Connor Bailey and Megan Eichelberger

Applying Machine Learning Techniques to Predict Future Large Magnitude Earthquakes Using Historical Magnitude and Depth Data

Abstract:

Earthquakes are some of the most significant natural disasters that humans have to navigate. They can leave massive casualties and damages that can impact countries for decades. In this paper, we aim to use historical earthquake data to predict earthquakes in the Philippines with a magnitude of greater than 5.5. The study is influenced by the article “Earthquake magnitude prediction in Hindukush region using machine learning techniques” [4]. We used earthquake data from the IRIS (Incorporated Research Institutions for Seismology) database ranging from 2001-2021 [12]. We observed data from smaller magnitude earthquakes (defined by the USGS as less than 5.5) to evaluate precursors that can lead to higher magnitude earthquakes (greater than 5.5 magnitude).

Utilizing neural networks and survival analysis on depth and magnitude of low magnitude earthquakes we were able to predict with high accuracy whether a region would experience higher magnitude earthquakes later, although the timing thereof proved inconclusive.

Introduction:

Earthquakes have affected mankind throughout history, and predictive methods have been developed from seismometers to observations of slug trails. The earliest recorded seismometer was developed in 132 A.D. by Zhang (sometimes Chang) Heng - a Chinese astronomer and philosopher [10]. This seismometer detected not only earthquakes, but their origin direction. Utilizing an urn with eight dragons, each with a ball in their mouth. When the ball dropped, the observer could identify that there was an earthquake and generally where it came from. While the device does not survive, historians are attempting to recreate the device. Two examples are presented in Figure 1. This illustrates the purpose of a seismometer - a device not to predict but to detect and record.

Figure 1a: One potential replica of the first seismometer
Science and Society Picture Library

Figure 1b: One potential replica of the first seismometer
Michigan Tech

Developed in 1935, the Richter Scale is a method to measure in layman's terms the earthquake's size, or Magnitude, as read by a seismometer. Magnitude is used to measure the amount of seismic energy that is released by an earthquake [5]. The numeric value is then binned and translated to easily understandable terms. Although there are other factors that contribute to damage such as infrastructure, the general scale is represented in Table 1 [8]. While significantly more sophisticated than Zhang Cheng's device, the information is still to detect and record rather than predict.

Magnitude	**Effects**
2.5-5.4	Felt, but slight damage
5.5-6.0	Slight damage to buildings or other structures
6.1-6.9	May cause significant damage in populated areas
7.0-7.7	Major earthquake with serious damage

Table 1 shows the breakdown of the Richter scale and corresponding consequences

The scientific community is divided on whether earthquakes can be predicted with confidence. While some assert that earthquakes cannot be predicted [1], other research demonstrates that they can be predicted through the analysis of precursory phenomena variations of electric fields, magnetic fields [3], animal behavior analysis [2], and historical earthquake data [13]. However, there is so little consensus that in 2012 six Italian scientists were convicted of manslaughter for telling citizens that earthquakes were 100% predictable and they had nothing to fear. Convinced of their safety, residents of L'Aquila, Italy did not evacuate and 309 people died in a 2009 earthquake. The scientists were exonerated in 2014 [11].

The Philippine Islands are located near some of the most likely earthquake regions such as Indonesia and Southeast Asia. From 2001-2021, there have been 443 earthquakes with a magnitude of greater than 5.5, which is the threshold for a 'significant' earthquake. This is due to its location in the Circum-Pacific belt, or "Ring of Fire", shown in Figure 2 [6,7]. This belt forms a horseshoe shape around the Pacific Ocean, following the edge of the Pacific plate.The Philippine plate and Pacific plate form a subduction zone (Figure 3). This occurs when one tectonic plate goes under another, causing volcanoes and earthquakes.

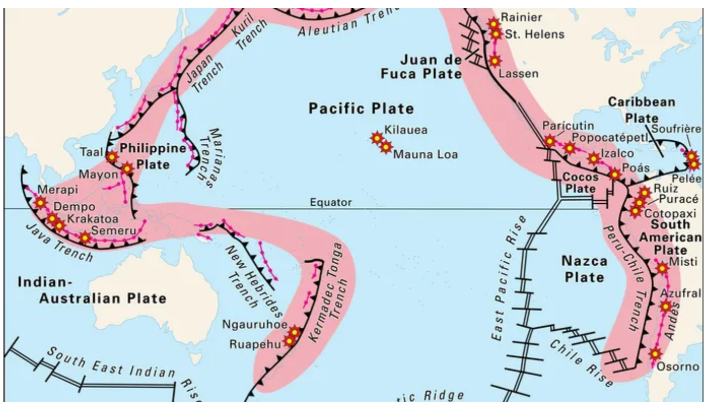

Figure 2: the Circum-Pacific Belt, responsible for approximately 90% of the world's earthquakes
Encyclopedia Britannica [7]

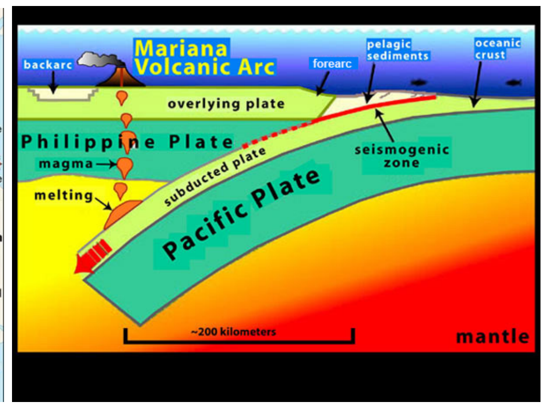

Figure 3: Philippine Subduction Zone
NOAA [9]

With the frequency of earthquakes in the region and the devastation that they cause, It would be a major step forward if there was a way to accurately predict high magnitude earthquakes. We apply machine learning techniques as outlined in section two to the problem, searching for a way to predict the next big earthquake to mitigate the devastation and loss of life.

Methods

We initially spent time exploring our datasets to learn a bit more about earthquakes. We also wanted to visualize some of the earthquake data in the Philippines specifically. First, we created Figure 5, which is a map of high magnitude earthquakes in the Philippines. There have been 443 major earthquakes in the Philippines from 2001-2021. Figure 5 displays the earthquakes, shaded according to their magnitude. This visual shows high magnitude earthquakes, as the sheer volume of lower magnitude earthquakes made the visual illegible.

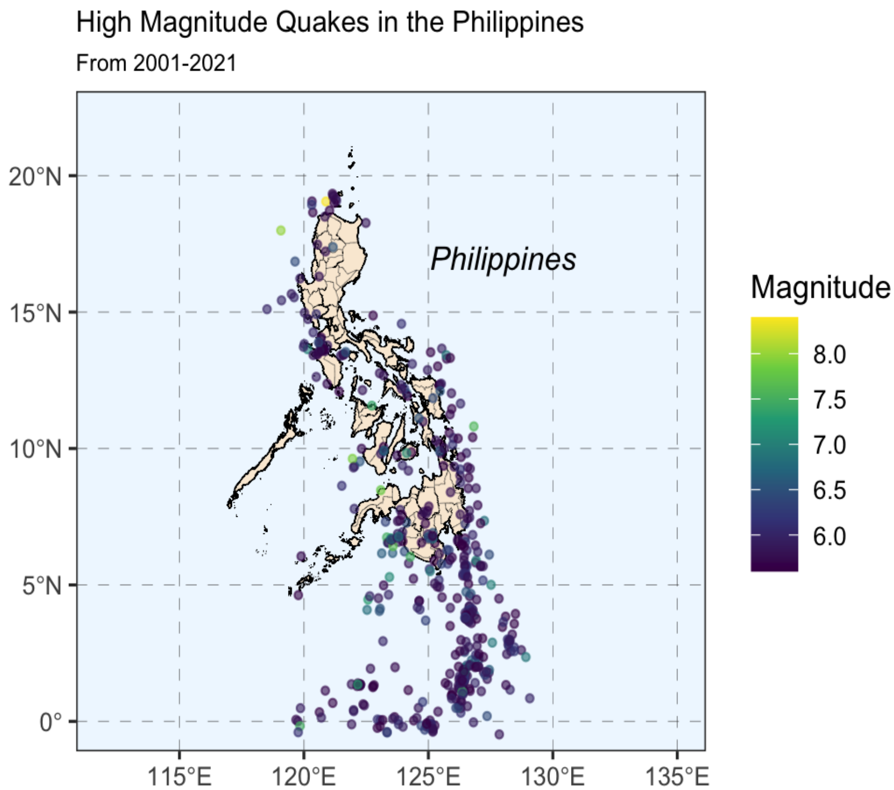

Figure 5: High magnitude earthquakes in the Philippines and surrounding areas since 2001. The points are mapped by magnitude. The brighter the color, the higher the magnitude

The dataset also contained a column that contained the depths of the earthquakes. Depths are broken down into three categories: shallow, intermediate, and deep. Intuitively, it may seem that higher magnitude earthquakes are more likely to be intermediate or deep. However, shallow earthquakes are by far the most destructive type. Shallow earthquakes tend to be more destructive seismic waves due to quakes traveling less distance to the surface than deep earthquakes, therefore losing much less energy, often resulting in more shaking. Figure 6 shows the distribution of depths.

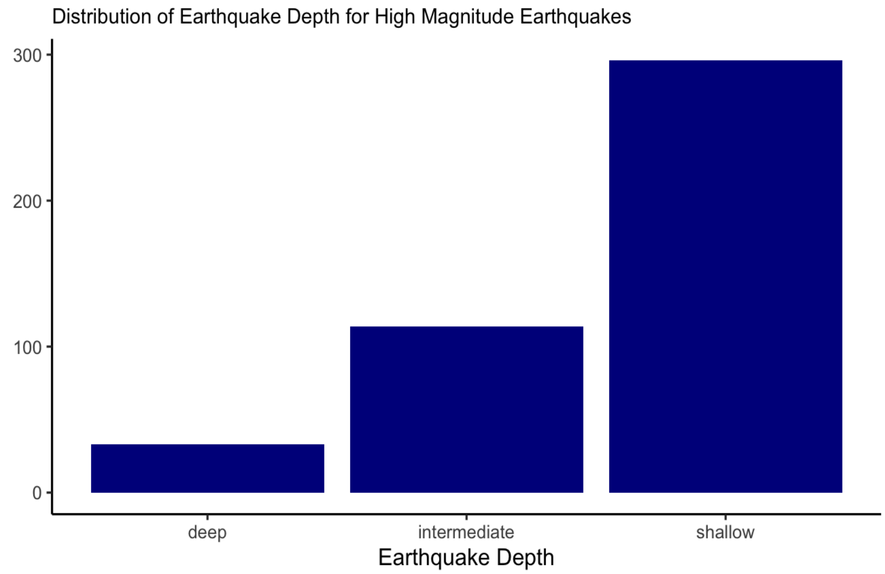

Figure 6: distribution of earthquake depths on high magnitude earthquakes in the Philippines and surrounding regions

We also looked at densities to see where there are spikes in either the depth of the earthquake or the magnitude. As seen in figure 7, there is a spike in shallow quakes seen in the Northern Molucca Sea just south of the Philippines. Our original capture also included part of Malaysia, and a spike in density for low magnitude quakes can be seen in Borneo.

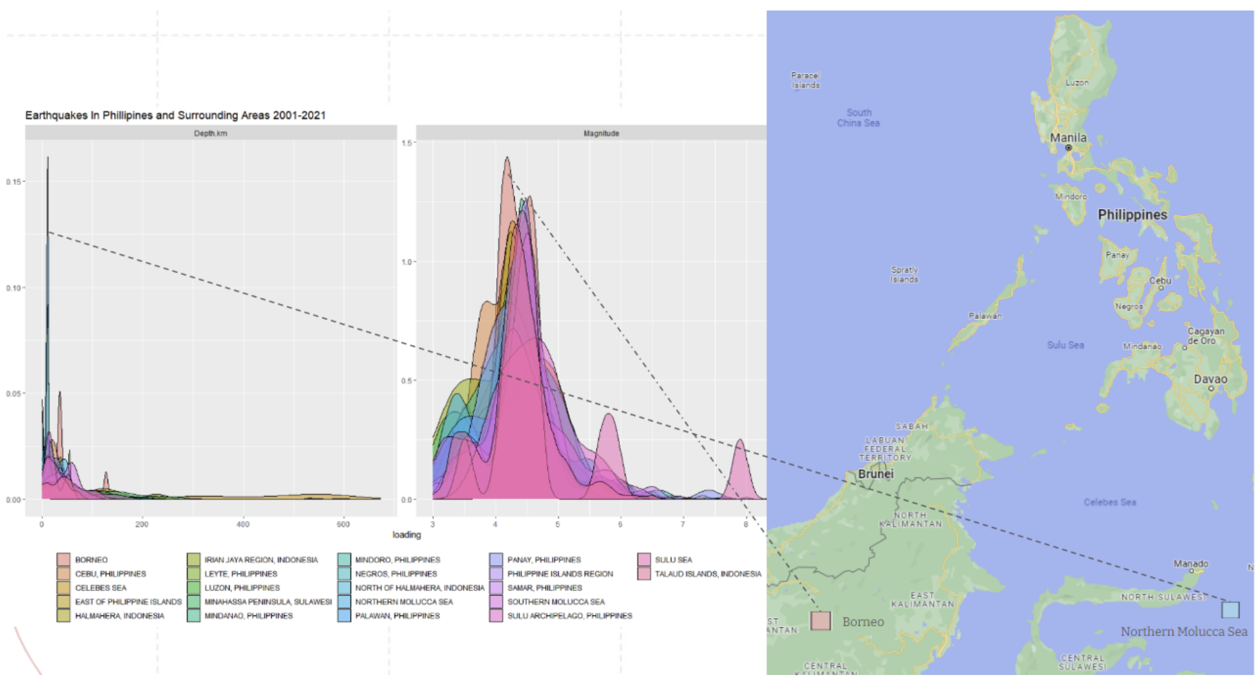

Figure 7: Density of Earthquakes in the Philippines and surrounding areas

From this visual, there does not immediately appear to be a correlation between the magnitude and depth of the earthquakes. This is borne out through our correlation matrix in figure 8. This figure was generated by comparing the depth and magnitude of earthquakes with a magnitude below 5.5 against a boolean identifying if there was a high magnitude earthquake in that region that occurred after the low magnitude earthquake. We can see that there is a high correlation between the magnitude, but not necessarily with the depth of the preceding quakes.

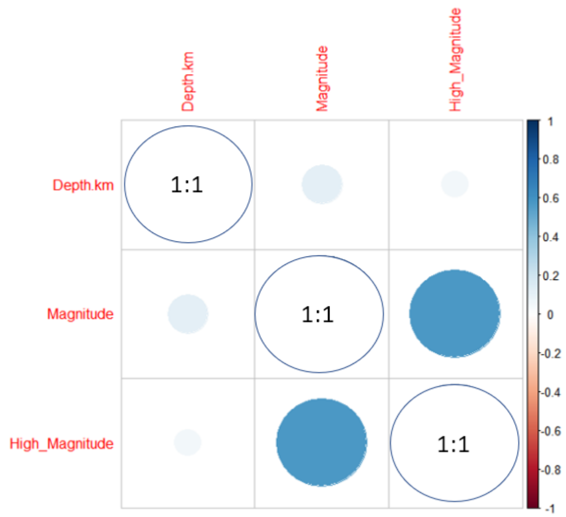

Figure 8: Correlation matrix using Magnitude and Depth of low magnitude Earthquakes to identify variable interdependence

Analysis

Neural networks are utilized extensively in the natural sciences due to the ability to execute complex pattern recognition routines quickly. This is helpful to not only analyze patterns, but to identify if a pattern exists at all. The network was trained on a subset of precursor data then tested on the remainder. Initially we planned to use this model for classification exclusively, but due to the outcomes discussed below we ran the data on additional models as well.

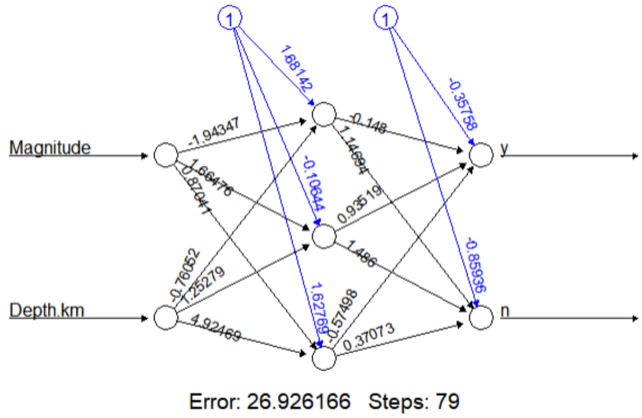

Figure 9: Neural network using Magnitude and Depth of low magnitude Earthquakes to predict whether or not large earthquakes happened later

We also utilized a random forest model, a decision tree that works through an iteration of yes/no decisions that lead to the most accurate outcome. A k nearest neighbor (KNN) model was applied as well. KNN is useful for identifying classifications: in this case whether a low magnitude earthquake was followed by a high magnitude earthquake in the same area later on.

While these models are adept at classification and predicting classification, they don’t necessarily do well for predicting when a future large earthquake might occur. For this model, we used what is called a survival probability model. This model is often used in medical testing for test group survival dropoff. In this case, we applied the model to the dates of the earthquakes and used it to predict the future trajectory. A survival probability of 1.0 says that there will be an earthquake during that time.

Results

There was a surprising return with the neural network - a kappa of 1. This denotes that the model was able to predict whether or not there was a future earthquake with near perfection. That high of a kappa is significant, and may speak to an issue in the model. For verification of findings we ran a random forest, which as well returned a kappa of 1. Next our KNN model, which was less accurate with a kappa of .67. While that is lower than the perfect 1 of the other models, a .67 is well in the bounds of substantial agreement.

Next we look to predict with some degree of certainty the timing of the large earthquakes. The survival plot in figure 10 failed to predict dates based on the provided criteria. Figure 10 shows the expectation that the ability to predict that there will be a future earthquake remains steady at least through the year 3119. There did not appear to be any predictive power for the actual date of a large earthquake following smaller ones.

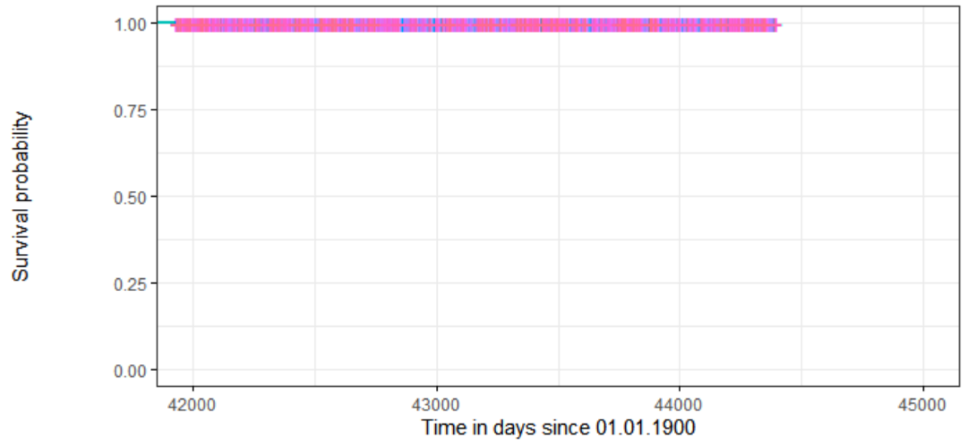

Figure 10: Predicted future probability that low magnitude earthquakes will precede high magnitude earthquakes

Conclusion

In light of the analysis performed in the scope of this study, it is highly unlikely that the date or time of a high magnitude earthquake can be predicted with any degree of certainty. The prediction that a high magnitude earthquake will follow a low magnitude earthquake is certain, but ultimately lacks significance. In an area with extensive seismic activity, the fact that there are low magnitude earthquakes in precedence of high seems to be correlative but does not denote causation. This paper has a relatively narrow scope with regards to predictors, and more research is needed with additional precursors. However, this analysis does uphold the conjecture that the timing of large earthquakes cannot be predicted.

Works cited

[1] Geller RJ, Jackson DD, Kagan YY, Mulargia F (1997) Enhanced: earthquakes cannot be predicted. Science 275(5306):1616–1620

[2] Grant RA, Raulin JP, Freund FT (2015) Changes in animal activity prior to a major (m = 7) earthquake in the Peruvian Andes. Phys Chem Earth Parts A/B/C 85:69–77

[3] Pulinets S, Ouzounov D (2011) Lithosphere-atmosphere-ionosphere coupling (LAIC) model-an unified concept for earthquake precursors validation. J Asian Earth Sci 41(4):371–382

[4] Asim, K. M., Martínez-Álvarez, F., Basit, A., & Iqbal, T. (2016). Earthquake magnitude prediction in Hindukush Region using machine learning techniques. *Natural Hazards*, *85*(1), 471–486. https://doi.org/10.1007/s11069-016-2579-3

[5] *CEA - earthquake measurements: Magnitude vs intensity*. How are Earthquakes Measured? Magnitude & Intensity Scales | CEA. (n.d.). Retrieved July 26, 2022, from https://www.earthquakeauthority.com/Blog/2020/Earthquake-Measurements-Magnitude-vs-Intensity

[6] *Earthquake glossary*. U.S. Geological Survey. (n.d.). Retrieved July 29, 2022, from https://earthquake.usgs.gov/learn/glossary/?term=Ring+of+Fire

[7] Encyclopædia Britannica, inc. (n.d.). *Ring of fire*. Encyclopædia Britannica. Retrieved July 29, 2022, from https://www.britannica.com/place/Ring-of-Fire

[8] *How are earthquakes studied?: UPSeis: Michigan Tech*. Michigan Technological University. (n.d.). Retrieved July 29, 2022, from https://www.mtu.edu/geo/community/seismology/learn/earthquake-study/

[9] *NOAA Ocean Explorer: Submarine ring of fire 2006*. NOAA Ocean Explorer Podcast RSS 20. (n.d.). Retrieved July 29, 2022, from https://oceanexplorer.noaa.gov/explorations/06fire/background/volcanism/media/subducting_plate.html

[10] Rigg, J. (2021, May 13). *The ancient earthquake detector that puzzled modern historians*. Engadget. Retrieved July 29, 2022, from https://www.engadget.com/2018-09-28-backlog-zhang-heng-seismoscope.html

[11] *Why Italian earthquake scientists were exonerated*. Science. (n.d.). Retrieved July 29, 2022, from https://www.science.org/content/article/why-italian-earthquake-scientists-were-exonerated

[12] Iris. (n.d.). Retrieved July 29, 2022, from https://www.iris.edu/hq/

[13] PANAKKAT, A. S. H. I. F., & ADELI, H. O. J. J. A. T. (2007). Neural network models for earthquake magnitude prediction using multiple seismicity indicators. International Journal of Neural Systems, 17(01), 13–33. https://doi.org/10.1142/s0129065707000890