Nowcasting Earthquakes with QuakeGPT: An AI-Enhanced Earthquake Generative Pretrained Tranformer

John Rundle[1]; Lisa Grant Ludwig[2]; Andrea Donnellan[3]; Geoffrey Fox[4], Ian Baughman[1]

[1]University of California, Davis (jbrundle@ucdavis.edu), [2]University of California, Irvine, [3] Purdue University, [4]University of Virginia, Charlottesville

1. Abstract

Our recent work on earthquake nowcasting has been concerned with the development of methods to track the time dependent state of earthquake risk using earthquake catalog data and standard machine learning techniques. We show the current state of these nowcasting calculations as they pertain to California. We also present a new approach to earthquake nowcasting based on science transformers (GC Fox et al., Geohazards, 2022). As explained in the seminal paper by Vaswani et al. (NIPS, 2017), a transformer is a type of deep learning model that learns the context of a set of time series values by means of tracking the relationships in a sequence of data, such as the words in a sentence. Transformers extend deep learning in the adoption of a context-sensitive protocol "attention", which is used to tag important sequences of data, and to identify relationships between those tagged data. Pretrained transformers are the foundational technology that underpins the new AI models ChatGPT (Generative Pretrained Transformers) from openAI.com, and Bard, from Google.com. In our case, we hypothesize that a transformer might be able to learn the sequence of events leading up to a major earthquake. Typically, the data used to train the model is in the billions or larger, so these models, when applied to earthquake problems, need the size of data sets that only long numerical earthquake simulations can provide. In this research, we are developing the Earthquake Generative Pretrained Transformer model, "QuakeGPT", in a similar vein. For simulations, we are using simulation catalogs from a stochastic physics-informed earthquake simulation model "ERAS", similar to the more common ETAS models. ERAS has only 2 uncorrelated parameters that are easily retrieved from the observed catalog. In the future, physics-based models such as Virtual Quake model could be used as well. Observed data, which is the data to anticipate with nowcasting, is taken from the USGS online catalog for California. In this talk, we discuss 1) recent results from our earthquake nowcasting machine learning methods; and 2) the architecture of QuakeGPT together with first results.

2. Nowcasting with Machine Learning

In most "earthquake prediction" methods, an "anomaly" preceding a major earthquake is proposed, and a few case studies in which the "anomaly" is observed followed by a major earthquake are offered as evidence of success. The problem with this approach, as pointed out over the years by Yan Kagan and Dave Jackson is that this procedure constitutes only 1 out of 4 possible cases.

Nomenclature for this case is "true positive", (TP). The other possible cases are "true negative" (TN), where no anomaly is observed, and no earthquake occurs; "false positive" (FP), where an anomaly is observed and no earthquake occurs; and "false negative" (FN), where no anomaly is observed but an earthquake is observed. This idea is embedded in the Receiver Operating Characteristic method we use here, in which we use an Exponential Moving Average filter on the monthly seismicity time series.

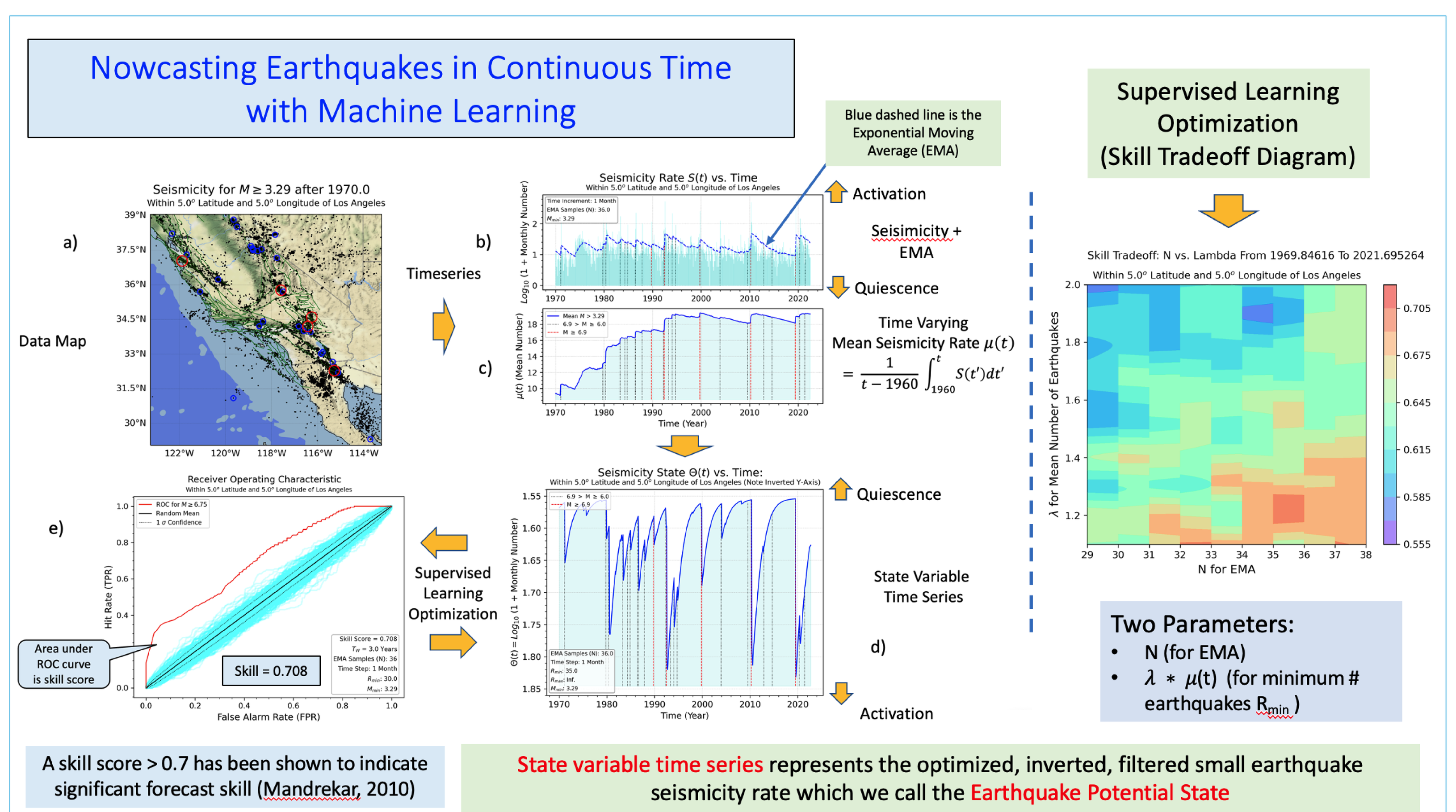

Figure 1. Workflow in which the seismicity is used to build the nowcast time curve as the optimized EMA, using the Receiver Operating Skill score as the loss function.

3. Nowcasting Movies

The nowcasting method has been used to produce movies detailing the current level of hazard in Southern California, due to the recent increase in small earthquake seismic activity. In the southern California video, the recent earthquakes in southern California are coincident with an enhanced spatial and temporal probability of a significant earthquake (M 6.75) in southern California. In the southern California area, the spatial probability density is in the vicinity of the Lamont earthquake, and generally lies along the Garlock fault, stretching from the epicentral region of the 2019 Ridgecrest earthquake, towards the intersection of the Garlock fault with the San Andreas (white box in Figure 3).

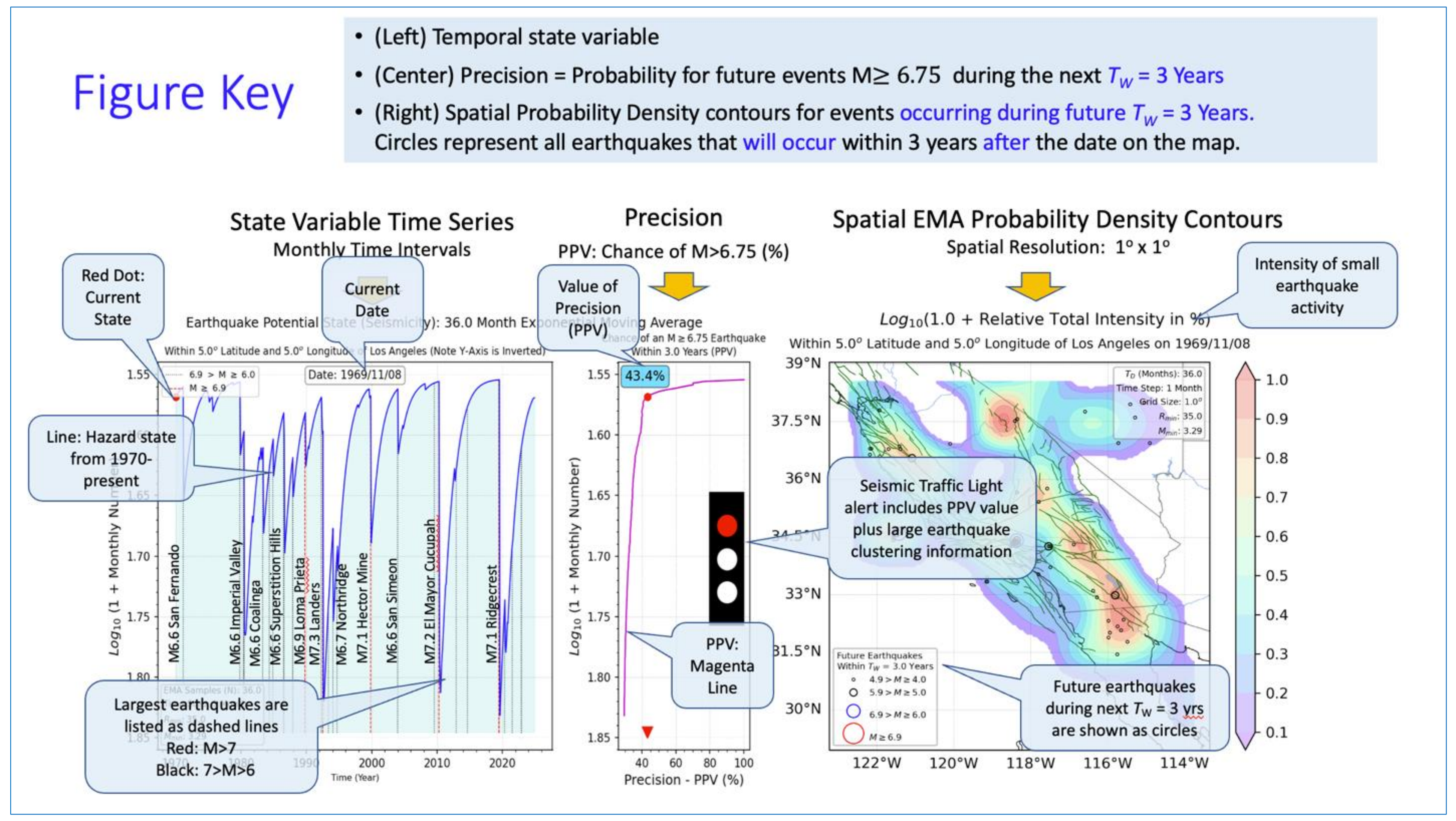

Figure 2: Figure key for a nowcast movie for Southern California. The various parts of the figure are as shown above. The left panel is the nowcast curve. The middle panel converts the nowcast curve into a probability for a M>6.75 earthquake to occur PPV_LogRTI_combined_image_0002147.png3 years after the time indicated on the nowcast curve. The traffic light indicator is the probability with the added condition allowing for earthquake clustering (foreshocks). The right panel shows the spatial contours for earthquakes that are expected to occur within 3 years. Small circles for this panel, which is for 1970, are the earthquakes that actually did occur within the following 3 years.

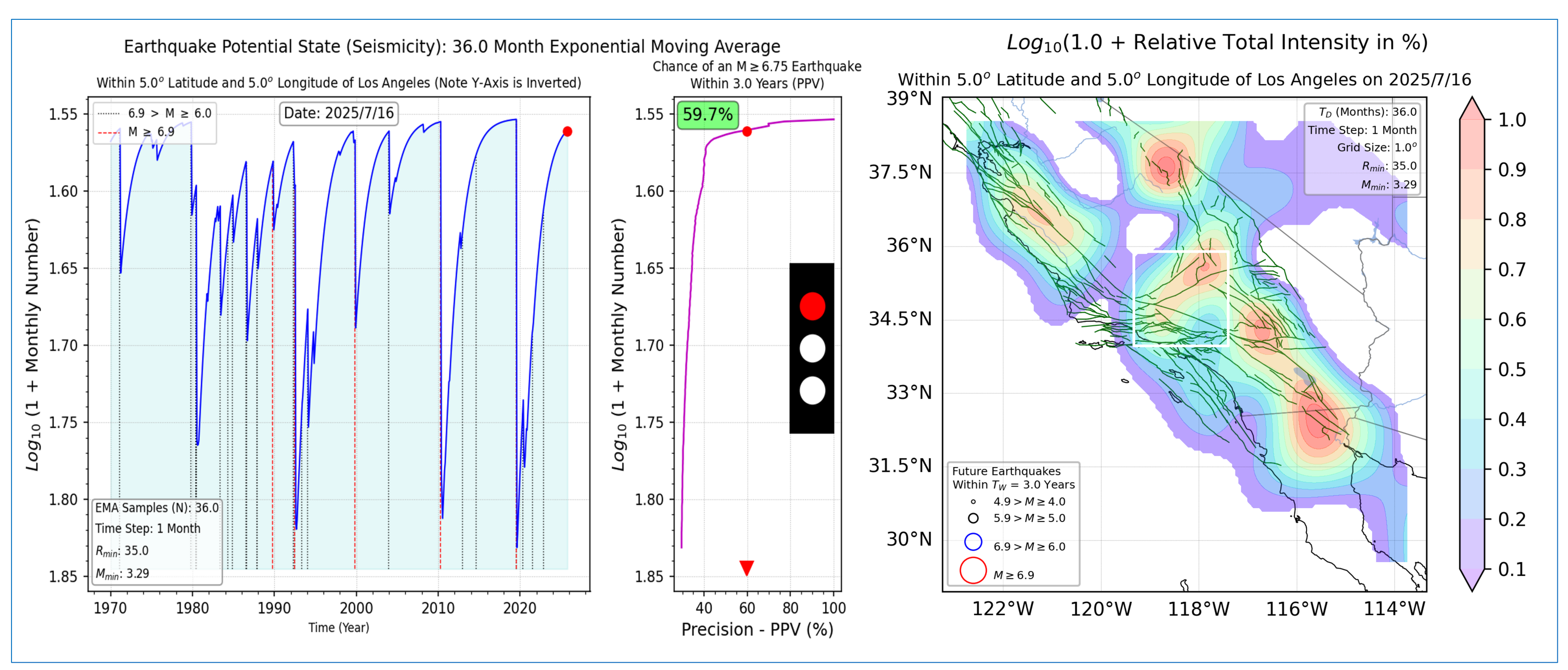

Figure 3: Final frame of a southern California video. Notice in particular the enhanced spatial probability extending from the epicentral area of the 2019 Ridgecrest earthquake in eastern California southwest towards the intersection of the Garlock and White Wolf faults with the San Andreas fault, the general area of the 8/9/2024 Lamont earthquake (circle). Notice also that in the center chart, the chance of a large earthquake is about to increase rapidly.The complete movie can be found at the following link: https://rundle.physics.ucdavis.edu/Movies/SouthernCalifornia.mp4. A movie for the entire state of California can be found at: https://rundle.physics.ucdavis.edu/Movies/California.mp4

4. Nowcasting with QuakeGPT

We have extended the nowcast idea into an AI framework, "QuakeGPT" using the idea of science transformers, which were introduced by Vaswani et al. (2017). Transformers combine deep learning neural networks with the idea of self-attention. This more general framework allows for a model to predict a sequence of data, such as languages, patterns, or in our case, nowcasting time series of data.

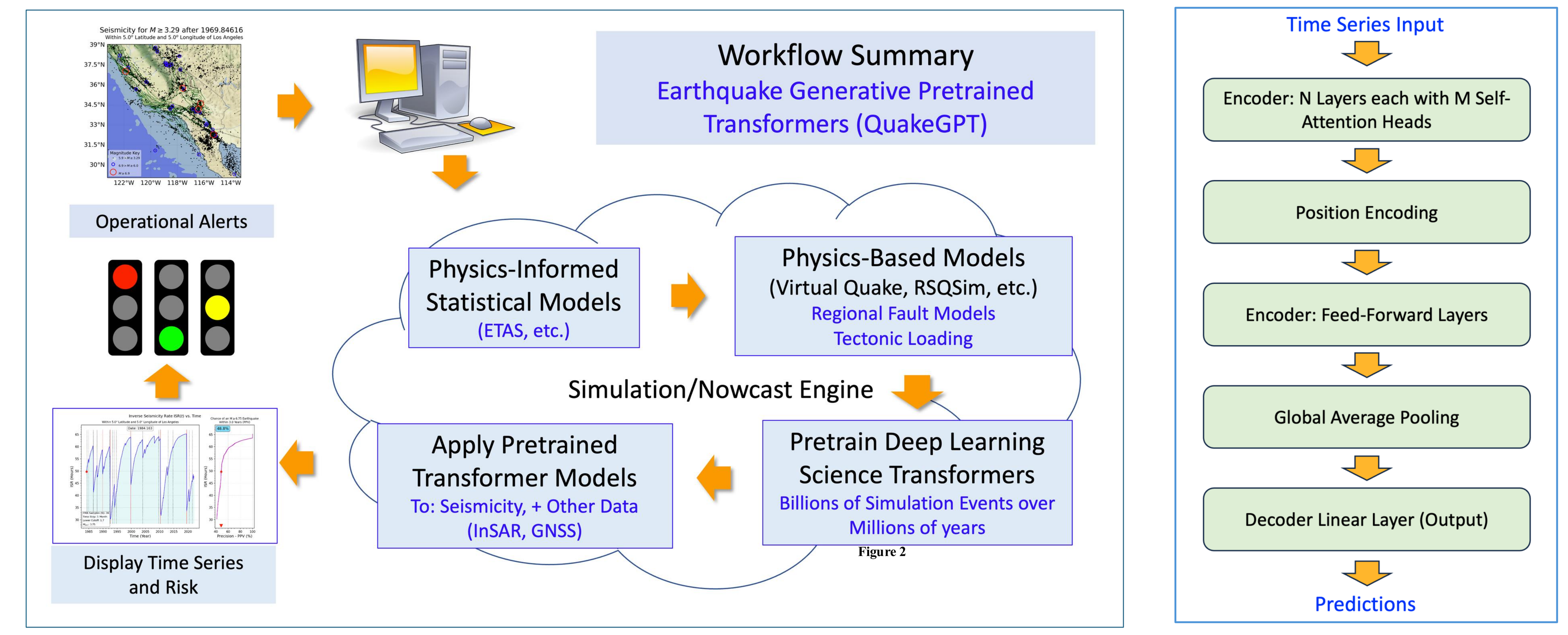

Figure 4: (Left) To instantiate the model, we apply the workflow as shown above to train the model and then apply it to observed data. Data is fed into the model in the form of physics-informed stochastic earthquake simulations (ERAS model), to add to physics-based dynamical fault models (not considered in this current paper). These long-time simulations are then used to pre-train the attention-based science transformer model. The pretrained transformer model is then used to predict the validation data, followed by prediction of values of the unobserved future time series. Potential hazard alerting can be carried out using a traffic-light or other alerting system. (Right) The structure of the transformer is shown, including encoder layers, position encoding, feed-forward layers, global average pooling layer, and an output (decoder) layer.

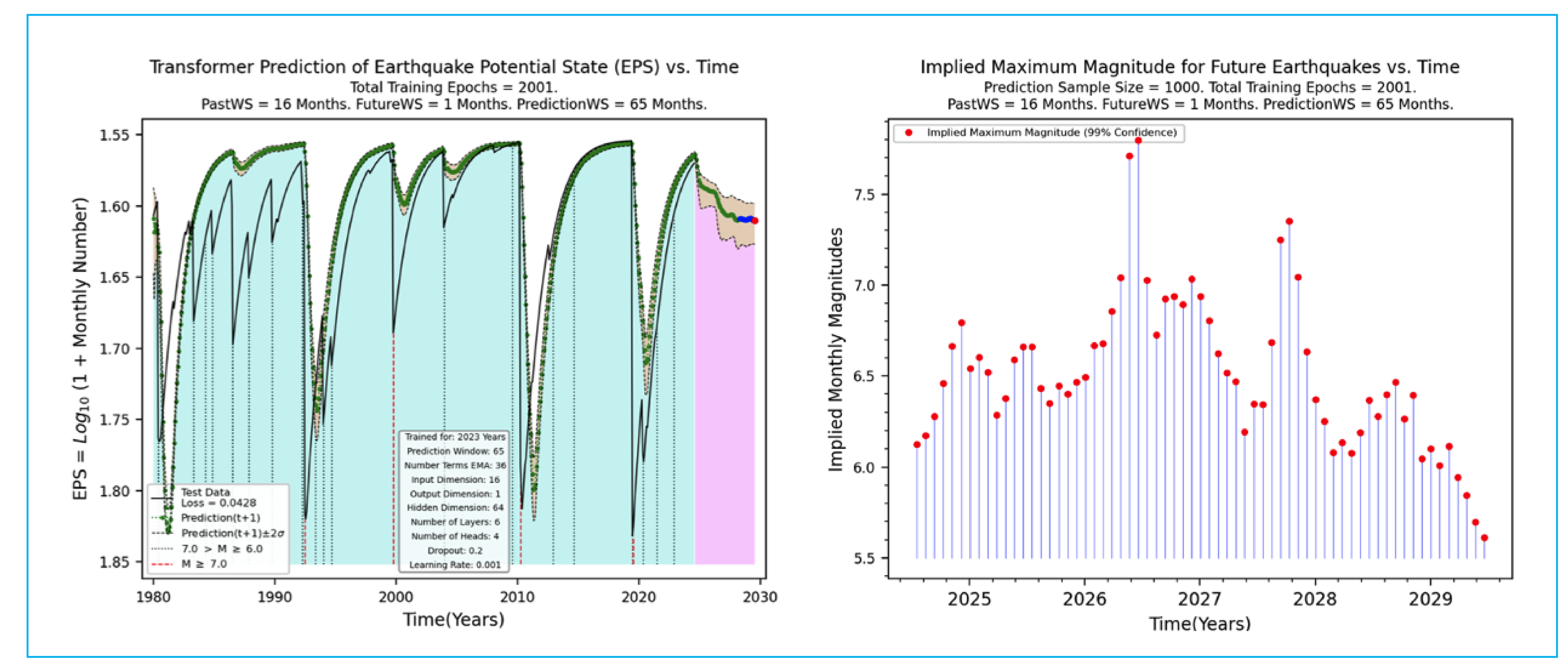

Figure 5: Results obtained by training the transformer model on 2000 years of stochastic simulation data, followed by application to the observed nowcast curve as shown in Figures 1-3 (Left). Simulations were carried out using the ERAS model. The observed nowcast curve can be seen overlaid by the green points, which represent predictions of the transformer model. In the cyan shaded region, the predictions use data from 16 previous points on the nowcast curve to predict the 17th following point. In the magenta region, which is the prediction beyond the current observations, 16 previous predictions are fed back into the model to predict the next point. The brown shaded region indicates the standard error. (Right) The transformer model is then used to estimate the maximum magnitude earthquake that might occur in the future. This magnitude is computed by identifying the earthquake magnitude from the training data that is found by comparing the magenta predicted nowcast with the training data.

References

Funding support was also provided by a generous gift from John Labrecque to UC Davis