NASA's Gateways to Blue Skies Competition
Advancing Aviation for Natural Disasters

Improving Earthquake Prediction with Artificial Intelligence and Machine Learning

Bowie State University

Team Members

Team Lead	**Team Member**	**Team Member**
Adewoye Olaoluwa	Ian Gabriel Mondares	Alivia Ross
Senior	Senior	Senior
Computer Technology	Computer Technology	Computer Technology

Faculty Advisor
Dr. Haydar Teymourlouei, Department of Technology and Security

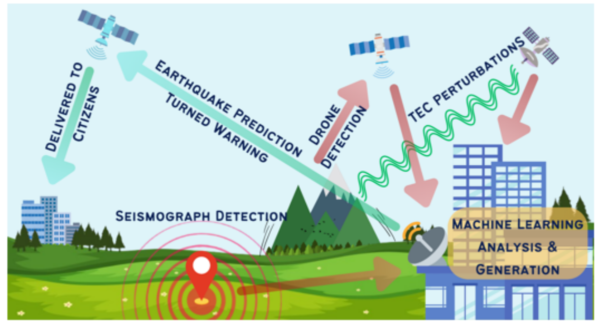
Quad Chart

Enhancing Earthquake Disaster Relief with Artificial Intelligence and Machine Learning

Project Summary:

- This research revolves around earthquake forecasting.
- The problem that this research is trying it resolve is earthquake prediction and detection, specifically the limitations of current systems such as ShakeAlert, which rely solely on seismic waves and may not provide timely warnings.
- The proposed solution begins with collecting data from satellites and other aircraft sensors. This data is then transmitted to a receiving station, where it is utilized by our proposed AI and machine learning technology for earthquake prediction.

Project Image:



Team Composition/Roles:

- Research Group (CyberSquad), composed of students within Computer Technology.
- Team Composition:
  - Adewoye Olaoluwa- Senior Computer Technology student.
  - Ian Gabriel Mondares- Senior Computer Technology student.
  - Alivia Ross - Senior Computer Technology student.
  - Dr. Haydar Teymourlouei- Faculty Advisor.
- Our team comprises computer technology students with background in machine learning, data analysis and programming. Alongside our advisor Dr. Haydar who guided the students throughout the project.

Proposed deployment timeline:

- Testing - 6 months to 1 year (Jan 2025 to July 2025 or Jan 2026)
  - Fine-Tuning on the program.
  - Ensuring secure practices with data transfer.
  - Large-scale software and hardware testing.
- Integration - 2–3 years (July 2025 (earliest) to Jan 2028)
  - Connecting USGS, USAID, and possible Satellite Communication Resources (i.e. NASA) together.
  - Obtaining equipment to accommodate AI needs.
- Deployment & Continuous Monitoring
  - Ongoing process. Up and running by 2030.
  - Steadily adding to system capabilities.
  - Edge computing Integration.

Abstract

Earthquakes are natural disasters that are difficult to predict and have caused a lot of major economic losses as well as devastating life losses. The current early detection of earthquakes does the job of informing the citizens of an incoming earthquake, but it only gives them limited time to respond and evacuate the affected area. Our proposed solution combines satellites and specialized drone sensors to collect real-time seismic data, coordinated with machine learning algorithms for earthquake prediction. Agencies such as the United States Geological Survey (USGS) and the United States Agency for International Development (USAID), as well as the National Aeronautics and Space Administration (NASA), will be essential for improving and implementing our proposed solution.

We studied and used prior work on earthquake phase/distance prediction to preprocess raw waveform data, and then predict earthquake magnitudes in addition to phase. The ML model analysis component had 85-90% success differentiating phase wave types and 0.002 loss when predicting magnitudes, resulting in predictions being within 0.4 magnitudes of true intensity on average. The generation component was able to create commonly looking waveforms after 150 training iterations.

The CONOPS for our prediction system will involve data being streamed in from satellite TEC perturbation measurements, drone geophones, and seismographs, all being used to forecast possible earthquakes. We estimate the total cost of the five years of development for the system is 98 million dollars. We anticipate the gain to be much larger due to reducing infrastructure damage, deaths, injuries, and expenditures from disaster response and recovery, and increasing productivity and economic participation. Through AI and ML, as well as advanced aviation technology, our research hopes to improve earthquake early detection and reduce the economic and life losses caused by earthquakes. We improve greatly on our previous research, completely revamping our machine learning strategy and expanding on previously ignored areas of our system, including integration and budget.

1 Introduction

1.1 Overview of Earthquakes

Earthquakes are natural phenomena that are caused by the movement of tectonic plates within the Earth’s crust. The root of this complicated geologic phenomenon is the separation of the lithosphere of the Earth into separate plates, where tension builds up at their borders because of interactions and releases as seismic waves [1]. Seismologists use advanced seismographic instruments to detect, analyze, and monitor these seismic waves, building a comprehensive understanding of earthquake dynamics on a global scale.

An earthquake is one of the most devastating natural disasters. There are a lot of communities that have suffered fatalities because of earthquakes. Roughly 2.32 million individuals were impacted by earthquakes, and almost 59% of earthquake-related fatalities resulted from a building collapsing [2]. In 2023, around 64 thousand lives were lost due to earthquakes and its aftershocks [64].

It is very challenging for seismologists to predict where and when the next earthquake could happen, even with the current technology. Furthermore, earthquakes can cause secondary effects due to the movement of land, which adds to the toll of this natural disaster. Earthquakes could result in tsunamis and landslides, which contribute to their injury and death rates along with a serious economic impact on the affected location.

1.2 Existing Earthquake Detection Methods and their Weakness

ShakeAlert is one of various earthquake early detection systems and is managed by USGS. ShakeAlert works by monitoring seismic waves within the area through seismographs to detect incoming earthquakes [3, 4]. When a large enough earthquake occurs that could cause shaking in a particular area, the system quickly sends out an alarm by estimating the location, the time till arrival to a given end user, and the strength. The alert is then effectively broadcast through a wide range of communication channels, such as smartphone apps, emergency alert systems, and other platforms, enabling people, organizations, and important infrastructure industries to take immediate protective and safety measures [3].

One issue with this system is that it might be too late when the alerts are sent. Since it solely relies on constant monitoring, the process of ingesting seismic data, realizing the activity is large enough to disrupt an area, and broadcasting alerts is racing against the seismic waves themselves, especially fast P-waves that can reach an area at the same time of an alert if the shaking is close enough. This does not allow some businesses and individuals to pack valuable, necessary tools and get out of the affected area.

1.3 A Modified ShakeAlert System

The goal of this project is to improve the current early detection system by implementing machine learning (ML) and seismic detection through aerial means. While the current earthquake early detection system, ShakeAlert, notifies individuals of impending earthquakes, it does not give them enough time to prepare and evacuate from it. By implementing artificial intelligence, we hope to shift focus from quick detection to advanced prediction of earthquakes, allotting more time for preparation with warnings. By using drones, we aim to advance how versatile seismic detection can be by using drone-mounted geophones that can be repositioned on a whim. By using satellites, we aim to show how environmental surveillance can be used to help predict earthquakes and how satellite communication can allow flexibility of communication to all stations for processing. We do not intend for this system to replace basic detection, but to enhance it with higher levels of nuance reachable through contemporary strategies.

2 Earthquake Prediction through Machine Learning & Aerial Strategies

Fig. 1 illustrates the components utilized within our suggested approach. This approach makes use of satellites and planes equipped with specialized sensors, among other aviation technologies. These sensors are responsible for collecting data, which our machine learning model is then trained with. The data is sent to a receiving station for storage after it has been gathered. The data is then analyzed, and an ML model is used to train the datasets so that predictions about the frequency and timing of earthquakes may be made.

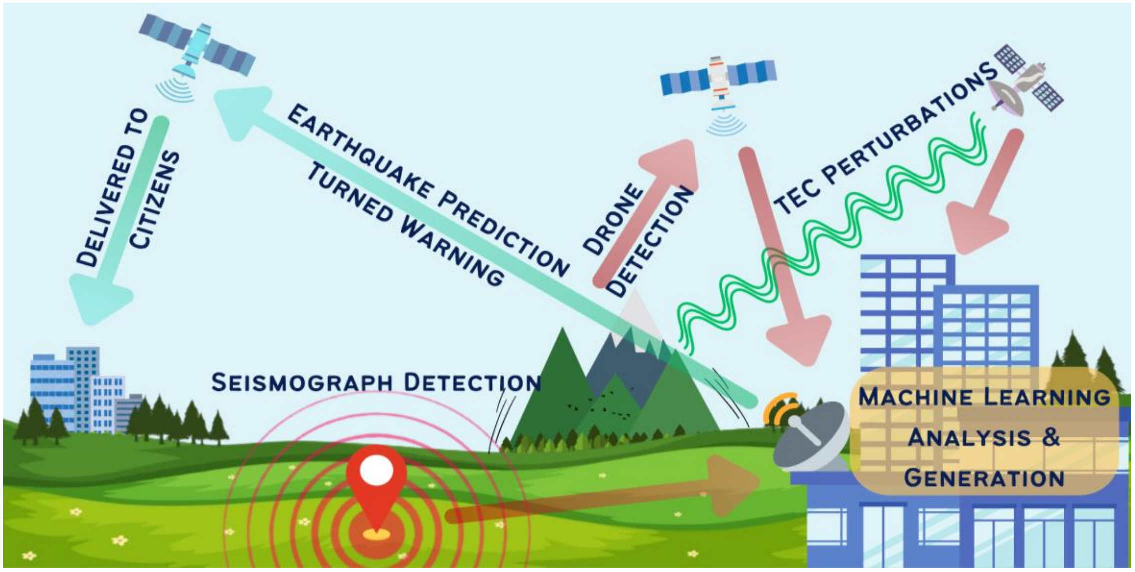


Fig. 1. A high-level overview of the modified ShakeAlert system. The diagram illustrates the seismic data flow from detection to delivery, where Machine Learning models are used to enable earthquake prediction and data synthesis.

2.1 Satellite transmission

Implementing satellites for further or remote data transmission offers many advantages, including a more established communication system and reliable data transfer that links across vast distances. In Figure 1, communication satellites transfer earthquake monitoring data to a receiving ShakeAlert station. Using satellites can provide global coverage for data transmission, which is crucial for flexible reach to every sensor and station.

2.2 Sensor Data and its Sources

Our solution seeks to expand the ways we can obtain seismic waveforms from traditional ground seismographs. One inherent drawback to these seismographs is the nature of their infrastructure. The best locations where seismographs and stations can operate are right next to faults as they can listen to where seismic activity is most likely to happen but if those locations are impossible to operate in, then that hampers their effectiveness. For example, many of California's faults are well within mountainous terrain and would be difficult to establish full seismograph stations at. [16, 17]

This is why sensor technology from the sky can help fix this issue. Mobile drones can enable geophones - the edge computing version of full-sized seismographs - to be deployed next to these faults, and for locations for these smaller devices to be adjusted on the fly if new developments are found [31].

Airplanes and satellites can then serve multiple purposes. Dynamic analysis requires a dynamic communication system, and both geophones and seismographs can be configured to communicate to aerospace hubs that can then forward their data back down to various stations. This will allow each station to have 1) seismic activity according to its own spatial context and 2) activity according to the context of the entire west coast or country. The second purposes is leveraging new methods being created in the recognition of seismic activity, such analyzing perturbations in the total electron content of the atmosphere due to seismic activity. In addition, Synthetic aperture radars satellites are satellites equipped with special sensors that can detect seismic activity [9]. These satellites can provide continuous ground monitoring and seismic activity patterns that can be used to predict earthquakes. By leveraging drones, airplanes, and satellites, we are allotted more freedom in monitoring the earth for earthquakes. [9]

2.3 Machine Learning

For our proposed solution, data is gathered by ground seismographs, satellites, and drones which we then feed to the ML models at the seismograph station.

Machine learning is the study of the computer's abilities to learn like a human being. There are a variety of strategies that have been developed to allow an algorithm to learn anything, but one of the most versatile methods is a neural network. Created to mimic the structure of a human brain, a neural network model is made up of multiple layers of nodes called neurons. Each neuron as a weight and an activation function connected to it, meaning a network of neurons can tailor itself to any outcome at a node level [40]. This means that if you can convert an abstract object to a series of numbers, and convert a given outcome to a set of numbers, you can theoretically create a model for anything, including 3D waveform data.

2.3.1 Phase & Magnitude Predictions

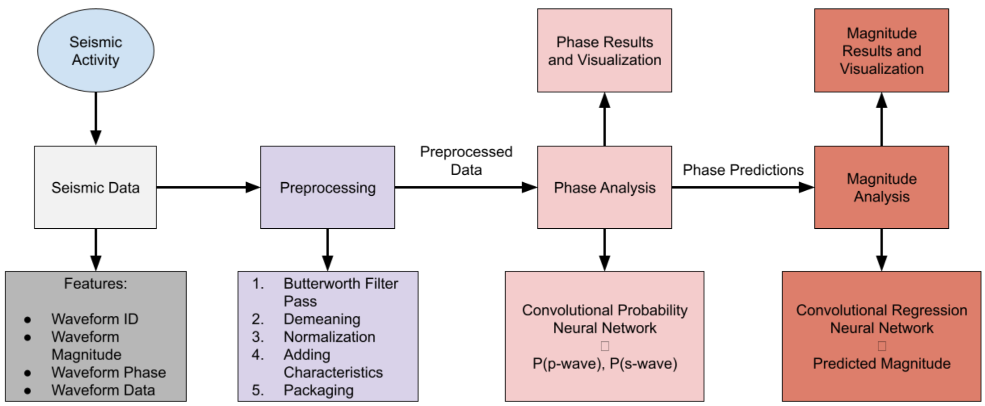


Fig. 2. A diagram of the process for seismic data analysis.

The objective of the ML analysis is to provide prediction models that can predict earthquakes before they happen. The process is illustrated in Figure 2. In summary, receiving stations will gather seismic data (including waves, previous earthquakes, and their magnitude levels). This seismic data will then be analyzed alongside the usual detection capabilities of the original system. Should the ML detect possible future activity, it will notify personnel to investigate, among other processes.

The method was tested on a dataset from the US Geological Survey's website [10]. William Yeck et al. obtained the data through the web services of the IRIS Data Management Center and then posted data in numerous compressed files, the largest being the P-Wave training waveforms, totaling at least 40GB worth of storage. Each earthquake waveform trace consists of three channels of one-minute samples conducted at 40Hz, resulting in 2400 measurements per channel and 7200 measurements per earthquake waveform. Each channel represents a degree of displacement (Vertical displacement, North-South displacement, and East-West displacement) to fully account for a 3D representation of an earthquake vibration [10].

Next features are computed: the earthquake's unique identifier, earthquake magnitude, the phase of the waveform, and the waveform measurements from the earthquake. This data will go through a preprocessing phase where we leverage Yeck's custom Butterworth filter and demeaning strategy to first prepare the waveforms, then package the data as training and testing dictionaries. After preprocessing the data, we then trained and tested four models. The first model was the phase classifier that will use probability to determine if a given waveform sample is a P-wave or S-wave. The next three models - one for specifically P-waves, one for S-waves, and one for all data - are magnitude models that predict the exact magnitude given a sample. There are three models as we wish to test the magnitude prediction ability after the data is classified by the phase model.

2.3.2 Synthetic Waveform Generation

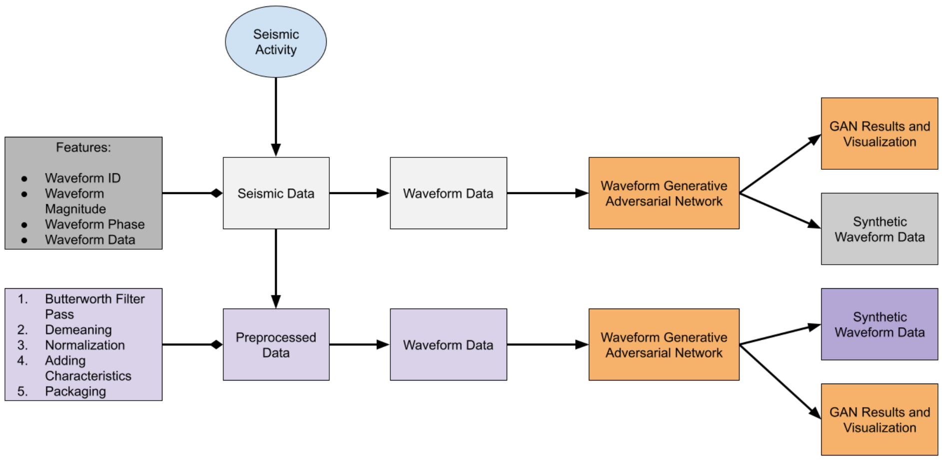


Fig. 3. A diagram of the process for synthetic seismic data generation.

To generate data off of the wave forms, we take advantage of the Generative Adversarial Network (GAN) method to create two different generators, illustrated in Fig. 3. This sort of network uses two autoencoders - a discriminator deciding whether an input value is real or fake, and a generator set on tricking the discriminator into predicting correctly. This setup forces the generator autoencoder to get better and better and fake real data until it becomes near indistinguishable [12]. We have two unique GANs set up to handle both raw and preprocessed seismic waveforms. Results from both can be sent into the ML analysis portion as if the data were collected through natural means.

3 Implementation Analysis

3.1 Python Program Setup

To implement all programs into a proof of concept, we developed a multi-file program in PyCharm. We utilize Python 3.11 as it is one of the latest stable builds of Python that can comfortably support deep learning libraries. The NumPy and Pandas libraries are used for data ingestion, manipulation, and saving both input and output data. The SciPy and Scikit-learn libraries are used for preprocessing and model analysis. The Seaborn, Matplotlib, and Imageio libraries are used for data visualizations.

The four files were organized into a main Python program, a preprocessing Python program, an analysis Python program, and a generation Python program. The main file will be the link between the other three programs so that the program can be run through one command or button press. It will first call the preprocessing file and ingest, then preprocess the waveforms that will be used throughout the program, along with associated phases and magnitudes (azimuths and distances aren't used but are also gotten too). This data will be first fed into the analysis program, which will use the phase data and magnitude data to train/test the phase model and magnitude models, respectively. This data is next fed into the generation program, which will first ingest its own data to generate raw waveforms, then use preprocessed data to generate preprocessed waveforms. The analysis and generation happen independent of each other.

3.2 Machine Learning TRL Evaluation

As a proof-of-concept, we set parameters in the main function to test the code. We utilize 2000 randomly selected waveforms for the analysis and preprocessed generation, with 1500 of them being P-waves and 500 of them being S-waves. We set the program to only use a 50-second window from each 60-second waveform. Lastly, we set the raw generation to use 5000 waveforms in total, and output 5000 synthetic waveforms.

The program took a total of 41 minutes and 35 seconds to run. Preparing the waveforms for analysis took just 34 seconds. The analysis section, which included the training of all four models, then the testing of all four models, took 6 minutes to conduct. The generation section, which includes the raw generation and the preprocessed generation, took around 34 minutes to complete.

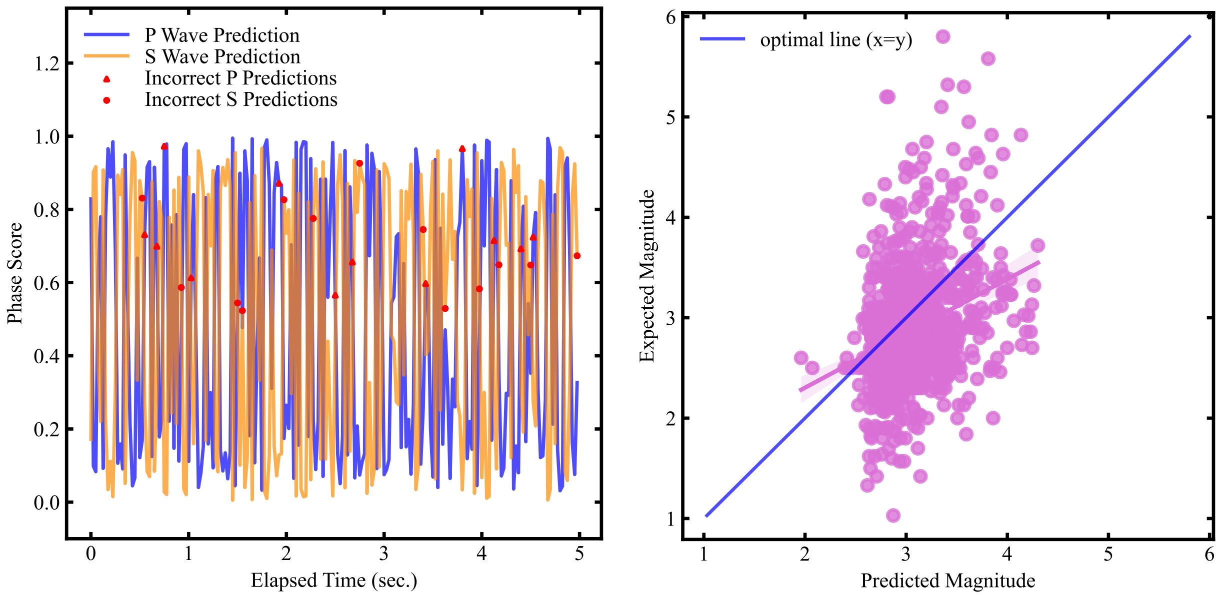


Fig. 4 & 5. Prediction Results of Phase Predictions (Left) and Magnitude Predictions (Right)

In terms of phase classification, the program averaged an 85-90% total accuracy rate throughout all the runs of the phase model. Figure 4 showcases the phase scores in terms of probability for a sample of 200 waveforms, with P-wave probability in blue and the S-wave probability in orange. The red dots showcase where the program predicted wrong, with 31 predictions incorrect within this 200-waveform sample.

In terms of magnitude classification, the program was able to achieve around a 0.002 loss rate when testing the general model's prediction capabilities. Figure 5 shows the magnitude results when scaled back to their true forms. When scaling up both the test magnitudes and the predicted magnitudes, the model was on average about 0.4 magnitudes off from the true number, and the regression line itself was very close to the optimal line in orange, showing that the predictions did align to the magnitudes. However, the coefficient of determination metric, or R^2, was around a low 0.15 meaning that the results were still very spread out from where it could be.

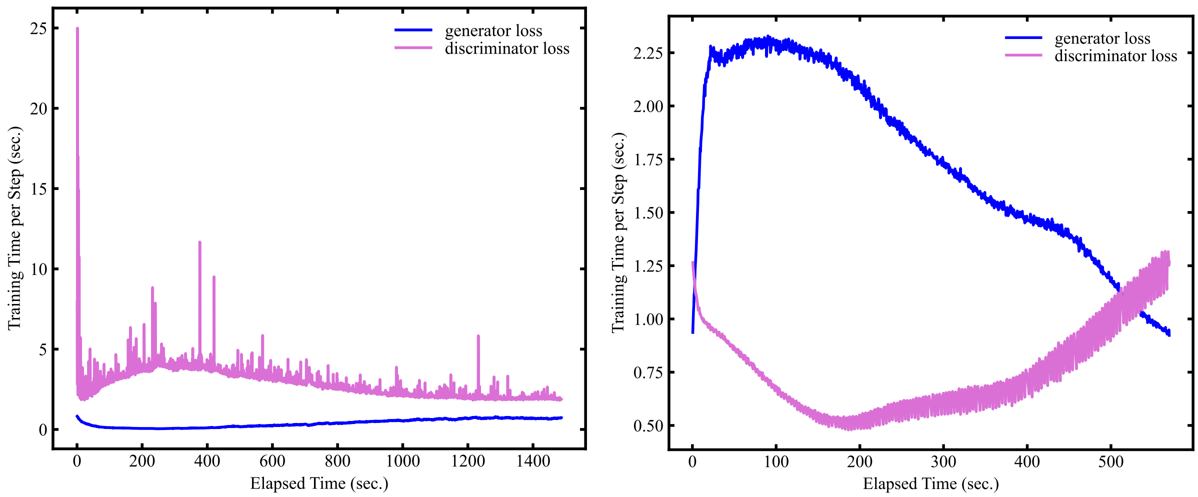


Fig. 6 & 7. Loss Evolutions of Generator and Discriminator for Raw GAN (Left) and Preprocessed GAN (Right)

For waveform predictions, we found that 150 epochs of training gave the GAN model plenty of time to learn on the provided waveform data. It took almost 25 minutes for the raw GAN model to train itself, while it took 9 minutes for the Preprocessed GAN model to train itself. This difference can be attributed to needing to train 5000 waveforms instead of 2000 waveforms, and the preprocessed data using only 50 seconds of each sample rather than using the full 60 seconds.

Figures 6 and 7 show the loss scores of the GAN for both raw and preprocessed data, respectively. For the raw data, the generator had a quiet runtime and only started improving once the discriminator was accurately differentiating waveforms. For the generated data, the two halves of the GAN model were in a constant fight with each other. Both outcomes resulted in waveforms that looked accurate to the most common waveforms, shown in Appendix A.

3.4 System Advantages & Limitations

We believe the system has several possible advantages. First and foremost, our solution will allow more advanced warning of an earthquake, enabling the necessary measures to be taken before the destructive forces can reach the site of impact. This can save the lives of residents, prevent unnecessary damage to various infrastructures, and enable a sense of safety and security that can allow communities to prosper. The next save we foresee is in manpower, as having artificial intelligence to handle predictions can allow a significant reduction in the resources needed for earthquake detection (mainly the setup of infrastructure) and emergency response. The last advantage is that the system largely uses components that are already well-established products in their own right. The majority of the work will not be the creation of the system, but leveraging components together in an integrated manner.

However, any solution that involves adding complexity will always have limitations. At the communication level, there is the worry of hackers that may attempt to take down the system or hold it for ransom, such as in the case of the Colonial Pipeline back in 2021 [42]. Another problem that can hamper communications is the environment itself. The terrain of the area where seismic activity is detected must allow for a clear line of sight between the sensor and the satellite, meaning heavy weather conditions can interrupt communications with drones and satellites alike. The analysis would have to fall back to ground communications with their local station. At the data processing level, the use of machine learning will require extra computational power and hardware modules, such as GPUs and TPUs made for conducting highly intensive computation, so an evaluation of current computer hardware will be done to see if they are suitable to handle running the models. Lastly, while most of the system itself just needs to be combined, there are no consumer-ready solutions available with geophones integrated into the drone, meaning some development will have to be focused on the design.

4 Deployment and Integration

4.1 Deployment Timeline

SUM of Cost	Year							
Cost Category	2025	2026	2027	2028	2029	2030	Yearly Maintenance	Grand Total
App Costs	$25,000	$25,000	$25,000	$25,000				$100,000
Com Contract Cost						$11,250,000	$11,250,000	$22,500,000
Cybersecurity Labor	$868,000	$868,000	$868,000	$868,000	$868,000	$19,220,000	$19,220,000	$42,780,000
Data Contract Cost						$4,706,551	$4,706,551	$9,413,102
Geophone Drone Fleet						$28,120,000		$28,120,000
Production Team					$761,000			$761,000
Seismograph & Satellite Installations						$259,000		$259,000
Software Engineering Labor	$761,000	$761,000	$761,000	$761,000		$25,847,000	$25,847,000	$54,738,000
Test Drones		$380,000						$380,000
Test Seismographs and Satellites		$14,000						$14,000
Grand Total	**$1,654,000**	**$2,048,000**	**$1,654,000**	**$1,654,000**	**$1,629,000**	**$89,402,551**	**$61,023,551**	**$159,065,102**

Fig. 8. Pivot Chart showcasing year-by-year costs of the development period, in addition to annual costs of maintaining the system.

The Deployment Timeline is split into four major phases of development and integration into current national earthquake detection systems. The phases are split up by years, where the cost for each year is displayed in Figure 8. How each cost was derived is shown in the cost explanation (sec. 5).

The first phase is called the Testing Phase, where the majority of work will be in raising the technology readiness level to TRL level 7. This will involve refining the speed, capacity, and performance of the solution itself, selecting hardware for data collection and satellite communication, ensuring proper cybersecurity measures work without inhibiting the performance of the solution, ensuring the geophone-mounted drones are able to send back seismic data without needing to physically return, and initializing the communications between all involved parties of the larger system. This phase should last from 2025-2026.

The second phase is called the Integration Phase, where the majority of the work will be ensuring organizational partners and all participating parties in the system have finalized roles. We expect the categories of interacting entities to be NASA as the provider of aerial and satellite communications, USGS as the owner of the national earthquake detection system, the alert providers that USGS works with to get seismic information out to consumers, the emergency response services that work based on alert information, the US citizens that we assume will either have the app or be in proximity to alerts, and the providers of hardware, such as satellite dishes, possibly new seismographs, drones, and various software solutions that will be needed. We expect the finalizing of this system to take longer, and thus give it a longer time range of 2026-2028.

The third phase is called the Validation Phase, where the majority of the work will be in raising the technology readiness level to TRL level 8, or “flight testing” the solution in an actual ShakeAlert environment. In this phase, a station would be used as a test station where the improved solution will run adjacent to the current system, along with all additional system components. This will allow us to create performance metrics and use experimental data from a real environment to compare the effectiveness of both solutions. Mainly, we will be seeking to ensure that our predictions will be able to report oncoming earthquakes faster than ShakeAlert can report a detection. We expect that the setup, testing, and possible tweaks resulting from testing should take from 2028-2030.

The last phase is the Deployment Phase, where resources will shift towards implementing a successful prediction system as the new national standard for earthquake monitoring. This will meaning outifitng each station with the necessary components and resources to do its own unique monitoring, while feeding into a central system that can aggregate results responsibly. This will start in 2030 and continue indefinitely as the system as a whole is improved and maintained.

4.2 Operational Integration

During the development phases, there will be overall objectives that we hope to achieve. The first main objective is for all parties of interest to be aware of the final "flight-ready" system and their roles in it by 2028. The second main objective is for specific hardware of the final "flight-ready" system by 2028. We shall achieve this through the themes of each phase. The testing phase will be where conversations shall begin for which organizations and potential companies should be in consideration for collaboration. The integration phase will be where this wide net shall be narrowed down only to companies that show interest in working on the development. The validation phase shall further narrow down only to the companies chosen to work within the system and should clarify all roles that each entity, including the end people, should have.

In the current proposal, the actors of interest are US National Aeronautics and Space Administration, United States Geological Survey, Federal/state/local disaster response entities and ShakeAlert partners (e.g., California Office of Emergency Services, US Federal Emergency Management Agency), Aerospace Data Providers (e.g., Maxar, Airbus), Satellite Communication Providers (SpaceX, BlackSky), and US Citizens that will be served by this effort.

The data provider shall provide real-time TEC perturbation information that should feed into USGS's ShakeAlert seismograph data. This data will be simultaneously analyzed by a seismic station. This station will be connected to a satellite network that can take in data from any seismograph and drone geophone. This data is both statically analyzed and fed into the machine learning model. The model's estimates will be used to forecast the magnitudes and possibly future magnitudes. If the anticipated magnitude is large enough and the model is certain of a serious earthquake, an alert will be sent through ShakeAlert's network to notify everyone of a highly likely seismic event.

4.3 Cost and Justification

Between the years of 2025-2030, we estimate the cost of our system to be around 98 million dollars, with every year after that being an additional 61 million in yearly costs. Our measurements assume that everything but the buildings of the stations is completely new and that dedicated contracts will be used for both TEC data and other atmospheric measurements and the satellite communication system, so this number will be lower than actually needed.

While this is a steep cost, we believe the anticipated cost from this system can be offset just by preventing the anticipated cost of earthquakes should we not move forward with the solution. A paper published in 2022 researched the numerous methods used to determine the cost of a human life. One of these methods, the value of statistical life (VSL), measured monetary value based on payouts, lost productivity, and other market factors. Papers that used the VSL method to determine the cost of a US person's life settled between 5 million and 15 million [21, 23]. On the low end, just losing 20-30 lives to an earthquake could result in as much potential value lost as the entire development phase of this solution.

Earthquake predictions can allow those same people to steel themselves, find cover, or even evacuate to open land should they be able to. This will increase the chance of someone surviving a disaster, or even avoiding much of the danger altogether. It will also decrease the pressure on disaster response teams, as-

- A more mentally prepared populous is easier to locate and direct towards shelters and emergency health care locations.
- More people in safe locations make the task of digging through rubble or hazardous areas less difficult.
- Recovery can be faster and efficient with more volunteers, and damaged places can return to normalcy quicker.
- Less people are injured throughout the seismic event.

-all of which decreases the money, effort, and workforce that must be expended during the aftermath of a seismic event.

An extra benefit should predictions get good enough that earthquakes can be anticipated far in advance is that homeowners and building owners can take strides to prepare their locations to be hit, resulting in less infrastructural damage. There is data to show that mitigation efforts, even with the small seconds before an earthquake hits, are as valuable as at least 10 times the amount of resources spent in the aftermath [26]. With a 2023 FEMA report showing 14.7 billion dollars lost annually in building damages attributed to earthquakes, any edge that can be taken to protect the infrastructure we use in our everyday lives should be taken [41].

5 Cost Explanation

The following costs are idealistic estimates assuming we are making modifications in support of the current infrastructure, we aren't adding or removing more stations, and we are not reusing equipment to save costs. This was to create a budget that accounts for all possible expenditures that directly affect our system.

Seismograph-Satellite Installations

Currently, there are 74 USGS stations that ShakeAlert leverages for data[16]. Through research with numerous vendors we found that high-end three-dimensional seismograph sensors fall between $1000-5000 [14, 15]. Commercial satellite communication providers such as Starlink charge around $500 to install physical satellites at a given location so that that location can access their network. [18, 19] Assuming we obtain the average high-end seismograph for $3000, along with a $500 dollar satellite installation to enable seismographs to communicate to a remote station, for each one of the 74 stations, yielding (3000 + 500) * 74 = $259,000 to equipment each ShakeAlert station with the needed seismograph equipment.

Labor costs for cybersecurity and software engineering/development

The United States Bureau of Labor Statistics reports the following annual salaries as of May 2023: The average software development employee makes $113,000 per year, and the average cybersecurity employee makes $124,000 per year. [20] The salaries we are looking at are umbrella terms. In reality, different positions may be used. We will require a main team consisting of a variety of specializations to work on the ML component of the solution. Assuming this time consists of seven people: 113,000 * 7 = $761,000 annually to pay salary for this team. We will require a main team of cybersecurity experts to ensure that the system at large is near invincible to possible attacks and vulnerabilities that can compromise its functionality. Assuming this team consists of seven people, it yields 124,000 * 7 = $868,000 annually to pay salary for this team. We may require satellite teams of developers and cybersecurity for each location, yielding 74 Locations * $113,000 per year * 3 software developers = $25,086,000 in development salary annually and 74 Locations * $124,000 per year * 2 cybersec personnel = $18,352,000 in cybersecurity salary annually.

Application costs

Application costs fall on a broad range, but large-scale apps that interact with a back-end source appear to cost at least $100,000. We assume this figure for the cost of investment into modifying end-user solutions that rely on ShakeAlert and spread it out across four years.

Satellite TEC Data

NASA's Earth Science Division (ESD) holds several contracts pertaining to data products collected by satellites owned by various companies. Most similar to the type of data we need are Maxar and Airbus. NASA's deal with Maxar's sat-constellation for $3,735,948 (or $933,785 per year) from 2018-2022 allows them access to "Collect Clouds, Aerosols, Vapor, Ice and Snow (CAVIS) imagery from WorldView-3 and WorldView-4 satellites." NASA's deal with Airbus's sat-constellation for $1,349,142 from 2021-2022 allows them access to “Earth-relevant data products obtained for scientific evaluation, including SAR satellite data products from TerraSAR-X, TanDEM-X, and PAZ constellations.” We took the average of all ESD contracts listed and came to an assumption of $4,706,551 annually to pay a data product including round-the-clock TEC perturbation monitoring data.

Satellite Communications

The Defense Information Systems Agency, or DISA, published a 2023 report of a 900 Million dollar contract (originally) split between 16 companies for the purpose of “Proliferated Low Earth Orbit (PLEO) Satellite-Based Services” for 5 years (2023-2028). 900 Million / 16 companies / 5 years = $11,250,000 annually on average per company. We use 11.25 million as the assumed annual cost of a Sat Com contract.

Drones and Geophones

A quadcopter autopilot drone fit for analyzing seismic activity in remote locations would cost $3,000 or more. Triaxial Geophones go for around $140-200. A drone with four geophones mounted for ground spiking yields 3000 + (200 * 4) = $3,800 per seismic autonomous drone. A test fleet of 100 seismic drones (cost: 3800 * 100 = $380,000). One fleet for each location, cost: 380,000 * 74 = $28,120,000 for all 74 stations.

6 Summary

We have presented an innovative solution for an improved earthquake prediction and alert system that utilizes contemporary strides in machine learning, aerial-based real-time earthquake monitoring, and satellite communication to improve efforts to inform citizens of incoming seismic events. We demonstrated that our main contribution to the system - a machine learning-based approach that predicts both the phase and magnitude of a wave - was able to complete a full run-through and predict the magnitude of an earthquake based on the waveform, with being half a magnitude off from the true value on average. Our financial analysis and justification showcased the disparity between the comparably small investments into the system, and the opportunity of saving and creating millions, if not billions of dollars in preventing deaths, further tailoring infrastructure, and creating an environment where disaster recovery efforts do not need to be as large. We believe that with full backing, we can usher in the next era of earthquake alerting, and create a world where earthquakes are yet another natural disaster that can long be anticipated and prepared for.

Appendix A - Waveform Examples

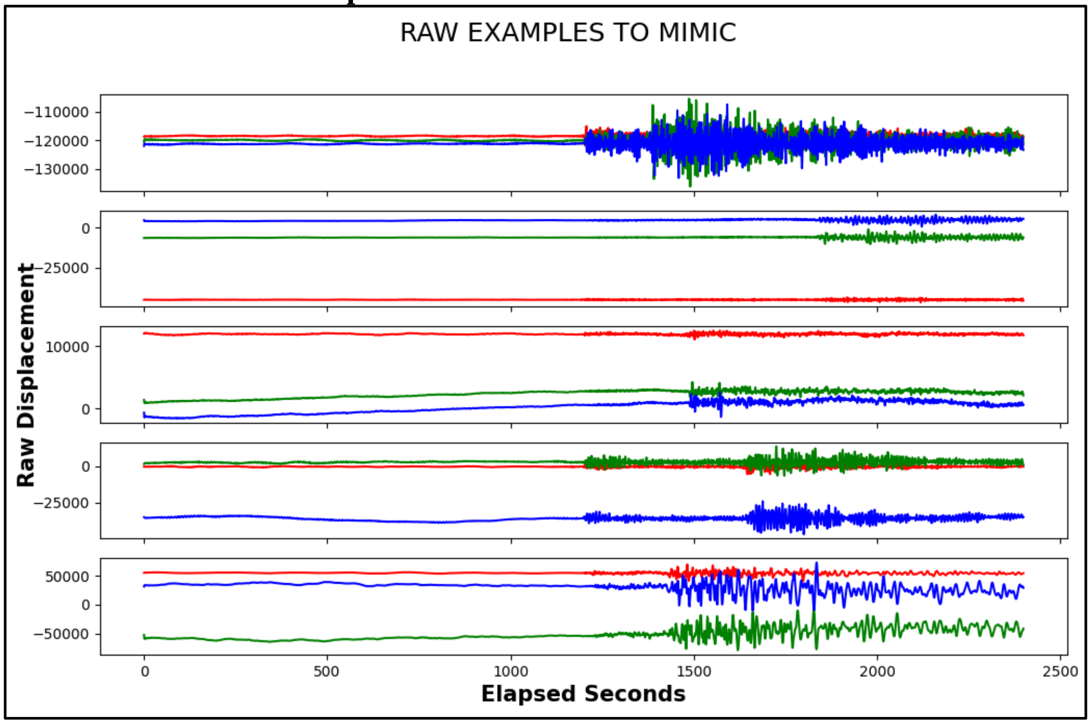


Fig. 9. Raw Waveform Samples fed into the Generative Adversarial Network

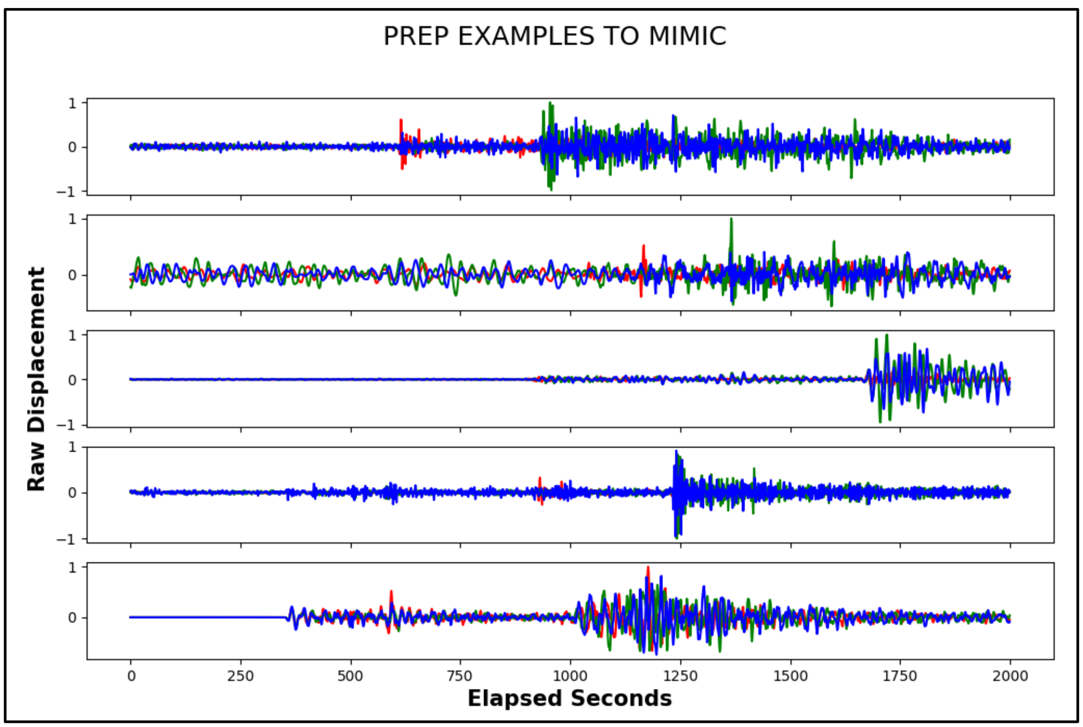


Fig. 10. Preprocessed Waveform Samples fed into the Generative Adversarial Network

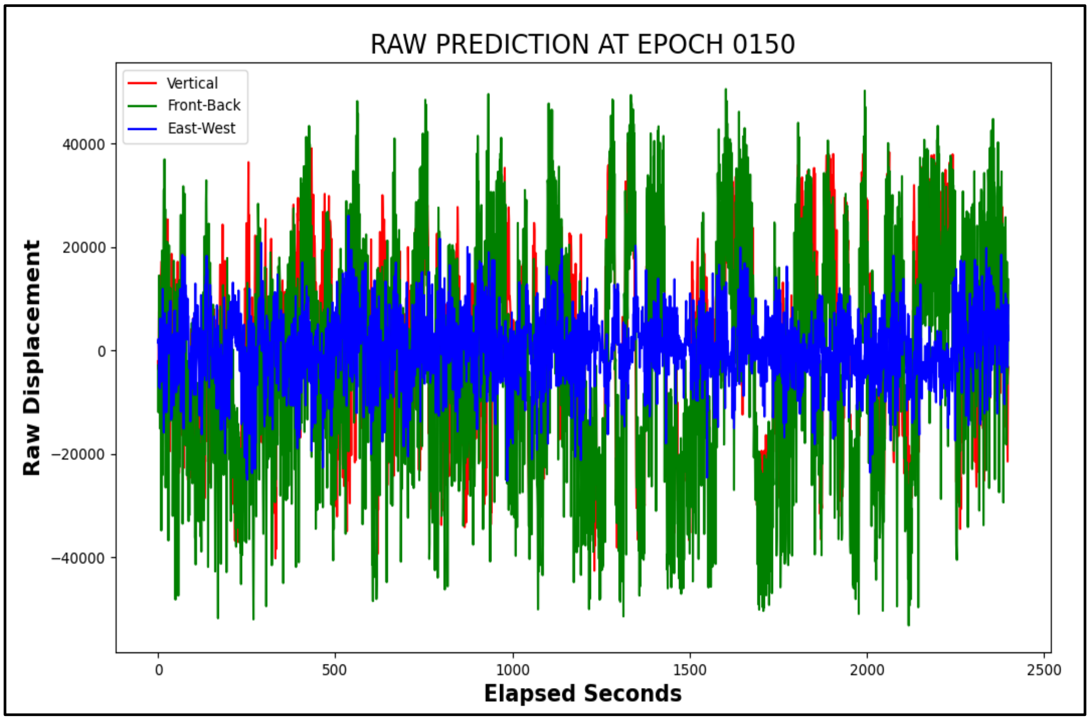


Fig. 11. Raw Waveform Prediction created after 150 training cycles.

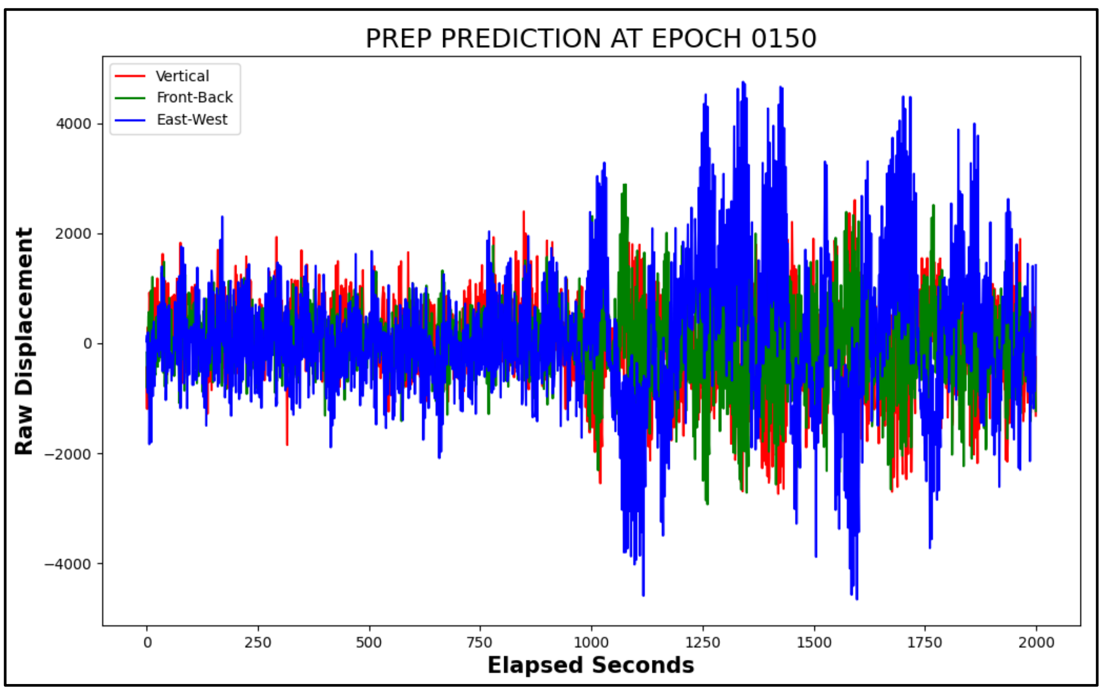


Fig. 12. Preprocessed Waveform Prediction created after 150 training cycles.

References

[1] H. Kanamori and E. E. Brodsky, “The physics of earthquakes,” *Reports on Progress in Physics*, vol. 67, no. 8, pp. 1429–1496, Jul. 2004, doi: https://doi.org/10.1088/0034-4885/67/8/r03.

[2] I. (SEK) Oberdorf, “KIT - KIT - Media - Press Releases - Archive Press Releases - PI 2016 - Natural Disasters since 1900: Over 8 Million Deaths and 7 Trillion US Dollars damage,” *www.kit.edu*, Jul. 11, 2023. https://www.kit.edu/kit/english/pi_2016_058_natural-disasters-since-1900-over-8-million-deaths-and-7-trillion-us-dollars-damage.php (accessed May 01, 2024).

[3] S. K. McBride *et al.*, “Developing post-alert messaging for ShakeAlert, the earthquake early warning system for the West Coast of the United States of America,” *International Journal of Disaster Risk Reduction*, vol. 50, p. 101713, Nov. 2020, doi: https://doi.org/10.1016/j.ijdrr.2020.101713.

[4] A. Nanfria, “Get an earthquake alert in Northern California? Here’s why ShakeAlert targeted you - CBS Sacramento,” *www.cbsnews.com*, Oct. 18, 2023. https://www.cbsnews.com/sacramento/news/did-you-get-an-earthquake-alert-heres-why-the-shakealert-targeted-you/

[5] “The United States Provides $12 Million in Additional Funding to Respond to the Devastating Earthquake in Afghanistan | Press Release,” *U.S. Agency for International Development*, Oct. 12, 2023. https://www.usaid.gov/news-information/press-releases/oct-12-2023-united-states-provides-12-million-additional-funding-respond-devastating-earthquake-afghanistan

[6] N. Casagli *et al.*, “Spaceborne, UAV and ground-based remote sensing techniques for landslide mapping, monitoring and early warning,” *Geoenvironmental Disasters*, vol. 4, no. 1, Mar. 2017, doi: https://doi.org/10.1186/s40677-017-0073-1.

[7] D. Contreras, T. Blaschke, D. Tiede, and M. Jilge, “Monitoring recovery after earthquakes through the integration of remote sensing, GIS, and ground observations: the case of L’Aquila (Italy),” *Cartography and Geographic Information Science*, vol. 43, no. 2, pp. 115–133, Apr. 2015, doi: https://doi.org/10.1080/15230406.2015.1029520.

[8] D. Contreras, S. Wilkinson, and P. James, “Earthquake Reconnaissance Data Sources, a Literature Review,” *Earth*, vol. 2, no. 4, pp. 1006–1037, Dec. 2021, doi: https://doi.org/10.3390/earth2040060.

[9] NASA Earth Science Data Systems, “What is Synthetic Aperture Radar?,” *EarthData*, 2019. https://www.earthdata.nasa.gov/learn/backgrounders/what-is-sar

[10] W. L. Yeck *et al.*, “Leveraging Deep Learning in Global 24/7 Real-Time Earthquake Monitoring at the National Earthquake Information Center,” *Seismological Research Letters*, vol. 92, no. 1, pp. 469–480, Sep. 2020, doi: https://doi.org/10.1785/0220200178.

[11] J. Woollam *et al.*, “SeisBench—A Toolbox for Machine Learning in Seismology,” *Seismological Research Letters*, vol. 93, no. 3, pp. 1695–1709, Mar. 2022, doi: https://doi.org/10.1785/0220210324.

[12] A. Radford, L. Metz, and S. Chintala, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks,” *arXiv.org*, 2015. https://arxiv.org/abs/1511.06434

[13] S. Polson, “/Unsupervised Machine Learning for Spacecraft Anomaly Detection in WebTCAD,” *Github*, Jun. 07, 2019. https://github.com/sapols/Satellite-Telemetry-Anomaly-Detection/blob/master/Paper/Unsupervised%20Machine%20Learning%20for%20Spacecraft%20Anomaly%20Detection%20in%20WebTCAD.pdf

[14] S. Derouin, “Desktop seismology: How a maker-inspired device is changing seismic monitoring,” *EARTH Magazine*, Nov. 08, 2017. https://www.earthmagazine.org/article/desktop-seismology-how-maker-inspired-device-changing-seismic-monitoring

[15] Raspberry Shake, “Professional Solutions,” *Raspberry Shake*. https://raspberryshake.org/professional-solutions/

[16] “Real-time Seismogram Displays,” *earthquake.usgs.gov*. https://earthquake.usgs.gov/monitoring/seismograms (accessed Apr. 28, 2024).

[17] “Quaternary Faults of West Coast | U.S. Geological Survey,” *www.usgs.gov*. https://www.usgs.gov/media/images/quaternary-faults-west-coast (accessed Apr. 29, 2024).

[18] B. Westover, “How to Get Starlink Satellite Internet and Set It Up the Right Way,” *PCMAG*, May 19, 2023. https://www.pcmag.com/how-to/starlink-starter-kit-how-to-get-it-and-set-it-up

[19] HomeAdvisor, "Learn how much it costs to Install a Satellite Dish.," *www.homeadvisor.com*, Oct. 24, 2022. https://www.homeadvisor.com/cost/home-theaters/install-a-satellite-dish/
[20] US Bureau of Labor Statistics, "May 2023 National Occupational Employment and Wage Estimates," *www.bls.gov*, Apr. 03, 2024. https://www.bls.gov/oes/current/oes_nat.htm#15-0000
[21] J. Daniell, A. Schaefer, F. Wenzel, and T. Kunz-Plapp, "The value of life in earthquakes and other natural disasters: historical costs and the benefits of investing in life safety," 2015. Available: https://aees.org.au/wp-content/uploads/2015/12/130-Daniell-James.pdf
[22] M. A. Salgado-Gálvez, G. A. Bernal, A. H. Barbat, M. L. Carreño, and O.-D. Cardona, "Probabilistic estimation of annual lost economic production due to premature deaths because of earthquakes," *Human and Ecological Risk Assessment: An International Journal*, vol. 22, no. 2, pp. 543–557, Oct. 2015, doi: https://doi.org/10.1080/10807039.2015.1095072.
[23] A. Kharb, S. Bhandari, M. Moitinho de Almeida, R. Castro Delgado, P. Arcos González, and S. Tubeuf, "Valuing Human Impact of Natural Disasters: A Review of Methods," *International Journal of Environmental Research and Public Health*, vol. 19, no. 18, p. 11486, Sep. 2022, doi: https://doi.org/10.3390/ijerph191811486.
[24] The Associated Press, "The world's Deadliest Earthquakes in the past 25 years, at a Glance," *NPR*, Feb. 09, 2023. https://www.npr.org/2023/02/09/1155836898/deadliest-earthquakes-list
[25] Multi-Hazard Mitigation Council (2019.). Natural Hazard Mitigation Saves: 2019 Report. Principal Investigator Porter, K.; Co-Principal Investigators Dash, N., Huyck, C., Santos, J., Scawthorn, C.; Investigators: Eguchi, M., Eguchi, R., Ghosh., S., Isteita, M., Mickey, K., Rashed, T., Reeder, A.; Schneider, P.; and Yuan, J., Directors, MMC. Investigator Intern: Cohen-Porter, A. National Institute of Building Sciences. Washington, DC. www.nibs.org
[26] BIDMC Communications, "Natural disaster recovery costs set to spiral, study shows," *Harvard Gazette*, Apr. 05, 2023. https://news.harvard.edu/gazette/story/2023/04/natural-disaster-recovery-costs-set-to-spiral-study-shows/
[27] CDC, "Stay Safe During an Earthquake | Natural Disasters and Severe Weather | CDC," *www.cdc.gov*, Sep. 24, 2020. https://www.cdc.gov/disasters/earthquakes/during.html
[28] A. Dogtiev, "App Development Cost," *Business of Apps*, Feb. 01, 2024. https://www.businessofapps.com/app-developers/research/app-development-cost/
[29] S. Burenko, "How much does software development costs?," *www.uptech.team*, Dec. 20, 2023. https://www.uptech.team/blog/software-development-costs
[30] U. S. G. A. Office, "National Security Space: Overview of Contracts for Commercial Satellite Imagery," *www.gao.gov*, Dec. 08, 2022. https://www.gao.gov/products/gao-23-106042
[31] Srikanth, L. Huang, C. Li, R. Stewart, and A. T. Becker, "Seismic surveying with drone-mounted geophones," *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, Aug. 2016, doi: https://doi.org/10.1109/coase.2016.7743566.
[32] "Drone Surveying: Best Survey Drones and More [New for 2023]," *UAV Coach*, 2024. https://uavcoach.com/drone-surveying/
[33] K. MARTINEZ, "Best Autopilot Drones [Updated 2020] Top 7 Drones With Autopilot Reviews," *Dronethusiast*, Nov. 18, 2019. https://www.dronethusiast.com/best-autopilot-drone-reviews/
[34] T. Boyd, "Geophones by Tom Boyd Colorado School of Mines," *Louisiana State University*, 2023. https://www.geol.lsu.edu/jlorenzo/ReflectSeismol03/Geophones_files/geophones.htm (accessed Apr. 28, 2024).
[35] GISCO, "Geophones," *giscogeo.com*. https://www.giscogeo.com/products-equipment/geophones (accessed Apr. 29, 2024).
[36] X. Meng, M. Ravanelli, A. Komjathy, and O. P. Verkhoglyadova, "On the North-South Asymmetry of Co-Seismic Ionospheric Disturbances During the 16 September 2015 Illapel M8.3 Earthquake," *Geophysical Research Letters*, vol. 49, no. 8, Apr. 2022, doi: https://doi.org/10.1029/2022gl098090.
[37] U.S. General Services Administration, "Satellite communications products and services," *www.gsa.gov*. https://www.gsa.gov/technology/it-contract-vehicles-and-purchasing-

programs/telecommunications-and-network-services/satellite-communications-products-and-services (accessed Apr. 29, 2024).
[38] S. Erwin, "SpaceX providing Starlink services to DoD under 'unique terms and conditions,'" *SpaceNews*, Oct. 03, 2023. https://spacenews.com/spacex-providing-starlink-services-to-dod-under-unique-terms-and-conditions/
[39] Office of Strategic Communication and Public Affairs, "DISA awards 16 contracts for Proliferated Low Earth Orbit Satellite-Based Services," *DISA*, Jul. 01, 2023. https://www.disa.mil/en/NewsandEvents/2023/DISA-awards-16-contracts
[40] IBM, "What Are Neural Networks?," *www.ibm.com*, 2023. https://www.ibm.com/topics/neural-networks
[41] FEMA, USGS, and NEHRP, "Hazus Estimated Annualized Earthquake Losses for the United States," FEMA, Apr. 2023. Available: https://www.fema.gov/sites/default/files/documents/fema_p-366-hazus-estimated-annualized-earthquake-losses-united-states.pdf
[42] J. Easterly and T. Fanning, "The Attack on Colonial Pipeline: What We've Learned & What We've Done Over the Past Two Years | CISA," *www.cisa.gov*, May 07, 2023. https://www.cisa.gov/news-events/news/attack-colonial-pipeline-what-weve-learned-what-weve-done-over-past-two-years
[43] "About Cal OES | California Governor's Office of Emergency Services," *Cal OES | CA.gov*. https://www.caloes.ca.gov/cal-oes/about-cal-oes/ (accessed Apr. 30, 2024).
[44] ShakeAlert, "ShakeAlert® | Earthquake Early Warning," *Shakealert.org*, 2019. https://www.shakealert.org/
[45] FEMA, "Home | FEMA.gov," *Fema.gov*, 2020. https://www.fema.gov/
[46] "Most Devastating and Largest Earthquakes in 2023 - Quantectum," *https://quantectum.com/*, Jan. 2024. https://quantectum.com/blog/most-devastating-and-largest-earthquakes-2023/