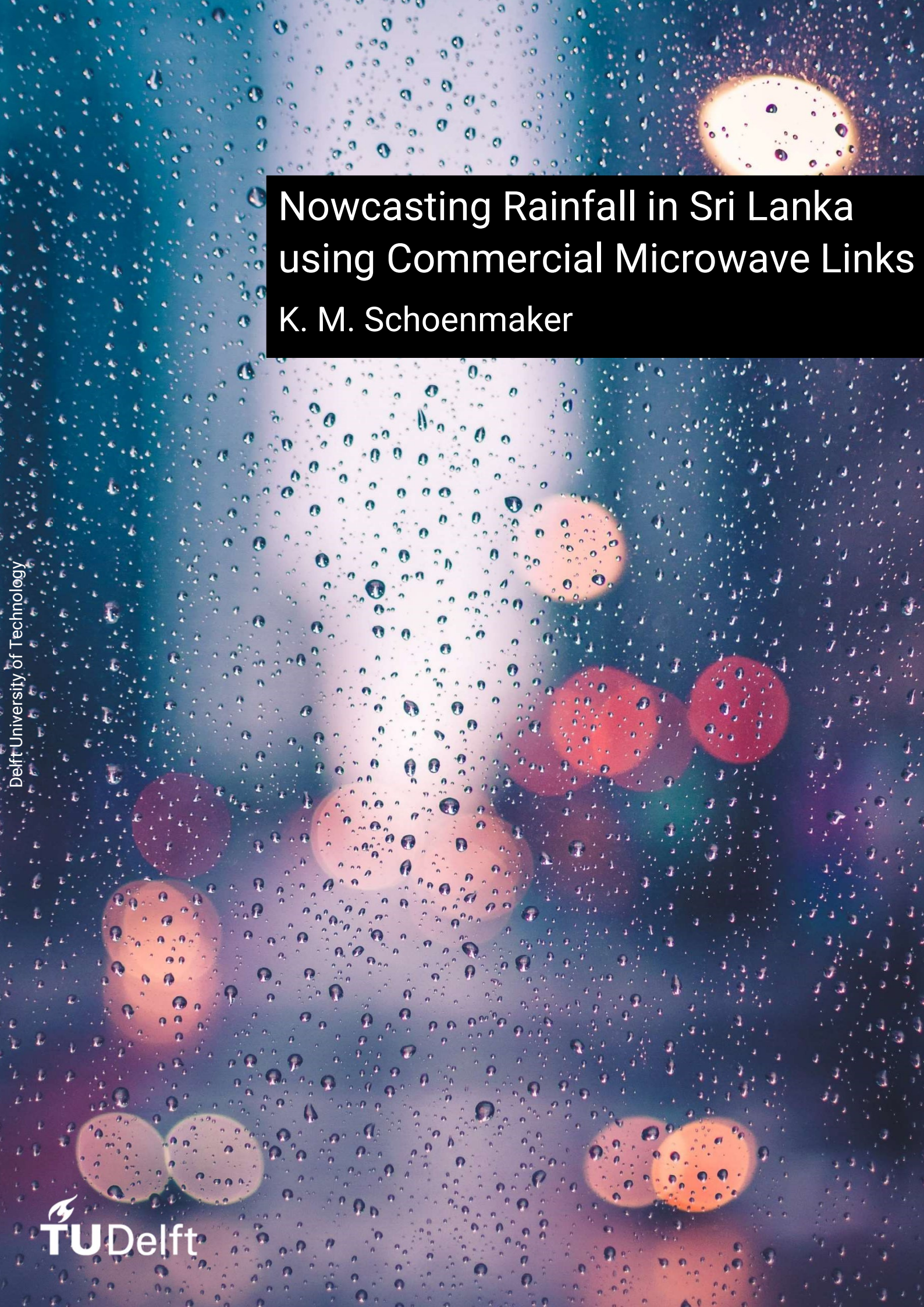


Nowcasting Rainfall in Sri Lanka using Commercial Microwave Links

K. M. Schoenmaker

Delft University of Technology

TU Delft

Nowcasting Rainfall in Sri Lanka using Commercial Microwave Links

by

K. M. Schoenmaker

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday July 6, 2022, at 11:00 AM.

Student number:	4570103	
Project duration:	December, 2021 - July, 2022	
Thesis committee	prof. dr. ir. R. (Remko) Uijlenhoet (Chair)	TU Delft (supervisor)
	ir. B. J. (Bas) Walraven	TU Delft (supervisor)
	dr. M. A. (Marc) Schleiss	TU Delft
	dr. A. (Aart) Overeem	KNMI
	R. O. (Ruben) Imhoff MSc	Deltares

Cover image: Max Bender, Chicago
https://unsplash.com/photos/1YHXFeOYpN0

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Preface

This thesis was written as part of my graduation for the master of Water Management at the faculty of Civil Engineering and Geosciences at the Delft University of Technology. This thesis was conducted between December 2021 and July 2022.

First, I would like to thank my supervisors, Remko Uijlenhoet and Bas Walraven, for all the time and support you have given me every week during our meetings. You provided me with great feedback, useful questions, and discussions to help me further with my research. I also want to thank Ruben Imhoff for your support with the Python installation and set-up. I would like to thank my other committee members, Marc Schleiss and Aart Overeem, for providing me with great feedback on my presentations and reports. Linda Bogerd, thank you for getting the GPM data on the correct grid.

I would like to thank my parents for their support throughout this journey. And also extra thanks to my mom, for reading this thesis for writing errors. Thanks to my friend, Josine van Marrewijk, for reading the thesis as a fellow Water Management student. I would finally want to thank the rest of my family and friends for all their help, support, and their distractions.

Karlijn Schoenmaker
Delft, June 2022

Abstract

Disasters like inland floods and landslides are caused by extreme rainfall. To increase the time to take early measures against such disasters it is of great importance to have access to accurate predictions of the rainfall. For the prediction of floods, Quantitative Precipitation Forecasts (QPFs) are used as input for hydrologic models. Numerical Weather Prediction (NWP) models are commonly used to generate QPFs, but for short lead times (less than 6 hours) the NWP forecasts are not accurate enough. For very short-term forecasting the nowcasting method is used. Nowcasting rainfall is a computational process of extrapolating the most recent rainfall observations and has great potential for lead times up to 6 hours. Rainfall estimations from weather radar are normally used for nowcasting, but these are not present in many parts of the world. A source for opportunistic sensing to generate rainfall estimations is Commercial Microwave Links (CMLs). Signals for telecommunication purposes will travel along a link path from one station to another. When a rainfall event occurs, the signal attenuates. This attenuation can be used for the estimation of path-averaged rainfall intensity estimations.

This thesis investigates the possibilities of using CMLs for estimating and nowcasting rainfall. The study area is Sri Lanka, a country without access to radar-based rainfall estimation. A 3.5-month CMLs data set with 2560 unique links from Dialog Sri Lanka is used and compared to 8 hourly and 12 daily rain gauge stations. The studied period is from September 12 to December 31 in 2019. To generate rainfall estimations from the CMLs, the algorithm RAINLINK is used. RAINLINK includes a set of default parameters, which was optimized for the Netherlands. For Sri Lanka, two new optimal parameters are derived. One is for the wet antenna attenuation (A_a), which is the reduction of the signal caused by a wet antenna. The second parameter is the coefficient (α), which determines how much of the minimum and maximum signal power is used to calculate the mean path-averaged rainfall intensity. For the optimal parameters daily CML estimations are calibrated with the 12 daily rain gauges. With the new optimal parameters, the rainfall estimations are calculated and used in Pysteps, an open-source library for nowcasting, to generate nowcasts for 20 rainfall events in the studied period. Deterministic and probabilistic nowcasts are generated. These nowcasts are compared with a benchmark nowcast, called Eulerian Persistance (EP).

RAINLINK gives more accurate rainfall estimations in Sri Lanka with the new parameters of the wet antenna attenuation (A_a) and the coefficient α compared to the default parameters. The daily values of the generated rainfall estimations from RAINLINK are compared with 12 rain gauges in a validation period (55 days), where the Nash-Sutcliffe Efficiency (NSE) is calculated. The validation resulted in an NSE of 0.85 for the new parameters and an NSE of 0.64 for the default parameters. The results show that the parameters are dependent on the path length of the link. Longer path lengths tend to have lower wet antenna attenuation and a higher value for the coefficient α. Nowcasts were made for the 20 selected rainfall events. The accuracy was calculated with two verification methods: the Fraction Skill Score and the Critical Success Index. From these results, the deterministic nowcasts with S-PROG and benchmark EP are accurate for lead times up to an hour. The deterministic nowcasts give the most accurate results, compared to the EP and the probabilistic nowcasts. The probabilistic nowcasts show the least accurate results and are with a lead time of one hour only accurate for low rainfall intensities (1 mm/hr). Overall, the nowcasting results are showing that with CML rainfall estimations, Pysteps can make accurate nowcasts in Sri Lanka. To improve rainfall estimations further, RAINLINK could implement a selection of the parameters based on the link length for each CML. Further research is necessary to include the other parameters of RAINLINK in the calibration. Finally, to improve the nowcasts, even more, a combination of rainfall estimations from CMLs, rain gauges, and satellite data could be considered.

Contents

Preface i

Abstract ii

Acronyms v

List of Figures 1

List of Tables 3

1 Introduction 5

2 Commercial Microwave Links 7
- 2.1 Introduction . . . 7
- 2.2 RAINLINK. . . . 8
  - 2.2.1 Reference signal level . . . 9
  - 2.2.2 Correction received signal powers . . . 9
  - 2.2.3 Mean path-averaged rainfall intensity . . . 9
- 2.3 Errors with CML rainfall estimation . . . 9
  - 2.3.1 Wet and dry classification . . . 9
  - 2.3.2 Wet antenna . . . 10
  - 2.3.3 Sampling strategy . . . 10

3 Nowcasting 13
- 3.1 Introduction nowcasting . . . 13
- 3.2 Deterministic nowcasting . . . 14
- 3.3 Probabilistic nowcasting . . . 14
- 3.4 Uncertainties in nowcasting . . . 14
- 3.5 Pysteps . . . 15

4 Study area and data 17
- 4.1 Study area. . . . 17
- 4.2 Rain gauge data . . . 18
  - 4.2.1 Uncertainties of rain gauge measurements . . . 18
- 4.3 Commercial Microwave Link data. . . . 19
- 4.4 Satellite data . . . 20

5 Methods 21
- 5.1 Calibration of the parameters A_a and α . . . 21
  - 5.1.1 Selection of rain gauges and CMLs . . . 21
  - 5.1.2 Calibration and validation period . . . 23
  - 5.1.3 Calibration with statistical values . . . 24
  - 5.1.4 Rainfall mapping. . . . 25
- 5.2 Build a nowcast with Pysteps . . . 25
  - 5.2.1 Selection of rainfall events . . . 25
  - 5.2.2 Setup of nowcasts . . . 26
- 5.3 Verification of the nowcasts . . . 27
  - 5.3.1 Fraction Skill Score. . . . 27
  - 5.3.2 Critical Success Index . . . 28
- 5.4 Comparison with rain gauges . . . 28

6 Results 29
6.1 Calibration parameters A_a and α . . . 29
6.1.1 Optimized parameters for daily values . . . 29
6.1.2 Comparison default and optimized parameters . . . 31
6.1.3 Relation between path length and the parameters A_a and α . . . 32
6.1.4 Impact of selected links and radius size on calibration . . . 33
6.1.5 Pre-processing QPE for nowcasting . . . 33
6.2 Nowcasting rainfall events . . . 34
6.2.1 Selection of 20 rainfall events . . . 34
6.3 Verification nowcasts for the 20 events . . . 35
6.3.1 Fraction Skill Score . . . 35
6.3.2 Critical Success Index . . . 37
6.4 Comparison rain gauges vs. QPF . . . 38
6.5 Comparison GPM vs. QPF . . . 39
6.6 Comparison performance nowcasts Sri Lanka vs. the Netherlands . . . 40

7 Discussion 41
7.1 Commercial Microwave Links . . . 41
7.2 Nowcasting . . . 42

8 Conclusion 43

References 45

Appendices 49

A Verification CML rainfall estimations 51

B Relation parameters with other factors 55
B.1 Dependency on link frequency . . . 55
B.2 Dependency on climatic zones . . . 56

C Verification nowcasting 57

D Comparison with GPM (Satellite data) 59

Acronyms

AR Auto-regressive

CML Commercial Microwave Link

CSI Crtical Success Index

DARTS Dynamic and Adaptive Radar Tracking of Storms

EP Eulerian Persistance

FSS Fraction Skill Score

GPM Global Precipitation Measurement

LINDA Lagrangian INtegro-Difference equation model with Autoregression

MAPLE McGill Algorithm for Precipitation Nowcasting by Lagrangian Extrapolation

MNOs Mobile Network Operator

MS Matched Swath

NEM Northeast Monsoon

NMS Network Management System

NS Normal Swath

NSE Nash-Sutcliffe efficiency

NWP Numerical Weather Prediction

QPE Quantitative Precipitation Estimation

QPF Quantitative Precipitation Forecasting

S-PROG Spectral Prognosis

SBMcast String of Beads Model for Nowcasting

SSEPS Short-Space Ensemble Prediction System

STEPS Short-Term Ensemble Prediction System

SWM Southwest Monsoon

VET Variational Echo Tracking

List of Figures

2.1 Illustration of two base stations and a Commercial Microwave Link (CML) . . . 7
2.2 Coefficient a and exponent b for different frequencies (International Telecommunication Union, 2005) . . . 8

3.1 A sketch of the relation between the quality and the lead time of forecasts of two methods: nowcasting (extrapolation of observations) and Numerical Weather Predictions (NWP) (Browning, 1980). . . . 13

4.1 Elevation and climate zones maps of Sri Lanka . . . 17
4.2 Locations of (a) the daily and hourly rain gauges and (b) the Commercial Microwave Links used in this study in Sri Lanka . . . 18
4.3 Frequencies and path lengths of all the available links received from Dialog Sri Lanka . . . 20

5.1 Links of Colombo (Radius = 1km) and location of rain gauge Colombo. . . . 23
5.2 Daily rainfall values for calibration and validation period. Red line indicates the average rainfall intensity from the 12 rain gauges and the blue dots give the rainfall intensities for every selected rain gauge. Zero rainfall values have been included. . . . 23
5.3 Workflow of making precipitation nowcasts with Python (Pulkkinen et al., 2019) . . . 27

6.1 Calculated NSE and relative bias for every combination of A_a and α for rain gauge Colombo. Red dots indicate the value with the maximum NSE and the blue dot indicates the location of the default parameters. . . . 30
6.2 Selected links around the rain gauges Colombo and Mannar . . . 31
6.3 Daily rainfall depths of rain gauges compared to the CML rainfall estimations at the locations of the gauges after interpolation of (a) default and (b) optimized parameters in only the validation period (November 6 to December 31, 2019). . . . 32
6.4 Hourly rainfall depths of rain gauges compared to the CML rainfall estimations at the locations of the gauges of (a) default and (b) optimized parameters in the calibration and validation period (September 12 to December 31, 2019). . . . 32
6.5 Information of all the rain gauges of (a) the path length spread and (b) the path length corresponding to the two optimized parameters (A_a and α). . . . 33
6.6 Selected rain gauge for each link in Sri Lanka based on the location of the link. . . . 34
6.7 Rainfall event of November 30, 2019 at 6:00 pm with (a) the QPE (b) Pysteps-D nowcast with a lead time of 15 minutes, (c) QPF with a lead time of 30 minutes, and (d) QPF with a lead time of 1 hour. . . . 35
6.8 FSS averaged values from all rainfall events of (a) EP, (b) Pysteps-D, and (c) the ensemble mean of Pysteps-P. . . . 36
6.9 CSI values with threshold of 1 mm/hr from (a) Pysteps-D, (b) Pysteps-P and (c) EP averaged over the 20 rainfall events. The red outline indicates the calculated area. . . . 37
6.10 CSI value distribution from the 20 rainfall events for a threshold of 1 mm/hr for Pysteps-D, Pysteps-P and EP from averaged lead times (15, 30, 45, and 1 hour), where the values of the Pysteps-P ensemble members are shown separately. . . . 38
6.11 Comparison from 20 rainfall events of (a) Pysteps-D nowcast and (b) QPE with the hourly rain gauge intensities on the same pixel, where the QPF is the accumulated rainfall intensity of four nowcasts made with lead times of 15, 30, 45, and 1 hour for each hour. . . . 39
6.12 A rainfall event on October 17, 2019 at 16:05 with (left) Pysteps-D nowcast at 16:00 (lead time of 1 hour), (middle) QPE at 16:00 and (right) GPM satellite rainfall estimation at 16:05. . . . 39

A.1 Verification of CML based interpolation rainfall depths at the locations of the (a) 8 hourly stations and (b) 12 daily stations in the period between September 12 and December 31, 2019. Zero values are also included. 51
A.2 Validation period of cumulative rainfall depths of the CML estimations with default and optimized parameters and the rain gauge data of every rain gauge separate. 52
A.3 NSE values of all gauges for all combinations of A_a and α. 54

B.1 Information of all the rain gauges of (a) the frequency spread and (b) the frequency corresponding to the two optimized parameters (A_a and α). 55
B.2 Relation of the two calibrated parameters A_a and α with three different climate zones: dry, wet and arid. 56

C.1 Comparison of two events with different CSI scores: (a) November 30 17:00-20:00 and (b) December 3 16:00-19:00 in 2019. 57
C.2 Comparison QPE and QPF (Pysteps-D) rainfall intensities with the rain gauges for the event of November 30 17:00-20:00. 58

D.1 Comparison QPF with GPM satellite product for 10 rainfall events with a lead time of 1 hour. . . 59
D.1 Comparison of a Pysteps-D nowcasts with a lead time of 1 hour with GPM satellite product for 10 rainfall events. 60

List of Tables

2.1 List of RAINLINK functions (Overeem et al., 2016a) 8

3.1 List of the Pysteps modules (Pulkkinen et al., 2019). 15

4.1 Data availability of hourly rain gauges in Sri Lanka (from highest to lowest). 19
4.2 Rainfall events captured by satellite (GPM) data in the period from September 12 to December 31, 2019. 20

5.1 Variables in RAINLINK (Overeem et al., 2016a) 22
5.2 Selected rain gauges for calibration in Sri Lanka. It shows the number of unique links for every radius at the location of each rain gauge. In yellow, the selected number of links is highlighted. . 22
5.3 Total rainfall and number of days without rain for the calibration (12 September - 6 November 2019) and validation period (7 November - 31 December 2019). Zero values are included in the calibration. 24
5.4 The 20 selected rainfall events to nowcast, with the start and end of each event and the maximum rainfall intensity from the CML rainfall estimations in mm/hr. 25
5.5 Setup for deterministic (Pysteps-D) and probabilistic (Pysteps-P) nowcasts made with Pysteps. 26

6.1 Rain gauges with selected radius (km) and local optimal parameters A_a (dB) (values between 0 and 4.0 dB) and α (-) (values between 0 and 1.0). 30
6.2 CSI values for different thresholds for Pysteps-D, Pysteps-P, and EP for the average of nowcasts over lead times of 15, 30, 45, and 1 hour and 20 rainfall events. 37
6.3 FSS scores (-) for CML nowcasts from the Netherlands (values from (Imhoff et al., 2020a)) and Sri Lanka (Pysteps-D and EP) for thresholds of 1 and 5 mm/hr and for lead times of 15, 30, and 1 hour. 40

A.1 Comparison of verification values CV and relative bias between study of Overeem et al. (2021) of this thesis both with default parameters (September 12 and December 31, 2019). 51

1

Introduction

Rainfall, and in particular extreme rainfall, is responsible for many disasters, for example, inland floods and landslides. The consequences of this extreme rainfall can be damage to infrastructure, buildings, and sometimes even loss of life. Quantitative Precipitation Estimation (QPE) can be used as input for hydrologic models. With these models, it is possible to estimate when and where a flood or other disaster takes place. If only QPEs are available, the lead time of a flood is equal to the basin response time. A Quantitative Precipitation Forecasting (QPF) is needed to increase the lead time. This is necessary to take earlier measures and to possibly reduce damage further.

Numerical Weather Prediction (NWP) is one method for short to medium-term rainfall forecasting. However, for short lead times (up to 6 hours) the forecasts made with NWP are not sufficient (Pierce et al., 2012). Nowcasting is the process of statistically extrapolating recent rainfall observations with lead times to 6 hours. It is a possible solution for the short-term forecasting of rainfall. Pysteps is an open-source library in Python for nowcasting (Pulkkinen et al., 2019).

Until recently, the most important instrument for forecasting rainfall in a short time frame is the weather radar. Radar is a good instrument for monitoring rainfall patterns. It can detect extreme weather conditions at a high spatial and temporal resolution (Fabry, 2015). A major constraint for this instrument is that in developing countries access to radar systems is not present, mostly because acquiring, maintaining, and operating them is too expensive (Wang et al., 2017; Saltikoff et al., 2019). Therefore, other instruments have to be used to generate high spatial and temporal resolution QPEs and QPFs.

One potential instrument that is used for rainfall estimation is satellites. The biggest advantage of satellites is that they cover the whole earth. Developing countries can benefit from satellite rainfall estimations. Geostationary products have a high temporal resolution. However, forecasting with satellites gives a low spatial resolution and is therefore not useful for hydrologic models as mentioned above. Next to this, the rainfall estimates made with satellites still have large uncertainties (Kidd & Huffman, 2011). A third type of instrument used for rainfall estimation is rain gauges on the ground. Rain gauges only sample a single point in the area. A lot of rain gauges are needed to obtain the same spatial accuracy comparable to weather radar or satellite images. In developing countries, the coverage is usually too low and can therefore not be used to forecast rainfall in the short term.

The function of a Commercial Microwave Link (CML), sometimes referred to as a link, is to propagate a signal with telecommunication purposes from one antenna to the other. To retrieve rainfall rates, the CMLs can be used as a source for opportunistic sensing. It is called opportunistic because the infrastructure of the links is already present for other purposes (telecommunication). During a rainfall event, in the path between two towers (one receiving and one transmitting), the signal of the link is attenuated. This rain-induced attenuation of the signal can be converted into a path-averaged rainfall intensity. An advantage of this source is that in many parts of the world links are already in use, especially in cities. Various studies showed already that a link network can be used for rainfall estimation (Leijnse et al., 2008; Overeem et al., 2011) and even for short-term forecasting (Imhoff et al., 2020a).

Through a collaboration with the GSM Association (GSMA), we have been able to gain access to data from CML networks. One of those networks is from the Mobile Network Operator (MNOs) Dialog Sri Lanka. Dialog Sri Lanka can provide a dataset for 4 months in 2019. Next to this, the Sri Lanka Department of Meteorology is providing the rain gauge data. The focus of this thesis is therefore on the country Sri Lanka. Enabling the possibilities of predicting (extreme) rainfall could be of great value to the country of Sri Lanka. Sri Lanka has a relatively low rain gauge density and there is no weather radar available to use for any rainfall forecasting. The biggest economic sectors of Sri Lanka are agriculture, energy, and tourism. Agriculture is highly dependent on rainfall. Extreme rainfall has an impact on agriculture, but also on other parts of the country: floods in urban areas (Manawadu & Wijeratne, 2021), landslides (Perera et al., 2018), and energy received from hydropower (World Bank Group, 2021).

Some areas in Sri Lanka have a high link density, mainly where the larger cities are located. In Overeem et al. (2021) the application of CMLs is analyzed in Sri Lanka to retrieve rainfall. This thesis studies the possibilities of optimizing certain parameters used in calculating the QPEs for the country of Sri Lanka. The calculations of the rain rates with CMLs are performed with RAINLINK (Overeem et al., 2016a). RAINLINK is an open-source package that can calculate rainfall from the transmitted and received signal levels of CMLs. The package uses default parameters based on calibration for the Netherlands (Overeem et al., 2013).

This thesis studies the possibilities of using rainfall retrievals from CMLs for rainfall nowcasting in a tropical country. This thesis will focus on the open-source library Pysteps to generate nowcasts. From the problem statement, the main research question follows:

How accurate is rainfall nowcasting using Pysteps and Commercial Microwave Links in Sri Lanka?

The subquestions that will lead to answering the research question are:

1. How well do the CML rainfall estimates compare with those from rain gauges in Sri Lanka using the default parameter values from RAINLINK?
2. Will optimizing the wet antenna attenuation A_a and the coefficient α of the algorithm give more accurate results for rainfall estimation in Sri Lanka?
3. What are the preparation steps to apply Pysteps to CML-derived rainfall fields in Sri Lanka?
4. How accurate are the nowcasting results compared with the CML rainfall estimation in Sri Lanka?
5. How accurate are the nowcasting results compared with the data received from the rain gauges in Sri Lanka?
6. How do the nowcasting results in Sri Lanka compare with nowcasting with CMLs in the Netherlands?

To answer the research questions, the thesis is organized as follows. Chapter 2 provides background information about CMLs and how to use them to estimate the rainfall. Chapter 3 gives background information about nowcasting and Pysteps. The study area of Sri Lanka and the data used in this thesis are described in Chapter 4. Chapter 5 contains the methodology to answer the different subquestions. The results of the thesis are given in Chapter 6. Chapter 7 provides a detailed discussion of these results and Chapter 8 gives conclusions and recommendations for future studies.

2

Commercial Microwave Links

2.1. Introduction

A Commercial Microwave Link (CML) is the connection between two communication stations (mostly telephone towers). A link path is a path between those two stations. Signals for telecommunication purposes will travel along a link path from one station to another. This often is present in both directions. The MNOs collect the data of the signal to analyze the performance of the signal. This signal data can also be used in an opportunistic way to estimate the rainfall. Provider Dialog Sri Lanka collects every 15 minutes the minimum and maximum received signal levels (dB). Next to this, Dialog Sri Lanka provides for each link the frequency (GHz), the coordinates of the start and end of the link path, and the link path length (km).

When a rainfall event occurs in the area of the link path (between the two stations) the signal loses some of its strength. Raindrops have about the same size like the microwave signal (millimeters), which causes the signal to lose strength. Attenuation of the signal increases with more raindrops and with increasing raindrop sizes. There is a connection between the attenuation of the signal and the rainfall rate along the path between transmitter and receiver. In Figure 2.1 the attenuation of the signal caused by rainfall is illustrated.

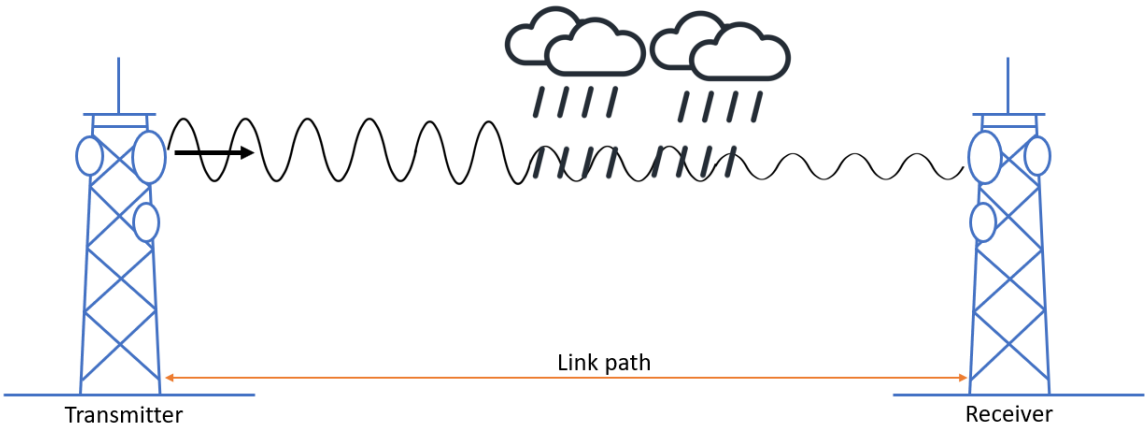


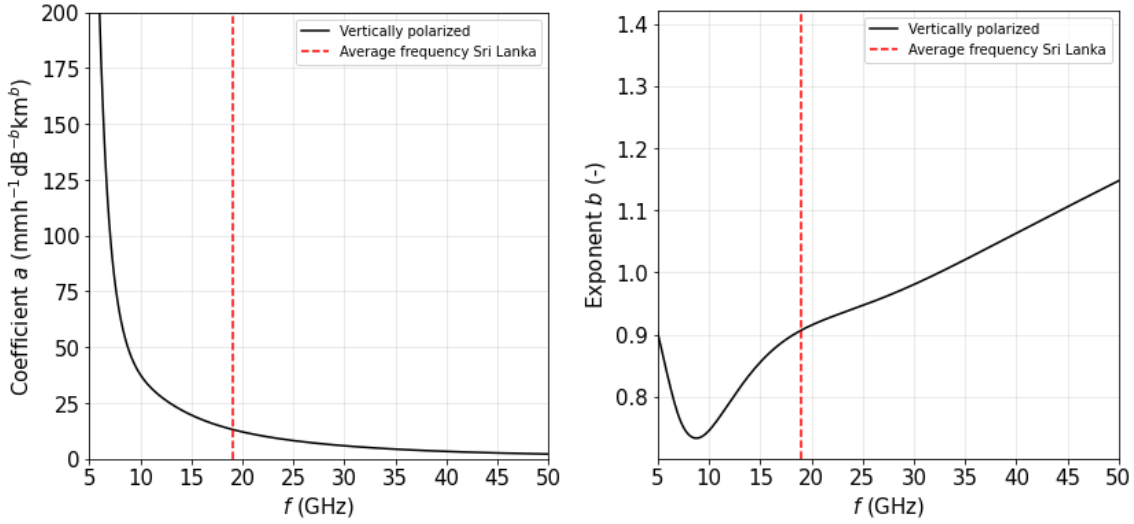
Figure 2.1: Illustration of two base stations and a Commercial Microwave Link (CML)

The attenuation of the microwave signal data, stored operationally in the Network Management System (NMS), can be used for rainfall estimates. The signal level for the dry weather conditions, called the reference signal level, must be determined to know how much the signal has changed. When rainfall events occur, the attenuation along the link path can be calculated. The rainfall intensity R (mmh^{-1}) can be calculated by using a power law (Atlas & Ulbrich, 1977):

$$R = ak^b \tag{2.1}$$

where k (dB km^{-1}) is the specific attenuation of a microwave signal with coefficient a (mmh^{-1}dB^{-b}km^{b}) and exponent b (-).

Both a and b depend on multiple factors. Those factors include the polarization, temperature, water phase, drop size distribution, drop shape, and canting angle distribution, but mainly on the frequency (Berne & Uijlenhoet, 2007; Leijnse et al., 2008). The specific attenuation model of International Telecommunication Union (2005) in Figure 2.2 shows the values for coefficient a and exponent b for different frequencies. It shows that the exponent b is close to 1 for frequencies close to 30 GHz, which is a frequency that is often used, for example in the Netherlands.



(a) Coeffcient a **(b)** Exponent b

Figure 2.2: Coefficient a and exponent b for different frequencies (International Telecommunication Union, 2005)

2.2. RAINLINK

In Overeem et al. (2016a) the open-source package called RAINLINK is introduced in scripting language "R". An overview of the RAINLINK functions is given in Table 2.1. Creating spatial rainfall maps from CMLs can be done in seven steps with RAINLINK: (1) Pre-processing, (2) Wet-Dry classification, (3) Determination of reference signal level, (4) Removal of outliers, (5) Correction signal powers, (6) Rainfall retrieval and (7) Interpolation. The default parameters of RAINLINK are based on a study of 12-day calibration in the Netherlands from Overeem et al. (2013).

Table 2.1: List of RAINLINK functions (Overeem et al., 2016a)

Step	Function	Subfunction	Description
1	`PreprocessingMinMaxRSL`	-	Preprocessing of linkdata
2	`WetDryNearbyLinkApMinMaxRSL`	-	Wet-dry classification with nearby link approach
3	`RefLevelMinMaxRSL`	-	Reference signal level determination
4	`OutlierFilterMinMaxRSL`	-	Remove outliers
5	`CorrectMinMaxRSL`	-	Correction of received signal powers
6	`RainRetrievalMinMaxRSL`	`MinMaxRSLToMeanR`	Compute mean path-averaged rainfall intensities Convert minimum and maximum to mean rainfall intensities
7	`Interpolation`	`IntpPathToPoint` `OrdinaryKriging`	Interpolate path-averaged rainfall intensities

2.2.1. Reference signal level

The next subsections describe how to determine the path-averaged rainfall intensities. First, the reference signal level must be determined, because the attenuation is the difference between the reference signal level and the received signal level:

$$P^C_{\text{min}} = \begin{cases} P_{\text{min}} & \text{if wet AND } P_{\text{min}} < P_{\text{ref}} \\ P_{\text{ref}} & \text{if dry OR } P_{\text{min}} \geq P_{\text{ref}} \end{cases} \tag{2.2}$$

$$P^C_{\text{max}} = \begin{cases} P_{\text{max}} & \text{if } P^C_{\text{min}} < P_{\text{ref}} \text{ AND } P_{\text{max}} < P_{\text{ref}} \\ P_{\text{ref}} & \text{if } P^C_{\text{min}} = P_{\text{ref}} \text{ OR } P_{\text{max}} \geq P_{\text{ref}} \end{cases} \tag{2.3}$$

where P^C_{min} is the corrected minimum signal power (dB), P^C_{max} is the corrected maximum signal power (dB) and P_{ref} is the reference signal level (dB).

2.2.2. Correction received signal powers

With the reference signal level, the corrected minimum and maximum attenuations are calculated:

$$\begin{aligned} A_{\text{min}} &= P_{\text{ref}} - P^C_{\text{max}} \\ A_{\text{max}} &= P_{\text{ref}} - P^C_{\text{min}} \end{aligned} \tag{2.4}$$

where A_{min} (dB) is the minimum rain-induced attenuation, A_{max} (dB) is the maximum rain-induced attenuation. The corrected minimum and maximum attenuations are used to calculate the minimum and maximum path-averaged rainfall intensities:

$$\langle R_{\text{min}} \rangle = a \left(\frac{A_{\text{min}} - A_{\text{a}}}{L} H\left(A_{\text{min}} - A_{\text{a}}\right) \right)^b \tag{2.5}$$

$$\langle R_{\text{max}} \rangle = a \left(\frac{A_{\text{max}} - A_{\text{a}}}{L} H\left(A_{\text{max}} - A_{\text{a}}\right) \right)^b \tag{2.6}$$

where R_{max} and R_{min} are the maximum and minimum path-averaged rainfall intensities, A_a (dB) is the wet antenna attenuation, L is the link path length and H is the Heaviside-function where $H = 0$ if the argument of H is less than zero, else $H = 1$.

2.2.3. Mean path-averaged rainfall intensity

With the minimum and maximum path-averaged rainfall intensities, the 15-minute mean path-averaged rainfall intensity is computed:

$$\langle R \rangle = \alpha \langle R_{\text{max}} \rangle + (1 - \alpha) \langle R_{\text{min}} \rangle \tag{2.7}$$

where R is the mean path-averaged rainfall intensity and α is the coefficient that determines the contributions of the minimum and maximum path-averaged rainfall intensities.

2.3. Errors with CML rainfall estimation

Various studies have investigated the main sources of error that can occur with estimating rainfall intensities using CMLs (Berne & Uijlenhoet, 2007; Leijnse et al., 2008; Rios Gaona et al., 2015; van Leth et al., 2018). Three important sources of error with CML rainfall estimates are signal fluctuations not related to rainfall, wet antenna attenuation, and sampling strategy of the CMLs. There are other uncertainties that one can think of with estimating rainfall with CMLs. For example, one can think of the errors and uncertainties of the $R - k$ power-law relation with coefficient a and exponent b. However, it has been shown that these errors are not dominant (Leijnse et al., 2008).

2.3.1. Wet and dry classification

Signal attenuation not only happens due to raindrops but also because of other causes unrelated to rainfall. One important cause of this type of error is dew formation on the antenna (Overeem et al., 2016b). Another cause for attenuation that all stations experience is atmospheric absorption. This depends on the factors of temperature, humidity, and air pressure (Upton et al., 2005). The error that follows from these causes is solved

by using a wet and dry classification before the calculation of the rainfall estimation.

There are two main approaches to classifying wet and dry periods. The first approach is called the link approach. The link approach is based on the reasoning that rainfall is correlated in space. The basic principles are based on the assumption that whenever it rains, all links within a certain area will experience rain-induced signal attenuation simultaneously. In RAINLINK a radius is chosen based on the spatial correlation of rainfall to. A link is selected and all links that are within the chosen radius with both endpoints of the selected link are selected as well. For each link, the attenuation and the specific attenuation are calculated. The mean values of the attenuations are calculated and with these values, each time interval is selected as either wet or dry. This is done by giving wet and dry specific thresholds of attenuation.

The second approach is called the radar approach. Radar data is used to identify if the conditions are wet or dry. This approach is only applicable if radar data is available in the studied area. It returns a classification of wet and dry periods for all the links that have at least three surrounding links selected within the chosen distance (Overeem et al., 2011).

2.3.2. Wet antenna

Wet antennas during rainfall can cause extra attenuation. During rainfall, a water film can be formed on the material of the antennas, which causes extra attenuation (Leijnse et al., 2008). The extra attenuation causes overestimation of the rainfall intensity and is referred to as the wet antenna attenuation (A_a). The study of Leijnse et al. (2008) showed that there is a significant wet antenna attenuation for all frequencies and its magnitude is dependent on the frequency. Correcting for the wet antenna attenuation is implemented in RAINLINK with a constant value (see Equations 2.5 and 2.6) of 2.3 dB. This value was estimated in the Netherlands for a 12-day calibration period (Overeem et al., 2013).

Multiple other studies have calculated the wet antenna attenuation for their study area. Minda and Nakamura (2005) calculated the wet antenna attenuation for one link with a frequency of 50-GHz and a path length of 0.82 km for 4 years in Japan. This resulted in a constant value for the wet antenna attenuation of 2.3 dB. In the study of Leijnse et al. (2007), a 27-GHz CML was used in the Netherlands with a path length of 5 km. The optimal value of A_a was 3.32 dB for a calibration period of 2 months (28 May 1999 - 23 July 1999).

In Overeem et al. (2011) the difference between radar and CML estimations is calculated for a range of values of A_a and α. The residual standard deviation and mean bias for combinations of A_a and α are calculated for low- and high-frequency classes (13-33 GHz and 38-39 GHz respectively). According to this study, both the low- and high-frequency classes have an optimal value for A_a of 1.30 dB for a 17-day dataset from 2009. In Pastorek et al. (2021) CMLs of 25-39 GHz frequencies and path lengths between 0.8 and 5.8 km were used to determine the wet antenna attenuation. The calibrated wet antenna attenuation was 3 dB from a dataset of 2 years (July 2014 - October 2016).

2.3.3. Sampling strategy

Signal powers are for example measured once every 15 minutes. The temporal variability of rainfall in 15 minutes can be large and this can cause errors. The temporal resolution is limited to how often the signal is measured and documented. Another cause of the error is that the MNOs record the received signal levels to the nearest 1 dB, or at best to the nearest 0.1 dB. which is causing rounding errors. When the path-integrated attenuation is smallest (usually with short links and with low frequencies) errors will be even larger, because the effect of rounding is also larger (Leijnse et al., 2008).

Also, the rainfall intensities calculated are path-averaged. This means that for each link the average rainfall intensity over the path length is calculated. The main uncertainty here is that the rainfall intensity most of the time is not the same over the entire path length.

Some datasets (including the dataset of Sri Lanka) contain only minimum and maximum received powers. The functions of RAINLINK apply to CML datasets that include minimum and maximum received power (P_{min} and P_{max}). The purpose of RAINLINK is to obtain in the end mean path-averaged rainfall intensities. Equation 2.7 showed that RAINLINK assumes that the mean path-averaged rainfall intensity can be calculated from a combination of R_{min} and R_{max}, where α gives determines their contributions. R_{min} is dominant

if α is closer to 0 and $R_{\max}$ is dominant if α is close to 1.

In two studies performed in the Netherlands an optimal value of α is chosen. In Overeem et al. (2011) the optimal value for α is found to be 0.334 for low-frequencies and 0.244 for high-frequencies. In Overeem et al. (2013) an optimal value for α of 0.33 is estimated, which is used by RAINLINK. In the algorithm of RAINLINK a single value for both parameters is chosen, and not for a value based on the frequency (as was done in Overeem et al. (2011)).

3

Nowcasting

3.1. Introduction nowcasting

Figure 3.1 shows a sketch of the relation between the lead time of the forecast and the skill for multiple forecasting methods. The lead time of a forecast is the time between the calculation of the forecast and the occurrence of the forecasting event. More recent studies still show similar results, but due to improvements in the Numerical Weather Prediction (NWP) models, the useful range of extrapolation-based forecasts is reduced to a few hours (Pierce et al., 2012).

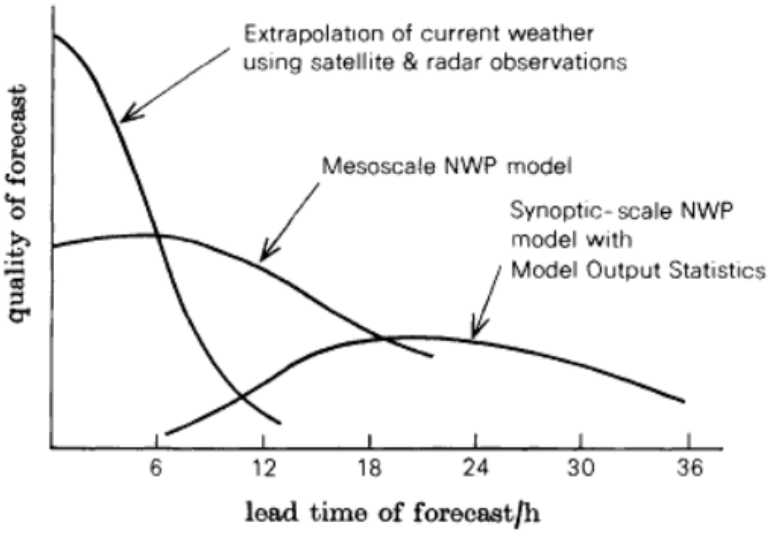


Figure 3.1: A sketch of the relation between the quality and the lead time of forecasts of two methods: nowcasting (extrapolation of observations) and Numerical Weather Predictions (NWP) (Browning, 1980).

For longer lead times, NWP models are used. These are physics-based models. The NWP models describe the most essential physical processes in the atmosphere and use them to predict the weather. The performance of NWP models with longer lead times is generally good. However, for shorter lead times, the NWP models do not provide sufficient forecast skills. One of the causes is that the computational limitations of NWP models only provide lead times of more than an hour. Next to this, the QPFs of NWP models exhibit relatively low accuracy in rainfall location, which makes them less suitable for the prediction of local rainfall (Pierce et al., 2012).

Nowcasting is the process of statistically extrapolating recent rainfall observations with lead times up to 6 hours (Pierce et al., 2012). It is shown that in particular for very short-term forecasting (0-3 hours ahead) the extrapolation-based forecasts (nowcasts) outperform these NWP models (Berenguer et al., 2012; Simonin et

al., 2017). Therefore, for short lead times, it is important to investigate the possibilities of nowcasting.

The first concept of rainfall forecasting with the method of extrapolating radar estimations is from the 1950s with the first proposal by Ligda (1951) and the early 1960s with the first application by Hilst and Russo (1960). From the first concepts of nowcasting, an important conclusion was already drawn. They found a correlation between the size of the precipitation features and predictability. Larger features live longer than smaller features (Pierce et al., 2012).

3.2. Deterministic nowcasting

Over time, several deterministic techniques and algorithms for precipitation nowcasting have been developed. The first group of techniques is called cell tracking algorithms. With an estimation of the advection velocity, a nowcast can be made (Pierce et al., 2012). An example of a cell tracking algorithm is TITAN (Dixon & Wiener, 1993). Cell tracking algorithms like TITAN are mainly used for thunderstorms and areas with severe convective weather conditions.

Next to cell tracking algorithms, there is another group of nowcasting techniques, namely field tracking algorithms. The difference between cell tracking and field tracking is that in cell tracking a spatially constant velocity is chosen while in the field tracking spatial and temporal changes in speed and direction are also taken into account (Pierce et al., 2012). According to a study of Germann and Zawadzki (2002), a semi-Lagrangian advection method is optimal.

Spectral Prognosis (S-PROG) is an example of a deterministic nowcasting model that is based on the conclusion from earlier studies that large features in a rain field evolve more slowly than small features (Seed, 2003). S-PROG consists of three main components:

1. The estimation of the field advection
2. The decomposition of the field into Fourier components and
3. An auto-regressive model for the scale-dependent Lagrangian evolution of the field

3.3. Probabilistic nowcasting

With probabilistic nowcasting, the uncertainties in the rainfall nowcast can be characterized by making ensembles (a set of forecasts). The nowcast can be displayed as the probability the rainfall intensity will exceed a certain threshold.

An example of a method for short-term ensemble rainfall nowcasts is Short-Term Ensemble Prediction System (STEPS) (Bowler et al., 2006). STEPS produces a defined number of ensemble members. The forecast cascade with a forecast rainfall pattern is calculated with the S-PROG scheme. In STEPS, each ensemble member has a different random cascade. The advection uncertainty is described by a field of velocities that contains a random component. With this method, each ensemble member represents a prediction to be obtained equally likely as the other ensemble members. The different ensemble members are obtained with different noise fields. With the ensemble members, uncertainties of the forecast are taken into account by adding spatially and temporally correlated noise (Bowler et al., 2006).

Next to STEPS, other probabilistic nowcast algorithms are the String of Beads Model for Nowcasting (SBMcast) (Berenguer et al., 2011) and the McGill Algorithm for Precipitation Nowcasting by Lagrangian Extrapolation (MAPLE) (Turner et al., 2004). The uncertainties in rainfall nowcast obtained by these probabilistic nowcast algorithms can also be used for the uncertainty in for example the discharge forecast (Heuvelink et al., 2020).

3.4. Uncertainties in nowcasting

Uncertainties in nowcasting are always present and these uncertainties can be categorized into two main sources: errors in input data of the nowcasts (i.e. the QPE) and model uncertainty. The uncertainty of quantitative precipitation estimations includes the spatial and temporal sampling errors of the measurements for the QPE. Also, the algorithm that transforms the signal into rainfall estimations is the main error source. A

more detailed explanation of errors and systematic biases in the rainfall amount related to QPE specifically obtained by CMLs can be found in Section 2.3.

The uncertainty of the model includes an assumption that is made by using persistence with the state of the atmosphere. But also the incorrect calculation of the field of motion, change in the motion, and change in the evolution of the precipitation fields can be present (Pierce et al., 2012; Foresti et al., 2019). Nowcasts are not based on physics and therefore it is expected that uncertainty increases with increasing lead times. The lead time for which a nowcast is still skillful is different for each nowcast, depending on the different uncertainties. It is difficult to distinguish the uncertainties between the error in the QPE and the error in the model itself (Germann et al., 2006).

3.5. Pysteps

The focus of Pysteps is on probabilistic nowcasting, but also deterministic nowcasting is an option. Pysteps allow for setups with different methods (the framework is modular). Pysteps is built based on the algorithms S-PROG and STEPS. Table 3.1 lists the 11 modules present in the Pysteps library, which all have different tasks.

Some of the modules have multiple methods that the user can choose from. The motion module (Table 3.1) consists of a couple of methods: a local Lucas-Kanade method (Lucas & Kanade, 1981), Variational Echo Tracking (VET) (Laroche & Zawadzki, 1994) and the Dynamic and Adaptive Radar Tracking of Storms (DARTS) algorithm (Ruzanski & Chandrasekar, 2012). Pulkkinen et al. (2019) showed that the three methods have almost no differences in terms of accuracy if only taking the precipitation field into account. Outside the precipitation fields, DARTS performs less compared to the other two methods. Furthermore, the computational requirements of VET are larger compared to Lucas-Kanade and DARTS.

The extrapolation module has the implementation of the semi-Lagrangian method (Germann & Zawadzki, 2002). This module also contains the Eulerian Persistence. The time-series module contains methods related to Auto-regressive (AR) models and correlation methods. The nowcasting module consists of the following methods: A simple advection-based extrapolation, Lagrangian INtegro-Difference equation model with Autoregression (LINDA), Lagrangian Probability, S-PROG, STEPS, and a localized form of STEPS called Short-Space Ensemble Prediction System (SSEPS) (Pulkkinen et al., 2019).

Apart from making the actual nowcast ensemble, Pysteps also contains a module for making statistical verification results. This contains forecast evaluation and skill scores for deterministic and probabilistic nowcasts.

Table 3.1: List of the Pysteps modules (Pulkkinen et al., 2019).

Module	Description
`io`	Reading radar composites and writing nowcast files
`motion`	Optical flow methods for motion field computation
`extrapolation`	Advection-based extrapolation
`timeseries`	Time series methods (e.g., AR models)
`noise`	Generation of stochastic noise to perturb precipitation and motion fields
`cascade`	Scale-based decomposition of precipitation fields
`nowcasts`	Implementation of nowcasting methods
`postprocessing`	Statistical post-processing of nowcasts
`verification`	Statistical verification of nowcasts and plotting the results
`visualization`	Plotting of precipitation and advection fields
`utils`	Miscellaneous utility functions

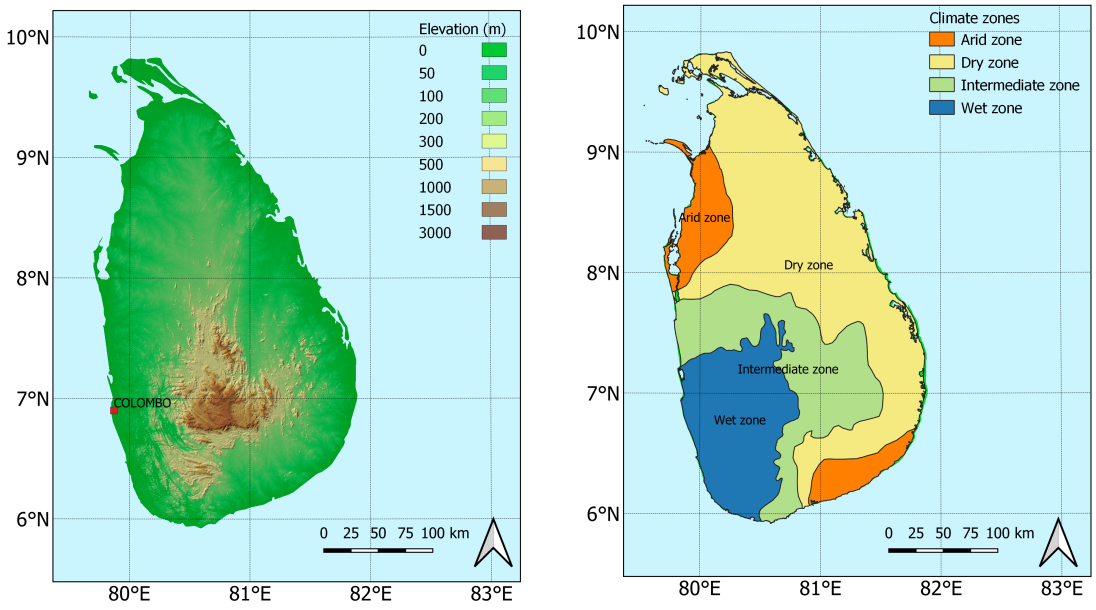
4

Study area and data

4.1. Study area

Sri Lanka is an island located in the Indian Ocean at the southeast tip of India with a landmass of 65,610 km^2. The capital of Sri Lanka is Colombo, with a population of 2.5 million inhabitants. The total population of Sri Lanka is 22.1 million inhabitants in 2021 (Department of Census and Statistics of Sri Lanka, 2021). Figure 4.1a shows the elevation map of Sri Lanka (NASA, 2013). The largest part of the island consists of plains up to 200 meters above sea level. The south-central part has mountains with peaks of 2,500 meters above sea level.

The location of Sri Lanka is in the tropical zone of the world, which gives little variation in temperature. The main variation in climate is the difference in the locations where rainfall events occur more often. Figure 4.1b shows the main climate zones in Sri Lanka (Harischandra et al., 2016; Sri Lanka Department of Meteorology, 2021). The climate in Sri Lanka is mainly influenced by the monsoon system in the Indian Ocean (Burt & Weerasinghe, 2014): Southwest Monsoon (SWM) from May to September and the Northeast Monsoon (NEM) from December to February. The annual rainfall varies from less than 900 mm in the Southeastern and Northwestern parts of the island to over 5,000 mm on the Western slopes of the mountains (Sri Lanka Department of Meteorology, 2021).



(a) Elevation (NASA, 2013)

(b) Climate zones (Harischandra et al., 2016)

Figure 4.1: Elevation and climate zones maps of Sri Lanka

4.2. Rain gauge data

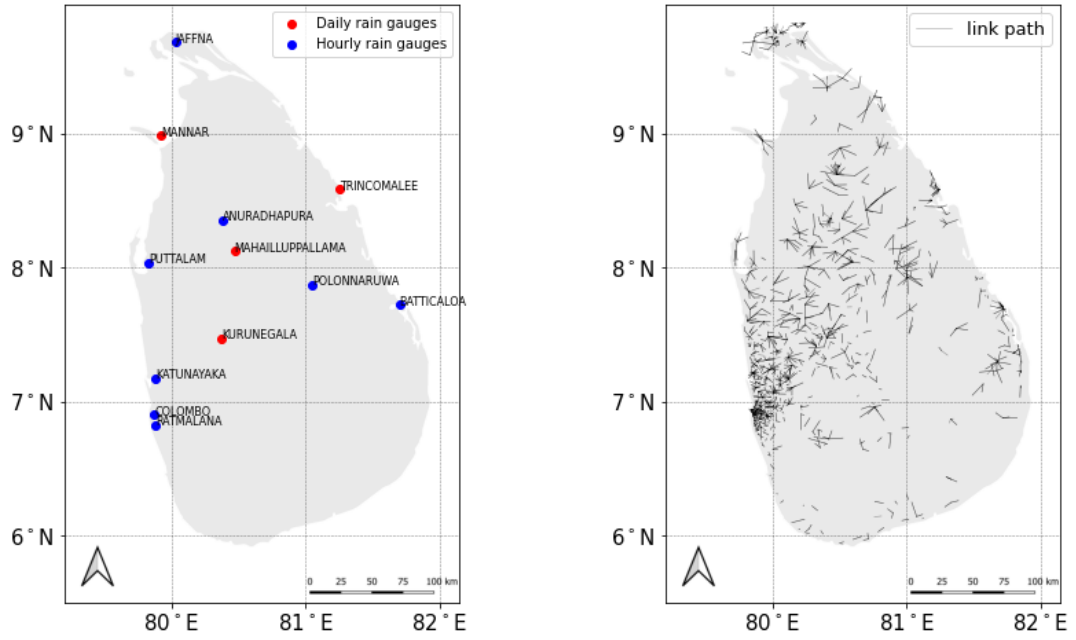
From the Sri Lanka Department of Meteorology data from 11 hourly and 12 daily rain gauges have been obtained. Figure 4.2a shows the locations of the gauge stations from the Sri Lanka Department of Meteorology that are used in this study. Other sections will explain why only these 8 hourly and 4 daily stations are used. The studied period is from September 12 to December 31, 2019.

Daily rainfall estimations from the rain gauges were using 3-hour rainfall depths. These were measured manually. These rainfall depths of 3 hours were accumulated to produce daily rainfall depths. The daily rainfall depths are given from 8:30 local time to 8:30 local time the next day. The measurements are obtained by emptying the gauge. Hourly rainfall depths were extracted from mechanical self-recording instruments (pluviographs). Hourly values start at the beginning of each hour (8:00, 9:00, etc).

4.2.1. Uncertainties of rain gauge measurements

Measuring the rainfall by emptying the gauge could be the cause of a measurement error. Another source of error from both daily and hourly gauge rainfall observations could be because of malfunctioning due to dirt (Sri Lanka Department of Meteorology, 2021).

For hourly rain gauges, the instrument needs to be manually read. Therefore, there is a chance of reading errors. The availability of hourly gauge data is lower compared to the daily gauge data due to mechanical problems and human errors (Overeem et al., 2021). For some days the hourly data values do not add up to the observed daily value. This can be explained by the errors of either the daily or hourly data observations.



(a) Rain gauges **(b)** Commercial Microwave Links

Figure 4.2: Locations of (a) the daily and hourly rain gauges and (b) the Commercial Microwave Links used in this study in Sri Lanka

The spatial differences between the rain gauge location and some of the CML locations could result in a difference in the timing of the rainfall event. If hourly values of rainfall intensities would be used to compare the rain gauges with the CML estimation, the differences could be very large due to this effect. If daily values are used, this effect would be much smaller, because the summation of all the rainfall intensities of that day is added. Therefore, daily rain gauge data is used to compare with the CML estimations and to calibrate the parameters A_a and α.

Hourly rain gauge data is used to compare the rainfall intensities of the rain gauges with the results of the nowcasts. Table 4.1 gives the data availability of hourly values for the 11 hourly stations. Some rain gauges have days or even weeks without hourly rainfall depths. To ensure that an hourly rain gauge can be used for comparison with the nowcast, a benchmark of 95% is added. Therefore, only the rain gauges with more than 95% data availability (8 of the 11 stations) for hourly rainfall depths are used: Anuradhapura, Batticaloa, Colombo, Jaffna, Katunayake, Polonnaruwa, Puttalam, and Rathmalana.

Table 4.1: Data availability of hourly rain gauges in Sri Lanka (from highest to lowest).

Rain Gauge	Longitude (°E)	Latitude (°N)	Number of hours without data	Availability [%]
Colombo	79.87	6.90	14	99.7
Anuradhapura	80.38	8.35	25	99.4
Puttalam	79.83	8.03	46	98.9
Katunayake	79.88	7.17	50	98.8
Polonnaruwa	81.05	7.87	50	98.8
Jaffna	80.03	9.68	85	97.9
Batticaloa	81.7	7.72	104	97.4
Rathmalana	79.88	6.82	216	94.6
Kurunegala	80.37	7.47	840	79.1
Mahailluppalama	80.47	8.12	855	78.8
Trincomalee	81.25	8.58	2262	43.8

4.3. Commercial Microwave Link data

The CML dataset with a total of 2560 unique links is received from the provider Dialog Sri Lanka. The dataset includes the following parameters: the frequency (GHz), minimum and maximum received signal levels (dB), the coordinates of begin and end of the link path, the link path length (km), and a unique ID for every link. The period of the dataset is from September 12 to December 31 in 2019 with an interval of 15 minutes.

Two days are missing in the CML dataset, October 3 and 4 in 2019. The period of the dataset must be adjusted to make a fair comparison between the rain gauge data and the link data. For daily comparison, both results are not taken into account from October 2 at 8:30 to October 5 at 8:15 local time in Sri Lanka (UTC+5:30). This removes a total of three days from the period. The last time interval in the dataset is December 31 at 23:45, 2019 local time in Sri Lanka (UTC+5:30). Because December 31 can not make a full 24-hour period (which would be until 8:15), this day is also not taken into account.

Figure 4.2b shows the locations of the link paths of the links from the dataset. Most links are located in, or close to cities. Figure 4.3 gives the frequency (GHz) and the link path length (km) of all the links in the dataset. Lower frequencies are less affected by attenuation. This is also why longer link paths often have lower frequencies. Longer link paths will have more attenuation due to the length of their path. In urban areas, more bandwidth is needed and therefore higher frequencies are required. The tropical climate in Sri Lanka causes significant attenuation by rainfall. Therefore, lower frequencies are more often used (compared to for example the Netherlands).

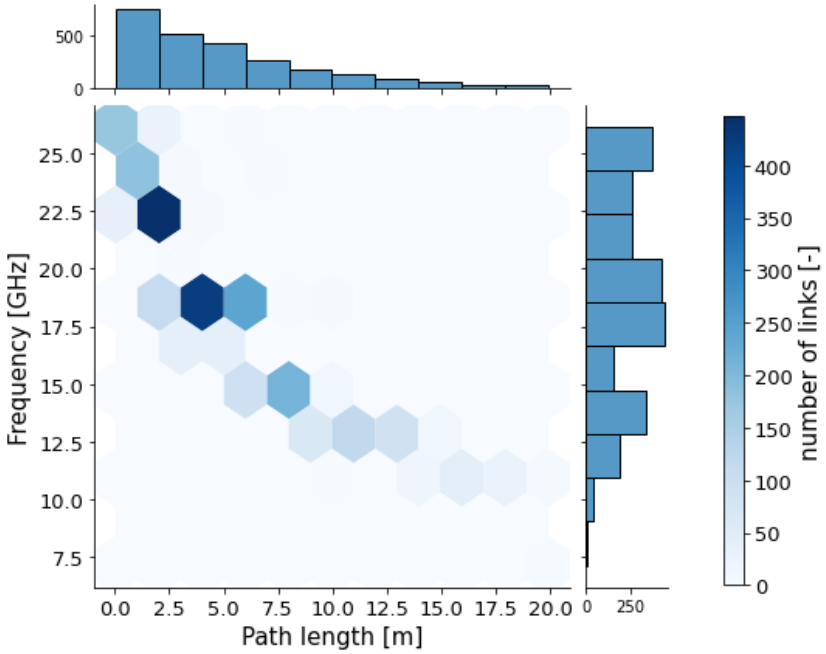


Figure 4.3: Frequencies and path lengths of all the available links received from Dialog Sri Lanka

4.4. Satellite data

Satellite data is used to be able to compare rainfall estimations of the satellite with the nowcasting results. For the satellite product, the Global Precipitation Measurement (GPM) combined precipitation product (GPM Combined Radar-Radiometer Precipitation Algorithm, version V06A) is used (Olson & Masunaga, 2018). A wide range of types of variables is included. In this study, the 'surfPrecipTotRate' is used. This rainfall intensity is provided in mm/hr.

The product contains two large datasets: Normal Swath (NS) and Matched Swath (MS). NS only used the Ku-band in comparison with MS which used a combination of Ku- and Ka-band. Ku-band is the lower frequency band (13.6 GHz), but this gives a wider swath. Ka-band is a higher frequency band (35.5 GHz) but results in a less wide swath. This study includes the NS because the swath is higher and therefore more events can be used for spatial comparison (Olson & Masunaga, 2018). Table 4.2 gives the date and time of these events. The scan time included is the local time in Sri Lanka (UTC+5:30). A total of 10 events from the period are included in this study.

Table 4.2: Rainfall events captured by satellite (GPM) data in the period from September 12 to December 31, 2019.

Event	Date (DD-MM-YYYY)	Local time (UTC+5:30)
1	28-09-2019	21:36
2	6-10-2019	19:20
3	9-10-2019	18:17
4	17-10-2019	16:05
5	28-10-2019	12:51
6	16-11-2019	18:51
7	24-11-2019	05:05
8	27-11-2019	15:32
9	5-12-2019	13:16
10	20-12-2019	21:22
11	31-12-2019	18:09

5

Methods

This chapter provides an overview of all the methods used in this study. First, the calibration of parameters in RAINLINK is explained in section 5.1. Next, the method for building a nowcast with Pysteps is described in Section 5.2. In Section 5.3 the verification methods used for the nowcasts are explained. Lastly, the method of the comparison with the rain gauges is given in Section 5.4.

5.1. Calibration of the parameters A_a and α

The default parameters in RAINLINK are optimized based on a study performed in the Netherlands (Overeem et al., 2013). Table 5.1 lists all the default parameters. Optimizing the parameters for the links in Sri Lanka could improve the rainfall estimation. This section describes how the parameters A_a and α of RAINLINK are calibrated, based on the CML and daily rain gauge datasets of Sri Lanka. In RAINLINK both parameters have a fixed value: 2.3 dB for A_a and 0.33 for coefficient α. This thesis also uses a fixed value for both parameters to compare the results with the default parameters. Section 2.2.1 showed that rainfall intensity can be calculated with Equations 2.5, 2.6 and 2.7. In the rainfall retrieval step of RAINLINK 4 parameters are needed for input: a, b, A_a and α. Parameters a and b are dependent on the frequency (Figure 2.2). The other two important parameters, the wet antenna attenuation A_a and coefficient α, are fixed parameters in RAINLINK.

The wet antenna attenuation A_a (dB) from Equations 2.5 and 2.6 is a correction for the bias resulting from a wet antenna. Coefficient α is used in Equation 2.7. This parameter is a fraction that is used to calculate the mean path-averaged rainfall intensity. The fraction is how much of the minimum and maximum path-averaged rainfall (R_{min} and R_{max}) are used. Both parameters influence the size of the rainfall intensity significantly. However, these two parameters are dependent on rainfall intensity, time variability, wet antennas, and other factors. It is hypothesized that the rainfall estimations in Sri Lanka could significantly improve if the parameters A_a and α are calibrated for an area with different climate and conditions.

5.1.1. Selection of rain gauges and CMLs

The calibration of the parameters is the comparison between the daily values of the rain gauges and the sum of all the 15-minute rainfall estimations of the CMLs of the day. For the calibration of the parameters, rain gauges have to be selected. Not all of the available rain gauges are used for the calibration because not all rain gauges have enough links near the location of the rain gauge. This is relevant because a large distance between the rain gauge location and the link will result in a difference in the timing and intensity of the rainfall event.

Only links within a 10 km radius are selected. Outside the 10 km, larger uncertainties are present when the rainfall estimations of the links are compared with the rain gauge data. If for calibration less than 15 links would be used, there is a larger possibility of having a biased outcome of parameters. Intermediate results showed that for some estimations a minimum of 15 links was needed (for some less). Other values could be chosen for the 10 km radius and the minimum of 15 links. However, for simplification and to have an as equal comparison as possible, the decision is made that for every rain gauge at least 15 links are selected.

Table 5.1: Variables in RAINLINK (Overeem et al., 2016a)

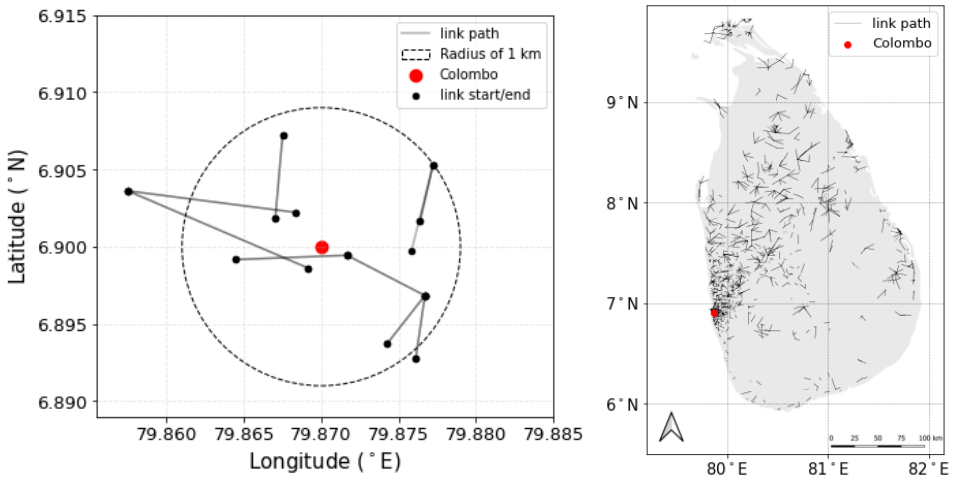
Variable description:	Symbol and unit	Value	Dependent on:
Wet-dry classification			
Radius	r (km)	15	Spatial correlation of rainfall
Minimum number of available (surrounding) links		3	
Number of previous hours over which max(P_{min}) is to be computed		24	
Minimum number of hours needed to compute max(P_{min})		6	
Threshold	median(ΔP_L) (dB km^{-1})	-0.7	Spatial correlation of rainfall
Threshold	median(ΔP) (dB)	-1.4	Spatial correlation of rainfall
Threshold (step 8 in Appendix C in AMT paper)	- (dB)	2	
Reference signal level			
Period over which reference level is to be determined	- (h)	24	
Minimum number of hours that should be dry in preceding period	- (h)	2.5	
Outlier filter			
Outlier filter threshold	F_t (dB km^{-1} h)	-32.5	Malfunctioning of links
Rainfall retrieval			
Wet antenna attenuation	A_a (dB)	2.3	Rainfall intensity, number of wet antennas, antenna cover
Coefficient	α	0.33	Time variability of rainfall
Coefficient of R-k power law	a (mm h^{-1}dB^{-b}km^b)	3.4-25.0	Drop size distribution, frequency
Exponent of R-k power law	b (-)	0.81-1.06	Drop size distribution, frequency

The rain gauges that are selected are those gauge stations that could meet the condition that a minimum of 15 Commercial Microwave Links must be located within a radius of 10 km around the location of the rain gauge. A CML is in this radius if the middle of the link path is within this 10 km. The radius that is selected is the smallest radius that contains at least 15 unique links in the area. Table 5.2 shows the rain gauges that have at least 15 links in their area within a 10 km radius. It also shows the number of unique links for every radius between 1 and 10 km that is available for each rain gauge. The chosen number of links for each rain gauge is highlighted.

The names of the selected rain gauges are Anuradhapura, Batticaloa, Colombo, Jaffna, Katunayake, Kurunegala, Mahailluppalama, Mannar, Polonnaruwa, Puttalam, Rathmalana, and Trincomalee. For example, Colombo has 23 links available within a 1 km radius of the rain gauge location. These 23 links are selected in the subset of the links and used for local calibration in Colombo (Figure 5.1a) Also, the location of the rain gauge Colombo in Sri Lanka is given in Figure 5.1b.

Table 5.2: Selected rain gauges for calibration in Sri Lanka. It shows the number of unique links for every radius at the location of each rain gauge. In yellow, the selected number of links is highlighted.

Radius (km)	Anur	Batt	Colo	Jaff	Katu	Kuru	Maha	Mann	Polo	Putt	Rath	Trin
1	0	6	23	6	2	4	0	0	0	0	4	0
2	1	8	68	12	6	6	0	2	0	1	13	3
3	1	18	128	24	12	10	2	9	0	3	31	10
4	5	20	183	34	19	14	2	9	4	5	42	10
5	16	22	241	34	36	22	6	9	8	7	74	24
6	21	25	330	38	48	30	8	9	10	7	78	30
7	26	29	365	45	64	30	10	15	10	15	107	30
8	28	29	403	47	79	42	10	15	14	17	128	30
9	29	32	430	53	93	47	14	21	14	17	170	32
10	29	33	464	61	105	57	16	23	22	19	200	34



(a) Links Colombo (Radius = 1km) **(b)** Links Sri Lanka

Figure 5.1: Links of Colombo (Radius = 1km) and location of rain gauge Colombo.

5.1.2. Calibration and validation period

For calibration of the parameters A_a and α, 50% of the 3.5 months of the dataset is used. The other 50% is used for validation. The calibration period is 55 days long and runs from September 12 2019 at 8:30 am until November 6 2019 at 8:30 am. There are three missing days for which the rainfall estimation is excluded from the period: October 2, 3, and 4 of 2019. The validation period is also 55 days and this period is from November 6, 2019, at 8:30 am to December 31, 2019, at 8:30 am. Figure 5.2 illustrates the daily rainfall intensities of both periods for all selected rain gauges. The total rainfall for each rain gauge of the calibration and validation period is given in Table 5.3. The table also contains the number of days without rainfall for both periods for every rain gauge. Days without rain are included in the calibration and validation. It shows that more days without rain are included in the validation period compared to the calibration period. Also, for most rain gauges the total rainfall of the validation period is less compared to the calibration period.

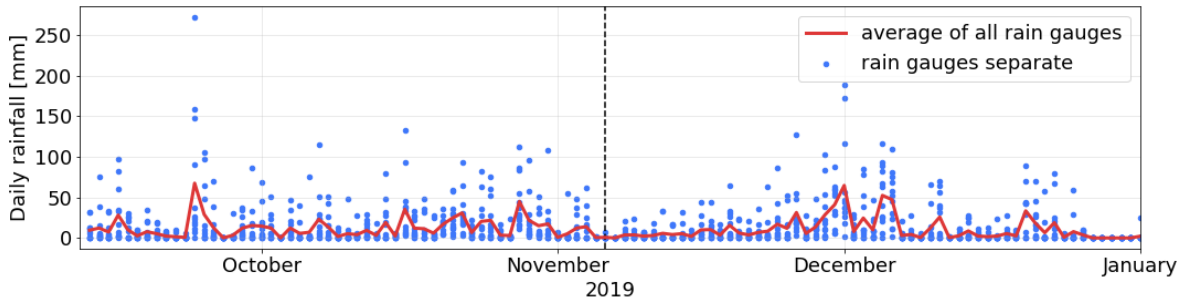


Figure 5.2: Daily rainfall values for calibration and validation period. Red line indicates the average rainfall intensity from the 12 rain gauges and the blue dots give the rainfall intensities for every selected rain gauge. Zero rainfall values have been included.

Table 5.3: Total rainfall and number of days without rain for the calibration (12 September - 6 November 2019) and validation period (7 November - 31 December 2019). Zero values are included in the calibration.

Rain Gauges	Calibration	Validation	Calibration	Validation
	Total rainfall [mm]		Number of days without rain [-]	
Anuradhapura	338	641	20	17
Batticaloa	619	1030	24	17
Colombo	925	611	8	23
Jaffna	676	459	18	18
Katunayake	959	395	7	28
Kurunegala	590	477	8	31
Mahailluppalama	710	593	17	21
Mannar	622	329	21	27
Polonnaruwa	442	905	24	15
Puttalam	895	442	14	27
Ratmalana	1083	580	9	18
Trincomalee	348	862	25	18

5.1.3. Calibration with statistical values

For every rain gauge, the first five steps from RAINLINK are followed once (these steps are discussed in Chapter 2) with the calibration subset. From these steps, the corrected received signal powers are calculated. These values are the same for every value of parameters A_a and α.

The calibration of parameters A_a and α use a range of values for both parameters. The wet antenna attenuation A_a is taken to range between 0 and 5.0 dB with steps of 0.5 dB. Coefficient α is taken to range between 0 and 1 with steps of 0.05. With step 6 of RAINLINK, the rainfall intensities are calculated. This step uses the wet antenna attenuation A_a and coefficient α. For every combination of the two parameters, the rainfall estimations are calculated every 15 minutes of the calibration period. These rainfall estimations for every 15 minutes are transformed into daily rainfall estimations from the links. These daily rainfall estimations are compared with the daily rainfall intensities from the rain gauges.

To assess which combination of parameter values is optimal two statistical values are calculated. The rainfall estimations are compared with the observations of the selected rain gauge. The Nash-Sutcliffe efficiency (NSE) is computed:

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N} \left(Q_{m,i} - Q_{o,i}\right)^2}{\sum_{i=1}^{N} \left(Q_{o,i} - \overline{Q_{o,i}}\right)^2} \tag{5.1}$$

where are Q_m (mmh^{-1}) the rainfall estimations from RAINLINK and Q_o (mmh^{-1}) the rainfall observations of the selected rain gauge.

When NSE = 1, it is a perfect fit between the rainfall estimations and the rain gauge. With NSE = 0 the rainfall estimations will give an equally good fit as the mean of the rain gauge observations.

For each combination of the two parameters the NSE value is calculated from the calibration period. The optimal combinations of the two parameters are the parameters that give the highest NSE value. For each rain gauge, one combination of parameters gives the highest NSE value and is chosen for the links close to that rain gauge. The optimal value of parameter A_a will have an accuracy of 0.5 dB and an accuracy of 0.05 for parameter α.

Next to the NSE, also the relative bias is computed:

$$\text{Relative Bias } = \frac{(\overline{Q}_m - \overline{Q}_0)}{\overline{Q}_0} \times 100\% \tag{5.2}$$

The relative bias is chosen to give insight into systematic errors that could be shown in the results. The relative bias will not be the decisive factor in determining the optimal parameters. It will help gain more insight into the performance of the rainfall estimations of all the different parameter combinations in comparison with the rain gauge data.

5.1.4. Rainfall mapping

After the calibration of the two parameters A_a and α for the links surrounding each rain gauge, the rainfall intensities can be calculated for all links. The two parameters A_a and α are used in step 6 of the RAINLINK algorithm as mentioned before. For each link, a fixed set from one of the 12 rain gauges needs to be selected. Thiessen polygons, polygons that represent the area of a point (rain gauge) where any location within that zone is closest to that point (rain gauge), are constructed around the 12 rain gauge locations. If the location of the middle of a link is within the polygon, those fixed values of A_a and α are assigned to that link. Once this is done for all the links, the rainfall intensity R (mm/hr) for all links at all time steps is calculated.

Using the calculated rainfall intensity for all links interpolated rainfall maps are constructed. RAINLINK uses the rainfall intensities from step 6 and interpolates with Ordinary Kriging in step 7. The result is interpolated rainfall estimations on the grid of Sri Lanka for every 15 minutes of the data period.

5.2. Build a nowcast with Pysteps

5.2.1. Selection of rainfall events

A total of 20 rainfall events are selected to assess the accuracy of the nowcasts. At least 4 hours are necessary for comparison with nowcasts of rainfall events of 3 hours with Pysteps. For the algorithm, three time steps with observations before the first time step of the rainfall event are used to nowcast. Next to this, the selected rainfall events have a wide spread of rainfall and/or high-intensity rainfall during the rainfall event. To check if the nowcasts have accurate results with higher intensity thresholds, the rainfall events that are analyzed should contain at least some high-intensity rainfall. To calculate the accuracy of the nowcast in different spatial scales, it is important to have multiple events with large spatial scales.

Table 5.4: The 20 selected rainfall events to nowcast, with the start and end of each event and the maximum rainfall intensity from the CML rainfall estimations in mm/hr.

Rainfall event	Date event	Start event	End event	Max. rainfall intensity [mm/hr]
1	23-9-2019	14:00	17:00	102
2	29-9-2019	20:00	23:00	74
3	30-9-2019	18:00	21:00	87
4	01-10-2019	17:00	20:00	131
5	06-10-2019	15:00	18:00	123
6	12-10-2019	17:00	20:00	138
7	13-10-2019	17:00	20:00	157
8	15-10-2019	15:00	18:00	162
9	16-10-2019	15:00	18:00	130
10	17-10-2019	17:00	20:00	141
11	27-10-2019	21:00	00:00	85
12	16-11-2019	17:00	20:00	123
13	28-11-2019	15:00	18:00	102
14	29-11-2019	21:00	00:00	92
15	30-11-2019	17:00	20:00	104
16	01-12-2019	18:00	21:00	100
17	03-12-2019	16:00	19:00	137
18	04-12-2019	14:00	17:00	141
19	05-12-2019	01:00	04:00	132
20	10-12-2019	11:00	14:00	59

The period of the data available for this study is between September 13 and December 31, 2019. One of the rainfall events in this period is from November 30 to December 11 in 2019. Due to this rainfall event, over 150,000 people were affected (Disaster Management Centre Sri Lanka, 2019). The start of this rainfall event, on November 30, will be used as an example for the nowcasting of the rainfall. Table 5.4 shows the 20 selected rainfall events that are used for the calculations of the nowcasts.

5.2.2. Setup of nowcasts

To build a nowcast in Python with Pysteps some essential settings need to be determined. These settings can be categorized into verification settings, forecast settings, the experiment set-up, and conditional parameters. The last step is the set-up of the nowcast and the output. A workflow scheme of a nowcast using Pysteps is illustrated in Figure 5.3 (Pulkkinen et al., 2019).

To nowcast the rainfall in Sri Lanka the open-source library Pysteps (version 1.5.1) is used in the programming language Python. The first step for nowcasting with QPEs received from Commercial Microwave Links and RAINLINK is to pre-process the data into the correct format. The io module from Pysteps (Table 3.1) has methods for reading radar-based composites only, so not for CMLs-based composites. Imhoff et al. (2020a) used CMLs from the Netherlands and calculated with RAINLINK rainfall to implement in Pysteps. They added an extra importer that allowed (adjusted) CML composites to import into Python to be used for Pysteps. The same importer is added to the io module in this study. After this, the pre-processed QPEs can be used for deterministic and probabilistic nowcasting.

The Eulerian Persistance (EP), also called the 'poor man' approach to nowcasting, is a method that returns the same sequence as the observation of the same initial field, and no extrapolation is applied to this method. In other words, QPF is equal to QPE_{t-1}. This is used as a benchmark to access if Pysteps-D and Pysteps-P are more skilled compared to EP.

Table 5.5 shows the setup for the deterministic nowcasts, Pysteps-D, made with the S-PROG algorithm. Furthermore, the Lucas-Kanade optical flow method is chosen because the combination of performance and computational requirements is the best for this method (see also Section 3.5). The semi-Lagrangian advection method is used in Pysteps as an extrapolation method. S-PROG is also used as a masking method and an AR model of order 2 is used.

Table 5.5 also gives the setup for the probabilistic nowcast, Pysteps-P. In this nowcast, the STEPS nowcasting algorithm is used (Bowler et al., 2006). For the masking method in Pysteps-P, a size of 20 ensemble members is chosen. The recommended ensemble size from Pulkkinen et al. (2019) was more than 24 members. Imhoff et al. (2020b) suggested that there are only marginal improvements between 24 and 48 members for low rainfall intensities, and thus 20 ensemble members are sufficient for low rainfall intensities. For high rainfall intensities improvements in the nowcasts could become visible. But the computational time becomes too long with 50 or 100 ensemble members. Therefore 20 ensemble members are still chosen for this study.

Table 5.5: Setup for deterministic (Pysteps-D) and probabilistic (Pysteps-P) nowcasts made with Pysteps.

	Pysteps-D	**Pysteps-P**
Nowcast	S-PROG	STEPS
Optical flow	Lucas-Kanade	Lucas-Kanade
Extrapolation	semi-Lagrangian	semi-Lagrangian
Cascade levels	6	8
Order of the AR(p) model	2	2
Precipitation threshold (mm/hr)	0.1	0.1
Mask method	S-PROG	incremental
Ensemble size	1	20
Seed number	24	24

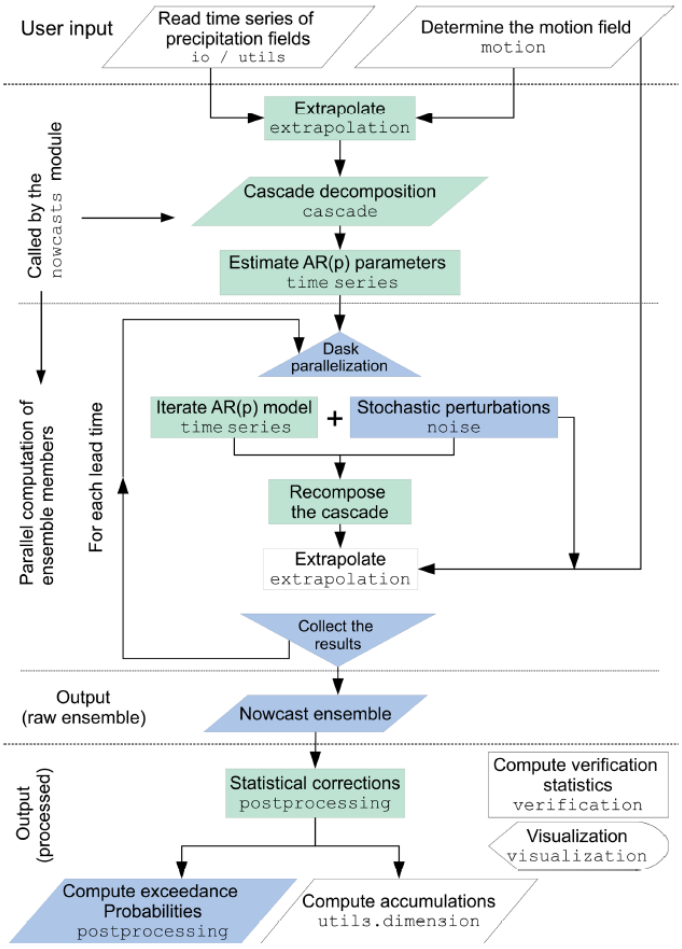


Figure 5.3: Workflow of making precipitation nowcasts with Python (Pulkkinen et al., 2019)

5.3. Verification of the nowcasts

For the assessment of the nowcasting skills, two verification methods are used in this study to assess the accuracy and skill of the nowcasts: The Fraction Skill Score (FSS) and the Crtical Success Index (CSI). These verification scores are calculated for EP, Pysteps-D, and Pysteps-P.

5.3.1. Fraction Skill Score

The Fraction Skill Score answers the question from which spatial scale the nowcast is skillful (Roberts & Lean, 2008). The range of FSS is from 0 to 1. A nowcast with an FSS score of zero has no skill and with a score of 1 has a perfect skill. The FSS is given by:

$$\mathrm{FSS}(n) = \frac{\mathrm{MSE}_{(n)} - \mathrm{MSE}_{(n)\mathrm{ref}}}{\mathrm{MSE}_{(n)\mathrm{perfect}} - \mathrm{MSE}_{(n)\mathrm{ref}}} = 1 - \frac{\mathrm{MSE}_{(n)}}{\mathrm{MSE}_{(n)\mathrm{ref}}} \tag{5.3}$$

where n is the neighborhood length, $\mathrm{MSE}_{(n)\mathrm{perfect}}$ is equal to zero, because it is for a perfect forecast and the MSE_n is the mean squared error of the fractions. The reference MSE represents the largest mean squared error possible and is given by:

$$\mathrm{MSE}_{(n)\mathrm{ref}} = \frac{1}{N_x N_y}\left[\sum_{i=1}^{N_x}\sum_{j=1}^{N_y} O^2_{(n)i,j} + \sum_{i=1}^{N_x}\sum_{j=1}^{N_y} M^2_{(n)i,j}\right] \tag{5.4}$$

where $O^2_{i,j}$ is the observed fraction and $M^2_{i,j}$ is the forecast fraction for every grid cell that exceeds a given

threshold of rainfall intensity, and where N_x and N_y are the number of pixels in x and y direction of the field. A nowcast is skillful if FSS is larger than $\text{FSS}_{\text{uniform}}$:

$$\text{FSS}_{\text{uniform}} = 0.5 + \frac{f_0}{2} \tag{5.5}$$

where f_0 is the fraction of observed points exceeding the threshold over the domain.

With increasing scale, the skill of the forecast will increase. With FSS the smallest scale of a forecast can be calculated for which the forecast is still skillful. Or, from a different perspective, the longest lead time for every nowcast for a certain threshold can be calculated. Whether or not the nowcast is skillful and also useful depends on the requirements of the scale of the nowcast the model needs to use. The FSS is calculated for five different thresholds (1,2, 5, 10, 20 mmh^{-1}), for spatial scales from 2 to 100 km, and for three different lead times (15, 30, 45, and 1 hour).

5.3.2. Critical Success Index

The Critical Success Index (Schaefer, 1990) is another way to get insight into the accuracy of the nowcast. This value is used often for nowcast research and can therefore be compared to other literature. The CSI gives insight into how well a nowcast corresponds with the observations. The CSI is the number of pixels in the nowcast that correctly corresponds with the observations above a certain threshold divided by the number of hits, misses, and false alarms linked to the observation:

$$CSI = \frac{N_{hits}}{N_{hits} + N_{misses} + N_{false}} \tag{5.6}$$

The CSI is calculated for three separate intensity thresholds (1, 5, and 10 mm/hr) and averaged over four lead times (15, 30, 45, and 1 hour) and all 20 rain events. For Pysteps-P the mean of the CSI of the 20 ensemble members is calculated. The CSI is calculated as a single number as well as a number for each of the four lead times. It is also displayed on the map of Sri Lanka to illustrate any spatial differences in accuracy.

5.4. Comparison with rain gauges

For the comparison of the nowcasts with rain gauges hourly rain gauge data is used. Hourly rain gauges are selected based on data availability (Table 4.1). For example, suppose a rainfall event is analyzed from 14:00 to 17:00. For every hour the rain gauges have rainfall depths available. For this event a nowcast is made at 13:45 for (14:00-15:00), at 14:45 for (15:00-16:00) and at 15:45 for (16:00-17:00). Each nowcast consists of nowcasts with 4 different lead times (15, 30, 45, and 1 hour), and these rainfall depths are combined to calculate the overall prediction of that hour. This rainfall depth is then compared at the location of the rain gauges with the hourly data of the selected rain gauges. For the 20 rainfall events, the rain gauges are compared with the nowcasts. All values are combined. The NSE and relative bias are calculated and these values are compared with the NSE and relative bias of the QPE for the same events.

6

Results

This chapter presents the results of this thesis. The chapter starts with the calibration of the parameters A_a and α in Section 6.1. With the calibrated parameters nowcasts are made for 20 rainfall events in Section 6.2. In Section 6.3 the results of the verification of the nowcasts are presented. A comparison of the nowcasts with hourly values of the rain gauges is presented in Section 6.4 and with the GPM data in Section 6.5. In Section 6.6 the nowcast performance of Sri Lanka is compared with the Netherlands.

6.1. Calibration parameters A_a and α

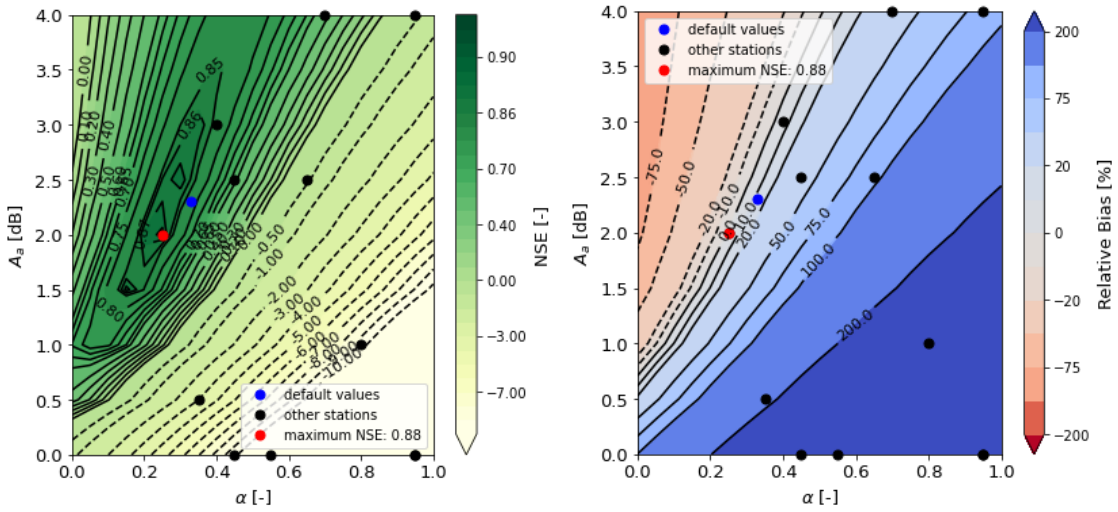
This section gives the results of the calibration of the parameters A_a and α. It also provides a comparison between the results of the default parameters and the optimal parameters. As an example, the result of one of the rain gauges, Colombo, is given. The results of the calibration of the parameters for all rain gauges are summarized in Table 6.1.

6.1.1. Optimized parameters for daily values

This section shows the results from optimizing the parameters of the 12 selected daily rain gauges. The parameters are calibrated by comparing daily rainfall accumulations of the rain gauges with accumulated daily rainfall totals of the links. The results from the links nearby the rain gauge Colombo are given as an example of the results of all the rain gauges.

Figure 6.1a shows for every combination of parameters A_a and α the calculated NSE (-) in Colombo. The red dot indicates the location of the maximum NSE score, which is in this case 0.89 at $A_a = 2$ and $\alpha = 0.25$. The blue dot indicates the NSE value of the default parameters ($A_a = 2.3$ and $\alpha = 0.33$) with a score of NSE = 0.86. The black dots indicate the optimal parameters from the calibration of other rain gauges. Results of the NSE values for all participating rain gauges are illustrated in Appendix A (Figure A.3).

Figure 6.1b shows for every combination of A_a and α the calculated Relative Bias (%) in Colombo. The red dot indicates the location of the maximum NSE and this combination gives a relative bias of 0.9%. The blue dot indicates the relative bias of the default parameters ($A_a = 2.3$ and $\alpha = 0.33$) with a relative bias of 12.6%. The relative bias shows the impact of the combination of the two parameters on the over-or underestimation of the rainfall depths in comparison to the rain gauges.



(a) NSE **(b)** Relative Bias

Figure 6.1: Calculated NSE and relative bias for every combination of A_a and α for rain gauge Colombo. Red dots indicate the value with the maximum NSE and the blue dot indicates the location of the default parameters.

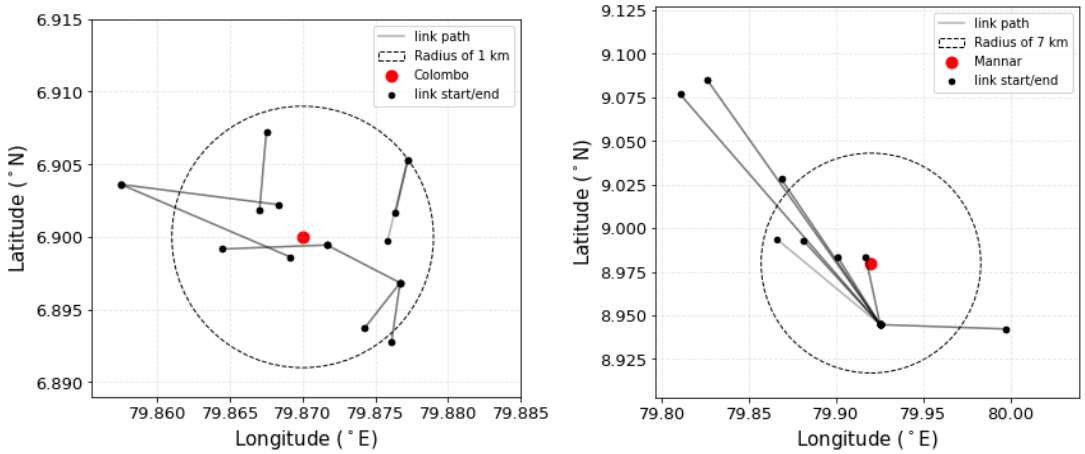
Parameters A_a and α are chosen at the location of NSE = NSE_{max}. Table 6.1 lists the optimal parameters for all the rain gauges. It also gives the radius and number of the selected links and the NSE_{max}. It shows that Mahailluppalama and Mannar have much lower maximum NSE values compared to other rain gauges. It also shows that there is a large variation of the two parameters between the different locations of the selected links.

Table 6.1: Rain gauges with selected radius (km) and local optimal parameters A_a (dB) (values between 0 and 4.0 dB) and α (-) (values between 0 and 1.0).

Rain Gauge	Anur	Batt	Colo	Jaff	Katu	Kuru	Maha	Mann	Polo	Putt	Rath	Trin
Radius (km)	5	3	1	3	4	5	10	7	10	7	3	5
Number of links	16	18	23	24	19	22	16	15	22	15	31	24
A_a (dB)	1.5	1	2	3	4	4	0	0	0	0	2.5	2.5
α (-)	0.45	0.8	0.25	0.4	0.95	0.7	0.95	0.95	0.45	0.55	0.45	0.65
NSE_{max} (-)	0.89	0.94	0.89	0.88	0.9	0.84	0.32	0.16	0.78	0.71	0.93	0.67

The selection of the links is different for each rain gauge for a couple of reasons. First, the number of links was different for each rain gauge. The condition was to have *at least* 15 links in the radius. For example, in rain gauge Mannar this resulted in a radius of 7 km with 15 links. For rain gauge Rathmalana, this resulted in a radius of 3 km with 31 links. The second difference between the rain gauges is that some rain gauges have links closer to the rain gauge than other rain gauges. The closer the links are to the radius, the more the certainty is that the rain gauge represents neighboring links well. For rain gauges with links further away, the uncertainty is higher. A third difference is that the patterns of the link networks are different for every rain gauge. Figure 6.2 shows the locations of the selected links for the rain gauges Colombo and Mannar. Colombo has links that have multiple different receiver/transmitter combinations. Mannar, on the other hand, has one tower where all the signals are transmitted to, or coming from. This makes Mannar very dependent on this one station for the calibration. If this station has any kind of error, all these links have this error as well and this will influence the calibration more.

Mannar is the most extreme case where all the links are connected to this one station. However, most of the rain gauges have similar patterns where part of the links all have the same station. This applies to all rain gauges except Colombo, Katunayake, Kurunegala, and Rathmalana.



(a) Colombo **(b)** Mannar

Figure 6.2: Selected links around the rain gauges Colombo and Mannar

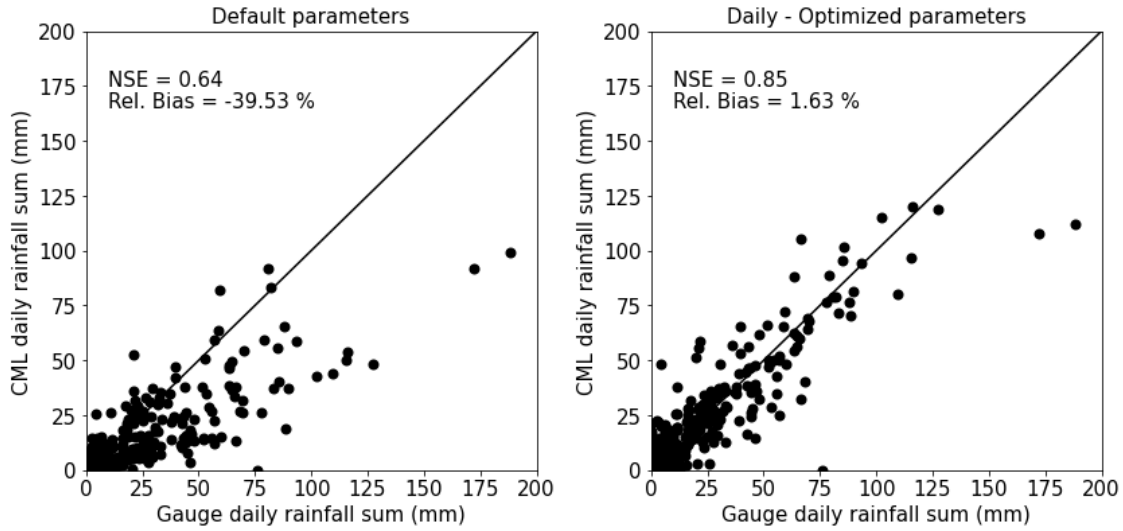
Rain gauges with longer path lengths had lower NSE values compared to the rain gauges with shorter path lengths. For example, Mannar has an NSE of 0.16. This maximum NSE score was obtained with $A_a = 0$ dB. A zero value means that no correction for wet antenna attenuation is added. With an α close to 1 ($\alpha = 0.95$) the maximum path averaged rainfall intensities is mostly used to calculate the mean path-averaged rainfall intensity. Next to this, for rain gauges with large path lengths, all the combinations of the two parameters still give an underestimation of the rainfall estimation compared to the rain gauges according to the relative bias.

One explanation for the underestimation of the rainfall from RAINLINK could be an error in the calculation of the wet-dry classification. Wet-dry classification is done with the link approach, where based on a radius of 15 km the links are selected. This radius of 15 km is chosen because it was assumed that all the links within this area will experience rain-induced signal attenuation in the same period. It could be that this assumption is incorrect. A smaller radius might improve the wet-dry classification. Another cause could be the a- and b-parameters that are only based on the frequency, but as mentioned before have other factors where the values could defer from those used in this study. A third cause that could be a factor in the difference in rainfall estimation could be the distance between the rain gauge and (most of) the links.

6.1.2. Comparison default and optimized parameters

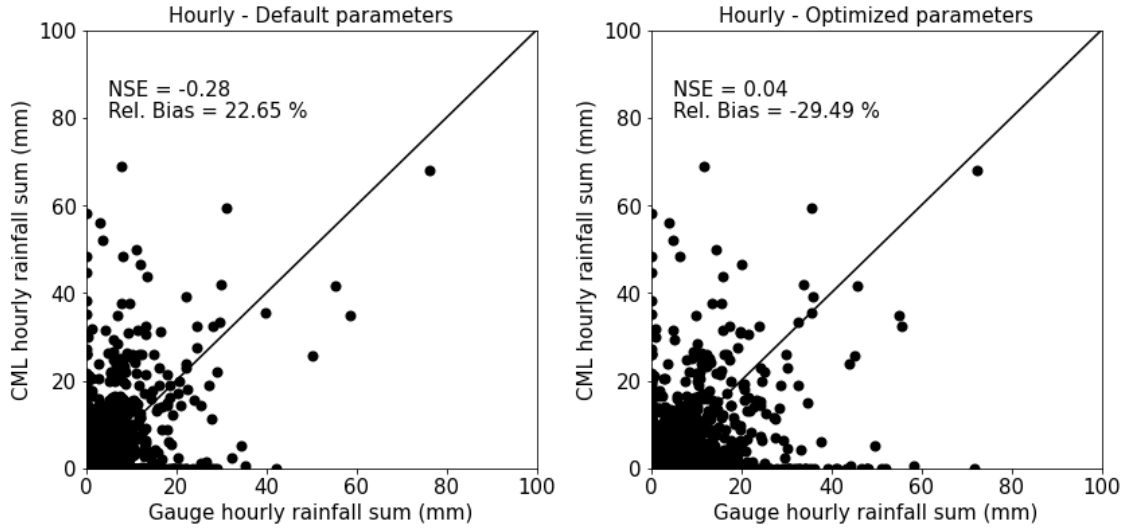
The default parameters from RAINLINK are used as a benchmark for calibration. For the hourly values, 8 rain gauges are selected and for the daily values, 12 rain gauges are selected. Figure 6.3 shows the comparison of the CML daily rainfall sum with the gauge daily rainfall sum for the default parameters as well as with the local optimal parameters. The NSE increased from 0.64 to 0.85 and the relative bias was −39.53% for the default parameters and is 1.63% after optimization. The figure of the default parameters illustrates that the CML rainfall retrievals often underestimate the rainfall compared to the rainfall depth of the rain gauges. This is confirmed by the large negative relative bias. In Appendix A a comparison with the results from Overeem et al. (2021) is made where the default parameters are used for the hourly and daily values for the period of September 12 to December 31, 2019.

For most of the rain gauges the optimization of the parameters improved the rainfall estimation according to the rain gauge data. However, for Katunayake the validation with default parameters scored somewhat better compared to the optimized parameters. However, the differences are quite small (NSE = 0.62 and NSE = 0.67, respectively). The same applies for Jaffna: NSE = 0.81 (optimal parameters) and NSE = 0.83 (default parameters). Figure 6.4 shows the verification of hourly CML rainfall estimations for both the default and the optimized parameters in the full period. The hourly values for the default parameters have an NSE of −0.28 and a relative bias of 22.65%. The figure illustrates that for hourly values there is much more overestimation of the CMLs in comparison to the rain gauge depth. Also, at some hours, the RAINLINK estimates rainfall while no rain has fallen according to the rain gauges.



(a) Default parameters **(b)** Optimized parameters

Figure 6.3: Daily rainfall depths of rain gauges compared to the CML rainfall estimations at the locations of the gauges after interpolation of (a) default and (b) optimized parameters in only the validation period (November 6 to December 31, 2019).



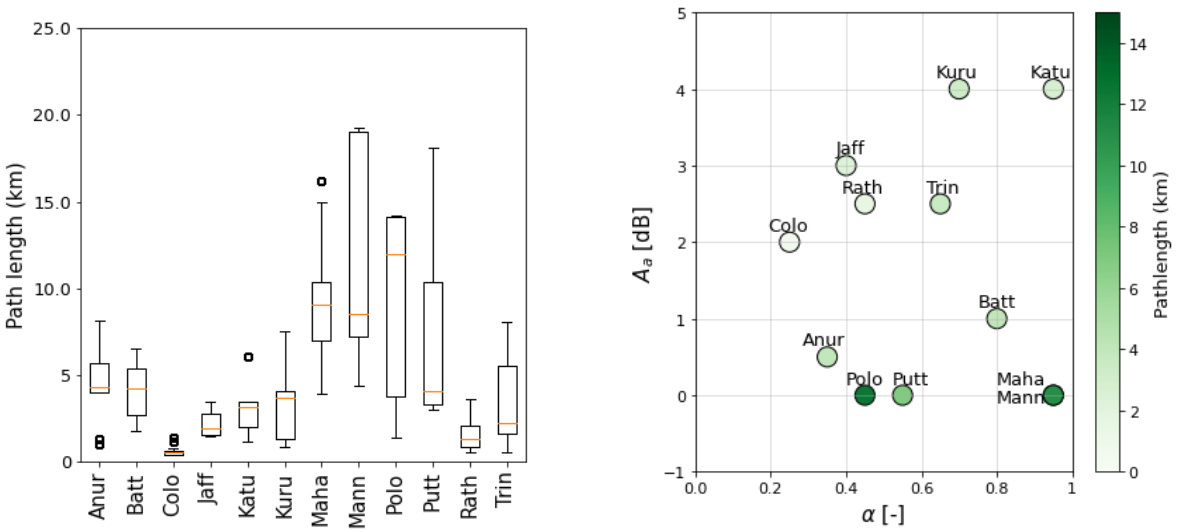
(a) Default parameters **(b)** Optimized parameters

Figure 6.4: Hourly rainfall depths of rain gauges compared to the CML rainfall estimations at the locations of the gauges of (a) default and (b) optimized parameters in the calibration and validation period (September 12 to December 31, 2019).

6.1.3. Relation between path length and the parameters A_a and α

Table 6.1 showed from the calibration the different rain gauges have different optimal values. This section analyses the potential causes for these differences. The selected links include a range of path lengths. As mentioned in Chapter 2, longer path lengths generally have lower frequencies and lower frequencies generally are less affected by attenuation. Figure 6.5a shows for each rain gauge the spread of the path length of the selected neighboring links. Figure 6.5b shows the average link path lengths of neighboring links at the gauge locations, located at the calibrated two parameters A_a and α.

It shows that for longer path lengths the wet antenna attenuation A_a is equal to zero and the coefficient α is close to 1. The NSE values of the optimization of the rain gauges with longer path lengths (Mahailluppalama and Mannar) are lower compared to the other NSE values. Also, these two rain gauges still have a large relative bias at the maximum NSE, while the relative bias of the other rain gauges is close to zero. Thus, for the rain gauges with longer path lengths, the optimization is not as optimal and remains with underestimation of the rainfall. In Appendix B two other possible dependencies are analyzed: the frequency and the locations of the rain gauges in the climate zones.



(a) Path length spread

(b) Relation path length, A_a and α

Figure 6.5: Information of all the rain gauges of (a) the path length spread and (b) the path length corresponding to the two optimized parameters (A_a and α).

6.1.4. Impact of selected links and radius size on calibration

The impact of the rain gauge on the calibration is investigated. Two rain gauges that are located relatively close to each other are Colombo and Rathmalana. The difference in results from the calibration of the parameters A_a and α between Colombo and Rathmalana are investigated, as they are very close to each other (9 km in a straight line). They are both in the same range for frequency and path lengths but still have different optimal values. This difference could be explained by a combination of factors. The radius of Colombo is smaller compared to the radius of Rathmalana. The selected links of Colombo are closer to the rain gauge and potentially have therefore more accurate estimations for the exact location of the rain gauge. Finally, the average of the frequency and path lengths of both rain gauges may be approximately the same, but Rathmalana does have links with longer path lengths in the selected data set (see the frequency ranges in Figure 6.5a).

The impact of choosing the radius with selected links is analyzed in Colombo. The number of selected links increases with increasing radius. The two parameters for the selected neighboring links for Colombo are calibrated with radius sizes of 1, 2, and 3 km. The number of links used for the calculation is 23, 68, and 128 links for the 1, 2, and 3 radius sizes respectively. All three calibrations showed that the optimal parameters of A_a are 2 dB and α is 0.25. For this particular case, changing the radius size did not influence the calibration of the parameters. This could be different for regions where fewer links are available and with longer distances of the links in connection to the rain gauges.

6.1.5. Pre-processing QPE for nowcasting

The links are assigned a certain parameter set based on their location. The parameters A_a and α of the rain gauge closest to the middle of the link path are selected for each link. Figure 6.6 shows the links in the color of the rain gauge the link is closest to. In this study, the selection of the parameters of the links is based on the geographical location of each link. This is not the only way to determine the parameters for each link. It could also be based on other factors, such as the link length. This distribution is used to determine the rainfall intensities with the locally optimized parameters. With the calculated rainfall intensities, RAINLINK

is also used to make QPE maps.

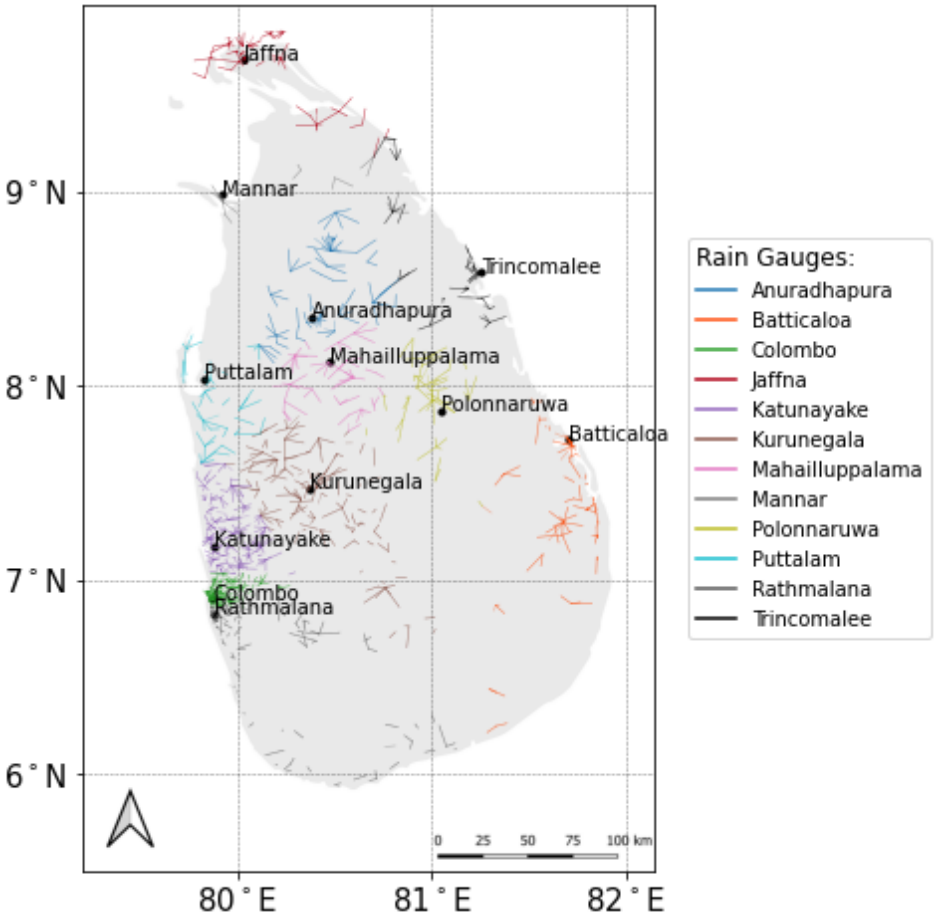


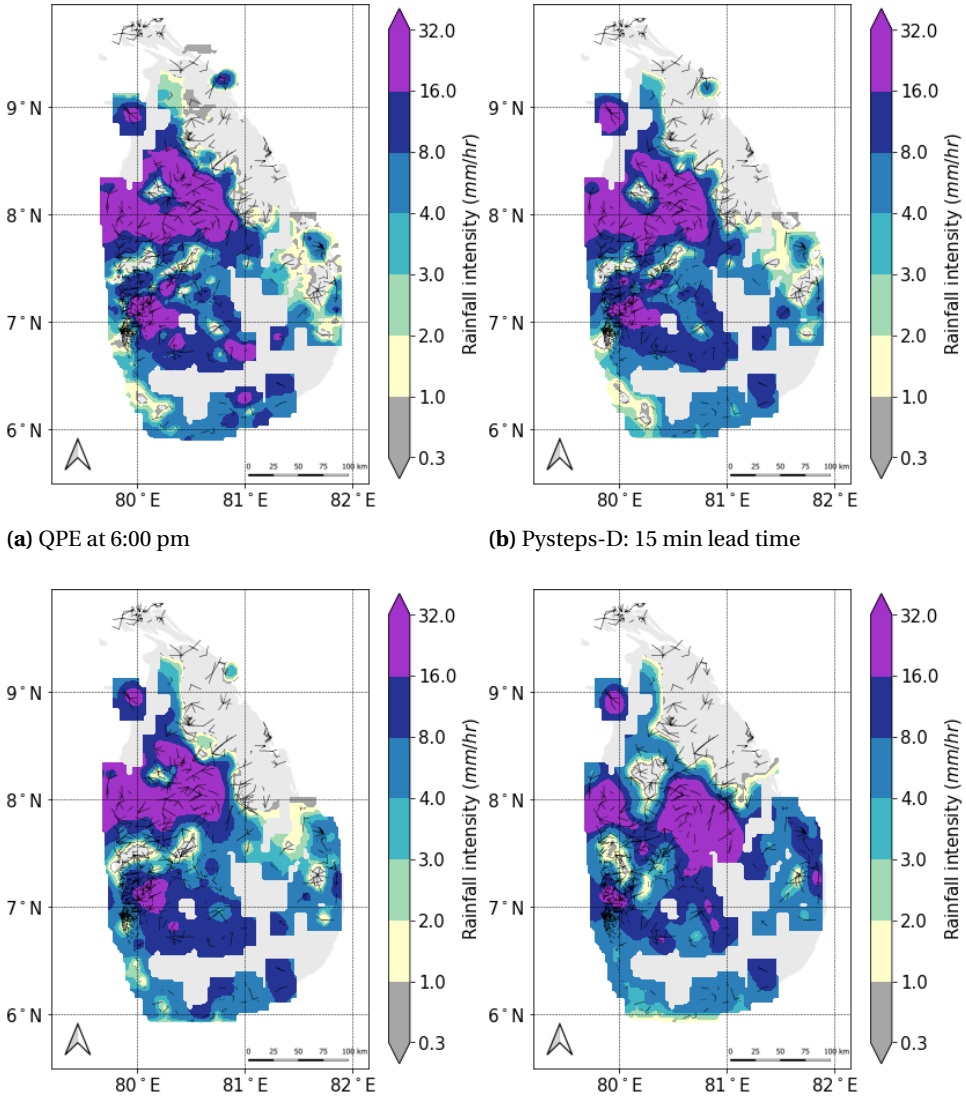
Figure 6.6: Selected rain gauge for each link in Sri Lanka based on the location of the link.

6.2. Nowcasting rainfall events

6.2.1. Selection of 20 rainfall events

The event of November 30 from 5:00 pm - 8:00 pm is used as an example throughout the next parts of the results. The rainfall event has a maximum rainfall intensity of 104 mm/hr. This is the average of the selected pixels. The rainfall event shows rainfall throughout the entire event and in a large area of Sri Lanka. Figure 6.7a shows the QPE of the rainfall event of November 30 at 6:00 pm. Figure 6.7b shows the Pysteps-D nowcast made for 6:00 pm with a lead time of 15 minutes. Figures 6.7c and 6.7d show Pysteps-D nowcasts made for 6:00 pm with lead times of 30 and 1 hour, respectively.

The figure shows that the locations and intensities of the rainfall event are well forecasted. The nowcasts made with lead times of 15 and 30 minutes have very similar results. The nowcast with a lead time of 1 hour does not capture the full rainfall event as well as the nowcasts with shorter lead times shown.



(a) QPE at 6:00 pm

(b) Pysteps-D: 15 min lead time

(c) Pysteps-D: 30 min lead time

(d) Pysteps-D: 60 min lead time

Figure 6.7: Rainfall event of November 30, 2019 at 6:00 pm with (a) the QPE (b) Pysteps-D nowcast with a lead time of 15 minutes, (c) QPF with a lead time of 30 minutes, and (d) QPF with a lead time of 1 hour.

6.3. Verification nowcasts for the 20 events

The 20 selected rainfall events are used to make nowcasts. The three methods (EP, Pysteps-D, and Pysteps-P) are tested. For Pysteps-P, the mean rainfall intensity of the 20 ensemble members is taken. For all of the nowcasts, the Fraction Skill Score and the Critical Success Index are calculated.

6.3.1. Fraction Skill Score

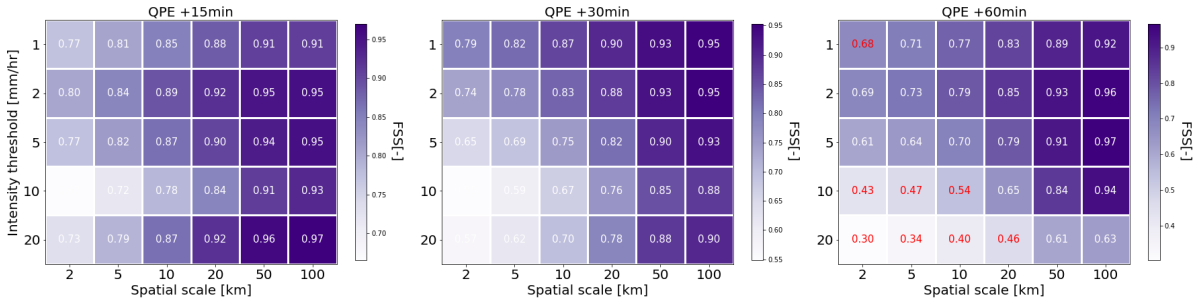
Figure 6.8 shows the FSS values for (a) Pysteps-D (b) Pysteps-P and (c) EP. For Pysteps-P the ensemble mean is taken for the calculation. FSS values are given for spatial scales from 2 to 100 km, intensity thresholds from 1 to 20 mm/hr, and lead times from 15 minutes (left), 30 minutes (middle), and 1 hour (right). The values indicated in red are the FSS values that score below $FSS_{uniform}$ (Equation 5.5). The red values indicate for which combination of intensity threshold and spatial skill the nowcast is not skillful anymore for that particular lead time.

The Pysteps-D nowcasts with 15 and 30 minutes lead times are skillful for all intensity thresholds and all spatial scales. Even for lead times up to an hour, the Pysteps-D nowcasts are still skillful. Only for intensities higher than 10 mm/hr, the nowcast is not skillful anymore for small spatial scales (< 10 km). For lead times of

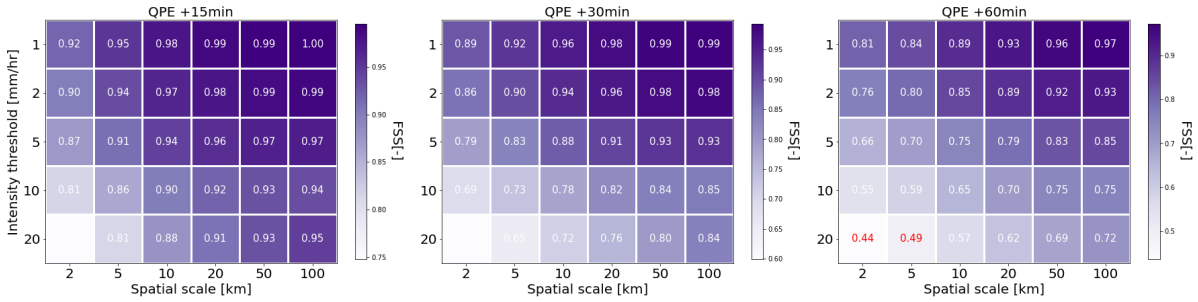
15 and 30 minutes, EP shows similar results. The FSS values for EP are lower compared to Pysteps-D, but still skillful for all intensity thresholds and spatial scales. However, for lead times of 1 hour, a clear difference is visible between the Pysteps-D nowcast and the EP. For EP, the nowcast is not skillful anymore for intensities above 5 mm/hr for spatial scales smaller than 20 km.

Overall, Pysteps-D scores especially better compared to the benchmark EP for longer lead times. A reason for this is that for longer lead times the advection and movement of the rainfall events are more important. The deterministic nowcast takes these processes into account and therefore has an advantage. With EP no extrapolation method is used, only the previous time step is taken, which loses its strength once lead times get longer. Still, the benchmark performed exceptionally well. One of the reasons for this is because the rainfall events are very steady and not much movement occurred, which made using the QPE maps of the previous time steps still effective.

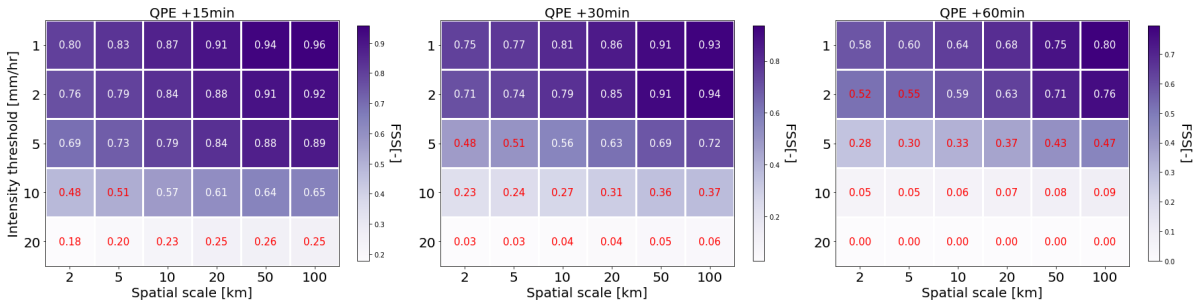
The ensemble mean of Pysteps-P shows worse results in comparison to the Pysteps-D and EP. The Pysteps-P nowcast is not skillful for high intensities (> 5 mm/hr). For low intensities, the nowcast of Pysteps-P remains skillful even for an hour. The reason for the poor result from the Pysteps-P nowcast is very likely since the mean of the nowcasts of the 20 ensemble members is taken. Therefore, especially for high intensities, the rainfall intensities are underestimated. In the next section with the CSI calculations, the ensemble members are also considered separately to show the spread in the accuracy of the ensemble.



(a) EP



(b) Pysteps-D



(c) Pysteps-P

Figure 6.8: FSS averaged values from all rainfall events of (a) EP, (b) Pysteps-D, and (c) the ensemble mean of Pysteps-P.

6.3.2. Critical Success Index

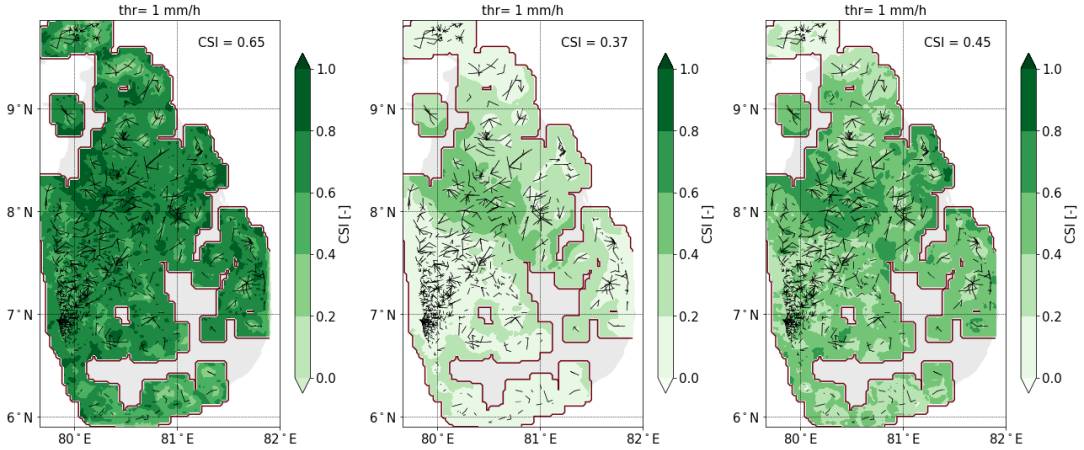
Table 6.2 shows the averaged CSI over the four lead times (15, 30, 45, and 1 hour) and the 20 rainfall events for thresholds of 1, 2, 5, 10, and 20 mm/hr. Figure 6.9 shows the average CSI of all the events for a threshold of 1 mm/hr for (a) Pysteps-D, (b) Pysteps-P, and (c) EP. The area used to calculate CSI is indicated by a red outline. The figures of Pysteps-P and EP illustrate that the CSI scores are not equal in the contributing area of Sri Lanka. The north and south and the area close to the sea have lower CSI values compared to the middle of the country.

Table 6.2: CSI values for different thresholds for Pysteps-D, Pysteps-P, and EP for the average of nowcasts over lead times of 15, 30, 45, and 1 hour and 20 rainfall events.

	Threshold (mm/hr)				
Algorithm	1	2	5	10	20
Pysteps-D	0.65	0.60	0.49	0.38	0.25
Pysteps-P	0.35	0.32	0.21	0.15	0.08
EP	0.45	0.40	0.30	0.19	0.12

Pysteps-D scores higher CSI values compared to the benchmark EP for all intensity thresholds. Thus, the nowcast of Pysteps-D is more accurate. The ensemble mean of Pysteps-P does not score better and this is likely for the same reasons as described in the FSS results.

The link density, especially in the south, is very low, and therefore the QPE maps might not be as accurate as in other places. There are only rainfall estimations available for the areas on the island, and not in the sea. This could be an explanation for lower CSI values in the areas close to the sea compared to the inland area. If a rainfall event starts near the sea and propagates towards the land this is not noted by the QPE. This could be part of the reason why for the Pysteps-P and the EP the values of CSI are lower in those areas.



(a) Pysteps-D **(b)** Pysteps-P **(c)** EP

Figure 6.9: CSI values with threshold of 1 mm/hr from (a) Pysteps-D, (b) Pysteps-P and (c) EP averaged over the 20 rainfall events. The red outline indicates the calculated area.

From the 20 rainfall events, the performance of the nowcasts for those different events also give variations in CSI values for all three nowcasts. Figure 6.10 illustrates the distributions of the CSI value from the 20 rainfall events for a threshold of 1 mm/hr for Pysteps-D, Pysteps-P, and EP from averaged lead times of 15, 30, 45, and 1 hour. The spread of CSI values of Pysteps-P is much higher compared to EP and Pysteps-D.

Appendix C (Figure C.1) shows the CSI values of the event of (a) November 30 17:00-20:00 and (b) December 3 16:00-19:00. Event (a) has a CSI score of 0.85 and event (b) a CSI score of 0.55 for Pysteps-D for a threshold of 1 mm/hr and an averaged nowcast of lead times of 15, 30, 45, and 1 hour.

To investigate the different ensemble members separately, the CSI value is computed for each ensemble member. Figure 6.10 shows the distribution of the CSI values for all the ensemble members, calculated for a threshold of 1 mm/hr and the average of the four lead times used also from previous calculations (15, 30, 45, and 1 hour). The red dot is the EP value per event. It shows that some of the ensemble members scored higher CSI values compared to the EP, and most of them had lower CSI values.

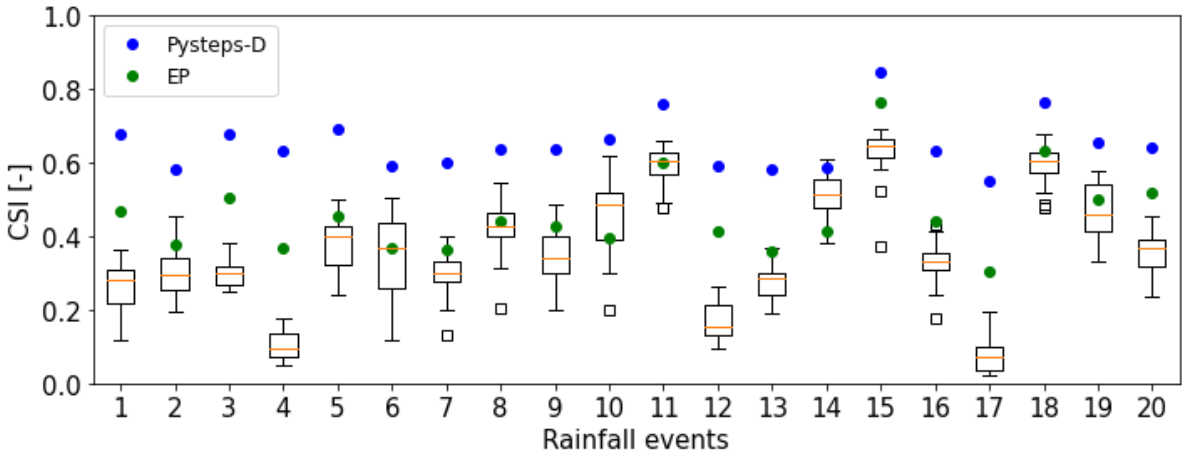


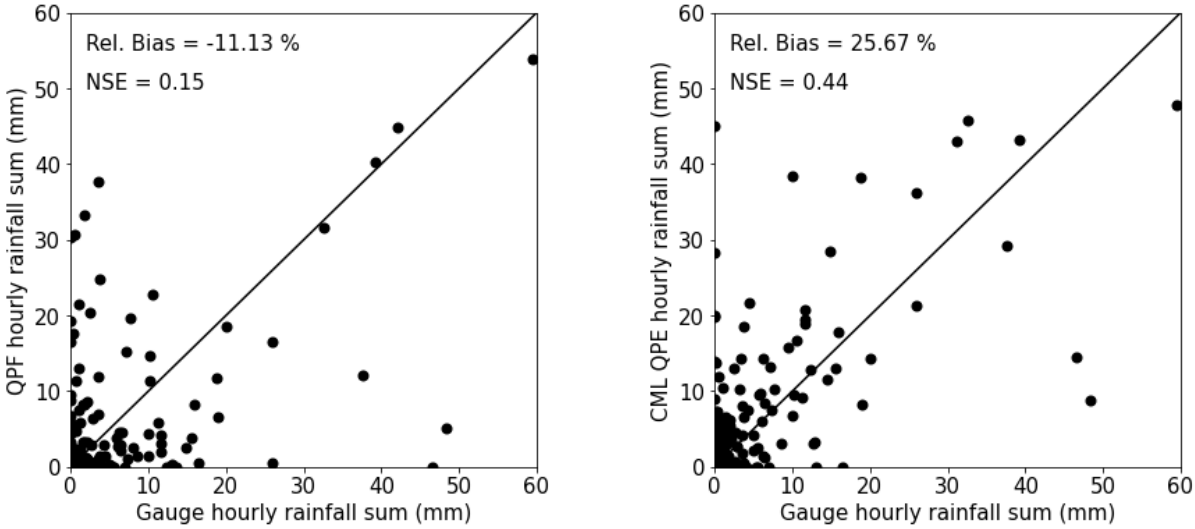
Figure 6.10: CSI value distribution from the 20 rainfall events for a threshold of 1 mm/hr for Pysteps-D, Pysteps-P and EP from averaged lead times (15, 30, 45, and 1 hour), where the values of the Pysteps-P ensemble members are shown separately.

6.4. Comparison rain gauges vs. QPF

The 20 rainfall events are also used for the comparison of the Pysteps-D nowcast with the rain gauges. In this comparison, the hourly rain gauge dataset is used (see Table 4.1 for the selected hourly rain gauges). For every hour the QPFs are calculated 15 minutes before that hour. For example, on September 23 from 14:00 to 15:00, a nowcast is made at 13:45 and the nowcast is made with lead times of 15, 30, 45, and 1 hour. Figure 6.11a shows the result of the accumulated QPF rainfall depths at the locations of the rain gauges compared to the hourly rainfall intensities.

Figure 6.11b shows for the same rainfall events the comparison of the CML-based rainfall depths compared with the rainfall depths of the rain gauges. The Pysteps-D nowcasts scored an NSE of 0.15 and the QPE scored an NSE of 0.44. The figure shows that sometimes rainfall is observed at the rain gauge location, but not by the links, and sometimes the other way around. As seen in the figure, the QPE shows similar behavior. Uncertainties that are caused by differences between the QPE and the rain gauges have been discussed in Section 6.1.

Appendix C (Figure C.2) shows the comparison between the rain gauge data and the QPF and QPE for one single event, the event of November 30. For each hour the rainfall depths of QPE, QPF, and the rain gauges for the hourly rain gauges are presented.



(a) Rain gauges vs. QPF

(b) Rain gauges vs. QPE

Figure 6.11: Comparison from 20 rainfall events of (a) Pysteps-D nowcast and (b) QPE with the hourly rain gauge intensities on the same pixel, where the QPF is the accumulated rainfall intensity of four nowcasts made with lead times of 15, 30, 45, and 1 hour for each hour.

6.5. Comparison GPM vs. QPF

The GPM satellite product is compared for 11 events (Table 4.2) to the QPF from Pysteps-D. This comparison is done for lead times of 15, 30, 45 minutes, and 1 hour. Figure 6.12 shows the event of October 17, 2019 at 16:05 and the nowcast of 1 hour ahead. Appendix D shows the results of the other 10 events. The times shown in the results are the local time in Sri Lanka (UTC+5:30).

The events show that the intensity of the rainfall events of Pysteps-D QPF and the QPE is most of the time higher in comparison to the GPM satellite product. Figure 6.12 shows that the spatial distributions of the rainfall event correspond well to each other. The pattern of the QPF is coarser compared to the satellite precipitation estimation.

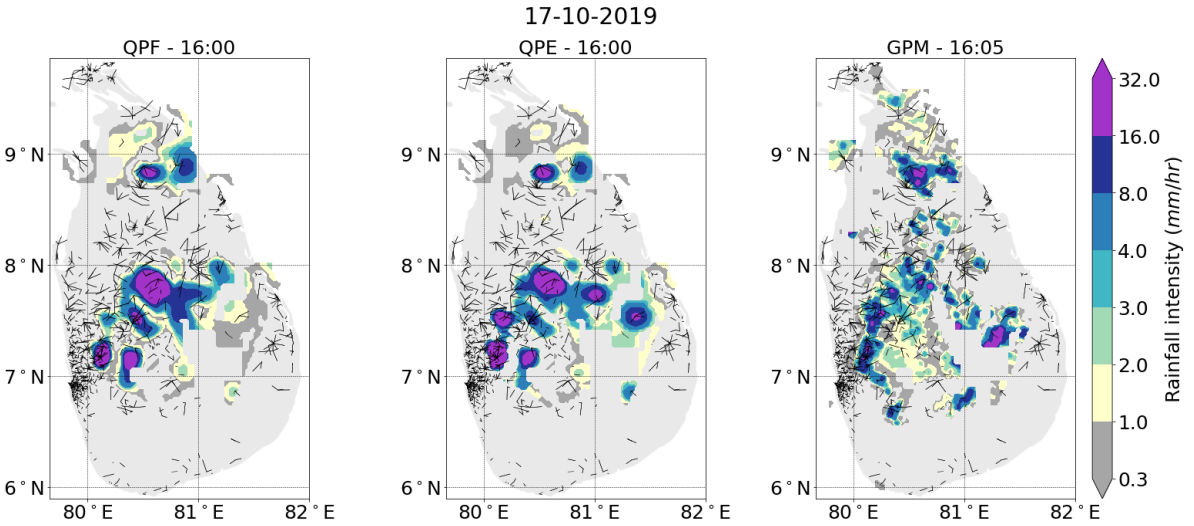


Figure 6.12: A rainfall event on October 17, 2019 at 16:05 with (left) Pysteps-D nowcast at 16:00 (lead time of 1 hour), (middle) QPE at 16:00 and (right) GPM satellite rainfall estimation at 16:05.

6.6. Comparison performance nowcasts Sri Lanka vs. the Netherlands

A study of Imhoff et al. (2020a) investigates the use of CMLs for nowcasting rainfall in the Netherlands. The study calculated the FSS for two thresholds: 1 and 5 mm/hr. Table 6.3 gives the comparison between the FSS scores of the Netherlands (NL) and Sri Lanka (SL) for the different thresholds and lead times for a spatial scale of 2 km (the smallest spatial scale obtained with CMLs in Sri Lanka). For Sri Lanka, the Pysteps-D nowcast is used. Again, red values indicate for which combination of intensity threshold and spatial skill the nowcast is not skillful anymore for that particular lead time.

Table 6.3: FSS scores (-) for CML nowcasts from the Netherlands (values from (Imhoff et al., 2020a)) and Sri Lanka (Pysteps-D and EP) for thresholds of 1 and 5 mm/hr and for lead times of 15, 30, and 1 hour.

Threshold	1 mm/hr				5 mm/hr			
lead time (min)	NL	SL-EP	SL-D	SL-P	NL	SL-EP	SL-D	SL-P
15	0.55	0.85	0.98	0.87	0.30	0.87	0.97	0.79
30	0.5	0.87	0.96	0.81	0.28	0.75	0.94	0.56
60	0.42	0.77	0.89	0.64	0.20	0.70	0.85	0.33

In the FSS calculation of Imhoff et al. (2020a) in the Netherlands a radar-based observation was used as the reference. The FSS calculation of this thesis in Sri Lanka used a CML-based observation as the reference because radar observations are not available in Sri Lanka. The FSS scores are therefore not calculated in the same way. Making a comparison in absolute numbers would not be fair.

However, the difference in skill between the two thresholds is visible in the scores for both the Netherlands and Sri Lanka. Even though the scores in the Netherlands show a larger drop in skill from intensities of 1 mm/hr to 5 mm/hr, it is also present for nowcasts in Sri Lanka. Furthermore, with increasing the lead times of the nowcasts the two studies show a similar reduction in FSS scores, both from 15 minutes to 30 minutes and from 30 minutes to 1 hour.

7

Discussion

7.1. Commercial Microwave Links

This thesis studied the calibration of two parameters A_a and α from the RAINLINK algorithm. However, there are other parameters used in RAINLINK that are also important for making path-averaged rainfall estimations. An example where all the parameters are taken into account is in the study of Wolff et al. (2022). That study used a 12-day calibration period for the Netherlands to improve the default parameter set from RAINLINK.

In this study, calibration is done by calculating the NSE for every combination of the parameters and choosing the parameter set with a maximal NSE value. A downside of using NSE as the measure of performance is that using the squared errors in NSE inflates the importance of high rainfall intensities. Maximizing the NSE forces the model to get the high rainfall intensities correct. For this study, the relative bias was also calculated. For future studies another statistical value (e.g. NSE_{log}) could be added to represent low rainfall intensities better.

The calibration period of this study was a period of 55 days from September 12 to November 6, 2019. Sri Lanka has strong seasonal differences due to the monsoons and also large differences in rainfall intensities between the monsoon and inter-monsoon seasons. Therefore, a recommendation would be to extend the calibration period to at least a year to include all the seasons of Sri Lanka.

This study used the calibrated parameters of selected links surrounding 12 rain gauges. With the calibrated parameters, each link was paired with one of the parameter sets from those 12 rain gauges. This was based on the shortest distance between the middle of each link and the location of the rain gauge. However, as also shown in Figure 6.6, some links have a large distance to their selected rain gauge. The uncertainty for those links is quite large because the correlation of the rainfall in space is less with larger distances.

This study already showed that the parameters are dependent on the link length. To improve the calculation of the rainfall estimations with RAINLINK, a possible solution could be to select the optimal values of parameters A_a and α for each link based on the path length of the links. Before this could be implemented, it is first recommended to perform a similar calibration for links grouped by path length. After this, for each link, a parameter set can be chosen based on the path length. This can be implemented in RAINLINK in a similar way as the values a and b are chosen for each link dependent on their used frequency.

For the interpolation of rainfall maps, default parameters from the Netherlands are used for the sill, range, and nugget of the variogram. Sri Lanka has a different climate compared to the Netherlands. Therefore, the parameters for the variogram are also expected to be different. This study used the default parameters for the variogram and did not improve these parameters. Optimizing these parameters for Sri Lanka could improve the interpolation and thus improve the rainfall maps even further.

Some locations in Sri Lanka do not have any available links. In addition to the areas inland that do not include any links, there are also no CMLs on the sea. For these areas, no QPE or QPF can be calculated with CMLs. It

is therefore important to investigate other sources of rainfall estimations (e.g. satellite data) to fill those gaps. Another advantage of this could be that the performance of the nowcasts could improve by adding information from the sea. Some rainfall events have their origin at sea and in this way, the rainfall events could be detected earlier.

This study only used the algorithm RAINLINK for the retrieval of rainfall maps from CML data. It would be interesting to compare the RAINLINK algorithm to other algorithms that also give rainfall maps from CML data. An algorithm that has recently been published is Pycomlink (Chwala, 2021). In Graf et al. (2021) a study in Germany has been performed with Pycomlink that showed promising results.

7.2. Nowcasting

There is an important limitation in the calculation of the verification values (FSS and CSI). The reference rainfall observations are the QPEs made with RAINLINK from the CML data. These QPEs are also used as input for the making of the nowcasts in Pysteps. The QPF and QPE are compared with each other because next to the satellite product, there are no spatial maps of rainfall intensities present in Sri Lanka. This thesis used a qualitative verification of the accuracy of the rain gauges and a spatial verification with the satellite data.

This study used 20 rainfall events for the calculation of the FSS and CSI scores. If one of the rainfall events is performing particularly well or poorly, the average of the scores will be affected. To perform a better verification, it is recommended to include more rainfall events with the calculations. For example, in the study of Imhoff et al. (2020b), a total of 1,533 events were selected and analyzed.

Next to analyzing more rainfall events, it would also be beneficial to analyze a longer period of rainfall events. This thesis showed results for nowcasts with lead times up to 60 minutes. Longer lead times could still be skillful. Although the results from this thesis show promising results for the deterministic nowcasts up to an hour, it is still uncertain how nowcasts with lead times beyond one hour will perform could be for the nowcasts.

This study used the algorithm Pysteps to make nowcasts, where both deterministic and probabilistic nowcasts are generated. The nowcasts were compared to the benchmark Eulerian Persistence. The algorithm Pysteps was the only library used in this thesis. However, there are numerous other libraries available that also can be used to make nowcasts. One example is Rainymotion (Ayzel et al., 2019), an open-source benchmarking algorithm.

8

Conclusion

This thesis investigates the possibility of using Commercial Microwave Links (CMLs) for nowcasting rainfall in Sri Lanka. The algorithm RAINLINK is used to generate QPEs from the CMLs. The QPEs are a vital part of the nowcast input. The default parameters of RAINLINK are calibrated for the Netherlands. This thesis investigated and calibrated two parameters for Sri Lanka. The calibration to the nowcasting of rainfall is made with the library Pysteps. The accuracy of the nowcasts produced with Pysteps is evaluated in this thesis.

First, this thesis focused on generating CMLs QPEs with the RAINLINK algorithm by using calibrated parameters. The results showed that by calibrating the parameters of the wet antenna attenuation (A_a) and the coefficient (α), RAINLINK gives more accurate QPEs in Sri Lanka. In a validation period of 55 days, the daily values of the generated QPEs from RAINLINK are compared with 12 rain gauges. This resulted in an NSE of 0.85. For comparison, the NSE of the generated QPEs from RAINLINK where default parameters are used resulted in an NSE of 0.64. Comparing the different parameter sets from the rain gauges, results showed that the parameters are dependent on, among other things, the path length of the link. Longer path lengths are more likely to have lower wet antenna attenuation and a higher value for the coefficient α compared to shorter path lengths. The thesis showed results of nowcasts made in a deterministic sense with S-PROG and a probabilistic sense with STEPS. The deterministic nowcasts are accurate and skillful for lead times up to an hour, even for high thresholds of 20 mm/hr and the smallest spatial scale (2 km). The probabilistic nowcasts are less accurate and skillful, although they are still skillful for the smallest spatial scale for the lowest intensity calculated (1 mm/hr) for lead times up to an hour. The nowcasting results showed that it is possible to generate accurate nowcasts from CML rainfall estimations.

To conclude, this thesis showed that by using calibrated parameters in RAINLINK the overall accuracy of the QPEs generated with RAINLINK is improved. Furthermore, the thesis showed that nowcasting rainfall in Sri Lanka with CML QPEs is accurate for lead times of at least an hour. Overall, this thesis showed promising results for the future of nowcasting with CMLs. Further research is needed to include more parameters in the calibration and to investigate other algorithms to generate CML QPEs. Based on the conclusions of the nowcasting, future studies could study the probabilistic nowcasts in more depth and compare the nowcasts generated with Pysteps with other nowcasting libraries available. The possibilities of combining CML rainfall estimations, satellite data, and rain gauges into one QPE and using this to nowcast the rainfall also are recommended.

References

Atlas, D., & Ulbrich, C. W. (1977). Path- and area-integrated rainfall measurement by microwave attenuation in the 1–3 cm band. *Journal of Applied Meteorology and Climatology*, *16*(12), 1322 - 1331. doi: https://doi.org/10.1175/1520-0450(1977)016<1322:PAAIRM>2.0.CO;2.

Ayzel, G., Heistermann, M., & Winterrath, T. (2019). Optical flow models as an open benchmark for radar-based precipitation nowcasting (Rainymotion v0.1). *Geoscientific Model Development*, *12*(4), 1387–1402. doi: https://doi.org/10.5194/gmd-12-1387-2019.

Berenguer, M., Sempere-Torres, D., & Pegram, G. G. (2011). SBMcast – an ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation. *Journal of Hydrology*, *404*(3), 226-240. doi: https://doi.org/10.1016/j.jhydrol.2011.04.033.

Berenguer, M., Surcel, M., Zawadzki, I., Xue, M., & Kong, F. (2012). The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part II: Intercomparison among numerical models and with nowcasting. *Monthly Weather Review*, *140*(8), 2689 - 2705. doi: https://doi.org/10.1175/MWR-D-11-00181.1.

Berne, A., & Uijlenhoet, R. (2007). Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability. *Geophysical Research Letters*, *34*(7). doi: https://doi.org/10.1029/2007GL029409.

Bowler, N. E., Pierce, C. E., & Seed, A. W. (2006). Steps: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society*, *132*(620), 2127-2155. doi: https://doi.org/10.1256/qj.04.100.

Browning, K. A. (1980). Review lecture: Local weather forecasting. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *371*(1745), 179–211. doi: https://doi.org/10.1098/rspa.1980.0076.

Burt, T., & Weerasinghe, N. (2014, 09). Rainfall distributions in sri lanka in time and space: An analysis based on daily rainfall data. *Climate*, *2*. doi: https://doi.org/10.3390/cli2040242.

Chwala, C. (2021). *Pycomlink*. https://github.com/pycomlink/pycomlink. GitHub.

Department of Census and Statistics of Sri Lanka. (2021). Mid-year population estimates by district & sex, 2016-2021. Retrieved from http://www.statistics.gov.lk/Resource/en/Population/Vital_Statistics/Mid-year_population_by_district.pdf

Disaster Management Centre Sri Lanka. (2019). Situation report - Sri Lanka. Retrieved from http://www.dmc.gov.lk/images/dmcreports/Situation_Report_on_2019__1570712541.pdf

Dixon, M., & Wiener, G. (1993). Titan: Thunderstorm identification, tracking, analysis, and nowcasting—a radar-based methodology. *Journal of Atmospheric and Oceanic Technology*, *10*(6), 785 - 797. doi: https://doi.org/10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2.

Fabry, F. (2015). *Radar meteorology: Principles and practice*. Cambridge University Press. doi: https://doi.org/10.1017/CBO9781107707405.

Foresti, L., Sideris, I. V., Nerini, D., Beusch, L., & Germann, U. (2019). Using a 10-year radar archive for nowcasting precipitation growth and decay: A probabilistic machine learning approach. *Weather and Forecasting*, *34*(5), 1547 - 1569. doi: https://doi.org/10.1175/WAF-D-18-0206.1.

Germann, U., & Zawadzki, I. (2002, 12). Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Monthly Weather Review - MON WEATHER REV*, *130*. doi: https://doi.org/10.1175/1520-0493(2002)130<2859:SDOTPO>2.0.CO;2.

Germann, U., Zawadzki, I., & Turner, B. (2006). Predictability of precipitation from continental radar images. Part IV: Limits to prediction. *Journal of the Atmospheric Sciences*, *63*(8), 2092 - 2108. doi: https://doi.org/10.1175/JAS3735.1.

Graf, M., El Hachem, A., Eisele, M., Seidel, J., Chwala, C., Kunstmann, H., & Bárdossy, A. (2021). Rainfall estimates from opportunistic sensors in Germany across spatio-temporal scales. *Journal of Hydrology: Regional Studies*, *37*, 100883. doi: https://doi.org/10.1016/j.ejrh.2021.100883.

Harischandra, I., Dassanayake, R., & De Silva, N. (2016, 01). Three sympatric clusters of the malaria vector Anopheles Culicifacies e (diptera: Culicidae) detected in Sri Lanka. *Parasites & Vectors*, *9*. doi: https://doi.org/10.1186/s13071-015-1286-3.

Heuvelink, D., Berenguer, M., Brauer, C. C., & Uijlenhoet, R. (2020). Hydrological application of radar rainfall nowcasting in the netherlands. *Environment International*, *136*, 105431. doi: https://doi.org/10.1016/j.envint.2019.105431.

Hilst, G. R., & Russo, J. A. (1960). An objective extrapolation technique for semi-conservative fields with an application to radar patterns. *The Travelers Research Center, Hardford*.

Imhoff, R. O., Brauer, C. C., Overeem, A., Weerts, A. H., & Uijlenhoet, R. (2020b). Spatial and temporal evaluation of radar rainfall nowcasting techniques on 1,533 events. *Water Resources Research*, *56*. doi: https://doi.org/10.1029/2019WR026723.

Imhoff, R. O., Overeem, A., Brauer, C. C., Leijnse, H., Weerts, A. H., & Uijlenhoet, R. (2020a). Rainfall nowcasting using commercial microwave links. *Geophysical Research Letters*, *47*. doi: https://doi.org/10.1029/2020GL089365.

International Telecommunication Union. (2005). Specific attenuation model for rain for use in prediction methods. Retrieved from https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.838-3-200503-I!!PDF-E.pdf

Kidd, C., & Huffman, G. (2011). Global precipitation measurement. *Meteorological Applications*, *18*(3), 334-353. doi: https://doi.org/10.1002/met.284.

Laroche, S., & Zawadzki, I. (1994). A variational analysis method for retrieval of three-dimensional wind field from single-doppler radar data. *Journal of Atmospheric Sciences*, *51*(18), 2664 - 2682. doi: https://doi.org/10.1175/1520-0469(1994)051<2664:AVAMFR>2.0.CO;2.

Leijnse, H., Uijlenhoet, R., & Stricker, J. (2008). Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation. *Advances in Water Resources*, *31*(11), 1481-1493. doi: https://doi.org/10.1016/j.advwatres.2008.03.004. (Hydrologic Remote Sensing)

Leijnse, H., Uijlenhoet, R., & Stricker, J. N. M. (2007). Hydrometeorological application of a microwave link: 2. precipitation. *Water Resources Research*, *43*(4). doi: https://doi.org/10.1029/2006WR004989.

Ligda, M. G. H. (1951). Radar storm observation. In T. F. Malone (Ed.), *Compendium of meteorology: Prepared under the direction of the committee on the compendium of meteorology* (pp. 1265–1282). Boston, MA: American Meteorological Society.

Lucas, B., & Kanade, T. (1981). *An iterative image registration technique with an application to stereo vision (IJCAI'81)*. San Fransisco, CA: Morgan Kaufmann Publishers Inc.

Manawadu, L., & Wijeratne, V. (2021). Anthropogenic drivers and impacts of urban flooding- a case study in lower kelani river basin, colombo sri lanka. *International Journal of Disaster Risk Reduction*, *57*, 102076. doi: https://doi.org/10.1016/j.ijdrr.2021.102076.

Minda, H., & Nakamura, K. (2005). High temporal resolution path-average rain gauge with 50-ghz band microwave. *Journal of Atmospheric and Oceanic Technology*, *22*(2), 165 - 179. doi: https://doi.org/10.1175/JTECH-1683.1.

NASA. (2013). Nasa shuttle radar topography mission global 3 arc second [data set]. *NASA EOSDIS Land Processes DAAC*.

Olson, W. S., & Masunaga, H. (2018). GPM combined radar-radiometer precipitation algorithm theoretical basis document (version 5). Retrieved from https://gpm.nasa.gov/sites/default/files/2020-05/Combined_algorithm_ATBD.V05.pdf

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2011). Measuring urban rainfall using microwave links from commercial cellular communication networks. *Water Resources Research*, *47*. doi: https://doi.org/10.1029/2010WR010350.

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2013). Country-wide rainfall maps from cellular communication networks. *Proceedings of the National Academy of Sciences*, *110*(8), 2741-2745. doi: https://doi.org/10.1073/pnas.1217961110.

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2016a). Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network. *Atmospheric Measurement Techniques*, *9*. doi: https://doi.org/10.5194/amt-9-2425-2016.

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2016b). Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks. *Water Resources Research*, *52*(10), 8039-8065. doi: https://doi.org/10.1002/2016WR019412.

Overeem, A., Leijnse, H., van Leth, T. C., Bogerd, L., Priebe, J., Tricarico, D., ... Uijlenhoet, R. (2021). Tropical rainfall monitoring with commercial microwave links in Sri Lanka. *Environmental Research Letters*, *16*. doi: https://doi.org/10.1088/1748-9326/ac0fa6.

Pastorek, J., Fencl, M., Rieckermann, J., & Bareš, V. (2021). Precipitation estimates from commercial microwave links: Practical approaches to wet-antenna correction. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-9. doi: https://doi.org/10.1109/TGRS.2021.3110004.

Perera, E. N. C., Jayawardana, D. T., Jayasinghe, P., Bandara, R. M. S., & Alahakoon, N. (2018, Aug 13). Direct impacts of landslides on socio-economic systems: a case study from aranayake, sri lanka. *Geoenvironmental Disasters*, *5*(1), 11. doi: https://doi.org/10.1186/s40677-018-0104-6.

Pierce, C., Seed, A., Ballard, S., Simonin, D., & Li, Z. (2012). Nowcasting. In J. Bech & J. L. Chau (Eds.), *Doppler radar observations* (chap. 4). Rijeka: IntechOpen. doi: https://doi.org/10.5772/39054.

Pulkkinen, S., Nerini, D., Pérez Hortal, A. A., Velasco-Forero, C., Seed, A., Germann, U., & Foresti, L. (2019). Pysteps: An open-source python library for probabilistic precipitation nowcasting (v1.0). *Geoscientific Model Development*, *12*. doi: https://doi.org/10.5194/gmd-12-4185-2019. (Background information for pysteps)

Rios Gaona, M., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2015). Measurement and interpolation uncertainties in rainfall maps from cellular communication networks. *Hydrology and Earth System Sciences*, *19*(8), 3571–3584. doi: https://doi.org/10.5194/hess-19-3571-2015.

Roberts, N. M., & Lean, H. W. (2008). Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, *136*(1), 78 - 97. doi: https://doi.org/10.1175/2007MWR2123.1.

Ruzanski, E., & Chandrasekar, V. (2012). An investigation of the short-term predictability of precipitation using high-resolution composite radar observations. *Journal of Applied Meteorology and Climatology*, *51*(5), 912–925. doi: https://doi.org/10.1175/JAMC-D-11-069.1.

Saltikoff, E., Friedrich, K., Soderholm, J., Lengfeld, K., Nelson, B., Becker, A., ... Tassone, C. (2019). An overview of using weather radar for climatological studies: Successes, challenges, and potential. *Bulletin of the American Meteorological Society*, *100*(9), 1739 - 1752. doi: https://doi.org/10.1175/BAMS-D-18-0166.1.

Schaefer, J. T. (1990). The critical success index as an indicator of warning skill. *Weather and Forecasting*, *5*(4), 570 - 575. doi: https://doi.org/10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.

Seed, A. W. (2003). A dynamic and spatial scaling approach to advection forecasting. *Journal of Applied Meteorology*, *42*(3), 381-388. doi: https://doi.org/10.1175/1520-0450(2003)042<0381:ADASSA>2.0.CO;2.

Simonin, D., Pierce, C., Roberts, N., Ballard, S. P., & Li, Z. (2017). Performance of met office hourly cycling NWP-based nowcasting for precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, *143*(708), 2862-2873. doi: https://doi.org/10.1002/qj.3136.

Sri Lanka Department of Meteorology. (2021). *Climate of Sri Lanka*. https://meteo.gov.lk/index.php?option=com_content&view=article&id=94&Itemid=310&lang=en&lang=en#2-southwest-monsoon-season-may-september.

Turner, B. J., Zawadzki, I., & Germann, U. (2004). Predictability of precipitation from continental radar images. Part III: operational nowcasting implementation (MAPLE). *Journal of Applied Meteorology*, *43*(2), 231 - 248. doi: https://doi.org/10.1175/1520-0450(2004)043<0231:POPFCR>2.0.CO;2.

Upton, G., Holt, A., Cummings, R., Rahimi, A., & Goddard, J. (2005). Microwave links: The future for urban rainfall measurement? *Atmospheric Research*, *77*(1), 300-312. doi: https://doi.org/10.1016/j.atmosres.2004.10.009. (Precipitation in Urban Areas)

van Leth, T. C., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2018). A measurement campaign to assess sources of error in microwave link rainfall estimation. *Atmospheric Measurement Techniques*, *11*(8), 4645–4669. doi: https://doi.org/10.5194/amt-11-4645-2018.

Wang, Y., Coning, E., Harou, A., Jacobs, W., Joe, P., Nikitina, L., … Sun, J. (2017). *Guidelines for nowcasting techniques*. World Meteorological Organization.

Wolff, W., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2022). Rainfall retrieval algorithm for commercial microwave links: stochastic calibration. *Atmospheric Measurement Techniques*, *15*(2), 485–502. doi: https://doi.org/10.5194/amt-15-485-2022.

World Bank Group. (2021). Climate risk country profile: Sri Lanka. *The World Bank Group and the Asian Development Bank*.

Appendices

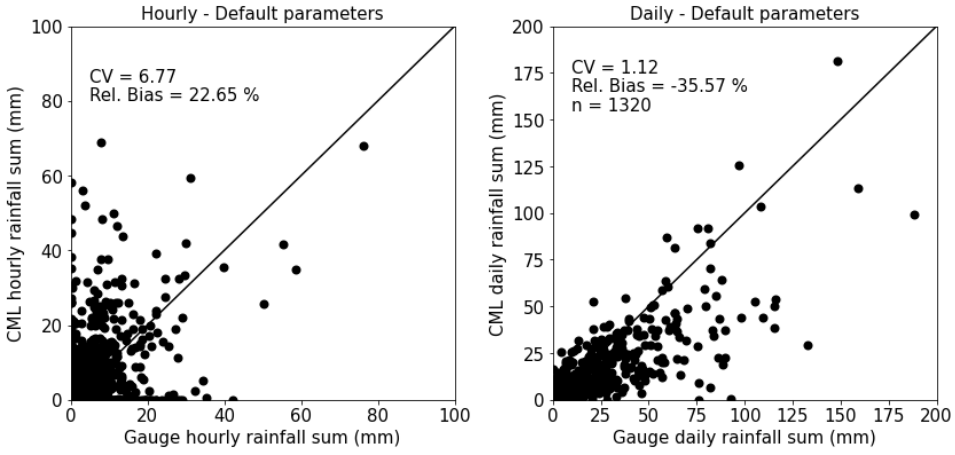
A

Verification CML rainfall estimations

Figure A.1 shows the verification of the CMLs for the full period (calibration and validation) with default parameters based on interpolated rainfall depths at the locations of the rain gauges with (a) hourly and (b) daily rainfall depths. The figure of the hourly values is the same figure as Figure 6.4a, because for the hourly values the full period was already used. Zero values have been included. The coefficient of variation (CV) and the relative bias are calculated. The CV is defined as the standard deviation of the residuals divided by the mean of the reference. The results are compared with the results obtained from the paper of Overeem et al. (2021). The compared results of the CV value and relative bias are lister in Table A.1. The differences in verification values between the two studies can be explained by the different number of rain gauges that were taken into account. In this study, 8 hourly and 12 daily stations are used. In the study of Overeem et al. (2021), 11 hourly and 16 daily rain gauges were used.

Table A.1: Comparison of verification values CV and relative bias between study of Overeem et al. (2021) of this thesis both with default parameters (September 12 and December 31, 2019).

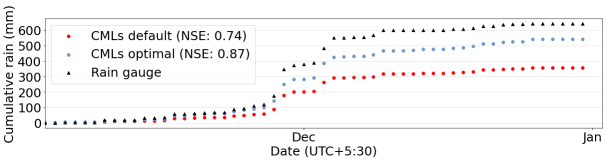
	Hourly values		Daily values	
Study	Overeem et al. (2021)	Current study	Overeem et al. (2021)	Current study
CV [-]	4.33	6.77	0.87	1.12
Rel. bias [%]	2.1	22.65	-17.6	-35.6



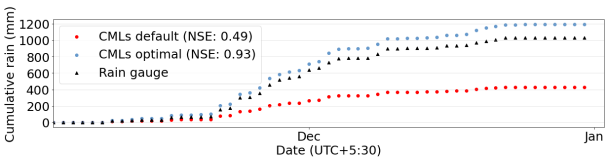
(a) Hourly **(b)** Daily

Figure A.1: Verification of CML based interpolation rainfall depths at the locations of the (a) 8 hourly stations and (b) 12 daily stations in the period between September 12 and December 31, 2019. Zero values are also included.

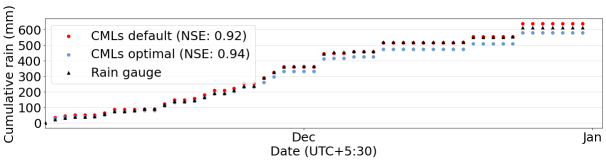
Figure A.2 gives the validation for all rain gauges and the rainfall estimation made with the selected neighbouring links. Figure A.3 gives for every rain gauge for every combination of parameters A_a and α the calculated NSE (-) with the selected neighbouring links.



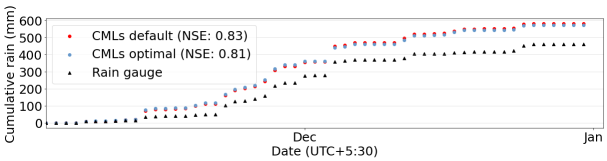
(a) Anuradhapura



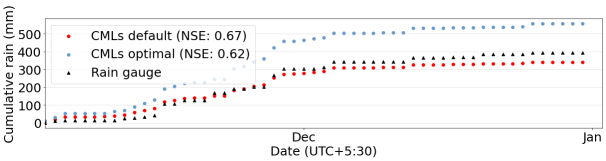
(b) Batticaloa



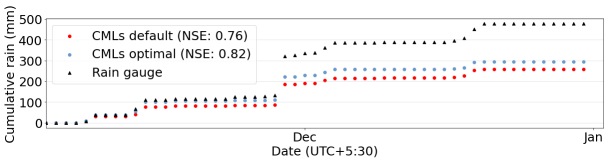
(c) Colombo



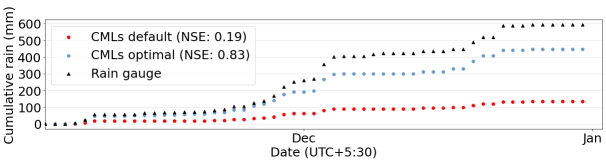
(d) Jaffna



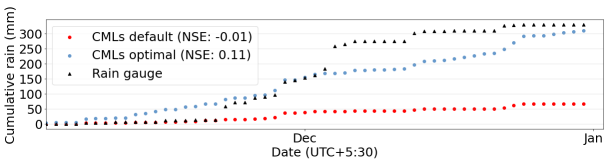
(e) Katunayake



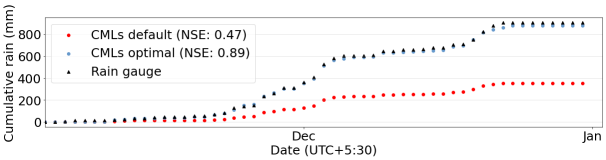
(f) Kurunegala



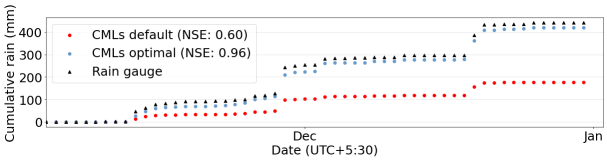
(g) Mahailluppalama



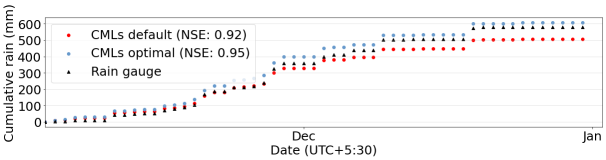
(h) Mannar



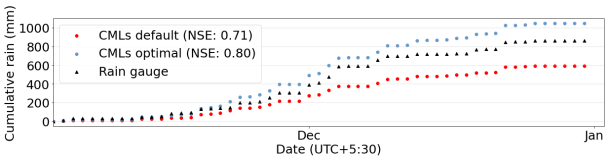
(i) Polonnaruwa



(j) Puttalam

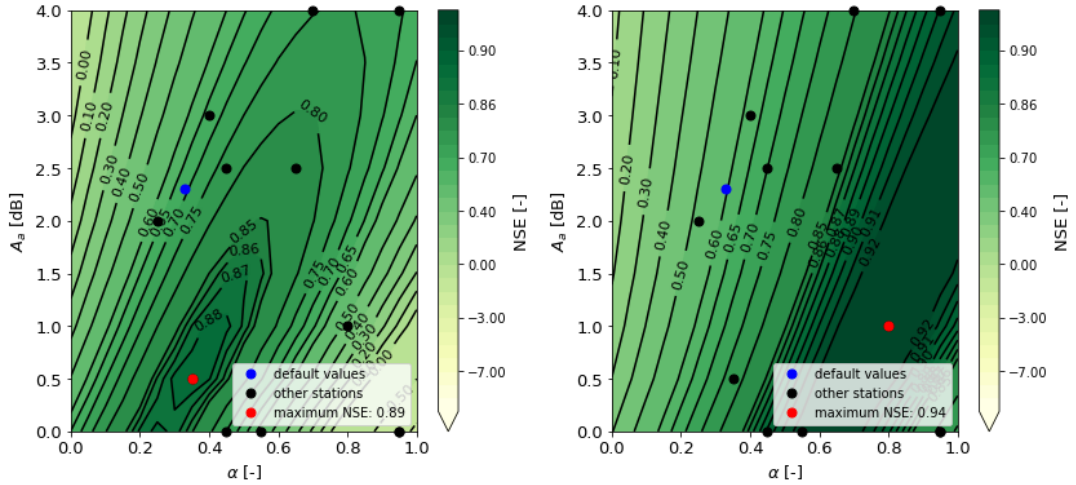


(k) Ratmalana



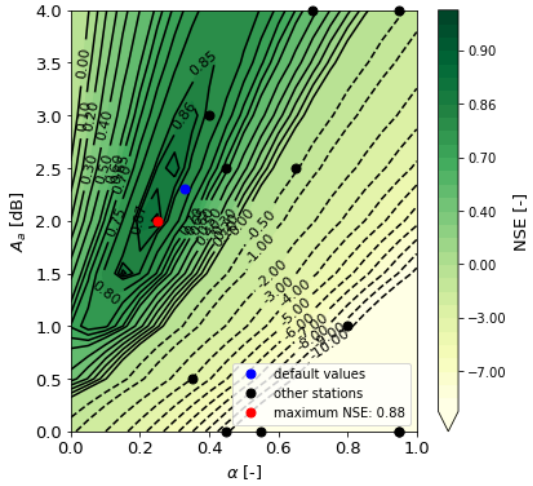
(l) Trincomalee

Figure A.2: Validation period of cumulative rainfall depths of the CML estimations with default and optimized parameters and the rain gauge data of every rain gauge separate.

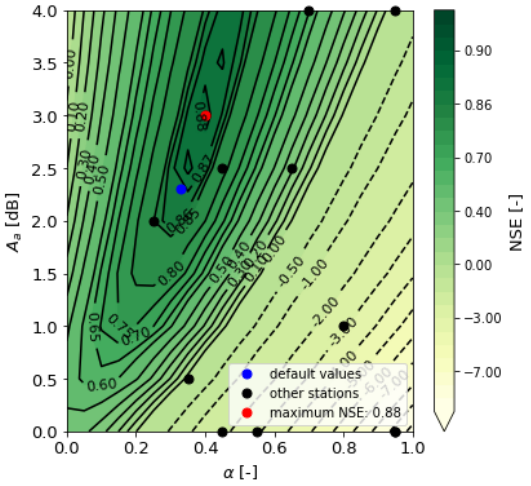


(a) Anuradhapura

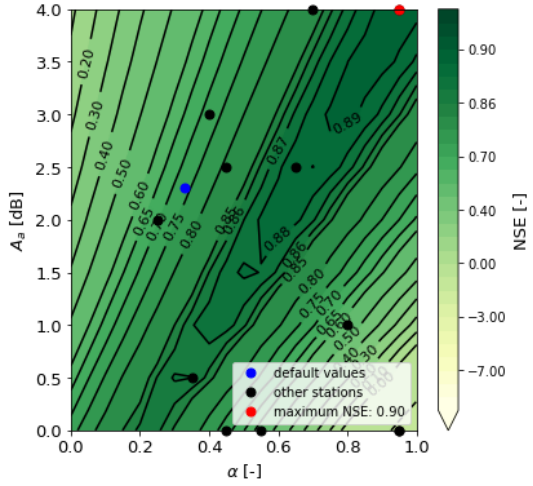
(b) Batticaloa



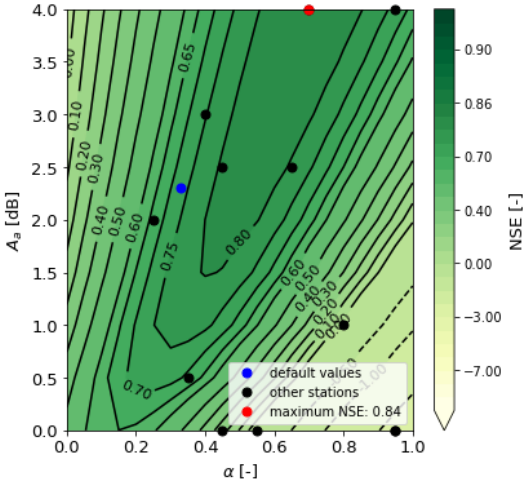
(c) Colombo



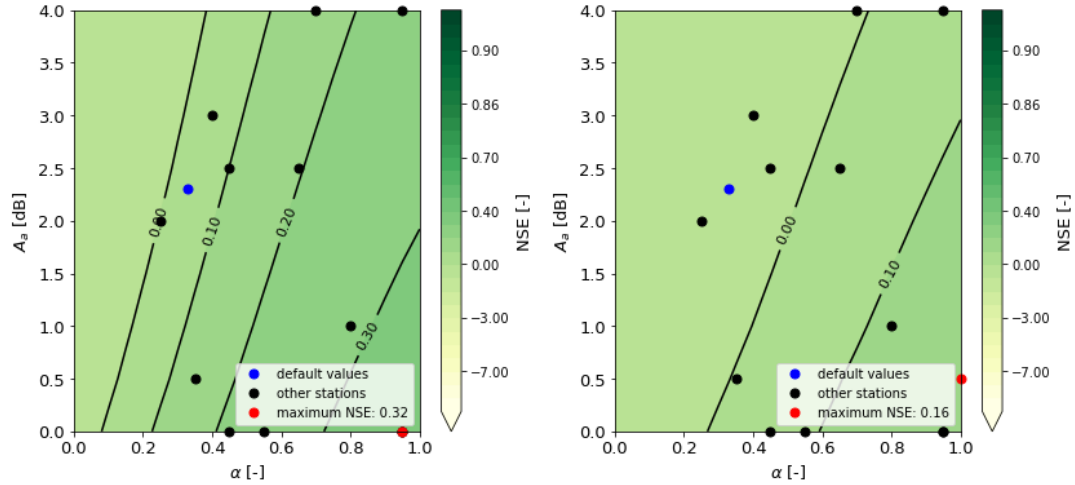
(d) Jaffna



(e) Katunayake

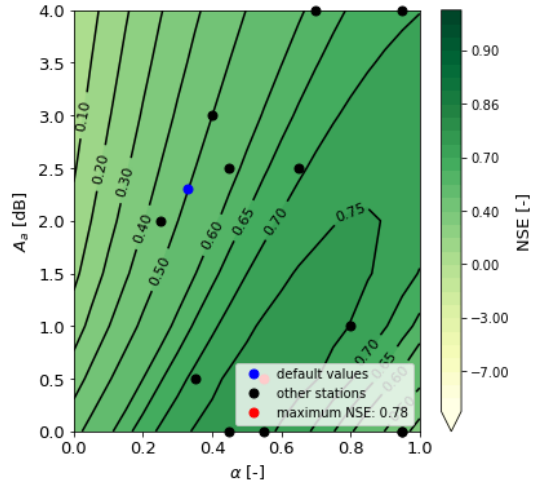


(f) Kurunegala

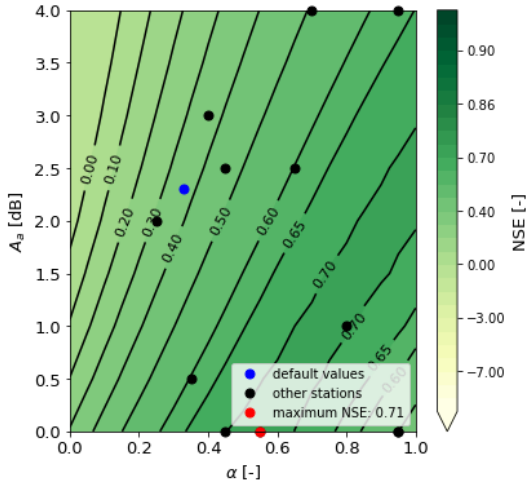


(g) Mahailluppalama

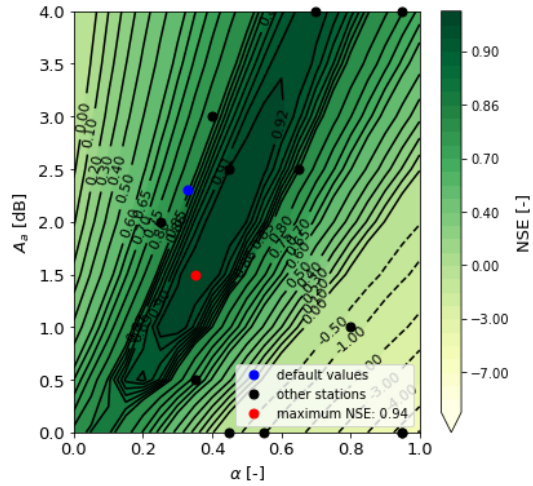
(h) Mannar



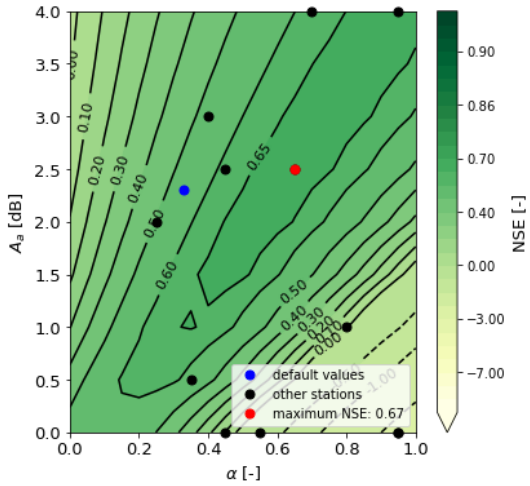
(i) Polonnaruwa



(j) Puttalam



(k) Ratmalana



(l) Trincomalee

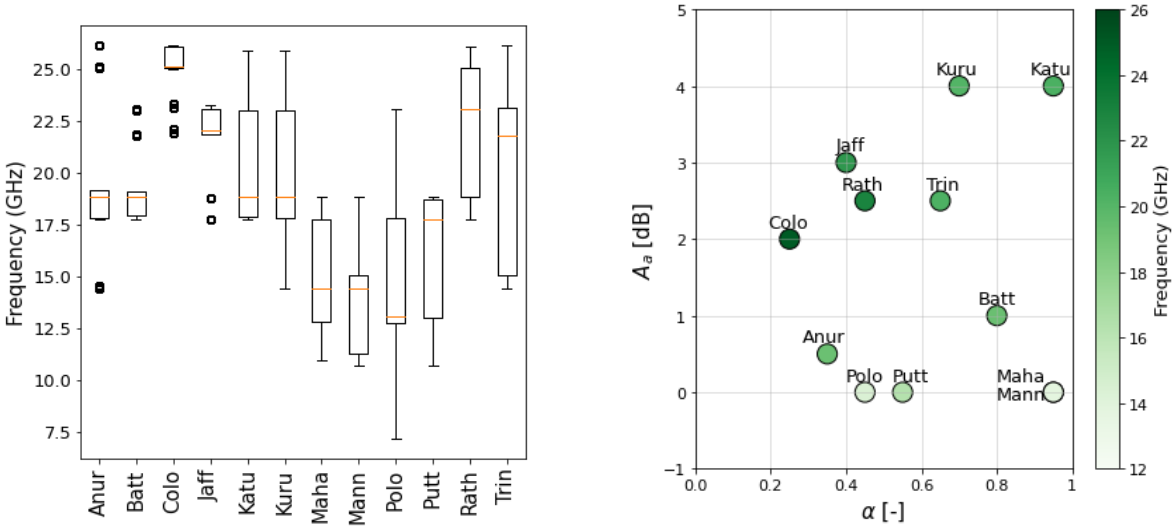
Figure A.3: NSE values of all gauges for all combinations of A_a and α.

B

Relation parameters with other factors

B.1. Dependency on link frequency

Figure B.1a illustrates the spread of the frequencies (GHz) of the selected neighbouring links for all the selected rain gauges. In Figure B.1b the average frequency of neighbouring links for each rain gauge location is shown including the values of the optimized two parameters A_a and α. It shows similar results to Figure 6.5b, where the path length of the selections of the links are given in relation to the two parameters.



(a) Frequency spread

(b) Relation frequency, A_a and α

Figure B.1: Information of all the rain gauges of (a) the frequency spread and (b) the frequency corresponding to the two optimized parameters (A_a and α).

B.2. Dependency on climatic zones

Figure B.2 illustrates the climate zones of the rain gauges and the location of the combination of the two calibrated parameters A_a and α.

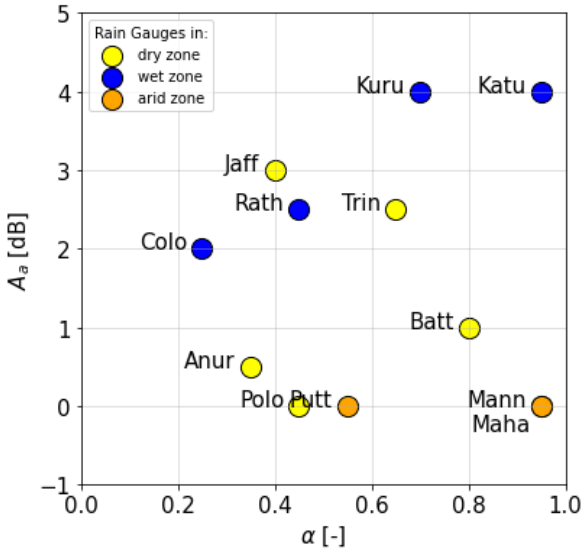
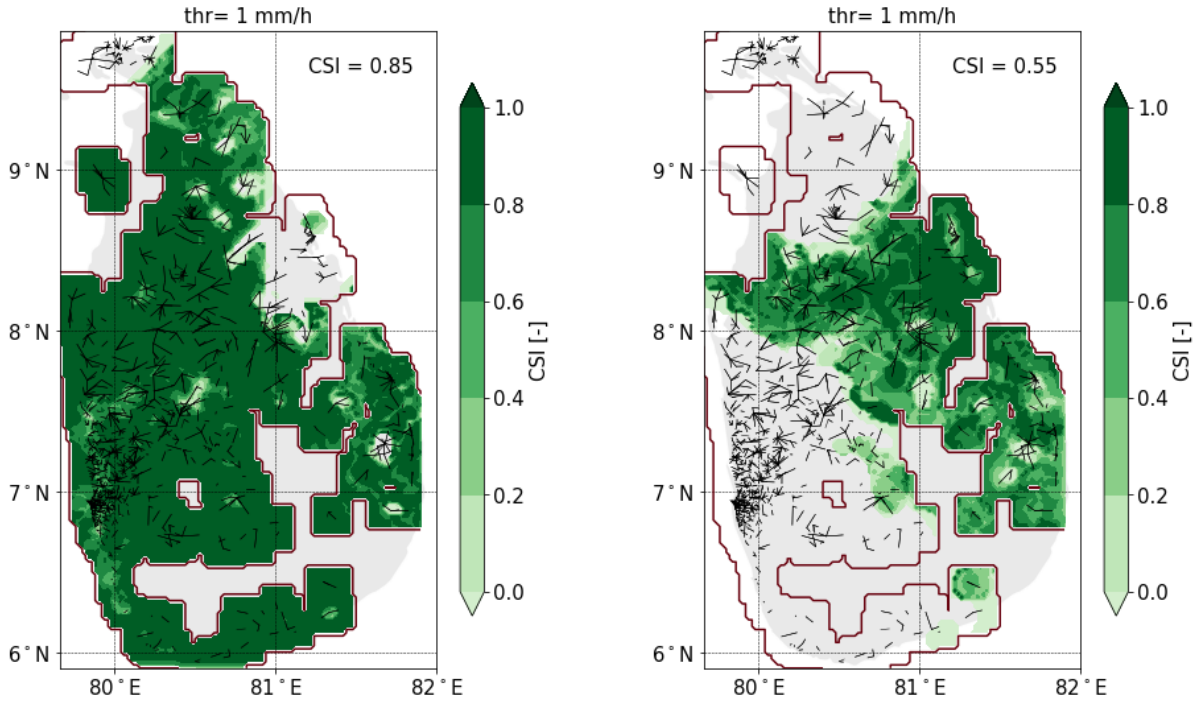


Figure B.2: Relation of the two calibrated parameters A_a and α with three different climate zones: dry, wet and arid.

C

Verification nowcasting

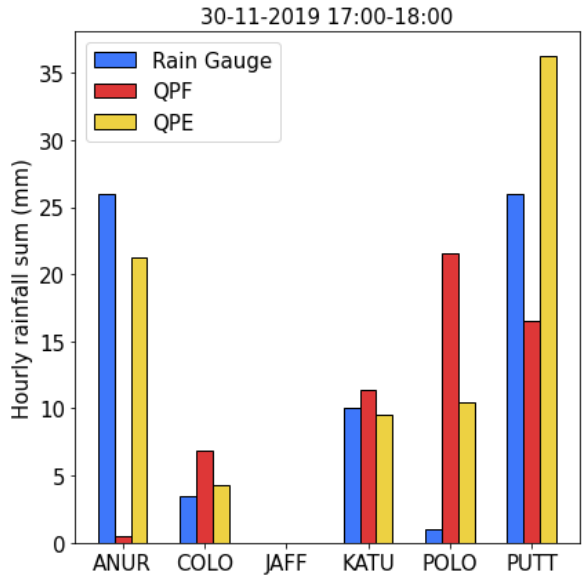
Figure C.1 gives the comparison of two events (November 30 17:00-20:00 and October 12 17:00-20:00) with the difference in CSI values. The average CSI value for the event on November 30 equals 0.78 and for October 12 this equals 0.27. Figure C.2 shows the event of November 30 and shows for each hour the QPE, QPF and rain gauge estimation for every hourly rain gauge.



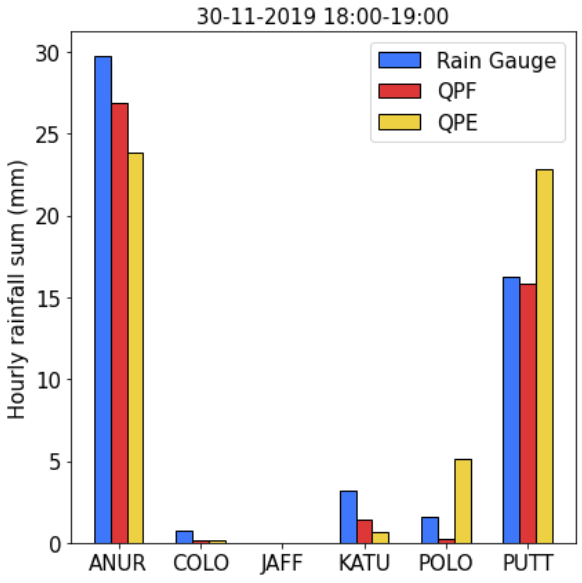
(a) November 30 (17:00-20:00)

(b) December 3 (16:00-19:00)

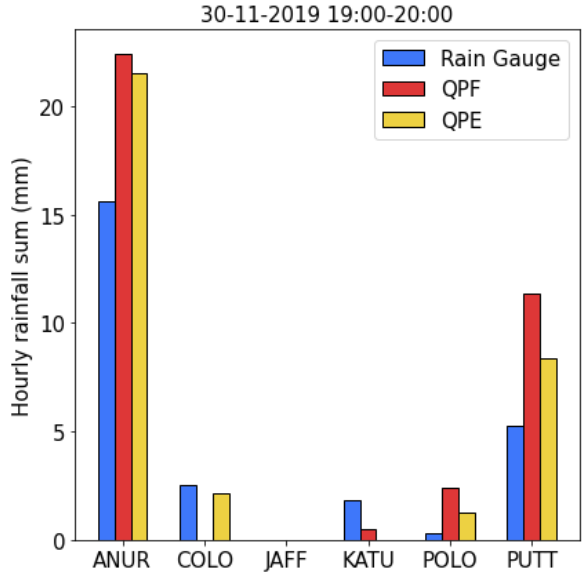
Figure C.1: Comparison of two events with different CSI scores: (a) November 30 17:00-20:00 and (b) December 3 16:00-19:00 in 2019.



(a) 17:00-18:00



(b) 18:00-19:00



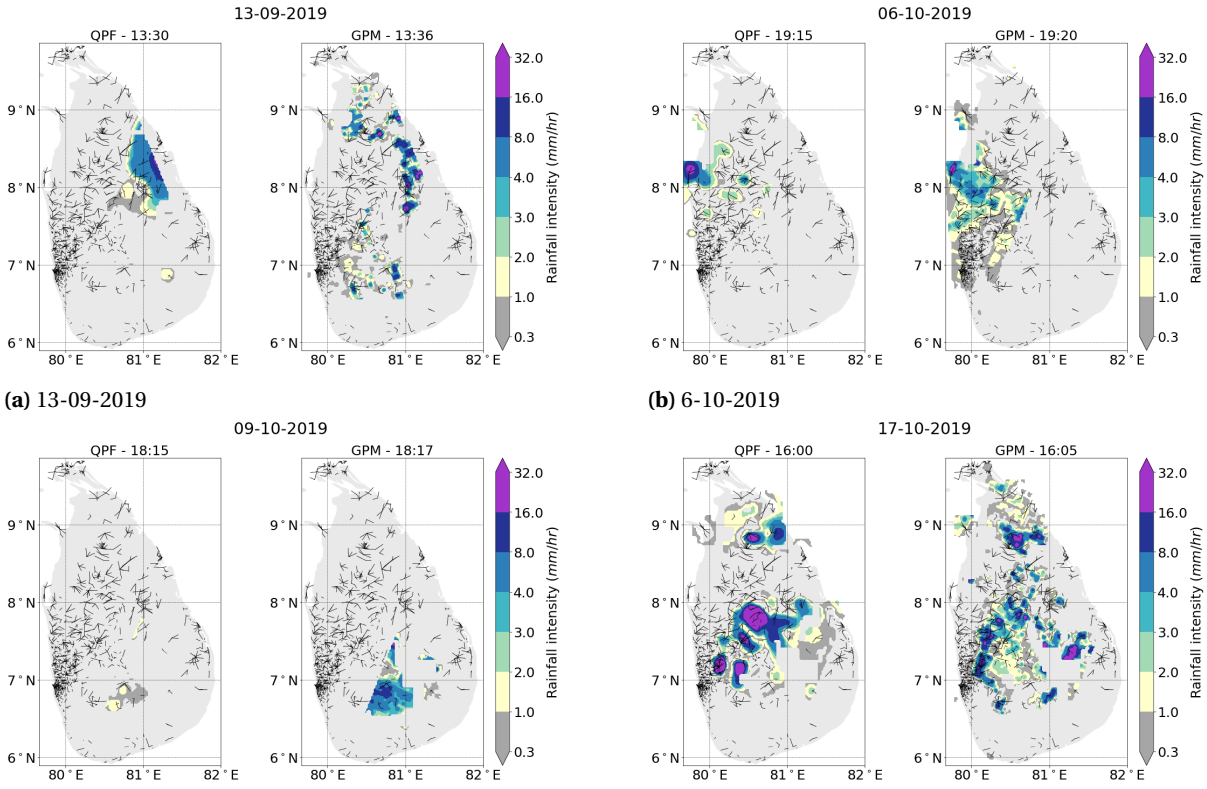
(c) 19:00-20:00

Figure C.2: Comparison QPE and QPF (Pysteps-D) rainfall intensities with the rain gauges for the event of November 30 17:00-20:00.

D

Comparison with GPM (Satellite data)

In Figure D.1 all the selected quantitative precipitation estimations from the satellite data are compared with the Pysteps-D nowcast of a lead time of 1 hour.



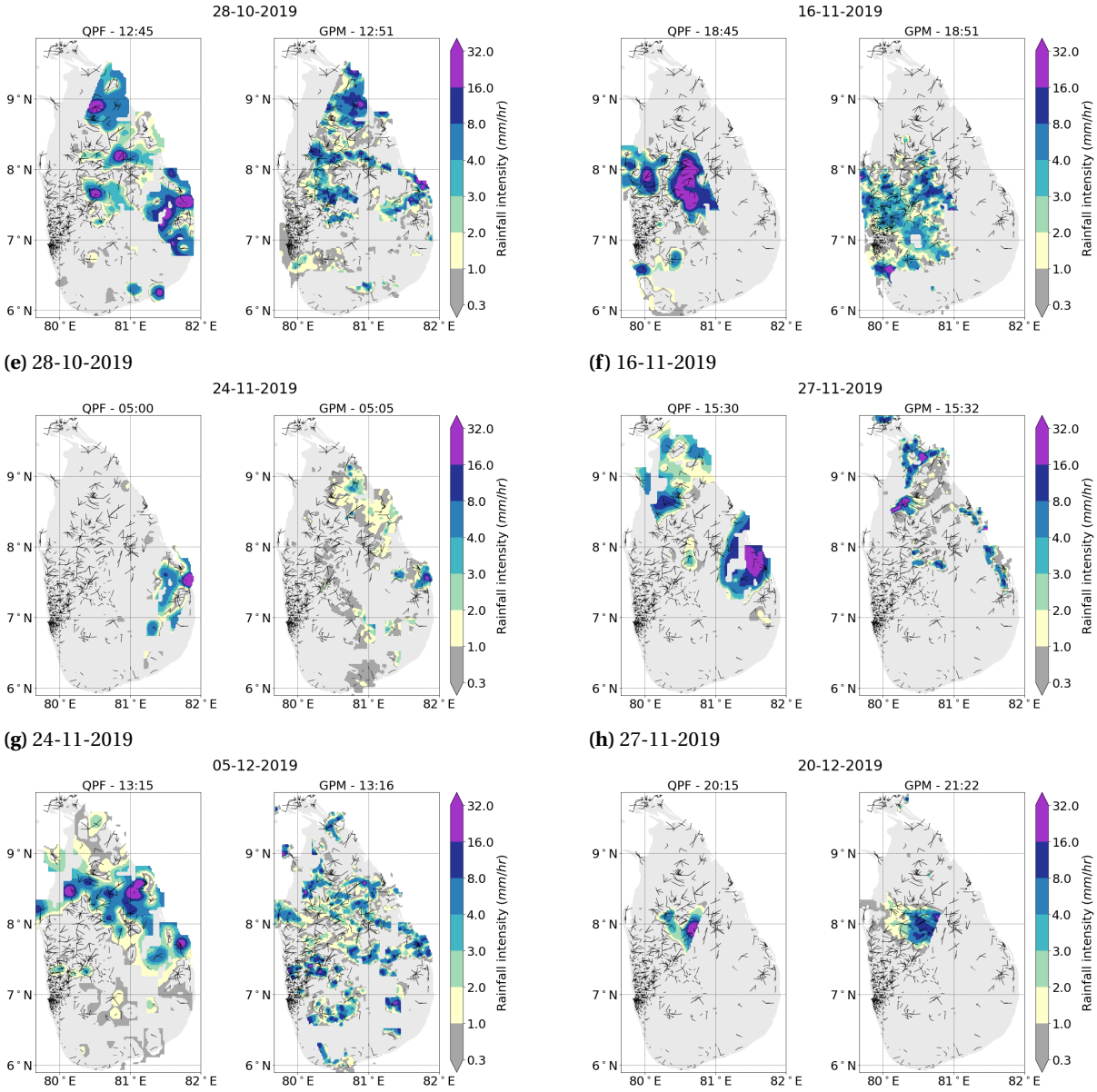
(a) 13-09-2019

(b) 6-10-2019

(c) 9-10-2019

(d) 17-10-2019

Figure D.1: Comparison QPF with GPM satellite product for 10 rainfall events with a lead time of 1 hour.



(e) 28-10-2019

(f) 16-11-2019

(g) 24-11-2019

(h) 27-11-2019

(i) 5-12-2019

(j) 20-12-2019

Figure D.1: Comparison of a Pysteps-D nowcasts with a lead time of 1 hour with GPM satellite product for 10 rainfall events.