Journal of Water Resources Management, Vol. 3 No. 2 (2025) p. 1-4

JOWRM

Journal of Water Resources Management
Journal homepage: https://journal.water.gov.my

e-ISSN: 2811-3578

Acoustic Rainfall Sensing: A Data-Driven Approach For Urban Rainfall Intensity Estimation

Mohammed I.I. Alkhatib[1a], Chow Ming Fai[1,2b*], Amin Talei[1c], Valentijn Pauwels[3d]

[1] Department of Civil Engineering, School of Engineering, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, Selangor 47500, Malaysia
[2] Monash Climate-Resilient Infrastructure Research Hub (M-CRInfra), School of Engineering, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, Selangor, 47500 Malaysia
[3] Department of Civil Engineering, Monash University, Clayton, Victoria, Australia

*Corresponding Author

Email: [a] mohammedisaalkhatib@gmail.com, [b] chow.mingfai@monash.edu , [c] amin.talei@monash.edu, [d] Valentijn.Pauwels@monash.edu

Received 07 March 2025; Accepted 19 September 2025; Available online 27 December 2025

Abstract: Accurate rainfall estimation in urban environments remains a significant challenge due to the limitations of traditional rain gauge networks and weather radar calibration issues. This study explores the potential of acoustic rainfall sensing, leveraging audio data recorded from rainfall impacting various urban surfaces. The rainfall audio data collection was conducted in Monash University Malaysia campus over two years, using professional recorders at five different locations. A data-driven approach was employed using artificial neural networks (ANN) and extreme gradient boosting (XGBoost) models to develop an acoustic rainfall estimation model. Results demonstrated that a combination of loudness, frequency, and cepstral domain features significantly improved prediction accuracy, with the ANN model outperforming XGBoost. The ANN model demonstrated consistent performance across the training (R^2 = 0.675, RMSE = 0.287 mm/min, MAE = 0.203 mm/min) and validation datasets (R^2 = 0.681, RMSE = 0.286 mm/min, MAE = 0.203 mm/min). The findings suggest that acoustic sensing, when integrated with urban IoT frameworks, can serve as a low-cost and scalable alternative for urban rainfall monitoring. Future research should focus on enhancing feature selection techniques and expanding real-world testing environments to improve model robustness.

Keywords: Acoustic sensing, rainfall estimation, urban hydrology, artificial intelligence, machine learning

1. Introduction

Providing accurate and timely spatial information on rainfall in tropical urban catchments remains a major challenge due to the high intensity, short duration, and spatial variability of convective rainfall. Traditional methods of rainfall measurement, including rain gauge networks and weather radars, face significant limitations in urban environments [1]. Rain gauges, although widely used, suffer from spatial gaps due to insufficient network density, while weather radars require extensive calibration and bias correction techniques, making them highly dependent on supplementary ground-based observations [2,3]. The limitations of these conventional approaches highlight the need for alternative, cost-effective, and scalable solutions for rainfall monitoring.

To address these limitations, this study proposes a novel approach using acoustic sensing for rainfall intensity estimation. Given the widespread availability of smartphones and internet connectivity, rainfall intensity can be estimated using audio recordings of rainfall impact on various urban surfaces. The concept of rainfall noise analysis has been successfully applied in marine environments, where tools like the Passive Aquatic Listener (PAL) have been used for raindrop size estimation and rainfall intensity inversion modeling [4]. Recent studies have explored rainfall sound classification in forest environments [5,6] and urban settings [7,8,9,10], demonstrating the feasibility of acoustic sensing for precipitation monitoring.

There are limited studies in the literature address which data-driven modelling technique is the most suitable for developing an acoustic rainfall-sensing model. However, the literature suggests a large pool of AI and ML models that have been successfully used in other applications and could be utilized for developing acoustic rainfall sensing models. These models include artificial neural networks (ANN), neuro-fuzzy systems (NFS), support vector machine (SVM), decision trees, boosted and bagged trees (e.g., XGBoost and random forest), long short-term memory (LSTM), and convolutional neural networks (CNN) [11,12,13]. Therefore, there is a need for a holistic approach to investigate the modeling capabilities of a diverse set of models. This research aims to establish a robust correlation between acoustic rainfall-generated noise and

Corresponding author: chow.mingfai@monash.edu

actual rainfall intensity, leveraging machine learning techniques to enhance urban rainfall monitoring capabilities. The findings of this study will contribute to the development of a scalable, low-cost, and real-time alternative to conventional rainfall monitoring techniques.

2. Methodology

This study was conducted at Monash University Malaysia and its surrounding urban areas, where five distinct locations (A–E) were selected for acoustic rainfall data collection (as shown in Figure 1). These locations were chosen based on their acoustic and physical diversity, representing common urban surfaces such as concrete, interlock tiles, steel, and glass. The primary criterion for site selection was to provide a broad loudness range, with Location A representing the loudest and Location B the quietest environment, while the remaining locations ensured coverage within these loudness boundaries to enhance model generalization. Additional constraints, such as constant electricity supply and minimal human interference, were also considered to ensure uninterrupted data collection.

The data collection spanned two years (from 1st September 2020 to 31st August 2022) to capture monsoon-season rainfall events in Peninsular Malaysia, covering both the Southwest monsoon (May–September) and the Northeast monsoon (November–March). Rainfall data were recorded using a Watchdog Spectrum 2000 weather station at a 1-minute resolution with a minimum detection sensitivity of 0.25 mm/min (15 mm/h). To ensure high-quality acoustic recordings and reduce uncertainties associated with smartphone-based data collection, professional Zoom H2n field recorders were used, saving uncompressed WAV files at a 44.1 kHz sampling frequency with 16-bit depth. The resulting dataset comprised 18,404 one-minute rainfall-audio pairs, with a maximum recorded rainfall intensity of 3 mm/min (180 mm/h). Unlike traditional event-based rainfall modeling, this study adopted a data-driven, non-time-series approach, treating each one-minute rainfall-audio pair as an independent data point rather than a sequential time step. This method enabled the model to learn from individual acoustic-rainfall interactions, improving its ability to generalize across different locations. The dataset was divided into training (80%), validation (10%), and testing (10%) subsets, ensuring statistical consistency across partitions. Training and validation datasets were used for model calibration, feature selection, and hyperparameter tuning, while the testing dataset remained unseen during training to provide an unbiased assessment of model performance. Additionally, the mixed dataset from five locations allowed the models to capture a wider range of acoustic variability, enhancing their ability to generalize to new urban environments.

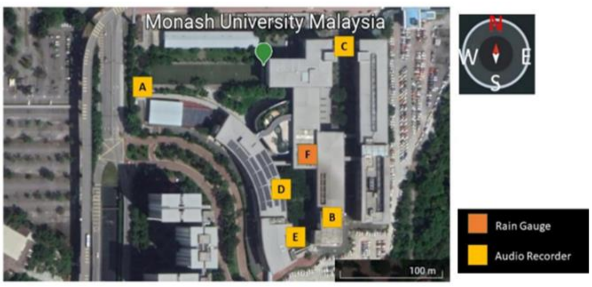

Fig. 1. Monash University Malaysia campus and the selected points for data collection

2.1 Artificial Neural Network (ANN) Model

The Artificial Neural Network (ANN) is a widely used machine learning model that employs interconnected processing units, or neurons, in a layered structure that mimics the human brain [14]. Each neuron carries weights and biases that are adjusted during the training process to minimize errors in predicting rainfall intensity [15]. In an ANN model, input neurons are connected to hidden layer neurons, which, in turn, are connected to output neurons. For a one-hidden-layer ANN, the linear combination of input variables is expressed as follows:

$$a_j = \sum(\omega_{ji}^{(1)} * x_i) + \omega_{j0}^{(1)} \quad (1)$$

where, $\omega_{ji}(1)$ are the weights, $\omega_{j0}(1)$ are the biases, and a_j are the activations. These activations are transformed using nonlinear activation functions such as the sigmoid, hyperbolic tangent, or rectified linear unit (ReLU), where ReLU is the most commonly used due to its efficient gradient propagation (Bishop, 2006). The output neurons follow another linear transformation:

$$a_k = \sum(\omega_{ki}(2) * z_i) + \omega_{k0}(2) \quad (2)$$

where $k = 1, \ldots$, k and K represents the number of outputs. In regression problems, the activation function for the output layer is an identity function, so that $yk = ak$.

Training the ANN involves finding the optimal weight vector to minimize the error between predicted and observed rainfall values. The loss function for regression problems is defined as:

$$E(w) = (1/2) * \sum ||y(x_n, w) - t_n||^2 \quad (3)$$

where xn represents the input vector, tn is the observed rainfall value, and $y(xn, w)$ is the predicted rainfall intensity. The optimization algorithm adjusts weights iteratively to minimize the error. Gradient descent optimization updates the weights by taking steps in the direction of the negative gradient of the loss function. Stochastic gradient descent (SGD) is an improved variant that updates weights based on a single data point at a time, ensuring efficient convergence. The backpropagation algorithm efficiently computes the gradients needed for weight updates. The error signal is propagated backward through the network using the equation:

$$\delta_j = (1 - z_j^2) * \sum(\omega_{kj} * \delta_k) \quad (4)$$

where δj represents the error for each output neuron. This iterative weight adjustment ensures optimal model learning.

2.2 Extreme Gradient Boosting (XGBoost) Model

XGBoost is a decision tree-based ensemble learning algorithm that employs a gradient-boosting framework [16]. It enhances traditional boosting techniques using parallel computation, allowing for efficient model training. The XGBoost model is mathematically expressed as:

$$\hat{y}_i = \varphi(x_i) = \sum f_k(x_i),\ f_k \in F \quad (5)$$

Where $\hat{yi}$ represents the predicted value, fk represents the k-th tree, $fk\ (xi\ )$ is the score of the i-th sample in the k-th tree, K is the total number of samples, xi is the ith input data, and F is all possible trees. The objective function consists of a loss function and a regularization term, formulated as follow:

$$L(\varphi) = \sum l(y_i, \hat{y}_i) + \sum \Omega(f_k) \quad (6)$$

Published by JPS Publishing
https://journal.water.gov.my

Where

$$\Omega(f) = \gamma T + 1\ 2\ \lambda \sum wj\ 2 \qquad (7)$$

Where yi is the observed value, $\hat{yi}$ is the predicted value, l is a loss function, Ω is the regularisation term to penalise the model complexity and avoid overfitting, λ is the complexity of each leaf, T total number of leaves in the decision tree, γ is a comprise parameter used to scaling the penalty, and w is the score on the j-th leaf.

Bayesian optimization was used to fine-tune hyperparameters for both models. The MATLAB bayesopt function was used for ANN, while the Python skopt library was employed for XGBoost. The objective function aimed to minimize RMSE over 100 iterations, using the expected-improvement-plus acquisition function to avoid local minima.

2.3 Model Performance Evaluation Metrics

Model performance was evaluated using three key metrics: the coefficient of determination (R^2), which measures prediction accuracy, with values closer to 1 indicating better performance; the root mean square error (RMSE), which quantifies overall prediction error while giving more weight to large deviations; and the mean absolute error (MAE), which measures the absolute difference between observed and predicted values, providing an unbiased assessment of error magnitude.

3. Result and discussion

To identify the most relevant acoustic features for rainfall intensity estimation, a Cross-Correlation Analysis (CCA) was performed on 44 extracted features from the time, frequency, and cepstral domains. Since the time-frequency domain feature is a two-dimensional representation, it was excluded from this analysis. The CCA results showed that the top-ranked features with absolute CC ≥ 0.5 include Mel Frequency Cepstral Coefficient-1 (MFCC1), root mean square of energy (RMNG), background noise (BGN), and average signal amplitude (ASA). MFCC1, belonging to the cepstral domain, exhibited the highest correlation with rainfall intensity (CC = +0.6), while the remaining three features were derived from the time domain. Features with absolute CC values between 0.3 and 0.5 included zero crossing rate (ZCR) from the time domain, along with spectral bandwidth (S_Bandwidth), spectral roll-off (S_Rolloff), spectral flux (S_Flux), spectral slope (S_Slope), spectral flatness (S_Flatness), spectral decrease (S_Decrease) from the frequency domain, and MFCC-2, MFCC-4, and MFCC-8 from the cepstral domain. In total, 14 features were shortlisted for further evaluation.

Following feature selection, the ANN and XGBoost models were trained using various feature combinations to identify the optimal input set. The dataset was divided into training, validation, and testing subsets, ensuring statistical consistency across data partitions. A systematic evaluation of 16383 non-repeating feature combinations was conducted, testing combinations from single-input models up to all 14-input models. The performance of these combinations was assessed using coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE).

As expected, MFCC1 emerged as the most effective single-input feature, given its highest correlation with rainfall intensity. Additional analysis revealed that loudness-related features (e.g., MFCC1, ASA, RMNG, and BGN) consistently contributed to model performance. However, performance gains diminished when multiple loudness features were combined, indicating potential feature redundancy. Notably, significant performance gains were observed when loudness features were combined with frequency or cepstral domain features (e.g., S_Decrease, S_Rolloff, S_Bandwidth, or higher-degree MFCCs). This suggests that complementary spectral and cepstral features enhance the ability of loudness features to differentiate rainfall intensities across varying urban surfaces.

Ultimately, the best feature combination was determined based on validation dataset performance, where adding additional features no longer yielded meaningful improvements. The optimal input set for ANN models consisted of six features, while the XGBoost model performed best with four features (Table 1). These findings highlight the importance of integrating at least one loudness feature (e.g., MFCC1 or RMNG) with spectral and cepstral features to maximize model efficiency in rainfall estimation tasks.

Table 1. Best feature combinations performance for ANN and XGBoost models

Model	Best Feature Combination	Dataset	R^2	RMSE ($mm.min^{-1}$)	MAE ($mm.min^{-1}$)
ANN (6-input)	MFCC1 ZCR MFCC4 MFCC8 S_Decrease S_Rolloff	Training	0.626	0.307	0.217
		Validation	0.625	0.309	0.219
XGBoost(4-input)	MFCC1 MFCC2 MFCC8 S_Bandwidth	Training	0.619	0.312	0.225
		Validation	0.604	0.315	0.229

To optimize model performance, a Bayesian search approach was applied to fine-tune hyperparameters for the six-feature ANN model and the four-feature XGBoost model. Table 2 presents the performance metrics for both models during training and validation. The fine-tuned ANN model demonstrated consistent performance across training and validation datasets, suggesting robust generalization with no overfitting. Conversely, the XGBoost model exhibited slight overfitting, as indicated by discrepancies between training and validation performance. Several attempts were made to reduce overfitting in the XGBoost model by adjusting regularization parameters during the Bayesian search process. However, these adjustments resulted in a trade-off between generalization and performance, with models achieving lower validation accuracy than the baseline fine-tuned model.

Table 2. Performance of fine-tuned ANN and XGBoost models

Model	Dataset	R^2	RMSE ($mm.min^{-1}$)	MAE ($mm.min^{-1}$)
ANN	Training	0.675	0.287	0.203
	Validation	0.681	0.286	0.203
XGBoost	Training	0.779	0.238	0.174
	Validation	0.630	0.304	0.216

The performance of the Artificial Neural Network (ANN) and Extreme Gradient Boosting (XGBoost) models in acoustic rainfall sensing for rainfall intensity estimation in Malaysia was evaluated based on coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) for both training and validation datasets. The ANN model

Published by JPS Publishing
https://journal.water.gov.my

demonstrated consistent performance across the training (R^2 = 0.675, RMSE = 0.287 mm/min, MAE = 0.203 mm/min) and validation datasets (R^2 = 0.681, RMSE = 0.286 mm/min, MAE = 0.203 mm/min). The minimal difference between training and validation results suggests that ANN generalizes well without significant overfitting, making it a robust and reliable approach for acoustic rainfall sensing. However, the moderate R^2 values indicate that ANN captures some variability in rainfall intensity but may require further tuning or additional feature extraction to enhance predictive accuracy.

The XGBoost model exhibited higher accuracy during the training phase (R^2 = 0.779, RMSE = 0.238 mm/min, MAE = 0.174 mm/min) compared to ANN, indicating its capability to fit complex relationships within the data. However, its validation performance declined significantly (R^2 = 0.630, RMSE = 0.304 mm/min, MAE = 0.216 mm/min), suggesting a degree of overfitting, where the model performs well on the training dataset but struggles to generalize to unseen data. The higher RMSE and MAE in the validation phase indicate that XGBoost has larger prediction errors than ANN when applied to independent datasets.

While XGBoost outperformed ANN during training, it suffered from overfitting, leading to reduced validation accuracy. In contrast, ANN maintained stable performance between training and validation datasets, making it the more generalizable model for acoustic rainfall estimation. XGBoost had lower RMSE and MAE in the training phase, indicating better initial fit, but these errors increased in the validation phase. Conversely, ANN maintained consistent error rates, suggesting more robust predictions across datasets. The results suggest that ANN is a more balanced model for rainfall intensity estimation, as it minimizes overfitting while maintaining steady performance in unseen datasets. While XGBoost captures rainfall intensity trends more effectively during training, its lower validation accuracy highlights the need for further regularization and hyperparameter tuning to improve generalization.

4. Conclusion

The comparison between ANN and XGBoost models highlights the trade-off between model complexity and generalization in acoustic rainfall sensing for rainfall intensity estimation in Malaysia. While XGBoost shows promise in capturing intricate relationships within training data, its higher validation error suggests the need for additional tuning. On the other hand, ANN delivers consistent and reliable performance, making it a preferable model for real-world deployment in rainfall intensity estimation applications. Future research should explore hybrid approaches or additional feature engineering techniques to further improve predictive accuracy.

Acknowledgement

The authors acknowledge the financial support provided by Monash University Malaysia through the Merit PhD Scholarship awarded to Mohammed I. I. Alkhatib. Additionally, this research was funded by a grant from the Ministry of Higher Education Malaysia (FRGS/1/2021/TK0/MUSM/02/6). The authors sincerely appreciate the support and resources made available for this study.

References

[1] Borga, M., Marra, F., Gabella, M., 2022. Chapter 5 - Rainfall estimation by weather radar, in: Rainfall. Elsevier, pp. 109–134.

[2] Smith, J.A., Baeck, M.L., Meierdiercks, K.L., Miller, A.J., Krajewski, W.F., 2007. Radar rainfall estimation for flash flood forecasting in small urban watersheds. Adv Water Resour 30, 2087–2097.

[3] Chumchean, S., Seed, A., Sharma, A., 2006. Correcting of real-time radar rainfall bias using a Kalman filtering approach. J Hydrol (Amst) 317, 123–137.

[4] Anagnostou, E. N., Kalogiros, J. A., Tarolli, M., et al. (2011). Passive aquatic listener tool for rainfall estimation. Journal of Hydrometeorology, 12(3), 456-472.

[5] Brown, A., Garg, S., Montgomery, J., 2019. Automatic rain and cicada chorus filtering of bird acoustic data. Appl Soft Comput 81, 105501.

[6] Bedoya, C., Isaza, C., Daza, J.M., López, J.D., 2017. Automatic identification of rainfall in acoustic recordings. Ecol Indic 75, 95–100.

[7] Avanzato, R., Beritelli, F., 2020. An innovative acoustic rain gauge based on convolutional neural networks. Information (Switzerland) 11.

[8] Avanzato, R., Beritelli, F., Franco, F. Di, Puglisi, V.F., 2019. A convolutional neural networks approach to audio classification for rainfall estimation, in: 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). pp. 285–289.

[9] Chen, Y., Wang, P., & Li, H. (2022). Deep learning-based acoustic rainfall classification for urban sensing. Environmental Monitoring and Assessment, 194(5), 1-15.

[10] Wang, X., Zhang, J., & Li, Q. (2022). Rainfall sound recognition in urban areas using CNN-based deep learning models. Applied Acoustics, 192, 108735.

[11] Zhang, Z., Xu, S., Zhang, S., Qiao, T., Cao, S., 2020. Attention based convolutional recurrent neural network for environmental sound classification. Neurocomputing.

[12] Himawan, I., Towsey, M., Roe, P., 2018. Detection and classification of acoustic scenes and events.

[13] Matsuoka, D., Nakano, M., Sugiyama, D., Uchida, S., 2018. Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model. Prog Earth Planet Sci 5, 80.

[14] Bishop, J., 2006. Pattern recognition and machine learning. Springer, Berlin.

[15] Haykin, S., 1998. Neural networks: a comprehensive foundation, 2nd ed. Pearson-Hall.

[16] Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, pp. 785–794.