A Decade of Prospective Evaluations of One-Day Seismicity Forecasts for California: First Results

Toño Bayona[1] (jose.bayona@bristol.ac.uk), Marcus Herrmann[2], William Savran[3], Fabio Silva[4], Philip Maechling[4], Warner Marzocchi[2] and Max Werner[1]

[1] School of Earth Sciences, University of Bristol, U. K., [2] University of Naples Federico II, [3] Nevada Seismological Observatory, University of Nevada, U.S.A., [4] Southern California Earthquake Center, University of Southern California, U.S.A.

Introduction

Since 2007, the Collaboratory for the Study of Earthquake Predictability (CSEP) has been prospectively evaluating one-day earthquake forecasting models for California with the goal of addressing seismological questions with important implications for time-dependent seismic hazards. Among others, the pool of 24 models includes various flavours of ETAS, STEP, non parametric models, and ensembles models (see Fig. 1). Here, we present present preliminary test results for 11 of these models using M3.95+ earthquakes observed in California from August 1, 2007 through June 30, 2018 (see Fig. 3). In addition, we compare the performance of these models with that of the time-independent smoothed seismicity HKJ model of Helmstetter et al. (2007). Thus, this unprecedentedly large dataset provides a unique opportunity to assess our ability to forecast earthquake clustering and advance operational earthquake forecasting in California.

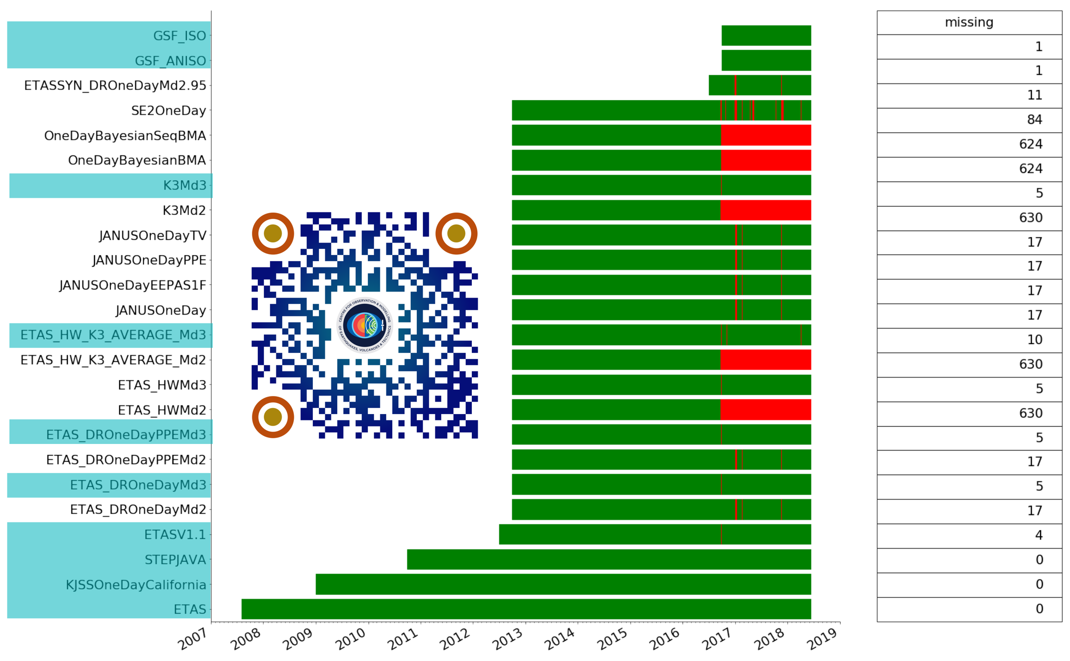


Fig. 1. Predictive pool of 24 one-day seismicity models forecasting M3.95+ earthquakes in California. In green, we show the daily forecasts that are available during the evaluation period and, in red, we show the opposite. In cyan, we highlight 11 models for which we present prospective test results here.

One-Day Seismicity Forecasts

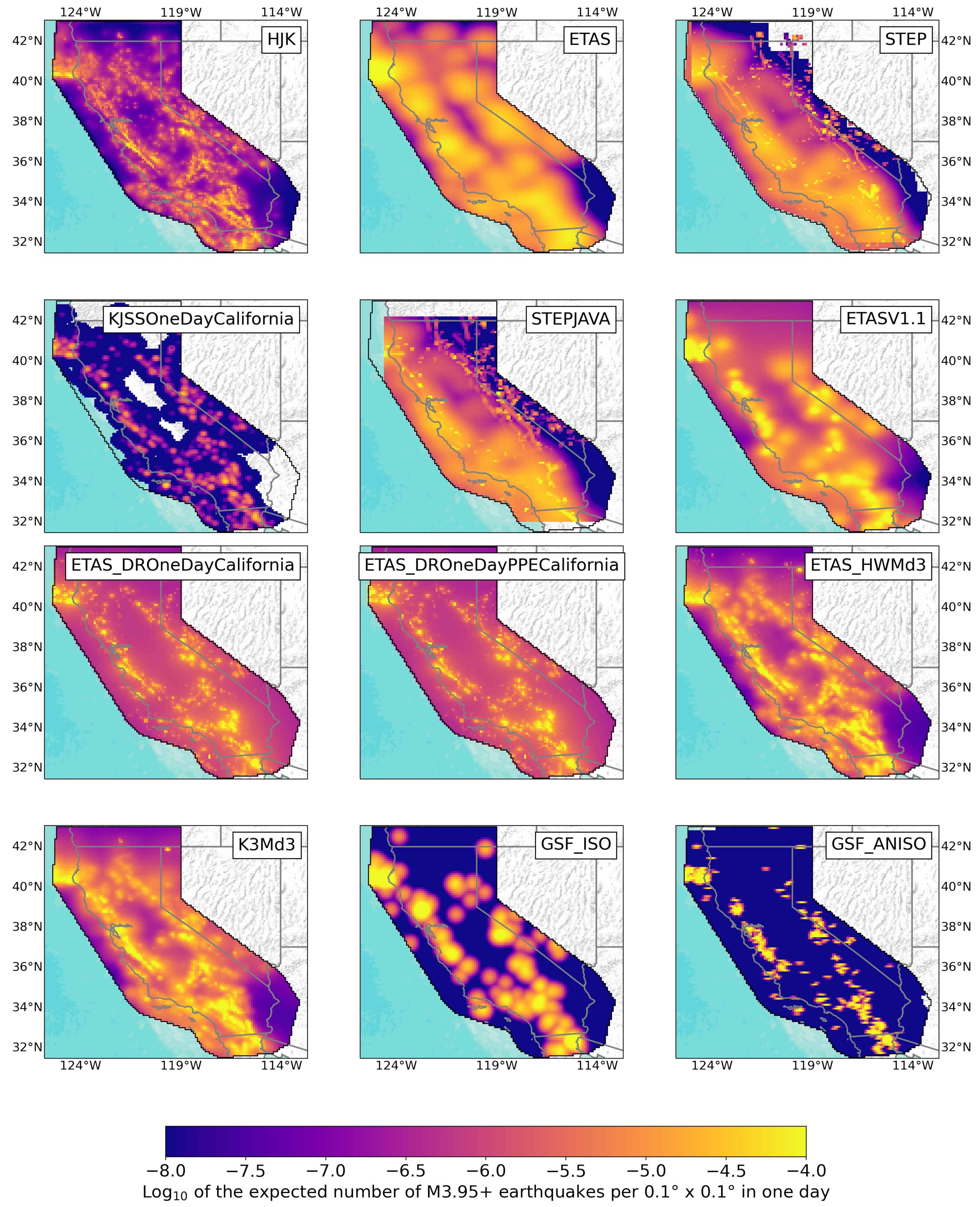
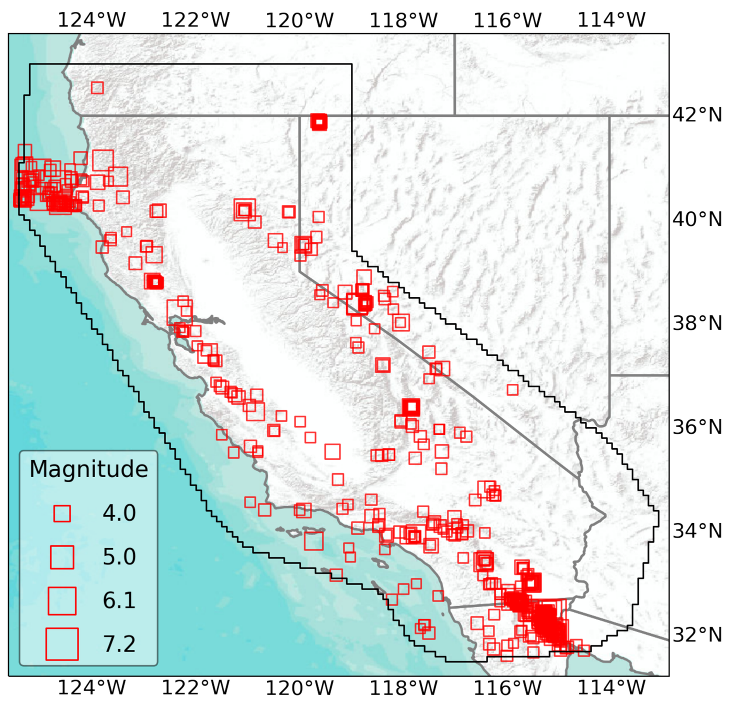


Fig. 2. Forecast maps provided by one-day time-dependent seismicity models for California. M3.95+ earthquake rate densities are expressed per 0.1° x 0.1° unit cell per day. Yellow and orange colours denote regions where expected earthquake rates are comparatively high, while blue-purple colours denote the opposite. With the exception of STEP, these forecasts were issued to forecast earthquakes in California on June 30, 2018 (the STEP forecast was issued on August 1, 2007).

Methods

We assess consistency between the expected and observed number and spatial cumulative distributions of M3.95+ earthquakes. For this purpose, we use the number (N) test of Zechar et al. (2010) and the spatial binary (S) test of Bayona et al. (2022). These consistency tests calculate the probability of observing the data under the models and were recently implemented in the so-called pyCSEP software toolkit (Savran et al., 2022). We also compare the performance of the one-day models with that of the HKJ seismicity model by means of information gains per activated bin (see Bayona et al., 2022). We select HKJ as our benchmark model, because after 15 years of prospective testing, it has been shown to be the most informative time-invariant CSEP earthquake forecasting model in California.

Evaluation Dataset



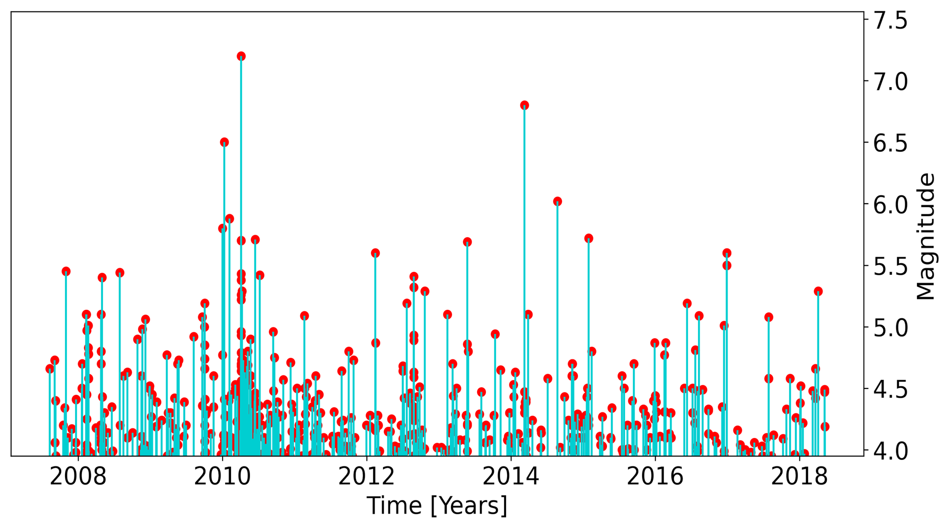


Fig. 3. a) Spatial distribution of M3.95+, d≤ 30 km earthquakes observed in California during the prospective evaluation period. This testing dataset contains nearly 600 target events, including the 2010 M7.2 el Mayor-Cucapah, 2014 M6.9 Cape Mendocino, and 2014 M6 South Napa earthquakes. b) Time vs magnitude plot of the prospective test dataset used in this study.

Consistency Test Results

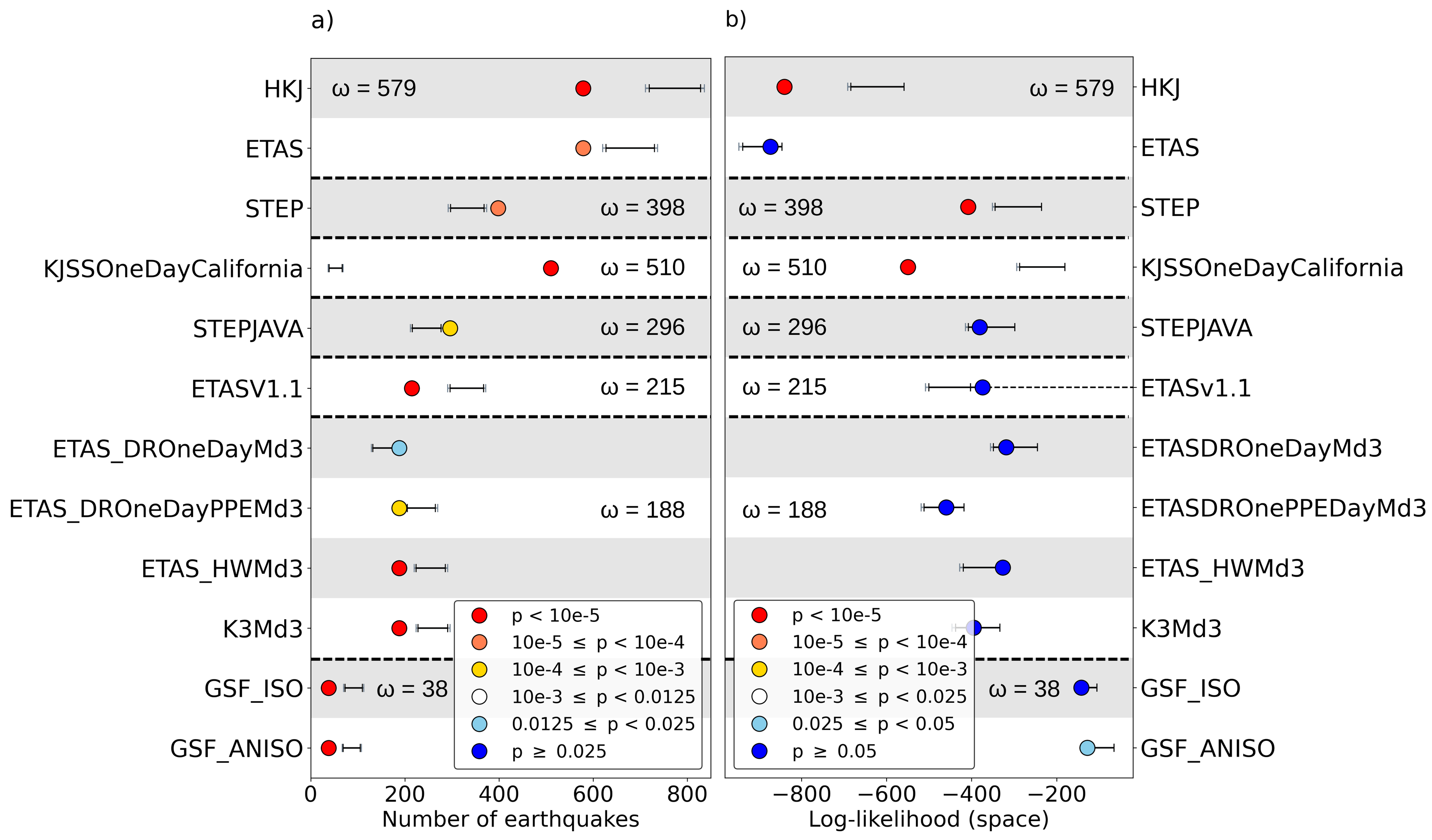


Fig. 4. Results of the cumulative N-test during the evaluation period. Circles represent the number of observed target earthquakes ω, and the colours denote the *p*-values calculated for the earthquake forecasting models. Blue colours indicate consistency between forecasts and observations and red-orange colours denote the opposite. Solid black and gray bars depict the 95% and 97.5% predictive intervals of the model's forecast distribution. b) Prospective results of the cumulative binary S-test for one-day seismicity models for California. Circles represent the cumulative observed joint probability of observing the data under the models. Colours and solid bars represent the same as in Fig. 4a. Horizontal dashed lines differenciate the number of earthquakes used to test each model.

Comparative Test Results

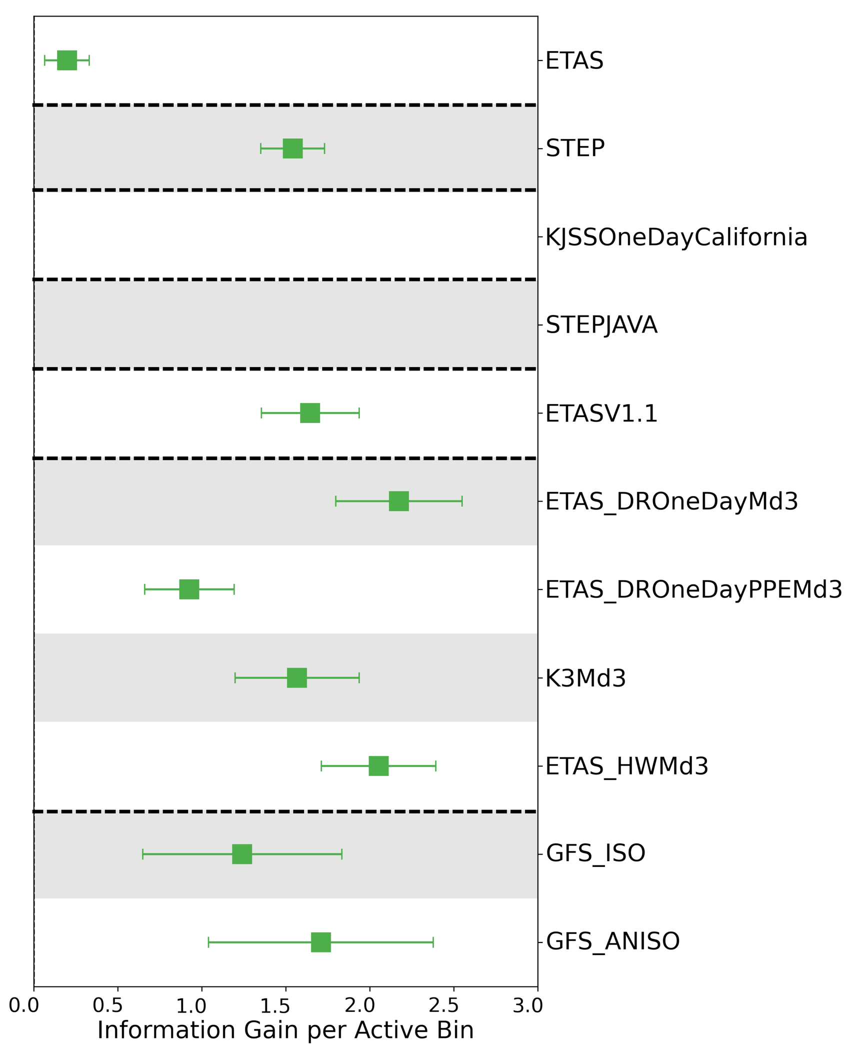


Fig. 3. Prospective T-test results for one-day seismicity models for California. We show Information Gains per Active Cell (IGPA) obtained by eleven time-varying models over HKJ, along with their calculated 95% confidence intervals. The green squares denote that the models that are statistically more informative than HKJ. Note that no IGPA values are displayed for STEPJAVA and KJSSOneDayCalifornia, because these models provide rates of 0 earthquakes in cells where seismicity has been observed during the evaluation period (see Fig. 2).

Conclusions

- Most models tend to overestimate the number of earthquakes in times of seismic "quiesence", while they underestimate seismicity during seismic sequences (see Fig. 6).

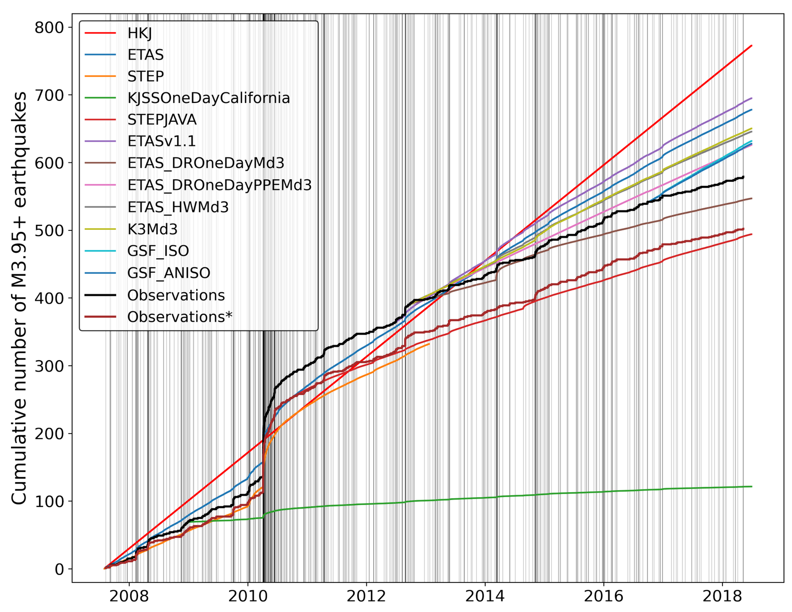


Fig. 6. Cumulative distribution of expected and observed M3.95+ earthquakes in California during the prospective evaluation period. In black, we show the cumulative number of earthquakes within the entire CSEP-California testing region, while in brown, we show cumulative target seismicity within the STEPJAVA testing region (see Fig. 2).

- Most models can adequately forecast the spatial distribution of observed quakes, especially in periods of seismic calm.

- The ensemble ETAS_DROneDayMd3 and ETAS_HWMd3 are the most informative, obtaining IGPAs of about 2.0 over HKJ.

- In the future, we will assess and compare the rest of the forecast models and study the variability of the test results over time.

References

Bayona et al. (2022); https://doi.org/10.1093/gji/ggac018
Savran et al. (2022); http://doi.org/10.21105/joss.03658
Zechar et al. (2010); https://doi.org/10.1785/0120090192