Indonesian Red Crescent Humanitarian Journal
published by Bulan Sabit Merah Indonesia

p-ISSN: 2829-8497, e-ISSN: 2829-8500
Vol.5, No.1, 2026, hal. 56-67
https://doi.org/10.56744/irchum.v5i1.127

When Should Disaster Managers Act? Flood Preparedness Window in Java, Indonesia

Yihuan Lin*[1], Juan Zhang[2]
School of Business and Economics[1], Institute for Mathematical Research (INSPEM)[2],
Universiti Putra Malaysia
43400 Seri Kembangan, Selangor, Malaysia
***Correspondence author:** linyihuan@outlook.com

Received: 02/07/2026	Revised: 18/08/2026	Accepted: 02/09/2026	Published Online: 20/09/2026

Abstract. This study analysed fifteen years of monthly rainfall and flood records from Java's six provinces to determine when disaster managers should act. We measured how long periods of elevated rainfall risk tend to persist and used machine learning to forecast how many flood events to expect each month. The results show that rainfall volatility in Java is persistent: a volatility shock loses half its effect over approximately 7.4 months. In practical terms, the October–November wet season onset is an important window for disaster preparedness: agencies that assess rainfall conditions during these two months can use that assessment to gauge whether the coming season may warrant high-alert mobilisation before the worst floods arrive, though this study does not test that window against other possible timings. Monthly flood forecasting performed reasonably well across all models, but adding the volatility estimate to the prediction models did not significantly improve forecast accuracy. This happens because the GARCH-derived variance is structurally collinear with the recent rainfall lags already used as predictors, so it adds little once recent rainfall is already represented in the model. This clarifies the roles of the two tools. Volatility monitoring answers when to prepare, while machine learning forecasts answer how many events to expect. We recommend Indonesian disaster management agencies adopt a two-step framework—assess rainfall volatility in October–November to set mobilisation levels, then update monthly flood forecasts throughout the wet season to guide operational response.

Keywords: flood disaster forecasting; rainfall volatility; machine learning; disaster preparedness; Java, Indonesia

Introduction

Floods are the most frequent and destructive natural hazard in Java. Despite covering less than seven percent of Indonesia's land area, the island accounted for nearly half of all flood-related disasters recorded nationally in 2021 (Farid et al., 2023). In that year alone, Java's floods killed 109 people, displaced over 127,000 from their homes, and affected approximately 2.3 million residents (Hamzah et al., 2023). Flood frequency has increased steadily over the past five decades, driven by urbanisation and population growth (Sejati et al., 2024). Climate projections indicate that rainfall extremes will intensify further through the coming decades. In practice, early mobilisation ahead of the wet season means pre-positioning relief logistics and personnel, preparing evacuation shelters, and activating community-based disaster risk management (CBDRM) networks before peak flooding arrives. These actions work best when agencies already know, in advance, when regional risk has become elevated. For disaster management agencies responsible for allocating emergency resources and issuing early warnings, the central question is not whether floods will occur but when risk will be high enough to justify action.

Researchers have pursued reliable flood risk forecasts through two main approaches. The first draws on classical time series models. Seasonal models capture the predictable rise and fall of monsoon rainfall reasonably well (Mishra et al., 2025). Researchers working in similar tropical contexts have shown that year-to-year rainfall risk tends to cluster: high-risk periods follow other high-risk periods, and this pattern can be estimated using statistical models of time-varying variance (Pandey et al., 2019; Uddin et al., 2020). The second approach draws on machine learning. Gradient boosting and deep learning networks can capture nonlinear relationships among rainfall, land use, and socioeconomic variables, and have been applied to flood risk assessment at regional scales (Chen et al., 2021; Nti et al., 2021), including near-term water-level prediction in Jakarta (Azi et al., 2024).

Each approach has a gap. Variance models characterise periods when risk is elevated but do not predict how many flood events will occur. Machine learning models can already include seasonal terms and lagged rainfall directly as features, so the open question is narrower than whether such models notice seasonality. It is whether a separate measure of rainfall volatility carries information beyond what rainfall, seasonality, ENSO, and lagged flood counts already provide. Rainfall on the island is highly seasonal: the wet season delivers over 375 mm per month on average while the dry season produces as little as 50 mm. Large-scale climate forces amplify year-to-year differences. La Niña phases increase annual rainfall across Central Java by up to 1,700 mm above normal (Firmansyah et al., 2022), and the influence of El Niño-Southern Oscillation on rainfall variability extends to coastal regions of East Java (Hasibuan et al., 2024). Whether a volatility feature adds anything once these seasonal and climate signals are already in the model is the question this study tests directly.

A growing body of research suggests that combining variance modelling with machine learning may help. A model can already account for seasonal risk through month indicators, harmonic terms, or lagged rainfall, without using GARCH at all. The narrower possibility this literature points to is that time-varying rainfall variance carries additional information that these conventional seasonal predictors do not capture. GARCH family models, which estimate time-varying variance in sequential data, have been used as feature inputs in machine learning pipelines in energy load forecasting (Zeng et al., 2022), financial volatility prediction (Amirshahi & Lahmiri, 2023; Araya et al., 2024), and drought forecasting in hydrology (Ghasempour et al., 2023). This approach has not yet been applied to monthly flood disaster risk forecasting in Indonesia.

This study applies that approach to monthly flood event forecasting across the Java region from 2010 to 2024. The study estimates monthly rainfall variance using three GARCH variants and combines each with three machine learning algorithms, yielding 12 model configurations evaluated under walk-forward validation. SHAP feature importance analysis is applied to examine whether and how the GARCH-derived variance feature contributes to model predictions. The goal is to determine whether modelling rainfall variance improves flood prediction and to clarify the distinct roles that volatility modelling and machine learning play in disaster risk assessment.

Methods

Study Area and Data

This study focuses on the Java region of Indonesia, comprising six provinces: DKI Jakarta, Jawa Barat, Banten, Jawa Tengah, Daerah Istimewa Yogyakarta, and Jawa Timur. Two monthly data series were constructed for January 2010 to December 2024 (n = 180 months).

Monthly rainfall estimates were obtained from the NASA Prediction of Worldwide Energy Resources (NASA POWER) system using the PRECTOTCORR bias-corrected precipitation parameter, averaged across the six provinces. We used an unweighted arithmetic mean and did not adjust for land area or population, so each province contributes equally to the regional average. Monthly flood event counts were compiled from the Data Informasi Bencana Indonesia (DIBI) database maintained by the Indonesian National Disaster Management Agency (BNPB), aggregated from 7,344 individual event records. Annual ENSO Niño 3.4 sea surface temperature anomalies were obtained from the NOAA Physical Sciences Laboratory and broadcast to monthly resolution. We repeated the single annual value for a given year across all twelve months of that year, without interpolation between years. Table 1 summarises all variables.

Table 1. Data Variables, Sources, and Coverage

Variable	Source	Period	n
Monthly rainfall (mm)	NASA POWER (PRECTOTCORR)	Jan 2010–Dec 2024	180 months
Monthly flood events	BNPB DIBI (individual records)	Jan 2010–Dec 2024	180 months
ENSO Niño 3.4 (°C)	NOAA Physical Sciences Laboratory	2010–2024 (annual)	15 years

Source: Authors' summary based on collected data.

Pre-Estimation Diagnostics

Four diagnostic tests were applied to the monthly rainfall series before model estimation. Stationarity was assessed using the Augmented Dickey-Fuller test (Dickey & Fuller, 1979) and the Kwiatkowski-Phillips-Schmidt-Shin test (Kwiatkowski et al., 1992). Seasonality was tested using the Kruskal-Wallis non-parametric test across twelve monthly groups. Conditional heteroskedasticity was assessed using the ARCH Lagrange Multiplier test (Engle, 1982) applied to both the raw rainfall series and the SARIMA residuals. Table 2 summarises results of these four diagnostic tests.

Table 2. Pre-Estimation Diagnostic Test Results

Test	Statistic	p-value	Decision
ADF (H_0: unit root)	−3.331	0.014	Stationary
KPSS (H_0: stationary)	0.029	≥0.100	Stationary
Kruskal-Wallis (H_0: equal monthly means)	123.803	<0.001	Strong seasonality
ARCH-LM on raw series (12 lags)	93.370	<0.001	ARCH effects present
ARCH-LM on SARIMA residuals (12 lags)	24.707	0.016	ARCH effects persist

Source: Authors' data analysis.

SARIMA Mean Equation

A SARIMA (1,0,0) (1,0,0) (12) model was fitted to the monthly rainfall series to remove the predictable seasonal mean before estimating the variance process (Box & Jenkins, 1976). The specification was selected by AIC-guided grid search (AIC = 1,961.62; BIC = 1,970.97). The residuals from this mean equation served as input to the GARCH variance estimation.

GARCH Family Models

Three GARCH variants were fitted to the SARIMA residuals. Standard GARCH (1,1) models conditional variance as a function of the squared previous residual and previous variance (Engle & Bollerslev, 1986). EGARCH (1,1) works in log-variance space and allows asymmetric responses to positive and negative shocks (Nelson, 1991). GJR-GARCH (1,1) introduces a threshold asymmetry term through an indicator variable (Glosten et al., 1993). GARCH-X, which adds the annual ENSO index as an exogenous variance regressor, is not analysed further here: its parameter estimates were identical to standard GARCH(1,1), so it adds nothing beyond the three variants above. Models were compared by AIC and BIC using a single fit on the full 180-month series. This full-sample fit describes the overall persistence of rainfall volatility across the study period. It is a description, not a forecast, so we use it only for that purpose. For the machine learning stage, we instead re-estimated SARIMA and each GARCH variant inside every walk-forward window, using only the 36 months before the target month, and took a one-step-ahead variance forecast as the predictive feature for that month. This keeps the variance feature free of information from later months.

Machine Learning Models and Hybrid Construction

Three machine learning algorithms were trained to predict monthly flood event counts: Random Forest (Breiman, 2001), XGBoost (Chen & Guestrin, 2016), and Support Vector Regression (Drucker et al., 1996). Long Short-Term Memory networks (Hochreiter & Schmidhuber, 1997) were not included, since sequence models typically need longer training histories than the 36-month walk-forward window provides. Baseline predictors included three monthly rainfall lags, the ENSO index from the prior year, three flood count lags, and sine-cosine seasonality components. We lagged rainfall and ENSO so that every prediction uses only information available before the target month. This makes the exercise a genuine one-month-ahead forecast rather than a same-month estimate. Hybrid models added the walk-forward GARCH conditional variance as a tenth feature. Hyperparameters were tuned using RandomizedSearchCV with 30 iterations and a 5-fold TimeSeriesSplit on the first 60% of observations.

Walk-Forward Validation and Evaluation

All 12 models were evaluated by walk-forward validation with a fixed window of w = 36 months, yielding 143 out-of-sample monthly predictions per model, from month 37 through month 179 of the 180-month series. Model performance was assessed by Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The Diebold-Mariano test (Diebold & Mariano, 1995) compared the best hybrid and best standalone configurations on equal predictive accuracy. We also ran a second, more direct Diebold-Mariano test on Random Forest alone against its GJR-GARCH hybrid counterpart, the same algorithm with and without the GARCH feature, to isolate that feature's contribution from any difference between algorithms. SHAP (SHapley Additive exPlanations) values were computed using TreeExplainer for both the standalone XGBoost model and the GJR-GARCH plus Random Forest hybrid, to measure how much each feature contributed to predictions.

Results and Discussion

Pre-Estimation Diagnostics and Rainfall Characteristics

Java's monthly rainfall is stationary by both ADF and KPSS tests (Table 2), requiring no differencing before GARCH estimation. Seasonality is pronounced: February averages 375.2 mm against August's 50.3 mm, a wet-to-dry ratio of 7.5. This amplitude is considerably more extreme than the South Asian monsoon systems studied by Pandey et al. (2019) and Uddin et al. (2020), where GARCH modelling identified persistent volatility clustering. Critically, volatility clustering persists after seasonal adjustment: the ARCH-LM test on SARIMA residuals remains significant ($p=0.016$), confirming that high-variance months cluster together independently of the seasonal pattern. All conditions for GARCH family estimation are satisfied.

Figure 1 illustrates these patterns across the 15-year study period. Monthly rainfall peaks and flood event spikes move together closely. This is an observed association, not a causal test, but it is consistent with seasonal rainfall driving most of the variation in flood frequency, while year-to-year variation in peak magnitude motivates the volatility modelling approach.

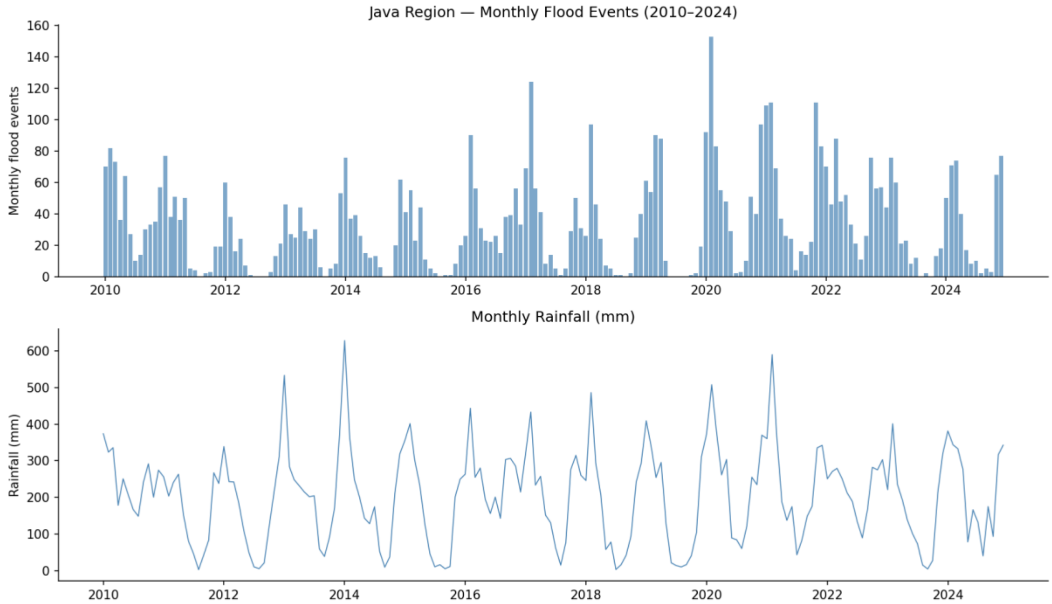

Figure 1. Monthly flood event counts (top) and regional rainfall in millimetres (bottom) across the Java region, January 2010 to December 2024. Both series show a pronounced wet season from October to April. The close co-movement between rainfall peaks and flood event spikes motivates the combined modelling approach used in this study. Source: authors' analysis of BNPB DIBI flood records and NASA POWER rainfall data.

GARCH Model Results

EGARCH (1,1) provides the best fit by both AIC (1,873.6) and BIC (1,883.2), followed by GARCH (1,1) and GJR-GARCH (1,1) (Table 3). The persistence parameter in EGARCH (1,1) is estimated at beta=0.911 and is highly significant. Using the standard half-life formula, ln(0.5) divided by ln(beta), this implies a volatility half-life of approximately 7.4 months. The ARCH shock coefficient is not significant (p=0.238), providing no evidence that variance responds to contemporaneous anomalies; the significant persistence parameter is consistent with variance building through accumulated past history. This pattern aligns with a climate system governed by large-scale oceanic forcing, where risk builds slowly and dissipates slowly across the monsoon cycle.

One additional result carries substantive content. The GJR-GARCH asymmetry parameter was not significant (gamma=0.025, p=0.939), providing no evidence of asymmetric variance responses to positive and negative rainfall anomalies. Unlike financial markets, where negative news raises volatility more than positive news of equal size (consistent with Nelson, 1991), rainfall is a physical process with no obvious asymmetric information mechanism. GARCH-X is excluded from the comparison, because it adds nothing beyond standard GARCH (1,1). This fits the wider literature: ENSO operates at the annual scale (Firmansyah et al., 2022; Hasibuan et al., 2024), so it belongs in the ML prediction stage as a feature, not in the monthly variance equation.

Table 3. GARCH Family Model Comparison

Model	LogL	AIC	BIC	Rank	Key Parameter (p-value)
EGARCH (1,1)	−933.815	1,873.629	1,883.208	1	beta=0.911 ($p<0.001$)
GARCH (1,1)	−934.285	1,874.570	1,884.149	2	beta=0.900 ($p<0.001$)
GJR-GARCH (1,1)	−934.251	1,876.503	1,889.275	3	gamma=0.025 ($p=0.939$)

Source: Authors' data analysis.

Figure 2 translates this into a seasonal profile, showing the mean flood event count for each calendar month across all 15 study years. Flood counts rise sharply from October, marked by the red band. This is a separate, descriptive pattern from the EGARCH persistence estimate above, not a confirmation of it. Both point toward the same operational moment: agencies that assess conditions at the wet season onset are making the decision that shapes preparedness for the months ahead.

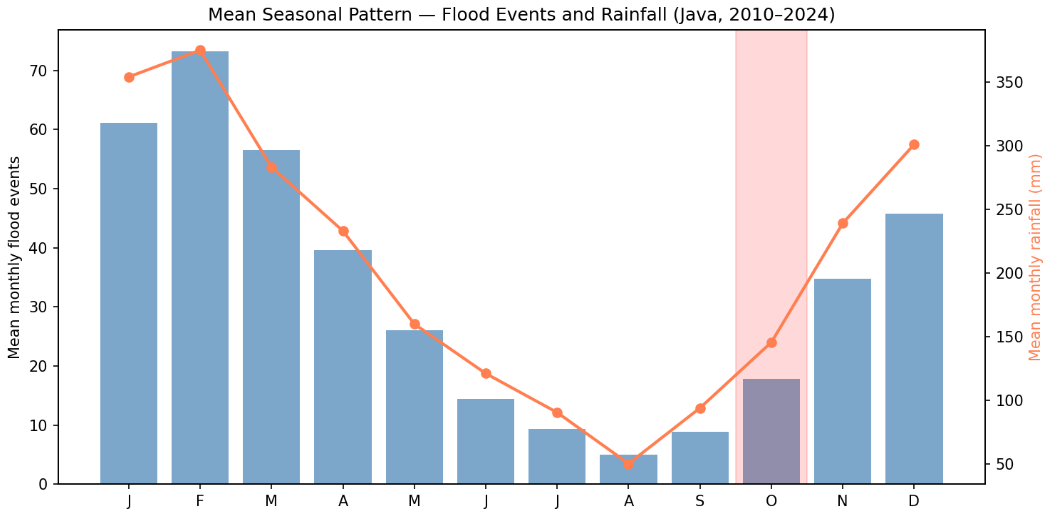

Figure 2. Mean monthly flood event counts by month across all 15 study years. The red band marks October, the start of the wet season. Flood frequency rises sharply from October onward. Together with the volatility persistence estimate above, this suggests the October-November period is a useful window for disaster preparedness planning, though this study does not test it against other possible timings. Source: authors' analysis of BNPB DIBI flood records and NASA POWER rainfall data.

Machine Learning and Hybrid Model Performance

Among standalone algorithms, XGBoost achieved the lowest RMSE (23.053), followed closely by Random Forest (RMSE = 23.072) and then Support Vector Regression (RMSE = 25.870). These error levels reflect a genuine one-month-ahead forecast: every prediction uses only rainfall and ENSO values from before the target month, not the target month's own rainfall. A model that could see the current month's actual rainfall would fit more closely, since rainfall is itself the dominant driver of flood counts, but that would not be a fair test of forecasting skill. MAPE values appear in Table 4 but should be read with caution. Fifteen zero-event dry-season months inflate percentage errors for every model. RMSE and MAE remain the primary comparison metrics.

GJR-GARCH (1,1) combined with Random Forest is the best model overall (RMSE = 22.987, MAE = 16.204), narrowly ahead of standalone XGBoost (RMSE = 23.053) and standalone Random Forest (RMSE = 23.072) (Table 4). GJR-GARCH ranks last by AIC among the three GARCH variants, yet its hybrid with Random Forest performs best. This suggests that AIC measures how well each model fits the conditional variance process in sample, whereas how useful the resulting variance is as a machine-learning feature is a separate question:

whether it adds information the other features do not already carry. Ghasempour et al. (2023) report a similar mismatch between GARCH model quality and feature usefulness in drought forecasting. The improvement from the best hybrid over the best standalone model is 0.29% in RMSE. This is small, and the Diebold-Mariano test below confirms it is not statistically significant.

Table 4. Top 10 Model Performance Rankings (Walk-Forward, n=143)

Rank	Model	Type	RMSE	MAE	MAPE (%)
1	GJR-GARCH (1,1) + Random Forest	Hybrid	22.987	16.204	119.8
2	XGBoost (standalone)	Standalone	23.053	16.544	119.0
3	Random Forest (standalone)	Standalone	23.072	16.440	127.7
4	EGARCH (1,1) + Random Forest	Hybrid	23.087	16.246	123.2
5	GARCH (1,1) + Random Forest	Hybrid	23.192	16.322	121.3
6	EGARCH (1,1) + XGBoost	Hybrid	23.200	16.547	120.8
7	GARCH (1,1) + XGBoost	Hybrid	23.263	16.835	127.3
8	GJR-GARCH (1,1) + XGBoost	Hybrid	23.361	16.904	127.9
9	GJR-GARCH (1,1) + SVR	Hybrid	25.709	17.616	114.6
10	SVR (standalone)	Standalone	25.870	17.599	113.1

Source: Authors' data analysis.

Diebold-Mariano Test

The Diebold-Mariano test found no statistically significant difference between the best hybrid (GJR-GARCH + Random Forest) and the best standalone model (XGBoost) over 143 walk-forward steps (DM = -0.103, p = 0.918). A non-significant result does not prove the two models perform identically. It means the test did not detect a difference large enough to rule out chance, given this sample size. As a more direct check, we also compared Random Forest with and without the GARCH feature, the same algorithm both times, so any difference can only come from that one feature. This test also found no significant difference (DM = -0.408, p = 0.684). Together, these results point the same way from two angles: adding the GARCH variance feature does not produce a detectable gain in monthly prediction accuracy. Zeng et al. (2022) make a relevant point beyond prediction accuracy: the GARCH parameters themselves, the persistence and climate drivers they describe, are a distinct output that no ML model produces.

SHAP Feature Importance Analysis

SHAP values test directly whether the GJR-GARCH conditional variance feature drives performance in the hybrid model. In the standalone XGBoost model, the seasonal cosine term contributes most (mean absolute SHAP = 9.830), closely followed by flood count lag-1 (9.184), rainfall lag-1 (4.329), and rainfall lag-2 (3.215). The GJR-GARCH plus Random Forest hybrid leans on a different lead feature: flood count lag-1 dominates (mean absolute SHAP = 9.205), followed by rainfall lag-1 (5.652) and the seasonal cosine term (3.390). The GJR-GARCH conditional variance ranks ninth of the ten features in the hybrid model, just above the third rainfall lag and well below every other predictor. Table 5 reports the full ranking for both models.

Table 5. SHAP Mean Absolute Feature Importance

Rank	Feature	XGBoost Standalone	GJR-GARCH + RF Hybrid
1	Seasonality (cos)	9.830	Flood count lag-1: 9.205
2	Flood count lag-1	9.184	Rainfall lag-1: 5.652
3	Rainfall lag-1	4.329	Seasonality (cos): 3.390
4	Rainfall lag-2	3.215	Rainfall lag-2: 1.581

5	ENSO (prior year)	2.215	Seasonality (sin): 1.314
6	Flood count lag-3	1.894	Flood count lag-3: 1.251
7	Flood count lag-2	1.753	Flood count lag-2: 0.512
8	Rainfall lag-3	1.322	ENSO (prior year): 0.436
9	Seasonality (sin)	0.559	GJR-GARCH variance: 0.364
10	-	-	Rainfall lag-3: 0.220

Source: Authors' data analysis.

Figure 3 shows the SHAP feature importance for both model configurations side by side. The GARCH-derived conditional variance sits near the bottom of the hybrid model ranking, ninth of ten features, well below every other predictor except the third rainfall lag. This matches what Table 5 reports numerically. The volatility estimate adds little independent predictive signal once the recent rainfall lags are already available as features.

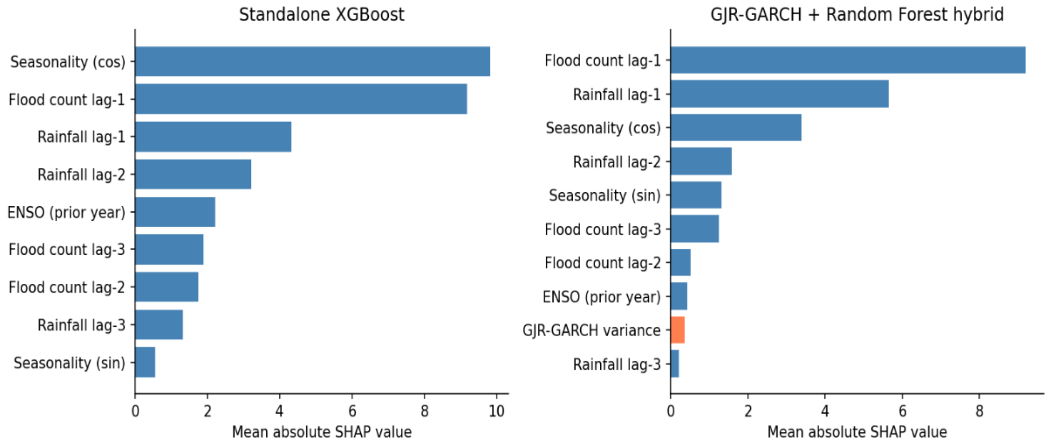

Figure 3. Mean absolute SHAP feature importance for the standalone XGBoost model (left) and the GJR-GARCH plus Random Forest hybrid model (right). The GARCH-derived conditional variance ranks ninth of ten predictors in the hybrid model, indicating that the volatility estimate adds little independent information beyond the rainfall lags already present in both configurations. Source: authors' model output.

This finding clarifies the source of the RMSE advantage in the best hybrid model. The improvement is not attributable to the GARCH volatility feature. The Diebold-Mariano results above are consistent with this reading. The collinearity is structural. The GARCH variance feature is estimated from the same recent rainfall values that already enter the ML feature set as rainfall lags. Once those lagged rainfall values are available to the model directly, a separate measure of how unusual they were adds little further information for predicting event counts. This contrasts with energy and financial domains, where GARCH features convey information genuinely absent from the baseline predictors (Zeng et al., 2022; Amirshahi & Lahmiri, 2023). The GARCH component nonetheless makes a substantive contribution through a different channel. EGARCH (1,1) reveals that Java's monsoon rainfall carries a persistent volatility structure with a half-life of approximately seven months. This characterisation of the risk cycle, when risk is elevated and for how long, is a distinct output that no ML model produces. Its practical value lies in informing preparedness timing. The two components serve complementary rather than integrated functions: GARCH characterises the seasonal risk cycle; ML predicts monthly event frequency. For BNPB and provincial BPBD offices, using both outputs separately provides more actionable information than either alone.

Conclusion

Research Conclusion

This study examined whether modelling rainfall volatility improves monthly flood disaster risk forecasting in Java, Indonesia. The results require a careful answer across two separate dimensions.

On volatility characterisation, EGARCH (1,1) best described the conditional variance structure with a persistence parameter implying a seven-month volatility half-life. ENSO operates at an annual scale rather than a monthly variance scale, so it functions here as a prediction feature rather than a variance driver. The October–November wet season onset looks like an important window for disaster preparedness mobilisation, though this study does not test it against other possible timings.

On monthly prediction, XGBoost achieved the lowest standalone error and GJR-GARCH combined with Random Forest produced the best hybrid performance. SHAP analysis found that the GARCH-derived conditional variance ranked near the bottom of all predictors. The 0.29% RMSE improvement from the best hybrid over the best standalone is not attributable to the GARCH feature. The conditional variance is structurally collinear with the recent rainfall values already in the feature set. Two Diebold-Mariano tests, one comparing the best hybrid against the best standalone model, and one comparing Random Forest with and without the GARCH feature, both found no statistically significant advantage. The central contribution of this study is a clarification. At monthly scale, GARCH and machine learning serve complementary rather than integrated functions in disaster risk assessment.

Limitations

Three limitations bound interpretation of these results. First, the temporal resolution mismatch between the variance model and ML target is structural at monthly scale: $\sigma^2 t$ is derived from the same rainfall series that dominates ML predictions. This limitation would not apply if the ML target were annual flood counts and the GARCH layer captured monthly variance across the preceding year. Second, treating Java's six provinces as a single region suppresses provincial heterogeneity that would be directly useful to local BPBD offices. As an interim step, local disaster managers can use the regional October to November volatility assessment as a general alert signal. They can then rely on local rainfall gauges and provincial BPBD monitoring networks to decide where and when to deploy resources within that wet season. Third, MAPE is unreliable for this data due to 15 zero-event dry-season months; RMSE and MAE are the reliable primary metrics.

Recommendations for Future Research

Three directions follow from these findings. The most immediate is testing the hybrid framework at annual temporal resolution, where $\sigma^2 t$ summarises 12 months of variance rather than duplicating the current month's rainfall, removing the collinearity identified here. A GARCH-MIDAS specification would be particularly well suited to bridge monthly variance dynamics to annual prediction targets. The second direction is provincial disaggregation using longer historical BNPB records to provide spatially targeted preparedness guidance. The third is incorporating additional climate drivers including the Indian Ocean Dipole and Madden-Julian Oscillation, which influence Java's intraseasonal rainfall variability and may improve prediction accuracy for wet-season onset months.

Acknowledgements

The authors acknowledge the Indonesian National Disaster Management Agency (BNPB) for providing the Data Informasi Bencana Indonesia (DIBI) flood event database, NASA for the POWER precipitation dataset, and NOAA Physical Sciences Laboratory for the ENSO Niño 3.4 index. No external funding was received for this research. Claude and DeepSeek were used for language polishing purposes; the authors take full responsibility for the scientific contributions and originality of this work.

References

Amirshahi, B., & Lahmiri, S. (2023). Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. *Machine Learning with Applications*, *12*, 100465. https://doi.org/10.1016/j.mlwa.2023.100465

Araya, H. T., Aduda, J., & Berhane, T. (2024). A hybrid GARCH and deep learning method for volatility prediction. *Journal of Applied Mathematics*, *2024*(1), 6305525. https://doi.org/10.1155/2024/6305525

Azi, A., Saleh, R. F., Ardana, W. M., & Kusrini, K. (2024). Flood prediction using support vector regression (Case study of floodgates in Jakarta). *Journal of Computer Network*s, *Architecture and High Performance Computing*, *6*(3), 1333–1344. https://doi.org/10.47709/cnahpc.v6i3.4360

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis forecasting and control—Rev. Ed.* Holden-Day. https://agris.fao.org/search/en/providers/122621/records/647396acce9437aa760043a3

Breiman, L. (2001). *Random forests*. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chen, J., Huang, G., & Chen, W. (2021). Towards better flood risk management: Assessing flood risk and investigating the potential mechanism based on machine learning models. *Journal of Environmental Management*, *293*, 112810. https://doi.org/10.1016/j.jenvman.2021.112810

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366), 427. https://doi.org/10.2307/2286348

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. *Advances in Neural Information Processing Systems*, *9*. https://papers.nips.cc/paper_files/paper/1996/hash/d38901788c533e8286cb6400b40b386d-Abstract.html

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987. https://doi.org/10.2307/1912773

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1–50. https://doi.org/10.1080/07474938608800095

Farid, M., Sihombing, Y. I., Kuntoro, A. A., Adityawan, M. B., Syuhada, M. M., Januriyadi, N. F., Moe, I. R., & Nurhakim, A. (2023). Development of flood hazard index under climate change scenarios in Java Island. *Progress in Disaster Science*, *20*, 100302. https://doi.org/10.1016/j.pdisas.2023.100302

Firmansyah, A. J., Nurjani, E., & Sekaranom, A. B. (2022). Effects of the El Niño-Southern Oscillation (ENSO) on rainfall anomalies in Central Java, Indonesia. *Arabian Journal of Geosciences*, *15*(24), 1746. https://doi.org/10.1007/s12517-022-11016-2

Ghasempour, R., Roushangar, K., & Alizadeh, F. (2023). Hybrid models for drought forecasting: Integration of multi pre-processing-data driven approaches and non-linear GARCH time series model. *Arabian Journal of Geosciences*, *16*(6), 361. https://doi.org/10.1007/s12517-023-11447-5

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

Hamzah, M. Z., Rico, R., & Sofilda, E. (2023). Flood disaster mitigation policy in strengthening social and economic resilience in Java Island. *OIDA International Journal of Sustainable Development*, *16*(10), 11–26. https://ssrn.com/abstract=4685776

Hasibuan, M. A., Shidiq, I. A., Asifin, H. A. Z., Sa'diyah, Q., & Giarno. (2024). Quantifying the relationship between ENSO-induced SST anomalies and rainfall variability in East Java's coastal regions. *Jurnal Geografi : Media Informasi Pengembangan Dan Profesi Kegeografian*, *21*(2), 94–102. https://doi.org/10.15294/jg.v21i2.16790

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, *54*(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

Mishra, P., Ray, S., Lal, P., Nair, S. B., Matuka, A., Tashkandy, Y., & Emam, W. (2025). Climate modeling for South Asia: Statistical and deep learning for rainfall and temperature prediction. *Scientific Reports*, *15*(1), 38235. https://doi.org/10.1038/s41598-025-22149-1

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, *59*(2), 347. https://doi.org/10.2307/2938260

Nti, I. K., Nyarko-Boateng, O., Boateng, S., Bawah, F. U., Agbedanu, P. R., Awarayi, N. S., Nimbe, P., Adekoya, A. F., Weyori, B. A., & Akoto-Adjepong, V. (2021). Enhancing flood prediction using ensemble and deep learning techniques. *2021 22nd International Arab Conference on Information Technology (ACIT)*, 1–9. https://doi.org/10.1109/ACIT53391.2021.9677084

Pandey, P. K., Tripura, H., & Pandey, V. (2019). Improving prediction accuracy of rainfall time series by hybrid SARIMA–GARCH modeling. *Natural Resources Research*, *28*(3), 1125–1138. https://doi.org/10.1007/s11053-018-9442-z

Sejati, A. W., Buchori, I., Lakshita, N. M., Wiratmaja, I. G. A., & Ulfiana, D. (2024). The spatial analysis of urbanization dynamic impacts in a 50-year flood frequency in Java, Indonesia. *Natural Hazards*, *120*(3), 2639–2659. https://doi.org/10.1007/s11069-023-06298-y

Uddin, M. A., Kamal, A. M., Shahid, S., & Chung, E.-S. (2020). Volatility in rainfall and predictability of droughts in Northwest Bangladesh. *Sustainability*, *12*(23), 9810. https://doi.org/10.3390/su12239810

Zeng, H., Shao, B., Bian, G., Dai, H., & Zhou, F. (2022). A hybrid deep learning approach by integrating extreme gradient boosting-long short-term memory with generalized autoregressive conditional heteroscedasticity family models for natural gas load volatility prediction. *Energy Science & Engineering*, *10*(7), 1998–2021. https://doi.org/10.1002/ese3.1122