Bulletin of the Seismological Society of America, Vol. 104, No. 6, pp. 3072–3083, December 2014, doi: 10.1785/0120140035

Regional Earthquake Likelihood Models II: Information Gains of Multiplicative Hybrids

by D. A. Rhoades, M. C. Gerstenberger, A. Christophersen, J. D. Zechar,*
D. Schorlemmer,* M. J. Werner,† and T. H. Jordan

Abstract The Regional Earthquake Likelihood Models experiment in California tested the performance of earthquake likelihood models over a five-year period. First-order analysis showed a smoothed-seismicity model by Helmstetter *et al.* (2007) to be the best model. We construct optimal multiplicative hybrids involving the best individual model as a baseline and one or more conjugate models. Conjugate models are transformed using an order-preserving function. Two parameters for each conjugate model and an overall normalizing constant are fitted to optimize the hybrid model. Many two-model hybrids have an appreciable information gain (log probability gain) per earthquake relative to the best individual model. For the whole of California, the Bird and Liu (2007) Neokinema and Holliday *et al.* (2007) pattern informatics (PI) models both give gains close to 0.25. For southern California, the Shen *et al.* (2007) geodetic model gives a gain of more than 0.5, and several others give gains of about 0.2. The best three-model hybrid for the whole region has the Neokinema and PI models as conjugates. The best three-model hybrid for southern California has the Shen *et al.* (2007) and PI models as conjugates. The information gains of the best multiplicative hybrids are greater than those of additive hybrids constructed from the same set of models. The gains tend to be larger when the contributing models involve markedly different concepts or data. These results need to be confirmed by further prospective tests. Multiplicative hybrids will be useful for assimilating other earthquake-related observations into forecasting models and for combining forecasting models at all timescales.

Introduction

The Regional Earthquake Likelihood Models (RELM) experiment in California tested the performance of a variety of earthquake forecasting models over a five-year period (Schorlemmer and Gerstenberger, 2007). The target earthquakes were those with magnitudes M ≥ 4.95 in the Advanced National Seismic System catalog. The models were based on a variety of data inputs and modeling techniques, including spatial smoothing of previous earthquake locations, geodetic estimates of strain rates from Global Positioning System data, identification of regions with fluctuating seismicity rates, geologic fault slip rates, and physics-based numerical earthquake simulation (Bird and Liu, 2007 [Bird and Liu]; Ebel *et al.*, 2007 [Ebel *et al.*]; Field, 2007; Helmstetter *et al.*, 2007 [Helmstetter *et al.*]; Holliday *et al.*, 2007 [Holliday *et al.*]; Kagan *et al.*, 2007 [Kagan *et al.*]; Shen *et al.*, 2007 [Shen *et al.*]; Ward, 2007 [Ward]; Wiemer and Schorlemmer, 2007 [Wiemer and Schorlemmer]). A first-order analysis of the results by the Collaboratory for the Study of Earthquake Predictability (CSEP) showed the HKJ smoothed-seismicity model by Helmstetter *et al.* (2007) to be the most informative model (Schorlemmer *et al.*, 2010; Zechar *et al.*, 2013).

Additive hybrids have sometimes proved effective in combining the information from disparate forecasting models to produce sizeable information gains over their component models (Rhoades and Gerstenberger, 2009; Rhoades and Stirling, 2012; Rhoades, 2013). However, in the case of the RELM experiment, a Bayesian analysis of additive hybrids found that a probability-weighted average of the RELM models does not outperform the best individual model to any appreciable extent (Marzocchi *et al.*, 2012).

A different approach to constructing hybrids is adopted here, using a multiplicative procedure. Hybrid models with multiplicative probability gains were suggested by the early work of Aki (1981) and Utsu (1982) on conditionally independent earthquake precursors. Later, Imoto (2006, 2007) generalized the notion of multiple earthquake precursors to

*Also at Southern California Earthquake Center, University of Southern California, 3651 Trousdale Parkway, MC-0740, Los Angeles, California 90089-0740.

†Department of Geosciences, Princeton University, Princeton, New Jersey 08544.

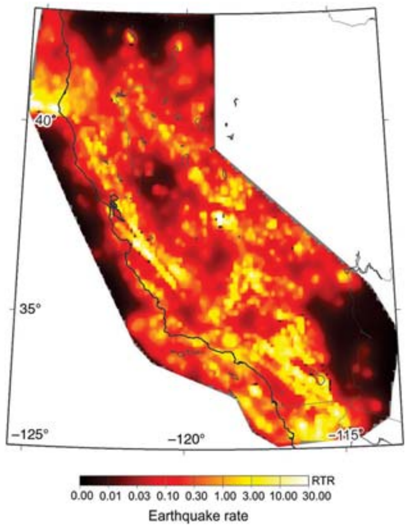

Figure 1. Map of earthquake rates, relative to reference (RTR), in the baseline Helmstetter *et al.* HKJ mainshock+aftershock model. In the reference model, one earthquake per year is expected to exceed any magnitude m in an area of 10^m km^2. The color version of this figure is available only in the electronic edition.

earthquake probabilities or rates estimated from multidisciplinary observations and showed that multiplicative probability gains could theoretically still be obtained without the independence assumption. See also Faenza and Marzocchi (2010) and Shebalin *et al.* (2014) for applications of a statistical model with multiplicative structure to earthquake forecasting. In the statistical literature, it is recognized that multiplicative models are often appropriate for modeling of count data (e.g., McCullagh and Nelder, 1989).

With these previous works in mind, we assume a multiplicative structure for the expected number of earthquakes in the cells of hybrid models formed from a selection of two or more models in the RELM experiment. Each selection is composed of a baseline model and one or more conjugate models. Because the main question of interest is whether the best individual model can be improved upon, only selections that include the best model as the baseline are considered. We aim to construct optimal multipliers that can be applied to the cell expectations of the baseline model by transforming the corresponding cell expectations of the conjugate models. We also aim to estimate the information gain of the hybrid model over the baseline model, taking into account the number of parameters involved in fitting the multipliers and the limited number of target earthquakes. A more dependable estimate of the information gain will be provided in due course by formal prospective testing of the hybrid models in the Southern California Earthquake Center (SCEC) CSEP testing center.

The RELM experiment included models in two different classes—the mainshock class and the mainshock+aftershock class. The five-year test period provided 20 target earthquakes in the former class and 31 target earthquakes in the latter class in the whole California test region. A majority of the models submitted applied only to southern California, with small variations between models in the regions actually covered. For these models, the test period provided no more than 11 earthquakes in the mainshock class and 22 in the mainshock+aftershock class. The same limitations apply to any hybrids that include these models. To have a satisfactory number of target events for the fitting of hybrid models, especially for southern California, the analysis here is focused on the mainshock+aftershock target earthquake set, that is, all earthquakes with M $\geq$4.95.

The spatial distribution of the rates of M 5.0 earthquakes in the Helmstetter *et al.* HKJ model is illustrated in Figure 1. In this and other figures, the earthquake rates are expressed relative to a reference (RTR) model in which one earthquake per year exceeding any magnitude m is expected in an area of 10^m km^2; the plotted rates are the result of dividing the forecast rates by the reference rates. In this form of presentation, introduced by Rhoades and Evison (2004), a common color scale can be applied for different magnitude selections; the rates in RTR units for a particular model are dependent on magnitude only if the Gutenberg–Richter b-value of the model differs from 1.

Method

Each cell of a RELM model applies to an earthquake location bin, indexed by j, and a magnitude bin, indexed by k. The expected number of earthquakes in each cell of the baseline model during the five-year period of the experiment is denoted by $\{\lambda_1(j,k)\}$, that of the conjugate models by $\{\lambda_i(j,k)\}$, $i = 2, \ldots, n_i$, and that of the hybrid model by $\{\lambda_{\rm H}(j,k)\}$, in which i indexes the conjugate model itself, $j = 1, \ldots, n_j$ is the number of location bins, and $k = 1, \ldots, n_k$ is the number of magnitude bins. For the baseline model, the expectation in each cell is preserved and becomes the first factor of the hybrid rate. For a conjugate model, the expected number in each spatial bin is first summed over all magnitude bins, and the summed expectation is denoted $\lambda_i(j,\cdot)$.

$$\lambda_i(j,\cdot) = \sum_{k=1}^{n_k} \lambda_i(j,k). \tag{1}$$

Implicit in this procedure is the assumption that a conjugate model contains no magnitude-specific information about earthquake occurrence beyond that already incorporated in the baseline model. This assumption is reasonable for models that follow the Gutenberg–Richter frequency–magnitude relation with a spatially invariant b-value. The summed expectation is treated as an index of earthquake occurrence, that is, as an alarm function in the sense of Zechar and Jordan (2008). The ordering of the summed expectations over the spatial cells is held to be important, but not the actual summed expectations themselves. This ordering is preserved in computing the model's multiplicative contribution to the hybrid. We use a smooth nonlinear transformation of flexible form, but with only a small number of adjustable parameters, to convert a summed expectation into a multiplier. The hybrid model has the form

$$\lambda_{\mathrm{H}}(j,k) = \lambda_1(j,k)\exp\left(a + \sum_{i=2}^{n_i} f_i[\lambda_i(j,\cdot)]\right), \quad (2)$$

in which f_i is an order-preserving function of the form

$$f_i[\lambda] = b_i(\log(1+\lambda))^{c_i} \qquad (b_i \geq 0; c_i > 0). \quad (3)$$

The adjustable parameters are the normalizing parameter a and the shape parameters b_i and c_i; $i = 2, \ldots, n_i$. The functional form of f_i is chosen because it is non-negative, monotone nondecreasing, and of flexible form. Within the range of values of cell expectations in this study, in which $\lambda \ll 1$ and therefore $\log(1+\lambda) \approx \lambda$, it can be approximately linear, convex downward, convex upward, or constant, depending on the parameters. The results will clearly depend to some extent on the form of the function used. When fitting a large set of target earthquakes, it would be feasible (and perhaps even desirable) to use a function with more parameters and, consequently, greater flexibility. However, the number of target earthquakes is not large in this study, and to avoid overfitting, a function with only two adjustable parameters for each conjugate model is used.

We compare the log-likelihood gains of multiplicative hybrids with those of additive hybrids. For this purpose, we construct optimal additive hybrids, in which the cell expectations are linear combinations of the corresponding expectations in the individual models that is

$$\lambda_{\mathrm{H}}(j,k) = \sum_{i=1}^{n} a_i\lambda_i(j,k) \qquad (a_i \geq 0; i = 1, \ldots, n), \quad (4)$$

in which there are n adjustable parameters (a_i, $i = 1, \ldots, n$).

The adjustable parameters of the hybrid models are fitted by maximum likelihood, using the downhill simplex method (Nelder and Mead, 1965). The aim is to produce hybrid models that are more informative than the individual models. To compute the information gain, we need to take into account the number of fitted parameters in each hybrid model and the number of target earthquakes. Accordingly, for the fitted hybrid models, we estimate the corrected information gain per earthquake (IGPEc) using the corrected Akaike information criterion (AICc) statistic (Hurvich and Tsai, 1989), which adjusts the standard AIC for choosing between competing fitted models (Akaike, 1974) taking into account the number of target earthquakes. The AICc statistic is of the form

$$\mathrm{AICc} = -2\ln L + 2p + \frac{p+1}{N-p-1}, \quad (5)$$

in which N is the number of target earthquakes, p is the number of fitted parameters, and $\ln L$ is the log likelihood of the fitted hybrid model, given by (Rhoades *et al.*, 2011)

$$\ln L = \sum_{n=1}^{N} \ln\lambda_{\mathrm{H}}(j_n,k_n) - \hat{N}_{\mathrm{H}}, \quad (6)$$

in which the N target earthquakes occur in cells $\{(j_n,k_n), n = 1, \ldots, N\}$, and $\hat{N}_{\mathrm{H}}$ denotes the total expected number, of target earthquakes of the hybrid model, that is

$$\hat{N}_{\mathrm{H}} = \sum_{k=1}^{n_k}\sum_{j=1}^{n_j} \lambda_{\mathrm{H}}(j,k). \quad (7)$$

The form of the log likelihood in equation (6) differs from that given by Schorlemmer *et al.* (2007) in that it is devoid of the Poisson penalty term for more than one target earthquake occurring in a single cell. However, this difference does not affect the optimization of parameters or the likelihood ratio of two different models.

If Δ is the change in AICc achieved by fitting the hybrid model $\{\lambda_{\mathrm{H}}(j,k)\}$, compared to the baseline model $\{\lambda_1(j,k)\}$, then the corrected information gain per earthquake is

$$\mathrm{IGPEc} = \frac{-\Delta}{2N}. \quad (8)$$

The product of the multipliers is unity for all values of λ_i if $a = 0$ and $b_i = 0$, for $i = 2, \ldots, n_i$. Therefore, the optimized hybrid model cannot have a lower log likelihood than the baseline model. However, the IGPEc can be negative, because of the penalties for p and N in AICc. The uncertainty of the corrected information gain can be assessed using an adaptation of the T-test (Rhoades *et al.*, 2011) for comparing one earthquake likelihood model to another based on independent data. From equations (5) and (6), we have

$$\mathrm{IGPEc} = \frac{\hat{N}_1 - \hat{N}_{\mathrm{H}}}{N} - \frac{1}{2N}\left[2p + \frac{p+1}{(N-p-1)}\right] + \frac{1}{N}\sum_{n=1}^{N}[X_n - Y_n], \quad (9)$$

in which $X_n = \ln\lambda_{\mathrm{H}}(j_n,k_n)$, $Y_n = \ln\lambda_1(j_n,k_n)$, and $\hat{N}_i$ denotes the total expected number of target earthquakes of the ith model.

Equation (9) is equivalent to the expression for the IGPE for prospective testing given in equation (17) of Rhoades *et al.* (2011), except for an added penalty in the second term for the number of fitted parameters and the finite number of target earthquakes. The adapted T-test is therefore the same as that of Rhoades *et al.* (2011) except for this added penalty. It must be emphasized that retrospective testing using a modified test statistic, however carefully constructed, is a poor substitute for prospective testing, in which all adjustable parameters are set independently of the test data. In retrospective testing, the problem of overfitting can never be eliminated with certainty.

The hybrid models were fitted using all earthquakes with M ≥ 4.95 in the test period as the target earthquakes. Some models were submitted in different versions to the mainshock and mainshock+aftershock classes and some only to one class (Zechar *et al.*, 2013). If a model was submitted to the mainshock+aftershock class, that version of the model is used for forming hybrids. This applies to the models by Bird and Liu (2007), Ebel *et al.* (2007), Helmstetter *et al.* (2007), Kagan *et al.* (2007), and Shen *et al.* (2007). Otherwise the model submitted to the mainshock class is used. This applies to the models by Holliday *et al.* (2007), Wiemer and Schorlemmer (2007), and the six different models by Ward (2007), although it emerged during the tests that the Holliday *et al.* model was actually intended for the mainshock+aftershock class (Zechar *et al.*, 2013). Because the fitting process includes a normalization parameter, no model is disadvantaged by the difference between the total expected number of earthquakes in these two classes. Ebel *et al.* (2007) submitted both an original and a corrected version of their model; only the corrected version is used here.

Table 1
Parameter Estimates and Other Statistics of Two-Model Multiplicative Hybrids

Conjugate Model	a	b_2	c_2	N	$\Delta \ln L$
Holliday *et al.* PI	−1.58	12.14	0.554	31	11.2
Bird and Liu Neokinema	−9.96	13.05	0.067	31	11.4
Wiemer and Schorlemmer ALM	−2.73	2.88	0.022	31	0.9
Ebel *et al.*	−1.26	1.40	0.050	31	0.08
Shen *et al.* geodetic	−9.05	19.79	0.153	22	16.2
Kagan *et al.*	−9.70	19.22	0.139	22	8.3
Ward geodetic8.1	−14.88	20.00	0.069	20	7.4
Ward geodetic8.5	−13.55	19.43	0.075	20	7.4
Ward seismic	−15.10	19.83	0.060	20	6.0
Ward combo	−13.58	19.77	0.086	20	6.5
Ward simulation	−15.16	18.08	0.030	20	3.6
Ward geologic	−14.03	16.99	0.038	20	4.4
Holliday *et al.* PI (S Cal)	−14.61	17.67	0.045	22	9.5
Bird and Liu Neokinema (S Cal)	−15.93	19.48	0.047	22	10.3

N, number of target earthquakes; $\Delta \ln L$, increase in log likelihood over baseline Helmstetter *et al.* HKJ model; S Cal, southern California.

Results

Two-Model Hybrids

The optimal hybrids were first computed for all pairs of models that include the Helmstetter *et al.* HKJ model as the baseline. Summary statistics of these two-model hybrids are shown in Table 1, including the fitted parameter values, the number of target earthquakes, and the increase in the log likelihood of the hybrid model compared to the baseline model. For the Bird and Liu Neokinema model and the Holliday *et al.* pattern informatics (PI) model, optimal hybrids were formed with the HKJ model both for the whole test region and for the smaller southern California region covered by the Shen *et al.* geodetic model. Both of these conjugate models give substantial log-likelihood increases when applied to the whole of California. Therefore, there is interest in knowing their value as conjugate models in comparison to the models that are defined only for southern California.

Table 1 shows that all hybrid models except those involving Wiemer and Schorlemmer Asperity-based Likelihood Model (ALM) and Ebel *et al.* give considerable gains in the log likelihood over the Helmstetter *et al.* HKJ model. The hybrid model with the largest gain is that involving the Shen *et al.* geodetic model for southern California. Among the hybrid models for the whole of California, the models involving the Bird and Liu Neokinema model and the Holliday *et al.* PI model have the largest gains, and these models also have the second and third ranked gains for southern California.

The log-likelihood gains of the best two-model multiplicative hybrids are much greater than those of the optimal additive models constructed, as in equation (4), by combining all models for the whole of California and for southern California. For the whole of California, the optimal additive hybrid model has a log-likelihood gain of 4.8, compared with 11.4 for the best two-model multiplicative hybrid. It has coefficients of 0.44 for Helmstetter *et al.* HKJ, 0.39 for Holliday *et al.* PI, 0.12 for Wiemer and Schorlemmer ALM, and less than 0.001 for the other models. The log-likelihood gain of 4.5 corresponds to an IGPEc of 0.06 if the penalties for the two parameters that make negligible contributions to the optimal additive hybrid are neglected and to a negative information gain if they are not. For southern California, the optimal additive hybrid has a log-likelihood gain of 0.5, compared to 16.2 for the best two-model multiplicative hybrid. It has coefficients of 0.90 for Helmstetter *et al.* HKJ, 0.28 for Holliday *et al.* PI, and less than 0.001 for all other models. The log-likelihood gain of 0.05 corresponds to a negative value of IGPEc for the additive hybrid model. The low gains from forming additive hybrids of these models are consistent with the results of the Bayesian analysis by Marzocchi *et al.* (2012).

Figure 2 illustrates the optimal transformation of the conjugate model in each two-model multiplicative hybrid and the cumulative distribution over spatial cells of the summed cell rate for each conjugate model covering the whole of California. Figure 3 does the same for the conjugate models for southern California. The optimal transformations differ noticeably for the hybrids involving the Holliday *et al.* PI model for the whole of California and for southern California, showing that the optimal transformations can be sensitive to a subset of the test region. The much higher multiplier for low cell expectations in the whole of California can be attributed to the influence of three target earthquakes in northern California for which the PI model had very low expectations (see tables 2 and S2 of Zechar *et al.*, 2013): the M 5.0 earthquake of 25 June 2007 at (41.12° N, 124.82° W),

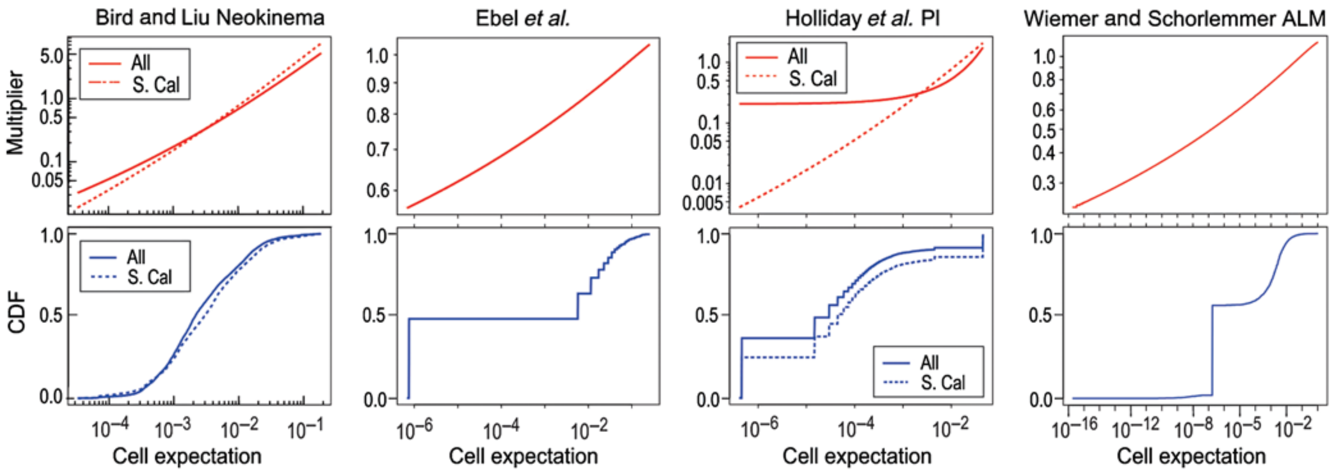

Figure 2. Conversion of cell expectation to multiplier for conjugate models defined for the whole of California in two-model hybrids with the Helmstetter *et al.* HKJ as the baseline model. For each conjugate model, the upper frame shows the multiplier $\exp(a + f_2[\lambda_2(j, \cdot)])$, and the lower frame shows the cumulative distribution of summed spatial expectations $\lambda_2(j, \cdot)$. For the Neokinema and pattern informatics (PI) models, the optimal multipliers for southern California only are also shown. The color version of this figure is available only in the electronic edition.

the M 5.0 earthquake of 26 April 2008 at (39.53° N, 119.93° W), and the M 5.4 earthquake of 30 April 2008 at (40.84° N, 123.50° W).

The cumulative distributions of the cell expectations, in conjunction with the optimal transformations of cell expectations, indicate what proportion of the cell expectations receives a multiplier in any given range. The cumulative distributions differ markedly between the conjugate models. In some cases, such as the Ebel *et al.* and Holliday *et al.* PI models (Fig. 2), the distribution function is composed mostly of jumps, indicating that the expectations are mostly concentrated at certain discrete values. In other cases, such as the Bird and Liu Neokinema and Kagan *et al.* models (Fig. 3), the distribution function increases smoothly. In yet other cases, such as the Wiemer and Schorlemmer ALM (Fig. 2) and Ward geologic (Fig. 3) models, the distribution function has both jumps and smooth sections.

For a hybrid forecast to have a large log-likelihood gain over the baseline model, it is necessary for the distribution of earthquake-cell expectations to be substantially different in the hybrid model than in the baseline model. If the multiplier varies over a wide range, the difference between the distribution of cell expectations in the hybrid and baseline models tends to be large. Therefore, it is not surprising that the conjugate models producing large log-likelihood gains (Table 1), such as the Shen *et al.* geodetic, Bird Liu Neokinema, and Holliday *et al.* PI, tend to have a wide range for the multiplier (Figs. 2 and 3), and that those producing small information gains, such as Ebel *et al.* and Wiemer and Schorlemmer ALM, tend to have a narrow range. However, having a wide range for the multiplier is not a sufficient condition to produce a large gain, because it is necessary also that the expected numbers in the conjugate model be well correlated with earthquake occurrence in some ways that the baseline model is not. Therefore, it is again not surprising that some models with a wide range for the multiplier, such as Ward simulation (Fig. 3), which has a multiplier varying over four orders of magnitude, can also have a low log-likelihood gain (Table 1).

The maximum value taken by the multiplier is indicative of the correlation of expected numbers in the baseline and conjugate models. Because the hybrid models in each class (the whole of California, or southern California) are all normalized to a similar number of target earthquakes, the total expected number of earthquakes, that is, the sum over all cells of the product of the baseline expected numbers and their associated multipliers, is approximately the same for all models in each class. The cells with the largest products contribute most to this sum. Conjugate models that are well correlated with the baseline model tend to have their highest expected numbers in the same cells as the baseline model, and the hybrid models are constructed so that the highest multipliers are in the cells with the highest conjugate model expectations. If high cell expectations in the baseline model all have high multipliers, then they all have very high products, and the expected number of earthquakes is strongly increased. Therefore, conjugate models that are well correlated with the baseline model will tend to produce multipliers with a relatively low maximum value, and vice versa. We deduce that the Ebel *et al.*, Holliday *et al.* PI, Wiemer and Schorlemmer ALM, the Ward geologic, Ward seismic, and Ward simulation, which have maximum multipliers not much larger than 1, are all well correlated with the Helmstetter *et al.* HKJ model, at least in the cells where their expected numbers are highest. Similarly, we deduce that the Shen *et al.* geodetic, Ward geodetic8.1, and Ward geodetic8.5 models, which have maximum multipliers of about 10, are not well correlated with the baseline model in the cells where their expected numbers are highest.

Figure 4 shows the adapted T-tests comparing the IGPEc of the HKJ model with each of the hybrid models for the whole

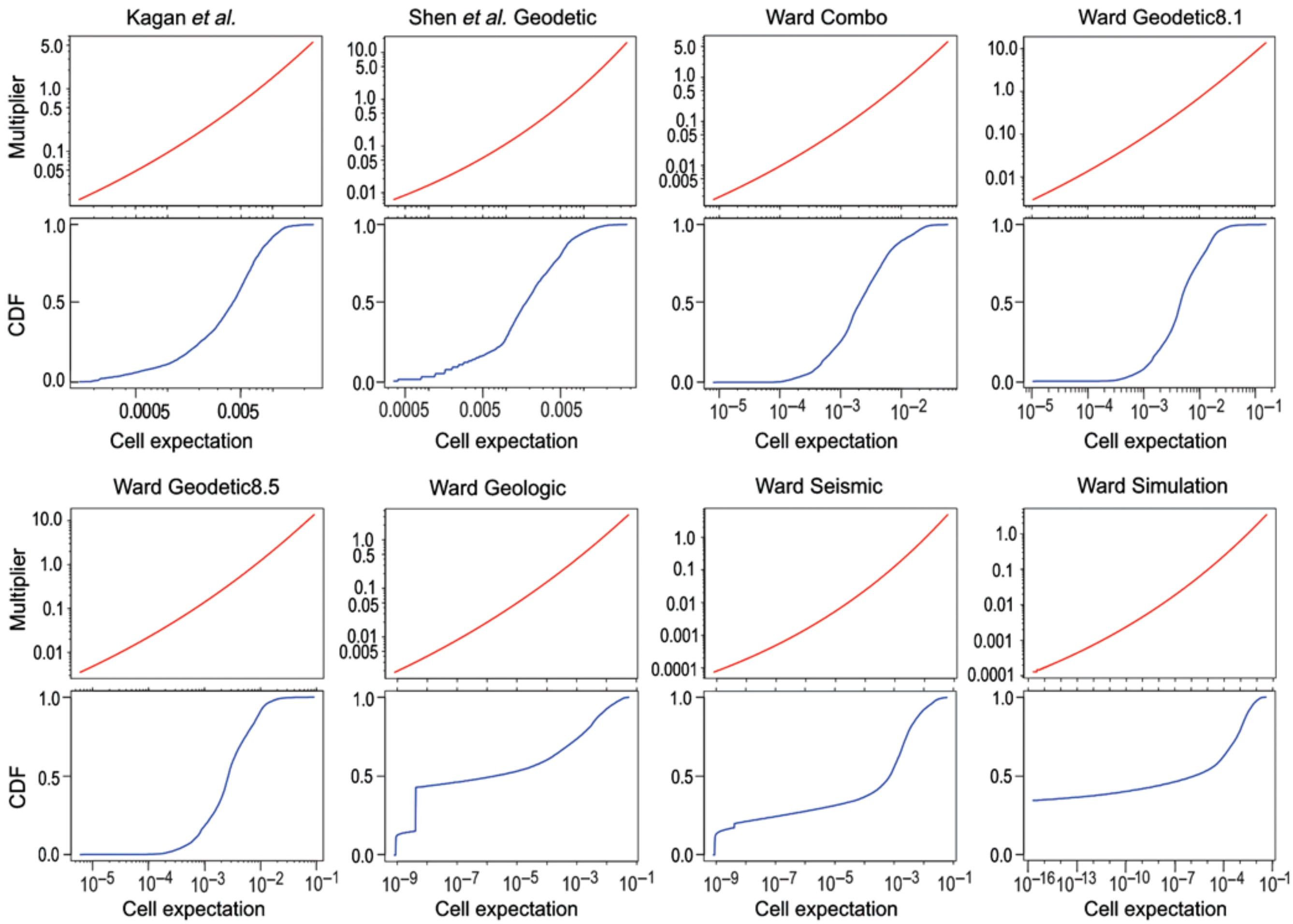

Figure 3. Conversion of cell expectation to multiplier for conjugate models defined for southern California only. For each conjugate model, the upper frame shows the multiplier $\exp(a + f_2[\lambda_2(j, \cdot)])$, and the lower frame shows the cumulative distribution of summed spatial expectations $\lambda_2(j, \cdot)$. The color version of this figure is available only in the electronic edition.

of California. In this figure, a negative IGPEc implies that the hybrid model is more informative than the HKJ model, and a 95% confidence interval wholly to the left of the vertical zero line indicates a statistically significant information gain at the 95% confidence level. The hybrid models are ranked according to information gain. Two hybrid models, involving Bird and Liu Neokinema and Holliday *et al.* PI, have an IGPEc of greater than 0.2 compared to the HKJ model, which is almost significant at the 95% confidence level. The other two, involving Ebel *et al.* and Wiemer and Schorlemmer ALM, are significantly less informative than HKJ, implying that the small likelihood gain of these hybrid models is insufficient to outweigh the penalty for the fitted parameters and the small target earthquake set.

Figure 5 shows the adapted T-tests of hybrid models for southern California, again ranked according to information gain. One model, the hybrid involving the Shen *et al.* geodetic model, has an IGPEc of 0.57, which is significant with more than 95% confidence. The information gain of the hybrid involving the Bird and Liu Neokinema model (0.30), which is ranked second, is not quite significant at the 95% level. However, the information gains of hybrids involving the following models, ranking third to sixth, are significant: Holliday *et al.* PI (0.26), Kagan *et al.* (0.20), Ward geodetic8.1 (0.18), and Ward geodetic8.5 (0.20). The hybrids involving the Ward combo, Ward seismic, Ward geologic, and Ward simulation models have smaller information gains relative to the HKJ model.

The estimates of information gain of the hybrid models, as shown in Figures 4 and 5, are generally greater than would be expected to result by chance. Given that these estimates are adjusted for the number of fitted parameters and the number of target earthquakes, the expected information gain from combining two models that contain no independent information on earthquake occurrence is zero or less. In fact, the IGPEc of most of the multiplicative models is positive, and for 5 out of 14 models (much more than the 1 out of 20 expected) it is significantly so.

Figure 6 illustrates the spatial distribution of expected numbers in the conjugate models and hybrid models for the whole of California. The hybrids with Bird and Liu Neokinema and Holliday *et al.* PI as the conjugate models have a

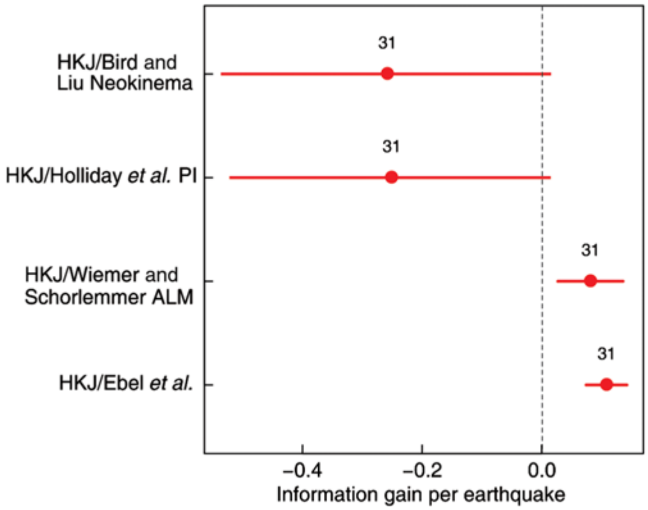

Figure 4. Information gain per earthquake (IGPEc) of the Helmstetter *et al.* HKJ model compared to two-model hybrids for the whole of California, corrected for parameter fitting and finite number of target earthquakes. Confidence intervals are 95% confidence limits. A negative information gain implies the hybrid model is better than the HKJ model. The color version of this figure is available only in the electronic edition.

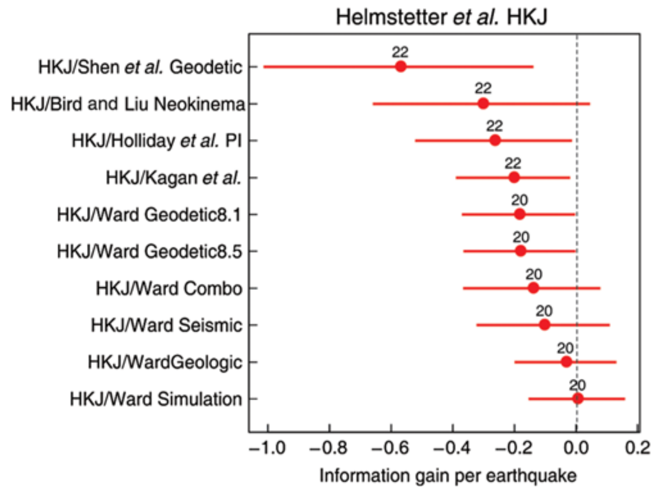

Figure 5. IGPEc of the Helmstetter *et al.* HKJ model compared to two-model hybrids for the southern California, corrected for parameter fitting and finite number of target earthquakes. Confidence intervals are 95% confidence limits. A negative information gain implies the hybrid model is better than the HKJ model. The color version of this figure is available only in the electronic edition.

similar IGPEc compared to the baseline HKJ model, but their spatial distributions are noticeably different from each other and from the HKJ model (Fig. 1). The spatial distribution of the Wiemer and Schorlemmer ALM and Ebel *et al.* conjugate models appear to be coarse versions of the spatial pattern represented in a more smoothly varying form by the HKJ model (Fig. 1). It is therefore not surprising that the spatial distributions of the corresponding hybrid models are visually similar to the HKJ model, and that these hybrid models are less informative than the HKJ model according to the adapted *T*-test (Fig. 4).

Figure 7 illustrates the spatial distribution of expected numbers in the conjugate models for southern California and their associated hybrids with the HKJ model. Again it is notable that models that give similar information gains in hybrid with the HKJ model, such as Kagan *et al.*, Ward geodetic8.1, and Ward geodetic8.5 (Fig. 5), have noticeably different spatial distributions in Figure 7.

Three-Model Hybrids

It is natural to search for the best multiple-model hybrids. However, the number of fitted parameters increases by two for each model added. Therefore, in view of the small number of target earthquakes and to avoid overfitting, we do not consider hybrids with more than three models. Starting with the best pair of models for the whole of California (Helmstetter *et al.* HKJ and Bird and Liu Neokinema) and for southern California (Helmstetter *et al.* HKJ and Shen *et al.* geodetic), three-model hybrids were formed with each remaining conjugate model. Table 2 shows the best three-model hybrids for the whole of California and for southern California. In both cases, the best three-model hybrid was formed by adding Holliday *et al.* PI as the third model. The two hybrids shown in Table 2 are the only three-model hybrids that had a higher IGPEc than the best two-model hybrid; they are also the only three-model hybrids having a log-likelihood increase $\Delta \ln L$ greater than one compared to the best two-model hybrid.

Two information gain statistics and their standard errors are given for each three-model hybrid in Table 2: IGPEc is the information gain per earthquake relative to the baseline HKJ model, and ΔIGPEc is the information gain per earthquake relative to the best two-model hybrid. These mean statistics can be assumed to be approximately normally distributed because of the central limit theorem. The IGPEc of 0.35 for the best three-model hybrid for the whole of California is slightly more than twice its standard error of 0.17, indicating statistical significance at about the 95% confidence level. The IGPEc of 0.79 for the best three-model hybrid for southern California is nearly three times its standard error of 0.29, indicating that it is easily significant at the 95% level. On the other hand, the ΔIGPEc of 0.092 for the best three-model hybrid for the whole of California is only 0.98 times its standard error, and the ΔIGPEc of 0.22 for the best three-model hybrid for southern California is only 1.6 times its standard error; therefore, neither of these information gains are statistically significant.

The spatial distributions of rates in the best three-model hybrid for the whole of California and southern California are mapped in Figure 8. In both cases, the best three-model hybrid has more contrasting earthquake rates than the corresponding best two-model hybrid. The areas with high rates are smaller and more intense, and the low rates (darker

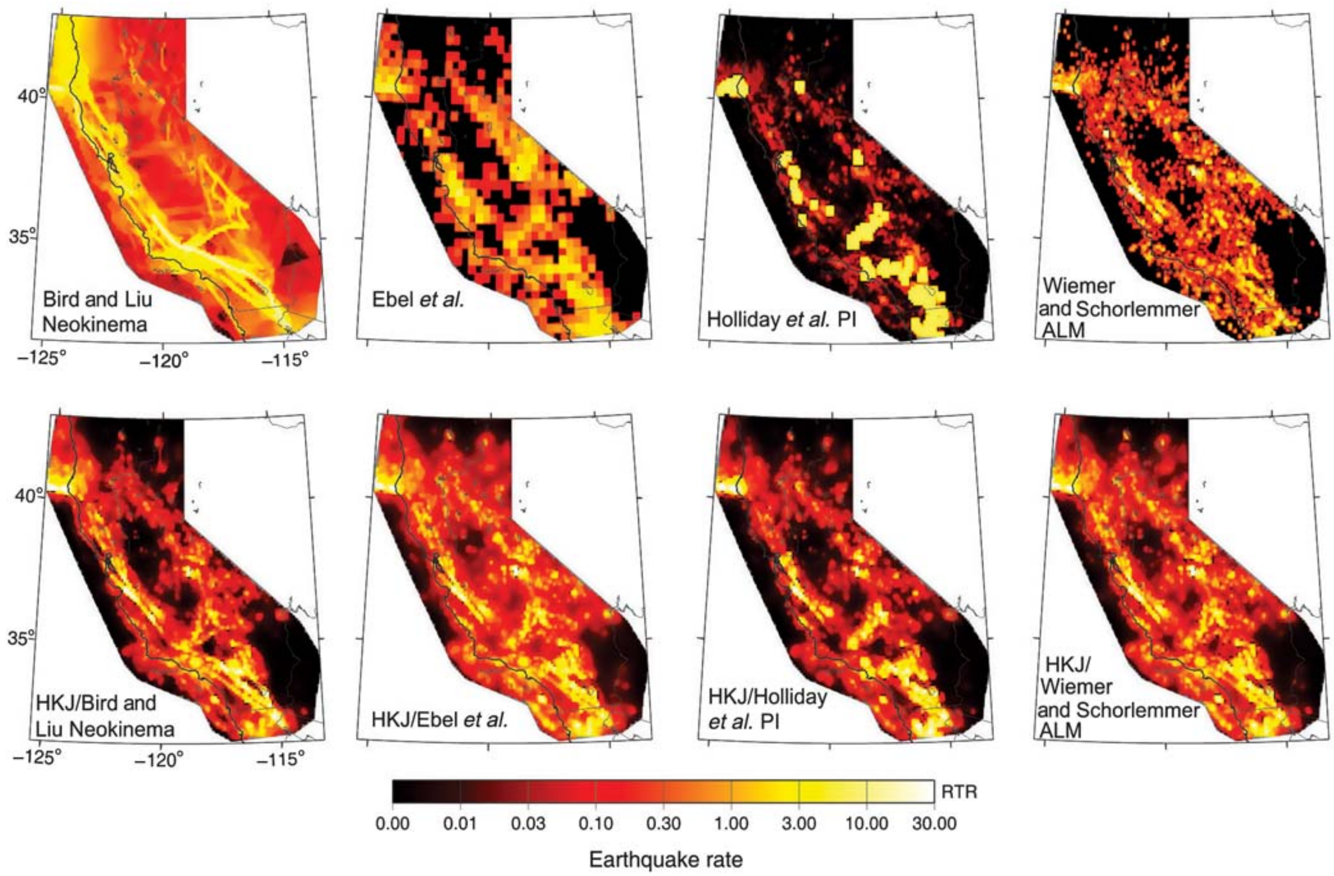

Figure 6. Map of earthquake rates, RTR, in conjugate models for the whole of California (on top) and in their hybrids with the baseline Helmstetter *et al.* HKJ model (underneath). In the reference model, one earthquake per year is expected to exceed any magnitude m in an area of 10^m km^2. The color version of this figure is available only in the electronic edition.

zones) occupy a larger proportion of the total area. These differences are due to the impact of the Holliday *et al.* PI model, which is itself a model of high contrasts, on the three-model hybrid rates.

Discussion

An observed trend is that larger information gains are obtained when the contributing models involve markedly different concepts or data. In particular, the models that make use of geodetic data, namely the Shen *et al.* geodetic, Bird and Liu Neokinema, Ward geodetic8.1, and Ward geodetic8.5 models, are effective as conjugate models with the Helmstetter *et al.* HKJ model, which makes use only of the earthquake catalog, as baseline. It has been noted above that these models are not closely correlated with the Helmstetter *et al.* HKJ model, and also that an effective conjugate model must be correlated with earthquake occurrences in a way that the baseline model is not. Further evidence that geodetic data could be used to improve earthquake forecasts based on catalog data was given recently by Wang *et al.* (2013). The effectiveness of the Holliday *et al.* PI model as a conjugate model in both two-model and three-model hybrids, is also worthy of note. Although based entirely on the earthquake catalog, this model is conceptually quite different from smoothed seismicity models. The information gains it provides in hybrid combinations are evidence that the seismicity patterns used by this model are useful for earthquake forecasting even if, as a standalone model, it may not be as informative as some others.

The larger information gains of the multiplicative hybrids used here, compared to additive hybrids, is partly due to the wider range of expectations that they entertain in any particular cell. Additive hybrids essentially produce weighted average expectations, modified by a normalization factor that is usually not much different from 1. Therefore, a cell expectation in an additive model is usually within the range of expectations of the individual models for the same cell. In contrast, in multiplicative hybrids, a cell expectation can be far outside the range of expectations of the individual models. However, this only occurs where the data as a whole support a much higher or lower multiplier than 1 for that cell, taking into account the constraints imposed on the multiplier by the parameterization of the model.

The challenge of continually increasing the information value of earthquake forecasting models can be met by learning how to integrate information from a variety of data and modeling inputs into hybrid models. The method presented here is a step in that direction. Although the parametric details of the

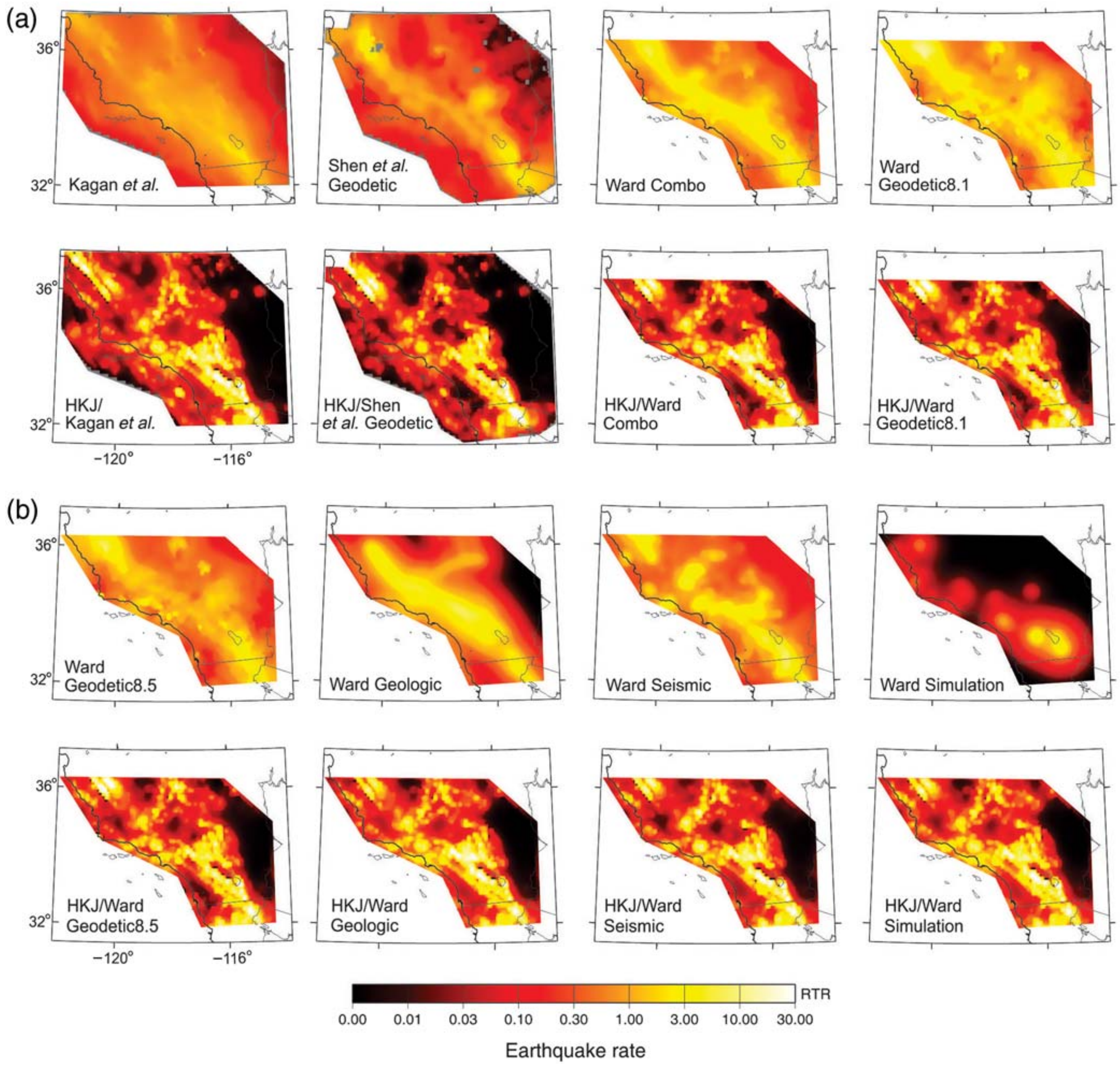

Figure 7. Map of earthquake rates, RTR, in conjugate models for southern California (on top) and in their hybrids with the baseline Helmstetter *et al.* HKJ model (underneath). In the reference model, one earthquake per year is expected to exceed any magnitude m in an area of 10^m km^2. (a) Kagan *et al.*, Shen *et al.* geodetic, Ward combo, and Ward geodetic8.1 models; (b) Ward geodetic8.5, Ward geologic, Ward seismic, and Ward simulation models. The color version of this figure is available only in the electronic edition.

method have no particular standing, the general approach is potentially very useful. It provides a systematic way of assimilating new elements, data or modeling inputs, into statistical forecasting models, by fitting a few extra parameters for each new element. This is similar to the way in which multiple regression analysis can be used to explain much more of the variation of a response variable than any individual explanatory variable could on its own.

In the present case, the inputs are all RELMs, but the method presented does not depend on that. The inputs could be any gridded data or modeled quantity, such as a strain map, a stress-change map, a binary variable indicating presence or absence of a proposed earthquake precursor, or a categorical variable indicating the degree of an alarm.

Just as multiple regression analysis is more robust when the number of observations is large and the number of explanatory variables is relatively small, so the present method is likely to be more robust when the number of target earthquakes is large and the number of input elements is relatively small. In any case, the results from retrospective fitting need to be confirmed by further prospective tests. All of the two-model hybrids and the best three-model hybrids derived here

Table 2
Parameter Estimates and Other Statistics of Best Three-Model Multiplicative Hybrids

	Whole of California	Southern California
Baseline (model 1)	Helmstetter *et al.* HKJ	Helmstetter *et al.* HKJ
Conjugate (model 2)	Bird and Liu Neokinema	Shen *et al.* geodetic
Conjugate (model 3)	Holliday *et al.* PI	Holliday *et al.* PI
a	−6.49	−10.38
b_2	7.77	19.72
c_2	0.094	0.230
b_3	16.70	8.64
c_3	0.764	0.188
N	31	22
$\Delta \ln L$	6.8	7.94
IGPEc	0.35	0.79
Std Error (IGPEc)	0.17	0.27
ΔIGPEc	0.092	0.22
Std Error (ΔIGPEc)	0.094	0.14

N, number of target earthquakes; $\Delta \ln L$, increase in log likelihood compared to best two-model hybrid; IGPEc, information gain per earthquake relative to baseline model; ΔIGPEc, information gain per earthquake relative to best two-model hybrid.

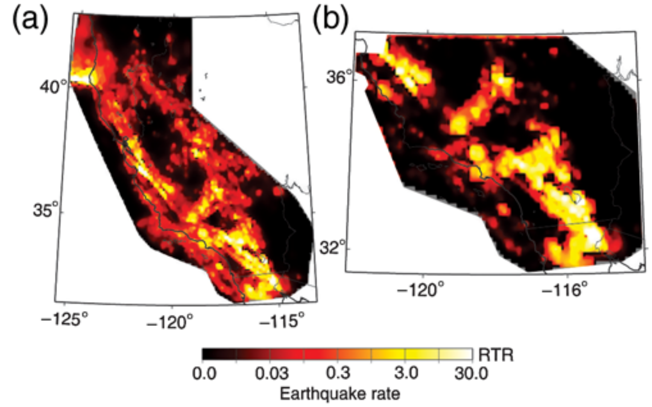

Figure 8. Map of earthquake rates, RTR, in the best three-model hybrids for (a) the whole of California: a hybrid of Helmstetter *et al.* HKJ, Bird and Liu Neokinema, and Holliday *et al.* PI; and (b) southern California: a hybrid of Helmstetter *et al.* HKJ, Shen *et al.* geodetic, and Holliday *et al.* PI. In the reference model, one earthquake per year is expected to exceed any magnitude m in an area of 10^m km^2. The color version of this figure is available only in the electronic edition.

have been submitted to the SCEC CSEP testing center for evaluation over a further five-year test period from 2011 to 2015.

These prospective tests also have limitations, because the forecasts were originally designed for a particular five-year period. They may have varying degrees of applicability to other periods, depending on the degree of time dependence in the data used to generate them. Ideally, a new prospective test of the hybrid models would use updated versions of the individual forecasts, produced by exactly the same methods as the original forecasts. Such updates can only be produced by the authors of the models.

There are many questions still to be answered by more widespread application of multiplicative hybrids. As well as the issue of how well retrospectively fitted hybrids will perform in prospective testing, there are questions of whether it is always advantageous to use the best individual model as the baseline, and whether a hybrid not including the best model might sometimes outperform those including it. Limited additional experimentation with the models in the RELM experiment, not described in detail here, has shown that for two-model hybrids for the whole of California, it is advantageous to use the best model as the baseline, and that all two-model hybrids that do not include the HKJ model have a negative corrected information gain over the HKJ model by itself.

Conclusions

The original aim of the RELM five-year experiment, and of subsequent CSEP experiments to date, was to prospectively test individual models for consistency with the target earthquakes occurring in a test region and to compare the performance of each model with that of other available models. However, an equally important aim, especially for operational earthquake forecasting, is to combine the available forecasting models and other relevant data to form the most informative model possible.

For the models in the RELM five-year experiment, multiplicative hybrids have been shown to be potentially more informative than additive hybrids constructed from the same individual models. Although a smoothed seismicity model was the most informative individual model in the RELM five-year experiment, other models, such as those making use of geodetic data and analysis of seismicity fluctuations, may have important contributions to make when constructing an improved forecast. Prospective testing of the multiplicative hybrid models derived here for a further period of five years or more will give an indication of the robustness of the fitting of the multiplicative hybrid models.

The technique of forming multiplicative hybrids described here will be useful for assimilating new and diverse earthquake-related datasets into forecasting models and for combining models from CSEP forecasting experiments at all timescales. Its routine application to other models and datasets already available, with priority given to combining models and data carrying different kinds of information, would give an indication of the information gains that are obtainable now.

Data and Resources

The Advanced National Seismic System catalog was obtained by a request to the Southern California Earthquake Center Collaboratory for the Study of Earthquake Probability Testing Center. Some plots were made using the Generic Mapping Tools version 4.2.1 (Wessel and Smith, 1998). The software used to optimize the multiplicative hybrid models is available on request from the first author, or from the SCEC CSEP testing center.

Acknowledgments

D. Harte, M. Stirling, P. Bird, and an anonymous reviewer provided helpful reviews of an earlier version of this manuscript. Maria Liukis aided this research through data processing and developing and running the Collaboratory for the Study of Earthquake Probability (CSEP) testing center software. This research was supported by the Southern California Earthquake Center (SCEC), which is funded under National Science Foundation (NSF) Cooperative Agreement EAR-1033462 and U.S. Geological Survey (USGS) Cooperative Agreement G12AC20038. The SCEC Contribution Number for this paper is 1837. D. A. R., M. C. G., and A. C. were supported by a grant from the New Zealand Natural Hazards Research Platform. D. S. was supported by the Global Earthquake Model project. J. D. Z. was supported by the REAKT Project via the Seventh Framework Program for Research (FP7), Contract Number 282862.

References

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Trans. Automat. Contr.* **AC-19,** 716–723.

Aki, K. (1981). A probabilistic synthesis of precursory phenomena, in *Earthquake Prediction*, D. W. Simpson and P. G. Richards (Editors), American Geophysical Union, 566–574.

Bird, P., and Z. Liu (2007). Seismic hazard inferred from tectonics: California, *Seismol. Res. Lett.* **78,** 37–48, doi: 10.1785/gssrl.78.1.37.

Ebel, J. E., D. W. Chambers, A. L. Kafka, and J. A. Baglivo (2007). Non-Poissonian earthquake clustering and the hidden Markov model as bases for earthquake forecasting in California, *Seismol. Res. Lett.* **78,** 57–65, doi: 10.1785/gssrl.78.1.57.

Faenza, L., and W. Marzocchi (2010). The proportional hazard model as applied to the CSEP forecasting area in Italy, *Ann. Geophys.* **53,** no. 3, 77–84, doi: 10.4401/ag-4759.

Field, E. H. (2007). Overview on the working group for the development of Regional Earthquake Likelihood Models (RELM), *Seismol. Res. Lett.* **78,** 7–16, doi: 10.1785/gssrl.78.1.7.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2007). High-resolution time-independent grid-based forecast for $M \leq 5$ earthquakes in California, *Seismol. Res. Lett.* **78,** 78–86, doi: 10.1785/gssrl.78.1.78.

Holliday, J. R., C. C. Chen, K. F. Tiampo, J. B. Rundle, D. L. Turcotte, and A. Donnellan (2007). A RELM earthquake forecast based on pattern informatics, *Seismol. Res. Lett.* **78,** 87–93, doi: 10.1785/gssrl.78.1.87.

Hurvich, C. M., and C.-L. Tsai (1989). Regression and time series model selection in small samples, *Biometrika* **76,** 297–307.

Imoto, M. (2006). Earthquake probability based on multidisciplinary observations with correlations, *Earth Planets Space* **58,** no. 11, 1447–1454.

Imoto, M. (2007). Information gain of a model based on multidisciplinary observations with correlations, *J. Geophys. Res.* **112,** no. 5, B05306, doi: 10.1029/2006JB004662.

Kagan, Y. Y., D. D. Jackson, and Y. Rong (2007). A testable five-year forecast of moderate and large earthquakes in southern California based on smoothed seismicity, *Seismol. Res. Lett.* **78,** 94–98, doi: 10.1785/gssrl.78.1.94.

Marzocchi, W., J. D. Zechar, and T. H. Jordan (2012). Bayesian forecast evaluation and ensemble earthquake forecasting, *Bull. Seismol Soc. Am.* **102,** 2574–2584, doi: 10.1785/0120110327.

McCullagh, P. and J. A. Nelder (Editors) (1989). *Generalized Linear Models*, Second Ed., Chapman and Hall/CRC, Boca Raton, Florida, ISBN: 0-412-31760-5.

Nelder, J. A., and R. Mead (1965). A simplex method for function minimization, *Comput. J.* **7,** 308–313.

Rhoades, D. A. (2013). Mixture models for improved earthquake forecasting with short-to-medium time-horizons, *Bull. Seismol Soc. Am.* **103,** no. 4, 2203–2215, doi: 10.1785/0120120233.

Rhoades, D. A., and F. F. Evison (2004). Long-range earthquake forecasting with every earthquake a precursor according to scale, *Pure Appl. Geophys.* **161,** 47–71, doi: 10.1007/s00024-003-2434-9.

Rhoades, D. A., and M. C. Gerstenberger (2009). Mixture models for improved short-term earthquake forecasting, *Bull. Seismol. Soc. Am.* **99,** 636–646, doi: 10.1785/0120080063.

Rhoades, D., and M. Stirling (2012). An earthquake likelihood model based on proximity to mapped faults and catalogued earthquakes, *Bull. Seismol. Soc. Am.* **102,** no. 4, 1593–1599, doi: 10.1785/0120110326.

Rhoades, D. A., D. Schorlemmer, M. C. Gerstenberger, A. Christophersen, J. D. Zechar, and M. Imoto (2011). Efficient testing of earthquake forecasting models, *Acta Geophys.* **59,** 728–747, doi: 10.2478/s11600-011-0013-5.

Schorlemmer, D., and M. C. Gerstenberger (2007). RELM testing center, *Seismol. Res. Lett.* **78,** 30–36, doi: 10.1785/gssrl.78.1.30.

Schorlemmer, D., M. C. Gerstenberger, S. Wiemer, D. D. Jackson, and D. A. Rhoades (2007). Earthquake likelihood model testing, *Seismol. Res. Lett.* **78,** 17–29, doi: 10.1785/gssrl.78.1.17.

Schorlemmer, D., J. D. Zechar, M. J. Werner, E. H. Field, D. D. Jackson, T. H. Jordan, and The RELM Working Group (2010). First results of the Regional Earthquake Likelihood Models experiment, *Pure Appl. Geophys.* **167,** 859–876, doi: 10.1007/s00024-010-0081-5.

Shebalin, P. N., C. Narteau, J. D. Zechar, and M. Holschneider (2014). Combining earthquake forecasts using differential probability gains, *Earth Planet Space* **66,** no. 1, article number 35.

Shen, Z.-K., D. D. Jackson, and Y. Y. Kagan (2007). Implications of geodetic strain rate for future earthquakes, with a five-year forecast of M 5 earthquakes in southern California, *Seismol. Res. Lett.* **78,** 116–120, doi: 10.1785/gssrl.78.1.116.

Utsu, T. (1982). Probabilities in earthquake prediction (the second paper), *Bull. Earthq. Res. Inst.* **57,** 499–524 (in Japanese).

Wang, T., J. Zhuang, T. Kato, and M. Bebbington (2013). Assessing the potential improvement in short-term earthquake forecasts fromincorporation of GPS data, *Geophys. Res. Lett.* **40,** 2631–2635, doi: 10.1002/grl.50554.

Ward, S. N. (2007). Methods for evaluating earthquake potential and likelihood in and around California, *Seismol. Res. Lett.* **78,** 121–133, doi: 10.1785/gssrl.78.1.121.

Wessel, P., and W. H. F. Smith (1998). New, improved version of the Generic Mapping Tools released, *Eos Trans. AGU* **79,** 579.

Wiemer, S., and D. Schorlemmer (2007). ALM: An asperity-based likelihood model for California, *Seismol. Res. Lett.* **78,** 134–140, doi: 10.1785/gssrl.78.1.134.

Zechar, J. D., and T. H. Jordan (2008). Testing alarm-based earthquake predictions, *Geophys. J. Int.* **172,** no. 2, 715–724, doi: 10.1111/j.1365-246X.2007.03676.x.

Zechar, J. D., D. Schorlemmer, M. J. Werner, M. C. Gerstenberger, D. A. Rhoades, and T. H. Jordan (2013). Regional Earthquake Likelihood Models I: First-order results, *Bull. Seismol. Soc. Am.* **103,** 787–798, doi: 10.1785/0120120186.

GNS Science
1 Fairway Drive
Lower Hutt 5010
P.O. Box 30-368
Lower Hutt 5040, New Zealand
d.rhoades@gns.cri.nz
m.gerstenberger@gns.cri.nz
a.christophersen@gns.cri.nz
(D.A.R., M.C.G., A.C.)

Swiss Seismological Service
ETH Zurich, Sonneggstrasse 5
8092 Zürich, Switzerland
jeremy.zechar@sed.ethz.ch
(J.D.Z.)

Helmholtz Centre Potsdam-GFZ German Research Centre for Geosciences
Telegrafenberg
14473 Potsdam, Germany
ds@gfz-potsdam.de
(D.S.)

School of Earth Sciences and Cabot Institute
University of Bristol
Wills Memorial Building, Queen's Road
Bristol BS8 1RJ, United Kingdom
max.werner@bristol.ac.uk
(M.J.W.)

Southern California Earthquake Center
University of Southern California
3651 Trousdale Parkway
MC-0740, Los Angeles
California 90089-0740
jordan@usc.edu
(T.H.J.)

Manuscript received 4 February 2014;
Published Online 30 September 2014