entropy

MDPI

Article

The Effect of Catalogue Lead Time on Medium-Term Earthquake Forecasting with Application to New Zealand Data

David A. Rhoades, Sepideh J. Rastin * and Annemarie Christophersen

GNS Science, 1 Fairway Drive, Avalon, P.O. Box 30–368, Lower Hutt 5040, New Zealand; d.rhoades@gns.cri.nz (D.A.R.); a.christophersen@gns.cri.nz (A.C.)
* Correspondence: s.rastin@gns.cri.nz; Tel.: +64-21-2810336

Received: 30 September 2020; Accepted: 4 November 2020; Published: 6 November 2020

Abstract: 'Every Earthquake a Precursor According to Scale' (EEPAS) is a catalogue-based model to forecast earthquakes within the coming months, years and decades, depending on magnitude. EEPAS has been shown to perform well in seismically active regions like New Zealand (NZ). It is based on the observation that seismicity increases prior to major earthquakes. This increase follows predictive scaling relations. For larger target earthquakes, the precursor time is longer and precursory seismicity may have occurred prior to the start of the catalogue. Here, we derive a formula for the completeness of precursory earthquake contributions to a target earthquake as a function of its magnitude and lead time, where the lead time is the length of time from the start of the catalogue to its time of occurrence. We develop two new versions of EEPAS and apply them to NZ data. The Fixed Lead time EEPAS (FLEEPAS) model is used to examine the effect of the lead time on forecasting, and the Fixed Lead time Compensated EEPAS (FLCEEPAS) model compensates for incompleteness of precursory earthquake contributions. FLEEPAS reveals a space-time trade-off of precursory seismicity that requires further investigation. Both models improve forecasting performance at short lead times, although the improvement is achieved in different ways.

Keywords: earthquake precursors; seismicity patterns; earthquake forecasting; statistical seismology

1. Introduction

Precursory seismicity to major earthquakes takes place over time scales ranging from less than a day to decades [1]. Precursory earthquakes are part of the general phenomenon of space-time earthquake clustering. Short-term clustering is the most obvious and easily identifiable type of earthquake clustering. It includes the well-known phenomena of "foreshocks" and "aftershocks". Models based on short-term clustering are mainly concerned with forecasting aftershocks [2–5]. While only a small proportion of major earthquakes have short-term foreshocks [6,7], most have longer term precursory seismicity. Systematic attempts to identify longer term precursory seismicity patterns, either activation or quiescence, have formed the basis of many proposed earthquake forecasting methods [8–13].

The 'Every Earthquake a Precursor According to Scale' (EEPAS) model, which is the focus of this paper, is based on an observed increase in the magnitude and rate of occurrence of minor earthquakes prior to a major earthquake and is known as the precursory scale increase (Ψ-) phenomenon [14].

The Ψ-phenomenon was proposed as a generalisation of the precursory swarm phenomenon [15,16]. It can be identified before most major earthquakes (mainshocks) in well-catalogued regions on time scales ranging from months to decades, depending on magnitude, within a region similar to that occupied by the consequent aftershocks [17]. The Ψ-phenomenon is

quantified by the cumulative magnitude anomaly (cumag) calculated for earthquakes with magnitudes greater than or equal to a chosen threshold magnitude, M_{thres}, in a region of interest over a time-period ending just before the occurrence of the major earthquake [17].

A recent example of the Ψ-phenomenon from the 2019 Ridgecrest, California, earthquake sequence [18] is shown in Figure 1. Figure 1a shows the area of the precursory earthquakes, the mainshocks and aftershocks, in different symbols. Figure 1b is a stem plot of earthquake magnitudes in time from the mid-1960s to late 2019. Figure 1c shows the cumag, obtained using

$$C(t) = \sum_{t_s \le t_i < t} [M_i - (M_{thres} - 0.1)] - k(t - t_s) \tag{1}$$

with

$$k = \sum_{t_s \le t_i < t_f} [M_i - (M_{thres} - 0.1)] / \left(t_f - t_s\right) \tag{2}$$

where t_s is the starting time and t_f is the finishing time [17].

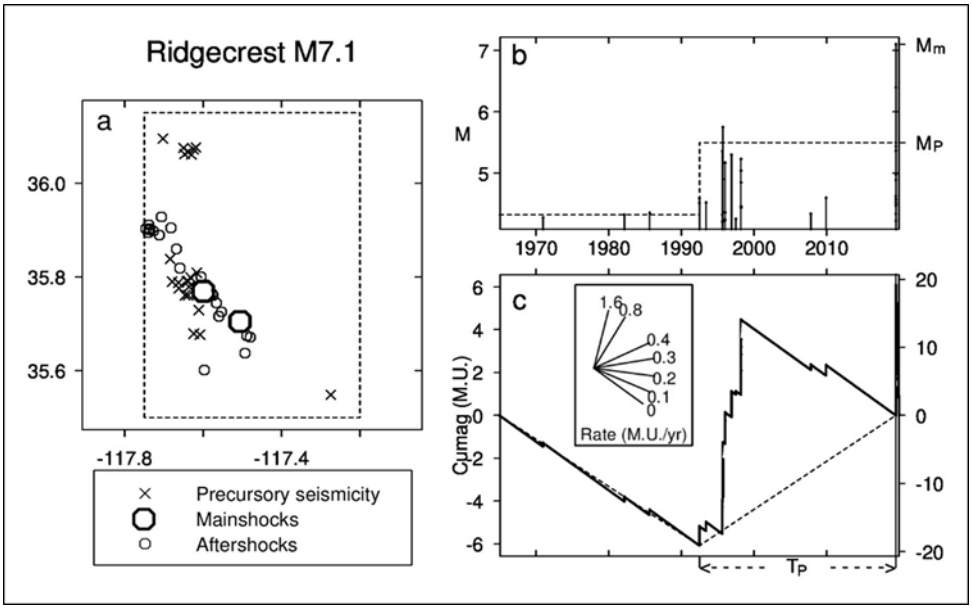

Figure 1. The Ψ-phenomenon for the 2019 Ridgecrest, California, earthquake, July 2019, magnitudes M 6.4 and 7.1. (**a**) Epicentres of the precursory seismicity, mainshocks and aftershocks. The rectangle containing them is the precursory area, A_P. (**b**) Magnitude versus time of prior and precursory earthquakes with the onset of Ψ in 1992. Dashed lines show precursory increase in magnitude level. M_m is main shock magnitude, M_P is precursor magnitude—the average magnitude of the three largest precursory earthquakes. (**c**) Changes in cumulative magnitude anomaly (cumag) with time. Dashed lines show precursory increase in seismicity rate. The protractor translates the cumag slope into seismicity rate in magnitude units per year (M.U. year^{-1}).

Each earthquake, i, with magnitude, M_i, larger than or equal to M_{thres}, leads to an upward jump in the $C(t)$ curve. M_{thres} was set to 4.3 in this example. $C(t)$ decreases during times of lower-than-average earthquake occurrence and increases during times of above-average earthquake occurrence. Changes in the slope of the cumag plot represent changes in the rate of seismicity. A large increase in seismicity leads to a sharp minimum of $C(t)$. The minimum of $C(t)$ is taken to mark the onset of Ψ [18].

Using many examples of the Ψ-phenomenon similar to Figure 1, predictive scaling relations were derived [17], as follows:

$$M_m = a_M + b_M M_P; \ \log T_P = a_T + b_T M_P; \ \ \log A_P = a_A + b_A M_P \tag{3}$$

These show that the precursor magnitude, M_P, is a predictor of the mainshock magnitude, M_m, the precursor time, T_P—the time between the onset of Ψ and the mainshock—and the area, A_P, within which the precursors, mainshock and aftershocks all occur. Figure 2 displays the Ψ-predictive scaling relations for the original 47 major earthquakes [14,17] and the recent Ridgecrest sequence (large symbols). Selected earthquakes represent the largest magnitudes without any regard to tectonic settings.

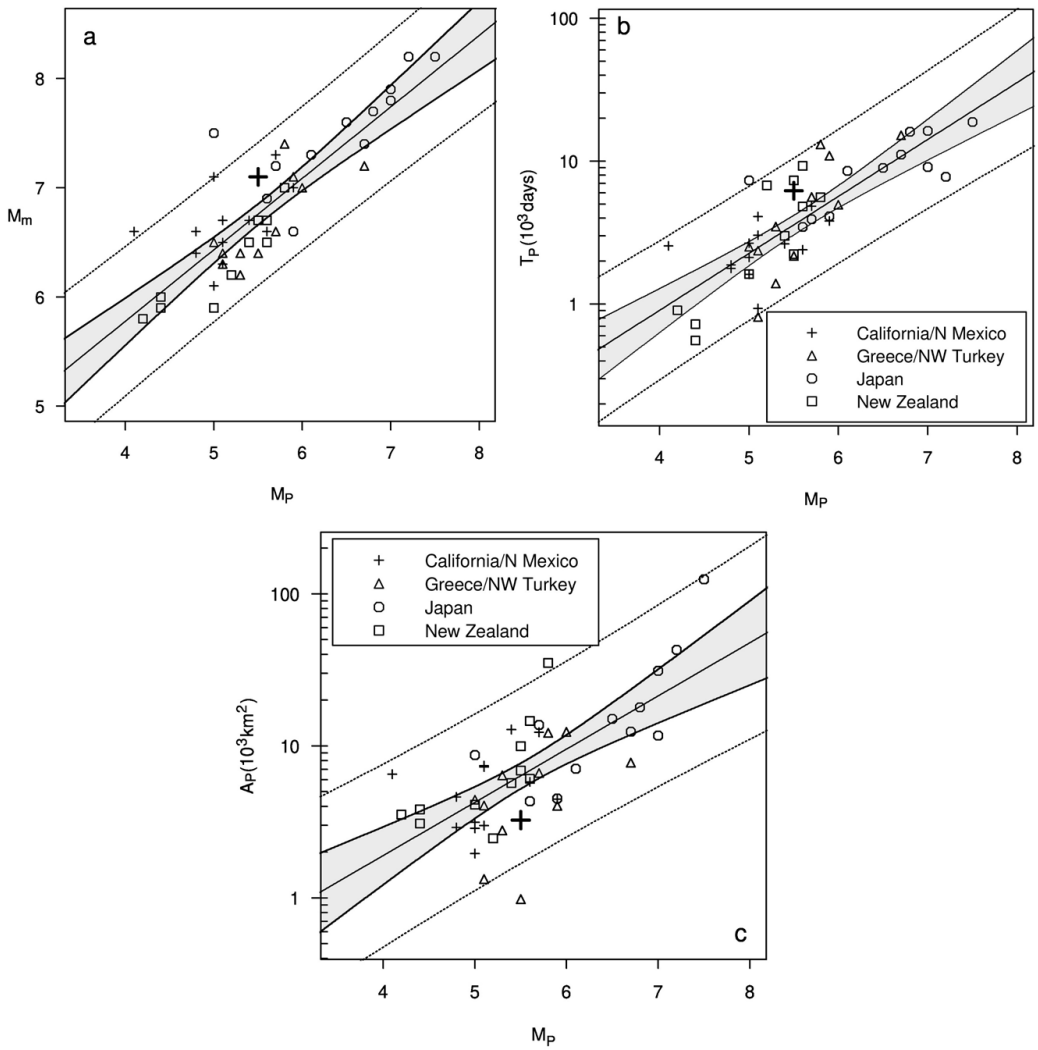

Figure 2. Ψ-predictive scaling relations between (**a**) mainshock and precursor magnitudes, M_m and M_P (Coefficient of determination R^2 = 71%), (**b**) precursor time, T_P and M_P (R^2 = 65%) and (**c**) precursor area, A_P and M_P (R^2 = 48%), for 47 major earthquakes from References [14,17] and the recent Ridgecrest sequence (large symbols). Dotted lines indicate 95 percent tolerance limits. The shaded regions are 95 percent confidence bands for the fitted relations.

The EEPAS model applies the Ψ-predictive scaling relations to every earthquake and was first developed for NZ and California [19]. In the model, each earthquake initiates a transient increase in the expected rate of occurrence of larger earthquakes to follow. EEPAS has been applied to Japan [20,21], Southern California [22] and Greece [23], and it performs well for these seismically active regions. It has been formally evaluated in several regional testing centres of the Collaboratory for the Study of Earthquake Predictability (CSEP) [24]. CSEP is an international effort to conduct and rigorously evaluate earthquake forecasting experiments [24,25].

The EEPAS model has provided the medium-term component of hybrid operational earthquake forecasts [26] and time-varying probabilistic seismic hazard models in NZ [27] for nearly a decade.

In this paper, we investigate the effect of catalogue lead time on the EEPAS model. The lead time for a given target earthquake is defined as the time between the start of the catalogue and the time of occurrence of the earthquake. For larger target earthquakes, the precursor time is longer and precursory seismicity may have occurred prior to the start of the catalogue and thus may not be completely recorded in the available data. Previously, EEPAS has been applied with a fixed catalogue starting time. The catalogue starting time is usually chosen as the earliest time at which the catalogue is considered sufficiently complete above a certain magnitude threshold—the minimum magnitude for contributing earthquakes. The model is then fitted to optimize its performance in forecasting target earthquakes above a certain larger magnitude threshold. Any precursory earthquakes that might have occurred before the catalogue starting time cannot contribute to the fitting of the model or to forecasts based on it. Using the same catalogue starting time for all target earthquakes gives different target earthquakes different lead times. This affects the fitting of the EEPAS model and its forecasting performance. To understand this effect, here, we derive a formula for the completeness of precursory earthquake contributions to a target earthquake as a function of its magnitude and lead time. We then develop a version of EEPAS that applies a fixed lead time to all target earthquakes. This means that different target earthquakes will have different catalogue starting times. We fit the fixed lead time EEPAS model to the NZ earthquake catalogue within a region of surveillance and over a time period to which EEPAS was previously fitted. Finally, we develop a further version of the EEPAS model that compensates for the lead time and thus can use the full available catalogue for fitting.

We review the EEPAS model formulation and introduce the NZ data in Section 2. We then derive the completeness of precursory earthquake contributions as a function of magnitude and lead time and the fixed lead time version of EEPAS is described. A method to compensate for incompleteness is described in Section 2.4. Results of applying the new methods to NZ data are given in Section 3. The discussion and conclusions make up Sections 4 and 5.

2. Method and Data

2.1. EEPAS Model Rate Density

In the EEPAS model, the rate density, $\lambda(t,m,x,y)$, of earthquake occurrence within a chosen depth range is defined for any time, t, magnitude, m, and location (x,y), where m exceeds a target threshold magnitude, m_c, and (x,y) is a point in a region of surveillance, R. Each earthquake (t_i,m_i,x_i,y_i), with t_i greater than a starting time, t_0, and m_i greater than a minimum magnitude, m_0, contributes a transient increment $\lambda_i(t,m,x,y)$ to the future rate density in its vicinity, given by

$$\lambda_i(t,m,x,y) = w_i f(t|t_i,m_i) g(m|m_i) h(x,y|x_i,y_i,m_i), \tag{4}$$

where w_i is a weighting factor to emphasise earthquakes that are most likely to be precursors, and f, g and h are densities of the probability distributions derived from the Ψ-predictive scaling relations (Equation (3)) for time, magnitude and location, respectively [19]. Following the notation of Reference [28], the magnitude density, g, is a normal density of the form

$$g(m|m_i) = \frac{1}{\sigma_M \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{m - a_M - b_M m_i}{\sigma_M}\right)^2\right], \tag{5}$$

where a_M, b_M and σ_M are parameters, with a_M and b_M based on the corresponding regression parameters in (3) and σ_M on the scatter of points around the regression line in Figure 2a. The time density, f, is a lognormal density of the form

$$f(t|t_i,m_i) = \frac{H(t-t_i)}{(t-t_i)\sigma_T \ln(10)\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log(t-t_i) - a_T - b_T m_i}{\sigma_T}\right)^2\right], \tag{6}$$

where $H(s) = 1$ if $s > 0$ and 0 otherwise, and a_T, b_T and σ_T are parameters, with a_T and b_T based on the corresponding regression parameters in (3) and σ_T on the scatter of points around the regression line in Figure 2b. The location density, h, is a bivariate normal density of the form

$$h(x, y|x_i, y_i, m_i) = \frac{1}{2\pi\sigma_A^2 10^{b_A m_i}} \exp\left[-\frac{(x - x_i)^2 + (y - y_i)^2}{2\sigma_A^2 10^{b_A m_i}}\right], \tag{7}$$

where σ_A and b_A are parameters, with b_A based on the corresponding regression parameter in (3) and $\log {\sigma_M}^2$ related to the regression parameter a_T in (3).

The rate density of the EEPAS model is obtained by summing over all past occurrences, including earthquakes outside R, that could affect the rate density within R:

$$\lambda(t, m, x, y) = \mu\lambda_0(t, m, x, y) + \sum\nolimits_{t_i \geq t_0, m_i \geq m_0} \eta(m_i)\lambda_i(t, m, x, y), \tag{8}$$

where μ is a mixing parameter representing the proportion of the forecast contributed by the background model component, λ_0 is the rate density of a background Poisson model with a location distribution based on proximity to the epicentres of past earthquakes (PPE) [19,21], t_0 is the starting time of the earthquake catalogue and η is a normalising function. For a given magnitude v, η is defined by

$$\eta(v) = \frac{b_M(1 - \mu)}{E(w)} \exp\left[-\beta\left(a_M + (b_M - 1)v + \frac{\sigma_M^2\beta}{2}\right)\right], \tag{9}$$

where $E(w)$ is the mean weight of earthquakes in the catalogue, and $\beta = b_{GR}$ ln10, with b_{GR} being the Gutenberg–Richter b-value [29]. Normalising over the whole fitting period and region of surveillance ensures that the number of earthquakes expected by the model approximately matches the actual number of target earthquakes. It also forces the forecasted magnitudes to approximately follow the Gutenberg–Richter relation, although the magnitude distribution can vary locally from this relation. The background Poisson model also conforms to the Gutenberg–Richter relation, both locally and in the whole region, R. The parameters of the EEPAS model are fitted to a chosen catalogue using the maximum likelihood method [19].

2.2. Data

To illustrate the fitting of the models, we use the NZ earthquake catalogue. The catalogue starting time is set to be 1951, based on an assessment of the quality and completeness of the NZ catalogue. A minimum magnitude $m_0 = 2.95$ was set for precursors. In order to fit the EEPAS model properly, a difference of about two units is needed between the minimum target earthquake magnitude and m_0. Therefore, the target earthquakes in the magnitude range between $m_c = 4.95$ and $m_u = 8.05$ were used with a fitting period of 1987–2006. This gives a minimum lead time of 36 years for target earthquakes. The region of surveillance is the NZ Earthquake Forecast Testing Centre test region [30], as shown in Figure 3. The depth range of 0–40 km was used. The selection of data is consistent with previous model fittings [30,31]. Although the catalogue is inhomogeneous and not complete above m_0 from 1951, it is considered adequate for EEPAS model fitting with these parameters [31].

Two versions of the EEPAS model denoted by EEPAS-0F and EEPAS-1F were previously fitted to this region and time period [30,31]. These two versions differ only in the weighting strategies applied to precursory earthquakes. The EEPAS-0F model weights all earthquakes equally and the EEPAS-1F model down-weights aftershocks, i.e., earthquakes are weighted by the probability that they represent independent events, rather than aftershocks, in an epidemic-type aftershock model [3,19]. For illustrating the effect of applying fixed lead times, we focus our attention on the EEPAS-0F model. EEPAS-0F was tested in the CSEP NZ Earthquake Forecast Testing Centre [24] from 2008 to 2017 while being updated every 3 months [26]. It is also routinely used in hybrid model earthquake forecasts published on the GeoNet website for the Canterbury, central NZ and Kaikoura

regions [32]. Here, we have refitted three parameters of this model to accommodate minor updating of the input catalogue since it was originally fitted. These parameters are the time-scaling parameter a_T, the spatial scaling parameter σ_A and the mixing parameter μ. We have also refitted the Proximity to Past Earthquakes (PPE) background model. The refitted parameters and those retained from previous fitting [31] are given in Table 1.

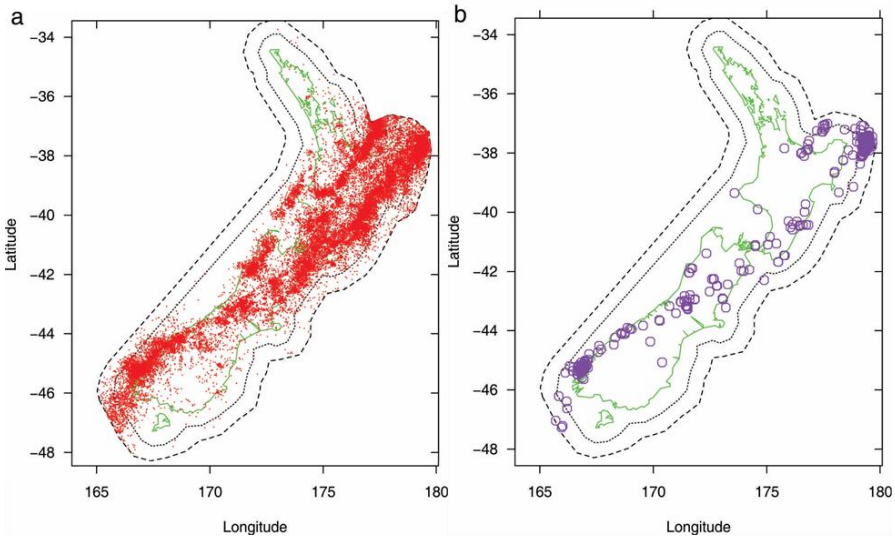

Figure 3. Maps of New Zealand (NZ) seismicity including test region (dotted inner polygon) and data collection region (dashed outer polygon), and earthquakes of magnitude (**a**) M > 2.95 from 1951 to 2006 with hypocentral depth $h \leq 45$ km (the precursor set of 44,998 earthquakes), and (**b**) M > 4.95 from 1987 to 2006, with $h \leq 40$ km (the target set, including 158 earthquakes in the test region).

Table 1. Every Earthquake a Precursory According to Scale (EEPAS) model Parameters.

Parameter	Details	EEPAS-0F
m_0	Minimum precursor magnitude	2.95 *
m_c	Minimum target magnitude	4.95 *
m_u	Maximum target magnitude	8.05 *
b_{GR}	Gutenberg–Richter b-value	1.16 †
a_M	Equation (5)	1.10 †
b_M	Equation (5)	1.0 †
σ_M	Equation (5)	0.39 †
a_T	Equation (6)	1.73
b_T	Equation (6)	0.39 †
σ_T	Equation (6)	0.60 †
b_A	Equation (7)	0.36 †
σ_A	Equation (7)	1.53
μ	Equation (8)	3.6×10^{-5}

* fixed; † fitted previously [31].

2.3. Completeness of Precursory Earthquake Contributions

An expression for the completeness of precursory earthquake contributions in the EEPAS model at a given time-lag, T, and a target magnitude, m, was presented in Reference [28]. The time-lag, T, is the interval between the end of the earthquake catalogue and the target time for which the forecast is being made. The main assumptions were to have precursory earthquakes equally likely to occur at any time and to have a magnitude distribution that follows the Gutenberg–Richter relation [29], with b-value b_{GR}. Using Bayes' rule, the expected rate of occurrence of precursors at time $t - \tau$ and

magnitude v with respect to target earthquakes at (t, m) is proportional to $f(t|t-\tau, v)g(m|v)10^{-b_{GR}v}$. The total contribution of precursory earthquakes expected to be present in a catalogue with a time-lag, T, to the EEPAS rate density at (t, m) is proportional to $c(T, L, m)$, given by

$$c(T, L, m) = \int_{m_0}^{m_u} \left[\int_{T}^{T+L} f(t|t-\tau, v) d\tau \right] g(m|v) 10^{-b_{GR}v} dv, \tag{10}$$

where L is the catalogue lead time and m_0 and m_u are the minimum and maximum magnitude thresholds for earthquakes in the catalogue. Note that the time density $f(t|t-\tau, v)$ depends on t only through the time interval $t-\tau$ (see Equation (6)). Therefore, the right side of Equation (10) is independent of t.

It was assumed in Reference [28] that L (at around 7 decades in the examples considered) was long enough so that the contributions to the EEPAS rate density at (t,m) from unrecorded precursory earthquakes that might have occurred before the start of the catalogue can be deemed negligible. Here, we make the reverse assumption: that the time-lag, T, is negligibly small and that the lead time, L, from the starting time, t_0, up to time, t, is short enough so that the contributions from unrecorded precursory earthquakes that might have occurred before t_0 are appreciable. By replacing T with 0 in (10), because f is log-normal, we have

$$c(0, L, m) = \int_{m_0}^{m_u} \Phi\left(\frac{\log L - a_T - b_T v}{\sigma_T} \right) g(m|v) 10^{-b_{GR}v} dv, \tag{11}$$

where Φ is the standard normal probability integral.

If the lead time is not constrained, the total contribution of precursory earthquakes expected to be present in a catalogue is $c(0, \infty, m)$, given by

$$c(0, \infty, m) = \int_{m_0}^{m_u} \eta(v) g(m|v) 10^{-b_{GR}v} dv. \tag{12}$$

Therefore, for a given lead time, L, the completeness of precursory earthquake contribution is estimated by $p(L, m)$, given by

$$p(L, m) = \frac{c(0, L, m)}{c(0, \infty, m)}. \tag{13}$$

The completeness is an increasing function of L and a decreasing function of m. Implicitly, it also depends on b_{GR} and the parameters of the EEPAS distributions for time and magnitude, namely a_M, b_M, σ_M, a_T, b_T and σ_T. Figure 4 displays contours of theoretical completeness for magnitudes of 5 to 8 and lead times, L, from 0 to 35 years. This was calculated using (13) for the EEPAS-0F model fitted to NZ earthquakes from 1987 to 2006.

In normal fitting of EEPAS, the lead time increases with time, t. Thus, the completeness of precursory earthquake contributions increases as we proceed through the target earthquakes in the catalogue. It is desirable to compensate for this variable completeness. To do so, we initially set the lead time of precursory earthquake contributions to a fixed value. Then, for a given target time, only earthquakes occurring within the lead time contribute to the time-varying component of the EEPAS rate density. This means using a moving data window, with a variable starting time for earthquake contributions depending on the target time. However, no lead time restriction is applied to the background model. This fixed lead time version of EEPAS is referred to hereinafter as FLEEPAS. FLEEPAS does not compensate for the incompleteness due to short lead times. However, partial compensation is achieved when FLEEPAS parameters are fitted to the data (see Section 3).

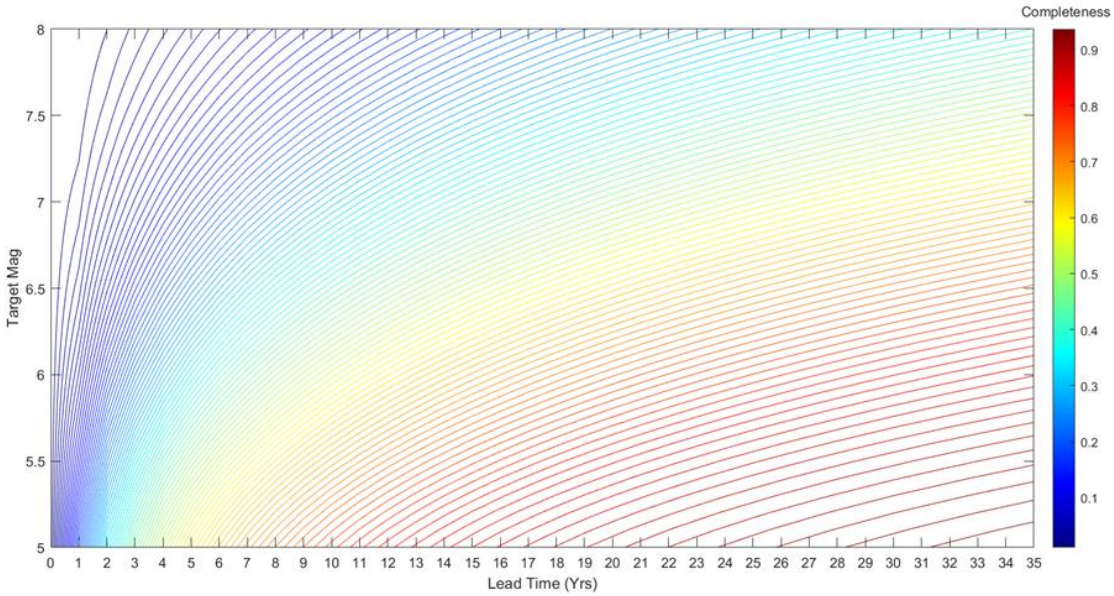

Figure 4. Contour plot of theoretical completeness for target magnitudes of 5 to 8 and lead times of 0 to 35 years for the EEPAS model fitted to the NZ catalogue (EEPAS-0F) in the 1987–2006 interval without aftershock down-weighting.

2.4. Compensating EEPAS for Incompleteness

2.4.1. Theory

Incompleteness of the precursory earthquake contributions caused by short lead times causes the EEPAS model to underestimate the expected number of target earthquakes. The shorter the lead time, the more severe the underestimation is expected to be. With a lead time of L, the contribution of the time-varying component of EEPAS to the earthquake occurrence rate density $\lambda(t, m, x, y)$ will be diminished at magnitude, m, by a factor of $p(L,\ m)$, on average.

Suppose that in the earthquake catalogue, earthquakes occur within the region of surveillance at a certain average rate per unit time denoted by γ and a magnitude distribution with a density of $\theta(m)$. Then, let $r(m) = \gamma\theta(m)$ denote the average rate density of occurrence of earthquakes with magnitude, m, in the region. If the EEPAS model is applied with fixed lead time, L, it is expected to forecast an average rate density of $r(L,m) < r(m)$ at magnitude m, where

$$r(L, m) = [\mu + (1 - \mu)p(L, m)]r(m), \tag{14}$$

where μ is the EEPAS mixing parameter. To compensate for the incompleteness, we can consider two end-members. In end-member A, we augment the background component, and in end-member B, the time-varying component. End-member A has a rate density of $\lambda_A(t, m, x, y)$, given by

$$\lambda_A(t, m, x, y) = [\mu + (1 - \mu)(1 - p(L, m))]\lambda_0(t, m, x, y) + \sum_{t_i \geq t_0, m_i \geq m_0} \eta(m_i)\lambda_i(t, m, x, y). \tag{15}$$

End-member B has a rate density of $\lambda_B(t, m, x, y)$, given by

$$\lambda_B(t, m, x, y) = \mu\lambda_0(t, m, x, y) + \frac{1}{p(L, m)} \sum_{t_i \geq t_0, m_i \geq m_0} \eta(m_i)\lambda_i(t, m, x, y). \tag{16}$$

From (13), it follows that the average earthquake rate density within R is equal to $r(m)$ for both end-members. This verifies that the end-members are correctly compensated for incompleteness.

For a given lead time, L, we consider convex linear combinations of these two end-members, with rate density $\lambda_C(t, m, x, y)$, given by

$$\lambda_C(t, m, x, y) = \varphi\lambda_A(t, m, x, y) + (1 - \varphi)\lambda_B(t, m, x, y), \tag{17}$$

where $0 \leq \varphi \leq 1$. All such linear combinations would retain the same expected average rate density, *r(m)*. The parameter φ can be fitted to the past earthquake catalogue using the maximum likelihood method, in which the log-likelihood is the objective function optimized. The log-likelihood (L) for a model with rate density λ is given by

$$\ln L = \sum_{i=1}^{N} \lambda(t_i, m_i, x_i, y_i) - E_\lambda(N), \tag{18}$$

where $E_\lambda(N)$ is the expected number of target earthquakes obtained by integrating the rate density over the time range (t_1,t_2,), magnitude range (m_c,m_u) and region of surveillance, R.

The log-likelihood is also used to measure the information gain of one model over another. Here, we estimate the information gain of one model X over another model Y [33,34], by

$$I(X, Y) = \frac{\ln L_X - \ln L_Y}{N}. \tag{19}$$

2.4.2. Fixed Lead Time Compensated EEPAS Model

We have implemented the theory above to compensate the EEPAS model for the fixed lead times. This is called the Fixed Lead Time Compensated EEPAS (FLCEEPAS) Model. FLCEEPAS compensates for the missing precursory earthquakes by fitting the parameter φ to optimize the mixture between the two end-members when the completeness $p(L, m)$ of precursory earthquake contributions for a given lead-time and target earthquake magnitude is estimated by (10).

3. Results

The FLEEPAS model is initially applied with the previously fitted EEPAS parameters (unfitted FLEEPAS). It is then used to fit new values of σ_A, a_t and μ for a given lead time (fitted FLEEPAS). Other parameters are kept at their values in Table 1. Systematically reducing the lead time from 35 to 3 years, we fit the FLEEPAS parameters. We then compute the theoretical completeness using the fitted FLEEPAS parameters, as shown in the contour plot of Figure 5. Comparing Figure 5 with Figure 4, it is evident that, for a given lead time and magnitude, a higher level of completeness is achieved with the fitted FLEEPAS.

The higher level of completeness results in a higher information gain. Figure 6 displays the information gain as a function of the lead time for both fitted and unfitted FLEEPAS models. The information gain is computed using (19) and is relative to the Stationary Uniform Poisson (SUP) model. SUP is a model of minimal information. Its rate density depends only on the number of target earthquakes and the Gutenberg–Richter b-value. Figure 6 shows that the fitted FLEEPAS models outperform the unfitted ones for lead times shorter than 11 years. For the unfitted models, the information gain declines strongly at lead times below 11 years. For the fitted models, the decline is much less strong. At the shortest lead time of 3 years, the decline is only about 0.2 for the fitted FLEEPAS model compared with nearly 0.7 for the unfitted model.

The fitted time and spatial distributions are strongly affected by the lead time. When fitting the FLEEPAS model, the mean of the time distribution (6) is proportional to 10^{a_t} and the area occupied by the spatial distribution (7) is proportional to σ_A^2, provided other parameters of these distributions are fixed. Therefore, 10^{a_t} and σ_A^2 are considered as time and spatial scaling factors to compare the change in the FLEEPAS time and spatial distributions. Figure 7 shows the time and spatial scaling factors of fitted FLEEPAS versus lead time. As the lead time is reduced, the time factor is seen to decrease,

and the spatial factor to increase. Such a time-space trade-off was previously observed in multiple identifications of the Ψ-phenomenon for individual earthquakes in a synthetic catalogue generated by a physics-based earthquake simulator [35].

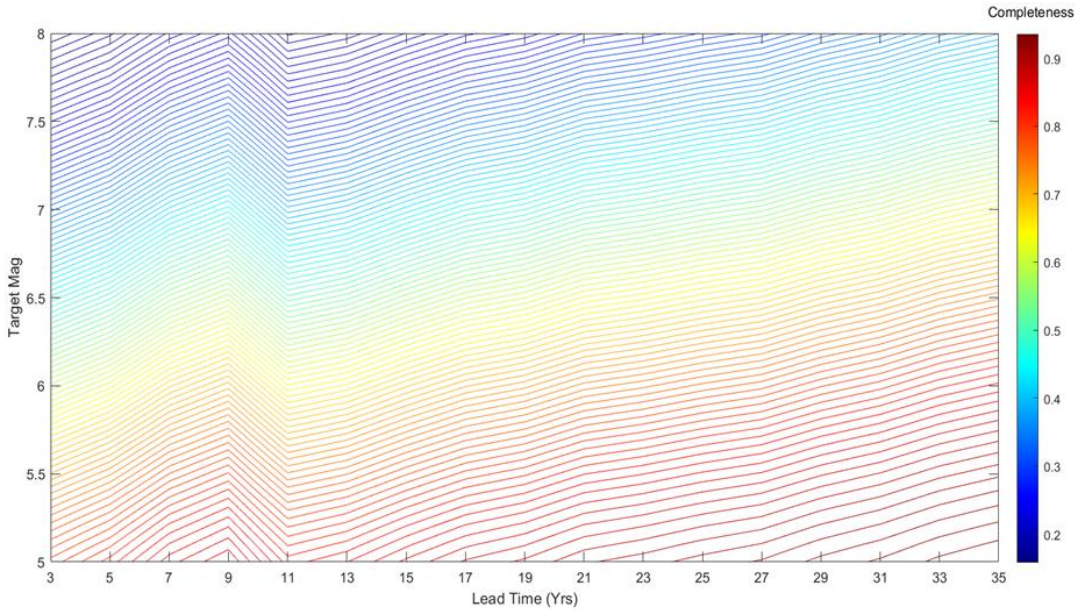

Figure 5. Contour plot of Fixed Lead time EEPAS (FLEEPAS) model completeness for magnitudes of 5 to 8 and lead times of 3 to 35 years fitted to the NZ data (see Section 2.2).

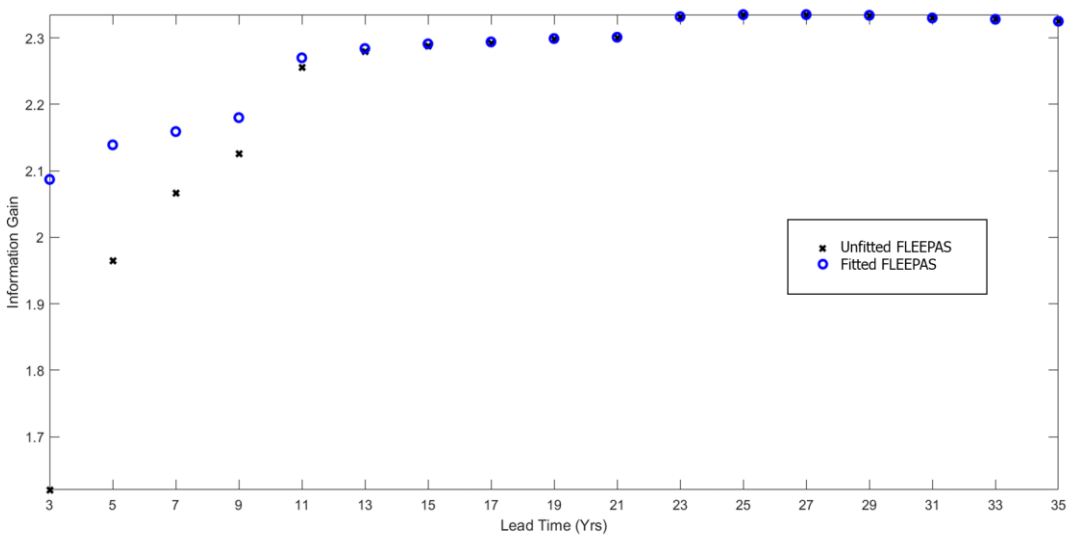

Figure 6. Information gain per earthquake compared to the stationary uniform Poisson model for the unfitted and fitted FLEEPAS models as a function of lead time.

The fitted mixing parameter is also affected by the lead time. Starting off from close to zero (Table 1), it initially remains close to zero but then increases strongly as the lead time is reduced below 11 years (Figure 8).

The FLCEEPAS model is used to fit φ for the same fixed lead times as above, with all other parameters kept at their values in Table 1. Figure 9 displays the optimal φ values versus lead time as well as information gain pre- and post-compensation. A value of φ close to 1 implies that most of the compensation is in the background model, and a value close to zero implies that most of

the compensation is in the time-varying component. As can be seen in Figure 9a, for short lead times, the background component has mainly contributed to the compensation, while for the lead times longer than 23 years, only the time-varying component contributes.

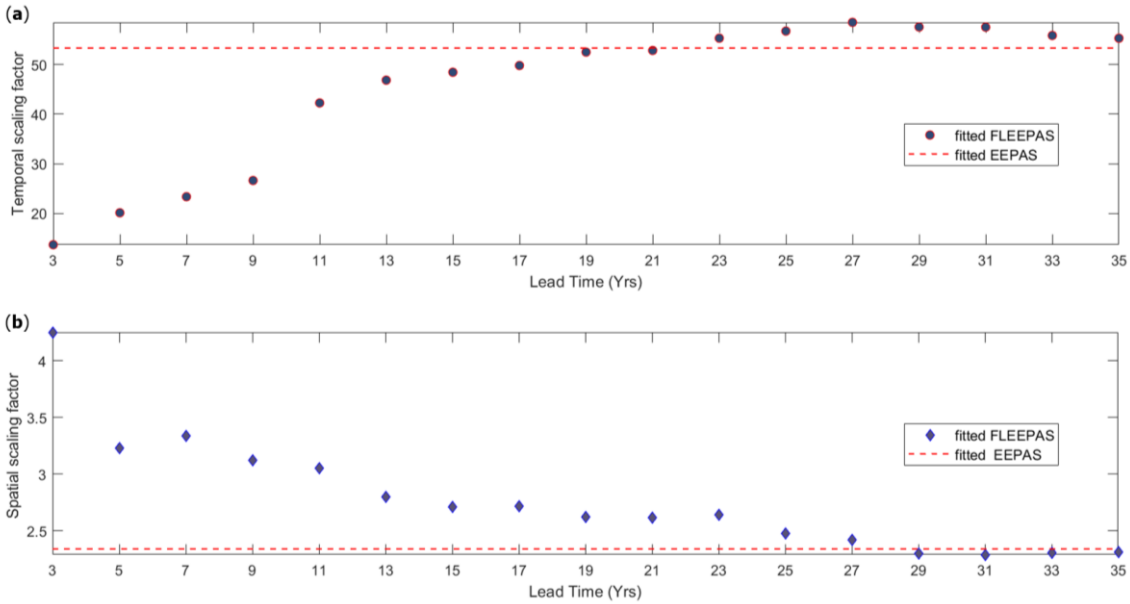

Figure 7. (**a**) Temporal scaling factor and (**b**) spatial scaling factor versus lead time in the FLEEPAS model fitted to NZ data. Dashed line shows corresponding factor for initial parameters from the fitted EEPAS-0F model (Table 1).

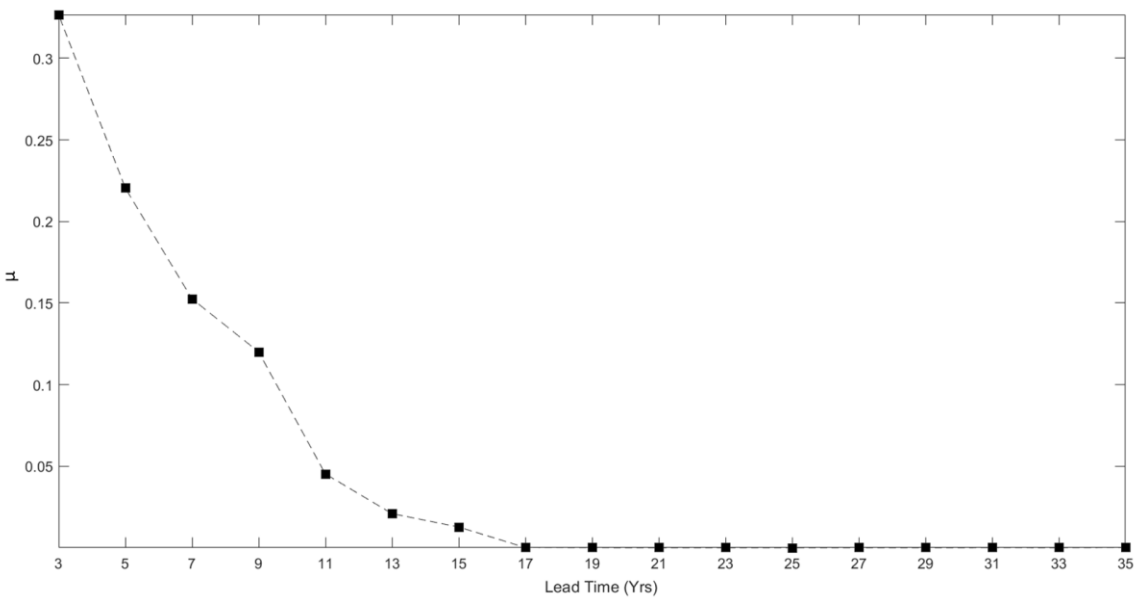

Figure 8. Fitted mixing parameter, μ, versus lead time in the FLEEPAS model fitted to NZ data as a modification of EEPAS-0F (Table 1).

As well as making a full compensation for the incompleteness, the FLCEEPAS model achieves similar information gains as the fitted FLEEPAS by only fitting a single parameter instead of three (Figures 6 and 9b). FLEEPAS and FLCEEPAS have different temporal and spatial distributions for short lead times due to the time-space trade-off.

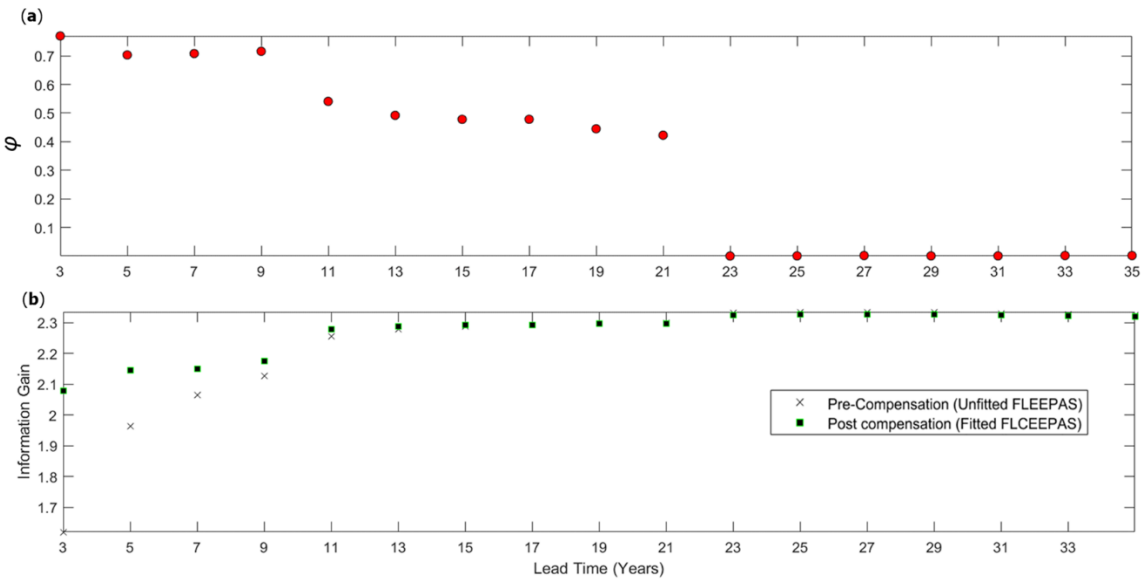

Figure 9. (**a**) Optimal φ versus lead time for the fitted FLCEEPAS model and (**b**) information gain per earthquake versus lead time for the unfitted FLEEPAS and fitted FLCEEPAS models. The information gain is relative to the Stationary Uniform Poisson model.

4. Discussion

The main objective motivating this work was to build understanding of precursory seismicity in order to improve forecasting of major earthquakes. The application of the FLEEPAS and FLCEEPAS models contributes to improved understanding of precursory seismicity, but the improved forecasting is yet to be fully realized.

A catalogue-based forecasting model like EEPAS can only use precursors within the available catalogue. In our application to NZ data, the available lead time starts from 36 years at the beginning of the fitting period. The precursor time for a large earthquake could be longer than 36 years. For example, the four longest precursor times amongst the 47 Ψ-identifications of Figure 2b range between 15,100 and 18,900 days, i.e., between 41 and 52 years. The fitted EEPAS time distribution is limited by the length of the available catalogue. Even a well-fitted EEPAS model to the available precursory information does not guarantee the adequacy of the time distribution f (Equation (6)). This is seen in the fitting of the FLEEPAS model at short lead times (Figure 7)—the shorter the lead time, the smaller the temporal scaling factor 10^{a_t} and the mean of the fitted time distribution f for a given magnitude. The predictive scaling relations (Figure 2) and the parameters of the fitted EEPAS-0F (Table 1) are likely to be similarly affected by the limited duration of the catalogues from which they are derived.

Magnitude thresholds, catalogue starting date and fitting period need to be chosen appropriately when the EEPAS model is applied to an earthquake catalogue. The values that are chosen ensure completeness and homogeneity of the catalogue above the lower magnitude thresholds over the fitting period within the region of surveillance. Another important consideration highlighted by this study is the completeness of the precursory earthquake contributions for the target earthquakes. There is evidence from the application of EEPAS to synthetic earthquake catalogues [35] that the precursor time for a given magnitude is longer if the regional strain rate is lower. Therefore, it is expected that in regions with low strain rates (or low seismicity), longer lead times will be required to achieve the same level of completeness of precursory contributions for a given target magnitude.

Fitting the FLEEPAS model to the NZ earthquake data for different lead times has revealed a time-space trade-off of precursory seismicity. In the EEPAS model, time and location have always been treated as independent distributions. However, the time-space trade-off shows that they are not independent. A time-space trade-off was previously noted in subjective identifications of the Ψ-phenomenon for a few major synthetic earthquakes generated by a physics-based earthquake

simulator [35]. The temporal and spatial scaling factors vary by about two across the range of lead times considered here (Figure 7). The previous synthetic examples suggest a trade-off of temporal and spatial scales ranging over an order of magnitude or more. Therefore, further systematic investigation is required to fully quantify the extent of the trade-off.

The fitted FLEEPAS and FLCEEPAS models are both seen to appreciably improve forecasting relative to the unfitted FLEEPAS model in the NZ data investigated here, for fixed lead times shorter than 11 years (Figures 6 and 9b). It can be easily understood that for lead times longer than 11 years, the improvement in forecasting should be small or non-existent. With a threshold magnitude $m_c = 4.95$ (Table 1), the target earthquakes are mainly of magnitudes 6 and below because of the Gutenberg–Richter relation. For magnitudes less than 6, the completeness of precursory earthquakes was shown to be high (>0.7) for lead times greater than 11 years (Figure 4). Therefore, applying FLEEPAS and FLCEEPAS should appreciably improve forecasting for only a few target earthquakes with the largest magnitudes for such lead times.

It is often found that hybrids of forecasting models perform better than their individual components, whether the hybrids are of additive [31,36], multiplicative [37–39] or maximum type [40]. The fitted FLEEPAS models for long and short lead times exploit precursors on different temporal and spatial scales. Ultimately, forming mixtures of models fitted with different lead times may bring larger improvements in forecasting.

5. Conclusions

A formula was developed for the completeness of precursory earthquake contributions to target earthquakes as a function of the catalogue lead time and target magnitude. Two new versions of the EEPAS model were developed to examine the effect of the lead time on medium-term earthquake forecasting and to compensate for incompleteness of precursory earthquake contributions.

The fixed lead time EEPAS model, FLEEPAS, provided insights on how the missing precursory earthquakes affect forecasting performance of the model and on the complexity of precursory seismicity. Applying FLEEPAS to the NZ catalogue improved model performance at short lead times and revealed a space-time trade-off of precursory seismicity which was not part of the original EEPAS model formulation. The fixed lead time compensated EEPAS model, FLCEEPAS, explicitly compensates for incomplete precursory earthquake contributions due to restricted lead times.

Fitting FLEEPAS and FLCEEPAS to the NZ catalogue improved model performance to about the same extent for a given lead time, although they achieved the improvement in different ways. Forming mixtures of models with different fixed lead times may ultimately lead to greater improvements in forecasting. Also, further systematic study of synthetic catalogues should lead to a better understanding of precursory seismicity patterns to support real data observations.

Author Contributions: Conceptualization, D.A.R., A.C. and S.J.R.; methodology, D.A.R.; software, D.A.R. and S.J.R.; writing—original draft preparation, D.A.R. and S.J.R.; writing—review and editing, D.A.R., A.C. and S.J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Strategic Science Investment Fund (SSIF) of the Ministry of Business, Innovation and Employment, NZ.

Acknowledgments: We thank David Burbidge and Rob Buxton for their constructive internal reviews of an earlier version of this manuscript. We also thank Robert Shcherbakov and three anonymous reviewers for many helpful comments and suggestions, which greatly improved the manuscript. We acknowledge the New Zealand GeoNet project and its sponsors EQC, GNS Science and LINZ, for providing the earthquake catalogue used in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rikitake, T. Classification of Earthquake Precursors. *Tectonophysics* **1979**, *54*, 293–309. [CrossRef]
2. Ogata, Y. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *J. Am. Stat. Assoc.* **1988**, *83*, 9–27. [CrossRef]

3. Ogata, Y. Space-Time Point-Process Models for Earthquake Occurrences. *Ann. Inst. Stat. Math.* **1998**, *50*, 379–402. [CrossRef]
4. Console, R.; Murru, M. A Simple and Testable Model for Earthquake Clustering. *J. Geophys. Res. Space Phys.* **2001**, *106*, 8699–8711. [CrossRef]
5. Gerstenberger, M.C.; Wiemer, S.; Jones, L.M.; Reasenberg, P.A. Real-Time Forecasts of tomorrow's Earthquakes in California. *Nat. Cell Biol.* **2005**, *435*, 328–331. [CrossRef] [PubMed]
6. Ende, M.V.D.; Ampuero, J. On the Statistical Significance of Foreshock Sequences in Southern California. *Geophys. Res. Lett.* **2020**, *47*, 2019–086224. [CrossRef]
7. Reasenberg, P.A. Foreshock Occurrence before Large Earthquakes. *J. Geophys. Res. Space Phys.* **1999**, *104*, 4755–4768. [CrossRef]
8. Keilis-Borok, V.; Kossobokov, V. Premonitory Activation of Earthquake Flow: Algorithm M8. *Phys. Earth Planet. Inter.* **1990**, *61*, 73–83. [CrossRef]
9. Kossobokov, V.G.; Keilis-Borok, V.I.; Smith, S.W. Localization of Intermediate-Term Earthquake Prediction. *J. Geophys. Res. Space Phys.* **1990**, *95*, 19763–19772. [CrossRef]
10. Sobolev, G. Precursory Phases, Seismicity Precursors, and Earthquake Prediction in Kamchatka. *Volcanol. Seismol.* **1999**, *20*, 615–627.
11. Jaume, S.C.; Sykes, L.R. Evolving Towards a Critical Point: A Review of Accelerating Seismic Moment/Energy Release Prior to Large and Great Earthquakes. *Pure Appl. Geophys. PAGEOPH* **1999**, *155*, 279–305. [CrossRef]
12. Nanjo, K.; Rundle, J.B.; Holliday, J.; Turcotte, D.L. Pattern Informatics and Its Application for Optimal Forecasting of Large Earthquakes in Japan. *Pure Appl. Geophys. PAGEOPH* **2006**, *163*, 2417–2432. [CrossRef]
13. Wyss, M.; Sobolev, G.; Clippard, J.D. Seismic Quiescence Precursors to Two M7 Earthquakes on Sakhalin Island, Measured by Two Methods. *Earth Planets Space* **2004**, *56*, 725–740. [CrossRef]
14. Evison, F.; Rhoades, D. Precursory Scale Increase and Long-Term Seismogenesis in California and Northern Mexico. *Ann. Geophys.* **2002**, *45*, 479–495.
15. Evison, F. The Precursory Earthquake Swarm. *Phys. Earth Planet. Inter.* **1977**, *15*, 19–23. [CrossRef]
16. Evison, F. Generalised Precursory Swarm Hypothesis. *J. Phys. Earth* **1982**, *30*, 155–170. [CrossRef]
17. Evison, F.F.; Rhoades, D. Demarcation and Scaling of Long-Term Seismogenesis. *Pure Appl. Geophys. PAGEOPH* **2004**, *161*, 21–45. [CrossRef]
18. Ross, Z.E.; Idini, B.; Jia, Z.; Stephenson, O.L.; Zhong, M.; Wang, X.; Zhan, Z.; Simons, M.; Fielding, E.J.; Yun, S.-H.; et al. Hierarchical Interlocked Orthogonal Faulting in the 2019 Ridgecrest Earthquake Sequence. *Science* **2019**, *366*, 346–351. [CrossRef]
19. Rhoades, D.; Evison, F.F. Long-Range Earthquake Forecasting With Every Earthquake a Precursor According to Scale. *Pure Appl. Geophys. PAGEOPH* **2004**, *161*, 47–72. [CrossRef]
20. Rhoades, D.; Evison, F. The EEPAS Forecasting Model and the Probability of Moderate-to-Large Earthquakes in Central Japan. *Tectonophysics* **2006**, *417*, 119–130. [CrossRef]
21. Rhoades, D.; Evison, F.F. Test of the EEPAS Forecasting Model on the Japan Earthquake Catalogue. *Pure Appl. Geophys. PAGEOPH* **2005**, *162*, 1271–1290. [CrossRef]
22. Rhoades, D. Application of the EEPAS Model to Forecasting Earthquakes of Moderate Magnitude in Southern California. *Seism. Res. Lett.* **2007**, *78*, 110–115. [CrossRef]
23. Console, R.; Rhoades, D.; Murru, M.; Evison, F.F.; Papadimitriou, E.; Karakostas, V.G. Comparative Performance of Time-Invariant, Long-Range and Short-Range Forecasting Models on the Earthquake Catalogue of Greece. *J. Geophys. Res. Solid Earth* **2006**, *111*, B09304. [CrossRef]
24. Zechar, J.D.; Schorlemmer, D.; Liukis, M.; Yu, J.; Euchner, F.; Maechling, P.; Jordan, T.H. The Collaboratory for the Study of Earthquake Predictability Perspective on Computational Earthquake Science. *Concurr. Comput. Pr. Exp.* **2009**, *22*, 1836–1847. [CrossRef]
25. Schorlemmer, D.; Werner, M.J.; Marzocchi, W.; Jordan, T.H.; Ogata, Y.; Jackson, D.D.; Mak, S.; Rhoades, D.; Gerstenberger, M.C.; Hirata, N.; et al. The Collaboratory for the Study of Earthquake Predictability: Achievements and Priorities. *Seism. Res. Lett.* **2018**, *89*, 1305–1313. [CrossRef]
26. Rhoades, D.; Christophersen, A.; Gerstenberger, M.C.; Liukis, M.; Silva, F.; Marzocchi, W.; Werner, M.J.; Jordan, T.H. Highlights from the First Ten Years of the New Zealand Earthquake Forecast Testing Center. *Seism. Res. Lett.* **2018**, *89*, 1229–1237. [CrossRef]
27. Gerstenberger, M.C.; Rhoades, D.A.; McVerry, G.H. A Hybrid. Time-Dependent Probabilistic Seismic-Hazard. Model. For Canterbury, New Zealand. *Seismol. Res. Lett.* **2016**, *87*, 1311–1318. [CrossRef]

28. Rhoades, D.; Christophersen, A. Time-Varying Probabilities of Earthquake Occurrence in Central New Zealand Based on the EEPAS Model Compensated for Time-Lag. *Geophys. J. Int.* **2019**, *219*, 417–429. [CrossRef]
29. Gutenberg, B.; Richter, C.F. Frequency of Earthquakes in California. *Bull. Seismol. Soc. Am.* **1944**, *34*, 185–188.
30. Gerstenberger, M.C.; Rhoades, D. New Zealand Earthquake Forecast Testing Centre. *Pure Appl. Geophys. PAGEOPH* **2010**, *167*, 877–892. [CrossRef]
31. Rhoades, D. Mixture Models for Improved Earthquake Forecasting With Short-to-Medium Time Horizons. *Bull. Seism. Soc. Am.* **2013**, *103*, 2203–2215. [CrossRef]
32. GeoNet. Earthquake Forecasts. Available online: https://www.geonet.org.nz/earthquake/Forecast/ (accessed on 14 September 2020).
33. Harte, D.; Vere-Jones, D. The Entropy Score and Its Uses in Earthquake Forecasting. *Pure Appl. Geophys. PAGEOPH* **2005**, *162*, 1229–1253. [CrossRef]
34. Rhoades, D.A.; Schorlemmer, D.; Gerstenberger, M.C.; Christophersen, A.; Zechar, J.D.; Imoto, M. Efficient Testing of Earthquake Forecasting Models. *Acta Geophys.* **2011**, *59*, 728–747. [CrossRef]
35. Christophersen, A.; Rhoades, D.A.; Colella, H.V. Precursory Seismicity in Regions of Low Strain Rate: Insights from a Physics-Based Earthquake Simulator. *Geophys. J. Int.* **2017**, *209*, 1513–1525. [CrossRef]
36. Marzocchi, W.; Zechar, J.D.; Jordan, T. Bayesian Forecast Evaluation and Ensemble Earthquake Forecasting. *Bull. Seism. Soc. Am.* **2012**, *102*, 2574–2584. [CrossRef]
37. Imoto, M.; Rhoades, D. Seismicity Models of Moderate Earthquakes in Kanto, Japan Utilizing Multiple Predictive Parameters. *Pure Appl. Geophys. PAGEOPH* **2010**, *167*, 831–843. [CrossRef]
38. Rhoades, D.; Gerstenberger, M.C.; Christophersen, A.; Zechar, J.D.; Schorlemmer, D.; Werner, M.J.; Jordan, T. Regional Earthquake Likelihood Models II: Information Gains of Multiplicative Hybrids. *Bull. Seism. Soc. Am.* **2014**, *104*, 3072–3083. [CrossRef]
39. Bird, P.; Jackson, D.D.; Kagan, Y.Y.; Kreemer, C.; Stein, R.S. GEAR1: A Global Earthquake Activity Rate Model Constructed from Geodetic Strain Rates and Smoothed Seismicity. *Bull. Seism. Soc. Am.* **2015**, *105*, 2538–2554. [CrossRef]
40. Rhoades, D.; Liukis, M.; Christophersen, A.; Gerstenberger, M.C. Retrospective Tests of Hybrid Operational Earthquake Forecasting Models for Canterbury. *Geophys. J. Int.* **2015**, *204*, 440–456. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).