Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Analyzing Cloud Size Using Weather Radar Data for Improved Flood Disaster Prediction

Atikah Balqis Basri [1*], Md Tanvir Hasan [2], Khairayu Badron [3], Dini Oktarina Dwi Handayani [4], Zainab S. Attarbashi [5], Norhafana Mohamed [6]
[1, 2] *Department of Information Systems, International Islamic University Malaysia, Kuala Lumpur, Malaysia*
[3] *Department of Electrical and Computer Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia*
[4, 5] *Department of Computer Science, International Islamic University Malaysia, Kuala Lumpur, Malaysia*
[6] *Department of Mechanical and Aerospace Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia*
**Corresponding Author Email: atikahbalqis@iium.edu.my*

ARTICLE INFO

Received: 10 Dec 2024

Revised: 15 Feb 2025

Accepted: 25 Feb 2025

ABSTRACT

Cloud size and rain cell analysis are essential to meteorological research and flood risk prediction, especially in regions vulnerable to heavy rainfall events and flooding. By leveraging weather radar data, which captures reflectivity values indicating precipitation intensity, researchers can derive cloud size and better understand rainfall's spatial and temporal patterns. This paper introduces a comprehensive approach for analyzing cloud size using weather radar data, incorporating a series of systematic steps that enhance the detection and evaluation of rain cells. The process begins with data acquisition, wherein raw radar data is obtained from weather monitoring stations or agencies. Following acquisition, preprocessing techniques are applied to convert dBZ values into reflectivity values, remove non-meteorological noise, and organize data into structured grids. These preprocessing steps ensure data accuracy and facilitate analysis across different spatial regions and time intervals. The next phase involves thresholding and cloud boundary definition, where a reflectivity threshold (e.g., 30 dBZ) is used to create a binary cloud mask, identifying significant rain cells within the radar scans. This binary mask provides a foundation for further analysis, allowing the delineation of cloud boundaries and the isolation of specific rain cell regions. Feature extraction is then performed to quantify critical attributes, such as cloud size, maximum reflectivity, and rain cell movement patterns, which are crucial for accurate flood prediction. Finally, visualization methods, including time series plots, allow for assessing rain cell evolution over time, providing real-time insights into rainfall dynamics. Collectively, these steps enhance the predictive accuracy of flood risk models and offer valuable data for disaster mitigation strategies, contributing to more effective and timely responses in flood-prone areas.

Keywords: Flood prediction, computational model, radar reflectivity, rain cell size, weather forecasting.

INTRODUCTION

Accurate flood forecasting is a critical challenge in meteorology, particularly in flood-prone regions where heavy rainfall can lead to sudden inundation. Traditional flood forecasting models typically rely on meteorological data such as rainfall intensity and historical river flow data. While these models are valuable, they are often limited in their ability to account for the spatial distribution and size of rain cells, which significantly influence the occurrence of flooding. Rain cells, which are clusters of convective storms or precipitation patterns, vary in size and duration. Larger, stationary rain cells or clusters of rain cells can cause localized flooding, while smaller, fast-moving rain cells can result in flash floods.

Weather radar systems provide a powerful tool for observing rain cells and assessing cloud sizes, which are key parameters in flood prediction models. Radar data captures the reflectivity values within rain cells, which can be

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

processed to estimate both the spatial extent and intensity of clouds. Analyzing cloud size and movement patterns can inform predictive models and improve flood warnings, providing early alerts that allow for timely preparation and response. This paper outlines a detailed approach to analyzing cloud size and rain cells, leveraging radar data for enhanced flood disaster prediction.

LITERATURE REVIEW

Flood Prediction in Malaysia

Flood prediction models have evolved, integrating various data types such as precipitation, river discharge, and hydrological models. In Malaysia, flood forecasting traditionally relies on rain gauges and hydrological models. However, these methods face challenges such as the uneven distribution of rain gauges and incomplete data coverage, especially in remote or less accessible areas [1].

Weather radar systems have become an essential tool for real-time flood prediction, as they provide high-resolution spatial and temporal data on rainfall. Radar reflectivity values (measured in decibels of Z or dBZ) correlate with rainfall intensity and offer a finer spatial resolution than rain gauges. Reflectivity values allow for a more accurate assessment of the rate of precipitation and its potential impact on flooding [2].

Rain Cell Size and Flooding

Research in cloud size and rain cell analysis has evolved substantially, driven by the need to understand precipitation dynamics and enhance flood prediction models [3]. Early studies primarily focused on the application of radar technology to observe precipitation patterns, with radar reflectivity as a key parameter for understanding rain intensity and distribution. The Z-R relationship, which converts radar reflectivity (dBZ) to rainfall rates, has been widely adopted to gauge precipitation levels and map rain cell characteristics across various spatial scales [4]. The use of radar data in meteorology has been particularly beneficial in regions prone to flash floods, where timely, high-resolution data is essential for effective disaster response and preparedness [5].

The thresholding method for cloud boundary identification, often based on reflectivity values (e.g., 30 dBZ or higher), has proven effective in isolating significant precipitation regions and estimating rain cell sizes. This thresholding approach, combined with connected component analysis, enables researchers to identify and classify contiguous areas of intense rainfall, referred to as rain cells. The extracted rain cells are valuable in understanding cloud dynamics, including cloud size, intensity, and movement patterns, which are essential features for flood risk models [6]. Additionally, studies have explored the integration of radar-derived features with machine learning, which allows for more complex, data-driven flood prediction models that incorporate historical flood data and rain cell characteristics [7].

More recent advancements involve the application of machine learning and data-driven approaches to radar data analysis. Machine learning algorithms, such as Random Forest and neural networks, have been employed to predict flood risks based on extracted rain cell features, offering improvements in predictive accuracy over traditional statistical models [8]. The integration of real-time radar data with predictive models enables near real-time flood risk assessments, enhancing early warning systems[9-11]. This paper builds on this foundation by presenting a systematic approach to cloud size and rain cell analysis, utilizing radar data to refine flood prediction and contribute to disaster mitigation strategies in high-risk regions[12].

Machine Learning for Flood Prediction

The integration of machine learning techniques in flood prediction is gaining attention due to their ability to process large and complex datasets. Machine learning models, including Random Forest, Support Vector Machines (SVM), and Neural Networks, have been shown to improve flood prediction accuracy by learning complex relationships between input features (e.g., radar data, rain cell size, and weather conditions) and flood occurrence [13]. By combining radar reflectivity data with rain cell size, machine learning models can better predict rainfall's spatial and temporal patterns that lead to flooding.

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

METHODOLOGY

Data Acquisition

Radar data is sourced from weather stations or meteorological agencies, typically stored in binary formats or specialized data structures such as NetCDF or HDF5. Reflectivity data from radar readings, given in decibels (dBZ), captures the intensity of precipitation within each grid cell, forming the basis for rain cell analysis.

Preprocessing

Reflectivity (Z) is derived from dBZ values using the formula

$R= a \cdot Z^b$

where R is the rainfall rate in mm/h, Z is the radar reflectivity in mm^6/m^3, and a and b are empirically derived constants.

which converts radar reflectivity readings into meaningful measurements of precipitation intensity [14]. To ensure data accuracy, non-meteorological signals, such as reflections from buildings or birds, along with low-intensity noise, are filtered out. The data is then organized into 2D or 3D grids representing spatial dimensions (latitude, longitude) and time, which facilitates analysis over specific regions and periods for enhanced interpretation of precipitation patterns.

Thresholding and Cloud Boundary Definition

A reflectivity threshold, such as 30 dBZ, is set to identify significant rain cells, marking areas with meaningful precipitation. Based on this threshold, a binary cloud mask is created, classifying grid cells within a cloud. Cells that exceed the threshold are flagged as part of the cloud, effectively delineating its boundaries and distinguishing areas of interest for further analysis.

Cloud Size Calculation

Connected component analysis or clustering algorithms are used to identify contiguous regions within the cloud mask, where each region corresponds to an individual rain cell. These regions are identified by grouping neighboring grid cells that share similar characteristics, such as exceeding the reflectivity threshold, thereby enabling the detection of separate rain cells.

Once the rain cells are identified, the total area covered by each cell is calculated by counting the number of grid cells that make up the rain cell. The physical area of the rain cell can then be derived if the spatial resolution of the radar data, such as kilometers per cell, is known. The cloud size in square kilometers is calculated using the formula:

Cloud Size (km^2) = Number of Grid Cells × Area per Cell.

For larger areas or over multiple periods, the areas of all identified rain cells can be aggregated to analyze overall cloud coverage. This can involve summing or averaging the cloud sizes to provide insights into the spatial extent of precipitation events, which can be useful for tracking storm developments and assessing potential flood risks.

Feature Extraction

Maximum reflectivity within each rain cell indicates precipitation intensity, with higher values signifying more intense rainfall. To aid in forecasting rain patterns, rain cells are tracked across successive radar scans to identify their movement, helping predict the path of precipitation. Additionally, if 3D radar data is available, a vertical structure analysis can offer valuable insights into the thickness of clouds and the altitudinal distribution of reflectivity, enhancing the understanding of how rainfall intensity varies with height and contributing to more accurate weather predictions.

Visualization and Interpretation

To analyze cloud size and movement, visualize the cloud mask with radar reflectivity overlays to inspect identified rain cells, and, if multiple time steps are available, create time series plots to track changes in cloud size and

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

movement over time, providing insights into the evolution of precipitation patterns. Using MATLAB code processes a series of radar data files to extract key information related to cloud size, rainfall rates, and rain cells. The radar data files are specified in the filenames array, which contains a list of file paths to the radar datasets. Each file is assumed to be in binary format, where the first 128 bytes contain a header that is skipped during the reading process. The remaining data is assumed to be in a specific format (32-bit floating-point), representing radar reflectivity measurements over multiple sweeps (scans) of the radar system. The code begins by initializing arrays to store results for each file: cloud_size_all to track the cloud size (number of cells with reflectivity greater than 30 dBZ), rainfall_rate_all to store rainfall rates, and rain_cells_all to keep track of the labeled rain cells (connected regions of high reflectivity). The loop iterates over each radar file, opening and reading the data, which is then reshaped into a 2D matrix where each row corresponds to a sweep and each column to a spatial point in the radar scan. This reshaped data is then used for further analysis.

Reflectivity values in the radar data are converted from dBZ (decibels of reflectivity) to linear reflectivity using the formula:

$$\text{Reflectivity} = 10^{(radar\_data/10)}$$

The code identifies regions with high reflectivity (greater than 30 dBZ) to estimate cloud size. The cloud size is determined by counting how many cells in the matrix exceed this threshold, providing a quantitative measure of the cloud's extent. Additionally, the reflectivity values are used to calculate rainfall rates using the Z-R relationship, where a formula involving constants a=200a = 200a=200 and b=1.6b = 1.6b=1.6 is applied to convert reflectivity into rainfall intensity (mm/h). Finally, the code extracts rain cells by labeling connected regions in the cloud mask (areas with reflectivity > 30 dBZ) using MATLAB's bwlabel function. Each labeled region represents a distinct rain cell. The script then generates a visual representation of the rainfall rate for the final sweep, displaying it as an image with a colorbar to indicate the rainfall intensity. The results for cloud size (number of cells) and the total number of rain cells are printed to the console for each file. The code thus provides both quantitative and visual outputs for analyzing radar data, specifically focusing on cloud structure and rainfall characteristics. The size of rain cells is extracted from the spatial distribution of rainfall patterns. Rain cells are identified by grouping contiguous areas with significant rainfall intensity (above a certain threshold) and analyzing their spatial extent and movement over time.

RESULT

Cloud Size vs. Rain Cells: As the cloud size increases (more radar cells with reflectivity > 30 dBZ), the number of rain cells also tends to increase. This is because a larger cloud will likely contain more distinct regions of high reflectivity, which are identified as individual rain cells. For example, the file with the largest cloud size has the most rain cells (1903). Variation Between Files: The first file represents a relatively small cloud with fewer rain cells, while the second file has a much larger cloud and more rain cells. The third file has the largest cloud and the highest number of rain cells, indicating a highly complex precipitation pattern with many distinct rain-producing areas.

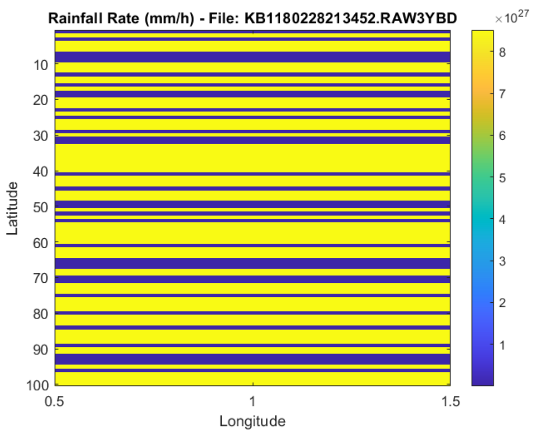

Figure 1. Clusters of radar cells for KB1180228213452.RAW3YBD

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

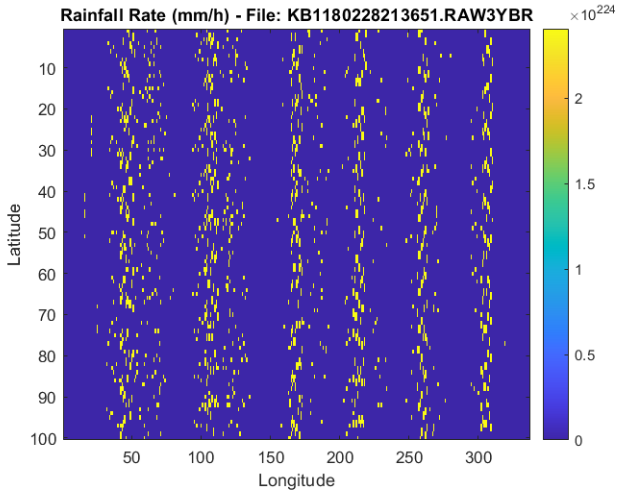

Figure 2. Clusters of radar cells for KB1180228213651.RAW3YBR

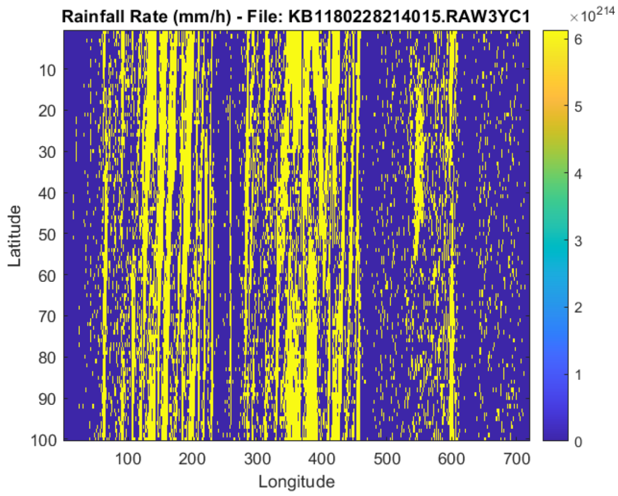

Figure 3. Clusters of radar cells for KB1180228214015.RAW3YC1

In conclusion, the results highlight the variability in cloud size and rain cell distribution across different radar datasets. The number of cells with high reflectivity and the number of distinct rain cells both increase as the cloud becomes larger and more complex.

DISCUSSION

The results presented in the output indicate significant variability in the cloud size and number of rain cells across the three radar data files, which reflects differences in the intensity, complexity, and spatial distribution of precipitation in each dataset. This variation provides insights into the characteristics of different weather systems captured by the radar and can be useful for understanding the dynamics of rainfall and cloud formation. While the results provide useful information about cloud and rain cell characteristics, they also raise some important considerations. For example, the accuracy of cloud size and rain cell counts depends on the threshold of 30 dBZ reflectivity, which may not always accurately represent the boundary between cloud and precipitation. In some cases, reflectivity just below this threshold might still contribute to rain cells or cloud formation. Additionally, the Z-R relationship used to convert reflectivity to rainfall rate has inherent assumptions and could lead to variations in rainfall estimates depending on the type of storm or weather system being analyzed.

Furthermore, the method of identifying rain cells assumes that contiguous high-reflectivity regions are directly associated with distinct rain cells. However, some rain cells may be separated by regions of low reflectivity or be part of a larger system that is more complex than a simple label-based approach can capture. Advanced techniques such as tracking rain cells over time or incorporating additional meteorological data could offer a more nuanced understanding of precipitation patterns.ile the results provide useful information about cloud and rain cell characteristics, they also raise some important considerations. For example, the accuracy of cloud size and rain cell counts depends on the threshold of 30 dBZ reflectivity, which may not always accurately represent the boundary between cloud and precipitation. In some cases, reflectivity just below this threshold might still contribute to rain cells or cloud formation. Additionally, the Z-R relationship used to convert reflectivity to rainfall rate has inherent assumptions and could lead to variations in rainfall estimates depending on the type of storm or weather system being analyzed.

CONCLUSION

Analyzing cloud size using weather radar data offers a robust approach to understanding rain cell characteristics and improving flood risk assessments. By leveraging radar reflectivity values, the described computational approach captures key features such as cloud size, maximum reflectivity, and cell movement. These features, in turn, enhance predictive models that aim to deliver timely and reliable flood alerts. This method provides a valuable addition to meteorological analysis, supporting disaster preparedness efforts in flood-prone regions. Future studies may further optimize these techniques through higher-resolution radar data and advanced machine learning models, refining flood prediction and mitigation strategies.

Acknowledgement

We express our sincere gratitude to the Fundamental Research Grant Scheme (FRGS) for their financial assistance, FRGS/1/2023/TK08/UIAM/03/2. To the Malaysian Meteorological Department (MetMalaysia) for their valuable collaboration and support, to the Department of Irrigation and Drainage, Ministry of Energy Transition and Water Transformation, which made this research possible.

REFERENCES

[1] Nadia, Ayoub., Azwa, Abdul, Aziz., Wan, Azani, Mustafa. (2023). 1. FloodIntel: Advancing flood disaster forecasting through comprehensive intelligent system approach. Journal of autonomous intelligence, doi: 10.32629/jai.v7i1.870 W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[2] Amrul, Faruq., Shamsul, Faisal, Mohd, Hussein., Aminaton, Marto., Shahrum, Abdullah. (2022). 2. Flood River Water Level Forecasting using Ensemble Machine Learning for Early Warning Systems. IOP conference series, doi: 10.1088/1755-1315/1091/1/012041

[3] Thanh, H., Nguyen., Sophie, Ricci., Andrea, Piacentini., Raquel, Rodriguez, Suquet., Gwendoline, Blanchet., Santiago, Peña, Luque., Peter, Kettig. (2022). 3. Enhancing Flood Forecasting with Dual State-Parameter Estimation and Ensemble-based SAR Data Assimilation.

[4] Efrat, Morin., Frederik, Kratzert. (2022). 4. Flood forecasting with machine learning models in an operational framework. Hydrology and Earth System Sciences, doi: 10.5194/hess-26-4013-2022

[5] Yi, Lin, Tew., Mou, Leong, Tan., Liew, Juneng., Kwok, Pan, Chun., Mohamad, Hafiz, bin, Hassan., Sazali, Osman., Narimah, Samat., Chun, Kiat, Chang., Muhammad, Humayun, Kabir. (2022). 5. Rapid Extreme Tropical Precipitation and Flood Inundation Mapping Framework (RETRACE): Initial Testing for the 2021-2022 Malaysia Flood. ISPRS international journal of geo-information, doi: 10.3390/ijgi11070378

[6] Roland, Löwe., Karsten, Arnbjerg-Nielsen., Rocco, Palmitessa., Dan, Wang., Heng, Jiang. (2022). 6. Improved flood predictions by combining satellite observations, topographic information and rainfall spatial data using deep learning. doi: 10.5194/egusphere-egu22-8823

[7] Foo-Hoat, Lim., Wei-Koon, Lee., Sazali, Osman., Amanda, Sean, Peik, Lee., Wei, Sze, Khor., Nurul, Hidayah, Ruslan., Nor, Hisham, M., Ghazali. (2022). 7. Multi-model approach of data-driven flood forecasting with error

Journal of Information Systems Engineering and Management
2025, 10(47s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

correction for large river basins. Hydrological Sciences Journal-journal Des Sciences Hydrologiques, doi: 10.1080/02626667.2022.2064754

[8] D., Vinod., Karunanithi, Senthamilselvi, Vijayanand., A., Sivanesh, Kumar. (2022). 8. A formal forecasting approach to predict flood disaster and recovery strategy using machine learning. Nucleation and Atmospheric Aerosols, doi: 10.1063/5.0120252

[9] Muhamad, Nur, Adli, Zakaria., Marlinda, Abdul, Malek., Maslina, Binti, Zolkepli., Ali, Najah, Ahmed. (2021). 10. Application of artificial intelligence algorithms for hourly river level forecast: A case study of Muda River, Malaysia. alexandria engineering journal, doi: 10.1016/J.AEJ.2021.02.046

[10] David, M., Romps., Andrew, M., Vogelmann. (2017). 3. Methods for Estimating 2D Cloud Size Distributions from 1D Observations. Journal of the Atmospheric Sciences, doi: 10.1175/JAS-D-17-0105.1

[11] D., Jolivet., A., J., Feijt. (2003). 4. Cloud thermodynamic phase and particle size estimation using the 0.67 and 1.6 µm channels from meteorological satellites. Atmospheric Chemistry and Physics, doi: 10.5194/ACPD-3-4461-2003

[12] Sergey, Y., Matrosov., Alexei, Korolev., Andrew, J., Heymsfield. (2002). 1. Profiling cloud ice mass and particle characteristic size from Doppler radar measurements. Journal of Atmospheric and Oceanic Technology, doi: 10.1175/1520-0426(2002)019<1003:PCIMAP>2.0.CO;2

[13] Coltin, Grasmick., Bart, Geerts., Jeffrey, R., French., Samuel, Haimov., Robert, M., Rauber. (2022). 6. Estimating Microphysics Properties in Ice-dominated Clouds from Airborne Ka – W band Dual-wavelength Ratio Reflectivity Factor in Close Proximity to in situ Probes. Journal of Atmospheric and Oceanic Technology, doi: 10.1175/jtech-d-21-0147.1

[14] Pragya, Vishwakarma., Julien, Delanoë., Susana, Jorquera., Pauline, Martinet., Frédéric, Burnet., Alistair, Bell., Jean-Charles, Dupont. (2023). 13. Climatology of estimated liquid water content and scaling factor for warm clouds using radar–microwave radiometer synergy. Atmospheric Measurement Techniques, doi: 10.5194/amt-16-1211-2023