Machine learning and artificial intelligence models development in rainfall-induced landslide prediction

Hastuadi Harsa[1], Anistia Malinda Hidayat[2], Adi Mulsandi[3], Bambang Suprihadi[3], Roni Kurniawan[1], Muhammad Najib Habibie[1], Thahir Daniel Hutapea[1], Yunus S. Swarinoto[1], Erwin Eka Syahputra Makmur[1], Welly Fitria[1], Rahayu Sapta Sri Sudewi[1], Alfan Sukmana Praja[1]

[1]Meteorology Division, Research and Development Center, Indonesia Meteorology Climatology and Geophysics Agency, Jakarta, Indonesia
[2]Ahmad Yani Meteorological Station, Indonesia Meteorology Climatology and Geophysics Agency, Semarang, Indonesia
[3]Department of Meteorology, School of Meteorology Climatology and Geophysics, Tangerang Selatan, Indonesia

Article Info	ABSTRACT
	In Indonesia, rainfall is one crucial triggering factor for landslides. This paper aims to build landslide event prediction models using several machine learning and artificial intelligence algorithms. The algorithms were trained with two different methods. The input of the algorithms was precipitation data obtained from the global satellite mapping of precipitation satellite observation, and the target was landslide event occurrence data obtained from the Indonesian National Board for Disaster Management. Each algorithm provided some model candidates with different parameter settings for each method. As a result, there were 52 and 72 model candidates for both methods. The best model was then chosen from each method. The result shows that the model generated by generalized linear model was the best model for the first method and deep learning for the second one. Furthermore, the best models at each method gained 0.828 and 0.836 for the area under receiver operating characteristics curve, and their log-loss were 0.156 and 0.154. The second method, which used input data transformation, provided better performance. *This is an open access article under the CC BY-SA license.*

Corresponding Author:

Hastuadi Harsa
Research and Development Center, Indonesia Meteorology Climatology and Geophysics Agency
Jl. Angkasa I No. 2, Kemayoran, Jakarta Pusat, DKI Jakarta, 10720, Indonesia
Email: hastuadi@gmail.com

1. INTRODUCTION

Classification is one of the significant tasks for machine learning (ML) and artificial intelligence (AI) [1]–[6]. The objective function of classification is to isolate regions in input feature space. The isolated regions are then labeled using guidance called target. Numerous algorithms with different approaches have been developed to determine how points in the input feature are associated with target labels. ML is designed to search this relationship using statistical schemes, while AI uses nature-inspired ones. This paper applies both ML and AI to classify some rainfall properties and whether they will potentially trigger a landslide or not. The classification is a binomial classification model derived from the best model chosen from many model candidates provided by ML and AI. the models were trained to recognize patterns in the rainfall properties data so that they were able to distinguish which combination of feature values would trigger a landslide.

Landslide is a devastating phenomenon for humans [7]–[13]. Some factors, e.g., climate conditions, human activities, geology structure, earthquakes, and topography, may trigger landslides. Rainfall is also a significant factor in landslide occurrences. Furthermore, as a maritime continent, Indonesia has abundant rainfall [14], [15]. Therefore, understanding the relationship between rainfall characteristics and the landslide

can contribute to the early warning of natural disasters in mitigation and management readiness. The integration of ML or AI with geo-information system (GIS) in landslide analysis has been studied previously [16]–[21]. Comparisons of several machine learning algorithms for landslide mapping were compared in [22]–[24]. The area under receiver operating characteristics curve (AUC) was the main metrics to assess the model performance in the literature [25]–[27]. The objectives of this paper are; assessing ML and AI performance in classifying whether some rainfall properties would trigger a landslide event or not; testing a data transformation prior to ML and AI model building and whether it can increase the performance of models.

2. METHOD
2.1. Data
The data used for the response variables in this study were reports of landslide events in the form of the date the landslide occurred from 2014 until 2018. These data were taken from the Indonesian National Board for Disaster Management (BNPB). The study area was Banjarnegara, a regency in the southwestern part of Central Java province in Indonesia. According to BNPB data, it is shown that the Banjarnegara regency is the most landslide-prone area in Central Java. There were 84 records of landslides out of 1826 dates in the data. The records were labeled "Yes" and "No" for the landslide that did and did not occur, respectively.

The availability of surface rainfall observation stations is minimal and unevenly distributed in most developing countries with highland and mountainous topography. This condition causes the forecasting of disasters and other extreme hydrometeorological phenomena to be challenging to perform [28]. These limitations make satellite rainfall estimation data are promising alternative solution [29]. Referring to this circumstance, the data used in this research were taken from the global satellite mapping of precipitation (GSMaP) as the explanatory variables. The data is presented in Spatio-temporal form. The spatial resolution of the data was 0.1°×0.1°, and the temporal resolution was in hourly format. The grid location at which we took the values from the GSMaP data was at the grid coordinate of 7°15'S and 109°45'E. Banjarnegara was in the coverage area of the selected GSMaP grid. We performed a pre-processing procedure to aggregate the GSMaP values from hourly to daily. The pre-processing procedure resulted in three explanatory variables, i.e., accumulation, intensity, and duration of rainfall at a particular date. The three derived variables were: i) accumulation: determined the total amount of rainfall on a date (mm); ii) intensity: calculated the average rainfall amount per hour for a particular date (mm/hour); and iii) duration: provided the total hours where rainfall happened on a date (hour).

The dataset included three explanatory variables, i.e., accumulation, intensity, and duration of the rainfall, and one response variable, i.e., landslide occurrence. First, the records were aligned in dates between the explanatory and the response variables. Next, we shifted all three explanatory variables backward to construct lagged values for the landslide variable. Finally, the shifting process was repeated to provide the two and three lagged values for all three main rainfall properties from the original dataset. In this stage, there were nine explanatory variables and one response variable. The nine explanatory variables were constructed from three rainfall properties one to three days before the landslide occurred. The dataset was composed of 1826 rows×10 columns in tabular form. The rows denoted the dates, and the column denoted nine explanatory variables and one response variable. Thus, in prediction terms, the system would predict whether the next day from a particular day would be potentially susceptible to a landslide by considering three rainfall properties (accumulation, intensity, and duration) from that day and one to two days before.

There were two methods examined in this study. The first method presented the dataset to the models directly. The second method applied two additional procedures: log-transformation and principal component analysis (PCA), before presenting the dataset to the models. Both datasets at each method were standardized just before being processed by the models. Eventually, in the second method's operational phase, input variables' values were scaled and subsequentially multiplied with the eigenvector provided by PCA. In addition, we added 1 to all values before calculating the log result. The reason is that the minimum value of all explanatory variables was 0. Otherwise, an infinitely small value would be returned by the logarithm function. This way resulted in the minimum values of all variables becoming 1, and all values were one greater. Although the addition of 1 changed the values, the data characteristics remained the same. The unchanged data characteristic was because the addition procedure was performed to all values without exception. A concise summary of the overall workflow is displayed in Figure 1. The data preparation phase, at the left part of the figure, consists of three main sections following input data. The model building phase, at the right part of the figure also has three sections after the data have been provided by the data preparation phase.

2.1. Model
There were four ML and one AI algorithm employed in this study. They are distributed random forest (DRF) [30], generalized linear model (GLM) [31], extreme gradient boosting (XGBoost) [32], generalized boosting machine (GBM) [33], and deep learning (DL) [34]–[36]. These algorithms were trained as binomial

classifiers as characterized by the training datasets of each method. However, the algorithms had different input variables for each method, as PCA in method-2 reduced the dataset's dimensionality. Several model candidates with the selected algorithms were built in each method. The model building phase utilized parallel computation with 15 cores and 20 gigabyte (GB) memory. The maximum of two hours of running time was determined for both methods. Each model candidate had different parameterization settings from another model candidate. After building the model candidate, two stacked ensembles (SE) model schemes were also carried out. The first SE was composed of every model candidate within the same algorithm. The second one was composed of all model candidates regardless of their algorithm.

The model candidates, including SE models, were sorted using their area under receiver operating characteristics curve (AUC) performance. In addition, log-loss performance metrics were also provided. The performances (AUC and log-loss) of all models were obtained from cross-validated test data. We provided 20 folds for model performance calculation with a random sampling scheme.
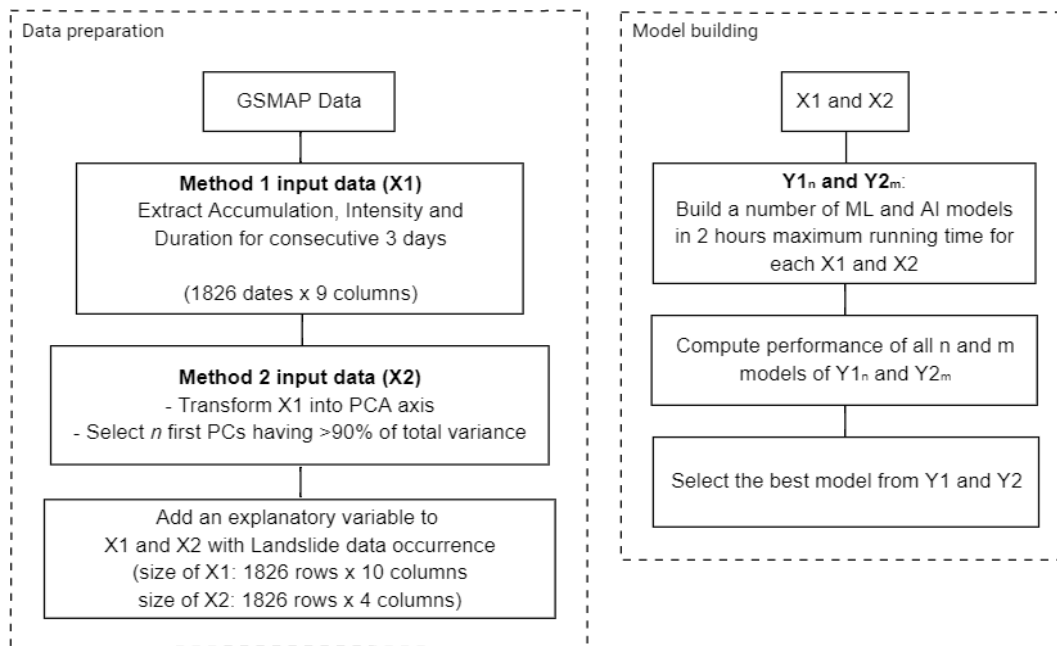
Figure 1. Workflow diagram

3. RESULTS AND DISCUSSION
3.1. Input data

The initial distributions of data are displayed in Figure 2(a) and their log-transformed by Figure 2(b). The landslide occurrence distributions are denoted by red area and labeled as 'Yes' while those without landslide are denoted by blue area and labeled as 'No' in both figures. It can be observed that after being transformed, the densities of values in each variable were more clearly visible. Thus, the characteristics of the data were also more exaggerated. For method-2, the second data format was subsequentially fed into PCA. Each datum in the input data has nine variables. Therefore, PCA re-mapped these data into its axis, called the principal component axis (PC), and there were also nine PCs in this case. These PCs were orthogonal to each other and ordered from the highest variance for a naming convention (PC1 and PC2). Since each PC was orthogonal, selecting only a few of them was often adequate to retain the variance in the original data.

The first three PCs had already gained more than 90% of the cumulative data variance explained, as displayed in Figure 3(a). Therefore, we only used three variables for method-2 instead of nine as in method-1. One thing to note is that PCA's axes are orthogonal. Hence, the multicollinearity among each variable was suppressed, as can be observed from the location of data in the PCA coordinate system in Figure 3(b).

3.2. Model performance

There were 57 models produced for method-1 and 72 for method-2 in two hours maximum running time for each method. The number of candidate models provided by each algorithm varied. Method-2 required less time for producing models since it had fewer explanatory variables than method-1. The performance of

models is represented by their location in log-loss and AUC coordinates, as displayed in Figure 4. AUC is a metric to obtain a binary (two-class) classifier model's performance. It ranges between 0 and 1, where 0 means a classifier model would predict a positive class as a negative class and vice versa.
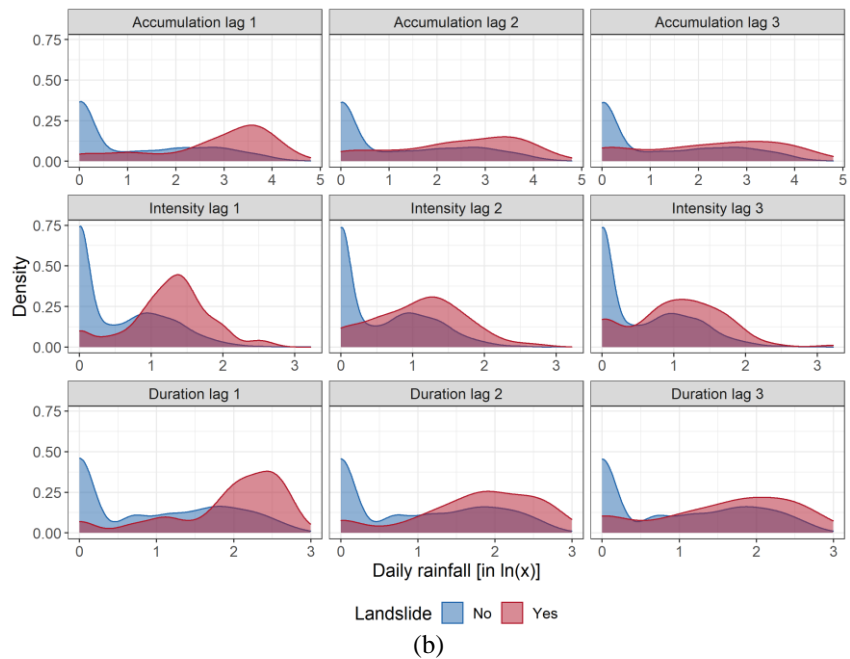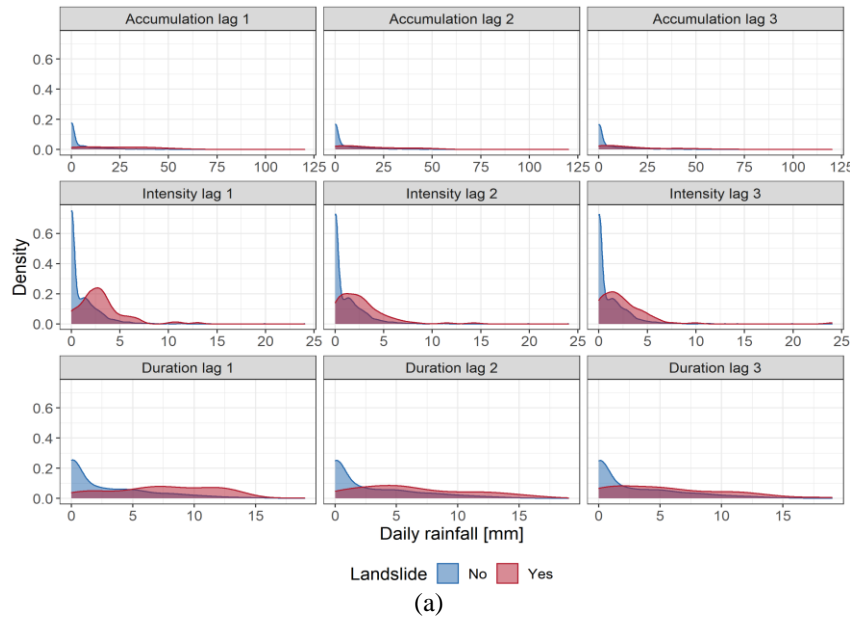

(a)

(b)

Figure 2. The input data density: (a) before and (b) after transformed

On the other hand, a model having an AUC value of 1 means that the model perfectly predicts the actual class. If an AUC value of a model is close to 0.5, then the model behaves as if it is a random guess model, as in a coin toss where the probability of head and tail is equal, i.e., 0.5. Log-loss is another performance metric for a binary classifier. It summarizes the mean difference between the logarithm of probability a datum is classified into a particular class with its actual class. The lower a log-loss value is, the better a model is.
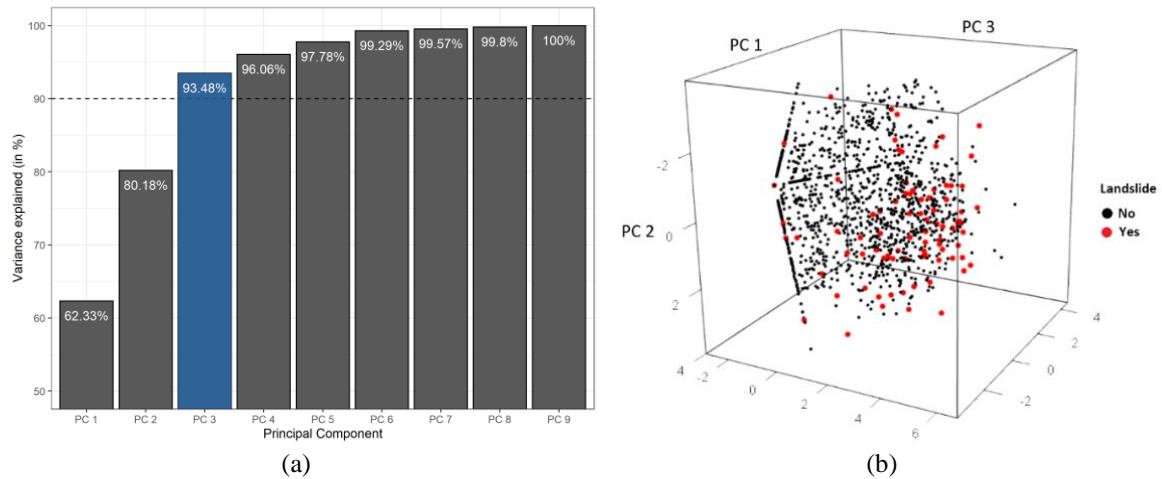
Figure 3. Data information retrieval using PCA denoted in (a) variance explained by each principal component axes and (b) the location of input data points in selected first-three PC axes: PC1, PC2, and PC3
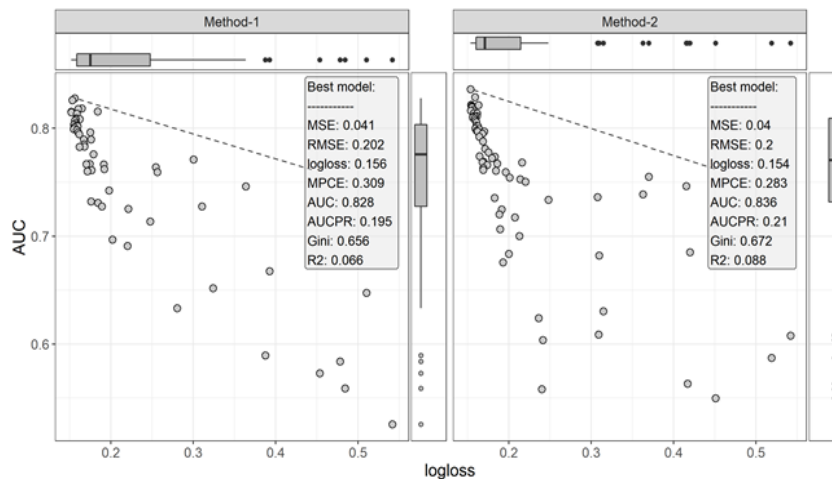
Figure 4. The performance of candidate models at each method. The performances are displayed as location at the AUC and log-loss axes, where a better model has higher AUC and lower log-loss

The best model is located in the top-left position. The top-left position has a combination of higher AUC and lower log-loss. Other performance metrics for the best models at each method are also displayed in the legend boxes. The legend boxes provide the information on mean squared error (MSE), root mean squared error (RMSE), mean per class error (MPCE), area under precision-recall (AUPCR), Gini index, and R-squared (R2). The performances of models at method-2 are quantitatively better than that of method-1. It can be observed from the fact that method-2 has higher values than method-1 for the greater-the-better metrics. The contribution of each explanatory variable is presented in Table 1. Method-1 has nine variables, while Method-2 has three. The critical values are presented in percentages. For example, the best model in method-1 considered that variable rainfall duration on a particular date had the most significant contribution to landslide occurrence the next day. In method-2, variable PC2 was regarded as the most significant variable by the best model.

Additionally, models in method-2 also have lower values for the lesser-the-better metrics. The log-loss values of models in method-2 are more concentrated at the lower part of the metric space. Therefore, it can be inferred that log-transformation and PCA were able to increase models' performance. Figure 5 shows the ROC curve of the best model in each method. The AUC values are also displayed at the bottom-right of the figure. The ROC of method-2 is slightly wider than that of method-1. Method-2 is more efficient than method-1 because it used fewer explanatory variables than method-1, but its AUC outperformed the AUC of

method-1. All metrics are calculated from the mean of 20 cross-validation datasets. Research related to the model development may also be performed in the future. For example, since this work only assessed the influence of rainfall on the landslide event, it is also essential to incorporate other variables such as soil type, land slope, and human activity.
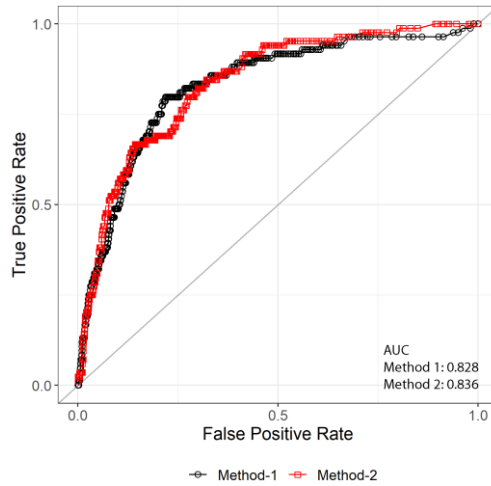

Figure 5. The ROC of the best models in both methods

Table 1. Variable importance, denoted in percent

Method-1		Method-2	
Variable	Importance percentage	Variable	Importance percentage
Duration lag 1	22.57	PC3	50.24
Accumulation lag 1	18.73	PC1	32.36
Duration lag 2	14.05	PC2	17.4
Accumulation lag 3	12.1		
Intensity lag 1	9.64		
Accumulation lag 2	6.17		
Duration lag 3	5.8		
Intensity lag 3	5.58		
Intensity lag 2	5.35		

4. CONCLUSION

Our study finds that proper pre-processing procedures for the input data increased ML and AI performance. We differentiate two ML and AI model development methods in the model development phase. The first one did not undergo pre-processing, while the second one did. Furthermore, the second method outperformed the first method, yet it only required fewer explanatory variables than the first method. We applied log-transformation and PCA, as pre-processing procedures, to our input data for the second method. The GLM provided the best model for the first method and DL for the second one. The models were a prediction of landslide occurrence one day ahead, given satellite-derived rainfall properties. Our work may contribute to the development of a rainfall-induced landslide warning system for disaster mitigation and management.

ACKNOWLEDGEMENTS

REFERENCES

[1] L. Hussain *et al.*, "Machine learning classification of texture features of MRI breast tumor and peri-tumor of combined pre- and early treatment predicts pathologic complete response," *BioMedical Engineering Online*, vol. 20, no. 1, 2021, doi: 10.1186/s12938-021-00899-z.

[2] Y. Gulzar, Y. Hamid, A. B. Soomro, A. A. Alwan, and L. Journaux, "A convolution neural network-based seed classification system," *Symmetry*, vol. 12, no. 12, pp. 1–18, 2020, doi: 10.3390/sym12122018.

[3] R. C. Maron *et al.*, "Artificial intelligence and its effect on dermatologists' accuracy in dermoscopic melanoma image classification:

web-based survey study," *Journal of Medical Internet Research*, vol. 22, no. 9, 2020, doi: 10.2196/18091.

[4] A. Zhang, X. Yang, L. Jia, J. Ai, and Z. Dong, "SAR image classification using adaptive neighborhood-based convolutional neural network," *European Journal of Remote Sensing*, vol. 52, no. 1, pp. 178–193, 2019, doi: 10.1080/22797254.2019.1579616.

[5] S. S. Gornale, P. U. Patravali, and P. S. Hiremath, "Automatic detection and classification of knee osteoarthritis using Hu's invariant moments," *Frontiers in Robotics and AI*, vol. 7, 2020, doi: 10.3389/frobt.2020.591827.

[6] G. Chen *et al.*, "Food/non-food classification of real-life egocentric images in low- and middle-income countries based on image tagging features," *Frontiers in Artificial Intelligence*, vol. 4, 2021, doi: 10.3389/frai.2021.644712.

[7] E. Karantanellis, V. Marinos, E. Vassilakis, and D. Hölbling, "Evaluation of machine learning algorithms for object-based mapping of landslide zones using UAV data," *Geosciences (Switzerland)*, vol. 11, no. 8, 2021, doi: 10.3390/geosciences11080305.

[8] S. Lainas, N. Depountis, and N. Sabatakakis, "Preliminary forecasting of rainfall-induced shallow landslides in the wildfire burned areas of western greece," *Land*, vol. 10, no. 8, 2021, doi: 10.3390/land10080877.

[9] O. Igwe and C. O. Una, "Landslide impacts and management in Nanka area, Southeast Nigeria," *Geoenvironmental Disasters*, vol. 6, no. 1, 2019, doi: 10.1186/s40677-019-0122-z.

[10] Z. Xia, M. Motagh, T. Li, and S. Roessner, "The June 2020 Aniangzhai landslide in Sichuan Province, Southwest China: slope instability analysis from radar and optical satellite remote sensing data," *Landslides*, vol. 19, no. 2, pp. 313–329, 2022, doi: 10.1007/s10346-021-01777-4.

[11] H. Jia, H. Zhang, L. Liu, and G. Liu, "Landslide deformation monitoring by adaptive distributed scatterer interferometric synthetic aperture radar," *Remote Sensing*, vol. 11, no. 19, 2019, doi: 10.3390/rs11192273.

[12] A. Dikshit, R. Sarkar, B. Pradhan, S. Segoni, and A. M. Alamri, "Rainfall induced landslide studies in indian himalayan region: a critical review," *Applied Sciences (Switzerland)*, vol. 10, no. 7, 2020, doi: 10.3390/app10072466.

[13] N. Sultana, "Analysis of landslide-induced fatalities and injuries in Bangladesh: 2000-2018," *Cogent Social Sciences*, vol. 6, no. 1, 2020, doi: 10.1080/23311886.2020.1737402.

[14] H. Irwandi, M. S. Rosid, and T. Mart, "The effects of ENSO, climate change and human activities on the water level of Lake Toba, Indonesia: a critical literature review," *Geoscience Letters*, vol. 8, no. 1, 2021, doi: 10.1186/s40562-021-00191-x.

[15] D. Muriyatmoko and S. M. Phuspa, "Analysis of rainy days and rainfall to landslide occurrence using logistic regression in Ponorogo East Java," *Geosfera Indonesia*, vol. 3, no. 2, p. 79, 2018, doi: 10.19184/geosi.v3i2.8230.

[16] K. Mandal, S. Saha, and S. Mandal, "Applying deep learning and benchmark machine learning algorithms for landslide susceptibility modelling in Rorachu river basin of Sikkim Himalaya, India," *Geoscience Frontiers*, vol. 12, no. 5, 2021, doi: 10.1016/j.gsf.2021.101203.

[17] M. Di Napoli *et al.*, "Rainfall-induced shallow landslide detachment, transit and runout susceptibility mapping by integrating machine learning techniques and gis-based approaches," *Water (Switzerland)*, vol. 13, no. 4, 2021, doi: 10.3390/w13040488.

[18] T. Qi, Y. Zhao, X. Meng, G. Chen, and T. Dijkstra, "Ai-based susceptibility analysis of shallow landslides induced by heavy rainfall in Tianshui, China," *Remote Sensing*, vol. 13, no. 9, 2021, doi: 10.3390/rs13091819.

[19] M. Kuradusenge, S. Kumaran, and M. Zennaro, "Rainfall-induced landslide prediction using machine learning models: the case of ngororero district, rwanda," *International Journal of Environmental Research and Public Health*, vol. 17, no. 11, pp. 1–20, 2020, doi: 10.3390/ijerph17114147.

[20] K. Nam and F. Wang, "An extreme rainfall-induced landslide susceptibility assessment using autoencoder combined with random forest in Shimane Prefecture, Japan," *Geoenvironmental Disasters*, vol. 7, no. 1, 2020, doi: 10.1186/s40677-020-0143-7.

[21] T. G. Nachappa, O. Ghorbanzadeh, K. Gholamnia, and T. Blaschke, "Multi-hazard exposure mapping using machine learning for the state of Salzburg, Austria," *Remote Sensing*, vol. 12, no. 17, 2020, doi: 10.3390/RS12172757.

[22] S. Zhao and Z. Zhao, "A comparative study of landslide susceptibility mapping using SVM and PSO-SVM models based on grid and slope units," *Mathematical Problems in Engineering*, vol. 2021, 2021, doi: 10.1155/2021/8854606.

[23] H. Wang, L. Zhang, K. Yin, H. Luo, and J. Li, "Landslide identification using machine learning," *Geoscience Frontiers*, vol. 12, no. 1, pp. 351–364, 2021, doi: 10.1016/j.gsf.2020.02.012.

[24] Z. Liu *et al.*, "Modelling of shallow landslides with machine learning algorithms," *Geoscience Frontiers*, vol. 12, no. 1, pp. 385–393, 2021, doi: 10.1016/j.gsf.2020.04.014.

[25] N. Tengtrairat, W. L. Woo, P. Parathai, C. Aryupong, P. Jitsangiam, and D. Rinchumphu, "Automated landslide-risk prediction using web GIS and machine learning models," *Sensors (Basel, Switzerland)*, vol. 21, no. 13, 2021, doi: 10.3390/s21134620.

[26] A. M. S. Pradhan and Y. T. Kim, "Rainfall-induced shallow landslide susceptibility mapping at two adjacent catchments using advanced machine learning algorithms," *ISPRS International Journal of Geo-Information*, vol. 9, no. 10, 2020, doi: 10.3390/ijgi9100569.

[27] V. H. Nhu *et al.*, "Landslide susceptibility mapping using machine learning algorithms and remote sensing data in a tropical environment," *International Journal of Environmental Research and Public Health*, vol. 17, no. 14, pp. 1–23, 2020, doi: 10.3390/ijerph17144933.

[28] M. A. Gebremedhin, M. W. Lubczynski, B. H. P. Maathuis, and D. Teka, "Novel approach to integrate daily satellite rainfall with in-situ rainfall, Upper Tekeze Basin, Ethiopia," *Atmospheric Research*, vol. 248, 2021, doi: 10.1016/j.atmosres.2020.105135.

[29] M. Zambrano-Bigiarini, A. Nauditt, C. Birkel, K. Verbist, and L. Ribbe, "Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile," *Hydrology and Earth System Sciences*, vol. 21, no. 2, pp. 1295–1320, 2017, doi: 10.5194/hess-21-1295-2017.

[30] G. Karakas, R. Can, S. Kocaman, H. A. Nefeslioglu, and C. Gokceoglu, "Landslide susceptibility mapping with random forest model for Ordu, Turkey," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 43, no. B3, pp. 1229–1236, 2020, doi: 10.5194/isprs-archives-XLIII-B3-2020-1229-2020.

[31] B. Kalantar, N. Ueda, V. Saeidi, K. Ahmadi, A. A. Halin, and F. Shabani, "Landslide susceptibility mapping: Machine and ensemble learning based on remote sensing big data," *Remote Sensing*, vol. 12, no. 11, 2020, doi: 10.3390/rs12111737.

[32] S. Albahli, "A deep ensemble learning method for effort-aware just-in-time defect prediction," *Future Internet*, vol. 11, no. 12, 2019, doi: 10.3390/FI11120246.

[33] J. A. Anochi, V. A. de Almeida, and H. F. de Campos Velho, "Machine learning for climate precipitation prediction modeling over South America," *Remote Sensing*, vol. 13, no. 13, 2021, doi: 10.3390/rs13132468.

[34] M. J. Gacto, J. M. Soto-Hidalgo, J. Alcala-Fdez, and R. Alcala, "Experimental study on 164 algorithms available in software tools for solving standard non-linear regression problems," *IEEE Access*, vol. 7, pp. 108916–108939, 2019, doi: 10.1109/ACCESS.2019.2933261.

[35] C. Mishra and D. L. Gupta, "Deep machine learning and neural networks: an overview," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 6, no. 2, p. 66, 2017, doi: 10.11591/ijai.v6.i2.pp66-73.

[36] M. S. Srividya, M. R. Anala, and C. Tayal, "Deep learning techniques for physical abuse detection," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 4, pp. 971–981, 2021, doi: 10.11591/IJAI.V10.I4.PP971-981.

BIOGRAPHIES OF AUTHORS

Hastuadi Harsa received his B.Sc. in Informatics Technology from Sepuluh Nopember Institut of Technology, Surabaya - Indonesia, in 2005 and his M.Sc. in Computer Science from Bogor Agricultural University, Bogor - Indonesia, in 2017. He has worked as a researcher for the Indonesia Meteorology Climatology and Geophysics Agency in Meteorological Division since 2006. His primary expertise are artificial intelligence and machine learning. He also performs research in image processing, signal processing, climate modeling and numerical weather prediction. He can be contacted at email: hastuadi@gmail.com.

Anistia Malinda Hidayat received her B.Sc. in Meteorology from the School of Meteorology, Climatology, and Geophysics, Jakarta - Indonesia, in 2020. She is a weather forecaster of Ahmad Yani Meteorological Station, Semarang - Indonesia. Her research interests are artificial intelligence, remote sensing, and weather forecasting. She can be contacted at email: anistia.malinda.hidayat@bmkg.go.id.

Adi Mulsandi received his B.Sc. in Physics from Cendrawasih University, Jayapura - Indonesia, in 2009 and his M.Sc. in Atmospheric Science from Bandung Institute of Technology (ITB), Bandung - Indonesia, in 2011. He is a PhD candidate in Climate Science at Bogor Agricultural University (IPB), Bogor - Indonesia. He has been an Assistant Professor in Department of Meteorology, School of Meteorology Climatology and Geophysics since 2011. His research interests are numerical weather prediction and remote sensing. He can be contacted at email: adi.mulsandi@stmkg.ac.id.

Bambang Suprihadi received his B.Sc. from the Faculty of Law, Indonesia University, Depok – Indonesia in 1989 and his M.Sc. in Environmental Education from State University of Jakarta, Jakarta – Indonesia in 1994. He is an Associate Professor of the School of Meteorology Climatology and Geophysics. He can be contacted at email: b_supri_hadi@yahoo.com.

Roni Kurniawan received his B.Sc. from Surabaya Hangtuah University and M.Sc. from the Indonesia University, Depok. He has been working for the Indonesia Meteorology Climatology and Geophysics Agency (BMKG) since 2006 as a Researcher at Meteorological Division of Research and Development Center BMKG. His research interest is ocean modeling. He can be contacted at email: ronikurniawan.313@gmail.com.

Muhammad Najib Habibie received his B.Sc. in Marine Science from the Diponegoro University, Semarang – Indonesia, in 2008 and M.Sc. in Geography from the Gadjah Mada University, Yogyakarta - Indonesia in 2017. He has been a junior researcher for the Indonesia Meteorology Climatology and Geophysics Agency (BMKG) in Meteorological Division since 2008 and was a lecturer at the School of Meteorology Climatology and Geophysics from 2011 until 2013. His primary expertise is oceanography. He is also interested in marine climate and extreme events. He is also on the study of intercomparison between the sea surface current observation using acoustic doppler current profiler and high-frequency radar. He can be contacted at email: najib.habibie@gmail.com.

Thahir Daniel Hutapea 🆔 ⑧ SC ◖ received his B.Sc. in Engineering from Pembangunan Panca Budi University, Medan - Indonesia, in 2009 and M.Sc. in Physics from Sumatera Utara University, Medan - Indonesia, in 2012. Currently, he works in Reasearch and Development Center of Meteorological, Climatological, and Geophysical Agency, Indonesia. He can be contacted at email: hutapea.df@gmail.com.

Yunus Subagyo Swarinoto 🆔 ⑧ SC ◖ was born in Blitar City, East Java Province, Indonesia. He was gratuated as Bachelor Degree of Atmospheric Physics from Univesity of Indonesia, Depok City, West Java Province; as Master Degree of Physical Geography from University of Indonesia, Depok City, West Java Province; and as PhD Degree of Applied Climatology from IPB University, Bogor City, West Java Province. He has a long experience working at Meteorological Climatological and Geophysics Agency (BMKG) since 1978 until now in the field of Meteorology, Climatology, and Databases. The last position of him was as a Senior Research Scientist at Research and Development Center in Head Office of BMKG Jakarta City since 2017. He can be contacted at email: yunusbmkg5757@gmail.com.

Erwin Eka Syahputra Makmur 🆔 ⑧ SC ◖ received his B.Sc. in Physics from Indonesia University, Depok - Indonesia, in 2000, M.Sc. in Atmospheric Science from Bandung Institute of Technology, Bandung - Indonesia, in 2008 and PhD in Applied Climatology from Bogor Agricultural University, Bogor – Indonesia, in 2014. His research interests are in atmospheric science, hydrology and oceanography. He can be contacted at email: erwinmakmur6@gmail.com.

Welly Fitria 🆔 ⑧ SC ◖ holds her Bachelor's degree in Statistics from Padjadjaran University, Jatinangor–Indonesia, in 2004 and a Master's degree in earth science from Bandung Institut of Technology, Bandung - Indonesia, in 2017. Currently, she works as a staff in Reasearch and Development Center of Meteorological, Climatological, and Geophysical Agency, Indonesia. Her main areas of research interest are statistics and oceanography. She can be contacted at email: wellyfitria234@gmail.com.

Rahayu Sapta Sri Sudewi 🆔 ⑧ SC ◖ holds her Bachelor's degree in Marine Science from The Faculty of Fisheries and Marine Science of Diponegoro University, Semarang - Indonesia, in 2005. Currently, she works as a staff of the Meteorological, Climatological, and Geophysical Agency (BMKG), Indonesia. Her main area of research interest is meteorology. She can be contacted at email: rahayu.sapta@gmail.com.

Alfan Sukmana Praja 🆔 ⑧ SC ◖ received his B.Sc. and M.Sc. (Mathematics) from Indonesia University of Education (UPI) and Bandung Institute of Technology (ITB) in 2004 and 2012, respectively, besides several professional certificates and skills. He has worked at Indonesia Meteorology Climatology and Geophysics Agency (BMKG) since 2012. He is also a Researcher at Meteorological Division, Research and Development Center BMKG from 2015 until now. His research interests are climate, meteorology, machine learning, statistics, mathematics, atmospheric science and remote sensing. He can be contacted at email: alfan.sukmana@gmail.com.