETAS models is of much less
658 consequence for the FMD than for the aftershock productivity.

659 According to Fig. 7(c), as expected, the anisotropic models predict more realistic spatial event
660 distributions. The spatial restriction leads to a much smaller improvement.

661 4.5 Sensitivity of Results

662 As a sensitivity study, we forecasted the Ridgecrest M7.1 sequence over a duration of 50 days. In
663 a longer time window, direct aftershock productivity has less dominance, and is being displaced
664 more and more by secondary triggering. The underestimation of direct aftershock productivity
665 (e.g. in the ETAS conventional model) typically goes along with more pronounced secondary trig-
666 gering, characterized by larger estimates of the productivity constant A , see table 2. Therefore,
667 we observed that the ETAS conventional model caught up the missing events over time. On the
668 other hand, this indicates a temporal distribution of aftershocks which is not in agreement with the
669 observations. Other sensitivity tests, such as the model estimation with varying cut-off magnitudes

670 M_c or different time windows for the generic California estimates showed generally stable results.

671 **5 Conclusion**

672 In this article, we combined an ETAS approach with generalized anisotropic and locally restricted
673 spatial kernels (Grimm et al., 2021) with the ETASI time model of Hainzl (2021). The new fea-
674 tures have the objective to solve the three major biases of the conventional ETAS model, which are
675 the isotropic and spatially unlimited kernel as well as the neglect of short-term incompleteness
676 in the fitted event records.

677

678 We estimated five different versions of the new ETASI time-space model to a generic, long-term
679 Californian event set and to the specific Ridgecrest M6.4 foreshock sequence. Then, we applied
680 the fitted model parameters to generate synthetic forecasts of the Ridgecrest M6.4 and the M7.1
681 sequences. The objective of this study is to analyze and compare the forecast quality regarding the
682 predicted size of the sequence, largest aftershock magnitude and spatial aftershock distribution.

683

684 The results indicate that the ETAS conventional model leads to a substantial underestimation of
685 the number of aftershocks due to its disproportionately small estimates of the direct aftershock
686 productivity for the M6.4 and M7.1 trigger events. The locally restricted ETAS models without
687 ETASI-extension provide more realistic, but still underestimated predictions, as they are affected
688 by the bias of short-term incomplete event sequences in the fitted event set. The combination of
689 ETASI model with locally restricted spatial kernel seems to solve the bias and provides the most
690 robust predictions in the forecasting experiments. The anisotropy of the spatial kernel has a posi-
691 tive impact on the model estimation, however, shows its strength more clearly in the prediction of

692 the spatial event distribution of aftershocks.

693

694 More as a by-product, we find that the Ridgecrest M6.4 foreshock sequence showed instable behav-
695 ior, favoring strong aftershocks by a small Gutenberg-Richter parameter $b < 0.8$. The instability
696 of the foreshock sequence can be interpreted as an indication of an imminent strong mainshock. In
697 consequence, models fitted on the long-term, stable Californian event records cannot adequately
698 predict the magnitude distribution of this sequence.

699

700 The new model provides a better understanding of spatio-temporal earthquake clustering and
701 solves three major biases of the conventional ETAS model at once. Particularly, it leads to bet-
702 ter estimates of the aftershock productivity and to improved forecasts of the size of a sequence and
703 the spatial distribution of aftershocks. These improvements may be of major interest for short-term
704 risk assessment during an on-going aftershock sequence, particularly for the risk of a second, dam-
705 aging earthquake following the initial trigger event. The anisotropic spatial forecast of aftershock
706 locations enables disaster response managers to take actions in areas at risk where particularly
707 high aftershock activity is expected.

708

709 Future work should test the forecast quality for other earthquake sequences. It would be interesting
710 to address the question whether the ETASI time-space model can be used to reliably detect the
711 instability of a live sequence, which could have positive impacts on emergency management during
712 on-going sequences. The temporal event distribution, which was neglected in our study, should be
713 included into such analyses.

714

715 **Data and Resources**

716 The earthquake event set for California has been downloaded from the Southern California Earth-
717 quake Data Center (<https://scedc.caltech.edu/data/alt-2011-dd-hauksson-yang-shearer.html>, last
718 accessed on October 25, 2021). Results and figures were produced using Matlab software.

719 **Acknowledgments**

720 Financial support for this work was provided by Munich Re through a scholarship granted to the
721 first author, and by the Department of Statistics at Ludwig-Maximilians-University Munich.

722 **Statements and Declarations**

723 Financial support for this work was provided by Munich Re through a scholarship granted to
724 the first author, and by the Department of Statistics at Ludwig-Maximilians-University Munich
725 (e.g. travel costs, conference and journal fees). The authors acknowledge that there are no relevant
726 financial or non-financial interests to disclose. This article does not contain any studies involving
727 human participants or animals performed by any of the authors.

728 **Contribution of the authors**

729 CG (first author) derived and programmed the model, conducted simulation experiments, prepared
730 figures and wrote the paper. SH provided theory of ETASI time model, consulted in programming
731 and conducted comparative calculations. CG, SH and MK designed the simulation experiments
732 and interpreted the results. HK consulted on statistical questions and was involved in the concept
733 of the study.

734 **References**

735 de Arcangelis, L., Godano, C., and Lippiello, E. (2018). The overlap of aftershock coda waves
736 and short-term postseismic forecasting. *Journal of Geophysical Research: Solid Earth*, **123**,
737 5661–5674. doi: 10.1029/2018JB015518.

738 Grimm, C., Käser, M., Hainzl, S., Pagani, M., and Küchenhoff, H. (2021). Improving Earthquake
739 Doublet Frequency Predictions by Modified Spatial Trigger Kernels in the Epidemic-Type Af-
740 tershock Sequence (ETAS) Model. *Bulletin of the Seismological Society of America*, (Xx). doi:
741 10.1785/0120210097.

742 Gutenberg, B. and Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the*
743 *Seismological Society of America*, **34**, 185–188. doi: 10.1038/156371a0.

744 Hainzl, S., Christophersen, A., and Enescu, B. (2008). Impact of earthquake rupture extensions on
745 parameter estimations of point-process models. *Bulletin of the Seismological Society of America*,
746 **98**(4), 2066–2072. doi: 10.1785/0120070256.

747 Hainzl, S. (2016a). Rate-dependent incompleteness of earthquake catalogs. *Seismological Re-*
748 *search Letters*, **87**(2A), 337–344. doi: 10.1785/0220150211.

749 Hainzl, S. (2016b). Apparent triggering function of aftershocks resulting from rate-dependent
750 incompleteness of earthquake catalogs. *Journal of Geophysical Research: Solid Earth*, **121**(9),
751 6499–6509. doi: 10.1002/2016JB013319.

752 Hainzl, S. (2021). ETAS-Approach Accounting for Short-Term Incompleteness of Earthquake
753 Catalogs. *Bulletin of the Seismological Society of America*. doi: 10.1785/0120210146.

754 Hainzl, S., Zakharova, O., and Marsan, D. (2013). Impact of aseismic transients on the estimation
755 of aftershock productivity parameters. *Bulletin of the Seismological Society of America*, **103**(3),
756 1723–1732. doi: 10.1785/0120120247.

757 Harte, D. S. (2013). Bias in fitting the ETAS model: A case study based on New Zealand seismic-
758 ity. *Geophysical Journal International*, **192**(1), 390–412. doi: 10.1093/gji/ggs026.

759 Harte, D. S. (2016). Model parameter estimation bias induced by earthquake magnitude cut-off.
760 *Geophysical Journal International*, **204**(2), 1266–1287. doi: 10.1093/gji/ggv524.

761 Hauksson, E., Yang, W., and Shearer, P. M. (2012). Waveform relocated earthquake catalog for
762 Southern California (1981 to June 2011). *Bulletin of the Seismological Society of America*, **102**
763 (5), 2239–2244. doi: 10.1785/0120120010.

764 Hutton, K., Woessner, J., and Hauksson, E. (2010). Earthquake monitoring in Southern California
765 for seventy-seven years (1932–2008). *Bulletin of the Seismological Society of America*, **100**(2),
766 423–446. doi: 10.1785/0120090130.

767 Jalilian, A. (2019). ETAS: An R package for fitting the space-time ETAS model to earthquake
768 data. *Journal of Statistical Software*, **88**(1). doi: 10.18637/jss.v088.c01.

769 Lippiello, E., Cirillo, A., Godano, G., Papadimitriou, E., and Karakostas, V. (2016). Real-time
770 forecast of aftershocks from a single seismic station signal. *Geophysical Research Letters*, **43**
771 (12), 6252–6258. doi: 10.1002/2016GL069748.

772 Marsan, D. and Ross, Z. E. (2021). Inverse Migration of Seismicity Quiescence During the 2019
773 Ridgecrest Sequence. *Journal of Geophysical Research: Solid Earth*, **126**(3), 1–13. doi: 10.
774 1029/2020JB020329.

- 775 Marsan, D. and Lengliné, O. (2008). Extending earthquakes' reach through cascading. *Science*,
776 **319**(5866), 1076–1079. doi: 10.1126/science.1148783.
- 777 Mizrahi, L., Nandan, S., and Wiemer, S. (2021). The effect of declustering on the size distribution
778 of mainshocks. *Seismological Research Letters*, **XX**, 1–10. doi: 10.1785/0220200231.
- 779 Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point
780 processes. *Journal of the American Statistical Association*, **83**(401), 9–27.
- 781 Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the*
782 *Institute of Statistical Mathematics*, **50**(2), 379–402.
- 783 Ogata, Y. (2011). Significant improvements of the space-time ETAS model for forecasting of
784 accurate baseline seismicity. *Earth, Planets and Space*, **63**(3), 217–229. doi: 10.5047/eps.2010.
785 09.001.
- 786 Ogata, Y. and Zhuang, J. (2006). Space-time ETAS models and an improved extension. *Tectono-*
787 *physics*, **413**(1-2), 13–23. doi: 10.1016/j.tecto.2005.10.016.
- 788 Omi, T., Ogata, Y., Hirata, Y., and Aihara, K. (2013). Forecasting large aftershocks within one day
789 after the main shock. *Scientific Reports*, **3**, 1–7. doi: 10.1038/srep02218.
- 790 Omi, T., Ogata, Y., Hirata, Y., and Aihara, K. (2014). Estimating the ETAS model from an early af-
791 tershock sequence. *Geophysical Research Letters*, (41), 850–857. doi: 10.1002/2013GL058958.
- 792 Page, M. T., van Der Elst, N., Hardebeck, J., Felzer, K., and Michael, A. J. (2016). Three in-
793 gredients for improved global aftershock forecasts: Tectonic region, time-dependent catalog

794 incompleteness, and intersequence variability. *Bulletin of the Seismological Society of America*,
795 **106**(5), 2290–2301. doi: 10.1785/0120160073.

796 Rhoades, D. A., Gerstenberger, M. C., Christophersen, A., Zechar, J. D., Schorlemmer, D., Werner,
797 M. J., and Jordan, T. H. (2014). Regional earthquake likelihood models II: Information gains
798 of multiplicative hybrids. *Bulletin of the Seismological Society of America*, **104**(6), 3072–3083.
799 doi: 10.1785/0120140035.

800 Seif, S., Mignan, A., Zechar, J. D., Werner, M. J., and Wiemer, S. (2017). Estimating ETAS: The
801 effects of truncation, missing data, and model assumptions. *Journal of Geophysical Research:*
802 *Solid Earth*, **122**(1), 449–469. doi: 10.1002/2016JB012809.

803 Utsu, T., Ogata, Y., and Matsu'ura, R. S. (1995). The centenary of the Omori formula for a decay
804 law of aftershock activity. *J. Phys. Earth*, **43**, 1–33.

805 Wells, D. L. and Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture
806 length, rupture width, rupture area, and surface displacements. *Bulletin of the Seismological*
807 *Society of America*, **84**(4), 974–1002.

808 Zhang, L., Werner, M. J., and Goda, K. (2018). Spatiotemporal seismic hazard and risk assessment
809 of aftershocks of M 9 megathrust earthquakes. *Bulletin of the Seismological Society of America*,
810 **108**(6), 3313–3335. doi: 10.1785/0120180126.

811 Zhuang, J., Ogata, Y., and Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake
812 occurrences. *Journal of the American Statistical Association*, **97**(458), 369–380. doi: 10.1198/
813 016214502760046925.

814 Zhuang, J., Ogata, Y., and Wang, T. (2017). Data completeness of the Kumamoto earthquake
815 sequence in the JMA catalog and its influence on the estimation of the ETAS parameters. *Earth,*
816 *Planets and Space*, **69**, 36. doi: 10.1186/s40623-017-0614-6.

817 **Contact to authors**

818 Christian Grimm (main author), *Ludwig-Maximilians-University Munich, Department of Statis-*
819 *tics, Ludwigstraße 33, 80539 Munich, Germany; Christian.Grimm@stat.uni-muenchen.de,*

820 Sebastian Hainzl, *GFZ German Research Centre for Geoscience, Physics of Earthquakes and Vol-*
821 *canoes, Helmholtzstraße 6/7, 14467 Potsdam, Germany,*

822 Martin Käser, *Ludwig-Maximilians-University Munich, Department of Earth and Environmental*
823 *Sciences, Geophysics, Theresienstraße 41, 80333 Munich, Germany, also at Munich Re, Section*
824 *GeoRisks, Königinstr. 107, 80802 Munich, Germany,*

825 Helmut Küchenhoff, *Ludwig-Maximilians-University Munich, Department of Statistics, Lud-*
826 *wigstraße 33, 80539 Munich, Germany*

827 **List of Tables**

- 828 1 Overview of the model variants tested in this paper. Non applicable columns are
829 filled with ”-”. Spatial restrictions R_i of event i are denoted in terms of the estimate
830 rupture length (RL_i).
- 831 2 Overview of model results for generic (long-term) California and Ridgecrest M6.4
832 parameter estimation.

Table 1: Overview of the model variants tested in this paper. Non applicable columns are filled with ”-”. Spatial restrictions R_i of event i are denoted in terms of the estimate rupture length (RL_i).

Name	Model Version	M_{aniso}	R_i (isotropic triggers)	R_i (anisotropic triggers)
ETAS conventional	ETAS	-	∞	-
ETAS iso-r	ETAS	-	$1 RL_i$	-
ETAS aniso-r	ETAS	6.0	$1 RL_i$	$0.5 RL_i$
ETASI iso-r	ETASI	-	$1 RL_i$	-
ETASI aniso-r	ETASI	6.0	$1 RL_i$	$0.5 RL_i$

Table 2: Overview of model results for generic (long-term) California and Ridgecrest M6.4 parameter estimation.

		Generic California Estimates					Ridgecrest M6.4 Estimates				
		ETAS			ETASI		ETAS			ETASI	
Parameter		conv.	iso-r	aniso-r	iso-r	aniso-r	conv.	iso-r	aniso-r	iso-r	aniso-r
μ	$\frac{1}{days}$	0.16	0.21	0.21	0.21	0.21	0.11	0.30	0.29	0.18	0.30
A		0.027	0.012	0.011	0.010	0.009	0.052	0.024	0.022	0.022	0.019
α	$\frac{1}{mag}$	1.30	1.87	1.92	1.98	2.05	1.13	1.71	1.75	1.76	1.83
c	$\frac{1}{days}$	0.004	0.010	0.010	0.005	0.005	0.008	0.015	0.014	0.010	0.007
p		1.06	1.08	1.08	1.09	1.09	1.16	1.09	1.06	1.07	1.04
D	km^2	0.085	0.037	0.110	0.037	0.107	0.135	0.085	0.469	0.080	0.399
γ	$\frac{1}{mag}$	1.60	1.86	2.09	1.88	2.10	1.15	1.43	1.55	1.44	1.57
q		1.51	1.03	2.14	1.07	2.20	1.93	1.73	8.98	1.72	8.79
T_b	sec				112.8	114.0				18.1	21.1
b		0.98	0.98	0.98	1.01	1.01	0.72	0.72	0.72	0.77	0.79
LL		20806	17478	18209	16321	17107	6524	6322	6433	6013	6131
ν_{branch}		0.73	0.60	0.59	0.61	0.60	1.38	1.76	1.89	1.54	1.52

833 List of Figures

- 834 1 (a) Magnitudes vs. event times of Ridgecrest Mw6.4 (red dots) and Mw7.1 (blue
835 dots) aftershock sequences. Event times are denoted in days before/after Mw7.1
836 mainshock, the dashed black line represents the time origin (M7.1 event time).
837 Light blue and light red dots mark aftershocks with magnitudes larger than 5. Yel-
838 low pentagram symbolizes Mw6.4 foreshock, and yellow hexagram marks Mw7.1
839 mainshock. (b) Aftershock locations of Ridgecrest Mw6.4 and Mw7.1 sequences.
840 Legend as in (a). (c) Magnitudes vs. logarithmic event times of Ridgecrest Mw6.4
841 sequence. The dashed red line represents a manually fitted estimate of the em-
842 pirical completeness function $M_c(t)$. (d) Magnitudes vs. logarithmic event times
843 of Ridgecrest Mw7.1 sequence. The dashed red line represents a manually fitted
844 estimate of the empirical completeness function $M_c(t)$.
- 845 2 Visualization of the spatial integral estimation needed for computing the log-likelihood
846 function (equation 6) for (a) isotropic triggers and (b) anisotropic triggers. The box
847 represents the boundary of the spatial target region (polygon), gridded into small
848 intervals. Red crosses symbolize the epicenters of the triggering events. In (a),
849 the red circle around the event represents the contour lines of an isotropic spatial
850 kernel and the shaded segments illustrate the radial partitioning method. In (b),
851 the red box and semi-circles symbolize the contour lines of the anisotropic spatial
852 kernel, the thin circular segment and the shaded segments illustrate the radial and
853 bin partitioning method.
- 854 3 Summary of forecasting experiments (from left to right): The five model versions,
855 listed in table 1, are fitted to a long-term California event sub-set (*Experiments 1*
856 and 2) and to the local M6.4 Ridgecrest sequence (*Experiment 3*). The estimated
857 parameters are applied to forward simulations of the Ridgecrest M6.4 sequence
858 (*Experiment 1*) and the Ridgecrest M7.1 sequence (*Experiments 2 and 3*). The pre-
859 dicted sequences are compared to the observed ones with respect to three attributes,
860 further described in the *Attributes and Measures* section.
- 861 4 Strike and relative rupture position optimization using initial ETAS parameter guesses
862 $D = 0.0025, \gamma = 1.78, q = 1.71$. (a), (b): Sum of forward trigger rate contribu-
863 tions to events within one hour against strike sample (x axis) and relative rupture
864 position sample (curves) for (a) the M6.4 Ridgecrest foreshock and (b) the M7.1
865 Ridgecrest mainshock. Text boxes show strike and relative rupture position esti-
866 mates at the curve maxima. (c), (d): Fitted rupture segments through cloud of
867 aftershocks after (c) the M6.4 Ridgecrest foreshock and (d) the M7.1 Ridgecrest
868 mainshock. Darker red and blue points represent aftershocks within the first hour
869 after the respective trigger event, lighter red and blue represent aftershocks within
870 the first 30 hours. Yellow pentagram symbolizes Mw6.4 foreshock, and yellow
871 hexagram marks Mw7.1 mainshock. Thick black lines represent estimated rupture
872 locations according to the strikes and relative rupture positions estimated in (a) and
873 (b).

- 874 5 Predicted cdfs of the number of aftershocks (a,c,e) and predicted pdfs of the largest
875 aftershock magnitude (b,d,f) for *Experiment 1* (a,b), *Experiment 2* (c,d) and *Exper-*
876 *iment 3* (e,f). Each predicted distribution is based on 10,000 simulated forecasts of
877 the Ridgecrest M6.4 sequence (a,b) and the Ridgecrest M7.1 sequence (c-f), using
878 the models indicated in the legend in the top left figure. Vertical gray lines show
879 the value of the observed sequence.
- 880 6 Predicted spatial event distributions for *Experiment 1* (a,b), *Experiment 2* (c,d) and
881 *Experiment 3* (e,f). Each predicted distribution is averaged over 10,000 simulated
882 forecasts of the Ridgecrest M6.4 sequence (a,b) and the Ridgecrest M7.1 sequence
883 (c-f), based on the ETASI iso-r model (a,c,e) and the ETASI aniso-r model (b,d,f) .
884 The color bar indicates the predicted, logarithmic probability that an event occurs
885 at the respective grid point.
- 886 7 Summary plots of forecasting results. Predicted probabilities per model that (a) the
887 number of aftershocks exceeds the observation (633 for Ridgecrest M6.4; 3,273
888 for Ridgecrest M7.1) and (b) the largest aftershock magnitude exceeds the obser-
889 vation (7.1 for Ridgecrest M6.4; 5.5 for Ridgecrest M7.1). Dashed horizontal lines
890 represent 2.5% and 97.5% quantiles. (c) Spatial information gains relative to the
891 ETAS conventional model prediction for the same experiment. Legend in (a) holds
892 for all plots.

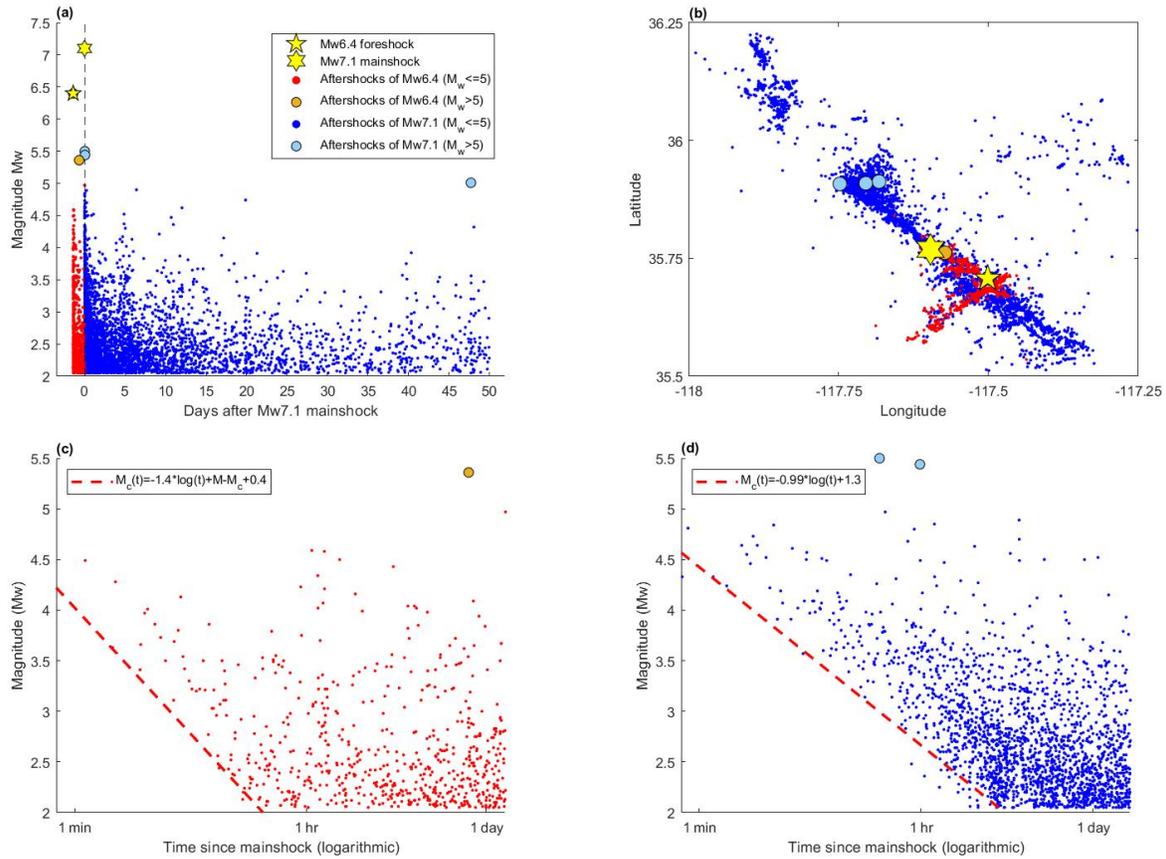


Fig 1: (a) Magnitudes vs. event times of Ridgecrest Mw6.4 (red dots) and Mw7.1 (blue dots) aftershock sequences. Event times are denoted in days before/after Mw7.1 mainshock, the dashed black line represents the time origin (M7.1 event time). Light blue and light red dots mark aftershocks with magnitudes larger than 5. Yellow pentagram symbolizes Mw6.4 foreshock, and yellow hexagram marks Mw7.1 mainshock. (b) Aftershock locations of Ridgecrest Mw6.4 and Mw7.1 sequences. Legend as in (a). (c) Magnitudes vs. logarithmic event times of Ridgecrest Mw6.4 sequence. The dashed red line represents a manually fitted estimate of the empirical completeness function $M_c(t)$. (d) Magnitudes vs. logarithmic event times of Ridgecrest Mw7.1 sequence. The dashed red line represents a manually fitted estimate of the empirical completeness function $M_c(t)$.

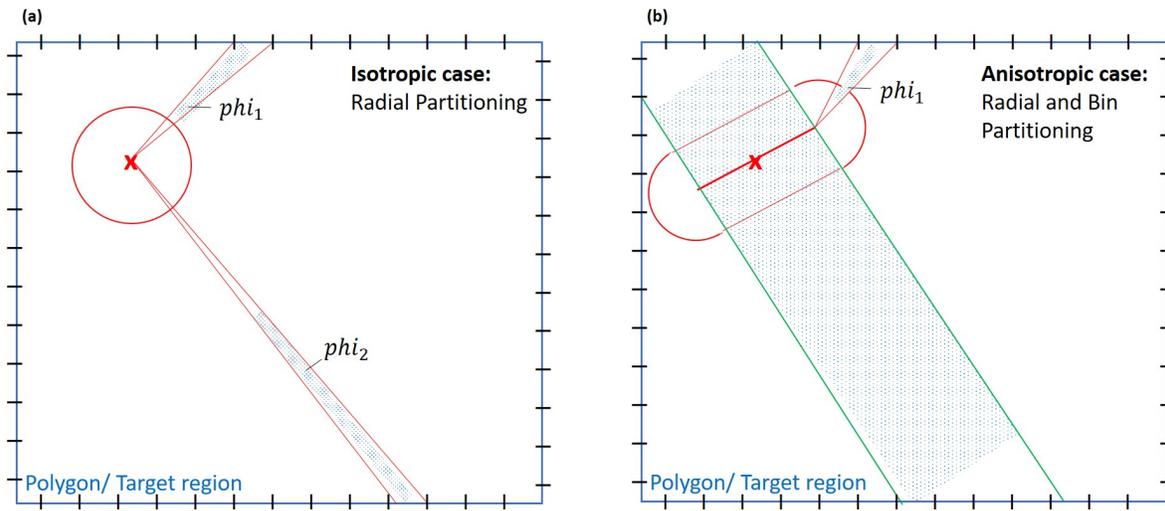


Fig 2: Visualization of the spatial integral estimation needed for computing the log-likelihood function (equation 6) for (a) isotropic triggers and (b) anisotropic triggers. The box represents the boundary of the spatial target region (polygon), gridded into small intervals. Red crosses symbolize the epicenters of the triggering events. In (a), the red circle around the event represents the contour lines of an isotropic spatial kernel and the shaded segments illustrate the radial partitioning method. In (b), the red box and semi-circles symbolize the contour lines of the anisotropic spatial kernel, the thin circular segment and the shaded segments illustrate the radial and bin partitioning method.

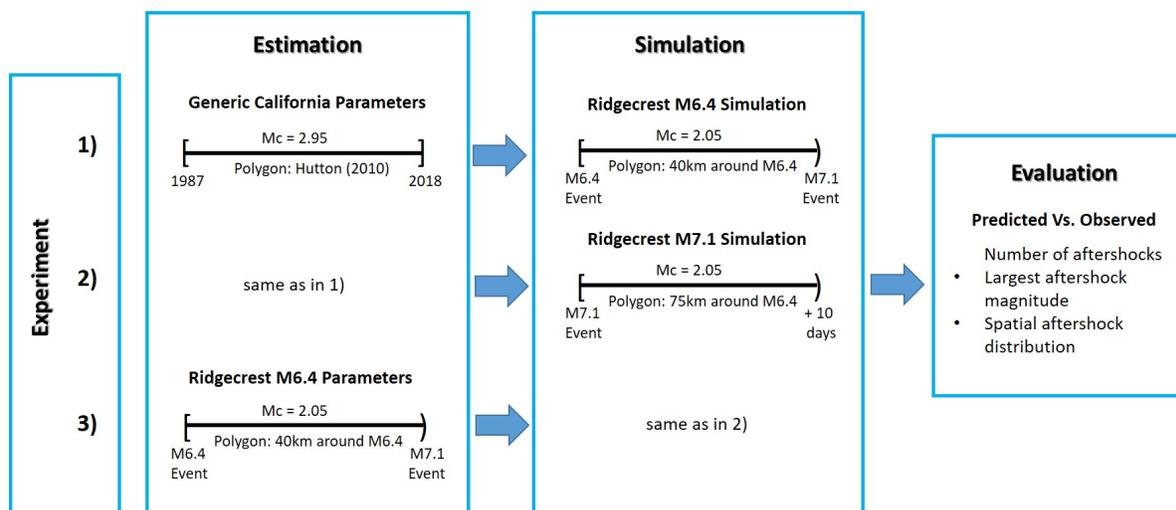


Fig 3: Summary of forecasting experiments (from left to right): The five model versions, listed in table 1, are fitted to a long-term California event sub-set (*Experiments 1* and 2) and to the local M6.4 Ridgecrest sequence (*Experiment 3*). The estimated parameters are applied to forward simulations of the Ridgecrest M6.4 sequence (*Experiment 1*) and the Ridgecrest M7.1 sequence (*Experiments 2* and 3). The predicted sequences are compared to the observed ones with respect to three attributes, further described in the *Attributes and Measures* section.

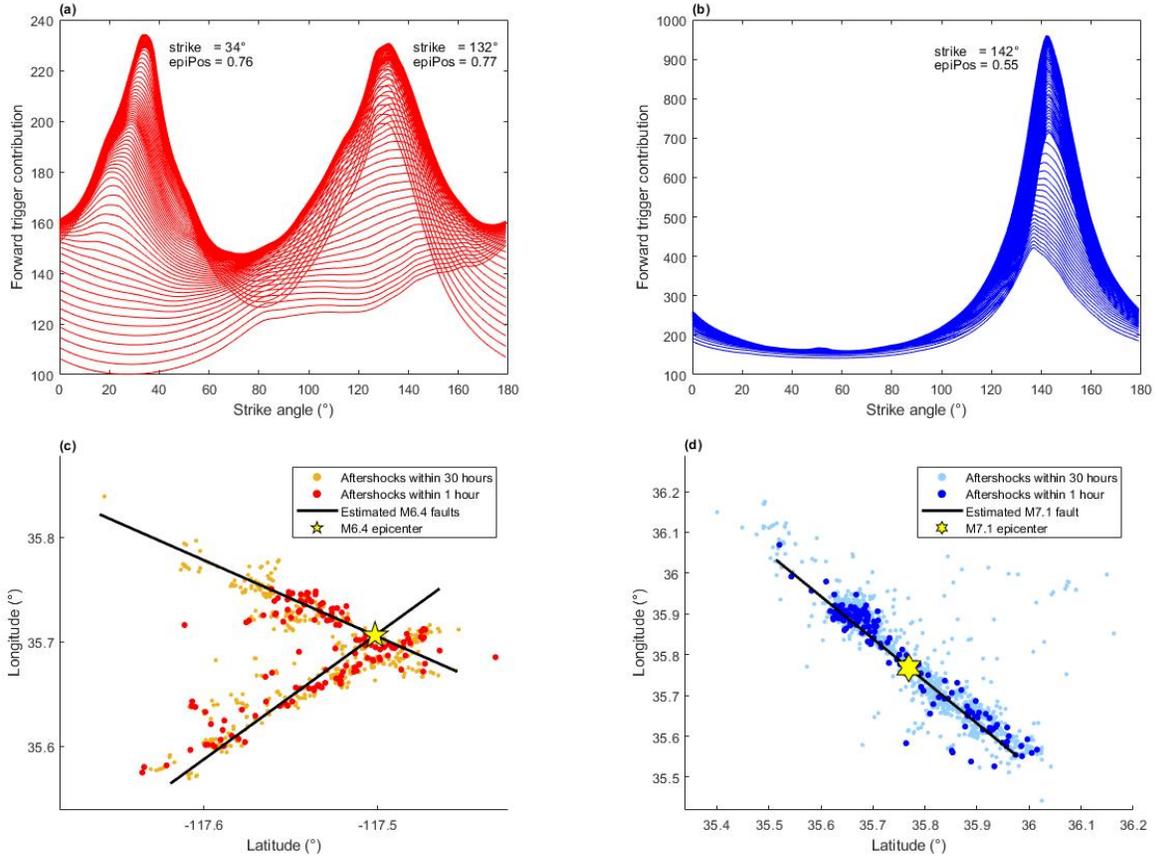


Fig 4: Strike and relative rupture position optimization using initial ETAS parameter guesses $D = 0.0025, \gamma = 1.78, q = 1.71$. (a), (b): Sum of forward trigger rate contributions to events within one hour against strike sample (x axis) and relative rupture position sample (curves) for (a) the M6.4 Ridgecrest foreshock and (b) the M7.1 Ridgecrest mainshock. Text boxes show strike and relative rupture position estimates at the curve maxima. (c), (d): Fitted rupture segments through cloud of aftershocks after (c) the M6.4 Ridgecrest foreshock and (d) the M7.1 Ridgecrest mainshock. Darker red and blue points represent aftershocks within the first hour after the respective trigger event, lighter red and blue represent aftershocks within the first 30 hours. Yellow pentagram symbolizes Mw6.4 foreshock, and yellow hexagram marks Mw7.1 mainshock. Thick black lines represent estimated rupture locations according to the strikes and relative rupture positions estimated in (a) and (b).

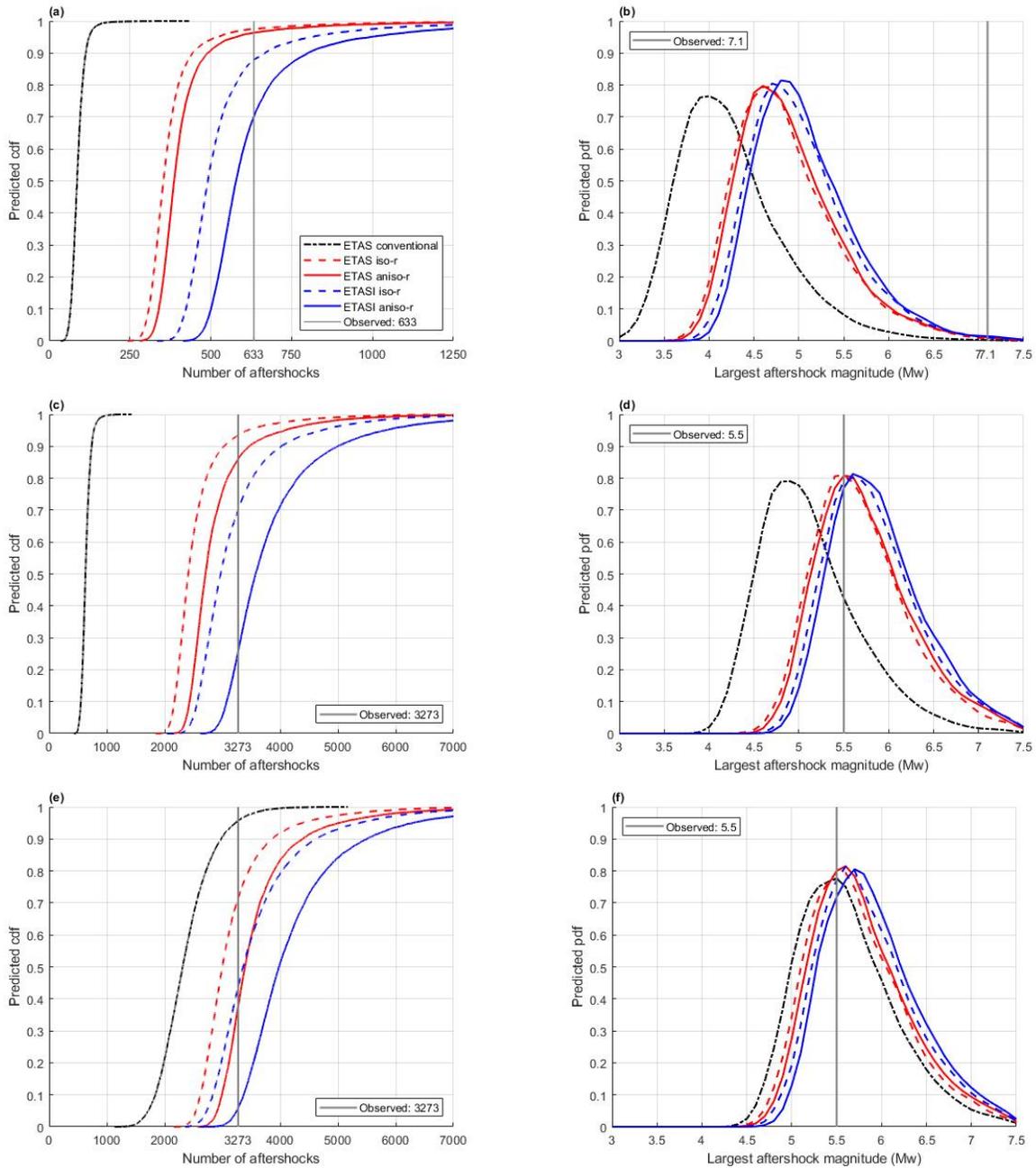


Fig 5: Predicted cdfs of the number of aftershocks (a,c,e) and predicted pdfs of the largest aftershock magnitude (b,d,f) for *Experiment 1* (a,b), *Experiment 2* (c,d) and *Experiment 3* (e,f). Each predicted distribution is based on 10,000 simulated forecasts of the Ridgecrest M6.4 sequence (a,b) and the Ridgecrest M7.1 sequence (c-f), using the models indicated in the legend in the top left figure. Vertical gray lines show the value of the observed sequence.

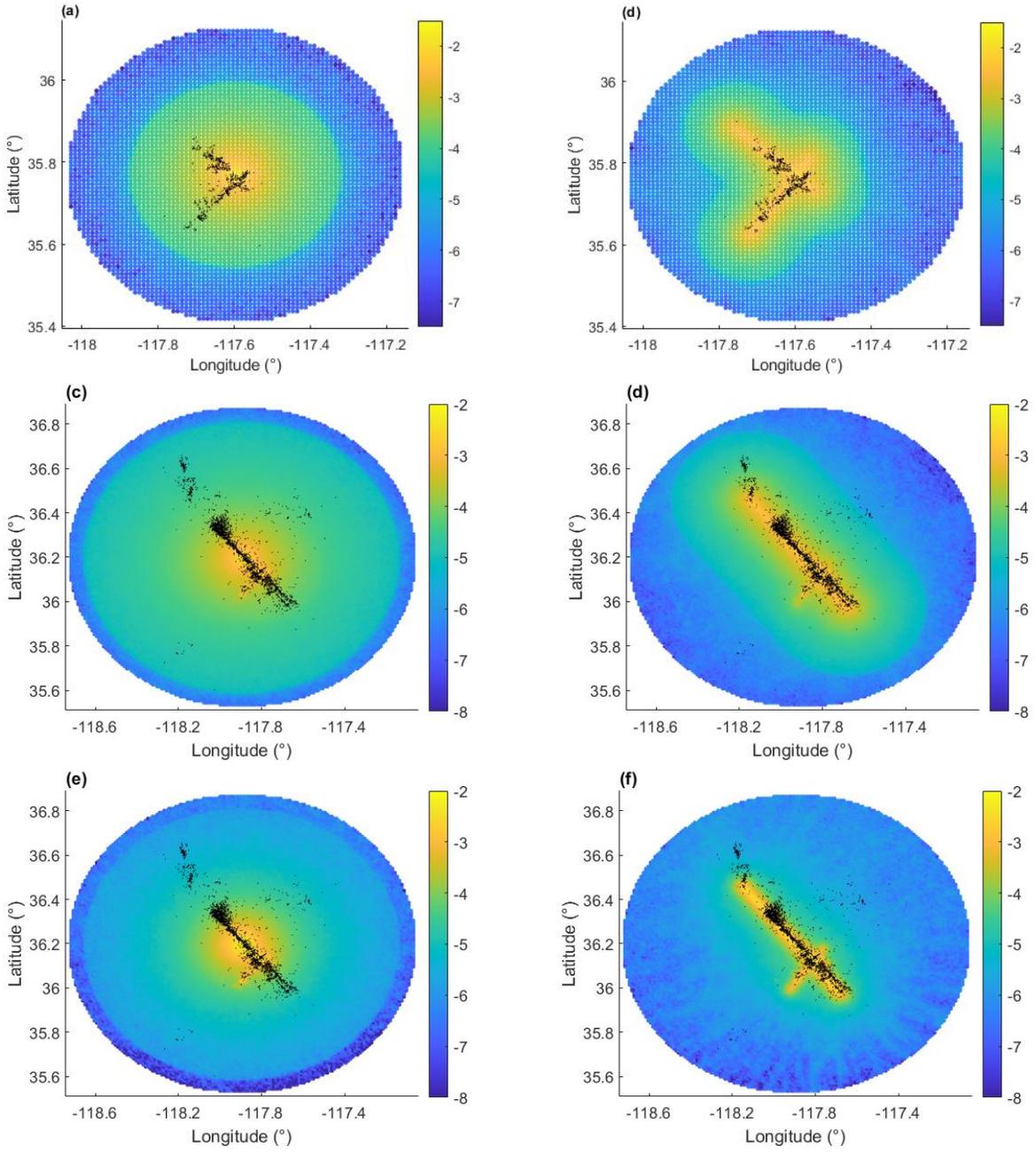


Fig 6: Predicted spatial event distributions for *Experiment 1* (a,b), *Experiment 2* (c,d) and *Experiment 3* (e,f). Each predicted distribution is averaged over 10,000 simulated forecasts of the Ridgecrest M6.4 sequence (a,b) and the Ridgecrest M7.1 sequence (c-f), based on the ETASI iso-r model (a,c,e) and the ETASI aniso-r model (b,d,f). The color bar indicates the predicted, logarithmic probability that an event occurs at the respective grid point.

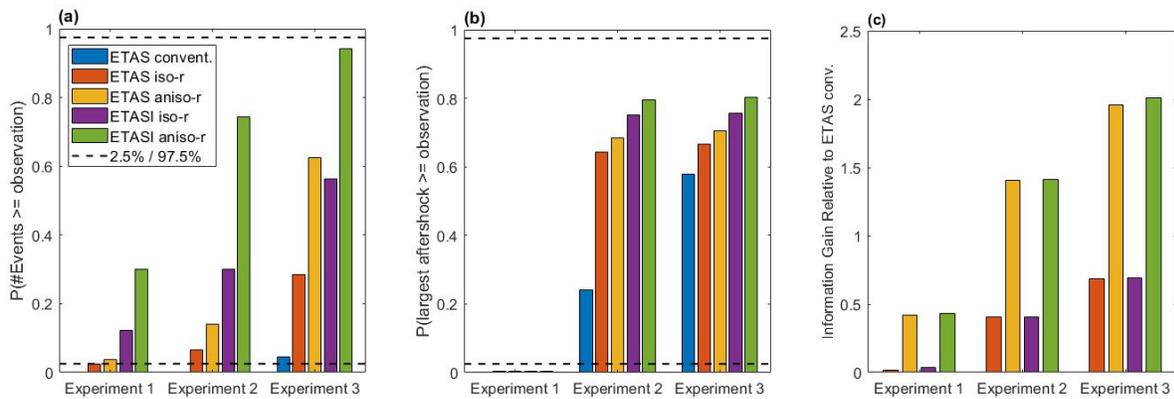


Fig 7: Summary plots of forecasting results. Predicted probabilities per model that (a) the number of aftershocks exceeds the observation (633 for Ridgecrest M6.4; 3,273 for Ridgecrest M7.1) and (b) the largest aftershock magnitude exceeds the observation (7.1 for Ridgecrest M6.4; 5.5 for Ridgecrest M7.1). Dashed horizontal lines represent 2.5% and 97.5% quantiles. (c) Spatial information gains relative to the ETAS conventional model prediction for the same experiment. Legend in (a) holds for all plots.