

Original Article

Leveraging Python and Google Earth Engine for Spatiotemporal Analysis to Develop a Synthetic Rainfall Distribution Pattern Using GSMaP Data

Nezar Farrag¹, Ashraf M. El-Moustafa¹, Morad H. Abdelsalheen¹

¹Department of Irrigation and Hydraulics, Faculty of Engineering, Ain-Shams University, Cairo, Egypt.

¹Corresponding Author : nezar.farrag20@gmail.com

Received: 07 February 2025

Revised: 09 March 2025

Accepted: 09 April 2025

Published: 30 April 2025

Abstract - Accurate representation of rainfall patterns is crucial for water resource management and flood mitigation, particularly in arid regions susceptible to flash floods. This study presents a novel methodology for generating synthetic rainfall distributions utilizing the GSMaP dataset to address the limitations of conventional hyetograph representations. The proposed algorithm comprises two stages: 1) Preprocessing hourly GSMaP precipitation data and employing the DBSCAN algorithm to identify individual storms; 2) Applying the Alternating Block Method to reorganize rainfall depths and generate a dimensionless synthetic distribution. A Python algorithm was developed to automate this entire process. The methodology was evaluated using a four-day storm in South Sinai, showing a strong similarity to the SCS Type-II distribution but revealing significant differences in hydrological modeling for Wadi Fieran. Specifically, the GSMaP distribution reduced peak discharge by 12% and increased flood volume by 0.27%. These findings highlight the utility of satellite-based precipitation data in enhancing hydrological simulations.

Keywords - Flood Hazard, Remote Sensing, Cloud Computing, Arid Climate, Precipitation, Synthetic Hyetograph, Sinai.

1. Introduction

Understanding and accurately modeling rainfall distribution is crucial for effective hydrological simulations, flood risk assessments, and water resource management, particularly in arid and semi-arid climates like the MENA region. Given the occurrence of multiple extreme events across the region, there is a pressing need to reevaluate and refine the existing synthetic rainfall distribution pattern.

This research focuses on a case study of a significant four-day storm event chosen for its potential to produce runoff. By systematically analyzing the spatial and temporal characteristics of this storm, the study seeks to demonstrate the utility of GSMaP data and Python-based processing in generating reliable synthetic rainfall patterns that can be used in hydrological modeling. The outcomes of this research are expected to contribute to the improvement of flood forecasting and risk assessment, as well as water resource management strategies in similar geographic contexts. It is worth mentioning that the Python algorithm developed in this research is intended to be broadly applicable and adaptable to any geographical region and any temporal period of interest.

The problem arises from the common practice of hydrologists using widely recognized synthetic rainfall distributions, such as the SCS Type II, in hydrological

simulations when a region-specific distribution is unavailable. However, in arid and semi-arid climates, the application of the SCS Type II rainfall distribution often results in a significant overestimation of runoff peak flows. This discrepancy stems from mismatches in temporal rainfall patterns, emphasizing the need for region-specific rainfall distributions or calibrated modifications to the SCS framework [1-3].

The research gap stems from the type of raw data used and the lack of spatiotemporal analysis capable of simulating storm movement. Traditional approaches for developing rainfall distribution patterns primarily rely on historical ground-based observations [4-6] However, these methods are often constrained by limited spatial coverage and temporal resolution, making it challenging to capture the dynamic nature of rainfall events accurately.

El-Sayed, 2018 developed rainfall distribution curves for the Sinai Peninsula using 127 recorded storms from 12 rain gauges. Storms were classified into four duration-based groups, and dimensionless hyetographs were created by centering the maximum rainfall depth. Average hyetographs and cumulative design curves were then generated. The WRRI distribution curves showed higher peak discharges than SCS profiles, especially for short storms, highlighting the importance of precise rainfall pattern modeling.



In recent years, advancements in remote sensing technology have provided alternative data sources, such as the Global Satellite Mapping of Precipitation (GSMaP) dataset, which offers high-resolution global precipitation measurements [7]. These datasets present an opportunity to enhance the precision and reliability of rainfall distribution models, especially in regions where ground-based hourly rainfall data is sparse or unavailable. This study aims to leverage the Python language to develop a synthetic rainfall distribution pattern using GSMaP data specifically tailored for hydrological simulations in arid and semi-arid regions. The methodology was applied to a significant storm event over South Sinai, and the resulting distribution was compared with the SCS type II distribution. Subsequently, both synthetic distributions will be utilized on a hydrological model for Wadi Fieran, and the variation in results will be demonstrated.

Ground-based precipitation measurements often encounter significant challenges and constraints, particularly in regions where gauge networks are sparse or absent altogether. For the purpose of generating synthetic rainfall distribution patterns, a large, hourly-resolution historical dataset is crucial. Thus, it becomes essential to identify dependable alternatives, such as satellite-based precipitation products offering continuous time series data and extensive global spatial coverage [8]. Satellite-based precipitation datasets have been developed to achieve higher spatial and temporal resolutions using combined data from Passive Microwave (PMW) sensors in low Earth orbit and Infrared (IR) radiometers in geostationary Earth orbit. Global Satellite Mapping of Precipitation (GSMaP) is a blended PMW-IR precipitation product. It has been developed in Japan for the Global Precipitation Measurement (GPM) mission [9] as the Japanese GPM standard product.

Recent studies have rigorously assessed the accuracy of the GSMaP dataset. For example, [10] demonstrated that GSMaP effectively captures the spatial patterns of summer precipitation over the United States. Similarly, [11] found that GSMaP exhibited considerable skill in the Indonesian Maritime Continent when compared against data from 152 gauge stations. Several studies have recently evaluated the performance of satellite precipitation products in Egypt. [12] analyzed the accuracy of five products (ARC, CHIRPS, GSMaP, TAMSAT, and PERSIANN-CCS) using daily gauge data from 30 stations over the period 2003-2018. The study concluded that GSMaP, followed by ARC, demonstrated strong performance, particularly for rainfall events with intensities of 1 mm/day or greater. As a result, the study recommended GSMaP for hydrological research in Egypt.

Another study focused on the performance of three prominent global satellite precipitation products: (PERSIANN-CDR), (TRMM3B42V7), and (GSMaPV6). This study evaluated their accuracy against daily gauge data from 23 stations in Egypt over the period 2003-2014 at both

daily and annual scales. The findings indicated that GSMaPV6 is the most suitable for hydrological applications in Egypt. Based on an extensive review of the literature on Satellite-Based Precipitation Data, GSMaP Version 6 (GSMaP V6) has been identified as the most suitable dataset for this study [12]. The GSMaP V6 is distinguished by the following characteristics [7]

- **Producer:** The GSMaP project was initiated and sponsored by the Japan Science and Technology Agency under the Core Research for Evolutional Science and Technology in November 2002 [13]
- **Spatial Resolution:** The data has a spatial resolution of approximately $0.1^\circ \times 0.1^\circ$, corresponding to about 11 km at the equator.
- **Temporal Resolution:** The dataset provides precipitation estimates at a 1-hour temporal resolution, making it suitable for capturing short-term precipitation events.
- **Coverage:** The GSMaP product covers the global region from 60°N to 60°S latitude.
- **Data Type:** This is an operational version, meaning it provides near-real-time precipitation estimates. The product includes both gauge-calibrated and non-gauge-calibrated versions, with the operational product often used for real-time monitoring.
- **Algorithm:** The GSMaP algorithm uses data from multiple satellite sensors, including passive microwave PMW radiometers, to estimate precipitation. The algorithm is regularly updated and calibrated to improve accuracy [14].
- **Version:** The dataset is version 6 (v6), indicating the iteration of the algorithm and processing chain used to generate the data, with improvements over previous versions in accuracy and coverage.

Figure 1 and Figure 2 depict an example of the output raster generated by GSMaP, capturing extreme storm events recently experienced in the Middle East, including the Jeddah storm of November 2022 and the Derna storm of September 2023.

2. Methodology

The methodology applied in this work is structured into two main phases: the initial phase involves the extraction of GSMaP data from GEE, while the second phase focuses on the analysis of this data to generate a synthetic rainfall distribution pattern. The process of extracting GSMaP data from GEE involves specifying both the temporal range and the geographic region of interest, which are key inputs determined by the user. The Python algorithm is programmed to automatically clip the extensive raw GSMaP dataset to the defined geographic area and filter it according to the selected time window, enabling further analysis. Accordingly, a prominent region and a significant storm event were selected to implement the methodology in this study. A brief description of the selected area and storm event is provided in subsequent sections.

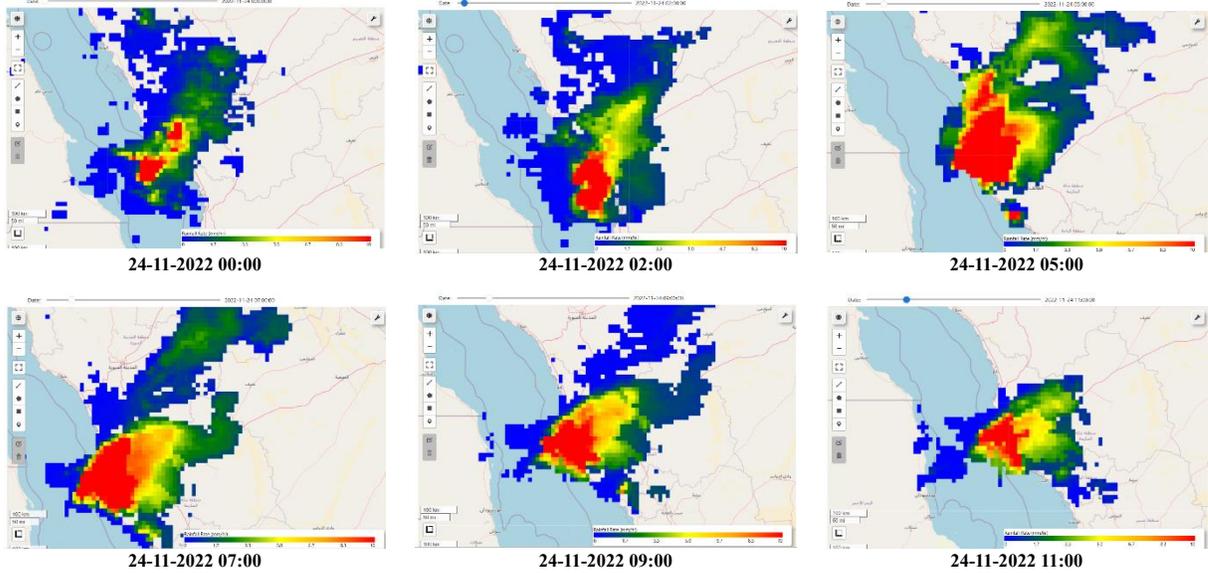


Fig. 1 Example of the output raster provided by GSMaP for the jeddah storm of november 2022

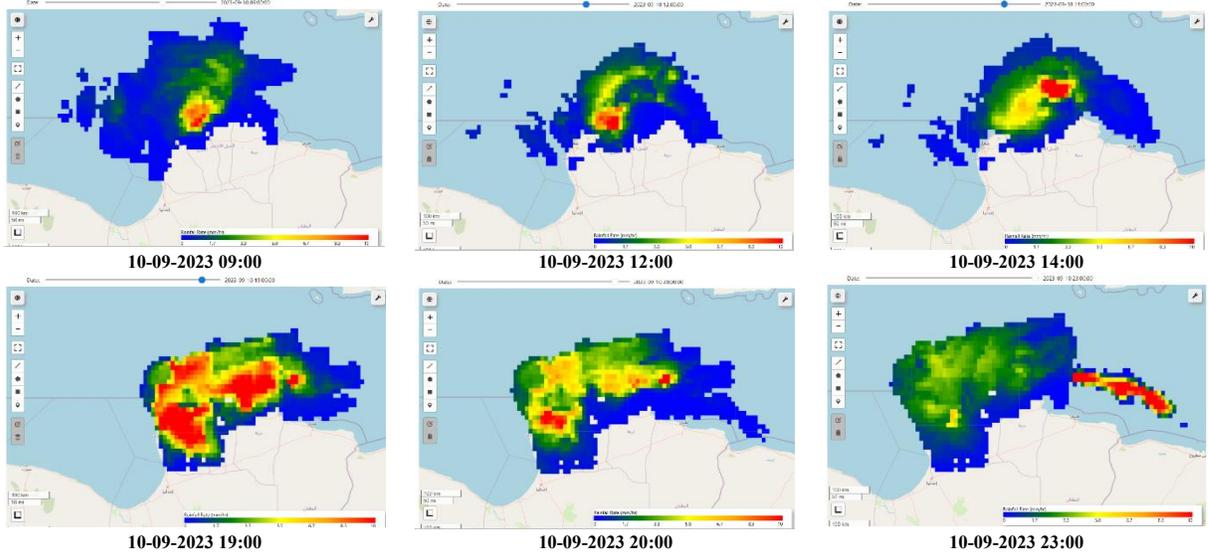


Fig. 2 Example of the output raster provided by GSMaP for the derna storm of september 2023

2.1. Extracting and Processing The GSMaP Data From GEE

The methodology employed in this study involves utilizing the Global Satellite Mapping of Precipitation (GSMaP) dataset to analyze precipitation patterns within a defined geographic region over a specific time period. A Python script was created in the Jupyter Hub environment to extract and process the necessary rainfall data. The workflow of the script is illustrated in Figure 3 and outlined as follows:

1. Initialization: The script begins by importing necessary libraries, including ee for Earth Engine operations, pandas and numpy for data manipulation, DBSCAN from sklearn. cluster for clustering, geopmap and ipyleaflet for geospatial visualization, and matplotlib for plotting.
2. Earth Engine Initialization: The Earth Engine API is initialized using ee.Initialize(), enabling access to satellite data and geospatial processing capabilities.
3. Defining the Study Area: A polygon geometry is defined by specifying the coordinates of its vertices. This polygon delineates the area of interest for precipitation analysis.
4. Temporal Filtering: The time period for analysis is set between March 13, 2020, and March 14, 2020. The GSMaP V6 dataset is filtered to include only the images captured during this period.
5. Data Processing:
 - Hourly Data Extraction: The script iterates over each image in the filtered dataset, extracts hourly precipitation data, and clips the data to the defined polygon.

- Maximum Pixel Value Calculation: For each clipped image, the maximum pixel value is computed. This value represents the highest recorded precipitation rate within the polygon.
 - Maximum Pixel Localization: A mask is created for pixels with the maximum value, and the centroid coordinates of these pixels are determined.
6. Data Aggregation: Results, including date-time, maximum pixel value, and the coordinates of the maximum pixel, are aggregated into a list and converted into a pandas DataFrame.
 7. Clustering Analysis: The DBSCAN clustering algorithm is applied to the coordinates of the maximum pixel values.
8. Advanced Rearrangements and Comparisons:
 - Rearranging rainfall values using the Alternating Block Method (ABM) to simulate synthetic distribution patterns.
 - Allows for visual and statistical comparison of the original and rearranged datasets.
 9. Accumulation Analysis: Cumulative rainfall depth and time are calculated using Pandas operations.
 10. Dimensionless Analysis: Calculate normalized (dimensionless) rainfall and time; matplotlib is used to plot Dimensionless rainfall vs. dimensionless time.

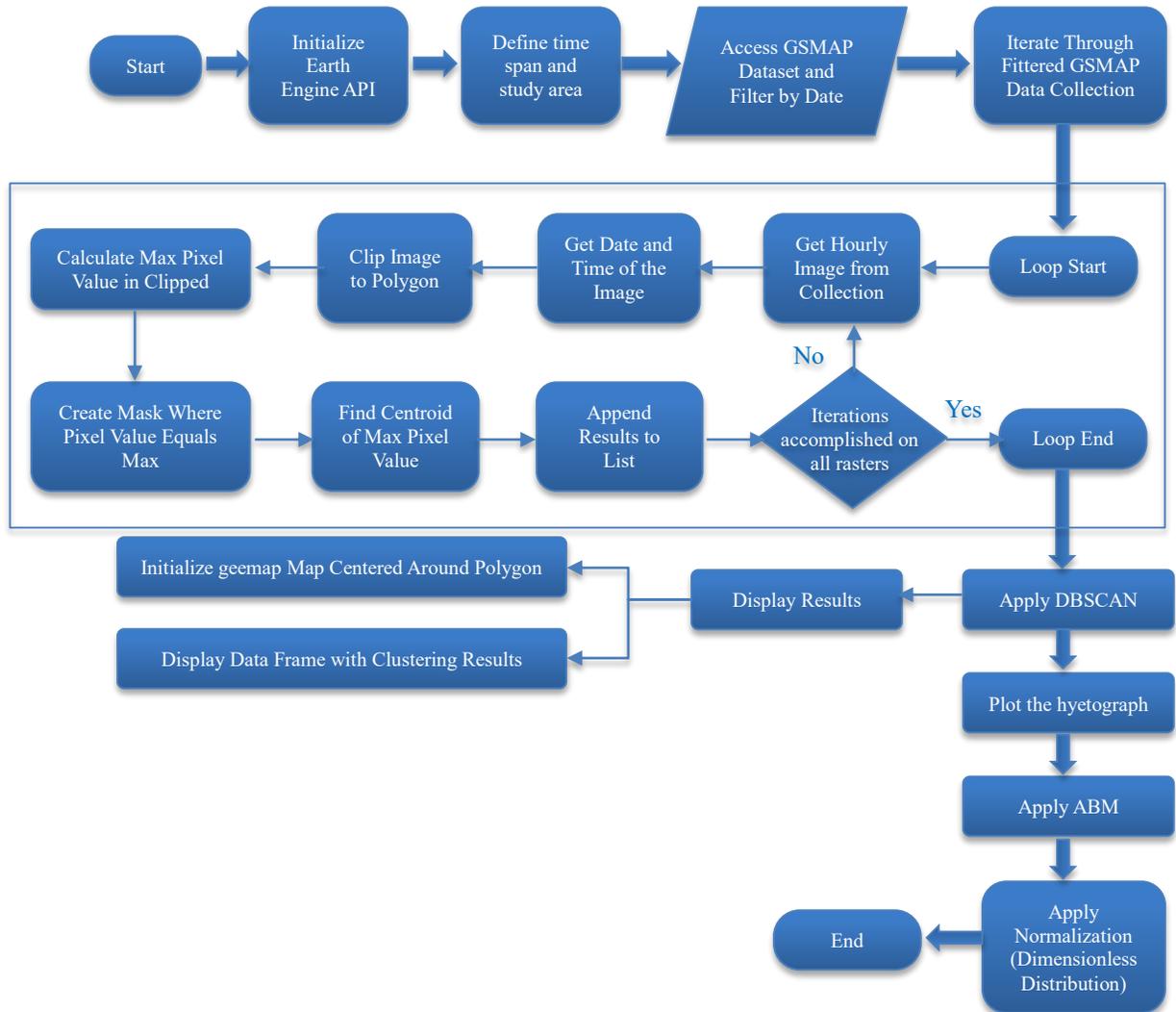


Fig. 3 The workflow of python algorithm

Given that the satellite product supplies rainfall data in raster format with an hourly time resolution, this allows us to track the movement of each storm’s core and monitor changes in location and intensity over the storm’s duration. The methodology centers on extracting the maximum pixel value

from each raster to construct the corresponding hyetograph (Figure 4). Therefore, the Python algorithm is designed to identify pixels with the highest precipitation rate and to determine their geographic coordinates, which is an essential step in applying the DBSCAN algorithm.

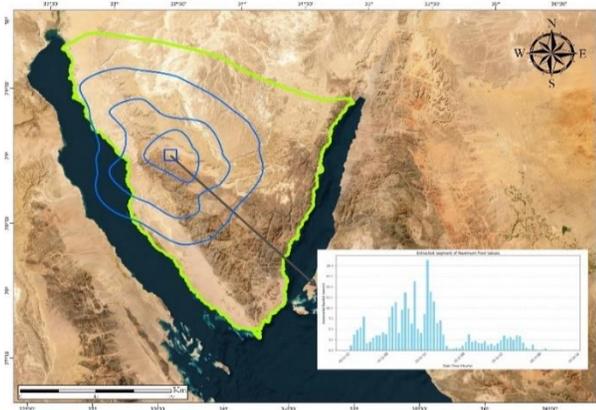


Fig. 4 Extracting the maximum pixel value from each raster to construct the corresponding hyetograph

2.2. DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

The algorithm’s methodology involves extracting the maximum pixel value from each raster across South Sinai. Confirming that each maximum pixel corresponds to the same storm is crucial. To identify cases where a different storm temporally overlaps with the original storm (Figure 5), an alert mechanism (DBSCAN algorithm) is employed to address this issue.

2.2.1. Technique of DBSCAN

The DBSCAN algorithm identifies clusters in a dataset by evaluating the density of data points. It defines clusters as areas of high point density separated by low density areas. The algorithm uses two parameters: epsilon (ϵ), the maximum distance between points to be considered neighbors, and minPts, the minimum number of points required to form a dense region. DBSCAN can detect irregularly shaped clusters and effectively specify noise (outliers) in the data, which is, in this case, the potential storm that might temporarily intersect with the original one [16].

2.2.2. Motivation

The condition of temporal overlap between storms, as illustrated in Figure 5, was the motivation for applying the DBSCAN algorithm. This issue happens when a subsequent storm enters the study area and contains a pixel with a value exceeding the maximum pixel value of the core of the original storm, causing the hyetograph to record it as part of the same event. By applying DBSCAN, the algorithm identifies multiple clusters (which, in this context, corresponds to the number of distinct storms at any time of interest), assigning the number of clusters corresponding to each record in the resulting table (Table 1). So, It acts as a warning system, alerting users to this condition so they can manually check and assign the correct value. Conversely, if the cluster field in the resulting table shows only the value “1,” it indicates that the entire run consists of a single storm, allowing the analysis to proceed without further adjustments.

This phenomenon has a low probability of occurrence, but it should be solved using the Python algorithm to make it robust and ready to be used by any hydrologist.

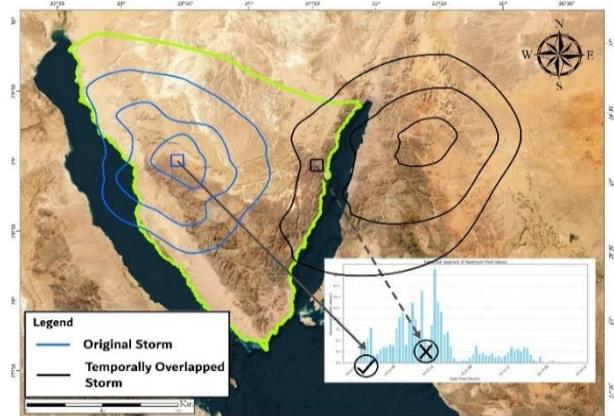


Fig. 5 A different storm temporally overlaps with the original storm

Table 1. The result of The DBSCAN algorithm

No.	date time	max pixel value	max pixel lat	max pixel lon	cluster
0	3/12/2020 19:00	3.638421297	29.91318784	32.78094697	1
1	3/12/2020 20:00	0.968537748	28.55074299	33.24657372	1
2	3/12/2020 21:00	0.297517508	29.83012362	32.66823834	1
3	3/12/2020 22:00	0.349309385	29.77095458	32.68263135	1
4	3/12/2020 23:00	0.525621474	29.64810603	32.69481902	1
...
25	3/13/2020 20:00	0.977441251	29.77095458	32.68263135	1
26	3/13/2020 21:00	1.834030271	29.64810603	32.69481902	1
27	3/13/2020 22:00	3.814805031	29.77095458	32.68263135	1
28	3/13/2020 23:00	1.906563282	29.27143684	32.88683928	1
29	3/14/2020 00:00	2.111501217	29.35024359	32.95012968	1

The output of DBSCAN Algorithm

Applying this algorithm to the study area during the time span of the March 2020 storm, it was evident that the extracted rainfall values correspond to a single, coherent storm in Table 1.

2.2.3. Applying the Alternating Block Method (ABM)

This approach is based on rearranging the rainfall depths over the entire storm duration so that the maximum rainfall depth occurs in the middle of the entire storm duration while the second greatest value occurs at the one-time step after the occurrence of the maximum one and so on [16], (Figure 6). The idea of alternating high- and low-intensity blocks reflects the natural variability of rainfall events, where high intensities are usually interspersed with lighter or no rain and intensities; hence it is applicable in a wide range of hydrological applications. Compared to more sophisticated methods like stochastic models or those based on complex meteorological theories, ABM requires less extensive rainfall data compared to some other methods. The Alternating Block Method (ABM) has notable limitations in hydrological applications. It often overestimates rainfall intensities, particularly in scenarios with repeated rainfall events, leading

to inaccuracies in hydrological assessments. While considered relatively simple compared to other techniques, it still requires a solid understanding of its methodology, which can be challenging for less experienced users. Additionally, ABM may not perform reliably in regions with unique rainfall patterns, limiting its applicability in diverse hydrological conditions. Its effectiveness also depends heavily on the quality and accuracy of the rainfall data used, with insufficient or inaccurate data potentially resulting in outputs that do not accurately reflect actual storm behavior. While ABM has its advantages, it is essential to compare its results with those obtained from other methods, such as Euler Type II and Huff’s curves [6] coherently studied this area, the author evaluated six approaches to develop the design of storm hyetographs and concluded that the ABM is the recommended approach to use in developing the design of storm hyetographs in all flood hazards and stormwater drainage studies. A function was conducted to automatically apply the Alternating Block Method (ABM) in the Python algorithm. Figure 6 illustrates an example of the Alternating Block Method from the literature [5].

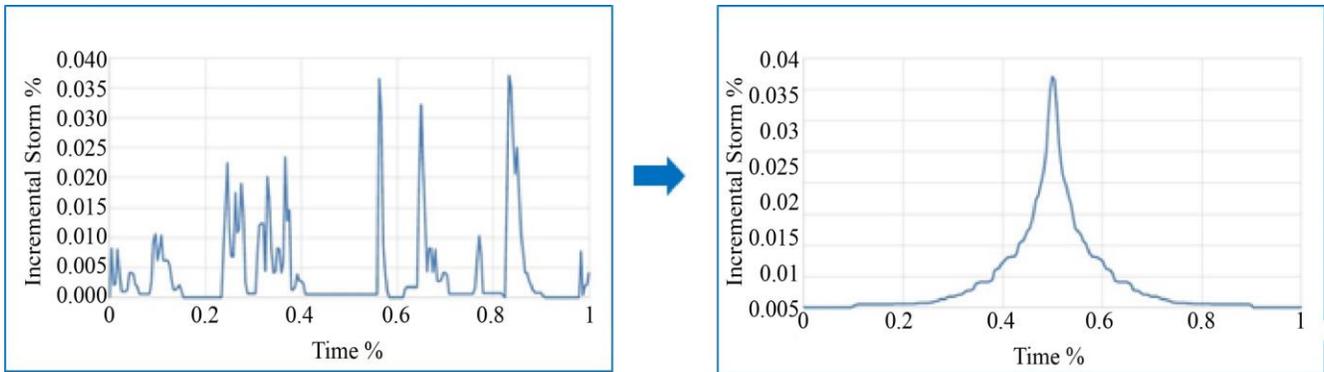


Fig. 6 Example of the alternating block method from the literature [5]

2.2.4. Developing of Synthetic Rainfall Distribution Pattern.

As the purpose of generating the synthetic rainfall distribution pattern is to be used in hydrological simulations and flood studies, only storms that could produce runoff are considered in further analysis. Took 10 mm as a threshold for runoff studies for Sinai Wadis.

So, only storms that have an accumulated rainfall depth equal to or more than 10 mm have been considered. After applying the ABM, a mass curve was created by plotting accumulated rainfall depth over time. To normalize this curve, the rainfall depth at each time was divided by the storm’s total rainfall, and the time elapsed was divided by the storm’s total duration, resulting in a dimensionless mass curve for easy comparison across storms [17]. The equations below show the normalization calculations.

$$R_{Normal}(t) = \frac{R(t)}{R_{total}} \quad \text{And} \quad T_{Normal}(t) = \frac{t}{T_{total}}$$

Where:

- $R_{Normal}(t)$ is the normalized (dimensionless) rainfall depth at time t ,
- R_{total} is the total rainfall over the storm duration,
- $T_{Normal}(t)$ is the normalized time,
- t is the time elapsed, and
- T_{total} is the total storm duration.

This normalization ensures that both rainfall depth and time are scaled from 0 to 1.

The Study Area

The South Sinai region, located between latitudes 32°30’00” N and 35°00’00” N and longitudes 27°30’00” E and 30°00’00” E, will be used as a pilot area to conduct the methodology. The geographic location and boundaries of South Sinai are depicted in the accompanying Figure 7. This study will use the prominent rainfall event that occurred in March 2020 to test the code, specifically from 13:00 on March 11 to 08:00 on March 14, 2020.

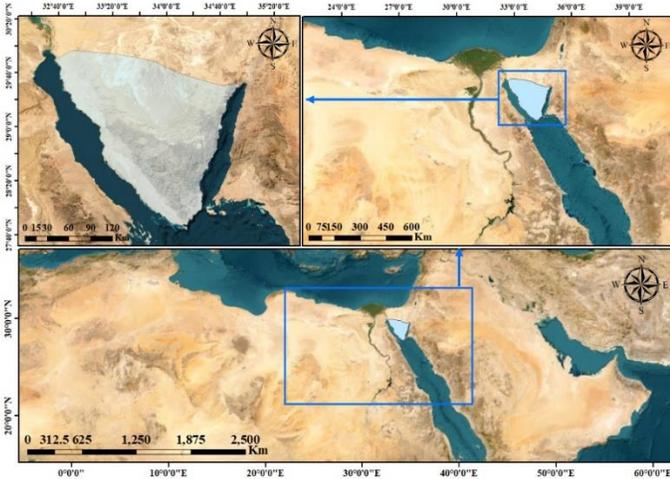


Fig. 7 The geographic location and boundaries of South Sinai

The Sinai Peninsula in Egypt is situated within an arid climatic zone that extends across northern Africa and southwestern Asia. Despite the prevalent dryness, the region occasionally experiences intense precipitation events. South Sinai, particularly characterized by its numerous extensive wadis, such as Wadi Feiran, Wadi Watir, and Wadi Dahab, has the potential to generate substantial flood volumes.

Historical records indicate that South Sinai has experienced several flash floods, leading to significant infrastructural damage, displacement of populations, and occasionally fatalities [18]. Despite their hazardous effects, flash floods in Sinai offer a potential source of unconventional freshwater. To effectively manage the impacts of flash floods and optimize the capture and utilization of this precious water source, it is crucial to understand and accurately model rainfall distribution.

The Study Storm

Egypt experienced one of the most severe meteorological events in over two decades, beginning late on March 11, 2020, with heavy rainfall persisting for two consecutive days, Nicknamed “the dragon” on social media. This atmospheric disturbance resulted in widespread flooding, severe winds, and sandstorms.

Meteorological instability has not been observed at this intensity since 1994. In response, the government advised citizens to remain indoors and enacted the closure of major interprovincial highways, disrupting governmental, public, and private sector activities.

The adverse conditions, characterized by torrential rains, powerful winds, and thunderstorms, led to extensive flooding across the nation, with a reported death toll exceeding 40 individuals. The Ministry of Social Solidarity (MoSS) confirmed that 10 fatalities and over 400 injuries occurred in Cairo alone, while Qena Governorate saw 3 deaths and 5 injuries. Additional casualties were reported in Giza, Ismailia,

Sharkeia, New Valley, Menofia, and South Sinai Governorates, with 12 individuals still missing. Furthermore, nationwide rail services were suspended after heavy rains caused a train collision in northern Giza, injuring 13 passengers [19]. This event underscores the significant vulnerability of infrastructure and emergency systems in Egypt to extreme weather phenomena, highlighting the need for improved resilience measures.

Figure 8 shows the map of accumulated multi-satellite precipitation with gauge calibration – Final Run [GPM_3IMERGHH v07] mm for the three days of the storm all over Egypt [20].

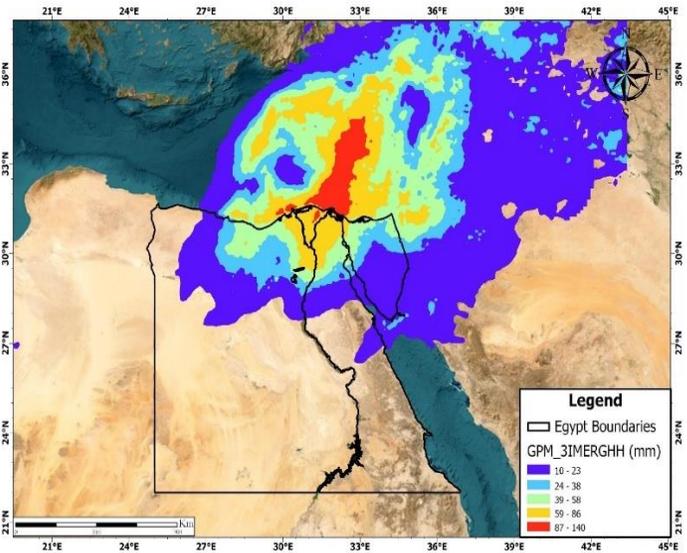


Fig. 8 The map accumulated multi-satellite precipitation with gauge calibration - Final Run for the severe storm of March 2020 all over Egypt [20]

3. Results and Discussion

The Python algorithm was applied to generate the hyetograph for the whole month of March 2020, and a case study of a prominent rainfall event that occurred specifically from 13:00 on March 11 to 08:00 on March 14 was selected to develop a synthetic rainfall pattern for South Sinai.

The chart below presents the resulting hyetograph (in blue bars) (Figure 9), while Table 2 provides a summary of key statistical characteristics for the case study storm event.

As evident from Table 2 , the majority of the storm’s rainfall was concentrated in the first half of the event’s duration, with more than half of the total precipitation occurring during the first third of the storm’s timeframe.

As previously mentioned in the methodology, the Alternating Block Method (ABM) was applied to rearrange the incremental hyetograph. Figure 9 shows the comparison between the resulting rearranged bar chart and the original hyetograph.

Table 2. Statistical characteristics of the resulting hyetograph

Total Rainfall	264.15 mm
Mean Rainfall	3.77 mm
Maximum Rainfall	21.25 mm at 2020-03-12 14:00:00
Rainfall Percentage in 1 st Third	55.69%
Rainfall Percentage in 2 nd Third	32.00%
Rainfall Percentage in 3 rd Third	12.31%
Rainfall Variability (Standard Deviation)	4.34 mm

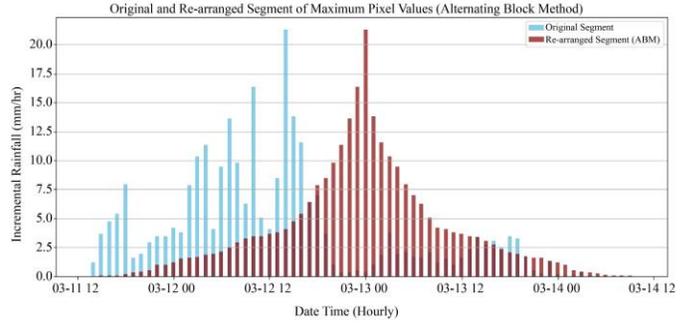


Fig. 9 The comparison between the resulting rearranged bar chart and the original hyetograph

Subsequently, a dimensionless mass curve was created by accumulating the incremental hyetograph and then performing normalization for the accumulated curve, producing the synthetic rainfall distribution pattern, as shown in Figure 10.

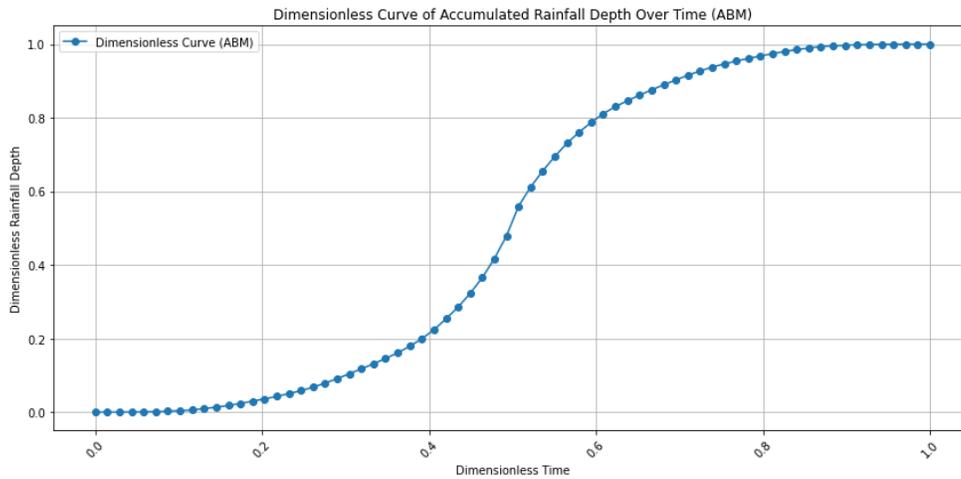


Fig. 10 The resulting Synthetic rainfall distribution pattern

To enable a thorough comparison between the synthetic distribution of the SCS Type II storm model and the GSMaP-derived synthetic distribution, the case study storm event is divided into three separate 24-hour intervals. This method ensures that each segment can be directly compared with the SCS Type II distribution, which is specifically designed to represent a 24-hour storm duration [4]. To precisely characterize the synthetic rainfall patterns derived from

GSMaP and identify the segment most analogous to the SCS Type II distribution, the total storm duration was subdivided into three portions (0-8, 8-16, and 16-24 hours). The percentages of rainfall amounts corresponding to each portion were quantified, and the outcomes of this analysis are presented in Table 3. While Figure 11 shows the comparison of all resulting patterns with the pattern of SCS type II.

Table 3. The percentage of rainfall corresponding to each portion

Rainfall Distributions	Portions of Storm Duration		
	1 st Third	2 nd Third	3 rd Third
GSMaP (11-03-2024 to 12-03-2024)	13.37%	70.78%	15.85%
GSMaP (12-03-2024 to 13-03-2024)	19.10%	59.86%	21.04%
GSMaP (13-03-2024 to 14-03-2024)	22.18%	53.14%	24.68%
SCS Type II	11.78%	75.98%	12.32%

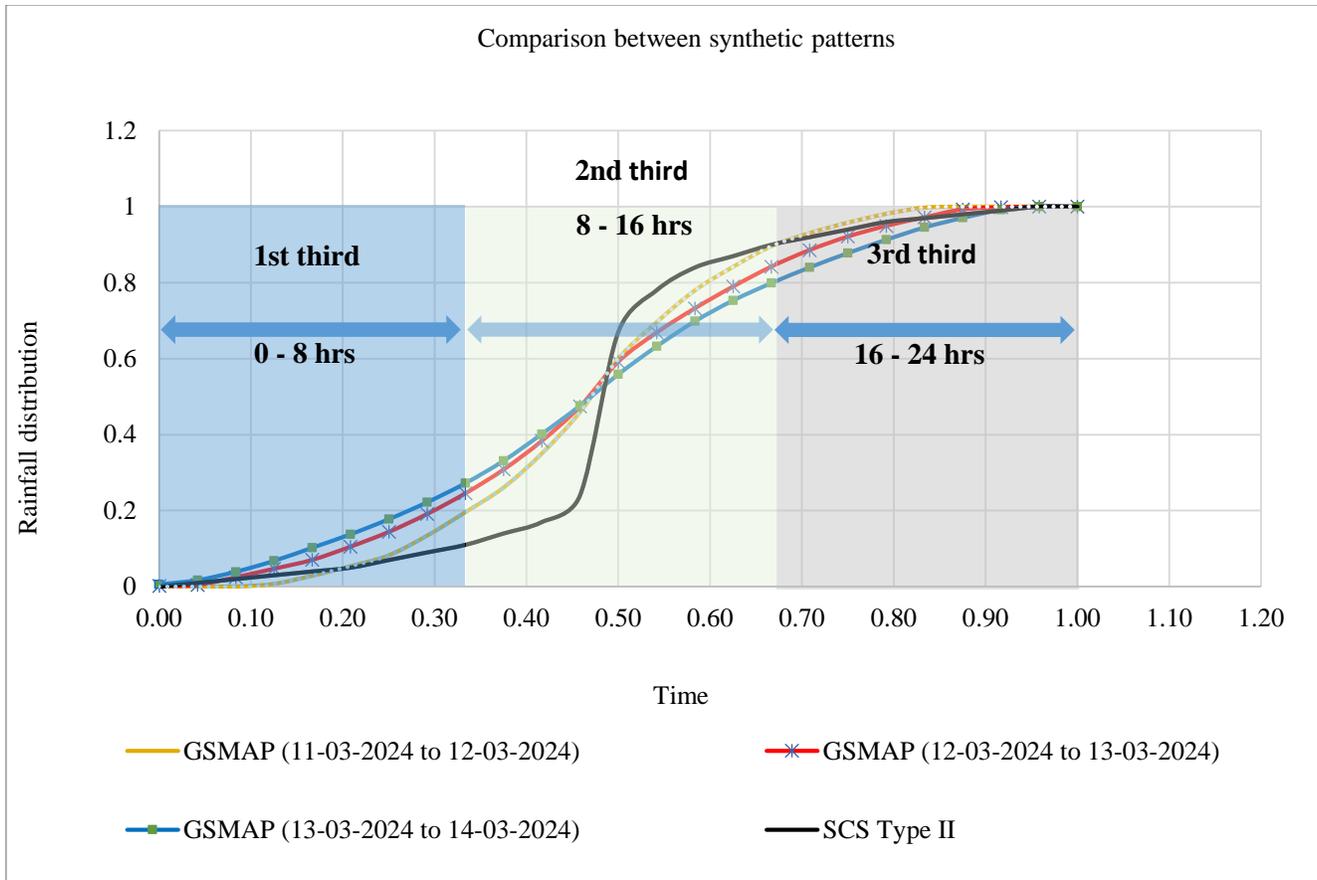


Fig. 11 Comparison of all resulting patterns with the pattern of SCS type II

The chart clearly indicates that the synthetic rainfall pattern observed during the initial 24-hour period (from March 11, 2024, to March 12, 2024) exhibits the closest resemblance to the SCS Type II distribution across all storm portions.

Conversely, This similarity progressively declines in the second and third 24-hour periods, where the patterns become increasingly smoother compared to the SCS Type II distribution, which is characterized by gradual slopes at the beginning and end and a dramatic increase in the central portion.

The level of similarity between the derived distributions and the SCS Type II distribution was determined by the Root Mean Square Error (RMSE), a statistical measure that expresses the average difference between two data sets. Among the analyzed distributions, GSMAP (11-03-2024 to 12-03-2024) showed the smallest RMSE value (0.0705), indicating the highest level of accurate overall agreement with the SCS Type II distribution.

A lower RMSE value indicates that the temporal patterns and scales of the GSMAP distribution show greater similarity to the reference distribution, further supporting the finding that this data set best represents the characteristics of the SCS

Type II curve. Therefore, based on the RMSE analysis, GSMAP (11-03-2024 to 12-03-2024) is considered the distribution with the highest level of similarity to the SCS Type II distribution among the analyzed data sets. The results of RMSE for each distribution are shown in the table below.

Table 4. RMSE for each distribution

Distribution	RMSE
GSMAP (11-03-2024 to 12-03-2024)	0.0705
GSMAP (12-03-2024 to 13-03-2024)	0.0914
GSMAP (13-03-2024 to 14-03-2024)	0.1106

To accentuate the impact of variations in rainfall distribution patterns, a comprehensive case study was undertaken in one of the prominent wadis in South Sinai, “Wadi Fieran.” A detailed hydrological analysis was performed on the Wadi Fieran watershed to estimate the resulting flood discharge associated with each rainfall distribution pattern.

Figure 12 illustrates all drainage lines in South Sinai and the geographical location of Wadi Fieran in relation to the boundaries of South Sinai on the background of the Digital Elevation Model (DEM).

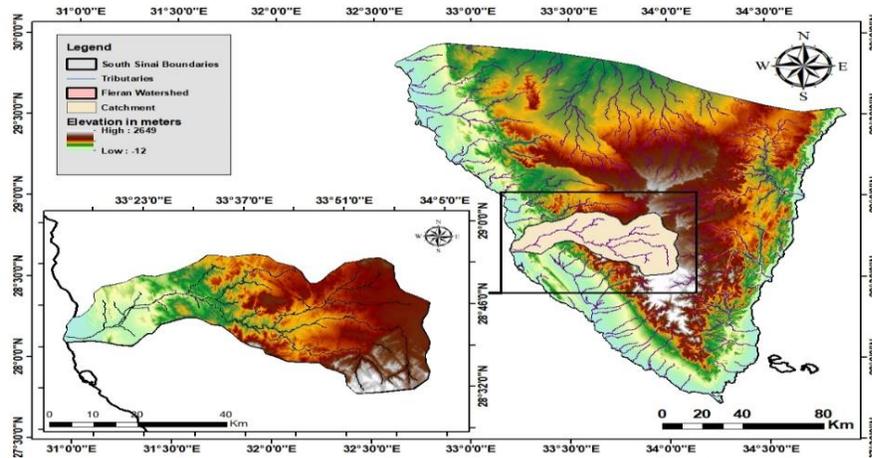


Fig. 12 The geographical location of Wadi Fieran in relation to the boundaries of South Sinai

A comprehensive hydrological model was developed for the Wadi Fieran watershed using HEC-HMS software. The significant storm event of March 2020 A three-day storm event was simulated in the meteorological model, with a total rainfall depth of 283 mm. Two hypothetical scenarios were considered. In the first scenario, the 283 mm rainfall was assumed to have been recorded by a ground-based rain gauge, yielding 56.24 mm,

180.55 mm, and 46.28 mm on the first, second, and third days, respectively. This scenario utilized the SCS Type II hypothetical storm distribution. The same 283 mm rainfall was distributed in the second scenario according to the actual rainfall pattern generated from the satellite-based GSMaP product. A comparison of the meteorological inputs and the resulting hydrological outputs for both scenarios is depicted in Figure 13.

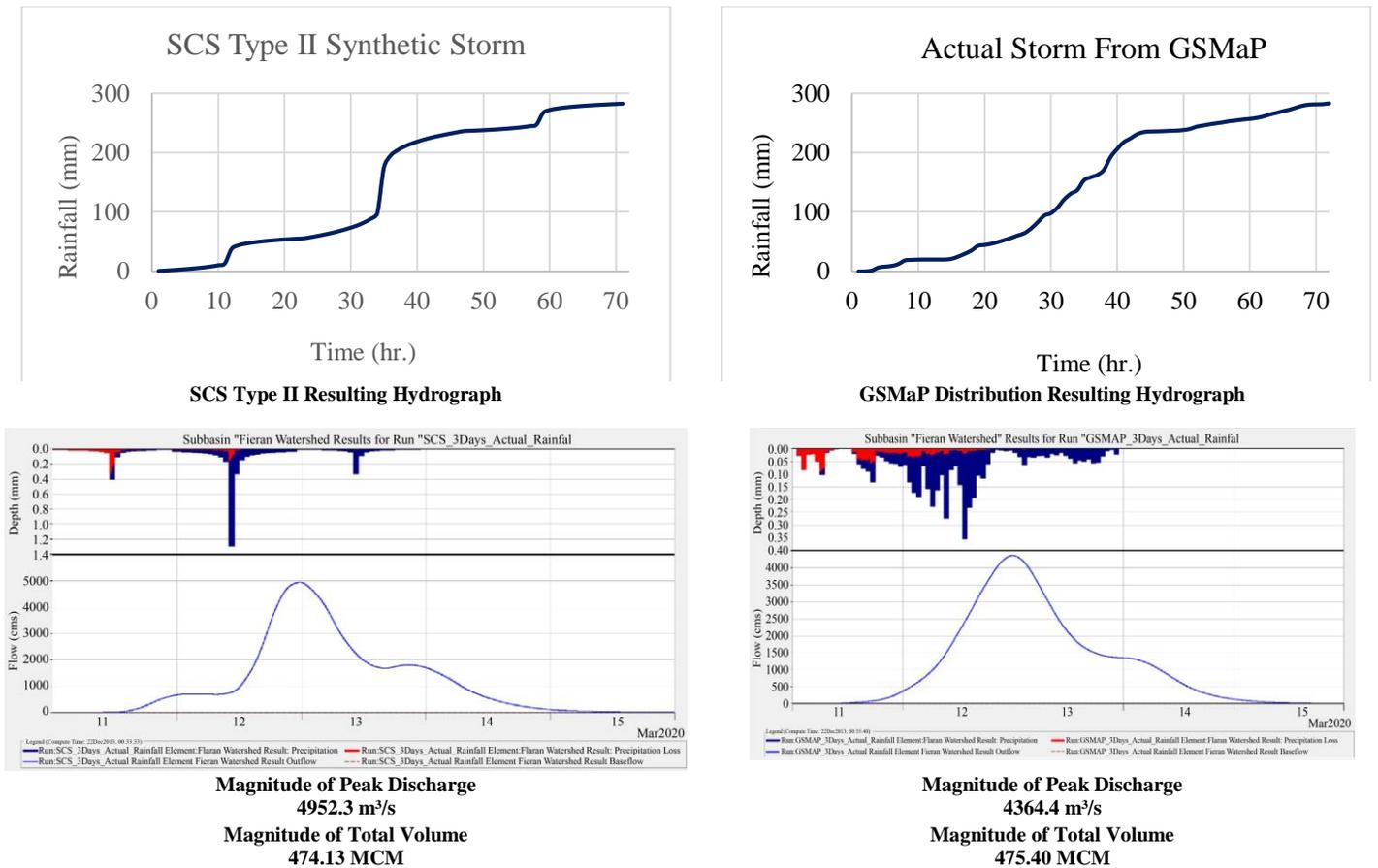


Fig. 13 Comparison of the meteorological inputs and the resulting hydrological outputs for both scenarios

The comparison of the resulting hydrographs that correspond to each rainfall distribution is depicted in Figure 14.

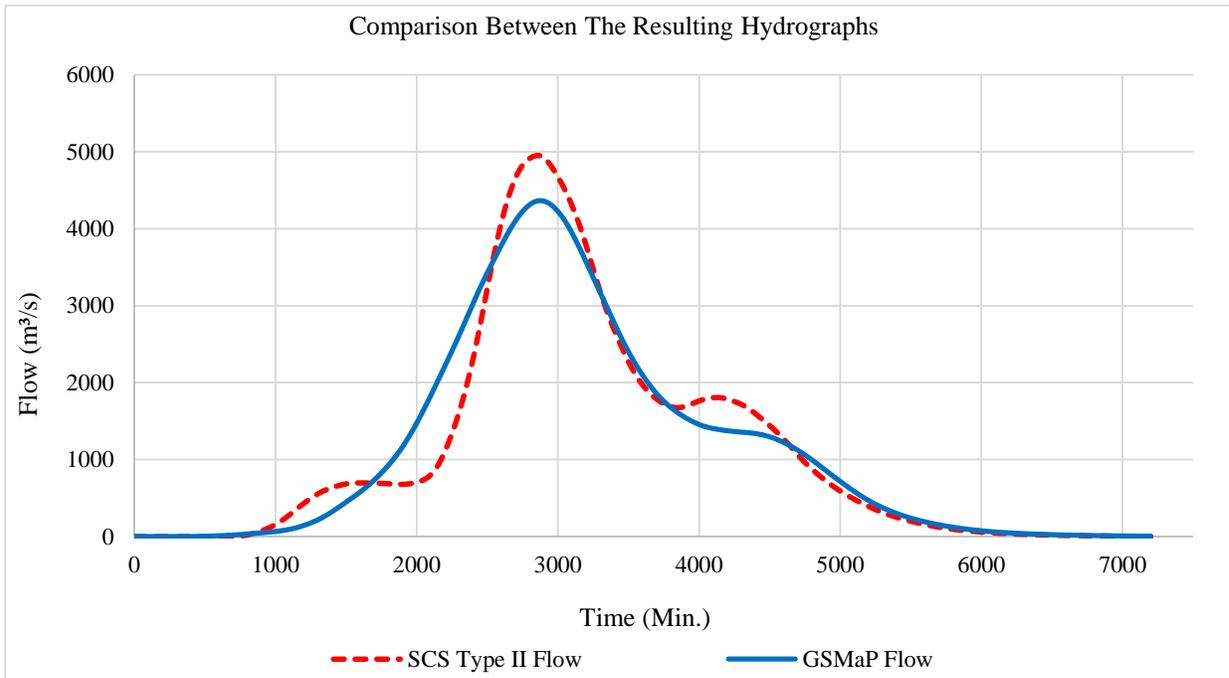


Fig. 14 Comparison of the resulting hydrographs

As demonstrated in the comparison of hydrographs, the divergence between the synthetic rainfall distribution patterns led to a significant variation in the resulting hydrographs. It was observed that the actual rainfall distribution from GSMaP produced a more realistic hydrograph and a lower magnitude of peak flood discharge in comparison to the SCS Type II distribution.

4. Conclusion

The methodology employed in this study effectively utilized GSMaP data to construct a synthetic rainfall distribution pattern. Advanced techniques, such as DBSCAN for spatial clustering of storm events and the Alternating Block Method (ABM) for rearranging rainfall depths, are utilized. A Python script was created to automate the entire process, including extracting GSMaP data, applying spatial and temporal filtering for the area of interest, and tracking the movement of storm cores to accurately generate corresponding hydrographs. This approach yields a dimensionless mass curve that effectively characterizes the rainfall distribution and could be applied in other regions with different climatic conditions. The results demonstrate that the synthetic rainfall pattern closely aligns with the SCS Type II distribution during the initial 24-hour period, with a gradual divergence in subsequent intervals.

A detailed hydrological model for the Wadi Fieran watershed was constructed using HEC-HMS software to demonstrate the impact of the variation in synthetic rainfall distribution. The results indicated that the differences in

synthetic rainfall distributions resulted in considerable variations in the hydrographs, with the GSMaP-derived distribution yielding a more realistic hydrograph with a 12% lower peak flood discharge and a 0.27% higher total flood volume in comparison to the SCS-type II distribution.

The study demonstrates the effectiveness of integrating satellite-based precipitation data with computational methods to produce accurate rainfall distribution patterns for hydrological simulations and flood risk assessments. Furthermore, it emphasizes the role of Python in automating data analysis and visualization. Overall, the distributions are expected to represent a kind of similarity when they correspond to the same climatic characteristics.

The user must visually validate the results of the DBSCAN algorithm to ensure that all records belong to the same cluster, representing a single storm event. Since the algorithm employs the Alternating Block Method to reorganize the rainfall distribution, slight overestimations in the resulting patterns are expected. Moreover, the algorithm may not perform reliably under diverse hydrological conditions, particularly in regions with unique or irregular rainfall patterns, and therefore, it is not recommended for such areas.

4.1. Recommendations

Since this study did not impose assumptions on the synthetic distribution or extracted data and is specifically derived for South Sinai, it is recommended to use the

developed pattern for hydrological analysis in South Sinai, where it has shown promising applicability.

Although GSMaP provides valuable rainfall estimates, biases can occur due to orographic effects, convectional local storms, and cloud cover interference. The GSMaP estimates are to be validated continuously against rain gauge and radar observations from ground data to warrant greater reliability. The comparisons enable the identification of biases, calibration improvement, and algorithm updates to better represent rainfall. Conducting statistical tests, correlation tests, bias diagnostics, and error diagnoses ensures GSMaP data is in close proximity to true observations, thus enhancing its use in hydrologic and climatologic studies.

The effectiveness of the developed algorithm depends heavily on the quality of the rainfall data used. Therefore, it is essential to thoroughly validate the chosen satellite product in future studies. While GSMaP was suitable for the study location and purposes of the present study, this is not necessarily a fact that will make it the most suitable product for all research. Insufficient or inaccurate data potentially results in outputs that do not accurately reflect actual storm behavior.

It is important to assess whether the Alternating Block Method (ABM) is suitable for the region’s specific rainfall characteristics. ABM may not perform reliably in areas with unique or highly variable precipitation patterns, requiring careful evaluation before implementation.

It is recommended to being informed about advancements in satellite-based precipitation products and evaluate their reliability for hydrological modeling. This is particularly relevant as updated versions often provide improved spatial resolution with smaller pixel sizes.

4.2. Future Research Directions

Cloud computing through Google Earth Engine is recommended for analyzing the spatial pattern of rainfall storms to identify the locations most susceptible to heavy precipitation in a given area of interest.

The study recommends validating the synthetic rainfall distribution pattern with additional storm events and upgrading the Python script for longer time frames and statistical analyses. Future research should extend this methodology to different regions and climatic conditions to

evaluate its broader applicability in flood risk management. Testing in diverse environments will help refine the model’s adaptability and accuracy.

Enhancing the algorithm to analyze storm movement trends along the area of interest with respect to the longest flow path of the catchment area would further improve its utility. Such improvements would enable hydrologists to simulate and evaluate multiple scenarios within hydrological models, ultimately enhancing the accuracy of their assessments and decision-making processes.

Acknowledgements

I would like to express my sincere gratitude to Professor Ashraf Mohamed El-Moustafa and Dr. Morad Hamid Abdeldayem for their invaluable support and expertise throughout this research project, which was instrumental to its success.

I extend special thanks to Professor Alfonso Torres for his guidance and encouragement, which set me on the right path in utilizing Python effectively with Google Earth Engine.

List of Abbreviations

GSMAP :	Global Satellite Mapping of Precipitation.
PMW-IR :	Passive Microwave and Infrared.
GPM :	Global Measuring Mission
ARC :	African Rainfall Climatology.
CHIRPS :	Climate Hazards Group Infrared Precipitation with Stations.
TAMSAT :	Tropical Applications of Meteorology Using Satellite and Ground-Based Observations.
PERSIANN-CCS :	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System.
NASA :	National Aeronautics and Space Administration
PERSIANN-CDR :	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record
TRMM :	Tropical Rainfall Measuring Mission
DBSCAN :	Density-Based Spatial Clustering of Applications with Noise
GEEMap :	Google Earth Engine Map.
SCS :	Soil Conservation Service.
MENA :	Middle East and North Africa
ABM :	Alternating Block Method

References

[1] Fahad Alahmadi, Norhan Abd Rahman, and Zulkifli Yusop, “Hydrological Modelling of Ungauged Arid Volcanic Environments at Upper Bathan Catchment, Madinah, Saudi Arabia, *Technology Journal*, vol. 78, no. 9-4, pp. 9-14, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

[2] Bhavin Ram et al., “Deriving Location-Specific Synthetic Seasonal Hyetographs using GPM Records and Comparing with SCS Curves,” *Journal of Water and Climate Change*, vol. 15, no. 2, pp. 747-758, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

[3] Wes Dick, and Ahmadreza Ghavasieh, “A 24-h Design Storm for the Fort McMurray Region,” *Canadian Journal of Civil Engineering*, vol. 42, no. 10, pp. 747-755, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [4] A.G. Awadallah, and N.S. Younan, “Conservative Design Rainfall Distribution for Application in Arid Regions with Sparse Data,” *Journal of Arid Environments*, vol. 79, pp. 66-75, 2012. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Eman Ahmed Hassan El- Sayed, “Development of Synthetic Rainfall Distribution Curves for Sinai Area,” *Ain Shams Engineering Journal*, vol. 9, no. 4, pp. 1949-1957, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Amro M. Elfeki, Hatem A. Ewea, and Nassir S. Al-Amri, “Development of Storm Hyetographs for Flood Forecasting in the Kingdom of Saudi Arabia,” *Arabian Journal of Geosciences*, vol. 7, pp. 4387-4398, 2014. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Vincenzo Levizzani et al., *Advances in Global Change Research, Satellite Precipitation Measurement*, Springer Cham vol. 1, pp. 1-66, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] Adel Bakheet, and Ahmed Sefelnasr, “Application of Remote Sensing data (GSMaP) to Flash Flood Modeling in an Arid Environment, Egypt,” *Proceeding of the International Conference on Chemical and Environmental Engineering*, vol. 9, pp. 252-272, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Arthur Y. Hou et al., “The Global Precipitation Measurement Mission,” *Bulletin of the American Meteorological Society*, vol. 95, no. 5, pp. 701-722, 2014. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Yudong Tian et al., “Evaluation of GSMaP Precipitation Estimates over the Contiguous United States,” *Journal of Hydrometeorology*, vol. 11, no. 2, pp. 566-574, 2010. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Fatkhuroyan and TrinhWati, “Accuracy Assessment of Global Satellite Mapping of Precipitation (GSMaP) Product Over Indonesian Maritime Continent,” *IOP Conference Series: Earth and Environmental Science: The 4th International Seminar on Sciences*, Bogor, Indonesia, vol. 187, pp. 1-12, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Mohamed Salem Nashwan, Shamsuddin Shahid and Xiaojun Wang, “Assessment of Satellite-Based Precipitation Measurement Products over the Hot Desert Climate of Egypt,” *Remote Sensing*, vol. 11, no. 5, pp. 1-18, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] K.I. Okamoto et al., “The Global Satellite Mapping of Precipitation (GSMaP) Project,” *Proceedings, 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005, IGARSS '05*, Seoul, Korea (South), pp. 3414-3416, 2005. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Guosheng Liu, “A Fast and Accurate Model for Microwave Radiance Calculations,” *Journal of the Meteorological Society of Japan, Ser. II*, vol. 76, no. 2, pp. 335-343, 1998. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] Martin Ester et al., “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” *KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, Portland Oregon, pp. 226-231, 1996. [[Google Scholar](#)] [[Publisher Link](#)]
- [16] K.M. Kent, “*A Method for Estimating Volume and Rate of Runoff in Small Watersheds*,” Report, U.S. Dept. of Agriculture National Agricultural Library, Current Serial Records, pp. 1-66, 1969. [[Publisher Link](#)]
- [17] J.V. Bonta, “Stochastic Simulation of Storm Occurrence, Depth, Duration, and Within-Storm Intensities,” *Transactions of the ASAE, American Society of Agricultural Engineers*, vol. 47, no. 5, pp. 1573-1584, 2004. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] E.A. El-Sayed and Emad Habib, “Advanced Technique for Rainfall-Runoff Simulation in Arid Catchments Sinai, Egypt,” *The 3rd International Conference on Water Resources & Arid Environment (2008) and First Arab Forum*, pp. 1-13, 2008. [[Google Scholar](#)]
- [19] Egypt: Flash Floods Emergency Plan of Action (EPoA) Final Report DREF n° MDREG015, (IFRC), ReliefWeb, 2020. [Online]. Available: <https://reliefweb.int/report/egypt/egypt-flash-floods-emergency-plan-action-epoa-final-report-dref-n-mdreg015#:~:text=/>
- [20] Giovanni the Bridge between Data and Science, Earth Data, 2024. [Online]. Available: <https://giovanni.gsfc.nasa.gov/giovanni/#service=AcMp&starttime=2020-03-11T00:00:00Z&endtime=2020-03-14T23:59:59Z&bbox=21.8144,21.0586,37.2832,32.748/>