Rapid modeling of catastrophic floods: a case study of the Padang flood on July 14, 2023

Yanuar Rizky Ramadhan[1*] and *Dennish* Ari Putro[1]

[1]PT Reasuransi MAIPARK Indonesia, Multivision Tower 8th Floor, DKI Jakarta, 12960, Indonesia

Abstract. This study presents a rapid flood modeling approach for catastrophic flood events in Padang City on July 14, 2023. Floods pose significant risks to urban areas, exacerbated by rapid urbanization and climate change. This research provides a solution to the urgent need for quickly and accurately flood mapping, crucial for effective disaster response and mitigation. Various data such as Digital Elevation Models (DEMs), rainfall, evapotranspiration, soil types, and land use were employed in this study. The modeling process involved DEM conditioning, bias correction of rainfall, defining the model domain using the Topographic Wetness Index (TWI), and applying hydrological and hydrodynamic models. Model verification was conducted using flood location and depth data collected from mass media and social media, showing that the model achieved 77.8% accuracy in mapping inundated areas, with Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) of flood depth at 38.093 and 27.584, before the model was calibrated. The calibration process significantly improved the model's accuracy.

1 Introduction

Hydrometeorological disasters, particularly floods, have become a serious concern in Indonesia over the past few decades [1]. Along with the rapid population growth and the rate of land-use changes, Indonesia has experienced an increasing trend in flood disasters in terms of both frequency and intensity. Data on flood events recorded by the Data Informasi Bencana Indonesia Badan Nasional Penanggulangan Bencana (DIBI BNPB) show an increasing trend of flood events in Indonesia from 2014 to 2021. Floods often have devastating impacts, including loss of life and property damage [2]. Therefore, research on flood management and mitigation has become increasingly important to protect communities and national assets.

The main challenge in flood risk mitigation is the ability to quickly and accurately map affected areas (extent of inundation and flood depth). In disaster situations, time is of the essence, and rapid response can reduce the losses caused by floods. However, rapid flood mapping is hindered by limitations in actual data and large computational resources, often taking a long time [3, 4].

To address the challenges faced in flood modeling, it is necessary to develop methods and tools that can accommodate data limitations and increase the speed of mapping affected areas [5]. This solution will enable authorities to respond to flood disasters more quickly and efficiently, thereby helping to reduce the losses caused by floods.

In facing these challenges, this research aims to develop a model that can accommodate data limitations and speed in mapping flood-affected areas. Thus, this research is expected to make a significant contribution to improving our ability to accurately and efficiently map flood-affected areas, and to help reduce the losses that can be caused by such disasters.

The catastrophic flood on July 14, 2023 inundated 10 districts in Padang City with water levels reaching 2 meters [6, 7]. This flood was caused by extreme rainfall triggered by a convergence pattern, where air currents from the north and south of the Indian Ocean met and extended from the Mentawai Islands to the Natuna Islands [8]. As a result of the flood, there was extensive damage to infrastructure, and it caused the deaths of two toddlers while approximately 101 families, or 403 individuals, were displaced [9]. This catastrophic flood event serves as the subject of this study.

2 Methodology

2.1 Study area

The study area of this research covers the entire region of Padang City, which is the national activity center and the capital of West Sumatra Province. The city has an administrative area of 1,414.96 km², including a significant addition from the sea area. Located between 00°44'00" and 01°08'35" South Latitude and 100°05'05" and 100°34'09" East Longitude, Padang City has a coastline of 68.126 km and the Barisan Mountains covering an area of 486.209 km². In 2022, the population of Padang City reached 919,145 people, with a population density of 1,324 people/km² [10].

* Corresponding author: yrramadhan@maipark.com

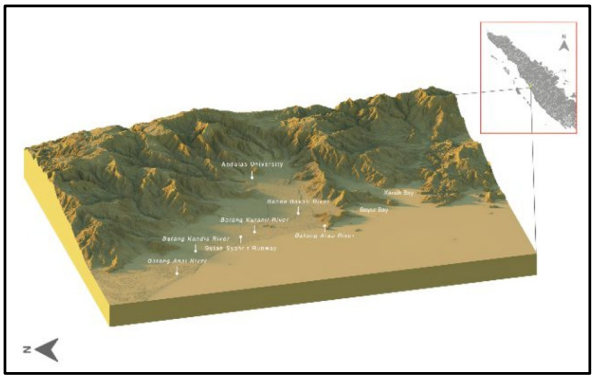

Fig. 1. Topographic map of Padang City.

Padang City is traversed by 23 rivers with a total length of 155.40 km, which are divided into 6 Watershed Areas (DAS), namely DAS Air Dingin, DAS Air Timbalun, DAS Batang Arau, DAS Batang Kandis, DAS Batang Kuranji, and DAS Sungai Pisang, (Fig. 1). These rivers generally have an elevation not much different from sea level, making many areas in Padang City prone to flooding, especially during high tides [11]. This situation is further exacerbated by the conversion of forest areas into residential and agricultural land, which reduces water absorption capacity, and the city's inadequate drainage system to handle heavy rainfall [12].

2.2 Data

In this research, we utilized several types of data to construct a rapid model of catastrophic floods. Firstly, we utilized an 8.5-meter resolution Digital Elevation Model (DEM) obtained from DEM Nasional (DEMNAS). This data was subsequently corrected using river network data from the Badan Informasi Geospasial (BIG) to produce HydroDEM.

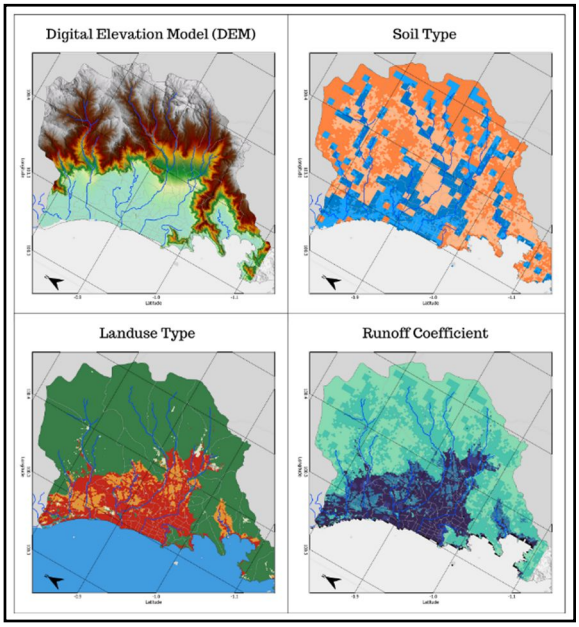

Fig. 2. Elevation data, soil type data, land use type data, and runoff coefficient data of Padang City.

Secondly, we employed meteorological data including rainfall, temperature, and wind speed. Rainfall data came from the Global Satellite Mapping of Precipitation (GSMaP) with a spatial resolution of 0.1 degrees, which was further corrected with rainfall data from the Meteorological Stations of the Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Temperature and wind speed data from BMKG Meteorological Stations were used to estimate evapotranspiration values.

Lastly, this study utilized soil data obtained from the Food and Agriculture Organization (FAO), as well as land use data with a spatial resolution of approximately 500 meters obtained from the Kementerian Lingkungan Hidup dan Kehutanan (KLHK). The incorporation of soil and land use data allows the model to simulate how different land covers and soil types influence water absorption and runoff, which is represented by the runoff coefficient. These data are shown in Fig. 2.

2.3 Method

2.3.1 DEM conditioning

DEM conditioning was performed to enforce the DEM to simulate flow conditions according to the actual river network. This process involved using several tools available in Archydro to modify the DEM, ensuring that elevation data reflects the actual topographic and river network conditions. This process is crucial because DEM data influences water movement in the analyzed area; thus, by conditioning the DEM, we can generate a flood model that is more accurate and relevant to real-world conditions.

2.3.2 Rainfall bias correction

GSMaP data is used to model and spatially distribute rainfall across the entire Watershed Area, (Fig. 3). However, GSMaP data often differs in values compared to direct surface observation data, necessitating a bias correction process. The method used for this correction is Quantile Mapping (QM) [13], which aims to align the spatial distribution of rainfall in GSMaP data to be more representative and consistent with more accurate surface observation data.

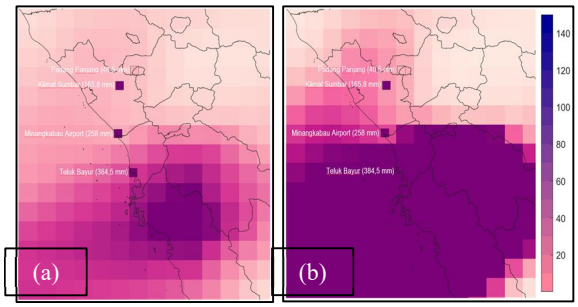

Fig. 3. GSMaP rainfall data. (a) before bias correction; (b) after bias correction.

The data is corrected using observation data from rain gauges at four BMKG stations located in Padang City and surrounding areas. By performing this correction, we can

obtain rainfall data that is more reliable for flood modeling.

2.3.3 Calculating evapotranspiration

Evapotranspiration in this study is calculated using the method recommended by the World Meteorological Organization (WMO) [14]. Based on the study by Celestin et al. (2020), which evaluated various simple empirical models for estimating evapotranspiration under different climatic conditions and data requirements, the WMO method has demonstrated excellent performance, making it a suitable alternative method. The calculation of evapotranspiration (ET0) using the WMO method is shown in the following formula:

$$ET0 = (1.298 + 0.934 \,.\, u_2) \,.\, (e_s - e_a) \quad (1)$$

, where u_2 is the wind speed measured at a height of two meters above ground level (m/s), e_s is the actual vapor pressure (kPa), and e_a is the saturation vapor pressure (kPa). $(e_s - e_a)$ represents the saturation vapor pressure deficit (kPa).

2.3.4 Determination of model domain

The model domain is determined by calculating the Topographic Wetness Index (TWI) [15] across the entire study area, in this case, Padang City as shown in Fig. 4. Areas identified by TWI as flood-prone will be selected as the model domain. The goal is to reduce the modeling domain as efficiently as possible so that the model can be produced in a shorter time. TWI for each grid data can be calculated based on elevation data using the following formula:

$$TWI = \frac{ln((\frac{TCA}{FW})}{tan(S))} \quad (2)$$

, where TCA is the total catchment area, FW is the flow width, and S is the slope. The following are the areas of Padang City identified by TWI as flood-prone areas, which were then used as the model domain in this study.

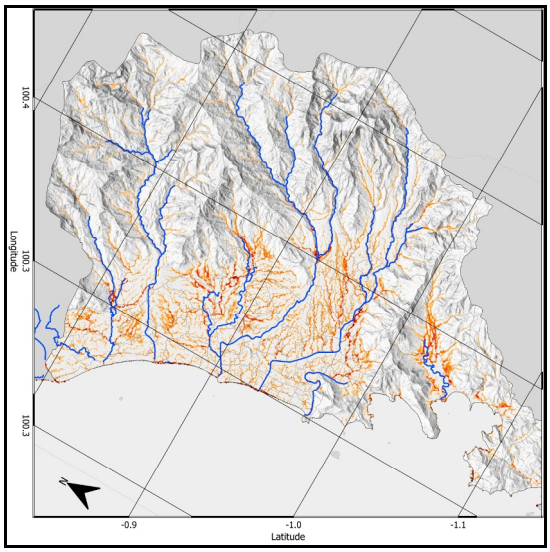

Fig. 4. Topographic Wetness Index (TWI) of Padang City.

2.3.5 Hydrological modeling

The hydrological model in this study refers to the Rational Method developed by Thomas Mulvaney in 1851 [16]. This method is widely used due to its simplicity and relatively low data requirements. Despite its simplicity, the method is quite accurate for estimating runoff discharge in response to rainfall, especially in small watershed areas [17]. Mathematically, the rational method can be written as follows:

$$Q = C \,.\, I \,.\, A \quad (3)$$

, where Q is the runoff discharge (m^3/s) C is the runoff coefficient, i is the rainfall intensity (mm/hour), and A is the area contributing to runoff discharge at a location (m^2). In its applicative context, to calculate runoff discharge at a data pixel, we sum up the runoff values from all contributing grids flowing into the target pixel, then multiply by the flow accumulation value at that pixel. Mathematically, this process can be expressed as:

$$Q = \Sigma R_o \,.\, Facc \,.\, \frac{1}{t};\; R_o = C \,.\, (P - ET0) \quad (4)$$

, where Q is the flow discharge at the target grid (m^3/s), ΣR_o is the sum of all contributing runoff to the target grid (m), F_{acc} is the flow accumulation value generated using Geographic Information System (GIS) software, and t equals 3600 (seconds), as the rainfall intensity used is per hour, hence t equals 1 = 3600/3600.

R_o is the difference between precipitation (P) and evapotranspiration (ET0) in meters, multiplied by the runoff coefficient (C).

2.3.6 Hydrodynamic modeling

Hydrodynamic modeling is based on the output of the hydrological model, specifically the runoff discharge at various locations along the river flow. These discharge values are divided by the cross-sectional area at each location to obtain the flow velocity of the river as shown in Fig. 5. The formula for calculating river flow velocity can be expressed as follows:

$$V = \frac{Q}{A} \quad (5)$$

, where V is the river flow velocity (m/s) at a cross section, Q is the river discharge at a cross section (m^3/s), and A is the cross-sectional area (m^2). Based on Bernoulli's principle, we can calculate the flow velocity at each cross section to determine the water level at the intersection points between the river layer and cross sections layer (river stations) using the following formula:

$$h = H - \left(\frac{V^2}{2g}\right) - z0 \quad (6)$$

, where h is the water height at the river station (m), H is the head energy (joule), V is the flow velocity (m/s), g is the gravitational acceleration (m/s^2), and z_0 is the elevation of the river bed (m).

Head Energy represents the total energy possessed by the fluid per unit weight, comprising kinetic energy head

and potential energy head. Mathematically, head energy can be written as:

$$H = KeH + PeH;\ KeH = \frac{1}{2}V^2;\ PeH = \Delta z\,.\,g \quad (7)$$

, where KeH is the kinetic energy head (joule) and PeH is the potential energy head (joule).

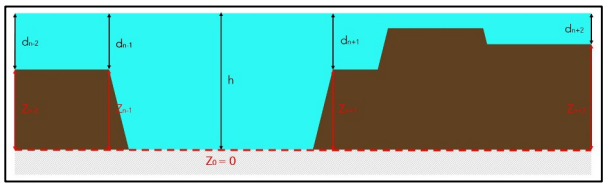

Fig. 5. Illustration of River Cross-section.

After obtaining the water heights at the river stations (h), the water heights at the intersection points between the bank line layer and cross sections layer ($d_{n\pm1}$), and at the intersection points between the cross section and the boundary of the model domain ($d_{n\pm2}$) can be calculated using a "bathtub model" approach with the following formula:

$$dn = h - zn \quad (8)$$

Following this process, we will have five water height data points at each cross section. These data are then interpolated across the entire TWI area calculated in the previous stage. Thus, we can generate a flood inundation map across the entire TWI area (model domain).

In the modeling results, sometimes separate inundation areas are produced apart from the main inundation areas (areas connected to the river). These separate areas are considered modeling errors and need to be removed.

2.3.7 Model output verification

In this study, we collected data on flood-affected locations through mass media and social media. We managed to gather 72 pieces of information about flood locations, 57 of which were quite specific, including residential names, street names, neighbourhood units (RT/RW), or certain landmarks, while the rest only included location information up to the sub-district or district level.

When the location information was very limited, we placed the geolocation point in the most likely area to be affected by floods. Of the 57 data points with specific location information, 17 included information on the flood inundation depth at that location.

We used all of this flood location information to qualitatively verify the model by calculating the Probability of Detection (POD). POD is a value that indicates how well a model can accurately predict floods. POD ranges from 0 to 1, with a value of 1 indicating perfect model accuracy. POD can be calculated using the following formula:

$$POD = \frac{Hits}{Hits+Misses} \quad (9)$$

"Hits" represent conditions where both the simulation results and actual conditions indicate the presence of flood inundation. "Misses" represent conditions where the simulation results indicate no flood inundation, while the actual conditions show the presence of flood inundation.

In addition to qualitative verification by calculating POD, we used the 17 pieces of flood location information containing inundation depth values to calculate the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) of the simulated flood depths (Fig. 6).

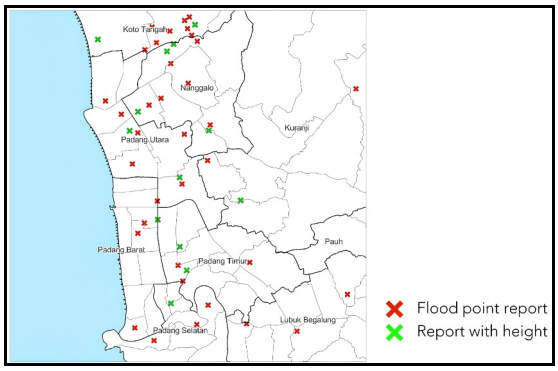

Fig. 6. Geolocated flood-affected locations data.

2.3.8 Model output calibration

Calibration was performed by calculating the ratio between observed flood heights and model output flood heights in grids where flood height observations were available. Each grid within a radius of 250-500 meters connected by a network order of the same level was corrected based on the flood height ratio of that grid (Fig. 7).

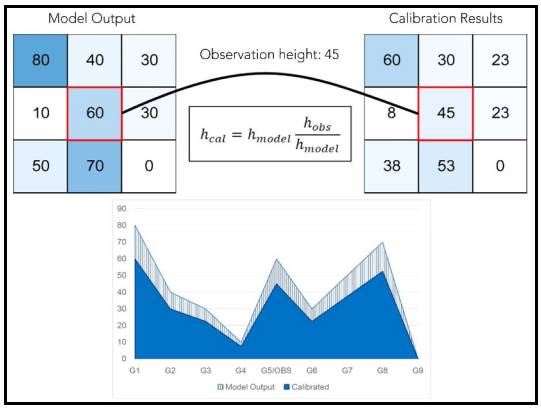

Fig. 7. Illustration of Calibration Method.

The stream order level was determined by calculating the stream order (Strahler number) as shown in Fig.8 [18].

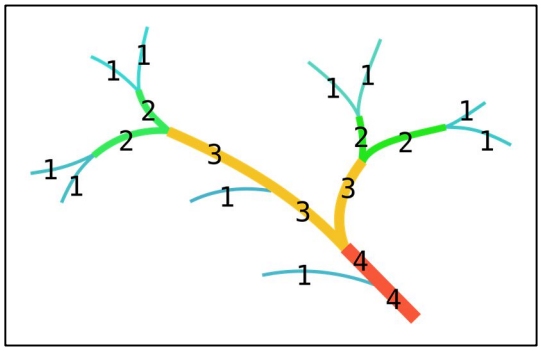

Fig. 8. Illustration of Stream Order [18]

The abundance of observational data plays a crucial role in the verification and calibration process. The contribution of the community in reporting flood inundation data, including specific details such as location, time, and flood height through social media or websites, is highly valuable for testing and improving the accuracy of the model. The more extensive and specific the flood inundation data collected and utilized, the greater the potential to produce a model that more accurately represents real-world conditions.

3 Results and discussion

Figure 9 shows the flood prediction model results overlaid with field verification data (red dots) and river flow lines (blue lines). The flood prediction model in the image reveals a pattern of inundation that is not concentrated in a single location but rather fragmented and dispersed across various areas. This pattern indicates that flooding in the region is not caused by a single source but by various local factors, such as topography, tributary flow, and drainage capacity in specific areas.

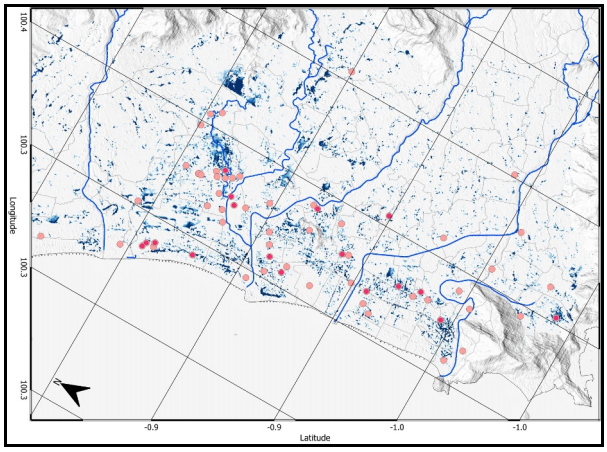

Fig. 9. Output model overlayed with verification data and river lines.

Table 1 presents the model verification results using actual data, with three main evaluation parameters. The Probability of Detection (POD) of 0.778 indicates that the model accurately mapped flood inundation in 77.8% of the 72 evaluated samples. The Root Mean Square Error (RMSE) of 38.093 indicates an average prediction error of approximately 38.093 cm across 17 samples, while the Mean Absolute Error (MAE) of 27.584 indicates an average absolute model error of approximately 27.584 cm across 17 samples.

Table 1. Model Output Verification Results

Parameter	Value	Sample Size
POD	0.778	72
RMSE	38.093	17
MAE	27.584	17

Model prediction errors, represented by RMSE and MAE values, are still quite significant. Therefore, verification was conducted to obtain a more representative model. Figure 10 shows a comparison of flood inundation maps before and after the calibration process.

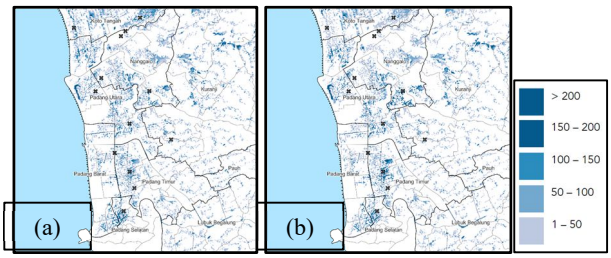

Fig. 10. Model Output Results. (a) before calibration; (b) after calibration.

The model results before and after calibration show a relatively similar flood extent. However, before calibration, the model tended to overestimate flood depth. The calibration process, which adjusted the model's flood depth based on observational data, resulted in a more realistic flood depth prediction.

Nonetheless, this calibration was conducted using only 17 observational flood depth samples, which is limited for representing flood events across the entire watershed. Although the model's depth error is reasonably good, there is still potential for more accurate results if supported by a larger and more evenly distributed set of observational data throughout the watershed area.

4 Conclusion

This study developed a rapid model for mapping flood-affected areas, using a case study of the flood in Padang on July 14, 2023. The model is designed to address data limitations and enhance the speed of flood-affected area mapping, which is crucial in disaster situations to mitigate losses. The research methodology includes the use of various types of data such as DEM, rainfall, evapotranspiration, soil type, and land use. The model process involves DEM reconditioning, rainfall bias correction, evapotranspiration calculation, determination of the model domain using the Topographic Wetness Index (TWI), as well as hydrological and hydrodynamic modeling. The model verification was conducted using flood location data collected from mass media and social media, showing that the model has an accuracy of 77.8% in mapping flood areas. Although there were significant prediction errors, the model results after calibration showed a significant increase in accuracy.

References

1. Indonesian National Board for Disaster Management (BNPB). National Disaster Risk Index. (2020).
2. Bubeck, P., Otto, A., and Weichselgartner, J. Societal impacts of flood hazards. Oxford Research Encyclopedia of Natural Hazard Science. (2017).
3. Supari, S., and Sofian, S. Flood early warning system in Indonesia: Status and challenges. Procedia Engineering, **212**, 444-451. (2018).

4. Bates, P. D., and De Roo, A. P. J. A simple raster-based model for flood inundation simulation. Journal of Hydrology, **236**(1-2), 54-77 (2000). https://doi.org/10.1016/S0022-1694(00)00278-X
5. Kalyanapu, A. J., Patil, V., and Ugwonali, O. A. High-performance computing framework for floodplain hydrodynamic modeling. Journal of Hydrologic Engineering, **21**(6), 04016010. (2016).
6. Hamzah, F. (2023, July 14). Heavy rain since yesterday, Padang City flooded. Tempo. Retrieved October 31, 2023, from https://www.tempo.co/read/hujan-deras-sejak-kemarin-kota-padang-terendam-banjir
7. Sastra, Y. (2023, July 14). Heavy rain since yesterday, Padang City hit by floods and landslides. Kompas. Retrieved October 31, 2023, from https://www.kompas.com/hujan-deras-sejak-kemarin-kota-padang-dilanda-banjir-dan-longsor
8. Antara. (2023, July 14). BMKG Jelaskan Penyebab Hujan Ekstrem di Balik Banjir dan Longsor Sumbar. DetikNews. Retrieved July 27, 2023, from https://news.detik.com/berita/d-6822313/bmkg-jelaskan-penyebab-hujan-ekstrem-di-balik-banjir-dan-longsor-sumbar
9. Badan Nasional Penanggulangan Bencana. (2023). Statistik bencana Indonesia. Retrieved July 27, 2023, from https://dibi.bnpb.go.id/home/index2
10. Padang City Government. General Overview of Padang City. Retrieved November 1, 2023, from https://padang.go.id/gambaran-umum-kota-padang
11. Zulvi, Y. Y., and HAR, R. Kajian laju infiltrasi akhir pada Daerah Aliran Sungai (DAS) Batang Arau Kota Padang ditinjau dari perbedaan litologi, tata guna lahan, dan sifat fisik tanah. Jurnal Bina Tambang, **3**(3), 1067. ISSN: 2302-3333. (2024).
12. Ikhvan, A., and Mera, M. Case Study: Significant factors in hazard and vulnerability assessments in flood mitigation in Padang City. IOP Conference Series: Earth and Environmental Science, **708**(1), 012026 (2021). https://doi.org/10.1088/1755-1315/708/1/012026
13. Panofsky, H. W., and Brier, G. W. Some Applications of Statistics to Meteorology. The Pennsylvania State University Press, Philadelphia. (1968).
14. World Meteorological Organization (WMO). (1966). Measurement and Estimation of Evaporation and Evapotranspiration. Technical Note No.83; Report of a Working Group on Evaporation Measurement of the Commission for Instrument and Methods of Observation. WMO: Geneva, Switzerland.
15. Beven, K. J., and Kirkby, M. J. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Journal, **24**, 43–69 (1979). https://doi.org/10.1080/02626667909491834
16. Mulvaney, T. On the use of self-registering rain and flood gauges in making observations of the relations of rainfall and flood discharges in a given catchment. Proceedings of the institution of Civil Engineers of Ireland, **4**(2), 18-33 (1851).
17. Grimaldi, S. and Petroselli, A. Do we still need the Rational Formula? An alternative empirical procedure for peak discharge estimation in small and ungauged basins. Hydrological Sciences Journal, **60**(1), 67–77 (2014). https://doi.org/10.1080/02626667.2014.880546
18. Strahler, A. N. Hypsometric (height-area) analysis of erosional topology. Bulletin of the Geological Society of America, **63**(11), 1117–1142 (1952). https://doi.org/10.1130/0016-7606(1952)63[1117]2.0.CO;2