Strain-, moment- and energy-based self-correcting models for probabilistic seismic hazard analysis in Italy

Varini E[1], Rotondi R[1], Basili R[2], Barba S[2], and Betrò B[1]

[1]Istituto di Matematica Applicata e Tecnologie Informatiche, Consiglio Nazionale delle Ricerche (Imati-CNR), Milano, Italy
[2]Istituto Nazionale di Geofisica e Vulcanologia (INGV), Roma, Italy

Abstract

This study presents a series of self-correcting models that were 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 carried out the analysis on a regional basis by subdividing the study area into eight tectonically coherent regions. In each region, we reconstructed the seismic history and statistically evaluated the completeness of the resulting seismic catalog. The conditional intensity function that characterizes each self-correcting model was estimated following the Bayesian paradigm and by applying Markov chain Monte Carlo methods. In this way, we obtain not only the parameter estimates given by their posterior mean, but also a measure of their uncertainty expressed by the simulated posterior distribution. The comparison of the four models through the Bayes factor indicates overall significant, albeit to different degrees, evidence in favor of the stress release model based on the scaled energy. Therefore, among the quantities considered, this turns out to be the most appropriate measure of the size of an earthquake for use in stress release models. At any instant, the 'time to the next event' is the realization of a Gompertz distributed random variable 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 the forecasting problem 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 date of the earthquakes that occurred after the end of the learning catalog. Four earthquakes of $M_w \geq 5.3$ occurred in the 2003-2012 decade, excluding their aftershocks, and all of these fall within the 75% highest probability density intervals of the forecast.

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 and Oakes*, 1974) that describe the temporal and spatial clustering of earthquakes [*Kagan* (1991), *Ogata* (1988), *Ogata* (1999), and references therein]. These models were originally proposed to explain the decay of secondary shocks following 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 and Chakrabarti et al.*, 2006); these operate typically on small space-time scales. The most popular models that attempt to incorporate physical knowledge 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 by Reid a century ago (*Reid*, 1910). As a first approximation, modern measurements using global positioning system (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 (*Vere-Jones*, 1978). Enriched versions of this model have been extensively adopted for over 20 years now [*Vere-Jones and Yonglu* (1988), *Zheng and Vere-Jones* (1991), *Zheng and Vere-Jones* (1994), *Bebbington and 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.

In this study, we analyze a series of self-correcting models by integrating the available information on the key components that are internationally considered among the most common input to probabilistic seismic hazard assessment. These are usually 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 Section 2, we propose four versions of the stress release 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 four models 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 independent datasets, one for each MR, including earthquakes ($M_w \geq 5.3$), and their most likely fault

sources. Statistical treatment of the possible incompleteness of the seismicity is also taken into account (Appendix A).

In Section 4, 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 mean, but also a measure of their uncertainty as expressed through the simulated posterior distribution of each parameter. The four models are compared one to the other through the Bayes factor, to determine which among the proposed measures of the size of an earthquake provides the best fit to the data. We find that the scaled energy is the measure of earthquake size that ensures the best performance of the stress release model.

We also deal with the forecasting problem by an analysis of the probability distribution $F(\omega_t|\mathcal{H}_t)$ of the 'time to the next event' conditioned on the previous history $\mathcal{H}_t$ of the process. In the case of the stress release model, this distribution turns out to be a Gompertz distribution, with two parameters that depend on the model parameters and the value of the hazard function at time t (Section 5). After summarizing its main properties, we examine the Gompertz distribution in the Bayesian framework by an 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). Retrospective validation is performed by evaluation of the expected time to the next event immediately after each earthquake in the datasets (Section 6). The mean square (and absolute) error between the forecast and the observed occurrence times further supports scaled energy as the best choice for fitting the data in most of the MRs (for a complete summary of results, see Appendix C). 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. These test events have been 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. 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 $\beta > 0$. 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. This model belongs to the class of self-correcting point processes of *Isham and 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 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 and Yonglu* (1988), *Zheng and Vere-Jones* (1991), *Zheng and Vere-Jones* (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 and 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 and 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, 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:

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

Another influential seismic parameter that gives information on the rupture behavior (*Kanamori and 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 stress release 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:

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

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

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

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

Hereinafter, we denote these models by $\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, 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 its records until 2006, and, from 2007 onwards, we used the Italian Seismic Instrumental and parametric Data-base (ISIDe; 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)).

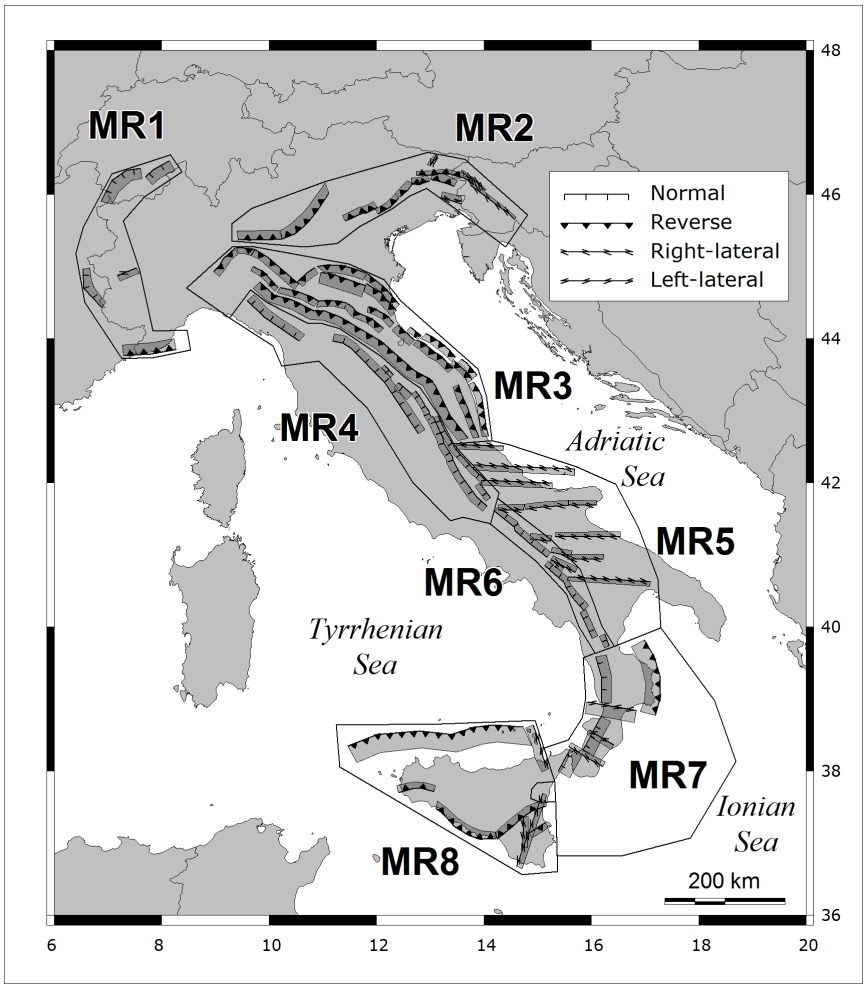

Figure 1: Map of the Composite Seismogenic Sources (CSS) from the DISS database, version 3.0.2 (*DISS Working Group*, 2007), classified according to the faulting mechanism. Shaded area: vertical projection of the fault plane to the ground surface. The outlined polygons are the MRs described in the text and Table 3.

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 a 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 1).

3.2 Earthquakes

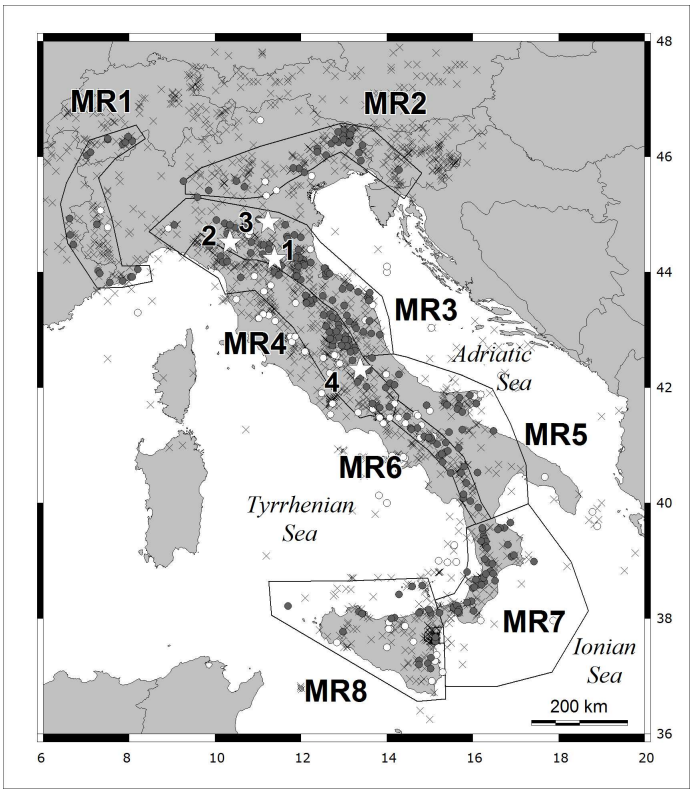

Figure 2: Map of earthquakes from the CPTI04 catalog (*CPTI Working Group*, 2004). Associations among earthquakes, MRs, and fault sources are listed in Tables 1-2. Stars indicate earthquakes that occurred after the end of the learning catalog, and were used to validate the forecast (see Section 6).

Table 1: List of 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_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	5.3				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		N	26	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		N	28	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	5.8
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 2: List of 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.

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	5.6		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

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, $I_0 = 5/6$ 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 2). The catalog is supplied in declustered form, such that events that occurred within 90 days and 30 km from the principal events (mainshocks) in seismic sequences have been removed. 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 half-monthly since April 16, 2005.

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 1 and 2), 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. 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).

To allow for potential underestimation of the earthquake magnitude, 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 and 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 in the eight MRs: $\hat{h}_2$ and $\check{s}$ are estimates of the occurrence rates in the complete part and of the changepoint.

Table 3: Faulting mechanisms in the macroregions.

ID	Name	Mechanism
MR1	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 affecting 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 affecting 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 4: Completeness of the learning datasets by MR.

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

$\check{s}$ = posterior mode of the position of the change-point, $\hat{h}_2$ = posterior mean rate, T_c = left end of the time interval under examination (see Equation (12)), * dataset considered as a complete set.

In spite of the criterion adopted in the choice of the initial time T_0 (see Appendix A), the model allows only poorly for the censored interval between the latest event t_n and the end of the catalog T_f; indeed, $\check{s}$ tends to be very close to t_1. We thus 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 inter-event time, which is calculated by 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. Notice 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 modeled by self-correcting models, which are sorted according to MR and fault source.

4 Model parameter estimation and model comparison

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 stress release 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 and Coppersmith*, 1994), where the parameters a and b depend on the faulting type of the associated fault source. Tables 1 and 2 provide the faulting types of each fault source.

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 model them through prior distributions, which are initially assigned based on the literature [*Votsi et al.* (2011), *Jiang et al.* (2011)] and the results of previous projects [*Rotondi and Varini* (2007)]. Then the prior distributions are better calibrated through pilot

runs of the estimation algorithm. In the Bayesian framework, the prior distribution of the parameter θ is denoted by π_0 and the log-likelihood function is given by:

$$\log \mathcal{L}(data \mid \theta) = \sum_{i=1}^{N} \log \lambda(t_i) - \int_{T_c}^{T_f} \lambda(s)\ ds. \tag{14}$$

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

$$\pi(\theta \mid data) = \frac{\mathcal{L}(data \mid \theta)\ \pi_0(\theta)}{\int_\Theta \mathcal{L}(data \mid \theta)\ \pi_0(\theta)\ d\theta} \tag{15}$$

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.

Table 5 gives the collected parameter estimates of the different models obtained through the McMC algorithm by generating a chain of $R = 250,000$ elements, after discarding the burn-in, and recording the output every $20th$ iteration, for each parameter.
The prior and posterior densities of the parameters of the four models for MR_3 and MR_4 are shown in Figures B.1 and B.2, respectively.

As an example, Figures 3 and 4 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 the quartiles of the pointwise estimate $\hat{\lambda}(t)$.

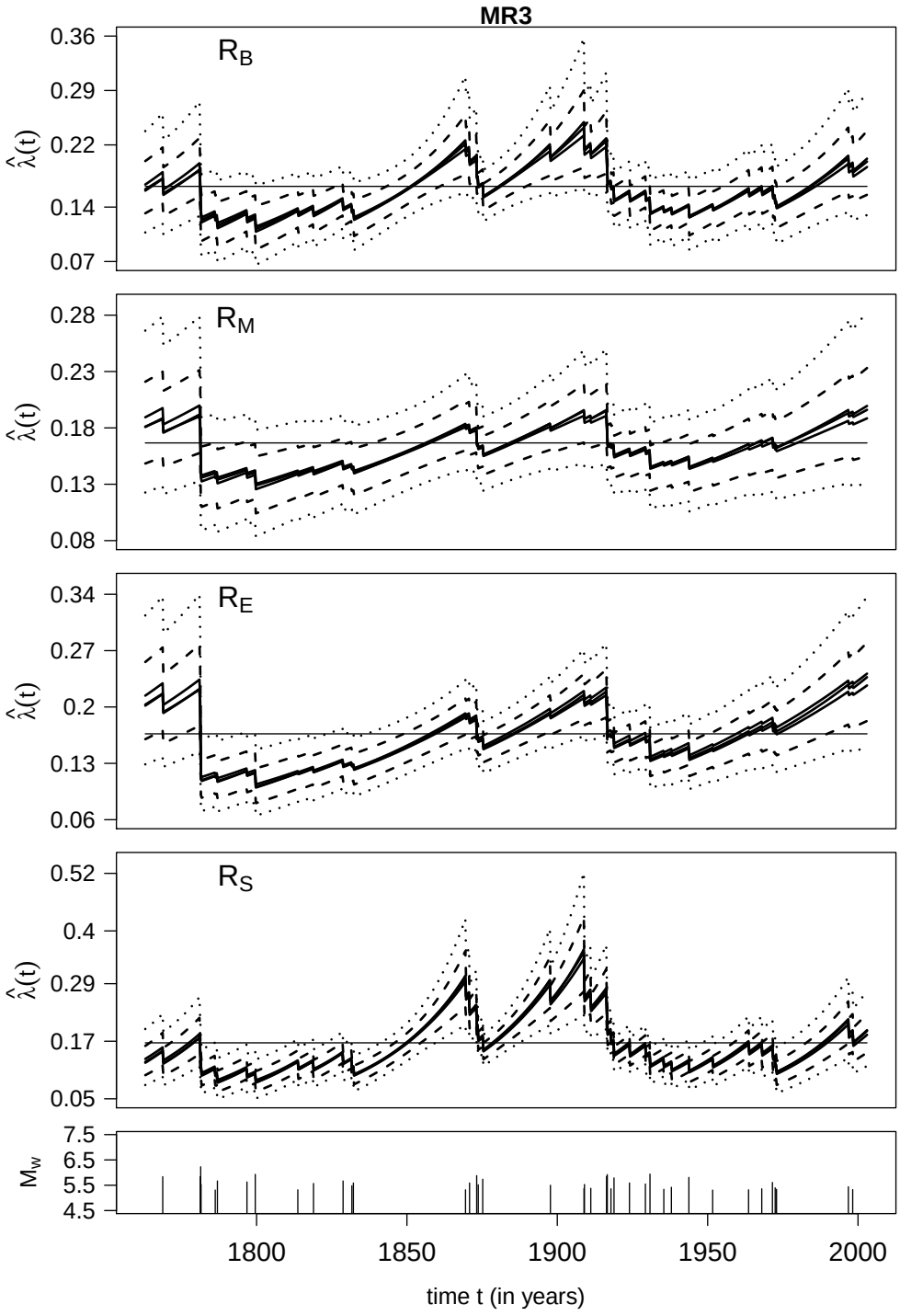

Figure 3: Conditional intensity function of the $\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, $\mathbf{R_S}$ models: ergodic mean, plug-in estimate, and median, are all represented by solid lines that are practically indistinguishable from each other; 1^{st} and 3^{rd} quartiles (dashed line), 10%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

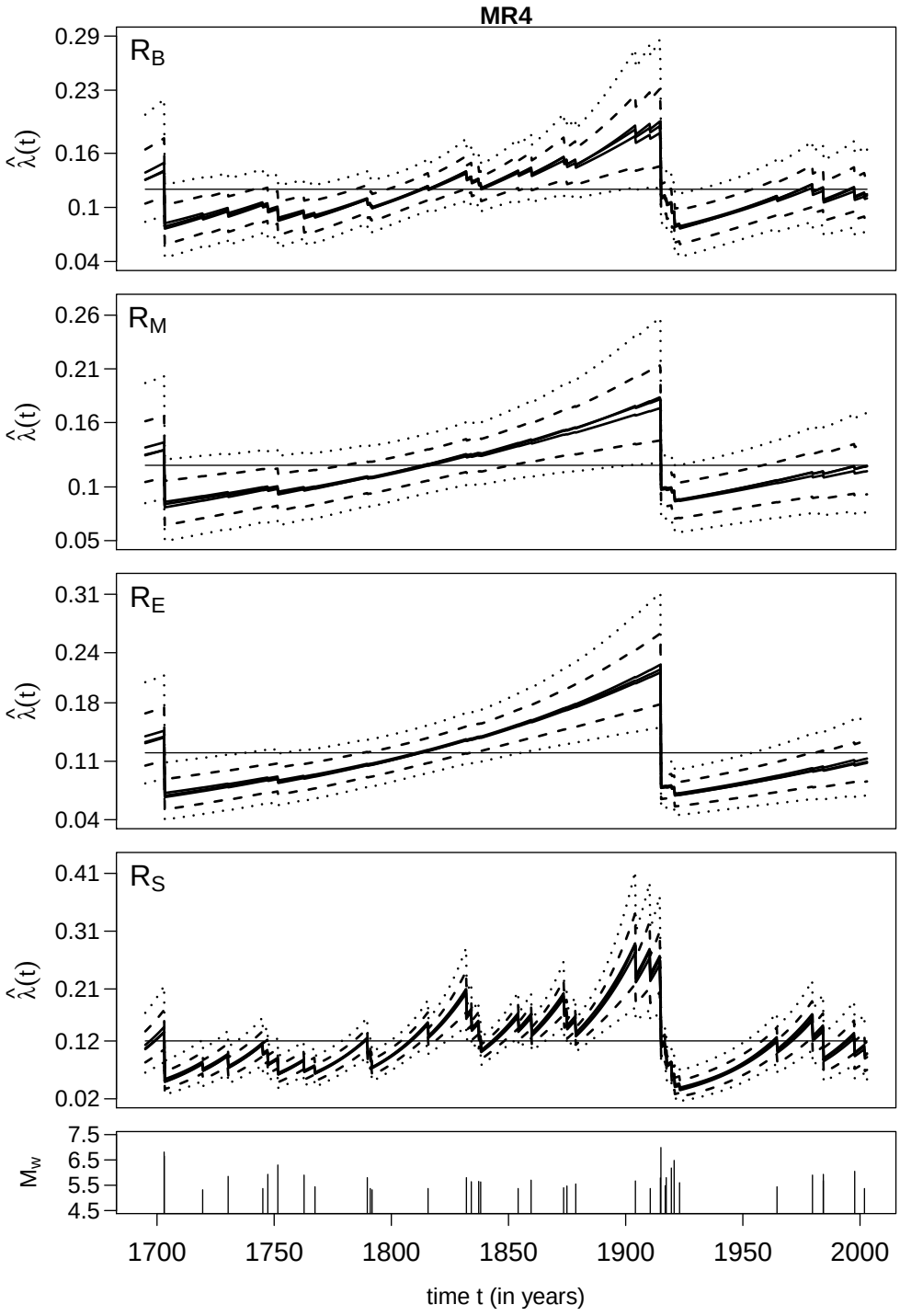

Figure 4: Same as Figure 3. Example taken from MR_4.

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

	$\mathbf{R_B}$			$\mathbf{R_M}$		
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
MR_1	-5.3508	0.2367	0.0571	-6.6818	0.1810	0.1413
MR_2	-3.0279	0.1020	0.1735	-2.9561	0.0115	0.6093
MR_3	-1.8252	0.0508	0.3098	-1.7104	0.0112	0.6514
MR_4	-2.0357	0.0257	0.3624	-2.0443	0.0018	2.6527
MR_5	-2.6620	0.0713	0.2410	-2.7102	0.0058	1.4626
MR_6	-2.6449	0.0469	0.2827	-2.6144	0.0045	3.0722
MR_7	-2.2067	0.0100	0.6617	-2.3179	0.0007	9.4255
MR_8	-3.4371	0.0276	0.2664	-3.3183	0.0010	6.3041
	$\mathbf{R_E}$			$\mathbf{R_S}$		
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
MR_1	-6.3043	0.3624	0.0560	-5.7141	2.0955	0.0079
MR_2	-3.2435	0.0476	0.2553	-3.2435	1.7120	0.0232
MR_3	-1.6010	0.0533	0.2257	-2.0779	0.5719	0.0528
MR_4	-2.0275	0.0044	1.4209	-2.2461	0.9074	0.0334
MR_5	-2.7300	0.0203	0.5817	-2.6639	2.5655	0.0240
MR_6	-2.4334	0.0116	1.7385	-2.6036	1.6341	0.0174
MR_7	-2.2292	0.0008	5.1907	-2.3543	0.4887	0.0449
MR_8	-3.2833	0.0007	6.2831	-4.0703	1.4187	0.0175

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:

$$B_{12} \;=\; \frac{pr(\mathbf{D} \mid \mathcal{M}_1)}{pr(\mathbf{D} \mid \mathcal{M}_2)} \;=\; \frac{pr(\mathcal{M}_1 \mid \mathbf{D})}{pr(\mathcal{M}_2 \mid \mathbf{D})} \div \frac{pr(\mathcal{M}_1)}{pr(\mathcal{M}_2)} \qquad . \tag{16}$$

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

$$pr(\mathbf{D} \mid \mathcal{M}_k) = \int pr(\mathbf{D} \mid \theta_k, \mathcal{M}_k)\; \pi(\theta_k|\mathcal{M}_k)\; d\theta_k \tag{17}$$

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 and Varini* (2007). 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.

To compare more easily these results, Table 7 shows the set of pairwise Bayes factors for each MR: according to the interpretation of Jeffreys' scale given by *Kass and Raftery* (1995), values in the three ranges (0, 0.5), (0.5, 1), (1, 2) of the $\log_{10} B_{12}$ indicate barely worth mentioning, positive, and strong evidence in favor of the model $\mathcal{M}_1$, respectively.

Table 6: Marginal $\log_{10}$-likelihood of the four stress release model versions. In bold: the maximum value indicating the best model in each MR.

region \ model	R_B	R_M	R_E	R_S
MR_1	-15.2526	-13.8646	**-13.8297**	-15.6588
MR_2	-27.3753	-27.6254	-28.1152	**-27.1912**
MR_3	**-48.9303**	-49.0073	-49.1549	-49.0412
MR_4	-50.4663	-50.7873	-51.5084	**-50.0828**
MR_5	-22.0331	-22.3722	-23.1623	**-21.4748**
MR_6	-28.3109	-28.2706	-28.5527	**-28.0712**
MR_7	**-43.3095**	-43.5711	-43.4778	-43.3455
MR_8	-35.6121	-36.4351	-35.7348	**-34.9708**

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

MR_1

M_2 \ M_1	R_B	R_M	R_E	R_S
R_B		1.39	1.42	0.41
R_M	1.39		0.04	1.79
R_E	1.42	0.04		1.83
R_S	0.41	1.79	1.83	

MR_2

M_2 \ M_1	R_B	R_M	R_E	R_S
R_B		0.25	0.74	0.18
R_M	0.25		0.49	0.43
R_E	0.74	0.49		0.92
R_S	0.18	0.43	0.92	

MR_3

M_2 \ M_1	R_B	R_M	R_E	R_S
R_B		0.08	0.23	0.11
R_M	0.08		0.15	0.03
R_E	0.23	0.15		0.11
R_S	0.11	0.03	0.11	

MR_4

M_2 \ M_1	R_B	R_M	R_E	R_S
R_B		0.32	1.04	0.38
R_M	0.32		0.72	0.70
R_E	1.04	0.72		1.43
R_S	0.38	0.70	1.43	

MR_5

M_2 \ M_1	R_B	R_M	R_E	R_S
R_B		0.34	1.13	0.56
R_M	0.33		0.79	0.90
R_E	1.13	0.79		1.69
R_S	0.56	0.90	1.69	

MR_6

M_2 \ M_1	R_B	R_M	R_E	R_S
R_B		0.04	0.24	0.24
R_M	0.04		0.28	0.20
R_E	0.24	0.28		0.48
R_S	0.24	0.20	0.48	

MR_7

M_2 \ M_1	R_B	R_M	R_E	R_S
R_B		0.26	0.17	0.04
R_M	0.26		0.09	0.23
R_E	0.17	0.09		0.13
R_S	0.04	0.23	0.13	

MR_8

M_2 \ M_1	R_B	R_M	R_E	R_S
R_B		0.82	0.12	0.64
R_M	0.82		0.70	1.46
R_E	0.12	0.70		0.76
R_S	0.64	1.46	0.76	

5 Probability distribution of the 'time to the next event'

In this Section, in an explicit way, we derive the probability distribution of the time to the next event for the specific class of stress-release models. At the instant t, let us consider the conditional intensity function:

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

of the general stress release 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:

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

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:

$$F_t(w \mid \theta) = 1 - \exp\{-\phi_t\ (e^{\eta w} - 1)\}\ , \tag{18}$$

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:

$$f_t(w \mid \theta) = \eta\phi_t e^{\eta w} e^{\phi_t} \exp(-\phi_t e^{\eta w}). \tag{19}$$

This function can take a large variety of shapes. It can be skewed to the right or to the left. To describe the characteristics of the Gompertz distribution (18), we give its summary statistics: mode, mean, variance, and quartiles. The mode of the density function (19) is as follows:

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

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

$$E(W_t \mid \theta) = -\frac{e^{\phi_t}}{\eta}\mathrm{Ei}(-\phi_\mathrm{t}), \tag{20}$$

where Ei() is the exponential integral $\mathrm{Ei(x)} = -\int_{-\mathrm{x}}^{\infty} (\mathrm{e}^{-\mathrm{u}}/\mathrm{u})\ \mathrm{du}$ (*Abramowitz and Stegun* (1972), p. 228).

For a given η, when ϕ_t (or equivalently $\lambda(t)$) gets close to 0, the equation (20) 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. At the same time, 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 of the hazard function computed at the occurrence time. Indeed, if an earthquake of size X_i occurs at time t_i, the drop of the hazard function, $\Delta\lambda(t_i) = \lambda(t_i^-)\ [1 - \exp(-\beta X_i)]$, depends on the value of the hazard function $\lambda(t_i^-)$ computed 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. Depending on the conditions of the system at that moment, the stress release model does not preclude a small waiting time, even immediately after a strong event.

The variance of W_t is such that:

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

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

The generic quantile of order q, and hence also the median corresponding to $q = 0.50$, is given by:

$$W_q = \eta^{-1}\log(1 - \phi_t^{-1}\log(1-q)).$$

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 $t+h$.

For all $h > 0$, we have $\phi_{t+h} = \phi_t e^{\eta h} \geq \phi_t$. Hence the expected waiting time of W_{t+h} decreases as h increases, and also $E(W_t \mid \theta) \geq E(W_{t+h} \mid \theta)$:

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

As ϕ_{t+h} tends to infinity as h increases, it holds (*Abramowitz and Stegun* (1972), p. 229)

that:

$$\frac{1}{2\eta}\ln\left(1+\frac{2}{\phi_{t+h}}\right) < E(W_{t+h} \mid \theta) < \frac{1}{\eta}\ln\left(1+\frac{1}{\phi_{t+h}}\right). \tag{21}$$

Therefore 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 the variance decreases to zero when h tends to infinity.

5.1 Forecast of the 'time to the next event' in the Bayesian framework

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:

$$F_t(w) = P(W_t < w) = \int_\Theta P(W_t < w \mid \theta)\pi(\theta \mid data)d\theta \ ,$$

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):

$$F_t(w) \approx \hat{F}_t(w) = \frac{\sum_{j=1}^{R} P(W_t < w \mid \theta^{(j)})}{R} \ . \tag{22}$$

The expected waiting time of W_t is estimated by the average of the expected waiting times $E(W_t \mid \theta^{(j)})$, $j = 1, ..., R$, as given by (20); 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 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 $T = t + W_t$, which links the time of the next event T with the corresponding waiting time W_t, 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.

6 Results

In this Section, we first illustrate the results obtained by performing a Bayesian comparison of the four versions of the stress release model introduced in Section 2, and then we show how these models can be tested through earthquake forecast procedures.

6.1 Bayesian model comparison

Based on the Bayes factor, we evaluated how well, comparatively, the four versions ($\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, $\mathbf{R_S}$) of the stress release model fit to the events that occurred in the eight MRs. Table 6 shows the marginal likelihood of each model for each MR. The frequency of the highest values indicates that the most plausible model is $\mathbf{R_S}$, and $\mathbf{R_B}$ appears to be the second best. More specifically, comparing the Bayes factors (Table 7) of all of the model pairs in each MR with Jeffreys' scale (*Kass and Raftery*, 1995), it can be seen that:

In MR_1, $\mathbf{R_E}$ behaves quite similarly to $\mathbf{R_M}$, and both of them show strong evidence against $\mathbf{R_B}$ and $\mathbf{R_S}$;

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

In MR_3, $\mathbf{R_B}$ performs just slightly better than the other models, with $\mathbf{R_E}$ being the worst again;

In MR_4, $\mathbf{R_S}$ behaves similarly to $\mathbf{R_B}$ and performs from substantially-to-strongly better than the other two models;

In MR_5, there is clear evidence in favor of $\mathbf{R_S}$, and substantial evidence against $\mathbf{R_E}$; $\mathbf{R_B}$ is the second best;

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

In MR_7, $\mathbf{R_B}$ performs slightly better, whereas $\mathbf{R_M}$ performs slightly worse than the remaining two models;

In MR_8, there is substantial-to-strong evidence in favor of $\mathbf{R_S}$, and the same against $\mathbf{R_M}$.

Note that MR_1 counts 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. In MR_4, MR_5, and MR_8, and also in MR_2, $\mathbf{R_S}$ performs substantially better than the other three models if we assimilate $\mathbf{R_B}$ to $\mathbf{R_S}$ in consideration of the low value of the Bayes factor for this couple of models. There is no substantial difference among models in MR_3, MR_6, and MR_7. With reference to Equations (8-11), these results indicate that a coefficient of 0.75 in the exponential term should be preferred rather than the other alternatives, because it is shared by the two best-performing models. The information on the faulting geometry provided through the rupture area (A) appears to improve the performance of the $\mathbf{R_S}$ model in all MRs in which the results are significant. Therefore, improvements in the fault-scaling laws used to estimate the rupture area could further improve the performance of the $\mathbf{R_S}$ model.

6.2 Retrospective- and prospective-forecast validation

The retrospective validation is carried out by forecasting the occurrence time of each earthquake (target event) included in each MR dataset, right after the occurrence of the event that precedes it; the discrepancy between the forecast time and the actual earthquake occurrence time is then calculated. The discrepancies are used to rate the various forms of the stress release model in each MR. To this end, we use the Gompertz distribution (Equation 22) and its statistical summaries: mean, median, 75% HPD interval, and 90% HPD interval. Figure 5 provides two forecast examples: one, (retrospectively) issued in MR_2 on 1812/10/25, the date of the occurrence of a M_w 5.70 earthquake, shows a waiting time to the next event that relatively closely predicts the occurrence date of the 1836/06/12, M_w 5.48, earthquake; the other is issued in MR_7 on 1975/01/16, the date of occurrence of a M_w 5.38 earthquake, and closely predicts the waiting time to the 1983/11/08, M_w 5.37, earthquake. Note the different shapes of the two density functions that characterize the expected inter-event times varying from almost 25 to less than 10 years.

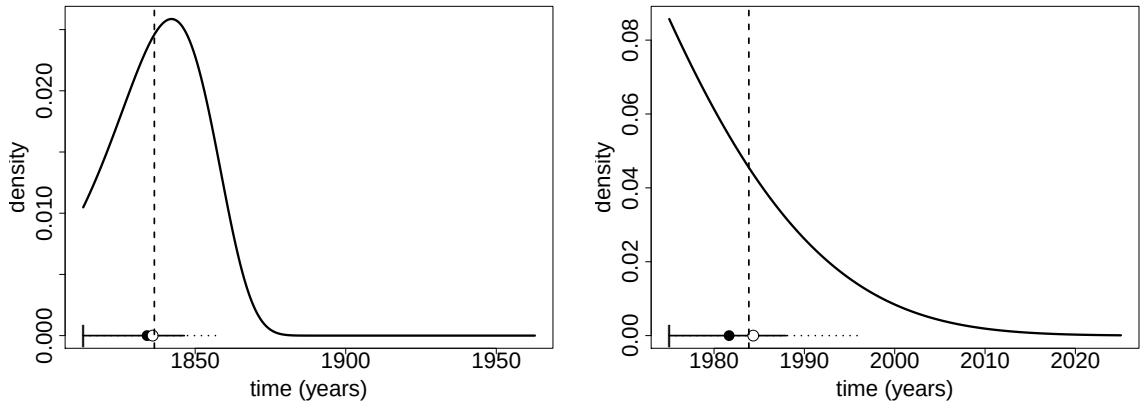

Figure 5: Examples of the estimated density functions of the time to the next event, and their statistical summaries. Legend: Gompertz density function (solid curve), mean (open circle), median (solid circle), 75% HPD (solid horizontal segment) and 90% HPD (dotted horizontal segment) intervals. The forecast 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 based on the $\mathbf{R_S}$ model and taken from MR_2 (left) and MR_7 (right).

Figure 6 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 5). The results for the other MRs are shown in Appendix C (Figures C.1-C.7). 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 inter-event 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 40 events examined, whereas for 33 events it is within the 75% HPD interval.

Table 8 summarizes the discrepancies of the forecasts for the four versions of the stress release models in the eight MRs, in terms of average length of the 75% HPD and 90% HPD intervals, as well as the mean square (absolute) error between the mean (median) and the observed time. The lowest values (or minimum discrepancy) are most often provided by the $\mathbf{R_S}$ model, which suggests the scaled energy as the most appropriate quantity to be used in stress release models.

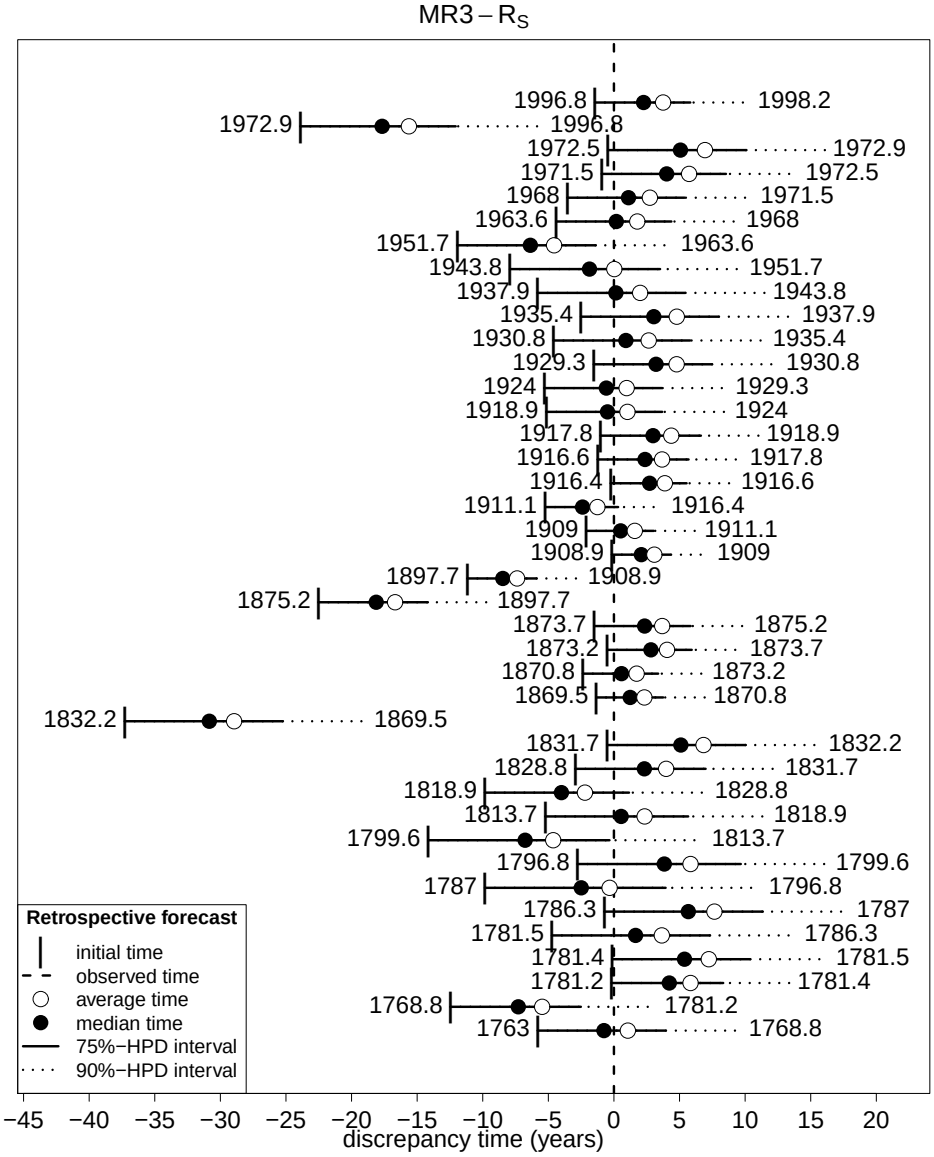

Figure 6: Time lines of 40 retrospective forecasts for MR_3, $\mathbf{R_S}$ model, in order of descending date from the top to the bottom. Each forecast is imagined to have been issued on the occurrence date (shown on the left, and marked by a vertical bar) of an event in the MR dataset, and to be aimed 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 5). The time lines are shifted laterally so that they intersect the vertical dashed line at the actual occurrence date of the target event.

To conduct a 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 the CPTI11 for the period from 2003-2006, and ISIDe for the period from 2007-2012 (see Section 3), and we found 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 ($M_w \geq 5.3$) 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, 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.

Table 9 summarizes the prospective forecasts provided by the $\mathbf{R_S}$ model for each MR. 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 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 9, first line in 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 ± 8.3 years; or by 2008.9,

2015.2, or 2022.6 with probabilities of 50%, 75%, and 90%, respectively. A little more than a year later (2003/01/01; Table 9, second line), by adding the information that no event had occurred in the meanwhile, the expected time to the next event moves forward by almost 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 9, 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 9, 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 2021, according to the mean value, or by 2019.5, 2025.4, and 2032.5, with probabilities of 50%, 75%, and 90%, respectively.

In MR_3, three earthquakes occurred in the period 2003-2012, and thus the forecasts are successively updated after each one of them. 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 1.57 years when considering the median values, whereas the mean square error is 7.14 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 7) they allow, and the similar values they achieve (Table B.1.

For completeness of information, Table C.1 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 stress release model in every MR.

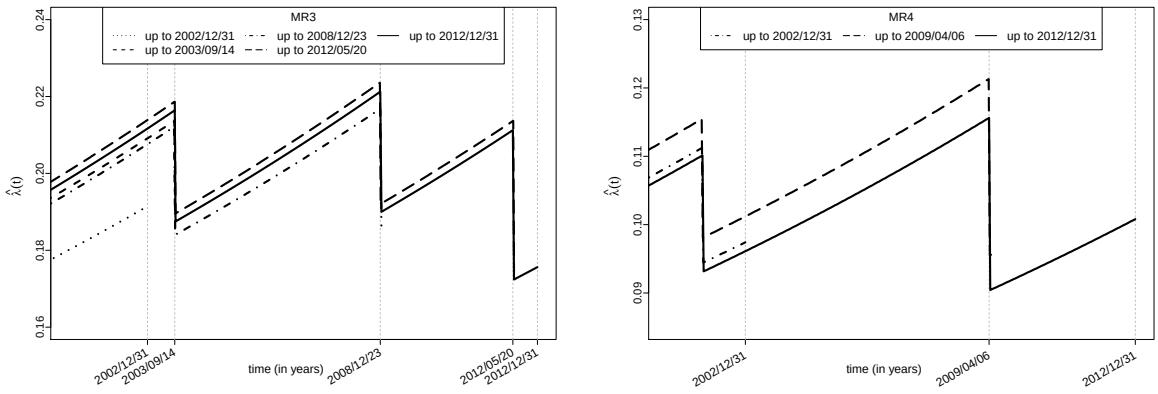

Figure 7: Estimate (ergodic mean) of the intensity function for the $\mathbf{R_S}$ model in MR_3 and MR_4, updated whenever new information (earthquake occurrence) is included in the relevant dataset.

7 Final remarks

We examined four different versions of the classic stress-release 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 - chosen to represent the physical process responsible for the generation of earthquakes. Equations (8)-(11) highlight the key elements (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 shape of the relevant conditional intensities (Figures 3 and 4). Despite the general similarity, note that the conditional intensity variation (equivalent to a hazard drop) is different in different stress release models, depending on the size of the intervening earthquake. With reference to Figure 4, 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 stress release model produces a hazard decrease that is relatively heightened for smaller earthquakes and abated for larger earthquakes.

As for the model comparison, the Bayes factor indicates (Table 7) that the $\mathbf{R_S}$ model performs from substantially ($\mathbf{R_S}$ vs. $\mathbf{R_B}$) to strongly ($\mathbf{R_S}$ vs. $\mathbf{R_M}$ and $\mathbf{R_E}$) better than the other models in MR_2, MR_4, MR_5, and MR_8. $\mathbf{R_E}$ and $\mathbf{R_M}$ perform considerably better than the others only in MR_1; nevertheless, we recall that results for this MR should be taken cautiously because of its reduced number of events (only seven) and its non-uniform tectonic characterization. Inconclusive results are instead obtained in MR_3, MR_7, and MR_6, in which the Bayes factor shows only slight evidence in favor of either $\mathbf{R_S}$ or $\mathbf{R_B}$. Considering that the $\mathbf{R_S}$ model performance prevails in the analyzed cases, we maintain that the scaled energy is likely to be the best way of measuring the size of earthquakes in stress release model applications.

The probability distribution of the time to the next event for the stress release model has been analytically identified as the Gompertz distribution (Section 5). This finding brings about an 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 stress release model versions. Different measures of the discrepancy between the expected occurrence time of an earthquake and the time of its actual occurrence (Table 8) have shown that the retrospective analysis supports the choice of the $\mathbf{R_S}$ model in most of the cases analyzed.

Table 8: 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, mean square (absolute) error between expected mean (median) and observed occurence times. In bold, the lowest values.

region	model	HPD average length		average discrepancy	
		90%	75%	median	mean
MR_1	R_B	101.7	68.5	29.4	42.7
	R_M	**67.0**	**46.3**	**16.2**	**28.8**
	R_E	68.3	47.4	18.5	29.7
	R_S	104.3	75.1	34.7	45.1
MR_2	R_B	33.7	**21.5**	9.2	12.6
	R_M	36.1	21.7	9.3	13.3
	R_E	37.1	23.0	10.0	14.6
	R_S	**30.8**	21.7	**9.1**	**11.7**
MR_3	R_B	14.5	8.9	4.6	7.4
	R_M	14.6	**8.8**	4.7	7.5
	R_E	14.9	9.1	4.6	7.5
	R_S	**14.4**	9.2	**4.5**	**7.2**
MR_4	R_B	20.4	**12.4**	6.6	8.8
	R_M	20.9	12.5	6.7	9.1
	R_E	22.0	13.2	6.9	9.3
	R_S	**19.6**	12.9	**6.5**	**8.4**
MR_5	R_B	31.7	20.1	8.9	12.1
	R_M	34.0	20.5	9.2	12.9
	R_E	37.6	23.4	10.2	14.7
	R_S	**26.9**	**20.0**	**8.2**	**10.2**
MR_6	R_B	51.3	**33.5**	12.7	17.6
	R_M	52.1	34.4	13.1	18.0
	R_E	53.0	37.4	13.8	18.9
	R_S	**47.0**	34.0	**12.5**	**16.3**
MR_7	R_B	21.6	**13.0**	6.9	8.5
	R_M	22.0	13.3	7.0	8.7
	R_E	22.2	13.3	7.0	8.7
	R_S	**21.2**	13.6	**6.8**	**8.4**
MR_8	R_B	52.7	33.0	15.4	20.6
	R_M	58.6	36.6	16.6	23.4
	R_E	54.7	33.0	15.2	21.2
	R_S	**45.2**	**31.7**	**14.0**	**17.6**

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 9). 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 5 years. However, no earthquake has occurred in MR_7, up until the end of 2012. By adding this information to the 2013 update, the forecast considerably postpones the expected occurrence time of the next event (by more than 10 years).

As we anticipated in Section 5, 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 9 confirm this general behavior, although the amount of delay and uncertainty gain remains variable, depending on repeated parameter estimates.

It is important to recall that both the time and space scales of the stress release 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 stress release 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 scaled energy as the best option in stress release models, and for the refining of the time-space limits of the stress release model applications in robust earthquake forecasting.

Depending on the data availability, possible future research directions could also aim at developing the *linked* (or *coupled*) version of the stress release model (e.g.: *Bebbington and Harte* (2003), *Kuehn et al.* (2008)) on the same Italian data, using the scaled energy as the measure of the sizes of the events.

Table 9: Prospective forecasts in each MR according to the R_S version of the stress release model. All 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).

region	date of forecast issue	HPD 75%	HPD 90%	median	mean (st.dev.)
MR_1	1887.2	1887.2-2011.7	1887.2-2056.9	1968.5	1977.8 (45.1)
	2003.0	2003.0-2048.9	2003.0-2084.5	2025.0	2038.8 (45.1)
	2013.0	2013.0-2060.8	2013.0-2098.4	2035.8	2050.1 (47.6)
MR_2	1977.7	1977.7-2014.9	1977.7-2028.3	2000.6	2003.4 (18.2)
	2003.0	2003.0-2023.5	2003.0-2034.1	2014.2	2017.4 (12.5)
	2013.0	2013.0-2030.7	2013.0-2040.5	2022.5	2025.6 (11.6)
MR_3	1998.2	1998.2-2006.2	1998.2-2011.1	2002.3	2003.9 (5.4)
	2003.0	2003.0-2010.0	2003.0-2014.5	2006.6	2008.0 (4.9)
	2003.7[(1)]	2003.7-2010.9	2003.7-2015.4	2007.4	2008.8 (5.0)
	2009.0[(2)]	2009.0-2016.1	2009.0-2020.6	2012.6	2014.1 (4.9)
	2012.4[(3)]	2012.4-2020.0	2012.4-2024.6	2016.3	2017.8 (5.1)
	2013.0	2013.0-2020.5	2013.0-2025.1	2016.9	2018.3 (5.1)
MR_4	2001.9	2001.9-2015.2	2001.9-2022.6	2008.9	2011.2 (8.3)
	2003.0	2003.0-2015.9	2003.0-2023.2	2009.8	2012.0 (8.1)
	2009.3[(4)]	2009.3-2022.4	2009.3-2029.8	2016.2	2018.5 (8.2)
	2013.0	2013.0-2025.4	2013.0-2032.5	2019.5	2021.7 (7.9)
MR_5	2002.8	2002.8-2016.7	2002.8-2023.9	2010.6	2012.8 (3.6)
	2003.0	2003.0-2016.8	2003.0-2023.9	2010.6	2012.9 (9.0)
	2013.0	2013.0-2024.2	2013.0-2031.4	2018.8	2021.3 (8.9)
MR_6	1998.7	1998.7-2025.9	1998.7-2039.7	2013.9	2018.0 (17.0)
	2003.0	2003.0-2028.0	2003.0-2041.4	2016.7	2020.9 (16.3)
	2013.0	2013.0-2035.2	2013.0-2048.4	2024.8	2029.2 (15.8)
MR_7	2001.4	2001.4-2009.5	2001.4-2015.2	2005.4	2007.4 (8.1)
	2003.0	2003.0-2010.9	2003.0-2016.5	2006.9	2008.8 (6.3)
	2013.0	2013.0-2021.9	2013.0-2028.4	2017.3	2019.6 (7.4)
MR_8	2002.7	2002.7-2038.6	2002.7-2055.1	2022.6	2027.2 (19.7)
	2003.0	2003.0-2038.8	2003.0-2055.2	2022.8	2027.4 (19.6)
	2013.0	2013.0-2045.2	2013.0-2060.2	2030.9	2035.1 (18.0)

[(1)] just after 2003/09/14 earthquake, M_w 5.3
[(2)] just after 2008/12/23 earthquake, M_w 5.4
[(3)] just after 2012/05/20 earthquake, M_w 5.9
[(4)] just after 2009/04/06 earthquake, M_w 6.1

A Completeness of the catalog: statistical analysis

Let us consider a catalog covering 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:

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

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 change-point 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 general considerations on 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 $f_0 = 5$. As for the time interval (T_0, T_f), we set $T_f = 2003$, the end of the CPTI04, while T_0 varies in each macroregion. 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 and 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. Suppose to have some transition kernel $q(\theta, \theta^*)$ (called the *proposal distribution*), which is easy to simulate from, 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 $E_{\theta|data}\{g(\theta)\}$ 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 $\left\{\theta^{(j)}\right\}_{j=0}^{R}$, i.e., its initial part (ca. 10%), 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 chose the open-source package BOA (*Smith*, 2005) for the R system for statistical computing (*R Development Core Team*, 2006), and checked that all of the generated sequences did not fail the following tests: Geweke test, Heidelberger & Welch test, and Raftery & Lewis test (*Smith*, 2007). Table B.2 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 to be the best in the statistical literature. As an example, Figure B.2 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 the MR_4 MR. We note that the posteriors are relatively peaked even when the priors, which are based on rather vague information, are very flat.

Table B.1: Parameter estimates of the $\mathbf{R_S}$ model updated by enlarging the history $\mathcal{H}_t$ on which the intensity function is conditioned - MR_3 and MR_4 macroregions.

			$\mathbf{R_S}$		
		t	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\rho}$
MR_3	(end of the catalog)	2002/12/31	-2.0779	0.5719	0.0528
	(event)	2003/09/14	-2.1317	0.5720	0.0541
	(event)	2008/12/23	-2.1317	0.5774	0.0541
	(event)	2012/05/20	-2.1539	0.5791	0.0545
		2012/12/31	-2.1472	0.5812	0.0543
MR_4	(end of the catalog)	2002/12/31	-2.2461	0.9074	0.0334
	(event)	2009/04/06	-2.2730	0.8982	0.0337
		2012/12/31	-2.2473	0.9215	0.0334

Table B.2: Prior and proposal distributions of the model parameters $\theta = (\alpha, \beta, \rho)$ adopted in the McMC estimation method. 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.

		prior distribution			proposal distribution		
model	region	α	β	ρ	α	β	ρ
$\mathbf{R_B}$	MR1	N(-5.50; 6.25)	Γ(0.20; 3.0E-2)	Γ(0.10; 8.0E-3)	N($*$; 2.00)	LogN($*$; 2.6E-3)	LogN($*$; 7.0E-4)
	MR2	N(-5.50; 6.25)	Γ(0.20; 3.0E-2)	Γ(0.20; 3.0E-2)	N($*$; 8.0E-1)	LogN($*$; 1.5E-3)	LogN($*$; 6.0E-3)
	MR3	N(-5.50; 6.25)	Γ(0.20; 3.0E-2)	Γ(0.30; 4.0E-2)	N($*$; 3.0E-1)	LogN($*$; 4.0E-3)	LogN($*$; 1.2E-3)
	MR4	N(-5.50; 6.25)	Γ(0.20; 3.0E-2)	Γ(0.35; 8.0E-2)	N($*$; 3.5E-1)	LogN($*$; 9.0E-4)	LogN($*$; 2.0E-2)
	MR5	N(-5.50; 6.25)	Γ(0.20; 3.0E-2)	Γ(0.30; 4.0E-2)	N($*$; 9.0E-1)	LogN($*$; 1.0E-2)	LogN($*$; 2.0E-2)
	MR6	N(-5.50; 6.25)	Γ(0.20; 3.0E-2)	Γ(0.20; 3.0E-2)	N($*$; 8.0E-1)	LogN($*$; 3.0E-3)	LogN($*$; 1.3E-3)
	MR7	N(-5.50; 6.25)	Γ(0.20; 3.0E-2)	Γ(0.90; 4.9E-2)	N($*$; 3.5E-1)	LogN($*$; 2.0E-4)	LogN($*$; 2.0E-1)
	MR8	N(-5.50; 6.25)	Γ(0.05; 6.0E-4)	Γ(0.20; 3.0E-2)	N($*$; 6.0E-1)	LogN($*$; 6.0E-4)	LogN($*$; 1.5E-3)
$\mathbf{R_M}$	MR1	N(-5.50; 6.25)	Γ(0.20; 2.3E-2)	Γ(0.5; 1.6E-1)	N($*$; 2.00)	LogN($*$; 5.0E-3)	LogN($*$; 1.5E-4)
	MR2	N(-5.50; 6.25)	Γ(0.02; 2.0E-4)	Γ(0.70; 3.6E-1)	N($*$; 8.0E-1)	LogN($*$; 2.0E-4)	LogN($*$; 2.5E-1)
	MR3	N(-5.50; 6.25)	Γ(0.02; 2.0E-4)	Γ(0.70; 3.6E-1)	N($*$; 3.0E-1)	LogN($*$; 2.0E-4)	LogN($*$; 1.5E-1)
	MR4	N(-5.50; 6.25)	Γ(0.02; 2.0E-4)	Γ(3.00; 8.00)	N($*$; 3.0E-1)	LogN($*$; 5.0E-6)	LogN($*$; 3.5)
	MR5	N(-5.50; 6.25)	Γ(0.02; 2.0E-4)	Γ(2.00; 3.00)	N($*$; 9.0E-1)	LogN($*$; 8.0E-5)	LogN($*$; 1.8)
	MR6	N(-5.50; 6.25)	Γ(0.02; 2.0E-4)	Γ(4.00; 4.00)	N($*$; 8.0E-1)	LogN($*$; 2.5E-5)	LogN($*$; 1.3)
	MR7	N(-5.50; 6.25)	Γ(0.02; 2.0E-4)	Γ(12.0; 4.9E+1)	N($*$; 3.5E-1)	LogN($*$; 1.0)	LogN($*$; 4.6E+1)
	MR8	N(-5.50; 6.25)	Γ(0.01; 2.5E-5)	Γ(5.00; 1.6E+1)	N($*$; 6.0E-1)	LogN($*$; 1.0)	LogN($*$; 1.0E+1)
$\mathbf{R_E}$	MR1	N(-5.50; 6.25)	Γ(0.30; 2.3E-2)	Γ(0.50; 1.6E-1)	N($*$; 2.00)	LogN($*$; 2.3E-2)	LogN($*$; 2.0E-4)
	MR2	N(-5.50; 6.25)	Γ(0.10; 2.5E-3)	Γ(0.50; 1.6E-1)	N($*$; 8.0E-1)	LogN($*$; 1.7E-3)	LogN($*$; 2.0E-2)
	MR3	N(-5.50; 6.25)	Γ(0.10; 2.5E-3)	Γ(0.50; 1.6E-1)	N($*$; 3.0E-1)	LogN($*$; 2.3E-3)	LogN($*$; 7.0E-3)
	MR4	N(-5.50; 6.25)	Γ(0.10; 2.5E-3)	Γ(1.50; 1.40)	N($*$; 4.0E-1)	LogN($*$; 2.0E-5)	LogN($*$; 6.0E-1)
	MR5	N(-5.50; 6.25)	Γ(0.10; 2.5E-3)	Γ(0.50; 1.6E-1)	N($*$; 8.0E-1)	LogN($*$; 7.0E-4)	LogN($*$; 1.5E-1)
	MR6	N(-5.50; 6.25)	Γ(0.10; 2.5E-3)	Γ(1.50; 1.40)	N($*$; 7.0E-1)	LogN($*$; 1.0E-4)	LogN($*$; 2.0E-1)
	MR7	N(-5.50; 6.25)	Γ(0.02; 2.0E-4)	Γ(5.00; 1.6E+1)	N($*$; 4.0E-1)	LogN($*$; 1.0E-6)	LogN($*$; 1.4E+1)
	MR8	N(-5.50; 6.25)	Γ(0.02; 2.0E-4)	Γ(5.00; 1.6E+1)	N($*$; 6.0E-1)	LogN($*$; 8.0E-7)	LogN($*$; 1.5E+1)
$\mathbf{R_S}$	MR1	N(-5.50; 6.25)	Γ(5.00; 9.00)	Γ(0.10; 8.1E-3)	N($*$; 1.40)	LogN($*$; 1.50)	LogN($*$; 8.0E-6)
	MR2	N(-5.50; 6.25)	Γ(5.00; 9.00)	Γ(0.10; 8.1E-3)	N($*$; 7.0E-1)	LogN($*$; 1.40)	LogN($*$; 1.5E-5)
	MR3	N(-5.50; 6.25)	Γ(5.00; 9.00)	Γ(0.10; 8.1E-3)	N($*$; 3.0E-1)	LogN($*$; 3.0E-1)	LogN($*$; 5.0E-5)
	MR4	N(-5.50; 6.25)	Γ(5.00; 9.00)	Γ(0.10; 8.1E-3)	N($*$; 4.0E-1)	LogN($*$; 8.0E-1)	LogN($*$; 2.0E-5)
	MR5	N(-5.50; 6.25)	Γ(5.00; 9.00)	Γ(0.10; 8.1E-3)	N($*$; 8.0E-1)	LogN($*$; 5.00)	LogN($*$; 2.0E-5)
	MR6	N(-5.50; 6.25)	Γ(5.00; 9.00)	Γ(0.10; 8.1E-3)	N($*$; 8.0E-1)	LogN($*$; 2.00)	LogN($*$; 1.0E-5)
	MR7	N(-5.50; 6.25)	Γ(5.00; 9.00)	Γ(0.10; 8.1E-3)	N($*$; 5.0E-1)	LogN($*$; 3.0E-1)	LogN($*$; 8.0E-5)
	MR8	N(-5.50; 6.25)	Γ(5.00; 9.00)	Γ(0.10; 8.1E-3)	N($*$; 5.0E-1)	LogN($*$; 5.0E-1)	LogN($*$; 5.0E-5)

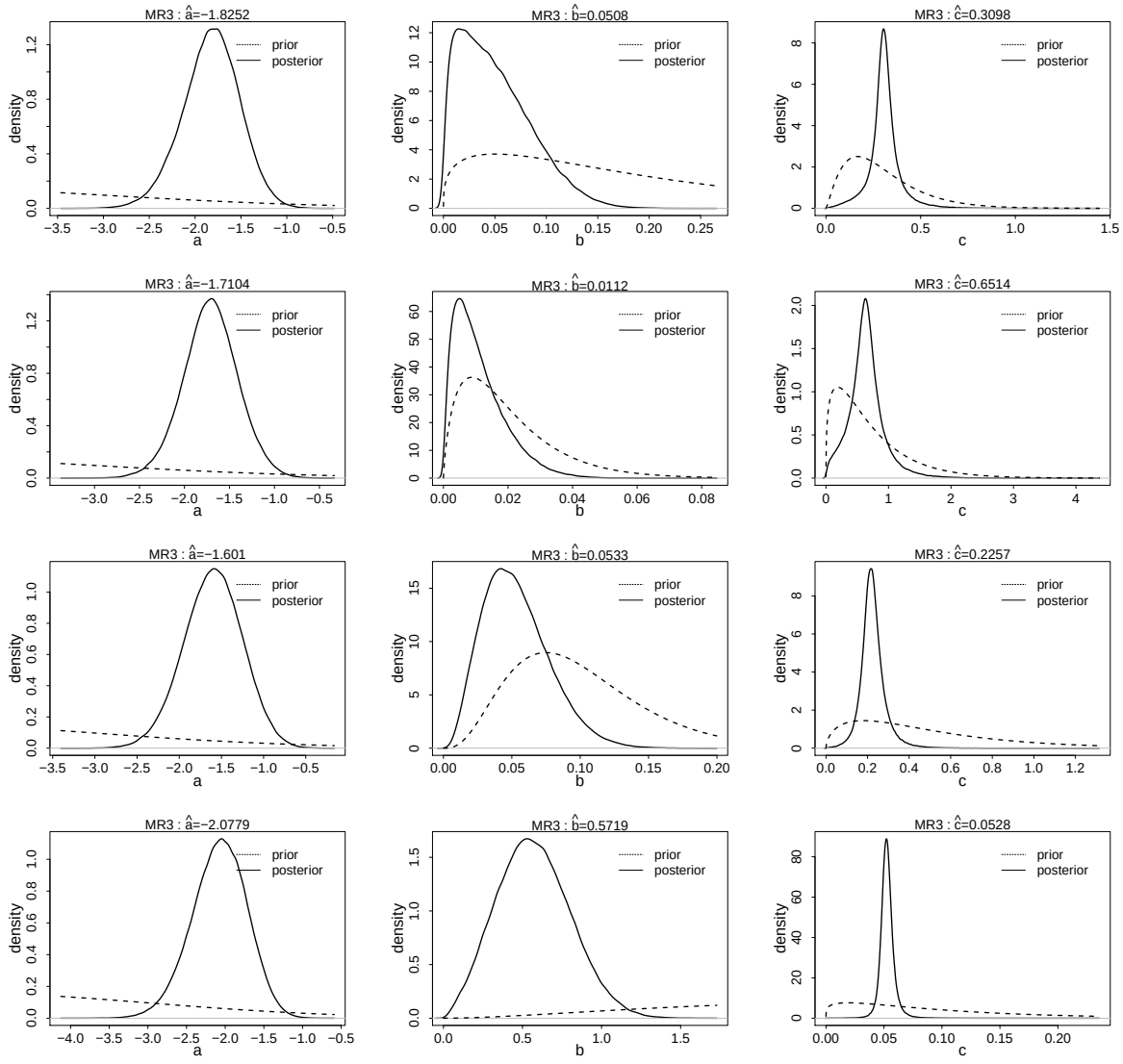

Figure B.1: From top to bottom, the $\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, $\mathbf{R_S}$ models. Prior (dotted) and posterior (solid) density functions of the parameters. Example taken from the MR_3 macroregion.

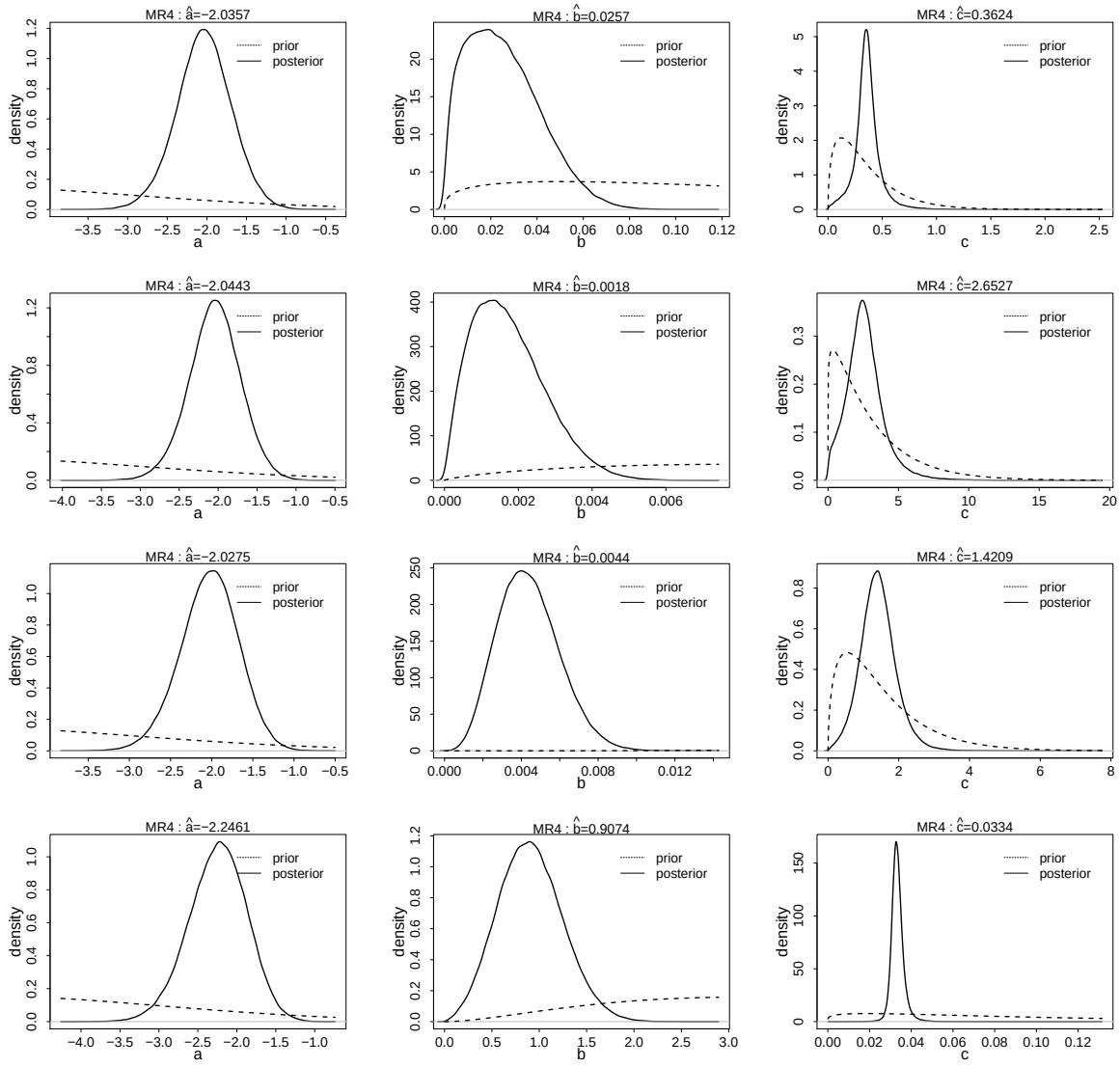

Figure B.2: Prior (dotted) and posterior (solid) density functions of the parameters of the $\mathbf{R_B}$, $\mathbf{R_M}$, $\mathbf{R_E}$, $\mathbf{R_S}$ models (from top to bottom). Example taken from the MR_4 macroregion.

C Retrospective validation

Figure C.1-C.7 summarize the retrospecive analysis of the forecasts issued at the occurrence time of every event in the datasets concerning the next event.

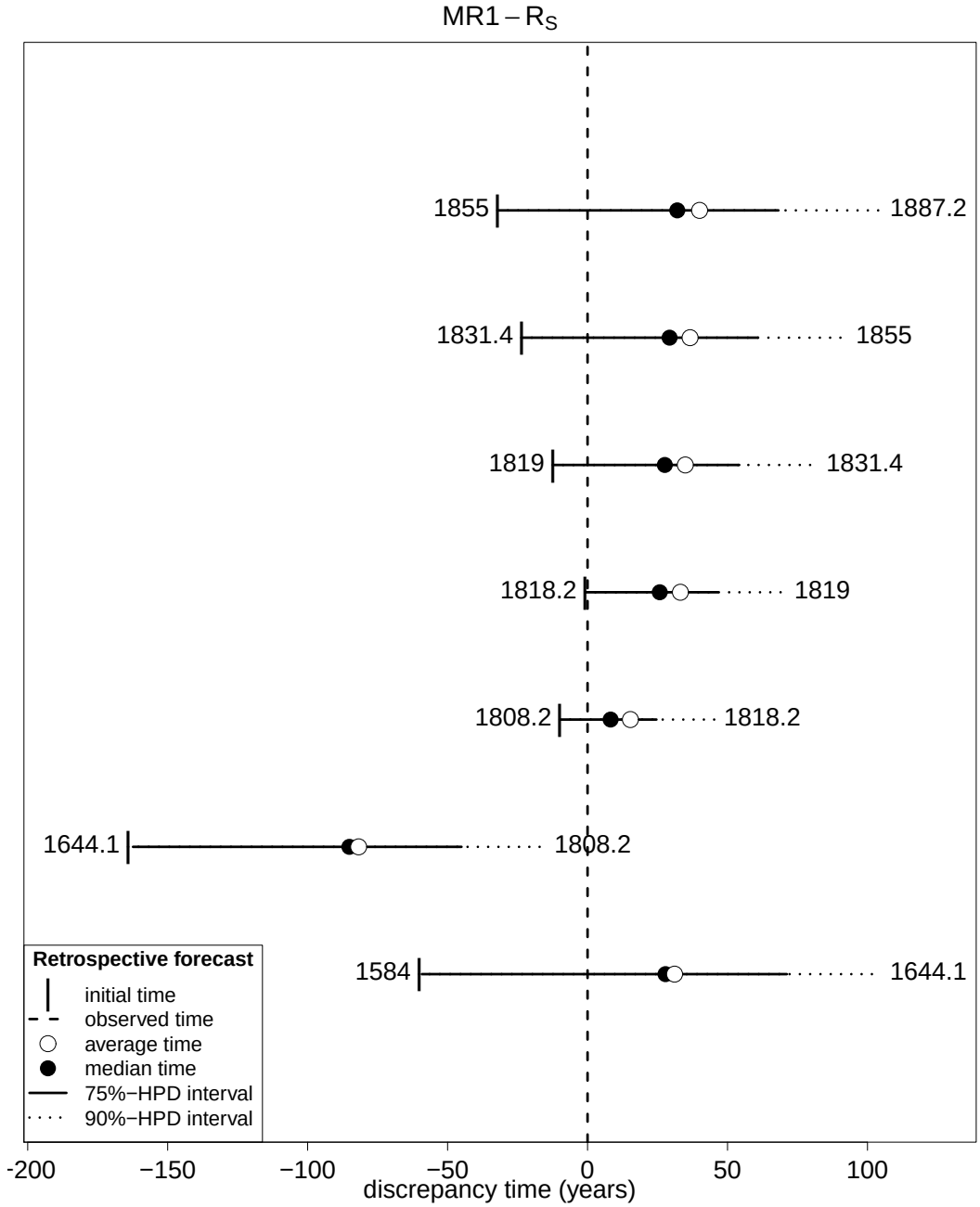

Figure C.1: As for Figure 6, validation results related to macroregion MR_1 - $\mathbf{R_S}$ model.

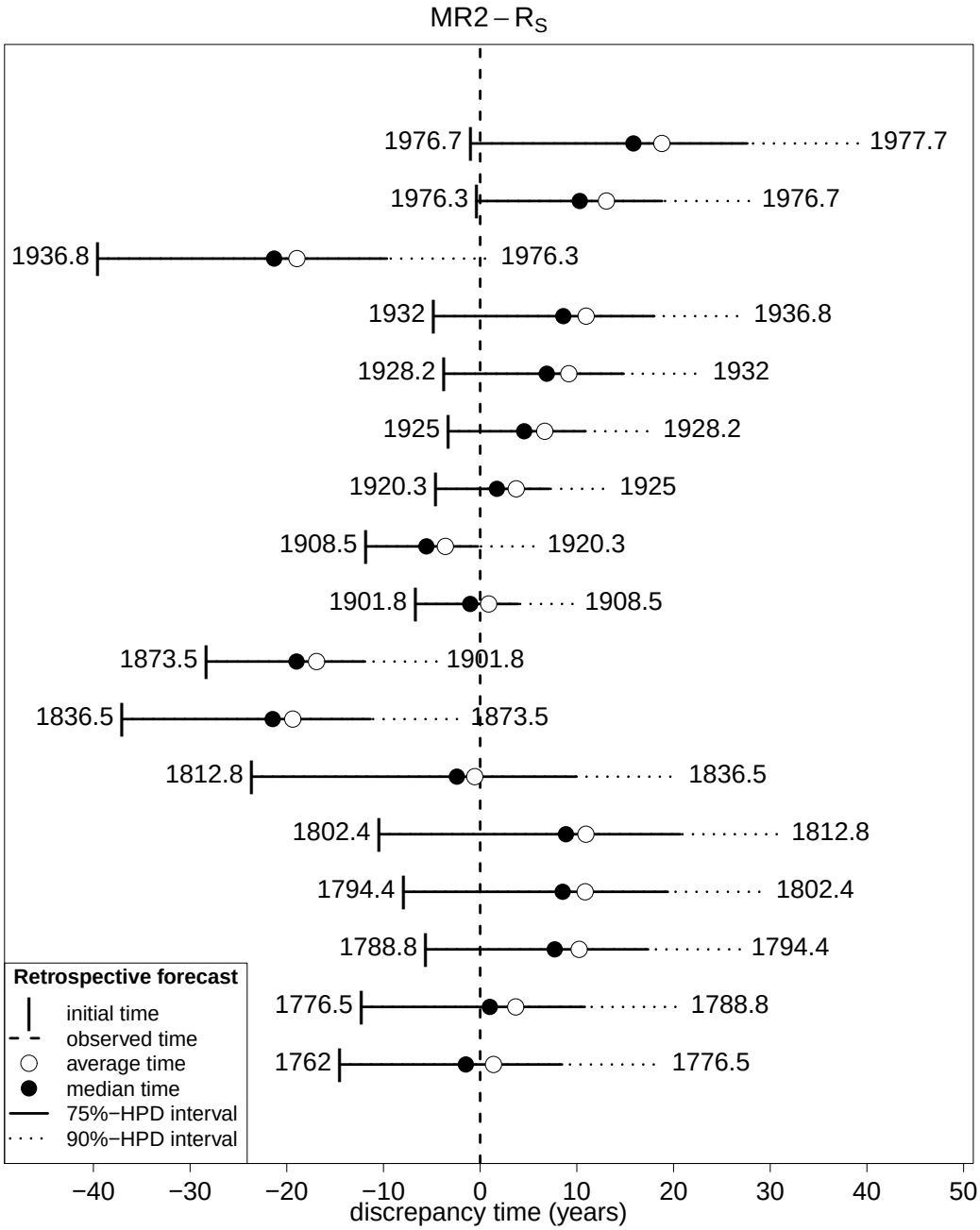

Figure C.2: As for Figure 6, validation results related to macroregion MR_2 - $\mathbf{R_S}$ model.

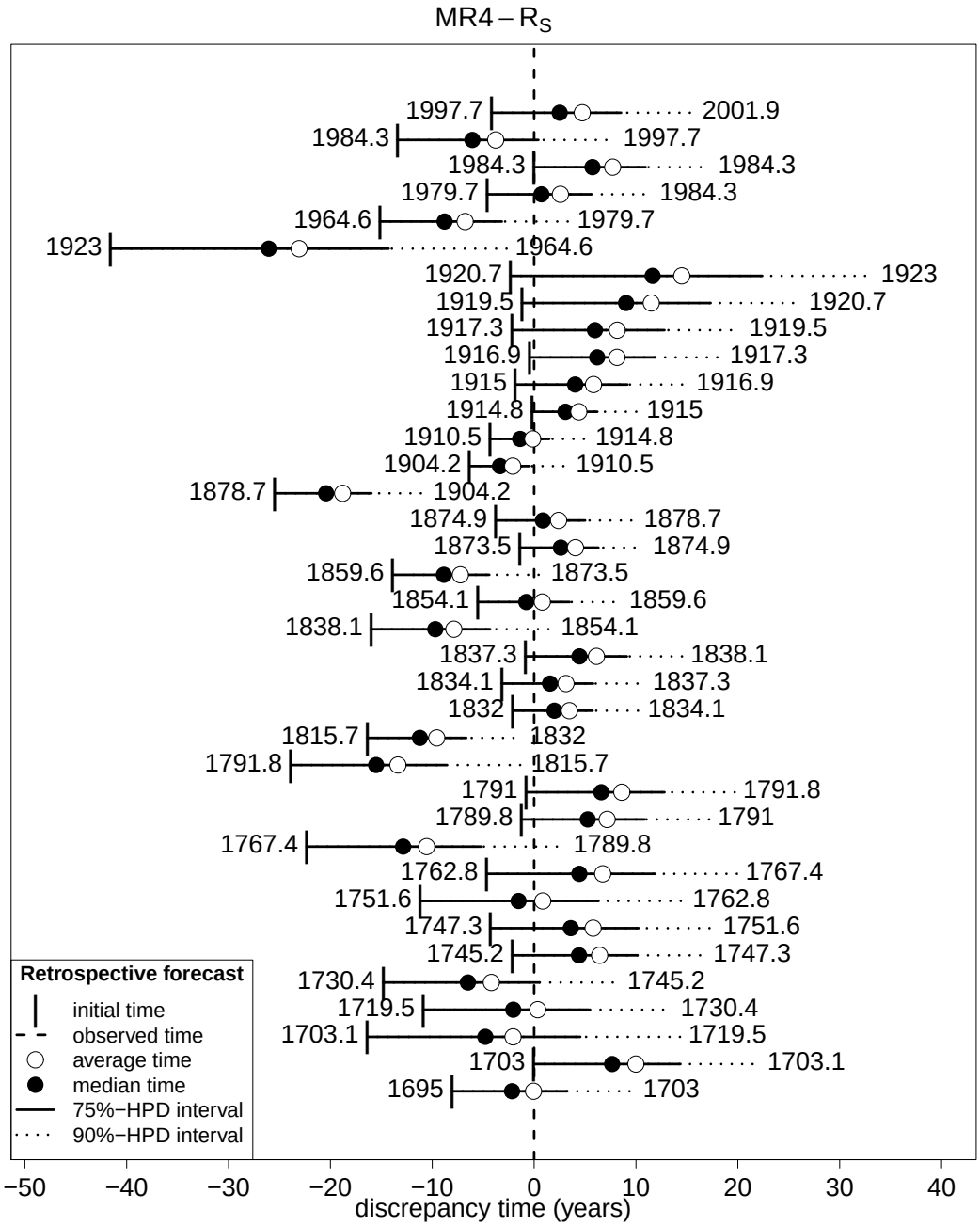

Figure C.3: As for Figure 6, validation results related to macroregion MR_4 - $\mathbf{R_S}$ model.

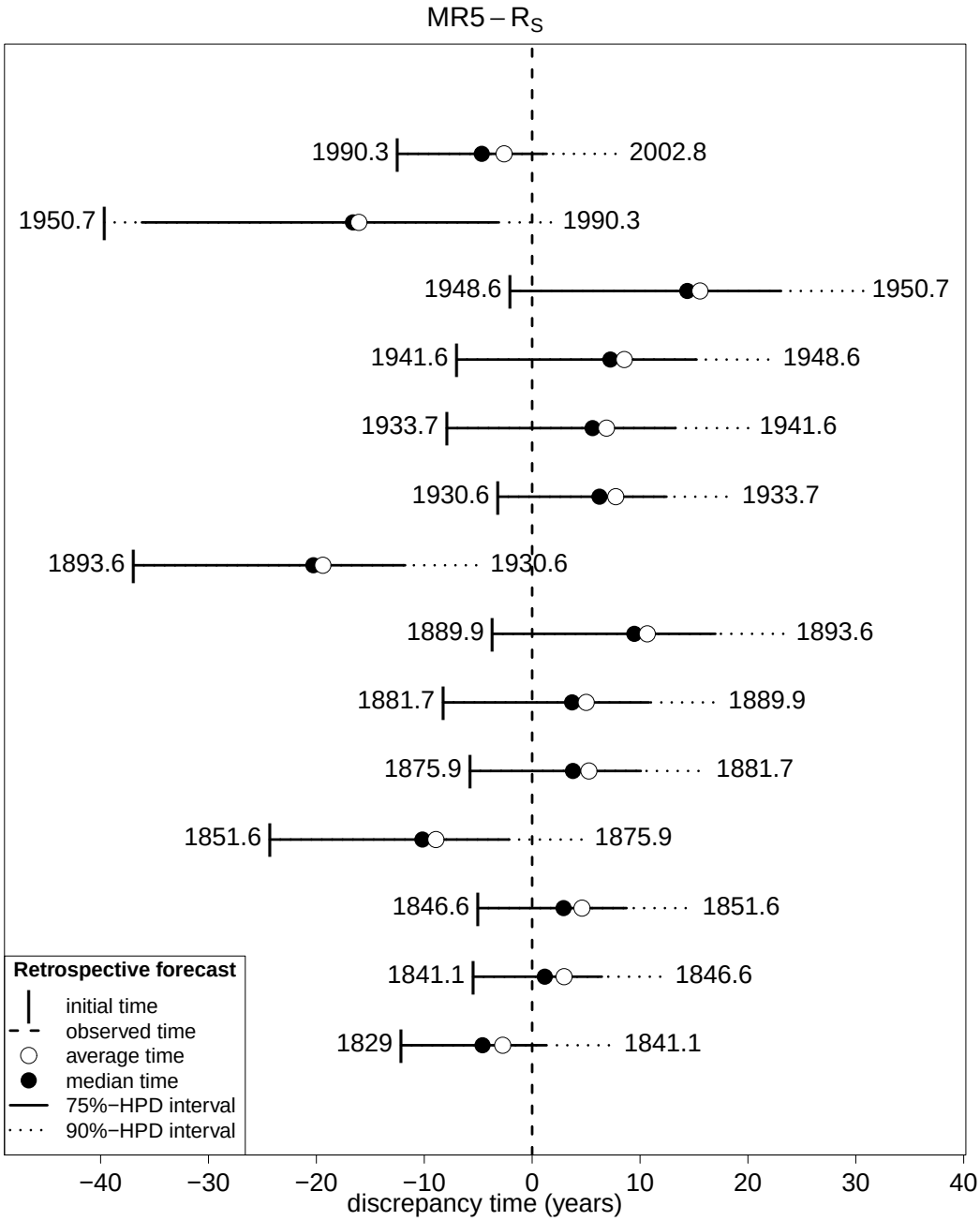

Figure C.4: As for Figure 6, validation results related to macroregion MR_5 - $\mathbf{R_S}$ model.

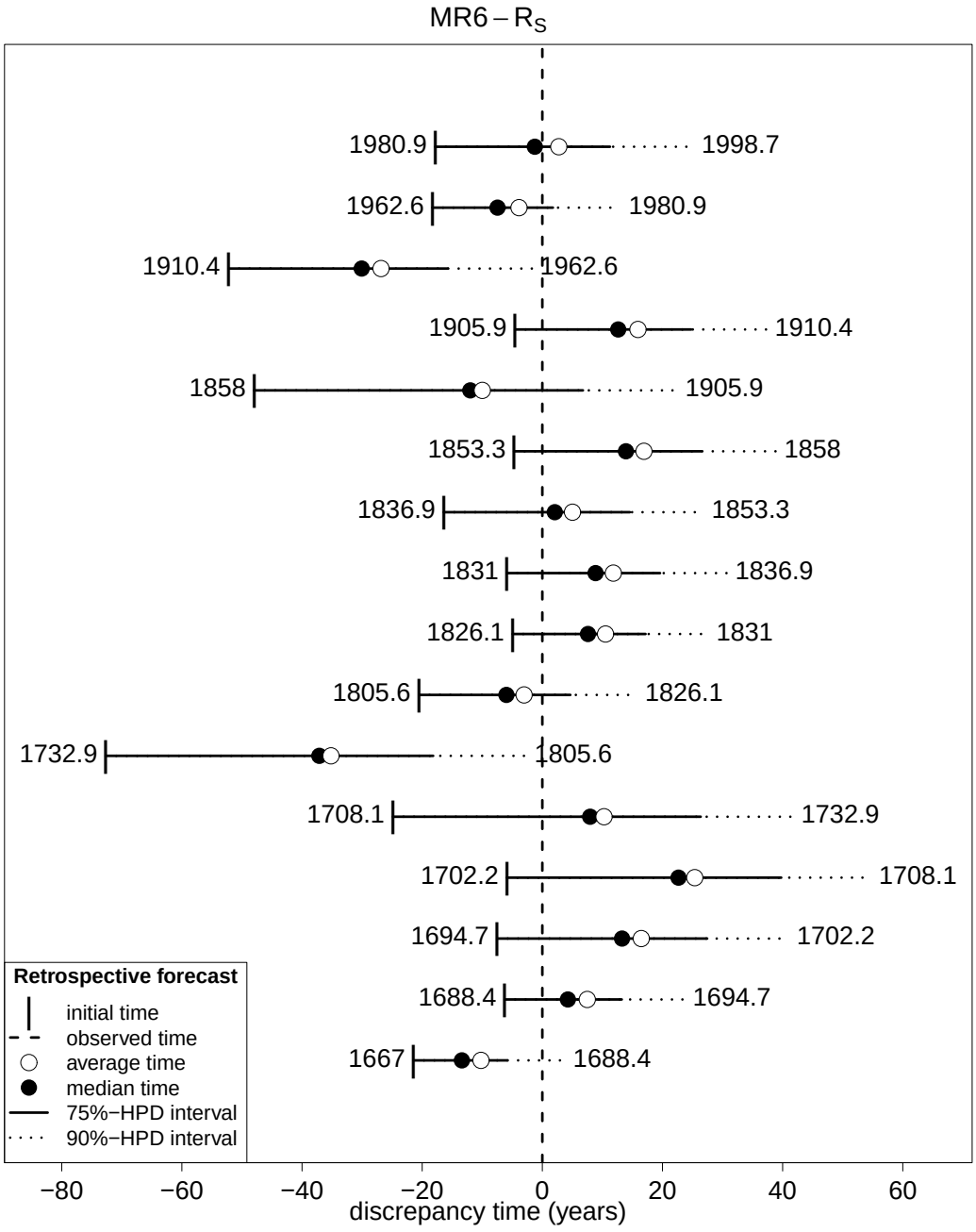

Figure C.5: As for Figure 6, validation results related to macroregion MR_6 - $\mathbf{R_S}$ model.

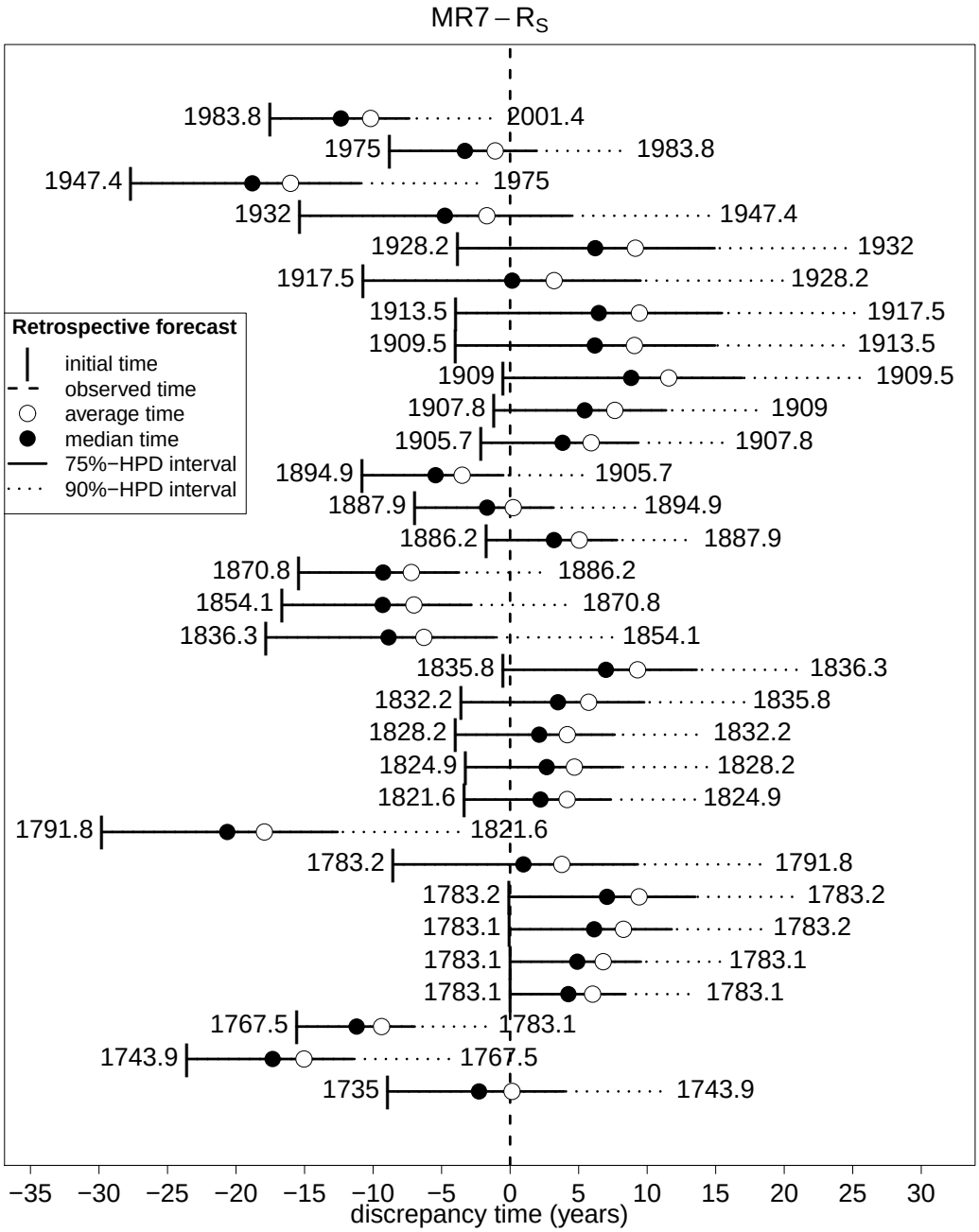

Figure C.6: As for Figure 6, validation results related to macroregion MR_7 - $\mathbf{R_S}$ model.

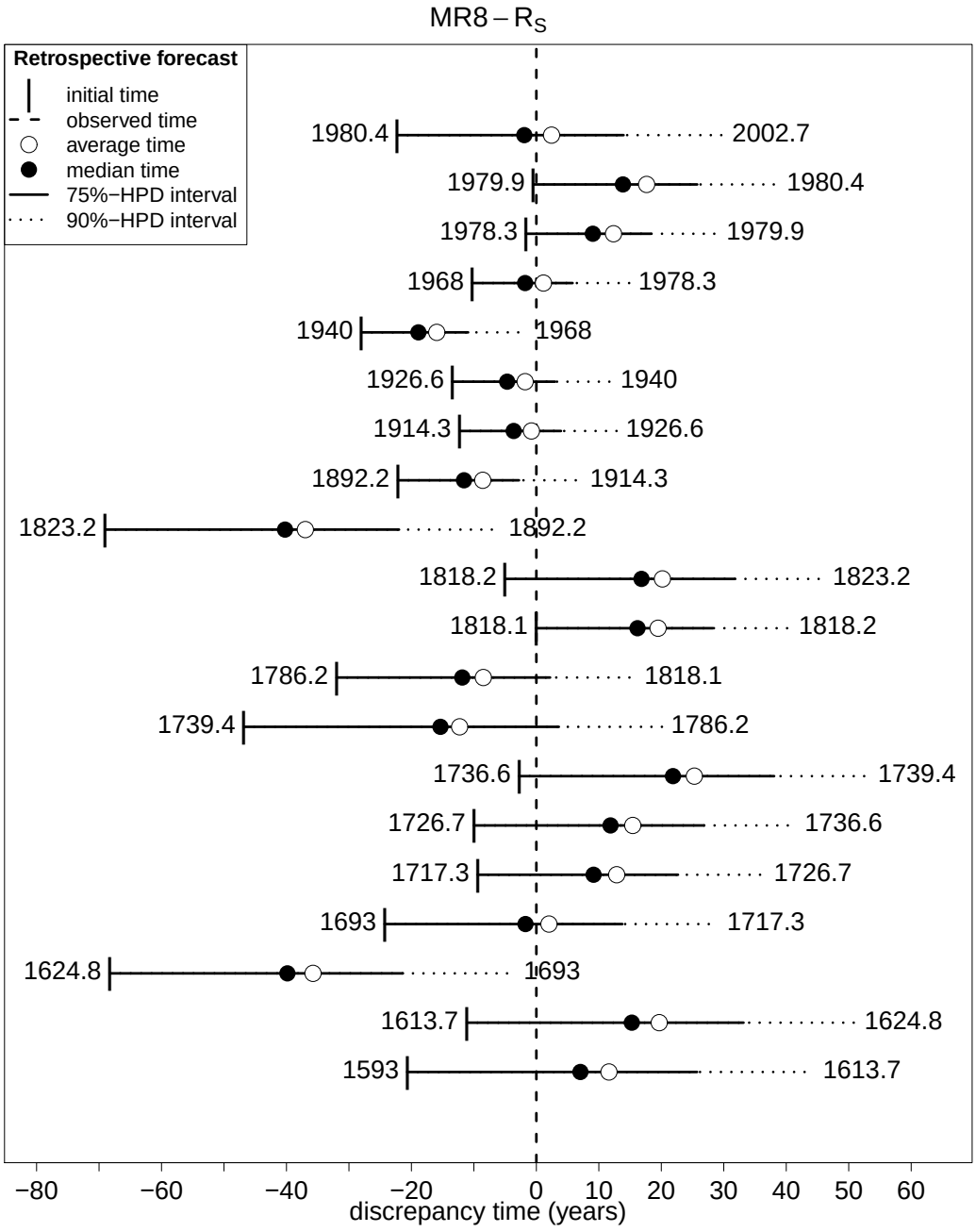

Figure C.7: As for Figure 6, validation results related to macroregion MR_8 - $\mathbf{R_S}$ model.

Table C.1: Prospective forecast after the ending date of the learning catalog. Summary of the estimated probability distribution of the time to the next event in each MR provided by all of the models: the 75% and 90% HPD intervals, median, mean and standard deviation.

region	model	HPD 75%	HPD 90%	median	mean (st.dev.)
MR_1	R_B	2003.0-2061.7	2003.0-2108.1	2031.1	2048.8 (59.2)
	R_M	2003.0-2127.0	2003.0-2187.2	2073.7	2091.0 (74.9)
	R_E	2003.0-2139.2	2003.0-2213.3	2077.3	2100.5 (90.2)
	R_S	2003.0-2048.9	2003.0-2084.5	2025.0	2038.8 (45.1)
MR_2	R_B	2003.0-2027.1	2003.0-2042.4	2015.2	2020.4 (17.5)
	R_M	2003.0-2028.4	2003.0-2046.9	2015.3	2021.9 (21.0)
	R_E	2003.0-2033.4	2003.0-2054.6	2017.9	2025.3 (23.8)
	R_S	2003.0-2023.5	2003.0-2034.1	2014.2	2017.4 (12.5)
MR_3	R_B	2003.0-2010.0	2003.0-2014.8	2006.4	2008.1 (5.3)
	R_M	2003.0-2010.2	2003.0-2015.4	2006.5	2008.3 (5.7)
	R_E	2003.0-2009.1	2003.0-2013.3	2006.0	2007.4 (4.7)
	R_S	2003.0-2010.0	2003.0-2014.5	2006.6	2008.0 (4.9)
MR_4	R_B	2003.0-2015.3	2003.0-2023.7	2009.1	2011.9 (9.2)
	R_M	2003.0-2015.1	2003.0-2023.6	2008.9	2011.9 (9.5)
	R_E	2003.0-2015.7	2003.0-2024.7	2009.2	2012.4 (10.0)
	R_S	2003.0-2015.9	2003.0-2023.2	2009.8	2012.0 (8.1)
MR_5	R_B	2003.0-2019.7	2003.0-2031.5	2011.2	2015.4 (13.5)
	R_M	2003.0-2020.5	2003.0-2033.7	2011.4	2016.2 (15.0)
	R_E	2003.0-2022.9	2003.0-2038.1	2012.5	2018.0 (17.4)
	R_S	2003.0-2016.8	2003.0-2023.9	2010.6	2012.9 (9.0)
MR_6	R_B	2003.0-2032.0	2003.0-2050.8	2017.7	2024.2 (22.1)
	R_M	2003.0-2034.5	2003.0-2053.7	2019.3	2025.6 (21.9)
	R_E	2003.0-2040.6	2003.0-2060.0	2023.7	2029.6 (23.7)
	R_S	2003.0-2028.0	2003.0-2041.4	2016.7	2020.9 (16.3)
MR_7	R_B	2003.0-2013.7	2003.0-2021.4	2008.2	2010.9 (8.5)
	R_M	2003.0-2014.0	2003.0-2021.7	2008.4	2011.0 (8.5)
	R_E	2003.0-2015.9	2003.0-2025.1	2009.3	2012.5 (10.2)
	R_S	2003.0-2010.9	2003.0-2016.5	2006.9	2008.8 (6.3)
MR_8	R_B	2003.0-2021.7	2003.0-2035.1	2012.2	2016.9 (15.1)
	R_M	2003.0-2017.9	2003.0-2029.0	2010.2	2014.1 (12.6)
	R_E	2003.0-2020.9	2003.0-2034.4	2011.5	2016.4 (15.1)
	R_S	2003.0-2038.8	2003.0-2055.2	2022.8	2027.4 (19.6)

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". The authors thank Fracassi and Valensise for providing the earthquake association with fault sources.

References

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

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

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

Basili, R., V. Kastelic, G. Valensise, and 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., and D.S. Harte (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. Seismol. Soc. Am.*, 41, 31-62.

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

Choy, G.L. and J.L. Boatwright (1995), Global patterns of radiated seismic energy and apparent stress, *J. Geophys. Res.*, **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., S. Richardson, and D.J. Spiegelhalter eds. (1996), Markov chain Monte Carlo in Practice. Chapman & Hall, London.

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

Herrmann, R., L. Malagnini, and I. Munafo (2011), Regional moment tensors of the 2009 L'Aquila earthquake sequence, *Bull. Seismol. 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., and M. Westcott (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 February 11, 2013

Jiang, M., S. Zhou, Y.J. Chen and Y. Ai (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.*, **82**, 2981-2987.

Kanamori, H., J. Mori, E. Hauksson, T.H. Heaton, L.K. Hutton, and M.L. Jones (1993), Determination of earthquake energy release and using M_L terrascope, *Bull. Seismol. Soc. Am.*, **81**, 330-346.

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

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

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

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

Maercklin, N., A. Zollo, A. Orefice, G. Festa, A. Emolo, R. De Matteis, B. Delouis, and A. Bobbio (2011), The effectiveness of a distant accelerometer array to compute seismic source parameters: The April 2009 L'Aquila earthquake case history, *Bull. Seismol. 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., F. Galadini, G. Valensise, M. Stucchi, R. Basili, S. Barba, G. Vannucci, and E. Boschi (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), 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.* 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. M. Wyss, K. Shimazaki and A. Ito) (Birkhäuser, Basel), *Pure Applied Geophys.*, 155, 471–507.

Pondrelli, S., S. Salimbeni, A. Morelli, G. Ekström, M. Olivieri and E. Boschi, 2010, Seismic moment tensors of the April 2009, L'Aquila (Central Italy) earthquake sequence, *Geophys. J. Int.*, 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 Statisticsal computing, Vienna, Austria, ISBN 3-900051-07-0, URL.http://www.R-prokect.org/

Reid, H.F. (1910), *T*he 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., and E. Garavaglia (2002), Statistical analysis of the completeness of a seismic catalogue, *Natural Hazards*, 25, 3, 245-258.

Rotondi, R., and E. Varini (2007), Bayesian inference of stress release models applied to some Italian seismogenic zones, *Geophysical Journal International*, **169**, 1, 301-314.

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

Scognamiglio, L., E. Tinti, A. Michelini, D. S. Dreger, A. Cirella, M. Cocco, S. Mazza, and A. Piatanesi (2010), Fast determination of moment tensors and rupture history: What has been learned from the 6 April 2009 L'Aquila earthquake sequence, *Seismol. Res. Lett.*, 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, *Physics of the Earth and Planetary Interiors*, 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 Statstical Software*, **21**, 11, 37 pp., available on http://www.jstatsoft.org/

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

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

Vere-Jones, D., and D. Yonglu (1988), A point process analysis of historical earthquakes from North China, *E*arthquake Research in China, **2**, 2, 165-181.

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

Wang, A., D. Vere-Jones, and X. Zheng (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.

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

Zheng, X., and D. Vere-Jones (1991), Application of stress release models to historical earthquakes from North China, *P*ageoph., **135**, 4, 559-576.

Zheng, X., and D. Vere-Jones (1994), Further applications of the stochastic stress release model to historical data, *T*ectonophysics, 229, 101-121.