

# **A SATELLITE-BASED ANALYSIS OF TROPICAL CYCLONE RAINFALL FOR IMPROVED FLOOD HAZARD ASSESSMENT, CASE STUDY IN DOMINICA**

NABUKULU CATHERINE

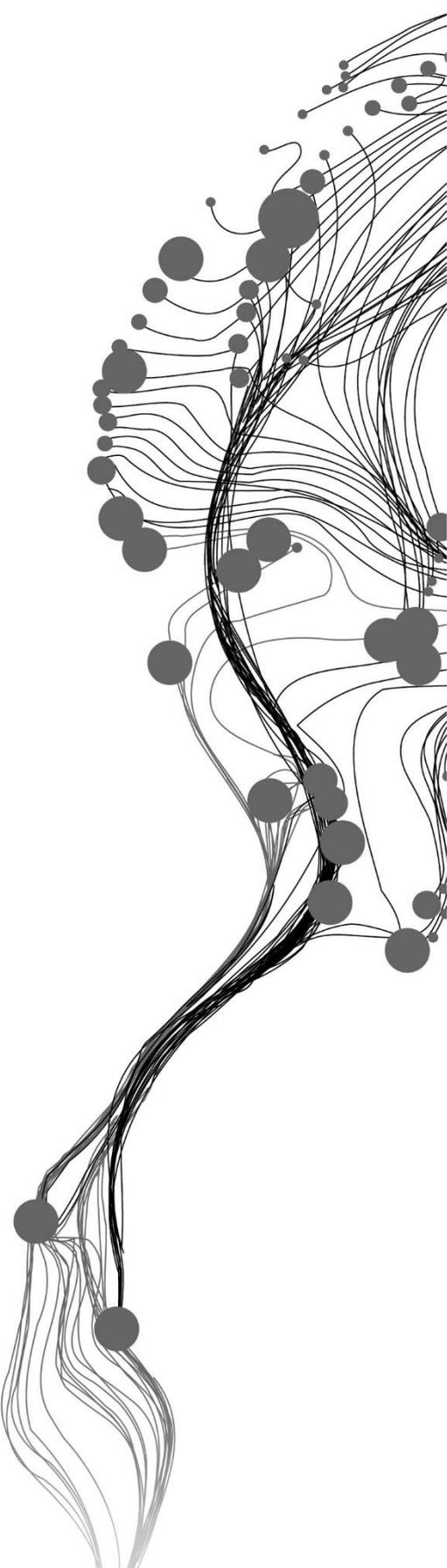
June 2021

SUPERVISORS:

Dr. Ir. Janneke Ettema

Prof. Dr. Victor Jetten





# A SATELLITE-BASED ANALYSIS OF TROPICAL CYCLONE RAINFALL FOR IMPROVED FLOOD HAZARD ASSESSMENT, CASE STUDY IN DOMINICA

NABUKULU CATHERINE

Enschede, The Netherlands, June 2021

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Natural Hazards and Disaster Risk Reduction

**SUPERVISORS:**

Dr. Ir. Janneke Ettema

Prof. Dr. Victor Jetten

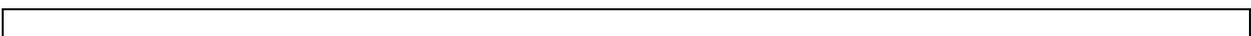
**THESIS ASSESSMENT BOARD:**

Prof. Dr. Norman Kerle (Chair)

Dr. Rein Haarsma (External Examiner, KNMI)

**DISCLAIMER**

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.



## ABSTRACT

The severity of weather events accompanying tropical cyclones (TC), such as torrential precipitation and strong winds, is changing because of increased global anthropogenic warming. The last 5-10 years have witnessed overwhelming flooding from tropical cyclone extreme rainfall causing damages that have burdened the economy, especially for developing countries in TC-prone zones. Due to the scarcity of long-term TC rainfall records, flood modelers in these countries use regional design storms for TC-related flood hazard assessment. However, design storms might fail to represent the intricate patterns of TC precipitation due to the few observational recordings of TCs by rain gauges, giving unrealistic estimations of the TC-related flooding hazard.

This research's solution is a new approach that attempts to categorize the structure of TC associate rainfall by utilizing satellite precipitation estimates from GPM-IMERG V06 data to improve TC-related flood hazard assessment. The method was tested on Tropical storm Erika (2015), which brought torrential rainfall to the study area in the vicinity of Dominica. The TC rainfall's distinct spatial-temporal behaviours were revealed using K-means in a time series clustering analysis. The research focused on the differences in the temporal distribution of the rainfall, also emphasized by the distinct flood responses modeled in openLISEM. First, Tropical storm Erika's rainfall temporal distribution was analyzed for three values of optimal clusters ( $K$ ), i.e., 5, 4, and 3. For each  $K$  value, one cluster is excluded from further analyses as its location away from TC, the precipitation amount, and intensity were significantly lower than the other clusters. The second step in the developed approach involved setting a 10mm/hr starting threshold to align the pixel times series. The third step was to derive cluster representative signals used as the precipitation input in the flood model. Rainfall signals resulting from  $K=5$  had similar quantified responses in flood extent, depth, volume, duration, and runoff ratio between the cluster signals. A final step in the form of an optimization approach was implemented to address these similarities and improve the generalization of the TC rainfall. At a reduced  $K$  value, the TC precipitation was divided into three (for  $K=4$ ) and two (for  $K=3$ ) levels of magnitude with distinct quantified flood responses. Concluding, rainfall signals resulting from  $K=4$  were selected as the TC associate rainfall dataset since they were associated with higher magnitudes in flood response.

We observed that different temporal behaviours of varying magnitude exist for precipitation accompanying a given tropical cyclone. Since flood characteristics change with intricate rainfall patterns, the consequences suffered in an area depend on which part of the TC passes that location. This study showed that flood hazard modelers and risk planners can utilize the developed approach to generate a reliable TC associate rainfall dataset to make better-informed decisions related to TC-induced flooding.

**Keywords:** Satellite precipitation estimates, Tropical cyclones, Time series clustering, Flood hazard assessment.

## ACKNOWLEDGEMENTS

Special thanks to my supervisors, Dr. Ir. Janneke Ettema and Prof. Dr. Victor Jetten for their support throughout the research. They have allowed me to exploit my potential in various ways especially through the discussions during our online meetings where we built ideas to make this research a success. I believe my scientific writing skill has improved because of their guidance while writing the thesis. I thank Prof. Dr. Norman Kerle, who has always chaired the assessment board during the different phases of the research. He was always critical (in a good way) and gave insightful comments that challenged me to improve my thinking and presentation of scientific ideas. I now finish this thesis as a better researcher than when I started.

I extend my gratitude to the ITC Foundation Scholarship committee for awarding me their scholarship to achieve my dream of studying my MSc from ITC.

I want to thank my friend Beatrice Kaijage who I now consider a sister. We have encouraged and supported each other in many ways, participated in church ministry, and shared many meals. I am glad that we are both completing our MSc with good grades, as we have always wanted. I also thank my classmates from the NHR class; you have made this journey so enjoyable.

In a special way, I thank my dear mum, Katusiime Jennifer, and my siblings, Nabakka Macklina and Bukenya Timothy, for their endless love and prayers throughout my education. They have been such a strong support system throughout the whole MSc journey. This was my first experience staying far from home for an extended period.

The best saved for last, I glorify the Almighty God for seeing me through this journey, for His favor and mercies throughout my life and my education. I say EBENEZER! For the far that the Lord has brought me. Webale Yesu, Webale Mukama!

# TABLE OF CONTENTS

---

1.	INTRODUCTION.....	1
1.1.	Background and Motivation.....	1
1.2.	Problem Statement.....	3
1.3.	Research Objectives and Questions.....	4
1.4.	Case Study Area.....	4
1.5.	Research Framework and Overview.....	5
2.	SPATIAL-TEMPORAL EVALUATION OF TROPICAL CYCLONE RAINFALL.....	7
2.1.	Tropical Cyclones and their Precipitation.....	7
2.2.	Tropical Cyclone Selection.....	8
2.3.	Satellite Rainfall Data Acquisition.....	8
2.4.	Rainfall Inspection for the Selected TC.....	9
2.5.	Rainfall Time Series Clustering.....	9
2.6.	Selection of Cluster Representative Signals.....	11
3.	FLOOD HAZARD MODELLING.....	13
3.1.	Flood Modelling in openLISEM.....	13
3.2.	Model Setup and Assumptions.....	14
3.3.	Investigated Flood Characteristics.....	15
3.4.	Optimization Approach based on the Flood Characteristics.....	17
4.	RESULTS: RAINFALL ANALYSIS.....	18
4.1.	Selected/Candidate Tropical Cyclones.....	18
4.2.	Optimal Number of Clusters ( $K$ ).....	20
4.3.	Spatial Clustering Statistics.....	21
4.4.	Temporal Clustering Statistics.....	22
4.5.	Cluster Representative Signals.....	26
5.	RESULTS: FLOOD MODELLING.....	32
5.1.	Using Five Temporal Clusters ( $K=5$ ).....	32
5.2.	Application of the Optimization Approach.....	39
5.3.	The Tropical Cyclone Associate Rainfall DataSet.....	40
6.	DISCUSSION AND CONCLUSION.....	43
7.	RECOMMENDATIONS.....	48
	LIST OF REFERENCES.....	50
	APPENDICES.....	56

## LIST OF FIGURES

---

Figure 1-1: Location of the studied catchment in the south of Dominica. ....	5
Figure 1-2: Flow chart of the research design and strategy .....	6
Figure 2-1: Structure of a tropical cyclone. ....	8
Figure 2-2: Illustration of the procedure for selection of cluster representative signals.....	12
Figure 3-1: Summary of openLISEM sediment and hydrological simulation. ....	14
Figure 3-2: Order of flow processes in openLISEM, including overland flow, channel flow, flooding, and flood recession. Adopted from Bout et al. (2018) .....	15
Figure 3-3: Illustration of the approach for optimizing the TC associate rainfall dataset. ....	17
Figure 4-1: Tracks for candidate TCs passing through the study area (500km diameter buffer around Dominica’s coast).....	18
Figure 4-2: Distribution of the pixel rainfall accumulation due to TS Erika’s precipitation over the study area from 26 <sup>th</sup> August 2015 0000UTC to 28 <sup>th</sup> August 2015 2330UTC. The four shapes, cross, star, triangle, and hexagon, are locations for the pixels whose time series are plotted in Figure 4-3.....	19
Figure 4-3: Time series plots for individual pixels selected randomly from the dataset. ....	20
Figure 4-4: Optimal $K$ values determined at locations where the graph starts to flatten as indicated by the blue arrow. a) Spatial clustering. b) Temporal clustering.....	21
Figure 4-5: A map representation of the spatial clustering of TS Erika’s rainfall.....	22
Figure 4-6: A map representation of the temporal clustering of TS Erika rainfall using $K=5$ .....	23
Figure 4-7: Rainfall distribution across all clusters. a) Pixel maximum rainfall intensity b) Cumulative rainfall. ....	24
Figure 4-8: A map representation of the temporal clustering of TS Erika rainfall using $K=4$ .....	25
Figure 4-9: A map representation of the temporal clustering of TS Erika rainfall using $K=3$ .....	26
Figure 4-10: a) Plot of T1’s spatially distinct time series with similar temporal behaviour. b) Timestep quantile series of T1 based on the original data (without applying a starting threshold).....	27
Figure 4-11: Plots of T1 timestep quantile series using varying starting thresholds.....	28
Figure 4-12: Rainfall time series (left) and cumulative plots (right) of the selected cluster representative signals when using $K=5$ . T1, T3, T4, and T5 plots are in rows 1, 2, 3, and 4, respectively.....	30
Figure 5-1: Scatter plots of the linear relationship between the maximum flood extent (Y-axis) and a) cumulative rainfall, b) maximum rainfall intensity. ....	33
Figure 5-2: Geographical representation of the maximum flood depths reached in the different temporal clusters due to rainfall of Q3. All the maps use the same scale of flood depths labelled Low, Medium, High, and Extra. ....	34
Figure 5-3: Scatter plots of the linear relationship between the maximum flood depth (Y-axis) and a) cumulative rainfall, b) maximum rainfall intensity.....	35
Figure 5-4: Scatter plots of the linear relationship between the maximum flood volume (Y-axis) and a) cumulative rainfall, b) maximum rainfall intensity.....	36
Figure 5-5: Linear relationship between the infiltration and a) cumulative rainfall, b) maximum rainfall intensity.....	36
Figure 5-6: Cumulative infiltration (mm) reached due to signal Q3 rainfall.....	37
Figure 5-7: Scatter plots of the linear relationship between the runoff ratio (Y-axis) and a) cumulative rainfall, b) maximum rainfall intensity.....	38

Figure 5-8: Precipitation time series and cumulative precipitation plots of the Tropical cyclone associate rainfall dataset for TS Erika resulting from  $K=4$ ..... 41

Figure 5-9: Flood depth (meters) distribution for all the three rainfall signals of the Tropical cyclone associate rainfall dataset for TS Erika resulting from  $K=4$ ..... 42

## LIST OF TABLES

---

Table 2-1: Labels and corresponding probabilities for the descriptive statistics calculated at each timestep. .....	11
Table 4-1: Statistics for the nine spatial clusters.....	21
Table 4-2: Summary statistics for the temporal clusters resulting from $K=5$ .....	23
Table 4-3: Summary statistics for the temporal clusters resulting from $K=4$ .....	24
Table 4-4: Summary statistics for the temporal clusters resulting from $K=3$ .....	25
Table 4-5: Observations in the characteristics of timestep quantile series at varying starting thresholds. ....	27
Table 4-6: Signal statistics computed for timestep quantiles series of T1 at varying thresholds.....	29
Table 4-7: Summary statistics for signals Q2, Q3, and Mean for the temporal clusters resulting from $K=5$ .....	29
Table 4-8: Summary statistics for signals Q2, Q3, and Mean for the temporal clusters resulting from $K=4$ .....	31
Table 4-9: Summary statistics for signals Q2, Q3, and the Mean for temporal the clusters resulting from $K=3$ .....	31
Table 5-1: Quantified maximum flood extent (km <sup>2</sup> ) caused by cluster representative signals resulting from $K=5$ .....	32
Table 5-2: Quantified maximum flood depth (meters) caused by cluster representative signals resulting from $K=5$ .....	33
Table 5-3: Quantified maximum flood volume (million m <sup>3</sup> ) caused by cluster representative signals resulting from $K=5$ .....	35
Table 5-4: Summary of the total infiltration (mm) for the cluster representative signals resulting from $K=5$ . .....	36
Table 5-5: Quantified runoff ratio caused by cluster representative signals resulting from $K=5$ .....	38
Table 5-6: Summary of the flood response time (hours) caused by cluster representative signals resulting from $K=5$ .....	39
Table 5-7: Summary of the average flood duration (hours) caused by cluster representative signals resulting from $K=5$ .....	39
Table 5-8: Comparison of flood characteristics resulting from rainfall signal Q3 for $K=4$ .....	40
Table 5-9: Comparison of flood characteristics resulting from rainfall signal Q3 for $K=3$ .....	40

## LIST OF ACRONYMS

---

CHARIM	Caribbean Handbook on Risk Information Management
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data
GDP	Gross Domestic Product
GFDRR	Global Facility for Disaster Reduction and Recovery
GIS	Geographic Information Systems
GPM	Global Precipitation Measurement
HURDAT2	The Revised Atlantic Hurricane Database
IMERG	Integrated Multi-satellitE Retrievals for GPM
IPCC	Intergovernmental Panel on Climate Change
LISEM	Limburg Soil Erosion Model
MERRA-2	second Modern-Era Retrospective analysis for Research and Applications
MOE	Ministry of Environment
N/A	Not Applicable
NHC	National Hurricane Centre
NOAA	National Oceanic and Atmospheric Administration
PERSIANN	Precipitation Estimation form Remotely Sensed Information using Artificial Neural Networks
SST	Sea Surface Temperature
SQKM	Square Kilometres
TC	Tropical Cyclone
TMPA	Tropical Rainfall Measurement Mission Multi-satellite Precipitation Analysis
UNFCCC	United Nations Framework Convention on Climate Change
UTC	Coordinated Universal Time
USD	United States Dollar
WMO	World Meteorological Organisation



# 1. INTRODUCTION

## 1.1. Background and Motivation

Tropical cyclone (TC) systems impact extended areas with their destructive power resulting directly from strong winds and indirectly from torrential rainfall amounts, landslides, flooding, and storm surges (Lenzen et al., 2019; Rahimi et al., 2015; Roux, 2019). A mature tropical cyclone consists of a low-pressure center, a convective eyewall, and active inner and outer spiral rainbands (Wang, 2012). The rainbands are located outside the eyewall; their spiraling nature influences the system's intensity and structure. The system affects areas between 200-500km diameter up to around 1000km (WMO, 2020). Experts of Munich Re-Insurance Company report that tropical cyclones dominated the global economic losses from natural disasters for the years 2017, 2018, and 2019 (Faust & Bove, 2017; Loew, 2019, 2020). According to Smith (2020), tropical cyclones were responsible for damages worth USD 945.9 billion during 1980-2019 in the U.S, which amounts to 53.9% of the total costs of the U.S billion-dollar natural disaster events. The Caribbean islands take 15 spots in the top 25 countries with the most TCs per square kilometer (Acevedo & Alleyne, 2016), hence the most vulnerable regions to these devastating weather systems. In some Caribbean countries, especially the smaller island states such as Dominica and Marie-Galante, economic damages suffered from TCs were observed to exceed their economy (Otker & Srinivasan, 2018).

The destruction from the recent very wet tropical cyclones has raised concern and interest in TC-induced rainfall in the Caribbean. For example, in 2015, Tropical storm Erika produced torrential rains with maximums up to 320.5mm (Pasch & Penny, 2015), resulting in catastrophic flash floods on Dominica and Guadeloupe. Two years later, the Caribbean islands were again hit by an episode of record-breaking precipitation brought by Hurricane Maria (Pasch et al., 2017) that led to rapid flooding, causing severe damages which halted the recovery efforts from the previous 2015 devastation. Puerto Rico and Dominica experienced the highest maximum totals above 558.8mm, followed by Guadeloupe and the Dominican Republic, with precipitation ranging from 254mm to 330.2mm. Knowing that the severest impacts in low developed countries are primarily due to overexposure, higher vulnerability, less coping capacity, and recovery, future TCs might have disproportionate consequences from their highly destructive force (Hallegatte et al., 2017).

Recently, most Caribbean islands are engaging in extensive flood hazard analysis from TCs and are worried that climate change will strongly affect hazard frequency and intensity (CHARIM, 2018). Already, the last five years of the annual Atlantic hurricane season have witnessed an increased occurrence of higher category TCs, landfalling TCs, precipitation extremes, and greater devastation, all linked to climate change impacts (Jacqueline, 2020; Stephenson & Jones, 2017; Thomas et al., 2017). In this context, many researchers are using climate models to produce scientific evidence on the likely influence of global anthropogenic warming induced by greenhouse gas emissions on future tropical cyclone activity (Bacmeister et al., 2018; Emanuel, 2013; Gallo et al., 2019; Mori, 2014; Murakami et al., 2012). The IPCC (2014) summarized findings from various studies on future climate change based on terminologies for assessing likelihood and levels of confidence, as in Mastrandrea et al. (2010).

The latest findings by a team from the WMO (Knutson et al., 2020) about the future TC activity for model simulations at 2°C global warmings reveal a range of likely (>66%probability) future impacts. 1) a 14% global average increase in TC precipitation rates, 2) higher coastal inundation due to the future sea-level rise resulting from the likely increase in moisture content due to projected increased warming of the sea surface, and 3) with a medium (about 5/10 chance) to high (about 8/10 chance) confidence, the global average TC intensities were projected to increase by a range of 1-10%. In addition, very intense TCs classified as categories 4 and 5 on the Saffir-Simpson Hurricane Wind Scale were projected to rise in their proportion at a median rate of 13%. Knutson et al.'s (2020) findings show a likelihood of a decrease in TC propagation speed; however, this is with low (about 2/10 chance) confidence hence the need for further research for an explicit link to climate change. Furthermore, IPCC (2014) reports that regional projections have low confidence as climate change's influence on tropical cyclones is region-specific. For instance, in the western North Pacific and North Atlantic basins, the most intense tropical cyclones will most likely not increase in frequency. Stephenson and Jones (2017) point out an existing knowledge gap for future TC projections specific to the Caribbean basin. In short, there is still a lot of uncertainty in the behaviour of these highly dynamic weather systems with current and future climate change.

Knowing that TC-induced rainfall poses a potential flood threat, future risk mitigation and management require accurate quantification of TC accompanying rainfall; however, long-term, high-quality precipitation measurements of TCs are scarce. The flood hazard assessment typically depends on the frequency and magnitude of peak discharges. However, long-term discharge records are non-existent for frequency analysis of flash floods, especially in hilly terrain with many rivers, which give a simultaneous and fast response. TC accompanying torrential precipitation is the primary cause of flash and riverine floods that have recorded devastating damages in the humid tropics over the years (Kostaschuk et al., 2001). What flood modelers are doing, they usually substitute the probability of the flood peak discharges with the probability of the meteorological forcing (i.e., extreme precipitation events) derived from long-term rainfall records.

Conventionally, long-term ground gauge measured daily rainfall records, for example, 30 years or more, are utilized to predict TC-related floods. Rain gauge records are collected as point measurements; however, floods result from accumulated areal rainfall over the study area (Rakhecha & Singh, 2009); therefore, measuring gauges must be distributed with a good density for adequate data capturing (Girons et al., 2015). It is possible to have no precipitation measurements during TC events because the gauges break down or may record an under catch due to the high-speed wind conditions, thus introducing a bias in the recorded rainfall estimations (Pollock et al., 2018). Considering weather radar-based measurements, they provide high spatial-temporal resolution rainfall estimates but are limited by acquisition backscatter and partial or whole blockage, especially in complex terrain regions (Gilewski & Nawalany, 2018). Furthermore, radar observations are limited in their operational period, implying insufficient long-term data to generate statistics for TC representative rainfall. Satellite precipitation estimates are more efficient for the continuous acquisition of long-time rainfall data and coverage over hard-to-reach areas (Elhamid et al., 2020; Yoshimoto & Amarnath, 2017). However, drawbacks have been observed due to the different products' precipitation retrieval algorithms (Jiang et al., 2019), their spatial and temporal resolution.

In 2014, NASA launched the GPM-IMERG (Integrated Multi-satellitE Retrievals for Global Precipitation Measurement) and has continuously upgraded its signature rainfall retrieval algorithm (Yong & Wang, 2020). The 0.1° x 0.1° grid precipitation data set is available in both near-real-time (Early and Late runs) and post real-time (Final run) with a time scale ranging from 30 minutes, three hourly and daily records. The current V06 supersedes the previous versions; its displacement vectors for precipitation motions are computed from

GEOS Forward Processing data for the Final run product and MERRA-2 for the near-real-time products (Huffman, Bolvin, et al., 2019; Tan & Huffman, 2019). The GPM-IMERG Final run product has been recognized to outperform the near-real-time products in extreme precipitation estimation and hydrological simulation (Zhong et al., 2017).

Other satellite precipitation datasets with global coverage exist, such as the Tropical Rainfall Measuring Mission (TRMM) (retired in 2015), Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), a 35+ year rainfall dataset (Chang et al., 2013; Funk et al., 2015; Huffman & Bolvin, 2018). Despite this, however, GPM-IMERG Final run V06 has been observed to provide the best estimates among these datasets. For instance, Le et al. (2020) evaluated the adequacy of precipitation estimates from satellite datasets including CHIRPS, TMPA, GPM-IMERG -V06, and PERSIANN over six basins (typhoons often affect some) in Vietnam; they concluded that GPM-IMERG had the highest performance. GPM-IMERG products are now used to estimate precipitation accompanying mesoscale convective systems such as tropical cyclones (Gutro, 2018; NASA, 2019). Omranian et al. (2018) and Tang et al. (2020) showed confidence in GPM-IMERG products' usefulness in applications such as natural disaster management, water resources management, and flood assessment. The Global Flood Monitoring System uses IMERG products as input for flood detection and forecasting around the world.

## 1.2. Problem Statement

Countries in TC prone zones commonly use design storms as the main rainfall inputs in hydrological models to simulate flood hazard characteristics; however, these synthetic curves were initially meant to determine peak discharge for channel and bridge design. By definition, a design storm is a synthetic estimation of the highest rainstorm over a catchment with a specific magnitude, frequency, and temporal distribution (Krvavica & Rubinić, 2020). Design storms are derived from Intensity-Frequency-Duration (IDF) curves generated from statistical analysis of long-period tipping bucket gauge measurements. Different design storms exist, from simple geometric shapes such as triangles (linear rise and fall from 0 to the peak rainfall) to shapes based on asymmetric probability density functions (Balbastre-Soldevila et al., 2019). The statistical distribution underlying the IDF curves raises the question; how representative are design storms for tropical cyclone rainfall? In addition, design storm assumptions such as the storm duration equalling to the rainfall duration (Berk et al., 2017), rainfall homogeneity across the catchment, and the subjective determination of the critical storm duration (De Paola et al., 2014; Winter et al., 2019) might fail to represent the complex structure of TC rainfall. Design storms are, by definition, single-peaked, while a TC can have multiple moments with peak rainfall. Therefore, these assumptions seem not to convey the TC rainfall complex structures as its spatial and temporal distribution is highly influenced by the TC intensity and motion.

Flood characteristics change with intricate rainfall patterns, making design storms' use for TC-associated flood hazard assessment questionable. Many developing countries in TC-prone zones lack IDF curves associated with TC rainfall (Lumbroso et al., 2011); therefore, they assume that it is represented in the design storm IDF curves. This research recognizes the gap for the need for a rainfall dataset that captures the complex structure of TC rainfall, one that can be assigned with a known probability. The research introduces an innovative approach that utilizes satellite precipitation estimates to generate a tropical cyclone associate rainfall dataset, an intermediate between real storms and design storms, critical for TC-related flood hazard assessment and risk mitigation. The method seeks to categorize the spatial-temporal characteristics of TC rainfall by conducting rainfall time-series clustering. The rainfall signal(s) from the clustering analysis is utilized as the precipitation input in a hydrological model to simulate the resultant flood characteristics. The research selects a case study in Dominica to demonstrate the application of this new approach. Since the

analysis requires a test area for the flood modelling part, a catchment in the south of the island of Dominica was selected for practical reasons because the area was severely affected by the TCs subject to this study. In addition, previous research includes multiple hazard analyses in Dominica, which practically means that the spatial dataset for flood modelling is available (Bout & Jetten, 2020; Serere, 2020; Westen, 2016).

### **1.3. Research Objectives and Questions**

#### **1.3.1. Main Objective**

This research aims to develop an approach that generates a tropical cyclone associate precipitation dataset for improved TC-related flood hazard assessment, solely utilizing TC rainfall data obtained from satellite precipitation estimates.

#### **1.3.2. Specific Objectives**

##### **Specific objective 1:**

To perform rainfall spatial-temporal pattern analysis for the most suitable tropical cyclones.

- Which tropical cyclones passed close to the case study in the period 2015 to 2019?
- Which tropical cyclone fulfills the selection criteria for being classified with the most devastating rainfall?
- What is the rainfall spatial-temporal pattern for the selected tropical cyclone?

##### **Specific objective 2:**

To perform spatial-temporal clustering of the TC rainfall time series to categorize the rainfall characteristics.

- What number of clusters can represent the characteristics of the rainfall time series in the spatial and temporal perspectives?
- What are the key differences between the clusters; in location, duration, amount, and intensity?
- How should the representative precipitation signals be derived from temporal clusters to be useful for flood modelling?

##### **Specific objective 3:**

To evaluate the tropical cyclone associate rainfall dataset by simulating the flood response using a flood characteristics prediction model.

- What selections of cluster representative signals have similarities or differences in their impacts on flood characteristics?
- How do the flood characteristics vary with the selection of cluster representative signals?
- How can the design rainfall dataset be optimized using information from the resulting flood characteristics?

### **1.4. Case Study Area**

The research focused on the Commonwealth of Dominica (capital city: Roseau), an island state in the eastern Caribbean that occupies about 750 square kilometers. Dominica is characterized by a steep and rugged landscape with the highest peak at 1,447 meters in Morne Diablotins (Paul-Rolle, 2014), receiving over 9,000mm annual rainfall averages. As a consequence of its terrain, this highly wet island experiences substantial variations in rainfall due to orographic effects (Barclay et al., 2019). These variations cause heavy pour that eventually impacts its population, mostly settled in the low elevation and coastal areas. In addition, the island is vulnerable to meteorological disasters such as torrential rainfall and powerful winds brought by

TCs that form in the North Atlantic Ocean, especially during the annual June to November hurricane season. In 2015-2019 alone, the lives and livelihoods of Dominica's already vulnerable population were devastated by two major extreme events, leaving thousands homeless and causing losses greater than the island's GDP. Nevertheless, the island is already building efforts for adaptation and resilience towards climate change-driven meteorological events and their consequences, especially the flooding hazard.

In this research, flood characteristics were modelled over a catchment in the south of Dominica (Figure 1-1), with steep and rugged terrain which determines where the flooding occurs. The catchment geomorphology confines the water input, especially in upstream areas, whereby the valley is likely to fill up from both sides. At some point, the river bends towards the east and back to the sea. The river flood plain is near the outflow point, and this is the only location where the floodwater can choose its own paths because of the diversion.

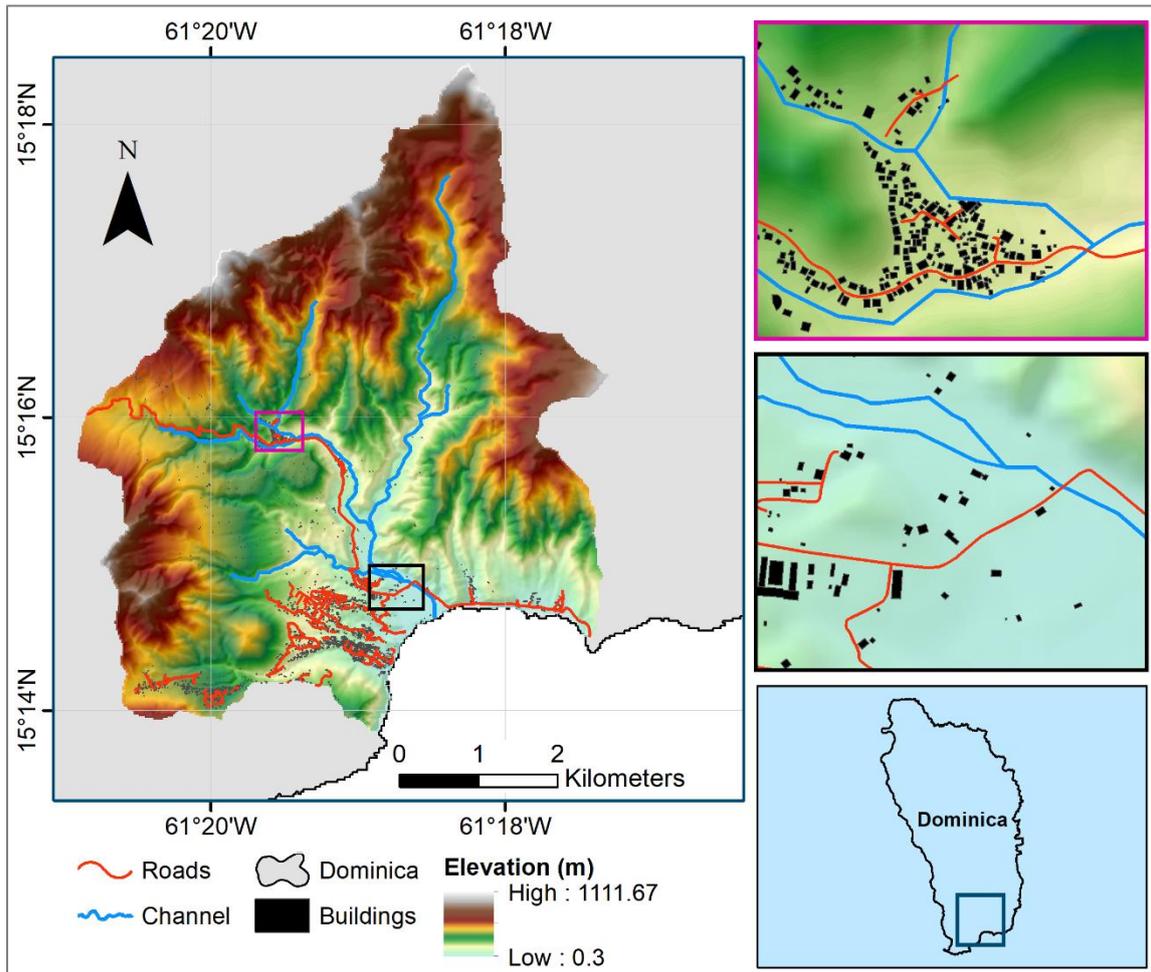


Figure 1-1: Location of the studied catchment in the south of Dominica.

The inserts show that most buildings are located in flood plains (low elevation areas) and are likely to suffer damages during a flood event. The buildings and road data were obtained from OpenStreetMaps.

## 1.5. Research Framework and Overview

The research strategy involves analyzing the spatial-temporal variability of only tropical cyclone precipitation to generate a rainfall dataset for improved TC-related flood hazard assessment. The research framework illustrated graphically in Figure 1-2 is divided into four stages: selecting the candidate TC, spatial-temporal analysis of the rainfall for the selected TC, time-series clustering, and flood characteristics prediction. The

output of the first stage forms the input of the next step; there is an iteration between the clustering and the flood modelling to optimize the building of the rainfall dataset. First (in Chapter 2), a case study location is selected, and all tropical cyclone tracks that passed in its vicinity are mapped; only those that satisfy the selection criteria are chosen. Next, the period of the candidate storm is used to guide the download of satellite precipitation estimates of GPM-IMERG half-hourly data, and an analysis is conducted to extract the pixel rainfall time series. Then, time-series clustering is performed to group the pixel rainfall time series both spatially and temporally. Next, the clusters are analyzed to derive representative signals used as the rainfall information for the flood characteristics prediction (in Chapter 3) over a catchment in the case study. The results of the TC rainfall analysis in stages 1-3 are in Chapter 4. In Chapter 5, the output flood characteristics are examined to guide running an iteration to improve the building of the TC associate rainfall dataset. The sixth chapter presents the discussions and conclusions; suggestions for further research are given in chapter seven.

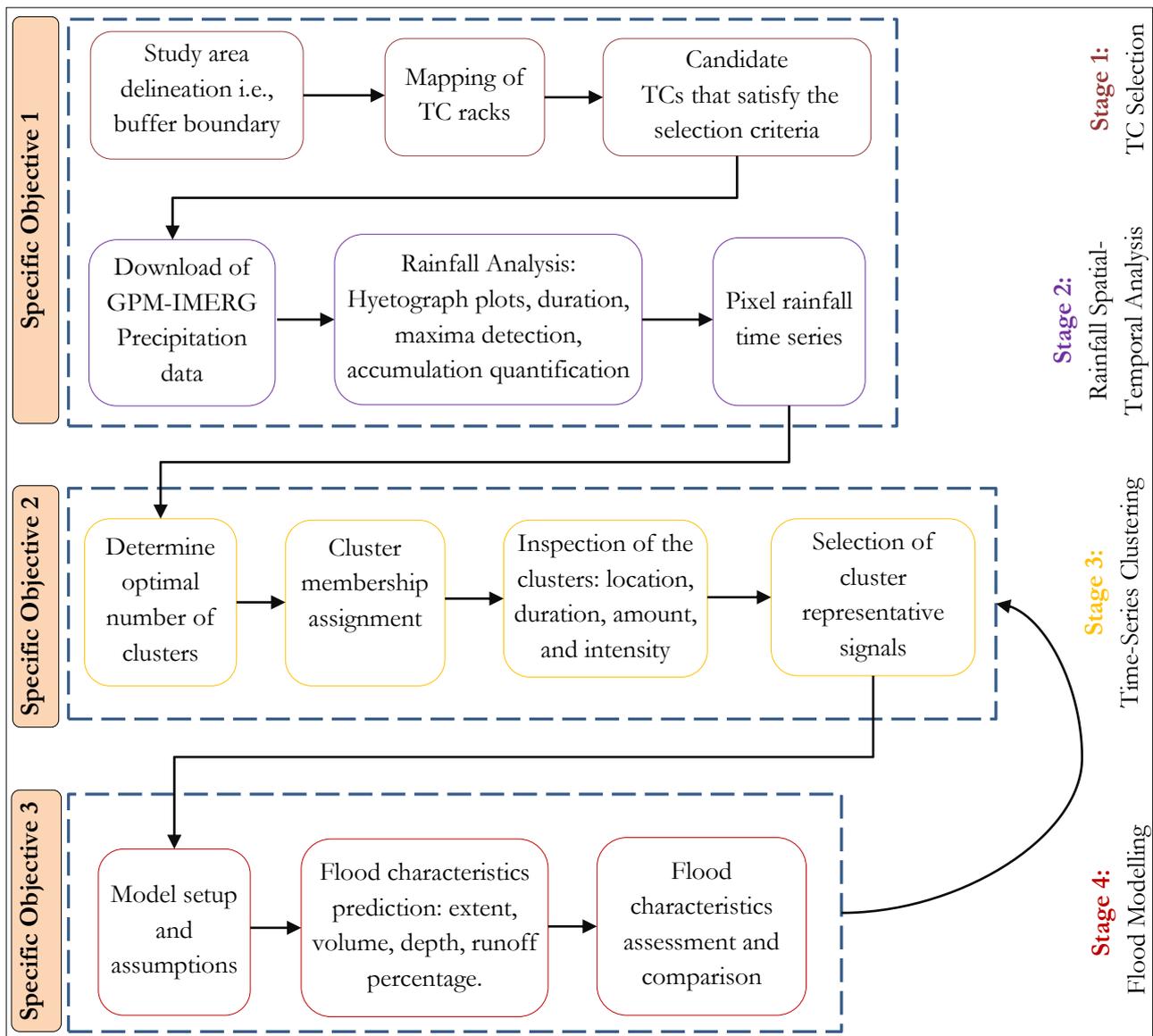


Figure 1-2: Flow chart of the research design and strategy

## 2. SPATIAL-TEMPORAL EVALUATION OF TROPICAL CYCLONE RAINFALL

The research's first and second objectives are to investigate tropical cyclone rainfall patterns and perform spatial-temporal clustering of the rainfall time series to classify their characteristics. This chapter starts with a literature review summarizing the knowledge on tropical cyclones and their accompanying precipitation in section 2.1. This knowledge is utilized to develop the approach for evaluating TC rainfall. Sections 2.2, 2.3, and 2.4 detail the methodology used to select the studied tropical cyclone and analyze its precipitation to obtain specific objective 1. Section 2.5 describes the method for performing the TC rainfall clustering analysis to obtain specific objective 2. The approach for deriving the cluster representative signals is explained in section 2.6

### 2.1. Tropical Cyclones and their Precipitation

NOAA's (2020) definition of a tropical cyclone emphasizes its characteristics of an organized storm system in a rapid circular motion with origin over the tropical oceans and closed low-level atmospheric circulation. The designations of TC systems differ by location (WMO, 2020); however, these very violent storm systems form in the same way. The warm moist air conditions that exist over a warm ocean surface (SST of at least 26°C) within 5° to 30° south and north of the Equator fuel the generation of tropical cyclones. Near-surface disturbances due to a rising column of moist warm air create a zone of low air pressure, and as the air continues to rise, it eventually cools off to form a system of clouds spinning around the low-pressure area (Kristen, 2019). The swirling movement, Figure 2-1a, is caused by the earth's Coriolis effect, whereby the clouds rotate anticlockwise and clockwise in the northern hemisphere and southern hemisphere, respectively. The system of clouds is blown over the ocean by trade winds causing it to grow bigger as it encounters more clouds and rotates faster while releasing more heat energy that powers the storm. The combination and persistence of these conditions for a long time drive tropical cyclone generation (Evans, 2017); on the other hand, the system weakens and dies out over cold water or land when their supply of warm moist air is cut.

The rainbands separated by gaps of no rain, as shown in Figure 2-1b, are capable of causing torrential rainfall over areas beneath the storm, stretching up to many kilometers away from the storm center. Several factors influence the distribution of tropical cyclone rainfall, such as the storm intensity, diurnal cycle, vertical wind shear, atmospheric moisture content, TC motion, and local terrain (Ayala, 2016; Cheung et al., 2018). Yu and Wang (2018) studied the asymmetric and axisymmetric distribution of rainfall accompanying landfalling TCs in China and other controlling environmental factors. They conclude that, unlike the extreme rainfall values, the axisymmetric TC rainfall area, average rate, and total volume increases with TC intensity. Their research mentions the likely association of large rainfall extremes to weaker storms than more intense tropical cyclones. Tropical cyclone intensity is classified based on the system's maximum sustained wind speed to determine the potential impact on society according to the Saffir-Simpson Hurricane Wind Scale (Timothy et al., 2019). However, this categorization does not provide information on the TC-associated rainfall, implying flood hazard assessments may focus on the highest category TCs leaving out those TCs associated with the heaviest rain.

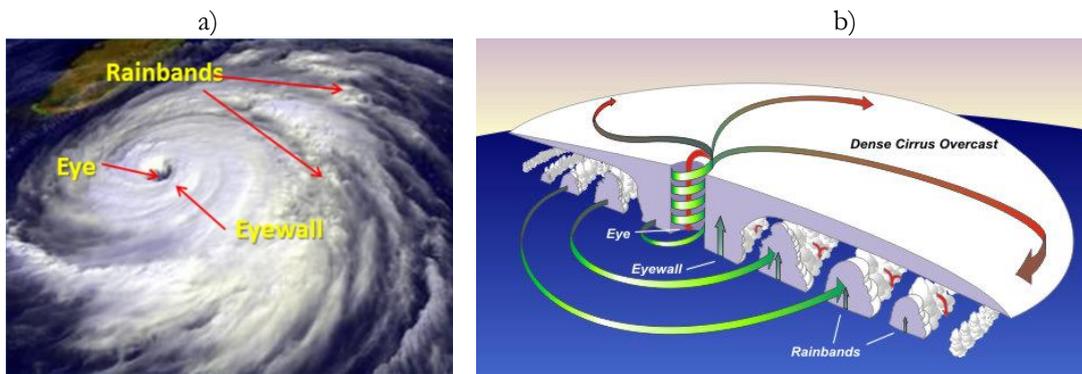


Figure 2-1: Structure of a tropical cyclone.

Satellite view (left) appearing as a thick body of clouds spiralling around the eye. Cross-section of the TC system (right), the upward-facing arrows represent the rising moist air. Source: <https://richhoffmanclass.com/chapter11.html>

## 2.2. Tropical Cyclone Selection

The research focused on tropical cyclone rainfall activity in the vicinity of the case study, Dominica, for 2015 to 2019. This study period was of interest considering the existing evidence of flooding devastation on the island due to tropical cyclone occurrences in the proximity of Dominica's coast. In addition, the availability of GPM-IMERG data (Huffman, Stocker, et al., 2019) was a crucial consideration in selecting the study period. The 500km diameter buffer drawn around Dominica's coast defined the study area boundary and was used as part of the criterion to exclude non-tropical cyclone rainfall (WMO, 2020).

TC best track information from the NHC-HURDRAT2 database was used to map all tropical cyclones which crossed the study area during the study period. The NHC-HURDAT2 database provides six-hourly UTC format best track positions of all tropical cyclones that passed in the Atlantic Ocean from 1851-2019 (Landsea & Beven, 2019). The TC positions and intensity are recorded at precisions of  $0.1^\circ$  latitude/longitude and 5knots, respectively (Landsea & Franklin, 2013). The database provided information on the TC timing, the TC system status changes based on its intensity, the eye location, and the wind radii while crossing the study area. Time information from the NHC-HURDAT2 database was used to select the range of GPM-IMERG files to acquire the rainfall for the selected candidate TCs.

Candidate storms satisfied the condition of traversing the 500km diameter buffer around Dominica with a track length more than the buffer boundary radius and the eye spending an extended time in the study area. The condition assumed that tropical cyclones that spent a long time within the study area could bring heavy rainfall volumes leading to massive flooding. The research's initial interest was in all the candidate storms; however, a decision to test the approach on one TC was made to reduce the data to be analyzed given the time constraints.

## 2.3. Satellite Rainfall Data Acquisition

The utilization of high spatial and temporal resolution satellite precipitation estimates was critical for the spatial-temporal pattern analysis of rainfall associated with the selected TC. Additionally, the precipitation data source needed to have a wide temporal coverage to capture precipitation data over the size of the study area (500km diameter buffer) for the whole period of the selected TC. NASA's GPM-IMERG Final Run Level-3 Half Hourly product (V06) was suitable for this purpose as the dataset provides precipitation estimates on a global coverage at  $0.1^\circ \times 0.1^\circ$  spatial resolution and 30minutes intervals. GPM-IMERG V06 half-hourly data files are stored at 2.8 MB in size and available in GeoTIFF format for GIS analysis. GPM-

IMERG V06 is the latest upgrade from V05; Mekonnen et al. (2021) commend the product's outstanding satellite extreme rainfall estimation and detection capabilities compared with other satellite rainfall products. Additionally, GPM-IMERG V06 can reproduce the diurnal cycle with very high performance (Tan et al., 2019; Tang et al., 2020) and capture and represent mesoscale convective systems (Cui et al., 2020).

In this research, GPM-IMERG V06 data was downloaded for a time window covering the period the TC eye spent in the study area plus two additional days, i.e., the day before the eye entered and a day after it left the study area. The selected time window enabled capturing all the rainfall brought by the TC when its low-pressure center passed the study area. The research only used data to the extent of the buffer boundary and not to the entire length of the TC track because the system's behaviour is likely to change with its lifetime and latitude. Therefore, concentration was only on TC rainfall that poured in the buffer. Half-hourly intervals for the time window of the selected TC implied the download of hundreds of GPM-IMERG V06 files; hence an autogenerated python script from the NASA website was utilized for this purpose.

#### 2.4. Rainfall Inspection for the Selected TC

The statistical examination of the downloaded precipitation data was necessary to generate pixel rainfall time series and gain insight for the clustering analysis. The investigation was performed in RStudio, an open-source tool with a rich library of statistical computation and spatially related packages to handle raster data (Hijmans, 2020). The scripts used in the analysis can be found in GitHub ([MSC-THESIS](#)). First, the raster files were clipped to the study area's extent using an automated process in ArcGIS Model Builder. The files were then combined as a raster stack in RStudio, and their pixel values (rainfall values in mm) were read in a data frame. The temporal rainfall patterns of the selected TC were investigated based on statistical computations for each pixel, including mean, sum, maximum, minimum, standard deviation, and quantiles. Time series were generated for each rainfall pixel to extract information such as the rainfall duration, maximum intensity, and total rainfall accumulation. The spatial patterns were analyzed based on what time the different pixels received rainfall as the selected TC moved across the buffer boundary.

#### 2.5. Rainfall Time Series Clustering

The second objective of this research is to perform spatial-temporal clustering of the TC rainfall pixel time series to make groups of distinct rainfall characteristics. The simplest form of clustering, the traditional one-way clustering (Wu et al., 2020), was implemented to classify the TC precipitation time series to form clusters based on either spatial or temporal similarity. In this research, K-means, the most commonly used traditional one-way clustering method, was employed to conduct spatial and temporal clustering of the TC rainfall. Previous research shows evidence of K-means' broad applicability for rainfall time-series clustering (Alam & Paul, 2020; Hadi et al., 2018; Machiwal et al., 2017). This partitioning clustering algorithm widely uses the Euclidean distance as its similarity measure to assign data to a predefined number of clusters ( $K$ ). Given time series  $X$  and  $Y$ , where  $X = (x_1, \dots, x_n)$  and  $Y = (y_1, \dots, y_n)$ , the Euclidean distance  $D(X, Y)$  between the series was calculated as:

$$D(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots 1$$

##### 2.5.1. Determining the Optimal Number of Clusters ( $K$ )

To determine the suitable value  $K$  of the optimal number of clusters needed to partition the data from the spatial and temporal perspectives, a heuristic approach, the elbow method (Naranjo-Fernández et al., 2020), was utilized. The method was implemented by running the K-means algorithm on the data over a range of

$K$  values, selecting an agglomeration coefficient, and plotting the coefficient against the varying  $K$  values (Ketchen & Shook, 1996). The output is a plot of a decreasing “arm-like” graph with an inflection point (elbow) which eventually flattens as  $K$  increases. The graph’s inflection point is a visual indicator of the most appropriate number of clusters to partition the data (Kodinariya & Makwana, 2013). However, the elbow may not be pronounced in some cases, introducing ambiguity in defining the optimal  $K$  values. In such scenarios, the starting point of a plateau or flattening of the elbow graph indicates the distinctiveness of the output clusters (Ketchen & Shook, 1996; Martin & Sinclair, 2007). The total within-cluster sum of squares, a measure of cluster compactness, was selected as the coefficient and plotted against the different  $K$  values (Syakur et al., 2018). The  $K$  value of optimal clusters should have a cluster compactness measure that is as small as possible.

**2.5.2. Spatial and Temporal Clustering**

For the spatial clustering, the K-means algorithm was run on the whole dataset to group locations with similar behaviour in the precipitation timing. Clustering in the spatial perspective treated locations (pixels) as objects and timestamps as attributes; then, locations were grouped in a single collection if their data elements (rainfall) behaved similarly in the time dimension (Wu et al., 2020). Only the cluster size and the timing of the precipitation were explored in the spatial clusters. The rain starting and ending times for the output clusters were approximated based on readings from their time series plots. The results of the spatial clustering will be shown; however, they were not investigated further because the research’s preference was in the temporal distribution of the TC rainfall.

Clustering in the temporal perspective treated timestamps as objects and the locations as attributes to group timestamps for data elements with similar behaviour across all areas (Wu et al., 2020). Lanfredi et al. (2020) only based on three parameters, including the lower quantile, median, and the upper quantile, to cluster pixel monthly precipitation estimates to assess their similarities in rainfall seasonality. However, in this research, temporal clustering was conducted by running the K-means algorithm over the upper quantile, maximum, mean, standard deviation, 90<sup>th</sup> percentile, and accumulated total. These statistics were selected based on the realization that the rainfall data for all the pixels were in the upper limits since the time series had long silent periods of little or no rainfall. The output clusters represented the varying temporal behaviour of the TC rainfall. Temporal cluster statistics were inspected for differences in rainfall location, duration, amount, and intensity. Temporal clusters characterized with low rainfall accumulation and intensity were regarded as not flood intense in this research; therefore, excluded from further analysis.

**2.5.3. Quality of the Clustering**

The two basic principles of clustering data into homogeneous groups include maximizing similarity within a cluster and minimizing similarity between the clusters (Esma, 2020; Warren, 2005). It was therefore of utmost importance to assess the goodness of the clustering based on these two principles. The K-means algorithm returns outputs, including the cluster size, an array of cluster centers, and cluster validation statistics used to assess the clustering goodness. To evaluate the clustering quality, the percentage ratio of the between sum of squares ( $BSS$ ) to the total sum of squares ( $TSS$ ) was calculated (Soetewey, 2020).

$$\frac{BSS}{TSS} * 100\% \dots\dots\dots 2$$

Where,  $BSS$  is a measure of how well the clusters are separated from each other, and  $TSS$  is a measure of the total variability in the data. The higher the percentage ratio, the more the variability in the data was

accounted for when assigning cluster membership to the observations, implying a higher value for *BSS*, spread out clusters, and therefore a higher quality of the clustering.

## 2.6. Selection of Cluster Representative Signals

The main challenge was to translate the temporal clusters into precipitation signals used as the input into the flood prediction model to evaluate the resultant impacts on the flooding hazard. The cluster representative signals needed to be selected in the most optimal format because they served as a link between the clusters (the different temporal behaviours of the TC rainfall) and the flood hazard modelling. Caution was taken to ensure that the cluster representative signals were realistic in rainfall accumulation, intensity, and duration as observed in the cluster statistics.

Rainfall time series of individual pixels were visualized (as illustrated in Figure 2-2a) to get insight into the range of variability within each cluster. Characteristics like duration, the number of peaks, peak intensity of these individual time series served as the basis for deciding on the most suitable method to derive cluster representative signals. The developed procedure to select the cluster representative signals involved three steps; 1) introducing thresholds to determine the beginning of the TC rainfall, 2) statistical aggregation with the timestep quantiles, and 3) deciding which quantile to use.

Starting thresholds were introduced to remove silent periods and the antecedent rainfall before the storm's start, as illustrated in Figure 2-2b. The use of thresholds was an essential step to align the rainfall series as the TC is a moving system, so the starting time of precipitation varies per pixel within a given cluster. In the second step, descriptive quantiles were computed at each time step, as illustrated in Figure 2-2c. Table 2-1 lists the labels and corresponding probabilities for the statistics computed at each timestep after applying the thresholds determining the onset of the TC rainfall.

Table 2-1: Labels and corresponding probabilities for the descriptive statistics calculated at each timestep.

Label	Q0	Q1	Q2	Q3	Q4	Q5
Probability	0	0.25	0.5	0.75	0.9	1

Finally, the timestep quantiles statistics were explored for each cluster to only select those quantile series with rainfall accumulation, maximum intensity, and duration comparable to the statistics of the temporal cluster. If more than one quantile series were selected in a given cluster, each of these would be used separately as the primary precipitation input into the flood model to simulate the landscape response for the flood hazard analysis. The decision on a cluster's most representative signal (quantile series) would then be the one with the highest impact on flood generation and hence used to build the final TC associate rainfall dataset.

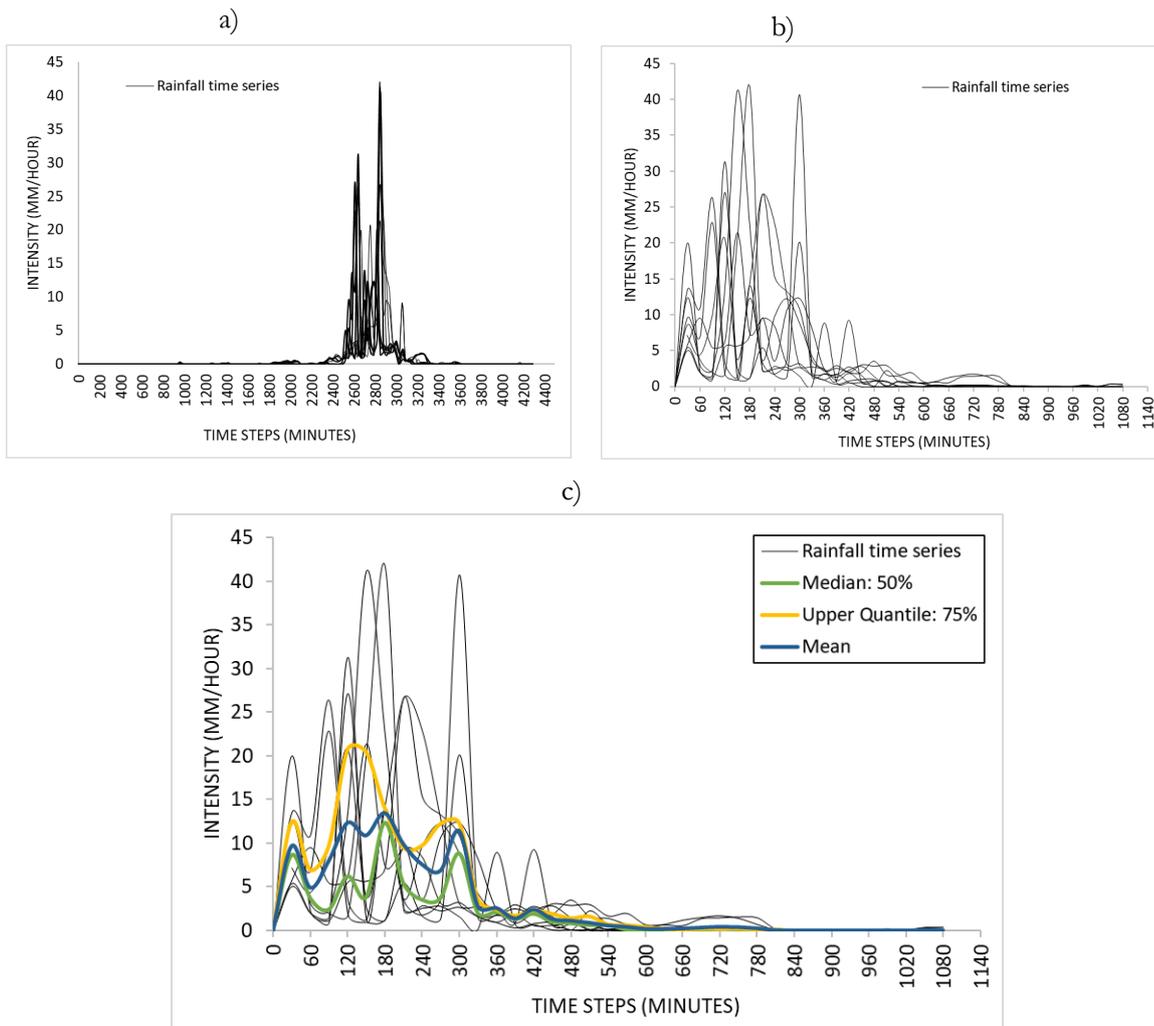


Figure 2-2: Illustration of the procedure for selection of cluster representative signals.  
 a) Visualization of pixel rainfall time series of a given cluster. b) Rainfall series after applying a starting threshold.  
 c) Example for timestep quantile series for probabilities 0.5, 0.75 calculated after applying the starting threshold.

## 3. FLOOD HAZARD MODELLING

This chapter presents the methodology used to obtain the third specific objective; evaluating the tropical cyclone associate rainfall dataset by simulating the flood response using a flood characteristics prediction model. The chapter starts with describing the utilized flood model, the motivation for its choice, and the model's data requirements in section 3.1. The model setup and assumptions applied in this research are in section 3.2. The investigated flood characteristics for the flood hazard assessment are in section 3.3. The last section, 3.4, describes the procedure for optimizing the TC associate rainfall dataset by using information from the flood response to the different rainfall signals.

### 3.1. Flood Modelling in openLISEM

To start with, the choice of the flood model was based on its ability to imitate hydrological water movement in the catchment during an extreme precipitation event (the cluster representative signal) because the interest of the research was on the prediction of the flooding, not the entire hydrological process. Many hydrological models exist, and researchers have evaluated their effectiveness for flood modelling and forecasting (Devia et al., 2015; Unduche et al., 2018; Wijayarathne & Coulibaly, 2020). This research utilized openLISEM, an open-source physically-based numerical model (Bout et al., 2018) developed by ITC-University of Twente. According to Jetten (2016), the model is used to simulate event-based spatial-temporal processes such as runoff, flooding, and sediments, at time steps less than 60seconds for periods between 1-24 hours at a spatial resolution less than 100m grid cells. Figure 3-1 summarizes the sediment and hydrological processes in openLISEM; however, this research only utilized the hydrological part. The sediment processes were, therefore, disabled when running the model. Previous research shows that the openLISEM model has already been used for flood mitigation (Pérez-Molina et al., 2017), flash flood modelling (Nurritasari et al., 2016), and exploration of catchment response to storms of varying magnitude (Baartman et al., 2012).

#### 3.1.1. Data Requirements for openLISEM

Like other flood forecasting models (WMO, 2013), openLISEM is parameterized by assembling several input spatial datasets that affect flood response in a catchment, including rainfall, Digital Elevation Model (DEM), soil characteristics, land use, and infrastructure. The research utilizes input parameter maps created by Bout et al. (2018). The procedure for assembling the model parameter maps involves converting the input spatial data to raster format and resampling the data to match the DEM resolution as it forms the mask for all the basic input maps. PCRaster GIS scripting (<https://pcraster.geo.uu.nl/>) is then employed to automatically create hydrological variable parameter maps from the raster data in a format that openLISEM can understand (Karssenberget al., 2010).

The DEM is used to delineate the catchment boundary and generate derivatives such as the local drainage direction, slope, flow accumulation, and gradient maps. Land use and infrastructure effects are implemented as an input layer because they influence the runoff behaviour and, consequently, the flood magnitude. Land use affects soil characteristics which in turn influences the soil hydrological variables and the infiltration capacity. The land use information is used to derive parameters including surface roughness, canopy storage, Manning's n, and cover. Saxton & Rawls' (2006) pedotransfer functions are combined with soil physical properties to derive parameter maps such as the porosity, average suction, and saturated hydraulic conductivity (ksat). The openLISEM model treats the parameter maps at a sub-grid cell basis, whereby the input layers are read as fraction maps of a given square cell (Bout & Jetten, 2018). All the input parameter

maps that trigger the different hydrological processes are overlaid; the soil physical information map and its derivatives form the base layer (Bout et al., 2018). Then, openLISEM reads through the overlay vertically during the model run to predict a single grid cell's hydrological response (Jetten, 2016).

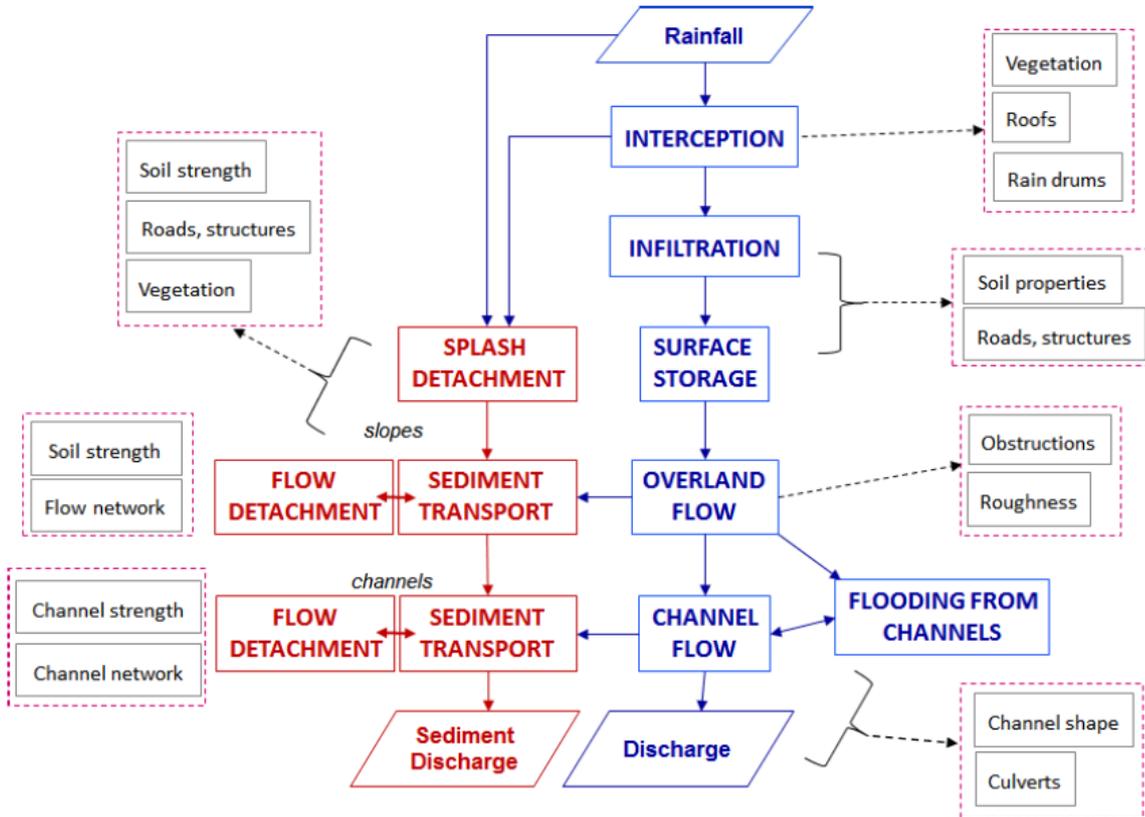


Figure 3-1: Summary of openLISEM sediment and hydrological simulation.

The dashed boxes indicate the main input variables needed in the processes. In this research, only the hydrological part (blue boxes) is utilized. Adopted from Bout et al. (2018) and Jetten (2016).

### 3.2. Model Setup and Assumptions

Dominica is characterized by thin volcanic soils (Rouse et al., 1986), which received close to 200mm of rainfall in the two weeks before Tropical storm Erika (the studied candidate TC). Considering these antecedent rains, Ogden (2016) set the initial soil moisture content at 90% of porosity when modelling floods due to Tropical storm Erika over the island. However, in this research, the initial soil moisture content was set at 85% of saturation to allow for more infiltration.

The research utilized the option where the DEM redirects the water flow downstream by setting the surface flow to a 2D dynamic wave for overland flow and flood. This option does not distinguish between flood and runoff; therefore, a threshold of 0.05m depth was used to serve this purpose. As shown in Figure 3-2, the model's flow processes are such that if runoff water is present at the surface, it is routed towards channels by either a 1D kinematic wave over a user-defined flow network or by a 2D dynamic wave over the DEM. Once in the channel, the discharge is routed with a kinematic wave towards the outlet. If the water level of the channel rises above the channel sides, flooding occurs, triggering a 2D flow of the Saint-Venant equations (Jetten, 2016). However, if the full 2D flow is used everywhere, there is no difference between runoff and floods in terms of flow calculations. Therefore, a user-defined (artificial) water level is set to

differentiate between runoff and flooding. The period of the flood recession usually occurs after the rain has stopped.

Lastly, the rainfall input for the flood modelling was the cluster representative signals stored in separate text files of intensities in mm/hr for time steps of 30 minutes corresponding to the GPM-IMERG temporal resolution. The openLISEM model was run separately for each input rainfall signal to output flood variables for the flood hazard assessment. For each of the runs, the model was set to time steps of 60 seconds.

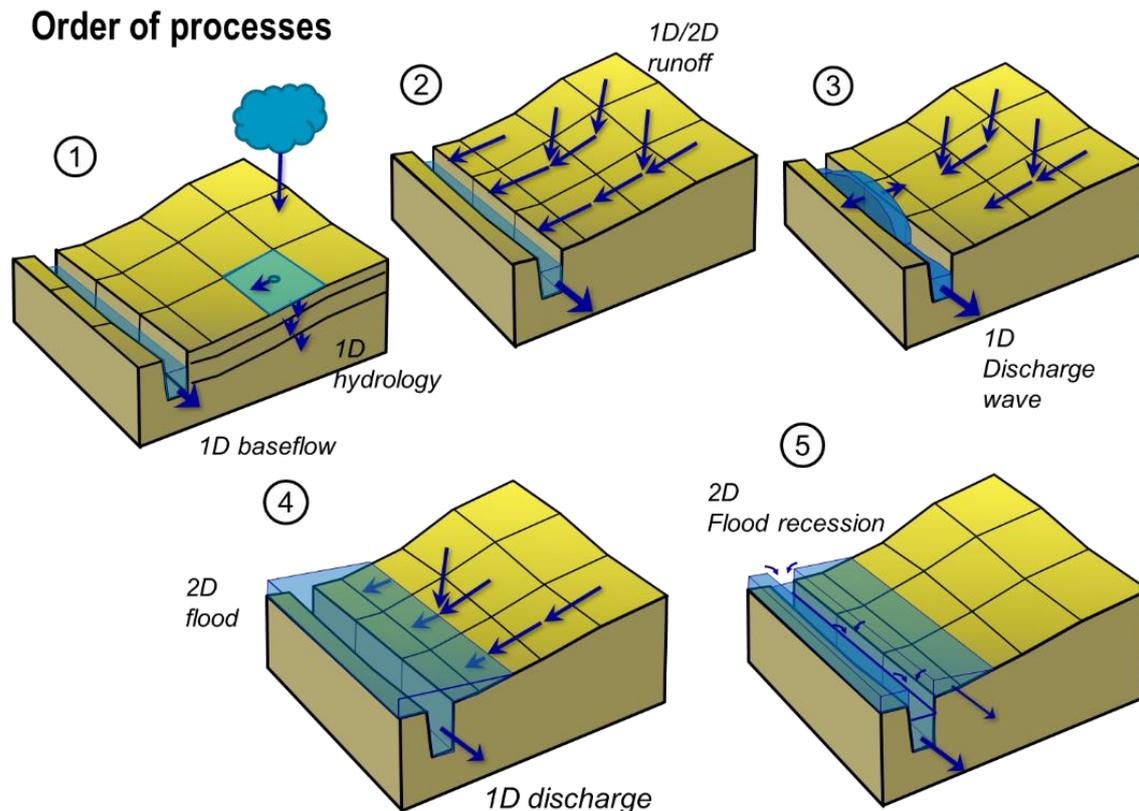


Figure 3-2: Order of flow processes in openLISEM, including overland flow, channel flow, flooding, and flood recession. Adopted from Bout et al. (2018)

### 3.3. Investigated Flood Characteristics

Knowing that the cluster signals represented the different temporal behaviours of the TC precipitation, it was necessary to investigate how these variations of the rainfall characteristics impact the catchment response and hence the flood hazard. The catchment characteristics, the model setup, and assumptions were not changing; only the rainfall input was changing. For each rainfall signal, the research examined a selection of flood characteristics (detailed in the subsections below) commonly considered essential indicators for assessing the impact of a flood hazard (Westen et al., 2011). As required in specific objective three, similarities and differences in the flood characteristics were quantified by examining outputs from the openLISEM model. An investigation based on linear regression of scatter plots was made to show how each flood characteristic varied with the rainfall accumulation and maximum intensity of a given cluster representative signal. The correlation value  $R^2$  and the trend line slope were used to assess the linear relationship. The insights gained from the analyzed flood characteristics guided the subsequent decisions on improving the TC associate dataset.

### **3.3.1. Flood Extent**

The research investigates flood extent by quantifying how much area of land flooded due to the different rainfall events. Depending on the severity of the forcing rainfall event, the flooded land can reach the extent of the flood plains. The nature of the catchment's terrain and rainfall characteristics influence where the flood happens in the catchment; it can be both upstream (valleys) and downstream. The flood extent can be represented on a map to show where the flood happened in the catchment. Knowledge of where and how much area flooded is essential for responders to plan rescue missions and issue safety guidelines to the affected communities.

### **3.3.2. Flood Depth**

Inundation depth is an essential variable in flood hazard assessment to measure how the water rises in the different parts of the catchment during the duration of the flood. The deeper the flood water level, the higher the risk of drowning and damages to housing, among other consequences. The model reports the highest flood level at any pixel during the event because flood hazard analysis is often based on this. However, this highest flood level is not reached everywhere simultaneously as the water moves downstream. Thus, the maximum flood level map is not something that actually exists at any one moment in a flood event. The highest flood level is calculated as the difference between openLISEM outputs of maximum flood volume and maximum flood area.

### **3.3.3. Flood Volume**

Accurate estimation of the flood volume is critical for flood hazard assessment because it is not enough to know where it will flood; the flood water accumulation should be quantified. Large flood volumes can float objects such as cars; they make roads unpassable or even breakthrough structures, especially when flowing at a high velocity. Information on the flood extent is combined with the DEM details to quantify the volume of floodwater flowing through a given area during the flood. The openLISEM model outputs information on the maximum flood volume and its distribution up to various depths; this information was used in the research to compare flood volume amounts resulting from the different rainfall events.

### **3.3.4. Infiltration**

If the infiltration rate is exceeded, all precipitation that falls will flow as surface runoff and trigger a flood. The infiltration rate is influenced by terrain characteristics, soil properties, and land-use changes in the catchment. For instance, the antecedent soil moisture conditions impact the amount of water that can penetrate the soil; the infiltration rate is faster for unsaturated soils and lower for partially saturated soils. Additionally, a catchment with steep and rugged terrain reduces the infiltrated water; therefore, it quickly flows downstream as surface runoff, eventually speeding up the onset of the flood event.

### **3.3.5. Runoff Ratio**

The ratio of rainwater that becomes runoff is influenced by the intensity of the incoming rainfall and catchment properties, including the terrain, soil moisture content, and land use. For instance, extreme precipitation falling on partially saturated soil with low infiltration rates implies more water available for runoff and flooding. The steep terrain of the catchment likely redirects the water into the channel, eventually flowing out into the sea. Fast-flowing runoff can break bridges, uproot trees in its path, and carry boulders and sediments from high to low elevation areas. The openLISEM model outputs information on the total outflow and total precipitation both in millimeters. The ratio of these two variables gives a dimensionless measure of the runoff percentage (Goel, 2011; Ratzlaff, 1994), necessary for watershed management and simulation of peak flow at the outlet.

### 3.3.6. Response time

TC accompanying heavy precipitation instigates rapid flooding with short lead times, consequently limiting the time available to evacuate the vulnerable population and increase the flood risk. As in Marchi et al. (2010), the study quantifies lag time as the measure of the response time, influenced by the catchment complex geomorphologic characteristics. The lag time is a widely used parameter computed to exist between the precipitation that fell on the catchment and the output hydrograph (Abdulkareem et al., 2019; Gericke & Smithers, 1935; Marchi et al., 2010). This time variable reflects the catchment's hydrological response, critical for flood hazard assessment regarding the water storage capacity, time of concentration, and peak discharge. Zhou et al. (2019) describe four ways of calculating the lag time. In this study, the response time was calculated between the peaks of the rainfall intensity and the discharge hydrograph extracted directly from openLISEM output.

### 3.3.7. Flood Duration

The flood duration is indicative of the time that the hazard lasts. Depending on the meteorological forcing and the catchment properties, floods may vanish faster or take an extended period to recede. In addition, the duration of the flood influences the consequences (usually indirect) in the aftermath of the hazard, for example, disruption of business, the spread of waterborne diseases, contamination of freshwater. In this research, average flood duration was investigated for all the rainfall signals for flood hazard assessment purposes.

## 3.4. Optimization Approach based on the Flood Characteristics

As this study develops a method with time series clustering to find TC representative precipitation signals for TC-related flood hazard assessment, optimization was essential to remove redundancy between clusters. In addition, we required that the rainfall signals of the final TC associate rainfall dataset have a distinct flood response. Therefore, we hypothesized that the elbow curve's inflection point was not well-defined, and the resultant flood characteristics were similar for representative rainfall signals of some clusters. In that case, an iteration was performed between the clustering analysis and the flood modelling, as illustrated in Figure 3-3. The approach involved reducing the  $K$  value, rerunning the temporal clustering, and then the flood characteristics prediction.

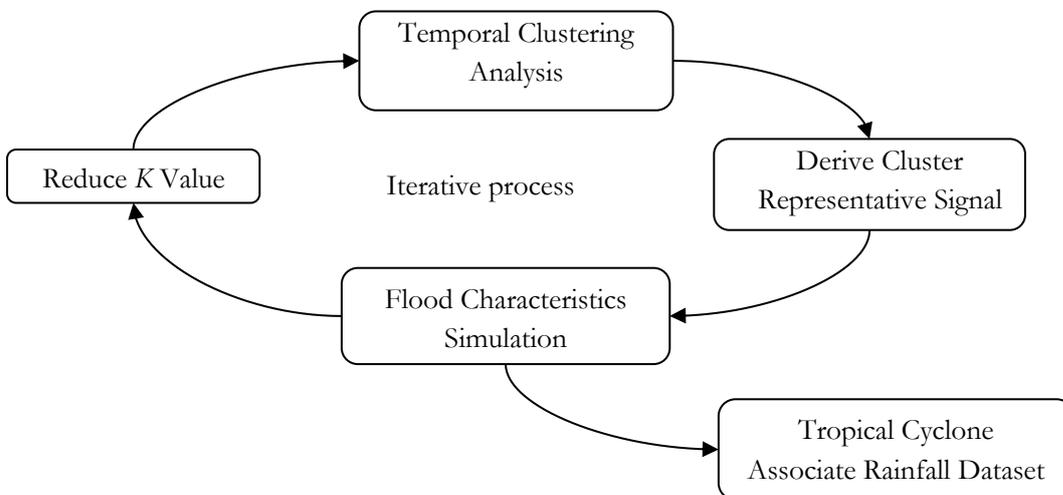


Figure 3-3: Illustration of the approach for optimizing the TC associate rainfall dataset.

## 4. RESULTS: RAINFALL ANALYSIS

The results presented in this chapter were attained by applying the methodology in chapter 2 to obtain the research's first and second specific objectives, i.e., rainfall spatial-temporal analysis and time series categorization. Section 4.1 details the results on the selected TC, including an analysis of its rainfall. The results of the elbow method for determining the most optimal number of clusters are in section 4.2. Sections 4.3 and 4.4 present the results of the spatial and temporal clustering, respectively. Finally, the derived cluster representative signals are in section 4.5.

### 4.1. Selected/Candidate Tropical Cyclones

Figure 4-1 shows that five TCs named by the NHC satisfied the selection criteria for suitable TCs; Tropical storm Erika (Pasch & Penny, 2015), Hurricane Maria (Pasch et al., 2017), Hurricane Isaac (Zelinsky, 2018), Hurricane Beryl (Avila & Fritz, 2018), and Hurricane Dorian (Avila et al., 2019). However, only Tropical storm Erika (hereinafter TS Erika) was selected for further analysis to reduce the data to be analyzed.

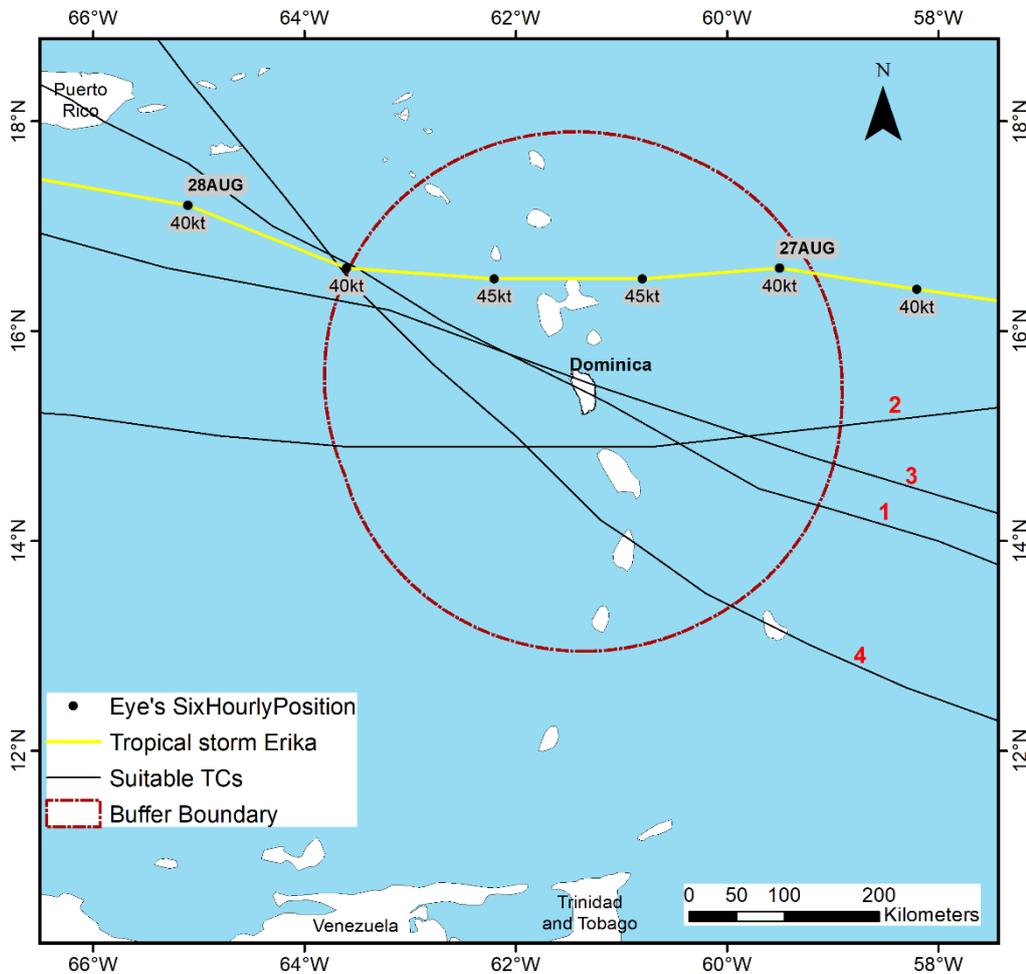


Figure 4-1: Tracks for candidate TCs passing through the study area (500km diameter buffer around Dominica's coast). The TC eye's six-hourly positions and corresponding intensity (knots) are for Tropical storm Erika. The other suitable TCs are 1) Hurricane Maria of 2017, 2) Hurricane Isaac of 2018, 3) Hurricane Beryl of 2018, and 4) Hurricane Dorian of 2019.

The mapped TC track in Figure 4-1 above shows that TS Erika developed over the eastern tropical Atlantic Ocean. The TC eye progressed westward, crossing through the Caribbean islands, the Caribbean Sea, south of the U.S Virgin Islands, before dispersing around the eastern tip of Hispaniola. TS Erika remained of category tropical storm throughout its lifetime, sustaining maximum wind speeds ranging 35-45knots. By 26<sup>th</sup> August 2015 1200UTC, the TC eye was already less than 250km away from the study area buffer, and it was still within 200km on 28<sup>th</sup> August 2015 0000UTC while it progressed further west. TS Erika’s eye entered the study area on 27<sup>th</sup> August 2015 with wind speeds of 40knots; however, the storm’s intensity increased by 5knots just before it touched the northern tip of Guadeloupe about 100km north of Dominica. TS Erika’s wind speeds later dropped back to 40knots just before the eye left the study area. TS Erika’s eye spent about 18hours (00:00-18:00 UTC) in the study area.

**4.1.1. Analysis of Tropical Storm Erika’s Rainfall**

The results in this section are for the spatial-temporal analysis of the selected TC (TS Erika) to obtain the first specific objective of this research. The downloaded 144 raster files of half-hourly precipitation estimate for TS Erika’s rainfall covered the time window between 26<sup>th</sup> August 2015 0000UTC and 28<sup>th</sup> August 2015 2330UTC, capturing all rainfall data including the day before the eye entered and after it left the study area. After the files were clipped to the buffer boundary’s extent, a total of 1911 pixels were analyzed. The number of pixels corresponds to the number of rainfall time series analyzed. As observed in Figure 4-2, two regions, both slightly south of the TC track, received much more total precipitation than others. The islands lie between or south of these two regions, suggesting that the heaviest rain fell over the ocean. The storm’s highest rainfall accumulation was 767.9mm.

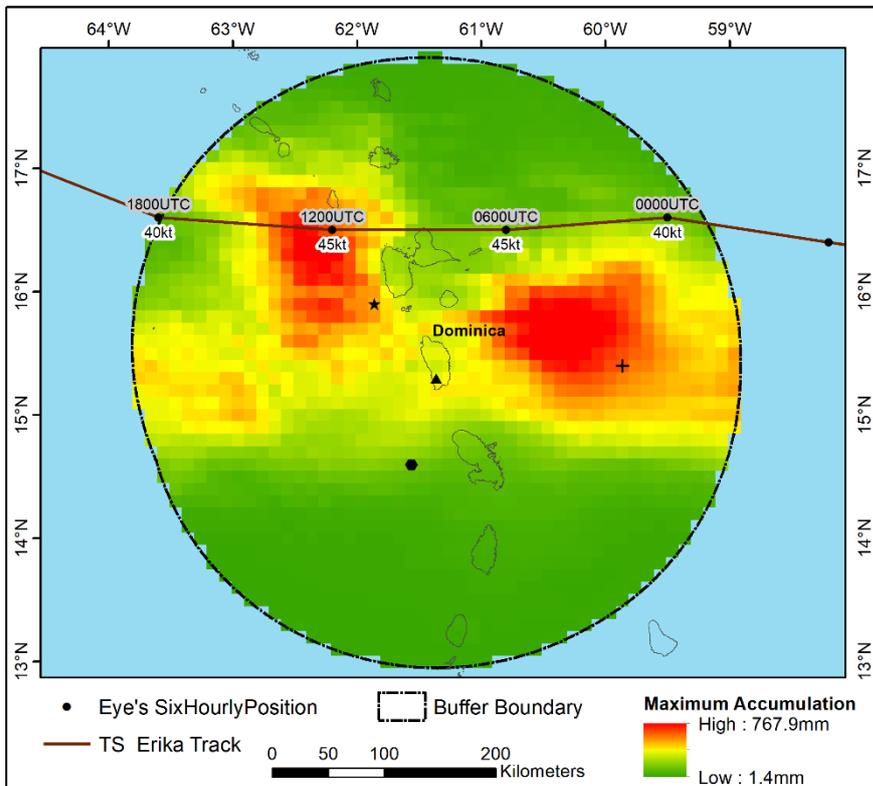


Figure 4-2: Distribution of the pixel rainfall accumulation due to TS Erika’s precipitation over the study area from 26<sup>th</sup> August 2015 0000UTC to 28<sup>th</sup> August 2015 2330UTC. The four shapes, cross, star, triangle, and hexagon, are locations for the pixels whose time series are plotted in Figure 4-3.

Examination of the rainfall time series for individual pixels revealed various rainfall characteristics, including multiple peaks, differences in duration, total amount, timing, and some pixels receiving antecedent rainfall way before TS Erika passed. For example, Figure 4-3 below are time series plots a), b), c), and d) for four randomly selected pixels marked with a cross, star, triangle, and hexagon, in Figure 4-2, respectively. The rainfall total in millimeters for each time series is indicated in the right-hand corner of each graph. Time series plot a) shows there were pixels that received relatively high rainfall intensities followed by a silent period just before the TS Erika's rainfall. The time series in plots a) and b) attained maximum intensities of 85.8mm/hr and 110.4mm/hr, respectively. However, the accumulated rainfall volume was higher for plot a) than for the time series in plot b). Multiple peaks are observed in plot c) and this pixel received relatively high rainfall volumes falling at a lower intensity. In plot d), very low rainfall totals and intensities were observed.

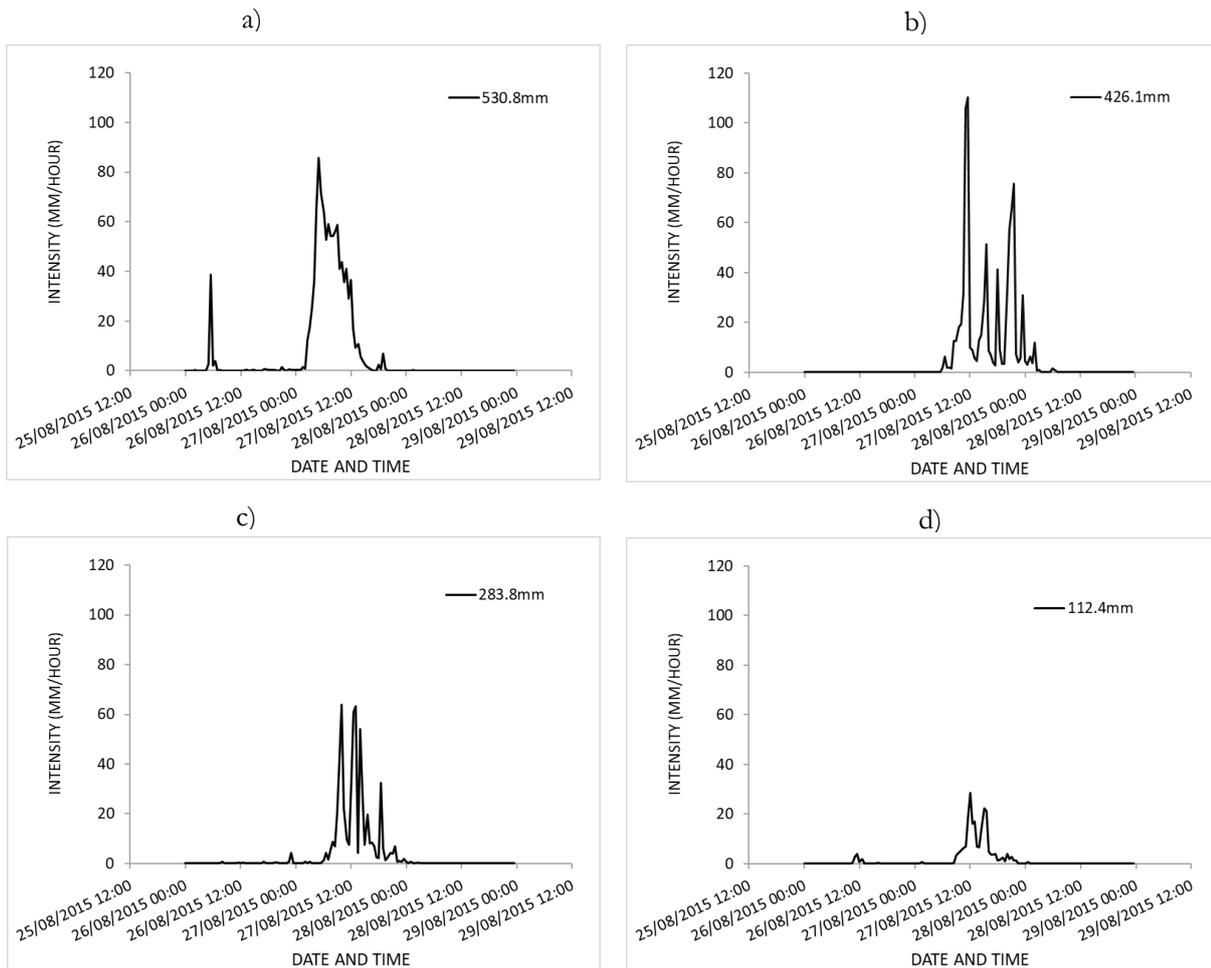


Figure 4-3: Time series plots for individual pixels selected randomly from the dataset.

On the Y-axis is the rainfall intensity (mm/hour). The date and time (UTC) information is plotted on the X-axis. Plots a), b), c) and d) are for the pixels highlighted in Figure 4-2 with a cross, star, triangle, and hexagon, respectively.

## 4.2. Optimal Number of Clusters ( $K$ )

The result obtained from the implementation of the elbow method was the optimal number of clusters needed to categorize the pixel rainfall time series based on the observed characteristics in section 4.1.1. Figure 4-4 plots the total within sum of squares, a measure of the cluster compactness, against varying  $K$  values ranging 1-15 for the spatial and temporal clustering. Examination of both graphs shows the curves' points of inflection (the elbow) were not so pronounced; therefore, the  $K$  values were determined at a

location where the graphs started to flatten as indicated by the blue arrows. As a result,  $K=9$  and  $K=5$  were selected as the optimal number of clusters for spatial and temporal clustering, respectively.

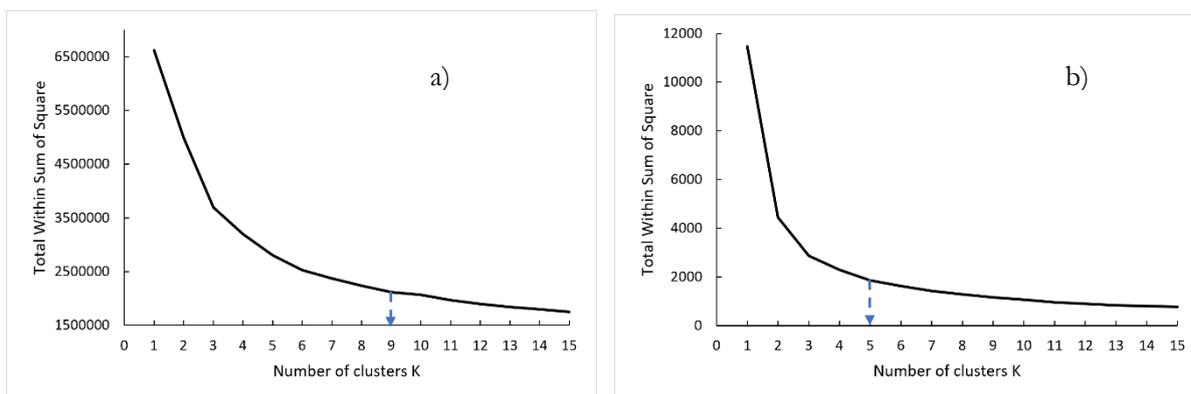


Figure 4-4: Optimal  $K$  values determined at locations where the graph starts to flatten as indicated by the blue arrow. a) Spatial clustering, b) Temporal clustering.

### 4.3. Spatial Clustering Statistics

Based on the elbow method (section 4.2), the spatial clustering partitioned the pixels into nine regions (Figure 4-5) with similar behaviour in the timing of their precipitation, as summarized in Table 4-1. In both the table and the map, the spatial clusters were labelled using the initial S; for instance, S1 is the same as spatial cluster 1. The precipitation timing (start and end) for each spatial cluster was read directly from the time series plots attached in Appendix 1. The spatial partitioning was completed with a clustering goodness measure of 61.8% for the ratio of  $BSS$  to  $TSS$ .

Table 4-1: Statistics for the nine spatial clusters.

The precipitation timing is read column-wise (the first date and time reading are the precipitation starting time and the second reading is the ending time).

Cluster		S1	S2	S3	S4	S5	S6	S7	S8	S9
Cluster size		76	99	100	215	1014	115	121	114	57
Maximum Intensity (mm/hour)	min	65.2	31.2	41.4	27.2	0.4	48	35.6	36.4	81.8
	max	120	95.8	119.4	115.4	111	120	105.4	90	120
Cumulative Rainfall (mm)	min	382	163.9	190.7	140.8	1.4	178.9	167.5	124.1	408.7
	max	767.9	435.5	564	483.6	225.2	583.3	570.2	372.1	724.4
	mean	584.7	333.7	367.4	283.9	51.14	368.7	314.5	240.0	541.1
Date and Time (UTC) August 2015	27 <sup>th</sup>	02:00	0:00			00:00	01:00	05:30		
		16:00	4:30	1:00	15:30		7:00	15:30	15:30	14:00
	28 <sup>th</sup>			2:00	02:30	01:00			00:03	03:00

As observed in Figure 4-5, spatial clusters east of Dominica (S1, S2, S6, S7) received rainfall mostly on 27<sup>th</sup> August 2015, while regions to the west (S3, S4, S8, S9) received rainfall between 14:00UTC on 27<sup>th</sup> up to around 03:00UTC on 28<sup>th</sup> August 2015. S5, the largest spatial cluster occupying 53.1% of the study area, is spread to the north and south of the TC track. Areas in S1 and S9 received the heaviest rainfall, while S5 regions accumulated the least precipitation. S2 are those areas that were first to receive rainfall immediately when the eye entered the buffer; however, by 14:30UTC, they received little or no rain, as shown in

Appendix 1. S4, the second-largest cluster, includes locations that only started receiving TC rainfall at around 15:00UTC on the 27<sup>th</sup>; their rain stopped after 14 hours at around 02:30UTC on 28<sup>th</sup> August. Spatial cluster 3 (S3) received rain on 27<sup>th</sup> and 28<sup>th</sup>; however, rainfall in this region started earlier than the other areas west of the buffer.

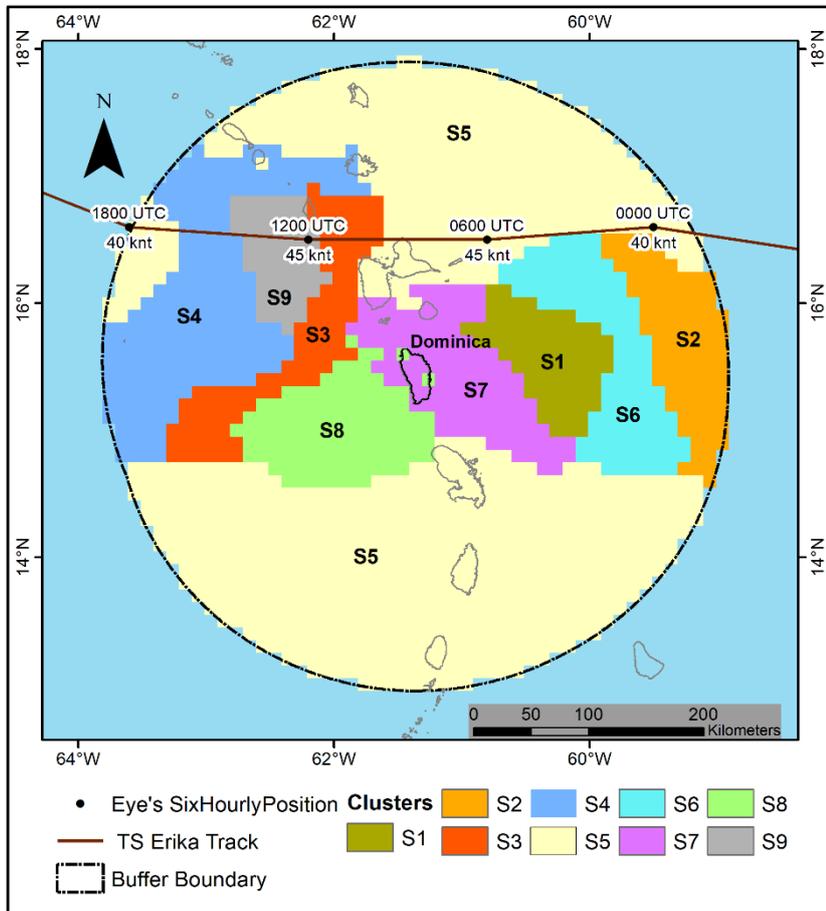


Figure 4-5: A map representation of the spatial clustering of TS Erika's rainfall.

#### 4.4. Temporal Clustering Statistics

Temporal clustering revealed the similarities in the temporal rainfall distribution across pixels in different locations of the buffer. Additionally, the individual temporal clusters represented the distinct temporal behaviours of the rainfall patterns accompanying TS Erika. Temporal clustering results are first presented for  $K=5$ , which was determined as optimal when using the elbow method in section 4.2. The temporal clustering results for a reduced value of  $K$ , i.e.,  $K=4$  and  $K=3$ , are presented to be later used in the optimization approach, as illustrated in Figure 3-3. The clustering goodness ( $BSS/TSS$ ) measures were 83.9% for  $K=5$ , 80.1% for  $K=4$ , and 75.1% for  $K=3$ . The temporal clusters are labelled with the initial T; for instance, T1 is the same as temporal cluster 1.

##### 4.4.1. Using Five Temporal Clusters ( $K=5$ )

The map in Figure 4-6 represents the geographical distribution of the five temporal clusters. It is observed that pixels in Dominica have a temporal distribution corresponding to T3, T4, and T5. The southern catchments of the island experienced precipitation distribution as in T3. Statistics in Table 4-2 show that temporal cluster 1 (T1) is the most extreme, comprising pixels that received the heaviest half-hourly rainfall

with maximum intensities ranging from 55-120mm/hr. Dominica was not hit by pixels from T1; most of the rain poured over the ocean. In addition, Figure 4-6 shows that the temporal clusters are spatially distinct; for example, T1 comprises blocs of pixels in the east and the west of the buffer. Likewise, T2 and T3 contain spatially distinct pixels of similar rainfall temporal distribution. T2 is the largest cluster; it comprises 41.8% of the rain pixels, mostly located away from the storm track, especially those far south of the study area (Figure 4-6). Further examination of corresponding statistics for T2 in Table 4-2 shows that these pixels received the lowest rainfall intensity and total volumes. Therefore, T2 was not considered flood-prone in this research and was excluded from further analysis.

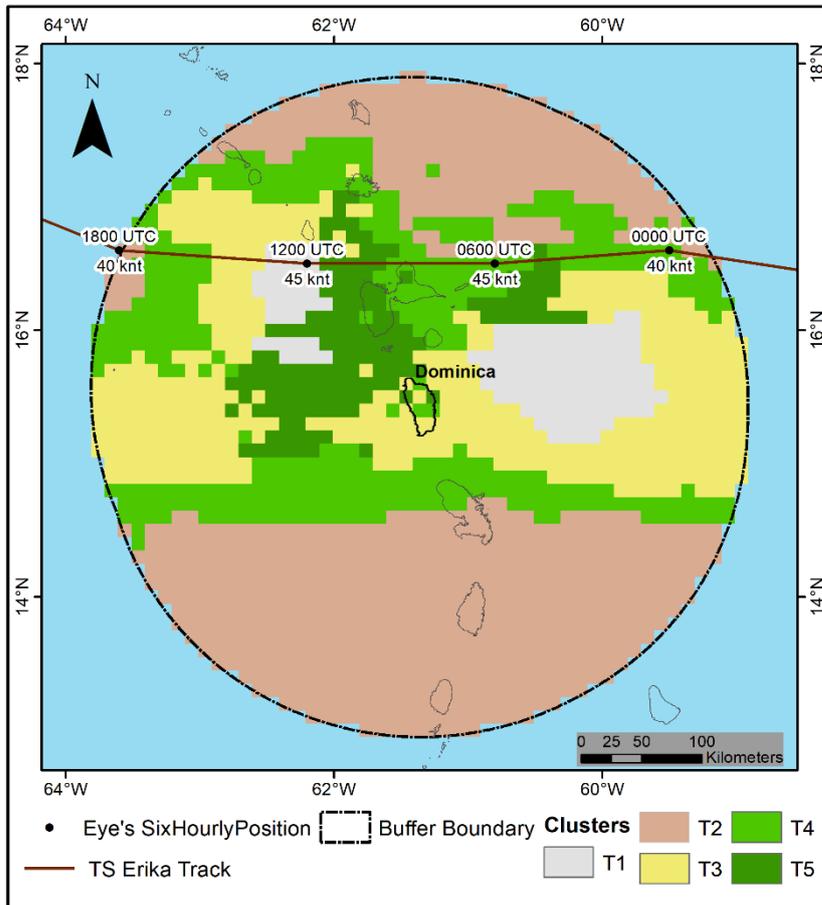


Figure 4-6: A map representation of the temporal clustering of TS Erika rainfall using  $K=5$

Table 4-2: Summary statistics for the temporal clusters resulting from  $K=5$

Cluster		T1	T2	T3	T4	T5
Cluster size (%)		132 (6.9%)	798 (41.7%)	433 (22.7%)	399 (20.9%)	149 (7.8%)
Maximum Intensity (mm/hour)	min	55	0.4	36.4	18.2	32
	max	120	50.2	120	111	120
Accumulated Rainfall (mm)	min	451.5	1.4	201.1	47.8	121
	max	767.9	112.4	571.5	275.9	603.2
	mean	583.94	29.41	344.17	167.46	331.4

Figure 4-7 are box plots showing the distribution of the maximum rainfall intensity and cumulative rainfall per cluster. Most pixels in T3 experienced maximum intensities above 40mm/hr but below 75mm/hr, also depicted in the corresponding range of relatively low accumulated volumes compared to T1. Most pixels in this T1 experienced maximum rainfall intensity above 80mm/hr, as shown in Figure 4-7a. T4 is the second less intense cluster, including pixels with low precipitation accumulation that pours at lower intensities. However, T4 has a few pixels that experience high maximum intensities compared to its members. The pixels in T5 experience relatively higher intensities than T3; however, their mean accumulated volumes are almost similar.

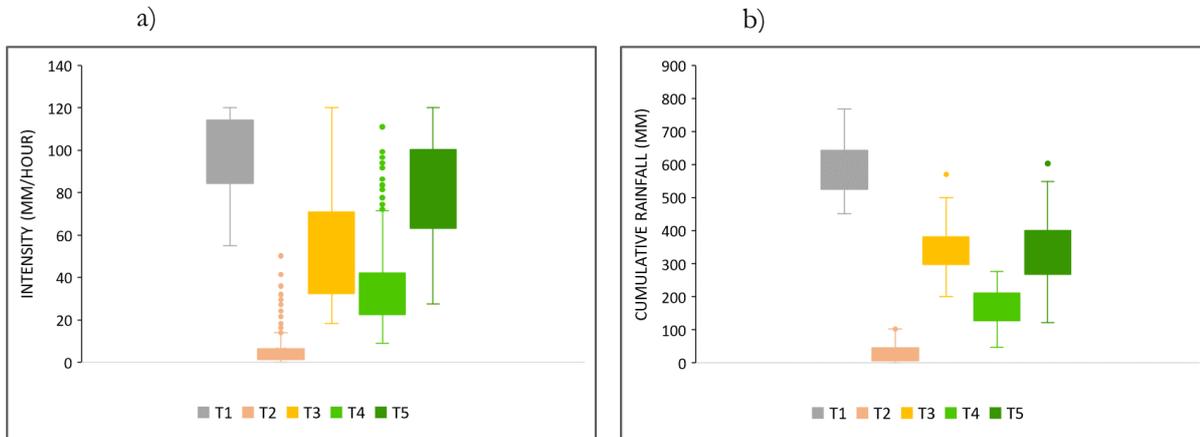


Figure 4-7: Rainfall distribution across all clusters. a) Pixel maximum rainfall intensity b) Cumulative rainfall.

For each box plot, the pixel distribution is such that the lower horizontal line (minimum, 0%), the bottom of the box (lower quartile, 25%), top of the box (upper quartile, 75%), upper horizontal line (maximum, 100%), points (outliers), and vertical lines (lower and upper 25% of the data).

#### 4.4.2. Using Four Temporal Clusters ( $K=4$ )

For four temporal clusters, the map and key statistics are shown in Figure 4-8 and Table 4-3. Compared to the result of five temporal clusters in subsection 4.4.1; we observe that pixels that made up T5 when using  $K=5$  were redistributed to make only four clusters. The boundaries of the least intense clusters in both cases, i.e.,  $K=5$  and  $K=4$ , vary by minor differences. Additionally, pixels over Dominica now belong to only two temporal clusters. From Table 4-3, we observe that T3 comprises some pixels that received high intensities up to 50.2 mm/hr; however, the cluster's range of accumulated rainfall is low and therefore regarded as not flood intense. The two spatially distinct blocs of the most extreme cluster have more pixels as compared to using  $K=5$ . Figure 4-8 shows that the clusters buffer each other in their order of decreasing magnitudes. For example, T2's pixels are surrounded by pixels of T4, T1 surrounds T4, and in turn, T1 is surrounded by T3.

Table 4-3: Summary statistics for the temporal clusters resulting from  $K=4$

Cluster		T1	T2	T3	T4
Cluster size (%)		434 (23.9%)	146 (7.6%)	818 (41.7%)	513 (26.8%)
Maximum Intensity (mm/hour)	min	18.8	55	0.4	42
	max	111	120	50.2	120
Accumulated Rainfall (mm)	min	47.8	433.7	1.4	202.2
	max	275.9	767.9	132.1	571.5
	mean	180.05	574.32	31.127	347.59

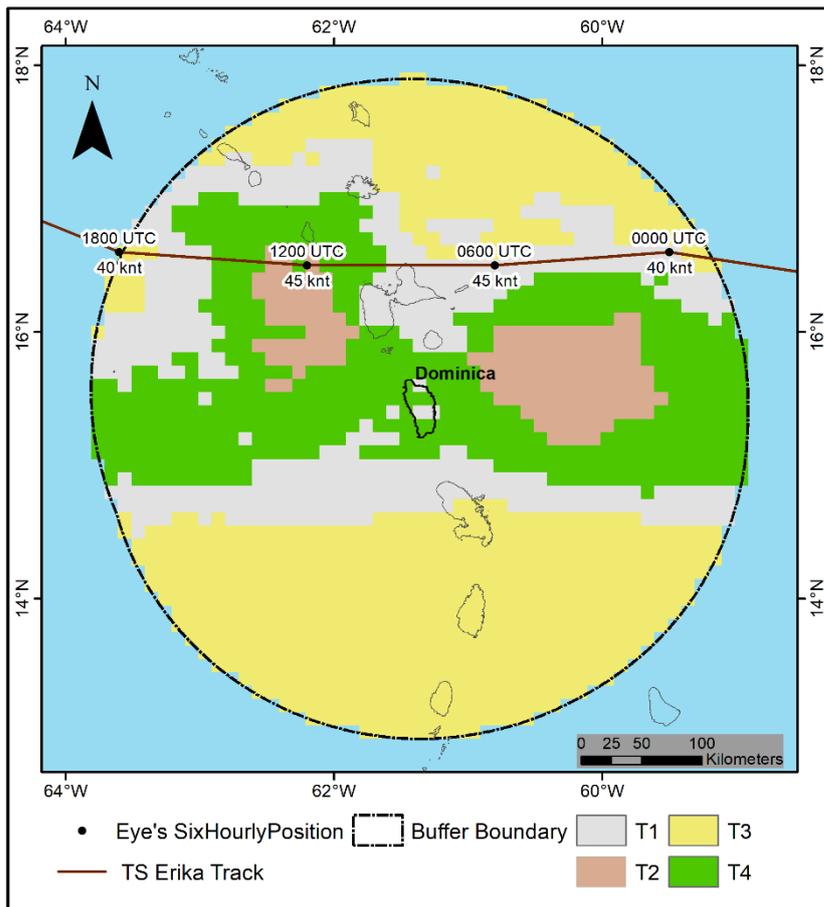


Figure 4-8: A map representation of the temporal clustering of TS Erika rainfall using  $K=4$

#### 4.4.3. Using Three Temporal Clusters ( $K=3$ )

For three temporal clusters, the map and key statistics are shown in Figure 4-9 and Table 4-4. Compared to the result from  $K=5$  and  $K=4$ , pixels over Dominica are only in one cluster. Most pixels of T3, T4, and T5 resulting from  $K=5$  were likely merged to form one big cluster when using  $K=3$  (Figure 4-9). Likewise, most pixels in T1 and T4 resulting from  $K=4$  were put together to create T3 when using  $K=3$ . It is still observed that the boundary of the least extreme cluster, in this case, T1 changes marginally. T1 was excluded from further analysis considering its statistics in Table 4-4, showing very low accumulated rainfall volumes compared to the other clusters. Figure 4-9 below shows that the pixels that make up T2 are surrounded by member pixels of T3. Consistently, the pixels of the T1 are surrounded by those in T3.

Table 4-4: Summary statistics for the temporal clusters resulting from  $K=3$

Cluster		T1	T2	T3
Cluster size (%)		937 (49%)	212 (11.1%)	762 (39.9%)
Maximum Intensity (mm/hour)	min	0.4	55	26
	max	68.6	120	120
Accumulated Rainfall (mm)	min	1.4	352.5	83.6
	max	171.9	767.9	489.2
	mean	42.702	530.79	279.82

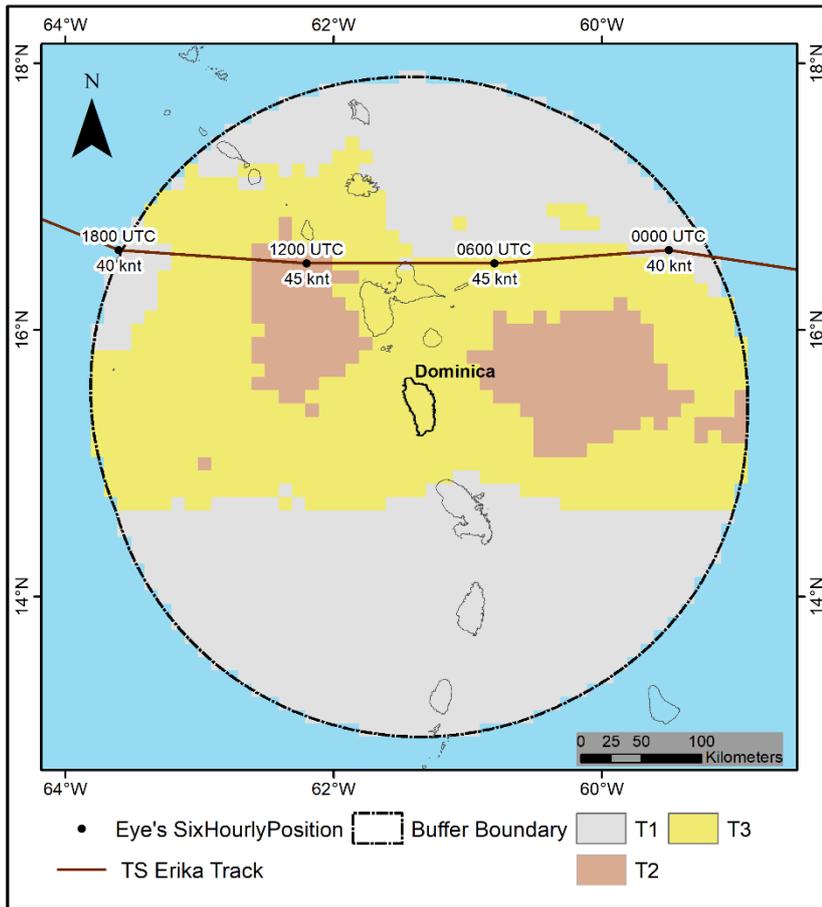


Figure 4-9: A map representation of the temporal clustering of TS Erika rainfall using  $K=3$

#### 4.5. Cluster Representative Signals

The derived cluster representative signals summarize the temporal characteristics of the individual clusters in terms of their rainfall accumulation, maximum intensity, and duration. The results for the selected representative signals are presented in detail first for the temporal clusters resulting from  $K=5$  (explained in detail in subsection 4.5.1). Then, the statistics for the representative signals of the temporal clusters resulting from  $K=4$  and  $K=3$  are presented in subsections 4.5.2 and 4.5.3.

##### 4.5.1. Using Five Temporal Clusters ( $K=5$ )

Cluster representative signals were derived for T1, T3, T4, and T5 since T2 was disregarded from further analysis (see subsection 4.4.1). The detailed explanation uses T1 as an example to demonstrate the procedure applied for all the clusters.

Figure 4-10a below is a plot of the 132-member time series for temporal cluster 1 (T1). This cluster comprises two blocs of spatially distinct pixel time series with similar behaviour in their temporal distribution. The black bloc has a rainfall duration lasting about 16 hours, and the blue block lasts a bit less than this duration. Observe that the black bloc has some pixel time series with antecedent rainfall of intensities reaching almost 40mm/hr. Referring to Figure 4-6, the black bloc are series for T1 pixels in the east and the blue bloc for pixels west of the study area. Calculating the time step quantiles for these series without applying a starting threshold outputs signals far from realistic regarding the cluster rainfall duration

and accumulated volumes since the series starting time differs. Figure 4-10b shows that the rainfall duration was doubled, and a second peak was observed because of the spatially distinct blocs of pixels.

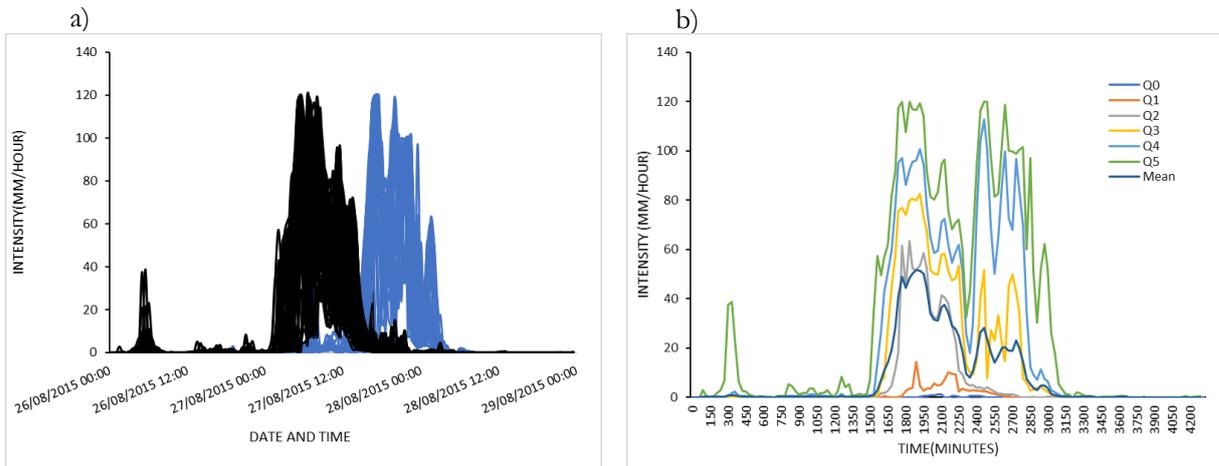


Figure 4-10: a) Plot of T1's spatially distinct time series with similar temporal behaviour. b) Timestep quantile series of T1 based on the original data (without applying a starting threshold).

As the storm's starting period was observed to vary, the beginning of the TC precipitation was tested for different intensities, including 2, 5, 10, and 20mm/hr. Table 4-5 summarises the observations when applying different starting thresholds (step 1 as explained in section 2.6) to determine the start of TS Erika's rainfall per pixel. The corresponding plots of the timestep quantiles series in Figure 4-11 show that the thresholds removed the silent periods and short antecedent rainfall before the start of the storm.

Table 4-5: Observations in the characteristics of timestep quantile series at varying starting thresholds.

Starting Threshold $\geq$	Remarks
1mm (2mm/hr)	Figure 4-11a: Q4 has two peaks, and the rainfall duration is still longer than the observed 16 hours in Figure 4-10a.
2.5 mm (5mm/hr)	Figure 4-11b: There are still pixels that receive antecedent rainfall intensities over 5mm/hr before the storm begins (see Figure 4-10a), causing the upper quantiles to have two peaks. Lower quantiles are almost aligned, and their duration is nearly comparable to the cluster duration (about or more than 6 hours).
5mm (10mm/hr)	Figure 4-11c: When the storm's start is raised to 10mm/hr, the second peak on Q4 is dissolved and almost the same period as the lower quantiles. However, Q5 is still double-peaked; there is a likelihood of a pixel(s) that receives antecedent rainfall above this threshold before the actual TC rainfall begins. This threshold attempts to bring back the cluster temporal behaviour, i.e., a steep increase, then the rain dying down with a duration of about or slightly more than 16hours.
10mm (20mm/hr)	Figure 4-11d: At this threshold, Q5 still shows a pixel(s) with a peculiar behaviour of antecedent rainfall intensities above 20mm/hr before the start of the storm. The starting rainfall value can be raised further to remove this effect.

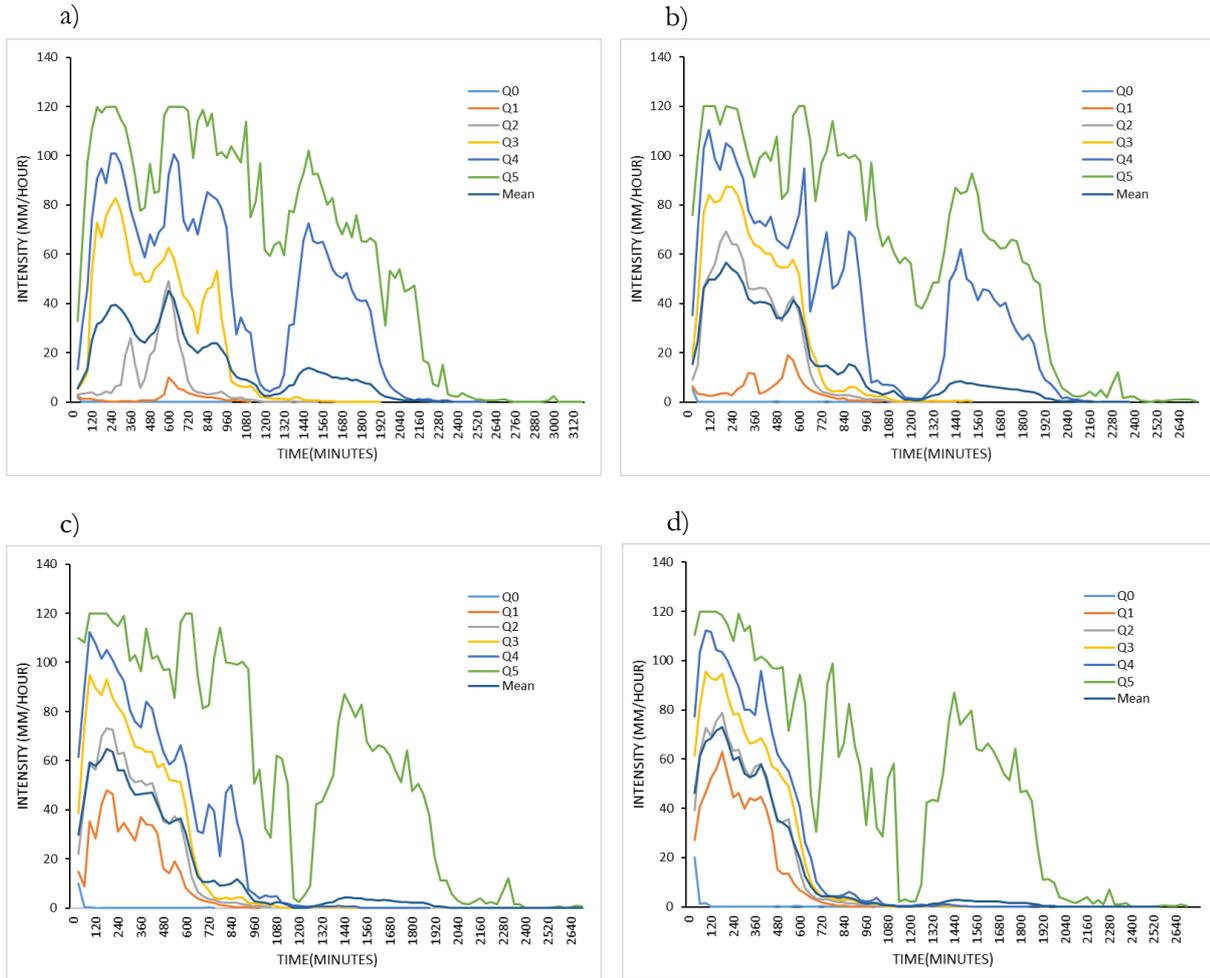


Figure 4-11: Plots of T1 timestep quantile series using varying starting thresholds.

a) Threshold  $\geq 1$  mm (2mm/hr). b) Threshold  $\geq 2.5$  mm (5mm/hr). c) Threshold  $\geq 5$  mm (10 mm/hr). d) Threshold  $\geq 10$  mm (20mm/hr). The plots for T3, T4 and T5 are in Appendix 3.

Figure 4-11 shows that varying the rainfall starting thresholds brings the characteristics of the calculated timestep quantiles (step 2 as explained in section 2.6) to realistic values. The resulting precipitation accumulation, duration, and maximum intensities become comparable to the cluster statistics, and the peaks of the timestep quantile series move to an almost similar timing. Table 4-6 summarizes the timestep quantile series statistics. From the table, we observe that rising the threshold values brings the accumulated totals up for the lower quantiles and down for the higher quantiles. The accumulated rainfall and duration for Q0, Q1, Q4, and Q5, are overly low or high and outside of the ranges of T1's statistics in Table 4-2. As Q2, Q3, and Mean were in line with the observed T1 characteristics; they were selected as cluster representatives to use as input for flood modelling in OpenLISEM (step 3 as explained in section 2.6). However, the selection was made only for the timestep quantiles series starting at 10mm/hr. This threshold better removes the non-TC antecedent rainfall and regenerates the cluster's temporal behaviour, as explained in Table 4-5. The procedure of varying starting thresholds for the pixel time series was repeated for T3, T4, and T5, and a similar selection of representative signals like in T1 was made. The time series plots for T3, T4, and T5 are in Appendix 2, and the graphs for the timestep quantiles at varying thresholds are in Appendix 3.

Table 4-6: Signal statistics computed for timestep quantiles series of T1 at varying thresholds. The statistics for T3, T4, and T5 are in Appendix 4.

Statistic		Q0	Q1	Q2	Q3	Q4	Q5	Mean
<b>Original Data (without threshold)</b>	Accumulation (mm)	4.0	60.1	407.3	905.7	1533.8	2240.5	583.9
	Maximum Intensity (mm/hr)	1.2	14.2	63.5	82.8	112.6	120.0	51.7
	Duration (Hours)	9.0	23.0	27.5	46.5	52.5	70.0	70.0
<b>Threshold &gt;=1 mm</b>	Accumulation (mm)	1.0	33.0	204.9	825.7	1820.3	3144.1	581.7
	Maximum Intensity (mm/hr)	2.0	9.9	49.0	83.0	100.8	120.0	45.1
	Duration (Hours)	0.5	18.5	25.5	32.0	43.0	53.0	39.5
<b>Threshold &gt;=2.5 mm</b>	Accumulation (mm)	2.7	83.8	490.3	722.1	1527.2	2707.0	578.1
	Maximum Intensity (mm/hr)	5.0	19.0	69.4	87.5	110.6	120.0	56.6
	Duration (Hours)	12.5	17.5	18.0	25.5	36.0	45.5	39.5
<b>Threshold &gt;=5 mm</b>	Accumulation (mm)	6.0	289.0	<b>518.9</b>	<b>729.9</b>	1029.1	2502.2	<b>574.6</b>
	Maximum Intensity (mm/hr)	10.0	48.0	<b>73.1</b>	<b>94.9</b>	112.4	120.0	<b>64.8</b>
	Duration (Hours)	12.5	16.5	<b>17.5</b>	<b>25.0</b>	31.5	45.0	<b>45.0</b>
<b>Threshold &gt;=10 mm</b>	Accumulation (mm)	11.7	374.0	544.2	721.6	879.7	2210.5	569.0
	Maximum Intensity (mm/hr)	20.0	63.0	78.7	95.5	112.4	120.0	73.1
	Duration (Hours)	12.0	16.5	17.0	24.0	31.5	45.0	38.0

Figure 4-12 below plots the time series and cumulative rainfall of the representative signals for T1, T3, T4, and T5 for the 10mm/hr starting threshold. T1, T3, and T4 signals are similar in shape; whereas, their duration and maximum intensities vary significantly. The representative signals for T5 behave very differently as its pixels experienced rainfall with multiple peaks but still at a lower magnitude than T3. It is observed in Table 4-7 that TS Erika rainfall generally poured for less than 24hours for Q2 and Q3, however, with varying intensities and total volumes across the study area. The duration of the signal at the Mean is slightly longer as the pixels that received antecedent rainfall of very high intensities close to an average of 12hours before the storm's start (for example, Figure 4-10a) likely affected its period. Based on the statistics in Table 4-7, it is observed that some representative signals behave almost similarly in duration, intensity, and total accumulation. Nevertheless, the rainfall signals were input in the flood model to reveal whether they are distinct enough to lead to different flood characteristics.

Table 4-7: Summary statistics for signals Q2, Q3, and Mean for the temporal clusters resulting from K=5

	T1			T3			T4			T5		
	Tr (mm)	Imax (mm/hr)	Dr (hr)									
<b>Q2</b>	518.9	73.1	17.5	307.1	50.2	15.5	107.6	20.2	14.5	159.7	24.6	18.0
<b>Q3</b>	729.9	94.9	25.0	440.0	63.8	17.5	209.9	38.9	20.5	427.3	61.6	22.5
<b>Mean</b>	574.6	64.8	45.0	334.6	45.4	37.0	155.9	24.8	27.5	317.7	37.2	38.0

Tr (mm): Accumulation (mm) Dr (hr): Period (Hours)

Imax (mm/hr): Maximum Intensity (mm/hr)

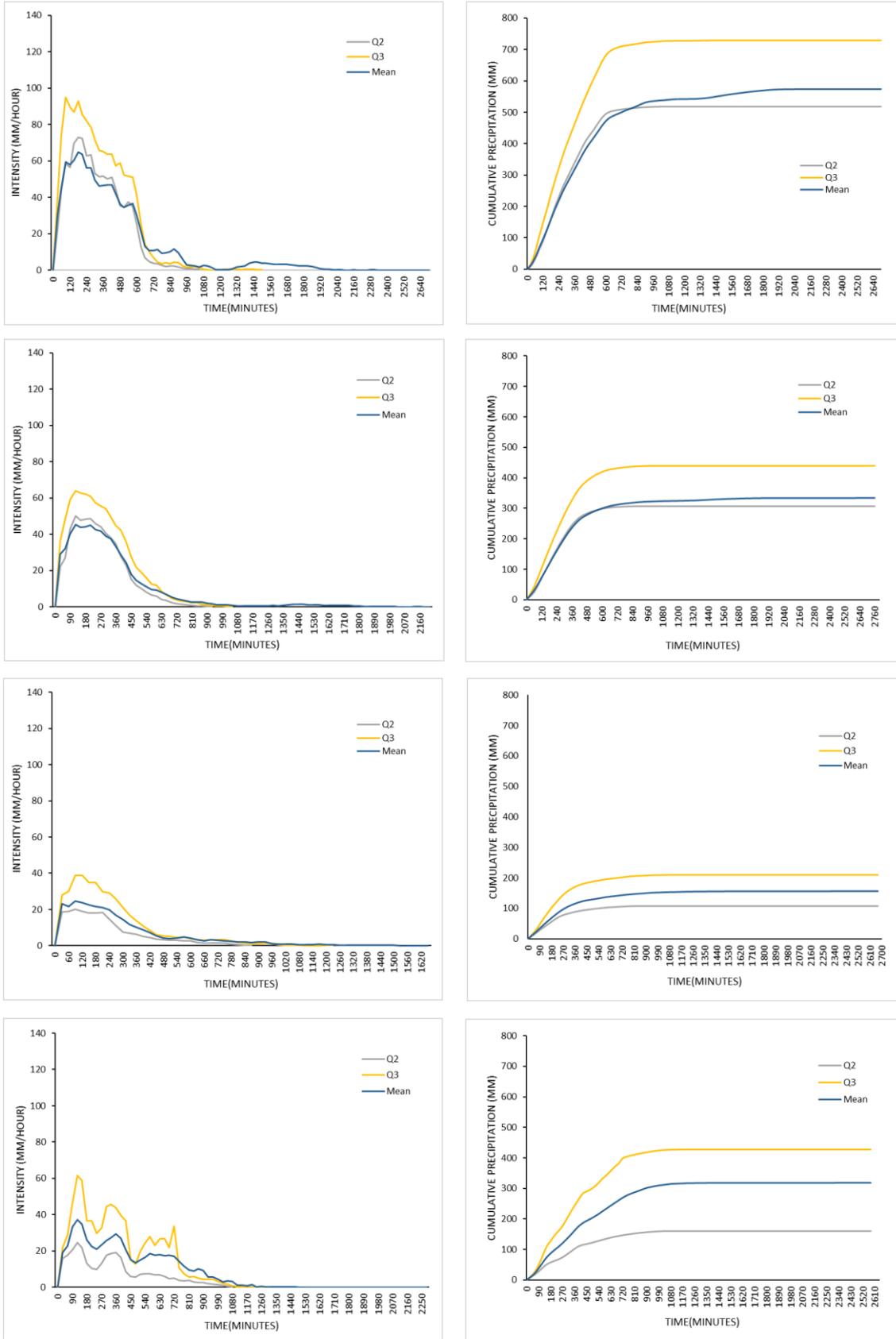


Figure 4-12: Rainfall time series (left) and cumulative plots (right) of the selected cluster representative signals when using  $K=5$ . T1, T3, T4, and T5 plots are in rows 1, 2, 3, and 4, respectively.

#### 4.5.2. Using Four Temporal Clusters (K=4)

When using  $K=4$ , representative signals were derived for T1, T2, and T3 by applying the procedure explained in subsection 4.5.1. The representative signals resulting from using  $K=4$  were used to optimize the building of the TC associate rainfall dataset when the flood characteristics of signals resulting from  $K=5$  were found to be similar. The statistics of the selected representative signals for  $K=4$  are summarized in Table 4-8, and their plots are in Appendix 5. The rainfall signals for all the clusters are similar in shape, showing sudden heavy rain pours for an extended time, followed by an abrupt drop in the rainfall intensity. The multiple peak behaviour observed in T5 resulting from  $K=5$  is dissolved. Except for T2, all signals at Q2 and Q3 have a rainfall duration lasting less than 24hours. The period of the Mean signal is still affected by pixels with higher antecedent rainfall before the storm starts.

Table 4-8: Summary statistics for signals Q2, Q3, and Mean for the temporal clusters resulting from  $K=4$

	T1			T2			T4		
	<i>Tr</i> (mm)	<i>I<sub>max</sub></i> (mm/hr)	<i>Dr</i> (hr)	<i>Tr</i> (mm)	<i>I<sub>max</sub></i> (mm/hr)	<i>Dr</i> (hr)	<i>Tr</i> (mm)	<i>I<sub>max</sub></i> (mm/hr)	<i>Dr</i> (hr)
<b>Q2</b>	116.4	21.8	15.0	500.8	71.1	18.0	293.8	49.6	16.0
<b>Q3</b>	230.4	42.2	18.0	727.3	91.4	25.0	446.2	65.0	18.0
<b>Mean</b>	169.0	26.1	27.5	564.6	61.2	38.5	336.6	45.4	33.0

*Tr* (mm): Accumulation (mm)

*Dr*(hr): Period (Hours)

*I<sub>max</sub>*(mm/hr): Maximum Intensity (mm/hr)

#### 4.5.3. Using Three Temporal Clusters (K=3)

For three temporal clusters, the representative signals were derived for T2 and T3. The statistics of the rainfall signals are in Table 4-9, and their plots are in Appendix 5. These rainfall signals were used in the optimization approach when generating the TC associate dataset. The maximum intensity and rainfall accumulation differ significantly for Q2 and Q3 in both clusters. The rainfall duration is now shorter as some signals have less than 18hours; however, the duration at the mean is still longer. The rainfall is shaped with a single peak, as observed earlier in the representative signals resulting from  $K=4$ .

Table 4-9: Summary statistics for signals Q2, Q3, and the Mean for temporal the clusters resulting from  $K=3$

	T2			T3		
	<i>Tr</i> (mm)	<i>I<sub>max</sub></i> (mm/hr)	<i>Dr</i> (hr)	<i>Tr</i> (mm)	<i>I<sub>max</sub></i> (mm/hr)	<i>Dr</i> (hr)
<b>Q2</b>	435.7	61.7	17.5	213.5	38.4	15.5
<b>Q3</b>	676.6	84.9	24.5	381.3	59.0	17.5
<b>Mean</b>	520.3	55.0	45.0	269.2	38.6	37.0

*Tr* (mm): Accumulation (mm)

*I<sub>max</sub>* (mm/hr): Maximum Intensity (mm/hr)

*Dr*(hr): Period (Hours)

## 5. RESULTS: FLOOD MODELLING

The results presented in this chapter were attained by applying the methodology in chapter 3 to obtain the research’s third specific objective of evaluating the tropical cyclone associate rainfall dataset by simulating the flood response using openLISEM model. The flood characteristics were examined to improve the building of the TC associate rainfall dataset. Section 5.1 presents the detailed assessment of the flood response for the rainfall signals resulting from using  $K=5$ . In section 5.2, the flood characteristics resulting from the optimization of the rainfall dataset are presented. Finally, the decision on the final Tropical Cyclone Associate Rainfall Dataset is in section 5.3.

### 5.1. Using Five Temporal Clusters ( $K=5$ )

The results are for the quantified flood response for the representative signals of T1, T3, T4, and T5. Temporal cluster 2 (T2) was not included in the flood modelling as it was regarded as not flood intense. The tables in this section summarize the quantified flood response for each of the rainfall signals. In the tables, the coloured cells are used to highlight similarities in the flood response for different cluster representative signals. The scatter plots in this section are for the examined flood variables vs. the precipitation variables of cumulative rainfall and maximum intensity. The points represent individual rainfall signals of the temporal clusters, where T1, T3, T4, and T5 are blue, gray, yellow, and red, respectively. For reporting purposes, the naming format of the cluster representative signals will be, for instance, Q2-T1 is the same as cluster representative signal Q2 of temporal cluster T1.

#### 5.1.1. Flood Extent

Some representative signals result in similar maximum flood extent even though they belong to different clusters. For example, as observed in Table 5-1, the flood extent ( $\text{km}^2$ ) resulting from rainfall of Q2-T3 is comparable to the flood area caused by Q3-T5. Likewise, Q3-T3 and Mean-T1 caused almost similar flood extents. The statistics in the table show that the rainfall signal Q3 causes the largest flood area for all the temporal clusters compared to Q2 and Mean. On the other hand, T1 has the most extensive flood extent, and T4 has the lowest.

Table 5-1: Quantified maximum flood extent ( $\text{km}^2$ ) caused by cluster representative signals resulting from  $K=5$ . Cells coloured similarly are for the flood extents that differ by less than  $0.005\text{km}^2$ . The values in the table are rounded off to four decimal places to show the very slight differences.

	Q2	Q3	Mean
T1	3.9844	4.8416	3.7620
T3	3.1204	3.7624	2.9836
T4	1.0176	2.1904	1.2992
T5	1.3976	3.1156	2.3816

Scatter plots of the quantified flood extents vs. the cumulative rainfall and flood extents vs. maximum intensity reveal a steady linear relationship (Figure 5-1). However, rainfall intensity is more dominant in driving the maximum flood extent. This finding is observed in the slope of the trend lines where maximum intensity has a slightly higher slope value and R-squared ( $R^2$ ) value. Signals with slightly lower cumulative rainfall volumes but higher maximum intensity caused more extensive flood areas than those with larger precipitation volumes received at lower intensities. For example, Q3-T3 and Mean-T1 experience a

considerable difference in their rainfall accumulation (the black circle in Figure 5-1a); however, they caused cause nearly the same flood extent (differing by only 0.0004km<sup>2</sup>). It can be observed that these two signals have almost similar maximum rainfall intensities (the black circle in Figure 5-1b).

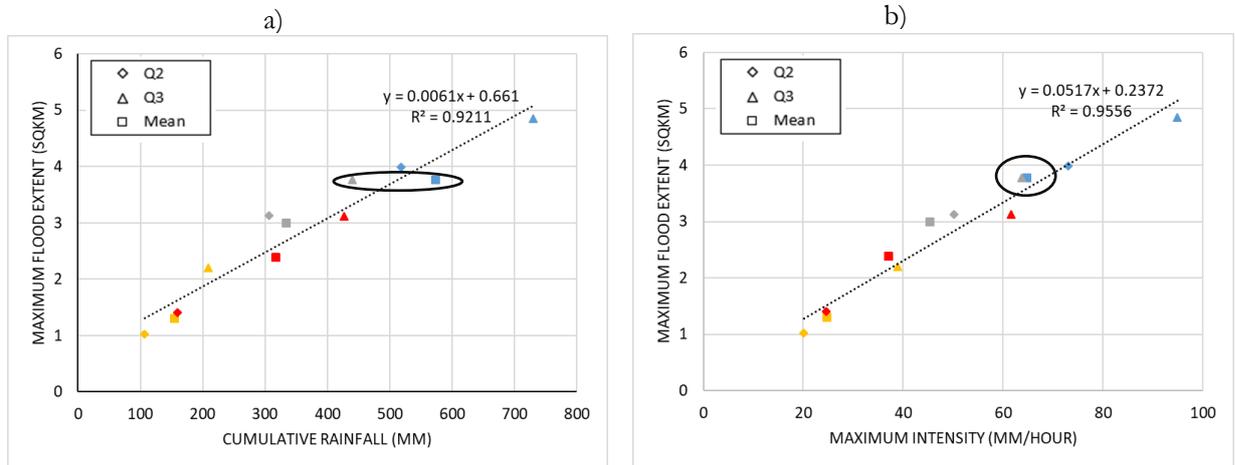


Figure 5-1: Scatter plots of the linear relationship between the maximum flood extent (Y-axis) and a) cumulative rainfall, b) maximum rainfall intensity.

### 5.1.2. Flood Depth Distribution

Table 5-2 summarizes the quantified maximum flood depths experienced in the catchment due to rainfall representative signals for the different temporal clusters. The table statistics show that some signals caused comparable maximum flood depths as highlighted by the colored cells; for example, Q3-T3 and Mean-T1 respond similarly. For all the clusters, maximum flood depths were highest due to rainfall of Q3, with T1 reaching water levels of 4.2meters at some point during the flood event. An area was considered not flooded if the water level was <=0.05meters, and these mostly corresponded to higher altitude locations in the catchment (Figure 5-2). A significant percentage of the flooded areas reached depths ranging from 0.05-0.5meters for all the cluster representative rainfall signals. Except for Q3-T1, flood depths above 2meters were fewer for all the rainfall signals.

Examination of Figure 5-3 shows a strong positive linear relationship observed between the flood depth and the rainfall characteristics. However, the correlation value R<sup>2</sup> for the maximum rainfall intensity is slightly higher (by about 0.027) than the value for the rainfall accumulation. The slope value of the trend line is higher for the maximum intensity showing that this rainfall variable has more influence on the maximum flood depth.

Table 5-2: Quantified maximum flood depth (meters) caused by cluster representative signals resulting from K=5.

Cells coloured similarly are for the flood depths that differ slightly. The values in the table are rounded off to four decimal places to show the very slight differences.

	Q2	Q3	Mean
T1	3.880	4.213	3.698
T3	2.788	3.697	2.612
T4	1.256	1.688	1.326
T5	1.319	2.846	1.846

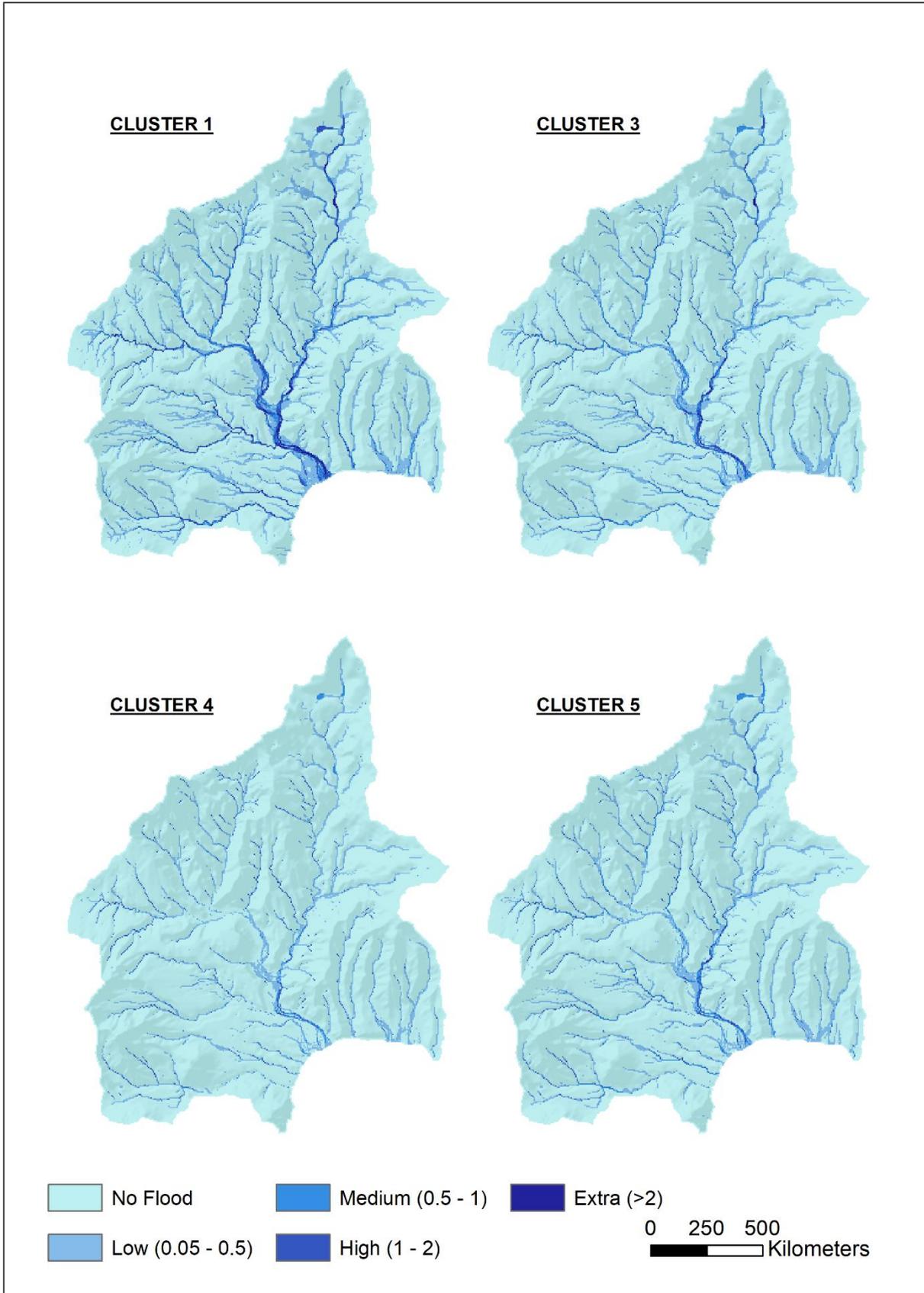


Figure 5-2: Geographical representation of the maximum flood depths reached in the different temporal clusters due to rainfall of Q3. All the maps use the same scale of flood depths labelled Low, Medium, High, and Extra.

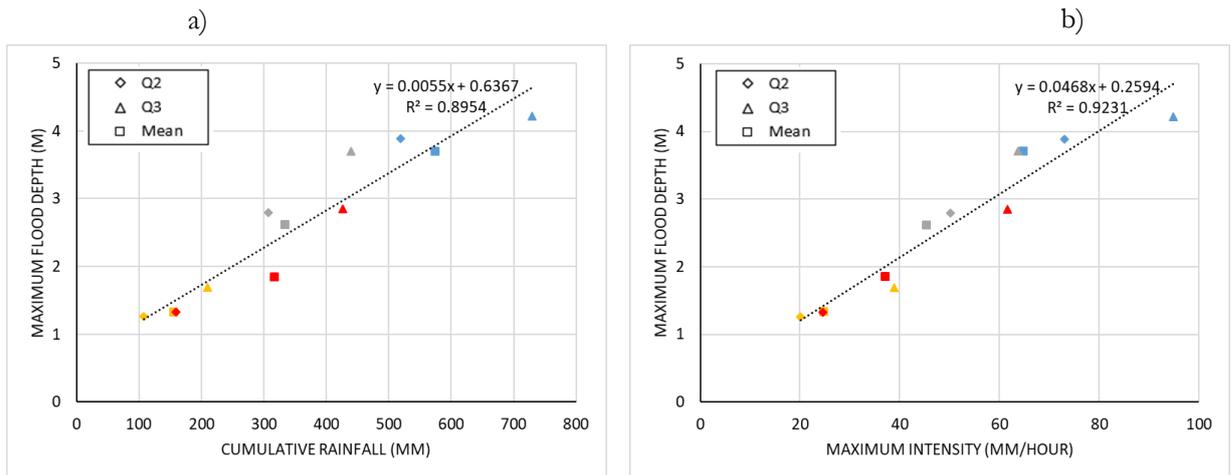


Figure 5-3: Scatter plots of the linear relationship between the maximum flood depth (Y-axis) and a) cumulative rainfall, b) maximum rainfall intensity.

### 5.1.3. Flood Volume

The statistics in Table 5-3 show that more rainfall signals from different clusters cause similar maximum flood volumes as highlighted by the coloured cells; for example, Q2-T5 and Mean-T4 have a similar response. T1 accumulated the most enormous flood volumes compared to other temporal clusters. The property of Q3 causing the largest flood response in all the clusters is again observed in the maximum accumulated flood volumes.

The scatter plots in Figure 5-4 show that while the maximum flood volume increased with the total rainfall, the maximum intensity was more influential in driving this flood variable due to the higher associated R<sup>2</sup> value of 0.97. Also, the slope of the trend line shows that the maximum flood volume increases by a slightly higher margin when the rainfall intensity increases than when the total rainfall increases. For example, Mean-T1 accumulates 134.3mm more than Q3-T3, but their maximum flood volume only differs by 5100m<sup>3</sup>, attributable to their very close maximum intensities.

Table 5-3: Quantified maximum flood volume (million m<sup>3</sup>) caused by cluster representative signals resulting from K=5. Cells coloured similarly are for the flood volumes that differ very slightly. The values in the table are rounded off to four decimal places to show the very slight differences.

	Q2	Q3	Mean
T1	1.6668	2.2789	1.5051
T3	1.0955	1.5000	1.0448
T4	0.4158	0.7662	0.5599
T5	0.5596	1.1218	0.8101

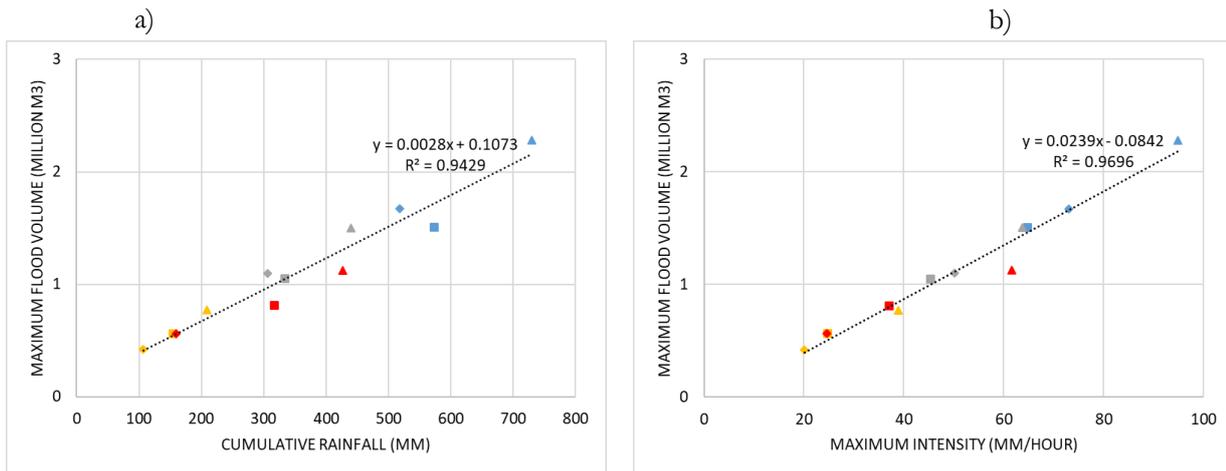


Figure 5-4: Scatter plots of the linear relationship between the maximum flood volume (Y-axis) and a) cumulative rainfall, b) maximum rainfall intensity.

**5.1.4. Infiltration**

As explained in section 3.3.4, the infiltration is influenced by the initialization of the soil moisture content. It should be noted that the initial soil moisture content was set at 85% of saturation. Statistics in Table 5-4 show that infiltration was generally low for all the cluster representative signals. Similarities in infiltration mainly were observed for signals within the same cluster, especially for Q3 and Mean, as highlighted in the table. The linear relationship between infiltration and the rainfall characteristics is generally weaker, as observed in the lower values for  $R^2$  (Figure 5-5). The cumulative rainfall has a slightly stronger influence on infiltration than the maximum rainfall intensity. Figure 5-6 is a map representation of the cumulative infiltration resulting from the representative signal Q3 for each temporal cluster.

Table 5-4: Summary of the total infiltration (mm) for the cluster representative signals resulting from  $K=5$ . Cells coloured similarly are for the values that differ by small margins.

	Q2	Q3	Mean
T1	92.1	97.4	99.5
T3	75.3	86.6	86.9
T4	60.8	73.0	73.4
T5	77.8	98.5	98.0

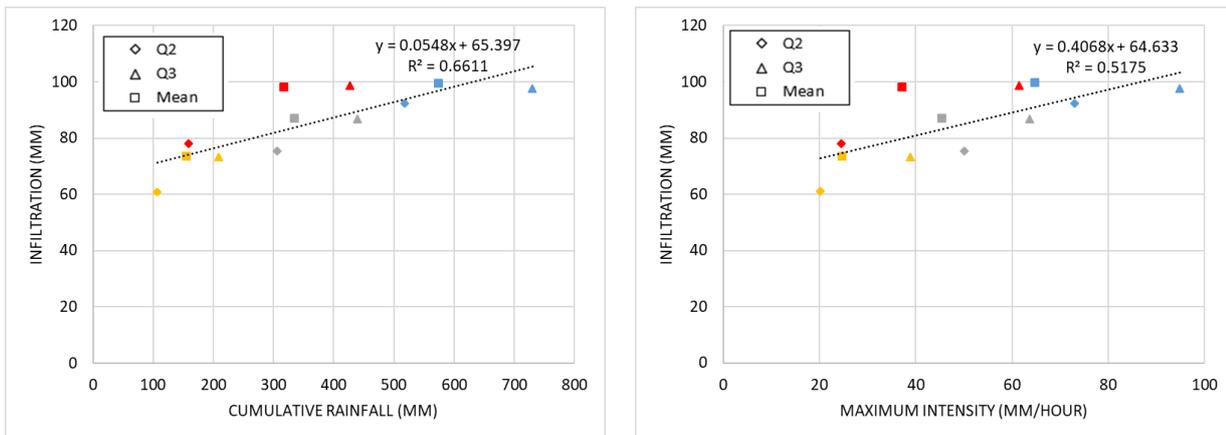


Figure 5-5: Linear relationship between the infiltration and a) cumulative rainfall, b) maximum rainfall intensity

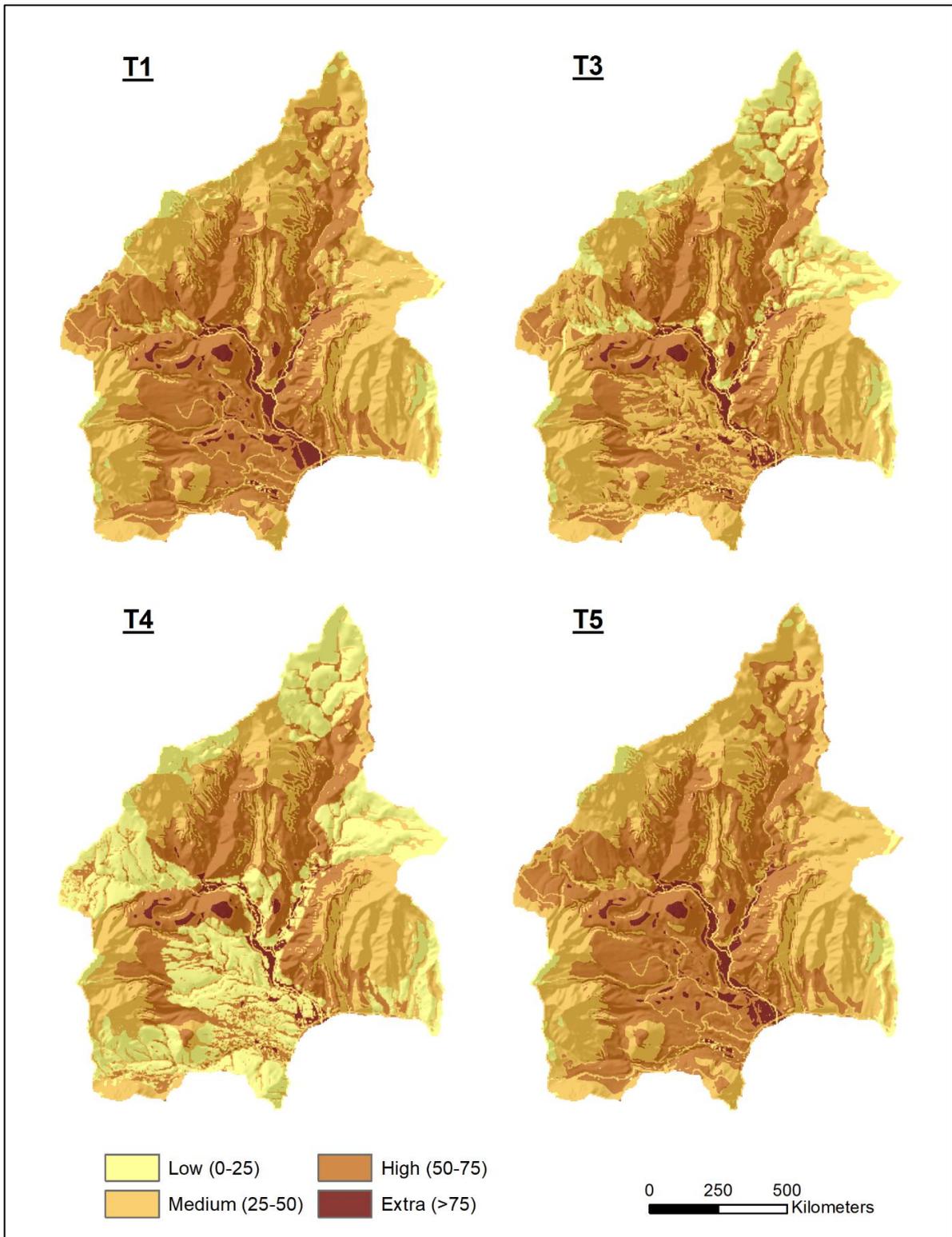


Figure 5-6: Cumulative infiltration (mm) reached due to signal Q3 rainfall.

**5.1.5. Runoff Ratio**

The statistics of the cluster runoff percentages are represented in Table 5-5 below. It is observed that some signals within the same cluster have almost similar runoff percentages, for example, Q2 and Mean of T1. Likewise, Q2 and Mean of T3 respond similarly. The runoff ratio for Q3 is more dominant for all the clusters, with T1 having the highest value of 0.84. The least runoff proportions are observed in T4, including areas that accumulate the least rainfall and experience the lowest maximum rainfall intensities. However, the runoff percentage for Mean-T4 differs minimally from that for Q2-T5.

From the scatter plots in Figure 5-7, it is observed that the maximum intensity is more dominant in influencing the runoff percentage than the cumulative rainfall since it has a higher R<sup>2</sup> value of 0.838. For example, the T1 signals have the highest maximum rainfall intensities and the highest runoff ratio. Thus, it shows, more water flows into the sea, meaning not all the incoming precipitation is converted to flood.

Table 5-5: Quantified runoff ratio caused by cluster representative signals resulting from K=5. Cells coloured similarly are for the values that differ by small margins. Again, the figures in the table are rounded off to four decimal places to show the very slight differences.

	Q2	Q3	Mean
<b>T1</b>	0.7958	0.8479	0.8034
<b>T3</b>	0.7113	0.7724	0.7003
<b>T4</b>	0.3336	0.5880	0.4435
<b>T5</b>	0.4296	0.7373	0.6492

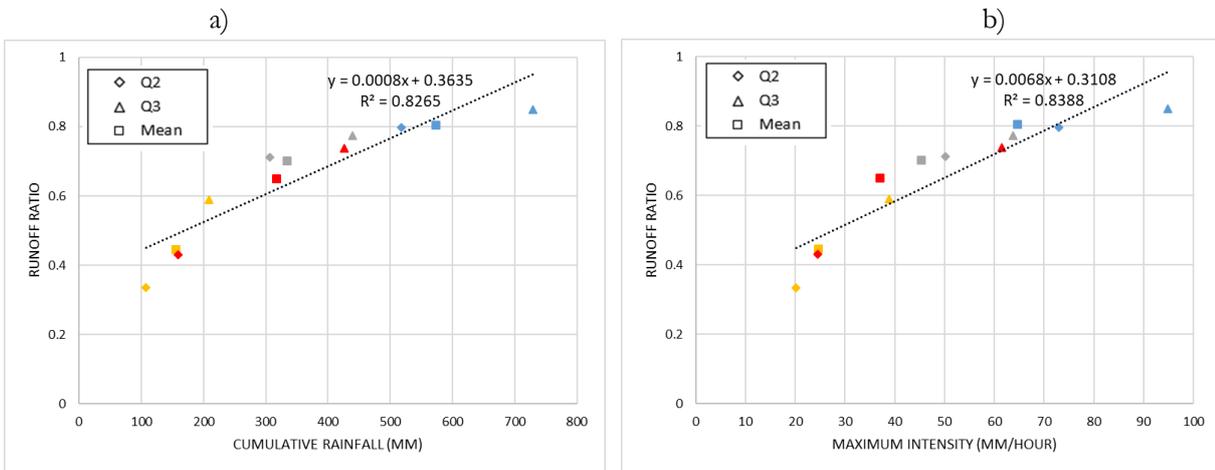


Figure 5-7: Scatter plots of the linear relationship between the runoff ratio (Y-axis) and a) cumulative rainfall, b) maximum rainfall intensity.

**5.1.6. Response Time**

As explained in section 3.3.6, response time was expressed as the difference between the timing of the peak discharge and the peak intensity. Table 5-6 below summarized the statistics for the cluster lag-to-peak time for all the representative signals. T1 has the shortest average lag time of about 2.5hours, and the most prolonged delay is experienced in T5. Flood response times for signals Q3-T1 and Q2-T3 are similar. Likewise, the response time for Q3-T3 and Mean-T4 differs by a small margin. Also, Q3-T4 and Mean-T3 respond similarly. The hydrographs for each representative signal of the four clusters are in Appendix 6.

Table 5-6: Summary of the flood response time (hours) caused by cluster representative signals resulting from  $K=5$ . Cells coloured similarly are for the values that differ by small margins (less than 0.5hours)

	Q2	Q3	Mean
T1	2.0	3.2	2.1
T3	3.2	3.0	3.5
T4	2.7	3.5	2.9
T5	5.0	4.4	5.2

### 5.1.7. Flood Duration

Similarities in the average flood duration for representative signals from the different clusters are observed as highlighted by the coloured cells in Table 5-7. In all the clusters, the average flood duration was most extended for the signals at Mean lasting over 25hours. The rainfall of Q2-T3 caused the shortest flood duration.

Table 5-7: Summary of the average flood duration (hours) caused by cluster representative signals resulting from  $K=5$ . Cells coloured similarly are for the values that differ by small margins (less than 0.5hours)

	Q2	Q3	Mean
T1	18.4	27.2	34.9
T3	15.7	18.9	30.3
T4	16.7	22.2	27.7
T5	23.6	18.6	27.4

### 5.1.8. Mini Discussion for the Observed Flood Response for Signals Resulting from $K=5$

The investigated flood characteristics show a similar flood response for the cluster representative signals. In some cases, signals of the same cluster have a similar flood response. In addition, the value  $K=5$  was an arbitrary choice because the elbow graph (see Figure 4-4b) was smooth and did not have a defined inflection point. However, the elbow method guided selecting the starting optimal number of clusters ( $K$  value). These findings imply a redundancy in the number of optimal clusters that were used to classify the rainfall time series. As explained in section 3.4, the optimization approach (see Figure 3-3) was applied to improve the building of the TC associate rainfall dataset. Since it is required that the final rainfall should have a distinct flood response, only the representative signal Q3 was be used to run the flood model during the optimization. Q3 was selected because the flood response for this signal was highest for most investigated flood characteristics, as observed in the previous subsections.

## 5.2. Application of the Optimization Approach

The iterative process of the optimization approach illustrated in Figure 3-3 was performed twice, i.e., when  $K=4$  and  $K=3$ . As explained in subsection 5.1.8, only cluster representative signal Q3 was used to run the flood model. The statistics of Q3 for  $K=4$  and  $K=3$  are in Table 4-8 and Table 4-9, respectively. The corresponding time series and cumulative plots of the Q3 signal are in Appendix 5. The results of the flood response are presented in the following subsections.

### 5.2.1. Using $K=4$

Compared to signals resulting from  $K=5$ , the quantified flood response for representative signals of  $K=4$  is significantly distinct, as observed in the calculated differences in Table 5-8. It is observed that the resulting

flood characteristics show three levels of the magnitude for TS Erika's rainfall which were labelled as Extreme (T2), Intermediate (T4), and Least Intense (T1).

Table 5-8: Comparison of flood characteristics resulting from rainfall signal Q3 for K=4

Cluster Rainfall Signal (Q3)	Flood Extent (km <sup>2</sup> )	Flood Depth (m)	Flood Volume (million m <sup>3</sup> )	Runoff Ratio	Infiltration (mm)	Flood Duration (hr)	Flood Response (hr)
T2	4.80	4.22	2.22	0.85	98.13	23.73	1.47
T4	3.70	3.65	1.42	0.77	88.95	19.16	3.13
<b>Diff</b>	<b>1.09</b>	<b>0.57</b>	<b>0.79</b>	<b>0.08</b>	<b>9.18</b>	<b>4.57</b>	<b>-1.67</b>
T2	4.80	4.22	2.22	0.85	98.13	23.73	1.47
T1	2.47	1.84	0.84	0.62	74.01	29.33	2.77
<b>Diff</b>	<b>2.32</b>	<b>2.39</b>	<b>1.38</b>	<b>0.23</b>	<b>24.06</b>	<b>-5.61</b>	<b>-1.30</b>
T4	3.70	3.65	1.42	0.77	88.95	19.16	3.13
T1	2.47	1.84	0.84	0.62	74.08	29.33	2.77
<b>Diff</b>	<b>1.23</b>	<b>1.82</b>	<b>0.59</b>	<b>0.15</b>	<b>14.87</b>	<b>-10.17</b>	<b>0.37</b>

### 5.2.2. Using K=3

This result divides TS Erika's rainfall into two levels of magnitude with a very distinct flood response, as shown in calculated differences in Table 5-9. Flood response of T2 from K=4 is comparable to T2 from K=3. Likewise, the flood response of T4 from K=4 is comparable to that of T3 from K=4. Based on this observation, the two levels of magnitude due to signals from K=3 can be labelled as Extreme (T2) and Intermediate (T3). However, quantified flood characteristics at K=4 are slightly higher than those for K=3 in both cases.

Table 5-9: Comparison of flood characteristics resulting from rainfall signal Q3 for K=3

Cluster Rainfall Signal (Q3)	Flood Extent (km <sup>2</sup> )	Flood Depth (m)	Flood Volume (million m <sup>3</sup> )	Runoff Ratio	Infiltration (mm)	Flood Duration (hr)	Flood Response (hr)
T2	4.62	4.18	2.05	0.84	97.69	22.76	1.75
T3	2.31	1.76	0.79	0.61	84.40	29.78	3.38
<b>Diff</b>	<b>2.31</b>	<b>2.43</b>	<b>1.26</b>	<b>0.23</b>	<b>13.29</b>	<b>-7.02</b>	<b>-1.63</b>

### 5.3. The Tropical Cyclone Associate Rainfall DataSet

The final decision on the Tropical cyclone associated rainfall dataset was based on the observed flood response for both cases of reducing the K value, i.e., K=4 and K=3. In both scenarios, the flood response is distinct, and the storm was divided into different levels of magnitude. However, the signals for K=4 have a higher impact than those for K=3. Additionally, K=3 divides the storm into only two levels of magnitude (Extreme and Intermediate), leaving out the third level (Least extreme) observed when using K=4. Thus, implying that K=3 is not a suitable choice for the optimal number of clusters as it starts to group pixels of different magnitude into the same cluster. Based on these observations, the cluster representative signals resulting from K=4 were selected as the final Tropical Cyclone Associate Rainfall Dataset for TS Erika (Figure 5-8).

A characteristic of rapidly increasing intensities with a maximum peak reached in less than 2 hours after the start of the storm is observed for the rainfall dataset. Additionally, about 50% of the accompanying rainfall volumes were accumulated less than 6 hours into the storm (Figure 5-8). When such high-intensity precipitation with heavy accumulation pours on steep terrain in a very short time; consequently, the flooding hazard is triggered.

Since the initial soil moisture content was at 85%, there was little room for infiltration, and hence the high values for the runoff percentage resulting in more discharge in the sea. The map in Figure 5-9 for the flood depth distribution shows that low and medium flood depths were mainly experienced along the main channel. Flood heights for classes High and Extra were primarily experienced in the river channel. There was no flood in the high elevation areas.

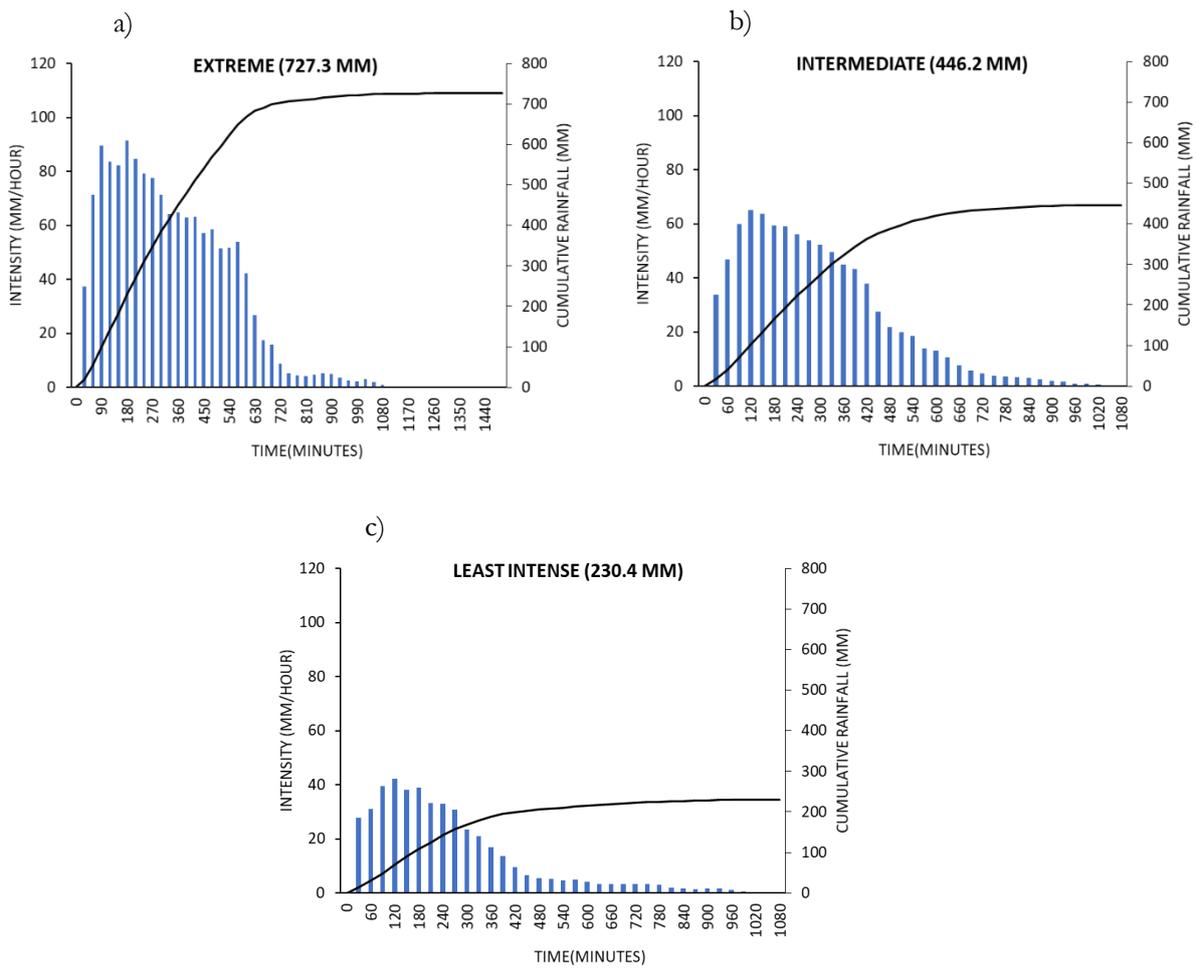


Figure 5-8: Precipitation time series and cumulative precipitation plots of the Tropical cyclone associate rainfall dataset for TS Erika resulting from  $K=4$ .

The blue bar graphs represent the half-hourly rainfall intensities, and the black line is for the cumulative precipitation. a) Extreme, b) Intermediate, and c) Least Intense

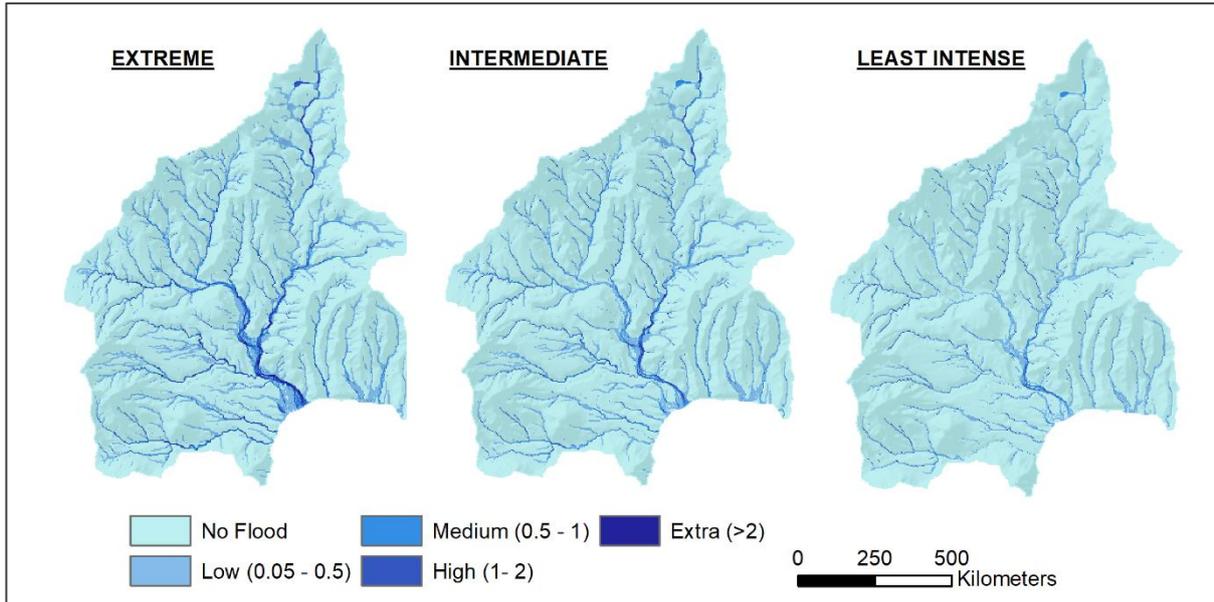


Figure 5-9: Flood depth (meters) distribution for all the three rainfall signals of the Tropical cyclone associate rainfall dataset for TS Erika resulting from  $K=4$ .

## 6. DISCUSSION AND CONCLUSION

### **Flood Hazard and the need for Rainfall Characterisation**

An important aspect of flood hazard assessment is knowledge of the local precipitation patterns because it is based on a combination of weather events with a given frequency (return periods) and associated magnitudes of peak discharges. However, most countries lack information for the frequency analysis of the flood hazard (Lumbroso et al., 2011); therefore, assumptions are made during the flood hazard assessment. A common assumption is to substitute the probability of the flood (the peak discharges) with the probability of the weather event causing the flood. The frequency and magnitude of the meteorological forcing (e.g., extreme precipitation events) has to be analyzed with relatively long rainfall time series (i.e., 30-50years); however, these records are non-existent for long-term detailed rainfall analysis. Another assumption is made to replace the required detailed weather information with the annual maximum 24hour rainfall to measure the total rainfall event. Long-term daily data series are present for most weather and rainfall stations. However, flood hazard modelling requires detailed sub-daily rainfall data to simulate correctly the catchment hydrology, flow dynamics, mass movement, and even debris flows. Therefore, there is a need for rainfall data detailed both in space and time, with a known frequency and magnitude.

Flood modelers translate precipitation records into design storms with a statistically derived shape and total storm depth that coincides with given return periods. However, design storms are originally not meant to do flood hazard modelling but to calculate peak discharge for channel and bridge design. Design storms have very simple shapes derived from IDF curves of long-term ground measurements, which may not provide a realistic estimate of rainfall accompanying TCs, because of extreme wind speeds and equipment breaking down during the event (Pollock et al., 2018). Additionally, observations of TC rainfall are sometimes regarded as outliers and excluded when fitting a distribution curve.

In this research, a new approach was developed to generate a TC associate rainfall dataset to provide a realistic estimate of the TC-related flooding hazard. The research utilized satellite precipitation estimates from GPM-IMERG to acquire the rainfall information needed. Unlike ground measurements, satellite estimates provide a continuous acquisition of rainfall data for the period of a given storm on a broad spatial coverage (Yoshimoto & Amarnath, 2017). Considering other satellite precipitation products, the TRMM has the most prolonged temporal coverage; however, this dataset does not have the resolution in space and time (Huffman & Bolvin, 2018) required for making good IDF curves. The GPM-IMERG data has a better spatial and temporal resolution; however, its temporal coverage is not long enough to derive statistics for IDF curves (Yong & Wang, 2020).

Knowing that precipitation measurement from just one GPM-IMERG pixel would not give enough information to generate IDF curves representative of the precipitation accompanying studied TC, the research used a more expansive study area to gather enough information. Previous research (Agustín Breña-Naranjo et al., 2015; Hernández Ayala & Matyas, 2016) utilized a buffer of a 500km radius around the TC low-pressure center to study accompanying rainfall. However, the interest of this research was in the TC rainfall activity in the vicinity of the island. How was the TC rain falling as the eye passed close to the island? Using a 500km diameter buffer around the island was suitable to define the required extent of the study area, assuming that the TC activity for far way locations would not affect the island.

**Specific objective 1: To perform rainfall spatial-temporal pattern analysis for the most suitable tropical cyclones.**

The devastation from tropical cyclone activity in the last 5-10 years suggests possible effects of climate change; therefore, the research was performed for a more recent period (2015-2019) for the possibility to include these effects and build resilience toward future episodes of TC precipitation. The study period saw landfalling TCs, higher category TCs and extreme precipitation accompanying the TCs (Blake & Zelinsky, 2017; Cangialosi et al., 2017). During the annual Atlantic hurricane seasons for the study period, the TC occurrences ravaged the Caribbean islands and the eastern coast of the United States, among other areas. For example, Dominica suffered twice during 2015-2019, first due to rain from Tropical storm Erika in 2015 and two years later due to landfalling Hurricane Maria in 2017.

The research investigated rainfall accompanying TS Erika of 2015. While the storm's eye passed at a distance from Dominica, TS Erika's convective region passed over the island. As a result, TS Erika brought the island state enormous amounts of rainfall on 27<sup>th</sup> August 2015 when the storm's eye touched the northern tip of Guadeloupe about 100km north of Dominica (Nugent & Rios-Berrios, 2018). When the GPM-IMERG precipitation estimates over TS Erika were compared to the radar precipitation estimates calibrated using ground gauge observations in Ogden (2016), it was found that GPM-IMERG data gave an acceptable estimate of the precipitation measurements. For example, the measure of the maximum rainfall intensity experienced in Dominica was well estimated. Additionally, the rainfall totals that were calculated from GPM-IMERG over the island fall in the ranges provided in Pasch & Penny (2015). Since precipitation information in this research was not averaged over just one pixel, but for a wide study area, it is concluded that GPM-IMERG provided reliable precipitation estimates needed for this research. Some previous studies commend the satellite dataset for its performance in estimating extreme rainfall; however, they mention likely uncertainties, and thus the product's algorithms need further improvement (Cui et al., 2020; Omranian et al., 2018).

The TC is a dynamic system with a unique rainfall structure; therefore, it was essential to know the location of the convective region and how precipitation was distributed over the study area. TS Erika precipitation was mainly south of the track; now, the rainfall magnitude was not reducing with increasing distance from the TC center but rather from the location of the heaviest rainfall. For instance, some sites of lower rainfall accumulation and maximum intensities, especially in the north of the study area, were closer to the storm track than some pixels in the east that experienced the highest rainfall (Figure 4-2). There were two spatially distinct sites (both over the ocean) that received enormous precipitation. A likely explanation is that the storm was in motion, for as long as it reached an area with favourable conditions (Evans, 2017), it was capable of intensifying and causing heavy rainfall to areas within its path.

The observed differences in the characteristics of the rainfall time series for individual pixels were attributed to the dynamics in the motion of the TC eye and the TC intensity. The multiple peak characteristics were likely influenced by the moment of the increased TC intensity just before the eye hit Guadeloupe. Depending on a pixel's location from the position of the TC eye, it was possible to experience single-peaked or multiple peaked precipitations. The findings of the rainfall characteristics were essential considerations for categorizing the pixels in the clustering analysis. In addition, since rainfall extremes could be experienced anywhere in the study area, knowledge of the timing and location of the varying rainfall severity is crucial to better prepare for likely future occurrences.

**Specific objective 2: To perform spatial-temporal clustering of the TC rainfall time series to categorize the rainfall characteristics.**

The choice of the clustering algorithm was based on literature; also, due to time constraints, the research was not focused on comparing the performance of the different clustering algorithms. Therefore, the decision to use K-means may have been biased; however, previous studies, as detailed in section 2.5, commend the algorithm's high performance for rainfall time-series clustering. Additionally, the high values obtained for the *BSS/TSS* ratio used to assess the clustering quality gave the confidence that the algorithm detected the distinct rainfall behaviour and output quality clusters. While the main interest was in the temporal distribution of the rainfall, the research investigated the spatial aspect to illustrate the differences in the rainfall timing while the storm was in motion. Indeed, the time at which different locations received rainfall varied as the storm progressed westward. The rainfall timing followed the eye movement, and some pixels received rainfall earlier because they were closer to the eye as it moved westward.

The temporal clustering revealed the temporal distribution of TS Erika's rainfall, i.e., reducing magnitudes as the distance from the location of the heaviest rain increased. The two storm extremes were delineated, and while they were spatially distinct, the clustering recognized the similarities in their temporal distribution. The extreme cluster has rainfall magnitudes that have not been measured before in Dominica (Jetten, 2016); even so, this finding is justifiable because none of these pixels hit the island. The temporal clustering showed that the heaviest rain fell over the ocean.

Since the precipitation of the cluster representative signals would later be used to run the flood model, the rainfall information before the starting threshold gave an insight into how to initialize the model (the initial soil moisture content, *ksat*). The antecedent rainfall (whether TC related or not) makes the catchment wetter, decreases the storage capacity, eventually inducing runoff water. Thus, thresholds to determine the start of the TC rainfall proved useful to removing the antecedent rainfall that is not TC-related. However, caution was taken to ensure that the storm's essential starting timesteps were not discarded to avoid introducing missing information in the data. For this reason, different starting points were tested before deciding to use 10mm/hr. Despite this, however, some storm information was likely excluded and considered non-TC rainfall. An option not to regulate the storm's ending intensity and period was taken to compensate for the likely excluded information of when the storm started; only that, this choice could have influenced the duration. Additionally, some pixels received antecedent rainfall higher than the selected threshold (see Figure 4-10a); these pixels eventually affected the period of the final rainfall dataset. It is possible to identify these pixels by conducting further detailed analysis; however, it was not done in this study. Implying that practitioners interested in the rainfall dataset's duration dimension need to be careful unless this limitation is corrected.

The optimal number of clusters was critical to the research approach because the rainfall characteristics were classified based on the selected *K* value. Pixels of the temporal clusters resulting from the five optimal clusters determined with the elbow method were redistributed when the value of *K* was reduced. Implying that there was a redundancy when using *K*=5; nevertheless, it was an arbitrary value read based on the visualization of the smooth elbow graph (see Figure 4-4b). The reduction of the *K* value shows that the temporal distribution of TC rainfall characteristics does not have defined boundaries, as some pixels moved from one cluster to another. This is not surprising because the TC system's dynamics cause an uneven distribution of the maximum rainfall intensities and total volumes.

The decision on the cluster representative signal was not based on the rainfall realization of just one pixel in a cluster but on quantiles. Considering that GPM-IMERG is not validated on a pixel-by-pixel basis but rather using spatial averages (Maranan et al., 2020), the research derived the cluster representative signal using timestep quantiles computed from the rainfall time series of all pixels in a given cluster. The rainfall signals for Q0, Q1, Q4, and Q5 were not selected because these were likely few pixels probably located on the cluster boundary (remember, the cluster boundaries were unstable and observed to change at a reduced  $K$  value). Instead, the timestep quantiles at Q2 (0.5), Q3 (0.75), and Mean were selected because these signals comprised more pixels and were more realistic regarding the cluster temporal statistics.

Representative signals for the temporal clusters (T1, T3, T4) resulting from  $K=5$  have shapes comparable to those of design storms (single peaked curve with a big volume, see Appendix 7). However, one cluster deviated from this shape; T5 has multiple peaks suggesting that with the temporal clustering, it was possible to discriminate pixels with single-peaked rainfall and numerous peaks. Thus, the findings of the temporal clustering show that the different temporal behaviours of the TC rainfall indeed exist and should be considered when building a rainfall dataset specific to tropical cyclones. However, reducing the  $K$  value to optimize the derivation of cluster representative signals dissolved the multiple peaked behaviours of the rainfall. Therefore, introducing a limitation for which a definite conclusion is yet to be drawn about the optimal number of clusters that reveal the TC rainfall distinct temporal behaviours.

**Specific objective 3: To evaluate the tropical cyclone associate rainfall dataset by simulating the flood response using a flood characteristics prediction model.**

The flood modelling was used to evaluate the cluster representative signals and see if they had distinct flood characteristics for purposes of flood hazard assessment. The openLISEM model was already set up; only the precipitation input was changing (Serere, 2020). The rainfall was assumed to be homogeneous across the 34km<sup>2</sup> catchment, meaning that the effect of spatial variability was not investigated. The catchment was smaller than a grid cell of the GPM-IMERG dataset ( $0.1^\circ \times 0.1^\circ$  spatial resolution). The catchment's landscape with steep slopes and narrow valleys determined a lot of the flood response because only the coastal plain is flatter, and that is where water can more spread out. More rainfall may not mean more extensive flooding for such a catchment because more water is also discharged into the ocean.

The catchment terrain may have influenced the similarity of flood response for some of the representative signals resulting from  $K=5$ . For example, while rainfall signals from T5 had a different shape, their output flood characteristics were comparable to those of some signals, mainly in T3 and T4. It is not surprising because, for this catchment, that rainfall shape is less important, as a strong link exists between rainfall amount, maximum rainfall intensity, and the magnitude of the flood hazard. However, the flood response might be different for catchments where there is more space for the water to spread.

The similar reaction of the model to clusters labelled as different when using  $K=5$  suggested a redundancy in the cluster signals. It was speculated that the number of optimal clusters was likely related to the catchment geomorphology, and differences in flood response might appear for reduced optimal clusters. Since the rainfall signal at 75% quantile gave the most significant flood magnitude for all the clusters resulting from  $K=5$ , a decision was made to only use this signal to run the flood model during the optimization approach (see Figure 3-3). Cristiano et al. 's (2018) found that the threshold at the upper quantile (75%) was more representative of a cluster's hydrological response; they used this threshold to represent cluster flood generating precipitation.

The optimization procedure was such that if the flood response was distinct for each cluster representative signal (in this case, the 75% quantile), no more iterations were conducted. Different flood responses were indeed observed for both cases of reducing the  $K$  value; however, the quantified flood statistics were higher for  $K=4$ , and therefore selected as the Final Tropical Cyclone Associate Rainfall Dataset.

## Conclusions

Instead of deriving a design storm from long-term temporal data on one location, the method developed here has the advantage that it is based on all available rainfall data for all pixels in the study area. Therefore, the method does not depend on short and potentially bad datasets from ground measurement or IDF curves that are borrowed from other regions.

The final TC associate rainfall dataset can be assigned a known probability to be used for flood hazard assessment in Dominica. However, the known probability is a choice of stakeholders/end users as some of them base their disaster prevention and mitigation measures on station-based 24h rainfall return periods. For instance, the Caribbean Catastrophe Risk Insurance Facility (<https://www.ccrif.org/>) determines insurance based on 5,10,15, and 20-year return periods and uses ground-based rainfall. Therefore, the probability assigned to the TC associate rainfall dataset must be linked to existing values to ensure that methodology does not deviate from what the stakeholders experience.

The developed approach has been tested on only one TC (TS Erika); however, other TCs may give different values. Nevertheless, GPM-IMERG is a rich dataset of 20 years of spatial data (Tan & Huffman, 2019) over dozens of tropical cyclones, implying a large scope for further investigation. The research approach is not limited to Dominica; the island was used as a proof of concept; therefore, the method can be tested and applied in other TC-prone zones.

The TC associate rainfall dataset also revealed that TC rainfall has different levels of magnitude and as a result, the flooding suffered in a location depends on the part of the storm that hits the area. The TC is a dynamic system; therefore, planners and responders need to prepare for all levels of flood likelihood. Dominica's actors and stakeholders in disaster risk reduction can use the rainfall dataset to make informed decisions on investments to mitigate damages from flooding. Furthermore, investigating TC rainfall patterns provides valuable insight for impact-based forecasting to prompt actions to minimize the likely damages before the disaster hits. The rainfall dataset is useful for early warning because the storm brings heavy rainfall to different locations at varying times; therefore, citizens can be alerted beforehand on what time to expect the torrential rains.

## 7. RECOMMENDATIONS

Some of the research's key findings were:

- Existence of different temporal behaviours (with varying magnitudes) in the rainfall if the same tropical cyclone system.
- The practicality of time series clustering to classify the different temporal behaviours of tropical cyclone rainfall.
- Effectiveness of thresholds to align rainfall time series of spatially distinct pixels with similar temporal behaviour.

However, there were some limitations:

- A definite conclusion cannot be made based on the analysis of only one tropical cyclone. It is not representative of the hundreds of the TCs that have devastated the Caribbean islands.
- Some choices, such as the optimization process, were based on expert judgment; further development is needed towards objective approaches.
- The time dimension of the final rainfall dataset was influenced by pixels that received high-intensity antecedent rainfall and the use of thresholds to determine the beginning of the TC rainfall.
- Uncertainty in the suitable number of optimal clusters to categorize the temporal characteristics of the TC rainfall.

Therefore, there is still room for the potential development of this work. Below are some of the recommendations for future researchers:

- The current research analyzed what rainfall was happening in the island's vicinity. However, not all the output signals of the TC associate dataset likely have a probability experienced on the island; there is still a gap in deriving the island's statistical values. Furthermore, since each tropical cyclone is different, different magnitudes are likely to be experienced on the island as the storm is in motion.
- A TC behaves differently over larger landmasses than over small islands, which likely affects the clustering and resulting design event shapes and magnitude. Therefore, further investigation of more TCs as they make landfall is advisable. Also, different types of catchments may influence the flooding result.
- The clustering analysis is a critical step in the procedure of categorizing the TC rainfall. How would the  $K$  value change when exploring multiple TCs? Also, is it likely that another clustering algorithm would output a better clustering of the rainfall data than K-means?
- The research aimed to generate a distinct rainfall for a tropical cyclone; however, this dataset should be linked to a return period for purposes of flood hazard analysis. Further research should focus on assigning this probability; should it be for the individual signals from the dataset or the whole storm? Key stakeholders/users need to be consulted when assigning the probability.
- It is unknown if the output rainfall dataset is applicable for the whole island or limited to the catchment. Further research can use the same method but explore different catchments in Dominica.
- TS Erika was generally a straight mover, and the location of the convective region did not change; it remained south of the storm for the whole period of the storm. However, looping tropical cyclones such as Hurricane Ivan (Stewart, 2011) may cause the convective region to shift with the

storm dynamics. Future research should therefore consider how rainfall is distributed for TCs that change course.

## LIST OF REFERENCES

- Abdulkareem, J. H., Pradhan, B., Sulaiman, W. N. A., & Jamil, N. R. (2019). Development of Lag Time and Time of Concentration for a Tropical Complex Catchment under the Influence of Long-term Land use/Land cover (LULC) Changes. *Arabian Journal of Geosciences*, 12(3), 101. <https://doi.org/10.1007/s12517-019-4253-z>
- Acevedo, S., & Alleyne, T. (2016). Gone with the Wind: Estimating Hurricane and Climate Change Costs in the Caribbean. *IMF Working Papers*, 16(199), 1. <https://doi.org/10.5089/9781475544763.001>
- Agustín Breña-Naranjo, J., Pedrozo-Acuña, A., Pozos-Estrada, O., Jiménez-López, S. A., & López-López, M. R. (2015). The Contribution of Tropical Cyclones to Rainfall in Mexico. *Physics and Chemistry of the Earth*, 83–84, 111–122. <https://doi.org/10.1016/j.pce.2015.05.011>
- Alam, M. S., & Paul, S. (2020). A Comparative Analysis of Clustering Algorithms to Identify the Homogeneous Rainfall Gauge Stations of Bangladesh. *Journal of Applied Statistics*, 47(8), 1460–1481. <https://doi.org/10.1080/02664763.2019.1675606>
- Avila, L. A., & Fritz, C. L. (2018). *Hurricane Beryl (AL022018)*.
- Avila, L. A., Stewart, S. R., Berg, R., & Hagen, A. B. (2019). *Hurricane Dorain (AL052019)*.
- Ayala, J. J. H. (2016). *Climatology of Tropical Cyclone Rainfall Over Puerto Rico: Processes, Patterns and Impacts*.
- Bacmeister, J. T., Reed, K. A., Hannay, C., Lawrence, P., Bates, S., Truesdale, J. E., Rosenbloom, N., & Levy, M. (2018). Projected Changes in Tropical Cyclone Activity under Future Warming Scenarios using a High-Resolution Climate Model. *Climatic Change*, 146(3–4), 547–560. <https://doi.org/10.1007/s10584-016-1750-x>
- Balbastre-Soldevila, R., García-Bartual, R., & Andrés-Doménech, I. (2019). A Comparison of Design Storms for Urban Drainage System Applications. *Water (Switzerland)*, 11(4). <https://doi.org/10.3390/w11040757>
- Barclay, J., Wilkinson, E., White, C. S., Shelton, C., Forster, J., Few, R., Lorenzoni, I., Woolhouse, G., Jowitt, C., Stone, H., & Honychurch, L. (2019). Historical Trajectories of Disaster Risk in Dominica. *International Journal of Disaster Risk Science*, 10(2), 149–165. <https://doi.org/10.1007/s13753-019-0215-z>
- Berk, M., Špačková, O., & Straub, D. (2017). Probabilistic Design Storm Method for Improved Flood Estimation in Ungauged Catchments. *Water Resources Research*, 53(12), 10701–10722. <https://doi.org/10.1002/2017WR020947>
- Blake, E. S., & Zelinsky, D. A. (2017). *Hurricane Harvey (AL092017)*.
- Bout, B., & Jetten, V. G. (2018). The Validity of Flow Approximations when Simulating Catchment-Integrated Flash Floods. *Journal of Hydrology*, 556, 674–688. <https://doi.org/10.1016/j.jhydrol.2017.11.033>
- Bout, B., & Jetten, V. G. (2020). Catchment-scale multi-process modelling with local time stepping. *Environmental Earth Sciences*, 79, 184. <https://doi.org/10.1007/s12665-020-08914-7>
- Bout, B., Jetten, V. G., De Roo, A., Wesseling, C., & Ritsema, C. (2018). *openLISEM- Multi-Hazard Land Surface Process Model*. 255. <https://blog.utwente.nl/lisem/>
- Cangialosi, J. P., Latto, A. S., & Berg, R. (2017). *Hurricane Irma (AL112017)*.
- Chang, L. T.-C., Cheung, K. K. W., & McAneney, J. (2013). Case Study of TRMM Satellite Rainfall Estimation for Landfalling Tropical Cyclones: Issues and Challenges. *Tropical Cyclone Research and Review*, 2(2), 109–123. <https://doi.org/10.6057/2013TCRR02.04>
- CHARIM. (2018). *CHARIM - GeoNode*. CHARIM. <http://charim-geonode.net/>
- Cheung, K., Yu, Z., Elsberry, R. L., Bell, M., Jiang, H., Lee, T. C., Lu, K.-C., Oikawa, Y., Qi, L., Rogers, R. F., & Tsuboki, K. (2018). Recent Advances in Research and Forecasting of Tropical Cyclone Rainfall. *Tropical Cyclone Research and Review*, 7(2), 106–127. <https://doi.org/10.6057/2018TCRR02.03>
- Cristiano, E., Ten Veldhuis, M. C., Gaitan, S., Ochoa Rodriguez, S., & Van De Giesen, N. (2018). Critical Scales to Explain Urban Hydrological Response: An Application in Cranbrook, London. *Hydrology and Earth System Sciences*, 22(4), 2425–2447. <https://doi.org/10.5194/hess-22-2425-2018>
- Cui, W., Dong, X., Xi, B., Feng, Z. H. E., & Fan, J. (2020). Can the GPM IMERG Final Product Accurately Represent MCSs' Precipitation characteristics Over the Central and Eastern United States? *Journal of Hydrometeorology*, 21(1), 39–57. <https://doi.org/10.1175/JHM-D-19-0123.1>

- De Paola, F., Giugni, M., Topa, M. E., & Bucchignani, E. (2014). *Intensity-Duration-Frequency (IDF) rainfall curves, for data series and climate projection in African cities*. <https://doi.org/10.1186/2193-1801-3-133>
- Elhamid, A. M. I., Eltahan, A. M. H., Mohamed, L. M. E., & Hamouda, I. A. (2020). Assessment of the Two Satellite-Based Precipitation Products TRMM and RFE Rainfall Records using Ground Based Measurements. *Alexandria Engineering Journal*, 59(2), 1049–1058. <https://doi.org/10.1016/j.aej.2020.03.035>
- Emanuel, K. A. (2013). Downscaling CMIP5 Climate Models shows Increased Tropical Cyclone Activity over the 21st Century. *Proceedings of the National Academy of Sciences of the United States of America*, 110(30), 12219–12224. <https://doi.org/10.1073/pnas.1301293110>
- Esma, E. Ö. (2020). Clustering of Time-Series Data. In *Data Mining - Methods, Applications and Systems*. IntechOpen. <https://doi.org/10.5772/intechopen.84490>
- Evans, J. L. (2017). Chapter Four 4 . Tropical Cyclone Intensity , Structure , and Structure Change. *Global Guide to Tropical Cyclone Forecasting*, 126–155. [https://www.meted.ucar.edu/•Met-EdFAQ:https://www.meted.ucar.edu/resources\\_faq.php](https://www.meted.ucar.edu/•Met-EdFAQ:https://www.meted.ucar.edu/resources_faq.php)
- Faust, E., & Bove, M. (2017). *The Hurricane Season 2017: A Cluster of Extreme Storms*. Munich Re. <https://www.munichre.com/topics-online/en/climate-change-and-natural-disasters/natural-disasters/storms/hurricane-season-2017.html>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The Climate Hazards Infrared Precipitation with Stations - A New Environmental Record for Monitoring Extremes. *Scientific Data*, 2(1), 1–21. <https://doi.org/10.1038/sdata.2015.66>
- Gallo, F., Daron, J., Macadam, I., Cinco, T., Villafuerte, M., Buonomo, E., Tucker, S., Hein-Griggs, D., & Jones, R. G. (2019). High-Resolution Regional Climate Model Projections of Future Tropical Cyclone Activity in the Philippines. *International Journal of Climatology*, 39(3), 1181–1194. <https://doi.org/10.1002/joc.5870>
- Gericke, O. J., & Smithers, J. C. (1935). Review of Methods Used to Estimate Catchment Response Time for the Purpose of Peak Discharge Estimation. *Sciences Journal*, 59(11). <https://doi.org/10.1080/02626667.2013.866712>
- Gilewski, P., & Nawalany, M. (2018). Inter-Comparison of Rain-Gauge, Radar, and Satellite (IMERG GPM) Precipitation Estimates Performance for Rainfall-Runoff Modelling in a Mountainous Catchment in Poland. *Water*, 10(11), 1665. <https://doi.org/10.3390/w10111665>
- Girons, L. M., Wennerström, H., Nordén, L. Å., & Seibert, J. (2015). Location and Density Of Rain Gauges for the Estimation of Spatial Varying Precipitation. *Geografiska Annaler, Series A: Physical Geography*, 97(1), 167–179. <https://doi.org/10.1111/geoa.12094>
- Goel, M. K. (2011). Runoff Coefficient. *Encyclopedia of Earth Sciences Series, Part 3*, 952–953. [https://doi.org/10.1007/978-90-481-2642-2\\_456](https://doi.org/10.1007/978-90-481-2642-2_456)
- Gutro, R. (2018). *IMERG Measures Tropical Cyclone Ava's Disastrous Rainfall*. <https://phys.org/news/2018-01-imerg-tropical-cyclone-ava-disastrous.html>
- Hadi, A. F., Yudistira, I., Anggraeni, D., & Hasan, M. (2018). The Geographical Clustering of the Rainfall Stations on Seasonal GSTAR Modelling for Rainfall Forecasting. *Journal of Physics: Conference Series*, 1028(1), 12238. <https://doi.org/10.1088/1742-6596/1028/1/012238>
- Hallegatte, S., Vogt-Schilb, A., Bangalore, M., & Rozenberg, J. (2017). *Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters*. Washington, DC: World Bank. <https://doi.org/10.1596/978-1-4648-1003-9>
- Hernández Ayala, J. J., & Matyas, C. J. (2016). Tropical cyclone rainfall over Puerto Rico and its relations to environmental and storm-specific factors. *International Journal of Climatology*, 36(5), 2223–2237. <https://doi.org/10.1002/joc.4490>
- Hijmans, R. J. (2020). *Raster Data — R Spatial*. Innovation Lab for Collaborative Research on Sustainable Intensification. <https://rspatial.org/raster/spatial/8-rastermanip.html>
- Huffman, G. J., & Bolvin, D. T. (2018). *TRMM and Other Data Precipitation Data Set Documentation George* (Issue April).
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Stocker, E. F., & Tan, J. (2019). V06 IMERG Release Notes. *NASA*, 1–6. <https://pmm.nasa.gov/data-access/downloads/gpm>
- Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2019). *GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06*. Goddard Earth Sciences Data and Information Services

- Center (GES DISC). [https://disc.gsfc.nasa.gov/datasets/GPM\\_3IMERGHH\\_06/summary](https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGHH_06/summary)
- IPCC. (2014). *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. <https://doi.org/10.1017/CBO9781107415379>
- Jacqueline, J. (2020). *The Impacts of Climate Change on the Atlantic Hurricane Season*. CDAP DISASTER BULLETIN. <https://www.carilec.org/the-impacts-of-climate-change-on-the-atlantic-hurricane-season/>
- Jetten, V. (2016). *CHARIM Project: Dominica National Flood Hazard Map Methodology and Validation Report*.
- Jiang, Q., Li, W., Wen, J., Fan, Z., Chen, Y., Scaioni, M., & Wang, J. (2019). Evaluation of Satellite-Based Products for Extreme Rainfall Estimations in the Eastern Coastal Areas of China. *Journal of Integrative Environmental Sciences*, 16(1), 191–207. <https://doi.org/10.1080/1943815X.2019.1707233>
- Karssenber, D., Schmitz, O., Salamon, P., de Jong, K., & Bierkens, M. F. P. (2010). A Software Framework for Construction of Process-Based Stochastic Spatio-Temporal Models and Data Assimilation. *Environmental Modelling and Software*, 25(4), 489–502. <https://doi.org/10.1016/j.envsoft.2009.10.004>
- Ketchen, D. J., & Shook, C. L. (1996). The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique. *Strategic Management Journal*, 17(6), 441–458. [https://doi.org/10.1002/\(sici\)1097-0266\(199606\)17:6<441::aid-smj819>3.0.co;2-g](https://doi.org/10.1002/(sici)1097-0266(199606)17:6<441::aid-smj819>3.0.co;2-g)
- Knutson, T., Camargo, S. J., Chan, J. C. L., Emanuel, K., Ho, C. H., Kossin, J., Mohapatra, M., Satoh, M., Sugj, M., Walsh, K., & Wu, L. (2020). Tropical cyclones and climate change assessment part II: Projected response to anthropogenic warming. *Bulletin of the American Meteorological Society*, 101(3), 303–322. <https://doi.org/10.1175/BAMS-D-18-0194.1>
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on Determining Number of Cluster in K-Means Clustering. *International Journal of Advance Research in Computer Science and Management Studies*, 1(6), 2321–7782.
- Kostaschuk, R., Terry, J., & Raj, R. (2001). Tropical Cyclones and Floods in Fiji. *Hydrological Sciences Journal*, 46(3), 435–450. <https://doi.org/10.1080/02626660109492837>
- Kristen, E. (2019). *How Do Hurricanes Form?* NASA Science. <https://spaceplace.nasa.gov/hurricanes/en/>
- Krvavica, N., & Rubinić, J. (2020). Evaluation of Design Storms and Critical Rainfall Durations for Flood Prediction in Partially Urbanized Catchments. *Water*, 12(7), 2044. <https://doi.org/10.3390/w12072044>
- Lanfredi, M., Coluzzi, R., Imbrenda, V., Macchiato, M., & Simoniello, T. (2020). Analyzing Space-Time Coherence in Precipitation Seasonality Across Different European Climates. *Remote Sensing*, 12(1). <https://doi.org/10.3390/rs12010171>
- Le, M. H., Lakshmi, V., Bolten, J., & Bui, D. Du. (2020). Adequacy of Satellite-derived Precipitation Estimate for Hydrological Modelling in Vietnam Basins. *Journal of Hydrology*, 586, 124820. <https://doi.org/10.1016/j.jhydrol.2020.124820>
- Lenzen, M., Malik, A., Kenway, S., Daniels, P., Leung Lam, K., & Geschke, A. (2019). Economic Damage and Spill Overs from a Tropical Cyclone. *Natural Hazards and Earth System Sciences*, 19(1), 137–151. <https://doi.org/10.5194/nhess-19-137-2019>
- Loew, P. (2019). *The Natural Disasters of 2018 in Figures*. Munich Re. <https://www.munichre.com/topics-online/en/climate-change-and-natural-disasters/natural-disasters/the-natural-disasters-of-2018-in-figures.html>
- Loew, P. (2020). *Tropical Cyclones Cause Highest Losses*. Munich Re. <https://www.munichre.com/topics-online/en/climate-change-and-natural-disasters/natural-disasters/natural-disasters-of-2019-in-figures-tropical-cyclones-cause-highest-losses.html>
- Lumbroso, D. M., Boyce, S., Bast, H., & Walmsley, N. (2011). The Challenges of Developing Rainfall Intensity-Duration-Frequency Curves and National Flood Hazard Maps for the Caribbean. *Journal of Flood Risk Management*, 4(1), 42–52. <https://doi.org/10.1111/j.1753-318X.2010.01088.x>
- Machiwal, D., Dayal, D., & Kumar, S. (2017). Long-Term Rainfall Trends and Change Points in Hot and Cold Arid Regions of India. *Hydrological Sciences Journal*, 62(7), 1050–1066. <https://doi.org/10.1080/02626667.2017.1303705>
- Maranan, M., Fink, A. H., Knippertz, P., Amekudzi, L. K., Atiah, W. A., & Stengel, M. (2020). A Process-Based Validation of GPM-IMERG and its Sources Using a Mesoscale Rain Gauge Network in the West African Forest Zone. *Journal of Hydrometeorology*, 21(4), 729–749. <https://doi.org/10.1175/JHM-D-19-0257.1>

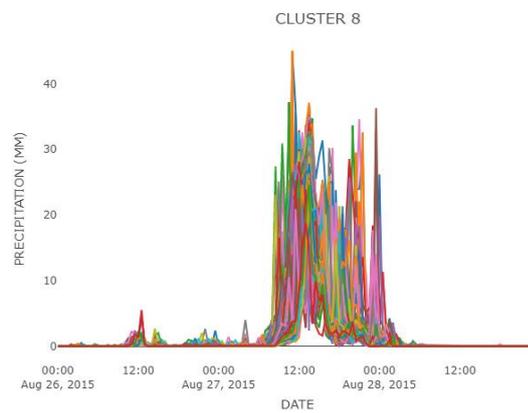
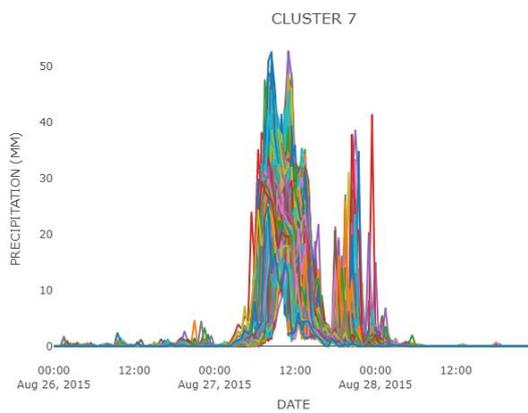
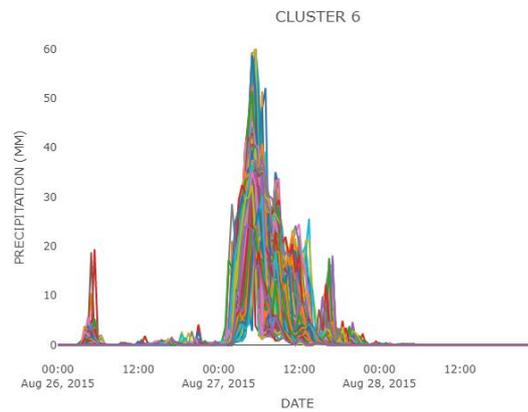
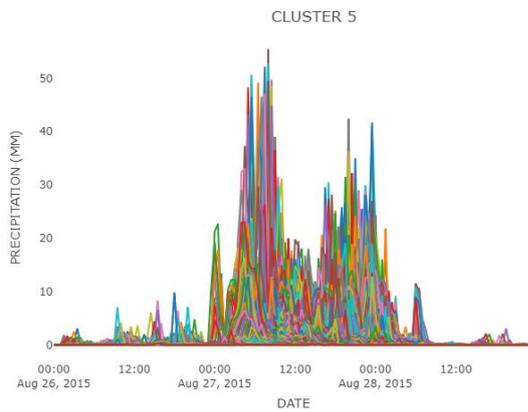
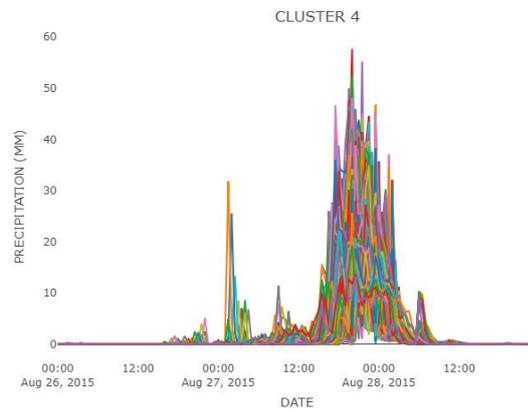
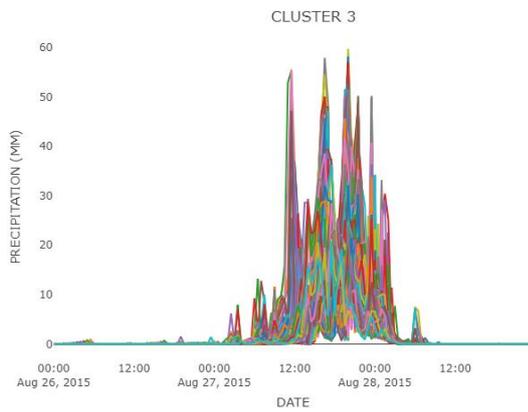
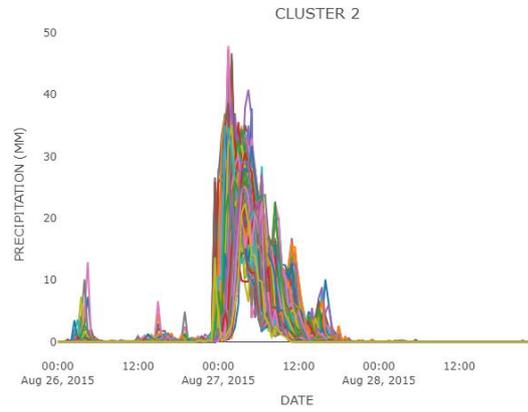
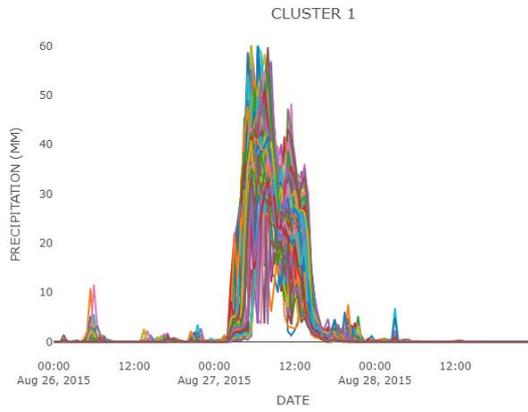
- Marchi, L., Borga, M., Preciso, E., & Gaume, E. (2010). Characterisation of Selected Extreme Flash Floods in Europe and Implications for Flood Risk Management. *Journal of Hydrology*, 394(1–2), 118–133. <https://doi.org/10.1016/j.jhydrol.2010.07.017>
- Martin, J. E., & Sinclair, R. R. (2007). A Typology of the Part-time Workforce: Differences on Job Attitudes and Turnover. *Journal of Occupational and Organizational Psychology*, 80(2), 301–319. <https://doi.org/10.1348/096317906X113833>
- Mastrandrea, M. D., Field, C. B., Stocker, T. F., Edenhofer, O., Ebi, K. L., Frame, D. J., Held, H., Kriegler, E., Mach, K. J., Matschoss, P. R., Plattner, G.-K., Yohe, G. W., & Zwiers, F. W. (2010). *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. <http://www.ipcc.ch>
- Mekonnen, K., Melesse, A. M., & Woldesenbet, T. A. (2021). Spatial evaluation of satellite-retrieved extreme rainfall rates in the Upper Awash River Basin, Ethiopia. *Atmospheric Research*, 249, 105297. <https://doi.org/10.1016/j.atmosres.2020.105297>
- Mori, N. (2014). Projection of Future Tropical Cyclone Characteristics Based on Statistical Model. In *Cyclones: Formation, Triggers and Control* (pp. 249–270). [https://www.researchgate.net/publication/264555463\\_Projection\\_of\\_Future\\_Tropical\\_Cyclone\\_Characteristics\\_based\\_on\\_Statistical\\_Model](https://www.researchgate.net/publication/264555463_Projection_of_Future_Tropical_Cyclone_Characteristics_based_on_Statistical_Model)
- Murakami, H., Wang, Y., Yoshimura, H., Mizuta, R., Sugj, M., Shindo, E., Adachi, Y., Yukimoto, S., Hosaka, M., Kusunoki, S., Ose, T., & Kitoh, A. (2012). Future Changes in Tropical Cyclone Activity Projected by the New High-Resolution MRI-AGCM. *Journal of Climate*, 25(9), 3237–3260. <https://doi.org/10.1175/JCLI-D-11-00415.1>
- Naranjo-Fernández, N., Guardiola-Albert, C., Aguilera, H., Serrano-Hidalgo, C., & Montero-González, E. (2020). Clustering Groundwater Level Time Series of the Exploited Almonte-Marismas Aquifer in Southwest Spain. *Water (Switzerland)*, 12(4). <https://doi.org/10.3390/W12041063>
- NASA. (2019). *NASA Estimates Heavy Rainfall in Hurricane Dorian*. <https://phys.org/news/2019-08-nasa-heavy-rainfall-hurricane-dorian.html>
- NOAA. (2020). *What is the Difference Between a Hurricane and a Typhoon?* National Ocean Service Website. <https://oceanservice.noaa.gov/facts/cyclone.html>
- Nugent, A. D., & Rios-Berrios, R. (2018). Factors Leading to Extreme Precipitation on Dominica from Tropical Storm Erika (2015). *Monthly Weather Review*, 146(2), 525–541. <https://doi.org/10.1175/MWR-D-17-0242.1>
- Ogden, F. L. (2016). *Indirect Peak Discharge Measurements on the Island of Dominica Associated with Extreme Rainfall from Tropical Storm Erika, August 27, 2015, with Modelling and Analysis*.
- Omranian, E., Sharif, H., & Tavakoly, A. (2018). How Well Can Global Precipitation Measurement (GPM) Capture Hurricanes? Case Study: Hurricane Harvey. *Remote Sensing*, 10(7), 1150. <https://doi.org/10.3390/rs10071150>
- Otker, I., & Srinivasan, K. (2018). Building Resilience in the Caribbean to Climate Change and Natural Disasters. *Finance and Development*, 55(1). <https://www.imf.org/external/pubs/ft/fandd/2018/03/otker.htm>
- Pasch, R. J., & Penny, A. B. (2015). Tropical Storm Erika (AL052015). In *National Hurricane Center*. [https://www.nhc.noaa.gov/data/tcr/AL052015\\_Erika.pdf](https://www.nhc.noaa.gov/data/tcr/AL052015_Erika.pdf)
- Pasch, R. J., Penny, A. B., & Berg, R. (2017). Hurricane Maria (AL152017). In *National Hurricane Center*.
- Paul-Rolle, A. (2014). *Commonwealth of Dominica Disaster Risk Reduction Country Profile*. <http://www.eird.org/>
- Pollock, M. D., O'Donnell, G., Quinn, P., Dutton, M., Black, A., Wilkinson, M. E., Colli, M., Stagnaro, M., Lanza, L. G., Lewis, E., Kilsby, C. G., & O'Connell, P. E. (2018). Quantifying and Mitigating Wind-Induced Undercatch in Rainfall Measurements. *Water Resources Research*, 54(6), 3863–3875. <https://doi.org/10.1029/2017WR022421>
- Rahimi, Y. G., Abbasi, E., & Farajzadeh, M. (2015). Analysis of the Effect of Tropical Cyclone Phet on the Occurrence of Heavy Rainfall and Floods in Chabahar, Iran. *Weather*, 70(12), 348–352. <https://doi.org/10.1002/wea.2582>
- Rakhecha, P. R., & Singh, V. P. (2009). *Applied Hydrometeorology*. Springer Netherlands. <https://doi.org/10.1007/978-1-4020-9844-4>
- Ratzlaff, J. R. (1994). Mean Annual Precipitation, Runoff, and Runoff Ratio for Kansas, 1971-1990. *Transactions of the Kansas Academy of Science (1903-)*, 97(3/4), 94. <https://doi.org/10.2307/3627776>
- Rouse, W. C., Reading, A. J., & Walsh, R. P. D. (1986). Volcanic Soil Properties in Dominica, West Indies.

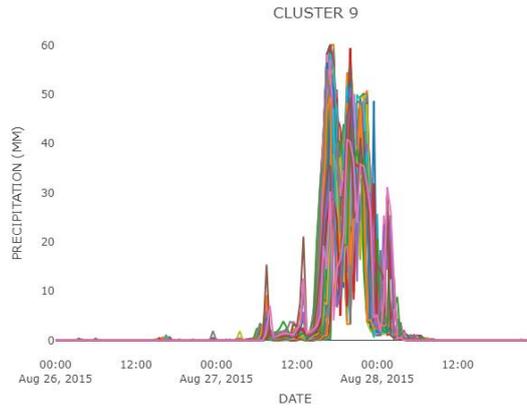
- Engineering Geology*, 23(1), 1–28. [https://doi.org/10.1016/0013-7952\(86\)90014-1](https://doi.org/10.1016/0013-7952(86)90014-1)
- Roux, F. (2019). *Tropical Cyclones: Impacts and Risks*. Encyclopedia of the Environment. <https://www.encyclopedie-environnement.org/en/air-en/tropical-cyclones-impacts-and-risks/>
- Saxton, K. E., & Rawls, W. J. (2006). Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions. *Soil Science Society of America Journal*, 70(5), 1569–1578. <https://doi.org/10.2136/sssaj2005.0117>
- Serere, H. (2020). Developing a Worst-case Tropical Cyclone Rainfall scenario for flood on Dominica. In *Masters Thesis*. <https://essay.utwente.nl/84941/>
- Smith, A. B. (2020). *2010-2019: A Landmark Decade of U.S. Billion-Dollar Weather and Climate Disasters*. <https://www.climate.gov/news-features/blogs/beyond-data/2010-2019-landmark-decade-us-billion-dollar-weather-and-climate>
- Soetewey, A. (2020). *The Complete Guide to Clustering Analysis. K-means and Hierarchical Clustering by Hand and in R*. Stats and R. <https://statsandr.com/blog/clustering-analysis-k-means-and-hierarchical-clustering-by-hand-and-in-r/>
- Stephenson, T. S., & Jones, J. J. (2017). Impacts of Climate Change on Extreme Events in the Coastal and Marine Environments of Caribbean Small Island Developing States (SIDS). In *Science Review*.
- Stewart, S. R. (2011). *Tropical Cyclone Report: Hurricane Ivan 2-24 September 2004* (Vol. 3). [http://www.nhc.noaa.gov/data/tcr/AL092004\\_Ivan.pdf](http://www.nhc.noaa.gov/data/tcr/AL092004_Ivan.pdf)
- Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, 336(1). <https://doi.org/10.1088/1757-899X/336/1/012017>
- Tan, J., & Huffman, G. J. (2019). Computing Morphing Vectors for Version 06 IMERG. *NASA*, 1–9.
- Tang, S., Li, R., He, J., Wang, H., Fan, X., & Yao, S. (2020). Comparative Evaluation of the GPM IMERG Early, Late, and Final Hourly Precipitation Products Using the CMPA Data over Sichuan Basin of China. *Water*, 12(2), 554. <https://doi.org/10.3390/w12020554>
- Thomas, A., Pringle, P., Pfliegerer, P., & Schleussner, C.-F. (2017). *Briefing Note: Tropical Cyclones: Impacts, the link to Climate Change and Adaptation*. [https://climateanalytics.org/media/tropical\\_cyclones\\_impacts\\_cc\\_adaptation.pdf](https://climateanalytics.org/media/tropical_cyclones_impacts_cc_adaptation.pdf)
- Timothy, S., Landsea, C., Hafele, G., Lorens, J., Taylor, A., Thurm, H., Ward, B., Willis, M., & Zaleski, W. (2019). *The Saffir-Simpson Hurricane Wind Scale*.
- Wang, Y. (2012). Recent Research Progress on Tropical Cyclone Structure and Intensity. *Tropical Cyclone Research and Review*, 1(2), 254–275. <https://doi.org/10.6057/2012TCRR02.05>
- Warren, L. T. (2005). Clustering of Time Series Data - A Survey. *Pattern Recognition*, 38(11), 1857–1874. <https://doi.org/10.1016/j.patcog.2005.01.025>
- Westen, van C. J. . (2016). *Caribbean Handbook on Risk Information Management National Scale Landslide Susceptibility Assessment for Dominica* (Issue May). <https://doi.org/10.13140/RG.2.1.4313.2400>
- Westen, van C. J., Alkema, D., Kerle, N., & Kingma, N. C. (2011). *Multi-Hazard Risk Assessment*.
- Winter, B., Schneeberger, K., Dung, N. V., Huttenlau, M., Achleitner, S., Stötter, J., Merz, B., & Vorogushyn, S. (2019). A Continuous Modelling Approach for Design Flood Estimation on Sub-Daily Time Scale. *Hydrological Sciences Journal*, 64(5), 539–554. <https://doi.org/10.1080/02626667.2019.1593419>
- WMO. (2013). Flood Forecasting and Early Warning. *Integrated FloodManagement Tools Series*, 19. [www.wmo.int](http://www.wmo.int)
- WMO. (2020). *Tropical Cyclones*. World Meteorological Organization. <https://public.wmo.int/en/our-mandate/focus-areas/natural-hazards-and-disaster-risk-reduction/tropical-cyclones>
- Wu, X., Cheng, C., Zurita-Milla, R., & Song, C. (2020). An Overview of Clustering Methods for Geo-Referenced Time Series: From One-Way Clustering to Co- and Tri-Clustering. In *International Journal of Geographical Information Science* (Vol. 34, Issue 9, pp. 1822–1848). Taylor and Francis Ltd. <https://doi.org/10.1080/13658816.2020.1726922>
- Yong, B., & Wang, H. (2020). Quasi-Global Evaluation of IMERG and GSMaP Precipitation Products over Land Using Gauge Observations. *Water*, 12(1), 243. <https://doi.org/10.3390/w12010243>
- Yoshimoto, S., & Amarnath, G. (2017). Applications of Satellite-Based Rainfall Estimates in Flood Inundation Modelling—A Case Study in Mundeni Aru River Basin, Sri Lanka. *Remote Sensing*, 9(10), 998. <https://doi.org/10.3390/rs9100998>
- Yu, Z., & Wang, Y. (2018). Rainfall Distribution in Landfalling Tropical Cyclones. In *Extreme Weather*. InTech. <https://doi.org/10.5772/intechopen.75910>

- Zelinsky, D. A. (2018). *Hurricane Isaac (AL092018)*.
- Zhong, R., Wang, Z., Lai, C., & Chen, J. (2017). Evaluation of the GPM IMERG Satellite-Based Precipitation Products and the Hydrological Utility. *Atmospheric Research*, 196, 151–163. <https://doi.org/10.1016/j.atmosres.2017.06.020>
- Zhou, M., Qu, S., Chen, X., Shi, P., Xu, S., Chen, H., Zhou, H., & Gou, J. (2019). Impact Assessments of Rainfall-Runoff Characteristics Response Based on Land Use Change via Hydrological Simulation. *Water (Switzerland)*, 11(4). <https://doi.org/10.3390/w11040866>

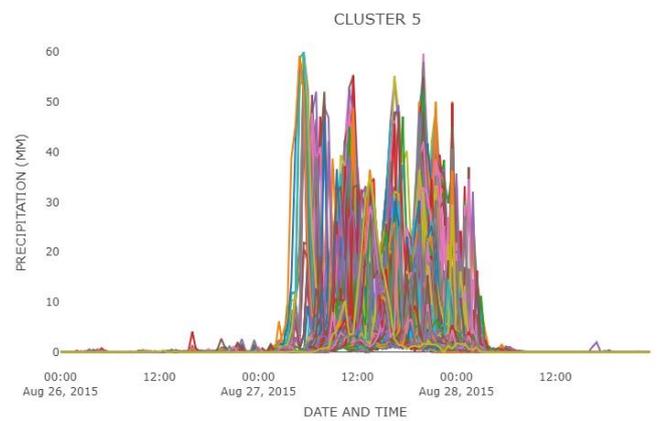
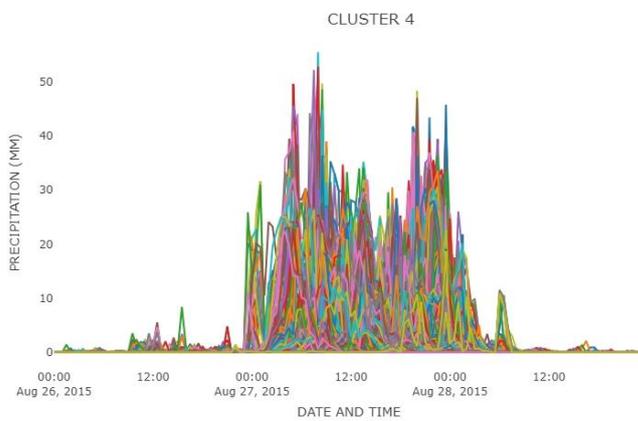
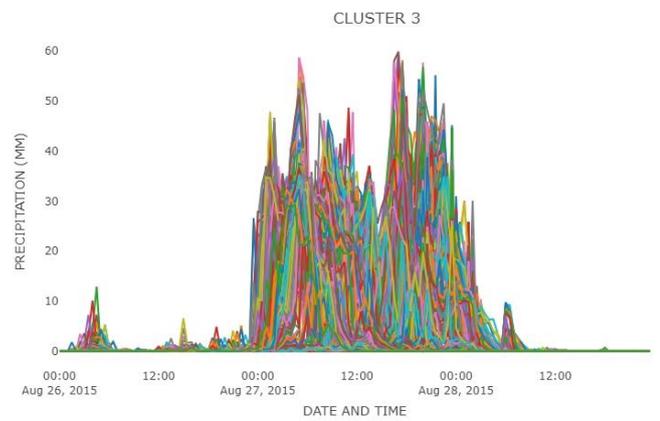
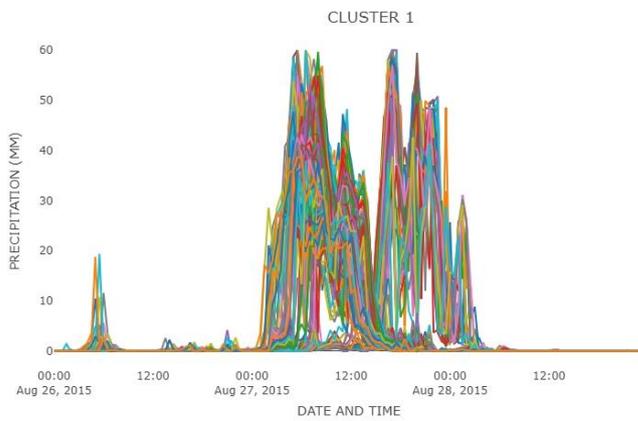
# APPENDICES

## Appendix 1: Time Series Plots for the Spatial Clusters





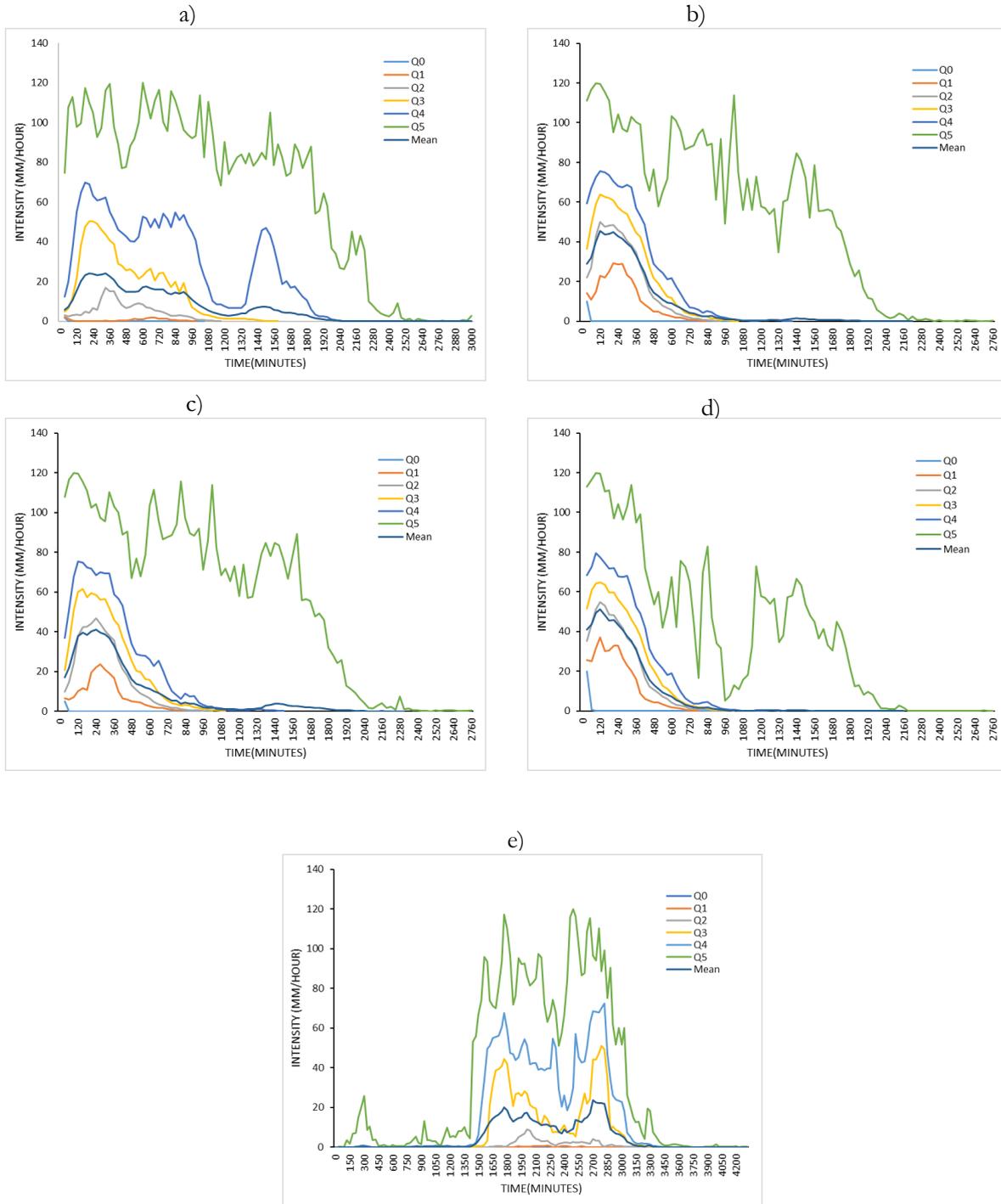
**Appendix 2: Time Series Plots for the Temporal Clusters T1, T3, T4, and T5 resulting from using K=5**



**Appendix 3: Plots of Time Step Quantiles Series for T3, T4, and T5 at varying Starting Thresholds. (Temporal Clusters result from using  $K=5$ )**

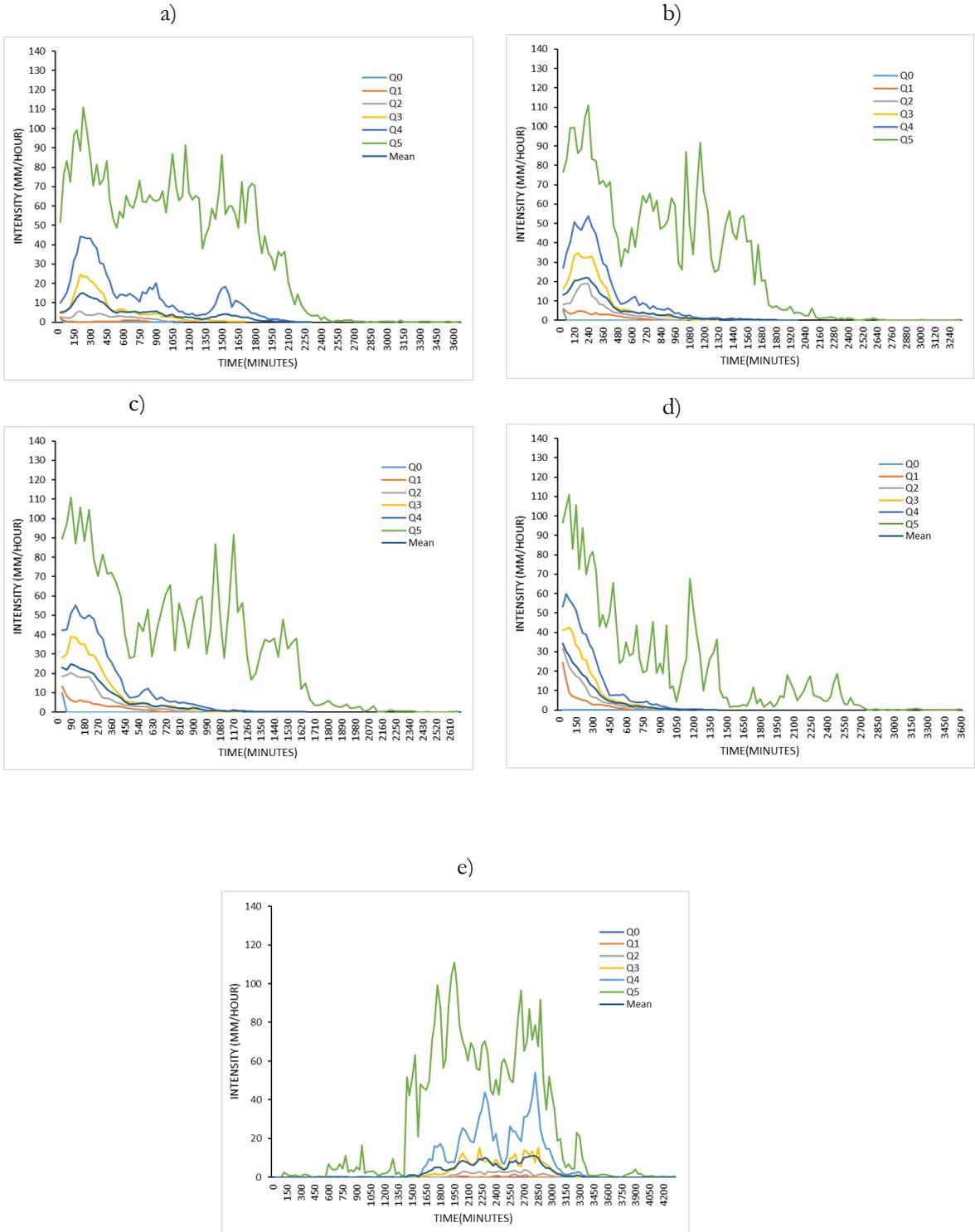
**A. CLUSTER T3**

a) Threshold  $\geq 1$  mm (2mm/hr). b) Threshold  $\geq 2.5$  mm (5mm/hr). c) Threshold  $\geq 5$  mm (10 mm/hr). d) Threshold  $\geq 10$  mm (20mm/hr). e) is no threshold



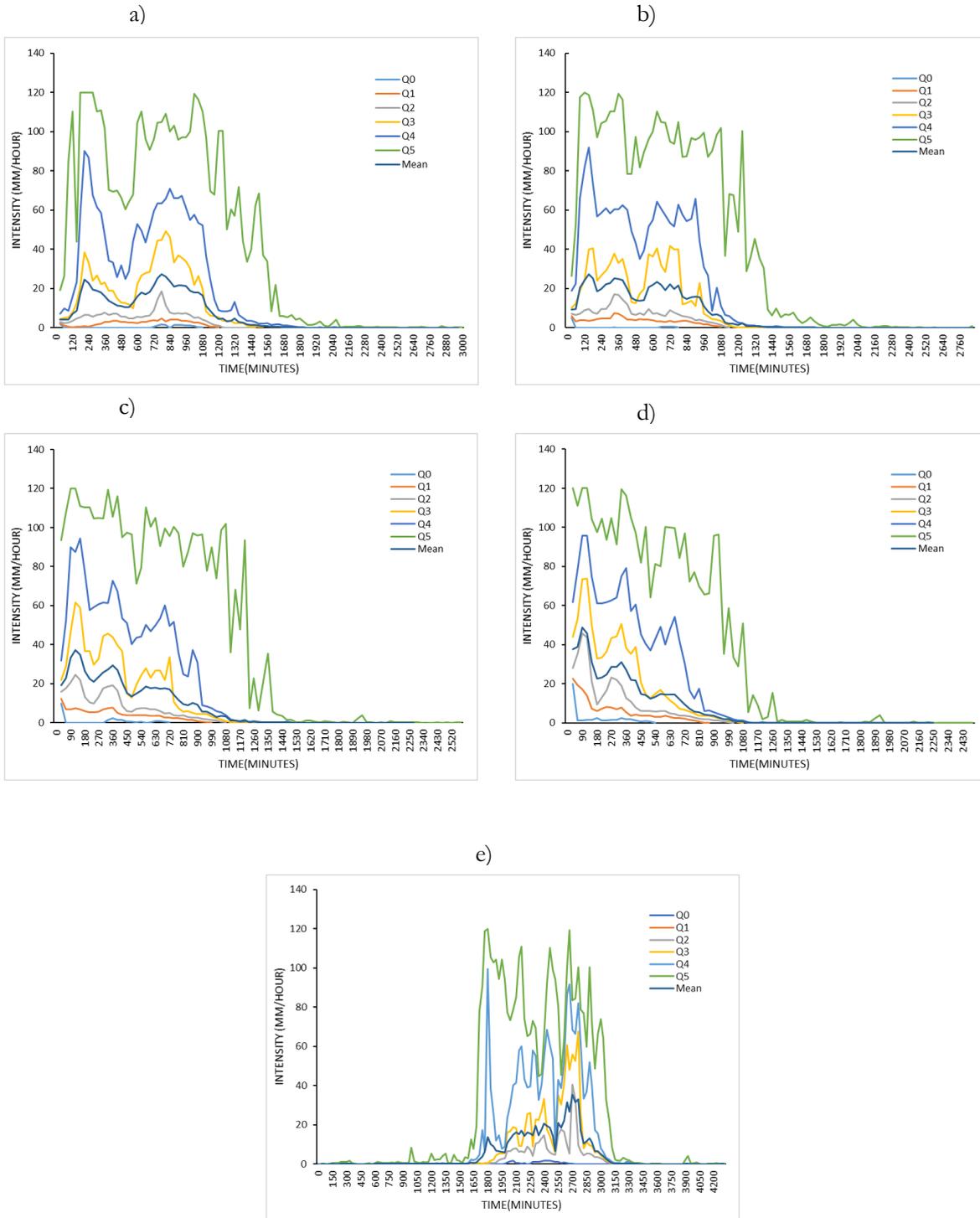
**B. CLUSTER T4**

**a)** Threshold  $\geq 1$  mm (2mm/hr). **b)** Threshold  $\geq 2.5$  mm (5mm/hr). **c)** Threshold  $\geq 5$  mm (10 mm/hr). **d)** Threshold  $\geq 10$  mm (20mm/hr). **e)** is no threshold



### C. CLUSTER T5

a) Threshold  $\geq 1$  mm (2mm/hr). b) Threshold  $\geq 2.5$  mm (5mm/hr). c) Threshold  $\geq 5$  mm (10 mm/hr). d) Threshold  $\geq 10$  mm (20mm/hr). e) is no threshold



**Appendix 4: Statistics Computed of Time Step Quantiles Series for T3, T4, and T5 at varying Starting Thresholds. (Temporal Clusters result from using K=5)**

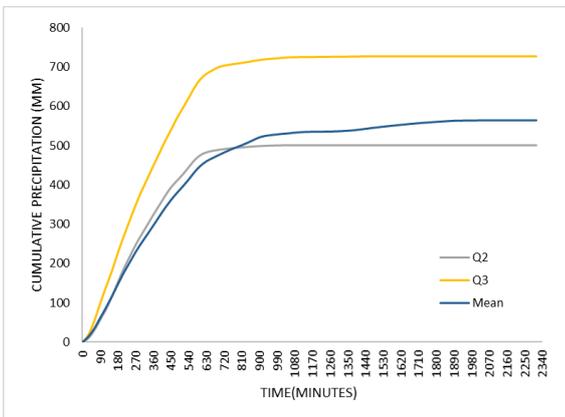
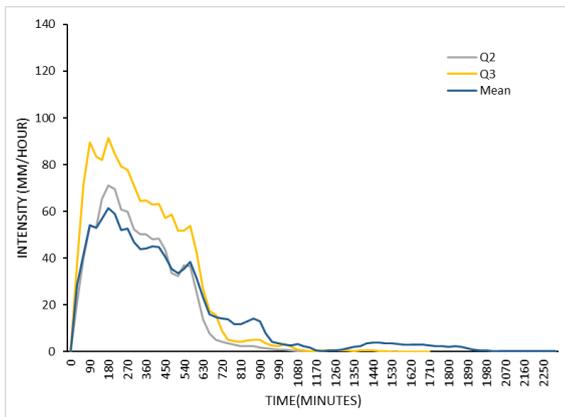
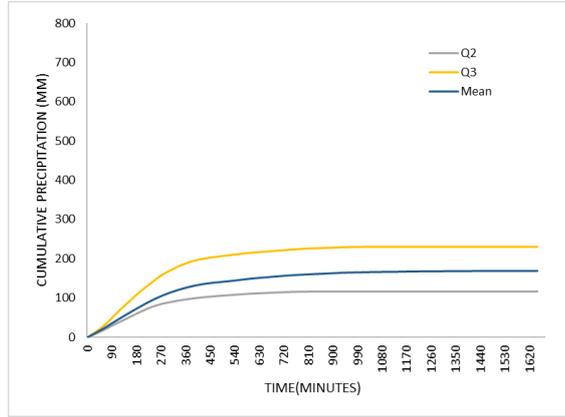
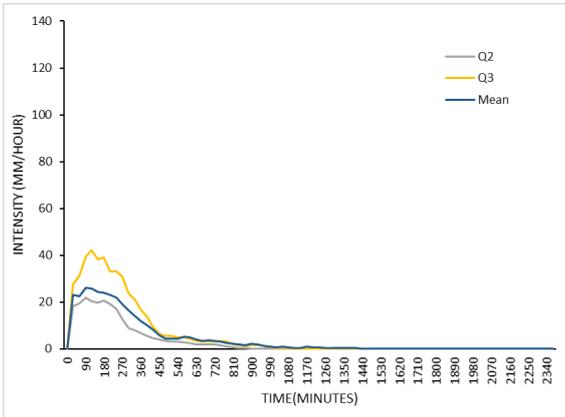
<b>T3</b>								
<b>Statistics</b>		<b>Q0</b>	<b>Q1</b>	<b>Q2</b>	<b>Q3</b>	<b>Q4</b>	<b>Q5</b>	<b>Mean</b>
<b>Original Data</b>	Accumulation (mm)	0.0	3.8	57.7	543.1	1157.8	2450.1	344.2
	Maximum Intensity (mm/hr)	0.0	0.8	9.0	51.2	72.5	120.0	23.6
	Period (Hours)	0.0	12.0	25.0	40.5	53.0	72.0	72.0
<b>Threshold &gt;=1 mm</b>	Accumulation (mm)	1.0	11.5	103.9	446.6	1096.2	3155.7	341.8
	Maximum Intensity (mm/hr)	2.0	2.2	16.8	50.2	69.8	120.0	24.3
	Period (Hours)	0.5	16.5	19.5	26.5	35.5	50.0	50.0
<b>Threshold &gt;=2.5 mm</b>	Accumulation (mm)	2.5	112.3	277.3	440.4	618.2	2592.8	338.2
	Maximum Intensity (mm/hr)	5.0	23.6	46.8	61.6	75.4	120.0	41.1
	Period (Hours)	0.5	14.0	16.0	18.5	25.0	46.0	38.5
<b>Threshold &gt;=5 mm</b>	Accumulation (mm)	5.0	155.5	<b>307.1</b>	<b>440.0</b>	587.0	2397.6	<b>334.6</b>
	Maximum Intensity (mm/hr)	10.0	29.2	<b>50.2</b>	<b>63.8</b>	75.8	120.0	<b>45.4</b>
	Period (Hours)	0.5	14.0	<b>15.5</b>	<b>17.5</b>	23.5	46.0	<b>37.0</b>
<b>Threshold &gt;=10 mm</b>	Accumulation (mm)	10.3	186.6	314.4	431.1	563.6	1803.0	328.0
	Maximum Intensity (mm/hr)	20.0	37.0	54.8	64.8	79.6	120.0	51.2
	Period (Hours)	1.0	13.0	15.0	17.0	22.0	46.0	36.0
<b>T4</b>								
<b>Statistics</b>		<b>Q0</b>	<b>Q1</b>	<b>Q2</b>	<b>Q3</b>	<b>Q4</b>	<b>Q5</b>	<b>Mean</b>
<b>Original Data</b>	Accumulation (mm)	0.0	8.1	45.1	172.2	510.8	1871.5	167.5
	Maximum Intensity (mm/hr)	0.0	1.6	3.8	15.3	54.0	111.0	11.2
	Period (Hours)	0.0	13.5	25.0	29.5	47.0	70.0	72.0
<b>Threshold &gt;=1 mm</b>	Accumulation (mm)	1.0	6.4	47.2	168.2	476.4	2248.7	165.1
	Maximum Intensity (mm/hr)	2.0	2.2	5.6	24.6	44.3	111.0	15.0
	Period (Hours)	0.5	14.0	19.0	28.5	36.0	61.0	38.5
<b>Threshold &gt;=2.5 mm</b>	Accumulation (mm)	2.5	33.7	98.3	212.5	365.5	1647.5	160.6
	Maximum Intensity (mm/hr)	5.0	6.0	18.8	34.7	53.7	111.0	21.8
	Period (Hours)	0.5	13.0	16.0	22.0	27.0	55.5	33.0
<b>Threshold &gt;=5 mm</b>	Accumulation (mm)	5.0	42.4	<b>107.6</b>	<b>209.9</b>	345.3	1484.4	<b>155.9</b>
	Maximum Intensity (mm/hr)	10.0	13.4	<b>20.2</b>	<b>38.9</b>	55.2	111.0	<b>24.8</b>
	Period (Hours)	0.5	13.0	<b>14.5</b>	<b>20.5</b>	23.5	44.5	<b>27.5</b>
<b>Threshold &gt;=10 mm</b>	Accumulation (mm)	0.0	50.9	106.9	201.6	313.7	1189.7	146.3
	Maximum Intensity (mm/hr)	0.0	24.3	31.6	42.8	59.8	111.0	34.4
	Period (Hours)	0.0	11.0	14.0	16.5	21.5	60.0	37.5
<b>T5</b>								
<b>Statistics</b>		<b>Q0</b>	<b>Q1</b>	<b>Q2</b>	<b>Q3</b>	<b>Q4</b>	<b>Q5</b>	<b>Mean</b>
<b>Original Data</b>	Accumulation (mm)	10.9	60.5	168.1	412.1	931.8	1977.7	331.4
	Maximum Intensity (mm/hr)	1.8	6.8	40.6	67.6	99.3	120.0	35.2
	Period (Hours)	12.0	20.0	22.5	28.5	46.0	72.0	72.0
<b>Threshold &gt;=1 mm</b>	Accumulation (mm)	7.1	46.7	122.2	466.5	957.4	2197.8	329.7
	Maximum Intensity (mm/hr)	2.0	4.6	18.4	49.2	90.1	120.0	27.3

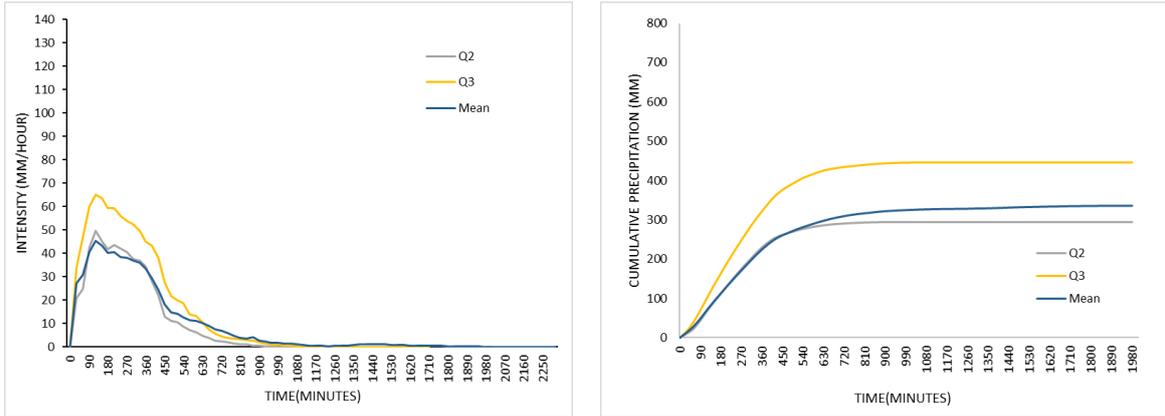
	Period (Hours)	18.0	19.5	24.5	25.5	30.0	50.0	49.5
<b>Threshold &gt;=2.5 mm</b>	Accumulation (mm)	3.7	65.3	135.3	426.1	929.1	2037.8	325.6
	Maximum Intensity (mm/hr)	5.0	7.4	17.0	41.8	92.1	120.0	27.2
	Period (Hours)	13.0	18.0	19.0	24.5	25.5	47.5	47.5
<b>Threshold &gt;=5 mm</b>	Accumulation (mm)	10.7	71.1	<b>159.7</b>	<b>427.3</b>	834.9	1969.2	<b>317.7</b>
	Maximum Intensity (mm/hr)	10.0	12.4	<b>24.6</b>	<b>61.6</b>	94.5	120.0	<b>37.2</b>
	Period (Hours)	12.0	16.5	<b>18.0</b>	<b>22.5</b>	25.0	43.0	<b>38.0</b>
<b>Threshold &gt;=10 mm</b>	Accumulation (mm)	22.3	87.1	189.3	404.4	751.2	1606.8	304.5
	Maximum Intensity (mm/hr)	20.0	22.8	45.8	73.8	95.9	120.0	48.8
	Period (Hours)	9.0	14.5	17.0	18.0	24.0	41.5	37.5

**Appendix 5: Cluster Representative Signals**

**A. Using temporal clusters resulting from  $K=4$**

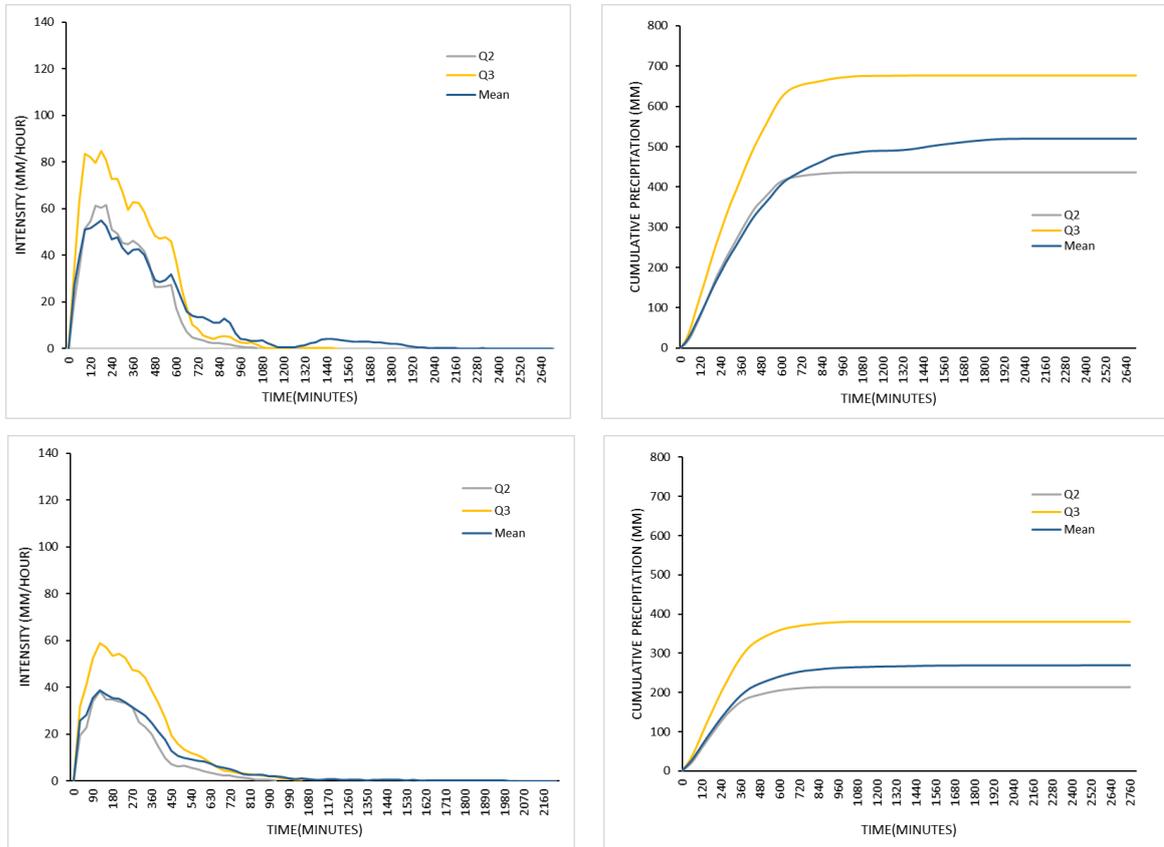
Rainfall time series (left) and cumulative plots (right) of the selected cluster representative signals. T1, T3, T4, and T5 plots are in rows 1, 2, 3, and 4, respectively.





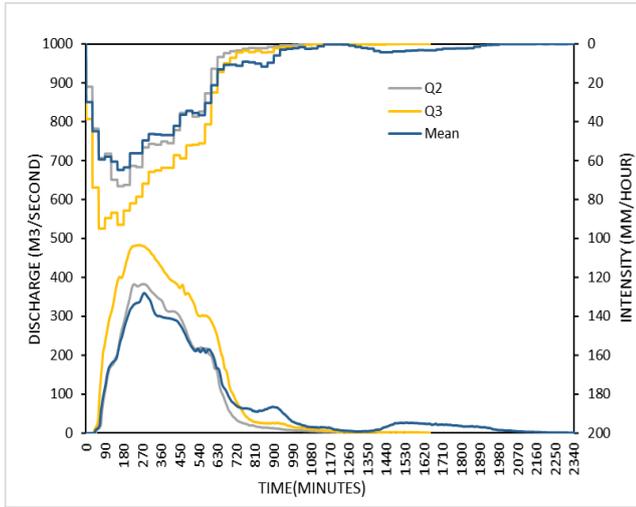
**B. Using  $K=3$**

Rainfall time series (left) and cumulative plots (right) of the selected cluster representative signals. T1, T3, T4, and T5 plots are in rows 1, 2, 3, and 4, respectively.

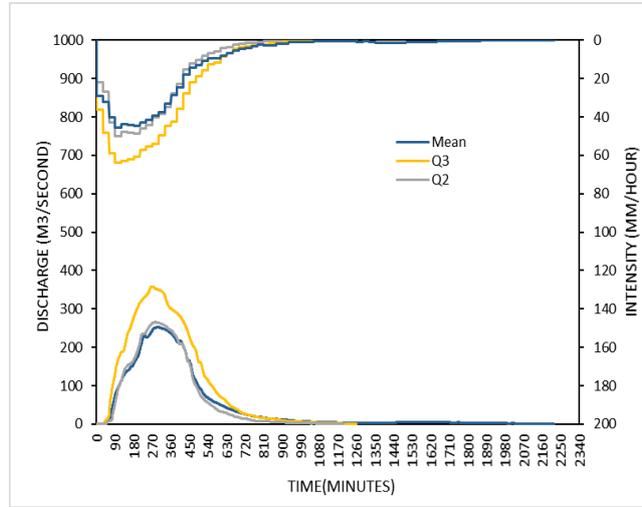


**Appendix 6: Discharge Hydrographs for the Representative Signals of the Clusters resulting from K=5**

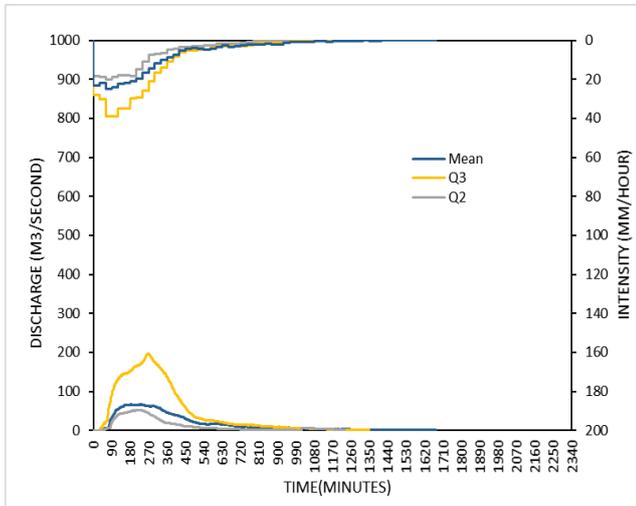
**Cluster T1**



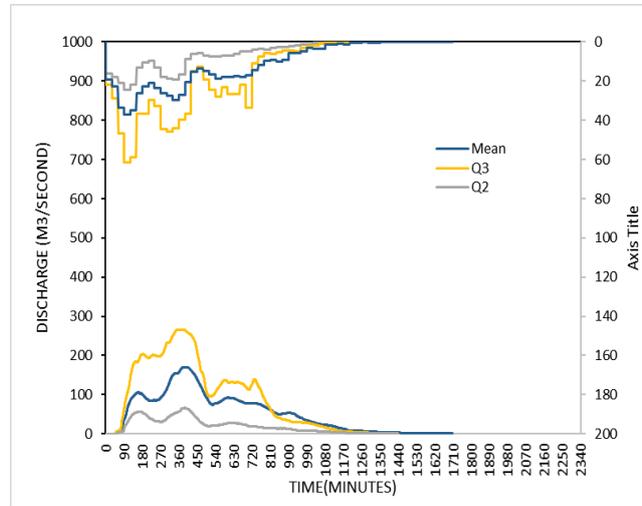
**Cluster T3**



**Cluster T4**



**Cluster T5**



**Appendix 7: Dominica Design Storms adopted from (Jetten, 2016)**

