# Grindulu fault cloud radon data for earthquake magnitude prediction using machine learning

**Thomas Oka Pratama[1,2], Sunarno[2], Agus Budhie Wijatna[2], Eko Haryono[1,3]**
[1]Doctorate Programe in Environmental Science, Graduate School, Universitas Gadjah Mada, Yogyakarta, Indonesia
[2]Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia
[3]Department of Environmental Geography, Faculty of Geography, Universitas Gadjah Mada, Yogyakarta, Indonesia

**Article Info**

***Article history:***

Received Feb 13, 2024
Revised Apr 24, 2024
Accepted Jun 5, 2024

***Keywords:***

Earthquake
Machine learning
Magnitude
Prediction
Radon

**ABSTRACT**

The study investigates the potential of integrating radon gas concentration telemonitoring systems with machine learning techniques to enhance earthquake magnitude prediction. Conducted in Pacitan, East Java, Indonesia, where the stations are near the active Grindulu fault, the research employs random forest (RF), extreme gradient boosting (XGB), neural network (NN), AdaBoost (AB), and support vector machine (SVM) methods. The study aims to refine earthquake magnitude prediction, utilizing real-time radon gas concentration measurements, crucial for disaster preparedness. The evaluation involves multiple metrics like mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), mean squared error (MSE), symmetric mean absolute percentage error (SMAPE), and conformal normalized mean absolute percentage error (cnSMAPE). XGB and SVM emerge as top performers, showcasing superior predictive accuracy with minimal errors across various metrics. XGB achieved MAE (0.33), MAPE (6.03%), RMSE (0.51), MSE (0.26), SMAPE (0.06), and cnMAPE (0.97), while SVM recorded MAE (0.34), MAPE (6.20%), RMSE (0.51), MSE (0.26), SMAPE (0.06), and cnSMAPE (0.97). The analysis reveals XGB as the most effective method, boasting the lowest error values. The study underscores the importance of expanding data availability to enhance predictive models, ultimately contributing to more precise earthquake magnitude predictions and effective mitigation strategies.



***Corresponding Author:***

Sunarno
Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada
St. Grafika No. 2, Sinduadi, Mlati, Sleman, Yogyakarta, Indonesia
Email: sunarno@ugm.ac.id

## 1. INTRODUCTION

The prediction of earthquake magnitude is a critical area of research due to its potential impact on public safety and infrastructure. Radon gas has been regarded as a possible precursor to earthquakes for a considerable time, and creating a comprehensive forecasting system that can accurately predict the date, time, magnitude, and exact epicenter of future earthquakes continues to be a challenge [1]–[11]. One approach to earthquake magnitude prediction involves using a telemonitoring system based on radon gas concentration coupled with machine learning techniques.

Presently, artificial intelligence (AI) methodologies play a crucial role in earthquake prediction, highlighting their capability to enhance readiness and response tactics in vulnerable regions [5], [11]–[17]. This approach leverages advanced technologies to analyze and interpret data for accurate predictions. Machine learning algorithms, such as support vector machines (SVM), extreme learning machines (ELM), and artificial neural networks (ANNs), have been applied in earthquake magnitude prediction models [14], [18]–[21]. Additionally, integrating adaptive neuro-fuzzy inference systems (ANFIS) has shown promise in enhancing the accuracy of earthquake magnitude predictions [22]. Furthermore, the use of ensemble learning methods, such as AdaBoost, has been explored to improve the robustness of earthquake magnitude prediction models [23].

Various machine learning methods have been employed in different studies to predict earthquake magnitudes. For instance, the minimum area of the alarm method was utilized to predict earthquakes with magnitudes greater than 6.0 and 5.5 in Japan and California [24]. Similarly, machine learning techniques were applied for earthquake magnitude prediction in the Hindukush region [25]. Furthermore, a probabilistic neural network (NN) was developed for earthquake magnitude prediction, achieving good earthquake results with magnitudes ranging from 4.0 to 6.0 [26]. Additionally, it proposed a deep-learning NN for large earthquake magnitude prediction in Taiwan, obtaining successful predictions for earthquakes with magnitudes between 4.0 and 6.0 [27]. These studies demonstrate the effectiveness of machine learning methods in earthquake magnitude prediction.

Earthquake magnitude prediction has been addressed through various machine learning methods, including probabilistic NN and deep learning NN. The incorporation of radon gas concentration data into the prediction models aligns with the findings of those who investigated the characteristics of geoelectric field signals prior to earthquakes [28]. This suggests that diverse geophysical data, including radon gas concentration, contributes to a comprehensive understanding of earthquake precursors. Moreover, Gómez *et al.* [29] presented a novel methodology for predicting large-magnitude earthquakes in chile using ensemble learning, emphasizing the potential of machine learning in forecasting seismic events with significant magnitudes. Additionally, Gitis and Derendyaev [30] discussed machine learning methods for spatial forecasting of maximum possible earthquake magnitudes, further highlighting the relevance of machine learning in seismic hazard assessments.

Considering the application of AI in earthquake prediction, Banna *et al.* [11] analysis underscores its success in forecasting earthquakes within certain magnitude thresholds (M3 to M5). However, challenges arise in predicting high-magnitude events due to their infrequency and unpredictable occurrence patterns. Significantly, notable discrepancies have been noted in predicting the timing and location of earthquakes, with deviations of up to 70 miles and considerable variations in prediction timeframes ranging from 20 days to 5 months [11]. According to Tehseen *et al*. [31], Table 1 presents the accuracy of the expert system suggested for forecasting earthquakes through an independent test dataset. These results highlight the intricacy and intrinsic uncertainties linked with earthquake forecasting, prompting continuous exploration and improvement of methodologies to boost predictive accuracy and dependability. The research focus has shifted toward machine learning and deep learning techniques since 2018, signifying a noteworthy evolution in earthquake prediction approaches.

Table 1. Accuracy is claimed in an expert system using an independent test set [31]

| References | Number of earthquake records | Accuracy (%) | Magnitude range |
|---|---|---|---|
| [32] | 9,531 | 69.8 | ≥2.0 |
| [33] | 12,690 | 50.14 | ≥3.0 |
| [34] | 337 | 63 | ≥3.0 |
| [35] | 10,567 | 40 | 0.1–5.9 |

Integrating radon gas concentration telemonitoring systems with machine learning techniques presents a promising approach to earthquake magnitude prediction. By leveraging advanced algorithms and diverse geophysical data, researchers aim to improve the accuracy and reliability of earthquake magnitude predictions, ultimately contributing to enhanced disaster preparedness and risk mitigation. In Indonesia, there is a method for predicting the time of an earthquake 1-4 days after the alarm goes off with a magnitude greater than 4.5M based on the radon gas concentration telemonitoring, but it does not specifically predict the magnitude of the earthquake [5], [36]. In this study, earthquake magnitude prediction is conducted based on real-time telemonitoring of radon gas concentration at stations in Pacitan, East Java, Indonesia near Grindulu fault using machine learning methods random forest (RF), extreme gradient boosting (XGB), NN, adaboost (AB), SVM. These stations already have an earthquake time prediction algorithm based on Pratama *et al*. [36] with sensitivity and precision of 78.79% and 70.27%, respectively. The development of the earthquake magnitude prediction system can help mitigate earthquake natural disasters so that it can provide warnings to the public, volunteers, or the government to minimize casualties.

## 2. METHOD

The real-time monitoring system for radon gas concentration is close to the Grindulu active fault in Pacitan, East Java, Indonesia, making it prone to seismic activity. The radon gas transducer is strategically positioned within the chamber room to ensure precise measurement of radon gas emissions, maintaining a maximum of 4.142 cm above ground level. Radon gas measurements are taken every 10 minutes to mitigate the influence of radiation emissions [37]. The earthquake prediction system's design is illustrated in Figure 1. Subsequently, data collected by the radon gas transducer is sent to the microprocessor and forwarded to the cloud server for real-time monitoring, contingent upon an active internet connection. Notably, the recorded radon gas data is securely stored in a designated data storage server and accessible via a dedicated web server. Moreover, earthquake-related data is acquired from Geofon Potsdam and the Indonesian Agency for meteorology, climatology, and geophysics, enriching the understanding of seismic events.

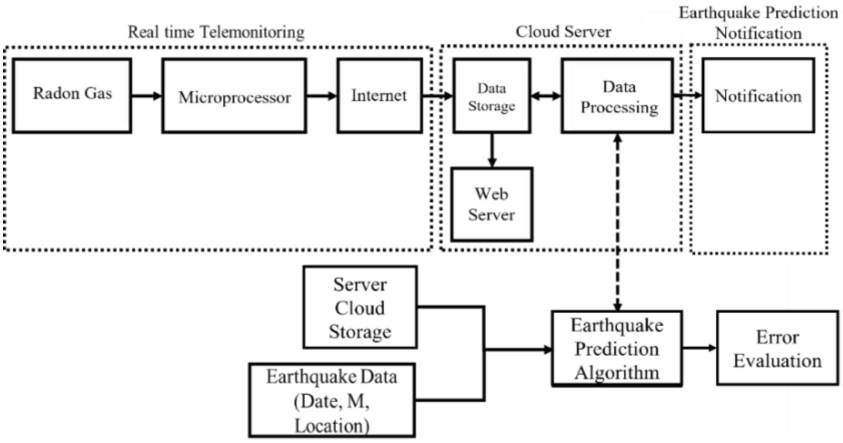



Figure 1. Earthquake prediction system block diagram [38]

The algorithm for predicting earthquake magnitude is created through supervised machine learning methods, utilizing data extracted from radon cloud data and historical Earthquake incidents. The model's performance assessment involves various metrics, including accuracy, mean squared error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE), symmetric mean absolute percentage error (SMAPE), conformal normalized mean absolute percentage error (cnSMAPE), and accuracy. Following evaluation, the most effective model will be implemented on a cloud server to deliver earthquake prediction alerts.

Table 2 illustrates the composition of the radon dataset [5]. Subsequently, Table 3 presents the systematic organization of information regarding radon gas concentrations and earthquake occurrences, adhering to Pratama approach [5]. The training dataset consisted of radon gas concentration data corresponding to the day of the earthquake prediction. In contrast, the test dataset included earthquakes within 1-4 days after the initial earthquake date prediction, with a minimum magnitude threshold of M4.5 between the Eurasia and Indo-Australia Plates. The data collection period ranged from 2/2/2022 to 22/1/2024, comprising 117 data points. Seventy percent of the dataset was allocated for training purposes, with the remaining portion reserved for evaluating and predicting earthquake magnitudes.

Table 2. Composition of the dataset [5]

| Variable | Description |
|---|---|
| d | The day on which the algorithm forecast was finalized utilizing the Pratama methodology [5] |
| Rd | Radon data average day d |
| R(d-1) | Radon data average day d-1 |
| R(d-2) | Radon data average day d-2 |
| . | . |
| . | . |
| R(d-6) | Radon data average day d-6 |
| R(d-7) | Radon data average day d-7 |
| . | . |
| . | . |
| XR(d-3) | Radon data average 3 days before R(d-2) = average R(d-3) to R(d-5) |
| XR(d-7) | Radon data average 7 days before R(d-2) = average R(d-3) to R(d-9) |
| XR(d-14) | Radon data average 14 days before R(d-2) = average R(d-3) to R(d-17) |

Table 3. An instance of a dataset

| Earthquake date prediction | XR (d-14) | XR (d-7) | XR (d-3) | R (d-7) | R (d-6) | R (d-5) | R (d-4) | R (d-3) | R (d-2) | R (d-1) | Earthquake date | Distance (km) | Actual Magnitude |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2/2/2022 | 58.02 | 58.81 | 68.17 | 45.24 | 64.41 | 77.03 | 71.48 | 56.00 | 33.43 | 49.93 | 2/4/2022 | 577.73 | 5.5 |
| 2/9/2022 | 53.88 | 48.96 | 57.33 | 40.83 | 46.57 | 43.88 | 55.47 | 72.62 | 74.74 | 47.90 | 2/10/2022 | 1,010.32 | 4.6 |
| 2/15/2022 | 56.54 | 66.49 | 70.63 | 47.90 | 58.30 | 63.77 | 60.03 | 88.08 | 78.26 | 47.48 | 2/17/2022 | 468.23 | 5.1 |

The machine learning process employed in this study follows supervised learning principles, utilizing a regression technique as depicted in Figure 2. The objective is for the model to discern underlying patterns or relationships within the dataset, enabling accurate predictions of earthquake magnitudes on unseen data based on the radon gas concentration from a telemonitoring station near the Grindulu fault. RF, XGB, NN, AB, and SVM are machine learning techniques used in deriving earthquake magnitude prediction algorithms [27]–[34]. The training dataset is utilized to construct the earthquake magnitude prediction model, which is subsequently evaluated using the test dataset to assess its efficacy.

This research implemented RF, XGB, NN, AB, and SVM methodologies using orange data mining version 3.36.2 software. The tuning is done on the features of each machine learning method to obtain the best model. By integrating the application of machine learning models with evaluations based on various metrics, this study aims to offer a comprehensive insight into the model's performance in predicting or analyzing the data under consideration.

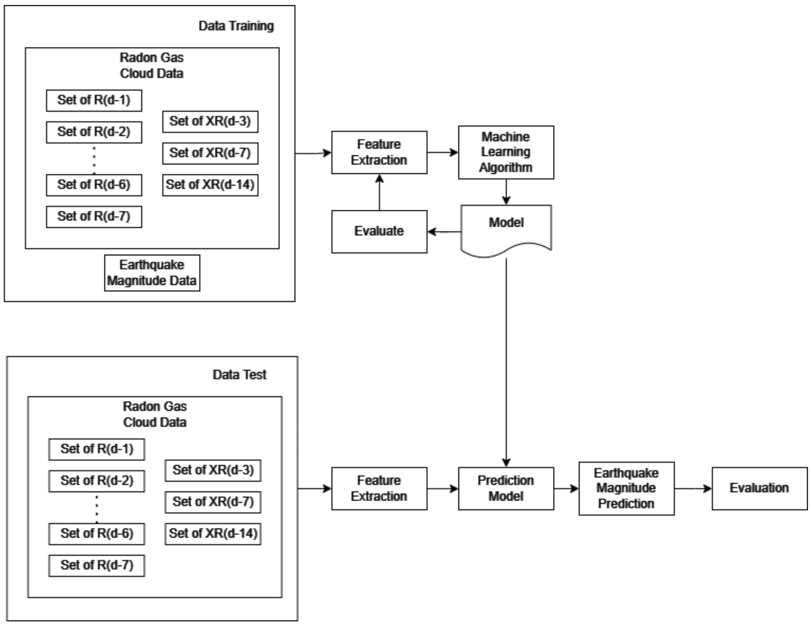



Figure 2. Structure of a supervised machine learning model [38]

## 3. RESULTS AND DISCUSSION

This study's earthquake magnitude prediction is based on radon gas concentration measurements near the active Grindulu fault. The radon gas concentration data storage is configured according to the predetermined dataset. The dataset for this study consists of 82 training data and 35 test data points. From this training data set, modeling is performed using RF, XGB, NN, AB, and SVM methods, each with different feature settings, until the best results are achieved. The outcome derived from this machine learning procedure yields the forecasted magnitude of forthcoming earthquakes based on the input test data.

Table 4 illustrates the forecast outcomes derived from the training dataset using a confusion matrix and metrics such as MAE, MSE, RMSE, and standard deviation representing the variance between actual and predicted magnitude. A true positive scenario emerges when the actual magnitude lies within the predicted magnitude range plus or minus the standard deviation error. Conversely, a false positive arises when the actual magnitude falls outside the predicted magnitude range plus or minus the standard deviation error. The evaluation results of the training data showed that the XGB method has the highest accuracy compared to other methods, with a true positive of 68, followed by the AB, NN, SVM, and RF with true positives consecutively

58, 53, 51, and 51. The XGB method has the highest accuracy with a value of 82.93% and the lowest standard deviation, MAE, MSE, and RMSE values compared to other methods. It indicates better accuracy compared to other models. The models vary in ability to correctly identify positive cases (true positives) and minimize false positives.

Table 4. Machine learning data training test result

| Parameter | Learning Methods | | | | |
|---|---|---|---|---|---|
| | RF | XGB | NN | AB | SVM |
| Standard Deviation | 0.29 | 0.15 | 0.5 | 0.2 | 0.44 |
| MAE | 0.21 | 0.05 | 0.37 | 0.08 | 0.34 |
| MSE | 0.08 | 0.02 | 0.27 | 0.04 | 0.19 |
| RMSE | 0.29 | 0.15 | 0.52 | 0.2 | 0.43 |
| True Positive | 50 | 68 | 53 | 58 | 51 |
| False Positive | 32 | 14 | 29 | 24 | 31 |
| Accuracy (%) | 60.98 | 82.93 | 64.63 | 70.73 | 62.20 |

Based on the evaluation results obtained for the predictive algorithm models, earthquake magnitude predictions were made using test data based on earthquake date predictions from Pratama [5] method, where earthquake predictions are valid 1-4 days after the predicted earthquake date. Table 5 illustrates the error evaluation of the machine learning-based earthquake magnitude prediction method subsequent to testing the training dataset and configuring features to generate optimal predictions. The RF and SVM method has the lowest standard deviation (0.38). The standard deviation of absolute error evaluation has a value that does not differ much between machine learning methods. The XGB has the lowest value of MAE (0.33), MAPE (6.03%), RMSE (0.51), MSE (0.26), SMAPE (0.06) and cnMAPE (0.97). Smaller values of these metrics suggest superior performance of the algorithm. SVM also performs consistently well, with competitive values across most error indices. RF, NN, and AB show slightly higher error indices than XGB and SVM, indicating slightly lower predictive accuracy. SVM has the highest magnitude prediction accuracy with a value of 77.14%, followed by AB (71.43%), NN (68.57%), XGB, and RF (65.71%). XGB and SVM appear to be the top-performing learning methods based on the provided error indices. According to our analysis, XGB has the lowest MAE and RMSE values, indicating good performance in minimizing the absolute error in regression tasks. However, SVM achieved the highest accuracy among the classification task algorithms. These results suggest that XGB performs better in tasks where absolute error minimization is important, whereas SVM performs better in tasks where classification accuracy is paramount. Researchers and practitioners can use these insights to select the most appropriate machine learning algorithms based on their specific goals and performance criteria.

Table 5. Earthquake magnitude prediction error evaluation

| Error index | Learning methods | | | | |
|---|---|---|---|---|---|
| | RF | XGB | NN | AB | SVM |
| St Dev of Absolute Error | 0.38 | 0.40 | 0.39 | 0.40 | 0.38 |
| MAE | 0.37 | 0.33 | 0.38 | 0.36 | 0.34 |
| MAPE (%) | 6.91 | 6.03 | 6.90 | 6.55 | 6.20 |
| RMSE | 0.53 | 0.51 | 0.54 | 0.53 | 0.51 |
| MSE | 0.28 | 0.26 | 0.29 | 0.29 | 0.26 |
| SMAPE | 0.07 | 0.06 | 0.07 | 0.07 | 0.06 |
| cnSMAPE | 0.96 | 0.97 | 0.96 | 0.97 | 0.97 |
| Accuracy (%) | 65.71 | 65.71 | 68.57 | 71.43 | 77.14 |

Figure 3 displays the dispersion errors utilizing boxplot representation for each method, elucidating their error characteristics. The NN method exhibits the greatest error dispersion, trailed by RF, AB, and XGB, with SVM demonstrating the lowest error dispersion based on the dataset. The XGB method has a median value closest to 0 compared to the other method. There are 2 outlier data points for all methods from the earthquake prediction evaluation based on radon gas concentration data. They represent data points that are unusually high compared to the rest of the dataset.

Figure 4 shows the histogram of errors for XGB in Figure 4(a), RF in Figure 4(b), NN in Figure 4(c), AB in Figure 4(d), and SVM in Figure 4(e) are scrutinized meticulously to examine the sign of deviation. All machine learning methods used in the study have the highest frequency of values at 0 to -0.25 M error. The XGB has the highest error frequency with a quantity of 14 at that frequency, followed by NN and SVM with 12 and RF and AB methods with 10. They produced a higher quantity of negative errors than negative bias, as can be noticed in the histograms of Figure 4. Negative bias refers to a systematic tendency for estimates to consistently underestimate the true value or exhibit a tendency toward lower values. In statistical terms, it means that the average of the estimates tends to be lower than the real value of magnitude.

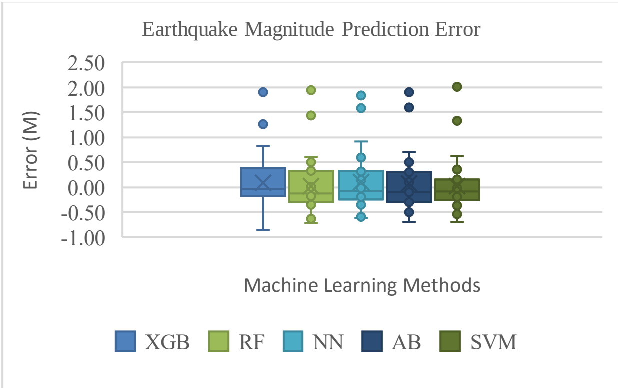



Figure 3. Boxplot generated by machine learning algorithms during the earthquake prediction of the 35 data test set

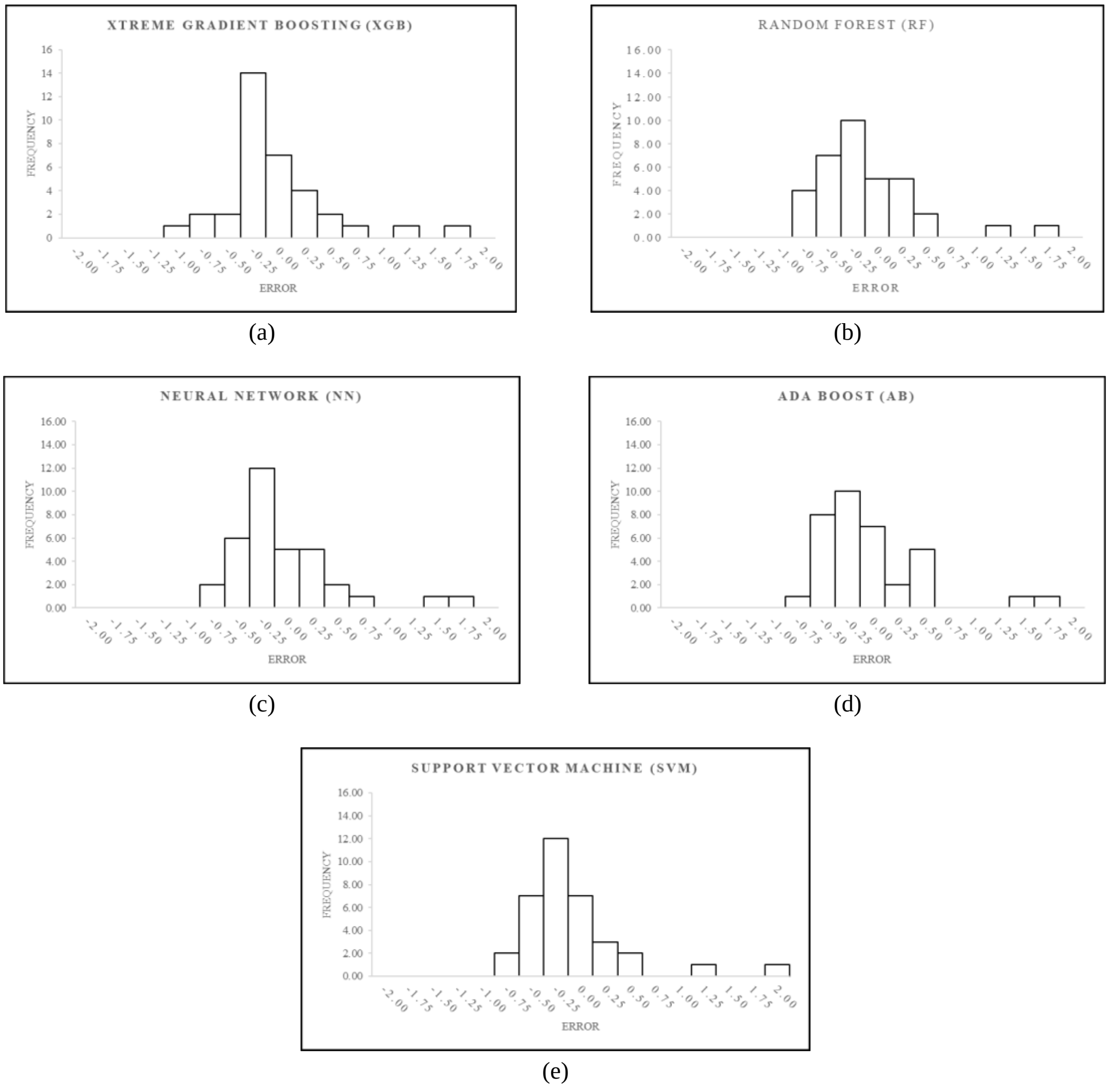
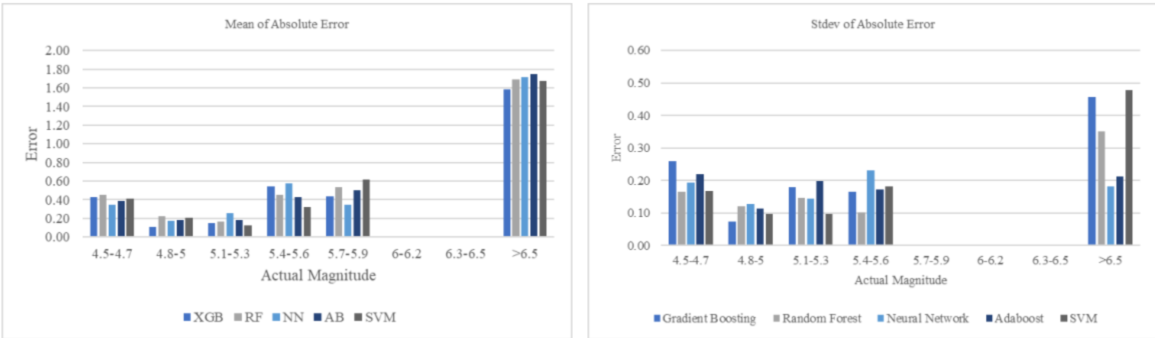



Figure 4. Histograms of the errors produced by (a) XGB, (b) RF, (c) NN, (d) AB, and (e) SVM algorithms when predicting the 35 data test set

In this research, we analyzed errors as depicted in Figure 5(a) for means of absolute error and Figure 5(b) for the standard deviation of absolute error. The results reveal that the XGB method exhibits the lowest MAE for the M4.8-M5 earthquake, registering at 0.11 and 1.58 for >M6.5. Additionally, the XGB method demonstrates a low standard deviation of absolute error across most magnitude ranges, with the lowest recorded at 0.07 for the M4.8-M5 magnitude range. The NN method displays the lowest MAE for M4.5-M4.7 and M5.7-M5.9, recording values of 0.35. Meanwhile, the SVM method showcases the lowest MAE for M5.1-M5.3 and M5.4-M5.6, with values of 0.13 and 0.32, respectively. Conversely, the RF method exhibits the highest MAE for M4.5-M4.7 and M4.8-M5, whereas the AB method shows the highest MAE for magnitudes above 6.5.




(a) (b)

Figure 5. Analysis of the (a) mean and (b) standard deviation of the absolute error of earthquake magnitude prediction error

Earthquakes above M5.5 are infrequent occurrences. Increasing the available data will enhance the system's learning capabilities, enabling more precise and accurate predictions of earthquake magnitudes. Continuous measurement data of radon gas concentration can assist researchers in designing earthquake prediction algorithms, thus aiding in earthquake mitigation efforts, particularly in Indonesia. Features studies may explore earthquake prediction based on radon gas concentration and total electron content (TEC) measurements measured at the exact location so as to design a reliable earthquake prediction system.

A unique methodology for earthquake magnitude prediction, which uses radon gas concentration measurements along the active fault of Grindulu in Pacitan, East Java, has shown better results than earlier study approaches in Table 6 (see appendix). MAE, MAPE, RMSE, MSE, and SMAPE were all lower than previously reported levels. Other studies have found limits in earthquake magnitude estimates, with numbers that are either too wide or too narrow. Long forecast durations provide issues, either resulting in lengthy periods of alert or insufficient time for escape. Previous research also lacked information about the expected location (such as which plate or fault section). Furthermore, Indonesia had few earthquake magnitude forecasts based on radon gas concentrations. This research has a cnSMAPE of 0.97 and an accuracy of 77.14%. The accuracy of the prediction approach used herein outperforms prior research, which failed to meet a 75% accuracy bar for earthquake magnitude forecasts while keeping precise prediction values and appropriate prediction durations.

## 4. CONCLUSION

The earthquake magnitude prediction was conducted using radon gas concentration measurements at the Pacitan station, East Java, Indonesia, situated near the active Grindulu fault utilized for modeling using RF, XGB, NN, AB, and SVM methods with various feature settings for 1-4 days earthquake prediction between Eurasia and Indo-Australia Plates. Error evaluations of machine learning methods were conducted, including Relative error, MAE, MAPE, RMSE, MSE, SMAPE, and cnSMAPE. XGB and SVM emerged as the best-performing methods, producing the lowest error values across these metrics. Specifically, XGB achieved MAE (0.33), MAPE (6.03%), RMSE (0.51), MSE (0.26), SMAPE (0.06), cnMAPE (0.97), and accuracy (65.71%) , while SVM recorded MAE (0.34), MAPE (6.20%), RMSE (0.51), MSE (0.26), SMAPE (0.06), cnMAPE (0.97), and accuracy (77.14%). A unique method for predicting earthquake magnitude, using measurements of radon concentrations along the active Grindulu fault in Pacitan, East Java, performs better than previous research methods on radon. Nevertheless, increasing data availability can enhance the learning capabilities of predictive models, contributing to more precise earthquake magnitude predictions and aiding in earthquake mitigation efforts.

## ACKNOWLEDGEMENTS

The author thanks the Sensor and Tele-Control Systems Laboratory research team at the Department of Nuclear Engineering and Engineering Physics, Universitas Gadjah Mada, Indonesia, for furnishing the radon gas telemonitoring system. Appreciation is extended to the Postdam Geofon and the Badan Meteorologi, Klimatologi dan Geofisika Indonesia for supplying the earthquake data. We also acknowledge the support of the RTA Program at Universitas Gadjah Mada with Grant Number 5075/UN1.P.II/Dit-Lit/PT.01.01/2023.

## APPENDIX

Table 6. Other research earthquake prediction results (*continued…)*

| Ref. | Variable (s) | Method | Limitation | Result |
|---|---|---|---|---|
| [12] | Animal behaviour, environmental dynamics, and chemical changes | Belief rule-based expert system (BRBES) | Time: 12-hour timeframe<br>Location: -<br>Magnitude: M>6.5 | AUC<br>BRBES: 0.969, FLBES: 0.789, ANN: 0.862 |
| [32] | Independent data test | Expert system | Time: -<br>Location: -<br>Magnitude: M0.1 – M5.9 | Accuracy below 70% |
| [5] | Radon and Groundwater level | Statistic | Time: 1-4 days<br>Location: Between Eurasia and Indo-Australia Plates<br>Magnitude: >4.5 | Sensitivity and precision >80% (time prediction) |
| [13] | Seismicity | Expert system | Time: 12-hour timeframe<br>Location: one-quarter of the Earth<br>Magnitude: M3.6 - M9.1 | Accuracy 100% |
| [14] | Meteorological and seismic data | Support vector regression | Prediction number of earthquakes in a month. The average magnitude of an earthquake in a month | Precision 96% predicting the mean magnitude. Accuracy 78% for the expected earthquake count in a month |
| [15] | Seismic parameters | Pattern recognition NN, recurrent neural network, RF and linear programming boost ensemble | Time: one month<br>Location: Hindukush<br>Magnitude: $\geq$5.5 | Training: Accuracy 79%<br>Test: Accuracy 65% |
| [16] | Cloud-based big data infrastructure | Regression algorithm | Time: 7 days<br>Location: California<br>Magnitude: M3- M7 | MAE:<br>0:59 ± 0:66 (M3-4)<br>0:25 ± 0:52 (M4-5)<br>0:27 ± 0:60 (M5-6)<br>0:28 ± 0:75 (M6-7)<br><br>MSE:<br>0:79 ± 1:53 (M3-4)<br>0:34 ± 1:42 (M4-5)<br>0:43 ± 2:06 (M5-6)<br>0:63 ± 2:73 (M6-7) |
| [18] | Seismic data | Multilayer perceptron NN | Time: days<br>Location: -<br>Magnitude: 4 Classification, M > 4 | Accuracy:<br>73.79% |
| [19] | Seismic Electric Signals (SES) | ANN | Time: days<br>Location: region of Greece<br>Magnitude: M2.1 – M5.2 and M>5.2 | Accuracy 84.01% (all), 58.02%, M $\geq$ 5.2 on the Richter scale). |
| [20] | Seismic data | ANFIS | Time: -Location: region of Iran<br>Magnitude: M > 5.5 | ANFIS by Grid Partition algorithm<br>$R^2$ = 0.94, MAE = 0.149, RMSE = 0.173 |
| [26] | Seismic data | PNN | Time: month<br>Location: California<br>Magnitude: 7 classification, M < 4.5 – M > 7.5 | RScore: 0.62- 0.78 (M 4.5 – M6.0) |
| [27] | Historical seismic events | NN model | Time: 30 days<br>Location: specific location or area in Taiwan<br>Magnitude: $\geq$ M6 | RScore is 0.303 |

Table 6. Other research earthquake prediction results

| Ref | Variable (s) | Method | Limitation | Result |
|---|---|---|---|---|
| [29] | Seismic data | Classifiers and ensemble learning | Time: 5 days<br>Location: Chile (Santiago, Valparaíso, Talca and Pichilemu)<br>Magnitude: M4 – M7 | Santiago:<br>Sensitivity 0.46<br>Specificity 0.98<br>Positive predictive value 0.65<br>F score 0.75<br>Auc 0.72<br><br>Valparaiso M5<br>Sensitivity 0.52<br>Specificity 0.89<br>Positive predictive value 0.36<br>F score 0.67<br>Auc 0.70<br><br>Talca<br>Sensitivity 0.9<br>Specificity 1<br>Positive predictive value 1<br>F score 0.98<br>Auc 0.97 |
| [39] | Seismic data | SVM and Naïve Bayes | Time: day<br>Location: Indonesia<br>Magnitude: - | SVM<br>RMSE: 0.751008, MAPE: 0.156257, MSE: 0.564013, MAE: 0.598473. |
| [40] | Climate data, seismic data | Long short-term memory (LSTM), bidirectional long short-term memory (Bi-LSTM), and transformer models | Time: day<br>Location: Japan, Indonesia, and the Hindu-Kush Karakoram Himalayan (HKKH) region<br>Magnitude: M3.8 – M5.8 | Indonesia<br>MAE = 0.066,<br>MSE = 0.007,<br>log cosh loss = 0.039,<br>MSLE = 0.003.<br>Jepang<br>MAE = 0.083,<br>MSE = 0.015,<br>log cosh loss = 0.054,<br>MSLE = 0.007.<br>HKKH<br>MAE = 0.083,<br>MSE = 0.011,<br>log cosh loss = 0.039,<br>MSLE = 0.005 |

## REFERENCES


[1] C. H. Scholz, L. R. Sykes, and Y. P. Aggarwal, “Earthquake prediction: A physical basis,” *Science*, vol. 181, no. 4102, pp. 803–810, 1973.

[2] M. Al-Hilal and M. Mouty, “Radon monitoring for earthquake prediction on Al-Ghab fault of Syria,” *Nuclear Geophysics*, vol. 8, no. 3, pp. 291-299, 1994.

[3] A. Alam, N. Wang, G. Zhao, and A. Barkat, “Implication of radon monitoring for earthquake surveillance using statistical techniques: a case study of wenchuan earthquake,” *Geofluids*, vol. 2020, no. 1, 2020, doi: 10.1155/2020/2429165.

[4] Sunarno, H. L. Firdaus, Y. F. Luckyarno, M. M. Waruwu, and R. Wijaya, “Detection system for deterministic earthquake prediction based on radon concentration changes in Indonesia,” *Journal of Engineering Science and Technology*, vol. 15, no. 3, pp. 1787–1798, 2020.

[5] T. O. Pratama, “Earthquake early warning system based on radon gas concentration and groundwater level fluctuation at Yogyakarta Region-Indonesia,” *M.Eng. Thesis*, Universitas Gadjah Mada, Yogyakarta, Indonesia, 2021.

[6] E. Dazé, “Earthquake precursors: radon gas,” *Slide Player*. [Online]. Available: https://slideplayer.com/slide/15798487/

[7] A. S. Bafti, “The relation of concentration radon gas in underground water with tectonic movements and Seismicity. Case Study: Jorjafk fault, Central Iran,” *Scientific Quarterly Journal of Geosciences*, vol. 30, no. 117, doi: 10.22071/gsj.2018.91418.1186.

[8] L. Cao and A. Gao, “Correlativity between short-term change of fault gas radon concentration and air temperature & atmospheric pressure,” *Acta Seismologica Sinica*, no. 4, pp. 719–729, 2014, doi: 10.3969/j.issn.0253-3782.2014.04.017.

[9] C. -C. Fu *et al.*, “Preseismic anomalies in soil-gas radon associated with 2016 M 6.6 Meinong earthquake, Southern Taiwan,” *Terrestrial, Atmospheric and Oceanic Sciences*, vol. 28, no. 5, pp. 787-798, 2017, doi: 10.3319/TAO.2017.03.22.01.

[10] Sunarno, M. M. Waruwu, and R. Wijaya, “Development of the real time telemonitoring system for earthquake prediction deduced from fluctuations in groundwater levels at Yogyakarta region-Indonesia,” *Journal of Theoretical and Applied Information Technology*, vol. 83, no. 1, pp. 95–99, 2016.

[11] Md. H. A. Banna, K. A. Taher, M. S. Kaiser, M. Mahmud, Md. S. Rahman, A. S. M. S. Hosen, and G. H. Cho, “Application of artificial intelligence in predicting earthquakes: state-of-the-art and future challenges,” *IEEE Access*, vol. 8, pp. 192880–192923, 2020.

[12] M. S. Hossain, A. A. Hasan, S. Guha, and K. Andersson, “A belief rule based expert system to predict earthquake under uncertainty,” *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 9, no. 2, pp. 26-41, 2018.

[13] A. Ikram and U. Qamar, “Developing an expert system based on association rules and predicate logic for earthquake prediction,” *Knowledge-Based Systems*, vol. 75, pp. 87–103, Feb. 2015, doi: 10.1016/j.knosys.2014.11.024.

[14] P. Hajikhodaverdikhan, M. Nazari, M. Mohsenizadeh, S. Shamshirband, and K.-W. Chau, “Earthquake prediction with meteorological data by particle filter-based support vector regression,” *Engineering Applications of Computational Fluid Mechanics*, vol. 12, no. 1, pp. 679–688, 2018, doi: 10.1080/19942060.2018.1512010.

[15] K. M. Asim, F. Martínez-Álvarez, A. Basit, and T. Iqbal, “Earthquake magnitude prediction in Hindukush region using machine learning techniques,” *Nat Hazards*, vol. 85, no. 1, pp. 471–486, Jan. 2017.
[16] G. A. -Cortés, A. M. -Esteban, X. Shang, and F. M. -Álvarez, “Earthquake prediction in California using regression algorithms and cloud-based big data infrastructure,” *Computers and Geosciences*, vol. 115, 2017, pp. 198–210, 2018.
[17] T. Khan, M. Rabbani, S. M. T. Siddiquee, and A. Majumder, “An innovative data mining approach for determine earthquake probability based on linear regression algorithm,” in *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India, 2019, pp. 1-4, doi: 10.1109/ICECCT.2019.8869286.
[18] J. Mahmoudi, M. A. Arjomand, M. Rezaei, and M. H. Mohammadi, “Predicting the Earthquake Magnitude Using the Multilayer Perceptron Neural Network with Two Hidden Layers,” *Civil Engineering Journal*, vol. 2, no. 1, pp. 1–12, Jan. 2016.
[19] M. Moustra, M. Avraamides, and C. Christodoulou, “Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals,” *Expert Systems with Applications*, vol. 38, no. 12, pp. 15032–15039, Nov. 2011.
[20] M. Mirrashid, “Earthquake magnitude prediction by adaptive neuro-fuzzy inference system (ANFIS) based on fuzzy C-means algorithm,” *Natural Hazards*, vol. 74, pp. 1577–1593, Dec. 2014.
[21] S. Zhang, Z. Shi, G. Wang, R. Yan, and Z. Zhang, “Application of the extreme gradient boosting method to quantitatively analyze the mechanism of radon anomalous change in Banglazhang hot spring before the Lijiang Mw 7.0 earthquake,” *Journal of Hydrology*, vol. 612, Sep. 2022, doi: 10.1016/j.jhydrol.2022.128249.
[22] X. Wu *et al.*, “Top 10 algorithms in data mining,” *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, Jan. 2008, doi: 10.1007/s10115-007-0114-2.
[23] W. Astuti, W. Sediono, R. Akmeliawati, A. M. Aibinu, and M. J. E. Salami, “Investigation of the characteristics of geoelectric field signals prior to earthquakes using adaptive STFT techniques,” *Natural Hazards and Earth System Science*, 2013.
[24] V. G. Gitis, A. B. Derendyaev, and K. N. Petrov, “Analyzing the performance of GPS data for earthquake prediction,” *Remote Sensing*, vol. 13, no. 9, 2021, doi: 10.3390/rs13091842.
[25] K. M. Asim, F. M. -Álvarez, A. Basit, and T. Iqbal, “Earthquake magnitude prediction in hindukush region using machine learning techniques,” *Natural Hazards*, vol. 85, pp. 471–486, 2016, doi: 10.1007/s11069-016-2579-3.
[26] H. Adeli and A. Panakkat, “A probabilistic neural network for earthquake magnitude prediction,” *Neural Networks*, vol. 22, no. 7, pp. 1018-1024, 2009, doi: 10.1016/j.neunet.2009.05.003.
[27] J. P. Huang, X. A. Wang, Y. Zhao, C. Xin, and H. Xiang, “Large earthquake magnitude prediction in taiwan based on deep learning neural network,” *Neural Network World*, pp. 149–160, 2018.
[28] D. Marchetti *et al.*, “Worldwide statistical correlation of eight years of swarm satellite data with m5.5+ earthquakes: new hints about the preseismic phenomena from space,” *Remote Sensing*, vol. 14, no. 11, 2022, doi: 10.3390/rs14112649.
[29] M. F. -Gómez, G. A. -Cortés, A. Troncoso, and F. M. -Álvarez, “Large earthquake magnitude prediction in chile with imbalanced classifiers and ensemble learning,” *Applied Sciences*, vol. 7, no. 6, Jun. 2017, doi: 10.3390/app7060625.
[30] V. Gitis and A. Derendyaev, “The method of the minimum area of alarm for earthquake magnitude prediction,” *Frontiers in Earth Science*, vol. 11, Nov. 2020, doi: 10.3389/feart.2020.585317.
[31] R. Tehseen, M. S. Farooq, and A. Abid, “Earthquake Prediction using expert systems: a systematic mapping study,” *Sustainability*, vol. 12, no. 6, Mar. 2020, doi: 10.3390/su12062420.
[32] M. Last, N. Rabinowitz, and G. Leonard, “Correction: Predicting the maximum earthquake magnitude from seismic data in Israel and its neighboring countries,” *PLoS one*, 2016, doi: 10.1371/journal.pone.0151751.
[33] S. Ghorbani, M. Barari, and M. Hoseini, “Presenting a new method to improve the detection of micro-seismic events,” *Environmental Monitoring and Assessment*, vol. 190, 2018, doi: 10.1007/s10661-018-6837-6.
[34] N. Tahernia, “Fuzzy-logic tree approach for seismic hazard analysis,” *International Journal of Engineering and Technology*, vol. 6, no. 3, pp. 182-185, 2014, doi: 10.7763/IJET.2014.V6.692.
[35] B. R. -Leduc, C. Hulbert, N. Lubbers, K. Barros, C. J. Humphreys, and P. A. Johnson, “Machine learning predicts laboratory earthquakes,” *Geophysical Research Letters*, vol. 44, no. 18, pp. 9276-9282, 2017.
[36] T. O. Pratama, M. M. Waruwu, and R. Wijaya, “Earthquake date prediction based on the fluctuation of radon gas concentration near grundulu fault,” *Jurnal Lingkungan dan Bencana Geologi*, vol. 14, no. 2, pp. 104-113.
[37] H. Cember and T. E. Johnson, *Introduction to health physics*, 4th ed. New York: McGraw-Hill Medical, 2009.
[38] T. O. Pratama, S. Sunarno, A. B. Wijatna, and E. Haryono, “Earthquake magnitude prediction in Indonesia using a supervised method based on cloud radon data,” *International Journal of Reconfigurable and Embedded Systems (IJRES)*, vol. 13, no. 3, pp. 577-585, Nov. 2024, doi: 10.11591/ijres.v13.i3.pp577-585.
[39] I. M. Murwantara, P. Yugopuspito, and R. Hermawan, “Comparison of machine learning performance for earthquake prediction in Indonesia using 30 years historical data,” *TELKOMNIKA*, vol. 18, no. 3, pp. 1331-1342, Jun. 2020, doi: 10.12928/telkomnika.v18i3.14756.
[40] B. Sadhukhan, S. Chakraborty, S. Mukherjee, and R. K. Samanta, “Climatic and seismic data-driven deep learning model for earthquake magnitude prediction,” *Frontiers in Earth Science*, vol. 11, Feb. 2023, doi: 10.3389/feart.2023.1082832.


## BIOGRAPHIES OF AUTHORS

**Thomas Oka Pratama** earned his Doctorate in Environmental Science from Universitas Gadjah Mada in 2021 and holds previous degrees, including a Bachelor of Engineering in Engineering Physics and a Master of Engineering in the same field, both obtained from Universitas Gadjah Mada in 2014 and 2021, respectively. Presently, he is lecturer in the Engineering Physics Program at the Department of Nuclear Engineering and Engineering Physics at Universitas Gadjah Mada, Indonesia. His research interests encompass developing telemonitoring systems and earthquake prediction methods utilizing radon gas concentration and groundwater levels. He can be contacted at email: thomas.o.p@ugm.ac.id.

**Sunarno** expertise lies in Engineering, Instrumentation & Control, and Engineering Physics. He is deeply engaged in Engineering Physics, Nuclear Engineering, and Instrumentation within the broader domain of Science and Technology. His research focuses on telemetry and telecontrol instrumentation, along with the developmendeveloping systems for natural disasters. He has also organized various training programs as part of his community service initiatives. He can be contacted at email: sunarno@ugm.ac.id.

**Agus Budhie Wijatna** research interests primarily focus on utilizing hydro isotope tracing techniques within the environmental and radiation domains. Some of his notable research works include "Geographic information system for mapping radon concentrations in Yogyakarta Special Region as a reference for public health policy analysis" [2021] and "Design of an ultraviolet monitoring system for early warning of extreme weather changes" [2018]. He has also contributed to intellectual property creation, particularly in developing a method for validating topographic, geological, and hydrogeological approaches using the Hydroisotope method to identify groundwater recharge areas [2023]. Additionally, he has published significant studies such as "Determination of groundwater recharge area using hydroisotope techniques in Sei Bingei and Surrounding Areas, Langkat Regency, North Sumatra" [2020], and "Study of aquifer connectivity between boreholes and groundwater resources below" [2013]. He can be contacted at email: budiwijatna@ugm.ac.id.

**Eko Haryono** specializes in karst geomorphology, karst hydrogeology, and landscape ecology. Holding a Ph.D. in Geography from Universitas Gadjah Mada, Indonesia, his doctoral thesis focused on the Karst Development Model in Karangbolong, Gunungsewu, Blambangan, and Rengel. Over the past fifteen years, he has devoted his research efforts to exploring various facets of karst landscapes. He has held key positions as the Head of the Doctorate Program in Environmental Science at Universitas Gadjah Mada. His active involvement in professional associations, including the Karst Commission Steering Committee at the International Geographical Union and the Karst Commission of the International Association of Hydrogeologists, underscores his dedication to the field. He has spearheaded numerous research projects, such as "Characterization of the Karst Sangkulirang-Mankalihat Area, East Kalimantan," funded by USAID-Yayasan KEHATI, and "Reconstruction of Sea Level Changes in Indonesia, and Evidence from the Karstic Island of Maratua". For further communication. He can be contacted at email: e.haryono@ugm.ac.id.