INTRODUCING SPATIAL HETEROGENEITY IN SEISMIC FORECASTS

Gina-Maria Geffers[1], **Christopher S. Sherman**[1], **Chaoyi Wang**[1] & **Kayla. A. Kroll**[1]
[1]Lawrence Livermore National Laboratory, Livermore, CA

NRAP
National Risk Assessment Partnership

BACKGROUND

Using the Illinois Basin Decatur Project (IBDP) dataset, we run a short-term forecast based on portions of the data to evaluate the performance of spatial areas (grid cells) where there is no seismicity and/or pressure data (i.e. Fig. 1) and how to extrapolate for these common occurrences.

We specifically look at the *seismogenic index (SI) model* – a statistical model defined entirely by seismotectonic features of a particular region[1]. A modified version includes use of the pressure data.[2]

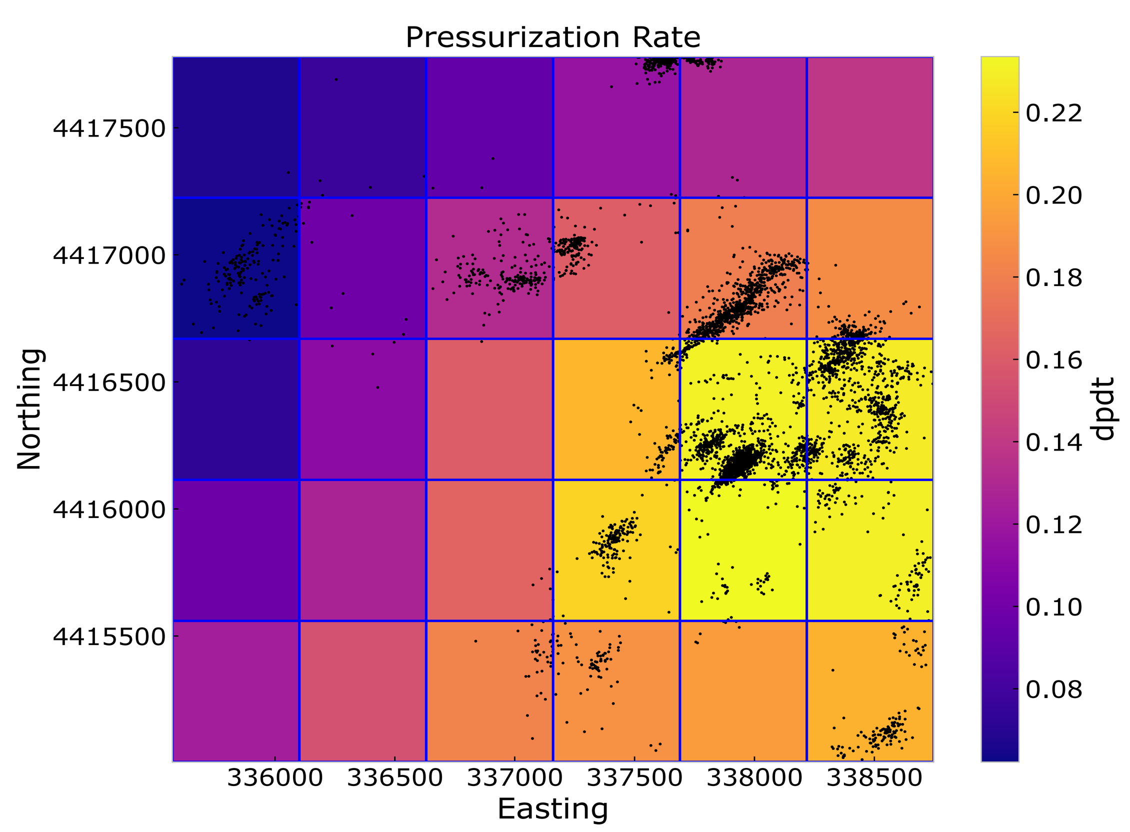


Fig 1. Spatial grid of the pressurization rate for the IBDP dataset. Black circles are locations of the earthquakes recorded in this area. Notably, there are multiple cells in which no seismicity occurs but the pressurization rate is non-zero.

METHOD

SI is calculated for every grid cell containing more than one event. We investigate the following methods for computing SI in cells with no events:

- leave empty cells as they are
- use the nearest neighbor cell's SI
- use the average of all existing SIs for the empty cells

For each of these, we run four scenarios:

- 25% of the data to train, forecasting the remaining 75%
- 50% of the data to train, forecasting the remaining 50%
- 75% of the data to train, forecasting the remaining 25%
- 100% of the data to train, forecasting the full 100% (complete overlap)

REFERENCES
1 Shapiro, S. A., Dinske, C. & Langenbruch, C. (2010), Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations, *Lead. Edge*, ***29***, *304-309*
2. Langenbruch, C., Weingarten, M. & Zoback, M. D. (2018), Physics-based forecasting of man-made earthquake hazards in Oklahoma and Kansas, *Nature comms.*, ***9***(3946)

Investigating impact of grid size on forecast

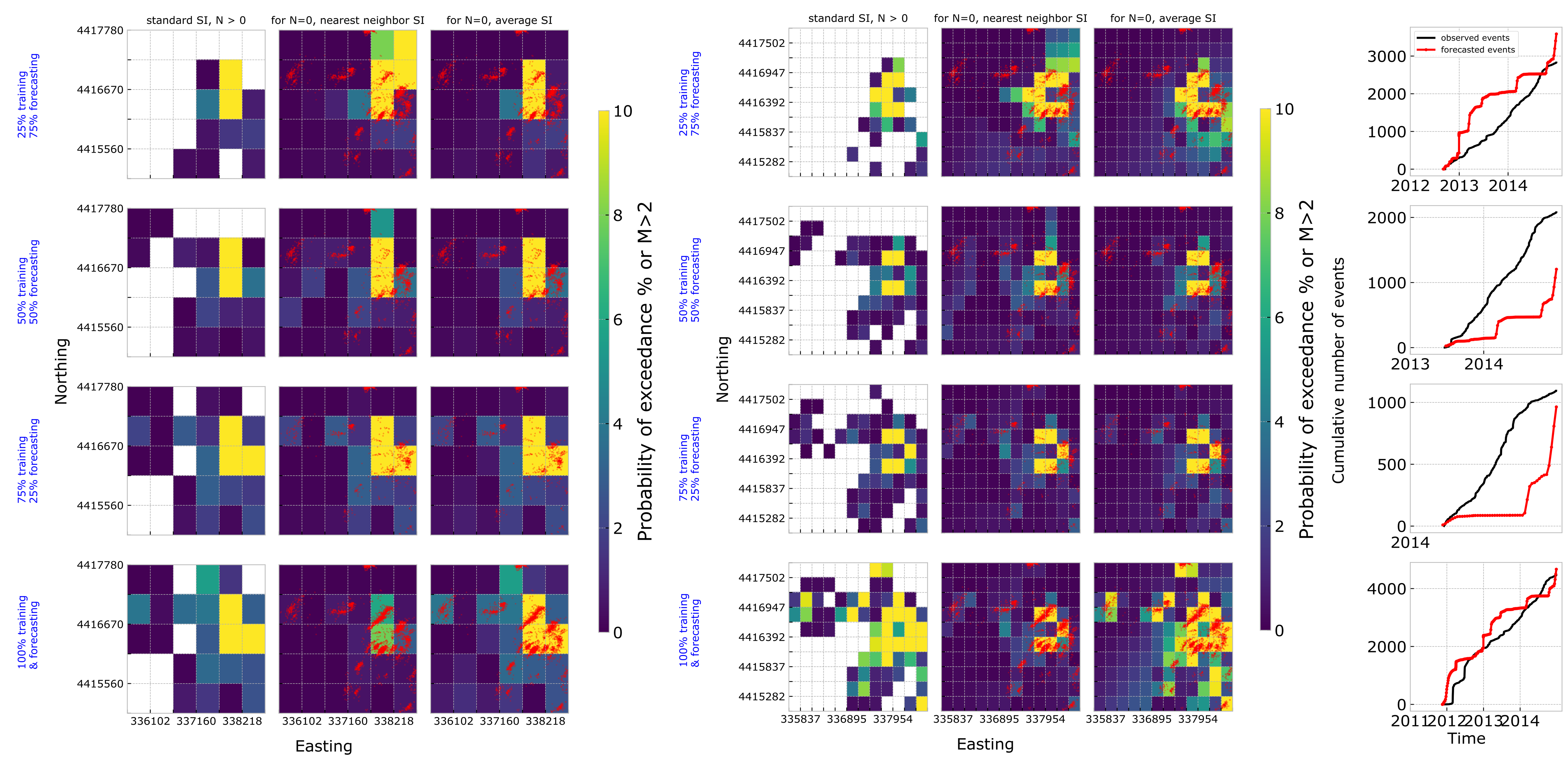


Fig 2. Spatial grid of the forecasted probability of exceedance for an M > 2 event for the coarse grid **(left)** as well as a finer grid **(center).** In each row, a different portion of the data is used to train and forecast the probability. Red circles indicate the events that are observed in the duration of the forecasted period. **Right:** The cumulative number of events over time for the observed events (black) and forecasted events (red) for the respective training to forecasting time period, over the entire region (using the average SI in empty cells).

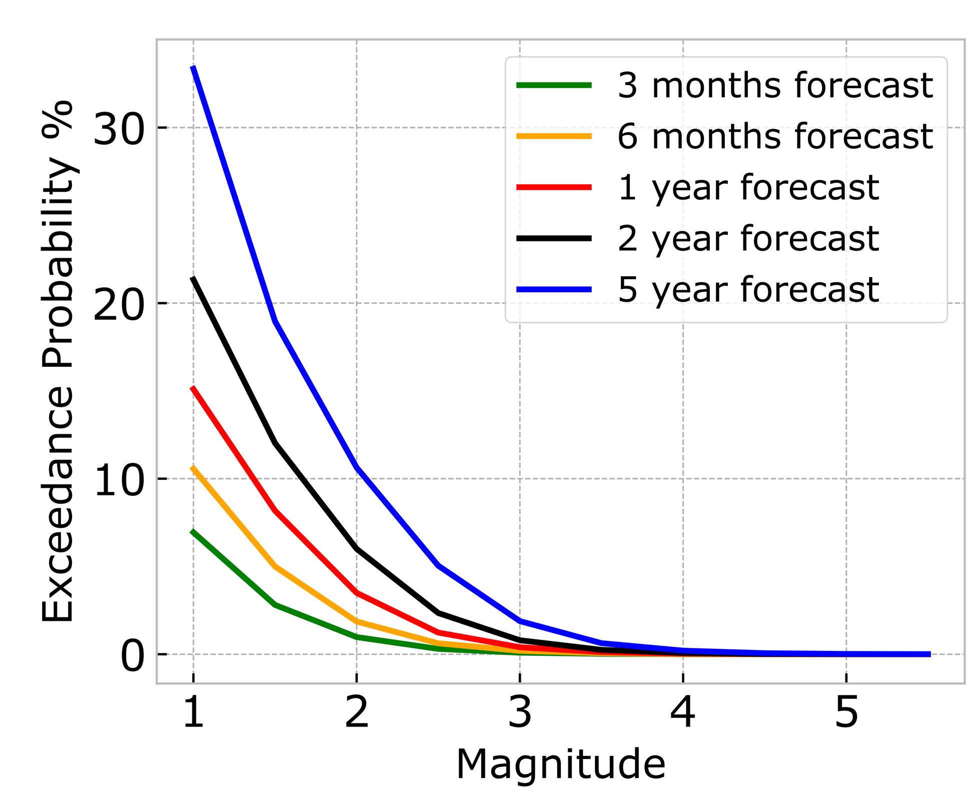
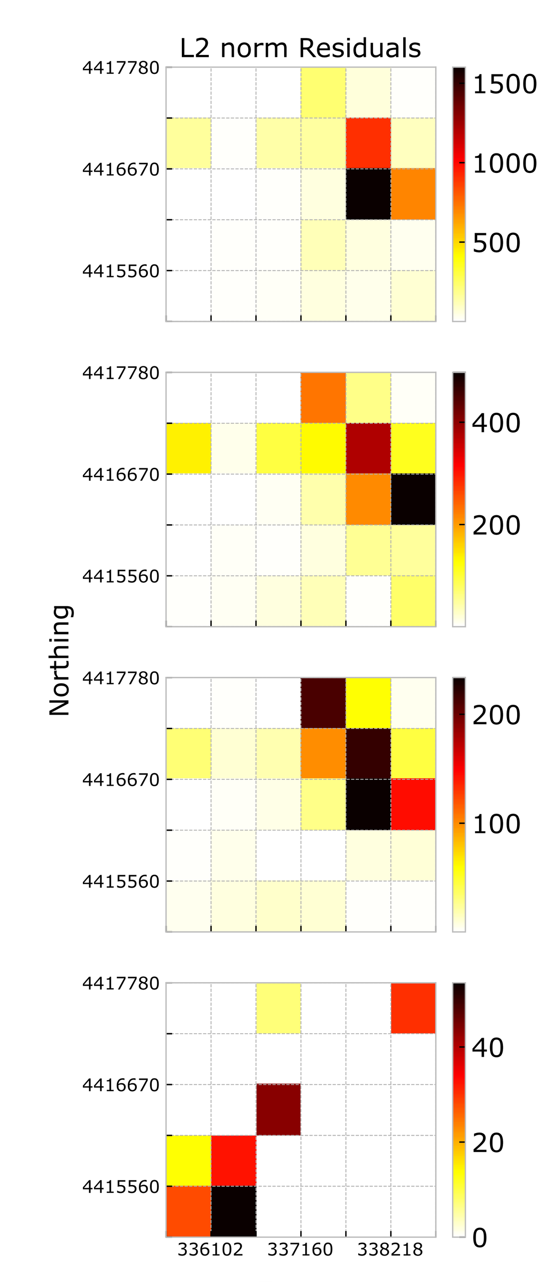


Fig 3. Probability of exceeding a M > 2 event for different forecast duration from 3 months to 5 years, using the average SI in empty cells.

Fig 4. The L2 norm residuals between the observed and forecasted number of events, using average SI in empty cells, for the same scenarios as in Fig. 2



SUMMARY

- Using the average of all existing SIs leads to better results than using the nearest neighbor SI
- if event clusters are only located within the training period, the forecast will then overpredict the number of events in these grid cells

LLNL-POST- 853469 This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

U.S. DEPARTMENT OF ENERGY
NATIONAL ENERGY TECHNOLOGY LABORATORY
Pacific Northwest NATIONAL LABORATORY
Lawrence Livermore National Laboratory
Los Alamos NATIONAL LABORATORY
BERKELEY LAB