**Machine Learning Predicts Aperiodic Laboratory Earthquakes**

Olha Tanyuk1, Daniel Davieau1, Dr. Michael L. Blanpied2, Dr. Charles South1 and Dr. Daniel W. Engels1

1 Southern Methodist University, Dallas, TX 75205, USA

2  U.S. Geological Survey Headquarters, Reston, VA 20192, USA

**Abstract.** Our goal is to find the pattern of aperiodic seismic signals that precede earthquakes. We apply machine learning to data set, that comes from a classic laboratory experiment involving repeated stick-slip displacement (“earthquake”) on a sliding interface, a type of experiment which has been studied in depth as a tabletop analog of seismogenic faults for decades. Here we show that by listening to the acoustic signal emitted by a laboratory fault, an algorithm tuned through machine learning can predict the time remaining before it fails with a 0.65 coefficient of determination and 1.61 sec mean absolute error. These predictions are based solely on the instantaneous physical characteristics of the acoustical signal and do not make use of its history. Los Alamos' initial work [1] showed that the prediction of laboratory earthquakes from continuous seismic data is possible in the case of quasi-periodic laboratory seismic cycles. In this work we use a much more challenging dataset with considerably more aperiodic earthquake failures with more realistic behavior.

# INTRODUCTION

Earthquakes cause mass destruction and loss of life. A traditional method to predict earthquakes is to look to past recurrence intervals. Because the recurrences are not constant, predictions can only be made within broad time windows. Over the last 15 years, there has been renewed hope that progress can be made regarding forecasting owing to tremendous advances in instrumentation quality and density. These advances have led to exciting discoveries of previously unidentified slip processes that include slow slip, low frequency earthquakes and Earth tremor, that occur deep in faults. These discoveries form a new understanding of fault slip and may lead to advances in predicting.[1]

In August 2017 Los Alamos National Laboratory (LANL) conducted an experiment [1] that illuminates the mechanics of slow-slip phenomena. They predicted the remaining time until laboratory earthquakes occur with 89% accuracy. In this paper we use acoustic data, which was provided by LANL in January 2019, as part of Kaggle project (reference?), which also represent laboratory slow-slip earthquakes. Data from this experiment exhibit a very aperiodic and more realistic behavior compared to the data they studied earlier, with earthquakes occurring very irregularly.[2] The results of this experiment are potentially applicable to the field of real world earthquakes. Other potential applications include avalanche prediction or failure of machine parts. “If this challenge is solved and the physics are ultimately shown to scale from the laboratory to the field, researchers will have the potential to improve earthquake hazard assessments that could save lives and billions of dollars in infrastructure.[2]” In this paper, given seismic signal data with considerably more a-periodic laboratory earthquake failures and modern computing hardware, we find the pattern of acoustic signals to predict the time at which laboratory earthquakes will occur.

# TUTORIAL MATERIAL

We hear about earthquakes mostly via news media when there is a large seismic event which is noticeable, causes death and destruction. These are stick–slip events that radiate seismic energy along the seams (fault lines) between tectonic plates. In this study we refer to these as Regular Earthquakes. Regular earthquakes are caused by a sudden slip on a fault. Tectonic plates are always slowly moving, but they get stuck at their edges due to friction. When the stress on the edge overcomes the friction, there is an earthquake that releases energy in waves that travel through the earth’s crust and cause the shaking that we feel. [5]

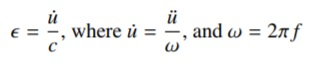
Another type of earthquake we refer to in this study is a Slow Slip Earthquake (SSE). SSE’s are fault behaviors that occur slowly enough to make them undetectable without instrumentation. They do not shake the ground and cause widespread destruction like regular earthquakes do. They occur near the boundaries of large earthquake rupture zones[3]. There is evidence to suggest that there is a relationship between slow slip earthquakes and more noticeable regular earthquakes[4].

LANL researchers discovered a way to successfully predict SSE in a laboratory experiment that simulates natural conditions. In 2017, this team trained a computer to pinpoint and analyze seismic and acoustic signals emitted during the movements along the fault. They processed massive amounts of data and identified a particular sound pattern previously thought to be noise that precedes an earthquake. The team was able to characterize the time remaining before a laboratory earthquake at all times.[6]

In the lab, the team imitated a real earthquake using steel blocks interacting with rocky material (fault gouge) to induce slipping that emitted seismic sounds. An accelerometer recorded the acoustic emission emanating from the sheared layers.[6] For the first time, researchers discovered a pattern that accurately predicted when a quake would occur. The team acknowledges that the physical traits of the lab experiment (such as shear stresses and thermal properties) differ from the real world but the application of the analysis to real earthquakes to validate their results is ongoing. This method can also be applied outside of seismology to support materials’ failure research in many fields such as aerospace and energy.[6] The team’s lab results reveal that the fault does not fail randomly but in a predictable manner. The observations also demonstrate that the fault’s critical stress state, which indicates when it might slip, can be determined using exclusively an equation of state.[6] So far seismologists and Earth scientists have mostly relied on catalogues of historical data to try to characterize the state of faults. These catalogues contain a minute fraction of seismic data, and remaining seismic data is discarded during analysis as useless noise. The authors discovered that—in the case of their laboratory faults--hidden in this noiselike data there are signals emitted by the fault that inform them of the state of the fault much more precisely than catalogues.[6] “Our work shows that machine learning can be used to extract new meaningful physics from a very well studied system,” said Bertrand Rouet-Leduc, Los Alamos Earth and Environmental Sciences Division scientist and the paper’s lead author. “It also shows that seismogenic faults are continuously broadcasting a signal that precisely informs us of their physical state and how close they are to rupture, at least in the laboratory.”

**2.1. Experimental setup [1] (was copy pasted from original LANL paper)**

A three-block assembly with two gouge layers is placed in a bi-axial stress configuration. Two 5 mm thick fault gouge layers are placed between the three blocks, which are held in place by a fixed normal load. The gouge material is comprised of Class IV beads with diameter 105-149 µm. The central block is sheared at constant displacement rate. The two data streams recorded for our purposes here are the shear stress and the acoustic signal. At some time while the gouge material is in a critical shear stress regime, the shear stress abruptly drops, indicating gouge failure. These large drops in shear stress are laboratory quakes. As applied load progressively increases, the inter-event time (recurrence) of laboratory earthquakes progressively decreases. At smaller applied loads the slips become aperiodic. In all cases, the rate of impulsive precursors accelerates as failure is approached. The acoustic particle acceleration Ü is measured on the central block and can be readily converted to dynamic strain used in the ML analysis:



with c ≈ 700m/s the average measured wave speed in the granular material, and f ≈ 40.3kHz. The sampling rate of the acoustic data is 330kHz. We are band-limited by the 20 accelerometer (the frequency response is poor above about 50kHz). Therefore we select one of the system mechanical resonances within this band occurring at 44 kHz. Using this peak we improve our signal to noise ratio. The ‘noise’ is of very different character when the piston is stopped, and reflects primarily the mechanical resonances of the system. In short, we are certain the signal we analyze is the acoustic signal emanating from the fault, and the electromagnetic and system noise play no role in the predictions. [1]

**2.2. Extra Trees Regressor overview**

Will be later

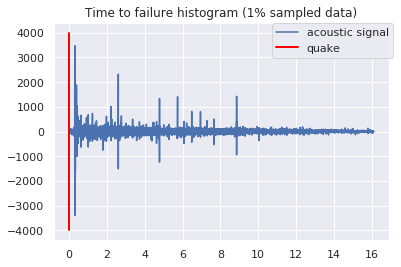
# DATA

The data used in this work is a chunk of 157.275 seconds of seismic data (ordered in time), which is recorded at 4MHz, hence 629,145,480 data points, and the output is time remaining until the following lab earthquake, in seconds.

The seismic data is recorded using a piezoceramic sensor, which outputs a voltage upon deformation by incoming seismic waves (henceforth we will use the term seismic signal or acoustic signal) . The seismic data of the input is this recorded voltage, in integers. (Table 1)

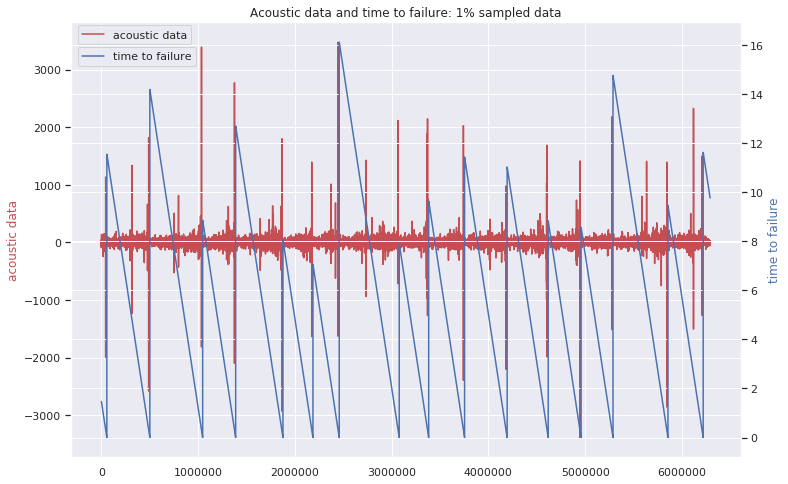
|  |  |  |
| --- | --- | --- |
| **Index** | **Voltage upon deformation by incoming seismic waves** | **Time remaining until the following lab earthquak**e |
| **0** | 12 | 1.469099998474121 |
| **1** | 6 | 1.469099998474121 |
| **2** | 8 | 1.469099998474121 |
| **3** | 5 | 1.469099998474121 |
| **4** | 8 | 1.469099998474121 |

The voltage rate of acoustic precursors accelerates as failure approaches, suggesting that upcoming laboratory earthquake timing could be predicted (Fig.1). In Fig. 1 we used 1% sample of the data. Red line indicates, that quake occurs, when time to failure approaches to 0. Minimum time remaining until the quake in the data is -5.5150e+03 sec.



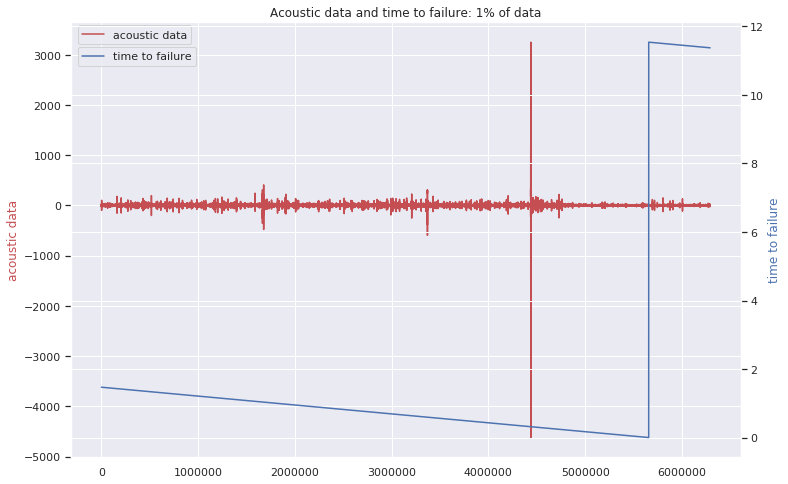
**Fig. 1.**

The data has 16 earthquakes. The shortest time to failure is 1.5 seconds for the first earthquake and 7 seconds for the 7th, while the longest is around 16 seconds. We can see that usually acoustic data shows huge fluctuations just before the failure and the nature of data is cyclical. Another important point: visually failures can be predicted as cases when huge fluctuations in signal are followed by small signal values (Fig.2). In Fig.2 we used 1% sample of the data, which will show the whole picture of the experiment (all 16 quakes). Each time, when time to failure close to 0, we have a quake.



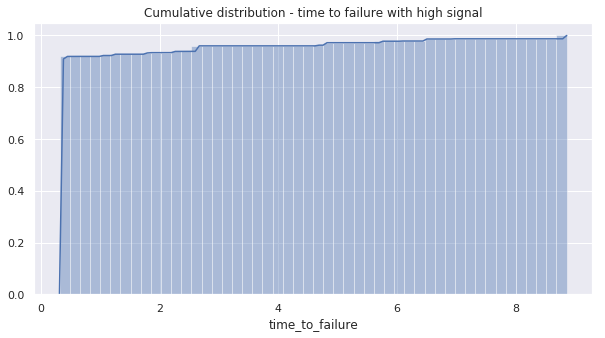
**Fig. 2**

On zoomed-in-time plot (first 1% of the data, not 1% sample that we used in Fig.2) we can see that the large oscillation before the failure is not in the last moment. There are trains of intense oscillations preceding the large one and also some oscillations with smaller peaks after it. (Fig. 3)

****

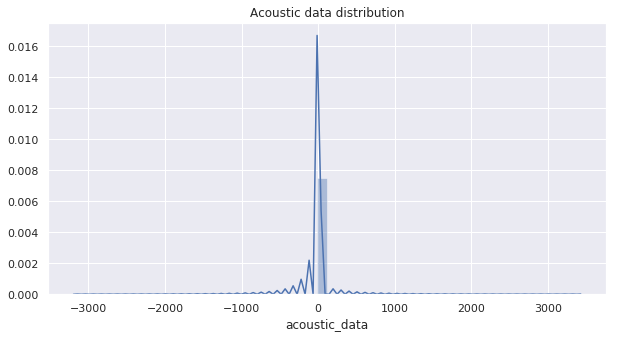
**Fig. 3**

We found that more than 90% of high acoustic values (absolute value greater than 1000) are around 0.31 seconds before an earthquake**!**

****

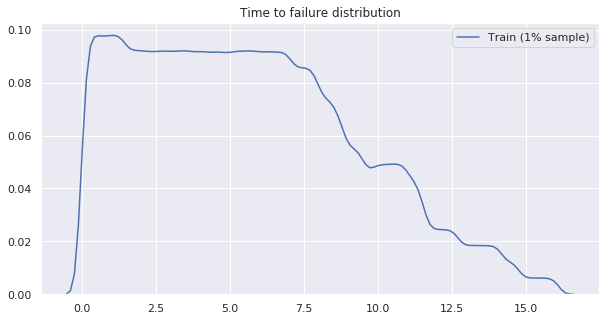
**Fig. 4**

Acoustic data distribution has a very high peak and we see outliers in both directions (Fig. 5).

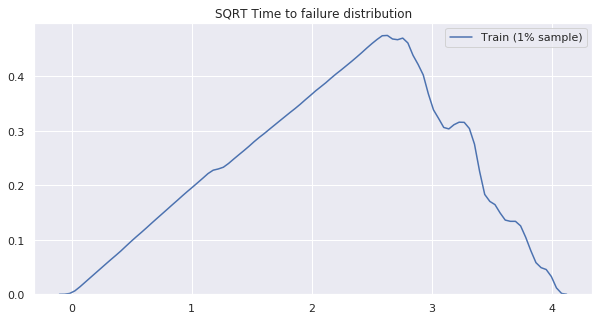
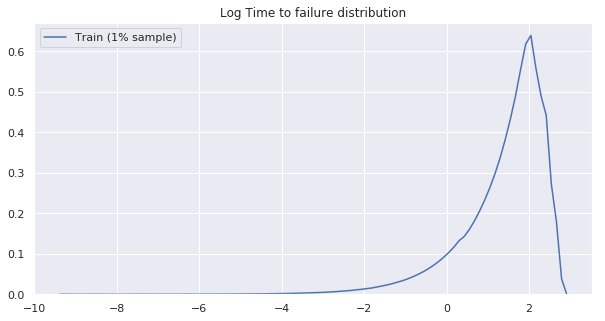


**Fig. 5.** Distribution of Seismic Signal Measurements by LANL

Distribution of time to failure seems to be right skewed (Fig.6). It should be normalized. We compared log and square root transformations visually (Fig. 7 and Fig. 8). The distribution of the time after square root transformation is still not ideally normal, but looks much better than the distribution of the time after log transformation. Log transformed time distribution looks more left skewed than normal. (Fig 8)

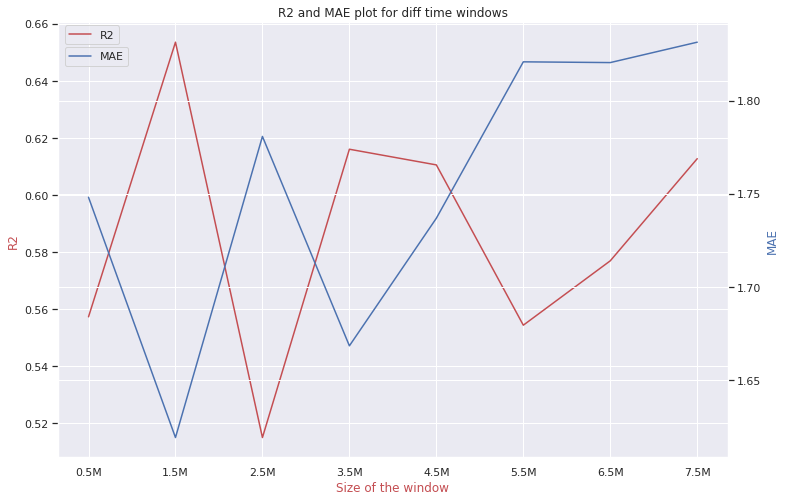


**Fig. 6**

**Fig. 7** **Fig. 8**

# METHODS AND EXPERIMENTS

LANL, working with quasi periodic seismic signals, achieved 0.89 coefficient of determination, using Random Forest technique and dividing data by 1.8 seconds time windows [1]. The most important features in LANL model were variance, kurtosis and threshold. We used a similar approach. Our goal is to predict the time remaining before the next failure using only moving time windows of the acoustic data. We divided our data into time windows that each contain 0.3 seconds of data (1,500,000 observations), which is small enough compared with lab quake cycle that we have 8 to 16 such time widows between each failure. As we discussed above, more than 90% of high acoustic values (absolute value greater than 1000) are around 0.31 seconds before an earthquake. It makes sense to divide our data by 0.3 sec windows to reduce error at the end of the quake cycle. We checked how sensitive our results are to the size of the window, and found that the highest R2 and smallest MAE we were able to get with 1.5M observations in each time window (Fig. 9).



**Fig.9**

So our new data set is 419 time windows. From each time window, we compute a set of 95 potentially relevant statistical features (e.g., mean, variance, kurtosis, min/max, threshold and so on), but using feature importance we found that only 35 were important. Here is the list of 35 features that were used in further analysis:

'std' - standard deviation

'q90' - 90% Quantile

'q95' - 95% Quantile

'q99'- 99% Quantile

'abs\_std'- absolute standard deviation

'ave\_roll\_std\_' - average rolling standard deviation for 100, 1000, 10000 observations

'std\_roll\_std\_' - variance of rolling standard deviation for 100, 1000, 10000 observations

'min\_roll\_std\_'- min rolling standard deviation for 100, 1000, 10000 observations

'q01\_roll\_std\_' - 1% Quantile of rolling standard deviation for 100, 1000, 10000 observations

'q05\_roll\_std\_' - 5% Quantile of rolling standard deviation for 100, 1000, 10000 observations

'q10\_roll\_std\_' - 10% Quantile of rolling standard deviation for 100, 1000, 10000 observations

'q90\_roll\_std\_' - 90% Quantile of rolling standard deviation for 100, 1000, 10000 observations

'q95\_roll\_std\_' - 95% Quantile of rolling standard deviation for 100, 1000, 10000 observations

'q99\_roll\_std\_' - 99% Quantile of rolling standard deviation for 100, 1000, 10000 observations

'std\_roll\_abs\_mean\_' - variance of rolling absolute mean for 100, 1000, 10000 observations

We apply different machine learning techniques such as the Random Forest Regressor, XGB Regressor, Decision Tree Regressor, LGBM Regressor, Extra Trees Regressor to the new continuous values that we got, analysing acoustic time series data.

In order to avoid correlation between new features we applied principal component analysis. Instead of using 35 features, we created just 5 that represented 99.9% of the full data variation.

We use a 50/50 continuous split of the full time series for use as training and testing data sets respectively. Contiguity of train and test data sets is important, since we want to minimize contamination of the training data with information about the test data.

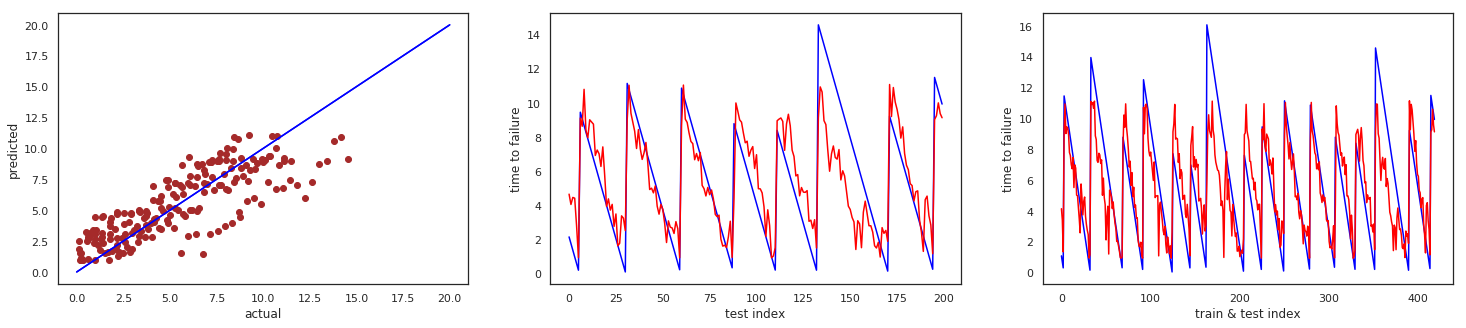
We computed regularization hyper-parameters for each machine learning predicting techniques by random grid search based on a 3-fold cross-validation.

# RESULTS

We run different techniques on a training data set (50% of the full data) before principal component analysis and after. Principal component analysis did not improve our findings significantly. We quantify the accuracy of our model using R2 (the coefficient of determination) and MAE (mean absolute error), applying predicting model on test data set. The best results we achieved using Extra Trees Regressor technique: MAE: 1.61, r2\_score: 0.65.

Hyper-parameters for this technique are 'max\_depth': 25, 'max\_features': 'auto', 'min\_samples\_leaf': 16, 'min\_samples\_split': 30, 'min\_weight\_fraction\_leaf': 0.06225661421078982, 'n\_estimators': 400. The Extra Trees Regressor overview is presented in the section 2.2.

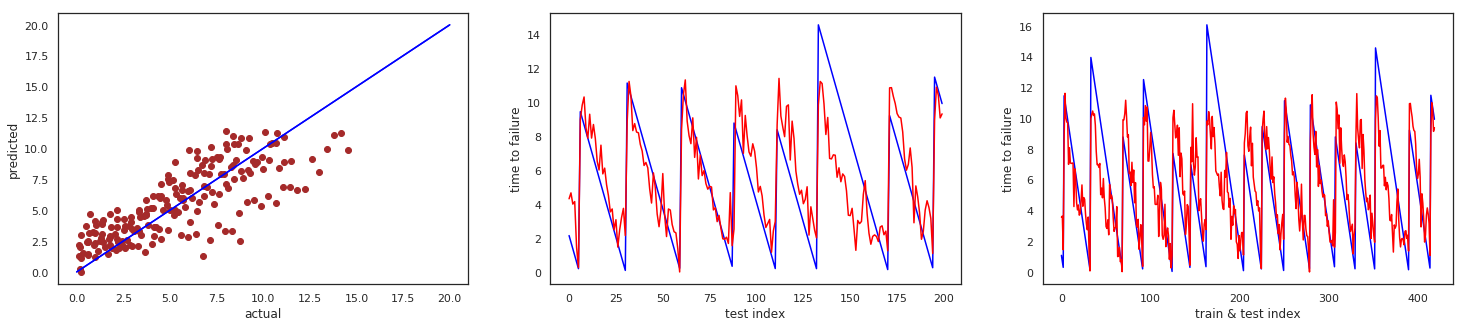
When making a prediction (red curve), we emphasize that there is no past or future information considered: each prediction uses only the information within one single time window of the acoustic signal (Fig. 10)



**Fig. 10**

Results we achieved using AdaBoostRegressor: MAE: 1.67, r2\_score: 0.62 (Fig.11).

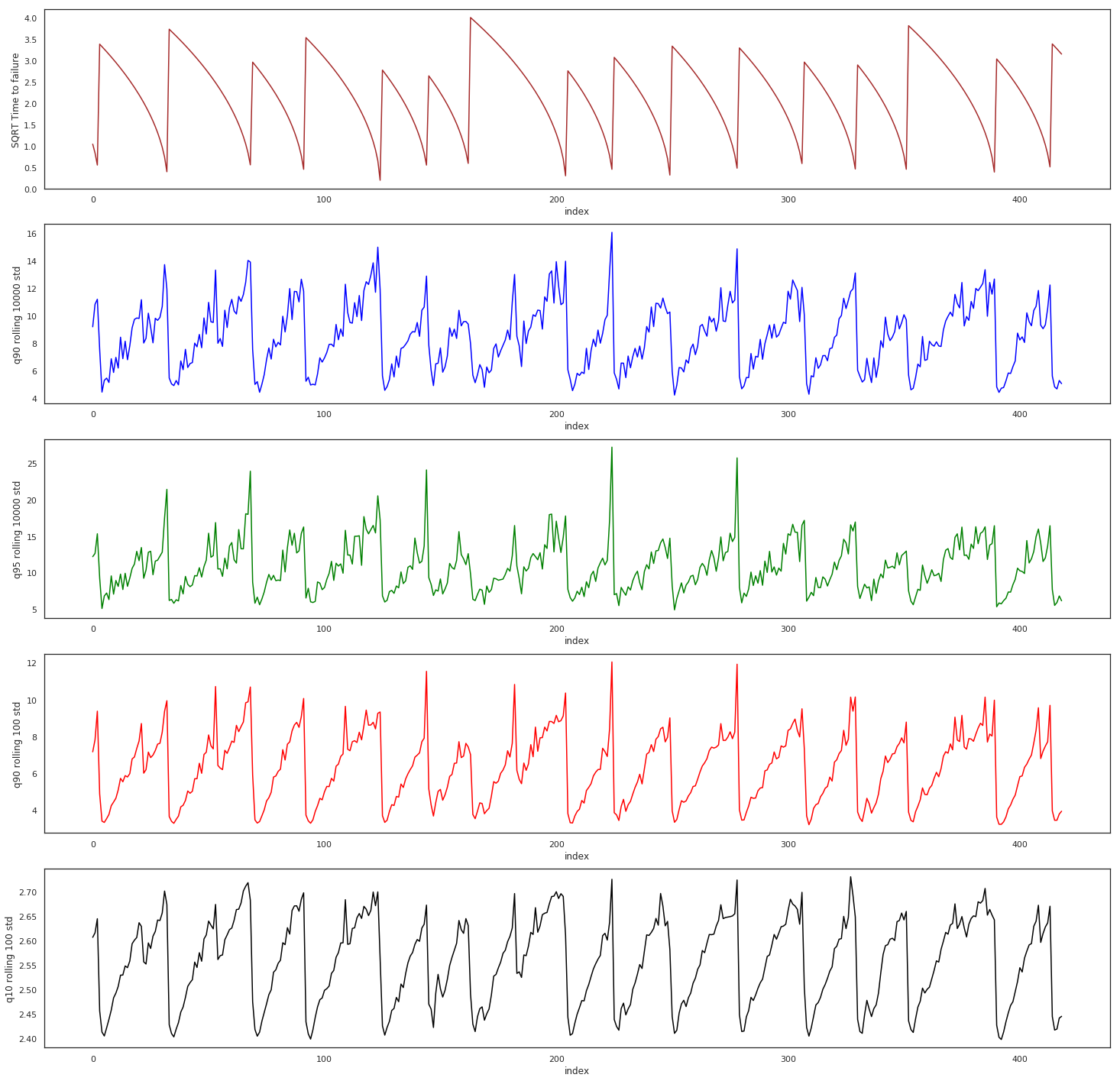
Hyper-parameters for this technique are: 'learning\_rate': 0.018261985736355728, 'loss': 'square', n\_estimators = 500, base\_estimator=Ridge(alpha=1).



**Fig. 11**

# ANALYSIS

The most accurate results with coefficient of determination 0.65 and mean absolute error 1.61 seconds we got using Extra Trees Regressor. The most important features are shown on Fig. 12. Top 5 are rolling standard deviations.



**Fig. 12**

# ETHICS

Scientists’ responsibility to inform the public about their results may conflict with their responsibility not to cause social disturbance by the communication of these results. A study of the well-known Brady-Spence and Iben Browning earthquake predictions illustrates this conflict in the publication of scientifically unwarranted predictions. Furthermore, a public policy that considers public sensitivity caused by such publications as an opportunity to promote public awareness is ethically problematic from (i) a refined consequentialist point of view that any means cannot be justified by any ends, and (ii) a rights view according to which individuals should never be treated as a mere means to ends. [7]

# CONCLUSION (I used LANL conclusion just for now and put our numbers. It is applicable to our work but should be saying in our words. Original LANL conclusion you can see here <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017GL074677>)

In our work we show that ML applied to this experiment provides failure forecasts with 1.61 sec mean absolute error and 0.65 coefficient of determination based on the instantaneous analysis of the acoustic signal at any time in the slip cycle.. These results should suffice to encourage ML analysis of seismic signals in Earth. This is one of the first applications of ML to continuous acoustic/seismic data with the goal of inferring failure times. These results suggest that previous analyses based exclusively on earthquake catalogs may be incomplete. In particular, ML‐based approaches mitigate human bias by automatically searching for patterns in a large space of potentially relevant variables. Our current approach is to progressively scale from the laboratory to the Earth by applying this approach to Earth problems that most resemble the laboratory system. An interesting analogy to the laboratory may be faults that exhibit small repeating earthquakes. For instance, fault patches located on the San Andreas Fault near Parkfield (Nadeau & McEvilly, [1999](https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017GL074677#grl56367-bib-0030); Zechar & Nadeau, [2012](https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017GL074677#grl56367-bib-0042)) exhibit such behavior. Repeaters at these fault patches may be emitting chattering in analogy to the laboratory. If so, can this signal be recorded by borehole and surface instruments? Whether ML approaches applied to continuous seismic or other geophysical data succeed in providing information on timing of earthquakes (not to mention the challenge of predicting earthquake magnitude), this approach may reveal unidentified signals associated with undiscovered fault physics. Furthermore, this method may be useful for failure prediction in a broad spectrum of industrial and natural materials. Technology is at a confluence of dramatic advances in instrumentation, machine learning, the ability to handle massive data sets and faster computers. Thus, the stage has been set for potentially marked advances in earthquake science.

# References

1. Bertrand Rouet-Leduc, Claudia Hulbert, N.L.K.B.C.J.H.P.A.J.: Machine learning predicts laboratory earthquakes
2. Kaggle, R.: Lanl earthquake prediction
3. Ikari Matt J, Marone Chris, S.D.M.K.A.J.: Slip weakening as a mechanism for slow earthquakes
4. Baptiste Rousset, Roland Burgmann, M.C.: Slow slip events in the roots of the san andreas fault
5. USGS: What is an earthquake and what causes them to happen? (March 2019)
6. Rouet-Leduc, B.: Los Alamos machine learning discovers patterns that reveal earthquake fault behavior.
7. Ayhan Sol, Halil Turan: The ethics of earthquake prediction. Volume 10, [Issue 4](https://link.springer.com/journal/11948/10/4/page/1), pp 655–666 (December 2004)