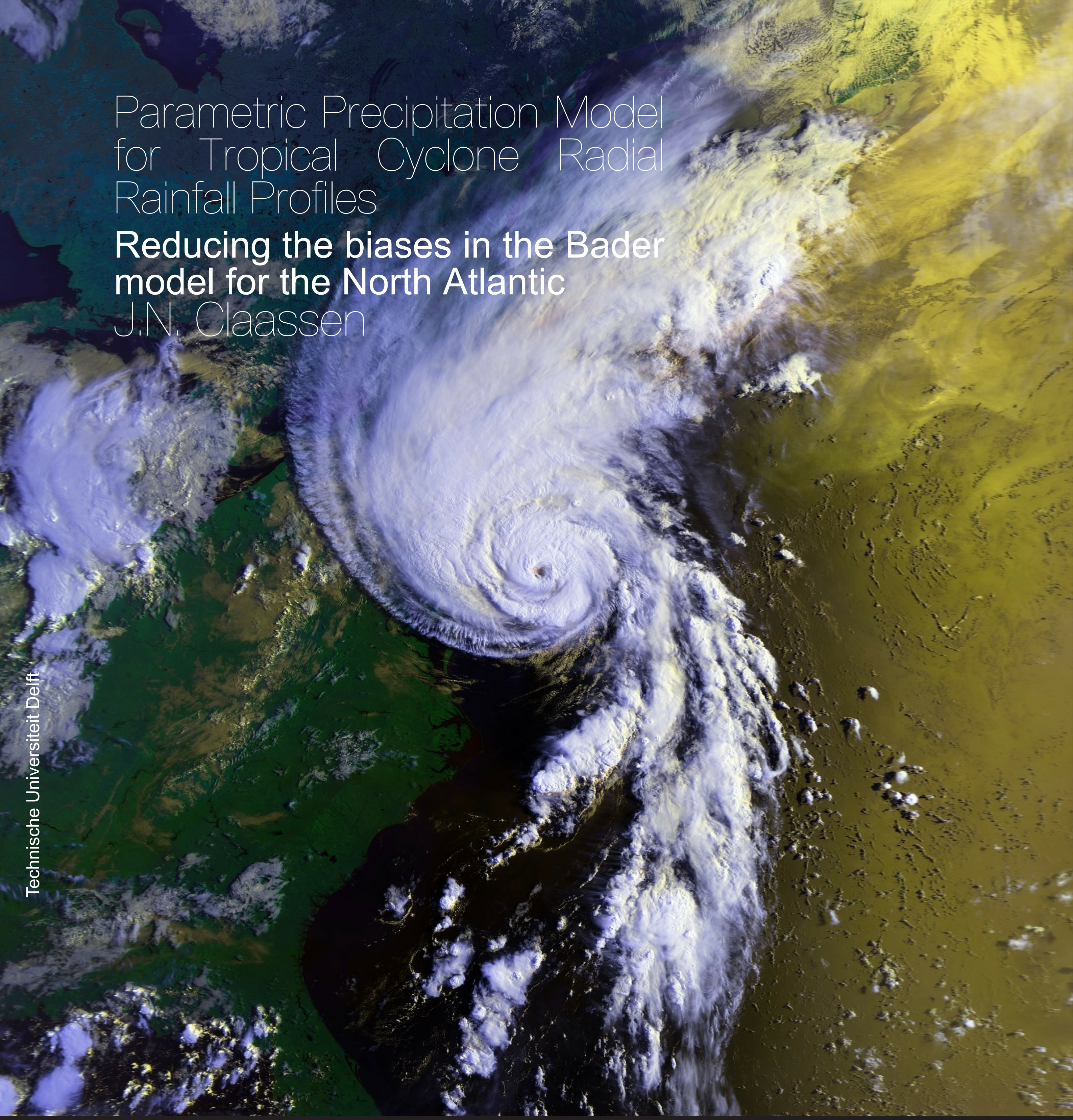
Parametric Precipitation Model for Tropical Cyclone Radial Rainfall Profiles

Reducing the biases in the Bader model for the North Atlantic

J.N. Claassen



Technische Universiteit Delft

Deltares

TUDelft

Parametric Precipitation Model for Tropical Cyclone Radial Rainfall Profiles

Reducing the biases in the Bader model for the North Atlantic

by

J.N. Claassen

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on July 8th, 2021 at 14:45.

Student number:	4490649	
Project duration:	November, 2020 – July, 2021	
Thesis committee:	Dr. M. A. Schleiss,	TU Delft, Chair
	Ir. T. W. B. Leijnse,	Deltares, Company Supervisor
	Dr. MC. ten Veldhuis,	TU Delft
	Dr. F. Glassmeier,	TU Delft

Preface

Four years ago I embarked on a journey across the ocean to find the American dream. Instead I found a passion for tropical cyclones. It was the year 2017, the fascinating year of Harvey, Irma and Ophelia. Next thing I knew I was in the office of the Weather and Climate Professor, full of excitement exclaiming I wanted to study hurricanes, just to find out I had zero technical skills to even begin to dream of such a project. Hence, with a liberal arts and science diploma in the pocket, I made my way to a place that would get me just one step closer, TU Delft.

Soon after my arrival to Delft I was introduced to Deltares during a lunch lecture where (and I quote literally) the Deltares researcher stated ”If you didn’t start programming yesterday, you’re already too late”. From that moment onward, I knew - this is the company that will turn me into a better engineer.

Looking back now, I couldn’t be happier to have experienced both the tropical cyclone related topic and the company in my thesis project. Yet none of this would have been possible without the help and guidance of some notable people.

First and foremost, I’d like to thank my supervisors, Marc Schleiss and Tim Leijnse. I want to acknowledge Marc for his time and effort to frequently provide me with excellent critical feedback. Without our routine weekly meetings I would not have been able to stick to my strict thesis schedule. I am grateful to Tim for sharing his knowledge on the topic, showing me the ins and outs of the Deltares work environment, and above all, for always making me feel like I had someone to talk to whenever I needed help. In addition, I want to thank the committee members, Marie-Claire ten Veldhuis and Franziska Glassmeier, for their time and feedback that greatly improved the quality of this thesis. Furthermore, I would like to thank Robert McCall for enabling me to conduct this research at Deltares and always actively participating at all progress meetings. I’d also like to thank Ferdinand Diermanse, for showing special interesting in my work.

With this thesis, my time as a student in Delft has come to an end. To the Efteling crew (you know who you are) I am grateful to have gone through the TU Delft master roller coaster experience with you all. Thank you for making Delft feel like home. I’d also like to thank the friends from my bachelor studies, my banda de las chicas, even when we are not studying at the same university anymore, our daily communication makes me feel your constant presence.

I would like to give special thanks to my family, most notably my parents, for supporting my education financially and reminding me to take moments for myself to slow down. Finally, I want to express my gratitude to Eric Lacoa Arends, for listening to me talk about nothing but the thesis the past months. Thank you for always believing in me, making me push myself to always be better.

J. N. Claassen
Delft, July 2021

Summary

Torrential rain from tropical cyclones can have a devastating impact, causing loss of life and billions in damages. To better understand the risk faced by coastal communities, it is important to estimate how often a tropical cyclone could occur and how much rainfall it will produce. One way to do this is by analyzing past storms and building parametric models of rainfall rates during tropical cyclone events. While many parametric precipitation models –such as the Bader model– exist, their accuracy remains limited and many challenges still need to be overcome. The most important challenges are output overestimation and a poor representation of rainfall over land. Therefore, this thesis aims to reduce these biases by answering the following research question:

How can the bias in the radial rainfall distributions of a tropical cyclone in Bader's parametrized model be reduced and be used for reliable rainfall estimates both above land and the ocean?

To answer this question, several new data sources were introduced from the TRMM/GPM satellites and Stage IV to improve the Bader model. While this original model only predicted precipitation based on maximum wind speed (v_{max}), the updated model also considers pressure deficit ΔP. The results suggest that ΔP can be a useful parameter to reduce bias and improve accuracy. However, it also leads to larger uncertainty ranges. Next, four precipitation profiles were proposed. A profile where precipitation is constant for low p_{max} values based on the predicted total rainfall (area under the graph) was selected for further exploration.

The new models are explored during a case study of Hurricane Florence. Both the ΔP- and v_{max}-based models produced satisfactory results, compared to the benchmark IPET model. Moreover, an alternative fit above land has been proposed, where the highest precipitation is simulated at the eye. The proposed land fit improved the median of the predictions based on both v_{max} and ΔP. The ΔP based model performed the best in the case study, however, no definitive conclusion could be reached upon which model is most suitable overall as more case studies would be required.

Finally the updated model has been compared to the original Bader model. The new data ensured better representation over land, the overestimation of precipitation was reduced, and the model was applied with more confidence outside of the training data set. Consequently, results showed an improvement on its prediction capabilities. As a concluding remark, this research project highlights the importance of having insightful data to enhance the decision-making and risk management of natural hazards: a model that accurately quantifies uncertainty and the risks associated with a TC, representing a valuable tool for better understanding flood risk. Nonetheless, there are still several ways to further improve the modeled profiles (e.g., by including more data, introducing asymmetry or adding temporal auto-correlation).

Contents

List of Figures

List of Tables

Acronyms

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CDF	cumulative distribution function
ConUS	contiguous United States
DPR	Dual-frequency Precipitation Radar
EBTRK	Extended Best Track Dataset
EOL	Earth Observing Laboratory
EORC	Earth Observation Research Center
GMI	GPM Microwave Imager
GPM	Global Precipitation Mission
GPROF	Goddard Profiling algorithm
HURDAT	Huricane Database
IPET	Interagency Performance Evaluation Task Force
JAXA	Japan Aerospace Exploration Agency
MAE	Mean absolute error
MSR	Modified Smith for Rainfall
MvCAT	Multivariate Copula Analysis Toolbox
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NHC	National Hurricane Center
NOAA	National Oceanic and Atmospheric Administration
NSE	Nash-Sutcliffe Efficiency
PHRaM	Parametric Hurricane Rainfall Model
PR	Precipitation radar
QSCAT-R	QuikSCAT Tropical Cyclone Radial Structure
R-CLIPER	Rainfall- CLImatology and PERsistence

RAMMB	Regional and Mesoscale Meteorology Branch
RFCs	River Forecast Centers
RMSE	Root Mean Squared Error
TC	Tropical Cyclone
TCWiSE	Tropical Cyclone Wind Statistical Estimation Tool
TD	tropical depression
TMI	TRMM Microwave Imager
TRMM	Tropical Rainfall Measuring Mission
TS	tropical storm

Symbols

N number of samples [-]

P_{eye} Pressure at the eye of the TC [hPa]

P_{outer} Pressure at the outer closed isobar of the TC [hPa]

R radius away from the eye [km]

R_{PHRaM} PHRaM precipitation [mm/hr]

$R_{R-CLIPER}$ R-CLIPER precipitation [mm/hr]

$R_{\textbf{shear mod}}$ shear modification [-]

$R_{\textbf{topography}}$ Orographic lifting component [-]

Rp_{max} radius of maximum precipitation [km]

Rv_{max} radius of maximum sustained wind speed [km]

TRR R-CLIPER predicted rain rate at particular radius [in/d]

T_0 Rain rate at the eye [in/d]

T_m maximum rain rate at r_m [in/d]

U Normalized maximum wind

V_m maximum sustained wind speed [kt]

ΔP Pressure deficit [hPa]

β_b BirnbaumSaunders scale parameter [-]

γ_b BirnbaumSaunders shape parameter [-]

λ_{ig} Inverse Gaussian shape parameter [-]

$\mathbf{V}_s$ Surface (10 m) wind field [m/s]

μ_l Mean of logarithmic values [-]

μ_{ig} Inverse Gaussian scale parameter [-]

σ_l Standard deviation of logarithmic values [-]

θ_1 Copula Parameter [-]

θ_{Di} Angle [°]

θ_{Ri} Angle [rad]

θ_a azimuthal angle [°]

a_i Fourier coefficients describing the azimuthal variations of the wavenumber-i fields [-]

a_n R-CLIPER operational constant [-]

b_i Fourier coefficients describing the azimuthal variations of the wavenumber-i fields [-]

b_n R-CLIPER operational constant [-]

b_s Holland fitting coefficient [-]

c Constant of proportionality [-]

h_s Gradient of elevation [-]

n_1 Fitting coefficient 1 [-]

n_2 Fitting coefficient 2 [-]

n_3 Fitting coefficient 3 [-]

p Random sample from $U(0, 1)$ [-]

p_{max} maximum precipitation [mm/hr]

p_r Precipitation [mm/hr]

r_e R-CLIPER curve fit parameter [-]

r_m radius of maximum rain [km]

u CDF of v_{max} [-]

v CDF of p_{max} [-]

v_{max} maximum sustained wind speed [m/])

x_{eye} Longitude of the eye [°]

x_n Holland fitting coefficient [-]

y_{eye} Latitude of the eye [°]

1

Introduction

1.1. Context

Tropical cyclones (TCs) can have devastating impact when they hit land, producing high social and economic losses. A well-known example is the 2017 Hurricane Harvey that cost approximately 68 lives and $125 billion in damage, caused predominately by heavy rainfall [1].

In order to properly prepare a coastal settlement against any TC, a reliable early warning system needs to be in place, responsible of providing real time information about surge levels, wind speed and precipitation intensity. As this information needs to be retrieved in a short period of time with high accuracy, the early warning system model is required to be computationally efficient and effective [2]. In addition, both climate variability assessments as well as flood risk assessments require accurate climatological and statistical models of rainfall rates to determine the likelihood of a TC event in an area of interest.

1.2. Challenges

Basic TC parameters such as storm motion, surface wind speeds and atmospheric pressure can be reproduced relatively well [3]. However, aside from a TCs characteristic high wind speed and storm surge, large amounts of precipitation can also lead to large flood-related damages such as overtopping riverbanks or lack of infiltration capacity [4]. Accurately predicting rainfall rates in tropical cyclones still represents a major challenge for numerical weather models. The large computational cost of numerical weather models can also be prohibitive when it comes to simulating thousands of different scenarios. Therefore, parametric models are investigated as a more efficient and simple alternative. Parametric rainfall models allow to estimate precipitation comparatively faster based on a set of TC parameters. Currently, several parametric TC precipitation models do exist, however, numerous factors continue to limit the potential of these tools. The most prominent issues with the current approaches are:

- Biased rainfall estimates: current models under- or overestimate total rainfall amounts, which leads to over- or underestimated flood risks. [5].
- Difficulty in reproducing realistic patterns of temporal variability: current models underestimate the variation in rainfall intensity over time.
- Poor generalization: some models proposed in literature turn to be unreliable when tested in new situations, decreasing performance.

- Poor representation of rainfall over land: training data sets have a good representation for precipitation above the ocean, yet data above land is sparse. Additional data with higher performance over land is required to better understand the flood risk [4]

.

1.3. Research questions

This thesis was born from a collaboration between TU Delft and Deltares - a leading institute for applied research on water and subsurface. Deltares provides TC risk assessment for the present and the future based on their synthetic cyclone simulator known as Tropical Cyclone Wind Statistical Estimation Tool (TCWiSE) [3]. The main objective of the research is to reduce the biases in Deltares's currently used Bader model and improve the accuracy of rainfall profiles in TCs above the North Atlantic ocean [4]. Hence, the research question is:

How can the bias in the radial rainfall distributions of a tropical cyclone in Bader's parametrized model be reduced and be used for reliable rainfall estimates both above land and the ocean?

The main research question is answered through the use of the following sub-questions:

1. What data set is most suitable to characterize radial rainfall profiles both over the ocean and over land?
2. What is the difference between tropical cyclone radial rainfall profiles over land and over the ocean?
3. How can parametric rainfall models for tropical cyclones be improved to better capture the observed variability over space and time?

Sub-question 1 is approached by introducing several data sets from the Tropical Rainfall Measuring Mission (TRMM) and Global Precipitation Mission (GPM), as well as Stage IV data. These data sets are blended together for optimal performance. Next, sub-question 3 is answered through exploring the pressure deficit (ΔP) as a predictive parameter, and testing a variety of radial rainfall fitting methods. Finally, sub-question 2 is investigated by performing a case study of Hurricane Florence (2018) to provide further insights on how the models performs over time.

1.4. Outline

To address the research questions above, the remainder of this report has been divided in five chapters. Chapter 2, the literature review, conceptualizes tropical cyclone fundamental processes, followed by a compilation of existing parametric models. Chapter 3 serves as an overview of all data sources used. Chapter 4 describes the methods to conduct the research. This includes a data blending technique, estimation of radial profiles, relevant toolboxes, performance metrics and case study content. Chapter 5, the results and discussion, presents the best copula fits and provides the model coefficients, followed by an interpretation of different model variations. The proposed model is further tested in a case study of Hurricane Florence, which hit the US in 2018. Here, the model performance is evaluated, while a distinction between land and ocean behaviour is made. Finally, the new model is compared to the original Bader model to assess whether the goal of the research has been met. The work is concluded in chapter 6, providing the key findings, limitations and future recommendations.

2

Literature Review

The literature review introduces the most relevant concepts for the thesis. Section 2.1 focuses on tropical cyclone fundamental processes including a more in depth understanding of tropical cyclone precipitation. Section 2.2 highlight the different parametric models that currently exist, showcasing both their advantages and limitations.

2.1. Tropical cyclone fundamentals processes

A Tropical Cyclone (TC) is a low pressure system that forms over tropical waters. A TC is characterized by a closed wind circulation and a defined center, called the eye. They are known for their associated high wind speeds and precipitation [6].

2.1.1. Formation & structure

According to the National Oceanic and Atmospheric Administration (NOAA) of the US, TCs only form under specific conditions. First of all, TCs are fueled by warm oceans, requiring water temperatures of above 26 °C for a depth of approximately 45 meters. Additionally, a large enough temperature gradient, and humid conditions in the atmosphere are needed in order for the warm moist air to rise and form clouds as well as thunderstorm conditions [6].

Furthermore, the cyclone needs a force that allows it to spin. This is known as the Coriolis force, that deflects movement to the right in the northern hemisphere and to the left in the southern hemisphere due to the rotation of the earth. The Coriolis force is zero at the equator, therefore, TCs can only form in the tropics, at least 300 km away from the equator. Finally, a low vertical wind shear can favor thunderstorm formation, a major energy source for TCs. Wind shear is the change of wind speed and direction with height, which has to at an adequate, as too much can inhibit convection [6].

A TC has an eye, eyewall, and spiral rain bands. These features can be recognized in fig.2.1. The eye is circular area characterized by low pressure, warm temperatures and relatively calm weather. In contrast, the eye wall, a ring of deep convective clouds around the eye, is usually where the highest winds and precipitation intensities are observed. The spiral bands are long, narrow bands of rain and thunderstorms that are oriented in the same direction as the wind. Like the eyewall, they are caused by convection, spiraling towards the center of the storm [6].

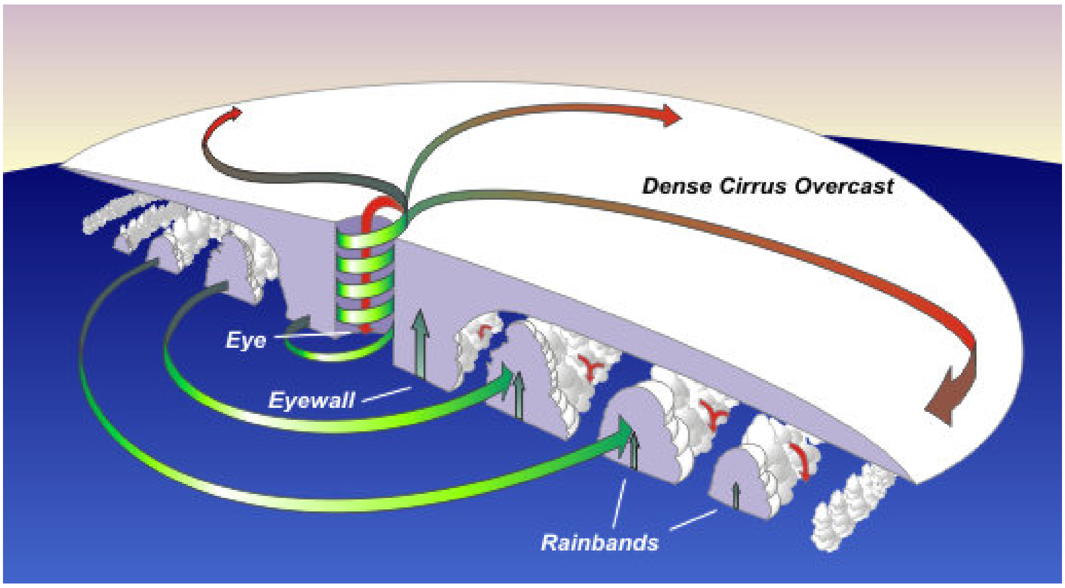


Figure 2.1: Overview of the Tropical Cyclone structure including the eye, eyewall and rain bands [7].

2.1.2. Wind & pressure

As previously mentioned TCs are low pressure systems where the central pressure deficit (pressure difference between the center of the storm and the outside) is large [8]. Generally, the stronger the TC, the higher the pressure deficit. The relationship between pressure deficit and strength can be explained by the wind-pressure relationship [9]. A larger pressure deficit means a steeper gradient in pressure from the outer closed isobar to the pressure at the eye. When the gradient is steeper the winds become more intense and the radius at which the maximum wind is measured occurs closer to the eye [8]. Both the maximum wind as well as the pressure deficit are of great interest to rainfall modellers as its a good indicator for the occurrence of precipitation.

Furthermore, as pressure and wind are closely related to the intensity of the TC, the TC strength has been classified by the maximum sustained wind speed. If the wind speed is under 34 knots, it is called a tropical depression (TD), between 34 and 64 knots it is a tropical storm (TS). Above 64 knots it is called a hurricane in the Atlantic and East Pacific, or a typhoon in the northern and West Pacific [6]. Within this denomination, it can further be classified in categories. This classification is known as the Saffir-Simpson Hurricane Wind Scale. The scale ranges from category 1 to 5, where 5 is the most intense [10]. The conditions for each category are summarized in table 2.1.

2.1.3. Precipitation

As introduced earlier, TCs can produce heavy rainfall. In some regions of the world this can contribute to 15-17% of the total annual rainfall, where many TCs are responsible for the highest rainfall accumulation on an hourly and daily time scale [11].

TC precipitation can be categorized by two main types, convective and stratiform. Convective precipitation is caused by the rising hydrometeors that grow with altitude until they fall down. Stratiform precipitation, on the other hand, is caused by weak vertical air motion and drifts down over a larger area of weak updraft, and grow slowly due to aggregation and deposition [11]. Convective precipitation can mostly be found in the inner eye wall, where there are high wind speeds, and stratiform precipitation forms further away from the eye (see fig.2.2).

In the 2004 paper on precipitation distribution of TCs using the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), the change in rainfall intensity based on radial distance from the eye

Table 2.1: Saffir-Simpson Hurricane Wind Scale

Category	**Sustained Winds**
1	4-95 mph 64-82 kt 119-153 km/h
2	96-110 mph 83-95 kt 154-177 km/h
3 (major)	111-129 mph 96-112 kt 178-208 km/h
4 (major)	130-156 mph 113-136 kt 209-251 km/h
5 (major)	157 mph or higher 137 kt or higher 252 km/h or higher

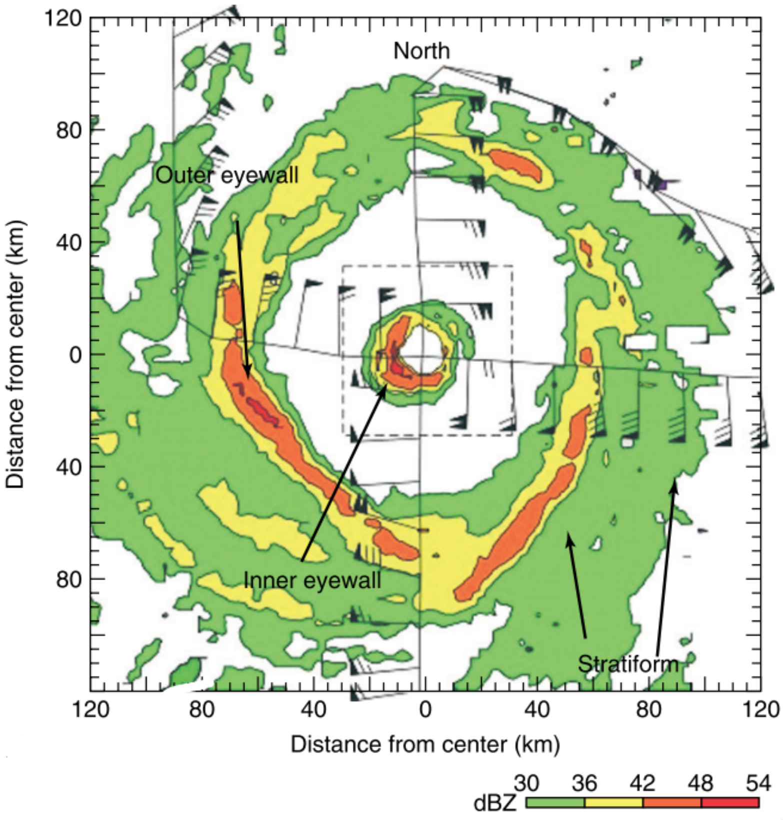


Figure 2.2: Tropical Cyclone distribution: this figure is Rogers et al. adaption of a figure originally presented by Dodge et al. showing Hurricane Gilbert on September 14th 1988 [11, 12].

is further highlighted [13].

Fig.2.3 shows the azimuthal mean rain rates of 260 TCs in all basins between 1998 and 2000. Where tropical storm (TS) are categorized by winds < 33 m/s, CAT12 are category 1–2 hurricanes ,and

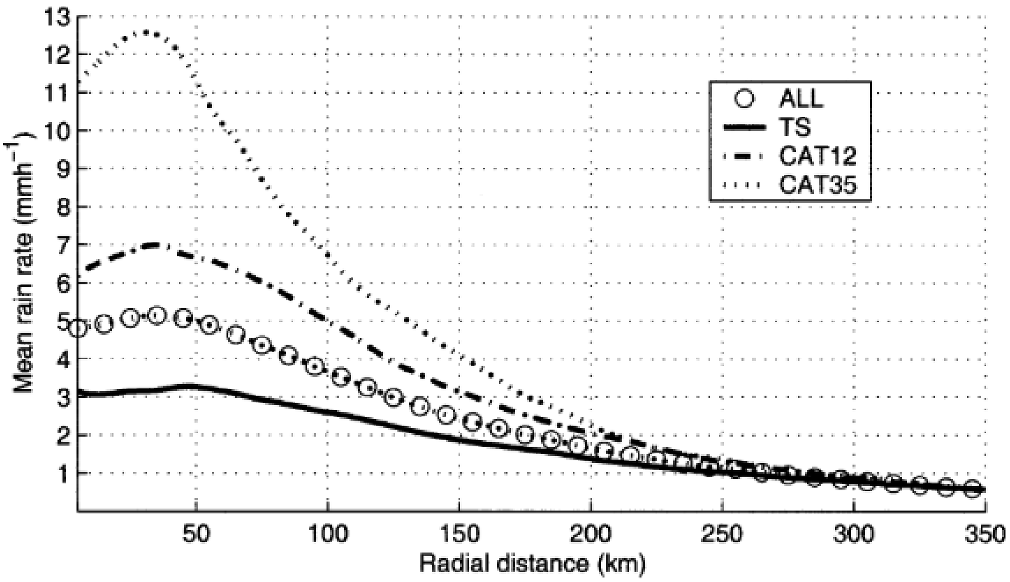


Figure 2.3: Azimuthal mean rain rate for different strengths in the TRMM data [13].

CAT35 are category 3–5 hurricanes. Here, it can further be interpreted that more intense precipitation is usually found within a 100 km away from the eye. Additionally, it is clear that higher storm categories are associated with both higher wind speeds and higher precipitation [13].

The distribution shown in fig.2.3 assumes a symetric radial profile, However, in reality precipitation varies greatly at equal distances from the eye. Most precipitation occurs in the direction of the storm motion ahead of the eye (in front of the storm) [13].

2.1.4. Distinction between precipitation above land and ocean

While Lonfat et al. observed the precipitation associated with TCs, no distinction was made between this event over ocean and over land. However, TC precipitation distribution can vary greatly between ocean and land conditions. When a storm makes landfall, additional factors, such as the presence of significant topography and extra tropical transition (transition from tropical to extra-tropical cyclone), can occur [14, 15].

Therefore, in the 2006 paper by Jiang et al. the authors test how TC rainfall varies over ocean and over land. To do so, they used TRMM data to observe 37 land falling TCs between 1998 and 2004.

In the study of Jiang et al. a TC is considered to be on land if 60% of the pixels are above land. Likewise, a TC is considered to be above the ocean if 60% of the pixels are above the ocean. Based on this division the azimuthal mean rain rates were calculated every 28 km away from the eye. the average of the azimuthal profiles have been computed for all land/ocean values (ALL), and for tropical storm (TS) with wind speeds < 17 m/s, tropical storm (TS) with wind speeds between 18 and 33 m/s, and hurricanes (HUR) with wind speeds > 34 m/s. The results are displayed in fig.2.4.

The profile shown in fig.2.4 suggests that mean rain rates increase with storm intensity at all radii for both over land and over ocean observations. The location of the peak rainfall also varies with intensity, as would be expected by the change in wind speed mentioned previously. For the HUR category, the peak rain rate for land observations is higher and the location of this peak is much closer to the storm center than those for ocean observations [14].

Furthermore, Jiang et al test the fundamental relationship between the maximum sustained wind

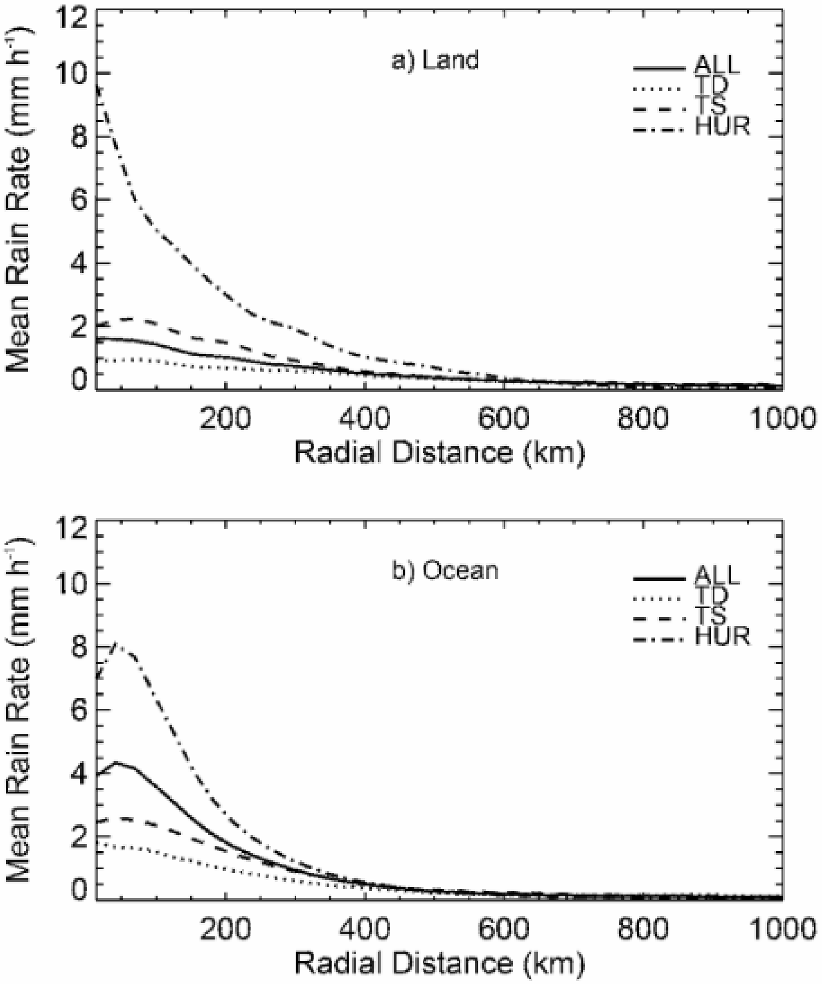


Figure 2.4: Azimuthal mean rain rate for different strengths: land vs. ocean [14].

speed and the maximum precipitation above land and above the ocean. As seen in fig.2.5, the correlation between maximum wind speed and maximum precipitation is higher over the ocean compared to over land. Therefore, there may be a reduced model performance for a parametric precipitation model based on wind speed above land.

2.2. Parametric precipitation models

A parametric precipitation model predicts precipitation based on other TC parameters, such as storm motion, surface wind speeds and atmospheric pressure. A good parametric rainfall model should be simple and need as little information as possible to achieve reasonable predictive performance at a low computational cost. Several parametric models already exist and will be highlighted in the following subsections. An overview of the different models can also be found in table 2.2.

2.2.1. R-CLIPER

The Rainfall- CLImatology and PERsistence (R-CLIPER) model, is said to be the first parametric model. Two TC R-CLIPER models were developed by Marks and DeMaria in 2003, based on the works of

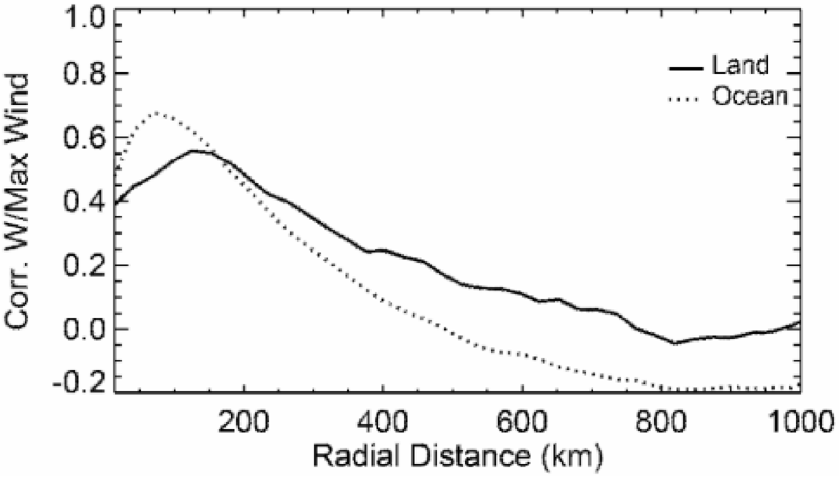


Figure 2.5: Correlations between the azimuthal mean rain rate and maximum wind speed [14].

Table 2.2: Existing parametric models including their training data, authors, pros, and cons.

Model	Precipitation training data	Authors	Pros	Cons
R-CLIPER 2003	TRMM TMI	Marks & DeMaria; Tuleya et al	- Little input variables needed	- Underestimates rainfall
PHRaM 2007	TRMM TMI	Lonfat et al	- Considers orographic lift - Considers assymetry	- Underestimates rainfall
IPET 2006	TRMM	US Army Corps of Engineers	- Very simple - Considers assymetry	- Overestimates rainfall
MRS 2009	TRMM PR	Langousis and Veneziano	- Good results over ocean	- Only valid over the ocean
Snaiki and Wu 2018	TRMM	Snaiki and Wu	-Good results -Attempts to incorporate rainbands	- Not fully parametric, part physical
Bader 2019	Qscat R	Bader	- Little input variables needed - Provides an uncertainty range	- Overestimates rainfall - Unable to capture spatial and temporal variability

Lonfat et al. [13, 16, 5]. The first was developed using hourly gauge data in order to evaluate TCs that make landfall. However, as the hourly gauge data was sparse, especially within the most intense 100 km from the storm center, a second model was proposed. The second R-CLIPER estimates rain from NASA's TRMM satellite. In the second R-CLIPER model, rainfall at different radii (TRR) away from the eye are calculated by several equations.

$$TRR(R) = \begin{cases} T_0 + (T_m - T_0)\left(\frac{R}{Rp_{max}}\right), & R < Rv_{max} \\ T_m e^{-\left(\frac{R - Rv_{max}}{r_e}\right)}, & R \geq Rv_{max} \end{cases} \tag{2.1}$$

$$U = 1 + \frac{V_m - 35}{33} \tag{2.2}$$

$$T_0 = a_1 + b_1 U \tag{2.3}$$

$$T_m = a_2 + b_2 U \tag{2.4}$$

$$Rv_{max} = a_3 + b_3 U \tag{2.5}$$

$$r_e = a_4 + b_4 U \tag{2.6}$$

Here, R is the radius of interest, T_0 is the rainfall at the eye, T_m is the maximum rain rate at the radius of maximum rain (Rp_{max}), and r_e is a curve fit parameter. T_0, T_m, r_m, r_e can be represented by a linear function of the normalized maximum wind (U). U is a function of V_m, the maximum sustained wind speed in knots (eq. 2.2). The constants a_n and b_n are fitted according to the data globally. R-Clipper solely depends on storm track and maximum wind speed. A symmetric distribution of rain rate inside the TC is assumed [4, 17, 5].

2.2.2. PHRaM

In order to incorporate additional relevant parameters, The Parametric Hurricane Rainfall Model (PHRaM) was introduced by Lonfat et al. to build on the original R-CLIPER algorithm. PHRaM does account for asymmetry by including vertical wind shear and topographic effects in the governing equations [4, 17, 5, 18].

$$R_{PHRaM} = R_{R-CLIPER} + R_{\text{shear mod }} + R_{\text{topography}} \tag{2.7}$$

$$R_{\text{shear mod }}(R, \theta_a) = \sum a_i(R)\cos(i\theta_a) + \sum b_i(R)\sin(i\theta_a) \tag{2.8}$$

$$R_{\text{topography }} = c\mathbf{V}_s \cdot \nabla h_s \tag{2.9}$$

The precipitation, R_{PHRaM} is the sum of $R_{R-CLIPER}$, the R-CLIPER precipitation (eq.2.1), the shear modification $R_{\text{shear mod}}$, and $R_{\text{topography}}$, the orographic lifting component. $R_{\text{shear mod}}$ is described by eq.2.8, where R is the radius away from the eye of the TC, a_i and b_i are Fourier coefficients describing the azimuthal variation of the wavenumber-i fields. The subscript i is the wavenumber considered between 1 and 2. θ_a is the azimuthal angle. $R_{\text{topography}}$ is described by eq.2.9, where c is a constant of proportionality, $\mathbf{V}_s$ is the surface (10m) wind field, and h_s is the ground elevation [18].

PHRaM was compared to R-CLIPER using Stage IV data for 2004 TCs showing that PHRaM almost doubles the predictive skill of the mean storm total rainfall [18].

2.2.3. IPET

Interagency Performance Evaluation Task Force (IPET) was introduced by the US Army Crops of Engineers to better evaluate hurricane protection systems. Like PHRaM and R-CLIPER, the IPET model, is based on the 2004 work of lonfat et al. However, instead of using the maximum sustained wind speed, precipitation is predicted by the central pressure deficit [19, 20].

$$p_r(R) = \begin{cases} 1.14 + 0.12\Delta P; \quad R \leq Rv_{\max} \\ (1.14 + 0.12\Delta P)e^{-0.3\left(\frac{R - Rv_{\max}}{Rv_{\max}}\right)}; \quad R > Rv_{\max} \end{cases} \tag{2.10}$$

The precipitation p_r at a certain radius R from the eye is linear with the pressure deficit ΔP when the R is closer to the eye compared to the radius of maximum wind Rv_{max}. Once R is larger compared to Rv_{max}, p_r decreases exponentially. In order to account for asymmetry, IPET additionally multiplies p_r by 1.5 if the coordinates of interest are to the right of the TCs motion, as this is where the higher precipitation levels can be expected.

The IPET model tends to overestimate the rainfall [5]. According to Brackins and Kalyanapu, for engineering purposes, a precipitation model which is biased slightly higher in order to ensure cautious results, however, the IPET model is yet to be suitable to use in flood models [5].

2.2.4. MSR

The Modified Smith for Rainfall (MSR) model creates asymmetric precipitation field that account for the asymmetry by including the storm motion [21, 22]. To validate the MSR data frames from TRMM's precipitation radar (PR) were used unlike the microwave imager (TMI) used in previous studies. According to Langousis and Veneziano the PR product is less bias compared to the TMI, especially closer to the eye where precipitation can be high.

While MSR shows relatively promising results, the model is only valid for open water sites and cannot be applied for land falling hurricanes. It also depends on a high number of variables, as seen in table.2.2, which are not always accessible [4]. Moreover, MSR is not reproducible from literature as it is designed for commercial use [5].

2.2.5. Snaiki & Wu

Snaiki and Wu made further improvements in the field, with a physics-based model that introduces the importance of rain-induced momentum flux at Earth's surface. The model is based on storm location, approach angle, translation speed, radius of maximum wind, pressure profile, surface drag coefficient, and turbulent diffusivity. The model output shows promising results appearing to be consistent with field measurements. Snaiki and Wu, like all other research demonstrate that rain intensity is shown to be highly correlated with the horizontal wind speed. In addition, they highlight that the central pressure difference can have a significant impact on the rain rate [17].

2.2.6. Bader

In a 2019 TU Delft thesis written by Daan Bader, in collaboration with Deltares, an approach to provide the radial rain profile based on sampling a variety of p_{max} values and using an adapted Holland wind fit is proposed. The foundation of the method implies that precipitation is a stochastic process, meaning that it contains a random component, which translates into uncertainty. Therefore, a variety of possible levels of precipitation are sampled for a particular TC parameter.

Table 2.3: Magnitude of dependence between maximum rainfall intensity (pmax) and other TC characteristics based on three different metrics.

Correlation Metric	Parameter	Latitude	Longitude	Storm Motion	Maximum Wind
Kendall's Rank	τ	-0.0207	0.0876	0.0173	0.2630
Spearman's Rank-Order	ρ	-0.0312	0.1317	0.0182	0.3719
Pearson Product Moment	r	-0.0573	0.1333	0.1777	0.4759

In Bader's research the Kendall's Rank, Spearman's Rank-Order and Pearson Product Moment were used as correlation metrics to identify which TC parameter had a sufficient relationship with the maximum precipitation (p_{max}). Longitude, latitude, storm motion and the maximum wind (v_{max}) were tested. Values of the correlation metric can range from –1 to +1, where a value of –1 indicates perfect negative correlation, while a value of +1 indicates perfect positive correlation. A value of 0 indicates no correlation between the two variables. Only v_{max} appeared to have a weak positive relationship and was therefore used to model p_{max} in Bader's approach (see table 2.3).

v_{max} and p_{max} were obtained from the QuikSCAT Tropical Cyclone Radial Structure (QSCAT-R) dataset. This dataset includes data between 1999-2009 for 804 TCs across various basins. v_{max} is the 10-minute averaged wind speed at 10-meter height above surface level in m/s. p_{max} is the highest precipitation measurement from the azimuthally averaged precipitation in mm/hr. For each time step in the data the p_{max} and v_{max} are obtained and are treated as a single independent observation. The resulting 8000 observations have been divided in a training set containing 70% training data and a 30% validation data randomly.

To sample p_{max} a "best-fit" copula was determined based on the training data. The Frank copula was selected to represent the joint cumulative distribution of p_{max} and v_{max},

$$C(u,v) = -\frac{1}{\theta_1} \cdot \ln\left(1 + \frac{\left(e^{-\theta_1 u} - 1\right)\left(e^{-\theta_1 v} - 1\right)}{\left(e^{-\theta_1} - 1\right)}\right) \tag{2.11}$$

where θ_1 is the copula parameter, which in the Bader model is set to 3.58. p_{max} can therefore be sampled based on a value of v_{max} using the conditional distribution.

$$u = -\frac{1}{\theta_1} \cdot \ln\left(1 + \frac{p\left(e^{-\theta_1} - 1\right)}{e^{-\theta_1 v} - p\left(e^{-\theta_1 v} - 1\right)}\right) \tag{2.12}$$

Here, p is a random sample from $U(0,1)$, v is the CDF of v_{max} and u is the CDF of p_{max}. Hence, the inverse CDF can be used to transform the results of a uniform distribution to the corresponding values of v_{max} and p_{max}.

In Baders work, p_{max} has been described by a Generalized Pareto marginal distribution with the following parameters:

- k = -0.0686 (shape parameter)
- σ = 45.2829 (scale parameter)
- θ = 10.002 location parameter)

v_{max} has been described by a Generalized Extreme Value marginal distribution with the following parameters:

- k = 0.346 (shape parameter)
- σ = 8.0676 (scale parameter)
- μ = 17.3637 (location parameter)

Fig.2.6 shows random samples over a range of v_{max} and p_{max} from the Frank copula (eq.2.11), including conditional sampling of p_{max} at v_{max} = 20 and 65 m/s (eq.2.12).

In order to use the sampled values for p_{max} to fit the radial rainfall profile of the TC an adjusted Holland wind profile has been proposed. The Holland wind-profile is designed to predict the wind at radii away from the eye [21].

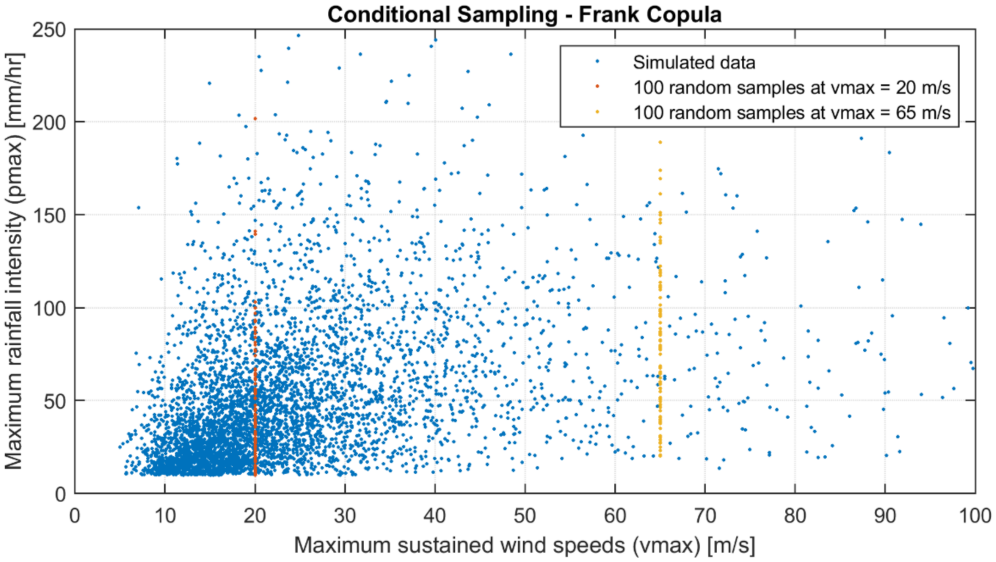


Figure 2.6: 5,000 random realizations of the Frank Copula in blue. The differently colored dots are 100 random samples retrieved with the conditional sampling for a maximum sustained wind speed of 20 m/s and 65 m/s respectively [4]

$$v(R) = \left(\frac{v_{max} * \left(\frac{Rv_{max}}{R}\right)^{b_s}}{\exp\left(\left(\frac{Rv_{max}}{R}\right)^{b_s}\right)} \right)^{x_n} \tag{2.13}$$

Here, v_{max} is the maximum wind speed (m/s),Rv_{max} the radius of where v_{max} is measured, and R is the radius of interest (km). b_s and x_n are fitting coefficients.

The above mentioned eq.2.13 is adapted to predict the radial precipitation as follows:

$$p_r(R) = \left(\frac{p_{max} * \left(\frac{Rp_{max}}{R}\right)^{b_s}}{\exp\left(\left(\frac{Rp_{max}}{R}\right)^{b_s}\right)} \right)^{x_n} \tag{2.14}$$

here, p_{max} is the maximum precipitation sampled from the copula, Rp_{max} the radius of where p_{max} is measured, and R is the radius of interest. b_s and x_n are fitting coefficients that have been fitted according to the data by a least-square fitting procedure as seen in fig.2.7, where the best fit provide eq.2.15-2.16 to compute x_n and b_s based on p_{max}. Lastly, as the actual Rp_{max} cannot be predicted, it is assumed that Rp_{max} = Rv_{max}.

$$x_n = 1.5 * pmax^{-0.031} \tag{2.15}$$

$$b_n = 0.22 * pmax^{0.033} \tag{2.16}$$

An example of a predicted radial rainfall distribution can be observed in fig.2.8. Fig.2.8a. shows the original adjusted Holland fit based on the sampled p_{max}. After analysis, the proposed radial rainfall profile tends to overestimate rainfall rates at larger radii. To prevent this overestimation a modification has been applied. All rainfall rates in the radial rainfall profile larger than 10 mm/hr are set to zero. This limits the radial rainfall profile to smaller radii as seen in fig.2.8b.

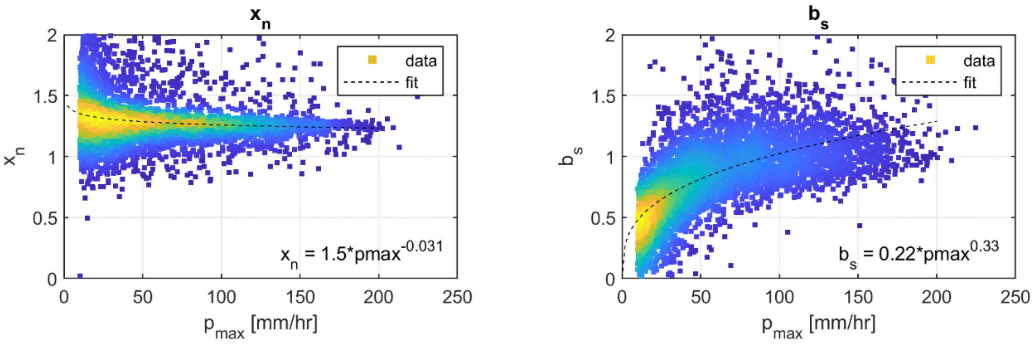


Figure 2.7: Best fit for the fitting coefficient x_n and b_s. The squares are color-coded according to the frequency of the observation, yellow squares indicate more frequent observations. [4]

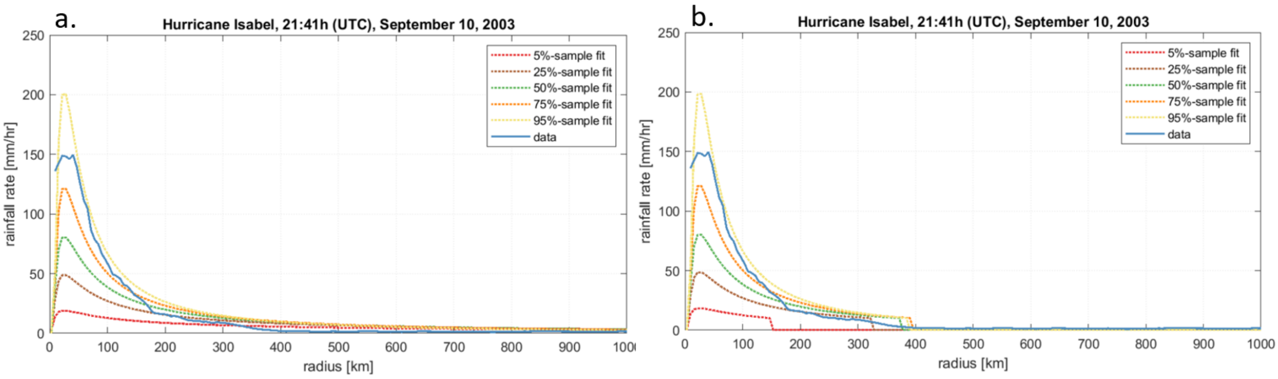


Figure 2.8: An example of rainfall intensity distributed over distance from the hurricane's eye for Hurricane Isabel (2003) at a moment in time. Fits for different sampled values for the maximum rainfall intensity are given.The solid blue line indicates the observed radial rainfall profile from the QSCAT-R dataset. a.is the original fit. b. is the adjusted fit where all rainfall bellow 10 mm/hr are set to 0 to reduce overestimation. [4]

3

Data

This chapter presents the data used for the study. There are two main types of data, namely, precipitation and best track. Several data sources for precipitation were obtained to achieve the highest resolution and the best representation above land and ocean. Table 3.1 shows an overview of the data used. Only data for the North Atlantic has been retrieved.

Table 3.1: Overview of the different data types and sources used in this research.

Product	Product type	Years active	Instrument	Spatial resolution	Temporal resolution	Source
TRMM	2A12:TMI	1997 - 2015	Microwave imager	5.1 x 5.1 km	varies	JAXA/EORC TC online database
TRMM	2A25:PR	1997 - 2015	Precipitation radar	4 x 4 km	varies	Nasa Earth Data online database
GPM	2AGROFGMI	2015 - present	Microwave imager	13 x 13 km	varies	JAXA/EORC TC online database
GPM	2ADPR	1997 - 2015	Precipitation radar	5 x 5 km	varies	Nasa Earth Data online database
STAGE IV	STAGE IV Hourly	2002 - present	Radar, gauge	4 x4 km	hourly	EOL online database
EBTRK	Atlantic dataset	1988-2018	-	-	6-hourly	RAMMB website

3.1. TRMM

The Tropical Rainfall Measuring Mission (TRMM) was a joint observatory by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA). The satellite spent 17 years in orbit from November 1997 till June 2015. This mission was specifically designed to gather tropical precipitation data between 35° north and south latitude [23].

TRMM used active sensors, passive sensors and the TRMM Microwave Imager (TMI) to quantify the precipitation [23].

The 2A12 ”TMI Profiling”, contains surface rainfall as well as vertical hydrometeor profiles on a pixel by pixel basis from the TMI instrument data using the Goddard Profiling algorithm GPROF2010. Surface rain is represented as the liquid portion of precipitation and are in mm/hr. The spatial resolution is approximately 5.1 x 5.1 km, except for the data originated before August 2001, whose resolution is 4.4 x 4.4 km due to a shift in the orbit [24].The 2A12 data has been acquired from the JAXA/EORC TC database [25].

The 2A25 ”PR Profile”, contains estimates of the vertical rainfal rate profile. Surface rain is represented as the liquid portion of precipitation and are in mm/hr. The spatial resolution is 4 x 4 km [24].The 2A25 data has been acquired from the NASA Earth database [26].

3.2. GPM

The Global Precipitation Mission (GPM) is the follow-up mission of TRMM. GPM only has two instruments on board, namely the Dual-frequency Precipitation Radar (DPR) and the GPM Microwave Imager (GMI), a newer version of TMI. The most significant upgrade from TRMM to GPM is its capability for a more global coverage. With a 65° north and south latitude reach it can capture storms that move to middle and high latitudes [23].

Similarly to 2A12, 2AGPROFGMI generates surface rainfall and vertical hydrometeor profiles on a pixel by pixel basis. However, 2AGPROFGMI uses an updated version of GPROF2010, GPROF2014. Surface rain is represented as the liquid portion of precipitation and are also in mm/hr. The spatial resolution of this product is 13 x 13 km [27].The 2AGPROFGMI data has been acquired from the JAXA/EORC TC database as well [25].

The 2ADPR data originates from the GPM on board DPR. Same as 2AGPROFGMI, the surface rain is represented in mm/hr. The resolution is 5 x5 km.The 2ADPR data has been acquired from the NASA Earth database [28].

3.3. Stage IV

Stage IV is a product of the American National Centers for Environmental Prediction (NCEP). Stage IV originates in late 2001, when NCEP started to mosaic from multi-sensor 1h and 6h analyses produced by the 12 contiguous United States (ConUS) River Forecast Centers (RFCs). The regional analyses completed at the RFCs uses an advanced multi-sensor analysis algorithm, followed by manual quality control performed by human analysts [29].

Stage IV has a 4 x 4 km spatial and hourly temporal resolution, merging data from 140 radars and approximately 5500 gauges over the ConUS [30]. The data has been retrieved through the Earth Observing Laboratory (EOL) online database [31].

3.4. Best track data set

The Extended Best Track Dataset (EBTRK) is a product of the Regional and Mesoscale Meteorology Branch (RAMMB), and builds upon the Huricane Database (HURDAT) developed by the National Hurricane Center (NHC). Originally HURDAT contains estimates of the latitude, longitude, 1-minute maximum sustained surface winds, minimum sea-level pressure at 6-hour intervals. EBTRK compliments this with additional information about the maximum radial extent of wind in four quadrants, radius of maximum wind, eye diameter, and pressure as well as radius of the outer closed isobar. The additional information has also been profided by the NHC [32]. The EBTRK has been obtained through the RAMMB website [33].

4

Methods

The methods section is divided into three parts. Section 4.1 explains how the different data sources introduced in chapter 3 are blended. Section 4.2 describes the methods used in order to compute and improve the parametrization of Bader's precipitation model. Finally, section 4.3 presents the case study.

4.1. Blending of satellite and Stage IV data

A predominant issue of the precipitation data introduced in chapter 3 is their coverage. For example, 3.1 shows that the temporal as well as spatial resolution of both TRMM and GPM can vary significantly. This can result in major TCs to only be captured a limited amount of times. Additionally, some time steps only captured part of the TC, as the area that the satellite captures does not always perfectly overlap the area of the TC. Therefore, the difference between two data frames capture the TC can differ several days. To reduce this limitation, both the data from the radar (PR/DPR) and the passive microwave sensor (TMI/GMI) are used. While the radar data is more reliable, the data obtained from passive microwave sensors has a larger spatial coverage. Moreover, TRMM this instrument is limited to the tropics, and both TRMM as well as GPM are not available or become less reliable above land. To overcome these limitations and recompile more reliable and frequent precipitation data over land, the satellite data is blended with the Stage IV product. To the best of the author's knowledge, this appears to be the first time that the value of such a blended satellite-radar-gauge data set for studying rainfall profiles in TCs over land/ocean is investigated.

An overview of the procedure used to merge the satellite data with Stage IV is provided in fig.4.1. In order to combine the different data sources, the data sets have been blended within a 550 km radius of the TCs eye, as this is approximately the radius of a TC. The longitude and latitude of the eye at each time step as well as the maximum sustained wind speed v_{max} and pressure deficit ΔP have been determined using the EBTRK best track data. Since the best track data is 6-hourly, linear interpolation was implemented to estimate the track at the time of interest.

From the longitude (x_{eye}) and latitude (y_{eye}) of the eye, points have been generated every 1/360 degrees (θ_{Di}) in radians (θ_{Ri}) and every 10 km in radius (R) outwards. The longitude and latitude of each point has been calculated as follows:

$$\Theta_{Ri} = \frac{\Theta_{Di}\pi}{180} \tag{4.1}$$

$$Lon_{i,j} = x_{eye} + \frac{1}{\sqrt{2}}(cos(\Theta_{Ri})r_j - sin(\Theta_{Ri})r_j) \tag{4.2}$$

$$Lat_{i,j} = x_{eye} + \frac{1}{\sqrt{2}}(cos(\Theta_{Ri})r_j + sin(\Theta_{Ri})r_j) \tag{4.3}$$

where $Lon_{1,1} = x_{eye}$ and $Lat_{1,1} = y_{eye}$.

For each longitude ($Lon_{i,j}$) and latitude point ($Lat_{i,j}$) the associated precipitation value is calculated for both data types of TRMM/GPM and Stage IV separately by linear interpolation. Next, each point is labelled as above land or above the ocean by testing whether it falls within the coastline polygon provided by MATLAB (coast.mat). following this, the rainfall at each point is selected based on the scheme in fig. 4.1, where $p_{i,j}$ is the precipitation at $Lon_{i,j}$ and $Lat_{i,j}$

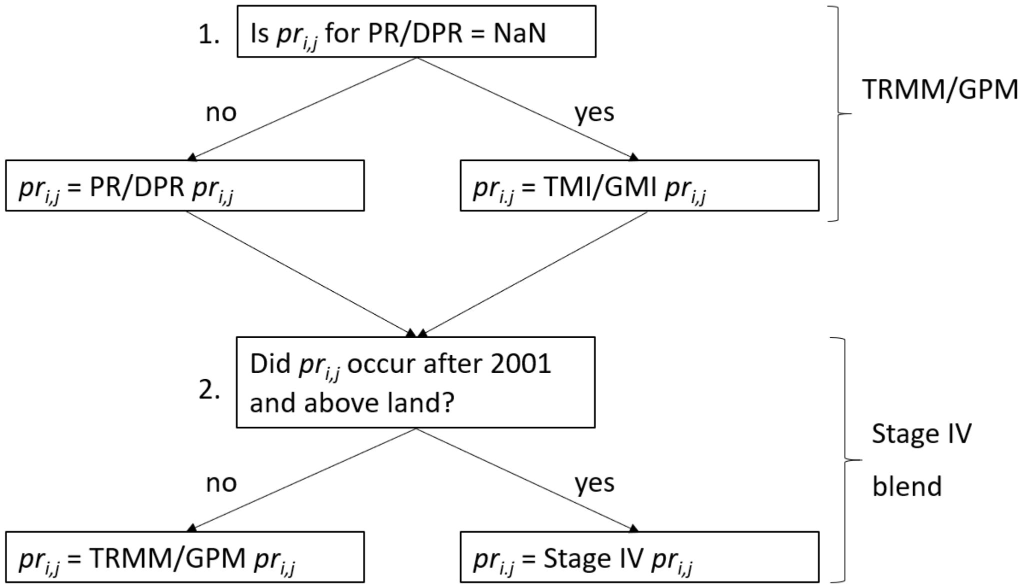


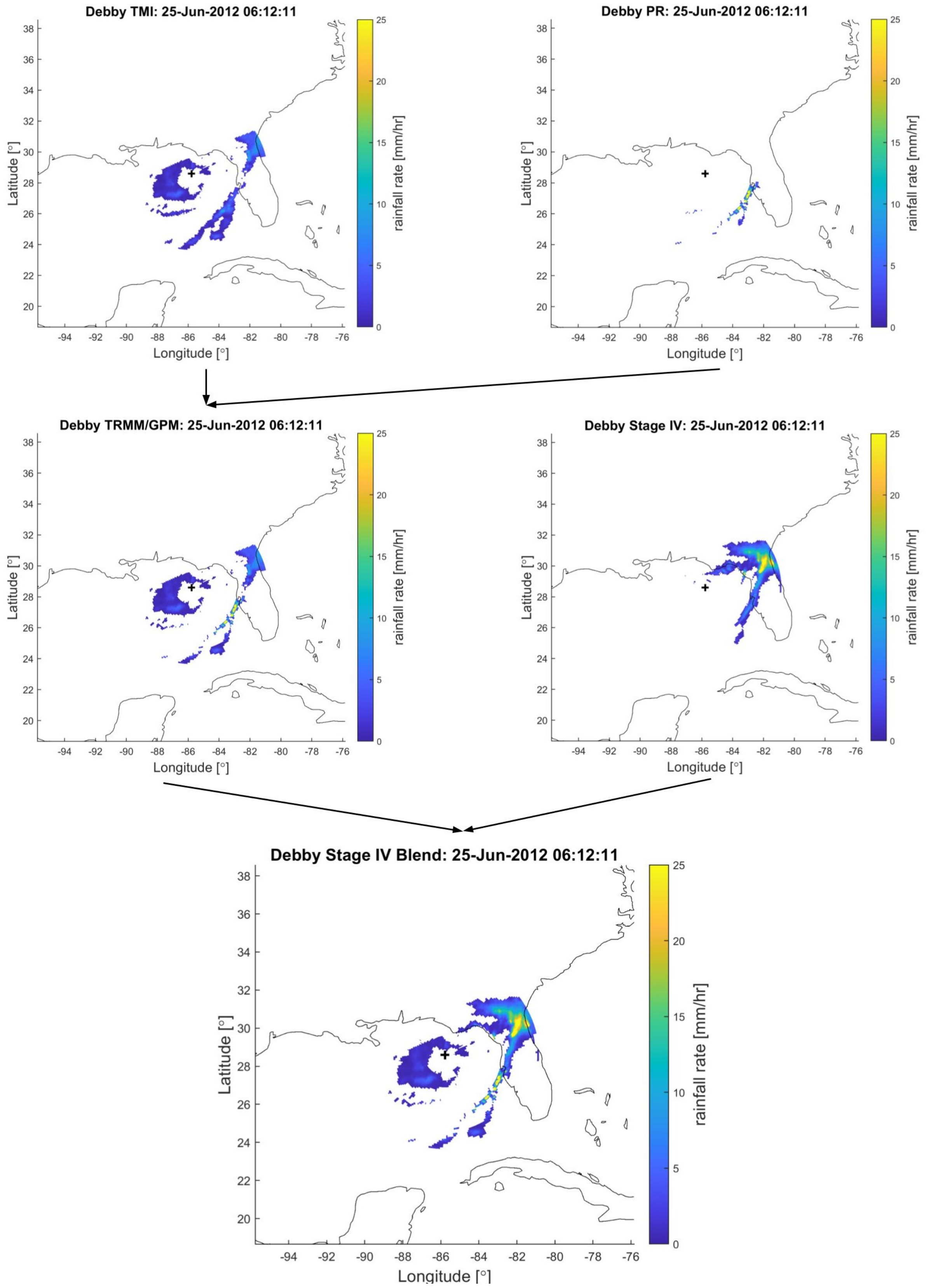
Figure 4.1: Data blending selection scheme.

In step 1 the PR/DPR data is favored over TMI/GMI. If PR/DPR is not available, TMI/GMI is used. The combination of the PR/DPR and TMI/GMI precipitation will be known as the TRMM/GPM data. A second data set is produced by blending the TRMM/GPM data with available Stage IV data. Stage IV data is hourly, therefore, the precipitation is averaged according to the nearest hour. Above land the Stage IV data is favored over the TRMM/GPM. For example, as seen in fig 4.2, a depiction of the separate data sources and the blended result for TC Debby is shown at a time frame corresponding to 06:12:11 AM UTC. Therefore, the resulting Stage IV data frame is 79.7% from the 06:00:00 AM UTC data frame and 20.3% from the 07:00:00 data frame.

4.2. Estimation of radial profiles

The azimuthally averaged rainfall rate is the mean precipitation at different radii. During the blending process, the precipitation is already calculated at 10 km increments away from the eye for every degree, resulting in 360 vectors. These vectors have been averaged, excluding **NaN** values, to obtain the mean rainfall every 10 km. These increments have been considered as a proper resolution for the study, ensuring computational efficiency.

Additionally, for each 10 km increment, the number of non-missing values are noted, as often not



Debby TMI: 25-Jun-2012 06:12:11

Debby PR: 25-Jun-2012 06:12:11

Debby TRMM/GPM: 25-Jun-2012 06:12:11

Debby Stage IV: 25-Jun-2012 06:12:11

Debby Stage IV Blend: 25-Jun-2012 06:12:11

Figure 4.2: Data blending example of Debby showing the individual satellite and Stage IV data as well as the blended products.

the entire TC is captured. To illustrate this consider the following examples: there are 360 for every distance, hence if exactly half of the TC is being captured, this means that 180 out of 360 data points are present at every distance. If on the other hand, 30% of the TC is captured, no points are being captured close to the eye. As this is likely the area where the highest precipitation is expected, it is possible that the true maximum was not captured. Consequently, taking note of the ratio of present and missing data provides an indication of its reliability.

4.3. Parametric precipitation model

The following subsections describes how Bader's model, introduced in section 2.2.6 is adapted (see fig.4.4). Next, the concept of copulas to compute p_{max} are detailed. Furthermore, the method for estimating the coefficients of the adapted Holland profile are explained. This model will be tested using the improved data described in the previous data section. Finally, the different performance metrics are introduced, as well as the method to identify the difference between above land and above ocean precipitation.

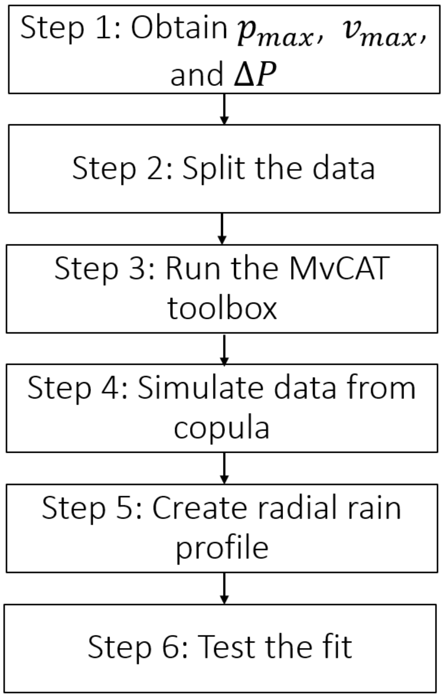


Figure 4.3: Conceptual map showing the steps to compute the parametric model.

4.3.1. Copulas

A Copula is a joint multivariate distribution used to model the dependence structure of two (or more) random variables. Copulas were first proposed in 1959 by Sklar, stating that the joint cumulative distribution function (CDF), $H(x, y)$ of a pair (X, Y) of continuous random variables can be written as:

$$H(x, y) = C\{F(x), G(y)\}, \quad x, y \in \mathbb{R} \tag{4.4}$$

where $F(x)$ and $G(x)$ are marginal distributions and $C\{F(x), G(y)\}$ is the copula. Hence, the selection of an appropriate model for the dependence between X and Y, represented by the copula, can then proceed independently from the choice of the marginal distributions [34].

The Multivariate Copula Analysis Toolbox (MvCAT) is a Copula toolbox developed in MATLAB [35].

The toolbox evaluates 26 copula families to describe the dependence structure of two variables. MvCAT rates the 26 copulas based on performance metrics: Likelihood, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Nash-Sutcliffe Efficiency (NSE), Root Mean Squared Error (RMSE). Likelihood, NSE and RMSE are used to minimize the residuals between observations and model simulations, while AIC weighs model complexity. BIC additionally considers the number of observations as well as complexity [4]. MvCAT additionally evaluates the the Kendall's Rank, Spearman's Rank-Order and Pearson Product that were introduced in section 2.2.6

By using the MvCAT toolbox, the most suitable copula for the GPM/TRMM data set is identified, as well as the Stage IV blended data set, which have been introduced in chapter 3 and section 4.1. Only the copula families included in the *copularnd*, a Matlab toolbox that draws random samples from a copula, have been included for the evaluation. These are the Gaussian, t, Frank, Gumbel, and Clayton Copula families [36]. The input of the toolbox is 70% of the v_{max} and p_{max} data, known as the training data. This 70% is randomly selected. Here, p_{max} is the maximum p_r from the azimuthally radial average. v_{max} is the maximum sustained wind speed obtained from the best track data, and has been converted from knots to metric units. The MvCAT identified the most favourable copula, hence the corresponding variable marginal distributions are used to sample values of p_{max}.

Additionally, the relationship between p_{max} and the pressure deficit (ΔP), which has not been tested by Bader, is tested as well, since it has been indicated as a suitable predictor of p_{max} by various studies, as seen in chapter 2. ΔP is calculated by:

$$\Delta P = P_{outer} - P_{eye} \tag{4.5}$$

where P_{eye} is the pressure at the eye of the TC, and P_{outer} is the pressure at the outer closed isobar of the TC in hPa.

4.3.2. Radial fit

In order to predict the azimuthal average of the TC precipitation, the adapted Holland wind-profile introduced in section 2.2.6 is adopted.

$$pr(R) = \left(\frac{p_{max} * R_p^{b_s}}{\exp\left(R_p^{b_s}\right)} \right)^{x_n} \tag{4.6}$$

Here, $R_p = \frac{Rp_{max}}{R} = \frac{Rv_{max}}{R}$, same as Bader assuming that $Rp_{max} = Rv_{max}$.

The fitting coefficients b_s and x_n have been fitted in four separate methods. These fitting methods have each been named according to their alphabetical bullet, **a**, **b**, **c**, and **d**.

a) Using a least square fitting method for each of the two data sets separately.

b) Using a least square fitting method for p_{max} bellow 5 mm/hr and above 5 mm/hr separately for each data set.

c) Calculate x_n such that the sampled p_{max} is the curves p_{max} by making $x_n = \frac{\log(p_{\max})}{\log\left(\frac{p_{\max}}{\exp(1)}\right)}$ (derivation in Appendix A). b_s is determined using a least square method.

d) Calculate x_n as in c. b_s according to the best fit of the area under the graph where b_s is tested for a variety of samples between 0 and 2 to find the optimal fit.

An overview of the different model versions is provided in fig.4.4

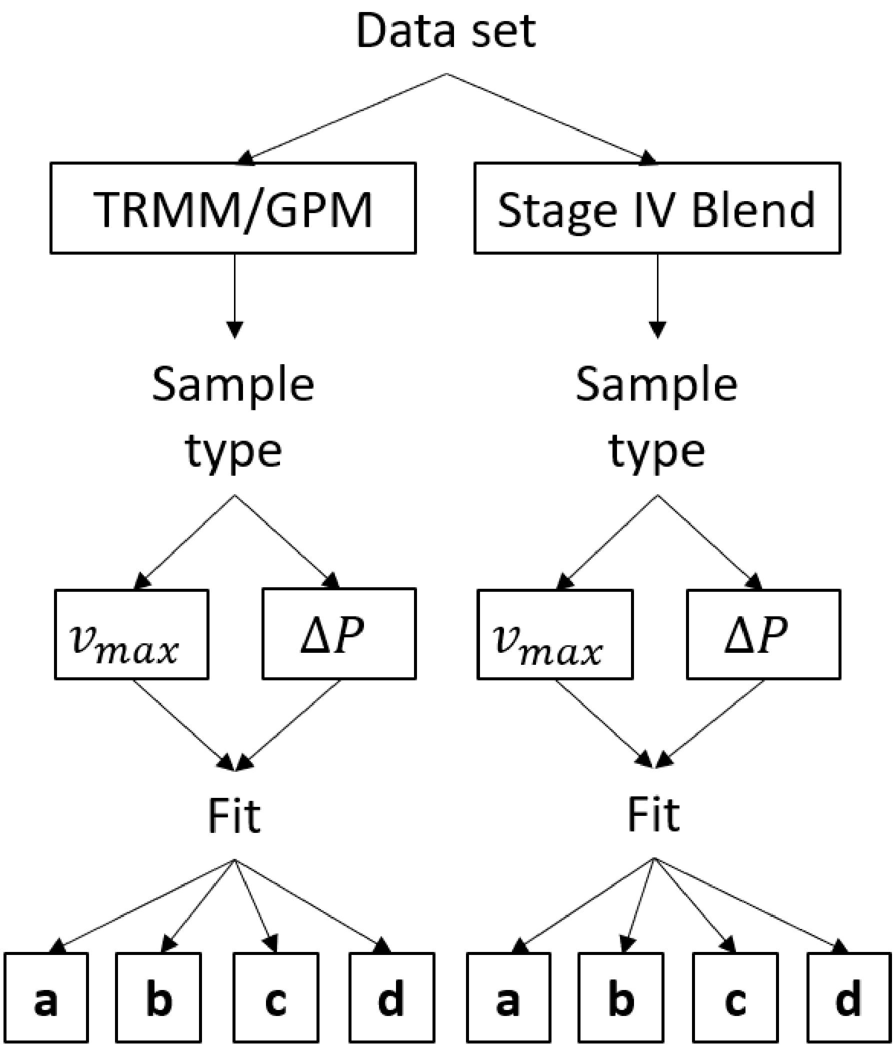


Figure 4.4: Conceptual map representing the different data sets, variables and fits tested for the model.

4.3.3. Performance metric

In order to test the performance of the models that calculate p_{max} based on v_{max} or ΔP, a performance metric is used. The performance metric used are the RMSE, the Mean absolute error (MAE) and the bias.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{p}_{max_i} - p_{max_i}\right)^2}{N}} \tag{4.7}$$

$$MAE = \frac{\sum_{i=1}^{N}\left|\hat{p}_{max_i} - p_{max_i}\right|}{N} \tag{4.8}$$

$$bias = \frac{\sum_{i=1}^{N}\hat{p}_{max_i} - p_{max_i}}{N} \tag{4.9}$$

Where N is the number of samples. $\hat{p}_{max_i}$ is the predicted sample for data frame i and p_{max_i} is the actual value of p_{max} from the data.

The metrics each give an indication how the models perform. Hence, a distinction can be made between the performance based on the TRMM/GPM and the Stage IV blended data set as well as the difference from sampling from v_{max} or ΔP. The 30% of the data that was not in the training data, known as the testing data, is used to compute the metrics.

4.4. Overview of the case study

For the case study, the model is tested to assess its potential to reproduce the precipitation of hurricane Florence 2018. Therefore, all time steps for Florence have been excluded from the training data set, as it is not allowed to influence the model.

First, the radial profiles of the time steps in the TRMM/GPM and Stage IV blend data set are modelled and compared to the data. Following this, an on-land case study area has been defined. Here, 80 time steps of the unblended hourly Stage IV precipitation data are investigated to compare how well the model can reproduce the mean and the cumulative mean within the case area. Finally, the results are compared to the benchmark IPET model introduced in chapter 2.10, as well as the original Bader model.

5

Results & Discussion

The following chapter discusses the results obtained in this research project. In the first section, the general parametric model is defined. Here, the results of the MvCAT toolbox are stated, and the most suitable copulas are highlighted. Next, an explanation of the different fits is provided, which are evaluated based on the performance metrics. The second section is focused on the case study of Florence (2018). First, the model's simulation of the radial profiles in the testing data is assessed. Alongside, the reproduction of precipitation within a specified case area is analysed and compared to both the benchmark IPET and the original Bader model.

5.1. Composition of the parametric model

The model consists of several components that are evaluated in this section. First, the most suitable copulas and marginal distributions are selected based on the Multivariate Copula Analysis Toolbox (Mv-CAT) toolbox. The copulas are used to conditionally sample p_{max} values. Next, the fitting coefficients for the profile have been computed in order to create radial rainfall profiles based on the simulated p_{max} provided by the copula. Once the copulas and profile coefficients have been defined, they are evaluated by the performance metrics. The most appropriate model versions are selected to be used in the following section.

5.1.1. Copula selection & correlation

To identify the best copula, the p_{max}, v_{max} and ΔP are determined for every time-step (fig.5.1). As prior studies have indicated, the highest rainfall is expected to occur within the first 100 km from the eye of the TC (see chapter 2.1.3). Therefore, if the majority of the values within the first 100 km are missing (NaN), the true p_{max} has not been likely captured by the satellite or Stage IV data. This is also apparent as the yellow points in fig.5.1 indicate a low p_{max} at a relatively high v_{max} or ΔP. Therefore, when more than seven values are missing in the first 100 km of the azimuthally averaged rain, these time-steps are removed to improve the data quality. This results in 1935 pairs of p_{max} and $v_{max}/\Delta P$ for the TRMM/GPM data set (See Appendix C.1 for supporting figure). When the Stage IV data is also included, the number of time-steps that has less than seven missing values within the first 100 km is 1965, showing that the stage IV data does cover the satellites blind spots above land, near the eye. As a consequence of this data pre-processing step, there is a stronger linear relationship after the removal of insignificant data.

The MvCAT toolbox, introduced in chapter 4, was used to perform a multivariate dependence analysis of the training data. The toolbox provided three important pillars of information: the correlation

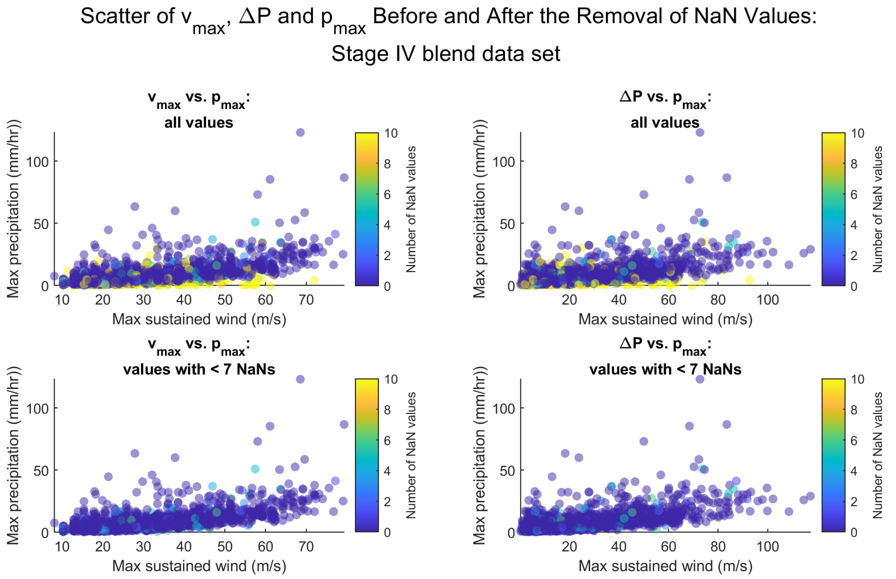


Figure 5.1: Scatter of v_{max}, ΔP and p_{max} before and after the removal of **NaN** values for the Stage IV data set.

Table 5.1: Magnitude of dependence between p_{max} and TC parameters v_{max} and ΔP based on three different metrics for the two data sets

data set	**-**	**TRMM/GPM**		**Stage IV Blend**	
Correlation Metric	**Parameter**	v_{max}	ΔP	v_{max}	ΔP
Kendall’s Rank	τ	0.5054	0.4579	0.5052	0.4815
Spearman’s Rank-Order	ρ	0.6825	0.6315	0.683	0.6587
Pearson Product Moment	r	0.6455	0.6125	0.6393	0.6262

coefficients, marginal distributions, and the most suitable copula (including its theta parameter, θ_1).

Table 5.1 provides the correlation between the parameters based on three different metrics described in section 2.2.6. Both the relationship with v_{max} and ΔP show a higher correlation compared to v_{max} in the QSCAT-R data set, used in the original Bader, as seen in table 2.3. The higher correlation score can be explained by the fact that Bader’s model did not consider the missing values in the radial rain profile, and indicates that the new data set is likely to lead to more precise results.

The marginal distribution of p_{max} lognormal for both data sets, and has a mean μ_l and standard deviation σ_l as indicated in table 5.2. Here, μ_l of the Stage IV blend data set is slightly higher compared to that of the GPM/TRMM data set, while σ_l is slightly lower. However, these slight differences are not significant.

Table 5.2: Lognormal marginal distribution for p_{max} and the mean as well as standard deviation of logarithmic values for each data set

data set	**GPM/TRMM**	**Stage IV**
μ_l	1.6635	1.6730
σ_l	0.8736	0.8676

The v_{max} can be described by an Inverse Gaussian marginal distribution with scale parameter μ_{ig} and shape parameter λ_{ig}. Finally ΔP has a Birnbaum-Saunders marginal distribution with scale parameter β_b and shape parameter γ_b. For both v_{max} and ΔP, the parameters only differ slightly between the data sets due to the additional data points when Stage IV is included as seen in tables 5.3 and 5.4.

Similar to the Bader model, the Frank copula provides the best fit for the majority of the data pairs. This is the case for the relationship between v_{max} and p_{max} in both data sets as well as the ΔP in the

Table 5.3: Inverse Gaussian marginal distribution for v_{max} with scale and shape parameter for each data set

data set	**GPM/TRMM**	**Stage IV**
μ_{ig}	30.1971	30.1223
λ_{ig}	147.8526	146.6439

Table 5.4: Birn-Baumsaunders marginal distribution for ΔP with scale and shape parameter for each data set

data set	**GPM/TRMM**	**Stage IV**
β_b	14.8101	14.8182
γ_b	1.0405	1.0385

TRMM/GPM data set. However, for the Stage IV blend data the Gaussian copula performs marginally better and has been used to simulate p_{max}. The copula types and parameters for each data set and parameter are summarized in table 5.5.

Table 5.5: Copula types and parameters for each data set

data set	**TRMM/GPM**		**Stage IV Blend**	
Parameter	v_{max}	ΔP	v_{max}	ΔP
θ_1	6.7661	4.1927	6.7510	0.6368

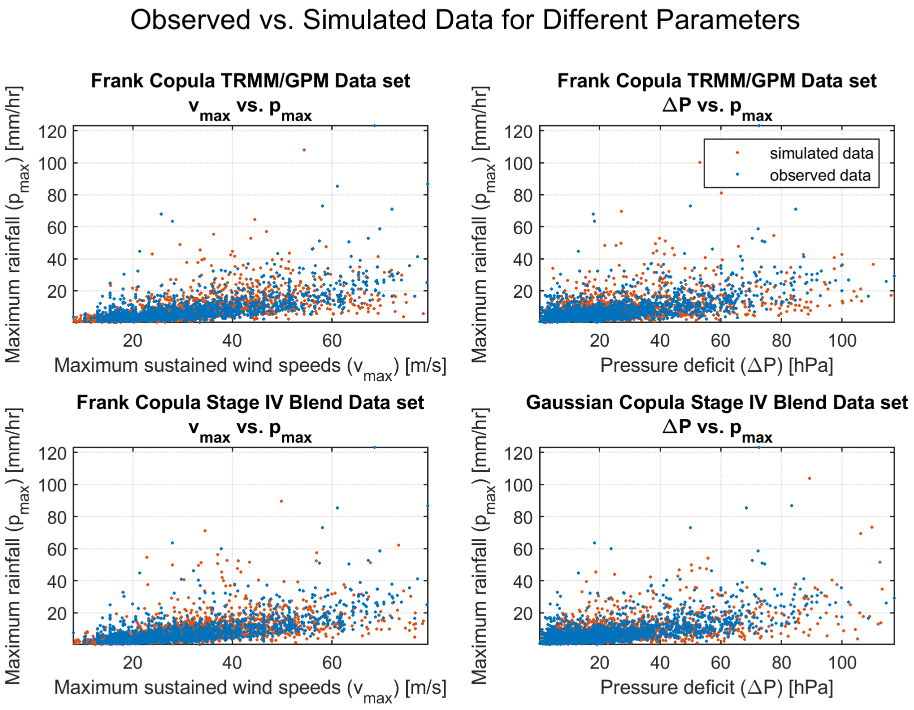


Figure 5.2: Scatter of $v_{max}/\Delta P$ and p_{max} for the different data sets and simulated data generated by the copula.

With the information provided by the MvCAT Toolbox, the most suitable copulas are used to generate random samples of the variables using the *copularnd* toolbox in MATLAB. The Frank copula can be described by eq.2.11 introduced in chapter 2.2.6. The Gaussian Copula is defined as:

$$\int_{-\infty}^{\phi^{-1}(u)}\int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\theta_1^2}} \exp\left(\frac{2\theta xy - x^2 - y^2}{2\left(1-\theta_1^2\right)}\right) dxdy^b \tag{5.1}$$

where v is the CDF of ΔP and u is the CDF of p_{max}.

Fig.5.2 shows the data as well as the simulated samples. Visually, these copulas seems to be able to capture the patterns of the data. To further assess how well the copulas reproduce the distribution in the data set, a quantile-quantile (Q-Q) plot has been generated (see fig.5.3). The Q-Q plot shows that the simulated data resembles the empirical distribution generated by the measurements, implying the suitability of these theoretical models. Nonetheless, these copulas do not fit the extreme values properly due to the fact that these are not well-represented in the training data.

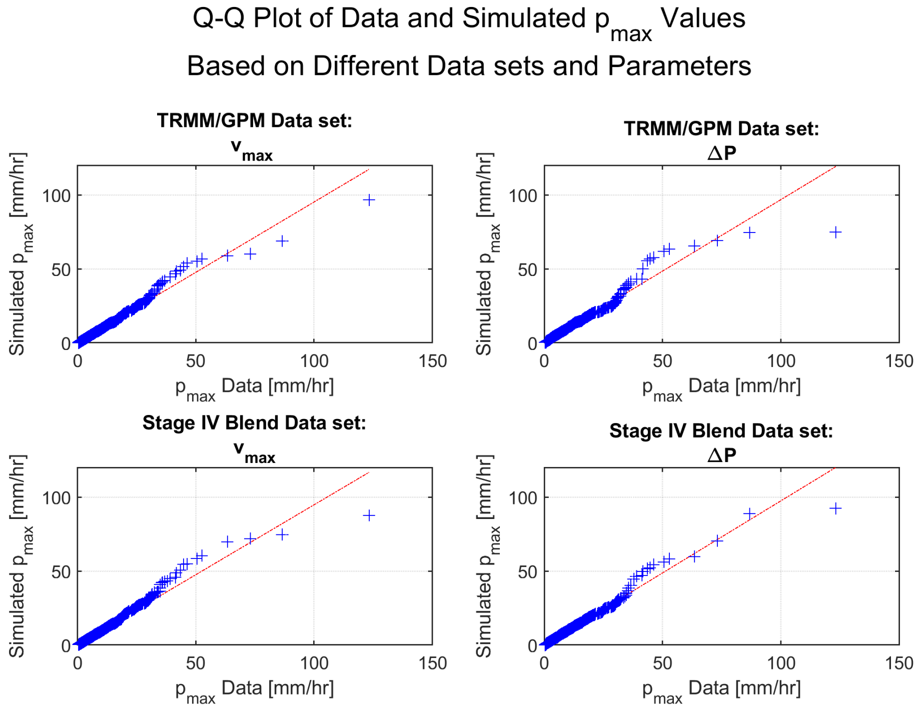


Figure 5.3: Quantile-Quantile plot that compares the distribution of the data with the theoretical distribution of the copulas based on 2000 simulated data points.

Next, to be able to further evaluate the copula fit on the basis of the performance metrics, a conditional sampling method is required. This allows the possibility to predict a range of p_{max} based on a range of v_{max} or ΔP. As explained in chapter 2.2.6, the conditional distribution is given by eq.2.12. Since there is no analytical expression for the Gaussian conditional, hence the values need to be derived numerically (see Appendix B).

The conditional samples generate low random precipitations for low values of $v_{max}/\Delta P$ and a larger range of p_{max} samples for higher $v_{max}/\Delta P$ inputs (fig.5.4). This suggests that the conditional samples provide a reasonably realistic representation of the p_{max} values that can be expected with each condition. However, as there is more training data available for low v_{max} and ΔP values, it become more certain that the conditional sampling technique performs as desired for this range compared to high v_{max} and ΔP data points (for supporting figures, see Appendix C.2 and C.3).

The analyses show that the addition of Stage IV data to TRMM/GPM has very little effect on the marginal distributions of p_{max}, indicating that the addition of Stage IV over land only has little influence on the sampling technique. Table 5.6 summarizes the main performance metrics for the different data sets. This shows that, in contrast, larger differences can be seen between the models based on v_{max} and the ones on ΔP. In particular, sampling from ΔP almost always leads to a larger range of values for p_{max} compared to sampling from v_{max}.

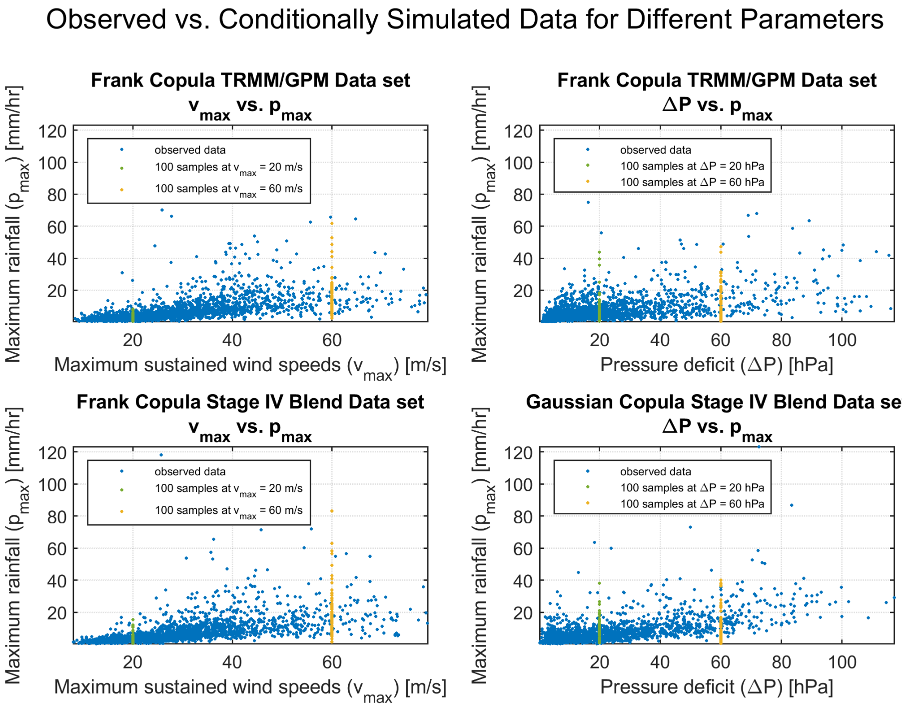


Figure 5.4: Conditional sampling of p_{max} based on v_{max} and ΔP.

Table 5.6: Performance metric of the different simulating methods: observed p_{max} vs. simulated p_{max}

data set & parameter type	**TRMM/GPM v_{max}**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	9.27	7.89	6.98	6.90	26.44
MAE	5.34	3.89	3.48	4.31	21.73
bias	-5.30	-3.33	-1.41	1.42	21.37

data set & parameter type	**Stage IV Blend v_{max} sample**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	9.34	7.94	7.04	6.95	25.79
MAE	5.45	3.99	3.56	4.43	21.38
bias	-5.40	-3.41	-1.49	1.33	21.00

data set & parameter type	**TRMM/GPM ΔP sample**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	9.94	8.51	7.48	7.35	29.93
MAE	5.93	4.37	3.76	4.86	27.51
bias	-5.89	-3.83	-1.59	1.91	27.24

data set & parameter type	**Stage IV Blend ΔP based**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	9.74	8.36	7.21	7.10	25.08
MAE	5.84	4.40	3.67	4.84	21.95
bias	-5.79	-3.89	-1.61	1.98	21.60

5.1.2. Comparison of radial profiles

The aim of the radial fit is to reproduce radial profiles of precipitation based on the radius where the maximum wind occurs (Rv_{max}) and p_{max}. The adapted Holland profile (eq.4.6) is used to simulate precipitation at different radii away from the eye. To define the x_n and b_s coefficients, four methods have been proposed, namely, **a**, **b**, **c**, and **d** (see chapter 4.3.2). In the following subsection each method is analyzed to determine which of them produces the most reliable fit as a function of p_{max}. Here, a final distinction is made between the inclusion and exclusion of Stage IV data, providing further evidence that the influence of Stage IV data is almost negligible.

Fit **a** uses an identical method to the original Bader model. The most favourable x_n and b_s values have been computed separately for each time step. Following this, a line of best fit is introduced in order to estimate the most suitable fitting coefficients n_1 and n_2, as seen in the following equations:

$$x_n = n_1 * p_{\max}{}^{n_2} \tag{5.2}$$

$$b_s = n_1 * p_{\max}{}^{n_2} \tag{5.3}$$

Based on the results of method **a** for the Stage IV data set, the best fit for x_n and b_s appear to be a good fit where there is a high density of low p_{max} values (see fig. 5.5). However, it may not be such an apporpriate fit for larger p_{max} samples where p_{max} > 15, hence, method **b** is introduced.

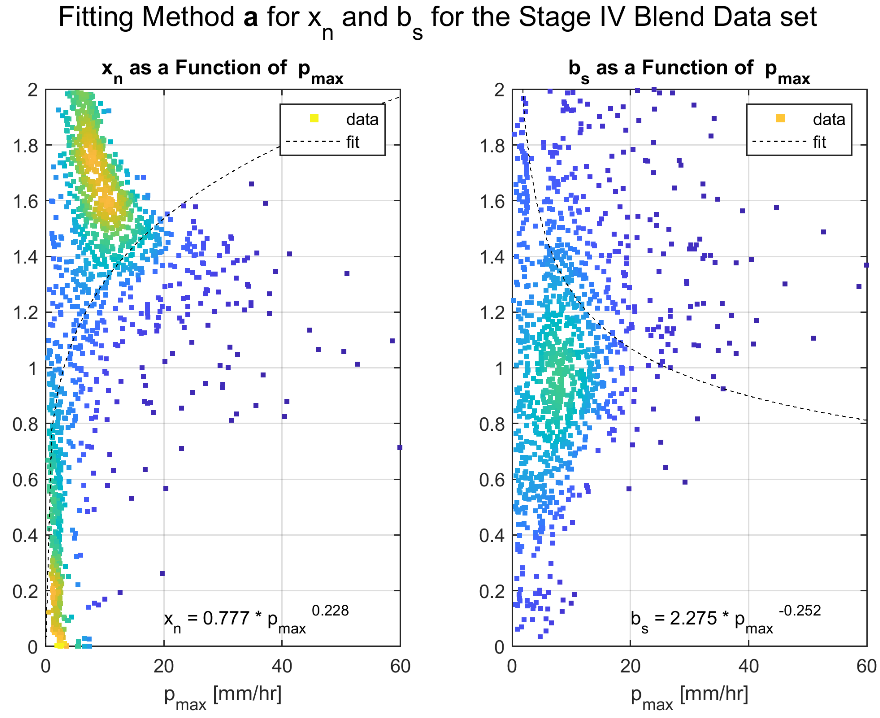


Figure 5.5: Best fit for x_n and b_s for the Stage IV blend data set method **a**. The scatter are colored according to density of the points (yellow is high, purple is low).

Method **b** for fitting the coefficients makes a distinction between time steps with a p_{max} above and below 5 mm/hr (see fig. 5.6). Again, a line of best fit produces the most favorable values for n_1 and n_2 as seen in eq.5.2-5.3. In comparison to **a**, method **b** produces a significantly better fit for the higher p_{max} values. Moreover, fig.5.6 indicates that there is no clear pattern in the values of x_n and b_s for lower p_{max} values, signalling that the Holland profile may not be appropriate under a particular threshold. Another indication that lower p_{max} values cannot be fit using the Holland profile is provided by fit **c** and **d**.

Method **c** for fitting the coefficients calculates x_n according to eq.A.4, ensuring the p_{max} of the fit is identical to the p_{max} of the sample. b_s is fit using a least square method similar to method **a** and **b** (see fig.5.7). While formulating method **c**, it was discovered that it was unsuitable for p_{max} values under 2.8 mm/hr. For the low values the match of x_n and b_s would result in an error, where p_{max} tends to extremely high values. Therefore, for the values below 2.8 mm/hr an alternative radial fit is introduced,

$$p_r(R) = p_{max} \tag{5.4}$$

where the precipitation is the same, low quantity, at all radii.

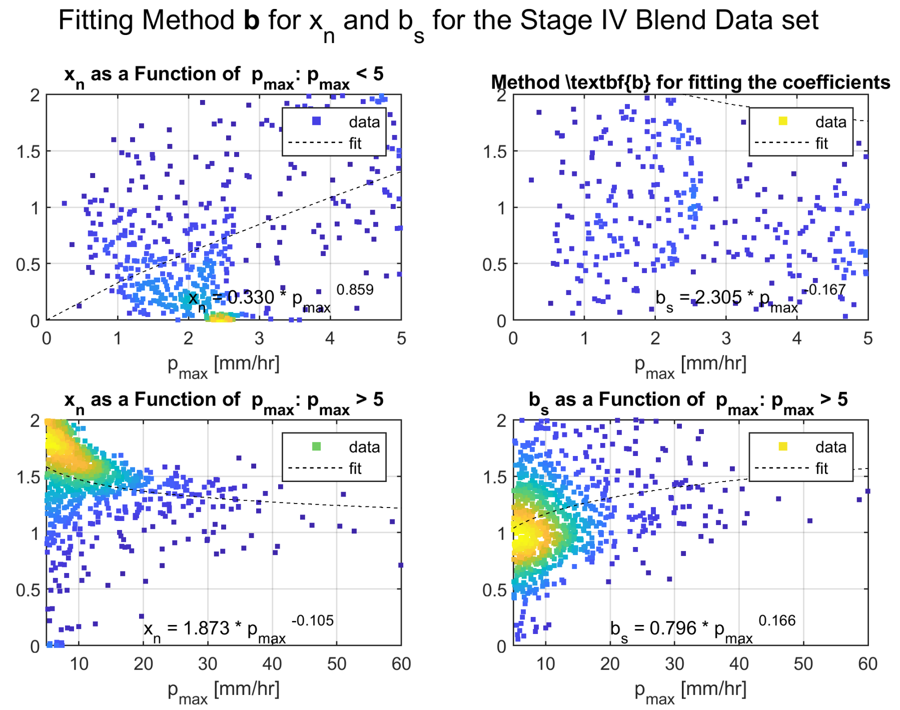


Figure 5.6: Best fit for x_n and b_s for the Stage IV blend data set method **b**. The scatter are colored according to density of the points (yellow is high, purple is low).

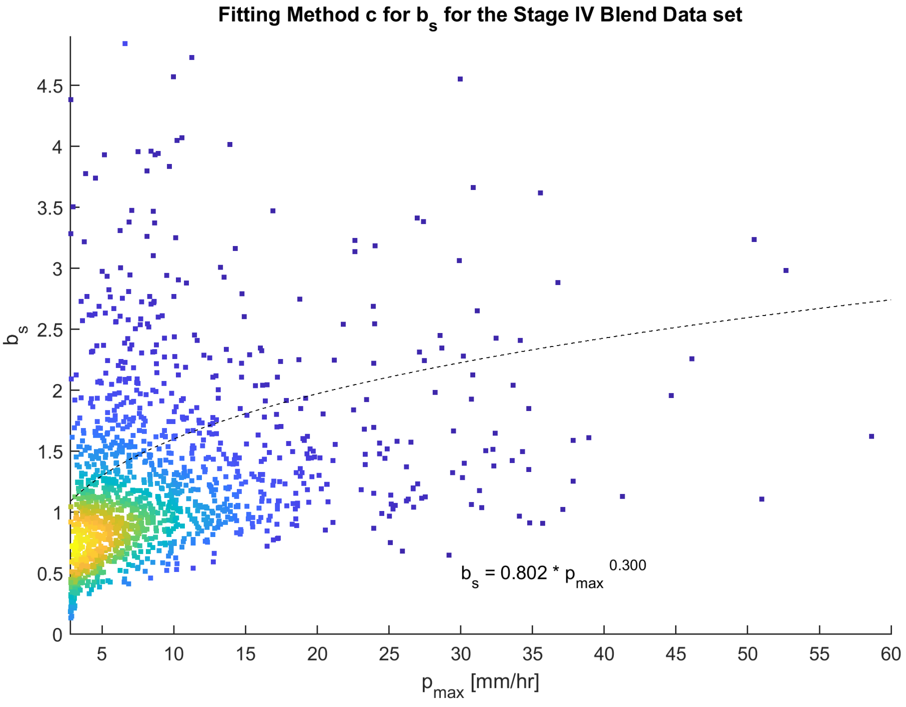


Figure 5.7: Best fit for b_s for the Stage IV blend data set method **c**. The scatter are colored according to density of the points (yellow is high, purple is low).

Method **d** measures x_n identically to **c**, however, it aims to fit b_s by matching it with the area under the curve. While the x_n coefficient determines the p_{max} of the curve, the b_s coefficient governs the slope at which the curve descents after reaching the fit's p_{max}, and hence has a large impact on the area. To define b_s, the area corresponding to each value is plotted (see fig.5.8). Following this, a line of best fit is determined such that the area under the curve can be predicted based on p_{max},

$$Area = n_1 + n_2 log(p_{\max}) + n_3 log(p_{\max})^{\ 2} \tag{5.5}$$

where n_1, n_2 and n_3 are fitting coefficients. b_s is calculated by testing values of b_s between 0 and 2 in order to identify which one best allows the Holland profile to match the calculated area. What can clearly be seen in fig. 5.8 is that the area under the curve becomes more sensitive at higher p_{max} values, creating a larger range of possibilities (heteroscedasticity). Here a smaller area indicates a steep peak of p_{max} and a larger area brings a more gradual decrease. This showcases a major weakness of all four methods, **a**, **b**, **c** and **d**, as a single fit will always fail to capture both extremes.

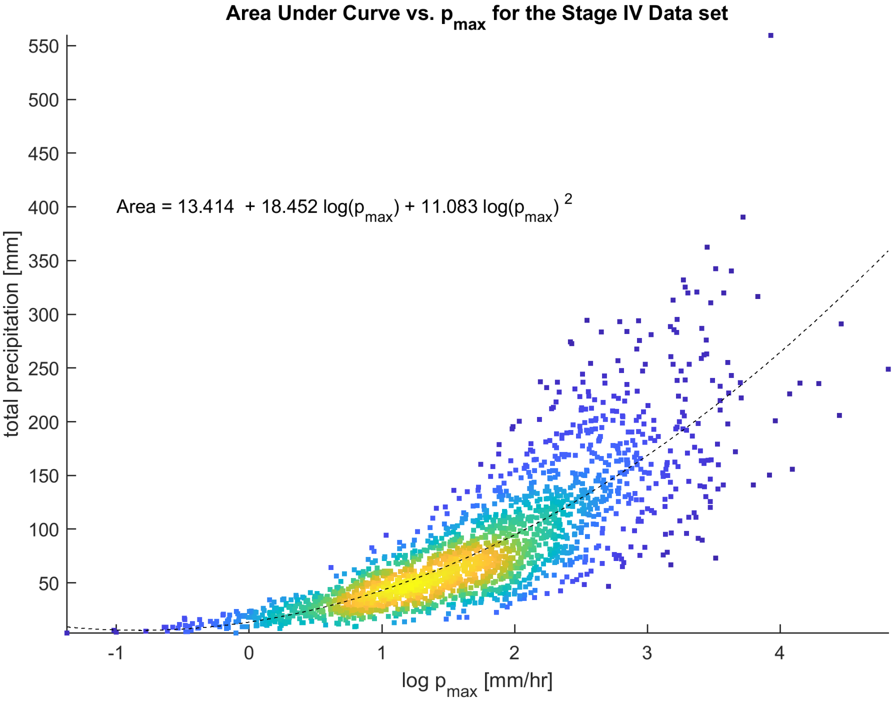


Figure 5.8: Best fit for the area under the graph (total rainfall [mm]) for the Stage IV blend data set method **d**. The scatter are colored according to density of the points (yellow is high, purple is low).

Table 5.7 provides an overview of the different variations of method **a**, **b**, **c** and **d** for both the Stage IV blend as well as the TRMM/GPM data set. As indicated, the corresponding graphs of the TRMM/GPM data set are displayed in Appendix C. Here, it is striking how similar the fits are for both data sets. The similarities indicate that the addition of the Stage IV data set for only a small number of time steps where the TC is above land, has almost negligible impact overall.

To further evaluate the difference between the data set fits as well as the fitting methods, they have been plotted alongside data according to the known p_{max} for random time steps (fig.5.9). Here, it is important that the p_{max} of the fit is close to the true p_{max} and that the area under the curve closely resembles that of the data for varying precipitation strengths.

As speculated previously, fig.5.9 confirms the notion that for low precipitation values the Holland profile is not the most suitable. The precipitation fluctuates greatly, reaching precipitation values close to p_{max} at large radii away from the eye, while the Holland profile always declines after reaching p_{max}. This is also clear when comparing the different methods. For low precipitation, both method **a** and **b** follow the Holland profile, declining after reaching p_{max}. Additionally, both **a** and **b** have a fitted p_{max} that is lower in comparison to the true p_{max}. Fit **c** and **d** on the other hand do have a fitted p_{max} that

Table 5.7: Overview of figures, equations and coefficients of the different fitting methods

Fit	Data set	Restriction	Variable & equation	Figure	n_1	n_2	n_3
a	TRMM/GPM	-	x_n: eq.5.2	fig.C.4	0.788	0.216	-
			b_s: eq.5.3		2.204	-0.231	-
	Stage IV blend		x_n: eq.5.2	fig.5.5	0.777	0.228	-
			b_s: eq.5.3		2.275	-0.252	-
b	TRMM/GPM	p_{max} <5	x_n: eq.5.2	fig.C.5	0.373	0.75	-
			b_s: eq.5.3		2.220	-0.136	-
		p_{max} >5	x_n: eq.5.2		1.842	-0.104	-
			b_s: eq.5.3		0.807	0.165	-
	Stage IV blend	p_{max} <5	x_n: eq.5.2	fig.5.6	0.330	0.859	-
			b_s: eq.5.3		2.305	-0.167	-
		p_{max} <5	x_n: eq.5.2		1.873	-0.105	-
			b_s: eq.5.3		0.796	0.166	-
c	TRMM/GPM	p_{max} >2.8	x_n: eq.A.4	-	-	-	-
			b_s: eq.5.3	fig.C.6	0.771	-0.309	-
		p_{max} <2.8	$p_r(R)$. eq.5.4	-	-	-	-
	Stage IV blend	p_{max} >2.8	x_n: eq.A.4	-	-	-	-
			b_s: eq.5.3	fig.5.7	0.802	-0.300	-
		p_{max} <2.8	$p_r(R)$. eq.5.4	-	-	-	-
d	TRMM/GPM	p_{max} >2.8	x_n: eq.A.4	-	-	-	-
			Area: eq.5.5	fig.C.7	13.208	19.101	9.952
		p_{max} <2.8	$p_r(R)$. eq.5.4	-	-	-	-
	Stage IV blend	p_{max} >2.8	x_n: eq.A.4	-	-	-	-
			Area: eq.5.5	fig.5.8	13.414	18.452	11.083
		p_{max} <2.8	$p_r(R)$. eq.5.4	-	-	-	-

match the true p_{max}, and do not decline, as the precipitation is constant. While it is clear that neither of these fits is suitable to capture the fluctuating precipitation, for such low values it is favorable to use method **c** and **d**, as it will allow p_{max} to be at the true radius of p_{max}, simultaneously covering peaks at later radii which provides a cautious overestimation.

For medium levels of precipitation the Holland fit appears to be more appropriate. Again, method **a** and **b** show a lower fitted p_{max} compared to the data (middle subplots of fig. 5.9). However, they each perform well at large radii, where the precipitation is close to zero. In the original Bader model the overestimation was predominately caused by relatively higher precipitation at larger radii, hence, this is an indication of improvement. Likewise, **c** also performs well at the larger radii and the fit reaches the true p_{max}. On the other hand, method **d** results in high precipitation values at larger radii, leading to a unreasonable overestimation of the total rainfall. The difference in method **c** and **d** is surprisingly less evident when using the Stage IV blend fit.

For high precipitation method **b**, **c** and **d** perform similarly to medium levels of precipitation (lower subplots of fig.5.9). However, method **d** has a significantly higher fitted p_{max} compared to the data. To further highlight this increase in fit, table 5.8 and 5.9 show the performance metrics for the v_{max} sampling of the TRMM/GPM data set using different fitting methods. Here, it is evident that at higher levels of p_{max}, such as that of the 95%, method **a** largely over estimates both p_{max} as well as the area under the graph. At the same time, the performance metrics indicate that **b** results to the largest underestimation of the area under the graph, as indicated by the 5th percentile. Therefore, out of all four methods, **a** is the least suitable.

Method **b** performs relatively well compared to the simulated p_{max} indicated in table 5.6. While it appears to be performing better at the 95%, showcasing a lower deviation with the data, this result must be interpreted with caution. High values of p_{max} that align with the 95% are rare cases, hence, the majority of the testing data will favor a lower fit, causing a smaller deviation. However, as **b** produces a

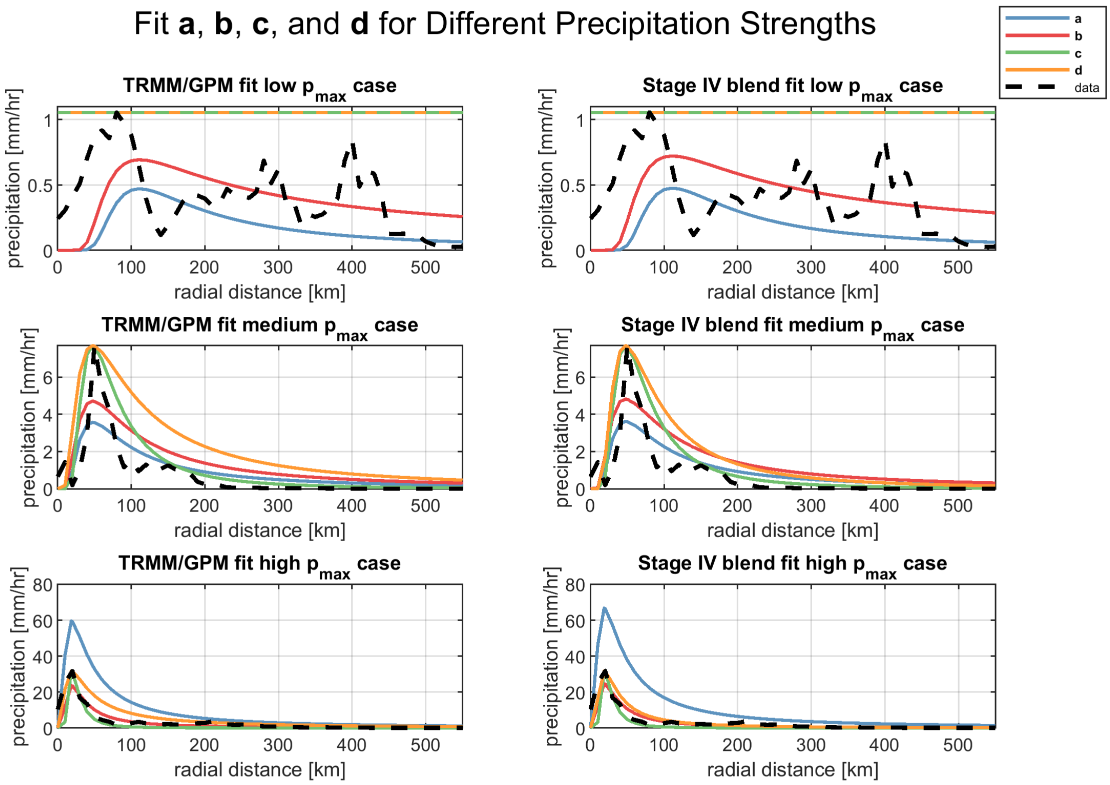


Figure 5.9: The fit from eq. 4.6 using method **a**, **b**, **c** and **d** to determine the coefficients for different precipitation magnitudes. The data is over ocean, hence no Stage IV data is included, however, both the TRMM/GPM fit and the Stage IV fit is displayed

Table 5.8: Performance metric of observed p_{max} vs. the p_{max} of the fit $p_r(Rv_{max})$ for **a** and **b**

data set, parameter type & fit	**Stage IV blend v_{max} based fit a**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	10.62	9.61	8.27	7.10	206.27
MAE	6.89	5.90	4.61	4.09	112.46
bias	-6.89	-5.89	-4.46	-1.13	112.06

data set, parameter type & fit	**Stage IV blend v_{max} based fit b**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	10.43	9.20	8.07	6.98	17.54
MAE	6.68	5.48	4.35	3.72	14.35
bias	-6.68	-5.45	-4.04	-1.76	13.52

lower fitted p_{max} compared to the sampled p_{max}, there is a risk of underestimating the extreme values. On the other hand, method **b** does present a competitive performance for simulating the area under the graph. Nevertheless, **c** and **d** have a p_{max} fit that is identical to the p_{max} sample (for **c** and **d** the observed p_{max} vs. the p_{max} of the fit is identical to the results shown in table 5.6). Consequently, they will not miss the higher p_{max} samples. Moreover, they have good representation of the area under the graph. Therefore, either **c** or **d** is favored.

Comparatively, **d** has a larger tendency to overestimate than **c**, and may therefore be more favourable when there is need for caution. The performance metrics for the sampling from ΔP and the TRMM/GPM set are provided in appendix D, and show similar trends as are observed in table 5.8 and 5.9. Again, the difference between both methods for the Stage IV blend fit are minimal, showing further indication that the additional Stage IV data has very little impact on both the sampling as well as fitting approach. This provides an answer to the first sub-question:

What data set is most suitable to characterize radial rainfall profiles both over the ocean and over land?

Table 5.9: Performance metric for the area under the graph for fit **a**, **b**, **c**, and **d**: v_{max} sampling Stage IV blend dataset

data set, parameter type & fit	**Stage IV blend v_{max} based fit a**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	88.00	71.74	54.00	87.89	3790.24
MAE	67.94	52.53	37.40	55.52	2112.53
bias	-67.87	-51.47	-27.89	27.68	2112.14

data set, parameter type & fit	**Stage IV blend v_{max} based fit b**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	83.82	64.34	53.84	64.13	301.83
MAE	61.62	45.37	37.77	45.65	251.56
bias	-60.63	-42.28	-20.17	15.81	251.03

data set, parameter type & fit	**Stage IV blend v_{max} based fit c**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	79.69	75.58	71.56	64.07	211.23
MAE	58.62	59.85	54.64	46.34	171.40
bias	-21.21	-6.59	3.24	9.25	167.97

data set, parameter type & fit	**Stage IV blend v_{max} based fit d**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	65.45	61.67	59.77	53.29	265.99
MAE	46.84	46.79	44.17	40.30	212.74
bias	-12.95	6.60	17.36	23.99	211.87

The Stage IV data does provide better coverage above land and the TRMM/GPM the best over ocean. However, the Stage IV data in the blend data set is not enough to make a clear difference, indicating that including only TRMM/GPM data would almost be equally effective to predict precipitation above land. As a result, there is no further benefit to make a distinction between the models based on either data set. However, for the case study, v_{max} and ΔP will be sampled from the Stage IV sampling technique as this should have a slight better representation over land. Finally, based on the previous analyses and understanding the advantages and disadvantages of each method, approach **d** is the only fit evaluated in the upcoming case study.

5.2. Case study

To gain a deeper understanding of the model's performance above land and above ocean, the model has been tested by attempting to reproduce the rainfall that fell during hurricane Florence in 2018.

Florence was a relatively large and slow moving TC that originated from a convectively active tropical wave, which was accompanied by a broad low pressure system that moved off the west coast of Africa on 30 August 2018 [37, 38]. It made landfall as a category 1 hurricane on the 14th of September at Wrightsville Beach, North Carolina, with a sustained maximum wind speed of 145 km/hr. After landfall, Florence's winds steadily weakened as it moved inland, however, torrential rain continued to fall for days. The storm produced record-breaking precipitation across North and South Carolina, exceeding the highest single-storm precipitation observed in this part of the country (approximately 10 inches) [37]. As a result, nine river gauges exceeded their 1-in-500 expected floods, and several dams breached, as reported by the USGS [39]. The flooding significantly damaged homes and other infrastructure, resulting in an estimated $16.7 billion in damages. As Florence produced relatively large amount of rain without being a major hurricane, and continues to produce precipitation with decreasing v_{max} as it travels inland, it serves as an appropriate case for exploration.

First the radial fits have been computed for the time-steps included in the TRMM/GPM and Stage

IV blend data sets, in order to identify how well it can reproduce the Satellite data that predominately originates above the ocean. As there is only a single time step where the TRMM/GPM data set varies from the Stage IV blend data set, a further analysis is computed on additional none blended Stage IV data, to evaluate the performance of the model above land in a designated area. Both the model skill of reproducing the mean at each time step as well as the cumulative precipitation are analysed and compared for the v_{max} and ΔP based models.

5.2.1. Reproducing the radial profiles from the testing data sets

All time steps from Florence 2018 were excluded from the training data set on purpose to evaluate the model performance without influencing the coefficients of the model. There are five time steps where the GPM satellite captured Florence's precipitation (see fig. 5.10). Only the 5th time step is close enough to the coast that there is Stage IV data added to Stage IV blend data set.

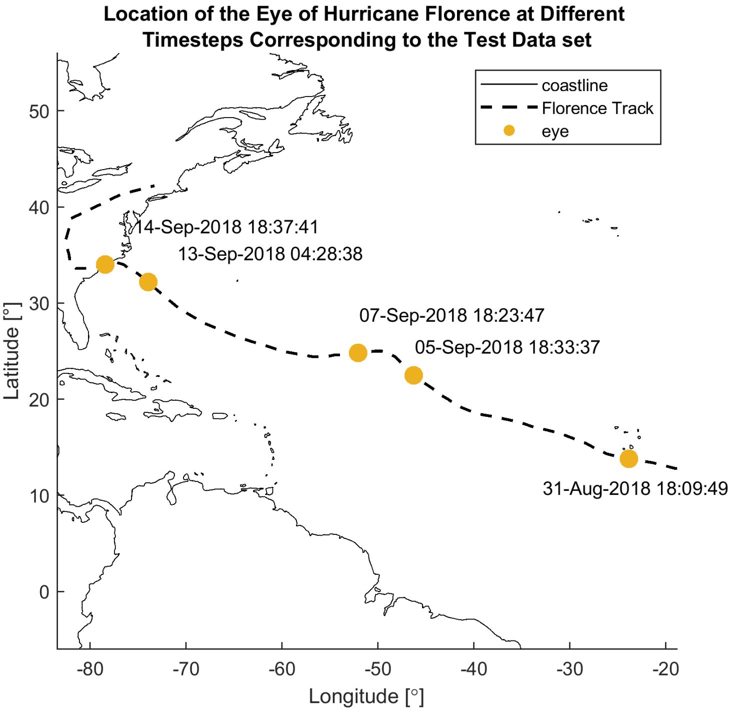


Figure 5.10: Location of the eye and time stamp for the data frames included in both data sets.

At the first time step, Florence is right of the coast of Africa and is yet to be a fully formed TC. However, GPM did measure a large peak in precipitation. The model is unable to capture this properly as there is both a low observed v_{max} and ΔP. Therefore, lower p_{max} values are simulated from the conditional Frank and Gaussian copula. The Gaussian Copula that simulates based on ΔP does have a 95% fit that is closer to the data, compared to the simulations of v_{max}. However, this can likely be attributed to the previously stated hypothesis that the ΔP samples a larger range of p_{max}. This is also evident in fig.5.2, where at low values of ΔP there is a larger possibility of p_{max} compared to low values of v_{max}.

In the second time step, the opposite is observed, where v_{max} simulates higher compared to ΔP, indicating that a larger p_{max} sample for ΔP does not always hold true. A possible explanation is that in this particular time-step, v_{max} is at the high end of all v_{max} values while ΔP is comparatively at a lower end of the ΔP range, as can be observed in fig.5.2.

Time step #3 is captured properly by the model. There is a good match with the observed and sampled p_{max}. After having reached p_{max} the data drops to low precipitation values at a higher rate

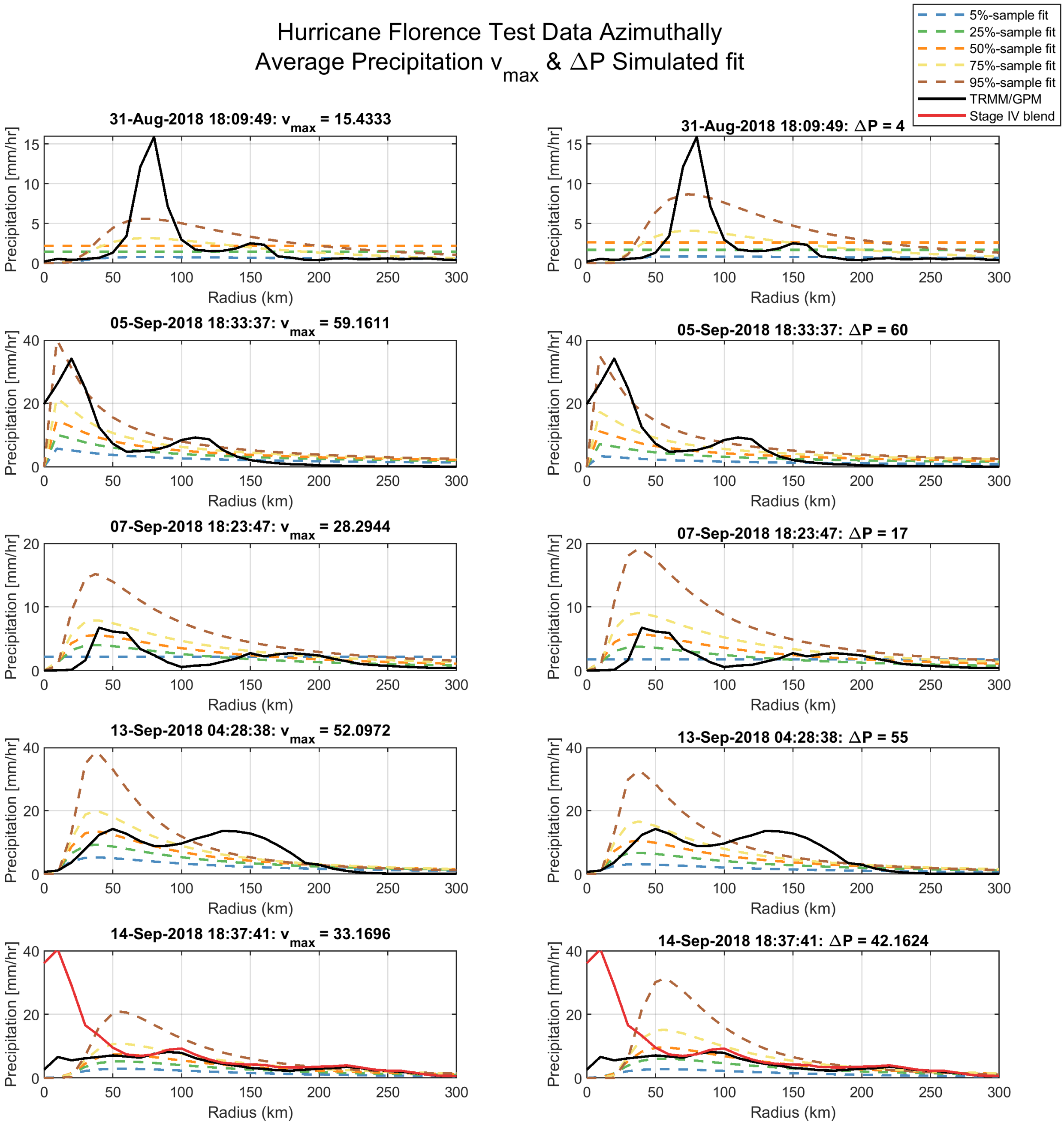


Figure 5.11: Florence testing data from GPM/TRMM and Stage IV blend with simulated radial rain profiles. The date and v_{max} [m/s] as well as ΔP [hPa] are displayed in the sub-titles.

compared to the fit. However, this does not pose a problem as there is another small peak in precipitation around 150 km that a less steep decline of the fit is able to capture.

In time step #4, the data's p_{max} is also well captured by the model, closely resembling the 50% based on v_{max} and the 75% based on ΔP. On the other hand, there is a second peak of equal p_{max} at a larger radii, which the Holland profile is unable to capture. This peak may be due to rain bands, that are not represented in the model. Like time-step two, the v_{max} results in a larger p_{max} samples compared to ΔP, likely for similar reasons.

The final time step stands out for several reasons. First of all, There is a remarkably large difference between the TRMM/GPM and the Stage IV blend data set till a distance of 50 km from the eye. When comparing the two data sets spatially, it is clear that there is a small area near the eye that the GPM satellite failed to capture (see fig.5.12). As the high precipitation in the Stage IV data is close to the eye, it carries a heavy weight in the azimuthally averaged precipitation. Beyond 50 km away from the eye, The pattern in the Stage IV blend and TRMM/GPM data matches well.

While the radius of maximum wind (Rv_{max}) appears to match the true radius of maximum rain (Rp_{max}) for time steps one to four, there is a clear difference in time step five. As can be observed in fig.5.10 and 5.12, the eye of Florence just hit land, which may be the reason why there is high precipitation near the eye. Therefore, while the Holland profile guarantees zero rainfall as the eye, an alternative fit where there is precipitation at the eye could provide a better representation of rainfall over land. To further investigate whether this is a reasonable hypothesis, the next subsection explores the precipitation behaviour and model performance over land.

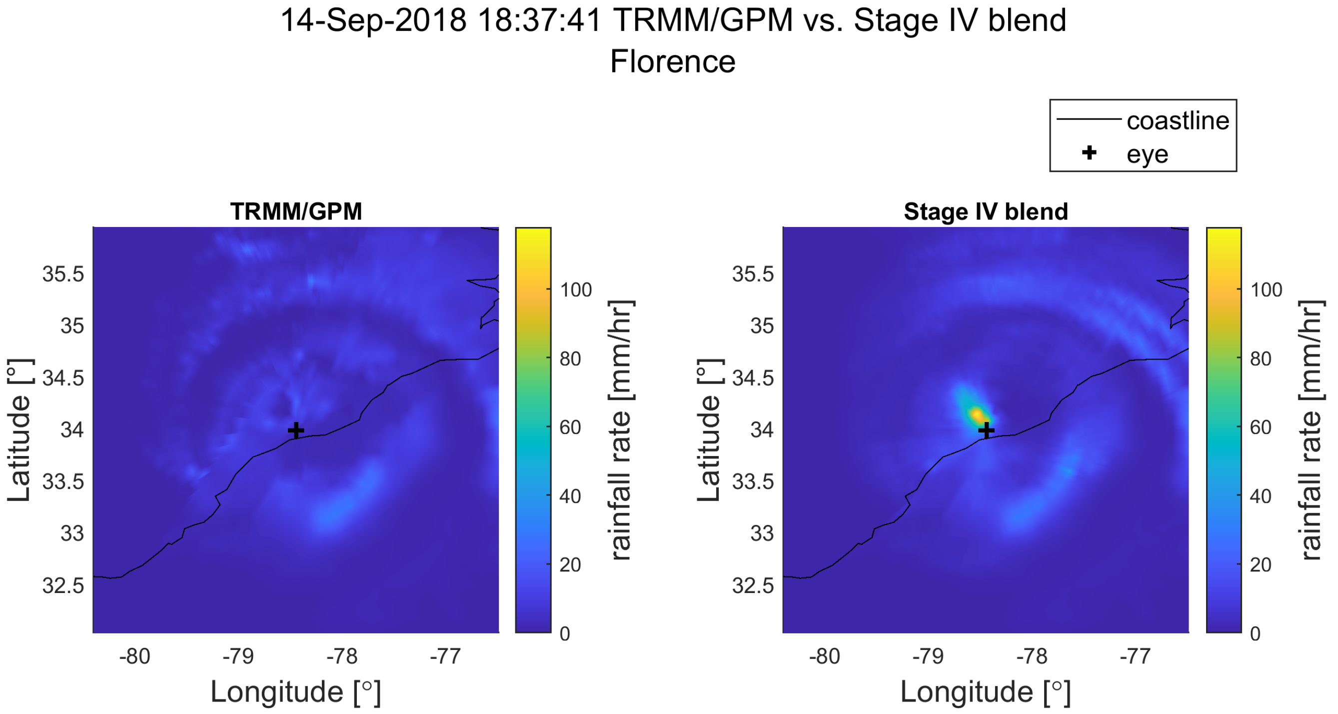


Figure 5.12: Figure showcasing the difference in spatial precipitation between the TRMM/GPM data set and the Stage IV blend data set

5.2.2. Modelled precipitation in case area

As has been introduced in chapter 3, Stage IV provides hourly coverage over the entire US. Therefore, there is a large amount of data that can be used to evaluate the model performance above land.

The case study area displayed in figure 5.13 has been defined based on the three conditions, which are:

1. There is data available within the entire polygon at all time steps to provide the most holistic precipitation overview.
2. There is a slight elevation difference within the polygon, in case orographic enhancement occurs.
3. Area on both the left and right side of Florence has to be present to evaluate possible asymmetry.

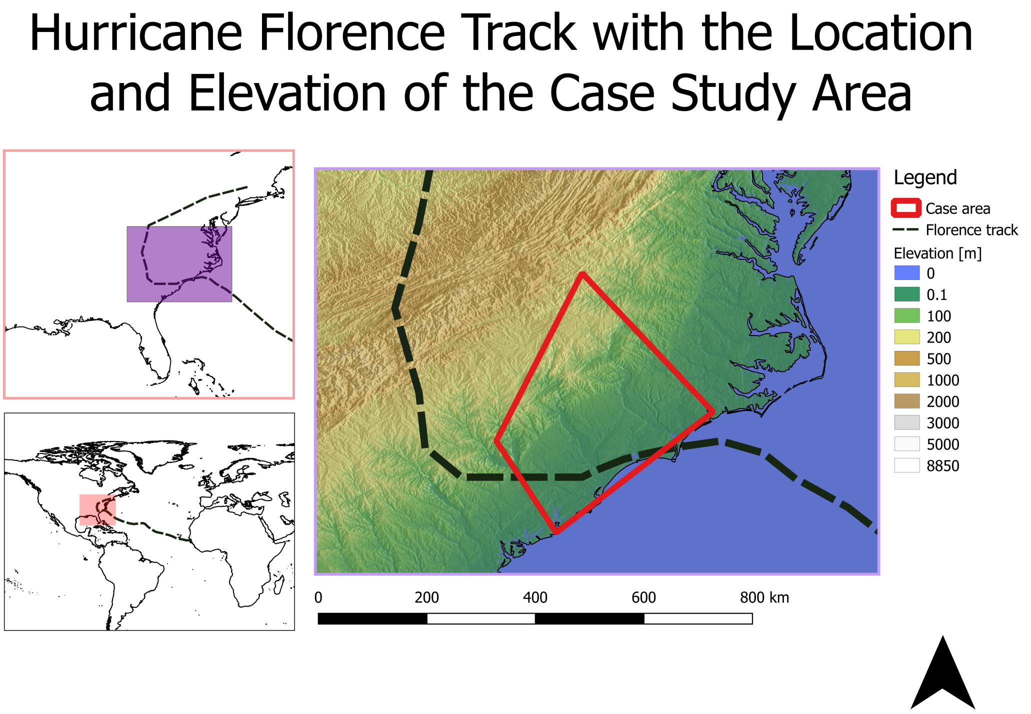


Figure 5.13: Map of the case study area including the track and elevation.

As the track of Florence moves south and turns back north after landfall, condition 2 and 3 are difficult to meet in combination with condition 1. Therefore, the case area extents further inland in the north. The case study area is approximately 94400 km^2.

Within the case study area, precipitation is interpolated to fixed longitude and latitude points for both the observed and modelled data. The mean of the precipitation at these points has been calculated and compared per time step.

Fig.5.15 provides the results obtained by computing the precipitation within the case study area using the v_{max} as a modelling parameter. The time steps 1 to 80 resemble those indicated in fig.5.14. At time step #1, where the eye is still approximately 200 km away from the coast, no precipitation is measured in the case area. The model is able to simulate the low rainfall intensity at this time step relatively well as the high v_{max} at this step results in a high sampled p_{max}. A high p_{max} results in a steep curve, hence the radial precipitation profile will have declined to low rainfall amounts 200 km from the eye. The same notion holds true for the modelled precipitation based on the ΔP (see fig.5.15).

As the eye approached the coast, the mean modelled precipitation increases, even though ΔP and v_{max} decrease. This can be explained by the fact that the location of Rv_{max} (the location of p_{max} in the Holland profile) is in, or closer to, the case study area. The data shows similar behaviour, indicating that peak rainfall can indeed be expected within approximately the first 100 km of the eye. As Florence moves through the case area, the observed precipitation fluctuates. This fluctuation is not captured

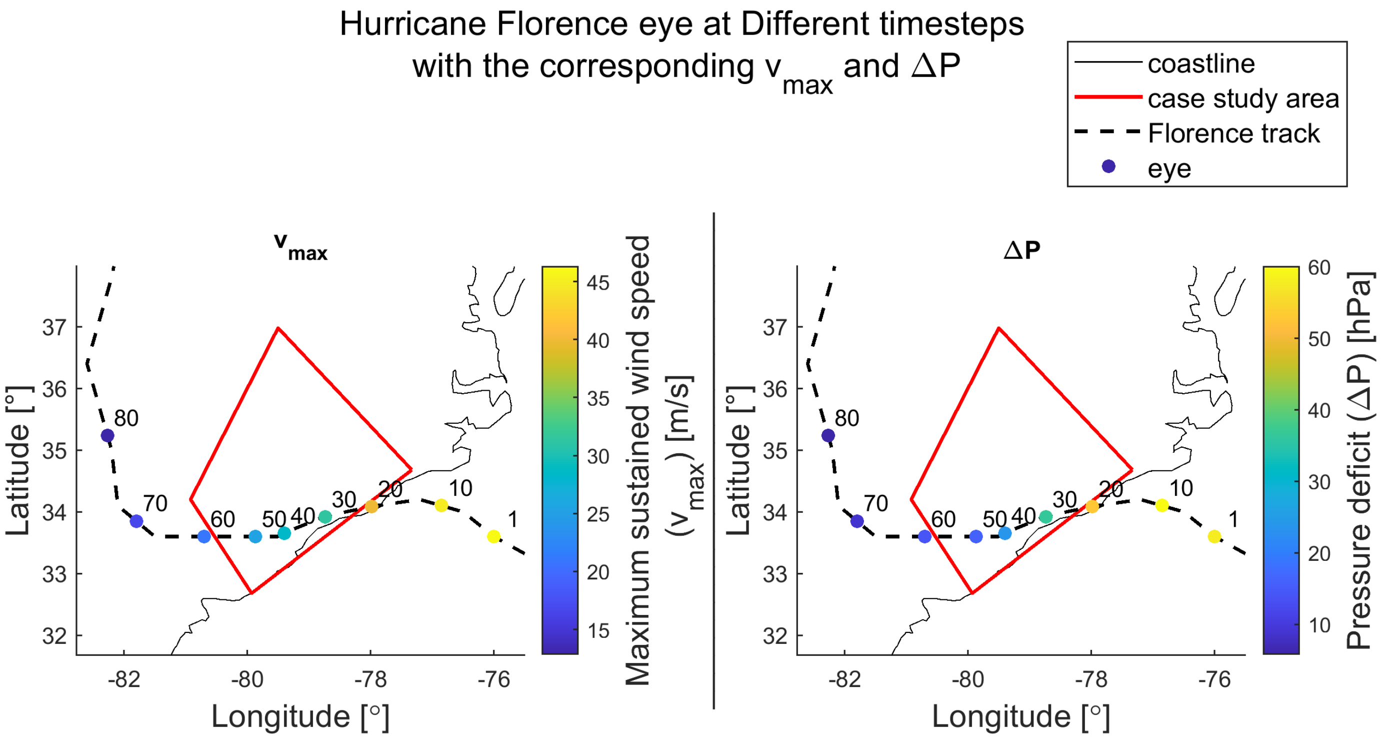


Figure 5.14: The track and time-steps corresponding to the Stage IV hourly data, including the corresponding v_{max} and ΔP.

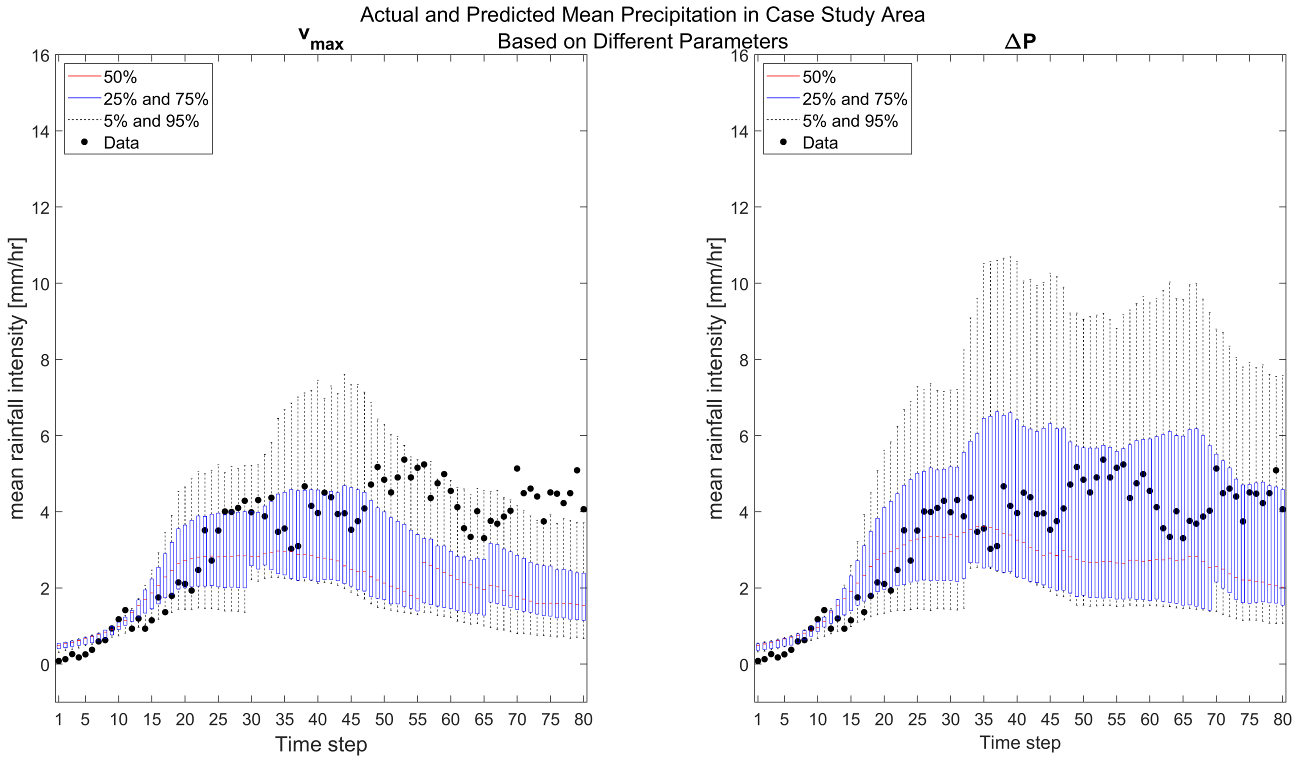


Figure 5.15: Mean precipitation within the case area compared to modelled precipitation based on the v_{max} and ΔP sample.

by the v_{max} modelled rainfall, which declines steadily after time step #45. The modelled rainfall by ΔP does show some fluctuation through time, however, the timing of these fluctuations does not always match those of the data. Around time step #55, a jump in the 50th percentile occur in the v_{max} modelled rainfall, however, this can be attributed to the nature of the Holland profile. When p_{max} is higher, the steep decline produces less precipitation further away from the eye compared to a profile with a low p_{max}. This results in the jump in time step #55, as the 50th percentile of the mean in the area is not always attributed to the 50th percentile of the simulated p_{max}.

In the last ten time steps, the observed data exceeds the 95% of the v_{max}, which is not surprising as the v_{max} at this time is very low and the eye is moving away from the case area. Despite that, the ΔP based model does allow for the last time steps to be captured within the error margins. Yet it should be noted that ΔP evidently results in larger error margin and higher 95% p_{max} sample.

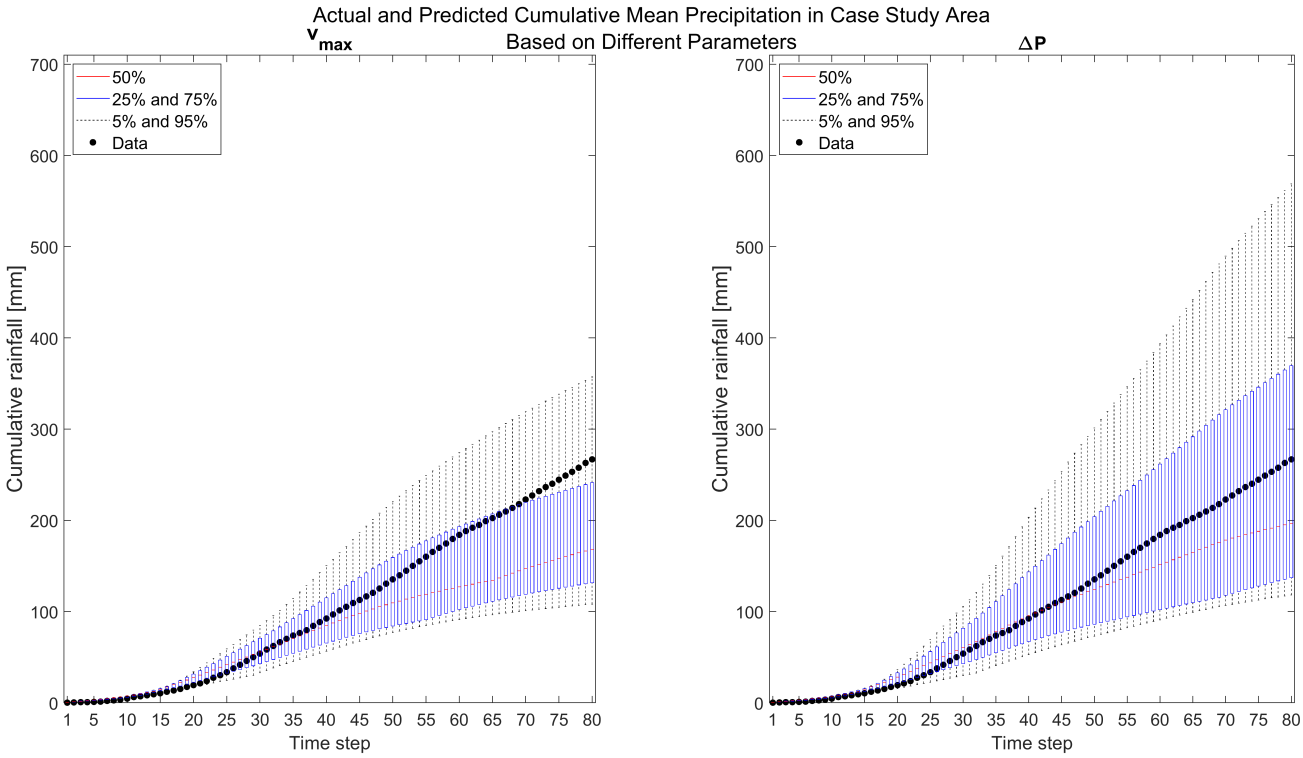
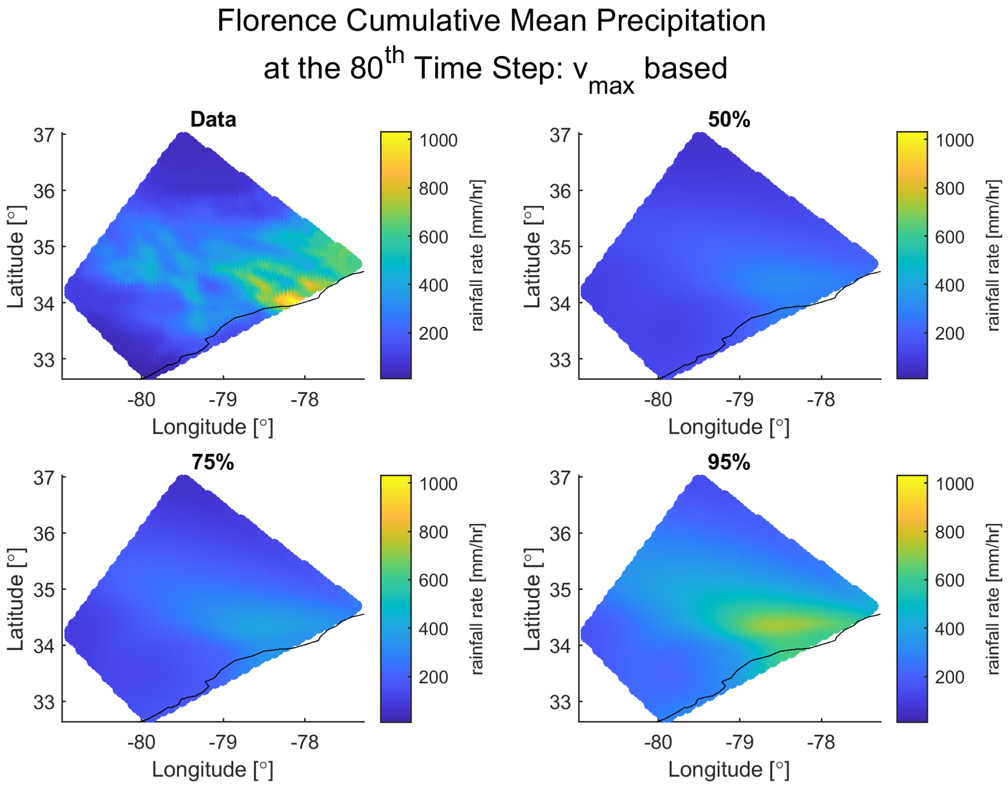


Figure 5.16: Mean cumulative precipitation within the case area compared to modelled precipitation based on the v_{max} and ΔP sample.

Aside from evaluating the predictive skill per time step, the model also has to be able to reproduce the cumulative precipitation. Both the v_{max} and ΔP modelled cumulative mean precipitation show promising results in the first half of the time steps as the 50% closely resembles the data (see fig.5.16 and 5.16). The ΔP model is able to sustain this trend longer compared to the v_{max} model. For both model versions the deviation between the data can likely be attributed to the fact that v_{max} and ΔP enter the lower spectrum in the second half of the time steps, hence will result in lower p_{max} samples. Unsurprisingly the v_{max} model again has smaller error margins, however, the 50% of the ΔP model more accurately resembles the cumulative mean precipitation at the 80th time-step.

While the cumulative precipitation within the case area can be reproduced well, this gives no indication on how well the spatial variation is captured. Therefore, the cumulative mean precipitation at the final time step has been plotted (see fig.5.17 and 5.18). It is clear that the ΔP model has higher precipitation for 75% and 95%, and deviates less from the v_{max} model for the 50%. Both the v_{max} model and ΔP model show similar spatial patterns, which can be attributed to the fact that they both use the same radial profile fitting method. This suggests that the fit of the profile is of more importance


Florence Cumulative Mean Precipitation at the 80th Time Step: v_{max} based

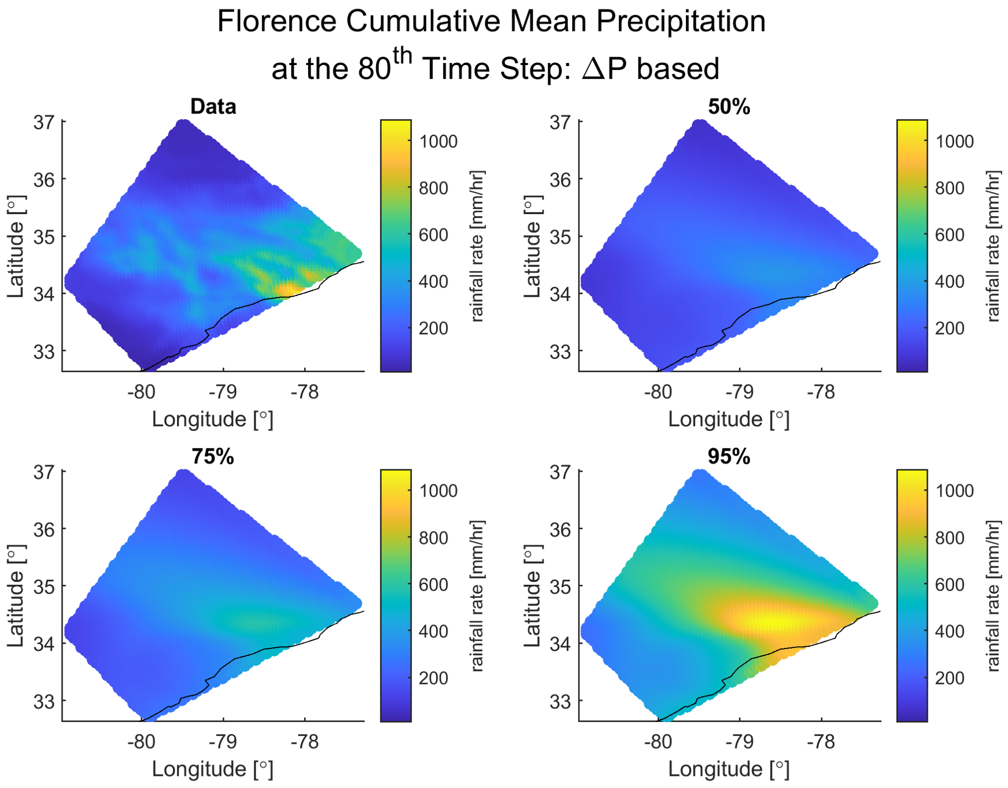
Data

50%

75%

95%

Figure 5.17: Cumulative mean rain at final time-step for the data and v_{max} modelled 50%, 70% and 95%.


Florence Cumulative Mean Precipitation at the 80th Time Step: ΔP based

Data

50%

75%

95%

Figure 5.18: Cumulative mean rain at final time-step for the data and ΔP modelled 50%, 70% and 95%.

to replicate the spatial rainfall, rather than the sampling method. The general precipitation gradient from north to south replicates that in the observed data. However, at different locations a different percentile may provide a better indication of the magnitude that can be expected. For example, when observing the ΔP model, the 50% appears to be the most appropriate for the far north and south of the case area, the 75% is the best for the north-west, while the 95% provides the peak observed to the south-east. It is interesting that the topography change in the north does not seem to have an impact on the precipitation pattern in this case. The location of the peak in the data is located where the eye hits land, resembling the peak observed in fig.5.11. Hence, this result reflects the hypothesis that zero precipitation at the eye over land may not hold true. A model which recognizes this is the IPET model which has been introduced in chapter 2.2.3. In the IPET model the p_{max} level rain occur from the eye until the radius of maximum wind. The IPET is compared to the ΔP based model in the next section, to understand whether it provides a better prediction.

5.2.3. IPET comparison

The IPET model is a highly simplified rainfall model. It is unique as it assumes high precipitation at the eye, unlike many other parametric models. IPET demonstrated the most skill in the 2020 study by Brackins and Kalyanapu [5]. Therefore, it is an interesting candidate compared to the ΔP based model, as IPET predicts precipitation based on ΔP as well.

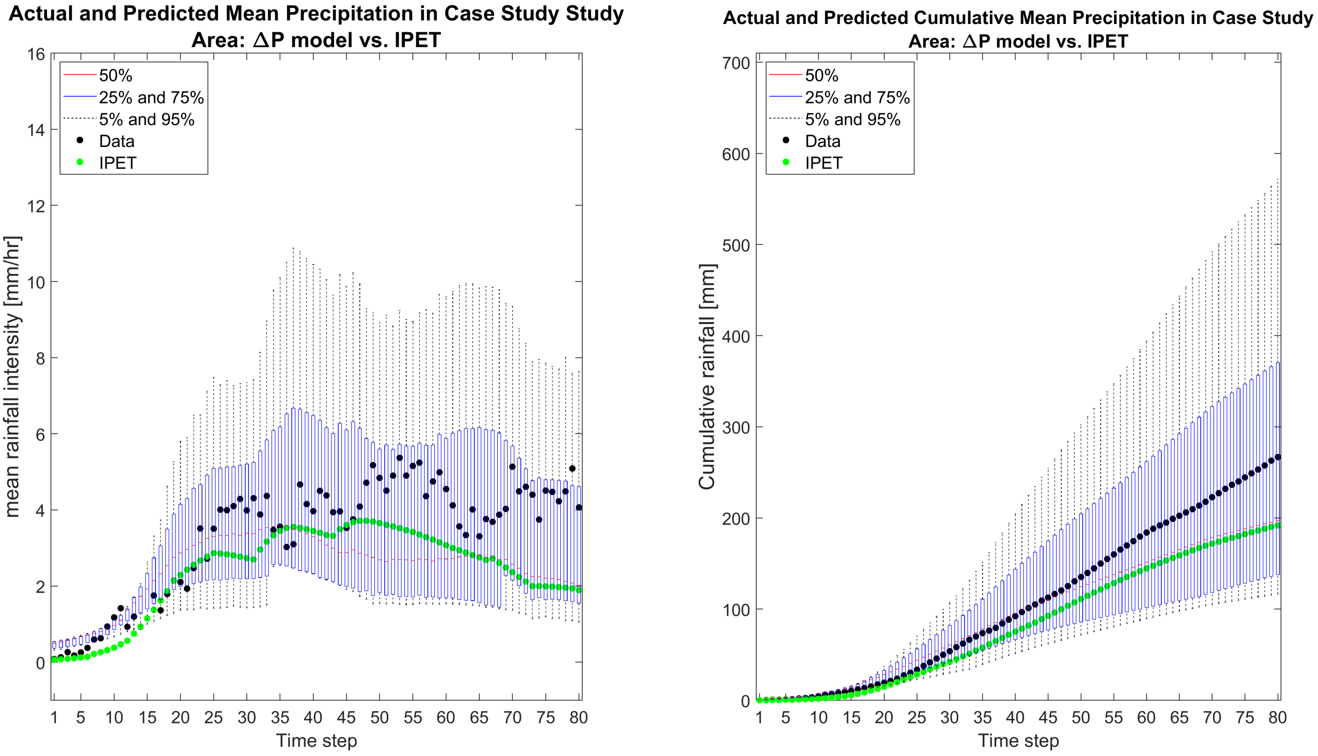


Figure 5.19: Comparison of the mean and cumulative mean precipitation within the case area computed by the IPET model and the ΔP land profile model.

In the first time steps, IPET has a lower precipitation compared to the ΔP model and the data (fig.5.19). This can be explained by the fact that IPET declines to zero faster after Rv_{max}. While the eye is still far away from the case area, relevant effects can still occur. As the eye moves through the case study area, the IPET model underestimates the precipitation. However, between time step #40 and #60, it performs better compared to the 50% of the ΔP model. Here, it slightly resembles the observed peak. Therefore, the model could be benefiting from the fit that does not assume zero precipitation at the eye. This is also visible in mean cumulative precipitation, where between time step #40 and #60 the cumulative IPET precipitation increases at a steep rate, while the ΔP modelled 50% precipitation trend starts to deviate more from the data, decreasing in slope, insinuating that at this point the ΔP model can also benefit from an IPET-like fit. Additionally, while the IPET model does underestimate the cumulative precipitation, it does spatially replicate the fact that a relatively high precipitation is measured where the eye hits the coast (see Appendix E.1). Unfortunately, IPET does fail to provide error margins, which is a major drawback of the model. Therefore, inspired by the IPET, a new land profile is proposed in the next section. The land fit aims to replicate the fit of IPET, with the uncertainty range of a copula based model.

5.2.4. Testing the land profile

The land profile models a radial precipitation of p_{max} from the eye until Rv_{max}, after which it follows the Holland profile descent (eq.4.6).

$$p_r(R) = p_{\max} \qquad R < Rv_{max} \tag{5.6}$$

When this fit is applied to the v_{max} based model, the general pattern of simulated precipitation appears to be similar to when the normal profile is used (see fig.5.20). However, the magnitude of the 50% and above precipitation is significantly higher. The 95% reaches similar precipitation values as the ΔP based model with a normal fit. However, the 50% precipitation for the land profile aligns better with the data. This is also the case for the ΔP model with the land fit (see fig.5.20). Here, the ΔP model with the land fit is able to replicate the mean precipitation skillfully until approximately time step #65. However, the land profile does not capture the high rainfall in the last time steps. As the eye has already moved out of the case study area, and the land fit only differs until the radius of maximum wind, there is no influence of this fit at such time steps.

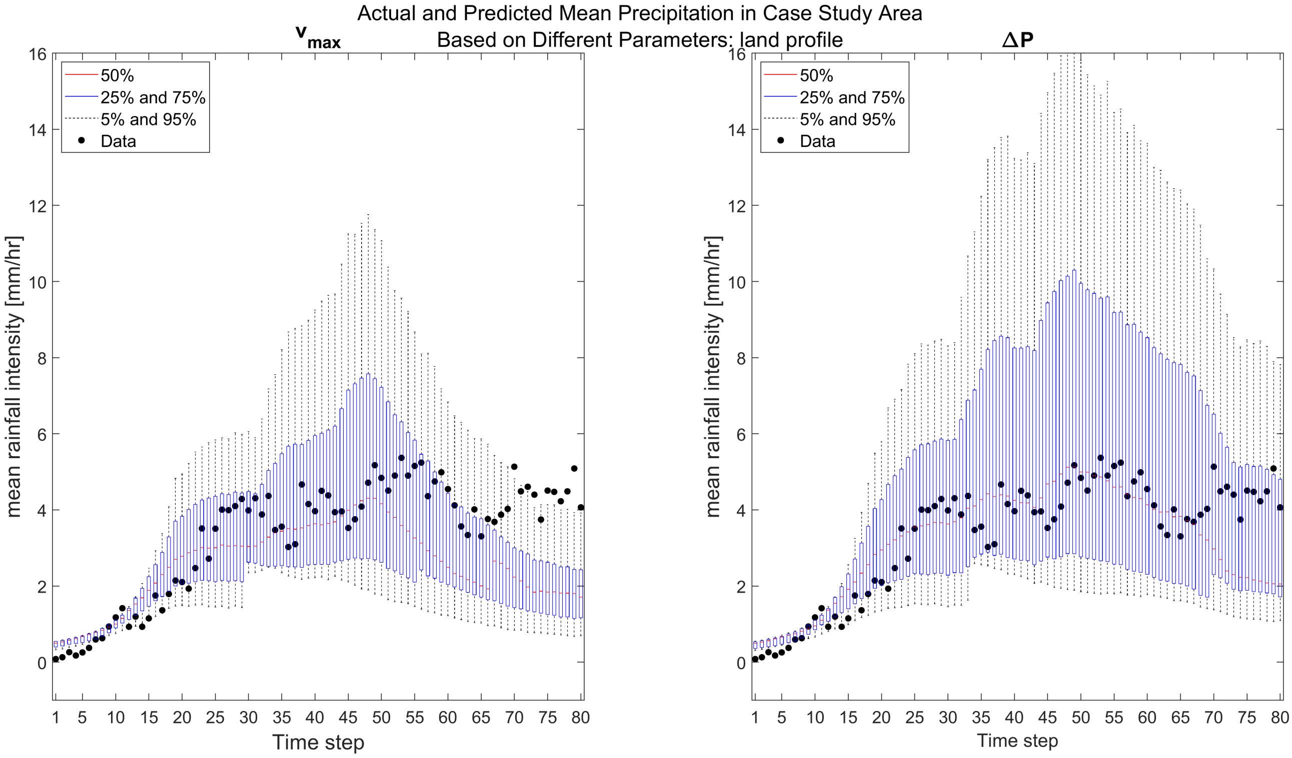


Figure 5.20: Mean precipitation within the case area compared to modelled precipitation based on the v_{max} and ΔP sample and land profile.

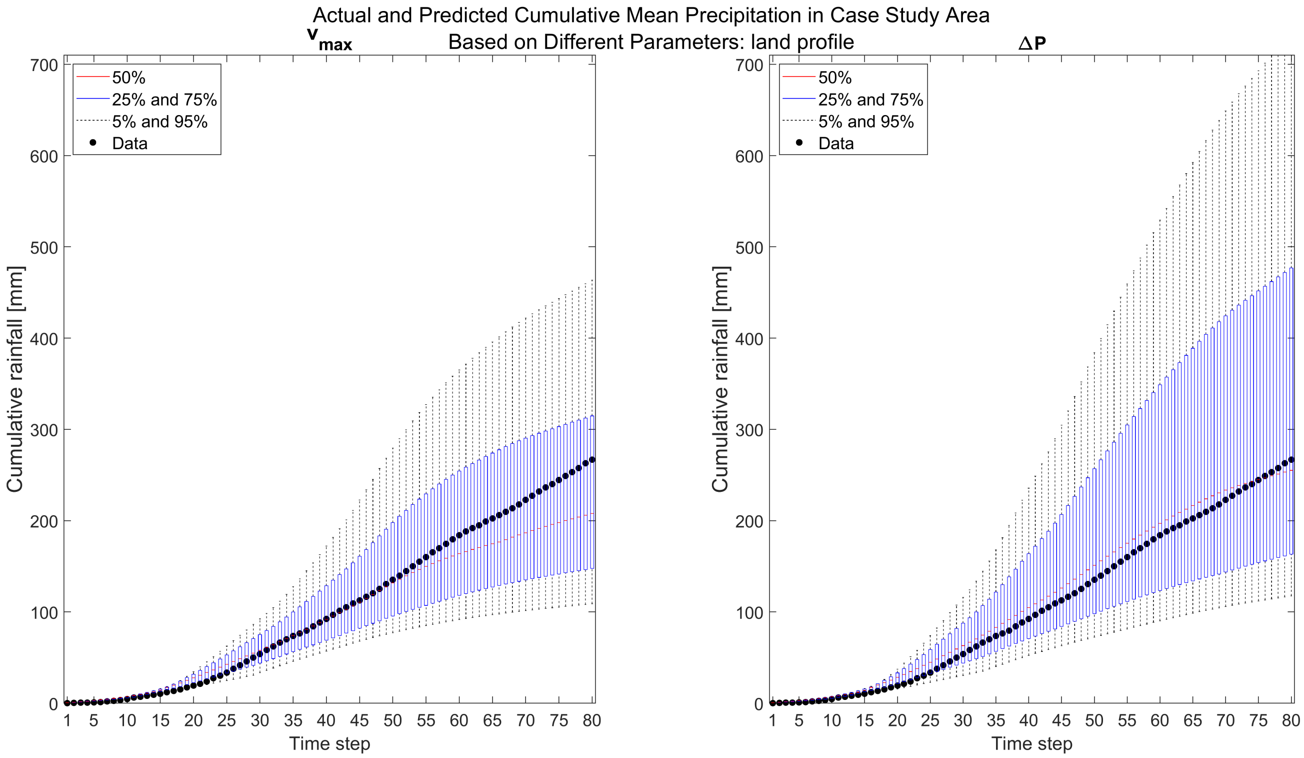


Figure 5.21: Mean cumulative precipitation within the case area compared to modelled precipitation based on the v_{max} and land profile.

The cumulative mean precipitation predicted my the v_{max} model with the land profile is competitive respect to the ΔP modelled normal profile (fig.5.21). While the cumulative mean precipitation varies equally at the final time steps, the error margins are smaller and it manages to replicate the observed rainfall for a longer period of time.

The cumulative mean precipitation predicted my the ΔP model 50% replicates the observed data better compared to the other model versions. Nonetheless, the error bars are the largest, therefore, it can not be established as the best model version, as large error margins imply major uncertainty.

In summary, the most suitable model is still open for interpretation. The findings of this case study do suggest that a v_{max} based model the land profile provides a significantly better fit. For the ΔP based model, the land profile also shows improvements over the normal profile, however, this is comes at the cost of the error margins.As the ΔP based model with the land fit does have the best 50%, this version is compared to the Bader model in the following subsections.

Based on these insights, the second sub-question can be answered.

What is the difference between tropical cyclone radial rainfall profiles over land and over the ocean?

This study confirms that over ocean, the relationship between v_{max} and p_{max} is well established as seen in literature [13, 14]. However, the case results give reason to believe that above land there is a better relationship with ΔP. Additionally, as it could also be observed in fig.2.4 by Jiang et al., precipitation levels can be high at the eye of the on land, while above ocean p_{max} is reached at a further distance. Therefore, providing a fit above land where precipitation is p_{max} at the eye can be a suitable solution. However, it should be noted that, while it does occur more often that rainfall peaks at the eye upon landfall, this cannot be easily generalized. Therefore, more case studies would need to be investigated to gain a deeper understanding of the most appropriate land fit.

5.2.5. Comparison to the original Bader model

As indicated by the research question, the main objective of this thesis is to reduce the biases in Bader's parametric model. To identify how well the objective has been met, the updated model is compared to the original Bader model.

The set up and coefficients of the original Bader model, explained in chapter 2.2.6, have been used to replicate the precipitation in the case study area. Since the Bader model has a tendency to overestimate rainfall, all precipitation below 10 mm/hr is set to zero.

From both the mean precipitation within the area and the cumulative mean, it is clear that the Bader model overestimates rainfall. The new model undoubtedly shows great improvements. The main reasons for these improvements provide and answer to sub research question three.

How can parametric rainfall models for tropical cyclones be improved to better capture the observed variability over space and time?

First of all, this research investigates the possibilities of enhancing the quality of the information fed to the model by blending different data sets. It is clear that constructing a novel data set provides better insights to build input-output relationships. Consequently, regardless of the method fit, this project highlights the importance of the pre-processing step of precipitation modelling. In addition, using the area under the graph to fit the Holland profile allowed to reduce overestimation at larger radii. Therefore, setting precipitation to 0 when it was above 10 mm/hr is no longer required.

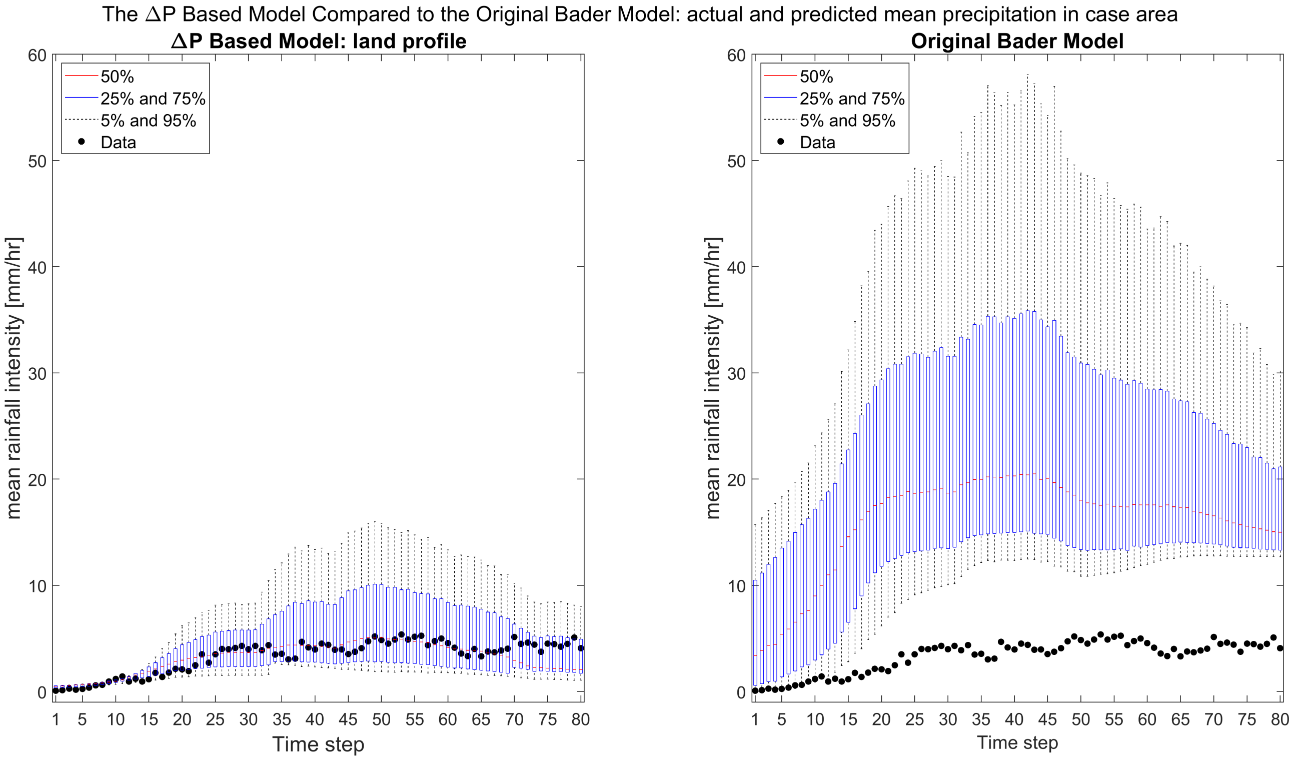


Figure 5.22: Comparison of the mean precipitation within the case area computed by the Bader model and the ΔP land profile model.

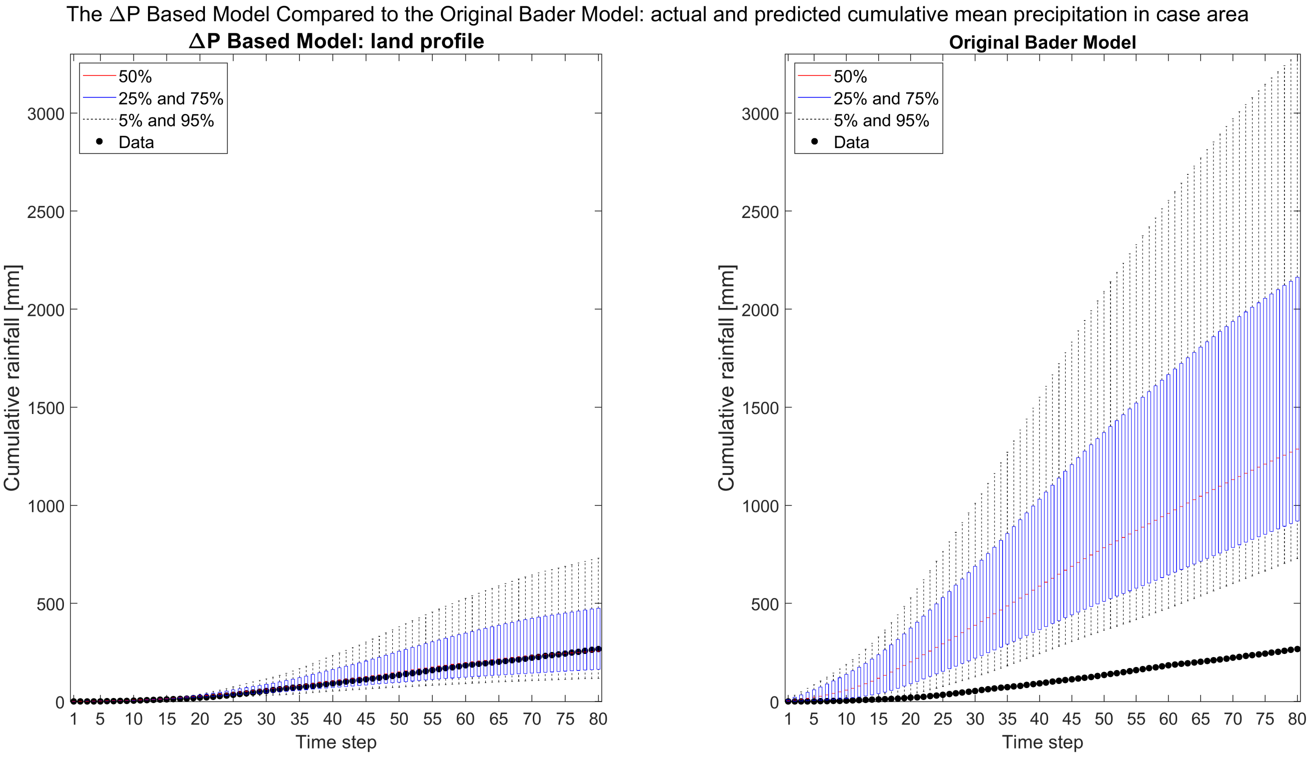


Figure 5.23: Comparison of the mean cumulative precipitation within the case area computed by the Bader model and the ΔP land profile model.

6

Conclusion

6.1. Synthesis

In this thesis an alternative improved parametric precipitation model has been proposed in order to answer the research question:

How can the bias in the radial rainfall distributions of a Tropical Cyclone in Bader's parametrized model be reduced and be used for reliable rainfall estimates both above land and the ocean?

In this study, the main biases of the Bader model have been identified as follows:

- Poor representation of rainfall over land in training data (QSCAT-R).
- Overestimation of precipitation.
- Poor generalization: when the model is applied to new data outside of the training set, performance decreases.

Three main components that are of importance to reducing the above mentioned biases in the model are identified, namely, the data, the copula and the precipitation profile.

Several data sources were explored to provide good coverage above land and ocean. Precipitation estimates from both the precipitation radar and microwave imager of the TRMM and GPM satellites were used to maximize spatial coverage. The coverage over land was improved with the help of Stage IV data, which is hourly data of the entire contiguous United States (ConUS). A distinction was made between TRMM/GPM data only and TRMM/GPM data with Stage IV data included, referred to as the Stage IV blend. This allowed to test whether including Stage IV over land influenced the trained model. The conclusion was that both approaches provided very similar results and that the inclusion of Stage IV data does not lead to significantly better performance, predominantly due to the fact that only TRMM/GPM time steps were included.

Two predictive parameters were tested, the previously used maximum sustained wind speed [m/s] (v_{max}), and the newly introduced pressure deficit [hPa] (ΔP). In both cases, the Frank copula provided a suitable approximation of the relationship between the maximum precipiation p_{max} [mm/hr] and bothv_{max}-ΔP. For ΔP, the Gaussian copula is also a suitable fit when using the Stage IV Blend data set. v_{max} has a marginally stronger relationship with p_{max}. Therefore, the p_{max} samples generated based on ΔP tend to have a higher uncertainty range. On the other hand, ΔP samples often generate higher p_{max} values, which may be an advantage for reproducing the extremes.

Four precipitation profiles were proposed and tested. The results showed that the Holland profile performs poorly for low p_{max}, as precipitation can remain at p_{max} level at larger radii. Therefore, a profile where precipitation is constant for low p_{max} values has been selected to be used in the case study of hurricane Florence (2018). This profile has been fitted based on the expected area under the curve (total rainfall [mm]).

The case study was conducted to interpret the predictive skill of the model when applied to data outside the training. Within the defined case area of approximately 94400 km^2, a cumulative average of 370 mm fell within 4 days. Most of the time, both the ΔP and v_{max} based models were able to reproduce a reasonable range of precipitation values compared with the observations. Here, the v_{max} based cumulative mean precipitation ranged 100-350 mm, and the ΔP based cumulative mean precipitation ranged from 110-570 mm. Nevertheless, single-parameter models are still too simplistic to reproduce the complex spatio-temporal rainfall patterns in tropical cyclones when they make landfall. For example, there are cases in which ΔP and v_{max} are low, yet the rainfall rates are high due other physical phenomena, such as topography. This is a real limitation of a model based on a single parameter.

A comparison with the IPET model showed that the new model is competitive as the 50th percentile of the cumulative mean precipitation based on the ΔP model almost identically matches that of IPET (200mm). However, the IPET model did provide inspiration for an alternative fit above land. The land fit improves the modelled 50% based on both v_{max} and ΔP. With the land fit, the ΔP based model performs the best in this particular case study. However, it is still unclear which model performs best over land, as one case study is not sufficient to derive a conclusive statement. Nonetheless, when comparing the updated model to the original Bader model, it shows improved results, as the Bader model overestimates precipitation by almost one order of magnitude. The new data ensured better representation over land, the overestimation of precipitation is reduced, and the model can be applied with more confidence outside of the training set.

In the field of risk management, having a model trained by proper pre-processed data that showcases the risks associated with extreme events, as a TC, can have real benefits in better understanding flood risk. However, the models are yet to be tested on flood case studies.

6.2. Limitations

While the updated model shows significant improvements compared to the original Bader model, it still contains several important limitations:

1. The model predictions have considerable uncertainty ranges (i.e. the model is overdispersed). The bias is relatively low when modelling the average of a larger area, however, the model cannot be used to predict specific rainfall intensities with high confidence at a specific location.
2. The model was trained based on data from the north Atlantic basin only and might not be directly transferable to other basins.
3. The model does not take into account possible asymmetry in rainfall rates. Hence, it unrealistically assumes that the rainfall is the same in all radial directions.
4. While the model can make a distinction between over ocean and over land, it does not fully consider the temporal variation of the rainfall profiles along the storm track (e.g. using an autoregressive process). Each profile is modeled independently from the others.

6.3. Recommendations for future research

The following recommendations are made for future research:

1. Test the model for a case study outside the North Atlantic. Alternatively, include data from other oceanic basins in the training data set, or replicate the model with only data from other oceanic basins.
2. Investigate possible ways to introduce asymmetry into the rainfall profiles. For example, test whether the model component used to produce wind profile asymmetry in Deltares's TCWiSE is suitable for precipitation and has the potential to improve the model's results[3]. Alternatively, provide a correction factor, like IPET, to simulate asymmetry.
3. Study the temporal correlation structure of rainfall profiles. For example, by using pdef/vmax over multiple time steps (multivariate input), applying a Kalman filter on the model parameters or adjusting the model coefficients based on the difference between the observed and modeled precipitation rates at the previous time step.

Bibliography

[1] Eric S. Blake and David A. Zelinksy. “Hurricane Harvey (AL092017)”. In: *National Hurricane Center Tropical Report* AL092017.2005 (2018), pp. 1–77. ISSN: 0893-2190. DOI: `10.1097/jpn.0000000000000424`.

[2] Seung Won Suh et al. “An efficient early warning system for typhoon storm surge based on time-varying advisories by coupled ADCIRC and SWAN”. In: *Ocean Dynamics* 65.5 (2015), pp. 617–646. ISSN: 16167228. DOI: `10.1007/s10236-015-0820-3`.

[3] Kees Nederhoff et al. “Simulating Synthetic Tropical Cyclone Tracks for Statistically Reliable Wind and Pressure Estimations”. In: *Natural Hazards and Earth System Sciences* (2020), pp. 1–30. ISSN: 1561-8633. DOI: `10.5194/nhess-2020-250`.

[4] D.J. Bader. “Including stochastic rainfall distributions in a probabilistic modelling approach for compound flooding due to tropical cyclones”. In: (2019). URL: `http://resolver.tudelft.nl/uuid:57b9e495-0c90-4cf5-ab22-e169fb908ac1`.

[5] John T. Brackins and Alfred J. Kalyanapu. “Evaluation of parametric precipitation models in reproducing tropical cyclone rainfall patterns”. In: *Journal of Hydrology* 580.March 2019 (2020), p. 124255. ISSN: 00221694. DOI: `10.1016/j.jhydrol.2019.124255`. URL: `https://doi.org/10.1016/j.jhydrol.2019.124255`.

[6] *Hurricane FAQ – NOAA’s Atlantic Oceanographic and Meteorological Laboratory*. URL: `https://www.aoml.noaa.gov/hrd-faq/#tc-formation`.

[7] *NWS JetStream - Tropical Cyclone Structure*. URL: `https://www.weather.gov/jetstream/tc_structure`.

[8] James Henry. “Tropical and equatorial climates”. In: *Encyclopedia of Earth Sciences Series* (2005), pp. 742–750. ISSN: 1871756X. DOI: `10.1007/1-4020-3266-8{\_}212`.

[9] Daniel R. Chavas, Kevin A. Reed, and John A. Knaff. “Physical understanding of the tropical cyclone wind-pressure relationship”. In: *Nature Communications* 8.1 (2017). ISSN: 20411723. DOI: `10.1038/s41467-017-01546-9`. URL: `http://dx.doi.org/10.1038/s41467-017-01546-9`.

[10] Timothy Schott et al. “The Saffir-Simpson Hurricane Wind Scale”. In: (2019), pp. 1–4.

[11] Robert Rogers, Frank Marks, and Timothy Marchok. “Tropical Cyclone Rainfall”. In: *Encyclopedia of Hydrological Sciences* June 2018 (2009). DOI: `10.1002/0470848944.hsa030`.

[12] Peter Dodge, Robert W. Burpee, and Frank D. Marks. “The kinematic structure of a hurricane with sea level pressure less than 900 mb”. In: *Monthly Weather Review* 127.6 I (1999), pp. 987–1004. ISSN: 00270644. DOI: `10.1175/1520-0493(1999)127<0987:TKSOAH>2.0.CO;2`.

[13] Manuel Lonfat, Frank D. Marks, and Shuyi S. Chen. “Precipitation distribution in tropical cyclones using the Tropical Rainfall Measuring Mission (TRMM) microwave imager: A global perspective”. In: *Monthly Weather Review* 132.7 (2004), pp. 1645–1660. ISSN: 00270644. DOI: `10.1175/1520-0493(2004)132<1645:PDITCU>2.0.CO;2`.

[14] Haiyan Jiang, Jeffrey B. Halverson, and Joanne Simpson. “Difference of rainfall distribution for tropical cyclones over land and ocean and rainfall potential derived from satellite observations and its implication on hurricane landfall flooding prediction”. In: *27th Conference on Hurricanes and Tropical Meteorology* (2006).

[15] Sarah C. Jones et al. “The extratropical transition of tropical cyclones: Forecast challenges, current understanding, and future directions”. In: *Weather and Forecasting* 18.6 (2003), pp. 1052–1092. ISSN: 08828156. DOI: `10.1175/1520-0434(2003)018<1052:TETOTC>2.0.CO;2`.

[16] Mark Demaria and Frank D Marks. “Development of a Tropical Cyclone Rainfall Climatology and Persistence (R-CLIPER) Model Principle”. In: December (2003).

[17] Reda Snaiki and Teng Wu. “A theoretical model for rapid estimates of rainfall during tropical cyclones”. In: *2017 Americas Conference on Wind Engineering, ACWE 2017* 2016 (2017).

[18] Manuel Lonfat et al. “A parametric model for predicting hurricane rainfall”. In: *Monthly Weather Review* 135.9 (2007), pp. 3086–3097. ISSN: 00270644. DOI: `10.1175/MWR3433.1`.

[19] US Army Corps of Engineers. “Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System: Final Report of the Interagency Performance Evaluation Task Force”. In: I.June (2009), p. 147. URL: `https://ipet.wes.army.mil/`.

[20] Manuel Lonfat, Frank D. Marks, and Shuyi S. Chen. “Precipitation distribution in tropical cyclones using the Tropical Rainfall Measuring Mission (TRMM) microwave imager: A global perspective”. In: *Monthly Weather Review* 132.7 (2004), pp. 1645–1660. ISSN: 00270644. DOI: `10.1175/1520-0493(2004)132<1645:PDITCU>2.0.CO;2`.

[21] Greg J. Holland, James I. Belanger, and Angela Fritz. “A revised model for radial profiles of hurricane winds”. In: *Monthly Weather Review* 138.12 (2010), pp. 4393–4401. ISSN: 00270644. DOI: `10.1175/2010MWR3317.1`.

[22] Andreas Langousis and Daniele Veneziano. “Theoretical model of rainfall in tropical cyclones for the assessment of long-term risk”. In: *Journal of Geophysical Research Atmospheres* 114.2 (2009), pp. 1–11. ISSN: 01480227. DOI: `10.1029/2008JD010080`.

[23] Josh (EOSDIS) Blumenfel. *From TRMM to GPM: The Evolution of NASA Precipitation Data | Earthdata*. URL: `https://earthdata.nasa.gov/learn/articles/tools-and-technology-articles/trmm-to-gpm`.

[24] NASA. *PRECIPITATION PROCESSING SYSTEM Metadata for TRMM Products*. 2017. URL: `https://pps.gsfc.nasa.gov/Documents/filespec.TRMM.V7.pdf`.

[25] JAXA. *JAXA/EORC Tropical Cyclone Database*. URL: `https://sharaku.eorc.jaxa.jp/TYP_DB/index.html`.

[26] *GES DISC Dataset: TRMM Precipitation Radar Rainfall Rate and Profile L2 1.5 hours V7 (TRMM_2A25 7)*. URL: `https://disc.gsfc.nasa.gov/datasets/TRMM_2A25_7/summary?keywords=2A25`.

[27] NASA. *PRECIPITATION PROCESSING SYSTEM GLOBAL PRECIPITATION MEASUREMENT File Specification for GPM Products*. 2020. URL: `https://gpm.nasa.gov/sites/default/files/2020-02/filespec.GPM__0.pdf`.

[28] *GES DISC Dataset: GPM DPR Precipitation Profile L2A 1.5 hours 5 km V06 (GPM_2ADPR 06)*. URL: `https://disc.gsfc.nasa.gov/datasets/GPM_2ADPR_06/summary?keywords=GPM_2ADPR_06`.

[29] Ying Lin and Kenneth E. Mitchell. “The NCEP stage II/IV hourly precipitation analyses: Development and applications”. In: *85th AMS Annual Meeting, American Meteorological Society - Combined Preprints* (2005), pp. 1649–1652.

[30] Hylke E. Beck et al. “Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS”. In: *Hydrology and Earth System Sciences Discussions* (2018), pp. 1–23. ISSN: 1027-5606. DOI: `10.5194/hess-2018-481`.

[31] Jun Du (NOAA). *EOL data archive – GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data*. DOI: `https://doi.org/10.5065/D6PG1QDD`. URL: `https://data.eol.ucar.edu/dataset/21.093`.

[32] *RAMMB: Tropical Cyclone Extended Best Track Dataset*. Feb. 2020. URL: `http://rammb.cira.colostate.edu/research/tropical_cyclones/tc_extended_best_track_dataset/index.asp`.

[33] *Atlantic EBRTK*. Feb. 2020. URL: `http://rammb.cira.colostate.edu/research/tropical_cyclones/tc_extended_best_track_dataset/data/ebtrk_atlc_1988_2018.txt`.

[34] Christian Genest and Anne-Catherine Favre. “Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask”. In: *JOURNAL OF HYDROLOGIC ENGINEERING* 12.4 (2007), pp. 347–368. ISSN: 10577408. DOI: `10.1016/j.jcps.2009.09.002`.

[35] M Sadegh, E Ragno, and A AghaKouchak. “Water resources research progress”. In: *Water Resources Research Progress* (2017), pp. 1–433. DOI: `10.1002/2016WR020242.Received`.

[36] *Copula random numbers - MATLAB copularnd - MathWorks Benelux*. URL: `https://nl.mathworks.com/help/stats/copularnd.html`.

[37] NWS and NOAA. *Hurricane Florence: September 14, 2018*. 2018. URL: `https://www.weather.gov/ilm/HurricaneFlorence`.

[38] Stacy R. Stewart and Robbie Berg. “National Hurricane Center tropical cyclone report: Hurricane Florence (AL062018)”. In: *National Hurricane Center* May (2019), pp. 1–98. URL: `https://www.nhc.noaa.gov/data/tcr/AL152017_Maria.pdf`.

[39] Toby D. Feaster et al. “Preliminary peak stage and streamflow data for selected U.S. Geological Survey streamgaging stations in North and South Carolina for flooding following Hurricane Florence, September 2018”. In: *Open-File Report 2018-1172* October (2018). ISSN: 2331-1258. URL: `https://pubs.er.usgs.gov/publication/ofr20181172`.

A

Deriving the x_n coefficient when the simulated p_{max} is equal to the fit p_{max}

$$p_r(R) = \left(\frac{p_{\max} * \left(\frac{Rp_{\max}}{R}\right)^{b_s}}{\exp\left(\left(\frac{Rp_{\max}}{R}\right)^{b_s}\right)} \right)^{x_n} = p_{\max} \tag{A.1}$$

$$p_r\left(Rp_{\max}\right) = p_{\max} = \left(\frac{p_{\max} * \left(\frac{Rp_{\max}}{Rp_{\max}}\right)^{b_s}}{\exp\left(\left(\frac{Rp_{\max}}{Rp_{\max}}\right)^{b_s}\right)} \right)^{x_n} \qquad \left(\frac{Rp_{\max}}{Rp_{\max}}\right)^{b_s} = 1 \tag{A.2}$$

$$p_r\left(Rp_{\max}\right) = p_{\max} = \left(\frac{p_{\max}}{\exp(1)}\right)^{x_n} \tag{A.3}$$

$$x_n = \frac{\log\left(p_{\max}\right)}{\log\left(\frac{p_{\max}}{\exp(1)}\right)} \tag{A.4}$$

B

Conditional Sampling from Gaussian Copula

1. Compute the CDF of the ΔP of interest using the BirnbaumSauders marginal distribution with the MvCAT computed β_b and γ_b as indicated in table 5.4. This is the v value
2. Compute the normal inverse of v (x_1) as well as the normal inverse of a uniformly distributed random number (x_2) and combine the two in a vector $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
3. Generate the lower Cholesky matrix **C** of the correlation matrix $\begin{pmatrix} 1 & \theta_1 \\ \theta_1 & 1 \end{pmatrix}$
4. multiply $\mathbf{X}$ and **C** resulting in a vector $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$
5. Compute the normal CDF of c_1 this is equal to u. coputing the normal CDF of c_2 would result in v from step 1
6. to obtain a value of p_{max}, u is converted back to the original data space by computing the inverse CDF of the lognormal family with μ_l and σ_l as indicated in table 5.2

C

Copula Selection & Correlation Supporting Graphs

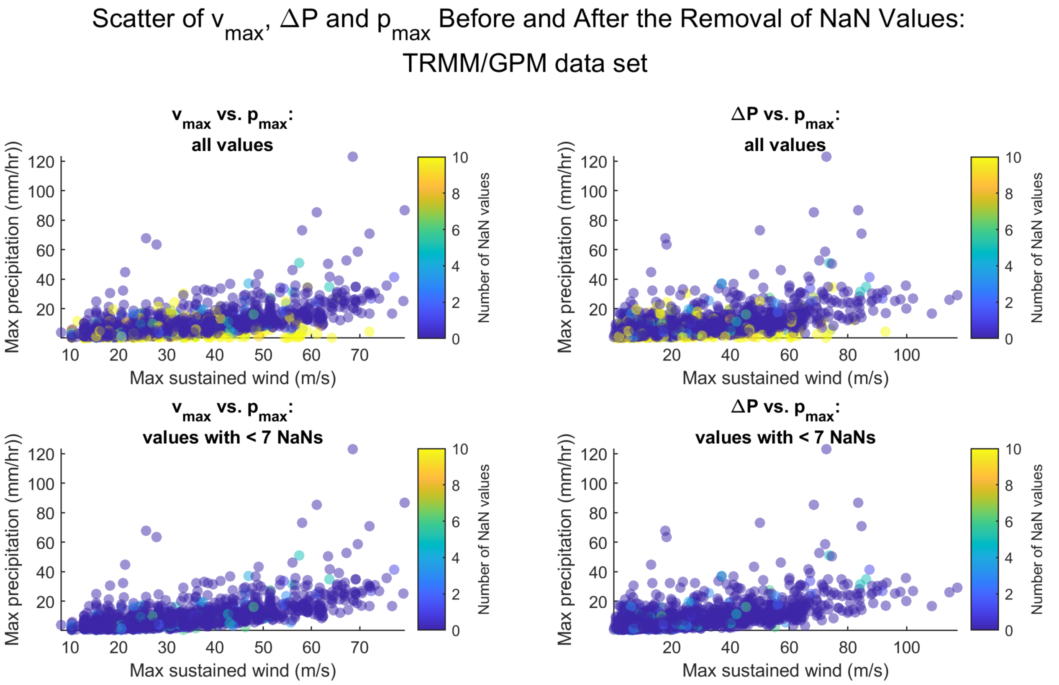
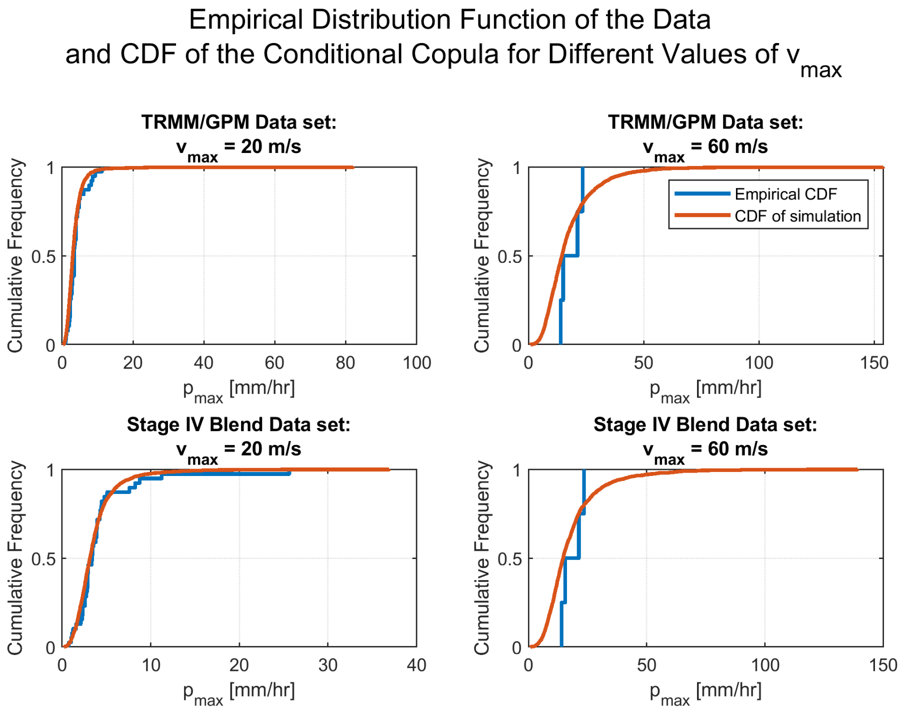


Figure C.1: Scatter of v_{max} and p_{max} before and after the removal of **NaN** values for the TRMM/GPM data set


Empirical Distribution Function of the Data
and CDF of the Conditional Copula for Different Values of v_{max}

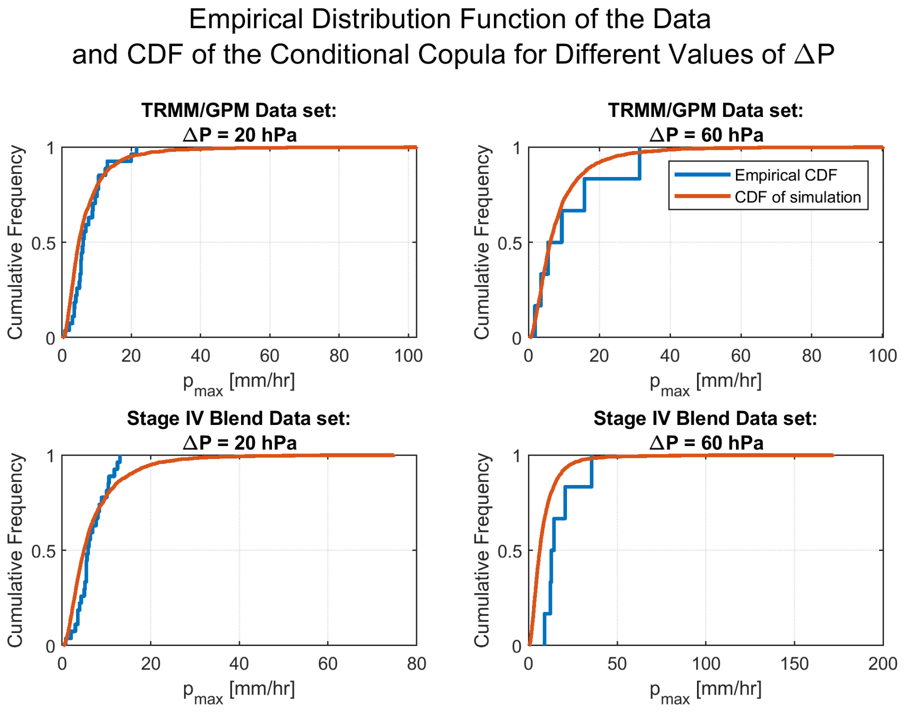
TRMM/GPM Data set: v_{max} = 20 m/s

TRMM/GPM Data set: v_{max} = 60 m/s

Stage IV Blend Data set: v_{max} = 20 m/s

Stage IV Blend Data set: v_{max} = 60 m/s

Figure C.2: Comparison of the empirical distribution of the data and the CDF if the conditional copula for p_{max} based on v_{max}


Empirical Distribution Function of the Data
and CDF of the Conditional Copula for Different Values of ΔP

TRMM/GPM Data set: ΔP = 20 hPa

TRMM/GPM Data set: ΔP = 60 hPa

Stage IV Blend Data set: ΔP = 20 hPa

Stage IV Blend Data set: ΔP = 60 hPa

Figure C.3: Comparison of the empirical distribution of the data and the CDF if the conditional copula for p_{max} based on ΔP

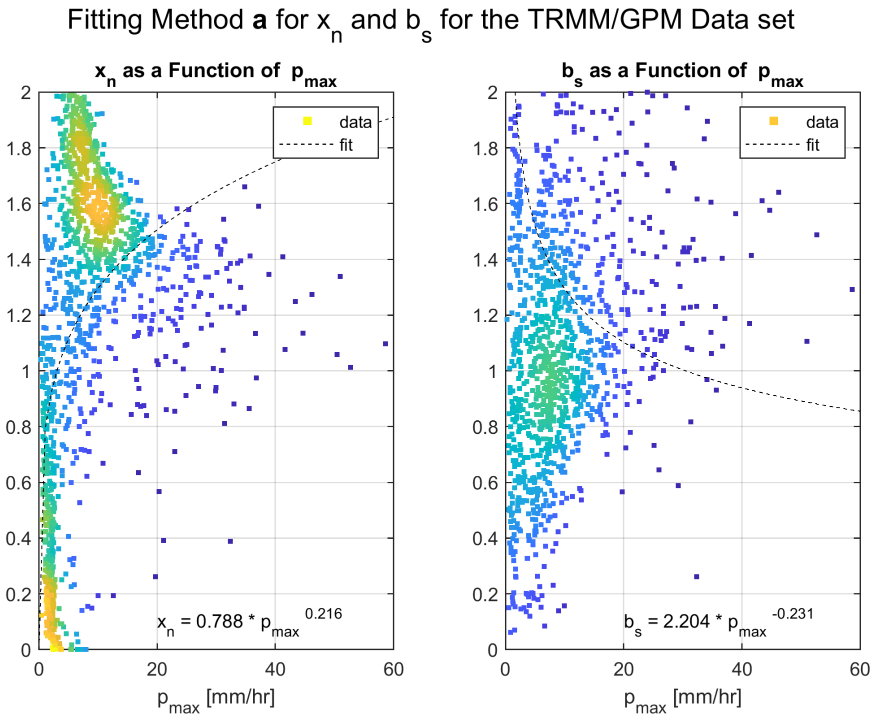


Figure C.4: best fit for x_n and b_s for the TRMM/GPM data set method **a**. The scatter are colored according to density of the points (yellow is high, purple is low).

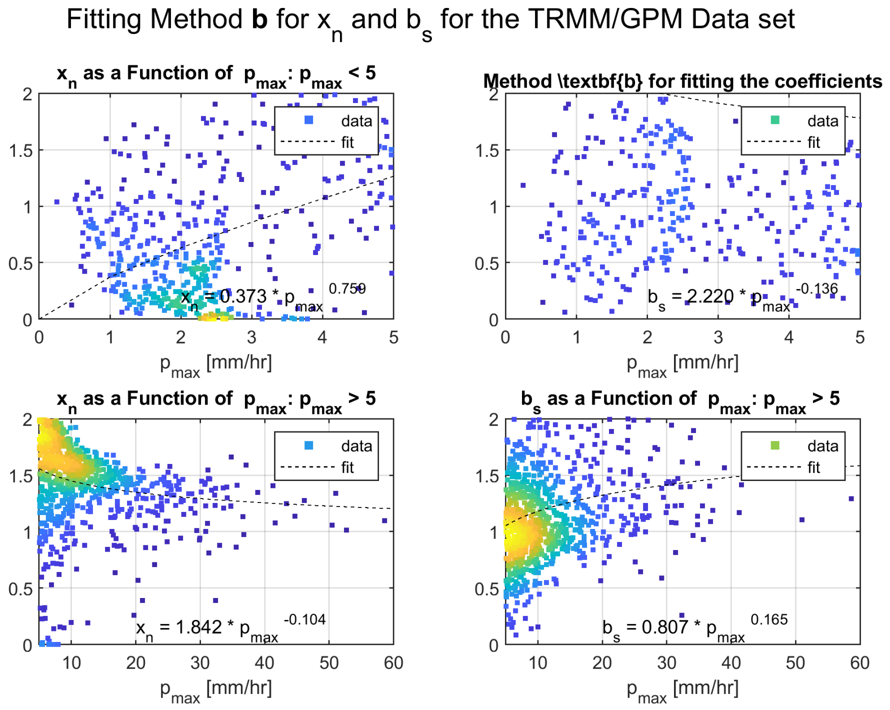


Figure C.5: best fit for x_n and b_s for the TRMM/GPM data set method **b**. The scatter are colored according to density of the points (yellow is high, purple is low).

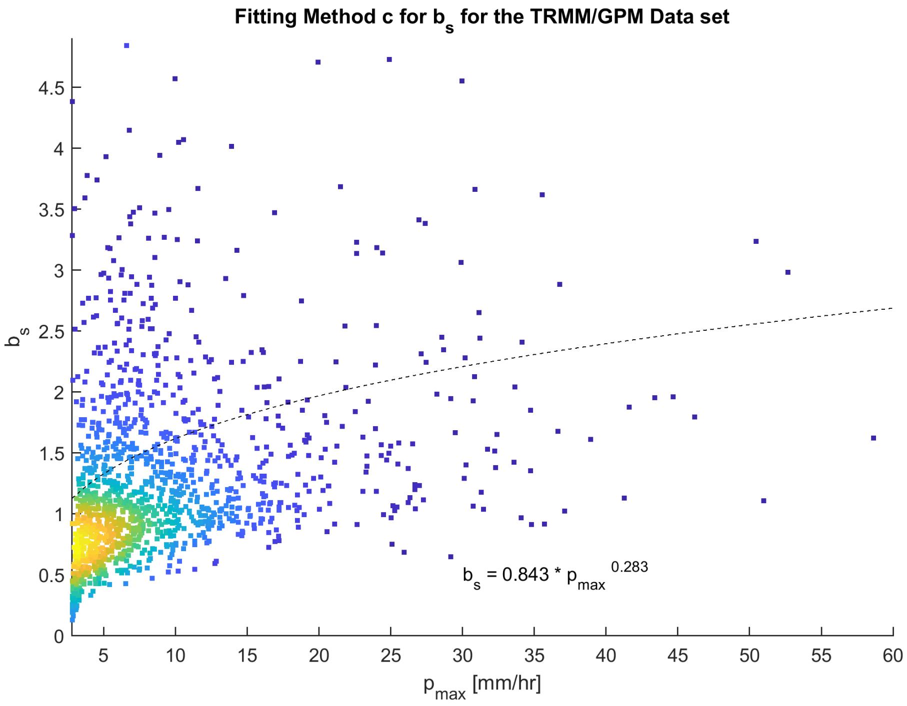


Figure C.6: best fit for x_n and b_s for the TRMM/GPM data set method **c**. The scatter are colored according to density of the points (yellow is high, purple is low).

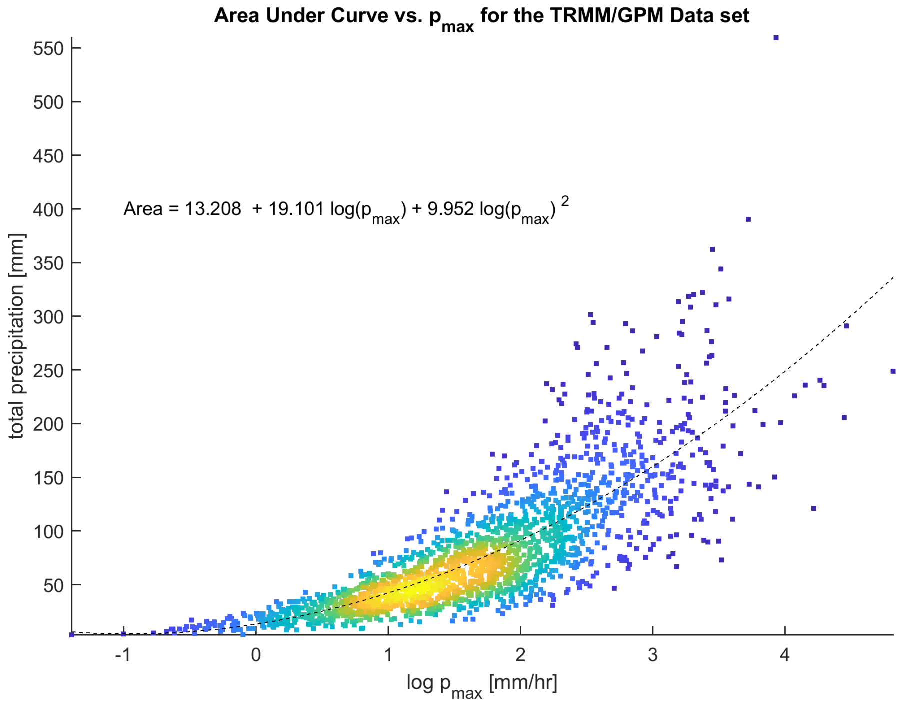


Figure C.7: best fit for the area under the graph for the TRMM/GPM data set method **d**. The scatter are colored according to density of the points (yellow is high, purple is low).

D

Performance Metrics for Varying Conditions

Table D.1: Performance metric of observed p_{max} vs. the p_{max} of the fit $p_r(Rv_{max})$ for **a** and **b**: v_{max} based for the TRMM/GPM data set

data set, parameter type & fit	**TRMM/GPM v_{max} based fit a**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	10.62	9.61	8.27	7.10	206.27
MAE	6.89	5.90	4.61	4.09	112.46
bias	-6.89	-5.89	-4.46	-1.13	112.06

data set, parameter type & fit	**TRMM/GPM v_{max} based fit b**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	10.43	9.20	8.07	6.98	17.54
MAE	6.68	5.48	4.35	3.72	14.35
bias	-6.68	-5.45	-4.04	-1.76	13.52

Table D.2: Performance metric of observed p_{max} vs. the p_{max} of the fit $p_r(Rv_{max})$ for **a** and **b**: ΔP based for the TRMM/GPM data set

data set, parameter type & fit	**TRMM/GPM ΔP based fit a**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	10.86	10.09	8.95	7.07	145.82
MAE	7.02	6.18	5.00	3.74	103.80
bias	-7.02	-6.16	-4.83	-1.63	103.78

data set, parameter type & fit	**TRMM/GPM ΔP based fit b**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	10.83	9.77	8.56	7.31	19.88
MAE	6.90	5.89	4.60	3.83	18.12
bias	-6.90	-5.85	-4.24	-1.44	17.56

Table D.3: Performance metric of observed p_{max} vs. the p_{max} of the fit $p_r(Rv_{max})$ for **a** and **b**: v_{max} based for the Stage IV Blend data set

data set, parameter type & fit	**Stage IV blend ΔP based fit a**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	10.82	10.00	8.62	7.10	218.02
MAE	7.06	6.24	4.91	3.90	101.19
bias	-7.06	-6.22	-4.74	-0.95	100.86

data set, parameter type & fit	**Stage IV blend ΔP based fit b**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	10.76	9.65	8.28	6.94	17.00
MAE	6.93	5.93	4.53	3.73	14.91
bias	-6.93	-5.90	-4.17	-1.21	14.11

Table D.4: Performance metric for the area under the graph for fit **a**, **b**, **c**, and **d**: v_{max} based for the TRMM/GPM data set

data set, parameter type & fit	**TRMM/GPM v_{max} based fit a**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	83.54	67.80	50.47	78.57	3199.45
MAE	64.52	49.73	35.42	50.63	1798.33
bias	-64.43	-48.73	-26.40	24.62	1798.25

data set, parameter type & fit	**TRMM/GPM v_{max} based fit b**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	79.60	60.59	50.45	60.36	291.90
MAE	58.79	42.72	35.54	43.46	243.75
bias	-57.95	-39.75	-18.61	15.45	243.34

data set, parameter type & fit	**TRMM/GPM v_{max} based fit c**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	75.01	72.58	69.16	63.72	214.01
MAE	55.07	57.68	53.16	46.51	175.10
bias	-18.84	-2.88	7.49	13.94	172.73

data set, parameter type & fit	**TRMM/GPM v_{max} based fit d**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	54.63	57.37	65.40	76.89	285.10
MAE	38.84	45.36	54.29	64.45	248.37
bias	-4.37	26.36	45.13	59.83	248.03

Table D.5: Performance metric for the area under the graph for fit **a**, **b**, **c**, and **d**: ΔP based for the TRMM/GPM data set

data set, parameter type & fit	**TRMM/GPM ΔP based fit a**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	87.98	74.87	58.38	76.31	2969.63
MAE	67.87	54.71	40.72	50.52	2125.13
bias	-67.77	-53.12	-29.54	28.10	2125.13

data set, parameter type & fit	**TRMM/GPM ΔP based fit b**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	85.66	68.19	57.33	77.58	388.75
MAE	62.48	49.51	40.56	55.17	338.21
bias	-61.67	-46.13	-17.70	30.93	337.75

data set, parameter type & fit	**TRMM/GPM ΔP based fit c**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	76.85	78.74	72.12	81.58	301.74
MAE	53.40	63.95	54.06	58.29	245.57
bias	-5.63	-5.43	1.20	29.12	243.17

data set, parameter type & fit	**TRMM/GPM ΔP based fit d**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	64.15	59.69	61.53	88.44	385.26
MAE	46.15	47.02	49.47	75.78	336.84
bias	-1.97	19.74	37.55	71.82	336.26

Table D.6: Performance metric for the area under the graph for fit **a**, **b**, **c**, and **d**: ΔP based for the Stage IV blend data set

data set, parameter type & fit	**Stage IV blend ΔP based fit a**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	90.82	76.89	56.01	91.87	3651.76
MAE	70.22	57.10	39.63	58.41	1960.52
bias	-70.11	-55.88	-30.00	37.95	1960.11

data set, parameter type & fit	**Stage IV blend ΔP based fit b**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	88.54	69.92	56.00	77.72	327.71
MAE	64.81	51.29	39.67	56.50	278.22
bias	-63.83	-49.06	-18.35	33.89	277.30

data set, parameter type & fit	**Stage IV blend ΔP based fit c**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	81.63	79.62	68.37	79.47	239.21
MAE	59.39	63.82	50.86	58.33	190.10
bias	-8.34	-13.15	-5.78	26.64	186.79

data set, parameter type & fit	**Stage IV blend ΔP based fit d**				
sample fit	**5%**	**25%**	**50%**	**75%**	**95%**
RMSE	71.21	64.37	53.35	65.91	285.71
MAE	53.84	48.70	38.69	49.66	228.33
bias	-3.43	0.09	8.79	36.72	227.85

E

Modelled Spatial Precipitation

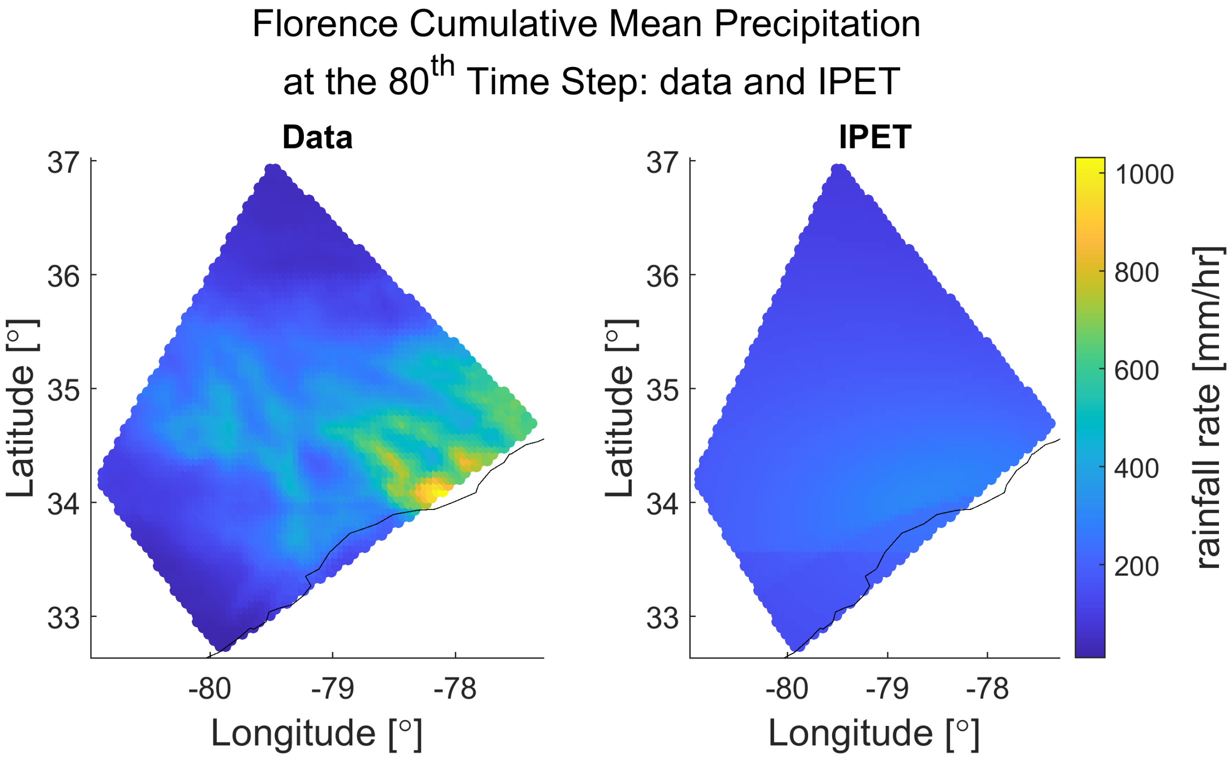


Figure E.1: Comparison of the mean cumulative precipitation within the case area of the data and computed by the IPET model

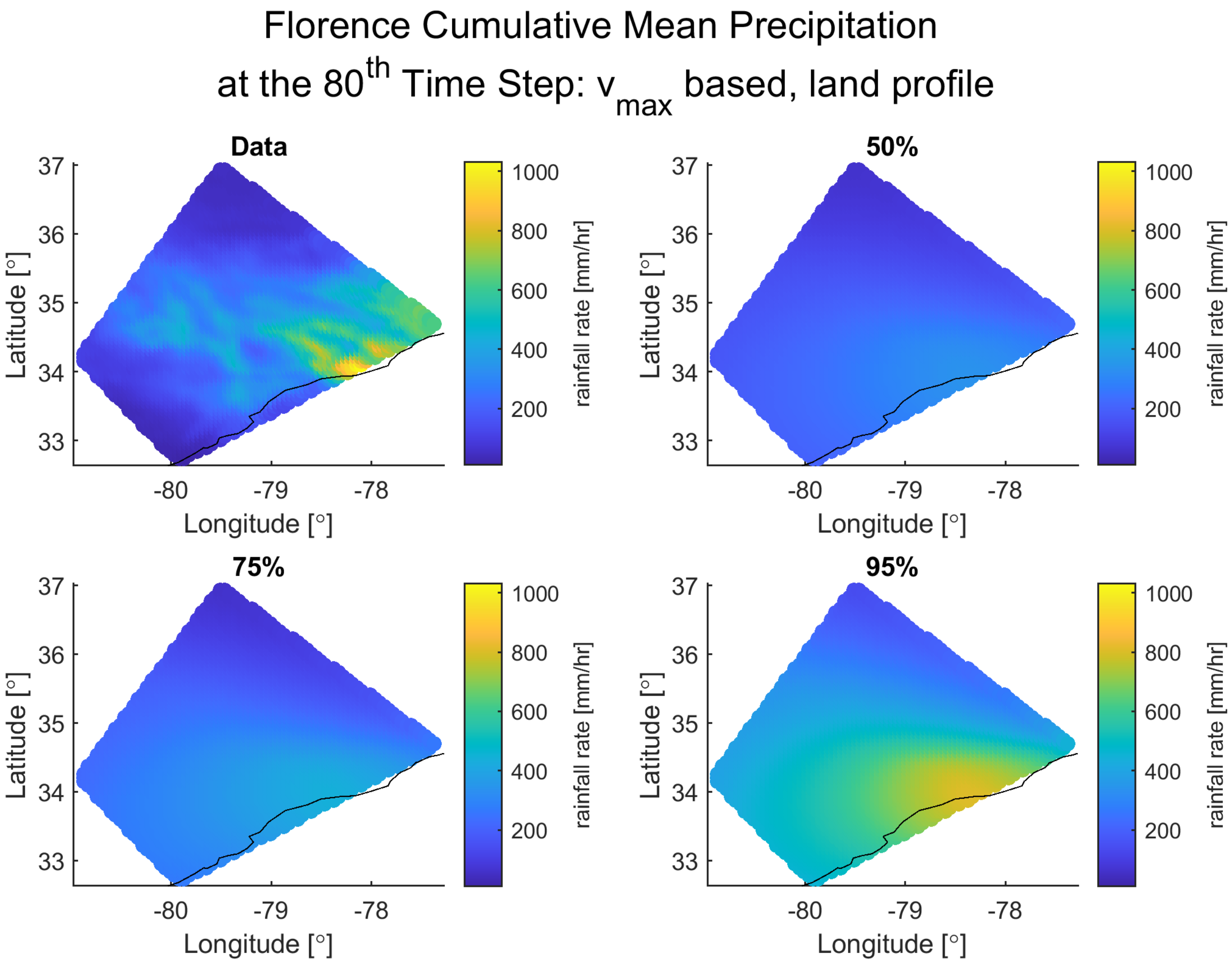


Figure E.2: Cumulative mean rain at final time-step for the dat and v_{max} modelled 50%, 70% and 95% for the land profile

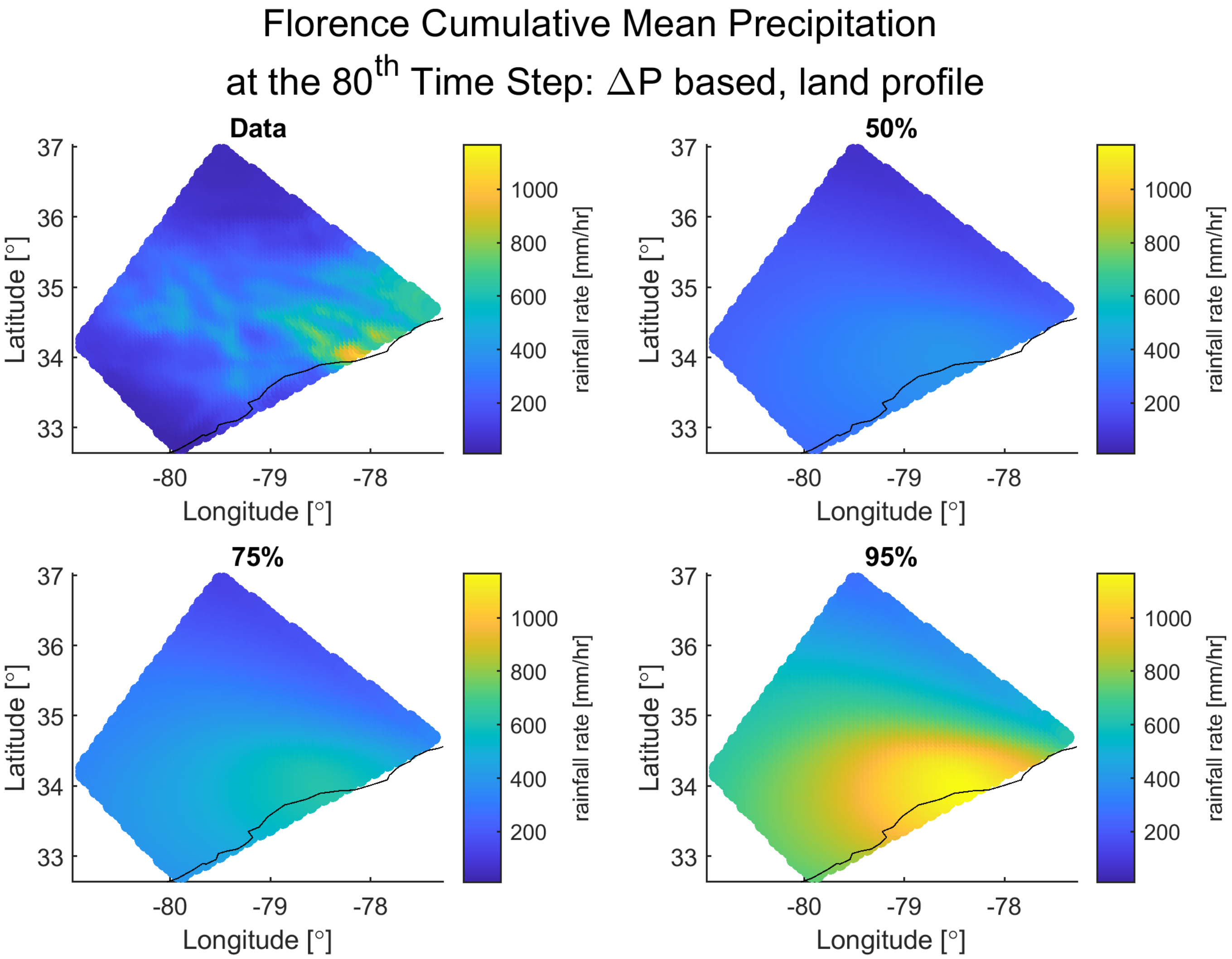


Figure E.3: Cumulative mean rain at final timestep for the dat and ΔP modelled 50%, 70% and 95% for the land profile