UC Irvine
UC Irvine Previously Published Works

Title

Optimizing Earthquake Nowcasting With Machine Learning: The Role of Strain Hardening in the Earthquake Cycle

Permalink

Journal

ISSN

Authors

Rundle, John B
Yazbeck, Joe
Donnellan, Andrea
et al.

Publication Date

DOI

Copyright Information

Earth and Space Science

Key Points:
- "Chaotic" seismicity contains hidden structure in the form of state variable time series
- Standard data science methods can be used to convert the time series to probabilities
- Both temporal and spatial probabilities can be computed

Correspondence to:
J. B. Rundle,
jbrundle@ucdavis.edu

Optimizing Earthquake Nowcasting With Machine Learning: The Role of Strain Hardening in the Earthquake Cycle

John B. Rundle[1,2,3,4], Joe Yazbeck[1], Andrea Donnellan[5], Geoffrey Fox[6], Lisa Grant Ludwig[4], Michael Heflin[5], and James Crutchfield[1]

[1]Department of Physics, University of California, Davis, CA, USA, [2]Santa Fe Institute, Santa Fe, NM, USA, [3]Department of Earth and Planetary Science, University of California, Davis, CA, USA, [4]Program in Public Health, University of California, Irvine, CA, USA, [5]Jet Propulsion Laboratory California Institute of Technology, Pasadena, CA, USA, [6]University of Virginia, Charlottesville, VA, USA

Abstract Nowcasting is a term originating from economics, finance, and meteorology. It refers to the process of determining the uncertain state of the economy, markets or the weather at the current time by indirect means. In this paper, we describe a simple two-parameter data analysis that reveals hidden order in otherwise seemingly chaotic earthquake seismicity. One of these parameters relates to a mechanism of seismic quiescence arising from the physics of strain-hardening of the crust prior to major events. We observe an earthquake cycle associated with major earthquakes in California, similar to what has long been postulated. An estimate of the earthquake hazard revealed by this state variable time series can be optimized by the use of machine learning in the form of the Receiver Operating Characteristic skill score. The ROC skill is used here as a loss function in a supervised learning mode. Our analysis is conducted in the region of 5° × 5° in latitude-longitude centered on Los Angeles, a region which we used in previous papers to build similar time series using more involved methods (Rundle & Donnellan, 2020, https://doi.org/10.1029/2020EA001097; Rundle, Donnellan et al., 2021, https://doi.org/10.1029/2021EA001757; Rundle, Stein et al., 2021, https://doi.org/10.1088/1361-6633/abf893). Here we show that not only does the state variable time series have forecast skill, the associated spatial probability densities have skill as well. In addition, use of the standard ROC and Precision (PPV) metrics allow probabilities of current earthquake hazard to be defined in a simple, straightforward, and rigorous way.

Plain Language Summary Earthquake nowcasting refers to the determination of hazard for major earthquakes at the present time, the recent past, and the near future. Nowcasting is an idea borrowed from economics, markets, and meteorology, where it has been frequently used. In this paper, we show that there is order hidden within chaotic earthquake seismicity using a very simple transformation of the data. Small earthquakes appear to transition from unstable stick-slip events that produce seismic waves, to stable sliding where no seismic waves are produced. Our hypothesis is that this transition is due to a material phenomenon called strain-hardening, that is frequently observed in laboratory rock mechanics experiments. The result is a state variable time series, computed over the last 51 years in California, that strongly resembles the long-anticipated cycle of stress accumulation and release. Using supervised machine learning techniques, we can optimize the two-parameter model. From that optimized model, we can rigorously calculate the probability of current hazard from major earthquakes. Extending these methods, we can also compute spatial hazard as well. The result is a new method for assessing earthquake hazard that may be useful for a variety of applications.

1. Introduction

Recent research has developed the idea of earthquake nowcasting, which uses state ("proxy") variables to infer the current state of the earthquake cycle (Luginbuhl et al., 2019; Pasari, 2019a, 2019b, 2020; Pasari & Sharma, 2020; Rundle et al., 2016, 2018, 2019, 2020). An approach such as this is needed since the cycle of stress accumulation and release is not observable (Rundle, Donnellan et al., 2021; Rundle, Stein et al., 2021; Scholz, 2019). These first approaches to nowcasting have been based on the concept of natural time (Sarlis et al., 2018; Varotsos et al., 2001, 2002, 2011, 2013, 2014; Varotsos, Sarlis, et al., 2020; Varotsos, Skordas, et al., 2020).

A review of the current state of earthquake nowcasting, forecasting, and prediction is given by Rundle, Donnellan et al. (2021); Rundle, Stein et al. (2021). Perez-Oregon et al. (2020) have also shown that nowcasting methods

Software: John B. Rundle, Joe Yazbeck, Geoffrey Fox, Michael Heflin
Supervision: John B. Rundle, Andrea Donnellan
Validation: Joe Yazbeck, Lisa Grant Ludwig
Visualization: John B. Rundle, Joe Yazbeck, Michael Heflin
Writing – original draft: John B. Rundle
Writing – review & editing: John B. Rundle, Andrea Donnellan, Geoffrey Fox, Lisa Grant Ludwig, Michael Heflin, James Crutchfield

can be extended to forecasting methods as well. These methods have begun to be applied in India (Pasari & Sharma, 2020), Japan (Nanjo, personal communication, 2020) and Greece (Chouliaras, personal communication, 2019).

In a series of recent papers, Rundle and Donnellan (2020) and Rundle, Donnellan et al. (2021), Rundle, Stein et al. (2021) used small earthquake seismicity, together with machine learning, to produce earthquake nowcasting curves that mimic the cycle of stress accumulation and release. These methods were based on identifying correlations present in the seismicity itself.

We note that other investigators have adopted a different approach to anticipating earthquakes using machine learning. Rouet-LeDuc et al. (2017) developed a prediction technique for acoustic emissions from events in laboratory experiments on regular, nearly periodic stick-slip friction. They also applied a similar method for episodic tremor and slip events in the Pacific Northwest (Rouet-LeDuc et al., 2019), which are also relatively regular in time.

2. Data

The data we used for the analysis was downloaded from the USGS earthquake catalog [1]. Shown in Figure 1 are events having magnitude $M > 3.29$ for a region of 5° latitude × 5° longitude centered on Los Angeles.

In Figure 1a, the small black dots are all the events during the time period 1 January 1970 to 31 December 2022, superposed on the geography of the state, and the earthquake fault system (dark green lines). The blue circles are the events having magnitudes $6.9 > M \geq 6.0$. The larger red circles are all events having magnitudes $M \geq 6.9$. The yellow star denotes a location for InSAR and GNSS observations that we discuss below.

In Figure 1b, we show the monthly number of earthquakes of all magnitudes as a function of time (light cyan vertical lines). In addition, Figure 1b shows the major earthquakes that occurred from January 1970 to 31 December 2022 as vertical lines and inverted triangles. Black dotted lines a black inverted triangles indicate earthquakes having magnitudes $6.9 > M \geq 6.0$, corresponding to the blue circles in Figure 1. Red dashed lines and red inverted triangles indicate earthquakes having $M \geq 6.9$, corresponding to the red circles in Figure 1.

Note that the earthquake numbers are shown in the form $\text{Log}_{10}(1 + \text{monthly number})$, the one being present to eliminate the singularity at $\text{Log}_{10}(0)$. Also, note that 1 month is defined as 4/52 years, equal to 0.07692 years. One can therefore think of this measure of time as a "lunar month."

Most importantly, Figure 1b shows a dashed blue line labeled as "EMA" for "Exponential Moving Average." This EMA curve is striking: it shows a cycle of higher activity following large mainshocks, followed by an increasing quiescence prior to the occurrence of the next mainshock. The higher activity just after mainshocks is of course due to the aftershocks. Quiescence prior to the mainshock has been pointed out by a variety of authors (e.g., Chouliaras, 2009; Kanamori, 1981; Rundle et al., 2011; Weimer & Wyss, 1994).

Furthermore, note that the EMA curve, discussed in more detail in the following section, uses a number of weights $N = 36$. Here $N = 36$ was selected based on a supervised machine learning method, using the area under the Receiver Operating Characteristic (ROC [2]) as a loss function.

The important point to note is that the EMA curve resembles an "upside down" version of the classic earthquake cycle of stress accumulation and release. This cycle has been discussed in many works on seismology, for example, Scholz, 2019. It is this filtered time series to which we now direct our focus.

We improved the method by adding an additional parameter to the model, R_{min}, which is an assumed minimum rate of small earthquakes occurring monthly. These two parameters, N and R_{min}, represent a filter applied to the seismicity time series. Henceforth we refer to time series of this type as "state variable" time series, denoted by $\Theta(t)$. As we show in the following, values of the time series indicate the current state of the earthquake cycle in the region.

The addition of the parameter R_{min} to the model is based on the hypothesis that a strain-hardening process occurs as the next mainshock is approached. This idea implies that there is a transition in small earthquake mechanics from unstable stick-slip early in the earthquake cycle to stable sliding ("silent earthquakes") late in the cycle. These "silent" small earthquakes would not be observable by seismographs, but would contribute to the overall
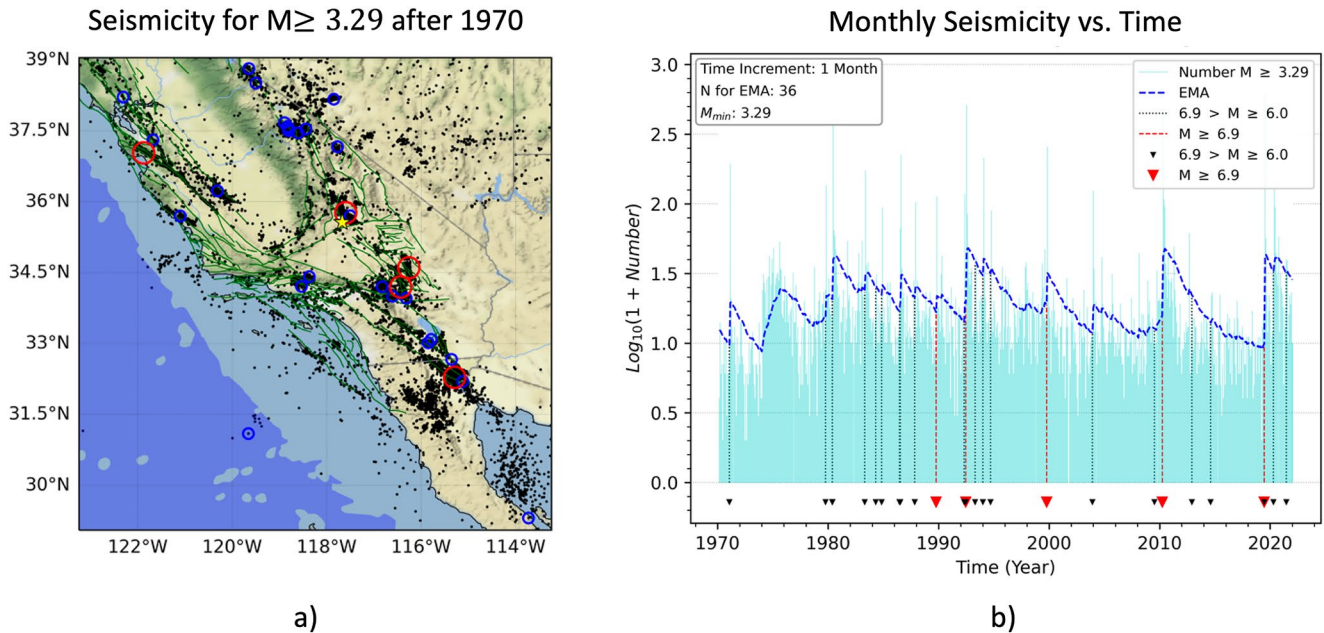
Figure 1. (a) Map of the region, 5° × 5° area centered on Los Angeles. Smallest blackdots represent earthquakes $M > 3.29$. Blue circles represent earthquakes $M \geq 6$, red circles represent earthquakes $M \geq 6.9$. Yellow star is the location of GNSS and InSAR data that is shown in Figure 4a. Monthly seismicity $M > 3.29$ as a function of time for the region in panel (a and b). Blue dashed line is the Exponential Moving Average (EMA) of the monthly seismicity rate with $N = 36$.

crustal straining that would be observable by GNSS and InSAR observations. We discuss these issues in a later section.

3. Results

We first show some of the most important results. We emphasize that the model in this paper is trained on all the data, so results do not represent back tested "predictions." This is in part due to the possibility of catalog completeness and uniformity issues prior to full automation of the seismic network, that occurred after about 1980 (Hutton et al., 2010), a point that we discuss later.

Figure 2 shows the optimized ROC curve (a), the time series $\Theta(t) \equiv \mathrm{Log}_{10}(1 + \text{Monthly Number})$ (b), and the "Precision" (PPV [2] c) of the time series. Note that $\Theta(t)$ is the optimized EMA time series from Figure 1b flipped "upside down" so that $\Theta(t)$ decreases at the time of the most recent major earthquake, then increases as the time of the next large earthquake is approached. This curve is very similar to other such time series found in our previous works (Rundle & Donnellan, 2020; Rundle, Donnellan et al., 2021; Rundle, Stein et al., 2021). In those works, we also pointed out that these time series are reminiscent of the cycle of tectonic stress accumulation and release.

The nowcast PPV [2]) is shown in Figure 2c. As discussed in the following, PPV is a quantitative measure of the success of nowcasting a large mainshock having $M \geq 6.75$ within a future time window $T_W = 3$ years.

The nowcasting process involves "predicting" that a large mainshock will occur during the future time window T_W. One then tabulates the fraction of such time windows in which a large mainshock did occur. Success in this case is defined as the fraction of windows T_W in which a large earthquake did occur divided by the total number of time windows, as a function of the current value of the state variable $\Theta(t)$. The two parameters in the model, N and R_{min}, are optimized by use of the skill score of the Receiver Operating Characteristic (ROC), as shown in Figure 2a and discussed below.

At each value of $\Theta(t)$, such as the red dot in Figure 2b, the value of PPV can be calculated. The value corresponding to the red dot in Figure 2b left is shown by the red dot in Figure 2c, with the value of 52.9% shown in the PPV plot at the right. Figures 2a and 2b are a frame from a movie that is available for downloading from the Supporting Information S1, and corresponds to the month beginning 27 March 1999, about 6 months prior to the M7.1 Hector Mine, California earthquake.
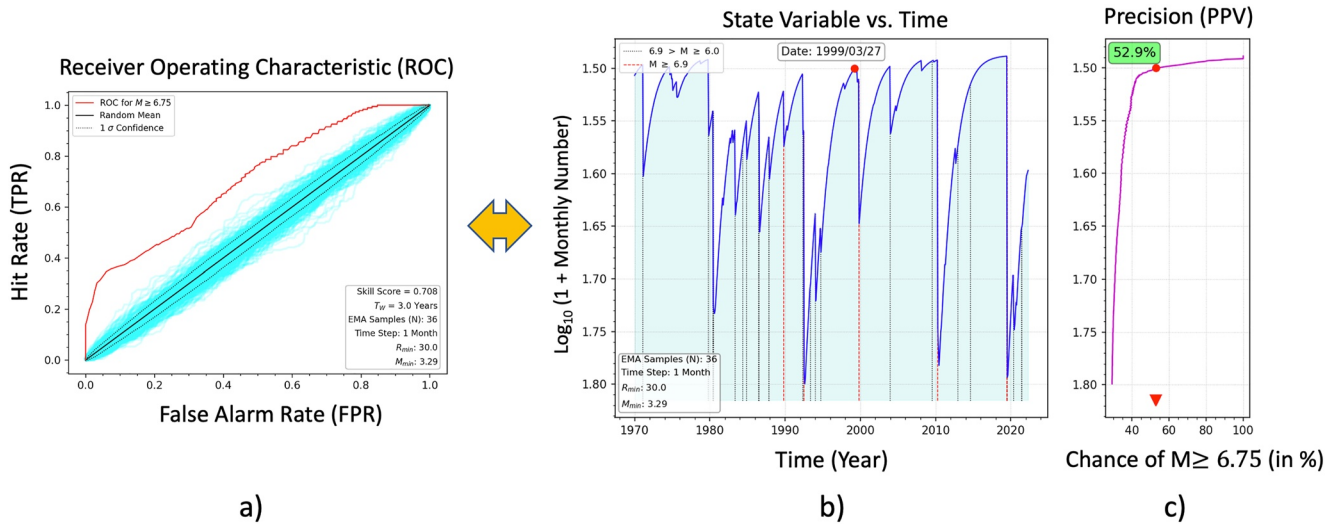
Figure 2. Receiver Operating Characteristic (ROC, (a), state variable time series (b), and Precision (PPV, (c). The figure represents frame 380 of a movie showing the time development of the nowcast. The movie is available in the supplementary material. The red dot represents the date for the frame, which was 27 March 1999. The vertical dashed red lines in (a) are the dates of the large earthquakes $M \geq 6.9$, the vertical black dotted lines are the earthquakes $M \geq 6.0$. Panel (c) represents the chance that an earthquake $M \geq 6.75$ will occur during the following 3 years. The chance of such an earthquake $M \geq 6.75$ at that moment is 52.9% as shown, both by the number in the green box and the red triangle at bottom. Skill is the area under the ROC curve in (a), which we find as 0.708. The no skill line is the diagonal line from lower left to upper right. Since the no skill area is obviously 0.5, this implies that the nowcast time series has significant skill. The cyan lines in (a) are ROC curves for 200 random time series, whose mean is the diagonal no skill line. The dotted lines bracketing the no skill line from lower left to upper right represent the $\pm 1\sigma$ value for the random time ROC series. More results are shown in Table 1.

Figure 2 shows that it is possible to build an optimized two-parameter model directly from the seismicity data, then compute the probability of major earthquakes by means of a rigorous process involving ROC methods. Improvements in the model can be made as we discuss in the following, but at this time, we propose a simple, transparent model.

Corresponding to the temporal model, we can also construct spatial probability density functions (PDFs) or Relative Total Intensity $I_R(\boldsymbol{x}_i,t)$, indicating geographically where large earthquakes are most likely to occur. The basic assumption for these spatial calculations is that large earthquakes tend to occur at locations with the largest number of small earthquakes (Holliday et al., 2005, 2006a, 2006b, 2007; Lee et al., 2011; Rundle et al., 2002).

To compute these $I_R(\boldsymbol{x}_i,t)$ functions, we tesselate the spatial region into a local grid of boxes. The results are shown in Figure 3b for spatial probability densities $I_R(\boldsymbol{x}_i,t)$ for two different grid box sizes, $0.5° \times 0.5°$, and $1° \times 1°$, for the same month as in Figure 2. Note that the two $I_R(\boldsymbol{x}_i,t)$ functions are both normalized over the entire region to 100%, after which the function $Log_{10}(1 + I_R(\boldsymbol{x}_i,t))$ is contoured.

A question arises as to which of the two $I_R(\boldsymbol{x}_i,t)$ functions is "better." Visually, it might seem that the finer-scale grid boxes convey more information. However, as we will discuss in the following section, this question can be answered at least partially by computing the spatial ROC, shown in Figure 3a. The result is (apparently) that finer-scale (smaller grid box) spatial partitions tend to have worse skill.

As discussed previously, skill is computed as the area under the spatial ROC curve when built using all future earthquakes having magnitudes $M \geq 3.29$ within the future $T_W = 3$ years. The future earthquakes are shown as dots and circles in the images. With the ROC method, we show below that the $I_R(\boldsymbol{x}_i,t)$ function in Figure 3a (bottom, grid size: $1° \times 1°$) has better skill than Figure 3b (top, grid size: $0.5° \times 0.5°$).

3.1. Details of Method

3.1.1. Exponential Moving Average (EMA)

Consider a time series $T(t_j)$ representing the number of small earthquakes occurring in a month, that is indexed at regular monthly intervals $j = 1, \ldots, J$. With justification discussed previously and below, we introduce a new parameter R_{min} such that, if $T(t_j) < R_{min}$, we set $T(t_j) = R_{min}$.
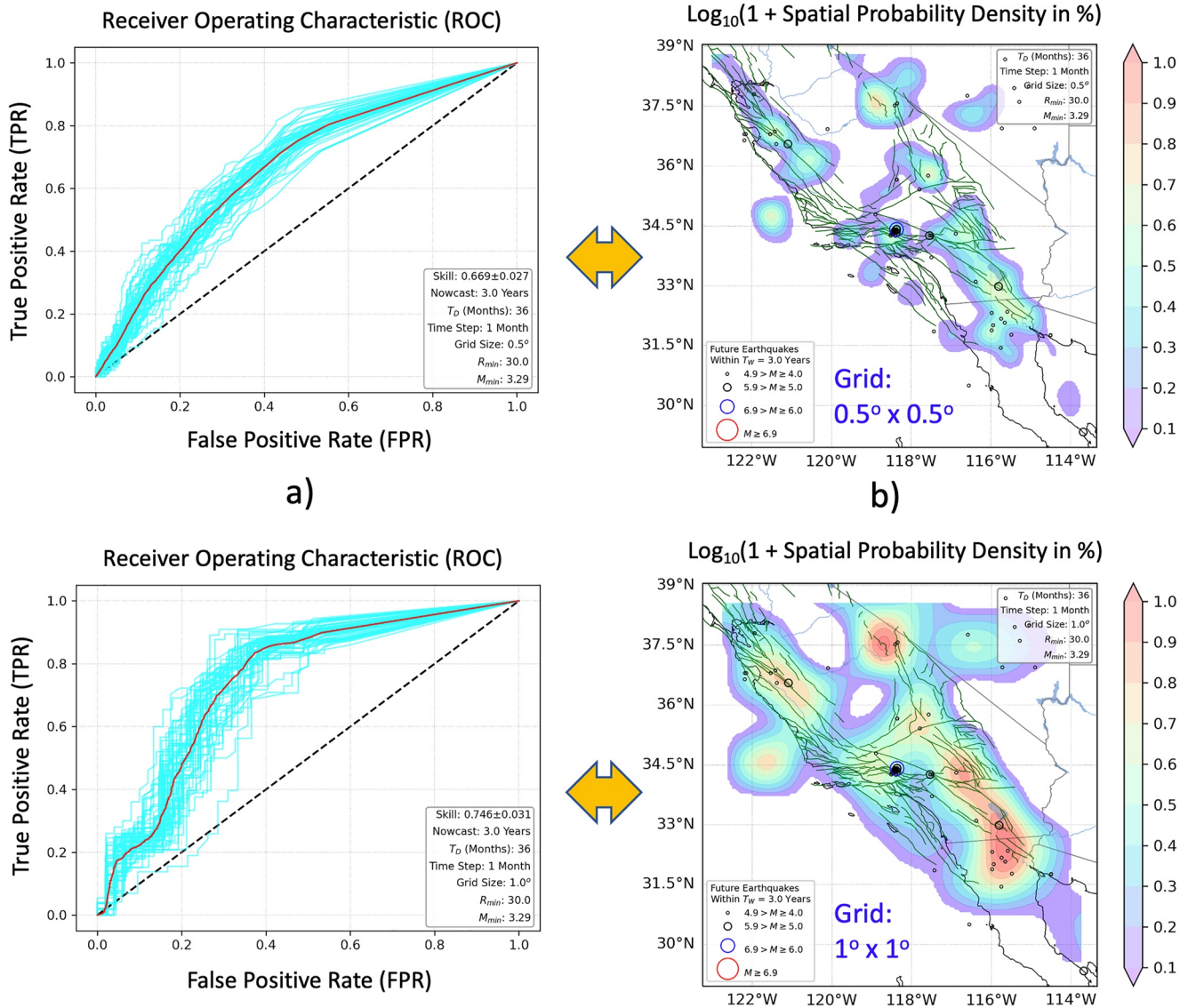
Figure 3. (a) Spatial ROC diagrams (a) for two grid box sizes, 0.5° × 0.5° and 1° × 1°. (b) Spatial probability densities (PDFs). For the PDFs, the future earthquakes during the 3 years after 27 March 1999 are shown as circles on the maps. Smallest circles are for $4.9 \geq M \geq 4.0$, next larger circles are for $5.9 \geq M \geq 5.0$, next larger circles (blue) are for $6.9 > M \geq 6.0$, and largest circles (red) are for $M \geq 6.9$. In the ROC curves in for the two PDFs in (a), the cyan curves are 48 spatial ROC curves evaluated at 1-year intervals beginning with 1970. The curves are evaluated using all earthquakes $M > 3.29$ for the following TW = 3 years. Thresholds are established for all values of the PDFs, and the values of TP, FP, FN, TN are computed, with the ROC curves being computed from those. Again, the dashed line from lower left to upper right is the no skill line. These and other data are listed in Table 1.

As described in Rundle and Donnellan (2021; [1]) the EMA averages over the preceding N times with a diminishing weight for more remote times in the past. More specifically:

$$T_{\mathrm{EMA}}(t) = \begin{cases} T(t_0) & t_j = 0 \\ \alpha T(t_j) + (1-\alpha)T(t_j - 1) & t_j > 0 \end{cases} \quad (1)$$

Here, $\alpha \leq 1$ represents the amount of weighting decrease. A higher value of α discounts older observations faster. The most common choice for α is in terms of the number of weights used is [1]:

$$\alpha = \frac{2}{1 + N} \quad (2)$$

3.1.2. Temporal Receiver Operating Characteristic

As discussed in Rundle and Donnellan (2020) and Rundle, Donnellan et al. (2021), Rundle, Stein et al. (2021), the ROC diagram, first developed with the advent of radar technology, uses a threshold varying from high to low (or low to high) values to classify the data into categories quantified in a "confusion matrix" or "contingency table." These categories are True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). The ROC diagram is a plot of True Positive Rate TPR ("Hit Rate") plotted as a function of False Positive Rate FPR ("False Alarm Rate"). The quantities TPR and FPR are defined as:

$$TPR = \text{TP}/(\text{TP} + \text{FN}); FPR = \text{FP}/(\text{FP} + \text{TN}) \tag{3}$$

These quantities are defined using the state variable time series from Figure 2b, $\Theta(t) \equiv \text{Log}_{10}(1 + \text{Monthly Number})$. Here we adopt the future time window $T_W = 3$ years and classify all points on $\Theta(t)$ for a given threshold value into the four possibilities, TP, FP, FN, TN using the confusion matrix. The threshold is then varied over many values to produce an ensemble of confusion matrices. This then yields $TPR(T_h)$, which is a function of the threshold value T_h, as an implicit function of $FPR(T_h)$.

The resulting temporal ROC is shown in Figure 2a for the time interval 1970–2022. The red curve is the ROC curve for $\Theta(t)$. The diagonal line from lower left to upper right is the "no skill" line. We also show as cyan lines ROC curves for 200 random time series by sampling randomly from $\Theta(t)$ with replacement (bootstrap method). The dashed black lines represent the standard deviation of the random time series.

ROC skill is represented by the area under the ROC curve. No skill is represented by the area of 0.5 under the diagonal no skill line. In Figure 3, the area under the red ROC curve for the time interval 1970–2022 is found by numerical integration to be 0.708, indicating the presence of significant skill.

Note that the values for $N = 36$ and $R_{min} = 30$ were found by using the ROC skill as a loss function, and then optimizing these parameters with a supervised learning method.

Table 1 (top) indicates that the ROC skill has improved significantly in more recent times, a point that we will return to below. Table 1 also includes an entry for 1970–2022 in which $R_{min} = 0$, demonstrating that even the one-parameter model (EMA with $N = 36$ weights) has ROC skill (0.665) better than random (0.5). The one-tailed p-statistics show that all skill scores are significant relative to random skill.

Table 1
Temporal and Spatial Receiver Operating Characteristic (ROC) Values for a Series of Nowcast Examples

Temporal and Spatial ROC Skill			
Time interval	Temporal skill	R_{min}	P-Stat
1950–2022	0.557	30	1.13×10^{-2}
1970–2022	0.708	30	$<10^{-5}$
1984–2022	0.759	30	$<10^{-5}$
1993–2022	0.925	30	$<10^{-5}$
1970–2022	0.657	0	$<10^{-5}$
Resolution	Spatial skill	R_{min}	P-Stat
0.5° × 0.5°	0.669 ± 0.027	30	$<10^{-5}$
0.667° × 0.667°	0.702 ± 0.020	30	$<10^{-5}$
1.0° × 1.0°	0.746 ± 0.031	30	$<10^{-5}$
0.5° × 0.5°	0.626 ± 0.029	0	$<10^{-5}$
1.0° × 1.0°	0.728 ± 0.035	0	$<10^{-5}$

Note. (Top) Optimal examples have two parameters ($N = 36$ and $R_{min} = 30$), but it can be seen that even one-parameter examples with $R_{min} = 0$ have skill significantly better than random, which is 0.5, as seen by the one-tailed P-statistics in the table. (Bottom) Spatial Receiver Operating Characteristic (ROC) values for a series of spatial probability density functions. It can be seen that there is a tradeoff between spatial resolution and skill, with lower spatial resolutions having better skill than higher spatial resolutions. All of the examples in the table have significant skill, as seen by the one-tailed P-statistics. It can also be seen that even one-parameter examples with $R_{min} = 0$ have skill significantly better than random. Another result from Table 1 is that the addition of the R_{min} parameter offers only a relatively modest improvement in skill.

3.1.3. Parameter R_{min}

Recall that the nowcast parameters were obtained by training on all the data from 1970 to 2018. R_{min} was computed by first computing:

$$\mu = \frac{1}{48} \int_{1970}^{2018} \Theta(t') \, dt' \tag{4}$$

We then defined a parameter λ, optimized using supervised learning with ROC skill as the loss function, to compute $R_{min} = \lambda \mu$. Optimization yielded the value $\lambda = 18$.

3.1.4. Relative Total Intensity Contours

Whereas the $\Theta(t)$ time series shown in Figure 2b uses all the $M \geq 3.29$ earthquakes in the defined geographic region, the spatially gridded relative intensity time series $I_R(x_i,t)$ consists of all time series in the ensemble of grid boxes. As shown by Rundle et al. (2002), Tiampo, Rundle, McGinnis, Gross, et al. (2002), Tiampo, Rundle, McGinnis, Klein, et al. (2002), and Holliday et al. (2005, 2006a, 2006b, 2007), large earthquakes tend to occur where the most small earthquakes occur. This fact provides the initial justification for the method.

We then defined a quantity $I_R(x_i,t)|_{EMA}$, where $I_R(x_i,t)|_{EMA}$ is the exponential moving average ($N = 36$) of $I_R(x_i,t)$ and $I_R(x_i,t)$ is required to have a lower bound $R_{min}(x_i) = \lambda\mu(x_i)$. Here $\mu(x_i)$ is the mean of $I_R(x_i,t)$ over the years 1970–2018. The values of $I_R(x_i,t)|_{EMA}$ were then plotted and contoured as in Figure 3. λ is the same parameter as used in the global R_{min}. We then normalized $I_R(x_i,t)|_{EMA}$ over space to 100%, so that it represents a spatial probability density.

We found that the addition of the lower bound $R_{min}(x_i)$ only modestly improves the skill of the spatial contours, as determined from the spatial ROC diagrams in Figure 3a. As for the temporal skill, we also found that using only the EMA for the spatial contours still produced a nowcast with skill better than random (Table 1 bottom).

3.1.5. Spatial Receiver Operating Characteristic

Using the same procedure as for the temporal ROC (Figure 3), we apply a varying spatial threshold ("waterline") and compute spatial *TPR* and *FPR*, as was done in Holliday et al. (2005). The results are shown for the two grid box sizes, $0.5° \times 0.5°$ and $1° \times 1°$ (Figure 4). It can be seen that the grid box size $1° \times 1°$ has higher skill than the $0.5° \times 0.5°$ grid box size.

The ROC curves in Figure 3 were computed at 1-year intervals from 1970 to 2018, producing the 48 cyan-colored curves. Similar to the temporal ROC, a future time interval $T_W = 3$ years was used, and the skill was computed using all future events with $M \geq 3.29$ within T_W. The red curve is the mean of the 48 cyan-colored curves. Using the skill obtained by integrating the area under the 48 cyan-colored curves, we find the mean skill and standard deviation over the 48 years.

ROC skill for the curves in Figure 3, together with some additional calculations, are shown in Table 1 (bottom). There we include calculations for 3 different grid sizes, including one-tailed *P*-statistics, from which it can be seen that the at higher spatial resolutions, the skill is worse. We find a tradeoff between skill and spatial resolution. The *P*-statistics show that all skill scores are significant relative to random (no) skill.

We also include calculations in which we assume $\lambda = 0$ for the two grid sizes. It can be seen that while skill is not as high as for nonzero λ, there is spatial skill nonetheless. Nonzero λ does improve skill, but only modestly.

4. Summary and Discussion

We again emphasize that we have trained the two-parameter model on all the data from 1970 to 2018. Thus, the results are meant to illustrate consistency of model with data, rather than being a validation study with disjoint training and test data.

As discussed previously, we have proposed a two-parameter model, using N weights for the exponential moving average, and the bias term R_{min}. The parameter R_{min} improves ROC skill significantly for the temporal skill, but only modestly for the spatial ROC skill. We now discuss the physical motivation for including R_{min}.
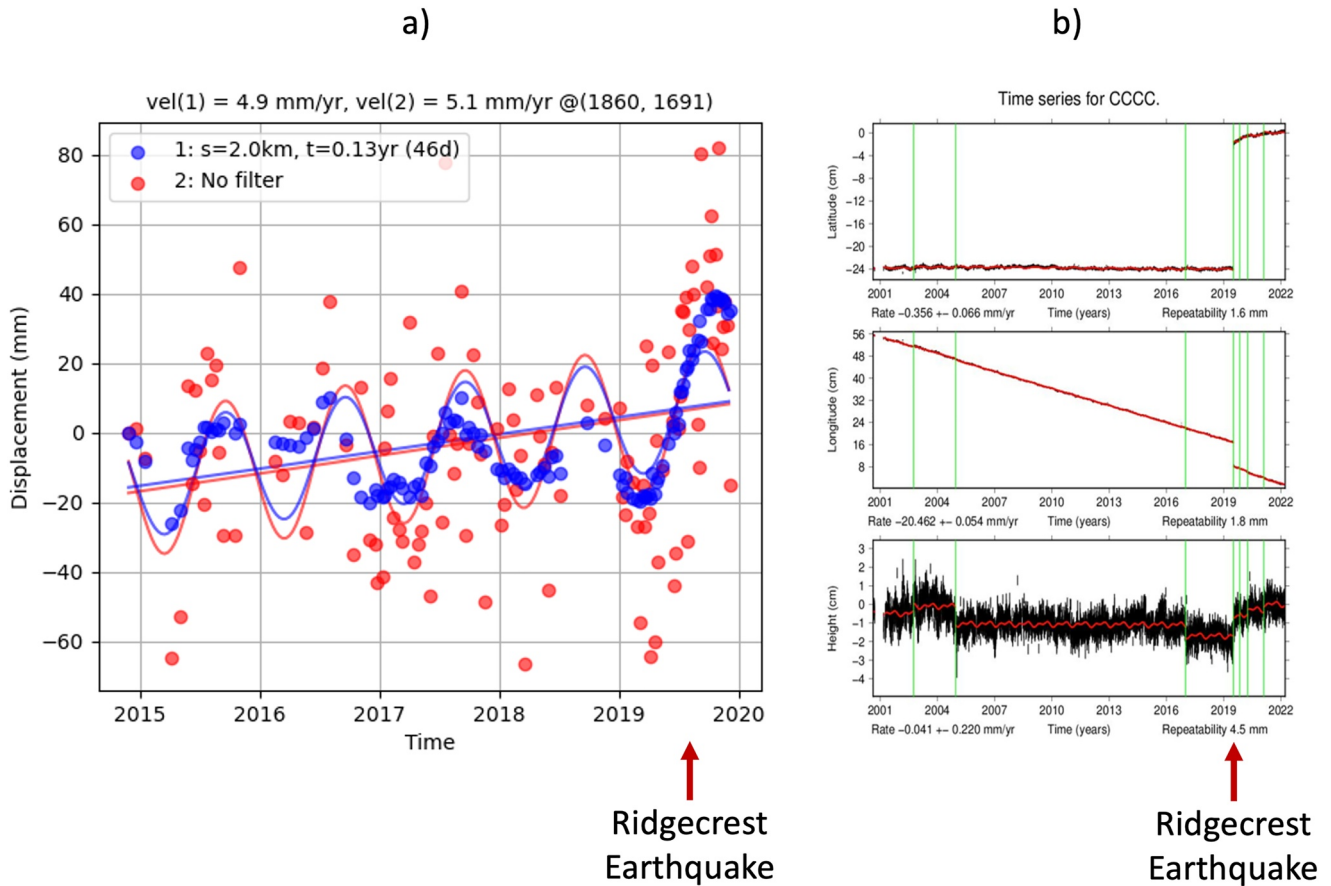
Figure 4. Surface deformation from Sentinel InSAR (a), and GNSS data (b) at a point having latitude 35.57, longitude −117.68, near the epicenter of the M7.1 Ridgecrest, California earthquake of 5 July 2019, over the years from 2015 to 2019 (yellow star in Figure 1, Cerro Coso Community College—CCCC). Using LiCSBAS, a spatial-temporal filter was applied that is high-pass in time and low-pass in space with a Gaussian kernel that is identical to filters used in other time series InSAR software packages such as StaMPS. The spatial filter width used is 2 km and the temporal filter width used is 0.16 years. After applying the filter, the velocity computed is 4.9 mm/yr. The fact that the deformation over the years prior to the Ridgecrest event is linear suggests that small earthquakes did not stop or decrease with time. When compared to the decrease in small seismic earthquakes prior to the large earthquakes seen in Figure 2, it suggests that there was a transition from unstable stick-slip events to stable sliding. This is because unstable events can only be seen with seismometers, whereas both unstable and stable slip can be observed with geodetic data.

We find that this minimum earthquake rate R_{min} becomes important in the lead-up to the large earthquakes as (presumably) the regional tectonic stress increases prior to the next large earthquake. This process is similar to the laboratory observed process of strain-hardening (Kandula et al., 2019) or strain-stiffening (Jiang & Srinivasan, 2013). In the former, more of the deformation is taken up by inelastic (plastic) straining as failure is approached. In the latter the laboratory rock sample becomes increasingly rigid, leading to a transition from unstable stick-slip on microcracks to stable sliding. In either case, the stable slip or plastic straining is modeled by the excess rate R_{min}.

Transition from unstable stick-slip to stable sliding has been observed in laboratory friction experiments (Beeler, 2004; Beeler et al., 2001; Chen & Lapusta, 2009; Dieterich, 1986; Heslot et al., 1994). It is seen as the stiffness of the testing machine increases, as well as in experiments in failure in triaxial stress experiments and for observations of seismicity.

Support for this application to earthquakes can be seen in both InSAR and GNSS observations of deformation prior to major earthquakes in California. In Figure 4 we show the surface deformation from Sentinel InSAR (a) and GNSS data (b) in at a point having latitude 35.57°, longitude −117.68°, near the epicenter of the M7.1 Ridgecrest, California earthquake of 5 July 2019, over the years from 2015 to 2019 (yellow star in Figure 1). The InSAR data were processed using the LiCSBAS package (Lazecký et al., 2020; Morishita, 2021; Morishita

et al., 2020; [3]). Note that the frame ID is 064A_05410_131313, and other details can be found at the COMET Sentinel InSAR portal [4]. The GNSS data were processed at JPL by Heflin et al. (2020) [5]. Additional GNSS data for the 4 April 2010 M7.2 El Mayor Cucupah earthquake also shows the same lack of anomalous precursory deformation. Examples of these data are shown on the ESSOAR archive [6] under the Supporting Information.

Aside from seasonal fluctuations in the InSAR displacement, the trends are linear, showing no sign of the type of quiescence implied by the seismically observable data. This fact suggests an approach to potentially locate the epicenter of an impending earthquake, by comparing the trend in surface deformation, observed by InSAR and/or GNSS, to the trend in Benioff strain due to the unstable slip-slip, small earthquakes. These observations might be focused on the areas of higher probability density seen in Figure 4.

Note that to calculate the velocity from the InSAR images, we adopt the following procedure. The velocity is computed using the method of small baseline inversion. Specifically, a stack of unwrapped interferograms is used to calculate the incremental displacements which are the displacements between two consecutive times. Then, in order to get the cumulative displacements and form the displacement time series, the incremental displacements are simply summed up for each acquisition. The velocity is then derived from the time series using the least squares method. Looking at the time series and ignoring the seasonal effects which manifest themselves as an increase in displacement during the Winter times and a decrease in the Summer time, we are able to observe a steady linear trend in displacement leading up to the earthquake.

As a final remark, we have found that the mean rate of small earthquakes has been highly variable in time, generally increasing. This time dependence could be explained either by the physical processes generating the earthquakes changing over time, or alternately by incompleteness of the catalog prior to the full implementation of the digital observation network in the mid-1990s. Hutton et al. (2010) have a discussion of the completeness level that is informative in this regard. Further refinements of the method presented here might include taking account of this time dependence, but this is a subject that we leave to future research.

Data Availability Statement

Python code that can be used to reproduce the results of this paper can be found at a GitHub repository (DOI: https://doi.org/10.5281/zenodo.7186635) [6]. Data for this paper was downloaded from the USGS earthquake catalog for California, and are freely available there. The Python code mentioned above can be used to download and model these data for analysis [6].

Note:

1. https://en.wikipedia.org/wiki/Moving_average#Exponential_moving_average (accessed 3/9/2022)
2. https://en.wikipedia.org/wiki/Receiver_operating_characteristic (accessed 3/9/2022)
3. https://github.com/yumorishita/LiCSBAS (accessed 3/9/2022)
4. https://comet.nerc.ac.uk/comet-lics-portal/ (accessed 3/9/2022)
5. https://sideshow.jpl.nasa.gov/post/series.html
6. https://github.com/jbrundle/OptimizingEarthquakeNowcasts/tree/main

References

Beeler, N. M. (2004). Review of the physical basis of laboratory-derived relations for brittle failure and their implications for earthquake occurrence and earthquake nucleation. *Pure and Applied Geophysics*, *161*, 1853–1876. https://doi.org/10.1007/s00024-004-2536-z

Beeler, N. M., Lockner, D. L., & Hickman, S. H. (2001). A simple stick-slip and creep-slip model for repeating earthquakes and its implication for microearthquakes at Parkfield. *Bulletin of Seismological Society of America*, *91*(6), 1797–1804. https://doi.org/10.1785/0120000096

Chen, T., & Lapusta, N. (2009). Scaling of small repeating earthquakes explained by interaction of seismic and aseismic slip in a rate and state fault model. *Journal of Geophysical Research*, *114*, B01311. https://doi.org/10.1029/2008JB005749

Chouliaras, G. (2009). Seismicity anomalies prior to 8 June 2008, Mw = 6.4 earthquake in Western Greece. *Natural Hazards and Earth System Sciences*, *9*, 327–335. https://doi.org/10.5194/nhess-9-327-2009

Dieterich, J. H. (1986). A model for the nucleation of earthquake slip. In S. Das, J. Boatwright, & C. H. Scholz (Eds.), *Earthquake source mechanics, Geophysical Monograph Series* (Vol. 37, pp. 37–49). AGU.

Heflin, M., Donnellan, A., Parker, J., Lyzenga, G., Moore, A., Ludwig, L. G., et al. (2020). Automated estimation and tools to extract positions, velocities, breaks, and seasonal terms from daily GNSS measurements: Illuminating nonlinear Salton Trough deformation. *Earth and Space Science*, *7*(7), e2019EA000644. https://doi.org/10.1029/2019EA000644

Heslot, F., Baumberger, T., Perrin, B., Caroli, B., & Caroli, C. (1994). Creep, stick-slip, and dry-friction dynamics: Experiments and a heuristic model. *Physical Review E*, *49*(6), 4973–4988. https://doi.org/10.1103/physreve.49.4973

Holliday, J. R., Chen, C. C., Tiampo, K. F., Rundle, J. B., Turcotte, D. L., & Donnellan, A. (2007). A RELM earthquake forecast based on Pattern Informatics. *Seismological Research Letters*, *78*, 87–93. https://doi.org/10.1785/gssrl.78.1.87

Holliday, J. R., Nanjo, K. Z., Tiampo, K. F., Rundle, J. B., & Turcotte, D. L. (2005). Earthquake forecasting and its verification. *Nonlinear Processes in Geophysics*, *12*(6), 965–977. https://doi.org/10.5194/npg-12-965-2005

Holliday, J. R., Rundle, J. B., Tiampo, K. F., & Donnellan, A. (2006). systematic procedural and sensitivity analysis of pattern informatics method for forecasting large (M>5) earthquake events in southern California. *Pure and Applied Geophysics*, *163*(11–12), 2433–2454. https://doi.org/10.1007/s00024-006-0131-1

Holliday, J. R., Rundle, J. B., Turcotte, D. L., Klein, W., & Tiampo, K. F. (2006). Space-time correlation and clustering of major earthquakes. *Physical Review Letters*, *97*, 238501. https://doi.org/10.1103/physrevlett.97.238501

Hutton, K., Woessner, J., & Hauksson, E. (2010). Earthquake monitoring in Southern California for seventy-seven years (1932–2008). *Bulletin of Seismological Society of America*, *100*(2), 423–446. https://doi.org/10.1785/0120090130

Jiang, C., & Srinivasan, S. G. (2013). Unexpected strain-stiffening in crystalline solids. *Nature*, *496*, 339–342. https://doi.org/10.1038/nature12008

Kanamori, H. (1981). The nature of seismicity patterns before large earthquakes. In D. W.Simpson, & P. G. Richards (Eds.), *Earthquake prediction: An international review, Maurice Ewing Series* (Vol. 4, pp. 1–19). American Geophysical Union.

Kandula, N., Cordonnier, B., Boller, E., Weiss, J., Dysthe, D. K., & Renard, F. (2019). Dynamics of microscale precursors during brittle compressive failure in Carrara marble. *Journal of Geophysical Research*, *124*, 6121–6139. https://doi.org/10.1029/2019JB017381

Lazecký, M., Spaans, K., González, P. J., Maghsoudi, Y., Morishita, Y., Albino, F., et al. (2020). LiCSAR: An automatic InSAR tool for measuring and monitoring tectonic and volcanic activity. *Remote Sensing*, *12*, 2430. https://doi.org/10.3390/rs12152430

Lee, Y. T., Turcotte, D. L., Holliday, J. R., Sachs, M. K., Rundle, J. B., Chen, C. C., & Tiampo, K. F. (2011). Results of the regional earthquake likelihood models (RELM) test of earthquake forecasts in California. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 16533–16538. https://doi.org/10.1073/pnas.1113481108

Luginbuhl, M., Rundle, J. B., & Turcotte, D. L. (2019). Statistical physics models for aftershocks and induced seismicity. *Philosophical Transactions of the Royal Society A*, *377*(2136), 20170397.

Morishita, Y. (2021). Nationwide urban ground deformation monitoring in Japan using Sentinel-1 LiCSAR products and LiCSBAS. *Progress in Earth and Planetary Science*, *8*, 6. https://doi.org/10.1186/s40645-020-00402-7

Morishita, Y., Lazecky, M., Wright, T. J., Weiss, J. R., Elliott, J. R., & Hooper, A. (2020). LiCSBAS: An open-source InSAR time series analysis package integrated with the LiCSAR automated Sentinel-1 InSAR processor. *Remote Sensing*, *12*, 424. https://doi.org/10.3390/RS12030424

Nanjo, K. Z. (2020). Were changes in stress state responsible for the 2019 Ridgecrest, California, earthquakes? *Nature Communications*, *11*, 3082. https://doi.org/10.1038/s41467-020-16867-5

Pasari, S. (2019a). Nowcasting earthquakes in the Bay of Bengal region. *Pure and Applied Geophysics*, *176*, 1417. https://doi.org/10.1007/s00024-018-2037-0

Pasari, S. (2019b). Nowcasting earthquakes in the Bay-of-Bengal region. *Pure and Applied Geophysics*, *23*, 537–559. https://doi.org/10.1007/s10950-019-09822-5

Pasari, S. (2020). Stochastic modeling of earthquake Interevent Counts (natural times) in northwest Himalaya and adjoining regions. In S. Bhattacharyya, J. Kumar, & K. Ghoshal (Eds.), *Mathematical modeling and computational tools, Springer proceedings in mathematics & statistics* (Vol. 320, pp. 495–501). Springer. https://doi.org/10.1007/978-981-15-3615-1_35

Pasari, S., & Mehta, A. (2018). Nowcasting earthquakes in the northwest Himalaya and surrounding regions. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-5*, 855–859. https://doi.org/10.5194/isprs-archives-xlii-5-855-2018

Pasari, S., & Sharma, Y. (2020). Contemporary earthquake hazards in the West-northwest Himalaya: A statistical perspective through natural times. *Seismological Research Letters*, *91*(6), 3358–3369.

Perez-Oregon, J., Angulo-Brown, F., & Sarlis, N. V. (2020). Nowcasting avalanches as earthquakes and the predictability of strong avalanches in the Olami-Feder-Christensen model. *Entropy*, *22*(11), 1228. https://doi.org/10.3390/e22111228

Rouet-Leduc, B., Hulbert, C., & Johnson, P. A. (2019). Continuous chatter of the Cascadia subduction zone revealed by machine learning. *Nature Geoscience*, *12*, 75–79. https://doi.org/10.1038/s41561-018-0274-6

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., & Johnson, P. A. (2017). Machine learning predicts laboratory earthquakes. *Geophysical Research Letters*, *44*, 9276–9282. https://doi.org/10.1002/2017GL074677

Rundle, J. B., & Donnellan, A. (2020). Nowcasting earthquakes in Southern California with machine learning: Bursts, swarms, and aftershocks may be related to levels of regional tectonic stress. *Earth and Space Science*, *7*, e2020EA0010. https://doi.org/10.1029/2020EA001097

Rundle, J. B., Donnellan, A., Fox, G., Crutchfield, J. P., & Granat, R. (2021). Nowcasting earthquakes: Imaging the earthquake cycle in California with machine learning. *Earth and Space Science*, *8*, e2021EA001757. 10.1029/2021EA001757

Rundle, J. B., Donnellan, A., Grant Ludwig, L., Gong, G., Turcotte, D. L., & Luginbuhl, M. (2016). Nowcasting earthquakes. *Earth and Space Science*, *3*, 480–486. https://doi.org/10.1002/2016EA000185

Rundle, J. B., Holliday, J. R., Yoder, M., Sachs, M. K., Donnellan, A., Turcotte, D. L., et al. (2011). Earthquake precursors: Activation or quiescence? *Geophysical Journal International*, *187*(1), 225–236. https://doi.org/10.1111/j.1365-246x.2011.05134.x

Rundle, J. B., Luginbuhl, M., Giguere, A., & Turcotte, D. L. (2018). Natural time, nowcasting and the physics of earthquakes: Estimation of risk to global megacities. *Pure and Applied Geophysics*, *175*, 647–660. https://doi.org/10.1007/s00024-017-1720-x

Rundle, J. B., Stein, S., Donnellan, A., Turcotte, D. L., Klein, W., & Saylor, C. (2021). The complex dynamics of earthquake fault systems: New approaches to forecasting and nowcasting of earthquakes. *Reports on Progress in Physics*, *84*(7), 076801. https://doi.org/10.1088/1361-6633/abf893

Rundle, J. B., Tiampo, K. F., Klein, W., & Martins, J. S. S. (2002). Self-organization in leaky threshold systems: The influence of near mean field dynamics and its implications for earthquakes, neurobiology and forecasting. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(Suppl. 1), 2514–2521. https://doi.org/10.1073/pnas.012581899

Sarlis, N. V., Skordas, E. S., & Varotsos, P. A. (2018). A remarkable change of the entropy of seismicity in natural time under time reversal before the super-giant M9 Tohoku earthquake on 11 March 2011. *EPL*, *124*, 29001. https://doi.org/10.1209/0295-5075/124/29001

Scholz, C. H. (2019). *The mechanics of earthquakes and faulting*. Cambridge University Press.

Tiampo, K. F., Rundle, J. B., McGinnis, S., Gross, S., & Klein, W. (2002). Mean field threshold systems and phase dynamics: An application to earthquake fault systems. *Europhysics Letters*, *60*, 481–487. https://doi.org/10.1209/epl/i2002-00289-y

Tiampo, K. F., Rundle, J. B., McGinnis, S. A., & Klein, W. (2002). Pattern dynamics and forecast methods in seismically active regions. *Pure and Applied Geophysics*, *159*, 2429–2467. https://doi.org/10.1007/s00024-002-8742-7

Varotsos, P., Sarlis, N. V., & Skordas, E. S. (2001). Spatiotemporal complexity aspects on the interrelation between Seismic Electric Signals and seismicity. *Practica of Athens Academy*, *76*, 294–321.

Varotsos, P., Sarlis, N. V., & Skordas, E. S. (2002). Long-range correlations in the electric signals that precede rupture. *Physical Review E— Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *78*, 059902. https://doi.org/10.1103/physreve.66.011902

Varotsos, P., Sarlis, N. V., & Skordas, E. S. (2011). Precursory seismic electric signals, earthquakes and other complex time-series. In *Natural time analysis: The new view of time*. Springer-Verlag. https://doi.org/10.1007/978-3-642-16449-1

Varotsos, P., Sarlis, N. V., & Skordas, E. S. (2014). Study of the temporal correlations in the magnitude time series before major earthquakes in Japan. *Journal of Geophysical Research: Space Physics*, *119*, 9192–9206. https://doi.org/10.1002/2014JA020580

Varotsos, P., Sarlis, N. V., Skordas, E. S., & Lazaridou, M. S. (2013). Seismic electric signals: An additional fact showing their physical interconnection with seismicity. *Tectonophysics*, *589*, 116–125. https://doi.org/10.1016/j.tecto.2012.12.020

Varotsos, P. A., Sarlis, N. V., & Skordas, E. S. (2020). Self-organized criticality and earthquake predictability: A long-standing question in the light of natural time analysis. *EPL*, *132*(2), 29001. https://doi.org/10.1209/0295-5075/132/29001

Varotsos, P. A., Skordas, E. S., & Sarlis, N. V. (2020). Fluctuations of the entropy change under time reversal: Further investigations on identifying the occurrence time of an impending major earthquake. *EPL*, *130*, 29001. https://doi.org/10.1209/0295-5075/130/29001

Wiemer, S., & Wyss, M. (1994). Seismic quiescence before the Landers (M = 7.5) and Big bear (M = 6.5) 1992 earthquakes. *Bulletin of Seismological Society of America*, *84*(3), 900–916.