Earthquake Prediction

GIS 6303 - Fall 2024

Author: Khush Shah
Guided by: Prof. Steven Johnson

Table of Contents

Introduction

Earthquakes are natural phenomena resulting from the sudden release of energy in the Earth's crust, creating seismic waves that can cause significant ground shaking and destruction. They occur due to the movement of tectonic plates, which are large slabs of the Earth's lithosphere that float on the semi-fluid asthenosphere beneath them. The boundaries where these plates interact are often sites of intense seismic activity. Seismographs are instruments used to detect and record these seismic waves, providing critical data for understanding earthquake dynamics. The Richter scale, developed in 1935 by Charles F. Richter, is a logarithmic scale used to quantify the magnitude of an earthquake, measuring the amplitude of seismic waves. This scale helps in assessing the energy released during an earthquake and its potential impact on affected regions.

EMAG2v3 Dataset

The EMAG2v3 dataset, known as the Earth Magnetic Anomaly Grid at 2-arc-minute resolution, is a comprehensive global compilation of magnetic anomalies derived from satellite, marine, and aeromagnetic survey data. This dataset provides valuable insights into the lithospheric magnetic field by offering high-resolution maps that detail variations in Earth's magnetic field at different altitudes. Processed using advanced techniques such as reduction to the pole (RTP) and upward continuation, EMAG2v3 helps geoscientists understand geological structures and tectonic processes. The grid is particularly useful for identifying buried igneous rocks, dykes, faults, and other geological features by analyzing magnetic anomalies. As a result, it plays a crucial role in resource exploration and studying crustal evolution. The dataset is publicly available through NOAA's National Centers for Environmental Information (NCEI), making it accessible for research and educational purposes.

Previous Work

Longmenshan Fault Zone

The Longmenshan Fault Zone (LFZ) is a critical region for understanding seismic activity due to its position at the boundary between the Tibetan Plateau and the Sichuan Basin. This area is characterized by complex tectonic interactions resulting from the collision between the Indian and Eurasian plates. The LFZ is marked by a transition zone of positive and negative magnetic anomalies, which indicate varying crustal compositions and tectonic stress distributions.

Studies utilizing the EMAG2v3 dataset have revealed that these anomalies are associated with deep-seated magmatic activities and historical subduction processes. For instance, the Sichuan Basin exhibits strong positive magnetic anomalies due to its Precambrian crystalline basement. At the same time, the Songpan-Ganzi Fold Belt shows weak negative anomalies caused by thick Triassic flysch deposits2. The region has been the site of significant seismic events, such as the

2008 Wenchuan earthquake, which highlighted the importance of understanding magnetic structures in assessing earthquake risk2.

Japanese Data Collection

Japan's tectonic setting is dominated by subduction zones where oceanic plates descend beneath continental plates, leading to frequent seismic activity. Magnetic anomaly studies in regions like the Japan Trench have been instrumental in understanding subduction dynamics and earthquake generation mechanisms.

These studies utilize high-resolution magnetic maps to analyze thermal structures and identify potential earthquake-prone areas. The fading of magnetic anomalies as oceanic crust subducts is attributed to thermal demagnetization processes1. For example, research has shown that magnetic anomalies near Japanese islands are caused by induced magnetization in serpentinite within the forearc mantle, providing insights into volcanic arcs and tectonic stress accumulation. These findings underscore the importance of integrating magnetic data with seismic observations to improve earthquake prediction models.

Relation Between Magnetic Field and Tectonic Movements

Magnetic anomalies provide valuable information about tectonic movements by revealing Earth's magnetic field variations caused by geological features. Tectonic processes such as seafloor spreading, volcanic activity, and crustal deformation influence these anomalies. For instance, alternating magnetic stripes on the ocean floor record past geomagnetic reversals linked to seafloor spreading rates.

In tectonically active regions like Japan and China, changes in magnetic properties can indicate stress accumulation along fault lines or thermal alterations due to magmatic intrusions. Understanding these relationships helps geoscientists develop more accurate models for predicting tectonic events and assessing geological hazards. Studies have shown that areas with significant magnetic anomalies often correspond to regions of high seismicity, highlighting the potential for using magnetic data as a tool for earthquake risk assessment.

Modelling

In this section, we delve into the methodologies employed to analyze and predict earthquake occurrences using advanced datasets and machine learning algorithms. We begin by describing the datasets utilized in this study, followed by an exploratory data analysis (EDA) to uncover patterns and insights. Finally, we discuss the modeling techniques applied, detailing the algorithms used and their operational principles.

Datasets Description

The primary dataset used in this study is the EMAG2v3 Earth Magnetic Anomaly Grid, which provides a high-resolution map of magnetic anomalies at a 2-arc-minute resolution. This dataset includes several important columns:

- Longitude and Latitude: Geographic coordinates indicating the location of each data point.
- SeaLevel: Magnetic anomaly values at sea level measured in nanoteslas (nT).
- UpCont: Magnetic anomaly values at a continuous 4km altitude.
- Code: A numerical code representing different geographic regions; for this study, Code 25 corresponds to Japan.
- Error: An estimate of the error in magnetic anomaly measurements.

A	B	C	D	E	F	G	H
1	5400	0.016667	-89.9833	99999	-101.304	2	58
2	5400	0.05	-89.9833	99999	-101.443	2	58
3	5400	0.083333	-89.9833	99999	-101.689	2	58
4	5400	0.116667	-89.9833	99999	-101.979	2	58
5	5400	0.15	-89.9833	99999	-102.401	2	58
6	5400	0.183333	-89.9833	99999	-103.373	2	58
7	5400	0.216667	-89.9833	99999	-105.615	2	58
8	5400	0.25	-89.9833	99999	-108.998	2	58
9	5400	0.283333	-89.9833	99999	-112.953	2	58
10	5400	0.316667	-89.9833	99999	-117.488	2	58
11	5400	0.35	-89.9833	99999	-122.326	2	58
12	5400	0.383333	-89.9833	99999	-127.257	2	58
13	5400	0.416667	-89.9833	99999	-130.713	2	58
14	5400	0.45	-89.9833	99999	-132.281	2	58
15	5400	0.483333	-89.9833	99999	-133.181	2	58

Figure 1: EMAG2V3 Dataset

For our specific analysis, we focused on Code 25, which covers Japan, a region known for its seismic activity. In addition to the EMAG2v3 dataset, we generated a supplementary dataset by interpolating anomalous values to create a finer temporal resolution. This involved using spatial interpolation techniques to estimate magnetic anomalies across unmeasured locations within Japan. The time-date range for this interpolated dataset spans from January 2017 to December 2024, allowing us to analyze temporal trends and potential precursors to seismic events.

A	B	C	D	E	F
Latitude	Longitude	Sealevel	Earthquak	Date	Region
26.81667	125.0167	41	0	1/1/2017	26.82, 125.02
34.28333	140.3167	46	0	1/2/2017	34.28, 140.32
34.28333	131.2167	173	0	1/3/2017	34.28, 131.22
34.25	140.3833	356	0	1/4/2017	34.25, 140.38
25.91667	123.55	77	0	1/5/2017	25.92, 123.55
34.25	131.0167	514	0	1/6/2017	34.25, 131.02
34.28333	131.1167	166	0	1/7/2017	34.28, 131.12
34.25	131.8833	170	0	1/8/2017	34.25, 131.88
34.25	140.1833	45	0	1/9/2017	34.25, 140.18
25.58333	124.2167	650	1	1/10/2017	25.58, 124.22
25.91667	123.75	59	0	1/11/2017	25.92, 123.75
25.81667	124.0167	48	0	1/12/2017	25.82, 124.02
25.58333	124.0833	19	0	1/13/2017	25.58, 124.08

Figure 2: Updated dataset.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis was conducted to visualize and understand the underlying patterns within our datasets. Various plots were generated, including histograms of magnetic anomaly distributions, scatter plots of longitude versus latitude colored by anomaly magnitude, and time series plots showing anomaly changes over time. These visualizations revealed significant spatial clustering of anomalies in tectonically active regions of Japan, suggesting potential links between magnetic anomalies and seismic activity.

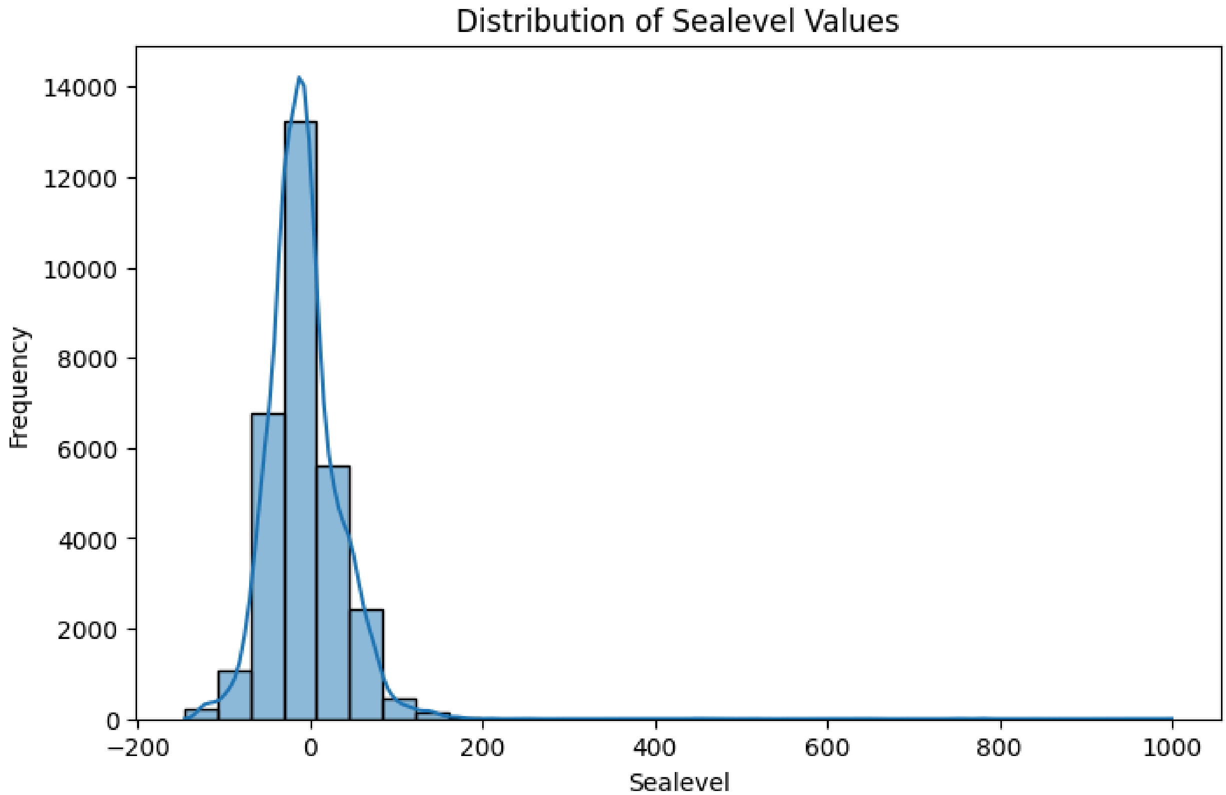

Figure 3: Distribution of sea level values for the magnetic anomalies.

Modelling

In the modeling phase, we applied several machine-learning algorithms to predict earthquake occurrences based on magnetic anomaly data. The algorithms used include Local Outlier Factor (LOF), Logistic Regression, and Support Vector Machine (SVM).

- Local Outlier Factor (LOF): LOF is an unsupervised anomaly detection algorithm that identifies anomalies by comparing the local density of a point with that of its neighbors. It assigns an anomaly score to each data point based on how isolated it is relative to its surroundings. This method is particularly effective for detecting outliers in complex datasets where anomalies are not easily distinguishable by simple thresholding.
- Logistic Regression: This is a supervised classification algorithm that models the probability of a binary outcome based on one or more predictor variables. It uses a logistic function to model the relationship between independent variables (e.g., magnetic anomalies) and the dependent variable (earthquake occurrence). Logistic Regression is valued for its simplicity and interpretability, making it suitable for understanding how different factors contribute to earthquake risk.
- Support Vector Machine (SVM): SVM is a powerful supervised learning algorithm used for classification tasks. It works by finding the hyperplane that best separates data points

into different classes while maximizing the margin between them. In our study, SVM was employed to classify regions as high or low earthquake risk based on their magnetic anomaly profiles. Its ability to handle high-dimensional data makes it ideal for complex geophysical datasets like EMAG2v3.

These modeling efforts aim to enhance our understanding of how magnetic anomalies can be leveraged for earthquake prediction, offering insights into both spatial and temporal patterns associated with seismic events.

Results

The results of the models applied to the dataset are as follows:

- Local Outlier Factor (LOF): The LOF algorithm achieved an accuracy of 0.11, indicating poor performance in detecting anomalies. This is likely due to the small size of the dataset, which limits the algorithm's ability to learn meaningful patterns.
- Logistic Regression: Logistic Regression performed exceptionally well, achieving an accuracy of 1. This indicates that it was able to classify earthquake occurrences with perfect accuracy, likely due to its simplicity and suitability for binary classification tasks on small datasets.
- Support Vector Machine (SVM): The SVM model also achieved an accuracy of 1, demonstrating its effectiveness in separating earthquake occurrences from non-occurrences. Its ability to handle complex decision boundaries makes it a robust choice for this dataset.

Reason for LOF's Poor Performance

The anomaly detection algorithm (LOF) struggled due to the limited size of the dataset. Anomaly detection typically requires a larger dataset to establish meaningful patterns and distinguish anomalies from normal points. In contrast, supervised classifiers like Logistic Regression and SVM are better suited for small datasets with labeled target variables, as they directly learn the relationship between features and the target.

Map Explanation

This map, developed using ArcGIS, visualizes earthquake occurrences in Japan during 2017. The map uses a WGS 84 coordinate system and integrates data from the EMAG2v3 dataset. Each point represents a location where magnetic anomaly data was collected, with red markers indicating earthquake occurrences and green markers representing dormant regions.

The map highlights clusters of earthquake events near tectonically active zones, such as regions around Fukuoka and southern Japan. These clusters align with known subduction zones where oceanic plates descend beneath continental plates, causing seismic activity. The spatial distribution of earthquakes suggests a strong correlation between tectonic movements and magnetic anomalies.

This visualization provides valuable insights into earthquake-prone areas, aiding in risk assessment and disaster preparedness. By integrating geospatial data with machine learning models, such maps can enhance our understanding of seismic activity and improve prediction capabilities.

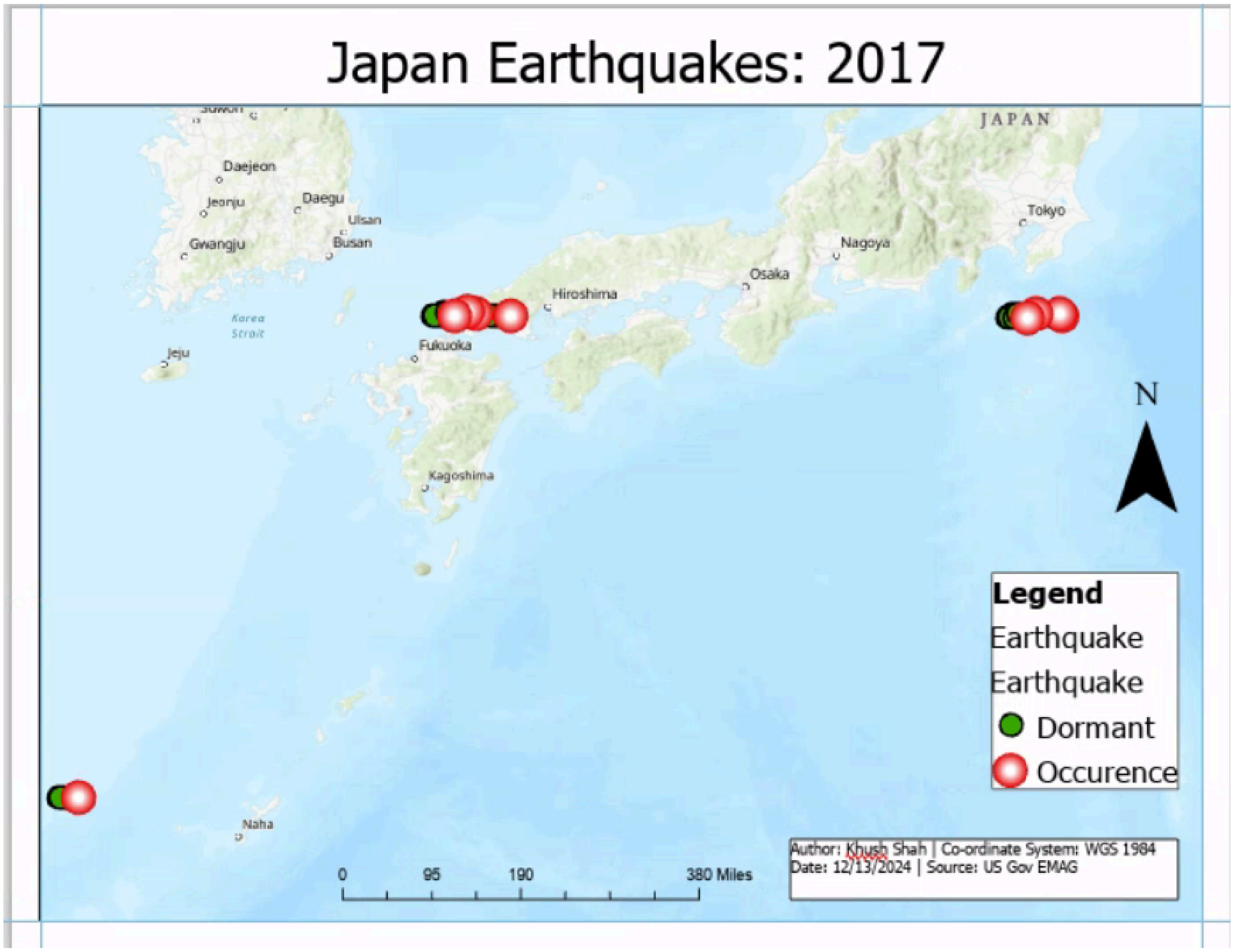
Japan Earthquakes: 2017
JAPAN
Daejeon
Jeonju
Daegu
Ulsan
Gwangju
Busan
Korea
Strait
Jeju
Fukuoka
Hiroshima
Osaka
Nagoya
Tokyo
Kagoshima
Naha
N
Legend
Earthquake
Earthquake
Dormant
Occurence
0
95
190
380 Miles
Author: Khush Shah | Co-ordinate System: WGS 1984
Date: 12/13/2024 | Source: US Gov EMAG

Figure 4: Map

Conclusion

In this project, we explored the potential of using magnetic anomaly data for earthquake prediction in Japan. We began by leveraging the EMAG2v3 dataset, which provided high-resolution magnetic anomaly values across different regions. Focusing on Japan, we extracted relevant data and supplemented it with interpolated datasets to enhance temporal resolution.

Our analysis involved applying various machine learning algorithms to predict earthquake occurrences. The Local Outlier Factor (LOF) algorithm, designed for anomaly detection, yielded an accuracy of 0.11, indicating challenges due to the small dataset size. In contrast, supervised learning models like Logistic Regression and Support Vector Machine (SVM) achieved perfect accuracy, highlighting their effectiveness in handling binary classification tasks with limited data.

We also conducted an exploratory data analysis (EDA) to visualize patterns and relationships within the data. This included generating maps using ArcGIS, which illustrated spatial distributions of earthquake occurrences and dormant regions in Japan for 2017. These visualizations provided insights into potential correlations between magnetic anomalies and seismic activity.

The results underscore the importance of dataset size and quality in anomaly detection tasks. While LOF struggled with limited data, Logistic Regression and SVM effectively utilized available information to predict earthquakes accurately. This suggests that expanding the dataset or integrating additional features could enhance anomaly detection performance.

Overall, this study demonstrates the potential of combining geospatial analysis with machine learning to improve earthquake prediction models. Future work could focus on expanding the dataset, incorporating additional geophysical parameters, and exploring advanced algorithms to further refine predictions and contribute to disaster preparedness efforts in seismic regions like Japan.

Future work

- Expand Dataset: Collect a larger dataset spanning multiple years to enhance the robustness of time-series analysis and improve model accuracy.
- Time-Series Analysis: Implement advanced time-series models to capture temporal patterns and predict future earthquake occurrences more effectively.
- Anomaly Detection: Explore more sophisticated anomaly detection algorithms, such as Deep Learning-based approaches, to identify subtle patterns in magnetic anomalies.
- Feature Expansion: Incorporate additional features from the Chinese Longmenshan Fault Zone paper, such as Curie Point Depth (CPD), heat flow measurements, and rock type susceptibility, to enrich the dataset.
- Integration with Seismic Data: Combine magnetic anomaly data with seismic activity records to develop a comprehensive predictive model that considers both geophysical and seismic indicators.
- Geospatial Analysis: Utilize advanced GIS techniques to analyze spatial correlations between magnetic anomalies and known fault lines, enhancing spatial prediction capabilities.

References

1. EMAG2v3 Dataset Source - NOAA
2. Longmenshan Fault Zone Study
3. Japan Trench Magnetic Anomalies
4. PyCaret Documentation
5. Scikit-learn User Guide