Nowcasting Earthquakes by Visualizing the Earthquake Cycle with Machine Learning:A Comparison of Two Methods

John Rundle[1], Andrea Donnellan[2], Geoffrey Fox[3], and James Crutchfield[1]

[1]University of California Davis
[2]Jet Propulsion Laboratory, California Institute of Technology
[3]Indiana University Bloomington

November 23, 2022

Abstract

The earthquake cycle of stress accumulation and release is associated with the elastic rebound hypothesis proposed by H.F. Reid following the M7.9 San Francisco earthquake of 1906. However, observing details of the actual values of time- and space-dependent tectonic stress is not possible at the present time. In previous research, we have proposed two methods to image the earthquake cycle in California by means of proxy variables. These variables are based on correlations in patterns of small earthquakes that occur nearly continuously in time. One of these is based on the construction of a time series by the unsupervised detection of small earthquake clusters. The other is based on expanding earthquake seismicity in PCA-derived patterns, to construct a weighted correlation time series. The purpose of the present research is to compare these two methods by evaluating their information content using decision thresholds and Receiver Operating Characteristic methods together with Shannon information entropy. Using seismic data from 1940 to present in California, we find that both methods provide nearly equivalent information on the rise and fall of earthquake correlations associated with major earthquakes in the region. We conclude that the resulting time series can be viewed as proxies for the cycle of stress accumulation and release associated with major tectonic activity. The figure shows the PCA patterns of small earthquakes associated with 5 major M>7 earthquakes in California since 1950.

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

by

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

[1] Department of Physics
University of California, Davis, CA

[2] Santa Fe Institute
Santa Fe, NM

[3] Department of Earth and Planetary Science
University of California, Davis, CA

[4] Division of Earth and Space Science
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA

[5] University of Virginia
Charlottesville, VA

[6] Program in Public Health
University of California, Irvine, CA

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 2-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 timeseries 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^o x 5^o in latitude-longitude centered on Los Angeles, a region which we used in previous papers to build similar timeseries using more involved methods (Rundle and Donnellan, 2020; Rundle et al., 2021). Here we show that not only does the state variable timeseries 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.

Key Points

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

Introduction

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

A review of the current state of earthquake nowcasting, forecasting, and prediction is given by Rundle et al. (2021). Perez-Oregon et al. (2020) have also shown that nowcasting methods can be extended to forecasting methods as well. These methods have begun to be applied in India (Pasari, 2019), Japan (K. Nanjo, 2020; personal comm., 2020) and Greece (G. Chouliaras, personal comm. 2019).

In a series of recent papers, Rundle and Donnellan (2020) and Rundle et al. (2021a,b,c) 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.

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 x 5° longitude centered on Los Angeles.

In **Figure 1a**, the small black dots are all the events during the time period 1/1/1970 – 12/31/2022, 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 1/1/1970 – 12/31/2022 as vertical lines and inverted triangles. Black dotted lines an 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 $Log_{10}(1 + \text{monthly number})$, the 1 being present to eliminate the singularity at $Log_{10}(0)$. 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, for example (Kanamori, 1981; Chouliaras, 2009; Weimer and Wyss, 1994; Rundle et al., 2011).

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

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 backtested "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 $\Theta(t) \equiv Log_{10}(1 + \text{Monthly Number})$ 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 and Donnellan; Rundle et al., 2021a,2021b). 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 ≥ 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 on the **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. **Figure 2a,b** is a frame from a movie that is available for downloading from the supplementary material, and corresponds to the month beginning 3/27/1999, about 6 months prior to the M7.1 Hector Mine, California earthquake.

Figure 2 shows that it is possible to build an optimized 2-parameter model directly from the seismicity data, then compute the probability of major earthquakes by means of a rigorous process involving Receiver Operating Characteristic (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 (Rundle et al., 2002; Holliday et al., 2005; 2006a,b; 2007; Lee et al., 2011)

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° x 0.5°, and 1° x 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 ≥ 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° x 1°) has better skill than **Figure 3b** (top, grid size: 0.5° x 0.5°).

Details of Method

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 $T(t_j) = R_{min}$.

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:

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

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]:

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

Temporal Receiver Operating Characteristic. As discussed in Rundle and Donnellan(2020) and Rundle et al. (2021a,b), 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:

$$TPR = TP/(TP + FN); \quad FPR = FP/(FP + TN) \qquad (3)$$

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 4 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 1-parameter model (EMA with N =36 weights) has ROC skill (0.665) better than random (0.5). The 1-tailed P-statistics show that all skill scores are significant relative to random skill.

Parameter R_{min}. Recall that the nowcast parameters were obtained by training on all the data from 1970-2018. 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 $\lambda = 18$.

Relative Total Intensity Contours. Whereas the $\Theta(t)$ time series shown in **Figure 2b** uses all the M ≥ 3.29 earthquakes in the defined geographic region, the spatially gridded relative intensity time series $I_R(\boldsymbol{x}_i,t)$ consists of all time series in the ensemble of grid boxes. As shown by Rundle et al. (2002; Tiampo et al., 2002a,b; Holliday et al., 2005; 2006a,b; 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 $I_R(\boldsymbol{x_i},\mathrm{t})|_{EMA}$, where $I_R(\boldsymbol{x_i},\mathrm{t})|_{EMA}$ is the exponential moving average (N = 36) of $I_R(\boldsymbol{x}_i,\mathrm{t})$ and $I_R(\boldsymbol{x}_i,\mathrm{t})$ is required to have a lower bound R_{min} ($\boldsymbol{x}_i$) = $\lambda\,\mu(\boldsymbol{x_i})$. Here $\mu(x_i)$ is the mean of $I_R(\boldsymbol{x_i},\mathrm{t})$ over the years 1970-2018. The values of $I_R(\boldsymbol{x_i},\mathrm{t})|_{EMA}$ were then plotted and contoured as in **Figure 3**. λ is the same parameter as used in the global R_{min}. We then normalized $I_R(\boldsymbol{x_i},\mathrm{t})|_{EMA}$ over space to 100%, so that it represents a spatial probability density.

We found that the addition of the lower bound R_{min} ($\boldsymbol{x}_i$) 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).

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° x 0.5° and 1° x 1° (**Figure 4**). It can be seen that the grid box size 1° x 1° has higher skill than the 0.5° x 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 ≥ 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 1-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 λ = 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.

Summary and Discussion

We again emphasize that we have trained the 2-parameter model on all the data from 1970-2018. 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 2-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}.

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 and 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 (Dieterich, 1986; Heslot et al., 1994; Beeler et al., 2001; Beeler et al., 2004; Chen and Lapusta, 2009),. 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^o, longitude -117.68^o, near the epicenter of the M7.1 Ridgecrest, California earthquake of July 5, 2019, over the years from 2015-2019 (yellow star in Figure 1). The InSAR data were processed using the LiCSBAS package (Morishita et al., 2020, 2021; Lazecký 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].

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.

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 1990's. 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.

Acknowledgements. Research by JBR and JY were supported under NASA grant NNX12AM22G to the University of California, Davis. Research by JBR was also supported in part by a grant from the Southern California Earthquake Center grant USC/USGS 131436517, and grant DE-SC0017324 from the US Department of Energy to UC Davis. GCF was supported by NSF CINES grant CINES 1835598. Portions of this research were also carried out at the Jet Propulsion Laboratory, California Institute of Technology under contract with NASA. "LiCSAR contains modified Copernicus Sentinel data [Year of data used] analysed by the Centre for the Observation and Modeling of Earthquakes, Volcanoes and Tectonics (COMET). LiCSAR uses JASMIN, the UK's collaborative data analysis environment (http://jasmin.ac.uk)

Notes

[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

References

Beeler, N. M., Lockner, D.L., and Hickman, S.H., A simple stick-slip and creep-slip model for repeating earthquakes and its implication for microearthquakes at Parkfield." *Bull. Seism. Soc. Am.* 91.6, 1797-1804. (2001)

Beeler, N.M., Review of the physical basis of laboratory-derived relations for brittle failure and their implications for earthquake occurrence and earthquake nucleation, *Pure App. Geophy*., 161, 1853-1876, 2004.

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

Chouliaras, G, Seismicity anomalies prior to 8 June 2008, Mw=6.4 earthquake in Western Greece, *Nat. Hazards Earth Syst. Sci.,* 9, 327–335 (2009)

Dieterich, J.H., A model for the nucleation of earthquake slip, in *Earthquake Source Mechanics*, Geophys. Monogr. Ser., vol. 37 (eds S. Das et al., (AGU, Washington, D.C. 1986), pp. 37–49.

Green, David M.; Swets, John A. (1966). Signal detection theory and psychophysics. New York, NY: John Wiley and Sons Inc. ISBN 978-0-471-32420-1.

Heflin, M, Donnellan, S., Parker, J, Lyzenga, G, Moore, A, Grant Ludwig, L, Rundle, JB, Wang, J, and Pierce, M., 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 (2020)

Heslot, F, Baumberger, T, Perrin, B, Caroli, B, and Caroli, C, Creep, stick-slip, and dry-friction dynamics: Experiments and a heuristic model." *Physical review E* 49.6, 4973 - 4988 (1994)

Holliday, J.R., Nanjo, K.Z., Tiampo, K.F., Rundle, J.B., and Turcotte, D.L., Earthquake forecasting and its verification, *Nonlin. Proc. Geophys.,* 12.6, 965-977(2005)

Holliday, JR, Chen, CC, Tiampo, KF, Rundle, JB, Turcotte, DL, and Donnellan, A, A RELM earthquake forecast based on Pattern Informatics, *Seism. Res. Let*t., 78, 87-93 (2007).

Holliday, JR, Rundle, JB, Tiampo, KF, and Donnellan, A systematic procedural and sensitivity analysis of pattern informatics method for forecasting large (M>5) earthquake events in southern California, *Pure Appl. Geophys.*, 10.1007/s00024-006-0131-1 v. 163, No. 11-12, 2433-2454 (2006a).

Holliday, JR, Rundle, JB, Turcotte, DL, Klein, W. and Tiampo, KF, Space-time correlation and clustering of major earthquakes, *Phys. Rev. Lett*., 97, 238501 (2006b)

Hutton, K, Woessner, J, and Hauksson, E., Earthquake monitoring in Southern California for seventy-seven years (1932–2008), *Bull. Seism. Soc. Am*., 100, No. 2, pp. 423–446, (2010)

Jiang, C and SG Srinivasan, *Nature*, 496, pages 339–342 (2013)

Kanamori, H., The nature of seismicity patterns before large earthquakes, pp. 1-19, in *Earthquake Prediction, An International Review*, ed. DW Simpson and PG Richards, American Geophysical Union, Maurice Ewing Series 4, (1981)

Kandula, N, Cordonnier, B., Boller, E., Weiss, J., Dysthe, DK, Renard, F., Dynamics of microscale precursors during brittle compressive failure in Carrara marble, *J. Geophys. Res*., 10.1029/2019JB017381, 6121-6139 (2019)

Lazecký, M.; Spaans, K.; González, P.J.; Maghsoudi, Y.; Morishita, Y.; Albino, F.; Elliott, J.; Greenall, N.; Hatton, E.; Hooper, A.; Juncu, D.; McDougall, A.; Walters, R.J.; Watson, C.S.; Weiss, J.R.; Wright, T.J. LiCSAR: An Automatic InSAR Tool for Measuring and Monitoring Tectonic and Volcanic Activity. *Remote Sens.* **2020**, *12*, 2430, https://doi.org/10.3390/rs12152430.

Lee, YT, Turcotte, DL, Holliday, JR, Sachs, MK, Rundle, JB, Chen, CC, and Tiampo, KF, Results of the Regional Earthquake Likelihood Models (RELM) test of earthquake forecasts in California, *Proc. Nat. Acad. Sci USA*, **108**, 16533-16538 (2011) DOI: 10.1073/pnas.1113481108

Luginbuhl, M., Rundle, J.B., Hawkins, A. and Turcotte, D.L. Nowcasting earthquakes: a comparison of induced earthquakes in Oklahoma and at the Geysers, California. *Pure Appl. Geophys.*, 175(1), 49-65 (2018a).

Morishita, Y.; Lazecky, M.; Wright, T.J.; Weiss, J.R.; Elliott, J.R.; Hooper, A. LiCSBAS: An Open-Source InSAR Time Series Analysis Package Integrated with the LiCSAR Automated Sentinel-1 InSAR Processor. *Remote Sens.* **2020**, *12*, 424, https://doi.org/10.3390/RS12030424.

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

Nanjo, K. Z. "Were changes in stress state responsible for the 2019 Ridgecrest, California, earthquakes?." *Nature communications* 11 (2020).

Omori, F., 1894. On the aftershocks of earthquakes: *Journal of the College of Science*, Imperial University of Tokyo.

Pasari, S. Nowcasting earthquakes in the Bay-of-Bengal region. *Pure Appl. Geophys.* 23, 537-559 (2019).

Pasari, S. Stochastic Modeling of Earthquake Interevent Counts (Natural Times) in Northwest Himalaya and Adjoining Regions. In: Bhattacharyya, S., Kumar, J. and Ghoshal, K. Mathematical Modeling and Computational Tools, *Springer Proceedings in Mathematics & Statistics*, 320, 495-501, Springer, Singapore (2020).

Pasari, S., and Mehta, A. 2018. Nowcasting earthquakes in the northwest Himalaya and surrounding regions. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-5, 855–859 (2018).

Pasari, S., and Sharma, Y. Contemporary Earthquake Hazards in the West-Northwest Himalaya: A Statistical perspective through Natural Times. *Seismol. Res. Lett.* (in print)

Pasari, S., Nowcasting Earthquakes in the Bay of Bengal Region, *Pure Appl. Geophys*. (2019) 176: 1417. https://doi.org/10.1007/s00024-018-2037-0

Perez-Oregon, Jennifer, Fernando Angulo-Brown, and Nicholas Vassiliou Sarlis. "Nowcasting Avalanches as Earthquakes and the Predictability of Strong Avalanches in the Olami-Feder-Christensen Model." Entropy 22.11 (2020): 1228. (2020).

Rouet-Leduc, B., Hulbert, C. and Johnson, P.A. Continuous chatter of the Cascadia subduction zone revealed by machine learning. *Nature Geoscience*, 12, 75-79 (2019).f

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C.J. and Johnson, P.A. Machine Learning Predicts Laboratory Earthquakes. *Geophys. Res. Lett.,* 44(18) (2017).

Rundle, John B., and David D. Jackson, Numerical simulation of earthquake sequences, *Bulletin of the Seismological Society of America* 67.5 (1977): 1363-1377.

Rundle, J. B., and Andrea Donnellan, 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* 7.9 (2020): e2020EA0010

Rundle, J.B., Holliday, J.R., Yoder, M., Sachs, M.K., Donnellan, A., Turcotte, D.L., Tiampo, K.F., Klein, W. and Kellogg, L.H., Earthquake precursors: activation or quiescence?, *Geophys. J. Int.* 187.1 225-236, (2011)

Rundle, J.B., Donnellan, A., Fox, GCF, Crutchfield, JP, and Granat, R., Nowcasting earthquakes: Imaging the earthquake cycle in California with machine learning, submitted to *Earth and Space Science*, (2021).

Rundle, J.B., Donnellan, A., Grant Ludwig, L, Gong, G., Turcotte, D.L. and Luginbuhl, M. Nowcasting earthquakes. *Earth and Space Science*, 3, 480-486 (2016).

Rundle, J.B., Luginbuhl, M., Giguere, A., and Turcotte, D.L. Natural time, nowcasting and the physics of earthquakes: Estimation of risk to global megacities. *Pure Appl. Geophys*., 175, 647-660 (2018).

Rundle, J.B., Luginbuhl, M., Khapikova, P. et al. Nowcasting Great Global Earthquake and Tsunami Sources, Pure Appl. Geophys. (2019b) doi:10.1007/s00024-018-2039-y

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

Rundle, JB, Giguere, A, Turcotte, DL, Crutchfield, JP, and Donnellan, A, Global seismic nowcasting with Shannon information entropy, Earth and Space Science, 6, 456-472 (2019a)

Rundle, JB, Luginbuhl, M, Giguere, A and Turcotte, DL, Natural time, nowcasting and the physics of earthquakes: Estimation of risk to global megacities, *Pure Appl. Geophys*., 175, 647-660 (2018)

Rundle, JB, Tiampo, KF, Klein, W and Martins, JSS, Self-organization in leaky threshold systems: The influence of near mean field dynamics and its implications for earthquakes, neurobiology and forecasting, *Proc. Nat. Acad. Sci. USA*, 99, Supplement 1, 2514-2521, (2002)

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

Rundle, JB, Turcotte, DL, Sammis, C, Klein, W, and Shcherbakov, R., Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems (invited), *Rev. Geophys. Space Phys*., **41**(4), DOI 10.1029/2003RG000135 (2003).

Sarlis, N.V., Skordas, E.S. and Varotsos, P.A. 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 (2018).

Scholz, C.H, *The Mechanics of Earthquakes and Faulting* (2019). Cambridge University Press.

Tiampo, K.F., Rundle, J.B., McGinnis, S.A. and Klein, W. Pattern dynamics and forecast methods in seismically active regions. *Pure Appl. Geophys.*, 159, 2429-2467 (2002a).

Tiampo, KF, Rundle, JB, McGinnis, S, Gross, S. and Klein, W, Mean field threshold systems and phase dynamics: An application to earthquake fault systems, *Europhys. Lett*., 60, 481-487, (2002b).

Varotsos, P. A., N. V. Sarlis, and E. S. Skordas, Self-organized criticality and earthquake predictability: A long-standing question in the light of natural time analysis, *EPL (Europhysics Letters)* 132.2 (2020b): 29001.

Varotsos, P., Sarlis, N.V. and Skordas, E.S. Long-range correlations in the electric signals that precede rupture. *Phys. Rev. E*, 66 (2002).

Varotsos, P., Sarlis, N.V. and Skordas, E.S. *Natural Time Analysis: The new view of time. Precursory Seismic Electric Signals, Earthquakes and other Complex Time-Series.* Springer-Verlag, Berlin Heidelberg (2011).

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

Varotsos, P., Sarlis, N.V. and Skordas, E.S. Study of the temporal correlations in the magnitude time series before major earthquakes in Japan. *J. Geophys. Res. Space Phys.*, 119, 9192-9206 (2014).

Varotsos, P., Sarlis, N.V., Skordas, E.S. and Lazaridou, M.S. Seismic Electric Signals: An additional fact showing their physical interconnection with seismicity. *Tectonophys.*, 589, 116-125 (2013).

Varotsos, P.A., Skordas, E.S. and Sarlis, N.V. Fluctuations of the entropy change under time reversal: Further investigations on identifying the occurrence time of an impending major earthquake. *EPL*, 130 (2020a).

Wiemer, S, and Wyss, M., Seismic quiescence before the Landers (M= 7.5) and Big Bear (M= 6.5) 1992 earthquakes., *Bull. Seism. Soc. Am.,* 84.3 , 900-916, (1994)

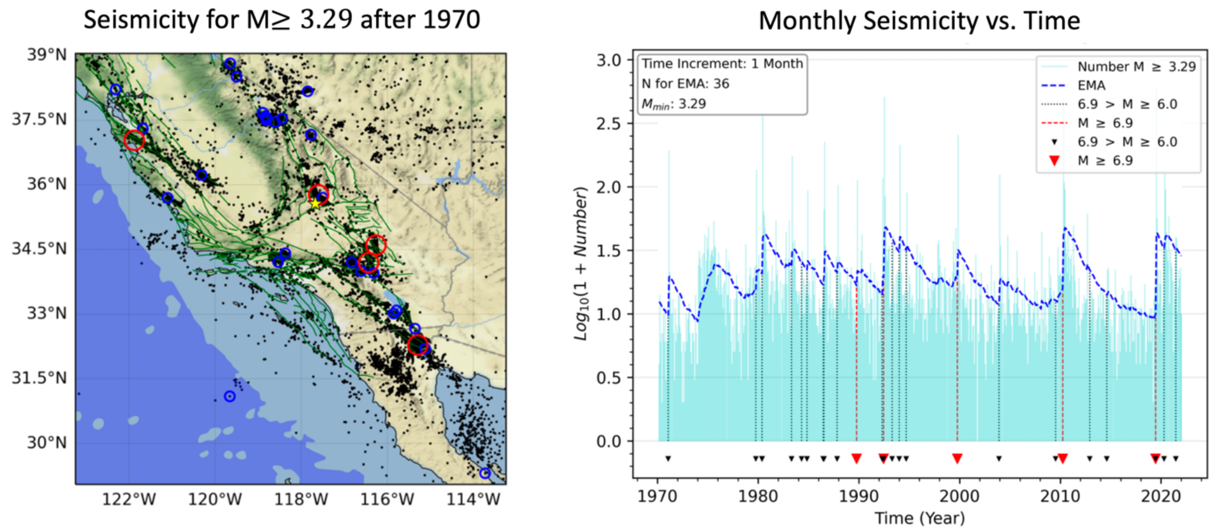
Table 1. Temporal and Spatial Receiver Operating Characteristic (ROC) values for a series of nowcast examples. (**Top**) Optimal examples have 2 parameters (N=36 and R_{min} = 30), but it can be seen that even 1 parameter examples with R_{min} = 0 have skill significantly better than random, which is 0.5, as seen by the 1-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 1-tailed P-statistics. It can also be seen that even 1 parameter examples with R_{min} = 0 have skill significantly better than random. Another result from Table 2 is that the addition of the R_{min} parameter offers only a relatively modest improvement in skill.

Table 1

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**
$0.5^{o} \times 0.5^{o}$	0.669 ± 0.027	30	$< 10^{-5}$
$0.667^{o} \times 0.667^{o}$	0.702 ± 0.020	30	$< 10^{-5}$
$1.0^{o} \times 1.0^{o}$	0.746 ± 0.031	30	$< 10^{-5}$
$0.5^{o} \times 0.5^{o}$	0.626 ± 0.029	0	$< 10^{-5}$
$1.0^{o} \times 1.0^{o}$	0.728 ± 0.035	0	$< 10^{-5}$

Figure 1. a) Map of the region, 5° x 5° area centered on Los Angeles. Smallest blackdots represent earthquakes M>3.29. Blue circles represent earthquakes M ≥ 6, red circles represent earthquakes M ≥ 6.9. Yellow star is the location of GNSS and InSAR data that is shown in Figure 9. b) Monthly seismicity M > 3.29 as a function of time for the region in Figure 1. Blue dashed line is the Exponential Moving Average (EMA) of the monthly seismicity rate with N = 36.



a) b)

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 3/27/1999. The vertical dashed red lines in a) are the dates of the large earthquakes M ≥ 6.9, the vertical black dotted lines are the earthquakes M ≥ 6.0. Panel c) represents the chance that an earthquake M ≥ 6.75 will occur during the following 3 years. The chance of such an earthquake M ≥ 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 ± 1 σ value for the random time ROC series. More results are shown in Table 1.

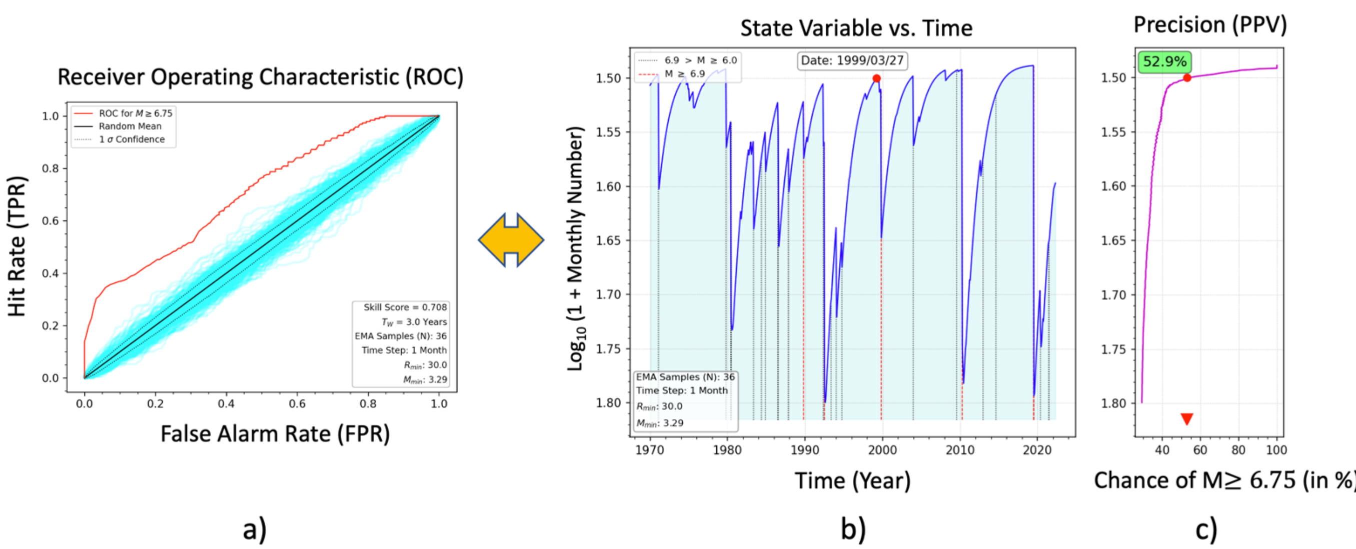


Figure 3. a) Spatial ROC diagrams a) for two grid box sizes, 0.5º x 0.5º and 1º x 1º. b) Spatial probability densities (PDFs). For the PDFs, the future earthquakes during the 3 years after 3/27/1999 are shown as circles on the maps. Smallest circles are for 4.9 ≥M ≥ 4.0, next larger circles are for 5.9 ≥M ≥ 5.0, next larger circles (blue) are for 6.9 >M ≥ 6.0, and largest circles (red) are for M ≥ 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 the following T_W = 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.

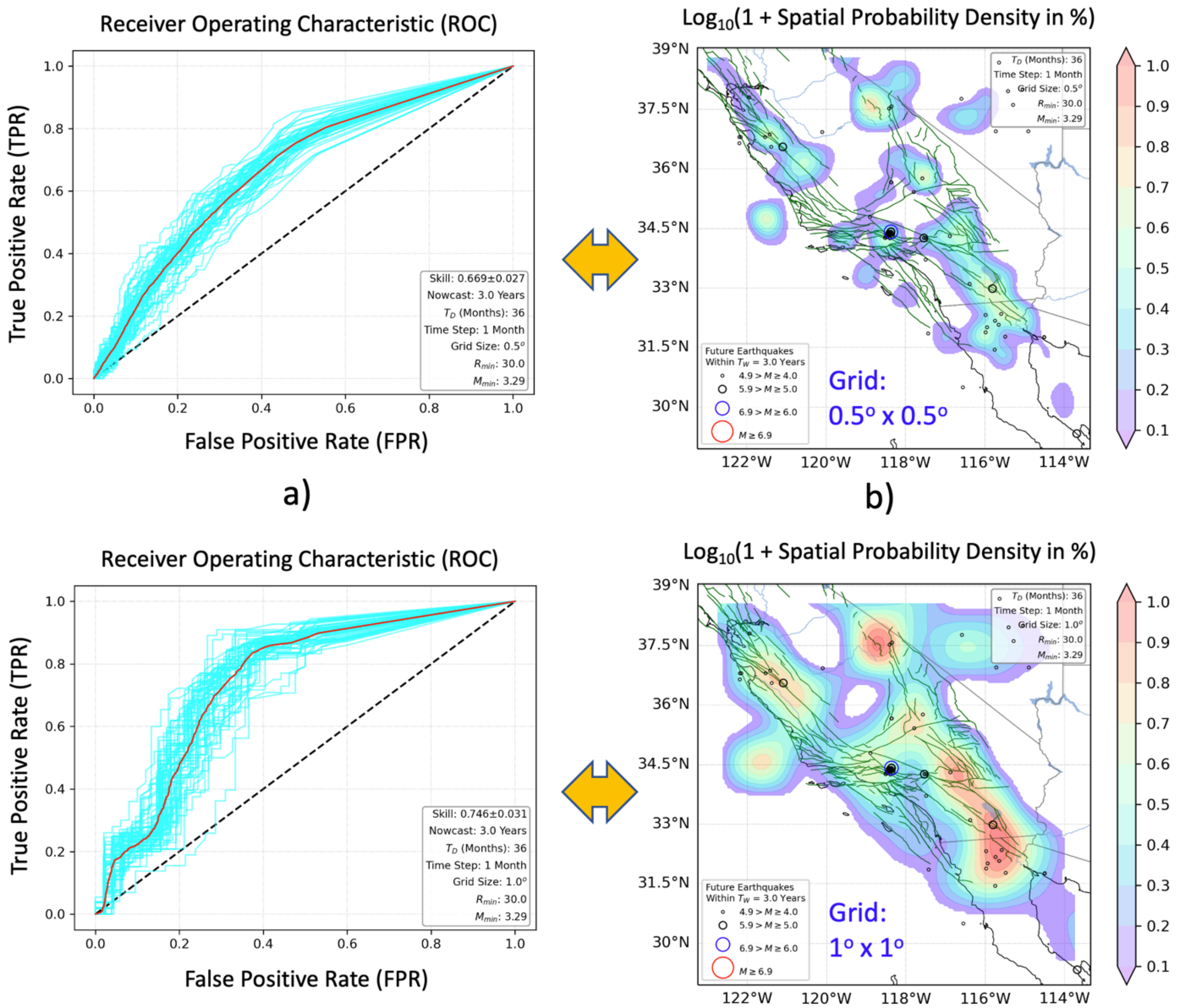
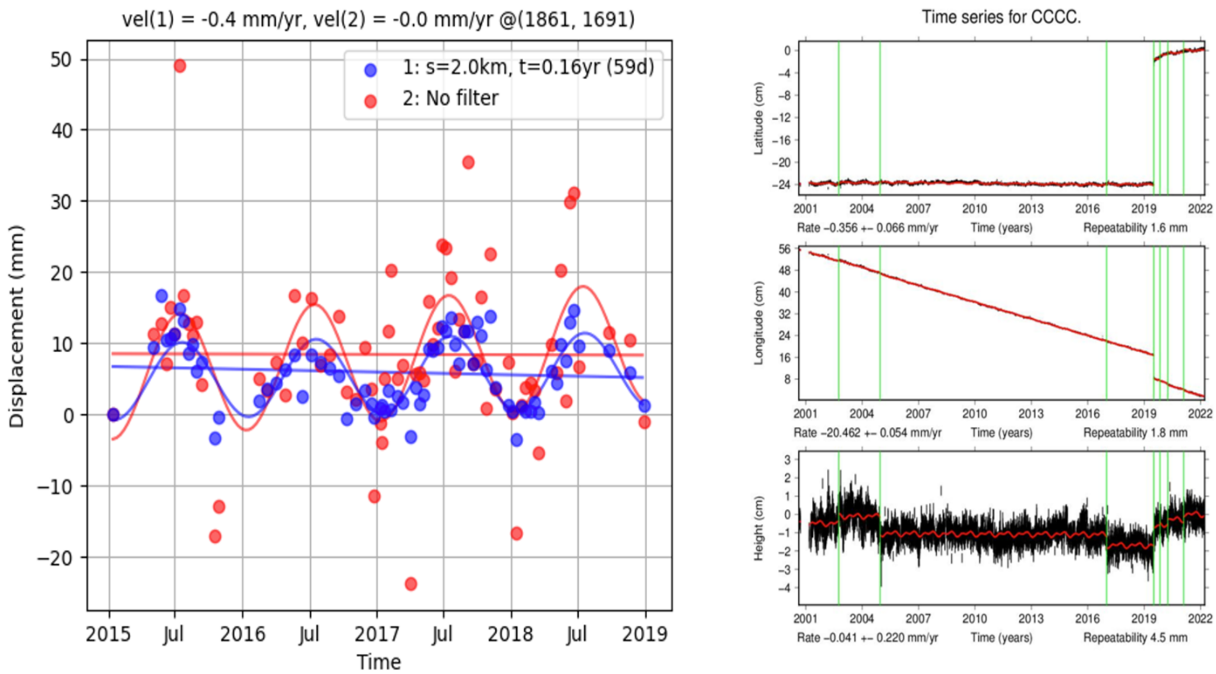


Figure 4. Surface deformation from Sentinel InSAR (a), and GNSS data (b) at a point having latitude 35.57, longitude -117.68, near the epicenter of the M7.1 Ridgecrest, California earthquake of July 5, 2019, over the years from 2015-2019 (yellow star in Figure 1, Cerro Coso Community College - CCCC). 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 2km and the temporal filter width used is 0.16 yr. After applying the filter, the velocity computed is -0.4 mm/yr from a velocity of 0mm/yr with no filter. 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.



a) b)