Laboratory Earthquake Prediction

Olha Tanyuk[1], Daniel Davieau[1], Michael L. Blanpied[2], Charles South[1] and Daniel W. Engels[1]

[1] Southern Methodist University, Dallas TX 75205, USA
[2] United States Geological Survey, Fort Worth, TX 76115

Abstract. In this paper we present a method for predicting the timing of laboratory earthquakes using machine learning. If a similar approach can be applied to improve natural earthquake prediction it will save lives. We use data collected from a laboratory experiment which exhibits similar behavior to natural earthquakes. We train a machine learning algorithm using half of the data, then predict using the other half. We compare predicted versus actual timings to measure the algorithm's accuracy. The result shows that the timing of laboratory earthquakes can be predicted up to 16 seconds in advance with 71 percent accuracy. The method and result demonstrates that machine learning can help if it can be scaled from the laboratory experiment to natural earthquakes.

1 Introduction

Earthquakes cause mass destruction and loss of life. Traditional earthquake prediction methods have relied on recurrence interval based models. Recurrence based predictions can only be made within decade spanning time windows and have not been successful. One such model predicted that a magnitude 6 earthquake would occur between 1985 and 1993 in the Parkfield California area but no significant event occurred until 2004 [1].

Researchers imitate natural earthquakes in the laboratory by placing rocky material between steel blocks and applying shear stress to induce slipping. Recent improvements in the instruments [2] used to measure signals have enabled the collection of larger volume data from more realistic and unpredictable laboratory earthquakes. However processing the data and detecting patterns in it has become more difficult to work with. In this paper we demonstrate that machine learning can be used to detect patterns in the more realistic data and predict laboratory earthquakes.

We use data which was collected by the Los Alamos National Laboratory and provided to the public via a Kaggle competition [4]. The data consists of 629 million acoustic signal observations and an accompanying record of the time remaining until a laboratory earthquake (failure) occurred [2]. We calculate additional statistical measures such as variance, kurtosis and skew for each observation. We use half of the data to train a machine learning algorithm. With the remaining half of the data, using only the acoustic signal as input we use the algorithm to calculate a prediction of the time remaining until failure. We measure the predictions accuracy by comparing to the actual remaining time to failure from the original data.

The result shows that the timing of laboratory earthquakes can be predicted up to 16 seconds in advance with 71 percent accuracy.

The data, hardware and software allows us to predict impending earthquakes accurately. However we only know 8-16 seconds before failure. Therefore practical applications may be limited. This may prove useful but only applies to laboratory experiments. This could be used in industry perhaps in researching materials for wallboard, machine parts and others.

2 Background

In August 2017 the Los Alamos National Laboratory (LANL) conducted an experiment [2] which predicted *laboratory* earthquakes with 90% accuracy. The team imitated an earthquake using steel blocks interacting with rocky material to induce slipping that emitted seismic signals. An accelerometer was used to record them. A random forest algorithm was trained on the signals and stick-slip failures (laboratory earthquakes). The trained algorithm was then used to generate predictions from separate signals (not used in the training). The predictions were measured against actual failures to determine the accuracy of the predictions [3].

We hear about earthquakes usually via news media when there is a large seismic event which is noticeable and causes death and destruction. These are stick–slip events that radiate seismic energy along the faults 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 move slowly but can get stuck at their edges due to friction. Stress builds gradually over time until it overcomes the friction resulting in a slip between the tectonic plates. Energy is released in waves that travel through the earth's crust. This is the shaking that we feel and know as an earthquake [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. We cannot feel slow slip earthquakes shake the ground. They occur near the boundaries of large earthquake rupture zones [6].

LANL researchers discovered a way to predict earthquakes in a laboratory experiment that simulates natural conditions. In 2017 they discovered a way to train a computer to pinpoint and analyze seismic signals emitted during movements along faults to predict an earthquake. They processed massive amounts of data and identified a particular sound pattern previously thought to be just white noise that preceded an earthquake. The team was able to characterize the time remaining before a laboratory earthquake at all times [3].

For the first time they discovered a pattern that accurately predicted when a laboratory earthquake would occur. The team acknowledged that the physical traits of the lab experiment 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 [3].

The results revealed that the faults failed in a predictable manner. The observations also demonstrated that the fault's critical stress state can be determined using exclusively an equation of state [3].

Traditionally scientists have relied exclusively on historical data to characterize the state of faults and predict regular earthquakes. Only a small fraction of seismic data had typically been collected and some was discarded during analysis as noise. The LANL team discovered that hidden in this noise are signals emitted by the fault that inform them of the state of the fault much more precisely [3].

> "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."

This suggests that there is a relationship between slow slip earthquakes and more noticeable regular earthquakes [7].

This *new* study is an analysis of the relationship between slow slip and regular earthquakes. We use this relationship information to predict regular laboratory earthquakes.

3 Data

The data we analyze is provided by LANL's 2019 Kaggle competition [4]. It was collected using a three-block assembly with two gouge layers placed in a bi-axial stress configuration. Two 5mm thick fault gouge layers were placed between the three blocks, which were held in place by a fixed normal load. The gouge material was comprised of beads with diameter 105-149 mm. The central block was sheared at constant displacement rate. The two data streams recorded were the shear stress and the acoustic signal. While the gouge material was in a critical shear stress regime, the shear stress abruptly dropped which indicated gouge failure(a laboratory earthquake). As applied load progressively increased, the recurrence of laboratory earthquakes progressively decreased. At smaller applied loads the slips became a-periodic. The acoustic particle acceleration was measured on the central block.

LANL is certain the signal recorded for analysis is the acoustic signal emanating from the fault

The acoustic data are integers ranging from -5515 to 5444 and have mean of 4.52. The time to failure is in seconds. We can see that acoustic data shows large fluctuations just before the failure and recurs cyclically. Failures can be predicted visually as cases when huge fluctuations in the signal are followed by

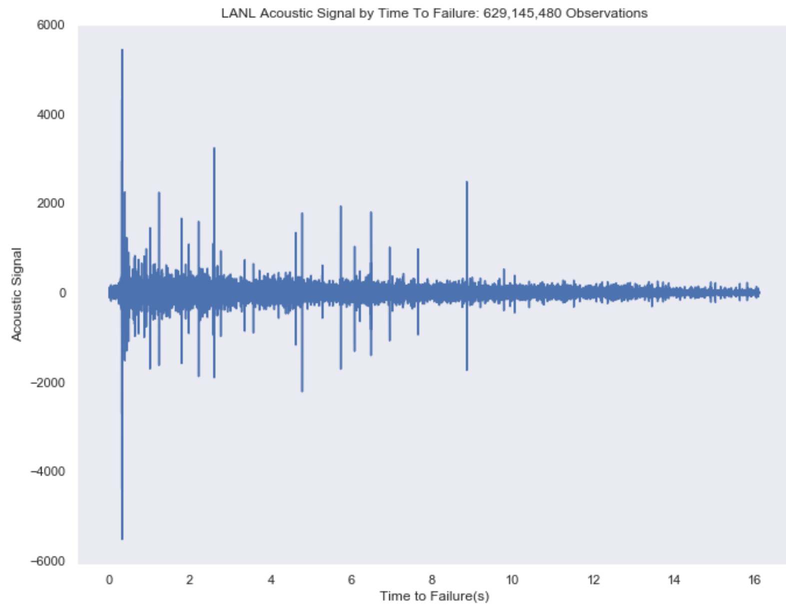


Fig. 1. The magnitude of each seismic signal and the time remaining before the next laboratory earthquake. There are 16 lab earthquakes. The shortest time to failure is 1.5 seconds for the first earthquake and 7 seconds for the 7th, while the longest is around 16 seconds.

Table 1. Sample of Data Provided

Index	Seismic Signal	Time to Failure
0	12	1.469099998474121
1	6	1.469099998474121
2	8	1.469099998474121
3	5	1.469099998474121
4	8	1.469099998474121

smaller signal values. This could be useful for predicting time to failure changes from 0 to high values.

4 Methods and Experiments

Our goal is to predict the time remaining before the next failure using only moving time windows of the acoustic data. We divide our data into windows containing 150,000 observations each (0.0375 seconds of seismic data)therefore

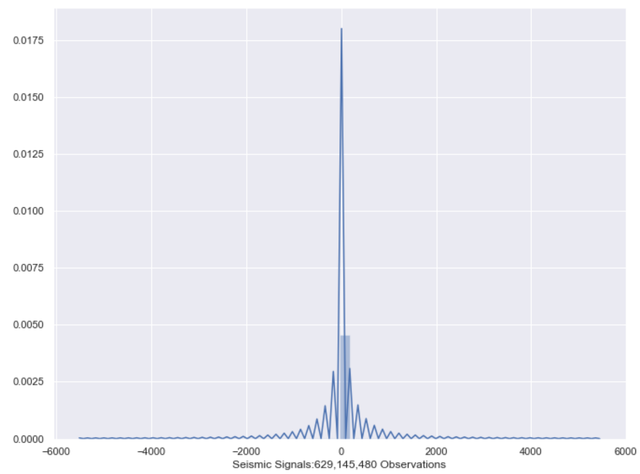


Fig. 2. The distribution of seismic signal measurements by LANL

Table 2. Seismic Signal Stats

Index	Seismic Signal
count	6.29145480
mean	4.47708428
std	2.61278939
min	9.55039650
25	2.62599707
50	5.34979773
75	8.17339516
max	1.61074009

4194 windows. From each time window, we compute a set of 98 potentially relevant statistical features (e.g., mean, variance, kurtosis). We apply machine learning techniques such as the Random Forest Regressor, XGB Regressor, Decision Tree Regressor, LGBM Regressor, Extra Trees Regressor to the new continuous windows of values.

Similar to the LANL study we create new features separated into two main classes: Distribution of signal's energy: we use couple of higher order moments of the acoustic data to capture the evolution of the signal's energy. Within each time window we compute the signal's normalized mean, minimum, maximum and higher moments (variance, skewness, kurtosis).

Precursors: the system enters a critical state when close to failure. We rely on different percentiles and thresholds to monitor this precursory activity. We use

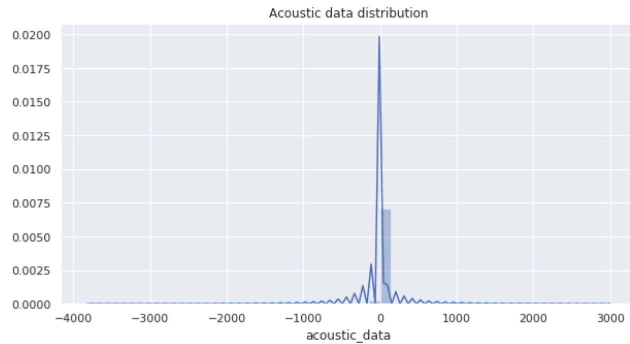


Fig. 3. 1% random sample from 629,143,480 observations

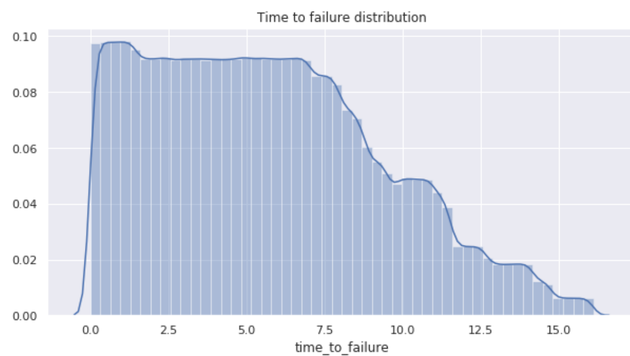


Fig. 4. The min value is very close to zero and the max is 16 seconds.

the 1st - 9th and 91th - 99th percentiles. Our thresholds measure the count of observations that the acoustic signal spends over a threshold value f0 and under a threshold value f1.

In order to avoid correlation between new features we applied principal component analysis. Instead of using 98 features, we created just 10 that represented 99.9 percent of the full data variation. We use a 70/30 random split of the full time series as training and testing data sets respectively. We compute regularization hyper-parameters for each machine learning predicting technique by random grid search based on a 3-fold cross-validation.

4.1 Data Transformations

The distribution of time to failure is right skewed. We apply a square root transformation to normalize it and improve the prediction models. It is still not ideally normal, but has improved.

Fig. 5. We checked how both variables changed over time. The red line is the acoustic data and the blue one is the time to failure. On a plot above we can see, that training data has 16 earthquakes. The shortest time to failure is 1.5 seconds for the first earthquake and 7seconds for the 7th, while the longest is around 16 seconds.

4.2 Algorithms

4.3 Recursive Neural Network

LSTM is a type of RNN that helps us deal with vanishing or exploding gradients (incorrect slope). For example most of our acoustic signal observations are within x however more extreme signals occur just before failure. The extreme differences between these observations can skew the accuracy of more traditional RNN algorithms.

LSTM separates the long term(fault slip) from short term (slow slip) Hidden state= memory from prior observations Typical RNN input+priorhiddenstate¿tanh (-1,1) activation function¿new hidden state LSTM is same but includes gates to determine what information is included in hidden state and what is not.

The gates are rnn's themselves.

input+priorhiddenstate¿sigmoid (0,1) activation function¿new hidden state. 0, 1 allows us to forget or remmeber Patrick Yam: LSTM cannot handle data with 150000 sequence length therefore we wavenet in the earlier layers as feature extraction and reduce the sequence length to 150.

Forget gate decides what to include 1, or exclude 0. https://www.youtube.com/watch?v=2GNbIKTKCfE https://www.youtube.com/watch?v=8HyCNIVRbSU

4.4 Auto Regressive integrated Moving Average (ARIMA)

The data shows evidence of non-stationarity (the mean, variance change over time). We use an ARMIA model to analyze the changing means and variance. This this model also can also account for white noise.

ARIMA(0,0,0)

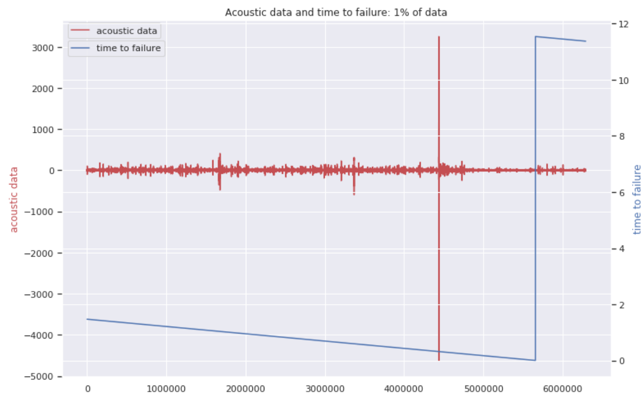


Fig. 6. On this zoomed-in-time plot we can see that actually the large oscillation before the failure is not quite in the last moment. There are also trains of intense oscillations proceeding the large one and also some oscillations with smaller peaks after the large one. Then, after some minor oscillations, the failure occurs. Interesting thing to check is the time between high levels of seismic signal and the earthquakes. We are considering any acoustic data with absolute value greater than 1000 as a high level

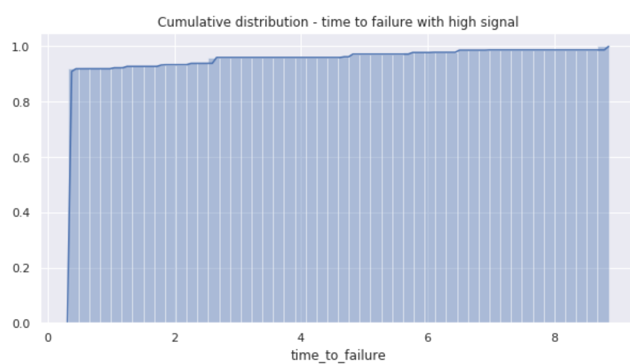


Fig. 7. More than 90% of high acoustic values are around 0.31 seconds before an earthquake

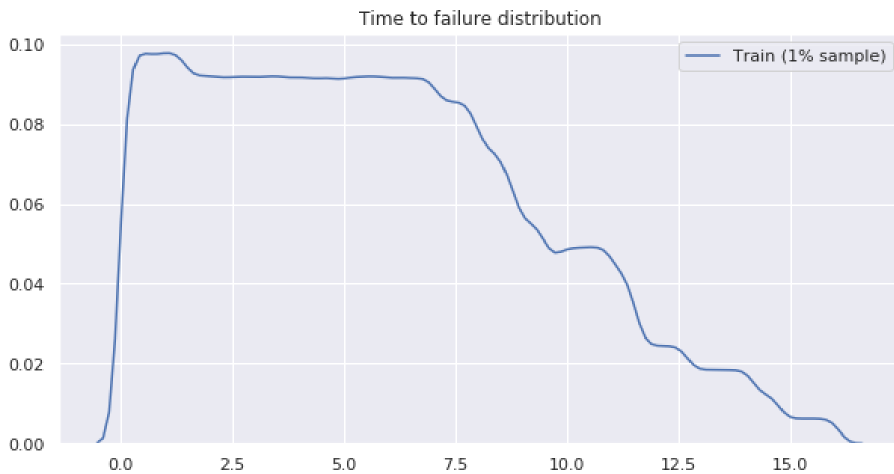


Fig. 8. Distribution of Time to Failure showing right skew

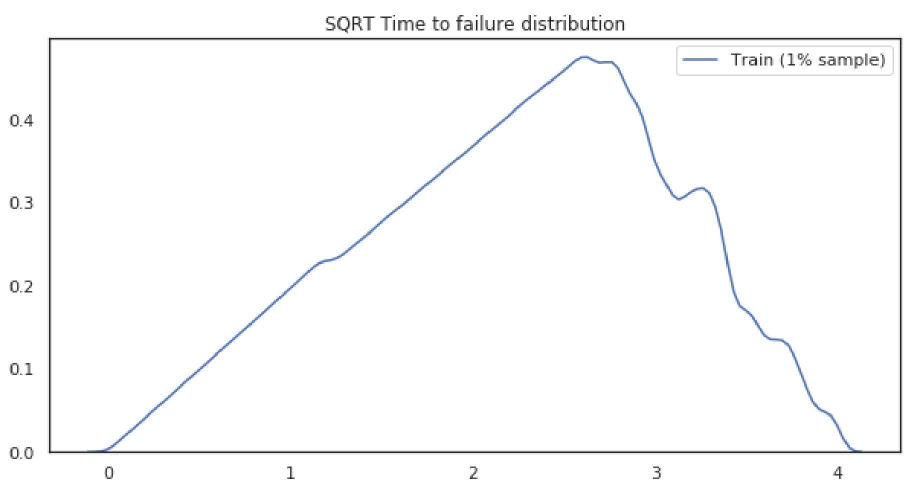


Fig. 9. Distribution of Time to Failure showing improved normality after applying square root

4.5 Gradient Boosting Decision Tree

5 Results

We run 5 different techniques on a train data set (70 percent of the full data) before principal component analysis and after. Principal component analysis did not improve our findings significantly, that is why we are not represent those results here. For each model we provide hyper-parameters details for future reproducibility (Table 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. We quantify the accuracy of our model using R2 (the coefficient of determination) and MAE (mean absolute error), applying predicting model on a 30 percent of the full data (test data).

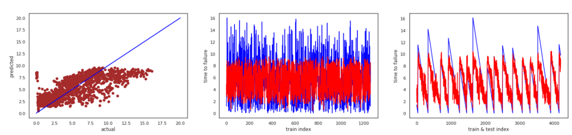
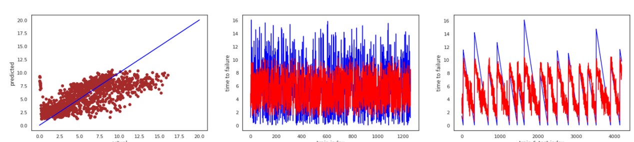
Method	Results
RandomForestRegressor 'min_samples_split': 12, 'min_samples_leaf': 4, 'max_features': 'sqrt', 'max_depth': 8 n_estimators=1000	mean_absolute_error: 2.0322193397002044 r2_score: 0.4953230388182709 
XGBRegressor 'colsample_bytree': 0.9059046706163597, 'gamma': 4.395361343977617, 'learning_rate': 0.22367318773362826, 'max_depth': 21, 'min_child_weight': 69.30113404571125, 'reg_alpha': 7.681525516492065, 'subsample': 0.8919123509453631 nthreads=-1, n_estimators=1000	mean_absolute_error: 2.0634676560951557 r2_score: 0.48347777795889124 

Fig. 10. Results by Model

6 Analysis

The most accurate results with coefficient of determination 0.5 and mean absolute error 2.03 we got using Random Forest Regressor. The most important

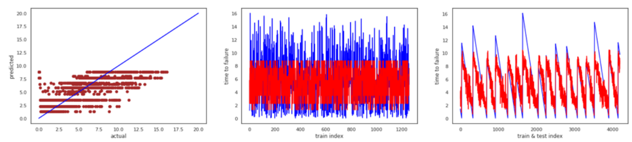
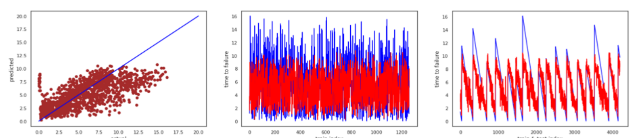
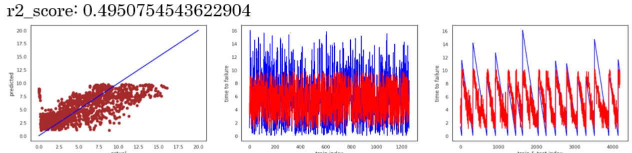
DecisionTreeRegressor 'max_depth': 4, 'max_features': 'sqrt', 'min_samples_leaf': 12, 'min_samples_split': 70	mean_absolute_error: 2.066672622791363 r2_score: 0.4703327918288054 
lgb.LGBMRegressor 'feature_fraction': 1, 'lambda_l1': 0.6, 'lambda_l2': 0.4, 'learning_rate': 0.05, 'max_depth': 8, 'min_data_in_leaf': 10, 'min_gain_to_split': 0.01, 'num_leaves': 44, 'random_seed': 13, 'subsample': 0.9 n_estimators=10000	mean_absolute_error: 2.089058482904736 r2_score: 0.43881088754719744 
ExtraTreesRegressor 'max_depth': 11, 'max_features': 'auto', 'min_samples_leaf': 40, 'min_samples_split': 200, 'min_weight_fraction_leaf': 0.010872755116371424, 'n_estimators': 200	mean_absolute_error: 2.0338440361453003 r2_score: 0.4950754543622904 

Fig. 11. Results by Model

features, shown on Fig. 7, suggest to check our data for correlated features, since rolling standard deviations show the same patterns. Plus in order to improve our predicting model, we need to get rid of features that have importance equal to 0 or close to 0, such as hypermoment, minimum, maximum and some of the rolling means. Currently we are working on improving of the model to make prediction for the beginning and the end of the one quake cycle. As you can see on a Fig. 1, seismic waves with small voltage are in our interest. Random forest model accurately predicts failure across load level, but hardly can predict outliers. It means that we still need to think about new features that we missing. Precursors that we choose during feature engineering were not good enough when the system enters a critical state close to failure.

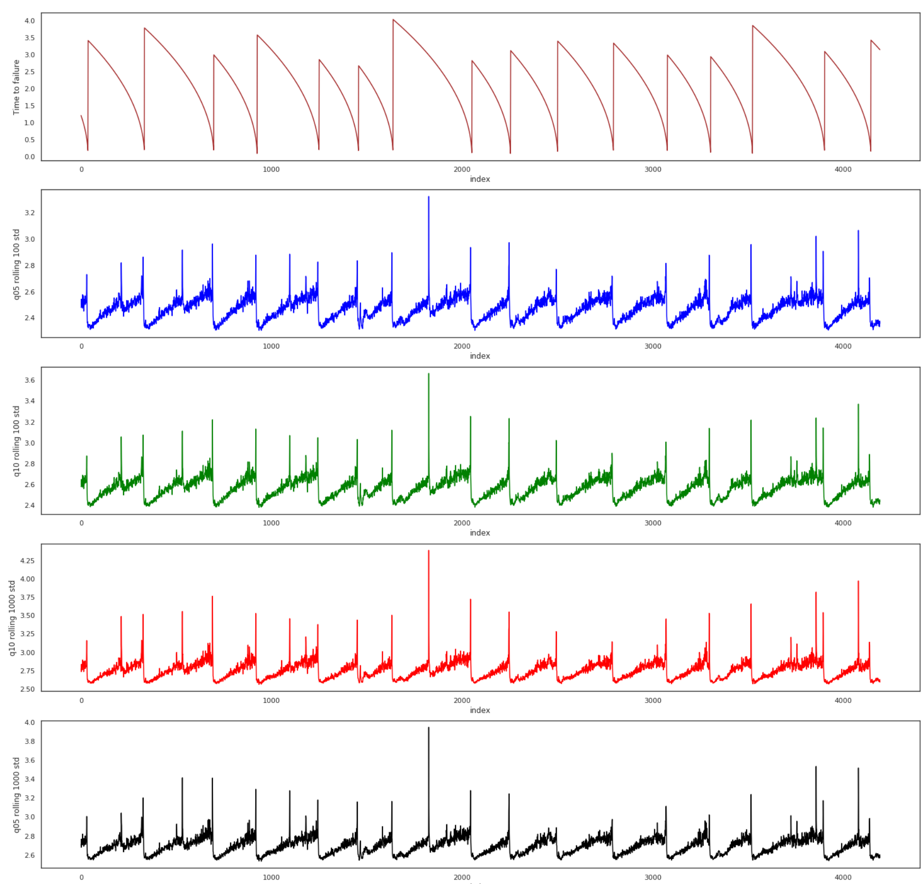


Fig. 12. Predictions by Model

7 Ethics

If people believe us and we are wrong; bad things can happen. If people believe us and we are right; good and bad things can happen.

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 [8].

8 Conclusions

The evidence in this new *observational* study suggests that *(Measure against prior study TBD. We have to state the fact that we cannot prove that the data collection technology improvements are better or that our algorithm is better unless we can accurately reproduce the prior model's results. Consider confounding variables such as different collection methods and unavailable details about the 2017 model.)*

Our application of machine learning predicts laboratory earthquakes with 2.03 seconds mean absolute error and 0.5 coefficient of determination. This is based on the analysis of the acoustic signal every moment in the slip cycle. However this is only 8-16 seconds in advance therefore practical applications may be limited. This only applies to laboratory experiments and not real life threatening earthquakes.

Machine learning searches for patterns that humans may not easily perceive. This method may be useful if LANL can scale it from the lab to the field. It can also be useful in researching industrial and natural materials.

References

1. Jackson, D.D.: The 2004 parkfield earthquake, the 1985 prediction, and characteristic earthquakes: Lessons for the future. Bulletin of the Seismological Society of America **96**(4B) (2006) 397–409 https://doi.org/10.1785/0120050821.
2. Rouet-Leduc, B.: Machine learning predicts laboratory earthquakes. Geophysical Research Letters **44**(18) (August 2017)
3. Rouet-Leduc, B.: Los Alamos machine learning discovers patterns that reveal earthquake fault behavior. Los Alamos National Laboratory, P.O. Box 1663 Los Alamos, NM 87545. Science briefs 2018 edn. (March 2018)
4. Kaggle: Lanl earthquake prediction. https://www.kaggle.com/lanl (March 2019)
5. Kious, W.J.: This dynamic earth: the story of plate tectonics. Rprt, U.S. Geological Survey, http://pubs.er.usgs.gov/publication/7000097 (1996)

6. Ikari, M.J.: Slip weakening as a mechanism for slow earthquakes. Nature Geoscience **6** (May 2013) 468 https://doi.org/10.1038/ngeo1818.
7. Rousset, B., Bürgmann, R., Campillo, M.: Slow slip events in the roots of the san andreas fault. Science Advances **5**(2) (2019)
8. Ayhan Sol, H.T.: The ethics of earthquake prediction. Science and Engineering Ethics **10**(4) (December 2004) 655–666