Stress release model and proxy measures of earthquake size. Application to Italian seismogenic sources

Elisa Varini[1], Renata Rotondi[1], Roberto Basili[2], and Salvatore Barba[2]

[1]Istituto di Matematica Applicata e Tecnologie Informatiche *Enrico Magenes*, Consiglio Nazionale delle Ricerche (CNR), Milan, Italy

[2]Istituto Nazionale di Geofisica e Vulcanologia (INGV), Rome, Italy

Corresponding author: Elisa Varini. Address: CNR-IMATI, via Bassini 15, 20133 Milan (Italy). E-mail: elisa@mi.imati.cnr.it

Abstract

This study presents a series of self-correcting models that are obtained by integrating information about seismicity and fault sources in Italy. Four versions of the stress release model are analyzed, in which the evolution of the system over time is represented by the level of strain, moment, seismic energy, or energy scaled by the moment. We carry out the analysis on a regional basis by subdividing the study area into eight tectonically coherent regions. In each region, we reconstruct the seismic history and statistically evaluate the completeness of the resulting seismic catalog.

Abbreviations:

SR	stress release
MR	macroregion
McMC	Markov chain Monte Carlo
ISS	Individual Seismogenic Sources
CSS	Composite Seismogenic Sources

Following the Bayesian paradigm, we apply Markov chain Monte Carlo methods to obtain parameter estimates and a measure of their uncertainty expressed by the simulated posterior distribution. The comparison of the four models through the Bayes factor and an information criterion provides evidence (to different degrees depending on the region) in favor of the stress release model based on the energy and the scaled energy. Therefore, among the quantities considered, this turns out to be the measure of the size of an earthquake to use in stress release models. At any instant, the time to the next event turns out to follow a Gompertz distribution, with a shape parameter that depends on time through the value of the conditional intensity at that instant. In light of this result, the issue of forecasting is tackled through both retrospective and prospective approaches. Retrospectively, the forecasting procedure is carried out on the occurrence times of the events recorded in each region, to determine whether the stress release model reproduces the observations used in the estimation procedure. Prospectively, the estimates of the time to the next event are compared with the dates of the earthquakes that occurred after the end of the learning catalog, in the 2003-2012 decade.

Keywords. Point process; Probabilistic forecasting; Interevent time distribution; Seismogenic sources; Bayesian inference.

1 Introduction

The formulation of stochastic models for seismic hazard assessment in probabilistic terms is essentially based on phenomenological analyses or physical hypotheses. Phenomenological analyses generate models that belong to the class of the self-exciting models (Hawkes & Oakes , 1974) that describe the temporal and spatial clustering of earthquakes (Kagan 1991; Ogata 1988, 1999; and references therein). These models were originally proposed to explain the decay of secondary shocks that follow a strong earthquake, and then they were applied for the detection of anomalies in seismic activity (Matsu'ura 1986; Ogata 1997). These empirical models aspire to provide a good descriptive fit to the data, but they do not necessarily strive for a context-specific physical explanation. Models based on physical hypotheses are more challenging, as these embody features that relate directly to the underlying scientific knowledge. Using these models, the aim is to explain how the evolution of the process depends on its history, in ways that can be interpreted in terms of the underlying mechanisms. Examples of such physical models are the block-slider, the branching for fractures, percolation, and cellular automata (Bhattacharyya & Chakrabarti et al. , 2006); these operate typically on small space-time scales. The most popular models that attempt to incorporate physical conjecture into the probabilistic framework and are concerned with large space-time scales are those included in the class of self-correcting models. In the seismological context, the elastic rebound theory still has the leading role, even though it was proposed a century ago by Reid (Reid , 1910). As a first approximation, modern measurements using global positioning systems (GPS) largely support the Reid theory as the basis of seismic movement along faults. Vere-Jones (1978) transposed this Reid theory into the framework of stochastic point processes, and in particular of the self-correcting models, through the first version of the stress release model. Enriched versions of this model have been extensively adopted for over 20 years now (Vere-Jones & Yonglu 1988; Zheng & Vere-Jones 1991, 1994; Bebbington & Harte 2003; Kuehn et al. 2008). One of their peculiarities is that they allow for possible interactions among neighboring fault segments as an explanation for the presence of clusters of even large earthquakes, in contrast to the quiescence that one would expect after a strong earthquake, according to the elastic rebound theory.

The stress release (hereinafter SR) model is based on a physical quantity that represents a proxy measure of the size of an earthquake, and that is generically indicated as 'stress'. Translating the 'elastic rebound theory' into stochastic terms, the occurrence probability in a SR model depends on the elastic stress stored on a fault, which is the result of its gradual accumulation due to tectonic forces, and of sudden releases during past earthquakes.

In this study, we focus on alternative choices for the proxy variable 'stress' to identify which physical quantity among those considered produces the best performance of the model. We propose four versions of the SR model in which the evolution of the system over time is represented by the amount of strain, seismic moment, seismic energy, or scaled energy. The values of these quantities for the events considered are obtained by integrating the available information on the most common input to probabilistic seismic hazard assessment, that is, the historical (macroseismic) and instrumental catalogs of seismicity, which are characterized by epicentral/ hypocentral location, origin time, and magnitude, and the map of seismogenic faults, as active faults deemed to be sources of large earthquakes and characterized by rupture parameters, such as area, mechanism, and magnitude.

In the literature the SR model was initially applied to strong earthquakes located in wide tectonic units, such as the northern China region (Vere-Jones & Yonglu , 1988). Then it turned out that the model fit can be improved by subdividing the region on the basis of seismicity, geophysical structure, and tectonic features, and by applying a different SR model to each subregion (Zheng & Vere-Jones 1991, 1994). Analogously, in Section 2, the four versions of the SR model are analyzed on a regional basis, by subdividing the Italian territory into eight large tectonically coherent zones, hereinafter called the macroregions (MRs). Using publicly available databases (Section 3), we put together eight datasets, one for each MR, that are constituted by earthquakes of $M_w \geq 5.3$ that are most likely associated with the fault sources that are included in each MR. Statistical treatment of the possible incompleteness of the recorded seismicity is also taken into account (Appendix A).

In Section 4, the model parameters are estimated following the Bayesian paradigm and applying Markov chain Monte Carlo (McMC) methods for sampling from the posterior probability distributions of the parameters. In this way, we obtain not only the parameter estimates, typically as their posterior means, but also a measure of their uncertainty, as expressed through the simulated posterior distribution of each parameter. In Section 4.2, the four models are compared one to the other through the Bayes factor and the Ando & Tsay information criterion (Ando & Tsay , 2010), to determine which among the proposed measures of the size of an earthquake provides the best fit to the data, and which resulting model shows the best predictive accuracy. We have also examined the various models in the light of the probability distribution $F(\omega_t|\mathcal{H}_t)$ of the 'time to next event' conditioned on the previous history $\mathcal{H}_t$ of the process. Results of the four SR models fitted to the data of each MR are shown in Section 5, and their performances are compared with each other and also with those of the Poisson model. Retrospective validation is performed by evaluation of the expected time to the next event immediately after each earthquake in the datasets (Section 5.2.2). The same analysis is then carried out in a prospective sense, which considers the earthquakes that occurred from the end of the learning catalog to the end of 2012 (Section 5.2.3). These test events were drawn from the available instrumental and parametric catalogs, while remaining as consistent as possible with the characteristics of the learning catalog.

All of the forecasts were carried out using data based on 2002 knowledge, as they were made available by the database compilers, so that our results are independent of subjective choices and only reflect the capability of the applied model in an actual context.

2 Self-correcting models

Let us take into account a region that can be considered as a seismic unit on the basis, for instance, of the kinematic context and the expected rupture mechanism, and with a sufficiently extensive historical record. Adopting the Reid elastic-rebound theory, we generically use the word 'stress' to indicate the quantity X that governs the state of the system in that region. We assume that X increases linearly with time at a constant loading rate ρ imposed by external tectonic forces, until it exceeds the strength of the medium. X then abruptly decreases each time an earthquake occurs. This hypothesis can be formalized by:

$$X(t) = X_0 + \rho\, t - S(t), \tag{1}$$

which expresses the variation of $X(t)$ over $t \in [0, T]$, where X_0 is the initial level of 'stress' and $S(t)$ is the accumulated 'stress' released by the earthquakes in the region at times $0 < t_i < t$, which is $S(t) = \sum_{i:t_i<t} X_i$. Assuming that the probability $\lambda(t)$ of instantaneous occurrence in $(t, t + dt)$ is a monotonic increasing function ψ of the 'stress' level, we have

$\lambda(t|\mathcal{H}_t) = \psi[X(t)]$ where $\mathcal{H}_t$ is the accumulated history of the process. In the original version of this model, given by Vere-Jones (1978), the form of the intensity function was $\lambda(t) = [\nu + \beta(t - \tau S(t)]^+$, where $[x]^+$ is 0 if $x < 0$; otherwise $[x]^+ = x$. Then, to guarantee the positivity of λ, an exponential function for ψ was chosen such that:

$$\lambda(t|\mathcal{H}_t) = \exp\{\nu + \beta X(t)\} = \exp\{\nu + \beta[X_0 + \rho\, t - S(t)]\} \tag{2}$$

with $\beta > 0$.

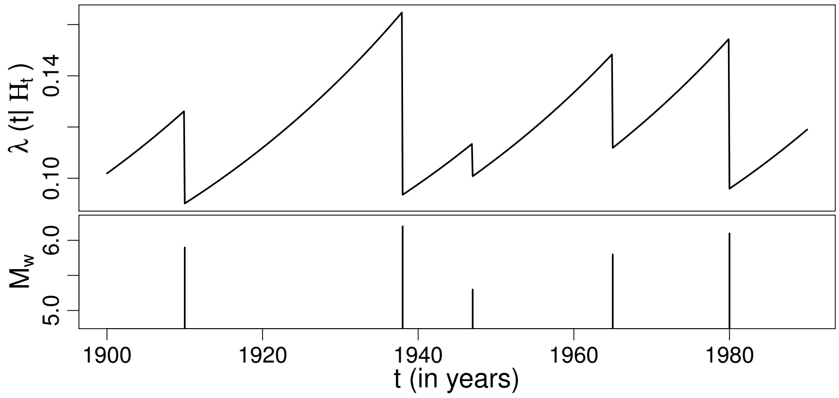

Figure 1: Representation of the conditional intensity function $\lambda(t|\mathcal{H}_t)$ of the stress release model (top); moment magnitude versus occurrence times of the related seismic dataset (bottom).

This implies that when $X(t)$ assumes a positive and larger value (i.e., low seismic activity), the intensity $\psi[X(t)]$ is also larger, and the occurrence probability increases; conversely, smaller negative values of $X(t)$ reduce the probability (Figure 1). This model belongs to the class of self-correcting point processes of Isham & Westcott (1979), with history-conditioned intensities. In other words, the model given by Equation (2) can be thought of in terms of the balance between the expected and observed values of the physical quantity X. In Equation (1), at each t_i, it can be seen that $X_0 + \rho t_i$ is the estimated 'stress' in the region, whereas $S(t_i)$ is the stress released by all of the earthquakes before t_i, and thus represents the lowest boundary of the stress estimate in the region. This line of reasoning implies that when the observed accumulated stress is lower than the expected, a seismic event is more likely to occur.

In Equation (2), X can be any physical parameter that constitutes a proxy measure of the strength of an earthquake, with the only constraint being that when dealing with long-term seismic hazard, this physical quantity can be evaluated from historical events. In the first applications of the stochastic model given by Equation (2) (Vere-Jones & Yonglu 1988; Zheng & Vere-Jones 1991, 1994), $X(t)$ is a scalar quantity - the Benioff strain - that can be calculated from:

$$\log_{10} X = \frac{1}{2} \log_{10} E = 0.75\ M_s + 2.4 \tag{3}$$

where E is the unknown seismic energy and M_s is the earthquake magnitude, which incorporates proportionality between the stress drop and the square root of the energy release (Benioff , 1951). To also take into account the contribution of energy lost to heat during an earthquake, the seismic moment M_0, given by:

$$\log_{10} M_0 = 1.5\ M_w + 9.1 \qquad (M_0 \text{ in } Nm), \tag{4}$$

(Kanamori & Brodsky , 2004) better represents the total seismic release. Note that M_s and M_w do not differ significantly for earthquakes with rupture lengths of 100 km or less (Kanamori , 1977).

The seismic moment depends on the coseismic displacement, and it is a static measure of the earthquake size related to its long-term tectonic effects. In contrast, the radiated energy is a dynamic measure of seismic potential for damage to anthropogenic structures. Hence energy and moment can be considered as complementary size measures in the estimation of seismic hazard. For recent earthquakes, however, the seismic energy computed through direct spectral analysis of broadband seismic waveforms can have significant regional and tectonic variations (Choy & Boatwright , 1995) that are largely neglected when using empirical formulae. In the case of historical earthquakes, ways to measure the amount of energy released that contain information on source, tectonic setting, and faulting mechanism can compensate for the inability to provide direct measurements of the energy.

Several studies have analyzed the scaling relationship for the apparent stress as a function of the seismic moment M_0, the rupture area A, and the average slip acceleration (Senatorski 2005, 2006). Considering different earthquake sets, from mining-induced, to small-to-moderate, and up to large earthquakes (Kanamori et al. , 1993), Senatorski (2007) deduced that the E-M_0 relationship is not linear, and the scatter in the $\log E$-$\log M_0$ plot can be noticeably reduced by taking into account the rupture area. Hence he proposed the relationship:

$$E \propto \frac{M_0^{1.5}}{\sqrt{A}}, \tag{5}$$

where A is the area of the fault surface that ruptures during an earthquake. Rupture area A is hereafter approximated by using the well-known regressions of Wells and Coppersmith (1994; see Section 4.1 for more details). Another influential seismic parameter that gives information on the rupture behavior (Kanamori & Heaton , 2000) is the scaled energy E_s, a non-dimensional radiated energy scaled with M_0, such that:

$$E_s = \frac{E}{M_0}. \tag{6}$$

Substituting the expression of Equation (5) for E in Equation (6), the following expression for the scaled energy is obtained:

$$E_s \propto \frac{{M_0}^{0.5}}{\sqrt{A}}. \tag{7}$$

In the present study, we examine the four different versions of the SR model (Eq. 2) that can be obtained by substituting X with the Benioff strain X_B (3), the seismic moment X_M (4), the seismic energy X_E (5), or the scaled energy X_S (7). The four models depend on the magnitude and threshold magnitude M_{th}, and are expressed by:

$$X_B = 10^{0.75\ (M_w - M_{th})}, \tag{8}$$

$$X_M = 10^{1.5\ (M_w - M_{th})}, \tag{9}$$

$$X_E = \frac{10^{2.25\ (M_w - M_{th})}}{\sqrt{A}}, \tag{10}$$

$$X_S = \frac{10^{0.75\ (M_w - M_{th})}}{\sqrt{A}}, \tag{11}$$

Hereinafter, we denote these models as $\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, and $\mathbf{R_S}$, respectively.

3 Databases

In the present study, we used two independently developed and publicly available databases (at the time this study was carried out): the Database of Individual Seismogenic Sources (DISS, version 3.0.2; DISS Working Group 2007), and the Parametric Catalog of Italian Earthquakes, version 2004 (CPTI04; CPTI Working Group 2004). These two databases reflect the level of knowledge at the end of 2002. To test our results we then used the most recent version of the Parametric Catalog of Italian Earthquakes, version 2011 (CPTI11; Rovida et al. 2011), which extends the records to 2006, and from 2007 onwards, we used the Italian Seismic Instrumental and parametric Data-base (ISIDe 2010; http://iside.rm.ingv.it/iside/standard/index.jsp).

3.1 Fault sources

DISS is a large repository of geological, tectonic and active fault data for Italy and the surrounding areas, which was compiled from first-hand experience of the authors and from a large amount of literature data (Basili et al. 2008; Basili et al. 2009). The database stores two main categories of parameterized crustal fault sources: Individual Seismogenic Sources (ISS) and Composite Seismogenic Sources (CSS), both of which are considered to be capable of releasing earthquakes of M_w 5.5 or greater. In most cases, the ISS represent the preferred source solutions of well-known large earthquakes of the past that ideally ruptured the fault from end to end (i.e., a fault segment). In recognition of the inherent difficulties in the identification of all possible fault segments in the tectonic record, however, in 2005 the DISS was extended to include the CSS, a source category that was also meant to expand the territorial coverage and completeness, and hence the capabilities, of the database. A CSS is essentially an active structure where the definition is based on regional surface and subsurface geological data that are exploited to identify and map entire fault systems. As opposed to the ISS, the termination of a CSS can be either an identified fault limit or a significant structural change. This implies that such fault sources can comprise an unspecified number of different potential ruptures, and can produce earthquakes of any size, at least in principle, up to an assigned maximum. The DISS (version 3.0.2) contains 81 such fault sources, most of which are located in Italy, whereas seven fault sources, which are not used in this study, are located in neighboring countries (Figure 2).

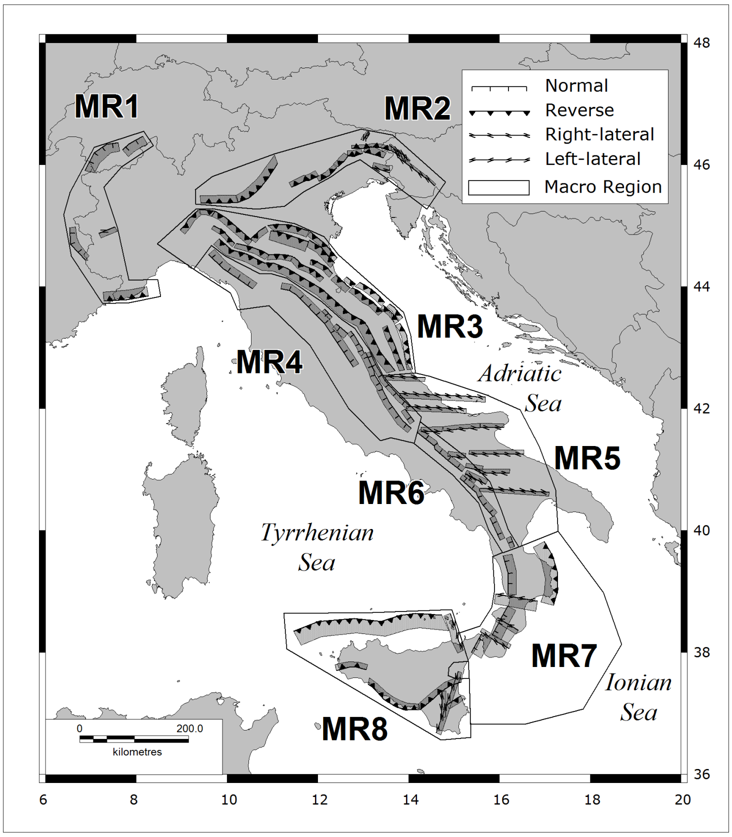

Figure 2: Map of the Composite Seismogenic Sources (CSS) from the DISS database, version 3.0.2 (DISS Working Group , 2007), classified according to the faulting mechanism. Shaded area: vertical projection of the fault plane to the ground surface. The outlined polygons are the MRs described in the text and Table 3.

3.2 Earthquakes

CPTI04 is a parametric catalog of earthquakes that exploits all of the sources of information that are available in historical documents and published scientific studies. The thresholds for including an earthquake in the catalog are as follows: for the pre-1980 section, macroseismic intensity I_0 = V-VI, evaluated through the Mercalli-Cancani-Sieberg scale (MCS), or $M_s = 4.0$; for the post-1980 section, $M_s = 4.15$; and for earthquakes located in the Etna volcano area, $M_s = 3.0$ (Figure 3).

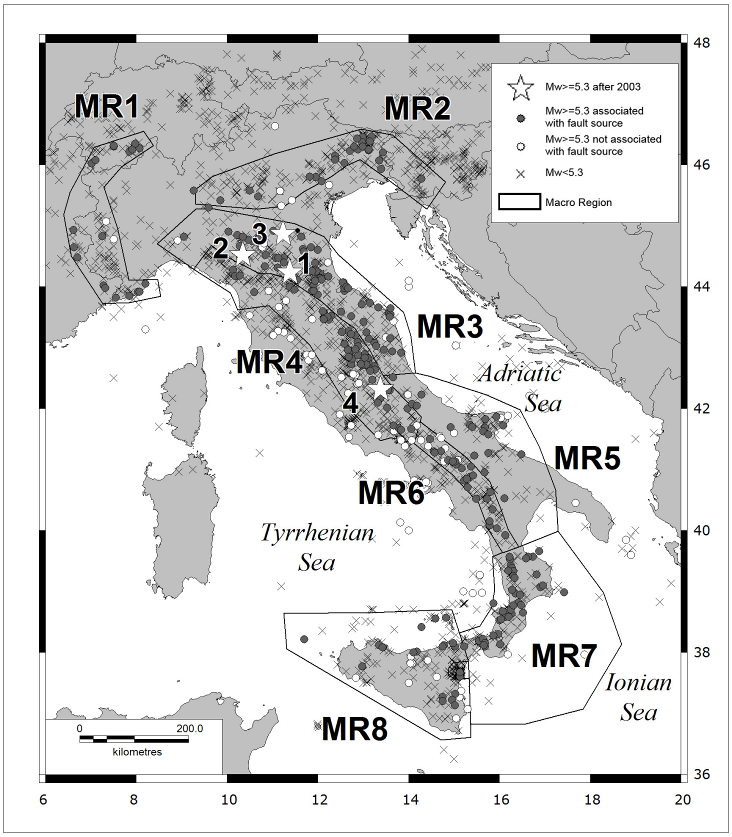

Figure 3: Map of earthquakes from the CPTI04 catalog (CPTI Working Group , 2004). The associations among the earthquakes, macroregions (MRs), and fault sources are listed in Tables 1-2. Stars indicate earthquakes that occurred after the end of the learning catalog, and were thus used to validate the forecast (see Section 5.2.3).

The catalog is supplied by the compilers in declustered form, such that the few historical events that were recorded within 90 days and 30 km from the principal events (mainshocks) in seismic sequences are not included. Each event in the catalog is characterized by its origin time, location, number of macroseismic intensity points, maximum and epicentral intensities, and moment and surface-wave magnitudes, which are based on empirical relationships for older events and on instrumental catalogs for modern events. ISIDe is a parametric catalog of seismicity that includes revised quasi-real-time earthquake locations based on data collected from the Italian National Seismic Network. The sizes of the events are given in the local magnitude scale (M_l). This catalog has been published twice a month since April 16, 2005.

region	CSS	fault type	date	M_w	region	CSS	fault type	date	M_w	region	CSS	fault type	date	M_w
MR_1	22	R	1644/02/15	5.88				1971/07/15	5.61				1791/01/00	5.37
			1818/02/23	5.55		12	R	1967/12/30	5.36				1815/09/03	5.37
			1819/01/08	5.34		18	R	1828/10/09	5.67				1859/08/22	5.70
			1831/05/26	5.54		20	R	1908/11/16	5.37				1873/07/12	5.40
			1854/12/29	5.77				1943/10/03	5.81				1874/12/06	5.47
			1887/02/23	6.29		27	R	1768/10/19	5.84				1904/02/24	5.67
	23	RL	1808/04/02	5.67				1781/06/03	6.23				1915/01/13	6.99
MR_2	2	R	1802/05/12	5.67				1799/07/28	5.93				1916/11/16	5.48
	7	R	1836/06/12	5.48				1869/06/25	5.32				1922/12/29	5.60
	48	R	1901/10/30	5.67				1873/03/12	5.88				1964/08/02	5.44
	61	R	1812/10/25	5.70				1873/09/17	5.52				1979/09/19	5.90
			1873/06/29	6.33				1918/11/10	5.79				1984/05/07	5.93
			1936/10/18	5.90				1937/12/10	5.42		26	N	1834/02/14	5.64
	62	R	1776/07/10	5.82				1951/09/01	5.31				1837/04/11	5.65
	64	R	1794/06/07	5.55				1972/11/26	5.34				1914/10/27	5.79
			1920/05/05	5.48				1998/03/26	5.33				1920/09/07	6.48
			1977/09/16	5.54		30	R	1916/05/17	5.85		28	N	1747/04/17	5.93
	66	R	1931/12/25	5.36				1916/08/16	5.92				1751/07/27	6.30
			1976/05/06	6.43		31	R	1972/06/14	5.40				1791/10/11	5.32
			1976/09/15	5.92		32	R	1897/09/21	5.50				1838/02/14	5.63
	67	LL	1788/10/20	5.71				1924/01/02	5.59				1997/09/26	6.05
			1908/07/10	5.34				1930/10/30	5.94		37	N	1745/03/00	5.37
			1924/12/12	5.53		39	R	1786/12/25	5.67				1767/06/05	5.44
			1928/03/27	5.75				1875/03/17	5.74				1789/09/30	5.80
MR_3	1	R	1781/04/04	5.84		44	R	1786/04/07	5.31				1832/01/13	5.80
			1781/07/17	5.53		46	R	1818/12/09	5.57				1854/02/12	5.37
			1813/09/21	5.32		47	R	1929/04/20	5.55				1878/09/15	5.55
			1870/10/30	5.59		49	R	1996/10/15	5.44				1910/06/29	5.37
			1911/02/19	5.38		51	R	1796/10/22	5.63				1917/04/26	5.80
			1935/06/05	5.34				1909/01/13	5.53				1919/06/29	6.18
			1963/08/09	5.32	MR_4	25	N	1703/01/14	6.81		40	N	1762/10/06	5.90
	8	R	1917/11/05	5.36				1703/02/02	6.65		41	N	2001/11/26	5.37
	9	R	1831/09/11	5.48				1719/06/27	5.32		56	N	1984/04/29	5.68
			1832/03/13	5.59				1730/05/12	5.85					

Table 1: List of the earthquakes in MR_1-MR_4 and their association to fault sources from DISS. Fault types: LL, left-lateral strike-slip; RL, right-lateral strike-slip; N, normal; R, reverse.

region	CSS	fault type	date	M_w	region	CSS	fault type	date	M_w	region	CSS	fault type	date	M_w
MR_5	3	RL	1941/08/20	5.37				1998/09/09	5.68		68	LL	1783/03/01	5.92
			2002/10/31	5.78		63	N	1694/09/08	6.87				1783/03/28	6.94
	5	RL	1846/08/08	5.33				1910/06/07	5.87				1821/08/02	5.37
			1990/05/05	5.84	MR_7	15	N	1767/07/14	5.83				1905/09/08	7.06
	58	RL	1841/02/21	5.40				1835/10/12	5.91				1947/05/11	5.71
			1875/12/06	6.07				1854/02/12	6.15				2001/05/17	5.60
			1889/12/08	5.55				1870/10/04	6.16		80	LL	1828/03/12	5.33
			1893/08/10	5.44				1886/03/06	5.56	MR_8	14	R	1613/08/25	5.57
			1948/08/18	5.58				1887/12/03	5.52				1726/09/01	5.61
	59	RL	1881/09/10	5.59				1913/06/28	5.65				1736/08/16	5.47
	75	RL	1950/09/05	5.73		16	N	1783/02/06	5.94				1739/05/10	5.54
	79	RL	1933/09/26	5.68				1908/12/28	7.24				1823/03/05	5.87
	84	N	1930/07/23	6.72				1909/07/01	5.55				1892/03/16	5.38
	89	RL	1851/08/14	6.33				1975/01/16	5.38				1940/01/15	5.34
MR_6	24	N	1688/06/05	6.72		19	R	1824/12/11	5.53				1979/12/08	5.44
			1702/03/14	6.32				1832/03/08	6.48				1980/05/28	5.71
			1732/11/29	6.61				1836/04/25	6.16				2002/09/06	5.89
			1805/07/26	6.57				1917/06/12	5.50		17	LL	1693/01/11	7.41
			1905/11/26	5.32				1932/01/02	5.62				1818/02/20	6.00
			1962/08/21	6.19				1983/11/08	5.37				1914/05/08	5.30
	34	N	1826/02/01	5.68		53	N	1743/12/07	5.79		21	R	1968/01/15	6.12
			1853/04/09	5.90				1783/02/05	6.91		35	LL	1624/10/03	5.57
			1857/12/16	6.96				1783/02/07	6.59				1818/03/01	5.63
			1980/11/23	6.89				1791/10/13	5.92		42	RL	1717/04/22	5.40
	38	N	1708/01/26	5.61				1928/03/07	5.90				1786/03/10	6.02
			1831/01/02	5.46		55	RL	1894/11/16	6.05				1926/08/17	5.32
			1836/11/20	5.83				1907/10/23	5.93				1978/04/15	6.06

Table 2: List of the earthquakes in MR_5-MR_8 and their association to fault sources from DISS. Fault types: LL, left-lateral strike-slip; RL, right-lateral strike-slip; N, normal; R, reverse.

3.3 Dataset construction

To carry out the model analysis in a regionalized way, we subdivided the Italian territory into eight large zones (see Table 3, Figures 2 and 3), which we refer to as the MRs (i.e., macroregions), because they are larger than the usual sizes of the zones in zonation models that are used for standard seismic hazard assessments in Italy.

ID	Name	Mechanism
MR_1	Western Alps	Mixed faulting mechanisms.
MR_2	Eastern Alps	Dominating south-verging thrust faulting mechanism with some strike-slip faulting in the easternmost portion of the MR (Slovenia).
MR_3	Central northern Apennines, east	Exclusively northeast-verging thrust faulting mechanism. Faulting depth is progressively shallower towards the northeast.
MR_4	Central northern Apennines, west	Exclusively normal faults with NE-SW extension axis that affect the crest of the Apennine mountain chain.
MR_5	Southern Apennines, Apulia	E-W trending right-lateral strike-slip faulting. Depth of faulting often deeper than in other regions.
MR_6	Southern Apennines, west	Exclusively normal faults with NE-SW extension axis that affect the crest of the Apennine mountain chain.
MR_7	Calabrian Arc	N-S to NE-SW trending normal faults, minor oblique-slip faults located inland, and thrust faults in the Ionian offshore. These last are mainly located in the overriding plate, and they are poorly mapped and difficult to associate with specific earthquakes.
MR_8	Sicily	Dominating thrust faulting, north-verging in the Tyrrhenian offshore, south-verging inland. Strike-slip faulting in the southwestern corner of Sicily.

Table 3: Faulting mechanisms in the MRs.

To construct these MRs, we aggregated zones from the seismic ZS9 zonation (Meletti et al. , 2008) based on their common tectonic characteristics, and refined the boundaries to include fault sources that belong to the same tectonic domain. Earthquakes from CPTI04 that are explicitly associated with an ISS based on geological/ geophysical studies in the DISS are also associated with the CSS, which contains the ISS. The remaining earthquakes are associated with the nearest CSS (Fracassi U. and Valensise G., personal communication). Hence each dataset represents the activity of a system of faults which belong to the same tectonic domain; this guarantees consistency with the assumptions underlying the SR model and agreement with the case studies proposed in the literature.

To allow for potential underestimation of the earthquake magnitudes, we considered all of the earthquakes with moment magnitude larger than 5.3. It is necessary to note that the algorithm used for the locating of historical events from macroseismic data used in CPTI04 cannot determine the hypocentral depth or reliably locate offshore events. The latter are automatically located near the coast, and can be mistaken for actual coastal events. To address the issue of the possible incompleteness of the catalog in the time span (T_0, T_f) covered by the data, we follow the statistical approach based on the detection of a changepoint in the occurrence rate function (Rotondi & Garavaglia , 2002); this point is meant as the beginning of the complete part of the catalog. The model and the estimation procedure are briefly recalled in Appendix A. Table 4 summarizes the results obtained for the eight MRs: $\hat{h}_2$ and $\check{s}$ are the estimates of the occurrence rates in the complete part and of the changepoint. The method adopted tends to place the estimate $\check{s}$ relatively close to t_1 (the time of the first earthquake occurred after T_0), where the unknown stress level could be high. This means that the analysis of the phenomenon started from a nonrandom point, but neglecting this piece of information. To overcome this issue we moved $\check{s}$ to T_c, so that the time interval that separates the beginning of the complete part of the catalog from the first event is equal to the average interevent time, which is calculated by also taking into account the censored observation related to the time elapsed between the latest event and T_f. Thus, we have the relationship:

$$T_c = t_1 - \frac{\sum_{i=1}^{n-1} (t_{i+1} - t_i) + (T_f - t_n)}{n-1}. \tag{12}$$

Extending the analyzed time interval in this way, no events are added to the original dataset. Thus, we start to observe the phenomenon when the stress level accumulated in the system is reasonably small, and a recharge period is roughly at the beginning. Note that the estimated changepoint of MR_1 falls beyond the most recent event (see 1887.15* in Table 4), which implies that the entire dataset can be considered as complete. Then, by applying Equation (12) to the data after 1600, we have the year 1584 as the initial time for the analysis.

Tables 1 and 2 list the earthquakes that make up the datasets analyzed below, which are sorted according to MR and fault source.

region	T_0	$\check{s}$	$\hat{h}_2$	T_c
MR_1	1448	1887.15*	0.0126	1584
MR_2	1197	1776.52	0.0676	1762
MR_3	1264	1781.25	0.164	1763
MR_4	1244	1703.03	0.120	1695
MR_5	1260	1841.13	0.0764	1829
MR_6	985	1688.42	0.0461	1667
MR_7	931	1767.53	0.108	1735
MR_8	1168	1613.64	0.0488	1593

Table 4: Completeness of the learning datasets according to MR: $\check{s}$ = posterior mode of the position of the changepoint, $\hat{h}_2$ = posterior mean rate, T_c = left end of the time interval under examination (see Equation (12)), * dataset considered as a complete set.

4 Bayesian inference and model comparison

A Bayesian approach to the analysis of the SR model is illustrated. Section 4.1 presents the Bayesian method for parameter estimation of the four versions of the SR model introduced in Section 2; then, Section 4.2 shows how these models can be tested through global summary measures of model performance and earthquake forecast procedures.

4.1 Parameter estimation

In this section, we deal with the problems of estimating the model parameters, and then of selecting the best model from the group of candidate models. Point processes are characterized by their intensity function $\lambda(t|\mathcal{H}_t)$ conditioned on the history $\mathcal{H}_t$ of the process itself. Hence, we have:

$$\lambda(t|\mathcal{H}_t) = \exp\left\{\nu + \beta[X_0 + \rho\, t - \sum_{i:t_i<t} X_i]\right\} \tag{13}$$

where X_i is the strain X_B (8), the seismic moment X_M (9), the seismic energy X_E (10), or the scaled energy X_S (11), depending on the version of the SR model under examination. The quantity X_i is released at time t_i by an earthquake where the magnitude is scaled by a threshold magnitude M_{th}. The rupture area involved in the expression of the seismic energy (5) and the scaled energy (7) is obtained as a function of the earthquake moment magnitude, by the regression $\log_{10} A_w = a{+}b\, M_w$ (Wells & Coppersmith , 1994), where the parameters a and b depend on the faulting type of the associated fault source. Specifically, $a = -2.87$ and $b = 0.82$ for normal fault (N), $a = -3.99$ and $b = 0.98$ for reverse fault (R), $a = -3.42$ and $b = 0.90$ for left/right-lateral strike-slip fault (LL/RL); Figure 4 represents the four proxy measures of the stress versus moment magnitude by taking into account the faulting types. Tables 1 and 2 provide the faulting types of each fault source.

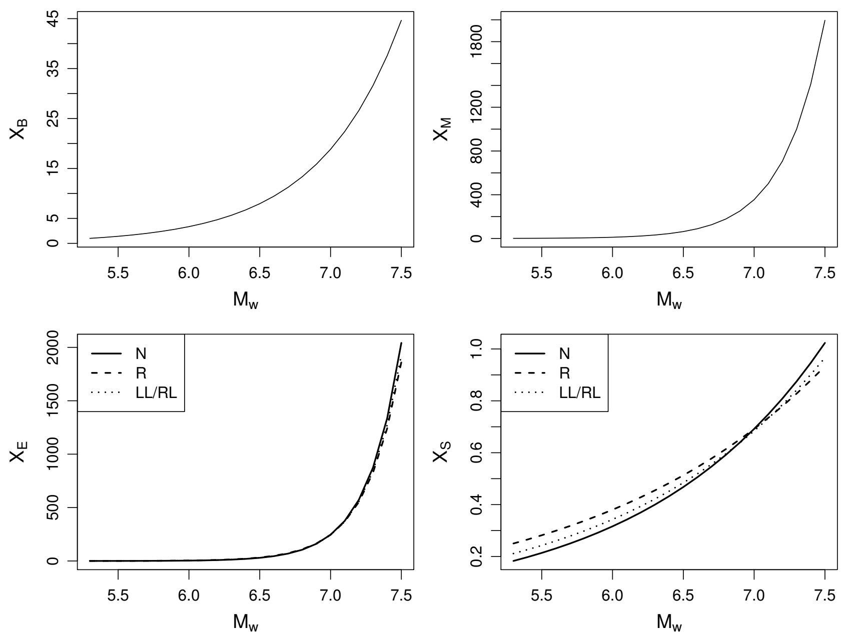

Figure 4: The strain X_B (top-left), the seismic moment X_M (top-right), the seismic energy X_E (bottom-left), and the scaled energy X_S (bottom-right) versus moment magnitude, where X_E and X_S are provided for different faulting types.

The parameter vector to be estimated is $\theta = (\alpha, \beta, \rho)$ where $\alpha = \nu + \beta\, X_0$ (see Equation (13)). According to the Bayesian paradigm, we assume the model parameters θ as random variables and formalize our beliefs about their variability, borrowed from the literature and previous experience, through prior distributions (e.g., as for the original version of the SR model, see Votsi et al. 2011; Jiang et al. 2011; Rotondi & Varini 2007). In our case, this information is not available because the SR model is here formulated in terms of moment and energy for the first time; moreover, the parameters α, β, ρ are not strictly related to easily measurable physical quantities. We then assign the prior distributions according to an objective Bayesian perspective, by combining the empirical Bayes method (Carlin & Louis , 2000) and the use of vague-proper prior distributions (Berger , 2006). We choose the families of the prior distributions according to the support of the parameters (β and ρ are positive parameters, and α lies on the real line), and we set the prior parameters (called *hyperparameters*) equal to the prior mean and variance of the corresponding model parameter; for instance, β follows *a priori* the Gamma distribution $Gamma(\xi, \nu)$ where $\xi = E_0(\beta)$ and $\nu = \mathrm{var}_0(\beta)$. According to the empirical Bayes method, preliminary values of the hyperparameters η are obtained by maximizing the marginal likelihood:

$$\eta_{\mathbf{EB}} = \arg\max_{\eta \in H}\; m(data \,|\, \eta) = \arg\max_{\eta \in H}\; \int_{\theta \in \mathbf{\Theta}} \mathcal{L}(data \,|\, \theta) \pi_0(\theta \,|\, \eta)\; d\theta \tag{14}$$

and by setting the standard deviations to 90% of the corresponding means, to avoid the estimates provided for the variances through the maximation (14) being too close to zero. This procedure clearly implies a double use of the data: in assigning the hyperparameters and in evaluating the posterior distributions. This philosophically undesirable double use can become a serious issue when the sample size is fairly small, as in our case. A solution is provided by choosing priors that 'span the range of the likelihood function' (Berger , 2006); that is, by varying the hyperparameters around their preliminary estimates $\eta_{\mathbf{EB}}$ and choosing those values that include most of the mass of the likelihood function, but that do not extend too far. For a graphic exemplification of this procedure we refer to Varini & Rotondi (2015).

In the Bayesian framework, the prior distribution of the parameter θ is given by π_0, and the log likelihood function is given by:

$$\log \mathcal{L}(data \mid \theta) = \sum_{i=1}^{N} \log \lambda(t_i) - \int_{T_c}^{T_f} \lambda(s)\ ds. \tag{15}$$

Through Bayes' theorem, the posterior distribution is given as:

$$\pi(\theta \mid data) = \frac{\mathcal{L}(data \mid \theta)\ \pi_0(\theta)}{\int_{\boldsymbol{\Theta}} \mathcal{L}(data \mid \theta)\ \pi_0(\theta)\ d\theta} \tag{16}$$

from which the estimate of the parameter can be obtained, which is typically given by the posterior mean, and measures of its uncertainty expressed through measures of location (median and mode), dispersion (variance and quantiles), and shape of the distribution (skewness and kurtosis). The explicit formulation of the posterior distribution generally requires the computation of multi-dimensional integrals. This can seldom be done in the closed form; numerical methods on integral approximations are a standard solution for this problem. Recently, methods based on the stochastic simulation of Markov chains have turned out to be highly efficient and flexible tools. McMC methods are a class of algorithms for sampling from probability distributions, which are based on constructing a Markov chain that has the desired distribution as its equilibrium distribution. The states of the chain after a large number of steps can be used as samples from the desired distribution. In the Bayesian context, the target distribution is the posterior distribution of the parameter θ. The algorithm applied to generate the Markov chains is summarized in Appendix B. Then diagnostic tools are applied to the sequences of the values generated for each parameter through pilot runs of the estimation algorithm, to test if it is safe to stop sampling and to use those sequences to estimate the characteristics of the posterior distributions, or if necessary, to vary the variance of the proposal distribution to reach the optimal acceptance rate so that a long run of the McMC algorithm guarantees the best estimates for the model parameters.

4.2 Model comparison

We provide an overview of the approaches for model comparison that are then applied in Section 5: the Bayes factor, the Ando & Tsay information criterion, and a retrospective analysis based on the probability distribution of the waiting time for the next event that is obtained from the SR model.

4.2.1 Bayes factor

We adopt the Bayesian approach to quantify the evidence in favor of one model in pairs of candidate models, through the Bayes factor. Given the models $\mathcal{M}_1$, $\mathcal{M}_2$, and the dataset $\mathbf{D}$, the Bayes factor is the ratio of the posterior odds of $\mathcal{M}_1$ to its prior odds; that is to say:

$$B_{12} = \frac{pr(\mathbf{D} \mid \mathcal{M}_1)}{pr(\mathbf{D} \mid \mathcal{M}_2)} = \frac{pr(\mathcal{M}_1 \mid \mathbf{D})}{pr(\mathcal{M}_2 \mid \mathbf{D})} \div \frac{pr(\mathcal{M}_1)}{pr(\mathcal{M}_2)} \quad . \tag{17}$$

When the prior probabilities of the two competing hypotheses are equal, the Bayes factor coincides with the posterior odds. The densities $pr(\mathbf{D} \mid \mathcal{M}_k)$, $k = 1, 2$ are obtained by integrating over the parameter space with respect to their prior distributions:

$$pr(\mathbf{D} \mid \mathcal{M}_k) = \int pr(\mathbf{D} \mid \theta_k, \mathcal{M}_k)\ \pi(\theta_k|\mathcal{M}_k)\ d\theta_k \tag{18}$$

where $\pi(\theta_k|\mathcal{M}_k)$ is the prior density of the parameter θ_k under $\mathcal{M}_k$, and $pr(\mathbf{D} \mid \theta_k, \mathcal{M}_k)$ is the likelihood function of θ_k. The quantity $pr(\mathbf{D} \mid \mathcal{M}_k)$ is a *marginal* (or *integrated*) *likelihood*; it is also referred to as *evidence* for $\mathcal{M}_k$. Details on the computational aspects concerning the evaluation of the Bayes factor can be found in Rotondi & Varini (2007).

4.2.2 Ando and Tsay information criterion

For each model, the Bayes factor considers the posterior probability induced by the prior distribution $\pi(\theta)$, and aims at the model comparison by looking for the best fit of the model to the data. Alternatively, one may be interested in the predictions from the various models and in choosing which model gives the best predictions of future observations generated by the same process as the original data. The predictive performance of a model $\mathcal{M}_k$ is assessed by scoring rules (Gneiting & Raftery , 2007); the most commonly used is the logarithmic score derived from the Kullback-Leibler distance between two distributions, the predictive distribution for new data $\mathbf{z}$ given the observations $\mathbf{y}$ and their true density $g(\mathbf{z})$:

$$\int \left[\log \frac{g(\mathbf{z}_n)}{pr(\mathbf{z}_n \mid \mathbf{y}_n, \mathcal{M}_k)}\right] g(\mathbf{z}_n)\ d\mathbf{z}_n$$

$$= \int \log[g(\mathbf{z}_n)]\ g(\mathbf{z}_n)\ d\mathbf{z}_n \ \ - \textstyle\int \log pr(\mathbf{z}_n \mid \mathbf{y}_n, \mathcal{M}_k)\ g(\mathbf{z}_n)\ d\mathbf{z}_n. \qquad (19)$$

The term relevant to the model $\mathcal{M}_k$ is the latter, which is the expected log-predictive likelihood where the unknown true density can be approximated by the empirical distribution $\tilde{g}(\mathbf{y}_n)$ constructed by the data so as to obtain as estimator the posterior predictive $\frac{1}{n}\ \log pr(\mathbf{y}_n \mid \mathbf{y}_n, \mathcal{M}_k)$. The accuracy of the predictions of future data is generally lower than the accuracy of the predictions of the same model for the observed data; then the resulting overestimation has to be corrected by applying some sort of bias correction. Following this approach, a variety of measures of predictive accuracy have been proposed in the literature, which are also referred to as information criteria; for instance, the Akaike information criterion (AIC) adopts the maximum likelihood estimate for θ, whereas the deviance criterion (DIC) uses the posterior mean $E(\theta \mid \mathbf{y}_n)$; for a review, we refer the reader to Vehtari & Ojanen (2012).

The Watanabe criterion (Watanabe , 2010) has the advantage of being fully Bayesian, because it averages the predictive distribution over the posterior distribution $\pi(\theta|\mathbf{y}_n)$ rather than conditioning on a point estimate. However, it is hardly applicable to data that, as in our case, are not independent given parameters. A solution is given by the Ando & Tsay criterion where the joint density can be decomposed into the product of the conditional densities $pr(\mathbf{y}_n \mid \theta) = \prod_{i=1}^{n} pr(y_i \mid y_{(1:i-1)}, \theta)$ (Ando & Tsay 2010, pgg. 747-748). The complete definition of this criterion is the following:

$$PL(\mathcal{M}_k) = \frac{1}{n}\ \left(\int \log pr(\mathbf{y}_n \mid \theta, \mathcal{M}_k)\ \pi(\theta|\mathbf{y}_n)\ d\theta\ -\ \frac{p}{2}\right), \qquad (20)$$

where, in the bias correction, p is the dimension of θ and the integral can be evaluated using draws from the posterior $\pi(\theta|\mathbf{y}_n)$ performed in the McMC estimation procedure, so that we have:

$$PL(\mathcal{M}_k) = \frac{1}{n}\left\{\log\left(\frac{1}{R}\sum_{j=1}^{R} pr(\mathbf{y}_n \mid \theta^{(j)}, \mathcal{M}_k)\right) - \frac{p}{2}\right\}. \tag{21}$$

To be on the same scale as the other criteria, we multiply Equation (21) by $-2n$.

4.2.3 Probability distribution of the 'time to next event'

For a more detailed analysis of the model performance we derive the probability distribution of the time to the next event for each class of SR model in an explicit way. This enables us to perform a retrospective analysis by comparing the occurrence time of each earthquake with its forecast value from the model. At the instant t, let us consider the conditional intensity function:

$$\lambda(t|\mathcal{H}_t) = \exp\{\alpha + \beta[\rho\, t - S(t)]\} \tag{22}$$

of the general SR model with parameter vector $\theta = (\alpha, \beta, \rho)$. Let W_t be the random waiting time for the next event given the history $\mathcal{H}_t$ up to t; hence, the occurrence time of the next event will be $T = t + W_t$. Hereinafter, for the sake of simplicity, we substitute the explicit indication of the conditioning on $\mathcal{H}_t$ with the subscript t.

The conditional cumulative distribution of W_t is given by:

$$\begin{aligned}
F_t(w \mid \theta) &= Pr(W_t \le w \mid \theta) = 1 - Pr(W_t > w \mid \theta) = 1 - Pr(N_{t+w} - N_t = 0 \mid \theta) \\
&= 1 - \exp\left(-\int_t^{t+w} \lambda(u)\, \mathrm{d}u\right) \\
&= 1 - \exp\left[-\frac{1}{\beta\rho}\left(e^{\alpha + \beta(\rho(t+w) - S(t))} - e^{\alpha + \beta(\rho t - S(t))}\right)\right] \\
&= 1 - \exp\left[-\frac{\lambda(t)}{\beta\rho}(e^{\beta\rho w} - 1)\right],
\end{aligned} \tag{23}$$

where N_s is the number of earthquakes recorded by time s. If we set $\phi_t = \lambda(t)/(\beta\rho)$ and

$\eta = \beta\rho$, then we have:

$$F_t(w \mid \theta) = 1 - \exp\{-\phi_t \ (e^{\eta w} - 1)\} \ , \tag{24}$$

which is a Gompertz distribution with shape parameter $\phi_t > 0$, scale parameter $\eta > 0$, and support $w \geq 0$. As the probability that an event occurs before a fixed time w increases with ϕ_t, the shape parameter ϕ_t can be interpreted as the propensity of the region to the occurrence. The probability density function is such that:

$$f_t(w \mid \theta) = \eta \phi_t e^{\eta w} e^{\phi_t} \exp(-\phi_t e^{\eta w}). \tag{25}$$

This function can take a large variety of shapes, and be skewed either to the right or the left. To describe the characteristics of the Gompertz distribution (24), we recall its summary statistics: mode, mean, variance, and quartiles (Lenart , 2014). The mode of the density function (25) is as follows:

$$w^* = \begin{cases} \dfrac{1}{\eta} \log \dfrac{1}{\phi_t}, \quad \text{with } 0 < \mathrm{F(w^*)} < 1 - \mathrm{e}^{(-1)} = 0.632 & \text{if } 0 < \phi_\mathrm{t} < 1 \\ 0 & \text{if } \phi_\mathrm{t} \geq 1. \end{cases} \tag{26}$$

The expected waiting time for the future event is such that:

$$E(W_t \mid \theta) = -\frac{e^{\phi_t}}{\eta} \mathrm{Ei}(-\phi_\mathrm{t}), \tag{27}$$

where Ei() is the exponential integral $\mathrm{Ei(x)} = -\int_{-\mathrm{x}}^{\infty} (\mathrm{e^{-u}/u})\ \mathrm{du}$, (Abramowitz & Stegun 1972, p. 228). On the one hand, according to the Reid theory, when ϕ_t (or equivalently $\lambda(t)$) gets close to 0, Equation (27) approaches ∞; i.e., after a large reduction in the hazard function $\lambda(\cdot)$ due to a very high 'stress' release, an unusually long waiting time should elapse before the next event. On the other hand, the expected waiting time can be short even when it is evaluated after relatively large earthquakes, because through the parameter ϕ_t it depends on the value that the hazard function has at the occurrence time. Indeed, if an earthquake of size X_i occurs at time t_i, the drop of the hazard function, $\Delta\lambda(t_i) = \lambda(t_i^-)\ [\exp(-\beta X_i) - 1]$, depends on the value of the hazard function $\lambda(t_i^-)$ computed immediately before the occurrence time. Consequently, variations in the hazard function caused by two events of the same size, but that occurred at different times, are typically different; hence, depending on the conditions of the system at that moment, the SR model does not preclude a small waiting time, even immediately after a strong event.

The variance of W_t is such that:

$$
\begin{aligned}
&V(W_t \mid \theta) \\
&= \frac{1}{\eta^2}\int_0^1 \log^2\left(1 - \frac{\log u}{\phi_t}\right) du - [E(W_t \mid \theta)]^2 \\
&= \frac{\phi_t e^{\phi_t}}{\eta^2}\left\{\frac{(\log^2\phi_t + 2\gamma\log\phi_t + \pi^2/6 + \gamma^2)}{\phi_t} - 2\ {}_3F_3\left[\begin{matrix}1,1,1\\ 2,2,2\end{matrix}\ ; -\phi_t\right]\right\} - [E(W_t \mid \theta)]^2
\end{aligned}
\tag{28}
$$

where $\gamma = 0.5772\ldots$ is the Euler-Mascheroni constant, and ${}_3F_3$ is the generalized hypergeometric function.

The generic quantile of order q is given by $W_q = \eta^{-1}\log(1-\phi_t^{-1}\log(1-q))$; hence, the median is equal to $\eta^{-1}\log(1-\phi_t^{-1}\log 0.5)$. Consistent with the definition of conditional intensity function, the hazard rate holds that $h_t(w \mid \theta) = f_t(w \mid \theta)/[1 - F_t(w \mid \theta)] = \phi_t\eta e^{\eta w} = \lambda(t)e^{\eta w} = \lambda(t+w)$, and hence it is an exponential increasing function.

In the case where additional time h has elapsed after the issue time t of the forecast, and no event has occurred during that time h, the distributions of the waiting times W_t and W_{t+h} can be compared. The second distribution is thus issued at time $(t+h)$, and it is enriched by the additional knowledge that no event has occurred between t and $t+h$. Since $\phi_{t+h} = \phi_t e^{\eta h} \geq \phi_t$ for all $h > 0$, the expected value of the waiting time W_{t+h} decreases as h increases; that is, $E(W_t \mid \theta) \geq E(W_{t+h} \mid \theta)$:

$$
\begin{aligned}
E(W_t \mid \theta) &= -\frac{e^{\phi_t}}{\eta}\mathrm{Ei}(-\phi_\mathrm{t}) = \frac{\mathrm{e}^{\phi_\mathrm{t}}}{\eta}\int_{\phi_\mathrm{t}}^{+\infty}\frac{\mathrm{e}^{-\mathrm{u}}}{\mathrm{u}}\,\mathrm{du} \overset{[\mathrm{u}=\phi_\mathrm{t}(\mathrm{z}+1)]}{=} \frac{1}{\eta}\int_0^{+\infty}\frac{\mathrm{e}^{-\phi_\mathrm{t}\mathrm{z}}}{\mathrm{z}+1}\,\mathrm{dz} \geq \\
&\geq \frac{1}{\eta}\int_0^{+\infty}\frac{e^{-\phi_{t+h}z}}{z+1}\,dz = E(W_{t+h} \mid \theta).
\end{aligned}
\tag{29}
$$

Moreover, it holds (Abramowitz & Stegun 1972, p. 229) that:

$$\frac{1}{2\eta} \ln\left(1 + \frac{2}{\phi_{t+h}}\right) < E(W_{t+h} \mid \theta) < \frac{1}{\eta} \ln\left(1 + \frac{1}{\phi_{t+h}}\right). \tag{30}$$

Therefore, as ϕ_{t+h} tends to infinity as h increases, the expected waiting time tends to zero as h grows to infinity and approaches its limit, with a convergence rate of $O(e^{-\eta h})$. Similarly, it can be shown that also the variance decreases to zero when h tends to infinity. For more details on the Gompertz distribution and further consideration of its application to other SR models we refer the reader to Varini & Rotondi (2015)

We recall that the Bayesian approach not only provides a point estimate of the parameters, but also a measure of their uncertainty in terms of the posterior distribution. Taking into account this uncertainty, the posterior predictive distribution of W_t is given by:

$$F_t(w) = P(W_t < w) = \int_{\boldsymbol{\Theta}} P(W_t < w \mid \theta)\ \pi(\theta \mid data)\ d\theta\ , \tag{31}$$

where the conditional Gompertz distribution of W_t is integrated with respect to the posterior distribution of the parameters. Pointwise approximation of the resulting probability distribution can be obtained by varying the model parameters into the Markov chains generated for their estimation (see Section 4.1):

$$F_t(w) \approx \hat{F}_t(w) = \frac{\sum_{j=1}^{R} P(W_t < w \mid \theta^{(j)})}{R}\ . \tag{32}$$

The expected value of the waiting time W_t is estimated by the average of the expected waiting times $E(W_t \mid \theta^{(j)})$, $j = 1, \ldots, R$, as given by (27); similarly for the variance of W_t, as the $\theta^{(j)}$ have negligible correlation, as indicated by the diagnostics on the convergence of the Markov chains. The mode of W_t can be evaluated through a numerical optimization algorithm (e.g., we use the direct search complex algorithm), which finds the waiting time in which the posterior predictive density function of W_t reaches the global maximum. The quantile of order q is the solution w_q of the equation $\hat{F}_t(w) = q$; we have solved this by the Müller method, as implemented in IMSL numerical libraries, version 4.0 (IMSL

, 2000). Through the quantiles, we then estimate the Highest Posterior Density (HPD) (or credible) interval of order q $(0 < q < 1)$ for the waiting time W_t, which is the time interval that satisfies the following two conditions: (a) the probability of that interval is q; and (b) the lowest density of any point within that interval is greater than or equal to the density of any point outside the interval. In other words, the most likely waiting times belong to the HPD interval, which turns out to be the smallest interval of order q.

The relationship $T = t + W_t$ links the time of the next event T with the corresponding waiting time W_t, and allows the estimation of the distribution $F(\cdot)$ of T and its summary statistics, so that it is possible to perform both retrospective and prospective validations.

5 Results

This section illustrates the results concerning both the parameter estimations and model comparisons related to the application of the four versions of the SR model to the data of each MR.

5.1 Parameter estimates

Details on the prior distributions used in the Bayesian inferential procedure are reported in Table B2. As illustrative examples, the prior and posterior densities of the parameters of the four models for MR_3 and MR_4 are shown in Figures B1 and B2, respectively. Table 5 collects parameter estimates of the different models obtained through the McMC algorithm by generating a chain of $R = 250,000$ elements, after discarding $50,000$ elements as burn-in, and recording the output every $20th$ iteration, for each parameter.

The α parameters for the four models of each MR are similar, and according to the order of their size, they are equal to the natural logarithm of the average number of events per year. The ρ parameters vary according to the stress proxy used in the model. Thus, e.g., in MR_4, for the middle value of the magnitude $M_w = 6.4$, the values of X_B, X_E, and X_S are about 16%, 42%, and 1%, respectively, of the value of X_M; analogously $\hat{\rho}_B$, $\hat{\rho}_E$, and $\hat{\rho}_S$ are 13.6%, 62%, and 1.3%, respectively, of $\hat{\rho}_M = 2.55$. As β and ρ behave inversely, $\hat{\beta}_E$ has the same order of size of $\hat{\beta}_M$, whereas $\hat{\beta}_B$ and $\hat{\beta}_S$ increases by one and two orders with respect $\hat{\beta}_M$.

	$\mathbf{R_B}$			$\mathbf{R_M}$		
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
MR_1	-5.65	3.14E-1	5.33E-2	-7.29	2.20E-1	1.38E-1
MR_2	-2.83	6.64E-2	1.66E-1	-2.87	1.09E-2	6.03E-1
MR_3	-1.72	3.02E-2	2.93E-1	-1.62	1.04E-2	6.04E-1
MR_4	-1.98	2.11E-2	3.48E-1	-2.02	1.21E-3	2.55
MR_5	-2.51	6.23E-2	2.26E-1	-2.57	3.78E-3	1.40
MR_6	-2.57	4.24E-2	2.66E-1	-2.58	3.49E-3	2.90
MR_7	-2.14	6.67E-3	6.36E-1	-2.18	3.56E-4	8.45
MR_8	-3.24	1.39E-2	2.69E-1	-2.18	3.51E-4	8.73
	$\mathbf{R_E}$			$\mathbf{R_S}$		
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
MR_1	-7.17	7.56E-1	4.24E-2	-4.96	1.38	7.31E-3
MR_2	-2.84	1.83E-2	2.47E-1	-2.92	9.92E-1	2.33E-2
MR_3	-1.63	2.30E-2	2.12E-1	-1.80	1.89E-1	5.20E-2
MR_4	-2.08	1.67E-3	1.59	-2.13	5.36E-1	3.31E-2
MR_5	-2.59	7.15E-3	6.96E-1	-2.45	1.63	2.27E-2
MR_6	-2.60	6.28E-3	1.68	-2.62	1.05	1.66E-2
MR_7	-2.19	4.44E-4	5.88	-2.17	1.70E-1	4.44E-2
MR_8	-3.27	3.72E-4	8.35	-3.61	7.89E-1	1.71E-2

Table 5: Parameter estimates for the $\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, and $\mathbf{R_S}$ models in each MR.

As an example, Figures 5 and 6 show the results for the estimate of the conditional intensity function that is obtained by applying the various models to the data from MR_3 and MR_4, which can be followed in two ways. The first is to replace the parameter estimates in the different versions of the expression (2), thereby obtaining the so-called plug-in estimate $\tilde{\lambda}(t) = \lambda(t \mid \hat{\theta}, \mathcal{H}_T)$, where $\hat{\theta}$ is the vector of posterior means. The second way is to estimate the conditional intensity through the ergodic mean $\hat{\lambda}(t) = \frac{1}{R}\sum_{j=1}^{R} \lambda(t \mid \theta^{(j)}, \mathcal{H}_T)$, where $\theta^{(j)}$ is the jth element of the Markov chain generated for each parameter by the McMC algorithm. Through the sequence $\left\{\lambda(t \mid \theta^{(j)}, \mathcal{H}_T)\right\}_{j=1}^{R}$, we can also obtain the median and quartiles of the pointwise estimate $\hat{\lambda}(t)$.

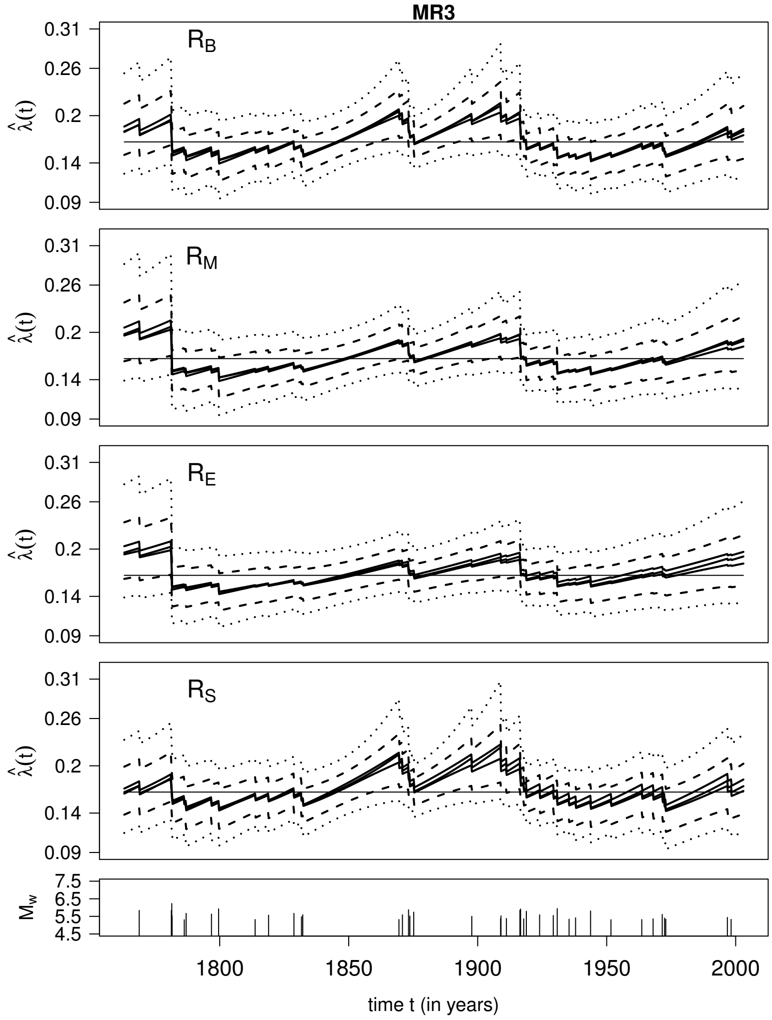

Figure 5: Conditional intensity function of the $\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, and $\mathbf{R_S}$ models. The ergodic mean, plug-in estimate, and median are all represented by solid lines that are practically indistinguishable from each other; 1^{st} and 3^{rd} quartiles (dashed line), 10% and 90% quantiles (dotted line). The Poisson rate is shown for comparison (horizontal thin line). The bottom panel shows the time history of the earthquakes scaled by their moment magnitudes (M_w). The example is taken from MR_3.

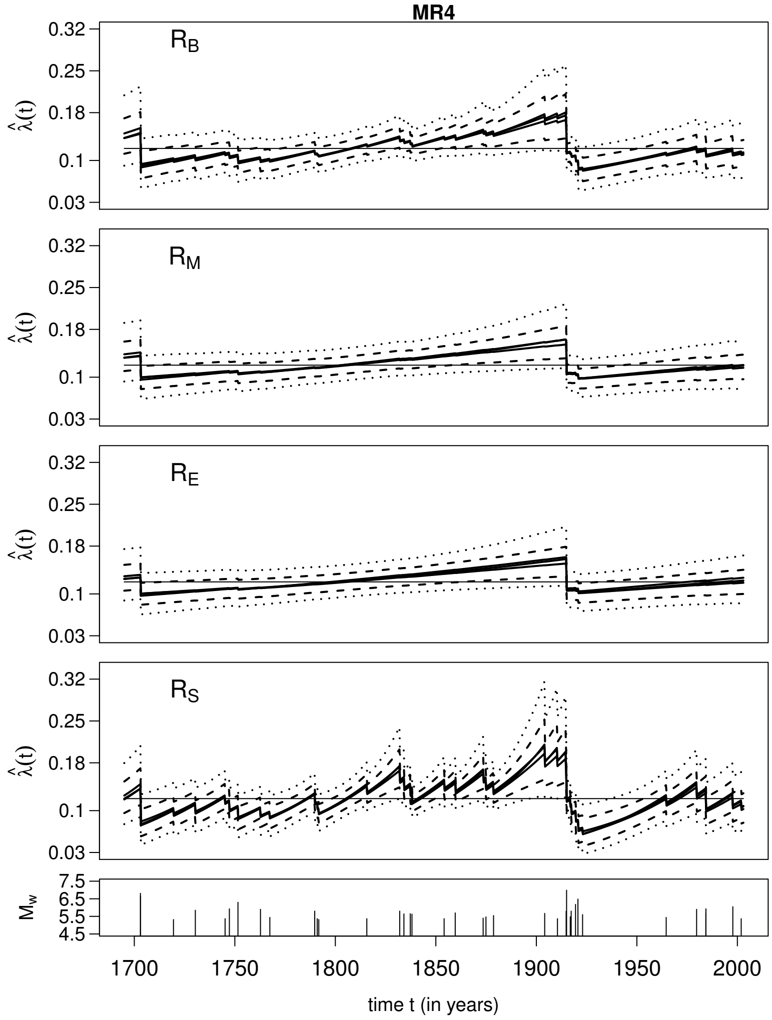

Figure 6: Same as Figure 5. The example is taken from MR_4.

5.2 Results of the model comparison

In this Section, we compare the four versions of the SR model to identify the best one; we note that what constitutes the "best" model is not uniquely defined, as it often depends on the goals of the user. Model testing can be performed for different purposes, such as the goodness of fit to the data of the learning set and the forecasting skill. To reach these aims, we propose two validation criteria: the Bayes factor that compares pairwise models through the ratio of their marginal densities with respect to the prior distributions of the parameters, and the information criterion by Ando and Tsay that averages the predictive distributions over the posterior distributions of the parameters.

5.2.1 Bayes factor

Table 6 shows the marginal $\log_{10}$ likelihood of each model, as applied to the various MRs, under the assumption that the prior probabilities of the models are equal; the maximum value represents the best model. In six out of eight MRs, the highest value is given by model $\mathbf{R_S}$, and in the remaining ones by model $\mathbf{R_E}$.

model / region	R_B	R_M	R_E	R_S
MR_1	-15.1469	-13.8686	**-13.5957**	-15.6580
MR_2	-27.3373	-27.5929	-27.5686	**-27.1243**
MR_3	-49.6949	-49.7956	**-48.9883**	-49.7344
MR_4	-50.3988	-50.6119	-50.6318	**-50.1548**
MR_5	-21.9602	-22.0761	-22.1250	**-21.4926**
MR_6	-28.2593	-28.2575	-28.2209	**-28.1026**
MR_7	-43.1532	-43.1471	-43.0928	**-43.0173**
MR_8	-35.3877	-35.4283	-35.3487	**-35.0062**

Table 6: Marginal $\log_{10}$ likelihood for the four stress release model versions. Bold: the maximum value, which indicates the best model in each MR.

More specifically, to evaluate the significance of these results, Table 7 shows the set of pairwise Bayes factors for each MR: according to the interpretation of Jeffreys' scale given by Kass & Raftery (1995), values in the three ranges of (0, 0.5), (0.5, 1), and (1, 2) of the $\log_{10} B_{12}$ indicate 'barely worth mentioning', 'positive', and 'strong' evidence, respectively, in favor of the model $\mathcal{M}_1$. Based on the Bayes factors, it can be seen that:

In MR_1, $\mathbf{R_E}$ behaves similarly to $\mathbf{R_M}$ ($\log_{10} B_{EM} = 0.27$ means that the evidence in favour of $\mathbf{R_E}$ is barely worth mentioning), whereas $\mathbf{R_E}$ shows strong evidence against $\mathbf{R_B}$ and $\mathbf{R_S}$;

In MR_2, there is slight evidence in favor of $\mathbf{R_S}$ compared to the other models, while $\mathbf{R_M}$ shows the worst performance;

In MR_3, $\mathbf{R_E}$ shows positive evidence against the other models, with $\mathbf{R_M}$ being the worst again;

In MR_4, there is slight evidence in favor of $\mathbf{R_S}$ compared to the other models, while $\mathbf{R_E}$ shows the worst performance;

In MR_5, $\mathbf{R_S}$ performs from slightly-to-moderately better than the other models; $\mathbf{R_E}$ shows the worst performance;

In MR_6, there is minimal evidence in favor of $\mathbf{R_S}$, and minimal evidence against $\mathbf{R_B}$;

In MR_7, as in MR_6;

In MR_8, $\mathbf{R_S}$ performs slightly better than the other models, with $\mathbf{R_M}$ being the worst.

MR$_1$

$\mathcal{M}_1$ \ $\mathcal{M}_2$	R_B	R_M	R_E	R_S
R_B	-	-1.2784	-1.5512	**0.5111**
R_M	**1.2784**	-	-0.2728	**1.7894**
R_E	**1.5512**	**0.2728**	-	**2.0623**
R_S	-0.5111	-1.7894	-2.0623	-

MR$_2$

$\mathcal{M}_1$ \ $\mathcal{M}_2$	R_B	R_M	R_E	R_S
R_B	-	**0.2556**	**0.2314**	-0.2130
R_M	-0.2556	-	-0.0243	-0.4686
R_E	-0.2314	**0.0243**	-	-0.4443
R_S	**0.2130**	**0.4686**	**0.4443**	-

MR$_3$

$\mathcal{M}_1$ \ $\mathcal{M}_2$	R_B	R_M	R_E	R_S
R_B	-	**0.1007**	-0.7067	**0.0394**
R_M	-0.1007	-	-0.8073	-0.0612
R_E	**0.7067**	**0.8073**	-	**0.7461**
R_S	-0.0394	**0.0612**	-0.7461	-

MR$_4$

$\mathcal{M}_1$ \ $\mathcal{M}_2$	R_B	R_M	R_E	R_S
R_B	-	**0.2131**	**0.2330**	-0.2440
R_M	-0.2131	-	**0.0199**	-0.4571
R_E	-0.2330	-0.0199	-	-0.4770
R_S	**0.2440**	**0.4571**	**0.4770**	-

MR$_5$

$\mathcal{M}_1$ \ $\mathcal{M}_2$	R_B	R_M	R_E	R_S
R_B	-	**0.1159**	**0.1648**	-0.4676
R_M	-0.1159	-	**0.0489**	-0.5835
R_E	-0.1648	-0.0489	-	-0.6324
R_S	**0.4676**	**0.5835**	**0.6324**	-

MR$_6$

$\mathcal{M}_1$ \ $\mathcal{M}_2$	R_B	R_M	R_E	R_S
R_B	-	-0.0018	-0.0384	-0.1567
R_M	**0.0018**	-	-0.0366	-0.1549
R_E	**0.0384**	**0.0366**	-	-0.1184
R_S	**0.1567**	**0.1549**	**0.1184**	-

MR$_7$

$\mathcal{M}_1$ \ $\mathcal{M}_2$	R_B	R_M	R_E	R_S
R_B	-	-0.0061	-0.0603	-0.1358
R_M	**0.0061**	-	-0.0542	-0.1297
R_E	**0.0603**	**0.0542**	-	-0.0755
R_S	**0.1358**	**0.1297**	**0.0755**	-

MR$_8$

$\mathcal{M}_1$ \ $\mathcal{M}_2$	R_B	R_M	R_E	R_S
R_B	-	**0.0406**	-0.0389	-0.3815
R_M	-0.0406	-	-0.0796	-0.4221
R_E	**0.0389**	**0.0796**	-	-0.3425
R_S	**0.3815**	**0.4221**	**0.3425**	-

Table 7: Bayes factors $\log_{10} B_{12}$ comparisons of the four stress release models, pair by pair (M_1 vs M_2), in every MR. Jeffreys' scale is used for rating the evidence in favor of M_1 models. Legend: bold, 0-0.5, 'barely worth mentioning'; gray striped, 0.5-1, 'positive evidence'; dark-gray striped, 1-2, 'strong evidence'.

Summarizing, we can say that the evidence in favor of $\mathbf{R_E}$ is sufficiently significant in MR$_1$ and MR$_3$, whereas in the other MRs, $\mathbf{R_S}$ performs just slightly better than the other models; anyhow, in all of the MRs the information on the faulting geometry provided through the rupture area (A) appears to improve the performance of the SR model. Note that MR$_1$ has only seven events associated with two fault sources and a poorly constrained tectonic setting; therefore, the results of this MR must be considered with caution. With reference to Equations (8-11), recalling that the rupture area is obtained by the regression $\log_{10} A_w = a + b\ M_w$ with $b \in \{0.82, 0.90, 0.98\}$ according to the faulting type (Wells & Coppersmith , 1994), it turns out that $X_B \propto 10^{0.75\ M_s}$, $X_M \propto 10^{1.5\ M_w}$, $X_E \propto 10^{(1.76,1.84)\ M_w}$, and $X_S \propto 10^{(0.26,0.34)\ M_w}$, where $(.,.)$ indicates the variability range of the magnitude coefficient. The same order of size of this coefficient in $\mathbf{R_B}$-$\mathbf{R_S}$ and $\mathbf{R_M}$-$\mathbf{R_E}$ can explain the similar performances of these models in the MRs where no or few events with $M_w \geq 6.5$ were recorded.

5.2.2 Retrospective forecast validation

Another tool to compare the performances of the four versions of the SR model is the analysis of their forecasting skill through retrospective forecast validation. Table 8 shows the value of the Ando & Tsay information criterion (Eq. (21)) for each model and for each MR. In the seven MRs of MR_2-MR_8, the highest value is given by model $\mathbf{R_S}$, and in the remaining MR_1, by model $\mathbf{R_E}$. These results agree with those provided by the Bayes factor, except for MR_3 where, however, the values are very similar. In all of the cases except MR_1, as the pairwise differences are less than 2, there is slight evidence in favor of these models.

region \ model	R_B	R_M	R_E	R_S
MR_1	70.7222	64.6214	**63.4842**	73.5678
MR_2	127.5156	128.6673	128.6785	**126.1706**
MR_3	226.8038	227.2734	227.3653	**226.5803**
MR_4	233.6563	234.7216	234.8703	**231.9754**
MR_5	102.7009	103.3292	103.4852	**100.2118**
MR_6	131.4636	131.4883	131.2725	**130.4955**
MR_7	200.3160	200.3123	200.1481	**199.6899**
MR_8	164.4944	164.6212	164.5529	**162.4373**

Table 8: Ando & Tsay information criterion (Eq. (21) times $-2n$) evaluated for the four stress release models. Bold: minimum value, which indicates the best model in each MR.

Another retrospective validation is carried out by evaluating the expected occurrence time of each earthquake (target event) included in each MR dataset, from right after the occurrence of the event that precedes it. The discrepancies between the expected times and the actual earthquake occurrence times can then be calculated. To this end, we use the Gompertz distribution (Equation 32) and its statistical summaries: mean, median, 75% HPD interval, and 90% HPD interval. Figure 7 provides two forecast examples: one, (retrospectively) issued for MR_1 on 1854/12/29, the date of the occurrence of a M_w 5.77 earthquake, shows a waiting time to the next event that relatively closely predicts the occurrence date of the 1887/02/23, M_w 6.29, earthquake; the other is issued for MR_2 on 1776/07/10, the date of occurrence of a M_w 5.82 earthquake, and closely predicts the waiting time to the 1788/10/20, M_w 5.71, earthquake. Note the different shapes of the two density functions that characterize the expected interevent times that vary from more than 30 years to about 12 years.

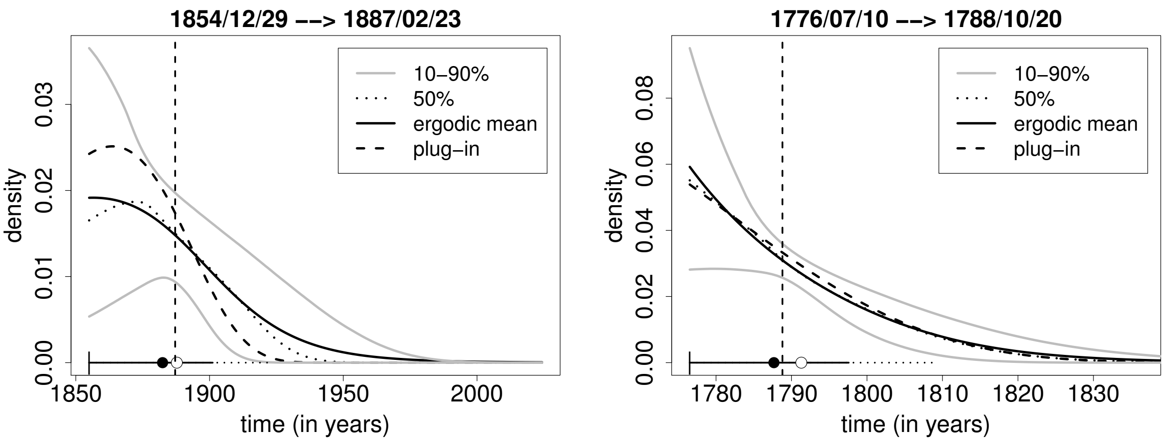

Figure 7: Examples of the estimated density functions of the time to next event, and their statistical summaries. Legend: Gompertz density function (solid curve), mean (open circle), median (solid circle), and 75% HPD (solid horizontal segment) and 90% HPD (dotted horizontal segment) intervals. The forecast issue date is denoted by a short vertical bar (|), and the occurrence time of the target event by a long, dashed, vertical line. The examples are taken from MR_1 (left) and MR_2 (right) and are based on the $\mathbf{R_E}$ and $\mathbf{R_S}$ models, respectively.

Table 9 summarizes the discrepancies of the forecasts for the four versions of the SR model for the eight MRs, in terms of the average lengths of the 75% HPD and 90% HPD intervals, as well as the mean absolute (root-mean-square) errors between the medians (means) and the observed times. For the absolute error, it is reasonable to compute its standard deviation, which is of the same order of magnitude as its mean in all of the MRs. In all of the MRs, the lowest values (or minimum discrepancy) essentially confirm the models chosen according to the Bayes factor (Table 7), except for MR_3; in this MR, even if the values of the indicators are very similar to each other, they support the model R_S in agreement with the Ando & Tsay information criterion (see Table 8). Hence, hereinafter we report the results provided by model $\mathbf{R_E}$ for MR_1, and by model $\mathbf{R_S}$ for the remaining MRs. The energy and the scaled energy again appear to be the appropriate quantities to be used in SR models.

Region	Model	HPD average length		Average discrepancy	
		90%	75%	Median	Mean
MR_1	R_B	94.4	65.9	27.9	40.7
	R_M	64.6	44.4	14.4	26.4
	R_E	**62.5**	**42.8**	**13.8**	**25.3**
	R_S	117.8	78.1	37.6	51.0
MR_2	R_B	33.8	**20.8**	9.1	12.6
	R_M	35.2	21.1	9.3	13.2
	R_E	35.3	21.0	9.2	13.1
	R_S	**31.8**	**20.8**	**8.8**	**11.7**
MR_3	R_B	14.0	**8.4**	**4.6**	**7.4**
	R_M	14.1	8.5	**4.6**	7.5
	R_E	14.1	8.5	4.7	7.5
	R_S	**13.9**	**8.4**	**4.6**	**7.4**
MR_4	R_B	19.9	12.0	**6.6**	8.8
	R_M	20.0	**11.9**	6.7	9.1
	R_E	20.0	**11.9**	6.7	9.1
	R_S	**19.3**	12.1	**6.6**	**8.6**
MR_5	R_B	31.1	19.3	8.7	12.1
	R_M	32.0	**19.0**	8.7	12.6
	R_E	32.1	**19.0**	8.8	12.7
	R_S	**27.8**	19.2	**8.1**	**10.7**
MR_6	R_B	50.5	32.4	12.8	17.7
	R_M	51.6	32.7	13.2	18.0
	R_E	51.9	33.1	13.2	18.1
	R_S	**48.1**	**32.3**	**12.2**	**17.0**
MR_7	R_B	21.1	12.6	6.9	8.5
	R_M	21.1	12.6	6.9	8.5
	R_E	21.0	**12.5**	6.9	8.5
	R_S	**20.8**	12.6	**6.8**	**8.3**
MR_8	R_B	50.6	30.4	14.6	19.9
	R_M	51.2	30.5	14.7	20.2
	R_E	51.1	30.5	14.7	20.1
	R_S	**46.4**	**30.1**	**14.2**	**18.3**

Table 9: Ability of retrospective forecasting of the four stress release models in each MR, in terms of the following indicators: average length of the 75% and 90% HPD intervals, the mean absolute (root-mean-square) error between the expected median (mean) and observed occurence times. Bold, lowest values.

Figure 8 shows the results of the retrospective validation of all of the data in MR_3 by representation of the statistical summaries of the estimated Gompertz density functions (see examples in Figure 7). The results for the other MRs are shown in Appendix C (Figures C1-C7). In these figures the reliability of the forecasts is expressed as the time discrepancy with respect to the actual occurrence of the targeted event. As a visual tip, for comparing the various discrepancies one with the other, time lines are vertically aligned with respect to the actual occurrence time of the target events. Forecasts to the right of the alignment thus correspond to overestimations of the interevent time, and the opposite for those to the left. In the case of MR_3, the actual event time is outside the 90% HPD interval only for 4 of the 39 events examined, whereas for 30 events it is within the 75% HPD interval.

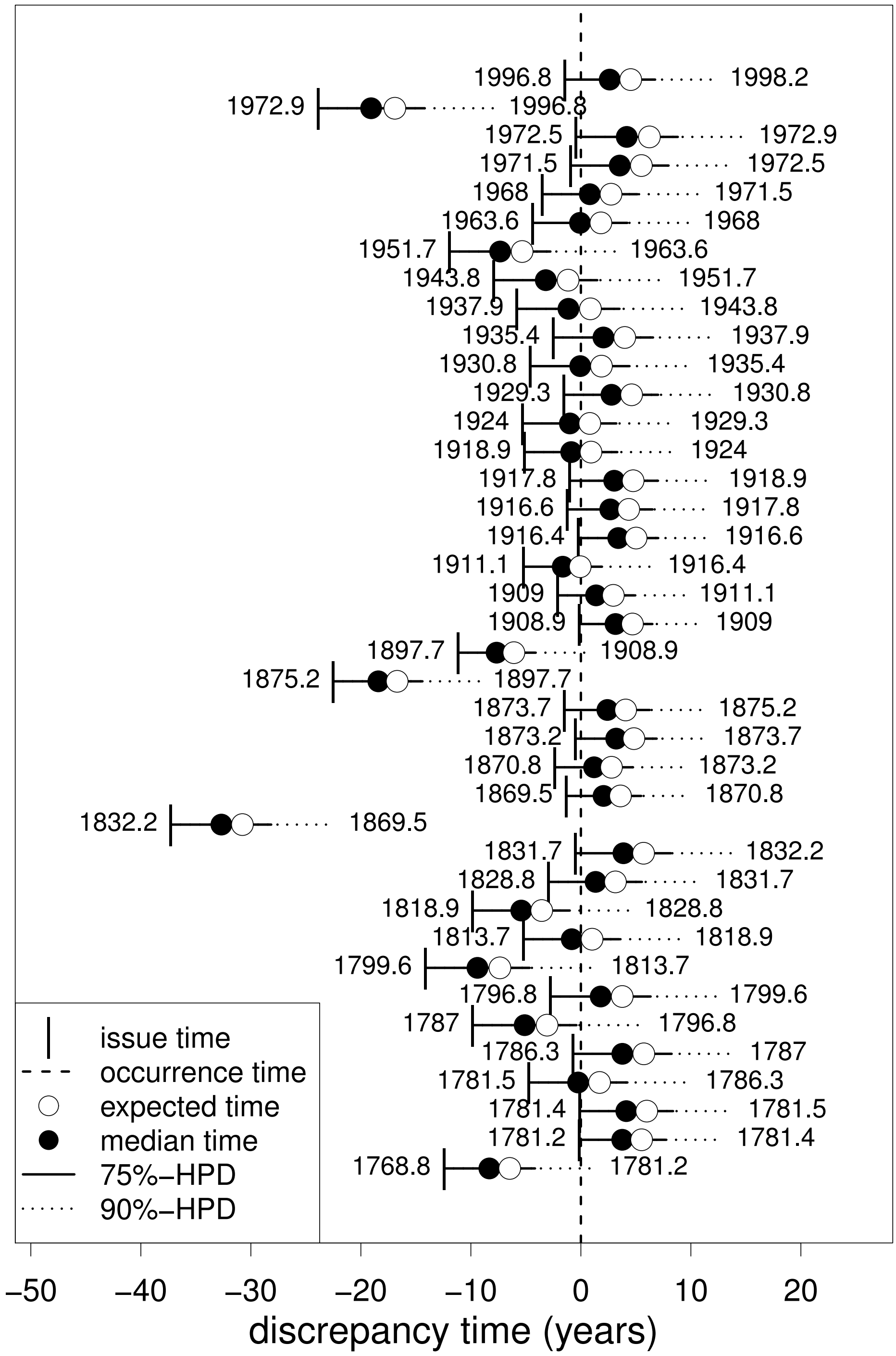

Figure 8: Time lines of 39 retrospective forecasts for MR_3, $\mathbf{R_S}$ model, in order of descending date from the top to the bottom. Each forecast is as would have been issued on the occurrence date (shown on the left, and marked by a short vertical bar) of an event in the MR dataset, and is aimed at predicting the date (on the right) of the next event (target). The forecasts are shown by the statistical summaries of their Gompertz density functions (see Figure 7). The time lines are shifted laterally so that they intersect the vertical dashed line at the actual occurrence date of the target event.

5.2.3 Prospective forecast validation

To conduct prospective validation, there is the need to first determine which earthquakes that occurred since the beginning of 2003 are consistent with the learning dataset used. To this end, we used CPTI11 for the period from 2003-2006, and ISIDe for the period from 2007-2012 (see Section 3), where there were the following four earthquakes:

1. 2003/09/14, $M_w = 5.29 \pm 0.09$ (from CPTI11), Bolognese Apennines, reverse faulting, MR_3;

2. 2008/12/23, $M_w = 5.4$, ($M_l = 5.2$, from ISIDe), Parma, reverse faulting, MR_3;

3. 2012/05/20, $M_w = 5.9$ ($M_l = 5.9$, from ISIDe), Finale Emilia, reverse faulting, MR_3;

4. 2009/04/06, $M_w = 6.1$ ($M_l = 5.9$, from ISIDe), L'Aquila, MR_4.

The CPTI11 catalog assigns earthquake #1 a magnitude that is very close to the threshold ($M_w \geq 5.3$) that we considered for the learning phase. However, Rovida et al. (2011) reported that the use of new empirical relations in CPTI11 decreases the magnitudes < 5.5 and increases those > 5.5, with respect to the CPTI04. Therefore, according to the rules of our learning catalog (CPTI04), the 2003/09/14 earthquake would be likely to be beyond the threshold, and we thus include it in the validation procedure with $M_w = 5.3$. The three earthquakes with $M_w \geq 5.3$ that occurred in the period 2007-2012 (#2, #3, and #4) are taken from ISIDe by exclusion of their aftershocks, i.e., for homogeneity with the CPTI04 declustering, the events that occurred within 30 km and 90 days are excluded. Note also that ISIDe uses local magnitude (M_l), and thus we obtain M_w values using the same conversion formula ($M_w = 0.812\ M_l + 1.145$) used for the compilation of CPTI04 (MPS Working Group 2004 , 2004).

The various magnitude determinations for earthquake #4 span a wide range that depends on the co-existence of source and path complexities and heterogeneities in the local seismic response (Ameri et al. , 2012). The most significant magnitude values are: $M_l = 5.9$, based on the INGV seismic bulletin from ISIDe; $M_w = 6.08$, based on the time-domain moment tensor (Scognamiglio et al. , 2010); $M_w = 6.13$, based on the regional moment tensor (Herrmann et al. , 2011); $M_l = 6.08 \pm 0.17$, based on the Huber mean of accelerometric determinations (Maercklin et al. , 2011); and $M_w = 6.3$, based on the regional centroid moment tensor (Pondrelli et al. , 2010). We thus adopt $M_w = 6.1$, as this appears to be the most frequent value.

Table 10 summarizes the prospective forecasts provided by the $\mathbf{R_E}$ model for MR_1, and by the $\mathbf{R_S}$ model for the other MRs. Note that the forecast issue dates considered here are: the date of the latest event in each MR learning dataset; the end date of the learning catalog (end of 2002, everywhere); the date when any earthquake occurred in each MR over the years 2003-2012 (in our case in MR_3 and MR_4); and the beginning of 2013. Forecasts are addressed in terms of the probability distribution of the time to the next event, as summarized by the median, the mean, and its standard deviation, as well as by the 75% HPD and 90% HPD intervals.

In MR_4, after the last observed event in the learning catalog (2001/11/26; Table 10, first line in the MR_4 block), it can be expected that the next earthquake with $M_w \geq 5.3$ will be in early 2011 according to the mean, with a standard deviation of ± 8.4 years; or by 2008.4, 2014.7, or 2022.7 with probabilities of 50%, 75%, and 90%, respectively. A little more than a year later (2003/01/01; Table 10, second line), by adding the information that no event had occurred in the meanwhile, the expected time to the next event moves forward by a year. This additional information not only lengthens the waiting time to the next event, but also reduces the uncertainty on the HPD interval length. After the 2009/04/06 earthquake (Table 10, third line), the estimation of the model parameters is fully repeated when the new earthquake is added to the dataset. Based on the seismic and tectonic knowledge available in 2002, and reinforced only with the addition of about 10 years of seismic history (Table 10, fourth line), the $\mathbf{R_S}$ model predicts that the next earthquake with $M_w \geq 5.3$ in MR_4 can be expected in 2022, according to the mean value, or by 2019.5, 2025.8, and 2033.7, with probabilities of 50%, 75%, and 90%, respectively.

Region	Date of forecast issue	HPD 75%	HPD 90%	Median	Mean (st.dev.)
MR_1	1887.2	2046.6-2327.3	1985.7-2401.5	2190.8	2198.1 (54.5)
	2003.0	2040.9-2304.0	2003.0-2359.0	2191.7	2204.2 (48.9)
	2013.0	2044.4-2301.0	2013.0-2360.1	2193.9	2207.3 (48.3)
MR_2	1977.7	1977.7-2009.8	1977.7-2025.9	1994.9	1999.9 (16.9)
	2003.0	2003.0-2024.0	2003.0-2036.9	2013.8	2018.1 (13.4)
	2013.0	2013.0-2033.0	2013.0-2045.9	2023.2	2027.7 (13.4)
MR_3	1998.2	1998.2-2006.7	1998.2-2012.5	2002.4	2004.4 (6.1)
	2003.0	2003.0-2011.1	2003.0-2016.7	2007.0	2008.9 (5.9)
	2003.7[(a)]	2003.7-2011.6	2003.7-2016.9	2007.6	2009.4 (5.7)
	2009.0[(b)]	2009.0-2016.8	2009.0-2022.1	2012.8	2014.7 (5.6)
	2012.4[(c)]	2012.4-2020.2	2012.4-2025.5	2016.3	2018.1 (5.6)
	2013.0	2013.0-2020.8	2013.0-2026.2	2016.9	2018.7 (5.6)
MR_4	2001.9	2001.9-2014.7	2001.9-2022.7	2008.4	2011.0 (8.4)
	2003.0	2003.0-2015.6	2003.0-2023.5	2009.4	2012.0 (8.3)
	2009.3[(d)]	2009.3-2022.3	2009.3-2030.4	2015.9	2018.5 (8.5)
	2013.0	2013.0-2025.8	2013.0-2033.7	2019.5	2022.1 (8.3)
MR_5	2002.8	2002.8-2019.2	2002.8-2029.5	2011.5	2015.0 (11.1)
	2003.0	2003.0-2019.3	2003.0-2029.6	2011.6	2015.2 (11.1)
	2013.0	2013.0-2028.4	2013.0-2039.9	2020.7	2024.8 (11.8)
MR_6	1998.7	1998.7-2029.0	1998.7-2047.3	2014.7	2020.8 (19.7)
	2003.0	2003.0-2031.7	2003.0-2049.6	2018.0	2024.1 (19.2)
	2013.0	2013.0-2040.5	2013.0-2059.0	2027.0	2033.5 (19.4)
MR_7	2001.4	2001.4-2012.5	2001.4-2020.4	2006.8	2009.6 (8.4)
	2003.0	2003.0-2014.0	2003.0-2021.9	2008.4	2011.1 (8.3)
	2013.0	2013.0-2025.0	2013.0-2033.7	2018.9	2021.9 (9.1)
MR_8	2002.7	2002.7-2035.0	2002.7-2053.8	2019.4	2025.5 (19.5)
	2003.0	2003.0-2035.3	2003.0-2054.0	2019.7	2025.7 (19.5)
	2013.0	2013.0-2043.9	2013.0-2061.9	2029.1	2035.0 (19.1)

[(a)] just after 2003/09/14 earthquake, M_w 5.3
[(b)] just after 2008/12/23 earthquake, M_w 5.4
[(c)] just after 2012/05/20 earthquake, M_w 5.9
[(d)] just after 2009/04/06 earthquake, M_w 6.1

Table 10: Prospective forecasts according to the R_E model in MR_1, and to the R_S model in the other MRs. All of the dates are expressed in decimal years. The estimated probability distribution of the time to the next event is expressed as: 75% and 90% HPD intervals, median, mean, and standard deviation (years).

In MR_3, three earthquakes occurred in the period 2003-2012, and thus the forecasts can be successively updated after each one of these. Note that all of these successive forecasts fall within the 75% HPD interval, and that the average absolute error of the forecast time for all three of these occurrences is 1.7 years when considering the median values, whereas the root-mean-square error is 3.29 years when considering the mean values.

We note that the model parameters are fully re-estimated after every new earthquake, by its inclusion in the learning dataset of the MR. The robustness of these parameter estimates is shown by the similar intensity functions (Figure 9) they allow, and the similar values they achieve (Table B1).

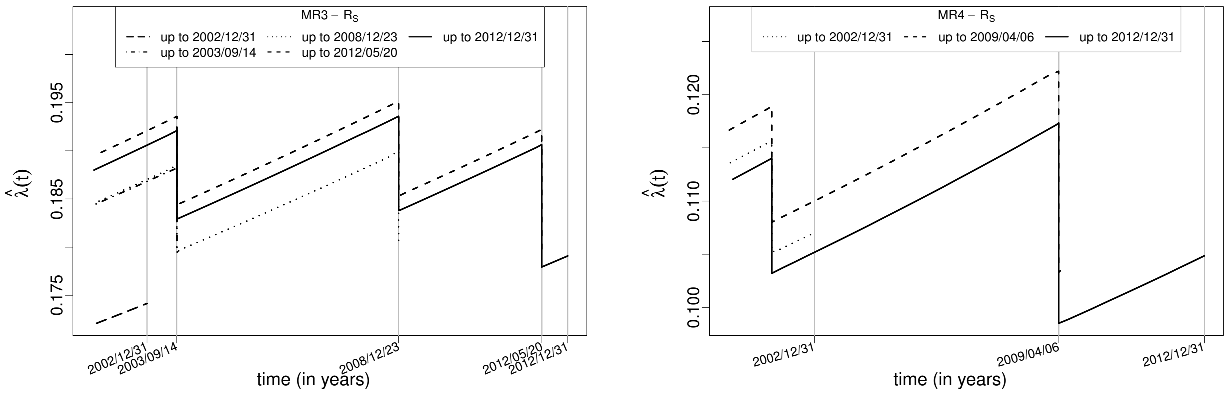

Figure 9: Estimate (ergodic mean) of the intensity function for the $\mathbf{R_S}$ model in MR_3 and MR_4, updated whenever new information (i.e., earthquake occurrence) is included in the relevant dataset.

For completeness of information, Table C1 provides a summary of all of the forecasts issued at the end of the learning catalog (end of 2002) for the four versions of the SR model for each MR.

5.3 Comparison with the Poisson model

The Poisson model is a time-independent point process that is defined by its conditional intensity function $\lambda(t) = e^{\alpha}$, where α is a real parameter; in particular, a SR model where the b parameter tends to zero is a Poisson model. In this view, it is apparent that the SR model is conceived as a time-dependent version of the Poisson model, and its conditional intensity function is expected to evolve in time around an average rate according to the variation of the level of 'stress' in the region. To compare the performances of the Poisson and SR models, the results on the Bayesian analysis of the Poisson model for each MR are summarized below. Similar to the results in Table 5, the Poisson parameter α is estimated for each MR; from MR_1 to MR_8 respectively, these estimates are -4.11, -2.67, -1.78, -2.12, -2.51, -3.05, -2.16, and -3.03. Similar to Tables 6 and 7, Table 11 shows the estimated values of the marginal log_{10} likelihoods and the Bayes factors between the versions of the SR and Poisson models.

For the marginal log_{10} likelihood, the Poisson model behaves worse than the best SR model for each MR. Based on the Bayes factor, we can note: positive/ strong evidence in favor of model R_E in MR_1 and MR_3; on the whole, positive/ strong evidence in favor of the SR models in MR_4, MR_7, and MR_8; and slight evidence in favor of R_S in the remaining MRs.

Region	marg. $\log_{10}\mathcal{L}$	$\log_{10} B_{12}$ R_B	R_M	R_E	R_S
MR_1	-15.8748	**0.7279**	**2.0063**	**2.2791**	**0.2168**
MR_2	-27.3635	**0.0262**	-0.2294	-0.2052	**0.2392**
MR_3	-49.7749	**0.0800**	-0.0207	**0.7866**	**0.0405**
MR_4	-52.3692	**1.9704**	**1.7573**	**1.7374**	**2.2144**
MR_5	-21.7897	-0.1705	-0.2864	-0.3353	**0.2971**
MR_6	-28.5106	**0.2513**	**0.2531**	**0.2897**	**0.4080**
MR_7	-43.7551	**0.6019**	**0.6080**	**0.6622**	**0.7377**
MR_8	-36.1995	**0.8118**	**0.7712**	**0.8507**	**1.1933**

Table 11: Global summary measures of the performance of the Poisson model in each MR: (marg.$\log_{10}\mathcal{L}$), marginal $\log_{10}$ likelihood; ($\log_{10} B_{12}$), logarithm of the Bayes factors of the four SR models, M_1, versus Poisson model, M_2. As for the Bayes factor, the Jeffreys' scale is used for rating the evidence in favor of M_1 models: bold, 0-0.5, 'barely worth mentioning'; gray striped, 0.5-1, 'positive evidence'; dark-gray striped, 1-2, 'strong evidence'.

Table 12 shows the results of the retrospective forecast validation by applying the Poisson model to each MR. We recall that according to the Poisson model, the waiting time to the next event is exponentially distributed with mean $e^{-\alpha}$, and consequently the forecast is time-independent. By comparing this with the results in Table 9, we note that the 90%-HPD intervals and all of the average discrepancies between the observed occurence times and the forecasted values estimated by the best SR model are less than those of the Poisson model, whereas the 75%-HPD intervals related to the Poisson model are narrower.

Taking the cue from this slightly larger uncertainty of the forecasts issued by the SR model, we highlight that the values in Table 9 are computed immediately after an event and that they can be updated as time passes and no occurrence happens, obtaining a reduction in the 75% and 90% HPD intervals of the waiting time variable (as shown in Table C1); of course this is not possible with the homogeneous Poisson model, for which the mean and variance of the waiting time do not depend on the time elapsed since the last event. This is more clearly depicted in Figure 10; through the model R_S, we calculate the forecasts issued immediately, and 10, 20 and 30 years since the 1922/12/29 earthquake in MR_4. We note that the forecasts are modified based on the additional information on nonoccurrence, and that the average waiting times and HPD intervals are shortened.

Region	HPD length		Average discrepancy	
	90%	75%	Median	Mean
MR_1	154.1	87.1	41.9	61.7
MR_2	34.5	**20.2**	9.1	12.4
MR_3	**13.9**	**8.3**	4.8	7.6
MR_4	19.6	**11.7**	6.7	9.2
MR_5	29.7	**17.3**	8.3	12.3
MR_6	50.8	**29.8**	14.4	20.2
MR_7	**20.4**	**12.1**	**6.8**	8.5
MR_8	49.1	**29.0**	14.7	20.3

Table 12: Ability of retrospective forecasting of the Poisson model in each MR, in terms of the following indicators: length of the 75% and 90% HPD intervals, and mean absolute (root-mean-square) error between the expected median (mean) and the observed occurence times. Bold, the lowest values for each MR compared to those in Table 9.

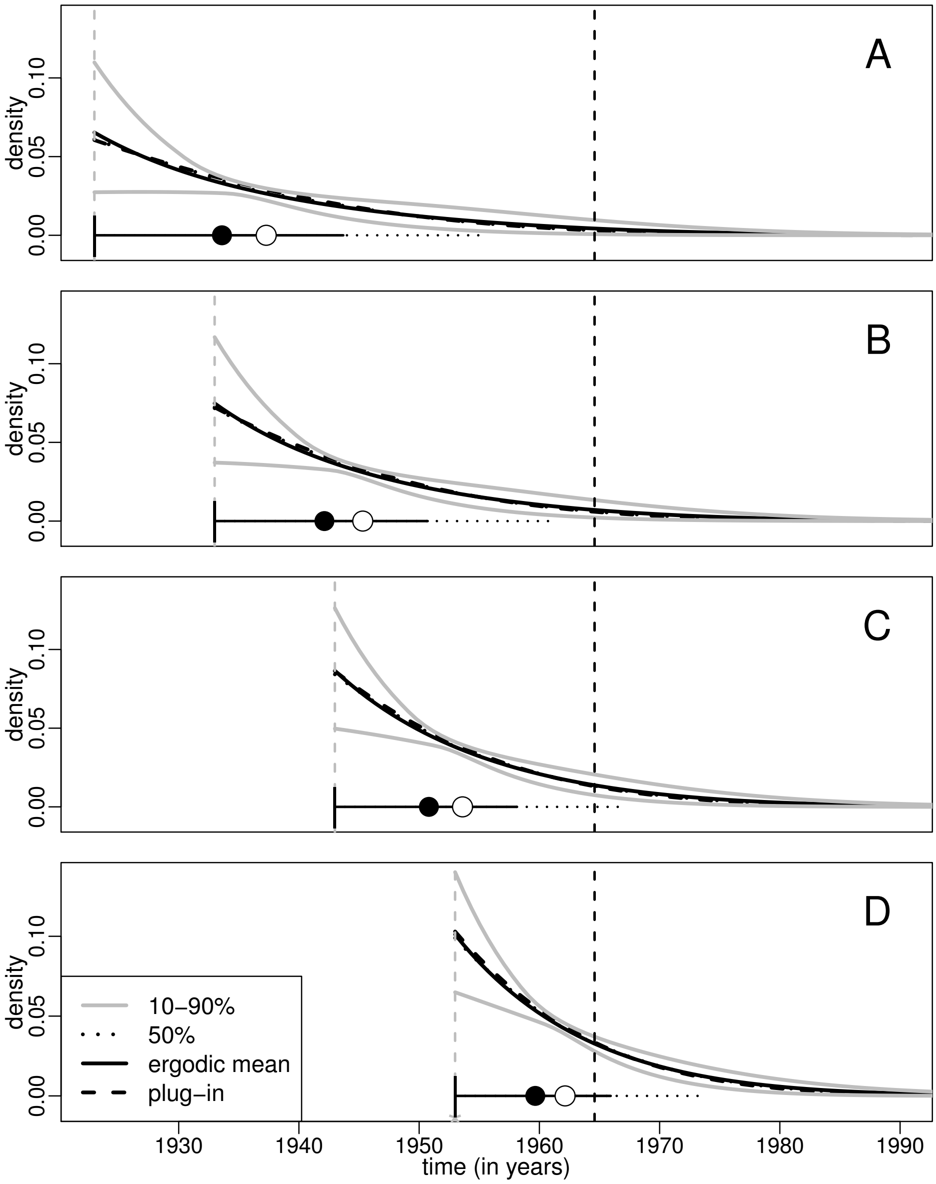

Figure 10: Density functions of the time to the next event and their statistical summaries, estimated at different issue times before the 1964/08/02 earthquake in MR_4 according to model $\mathbf{R_S}$. The forecast issue dates are for immediately after the 1922/12/29 earthquake (A), and for 10 years (B), 20 years (C), and 30 years (D) after this event. Legend: Gompertz density function (solid curve), mean (open circle), median (solid circle), 75% HPD (solid horizontal segment), and 90% HPD (dotted horizontal segment). The forecast issue date is indicated by the gray dashed vertical line, and the occurrence time of the target event by the black dashed vertical line.

6 Final remarks

We examined four different versions of the classic SR model, based on the probabilistic translation of the elastic rebound theory and including the contribution of the tectonic information. All of these model versions imply a sudden hazard reduction right after a strong earthquake (threshold set at $M_w \geq 5.3$) and an exponentially increasing hazard function between two consecutive earthquakes (excluding the aftershock sequences).

The four model versions, however, differ from one to the other in the quantity - strain, moment, energy, and scaled energy - as chosen to represent the physical process responsible for the generation of earthquakes. Equations (8)-(11) highlight the key elements (i.e., earthquake magnitude, fault rupture area, exponential coefficient) that quantify the abrupt change in the system when an earthquake occurs. The affinity among these elements is reflected in the similarity of the shapes of the relevant conditional intensities (Figures 5 and 6). Despite the general similarity, note that the conditional intensity variation (equivalent to a hazard drop) is different in the different SR models, depending on the sizes of the intervening earthquake. With reference to Figure 6, take for example the amount of the vertical drop in the conditional intensity after the 1915/01/13, $M_w = 6.99$, earthquake and the vertical drop after all of the other moderate earthquakes ($M_w < 6$). The ratio between these two values for the $\mathbf{R_S}$ model is much smaller than the same ratio in any of the $\mathbf{R_B}$, $\mathbf{R_M}$, and $\mathbf{R_E}$ models. In other words, when the scaled energy is adopted, the SR model produces a hazard decrease that is relatively heightened for smaller earthquakes and abated for larger earthquakes.

As for the model comparisons, the Bayes factor indicates (Table 7) that the $\mathbf{R_S}$ model performs, slightly in MR_2, MR_4, MR_6-MR_8, and moderately, in MR_5, better than the other models. $\mathbf{R_E}$ performs considerably better than the others in MR_1 and moderately so in MR_3. However, we note that the results for MR_1 should be taken cautiously because of the lower number of events (only seven) and its nonuniform tectonic characterization. For the predictive performance, the Ando & Tsay information criterion supports (Table 8) the conclusions reached by the Bayes factor, except for MR_3, where the criterion assigns slight evidence in favor of $\mathbf{R_S}$. Overall, although the differences among the model performances are not clearly significant, we suggest that adopting the energy or the scaled energy as a proxy measure of earthquake size is advisable. Indeed, the scaled energy allows the model to be enhanced with information on the rupture parameters, such as the area and the mechanism, which can be expected to become progressively less uncertain, in the future, as knowledge of earthquake faulting improves (e.g., fault scaling relationships).

The probability distribution of the time to next event for the SR model has been analytically identified as the Gompertz distribution (Section 4.2.3) with two parameters that depend on the model parameters and on the value of the hazard function at time t (Section 4.2.3). After summarizing its main properties, we examined the Gompertz distribution in the Bayesian framework by evaluation of its posterior predictive distribution through the Markov chains generated from the posterior distributions of the model parameters in the estimation procedure (the McMC algorithm is detailed in Appendix B). These findings bring about immediate benefit, by allowing modelers to avoid approximating this distribution through numerical simulations (e.g., Wang et al. 1991). We thus used the Gompertz distribution and its statistical summaries to run a set of retrospective and prospective forecasts of the occurrence times of the main shocks, and then we validated the procedure against the data observed.

Retrospective forecasts have also been used as a further criterion for supporting the selection of the best SR models. Different measures of the discrepancy between the expected occurrence time of an earthquake and the time of its actual occurrence (Table 9) have shown that the retrospective analysis supports the choice of the $\mathbf{R_S}$ model in most of the cases analyzed here.

Based on the knowledge available in 2002 in terms of the seismicity and tectonics, prospective forecasts issued at the very beginning of 2003 indicated that in decreasing order of immediacy, MR_3, MR_7, and MR_4 were the most prone areas to be hit by earthquakes of $M_w \geq 5.3$ in the following decade (Table 10). Of these MRs, earthquakes have actually occurred in MR_3 (three events) and MR_4 (one event) with forecasts in terms of median and mean with an average accuracy of about 6 years. However, no earthquake has occurred in MR_7 up to the end of 2012. By adding this information to the 2013 update, the forecast postpones the expected occurrence time of the next event considerably (by more than 10 years).

As we anticipated in Section 4.2.3, updating a forecast during the waiting time by adding the information that no earthquake has occurred tends to postpone the time to the next event and to reduce the uncertainty around that value. This effect is achieved through the shortening and peaking of the probability density function of the time to the next event. The prospective forecasts reported in Table 10 confirm this general behavior, although the amount of delay and uncertainty gain remain variable, and depend on repeated parameter estimates.

It is important to recall that both the time and space scales of the SR models and their associated uncertainties that we have investigated here depend on the characteristics of the available datasets. Note that there is a trade-off between the size of the region to be investigated and the length of the learning dataset. On the one hand, a reduction in the size of the region would be likely to improve its tectonic characterization, which would allow the analyst to single out homogeneous faults and avoid mixing tectonic structures that obey mechanically different stress-loading systems. It would also imply a smaller spatial domain within which the forecasted earthquakes can occur. On the other hand, a smaller area would capture fewer earthquakes for building the learning dataset, thereby worsening the robustness and overall quality of the SR model. The balancing of these factors (tectonics and seismicity) in the Italian case allowed us to investigate only a limited number of cases (the eight MRs). Additional studies are thus needed for the exploration of more fault systems with different seismic histories, to further test the energy and scaled energy as the best option in SR models, and for the refining of the time-space limits of the SR model applications in robust earthquake forecasting. Similar limitations hold for the application of the interesting extension of the SR model which was presented by Jiang et al. (2011) and which requires knowledge of source parameters that are rarely available for historical Italian earthquakes.

Depending on the data availability, possible future research directions can also be aimed at developing the *linked* (or *coupled*) versions of the SR models (e.g., Bebbington

& Harte 2003, Kuehn et al. 2008) on the same Italian data, using the (scaled) energy as the measure of the sizes of the events.

Acknowledgements

We are grateful to two reviewers for their helpful comments. The authors also thank Fracassi and Valensise for providing the earthquake association with fault sources. This study was partially funded by the Italian Dipartimento della Protezione Civile in the framework of the 2007-2009 Agreement with Istituto Nazionale di Geofisica e Vulcanologia (INGV), project S1: "Analysis of the seismic potential in Italy for the evaluation of the seismic hazard".

References

Abramowitz, M. & Stegun, I.A., 1972. *Handbook of Mathematical Functions*, Dover Publications, New York.

Ameri, G., Gallovic, F. & Pacor, F., 2012. Complexity of the Mw 6.3 2009 L'Aquila (central Italy) earthquake: 2. Broadband strong motion modeling, *J. Geophys. Res.: Solid Earth*, **117**, B04308, doi:10.1029/2011JB008729.

Ando, T. & Tsay, R., 2010. Predictive likelihood for Bayesian model selection and averaging, *International Journal of Forecasting*, **26**, 744-763.

Basili, R., Valensise, G., Vannoli, P., Burrato, P., Fracassi, U., Mariano, S., Tiberti, M.M. & Boschi, E., 2008. The Database of Individual Seismogenic Sources (DISS), version 3: summarizing 20 years of research on Italy's earthquake geology, *Tectonophysics*, **453**, 1-4, 20-43, doi: 10.1016/j.tecto.2007.04.014

Basili, R., Kastelic, V., Valensise, G. & DISS Working Group 2009, 2009. DISS3 tutorial series. Guidelines for Compiling Records of the Database of Individual Seismogenic Sources, Version 3, *Rapporti Tecnici INGV 108*, 20 pps., http://portale.ingv.it/produzione-scientifica/rapporti-tecnici-ingv/archivio/rapporti-tecnici-2009/

Bebbington, M. & Harte, D.S., 2003. The linked stress release model for spatio-temporal seismicity: formulations, procedures and applications, *Geophys. J. Int.*, **154**, 3, 925-946.

Benioff, H., 1951. Earthquakes and rock creep, Part I: Creep characteristics of rocks and the origin of aftershocks, *Bull. Seism. Soc. Am.*, 41, 31-62.

Berger, J., 2006. The case for objective Bayesian analysis, *Bayesian Anal.*, **1**, 3, 385-402.

Bhattacharyya, P. & Chakrabarti, B.K. et al., 2006. *Modelling Critical and Catastrophic Phenomena in Geoscience*, Lect. Notes Phys., 705, Springer, Berlin Heidelberg, DOI 10.1007/b11766995

Carlin, B.P. & Louis, T.A,. (2000). *Bayes and emprical Bayes methods for data analysis*, Chapman & Hall, London.

Choy, G.L. & Boatwright, J.L., 1995. Global patterns of radiated seismic energy and apparent stress, *J. Geophys. Res.: Solid Earth*, **100**, B9, 18,205-18,228.

CPTI Working Group, 2004. Catalogo Parametrico dei Terremoti Italiani, version 2004 (CPTI04), INGV, Bologna, available on http://emidius.mi.ingv.it/CPTI04/

DISS Working Group, 2007. Database of Individual Seismogenic Sources (DISS), Version 3.0.2: A compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas, http://diss.rm.ingv.it/diss/, © INGV 2007 - Istituto Nazionale di Geofisica e Vulcanologia, Rome , Italy, DOI:10.6092/INGV.IT-DISS3.0.2

Gilks, W.R., Richardson, S. & Spiegelhalter, D.J., eds., 1996. *Markov chain Monte Carlo in Practice*, Chapman & Hall, London.

Gneiting, T. & Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation, *J. Am. Stat. Assoc.*, **102**, 359-378.

Hawkes, A.G. & Oakes, D.A., 1974. A cluster process representation of a self-exciting process, *J. Appl. Probab.*, 11, 493-503.

Herrmann, R., Malagnini, L. & Munafo, I., 2011. Regional moment tensors of the 2009 L'Aquila earthquake sequence, *Bull. Seism. Soc. Am.*, 101, 975-993, doi:10.1785/0120100184.

International Mathematics and Statistics Library (IMSL) Numerical Libraries, Version 4.0, 2000. Rogue Wave Software, Inc.

Isham, V. & Westcott, M., 1979. A self-correcting point process, *Stochastic Processes and Their Applications*, **8**, 335-347.

Italian Seismological Instrumental and parametric Data-basE, 2010. *Italian Seismic Bulletin*, Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italy, available on http://iside.rm.ingv.it. Last accessed date June 3, 2015.

Jiang, M., Zhou, S., Chen, Y.J. & Ai, Y., 2011. A new multidimensional stress release statistical model based on coseismic stress transfer, *Geophys. J. Int.*, **187**, 3, 1479-1494.

Kagan, Y.Y., 1991. Likelihood analysis of earthquake catalogue, *Geophys. J. Int.*, 106, 135-148.

Kanamori, H., 1977. The energy release in great earthquakes, *J. Geophys. Res.: Solid Earth*, **82**, 2981-2987.

Kanamori, H. & Brodsky, E.E., 2004. The physics of earthquakes, *Reports on Progress in Physics*, **67**, 1429-1496.

Kanamori, H. & Heaton, T.K., 2000. Microscopic and macroscopic physics of earthquakes, *Geocomplexity and the Physics of Earthquakes*, Geophysical Monograph 20, AGU, 127-141.

Kanamori, H., Mori, J.,Hauksson, E., Heaton, T.H., Hutton, L.K. & Jones, L.M., 1993. Determination of earthquake energy release and M_L using terrascope, *Bull. Seism. Soc. Am.*, **83**, 330-346.

Kass, R.E. & Raftery, A.E., 1995. Bayes factor, *J. Am. Stat. Ass.*, **90**, 430, 773-795.

Kuehn, N.M., Hainzl, S. & Scherbaum, F., 2008. Non-Poisson earthquake occurrence in coupled stress release models and its effect on seismic hazard, *Geophys. J. Int.*, **174**, 649-658.

Lenart, A., 2014. The moments of Gompertz distribution and maximum likelihood estimation of its parameters, *Scand. Actuar. J.*, 2014:3, 255-277, doi:10.1080/03461238.2012.687697

Maercklin, N., Zollo, A., Orefice, A., Festa, G., Emolo, A., De Matteis, R., Delouis, B. & Bobbio, A., 2011. The effectiveness of a distant accelerometer array to compute seismic source parameters: The April 2009 L'Aquila earthquake case history, *Bull. Seism. Soc. Am.*, 101, 354-365, doi:10.1785/0120100124.

Matsu'ura, R.S., 1986. Precursory quiescence and recovery of aftershock activities before some large aftershocks. *Bulletin of the Earthquake Research Institute*, University of Tokyo, 61, 1-65.

Meletti, C., Galadini, F., Valensise, G., Stucchi, M., Basili, R., Barba, S., Vannucci, G. & Boschi. E., 2008. A seismic source zone model for the seismic hazard assessment of the Italian territory, *Tectonophysics*, **450**, 85-108, doi.10.1016/j.tecto.2008.01.003.

MPS Working Group 2004, 2004. Redazione della mappa di Pericolosità Sismica Prevista dall'Ordinanza PCM 3274 del 20 Marzo 2003. *Rapporto Conclusivo per il Dipartimento della Protezione Civile*, Milano-Roma, INGV, 2004 April. 65 pps., 5 appendixes; http://zonesismiche.mi.ingv.it.

Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes, *J. Amer. Statist. Assoc.*, **83**, 401, 9-27.

Ogata, Y., 1997. Detection of precursory relative quiescence before great earthquakes through a statistical model, *J. Geophys. Res.: Solid Earth*, 97, 19, 845-19,871.

Ogata, Y., 1999. Seismicity analysis through point-process modeling: a review, In *Seismicity Patterns, Their Statistical Significance and Physical Meaning* (eds. Wyss, M., Shimazaki, K. &Ito, A.) (Birkhäuser, Basel), *Pure Appl.Geophys.*, 155, 471–507.

Pondrelli, S., Salimbeni, S., Morelli, A., Ekström, G., Olivieri, M. & Boschi, E., 2010. Seismic moment tensors of the April 2009, L'Aquila (Central Italy) earthquake sequence, *Geophys. J. Int.*, **180**, 238-242, doi: 10.1111/j.1365-246X.2009.04418.x

R Development Core Team, 2006. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical computing, Vienna, Austria, ISBN 3-900051-07-0, URL www.r-project.org

Reid, H.F., 1910. *The Mechanics of the Earthquake. The California Earthquake of April 18, 1906*, Report of the State Investigation Commission, Vol. 2, Carnegie Institution of Washington, Washington, D.C.

Rotondi, R. & Garavaglia, E., 2002. Statistical analysis of the completeness of a seismic catalogue, *Natural Hazards*, 25, 3, 245-258.

Rotondi, R. & Varini, E., 2007. Bayesian inference of stress release models applied to some Italian seismogenic zones, *Geophys. J. Int.*, **169**, 1, 301-314.

Rovida, A., Camassi, R., Gasperini, P. & Stucchi. M., eds., 2011. CPTI11, the 2011 version of the Parametric Catalogue of Italian Earthquakes. Milano, Bologna, http://emidius.mi.ingv.it/CPTI

Scognamiglio, L., Tinti, E., Michelini, A., Dreger, D.S., Cirella, A., Cocco, M., Mazza, S. & Piatanesi, A., 2010. Fast determination of moment tensors and rupture history: What has been learned from the 6 April 2009 L'Aquila earthquake sequence, *Bull. Seism. Soc. Am.*, 81, 892-906, doi:10.1785/gssrl.81.6.892.

Senatorski, P., 2005. A macroscopic approach towards earthquake physics: the meaning of the apparent stress, *Physica A*, 358, 551-574.

Senatorski, P., 2006. Fluctuations, trends and scaling of the energy radiated by heterogeneous seismic sources, *Geophys. J. Int.*, **166**, 267-276.

Senatorski, P., 2007. Apparent stress scaling and statistical trends, *Phys. Earth Planet. Inter.*, **160**, 230-244.

Smith, B.J., 2000. Bayesian Output Analysis Program - (BOA) Version 1.1 User's Manual, Department of Biostatistics, School of Public Health, University of Iowa.

Smith, B.J., 2007. boa: An R package for MCMC output convergence assessment and Posterior inference, *Journal of Statistical Software*, **21**, 11, 37 pp., available on http://www.jstatsoft.org/

Stucchi, M., Albini, P., Mirto, C. & Rebez, A., 2004. Assessing the completeness of Italian historical earthquake data, *Annals of Geophysics*, **47**, 2/3, 659-673.

Varini, E. & Rotondi, R., 2015. Probability distribution of the waiting time in the stress release model: the Gompertz distribution, *Environmental and Ecological Statistics*, **22**, 3, 493-511, doi:10.1007/s10651-014-0307-2

Vere-Jones, D., 1978. Earthquake prediction - A statistician's view, *J. Physics Earth*, **26**, 129-146.

Vere-Jones, D. & Yonglu, D., 1988. A point process analysis of historical earthquakes from North China, *Earthquake Research in China*, **2**, 2, 165-181.

Vehtari, A. & Ojanen, J., 2012. A survey of Bayesian predictive methods for model assessment, selection and comparison, *Stat. Surv.*, **6**, 142-228.

Votsi, I., Tsaklidis, G.M. & Papadimitriou, E.E., 2011. Seismic hazard assessment in central Ionian Islands area (Greece) based on stress release models, *Acta Geophysica*, **59**, 4,701-727.

Wang, A., Vere-Jones, D. & Zheng, X., 1991. Simulation and estimation procedures for stress release models, in *Stochastic Processes and Their Applications, Lecture Notes in*

Econometrics and Mathematical Systems, Vol. 370, pp. 11-27, eds. Beckmann, M.J., Gopalan, M.N. & Subramanian, R., Springer, Berlin.

Watanabe, S., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory, *J. Mach. Learn. Res.*, **11**, 3571-3594.

Wells, D.L. & Coppersmith, K.L., 1994. New relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seism. Soc. Am.*, **84**, 4, 974-1002.

Zheng, X. & Vere-Jones, D., 1991. Application of stress release models to historical earthquakes from North China, *Pure Appl.Geophys.*, **135**, 4, 559-576.

Zheng, X. & Vere-Jones, D., 1994. Further applications of the stochastic stress release model to historical data, *Tectonophysics*, 229, 101-121.

A Completeness of the catalog: statistical analysis

Let us consider a catalog that covers the time interval (T_0, T_f), and suppose that there is a point s in this interval in which the seismicity rate changes, so that the global model for the number of events within the given time interval is the mixture of two Poisson processes, with the intensity function $\lambda(t)$ given by:

$$\lambda(t) = h_1 \ I_{t<s}(t) + h_2 \ I_{t \geq s}(t) \tag{33}$$

where h_1 and h_2 are the seismicity rate of the pre-complete and complete parts, respectively. According to the Bayesian approach, both the rates and the position of the changepoint s are random variables; we assume that both h_1 and h_2 follow the prior distribution $Gamma(a_0, b)$, with density function $b^{-a_0} e^{-h/b} h^{a_0 - 1}/\Gamma(a_0)$, while s is uniformly distributed on (T_0, T_f). *A priori* information on the variability of the yearly occurrence rate is inferred from the general considerations of the average number of events under examination. In the present study, we considered the shocks with $M_w \geq 5.3$ recorded in the CPTI04 for 1600-2002, a period generally considered sufficiently complete in the literature on Italian seismicity (Stucchi et al. , 2004). The uncertainty on the occurrence rate is then incorporated in the model through a further hierarchical level by considering b as an $InvGamma(c_0, f_0)$ distributed random variable. In our case, parameter a_0 and hyperparameters c_0 and f_0 are set as $a_0 = 0.1$, $c_0 = 3$, and $f_0 = 5$. For the time interval (T_0, T_f), we set $T_f = 2003$, as the end of the CPTI04, while T_0 varies in each MR. To balance the final gap between T_f and the time t_n of the last event, we approximately set T_0 back by $(T_f - t_n)$, so we have $T_0 = t_1 - T_f + t_n$, with t_1 as the time of the first event in the dataset.

We estimate the model parameters h_1, h_2, s, and b through Gibbs sampling, one of the most popular McMC methods, which is a class of methods that are based on the simulation of samples of dependent values that constitute a realization of a stationary Markov chain asymptotically convergent in distribution to the quantity to estimate (Gilks et al. , 1996). For a detailed description of the algorithm, see Rotondi & Garavaglia (2002).

Model estimations provide the posterior probability distributions of the parameters; the most probable value (mode) of s is assumed as the beginning of the complete part of the dataset, whereas the posterior mean of h_2 gives the estimate of the corresponding seismicity rate. We recall that measures of the uncertainty of the estimates, expressed through measures of location (mean, mode) and dispersion (variance, quantiles) can be drawn from the posterior distribution of the parameters.

B McMC methods

We implemented the Metropolis-Hastings algorithm to generate a Markov chain for each parameter, as summarized below. Assuming some transition kernel $q(\theta, \theta^*)$ (called the *proposal distribution*), from which it is easy to simulate, such that:

1. Initialize the chain by simulating $\theta^{(0)}$ from the prior distribution $\pi_0(\theta)$, and set the iteration counter $j = 1$.

2. Generate a *proposed* value θ^* using the kernel $q(\theta^{(j-1)}), \theta^*)$.

3. Evaluate the *acceptance probability* $\alpha(\theta^{(j-1)}, \theta^*)$ of the proposed move, where:

$$\alpha(\theta^{(j-1)}, \theta^*) = \min\left\{1, \frac{\pi(\theta^*|data)\ q(\theta^*, \theta^{(j-1)})}{\pi(\theta^{(j-1)}|data)\ q(\theta^{(j-1)}, \theta^*)}\right\} .$$

4. Put $\theta^{(j)} = \theta^*$ with probability $\alpha(\theta^{(j-1)}, \theta^*)$, otherwise retain the current value of θ: $\theta^{(j)} = \theta^{(j-1)}$.

5. Change the counter from j to $j + 1$ and return to step 2.

Given a function $g(\theta)$, under suitable regularity conditions it has been shown that the ergodic mean $\frac{\sum_{j=1}^{R} g(\theta^{(j)})}{R}$ converges almost surely to $E_{\theta|data}\left\{g(\theta)\right\}$ as $R \to \infty$; therefore, if we set $g(\theta) = \theta$ or $g(\theta) = [\theta - E(\theta)]^2$, by applying this theorem, we obtain the estimate of the mean and variance of θ respectively. It is important to note that the density of interest $\pi(\cdot \mid data)$ only enters in the acceptance probability as a ratio, and so the method can be used when this density is known up to a normalizing constant, for instance

$\pi(\theta \mid data) \propto \mathcal{L}(data \mid \theta)\ \pi_0(\theta)$. The Markov chain generated through the algorithm is reversible and has a stationary distribution $\pi(\theta \mid data)$ irrespective of the choice of the proposal distribution. The critical point of this method is how to assess the convergence of the sampler; to solve this issue, we first discard the 'burn-in' of the simulated sequence $\{\theta^{(j)}\}_{j=0}^{R}$, i.e., its initial part (ca. 10%-20%), to reduce the dependence on the initial value; then we apply one of the software tools that are available for McMC convergence diagnostics. In particular, we choose the open-source package BOA (Smith , 2005) for the R system for statistical computing (R Development Core Team , 2006), and check that all of the generated sequences do not fail the following tests: Geweke test, Heidelberger & Welch test, and Raftery & Lewis test (Smith , 2007). Table B2 reports the prior and proposal distributions used in the McMC algorithm for the parameter estimation: we note that the mean of every proposal is given by the current value of the chain, whereas the value of the variance is assigned through some pilot runs of the algorithm so that the acceptance probability varies in the range of 25% to 40% - a range that has been suggested in the statistical literature to be the best. As an example, Figure B2 shows the prior density and the kernel density estimates of the posterior density of each parameter of the various models obtained by analyzing the data from MR_4.

		t	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
MR_3	(end of the catalog)	2002/12/31	-1.80	1.89E-1	5.20E-2
	(event)	2003/09/14	-1.83	1.90E-1	5.52E-2
	(event)	2008/12/23	-1.83	1.93E-1	5.53E-2
	(event)	2012/05/20	-1.85	1.94E-1	5.68E-2
		2012/12/31	-1.84	1.95E-1	5.62E-2
MR_4	(end of the catalog)	2002/12/31	-2.13	5.36E-1	3.31E-2
	(event)	2009/04/06	-2.15	5.32E-1	3.35E-2
		2012/12/31	-2.13	5.52E-1	3.29E-2

Table B1: Parameter estimates of the $\mathbf{R_S}$ models for MR_3 and MR_4, updated by enlarging the history $\mathcal{H}_t$ on which the intensity function is conditioned.

		Prior distribution			Proposal distribution		
Model	Region	α	β	ρ	α	β	ρ
$\mathbf{R_B}$	MR_1	N(-4.00; 13.0)	Γ(0.50; 2.0E-1)	Γ(0.10; 8.1E-3)	N(∗; 1.7)	LogN(∗; 4.0E-2)	LogN(∗; 4.0E-4)
	MR_2	N(-2.50; 5.0)	Γ(0.10; 8.1E-3)	Γ(0.20; 3.2E-2)	N(∗; 8.0E-1)	LogN(∗; 7.5E-3)	LogN(∗; 1.5E-2)
	MR_3	N(-1.00; 8.0E-1)	Γ(0.05; 2.0E-3)	Γ(0.40; 1.3E-1)	N(∗; 3.0E-1)	LogN(∗; 1.8E-3)	LogN(∗; 3.0E-2)
	MR_4	N(-1.50; 1.8)	Γ(0.05; 2.0E-3)	Γ(0.40; 1.3E-1)	N(∗; 3.5E-1)	LogN(∗; 5.0E-4)	LogN(∗; 4.0E-2)
	MR_5	N(-2.00; 3.2)	Γ(0.20; 3.2E-2)	Γ(0.40; 1.3E-1)	N(∗; 9.0E-1)	LogN(∗; 1.0E-2)	LogN(∗; 3.0E-2)
	MR_6	N(-2.50; 5.0)	Γ(0.10; 8.1E-3)	Γ(0.50; 2.0E-1)	N(∗; 8.0E-1)	LogN(∗; 2.0E-3)	LogN(∗; 2.3E-2)
	MR_7	N(-2.00; 3.2)	Γ(0.02; 3.2E-4)	Γ(1.00; 8.1E-1)	N(∗; 4.0E-1)	LogN(∗; 1.6E-4)	LogN(∗; 5.0E-1)
	MR_8	N(-3.00; 7.0)	Γ(0.03; 7.0E-4)	Γ(0.20; 3.2E-2)	N(∗; 7.0E-1)	LogN(∗; 4.0E-4)	LogN(∗; 6.0E-2)
$\mathbf{R_M}$	MR_1	N(-5.00; 20.2)	Γ(0.50; 2.0E-1)	Γ(0.30; 7.0E-2)	N(∗; 2.0)	LogN(∗; 6.0E-3)	LogN(∗; 1.0E-3)
	MR_2	N(-2.50; 5.0)	Γ(0.03; 7.0E-4)	Γ(0.80; 5.0E-1)	N(∗; 8.0E-1)	LogN(∗; 3.0E-4)	LogN(∗; 4.0E-1)
	MR_3	N(-1.00; 8.0E-1)	Γ(0.02; 3.2E-4)	Γ(0.80; 5.0E-1)	N(∗; 3.0E-1)	LogN(∗; 2.0E-4)	LogN(∗; 2.5E-1)
	MR_4	N(-1.50; 1.8)	Γ(0.003;7.0E-6)	Γ(3.00; 7.0)	N(∗; 3.0E-1)	LogN(∗; 3.0E-6)	LogN(∗; 5.0)
	MR_5	N(-2.00; 3.2)	Γ(0.01; 8.1E-5)	Γ(2.00; 3.2)	N(∗; 9.0E-1)	LogN(∗; 5.0E-5)	LogN(∗; 3.5)
	MR_6	N(-2.50; 5.0)	Γ(0.01; 8.1E-5)	Γ(6.00; 30.0)	N(∗; 8.0E-1)	LogN(∗; 1.5E-5)	LogN(∗; 2.4)
	MR_7	N(-2.00; 3.2)	Γ(0.001;1.0E-6)	Γ(12.0; 1.1E+2)	N(∗; 4.0E-1)	LogN(∗; 4.0E-7)	LogN(∗; 1.0E+2)
	MR_8	N(-3.00; 7.0)	Γ(0.001;1.0E-6)	Γ(8.00; 5.0E+1)	N(∗; 7.0E-1)	LogN(∗; 3.0E-7)	LogN(∗; 8.0E+1)
$\mathbf{R_E}$	MR_1	N(-5.00; 20.2)	Γ(1.50; 1.8)	Γ(0.05; 2.0E-3)	N(∗; 2.0)	LogN(∗; 8.0E-2)	LogN(∗; 1.0E-4)
	MR_2	N(-2.50; 5.0)	Γ(0.04; 1.3E-3)	Γ(0.30; 7.0E-2)	N(∗; 8.0E-1)	LogN(∗; 8.0E-4)	LogN(∗; 8.0E-2)
	MR_3	N(-1.00; 8.0E-1)	Γ(0.04; 1.3E-3)	Γ(0.30; 7.0E-2)	N(∗; 3.0E-1)	LogN(∗; 1.0E-3)	LogN(∗; 4.0E-2)
	MR_4	N(-1.50; 1.8)	Γ(0.004;1.3E-5)	Γ(2.00; 3.2)	N(∗; 3.0E-1)	LogN(∗; 8.0E-6)	LogN(∗; 3.0)
	MR_5	N(-2.00; 3.2)	Γ(0.02; 3.0E-4)	Γ(1.00; 8.1E-1)	N(∗; 9.0E-1)	LogN(∗; 1.5E-4)	LogN(∗; 9.0E-1)
	MR_6	N(-2.50; 5.0)	Γ(0.02; 3.2E-4)	Γ(3.00; 7.0)	N(∗; 8.0E-1)	LogN(∗; 5.0E-5)	LogN(∗; 8.0E-1)
	MR_7	N(-2.00; 3.2)	Γ(0.001;1.0E-6)	Γ(8.00; 5.0E+1)	N(∗; 4.0E-1)	LogN(∗; 1.0E-6)	LogN(∗; 4.8E+1)
	MR_8	N(-3.00; 7.0)	Γ(0.001;1.0E-6)	Γ(8.00; 5.0E+1)	N(∗; 7.0E-1)	LogN(∗; 3.0E-7)	LogN(∗; 6.0E+1)
$\mathbf{R_S}$	MR_1	N(-3.50; 1.0E+1)	Γ(3.00; 7.0)	Γ(0.01; 8.1E-5)	N(∗; 1.7)	LogN(∗; 1.7)	LogN(∗; 2.0E-5)
	MR_2	N(-2.50; 5.0)	Γ(2.00; 3.2)	Γ(0.04; 1.3E-3)	N(∗; 8.0E-1)	LogN(∗; 1.5)	LogN(∗; 8.0E-5)
	MR_3	N(-1.00; 8.1E-1)	Γ(0.30; 7.0E-2)	Γ(0.08; 5.0E-3)	N(∗; 3.0E-1)	LogN(∗; 6.0E-2)	LogN(∗; 1.0E-3)
	MR_4	N(-1.50; 1.8)	Γ(1.00; 8.1E-1)	Γ(0.04; 1.3E-3)	N(∗; 3.0E-1)	LogN(∗; 3.8E-1)	LogN(∗; 8.0E-5)
	MR_5	N(-2.00; 3.2)	Γ(3.00; 7.0)	Γ(0.04; 1.3E-3)	N(∗; 9.0E-1)	LogN(∗; 4.0)	LogN(∗; 6.0E-5)
	MR_6	N(-2.50; 5.0)	Γ(2.00; 3.2)	Γ(0.03; 7.0E-4)	N(∗; 8.0E-1)	LogN(∗; 1.2)	LogN(∗; 3.0E-5)
	MR_7	N(-2.00; 3.2)	Γ(0.40; 1.3E-1)	Γ(0.08; 5.0E-3)	N(∗; 4.0E-1)	LogN(∗; 7.0E-2)	LogN(∗; 1.0E-3)
	MR_8	N(-3.00; 7.0)	Γ(1.50; 1.8)	Γ(0.01; 8.1E-5)	N(∗; 7.0E-1)	LogN(∗; 5.0E-1)	LogN(∗; 3.0E-5)

Table B2: Prior and proposal distributions of the model parameters $\theta = (\alpha, \beta, \rho)$ adopted in the McMC estimation method. The mean and variance of every prior/ proposal distribution are reported, so that, e.g., for the Gamma distribution, the shape and scale parameters can be derived. The mean of each proposal distribution is set equal to the current value of the corresponding parameter in the Markov chain.

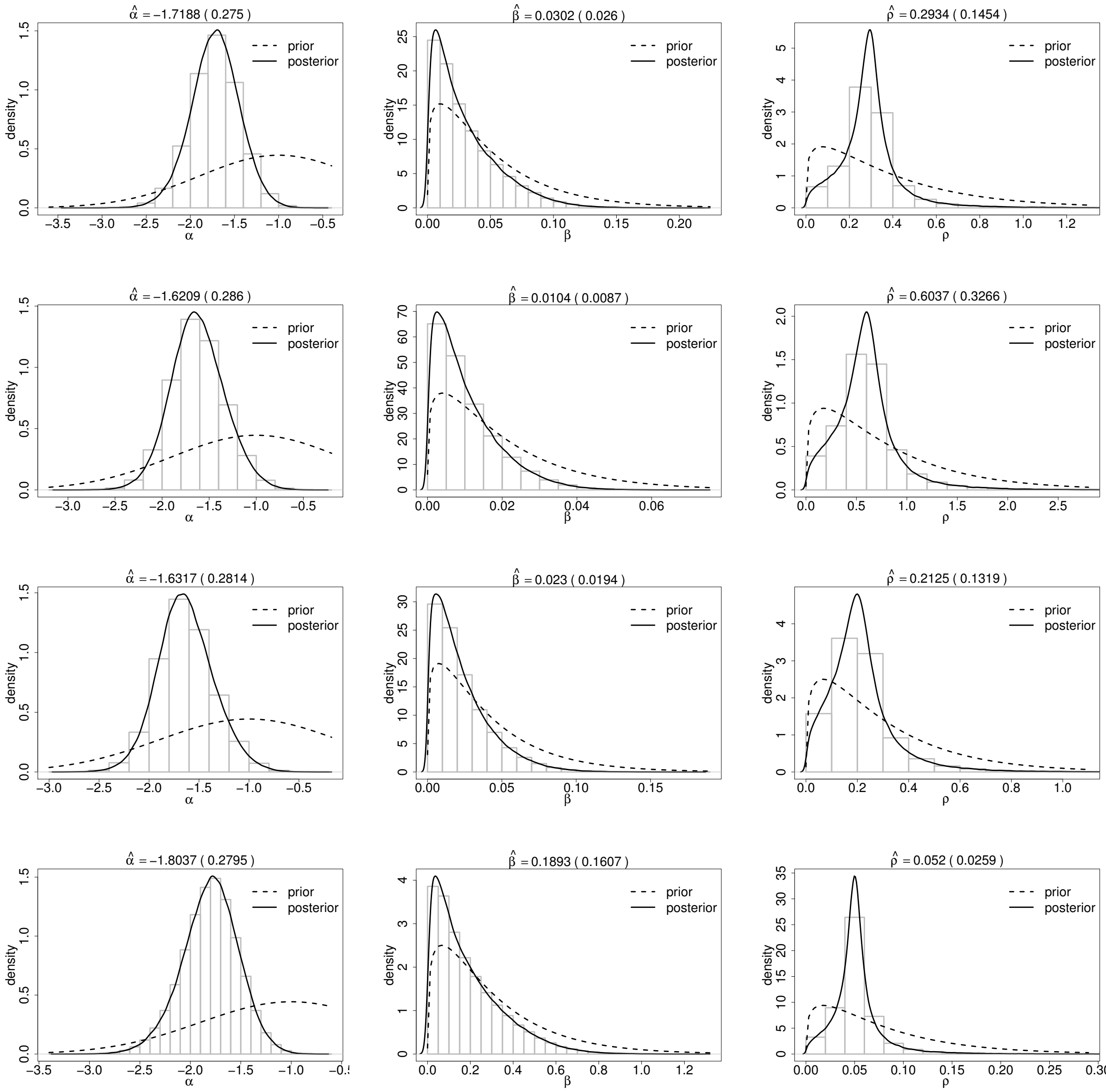

Figure B1: From top to bottom, the $\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, and $\mathbf{R_S}$ models. Prior density functions (dotted line); histograms and kernel posterior density estimates (solid line) computed from the values of the Markov chain of each parameter α, β, and ρ. Example taken from MR_3.

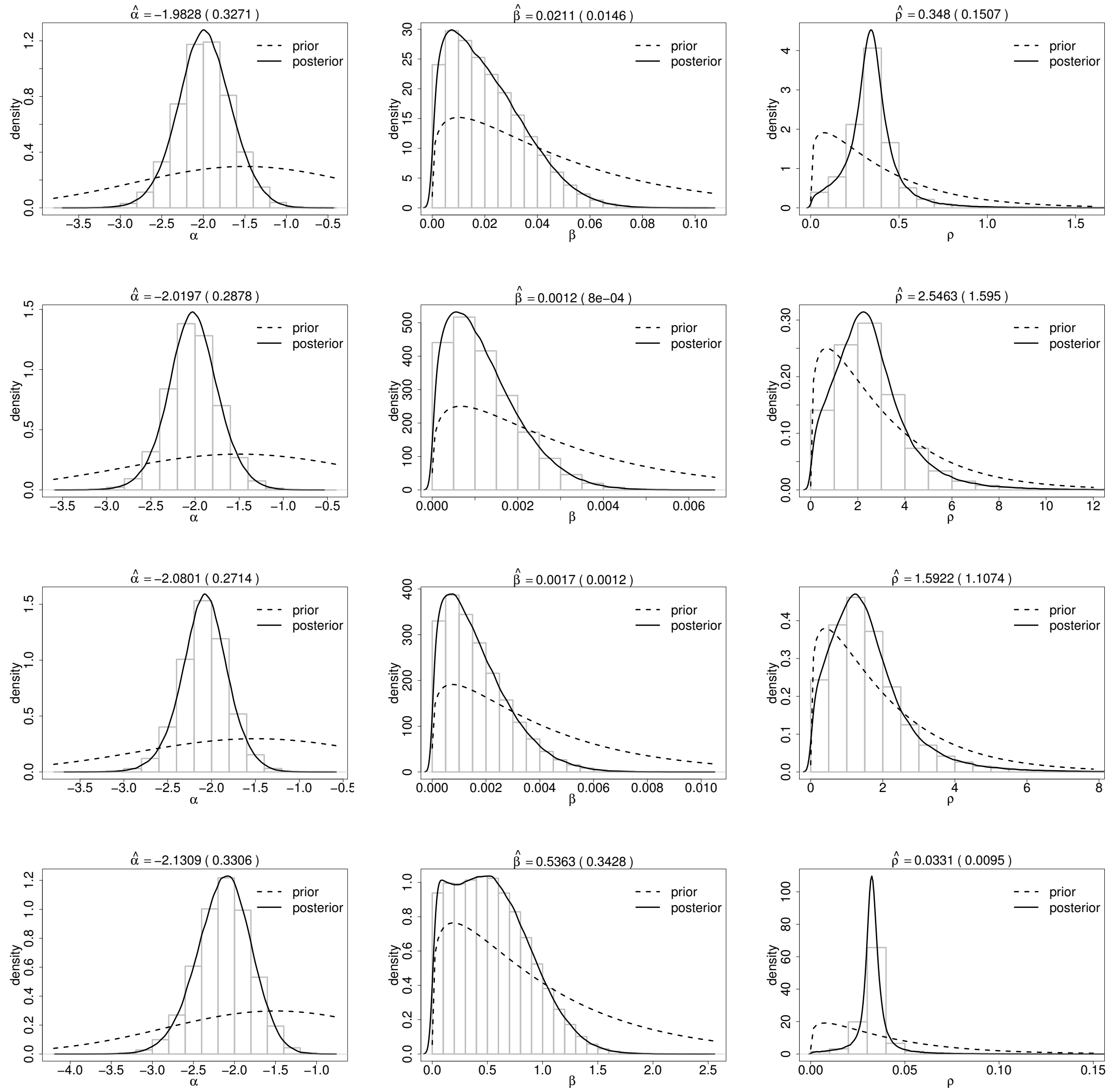

Figure B2: Same as Figure B1. Example taken from MR_4.

C Retrospective validation

Figures C1-C7 summarize the retrospective analyses of the forecasts issued at the occurrence time of every event in the datasets.

Region	Model	HPD 75%	HPD 90%	Median	Mean (st.dev.)
MR_1	R_B	2003.0-2061.6	2003.0-2106.3	2031.4	2048.3 (41.4)
	R_M	2003.0-2139.8	2003.0-2201.6	2085.0	2101.2 (39.7)
	R_E	2003.0-2304.0	2003.0-2359.0	2191.6	2204.2 (48.9)
	R_S	2003.0-2066.5	2003.0-2119.2	2032.7	2053.1 (51.2)
MR_2	R_B	2003.0-2026.8	2003.0-2043.2	2014.8	2020.5 (16.7)
	R_M	2003.0-2028.4	2003.0-2047.1	2015.3	2022.0 (18.9)
	R_E	2003.0-2027.7	2003.0-2046.0	2014.9	2021.5 (18.9)
	R_S	2003.0-2024.0	2003.0-2036.9	2013.8	2018.1 (13.4)
MR_3	R_B	2003.0-2010.8	2003.0-2016.2	2006.8	2008.7 (5.7)
	R_M	2003.0-2010.5	2003.0-2015.8	2006.7	2008.5 (5.6)
	R_E	2003.0-2010.5	2003.0-2015.8	2006.7	2008.5 (5.6)
	R_S	2003.0-2011.1	2003.0-2016.7	2007.0	2008.9 (5.9)
MR_4	R_B	2003.0-2015.5	2003.0-2024.1	2009.2	2012.1 (9.0)
	R_M	2003.0-2015.0	2003.0-2023.5	2008.9	2011.8 (9.0)
	R_E	2003.0-2014.7	2003.0-2022.9	2008.8	2011.6 (8.8)
	R_S	2003.0-2015.6	2003.0-2023.5	2009.4	2012.0 (8.3)
MR_5	R_B	2003.0-2021.0	2003.0-2034.1	2011.8	2016.5 (13.5)
	R_M	2003.0-2020.9	2003.0-2034.1	2011.6	2016.3 (13.8)
	R_E	2003.0-2020.8	2003.0-2033.9	2011.6	2016.3 (13.7)
	R_S	2003.0-2019.3	2003.0-2029.6	2011.6	2015.2 (11.1)
MR_6	R_B	2003.0-2033.8	2003.0-2054.5	2018.6	2025.8 (21.6)
	R_M	2003.0-2037.1	2003.0-2059.6	2020.2	2028.0 (23.5)
	R_E	2003.0-2039.3	2003.0-2062.8	2021.4	2029.5 (24.4)
	R_S	2003.0-2031.7	2003.0-2049.6	2018.0	2024.1 (19.2)
MR_7	R_B	2003.0-2014.5	2003.0-2022.7	2008.6	2011.5 (8.7)
	R_M	2003.0-2015.1	2003.0-2023.7	2009.0	2011.9 (9.1)
	R_E	2003.0-2015.4	2003.0-2024.1	2009.1	2012.1 (9.3)
	R_S	2003.0-2014.0	2003.0-2021.9	2008.4	2011.1 (8.3)
MR_8	R_B	2003.0-2025.7	2003.0-2042.5	2014.0	2019.9 (17.5)
	R_M	2003.0-2023.1	2003.0-2038.2	2012.7	2018.0 (15.7)
	R_E	2003.0-2022.9	2003.0-2037.7	2012.6	2017.8 (15.5)
	R_S	2003.0-2035.3	2003.0-2054.0	2019.7	2025.7 (19.5)

Table C1: Prospective forecast after the end date of the learning catalog. Summary of the estimated probability distributions of the times to next event in each MR provided by all of the models: 75% and 90% HPD intervals, median, mean, and standard deviation.

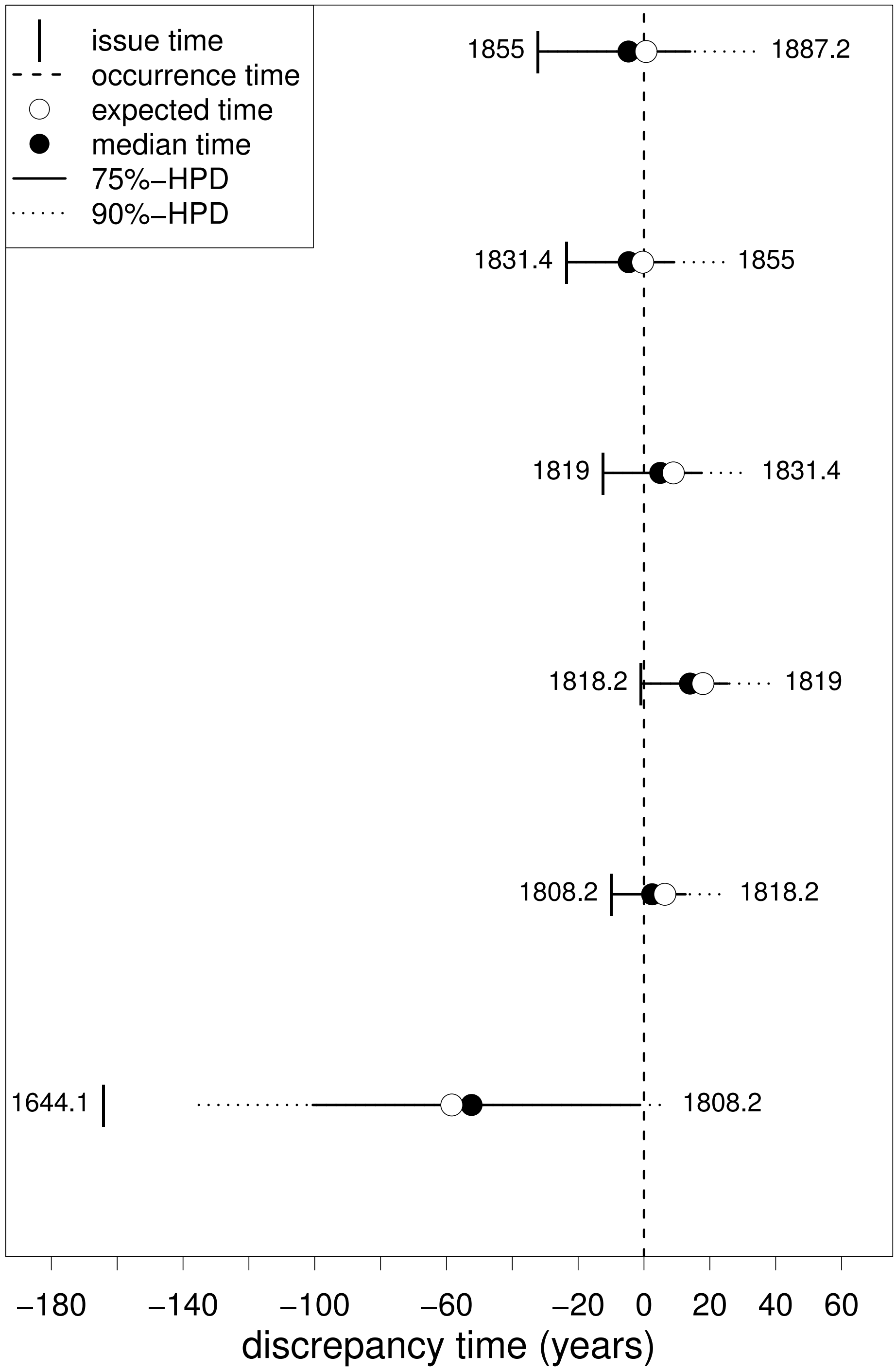

Figure C1: As for Figure 8, validation results related to macroregion MR_1 - $\mathbf{R_E}$ model.

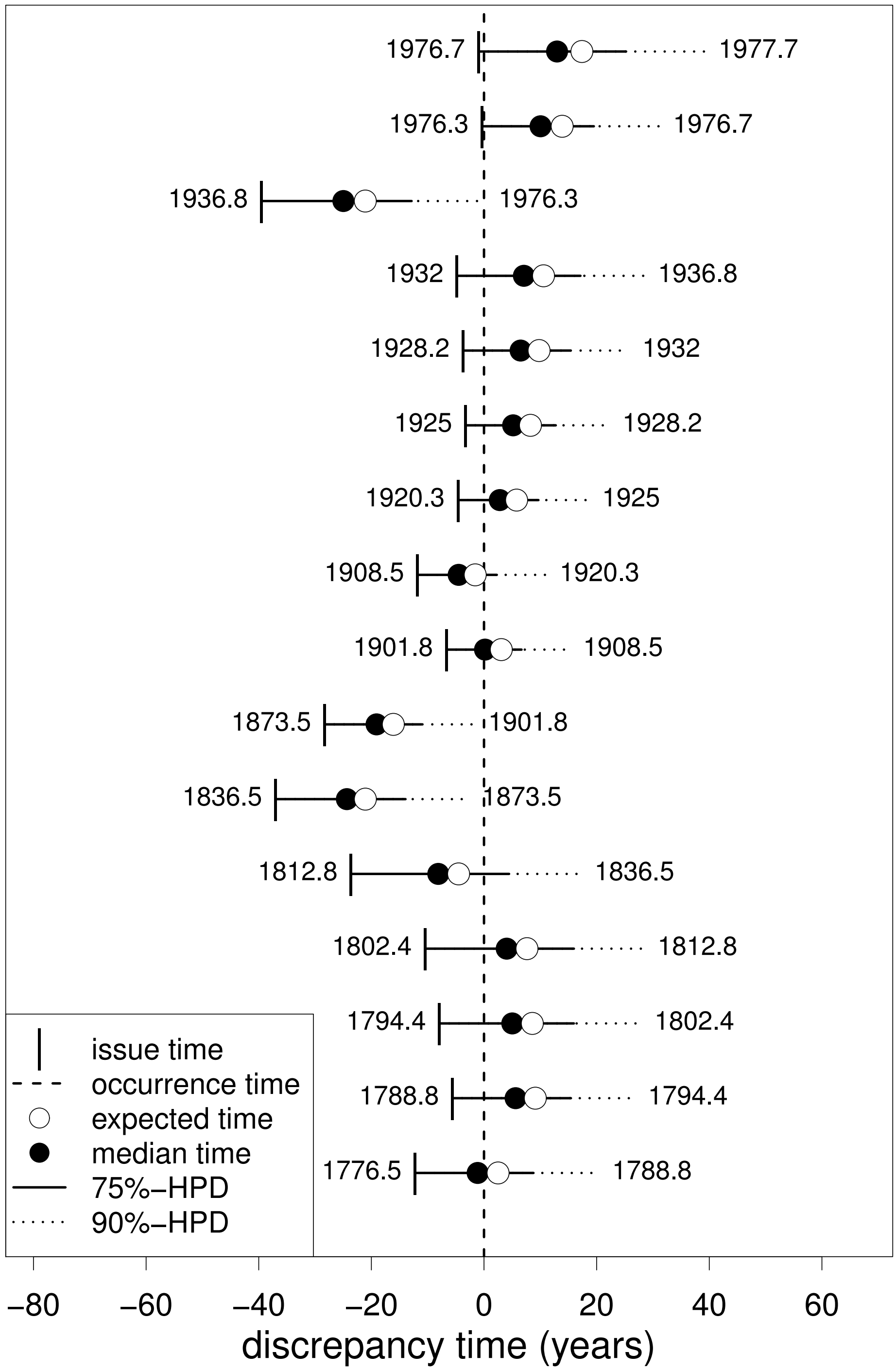

Figure C2: As for Figure 8, validation results related to macroregion MR_2 - $\mathbf{R_S}$ model.

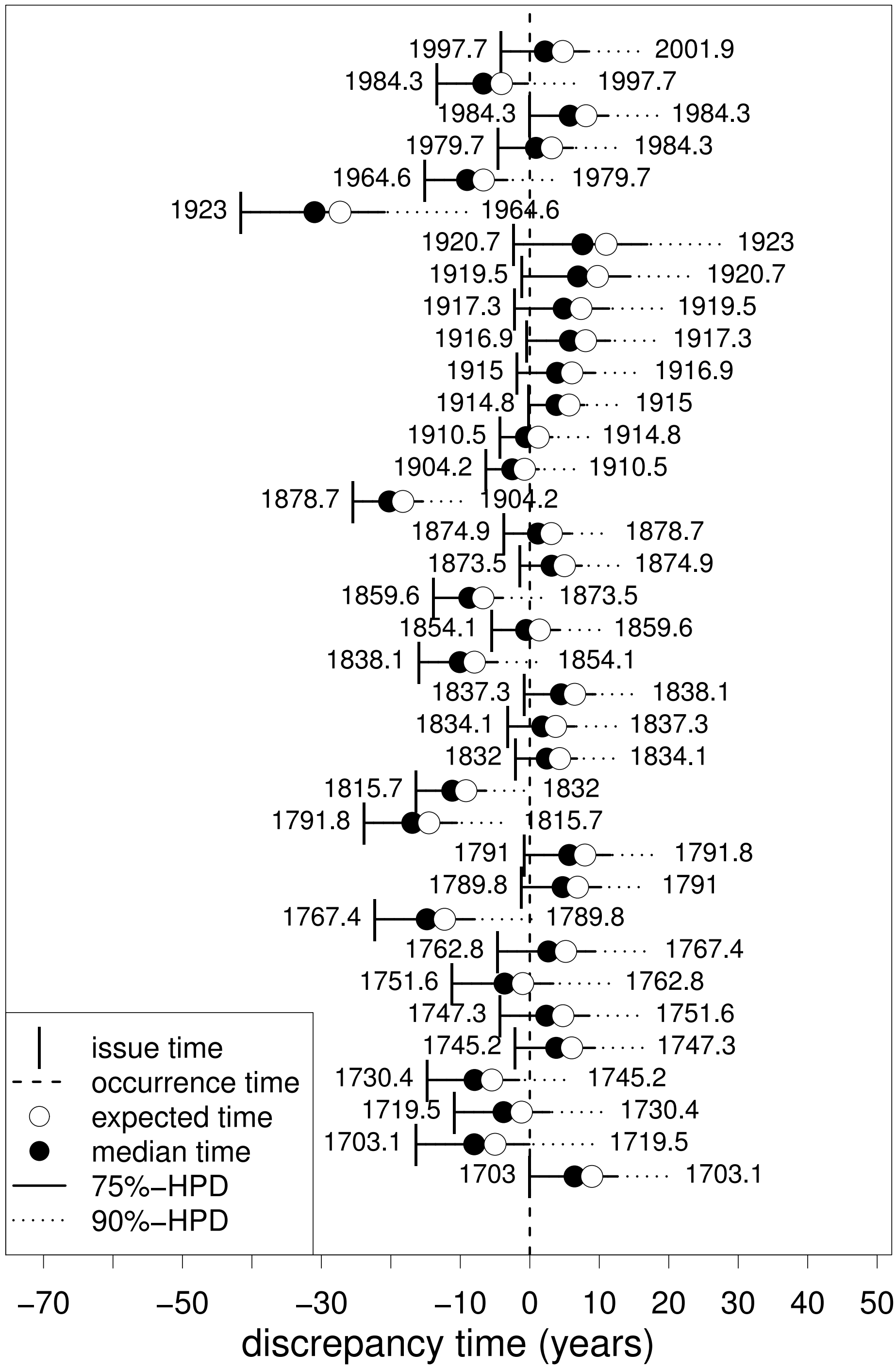

Figure C3: As for Figure 8, validation results related to macroregion MR_4 - $\mathbf{R_S}$ model.

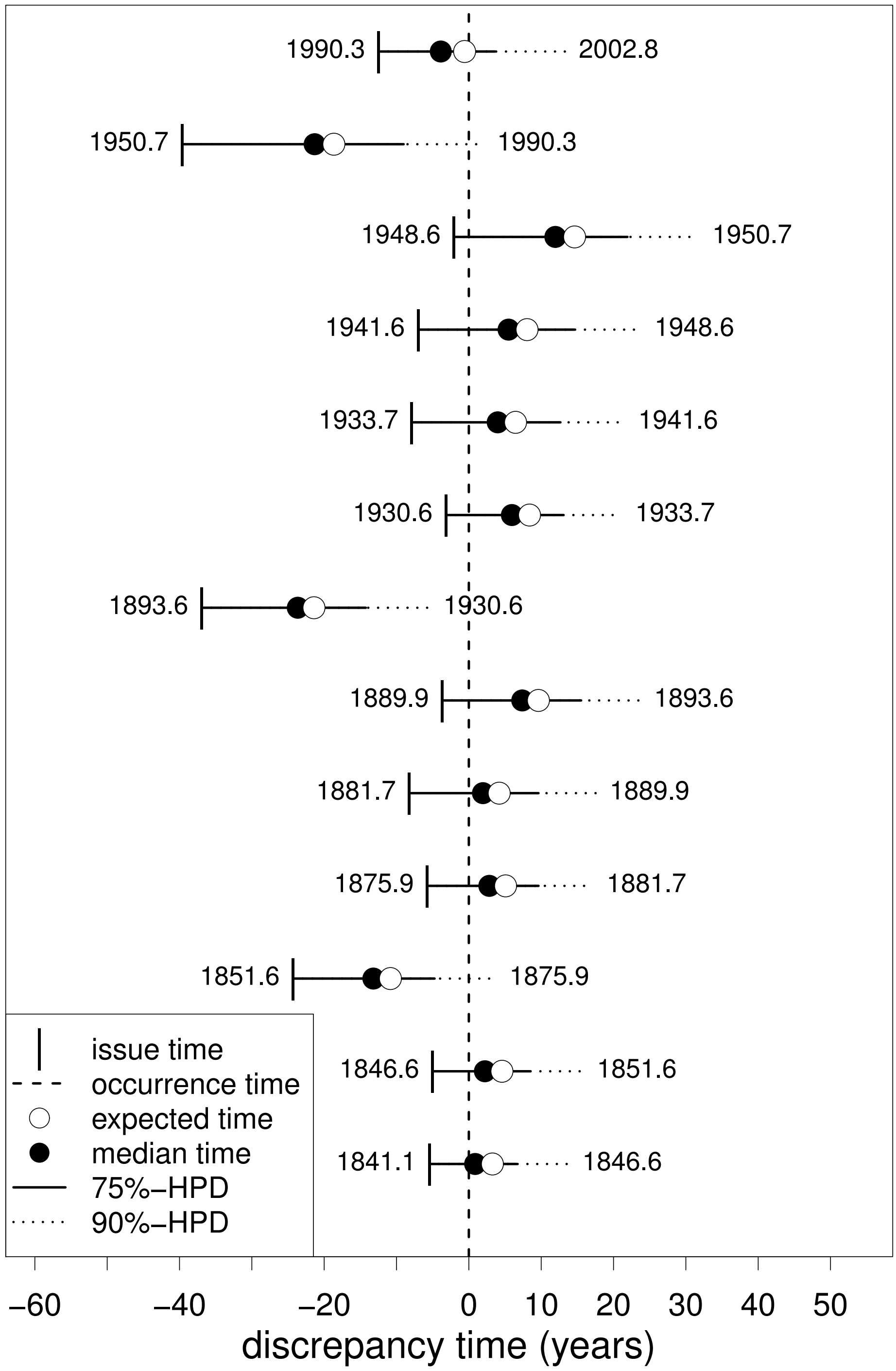

Figure C4: As for Figure 8, validation results related to macroregion MR_5 - $\mathbf{R_S}$ model.

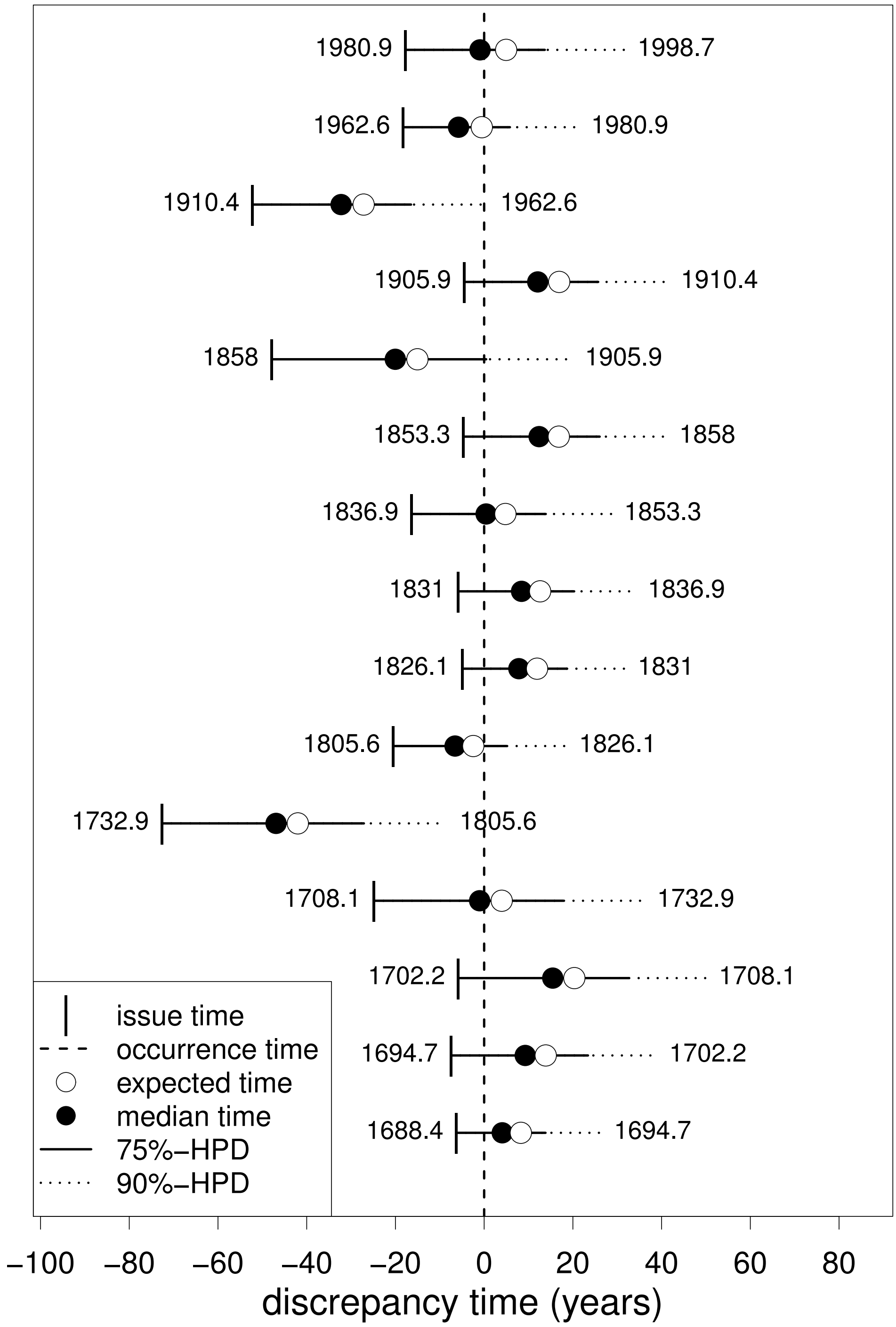

Figure C5: As for Figure 8, validation results related to macroregion MR_6 - $\mathbf{R_S}$ model.

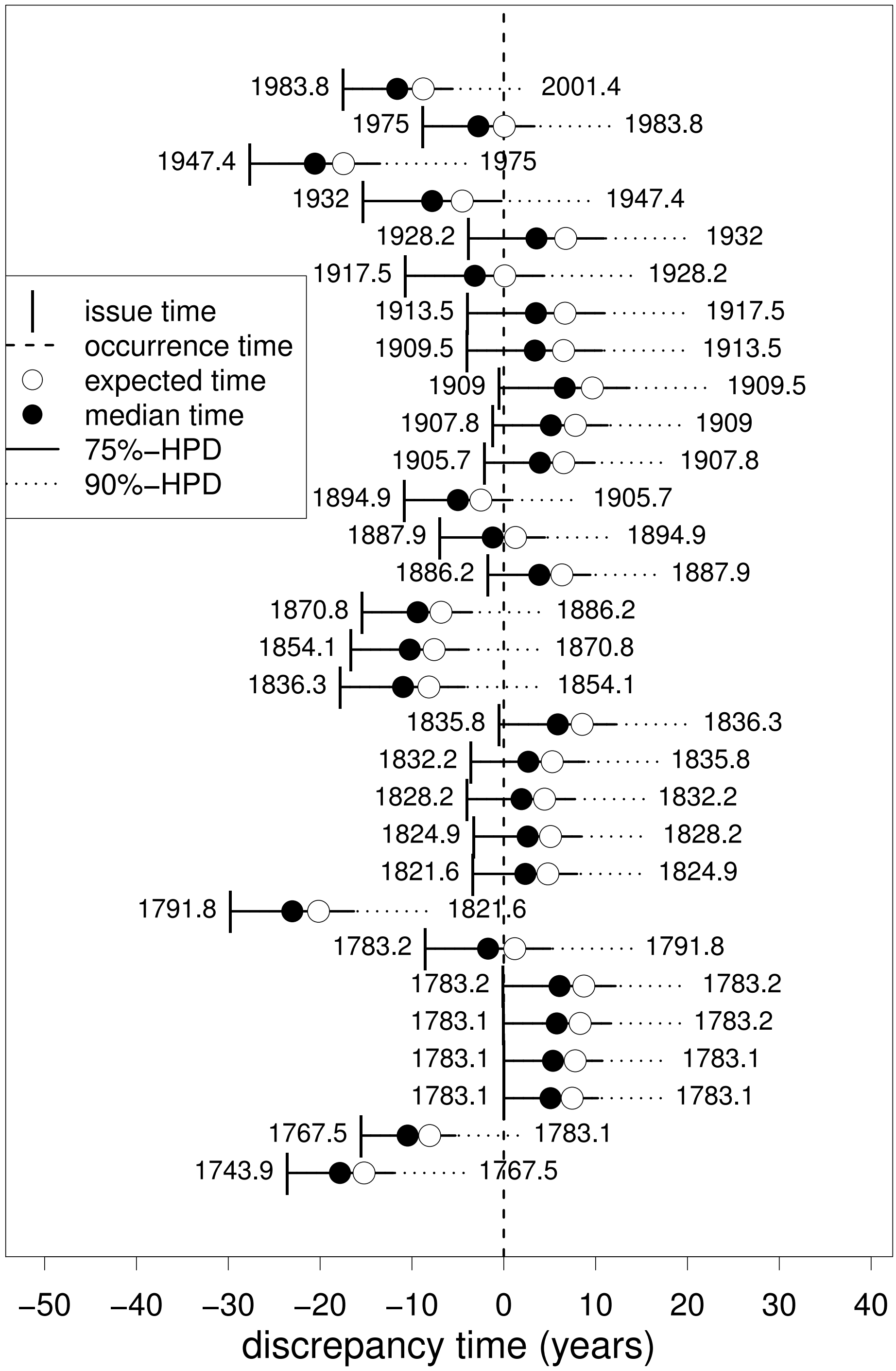

Figure C6: As for Figure 8, validation results related to macroregion MR_7 - $\mathbf{R_S}$ model.

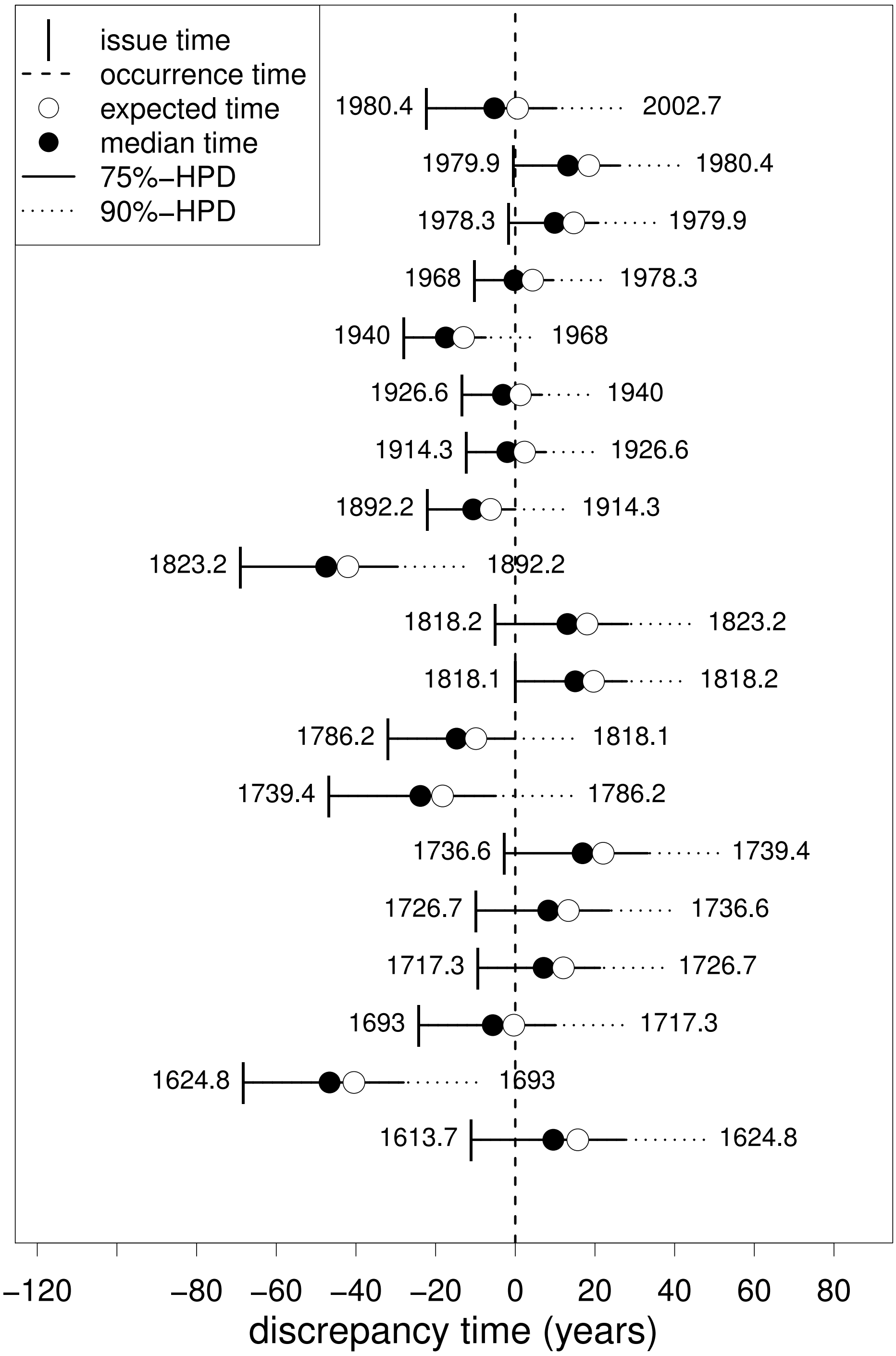

Figure C7: As for Figure 8, validation results related to macroregion MR_8 - $\mathbf{R_S}$ model.