



MONASH University

DOCTORAL THESIS

Improved rainfall measurement using microwave links

Author:

Jayaram PUDASHINE

Master's in Water Engineering

Supervisors:

Prof. Jeff WALKER

Assoc. Prof. Valentijn PAUWELS

Dr. Adrien GUYOT

Dr. Mahesh PRAKASH

A thesis submitted for the degree of Doctor of Philosophy at

Monash University in March 2021

MONITORING, PREDICTION AND PROTECTION

Department of Civil Engineering

Copyright notice

© Jayaram Pudashine (2021). Except as provided in the Copyright Act 1968, this thesis may not be reproduced in any form without the written permission of the author.

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Synopsis

Precise and timely rainfall estimates are crucial not only for real-time flood forecasting but also for various water management activities. Microwave links from telecommunication operators have been proposed as a complementary and cost-effective means to provide valuable information concerning the space-time variability of rainfall. Especially in urban areas, where there is a dense network of such microwave links, they can provide an untapped resource of rainfall estimates at an intermediate scale between point measurement from rain gauges and weather radar data. Although the theoretical basis of such rainfall estimation from microwave links is well studied and understood, real-world applications in an operational setting have been very limited due to uncertainties and errors associated with this technique. Therefore, the focus of this thesis is to develop a better understanding of uncertainties associated with such a technique and ultimately propose a methodology to improve the rainfall estimation using such opportunistic data.

First, a dedicated experimental microwave link was set up close to the city of Melbourne. This setup was complemented by 2 laser disdrometers, 5 automated rain gauges and 2 weather stations positioned at several locations along the path. Two and half years of the data collected from the 2 laser disdrometers were used for deriving the parameters for the power relationship between rainfall and attenuation at different frequencies from 1 to 75 GHz. This was later used for validating the rainfall retrieval algorithm for the selected 72 different rainfall events. The overall result showed a good correlation between the observed and estimated rainfall rates, however there was an overestimation of about 90%. Later, the same experimental data were sampled to three different commonly used sampling strategies to understand the impact on the rainfall retrieval. It was found that the average sampling strategy for both the 5 and 15-minute periods provided rainfall estimates closer to the ground observation compared with the minimum/maximum and instantaneous strategies. Also, the impact of data quantization for the 24 GHz frequency microwave link was found to be negligible. Based on a simple wet antenna experiment, an additional attenuation in the order of up to 3.5 dB was observed. When comparing the three different wet antenna attenuation models, there was a substantial improvement in the rainfall estimates. However, when the parameters derived based on the experimental microwave link were used in the commercial microwave link, the constant wet antenna attenuation model performed better.

Second, rainfall retrieval using the commercial microwave link (CML) was assessed. More than 100 CMLs with their microwave frequency ranging between 10 and 40 GHz were used for rainfall retrieval. The 15-minute received signal levels (RSLs) for each CML for 2 years provided a unique dataset to compare the performance of the rainfall retrievals using the two sampling strategies (average and minimum/maximum). The open-source algorithm RAINLINK was used for deriving rainfall from the 15-minute RSL data. From two years of data, a subset of 30 rainy days distributed across this period was used for calibrating the RAINLINK parameters, with the remaining data used for validation. For this study, only path-averaged rainfall intensities were evaluated based on a gauge-adjusted radar product serving as the reference. The result of the wet-dry classification showed that the minimum and maximum RSL data performed better, with a lower false alarm ratio and a higher Matthews correlation coefficient than average RSL data. For the rainfall retrieval, both datasets showed a similar correlation with the gauge adjusted radar product. However, based on other statistics (RMSE, bias and CV) the minimum and maximum RSL data outperformed the average for the rainfall retrieval.

In the last part of this research, a deep learning approach was developed to improve rainfall retrieval. A two-layer Long Short-Term Memory (LSTM) network model with 30 neurons each was designed and implemented. This model was trained with a one-year disdrometer dataset comparable to equivalent data provided by telecommunication companies. For the test dataset, the Root Mean Square error between the observation and the results from the LSTM model was reduced from 4.4 mm h^{-1} to 0.67 mm h^{-1} , the relative bias reduced from 125% to 3%, and the coefficient of determination increased from 0.60 to 0.96. This trained LSTM model was then applied to data from a nearby CML with substantial improvements in rainfall estimation.

Declaration of Authorship

I, Jayaram Pudashine, declare that this thesis titled, ‘Improved rainfall measurement using microwave links’ and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at the University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always cleared attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date: 16/03/2021

Acknowledgements

This research work would not have been possible without the help of many people in various ways. Therefore, I would like to express my gratitude to all the helping hands who supported me during my whole PhD journey.

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Jeffrey Walker for his continuous support and encouragement for my PhD studies. During the past four years, he was always there to provide support and help whenever I had problems. I deeply appreciate for his dedication, patience and the time he spent reading and providing comments on my every draft manuscript. Beside my supervisor, I would like also to like express my heartfelt gratitude to my co-supervisors, Dr. Adrien Guyot, Assoc. Professor Valentijn Pauwels and Dr. Mahesh Prakash. I am always indebted to Dr. Guyot for his quick feedbacks and innovative ideas to handle problems during my PhD studies.

I would also like to thank Professor Remko Uijlenhoet who supported for my exchange program to Wageningen University, The Netherlands. Also, thankful to Dr. Aart Overeem for his constructive comments for drafting the article.

I would especially like to thank those who have assisted me in setting up the experimental site. A special thanks goes to Kiri Mason, Long Goh, Richard Williamson and Saravnan Mani for helping me to prepare necessary equipment, their invaluable efforts in the installation and data collection. I would like to like to thank Michael O'Hagan, Geoff Eley for providing data. Also, I would like to thank Dr. Alan Seed and Dr. Carlos Velasco from the Bureau of Meteorology for their support and encouragement to finish my thesis on time.

I am also grateful to the financial support I received for my research from Monash University, Australian Research Council and a top up scholarship from the CSIRO – Data 61. Also, thanks to my wonderful Nepalese friends Siddhartha Shrestha, Grishma Khadka, Lavaraj Devkota, and Prakash Aryal for their continuous support throughout the whole journey.

Finally, my deepest thanks goes to my beloved wife, my lovely son and my parents for their love, care, comfort, understanding and encouragement throughout the PhD journey. They always had confidence in me and stood behind me no matter what happened. I would like to dedicate this thesis to my parents for the greatest support they have been giving to me ever since I was born.

Table of Contents

Copyright notice	i
Synopsis.....	ii
Declaration of Authorship	iv
Acknowledgements	v
Table of Contents	vi
List of Figures	x
List of Tables.....	xvii
Abbreviations	xviii
Chapter 1 Introduction.....	1
1.1 Importance of rainfall measurement	1
1.2 Statement of the problem.....	2
1.3 Objectives and scope of research.....	4
1.4 Outline of the approach	5
1.5 Structure of the thesis	6
Chapter 2 Literature review.....	8
2.1 Conventional rainfall measurement techniques	8
2.1.1. Rain gauges	8
2.1.2 Weather radar	8
2.1.3 Satellite.....	9
2.2 Rainfall measurement using microwave link	9
2.3 Theoretical background on microwave attenuation.....	12
2.3.1 Microwave attenuation in the near-surface atmosphere	12
2.3.2 Review of the rainfall retrieval algorithm.....	16
2.4 CML studies around the world	20
2.5 Knowledge gap and the proposed approach	27

2.6 Chapter Summary	28
Chapter 3 Experimental setup and dataset used in this study.....	29
3.1 Experimental setup	29
3.1.1 Research microwave link	29
3.1.2 Auxiliary instruments.....	34
3.2 Other data.....	36
3.2.1 Commercial microwave link data	36
3.2.2 Weather radar data	38
3.3 Chapter Summary	38
Chapter 4 Uncertainties associated with microwave link rainfall estimation	39
4.1 Background.....	39
4.2 Study area and data used.....	41
4.3. Data Processing	41
4.3.1 Disdrometer data	41
4.3.3 Rainfall events.....	44
4.3.4 Data sampling and quantization.....	46
4.3.5. Wet antenna attenuation.....	46
4.4 Results and discussion	49
4.4.1 Rainfall -specific attenuation relationship	49
4.4.2 Validation of the rainfall retrieval algorithm	49
4.4.3 Impact of sampling strategies	61
4.4.4 Impact of quantization errors	66
4.4.5 Wet antenna attenuation experiment.....	67
4.4.6 Modelling wet antenna attenuation	69
4.5 Chapter summary.....	74
Chapter 5 Validation of rainfall retrieval using commercial microwave links	75
5.1 Background.....	75

5.2 Study area and data used.....	77
5.2.1 Study area.....	77
5.2.2 Data used.....	77
5.3 Methodology.....	77
5.3.1 Preliminary data processing and quality check.....	77
5.3.2 Use of RAINLINK.....	78
5.3.3. Calibration of the RAINLINK parameters.....	81
5.3.4 Performance metrics	84
5.4. Results.....	86
5.4.1 Calibration.....	86
5.4.2 Validation result	87
5.5 Discussion.....	99
5.5.1 Optimized parameters for RAINLINK	99
5.6 Chapter summary.....	101
Chapter 6 Deep Learning model for improving rainfall estimates.....	102
6.1 Background.....	102
6.2 Study area and data.....	104
6.3 Methodology.....	104
6.3.1 Long short-term memory network	104
6.3.2 Modelling approach	106
6.3.3 Rainfall retrieval from CML data	107
6.4 Results.....	108
6.4.1 Verification result	108
6.4.2 Performance of various architectures of the deep learning model.....	110
6.4.3 Sensitivity analysis.....	112
6.4.4 Application of the disdrometer-trained LSTM model to the CML dataset.....	114
6.5 Summary.....	114

Chapter 7 Conclusions and future work	115
7.1 Conclusions.....	115
7.1.1 Uncertainties associated with microwave link rainfall estimates	115
7.1.2 Rainfall retrieval from commercial microwave links	116
7.1.3 Deep learning for improved rainfall estimates.....	117
7.2 Limitations and Future work	117
Appendix	120
References	121

List of Figures

Figure 1-1:	Example of a commercial microwave link antenna tower, used in telecommunication networks.	3
Figure 2-1:	Use of microwave links for the mobile communication.	10
Figure 2-2:	Specific attenuation of a standard atmosphere and dry air (ITU-R, 2016)...	13
Figure 2-3:	Frequency dependency of parameters a and b as proposed by ITU-R and Leijnse et al. (2007) in the power-law relationship.	16
Figure 3-1:	(a) Location of the experimental setup (the orange shaded circle provides the indicative location of the experimental site in the Greater Melbourne region). (b) Experimental setup including transmitter and receiver antennas. Disdrometers were installed at each end of the microwave link. Three tipping bucket rain gauges were installed along the length of the link.	30
Figure 3-2:	(a) Transmitter antenna (antenna within the red circle) installed at the Melbourne Water Reservoir placed on a 25 m high water tower, and (b) receiver antenna installed at Lakeside drive on a 2 m mast together with the disdrometer.	31
Figure 3-3:	Received signal power vs received voltage data used for calibration of the detector board for RAL link (24 GHz)	32
Figure 3-4:	Profile of the experimental link. The hatch lines indicate the ground profile. The transmitter antenna was installed at the 0 km mark and the receiver antenna at the 3.79 km mark.	32
Figure 3-5:	Time series of a rainfall event on the 9th to 10th May 2018 showing: (a) received signal level; (b) Specific attenuation measured by two disdrometers.	33
Figure 3-6:	The OTT PARSIVEL1 disdrometers installed at each end of the experimental link path. (a) Disdrometer installed at the Melbourne Water Mount View reservoir and (b) Disdrometer installed at Lakeside Driver, Burwood.	34
Figure 3-7:	Tipping bucket rain gauge installed along the experimental path; (a) and (b) RG3-M rain gauges installed on the edge of one-story rooves, (c) TB3 rain gauge installed on the top of a one-story building. All three rain gauges were installed following the guidelines of the WMO.	35

Figure 3-8:	Commercial microwave links (CML) from one of the operators for Greater Melbourne. Red lines indicate microwave links having a microwave frequency greater 10 GHz (which were used for this study). Green lines indicate microwave links with frequencies less than 10 GHz (which were not used for this study). On the top left is a plot of the frequency of the microwave links f (GHz) against the path length L (km) for the 144 links used for this study.....	36
Figure 3-9:	Time series of a rainfall event on the 16 th to 19 th June 2018 showing: (a) 15-min average received power level; (b) 15-min minimum and maximum received power levels; and, (c) Path-averaged observed rainfall rate along the microwave link.	37
Figure 4-1:	Distribution of rainfall depth along the RAL path for 72 rainfall events. (Distances from each of the rain gauges to the Site T (transmitter antenna) were: 0.88 km from TB1, 2.16 km from TB2, and 3.01 km from TB3).	45
Figure 4-2:	Time series of an event on the 3 rd and 4 th May 2018 of (a) received signal level and the reference signal (green dashed line); the shaded colour shows the 5 th and 95 th percentile of the reference signal during the dry period; (b) specific attenuation derived based on the RAL data and theoretical specific attenuation derived for the RAL frequency using the processed disdrometer data; (c) rainfall intensity derived from the RAL data using the $R-k$ power law and path-averaged rainfall intensity using three tipping bucket rain gauges and two disdrometers; (d) variation of relative humidity (green line) and average temperature (red line) for the same event; (e) average wind speed (blue) and wind direction (brown) at the transmitter site.....	51
Figure 4-3:	Accumulated rainfall plot for an event on 3 rd and 4 th May 2018.....	52
Figure 4-4:	Time series of an event on the 3 rd and 4 th May 2018 using the 15-minute dataset of (a) received power levels (solid lines) from the CML, the RAL and the reference signal (dotted line); (b) rainfall intensities derived from the link attenuation using the $R-k$ power law.....	53
Figure 4-5:	Same as Figure 4-2 but for an event on 18 th August 2018.....	54
Figure 4-6:	Same as Figure 4-4 but for an event on the 18 th August 2018.....	55
Figure 4-7:	Same as Figure 4-2 but for an event on the 16 th and 17 th June 2018.....	56
Figure 4-8:	Same as Fig 4-4 but for an event on the 16 th and 17 th June 2018.	58

Figure 4-9:	(a) Scatterplot of rainfall rate estimated for the RAL versus path-averaged ground observation based on the 1-minute dataset; (b) double mass curve derived based on the RAL and path-averaged ground observation; and, (c) scatter plot of the specific attenuation measured from the RAL versus based on two disdrometers at two ends.	59
Figure 4-10:	Scatter plot of total rainfall depth from the RAL and ground observation for all 72 rainfall events; each data points on the scatter plot represent one rainfall event.....	59
Figure 4-11:	Scatterplots of the link derived rainfall intensities versus path-averaged ground observed rainfall as obtained from two disdrometers and three tipping bucket rain gauges for (a) 1-minute, and (b) 15-minute.	60
Figure 4-12:	Complementary cumulative distribution function of path-average ground observation and derived based on the RAL and CML.	61
Figure 4-13:	(a)-(c) Scatter plot for 5-minute sampling derived based on the RAL versus path-averaged ground observation; (d) Average versus MinMax; (e) Average versus Instantaneous; and (f) MinMax versus Instantaneous.	62
Figure 4-14:	Same as Figure 4-13 but for 15-minute sampling.	63
Figure 4-15:	Complementary cumulative distribution plot of rainfall depth for (a) 5 and (b) 15-minute sampling data.....	64
Figure 4-16:	Scatter plot of (a) 5-min Average vs 1-min Average; (b) 5-min MinMax vs 1-min Average; (c) 5-min Instantaneous vs 1-min Average; (d) 15-min Average vs 1-min Average; (e) 15-min MinMax vs 1-min Average; (f) 15-min Instantaneous vs 1-min Average; (g) 15-min Average vs 5-min Average; (h) 15-min MinMax vs 5-min MinMax; and (i) 15-min Instantaneous vs 5-min Instantaneous	65
Figure 4-17:	Distribution of rainfall based on three sampling strategies for 5 and 15-minute compared with observed and reference (1-minute).	66
Figure 4-18:	(a) Scatter plot of rainfall depth based on the RAL data with no quantization vs 1 dB quantized data; and (b) Complementary cumulative distribution function for various level of quantized data.	66
Figure 4-19:	Photograph of the receiver antenna just after receiver antenna sprayed with water during the wet antenna experiment. Water droplets formation was observed on the antenna film.	67

Figure 4-20:	Time series of the wet antenna simulation experiment conducted on 23rd August 2019. Three experiments were conducted at different times (dotted line on the plot indicate the start time of the water spray [I :10:16 AM, II :10:58 AM and III: 11:14 AM]); (a) received signal level with the baseline signal plotted on orange colour; (b) specific attenuation of the RAL link; and (c) rainfall equivalent using the R-k relationship.	68
Figure 4-21:	Complementary cumulative distribution function of (a) Rainfall measured by the RAL and ground observation, and (b) total path attenuation. For attenuation, the ground observation was obtained based on the theoretical raindrop attenuation.	69
Figure 4-22:	Measured attenuation related to the path-averaged rainfall intensities shown at log-scale. Pearson’s correlation (ρ) and Spearman rank correlations (ρ_s) are displayed on the plot.	71
Figure 4-23:	(a) Fitting of WAA of antenna pairs on rate for RAL using the least squares method, rain rate plotted on a log scale and; (b) CCDFs of attenuation based on the RAL and ground observation based on the theoretical raindrop attenuation with the correction applied.....	71
Figure 4-24:	Scatter plot of rainfall intensity derived from the RAL and path-averaged ground observation after applying the wet antenna correction based on (a) the power-law model of Valtr et al. (2019); (b) the exponential decay model of Garcia-Rubia et al. (2011); and (c) a constant wet antenna attenuation of 1.1 dB.....	72
Figure 4-25:	Double mass curve of the RAL derived rainfall versus path-averaged ground observation after applying WAA.....	72
Figure 4-26:	Scatter plot of the CML data after applying wet antenna correction using (a) the power law model; (b) the exponential decay model; and (c) the constant attenuation model.....	73
Figure 4-27:	Distribution of accumulated rainfall for 72 rainfall events based on analytical method compared with the empirical approach and observed data.	74
Figure 5-1:	Example of some suspicious microwave link data: (a) There is noise observed in the minimum (green) and maximum (yellow) RSL; (b) There is a sudden drop in the received power level even if the transmitted power remains constant. Blue lines indicate observed rainfall rates from <i>Rainfields</i>	78

Figure 5-2: (a) Coefficients a and (b) exponent b of the power-law relation between R and k for both horizontally and vertically polarized signals for frequencies ranging from 1 to 70 GHz. The values recommended by the International Telecommunication Union, Radio communication (ITU-R, 2005) for computing specific attenuation for given rain rates for world wide application are shown in dotted and dashed lines. 81

Figure 5-3: Sensitivity analysis using MinMax RSL data for: (a) Threshold median; (b) Threshold median per unit length; (c) Wet antenna attenuation; and (d) Alpha coefficient. 83

Figure 5-4: Time-series of (a) average RSL; (b) minimum and maximum RSL; and (c) gauge adjusted radar rainfall intensities for a selected rainfall event for LinkID 62. The wet-dry classification using the calibrated parameters is shown as shaded colours. 89

Figure 5-5: Normalized confusion matrix for wet-dry classification for (a) Average data; and (b) MinMax data 90

Figure 5-6: Box plots showing the performance criteria for the wet-dry classification using the nearby links methods with: (a) Probability of detection (POD); (b) False alarm ratio (FAR); (c) Matthew correlation coefficient (MCC); and (d) Error rate (ERR) for Average and MinMax data. 91

Figure 5-7: Probability of detection (POD) based on a reference rainfall rate. Here five thresholds (0, 0.1, 0.5, 1.0 and 5.0-mm h⁻¹) for the gauge adjusted radar data are used to filter the data; thus this result provides POD result for only wet intervals..... 92

Figure 5-8: Validation criteria showing the ability of CMLs to detect rainy periods against reference data; (a) POD and FAR; and. (b) Matthew correlation coefficient (MCC) and Error rate (ERR). All four parameters were calculated for Average and MinMax for various accumulation intervals, using a threshold of 0.1 mm h⁻¹ to detect rain occurrence..... 92

Figure 5-9: Validation of path-average CML-rainfall against gauge-adjusted radar rainfall. Scatter density plot of link-derived rainfall with radar over intervals of (a) 15 minutes; (b) 1 hour; and (c) 3 hour. Only gauge-adjusted radar rainfall depths greater than zero were selected. 94

Figure 5-10:	Scatter density plots of path-average CML-rainfall against gauge-adjusted radar rainfall for 342 rainfall events.	96
Figure 5-11:	Double mass curves per event for all links: (a) Link-derived cumulative rainfall using average data versus cumulative gauge-adjusted radar rainfall; and (b) Link-derived cumulative rainfall using min/max RSL data versus cumulative gauge-adjusted rainfall data.	96
Figure 5-12:	Density plot of the double mass curve for all links during only the wet period (no false alarm included) for: (a) Average, and (b) MinMax data.	97
Figure 5-13:	Comparison of link-derived rainfall retrievals using 15-min Average vs MinMax RSL data for: (a) wet periods only; and (b) all time intervals with false alarm only.....	98
Figure 5-14:	Spatial distribution of the correlation coefficient (circle size) between the reference and CML-derived hourly rainfall accumulation at different frequencies (circle colour) for (a) Average; and (b) MinMax data.	99
Figure 5-15:	As for Figure 5-14 but for room mean square error (RMSE).....	99
Figure 6-1:	(a) General architecture of a two-layer sequence-to-sequence recurrent neural network. The output of the second layer for each of the time steps is fed into a dense layer to calculate the prediction ($y_1, y_2 \dots y_n$). (b) The internal architecture of the LSTM recurrent cell [Adapted from: Hu et al. (2018)].	105
Figure 6-2:	Comparison of the predicted to the observed (a, b) average attenuation based on disdrometer derived attenuation and (c,d) rainfall rate based on disdrometer derived attenuation estimates for the weighted average method with $\alpha=0.21$; and for the LSTM Model.	108
Figure 6-3:	(a) Time series of predicted and observed rainfall intensity based on disdrometer data from conventional models and LSTM for 2 and 3 December 2017. Bottom panels: Comparison of the predicted to the observed rainfall for the same days for (b) the conventional approach with $\alpha=0.31$; (c) the conventional approach with $\alpha=0.21$; (d) the LSTM model.	110
Figure 6-4:	LSTM model performance for (a) RMSE; (b) Relative percentage bias; (c) Coefficient of variation vs No. of hidden neurons; (d) RMSE; (e) Relative percentage bias; and (f) Coefficient of variation vs no. of hidden layers...	111

Figure 6-5: Scatter plot of the predicted *versus* observed rainfall intensity with +0.4 dB/km constant bias for (a) with $\alpha=0.31$; (b) with $\alpha=0.21$; and (c) LSTM model 113

Figure 6-6: Scatter plot of the predicted *versus* observed rainfall intensity with -0.4 dB/km constant bias (a) with $\alpha=0.31$; (b) with $\alpha=0.21$; and (c) LSTM model 113

Figure 6-7: Scatter plot of the predicted *versus* observed rainfall intensity with added random noise from -0.3 to +0.3 dB km⁻¹ for (a) a weighting factor with $\alpha=0.31$; (b) a weighting factor with $\alpha=0.21$; and (c) LSTM model. 113

Figure 6-8: Scatter plots of the predicted *versus* observed rainfall intensity using commercial microwave link data for: (a) the weighted average method with $\alpha=0.30$; (b) the weighted average method with $\alpha=0.21$; and (c) the disdrometer-trained LSTM Model..... 114

List of Tables

Table 2-1:	Published works on rainfall measurement using microwave links.....	21
Table 3-1:	Characteristics of experimental link antenna.....	31
Table 3-2:	Data storing frequency for the different instruments used in the experimental setup.....	33
Table 4-1:	Location of the instruments used for this study.....	42
Table 4-2:	Summary of the 72 rainfall events (rainfall rates were based on 1-minute stored frequency).	45
Table 4-3:	Coefficients and exponents (a and b parameters) of the $R-k$ relationship derived for both the RAL link (f : 24 GHz) and the CML (f : 22.7 GHz) for both horizontally and vertically polarised radiations [Unit of a is $\text{mm h}^{-1}\text{dB}^{-b}\text{km}^b$, b is unitless].	49
Table 4-4:	Comparison of the a and b parameters of the rainfall-attenuation relationship for three rainfall events.	57
Table 5-1:	List of variables used for calibration of RAINLINK. Here, the alpha coefficient provides the weightage between the minimum and maximum rainfall. The parameters a and b used in this relationship are based on a local disdrometer data.....	82
Table 5-2:	Confusion matrix based on gauge adjusted radar and CML.....	84
Table 5-3:	Calibration results for a selection of four of the RAINLINK parameters and comparison with values obtained for The Netherlands.	87
Table 5-4:	Performance criteria for the calibration period for Average and MinMax... ..	88
Table 5-5:	Performance of wet-dry classification for 15-min average and 15-min minimum/maximum RSL data.	89
Table 5-6:	Validation of 15-min and 1-hour accumulation link-derived rainfall against gauge-adjusted radar rainfall (reference) on a seasonal basis.....	95
Table 6-1:	Comparison of the performance of various model architectures for rainfall prediction for the training and test datasets based on the disdrometer.	111

Abbreviations

ACMA	Australian Communication and Media Authority
CCDF	Complementary Cumulative Distribution Function
CML	Commercial Microwave Links
CNN	Convolutional Neural Network
CV	Coefficient of Variation
DSD	Drop Size Distribution
ERR	Error Rate
FAR	False Alarm Rate
GEV	Generalized Extreme Value
GRU	Grated Recurrent Unit
ITU-R	International Telecommunication Union Radiocommunication
LSTM	Long Short-Term Memory
MCC	Matthews Correlation Coefficient
MSE	Mean Squared Error
NMS	Network Management Systems
POD	Probability of Detection
RAL	Rutherford Appleton Library
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RSL	Received Signal Level
TN	True Negative
TSL	Transmitted Signal Level
WAA	Wet Antenna Attenuation

Chapter 1

Introduction

This thesis demonstrates the potential to use commercial microwave links for rainfall retrieval in Australia. Moreover, it contributes to the global body of knowledge using commercial microwave links for rainfall retrieval and underpins the development of a more accurate operational rain measurement capability, including measurement of rainfall and drop size distribution. This study also develops a better understanding of the uncertainties in rainfall estimation associated with this measurement technique.

1.1 Importance of rainfall measurement

Rainfall is one of the dominant factors in terrestrial and atmospheric mass and energy balance over a wide range of spatiotemporal scales (Testik and Gebremichael, 2013). It is a primary source of freshwater, which plays a critical role in Earth's hydrological cycle. Thus, accurate and timely rainfall measurement is crucial for a wide range of applications in earth, environment, agricultural and other science and engineering disciplines. Following are some of the examples where quantitative rainfall measurement is considered important:

- Rainfall is arguable the most critical input in any hydrological model (Alemayehu et al., 2017; Skoulikaris et al., 2020). This is the primary input forcing to drive the simulation of the hydrological cycle in the catchment.
- Accurate and timely rainfall information is critical for providing real-time flood forecasting and issues early warning to reduce life and property loss (Michaelides, 2008).
- Rainfall is one of the most important climatic factors influencing the growth characteristics of the crops (Watson and Challinor, 2013). It provides the water that serves as a medium through which nutrients are transported for crop development. Thus, having this information known would help farmers to increase productivity.
- Information of rainfall is critical for water resource management as reservoir operation, irrigation planning and management etc.
- Accurate quantitative rainfall information is crucial for both Short term weather prediction (often called as *nowcasting*) or long-term weather forecasting.

- Rainfall is one of the dominant factors that control the ecosystem dynamics like water and energy fluxes, landslide activity etc.

1.2 Statement of the problem

In spite of its importance, there are a number of challenges with providing accurate and routine spatial information on rainfall. Importantly, rain gauges are point measurements that always underestimate the true rainfall, and are difficult to site and maintain in urban and mountainous areas (Muller and Kidd, 2006). Moreover, weather radar has limited spatial coverage, is impacted by clutter in urban and mountainous areas, and suffer numerous errors due to beam spreading and angle (Joss et al., 1990; Germann et al., 2006; Berne and Krajewski, 2013). Furthermore, satellites are challenged by spatial resolution, temporal repeat, and the ability to accurately relate the observations to rainfall amounts, especially over land (Kucera et al., 2013). Finally, while weather prediction models can give high spatial and temporal resolution estimates of rainfall, the accuracy of these estimates remain inadequate for most applications (Sene, 2010). Therefore, accurate and timely high spatial resolution information on rainfall is rarely available in the areas where it is needed most. However, there is a potential solution. Continental areas are becoming increasingly crisscrossed with commercial microwave links for pushing data around, particularly in our urban areas, and these have been shown to provide great potential for supplementing the above approaches with rich information on rainfall estimates (Messer, 2018; Uijlenhoet et al., 2018; Chwala and Kunstmann, 2019). Networks of commercial microwave links are installed and maintained for the purpose of telecommunication however rainfall can also be measured as a by-product. Rainfall can be estimated based on the loss of microwave signal over a microwave link path between a transmitting and receiving antenna (see Figure 1-1 for an example). The strength of the signal at the receiving and transmitting end of the link path are usually monitored by telecom operators for quality purposes. Upton et al. (2005) introduced the idea to use microwave link observations to determine path-averaged rainfall, which was later proven to be successful with actual commercial microwave link data.



Figure 1-1: Example of a commercial microwave link antenna tower, used in telecommunication networks.

Although the theoretical basis of rainfall estimation from microwave links has been well-studied and understood, real-world applications have been minimal (Overeem et al., 2016b). The parameters used in the rainfall-attenuation relationship provided by the International Telecommunication Union – Radiocommunication Sector (ITU-R) are based on global weather data, and so there is a need to derive these parameters for local weather condition. Theoretically, received power can be measured with a fine temporal resolution and precision, but in reality, these data are often made available at a relatively coarse resolution and precision (typically every 15-minute with 1 to 0.1 dB precision), leading to uncertainties in derived rainfall estimates. Thus, the effect of these sources of uncertainty on rainfall observation needs close investigation. Moreover, wet antenna attenuation remains one of the largest sources of uncertainty in the existing rainfall retrieval algorithm from commercial microwave links, and so this needs to be characterised and carefully understood.

Furthermore, rainfall measurements using CMLs largely depends on the sampling strategy and the frequency of the received signal level adopted by the mobile operators. Most of the operators only record the minimum and maximum received signal level (RSLs), but some others also record the average and instantaneous RSL data sampled over 15 minutes. Thus,

the difference in the performance for the rainfall retrieval using various sampling data need further investigation.

Also, at instances when only 15-minute minimum and maximum RSL data are available, a constant distribution between the minimum and maximum for the rainfall retrieval is usually assumed. However, the distribution of rainfall or attenuation is generally not consistent across any 15-minute period. A more robust and accurate methodology of estimating rainfall when using limited data such as minimum and maximum RSL data is therefore necessary.

1.3 Objectives and scope of research

The main objective of this research was to demonstrate the potential for an operational application of rainfall measurement from commercial microwave links. Consequently, the research questions and hypotheses for this thesis were as follows:

- 1. What are the errors associated with microwave link rainfall estimates, and how can these errors be minimised for more accurate rainfall estimates using microwave links?** The accuracy of the rainfall estimates depends on the type of rainfall event (eg. convective, stratiform). This also largely depends on the data sampling strategy of the attenuation data. The higher the sampling frequency of the data, the shorter could be the averaging interval and the better the temporal accuracy. Quantisation errors could also have a significant impact on the estimated rainfall amounts. Similarly, wet antenna attenuation could be characterised based on the rainfall intensity.
- 2. How is the performance of rainfall retrieval from commercial microwave links impacted by sampling strategy?** The most common sampling strategy, which telecommunication operators use to store the received signal level, is the 15-minute minimum and maximum. However, in some cases, average sampling data are also available. These additional average sampling data could provide extra information for the wet-dry classification and the rainfall retrieval, which could eventually be helpful for the overall performance of the rainfall retrieval using CML data.
- 3. Can a data-driven technique be used to improve rainfall estimates from CMLs?** Rainfall estimates based on a constant weighting using minimum and maximum attenuation induces uncertainties in the rainfall estimates as the distribution of the rainfall or attenuation is not consistent across any 15-minute period. However, data-driven

techniques such as deep learning models could capture such relationships from historical data. A deep learning model would make the most of the dynamics of the minimum/maximum time-series, while a standard approach would consider single timesteps of minimum/maximum separately to retrieve attenuation.

1.4 Outline of the approach

The research conducted for this thesis was comprised of the following tasks that can be broken down as follows:

- 1) Understanding uncertainties associated with microwave link rainfall estimation**
 - a. Set up a dedicated experimental microwave link including disdrometers, weather stations and rain gauges along the link path.
 - b. Analysis and estimation of rainfall of different types and report issues associated with each of these different rainfall types.
 - c. Compare the performance of three common sampling strategies, namely average, minimum/maximum and instantaneous, for various time intervals.
 - d. Compare existing wet antenna attenuation models and propose local parameters based on the experimental link data.
- 2) Evaluation of rainfall retrievals from commercial microwave link data for Melbourne**
 - a. Calibration of the rainfall retrieval model named RAINLINK for average and minimum/maximum RSL data obtained for the Greater Melbourne region.
 - b. Analyse and compare the performance of wet-dry classification and rainfall retrieval for average and minimum/maximum RSL data.
 - c. Present the analysis results and discuss the differences in outcomes.
- 3) Deep learning techniques for improving rainfall measurement from CML**
 - a. Develop various machine learning models and evaluate their performance using simulated specific attenuation based on disdrometer data.
 - b. Conduct an independent test of the best-identified model for its performance.
 - c. Validate the model using actual CML data for rainfall retrieval.
 - d. Present the analysis and discuss its implication for retrieval.

1.5 Structure of the thesis

This thesis is structured into seven chapters, including this Introduction chapter. A number of these chapters have also been the basis of journal publications, as indicated below. The remaining six chapters are:

- Chapter 2: Literature review

The literature review provides a broad overview of the theoretical background of using microwave link attenuation for rainfall measurement. This includes a review of and comparison with the existing conventional rainfall measurement techniques. Existing knowledge gaps are identified, and research questions to fill these gaps are formulated. Lastly, opportunities for utilisation of the research are presented.

- Chapter 3: Study area and data used

This chapter provides a description of the study area and data used for this study. This includes a detailed description of the setup of a dedicated experimental microwave link alongside a regular commercial microwave link for this study, including the addition of disdrometers, weather stations and rain gauges. This chapter also provides detailed information on the collection of commercial microwave link data across Melbourne, including reference rainfall data from radar and rain gauges.

- Chapter 4: Uncertainties associated with microwave link rainfall estimates

This chapter uses the datasets from the experimental link to investigate the uncertainties associated with microwave link rainfall estimates. This work has formed the basis of an article to be submitted to *Advances in Water Resources*:

- **Pudashine, J.**, Guyot, A., Pauwels, V.R., Overeem, A., Seed, A., Uijlenhoet, R., Prakash, M., and Walker, J.P. (2020). Uncertainties associated with microwave link rainfall estimates: Insights from an experimental setup in Melbourne, Australia. *Advances in Water Resources*. Manuscript under internal review.

- Chapter 5: Validation of commercial microwave links

This chapter provides a comparison of the two-sampling data commonly used by the telecommunication operators for the Greater Melbourne region as a case study. This work has formed the basis of an article submitted to the *Journal of Hydrology*:

- **Pudashine, J.**, Guyot, A., Pauwels, V.R., Overeem, A., Seed, A., Uijlenhoet, R., Prakash, M., and Walker, J.P. (2020). Urban rainfall retrieval using commercial microwave links: Effect of sampling strategy on retrieval accuracy, *Journal of Hydrology*. Manuscript under the second review.

- Chapter 6: Deep learning for rainfall improvement

This chapter provides a formulation of an innovative deep learning technique for improving the rainfall estimation if only minimum and maximum received signal level data are obtained. The results from this study have formed the basis of an article published in *Water Resources Research*:

- **Pudashine, J.**, Guyot, A., Petitjean, F., Pauwels, V.R., Uijlenhoet, R., Seed, A., Prakash, M., and Walker, J.P. (2020). Deep Learning for an Improved Prediction of Rainfall Retrievals From Commercial Microwave Links. *Water Resources Research*, 55, e2019WR026255. <https://doi.org/10.1029/2019WR026255>

- Chapter 7: Conclusion and future recommendations

This chapter presents a summary of the discussion of the main conclusions that can be drawn from the research and provides some future direction of work.

The following co-authored paper also contributed to the work of this thesis. For this paper, my contribution was in data processing and analysing the disdrometers data.

- Guyot, A., **Pudashine, J.**, Protat, A., Uijlenhoet, R., Pauwels, V.R., Seed, A., and Walker, J.P. (2019). Effect of disdrometer type on rain drop size distribution characterisation: a new dataset for south-eastern Australia. *Hydrology and Earth System Sciences*, 23,4737-4761. <https://doi.org/10.5194/hess-23-4737-2019>

Chapter 2

Literature review

This chapter provides an overview on literature related to the various rainfall measurement techniques; their drawbacks, basic theoretical background on microwave attenuation and its application for the estimation of rainfall rate, previous experimental studies for rainfall estimation using microwave link and research gaps in the existing literature.

2.1 Conventional rainfall measurement techniques

2.1.1. Rain gauges

Rain gauges are one of the conventional techniques for measuring rainfall all over the world, even though it provides only a point measurement (Wood et al., 2000). The tipping buckets are the most commonly used worldwide but are costly due to regular maintenance and their method of operation (Sevruk, 2002). Moreover, gauges are difficult to site in dense urban areas and mountain valleys. They need careful positioning, and they are prone to vandalism in urban areas; and provide accurate rainfall measurement only during low to intermediate intensity rainfall (Muller and Kidd, 2006). They are prone to errors in recording very heavy intensity rainfall due to inherent characteristics of the counting device (La Barbera et al., 2002; Muller and Kidd, 2006). Rain gauges also suffer from biases related to both the intensity and the wind field around the site. They are also prone to errors subjected from losses due to wind effect, wetting, splashing and evaporation (Emad et al., 2008). Since the rainfall field has a strong spatial gradient, particularly in the case of convective storms, single-point measurements cannot reflect the areal rainfall distribution in the large catchment (Michaelides et al., 2009).

2.1.2 Weather radar

In contrast, operational weather radars can provide broad spatial coverage of up to 300 km, by emitting a beam through the atmosphere at some angle above the horizon and measuring the reflected energy. Thus, the presence of large objects such as buildings or mountains in the beam renders the use of radar impossible. This is called clutter and is a serious problem, especially in urban areas. Radar measurement is based on the empirical relationship between radar reflectivity and precipitation which can be written as $Z = aR^b$, with Z the radar

reflectivity ($\text{mm}^6 \text{m}^{-3}$), R the rainfall (mm h^{-1}) and a and b parameters depending on the type of precipitation (snow, hail, rain, convective or stratiform). In many cases, the type of rainfall and its drop size distribution are unknown, causing an error in the rainfall retrieval (Morin et al., 2003). Most commonly, operational weather radar uses the frequency in C- and S-band range with wavelengths of about 5 and 10 cm, respectively, which are proven to be less sensitive to lighter rainfall. Similarly, X-band radars (wavelength of about 3 cm) have been used in research for many years (Sene, 2013), but these provide a shorter range and suffer from a greater signal attenuation than other radars due to the shorter wavelength (Germann et al., 2006). One rainfall event behind another can be almost completely blocked out.

2.1.3 Satellite

Satellite rainfall has the advantage of providing global estimates of rainfall, which is especially useful for remote areas and areas in developing countries where there is a lack of ground observations (Kidd and Levizzani, 2011). Currently, a number of satellite products are available using a wide range of techniques and sources from real-time information. Most of these products combine visible and infrared observations from geostationary satellites, with microwave observations from polar-orbiting satellites as well as outputs from Numerical Weather Prediction models in some cases (Sene, 2013). Even though the precipitation products derived from the satellite observations have already reached a good level of maturity, there still exist issues of accuracy and resolution (both temporal and spatial), which limit the stand-alone application in an operational setting (Brocca et al., 2014).

2.2 Rainfall measurement using microwave link

The innovative technique of utilizing attenuation data from microwave links to measure rainfall rate has recently gained the interest of a number of researchers across the globe. It is now considered a reliable means for continuous rainfall monitoring, complementing conventional rain measurement techniques (Messer et al., 2006). The major advantage of microwave links is that they can provide path-integrated measurements of attenuation by rainfall close to the surface at a range of a couple of meters above the ground level (Fencia et al., 2012). Figure 2-1 shows how microwave links are used in a mobile network. Mobile network operators (MNO) uses backhaul towers (base stations) to relay the high volume of the data from one place to another. As the link between these back-haul towers required high bandwidth, most of these links are of higher frequency which is usually attenuated by rain.

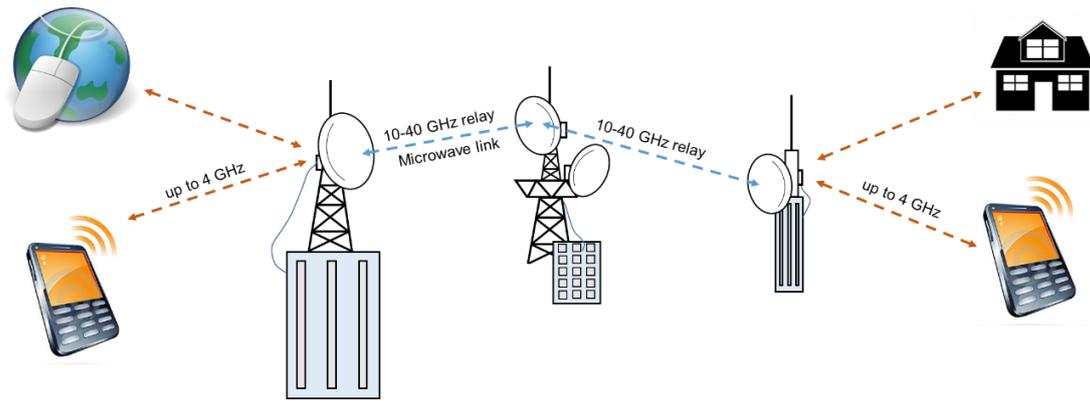


Figure 2-1: Use of microwave links for the mobile communication.

The rainfall measurement using a microwave link is based on the fact that microwave signals are absorbed and scattered when passing through the rain in the microwave region of wavelengths between 3 cm to 0.3 cm, corresponding to frequencies between 10 GHz to 100 GHz. The theoretical relationship of this attenuation phenomenon was studied a long time ago by Stratton (1930) and Mueller (1946). In these earlier days, the aims were to study the processes in order to establish a reliable microwave transmission along with the link, focusing on the prediction of unwanted interference. After the 1960s, weather radar technology used the attenuation and scattering that rainfall has on microwave signals to measure the rainfall intensity (Atlas and Ulbrich, 1977; Giuli et al., 1991; Jameson, 1991). Later, several experimental microwave links with dual and multi frequencies were studied to measure the path integrated rainfall rate (Rincon et al., 1996; Rincon, 2002; Holt, 2003; Rahimi et al., 2003).

The use of commercial microwave links (CML) for rainfall estimation was first recommended by Messer (2006). Received signal level (RSL) obtained from commercial telecommunication companies were processed to obtain rain-induced attenuations and then converted to average rainfall rates using an empirical power-law relationship. Leijnse et al. (2007) also used a commercial microwave link operating at 38 GHz data for a 2-month period to demonstrate its usefulness in the climatic setup of The Netherlands. They have considered eight rainfall events for obtaining the rainfall estimates from the link, which is well represented when compared with rain gauges and C-band radar data. Results showed that the link-derived rainfall compared better during the uniform rainfall event, whereas in variable rain, it showed a good comparison with radar rainfall. Since then, several studies have

demonstrated the feasibility of using microwave links from the commercial telecommunication network to measure rainfall rates (Messer, 2006; Leijnse et al., 2007; Blazquez et al., 2009; Goldshtein; et al., 2009; Chwala et al., 2012; Overeem, 2013; Doumounia, 2014; Overeem, 2016b).

Zinevich et al. (2008) demonstrated the robustness of microwave links to provide accurate rainfall estimates over an area of 3200 km² in Israel, providing more precise rainfall in the urban areas with an average correlation of 0.89. Their findings suggested that there is a significant impact on rainfall estimation due to quantization (rounding and truncation) errors, which need to be addressed. Overeem (2013) used CML data to produce country-wide rainfall maps for the Netherlands. Doumounia (2014) used a 29 km long link at 7 GHz to derive rainfall depths during the monsoon season in West Africa; 95% of rainfall was detected with a daily correlation coefficient of 0.8 and a bias of 6%.

The CML network is constantly expanding worldwide, and it exists wherever there is a reach of mobile phones. There is no additional investment in the infrastructure to make use of it to measure rainfall. These networks are denser in urban areas where there remains a challenge to improve the accuracy of observations of rainfall using weather radar and rain gauges. Mostly in developing countries where there is no radar measurement or rain gauges, CML could be the best alternative option for measuring rainfall. Besides this, CML has been used in various studies related to hydrological applications and flash flood early warning systems (Hoedjes et al., 2014). Microwave link attenuation data and rainfall derived from it are found useful for improving X-band radar estimates (Kramer, 2005) in Germany. They found that an error margin of 30 percent can be corrected using the microwave link data compared with gauge-based rainfall observation. Similarly, other studies also conducted in Germany found that microwave link attenuation data are helpful for optimizing the performance of weather radars by providing adequate estimates of the attenuation correction factor (Tromel et al., 2014). Further, Bianchi et al. (2013) recently used microwave link rainfall to spot fault rain gauges in Switzerland.

2.3 Theoretical background on microwave attenuation

2.3.1 Microwave attenuation in the near-surface atmosphere

When a radio wave propagates through the atmosphere due to the presence of mainly oxygen and water vapour, it suffers from refraction/reflection, scattering and absorption/attenuation (Seybold, 2005). Except refraction, all other effects are insignificant below 30 MHz but rain attenuation starts to be a significant factor above 10 GHz. Microwave attenuation happens mostly through two main mechanisms.

a) Attenuation due to gaseous molecules in the atmosphere

Atmospheric gases (mainly oxygen and water vapour) absorb electromagnetic energy at microwave frequencies, which causes loss in the received signal power. This loss is very minimal for short distances and low frequencies. For example, there is a specific attenuation of 0.05 dB km⁻¹ for 1 GHz. For the frequency range from 2 to 40 GHz, O₂ and H₂O are the predominant attenuation factors in the atmosphere. There has been well-established theory and models for the understanding of these processes as described by Liebe (1981).

The specific gaseous attenuation is given as:

$$k_{gas} = k_0 + k_w = 0.1820 f \left(N''_{Oxygen}(f) + N''_{WaterVapour}(f) \right), \quad (2.1)$$

where k_0 and k_w are the specific attenuation (dB km⁻¹) due to dry air and water vapour, respectively, f is the frequency (GHz) and $N''_{Oxygen}(f)$ and $N''_{WaterVapour}(f)$ are the imaginary parts of the frequency-dependent complex refractivity which is given by:

$$N''_{Oxygen}(f) = \sum_i(Oxygen) S_i F_i + N''(f), \quad (2.2)$$

$$N''_{WaterVapour}(f) = \sum_i(WaterVapour) S_i F_i. \quad (2.3)$$

Here, S_i is the strength of i -th oxygen or water vapour line, F_i is the oxygen or water vapour line shape factor. Figure 2-2 shows the specific attenuation for 15°C for a standard atmosphere containing water vapour density of 7.50 g/m³ and dry air.

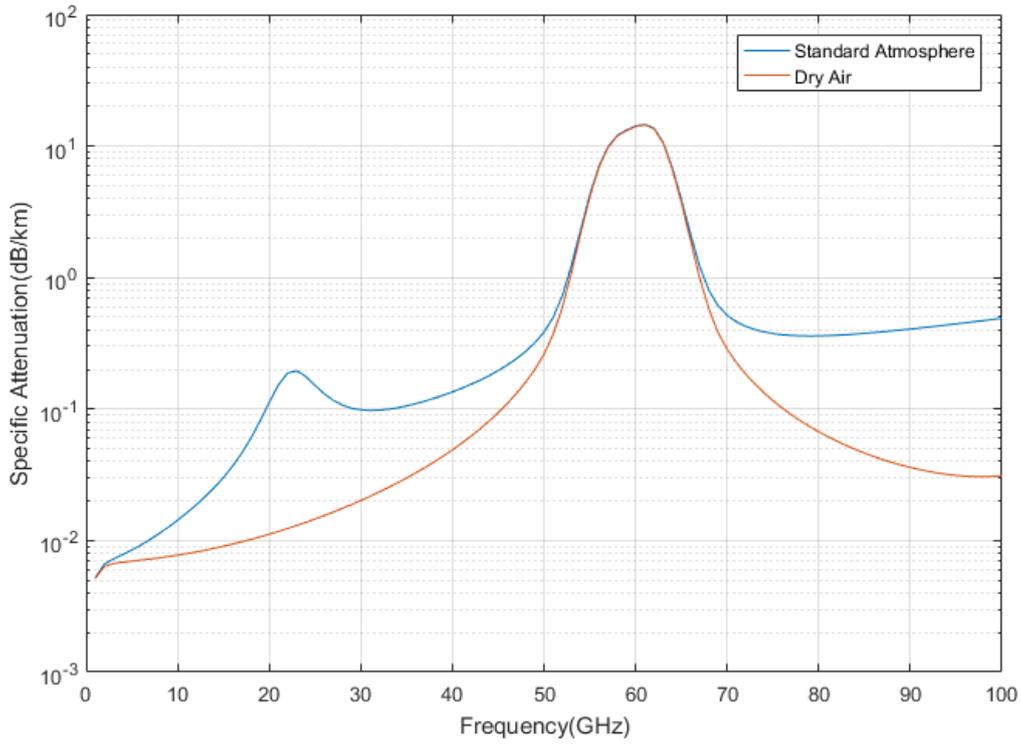


Figure 2-2: Specific attenuation of a standard atmosphere and dry air (ITU-R, 2016).

b) Attenuation due to rain

According to Vande (1981), if an electromagnetic wave with a wavelength λ and a wave number $k_0 = \frac{2\pi}{\lambda}$ passes through a region containing hydrometeors (with $N(D)dD$ the number of drops of equi-volume diameter $D, D + dD$ and forward scattering amplitude of $f(D)$) then:

$$K = k_0 + \frac{2\pi}{k_0} \int_0^{D_{max}} f(D)N(D)dD. \quad (2.4)$$

Then attenuation due to rain A (dB km^{-1}) is given as:

$$k = 8.686 \times 10^5 \text{Im}(K), \quad (2.5)$$

where D and f are in cm, K and k_0 are in cm^{-1} and $N(D)$ in cm^{-4} .

The relationship between attenuation and rainfall has been investigated by several authors. Using the Mie scattering solution, Olsen et al. (1978) proposed an empirical power relationship between rainfall rate R (mm h⁻¹) and specific attenuation k (dB km⁻¹) given as:

$$k = cR^d, \quad (2.6)$$

where c and d are the power-law parameters depends on the frequency f , polarization, raindrop size distribution and temperature. The total attenuation A (dB) can be obtained by integrating the specific attenuation over the path length (L) as:

$$A = \int_0^L k(x)dx, \quad (2.7)$$

Substituting eq (2.7) and (2.6) can be written as:

$$A = c \int_0^L R(x)^d dx, \quad (2.8)$$

For the sake of simplicity, the dependence of time (t) and frequency (f) is not considered in the notation. Now, considering the constant transmit power, the total attenuation A due to rainfall on a microwave link of length L , which is related to the received power level during the rainfall event (expressed in watts) and the received power during the dry period P_0 (W) is given as :

$$A = 10 \log \left[\frac{P_0}{P} \right] = 10 \log(P_0) - 10 \log(P). \quad (2.9)$$

This can be written using the natural log as:

$$A = \frac{10}{\ln(10)} \ln \left[\frac{P_0}{P} \right]. \quad (2.10)$$

Taking exponential on both sides of the equation:

$$\frac{P}{P_0} = \exp \left[-\frac{\ln(10)}{10} A \right]. \quad (2.11)$$

Again, substituting the value of equation 2.8 in equation 2.11:

$$\frac{P}{P_0} = \exp \left[-\frac{\ln(10)}{10} \int_0^L k(x)dx \right], \quad (2.12)$$

which is a form of Lambert-Beer's law of exponential extinction that is one of the fundamental equations used in physics. This equation is time and frequency-dependent which expresses the basic principle of the electromagnetic signal that it experiences attenuation as they pass through the rain medium. However, there is a need to invert the problem for the measurement of total rain-induced attenuation, A using equation (2.9) from the received power during the rainfall event and that just before the rainfall event. And once A is estimated, equation (2.6) can be used to invert to get the rainfall measurement. An important steps was the nearly simultaneous findings by several research groups about a decade ago that, for typical frequencies employed by microwave links in cellular communication networks, the exponent d of the power-law relation between specific attenuation k and rainfall rate R is very close to the unity. Thus, equation 2.6 can be approximated as:

$$A \approx c \left[\int_0^L R(x) dx \right]^d, \quad (2.13)$$

Dividing by the total path length L both sides,

$$\bar{k} \approx c \bar{R}^d, \quad (2.14)$$

Where $\bar{k} = A/L$ which is the path specific attenuation and $\bar{R} = \frac{1}{L} \int_0^L R(x) dx$

Now, inverting equation 2.14 yields the desired rain rate estimator given as:

$$\bar{R} \approx a \bar{k}^b, \quad (2.15)$$

where, $a = (1/c)^{1/d}$ and $b = 1/d$. This equation (2.15) forms the basis of any rainfall retrieval algorithm for microwave links. Equation 2.15 is exact if d is unity

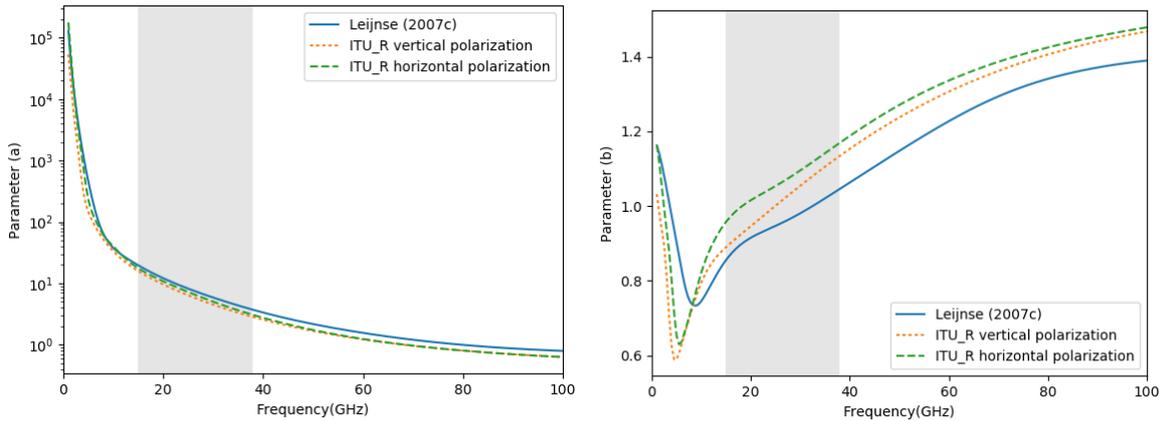


Figure 2-3: Frequency dependency of parameters a and b as proposed by ITU-R and Leijnse et al. (2007) in the power-law relationship.

2.3.2 Review of the rainfall retrieval algorithm

The empirical power-law proposed by Olsen et al. (1978) as presented in equation (2.15) is the basic relationship used for deriving rainfall from microwave link attenuation data, where R is the rainfall rate (mm h^{-1}), k is the specific attenuation (dB km^{-1}), and a ($\text{mm h}^{-1} \text{dB}^{-b} \text{km}^b$) and b are the parameters also related to frequency, polarization, temperature, drop size distribution, drop shape and canting angle. Figure 2-3 shows the parameters a and b for the frequency range 0 to 100 GHz which have been recommended for use worldwide by ITU-R as well as the parameters used by Leijnse et al. (2007) for The Netherlands.

To estimate rainfall depths from attenuation data, the following steps in the CML data processing have been proposed in the previous literature, as discussed in the following section.

a) Distinction between wet and dry periods

There is a decrease in received signal level during dry weather condition due to absorption and scattering by atmospheric gases, water films or dew formation on the antenna horn, and even due to the variation in temperature in the electronic components of the receiver itself (Liebe, 1981; Holt, 2003; Upton et al., 2005). This fluctuation in signal level could be misinterpreted as being caused by rainfall even in the absence of any form of precipitation. Thus, various methods have been proposed to classify wet from dry periods in order to quantify the correct baseline/reference signal level.

i) Nearby link approach: One of the approaches to identify whether the timestep is wet or dry is based on a temporal correlation with the time series from the microwave links in the proximity. This method relies on the fact that the decrease in received signal power level in the proximity of the links is due to alteration of weather conditions on those links. In this method, the interval of 15 min for the selected link is considered as wet if there is also a mutual decrease in the received signal power in the proximity links within a certain radius (Overeem et al., 2011; Overeem et al., 2016a). This approach was successfully used to process 2 years of microwave link data covering the Netherlands. The code is available in a package called RAINLINK (Overeem et al., 2016a).

ii) Radar approach: Overeem (2011) proposed this method which uses the unadjusted radar data to distinguish between the dry and wet conditions. For each link, overlapping radar pixels are checked for the rainfall intensity higher than 1 mm h^{-1} . If the intensity is greater for the selected link, the current and following time step is also categorized as wet. Since radar measures rainfall at greater heights, it will take additional time to reach the ground. It was found that in the case of the Rotterdam region, the fall time ranges from 5 to 12 min for a height of 1.5 km (Overeem, 2011, 2016b, a).

iii) Time series analysis: Besides these, there have been further approaches to analyse the individual time series of CMLs. (Chwala et al., 2012) have suggested using a spectral time series analysis. Similarly, Wang (2012) used Markov switching models to distinguish between dry and wet periods from the commercial telecommunication microwave links. As these methods do not require any empirical threshold value, they performed better compared with existing methods. Schleiss and Berne (2010) also used a time series analysis to classify between dry and wet periods. Their result performed well with the reliability of capturing 93% of the total rain amount on average. The only disadvantages of this method are that they require high-frequency data.

iv) Satellite-derived products to inform on possible rain occurrence: Recently, vanhet et al. (2017) proposed a new approach of using geostationary precipitation data from the Precipitation Clouds (PC) and Cloud Physical Properties (CPP) products derived from EUMETSAT. Each of the pixels is considered wet if one or both of these satellite product has rain with the probability of precipitation higher than 20 % for the PC product or a rainfall intensity greater than 0.1 mm/hr for the CPP product. Each of the links overlapping these

classified pixels are used to differentiate between dry and wet periods for each of the time steps of the links. This method was found to be prominent for the daily rainfall estimates.

v) Data-driven techniques: Various author proposed data-driven techniques like deep learning models to distinguish the time steps between the wet and dry period from the microwave link data. Habi and Messer (2018) developed a long short-term memory (LSTM) using a multi-variable time series to monitor the wet period in the time series. Recently, Polz et al. (2020) proposed a convolutional neural network (CNN) to recognize the attenuation pattern and then distinguish between the dry and wet periods. The model was trained using 4 months of data from randomly chosen CMLs which was later verified on 2 different months of data for all 3904 CMLs. This proposed method was outperforming the reference method; however, the data driven techniques need a large CML data set for the training the model.

Besides these above approaches, there are also other methods, such as using a decision tree method as demonstrated by Cherkassky et al. (2014) to distinguish between rain and sleet events with several classification features.

b) Reference signal level determination

Rainfall is the principal cause of attenuation of microwave links but is not the only factor causing signal fluctuations. Received signal power level will vary due to changes in atmospheric gaseous concentrations, translated as fluctuations of variables such as air relative humidity, atmospheric pressure, air temperature and wind velocity. Even though such fluctuations are not as large as those caused by rainfall, these need to be accounted for in the data processing to obtain a specific attenuation in order to derive a rainfall depth (Upton et al., 2005). Various authors suggested different methodologies to remove the baseline signal from the observed attenuation data. These include simple methods such as using a constant threshold to more complex methods, using advanced signal processing techniques (such as Fast Fourier Transforms and wavelet analysis (Minda and Nakamura, 2005; Upton et al., 2005; Overeem, 2011; Chwala et al., 2012). Fenicia et al. (2012) tested a linear low-pass filter to estimate the baseline signal with the assumption that the reference signal do not remain constant for any rainfall even. This linear loss pass filter baseline method provides a more realistic reference level than the constant baseline method.

c) Identifying erratic signal behaviour

Rainfall is not only the source of attenuation in the microwave link. Additional attenuation can be caused due to various atmospheric constituents like water vapour and oxygen. It can also be caused due to refraction, obstruction, multipath propagation. Thus, these additional attenuations should be identified and removed before converting the signal to the rainfall intensity. Overeem (2016a) introduced a pre-processing step for the minimum-maximum power level data to remove the malfunctioning and noisy links. This is based on a filter that relies on the principle of rainfall distribution over space. The filter discards a time interval of a link for which the cumulative difference between its specific attenuation and the surrounding links becomes lower than the outlier filter threshold. The implementation of this algorithm is included in a RAINLINK package. Recently, Graf et al. (2020) applied several quality checks on CML data using the filter threshold based on two criteria. The first is based on a 5 h moving window standard deviation exceeding the threshold of 2.0 for more than 10 % of a month, and the second is a 1 h moving window standard deviation exceeding the threshold of 0.8 more than 33% of the time in a month. These filters remove all the links with missing, noise and unrealistic data.

d) Estimating wet antenna attenuation

Wet antenna attenuation (WAA) is an essential factor that needs to be estimated correctly to obtain accurate rainfall measurement using microwave links (Fencl et al., 2018; Chwala and Kunstmann, 2019; Valtr et al., 2019) Various theoretical models and pragmatic approaches have been proposed to address this issue. WAA mainly depends on the type and material of the antenna cover. Blevis (1965) first modelled WAA as a uniform water thickness water film depending on the rain rate following a power-law relationship. A rain rate dependence with WAA was also obtained from the experimental work conducted by (Islam and Tharek, 2000; Kharadly and Ross, 2001; Minda and Nakamura, 2005). Leijnse et al. (2008) proposed a semiempirical model where WAA was considered a function of water film thickness on the antenna surface dependent on rainfall intensity. Furthermore, Fencl et al. (2018) used a 38 GHz microwave link to demonstrate a rain-rate dependence of WAA reaching the maximum value of 9 dB in extreme cases. Recently Valtr et al. (2019) proposed a model describing the wet antenna attenuation as a function of the rain rate, which also follows a power relationship between the WAA and rainfall intensity. On the other hand, Schleiss et al. (2013) proposed an exponential decay WAA model, which determines how fast the WAA reaches the

maximum value over the course of the rainfall event. The modelled WAA value increases sharply at the beginning of the rainfall event reaching the maximum value of 2.3 dB during the rainfall event. As a pragmatic approach, when only minimum and maximum RSL data is available, Overeem et al. (2011) suggested a constant wet antenna attenuation during wet periods, estimated for the transmitter and receiver antennas together.

2.4 CML studies around the world

Even though the near linearity on the power relationship was known for a long time to the meteorological community, this was limited to the dedicated experimental microwave link path for the rainfall estimation until the early 2000s, when Messer et al. (2006) and Leijnse et al. (2007) concomitantly demonstrated the use of CML signal attenuation for rainfall measurement. This was a major breakthrough toward demonstrating the potential to use the more than 4 million commercial microwave links in the world (Ericsson, 2017) for rainfall monitoring purposes. Since then, several studies have demonstrated the feasibility of using microwave links from the commercial telecommunication network to measure rainfall rates around the world, including Brazil (Rios Gaona et al., 2015), Burkina Faso (Doumounia et al., 2014), Czech Republic (Fencl et al., 2013; Fencl et al., 2017), Germany (Chwala et al., 2012; Chwala et al., 2016; Smiatek et al., 2017; Graf et al., 2020), Israel (Messer et al., 2006; Zinevich et al., 2008; Goldshtein et al., 2009), Italy (Roversi et al., 2020), The Netherlands (Leijnse et al., 2007; Overeem et al., 2011, 2013, 2016b; de Vos et al., 2019), Pakistan (Sohail Afzal et al., 2018) and Switzerland (Bianchi et al., 2013).

These validation studies have been conducted based on a few links to a couple of thousand links covering the entire country, such as The Netherlands (Overeem et al., 2013) and Germany (Graf et al., 2020). The temporal resolution of such CML rainfall estimates typically varies from a few seconds to 15 minutes, with most telecommunication operators sampling the received signal level (RSL) at 10 Hz but storing it at a much coarser temporal resolution. The summary of such studies using microwave links for the rainfall measurement conducted worldwide is summarized in Table 2-1.

Table 2-1: Published works on rainfall measurement using microwave links

Authors	Country	Dataset	Summary
Leijnse et al. (2007)	The Netherlands	2 CML 38GHz links 15 minutes instantaneous RSL data	Eight rainfall events were evaluated. Results were compared with rain gauges and C-band radar data.
Blazquez et al. (2009)	France	4 microwave links (26 GHz dual polarization links) 30-sec instantaneous data	Two rainfall retrieval models were compared. Results showed a correlation coefficient of about 0.8 when compared with the rain gauge close to the links.
Goldshstein et al. (2009)	Israel	22 microwave links Frequency range 17 to 24 GHz RSL recoded at a temporal resolution of 1 minute with 1 dB resolution.	An algorithm to construct a spatial distribution of rainfall map obtained from the CML was proposed.
Watson and Hodges (2009)	United Kingdom	38 GHz and two simulated links of 35 and 45 GHz	A reconstruction algorithm to estimate the rainfall field was proposed.
Schleiss and Berne (2010)	France	4 operational telecommunication links 19 and 26 GHz 10 milliseconds instantaneous transmitted and received powers	A method to distinguish between dry and wet periods was presented. Result detected about 92% of wet periods and about 93% of the total rainfall amounts.

Authors	Country	Dataset	Summary
Overeem et al. (2011)	The Netherlands	57 microwave links Minimum and maximum receiver power over 15-min interval with the resolution of 0.1dB	Rainfall derived from microwave links were compared with gauge-adjusted radar rainfall estimates over the path.
Rayitsfeld et al. (2012)	Israel	70 CMLs, 16-24 GHz with lengths varying from 0.8 to 18 km RSL logged every minute at a resolution of 1 dB	Results obtained from the CMLs were compared with the rain gauges in the vicinity of the links. 1-minute temporal resolution showed a correlation up to 0.75; however using a 10-minute resolution showed a correlation up to 0.85.
Chwala et al. (2012)	Germany	5 CMLs (15, 18.7 and 23 GHz), Data logger installed on the tower to record the RSL data. Resolution < 0.05 dB	A new wet-dry classification algorithm was proposed using short-time Fourier transform. Rainfall derived for 4 months period showed a good correlation up to 0.81 between link and rain gauge data.
Bianchi et al. (2013)	Switzerland	14 CMLs 23, 38, 58 GHz, 0.3 – 8.4 km links Temporal resolution from 2 to 5 minutes	CMLs data were used to identify the malfunctioning of rain gauges. An algorithm to detect error occurrence and quantitative errors in rain gauge measurement was proposed.
David et al. (2013)	Israel	36 CMLs (17-23 GHz), 15-min minimum and maximum RSL data with 0.1 dB quantization	Results from the CMLs suggested the probability of early detection of convective cells as compared with rain gauges.

Authors	Country	Dataset	Summary
Fencel et al. (2013)	Czech Republic	14 CMLs ~ 38 GHz	CMLs derived rainfall showed better information on spatio-temporal rainfall variability that has the potential to use for urban flood modelling.
Overeem et al. (2013)	The Netherlands	2400 microwave links Minimum and maximum received power with 1 dB resolution over 15-minutes	Countrywide validation of rainfall retrieval using CMLs providing the usefulness of such data for real-time rainfall monitoring over large areas.
Cherkassky et al. (2014)	Israel	3 CMLs 18.36 (11.9 km), 19.37 GHz (12.8 km)	The classification of wet period based on the kernel Fisher discriminant analysis. Results showed that the classification is in good agreement (~ 85%) with the data obtained from the disdrometer.
Doumounia et al. (2014)	Burkina Faso	7 GHz, 29 km microwave link Temporal resolution of 1 sec	Rainfall estimation obtained from the microwave link was compared with the rain gauge data, showing a correlation of 0.8. Accuracy of 95% in detection rainy days.
Liberman et al. (2014)	Israel	96 CMLs 18-23 GHz RSLs are stored at 15-minute with a resolution of 0.1 dB.	A novel approach of 2-D rainfall reconstruction was proposed using commercial microwave link rainfall estimation. These results outperformed the radar measurement both in intensity and the total amount of rainfall.

Authors	Country	Dataset	Summary
Fencl et al. (2015)	Czech Republic	14 microwave links 25-39 GHz 10-sec RSL data Quantization of transmitted and RSL are 1 and 1/3 dB	Microwave links derive rainfall estimates well-captured microscale spatio-temporal rainfall dynamics during all 9 rainfall events.
Ostrometzky et al. (2015)	Israel	4 CMLs 18.36 GHz and 19.37 GHz, 12.08 km 15-minute minimum and maximum RSL data with 1 dB quantization	A new approach for monitoring other than pure rain precipitation using microwave links was proposed taking into the advantage of having multiple links at the same location. The result suggested the outperformance during the mixed precipitation compared with standard power-law based estimation.
Ostrometzky et al. (2016)	Israel	18.6 GHz (16 km) microwave link 15-minute minimum and maximum RSL data with a resolution of 0.3 dB	This study proposed a methodology to calibrate the power-law parameters locally using the standard measurement from rain gauges and existing CML.
Overeem et al. (2016b)	The Netherlands	More than 3000 CMLs 15-minute minimum and maximum RSL data	Countrywide rainfall maps using two and half years of commercial microwave link was presented. The evaluation was done with the comparison of gauge-adjusted radar data.

Authors	Country	Dataset	Summary
Fencel et al. (2017)	Germany and Czech Republic	Dubendorf, Germany (38 GHz dual- polarized CML) Prague, Czech Republic (4 CMLs at frequencies 25,32 and 38 GHz)	This study proposed a method to reduce the bias on the rainfall estimation from CMLs by adjusting quantitative precipitation estimates to existing rain gauges.
Schip et al. (2017)	The Netherlands	~ 2400 CMLs 13 to 40 GHz 15-minute minimum and maximum RSLs were collected with 1 dB resolution	Meteosat Second Generation satellite precipitation product was used for the wet/dry classification.
Fencel et al. (2018)	Czech Republic	8 CMLs 37.3 to 39.2 GHz (48 -497 m) Data are recorded using a specially designed server-side software that pools data once per 10 seconds.	Wet Antenna Attenuation distribution and the upper limits based on the rainfall climatology was identified.
Rios Gaona et al. (2018)	Brazil	145 CMLs (8 to 23 GHz) Quantization level of 0.1 dB	The open-source algorithm RAINLINK was applied for rainfall retrieval. Results were evaluated against the automatic gauge network.

Authors	Country	Dataset	Summary
de Vos et al. (2019)	The Netherlands	1936 CMLs 12.8 to 39.3 GHz Instantaneous 15-minute data with quantization of 0.1 dB for RSL and 1 dB for TSL.	Rainfall intensities were validated with data obtained from gauge-adjusted radar product. A good correlation up to 0.63 was obtained for daily estimates of rainfall accumulations.
Graf et al. (2020)	Germany	3904 CMLs 10 to 40 GHz, 0.1 to 30 km Both RSL and TSL at a temporal resolution of 1 minute with a quantization of 0.3 dB for RSL and 1 dB for TSL.	This study evaluated the rainfall retrieval of CML data compared with a gridded gauge-adjusted hourly radar product with a good agreement for hourly, monthly and season rainfall sums expect for the winter season.
Polz et al. (2020)	Germany	Same data set as Graf et al. (2020)	Convolution Neural Networks was developed for the wet-dry classification of the CML data, which was later compared with the rolling window standard deviation method.
Roversi et al. (2020)	Italy	357 CMLs 6 to 42.6 GHz, 162 m to 30 km 15-min minimum and maximum with 1 dB resolution	The accuracy of the rainfall retrieval from CML was evaluated on different scales, including a single link, 5 km grid and catchment scale. An open-source algorithm RAINLINK was used for the analysis.

2.5 Knowledge gap and the proposed approach

After reviewing the literature, it is clear that the empirical power law between the signal attenuation and rainfall rate is the basis for rainfall retrieval from microwave link. The parameters of this power law depend mainly on the drop size distribution, which will vary significantly from one place to another. The International Telecommunication Union – Radiocommunication Sector (ITU-R) provides standard parameter value for worldwide use, which do not consider the local weather condition. Thus, there is a need to derive these parameters for the local weather condition for improved rainfall estimation. Furthermore, there have been not any validation studies using the microwave links in Australia. Theoretically, received power can be measured with a fine temporal resolution and precision, but in reality, these data are often made available at a relatively coarse resolution and precision (typically every 15-minute with 1 to 0.1 dB precision), leading to uncertainties in derived rainfall estimates. Thus, the effect of these sources of uncertainty on rainfall observation needs close investigation. Furthermore, rainfall measurements using CMLs largely depends on the sampling strategy and the frequency of the received signal level adopted by the mobile operators. Most of the operators only record the minimum and maximum received signal level (RSLs), but some others also record the average and instantaneous RSL data sampled over 15 minutes. Thus, the difference in the performance for the rainfall retrieval using various sampling data need further investigation. Also, at instances when only 15-minute minimum and maximum RSL data are available, a constant distribution between the minimum and maximum for the rainfall retrieval is usually assumed. However, the distribution of rainfall or attenuation is usually not consistent across any 15-minute period. A more robust and accurate methodology of estimating rainfall when using limited data such as minimum and maximum RSL data is therefore necessary.

Therefore, this research aims to validate the existing rainfall retrieval algorithms for the Australian climate, to develop an understanding of uncertainties in rainfall retrieval using different sampling strategies, and developing a new data data-driven rainfall methodology. The research is divided into three steps: (1) establish and collect data from a dedicated experimental link and later use that data to validate the rainfall retrieval algorithm and further investigate the uncertainties on the retrieval; (2) evaluation of the rainfall retrieval from commercial microwave link data, and (3) develop a deep learning model to improve the

rainfall estimation when there is only limited information available. These three steps will be covered in Chapter 4, 5 and 6, respectively.

2.6 Chapter Summary

The innovative technique of rainfall measurement using commercial microwave links has been discussed. However, reliable estimation of rainfall from such an opportunistic source remains a challenging task. The power-law model is the basis of the rainfall retrieval algorithm, which requires parameters derived based on a local climatic condition. There are other uncertainties in the rainfall estimates due to data sampling strategies adopted by the mobile operators, which need to be investigated. One of the knowledge gaps in getting accurate rainfall measurement from a commercial microwave link is due to the availability of only limited information. Therefore, this research proposes a methodology of using deep learning models to improve the rainfall estimates if only minimum and maximum received signal level data are available.

Chapter 3

Experimental setup and dataset used in this study

This chapter provides an overview of the study area, together with the experimental setup established for this study and the data used for this research. The experimental setup includes an experimental microwave link (microwave scintillometer) along with two disdrometers, two weather stations and three rain gauges located along the path of the experimental link. In addition to the data from the experimental link, this study also includes microwave link attenuation data received from one of the telecommunication operators in metropolitan Melbourne and reference rainfall data obtained from the Bureau of Meteorology. These data were used to understand the uncertainties in rainfall retrieval from microwave link attenuation in Chapter 4, validate the microwave link attenuation rainfall retrieval algorithm in Chapter 5, and demonstrate the application of a machine learning model for improving the microwave link attenuation derived rainfall in Chapter 6.

3.1 Experimental setup

3.1.1 Research microwave link

The experimental setup was located in suburban Melbourne, Australia with the core of the experimental setup formed by a dual polarisation microwave research link operating at 24 GHz with a path length of 3.79 km. The transmitter was installed at the Mount Waverly Reservoir [37°53'26.17"S, 145°10'21.10"E] and the receiver was installed at a rooftop in Lakeside Drive [37°51'21.74"S, 145°10'5.85"E] as shown in Figure 3-1. The transmitter antenna was placed on a 25 m high water tower, and the receiving antenna was placed on a 2 m high mast above the rooftop of a 3-story building, as shown in Figure 3-2.

This 24 GHz research microwave link is a custom-built instrument by Rutherford Appleton Laboratories (RAL), UK. The receiver contains both linear and logarithm detector boards, for both horizontal and vertical polarisations. For this research work, only data from the logarithm detector were analysed, as these are similar to the data stored by the commercial telecom operators. Additional characteristics of the experimental link antenna are given in Table 3-1.

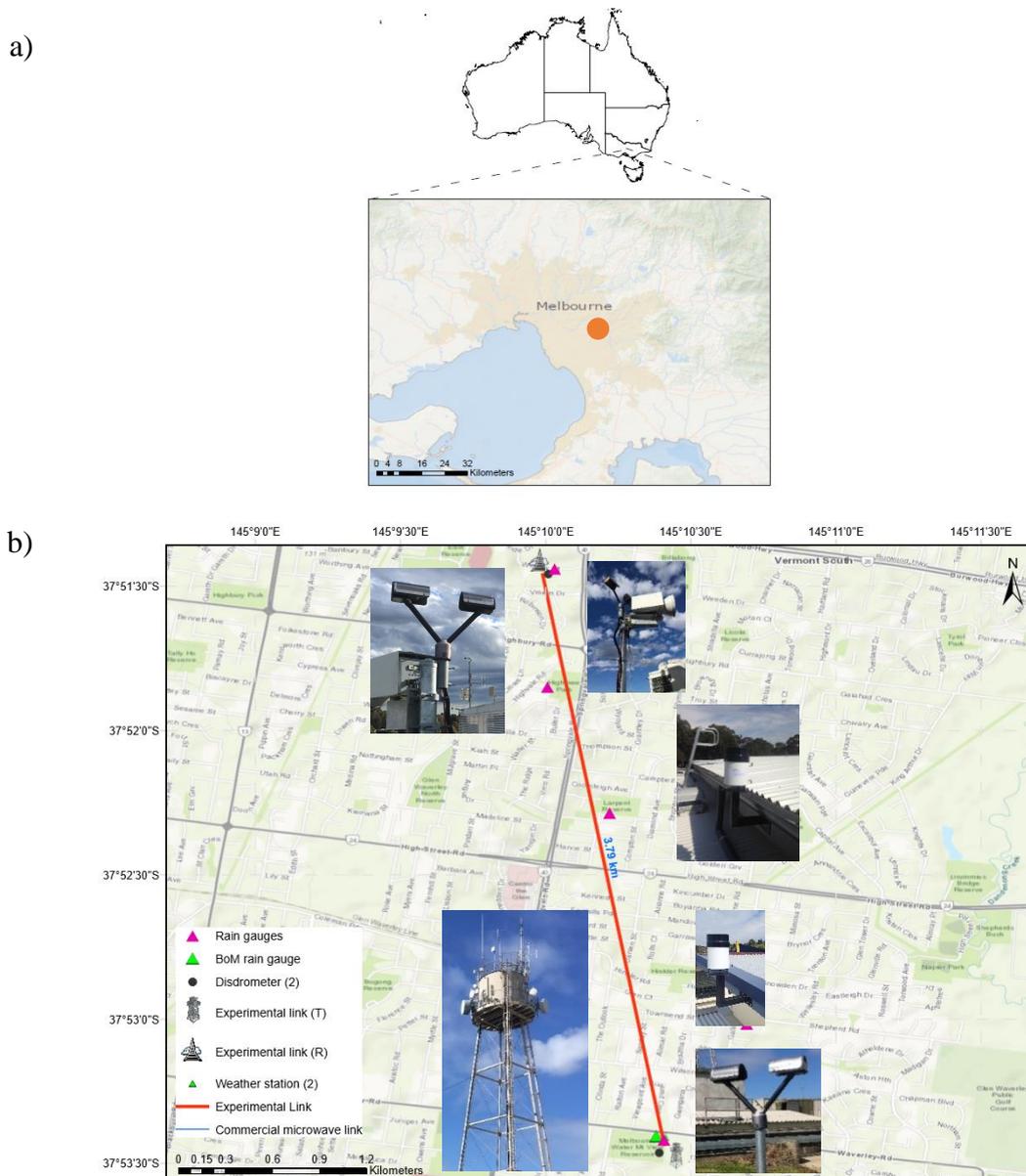


Figure 3-1: (a) Location of the experimental setup (the orange shaded circle provides the indicative location of the experimental site in the Greater Melbourne region). (b) Experimental setup including transmitter and receiver antennas. Disdrometers were installed at each end of the microwave link. Three tipping bucket rain gauges were installed along the length of the link.

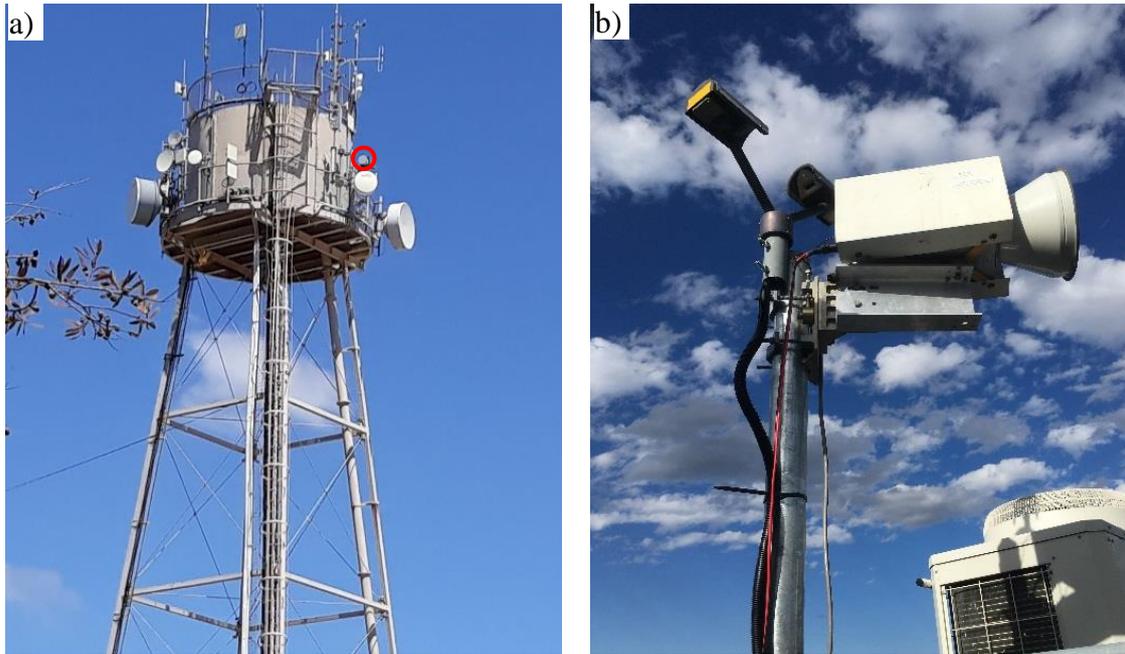


Figure 3-2: (a) Transmitter antenna (antenna within the red circle) installed at the Melbourne Water Reservoir placed on a 25 m high water tower, and (b) receiver antenna installed at Lakeside drive on a 2 m mast together with the disdrometer.

Table 3-1: Characteristics of experimental link antenna

Beamwidth	3.0°
Antenna diameter	250 mm
Antenna gain	34.5 dBi
Bandwidth	~ 1.2 KHz
Peak Power	0.2 W/m ²
Frequency	24.00 GHz

Before the instrument was set up at the site, Rutherford Appleton laboratories calibrated the instrument and provided the calibration curve shown in Figure 3-3. This enabled conversion of the received voltages (V) to received power in decibels (dB) for the logarithm detector board. The transmitted power was kept constant throughout the entire experiment.

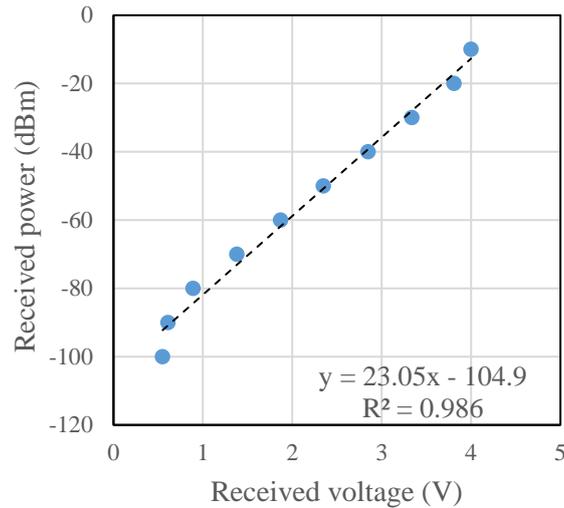


Figure 3-3: Received signal power vs received voltage data used for calibration of the detector board for RAL link (24 GHz)

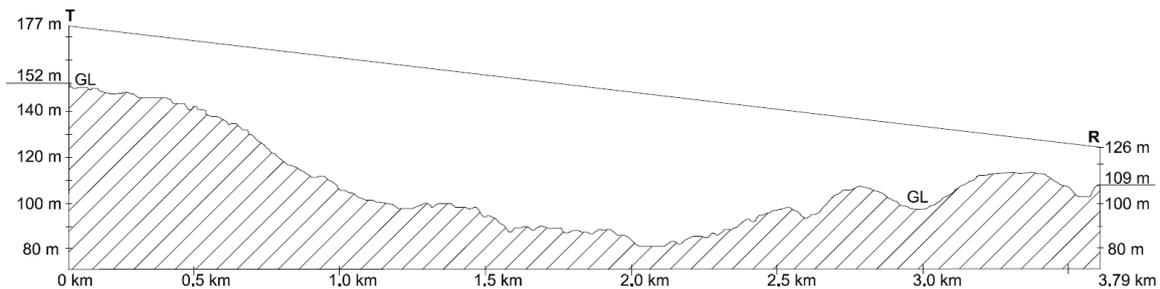


Figure 3-4: Profile of the experimental link. The hatch lines indicate the ground profile. The transmitter antenna was installed at the 0 km mark and the receiver antenna at the 3.79 km mark.

Most of the area along the experimental path is a residential urban area, with few small parks and urban reserves having a number of 20 to 30 m high trees, typical of the suburban Australian landscape. The terrain in between the transmitter and receiver consists mostly of one-story residential houses with a few buildings being of two or three stories construction (profile of the experimental link is shown in Figure 3-4). Similarly, there are no big trees or any other obstruction that exist along the path length. The maximum width of the Fresnel zone (halfway along the path) at the featured frequency is 3.43 m. Thus, considering the height of the antenna, locations compared to the intermediate terrain and height of the trees, there were no permanent obstructions affecting the beam, and no multipath effect was

Table 3-2: Data storing frequency for the different instruments used in the experimental setup.

Instrument	Sampling frequency	Storing frequency
Experimental link	100 Hz	1 minute
OTT PARSIVEL ¹ disdrometer	25 kHz	30 sec
Weather stations	30 sec	30 sec
Rain gauges	1 min	1 min
Commercial microwave link	10 Hz	15 minutes (minimum, maximum and average)

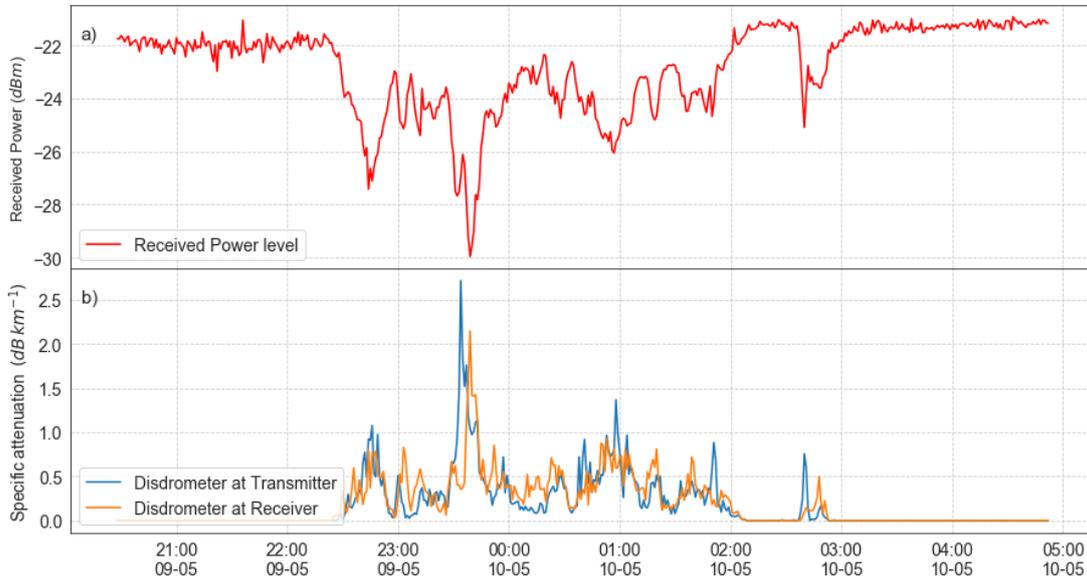


Figure 3-5: Time series of a rainfall event on the 9th to 10th May 2018 showing: (a) received signal level; (b) Specific attenuation measured by two disdrometers.

observed. Sample of the data collected from the RAL for one of the rain event is shown in Figure 3-5.

A duplex link from a commercial telecommunication operator (frequency: 23.33 GHz/22.71 GHz) also operated on the same path as our RAL. Received signal level (RSL) data were stored by one mobile operator (name not disclosed due to a Non-Disclosure Agreement). This provided us with actual data to compare against the experimental high-resolution data. The commercial RSL data were sampled at 10 Hz but stored every 15-minutes, providing average, minimum and maximum RSL during each 15-minute interval over the study period.



Figure 3-6: The OTT PARSIVEL1 disdrometers installed at each end of the experimental link path. (a) Disdrometer installed at the Melbourne Water Mount View reservoir and (b) Disdrometer installed at Lakeside Driver, Burwood.

This experimental setup was complemented by two laser disdrometers (OTT¹ PARSIVEL), three tipping bucket rain gauges, and two weather stations positioned at several locations along the experimental link path, as shown in Figure 3-1. This experimental setup was operational from 1 March 2018 until 28 November 2019. All these data were stored on Campbell Scientific dataloggers (CR1000 and CR300) and remotely collected on a daily basis. The sampling and storing frequencies for the various instruments used in this experimental setup are shown in Table 3-2.

3.1.2 Auxiliary instruments

a) Disdrometers

An OTT PARSIVEL¹ laser disdrometer was installed at both ends of this experimental link (receiver and transmitter side of the RAL link) as shown in Figure 3-6. Both disdrometers were placed on flat surfaces 2 m above ground level. These disdrometers are intended for measurement of the hydrometeor size and fall-speed. They are capable of measuring raindrop sizes up to about 25 mm and use 32 size classes of different widths, spread over 0-26 mm. However, the lowest two size classes were not used due to their low signal to noise ratio. The velocity was also subdivided into 32 non-equidistant classes, meaning that these instruments

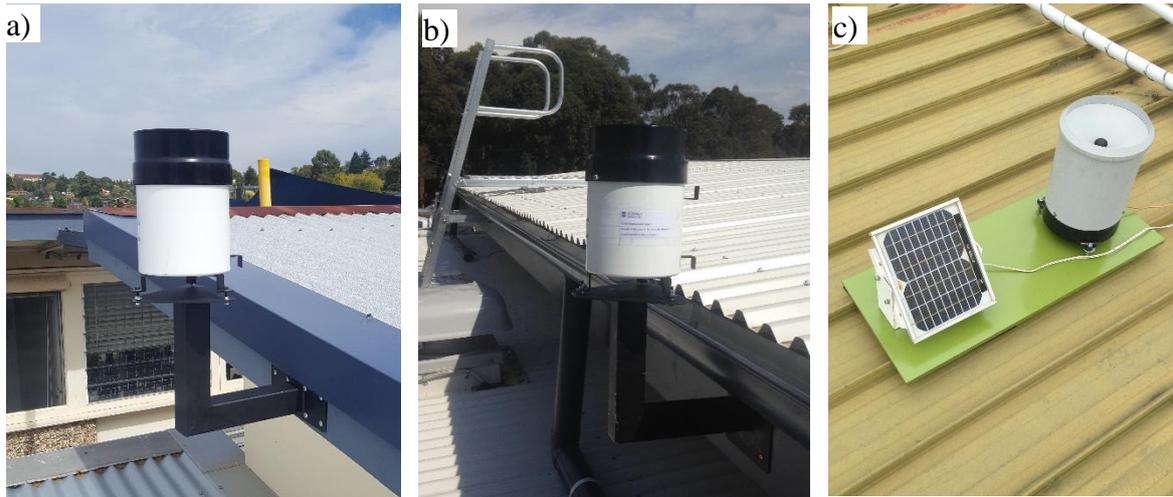


Figure 3-7: Tipping bucket rain gauge installed along the experimental path; (a) and (b) RG3-M rain gauges installed on the edge of one-story rooves, (c) TB3 rain gauge installed on the top of a one-story building. All three rain gauges were installed following the guidelines of the WMO.

stored particles in 32×32 matrices with a temporal resolution of 30 s. These disdrometers have built-in processing units which sample the raw laser signals and convert them into hydrometeor counts using an algorithm (which is not disclosed by OTT) and aggregates the samples to 30 s.

b) Weather stations

At both ends of the experimental link, weather variables such as wind speed and direction, air temperature and humidity, and barometric pressure were measured and recorded using a compact built-in weather sensor (WXT520, Vaisala). These variables were measured every 1 s and stored every 30 s. Data were collected remotely on a daily basis.

c) Rain gauges

Three tipping bucket rain gauges were also installed (see Figure 3-1) in order to provide a finer spatial resolution along the RAL pathlength. Among these, two tipping bucket rain gauges were manufactured by Hobo (RG3-M), while the remaining one was manufactured by HSA (TB4). These rain gauges were installed on the edge of one-story building rooves, as shown in Figure 3-7. All three rain gauges recorded rainfall amounts with a resolution of 0.2 mm and stored cumulative rain amounts at 1-minute time steps.

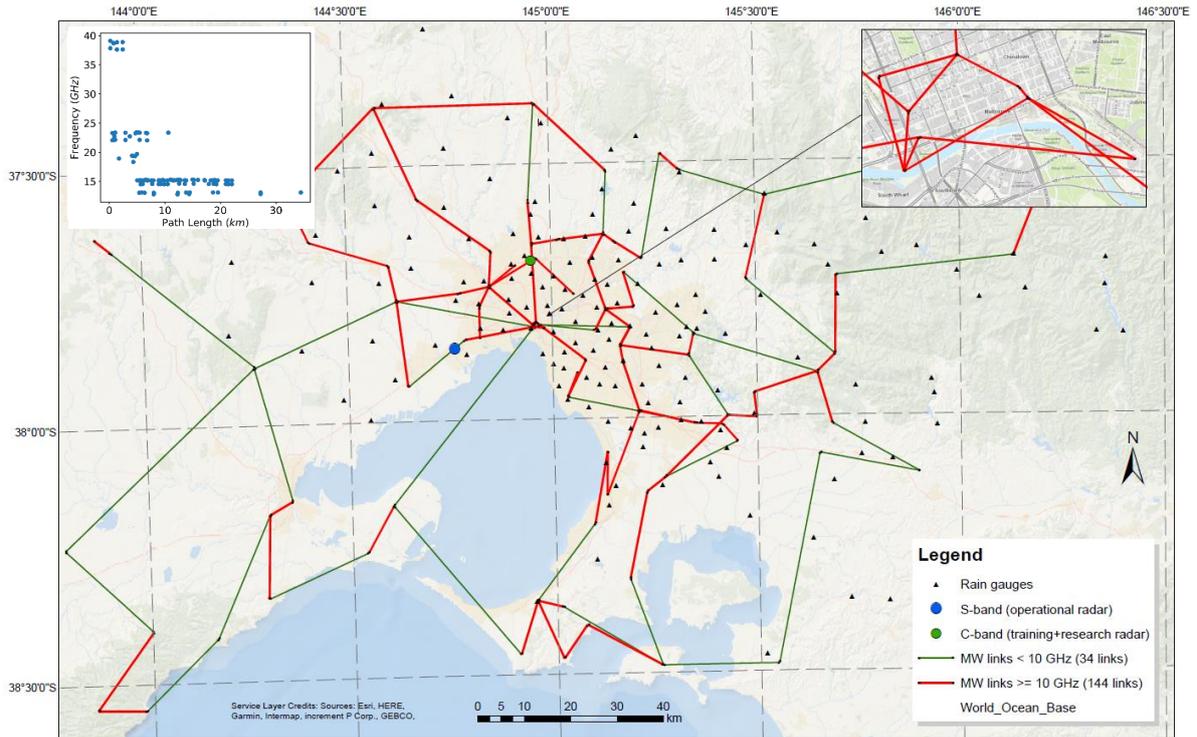


Figure 3-8: Commercial microwave links (CML) from one of the operators for Greater Melbourne. Red lines indicate microwave links having a microwave frequency greater 10 GHz (which were used for this study). Green lines indicate microwave links with frequencies less than 10 GHz (which were not used for this study). On the top left is a plot of the frequency of the microwave links f (GHz) against the path length L (km) for the 144 links used for this study.

3.2 Other data

3.2.1 Commercial microwave link data

Received signal level (RSL) data from one of the mobile operators in the area were collected for this study within a radius of approximately 200 km around the Melbourne Central Business District. Data from a total of 178 microwave links (64 duplex links and 50 single links) for the period ranging from 15 July 2017 to 31 July 2019 were collected as shown in Figure 3-8. This dataset contained minimum, maximum and average RSLs over 15-minute intervals with a resolution of 0.1 dB, based on a 10 Hz sampling rate. These 178 links had frequencies ranging from 6 to 39 GHz with corresponding path lengths ranging from 0.2 km to 57 km. Among the 178 microwave links, only links with frequencies above 10 GHz (144

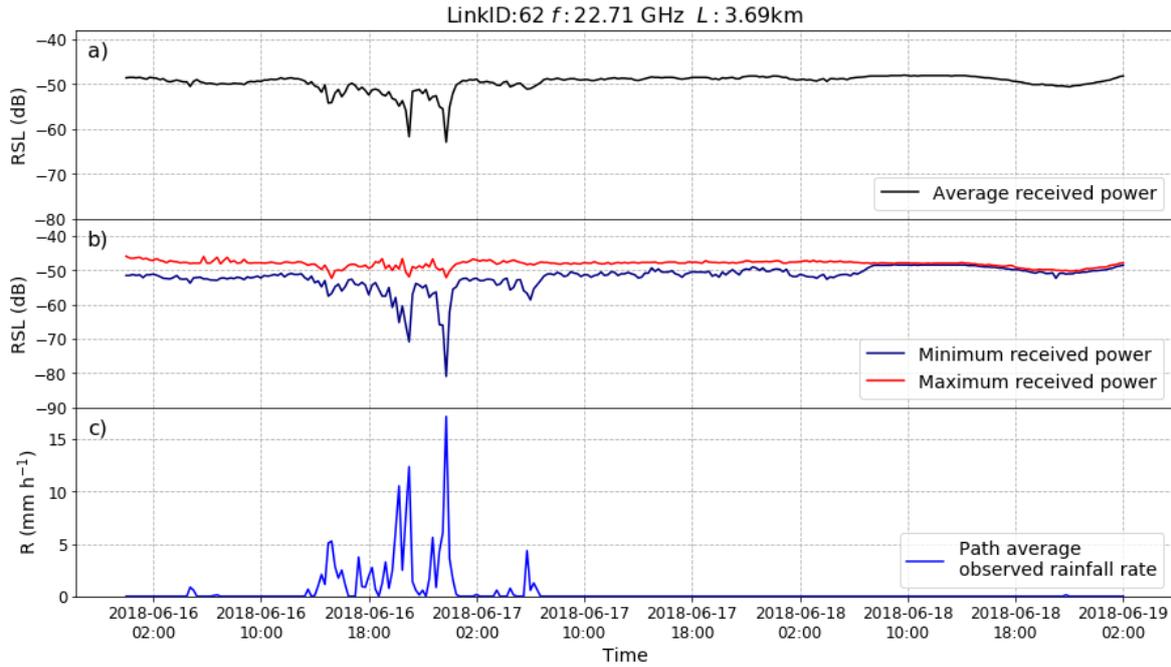


Figure 3-9: Time series of a rainfall event on the 16th to 19th June 2018 showing: (a) 15-min average received power level; (b) 15-min minimum and maximum received power levels; and, (c) Path-averaged observed rainfall rate along the microwave link.

links) were retained for this study. Among these selected 144 links, 9 links were horizontally polarised while all other links were vertically polarised. The majority of the links (128 links) were duplex (transmitter and receiver at both ends), with only 16 links being a single direction, thus forming 80 unique link paths. The transmitting powers of all retained links were constant through time. Figure 3-9 shows an example of a time series of 15-minute Average and MinMax RSL for a selected event. In both cases, the received power levels decreased proportionally with the observed rainfall rate; however, the power level still fluctuated during the dry periods. This fluctuation is seen more in the minimum and maximum compared with the average power level. The values of the minimum and maximum power levels over the 15-minute intervals obviously reached more extreme values (larger maximum and lower minimum). Since the RSL was sampled at 10 Hz, the minimum and maximum are the extremes of a distribution of values or the outlier of that distribution (Pudashine et al., 2020).

3.2.2 Weather radar data

An S-band weather radar operated by the Australian Bureau of Meteorology (BoM) produced data over the study area during the 2 years of the experimental link data collected in this study. This radar was located at Laverton (37°51'36" S, 144°45'36" E), 44 m above sea level. A gauge-adjusted radar data product named *Rainfields* version 2 was obtained from the BoM with a spatial resolution of 0.5 km × 0.5 km and a temporal resolution of 5 minutes (Seed et al., 2008). *Rainfields* is a comprehensive framework, which provides real-time quality-controlled quantitative precipitation for the operational Australian weather radar network. This framework follows a series of quality control measures, including removal of ground and sea clutter, interferences, bright band correction and partial beam blocking. This filtered observation was then converted to surface rainfall maps by combining existing rain gauges using kriging to interpolate three-dimensional observations onto a surface that is 1000 m above the radar.

3.3 Chapter Summary

This chapter has presented an overview of the data sets used in this study, including the setup of a dedicated experimental microwave link with auxiliary instruments, data collected by commercial microwave links, and ground reference rainfall data. Data collected from the experimental link, disdrometers and weather stations data are made publicly available on <https://doi.org/10.5281/zenodo.4442322>. However, due to the non-disclosure agreement with the mobile network operator, the CML data are not able to be made available.

Chapter 4

Uncertainties associated with microwave link rainfall estimation

This chapter provides an evaluation of the rainfall retrieval algorithm at the scale of a single microwave link, using ground-based observations, including rain-gauges and disdrometers as the ground-based rainfall benchmark. The primary goal of this chapter was to investigate the impact of data sampling strategy and data quantization scheme used for rainfall retrieval using a microwave link. A relatively simple retrieval algorithm, including a minimum number of corrections, was used. In parallel, the dedicated experimental link was also compared to a real commercial microwave link operating on the same path for rainfall retrieval.

4.1 Background

Rainfall retrieval from attenuation values collected from an individual link involves various steps, including (a) dry/wet classification of each time step; (b) baseline/reference signal determination; (c) identification of faulty/noisy received signal level; (d) correction for wet antenna attenuation, data quantization and other physical phenomena such as dew formation; and (e) conversion of specific attenuation to rainfall intensity. Each of these steps is subject to uncertainties arising from assumptions in the chosen model and its parameterisation, but also from the precision of the hardware and data generation (Fenicia et al., 2012).

All these uncertainties can be broadly classified into two categories (a) technological and (b) environmental (Zinevich et al., 2010). The technological uncertainties include errors and uncertainties associated mainly due to the hardware, which remains a major challenge, as the operational microwave links used by telecommunication operators are optimized for communication rather than for rainfall monitoring. Similarly, environmental uncertainties include errors associated to the various physical phenomena occurring along the link path, such as the variability of the drop-size distribution of rainfall and dew formation on the antenna.

Various experimental setups have been designed to understand underlying uncertainties at the different steps of the rainfall retrieval algorithm (Christopher et al., 1996; Rincon et al., 1996; Rincon and Lang, 2002; Rahimi et al., 2003; Fenicia et al., 2012; van Leth et al., 2018). Leijnse et al. (2008) reported wet antenna attenuation as the dominating source of error at

frequencies lower than 30 GHz and for links path lengths greater than 2 km. They also reported uncertainties due to variability in rainfall drop size distribution along the link path. However, links operating in the 10 GHz to 40 GHz frequency band are in the region of a near-linear relationship between rainfall and microwave attenuation, where the impact of the drop size distribution is minimal. Recently, van Leth et al. (2018) showed that the attenuation caused in a microwave link is not only due to rainfall itself, with fog and other weather phenomena also causing the same amount of attenuation. They also found that different hardware and materials used for the antenna pose some uncertainties; however, a consistent phenomenon was observed among dedicated experimental links and real commercial microwave links.

As the primary purpose of commercial microwave links is the provision of reliable communication rather than weather monitoring, the opportunistic nature of the data causes uncertainties in the rainfall retrieval process. Most of the telecommunication operators store the data at 15-minute time intervals, even if they are sampled at 10 Hz. But the type of statistics extracted from high-frequency data stored within this 15-minute interval varies across operators. Most of the operators provide the minimum and maximum of the received signal level (RSL) at 15-minute intervals, while others provide only the 15-minute average or in some cases, the 15-minute instantaneous RSL. Thus, the impact and uncertainties related to the data sampling techniques need to be assessed and incorporated in the rainfall retrieval model. Moreover, telecommunication operators employ different quantization (rounding off the raw data) on these RSL data to reduce the load of the storage on their operational system; also having a potential impact on the rainfall retrieval (Leijnse et al., 2008).

Similarly, excess attenuation is a critical issue with the microwave link rainfall estimation (Fencel et al., 2019; Valtr et al., 2019). One of the main causes of this excess attenuation is wet antenna attenuation (WAA). Various theoretical models and pragmatic approaches have been proposed to address this issue. Leijnse et al. (2008) proposed a semiempirical model where WAA was considered a function of water film thickness on the antenna surface dependent on rainfall intensity. Similarly, Schleiss et al. (2013) proposed an exponential WAA model which increased attenuation during the rainfall event, reaching a maximum value. However, this model was shown to poorly correlate with the rainfall intensity. As a pragmatic alternative, Overeem et al. (2011) suggested a constant wet antenna attenuation during wet periods, estimated for the transmitter and receiver antennas together. Recently

Valtr et al. (2019) proposed a model describing the wet antenna attenuation as a function of the rain rate, which also follows a power relationship between the WAA and rainfall intensity. Thus, there is a need to compare these different models for rainfall retrieval. Also needing testing is whether parameters derived based on a research experimental link can be directly applied on real commercial microwave links.

4.2 Study area and data used

This study was based on the experimental setup described in Chapter 3. Besides the experimental link and auxiliary data (disdrometers, weather stations and rain gauges), a co-located commercial microwave link along the same experimental path was also used for this analysis. Herein, data collected from the custom-built microwave link, designed and produced at Rutherford Appleton Laboratory, United Kingdom is referred to as RAL, while the commercial microwave link from a telecommunications operator (name not disclosed due to a non-disclosure agreement) along the same path is abbreviated as CML. In short, provides the locations of the various instruments used in this chapter. For more detailed information regarding the data sampling and storing frequency for each of the instrument, please refer to Chapter 3.

4.3. Data Processing

4.3.1 Disdrometer data

Although measuring the drop sizes and fall velocity using a disdrometer, there are various spurious and physically unrealistic particles that can be recorded in the measurement, which need to be removed before further analysis. Accordingly, a series of filtering procedures were applied, starting with a filter to remove the outliers as suggested by Jaffrain and Berne (2011), requiring particle size and velocity distribution to fall between $\pm 50\%$ of the empirical fall velocity diameter suggested by Atlas et al. (1973). Similarly, the filter by Tokay et al. (2013), using the number of particles for each time step in three different bins, was used to remove unrealistic bins with a significantly less number of particles. Hence, for the disdrometer data sampling interval of 30 seconds, a threshold of 10 particles per 30 s intervals was chosen. Very small drizzle, e.g. recorded rain rates per timestamp that were less than 0.1 mm h^{-1} were also excluded from the analysis (Jaffrain and Berne, 2011). All this post-processing was

Table 4-1: Location of the instruments used for this study.

Location	Coordinates	Instrument	Abbreviation used for this study
Mount view Reservoir	37°53'26.17"S, 145°10'21.10"E	Transmitter (RAL and CML), Weather station, OTT ¹ disdrometer	Site T
Burwood East	37°51'21.74"S, 145°10'5.85"E	Receiver (RAL and CML), Weather station, OTT ¹ disdrometer	Site R
Larpent reserve	37°52'16.92"S, 145°10'12.75"E	Tipping bucket rain gauge	TB1
Mount View Primary School	37°53'0.83"S, 145°10'42.84"E	Tipping bucket rain gauge	TB2
Private House property at Balfour Court	37°51'50.90"S, 145°10'0.41"E	Tipping bucket rain gauge	TB3

conducted using a script written in python language using some of the existing libraries as pyDSD and pyTmatrix (Leinonen, 2014). The drop size distribution was calculated using:

$$N(D) = \sum_{j=1}^{32} \frac{n_{ij}}{A_i \Delta t v_j \Delta D_i}, \quad (4.1)$$

where, n_{ij} is the number of droplets recorded for measured fall velocity v_j (m s^{-1}) for the velocity bin j , A_i (m^2) is the effective sampling area for the i_{th} size bin and Δt (s) is the sampling interval which was 30 s for this study. The effective sampling area A_i is calculated using

$$A_i = A \left(1 - \frac{D_i}{2\omega}\right), \quad (4.2)$$

where ω (m^2) is the width of the laser beam.

The rainfall intensity R (mm h^{-1}) and specific attenuation k (dB km^{-1}) are computed by integrating the drop size distribution $N(D)$, weighted by appropriate functions, such that

$$R = 6\pi \times 10^{-4} \int_0^{\infty} D^3 v(D) N(D) dD, \quad (4.3)$$

$$k = \frac{1}{\ln 10} \int_0^{\infty} Q_{ext} N(D) dD, \quad (4.4)$$

where D is the volume equivalent raindrop diameter (mm), Q_{ext} is the extinction cross-section (cm^2) of a rainfall drop with equi-volumetric diameter D (mm), N is the drop size distribution, $N(D)dD$ (m^{-3}) is the total number of drops in the diameter interval of $(D, D+dD)$ per unit volume, and R is the rainfall rate (mm h^{-1}). The extinction cross-section Q_{ext} was derived using a python implementation (pyTmatrix) of the T-matrix approach (Leinonen, 2014). Here, the shape of the rainfall was approximated by an oblate spheroid, with the axis ratio dependent on the volume equivalent-diameter. The default parameters, including a canting angle of 15° , and the drop shape model of Brandes et al. (2003) were used for the T-matrix calculations. Both horizontal and vertical attenuations were calculated for both the RAL (24 GHz) and CML (22.7 GHz).

4.3.2 Microwave links

A basic algorithm without applying any correction of the phenomenon was first used to calculate the rainfall intensity from the RSL of the microwave links. Following are the steps that were applied:

a) Dry-wet classification: This was the very first step in calculating rainfall intensity from microwave links. Overeem et al. (2016a) used a spatial correlation looking at nearby links to identify the given time steps as either dry or wet. As only a single path was used in this analysis, the nearby link approach was not applicable. Rather, the rainfall information collected from the two disdrometers and the three rain gauges were used to distinguish dry and wet periods. Time steps for which the rainfall rate was observed by either of the instruments as larger than or equal to 0.1 mm h^{-1} were classified as ‘wet’ with the remaining time steps classified as ‘dry’.

b) Estimating reference signal level: The reference signal level was calculated based on a moving median overall measurements of the signal level classified as dry in a centred window of 24 h. The specific attenuation was then calculated as

$$k = \max(R_{ref} - R_x, 0), \quad (4.5)$$

where R_{ref} is the baseline/reference signal level and R_x is the received signal level of the microwave link.

c) Applying attenuation correction: Additional attenuation caused by the wet antenna, fog and dew formation were removed in this step. For investigating the underlying uncertainties associated with rainfall retrieval, no correction factors were applied.

d) Computation of the rainfall rate: The rainfall rate was finally calculated using the power relationship between rainfall intensity (R) and specific attenuation (k), (Olsen et al., 1978) as

$$R = ak^b, \quad (4.6)$$

where, R is the rainfall intensity (mm hr^{-1}), k is the attenuation of the signal (dB km^{-1}), and a and b are parameters depending on the frequency, polarization, drop size distribution, drop shape and canting angle. These a and b used in this equation were derived based on the processed disdrometer data. The rainfall intensities and specific attenuation were derived for two frequencies (24 GHz: RAL link and 22.7 GHz: CML) and for both polarizations using equation 4.3 and 4.4, respectively. These data were then later used to fit the R - k power-law models using a non-linear square fit. As there were disdrometers at both ends of the link, the average of both of these disdrometers was utilised for the rainfall retrieval. Later, path-averaged ground observation (used as reference rainfall rate) was obtained by calculating the mean rainfall intensity measured by two disdrometers and three tipping bucket rain gauges.

4.3.3 Rainfall events

For this analysis, several rainfall events were identified based on the three tipping bucket rain gauges and two disdrometers installed along the path of the link. As the data frequencies of these instruments were different, all the data were resampled to a one-minute interval. Moreover, all the following conditions had to be met to conform as a rainfall event:

- A minimum rainfall intensity measured by any of these five instruments larger than 0.1 mm h^{-1} .
- A minimum rainfall duration of 1 hour.
- A maximum gap of 1 hour with no rain within the storm.

Table 4-2: Summary of the 72 rainfall events (rainfall rates were based on 1-minute stored frequency).

Longest duration (h)	10
Shortest duration (h)	1
Minimum rainfall depth (mm)	0.8
Maximum rainfall depth (mm)	38.09
Minimum rainfall intensity (mm h ⁻¹)	0.1
Maximum rainfall intensity (mm h ⁻¹)	150
Total cumulative minutes for all records (min)	46080

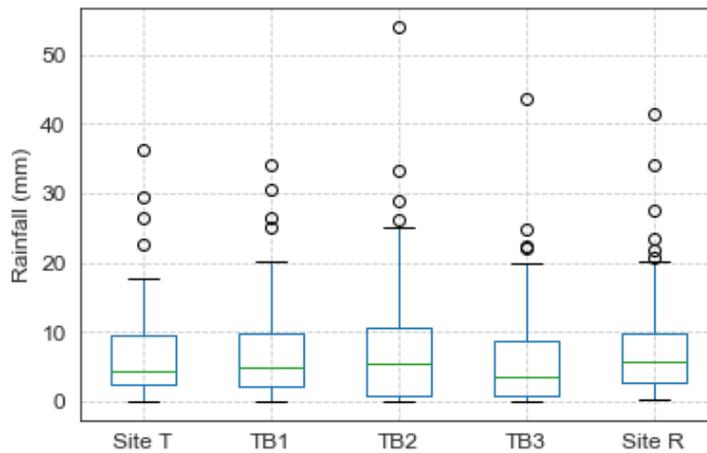


Figure 4-1: Distribution of rainfall depth along the RAL path for 72 rainfall events. (Distances from each of the rain gauges to the Site T (transmitter antenna) were: 0.88 km from TB1, 2.16 km from TB2, and 3.01 km from TB3).

Based on the above-mentioned conditions and availability of data from all the instruments, there were 72 rainfall events identified during the 2 years of data collection. Table 4-2 shows the summary of the rainfall events.

Figure 4-1 shows the distribution of the rainfall depth measured by the 5 different instruments across these 72 rainfall events. Site T and Site R refer to the rainfall measurement recorded by the OTT¹ disdrometer at the transmit and receive ends of the link, while TB1, TB2 and TB3 refer to the three tipping bucket rainfall measurements along the link path.

4.3.4 Data sampling and quantization

In order to evaluate the impact of various data sampling strategies on the overall rainfall retrieval, raw RAL data were resampled to 5 and 15 minutes. For each of the sampling periods, the three commonly used sampling strategies were used: (a) Average; (b) Minimum and maximum (hereafter called as MinMax); and (c) Instantaneous. For each of these strategies, RAL received power level data were sampled at the start of the period. For the instantaneous strategy, only one snapshot of data at the start of the period was taken. Once the RSL data were resampled to 5 and 15-minute for all three different strategies, a similar algorithm was applied for the rainfall retrieval. To have a fair comparison of the results from different strategies sampled at two different time periods, obtained rainfall results were finally all resampled to 15-minute interval. Similarly, to further understand the impact of data quantization on the overall rainfall retrieval process, raw RAL data were resampled with quantization levels ranging from 0.1 to 1dB. A similar rainfall retrieval algorithm was applied using each quantized received signal level data.

4.3.5. Wet antenna attenuation

Three different wet antenna models were compared for this study. A model proposed by Garcia-Rubia et al. (2011) with wet antenna attenuation A_a (dB) expressed as a function of the measured attenuation A_p (dB) was used, given as

$$A_a = C(1 - \exp(-dA_p)), \quad (4.7)$$

where C is the highest expected wet antenna attenuation (WAA), and d is a coefficient to be determined from fitting the model to the experimental data. This model is an improved version of the model proposed by Kharadly and Ross (2001) which is expressed as a function of measured attenuation A_m (dB) given as

$$A_a = C(1 - \exp(-dA_m)). \quad (4.8)$$

This model was used purely to separate the wet antenna effect from measured values, making it not applicable for estimating WAA before actual measurements are available.

Similarly, the model proposed by Valtr et al. (2019) was also used, given as

$$A_a = k'R^{\alpha'}, \quad (4.9)$$

where k' and α' are coefficients determined by fitting the left-hand side of equation (4.9) to the predicted wet antenna attenuation using the least square method. Similarly, a simple constant attenuation model was also used. For the constant attenuation model, a constant wet antenna attenuation (dB) was identified based on the optimization of RMSE and percent bias between the rainfall estimation from the experimental link and the path-averaged ground observation.

4.3.6 Rainfall retrieval using the analytical approach

Among various analytical approaches, two approaches proposed by Ostrometzky (2017) were used for this study. These are based on a statistical signal processing tool in combination with the Extreme Value Theory (EVT); thus, only data sampling based on the extreme values i.e. minimum and maximum RSL data, were feasible for this analysis. Following are the two approaches used for this study:

a) Generalized Extreme Value (GEV) approximation method: This analytical method was based on the assumption that for certain types of parent distribution, such as the exponential distribution or log-normal distribution, the parameter vector of the original parent distribution $\hat{\phi}$ can be estimated from the parameter derived from the GEV vector $\hat{\psi}$ (Ostrometzky and Messer, 2014). The following workflow suggested by Ostrometzky and Messer (2014) and Ostrometzky (2017) was used:

1. Pre-processing: This was done in two stages.
 - (a) Applying bias correction: This bias was mainly due to the data quantization of the collected raw data. This was assumed as a constant for a long period for each CML with the bias-corrected minimum (A_c^{min}) and maximum attenuation (A_c^{max}) calculated as:

$$A_c^{min} = A^{min} + \left(\frac{\delta_T + \delta_R}{2} \right), \quad (4.10)$$

$$A_c^{max} = A^{max} + \left(\frac{\delta_T + \delta_R}{2} \right), \quad (4.11)$$

where, A^{min} and A^{max} are the unbiased raw minimum and maximum attenuation and δ_T and δ_R are hardware quantizer values, respectively.

- (b) Filtering data other than rain: In this stage, corrected minimum and maximum attenuation were used to estimate the attenuation only due to rain (A_r^{max}) in dB. This was calculated as:

$$A_r^{max} = \{ A_c^{max}[i] - (A_c^{min}[i], \dots \dots A_c^{min}[i - T]) , 0\}, \quad (4.12)$$

where i indicates the i^{th} timestep and T is a constant parameter obtained after the optimization. For this study a value of $T=10$ was used.

2. The power-law based on eq (4.6) was used on A_r^{max} to estimate the maximum observed rainfall rate vector defined as r^{max} .
3. The maximum likelihood estimation was performed on the vector of r^{max} obtained in the previous step to estimate the GEV parameter vector $\hat{\psi}$.
4. The estimated GEV vector ψ was later used to estimate the $\hat{\mu}$ and $\hat{\sigma}$.
5. As the final step, using the $\hat{\mu}$ and $\hat{\sigma}$, the accumulated rainfall was calculated using:

$$\hat{R}(t) = t \cdot e^{\hat{\mu} + \frac{\hat{\sigma}^2}{2}}, \quad (4.13)$$

where, $\hat{R}(t) = \int_0^t r(t)dt \approx E[r|mRSL]$ and $E[r]$ is the expected value of the rain r which is given by the Log-Normal distribution properties.

b) Calibrated power law

After conducting the pre-processing according to the GEV approximation method, A_c^{max} was directly used to calculate R_{avg} using:

$$R_{avg} = \left(\frac{A_c^{max}}{a_{cal}^{max} \cdot L} \right)^{\frac{1}{b}}, \quad (4.14)$$

where, A_c^{max} is the maximum attenuation, L is the path length of a microwave link and a_{cal}^{max} is the calibrated/adjusted a parameter given as:

$$a_{cal}^{max} = a \cdot (\ln \ln (K) + \gamma)^b, \quad (4.15)$$

where, K is the number of samples from which the maximum attenuation sample is extracted, which is hardware dependent (for this study the value of $K = 90$), γ is Euler's constant, which equals $\gamma = 0.57722$ and a and b are the power-law parameters. The detailed description and the assumption made to derive equations (4.14) and (4.15) can be found in

Table 4-3: Coefficients and exponents (a and b parameters) of the R - k relationship derived for both the RAL link (f : 24 GHz) and the CML (f : 22.7 GHz) for both horizontally and vertically polarised radiations [Unit of a is $\text{mm h}^{-1}\text{dB}^{-b}\text{km}^b$, b is unitless].

	a_{24H}	b_{24H}	a_{24V}	b_{24V}	$a_{22.7V}$	$b_{22.7V}$
Site T	7.83	0.93	8.66	0.96	9.56	0.95
Site R	7.75	0.94	8.62	0.98	9.60	0.97
ITU-R (2005)	6.88	0.99	7.79	1.04	9.69	0.93

Ostrometzky and Messer (2020). For this method, both minimum and maximum attenuation values can be used; however, minimum attenuation contains a higher signal-to-noise ratio (SNR) and is likely to be affected by the other-than-rain phenomenon. Thus, this was not included in this study.

4.4 Results and discussion

4.4.1 Rainfall -specific attenuation relationship

Table 4-3 shows the resulting a and b parameters for the RAL (24 GHz) and the CML link (22.7 GHz). The power exponents for both the 24 GHz and 22.7 GHz links were close to unity, indicating that the specific attenuations and rainfall intensities were linearly proportional; however, when compared with the result obtained by International Telecommunication Union – Radiocommunication (ITU-R) there were some differences in both the a and b parameters. The parameter a based on ITU-R was underestimated by about 13% and 10% for horizontal and vertical polarizations respectively at 24 GHz; however, there was only about 1% overestimation at 22.7 GHz. Similarly, the parameter b based on ITU-R was overestimated by about 6% for both horizontal and vertical polarizations at 24 GHz and underestimated by about 3% at 22.7 GHz for vertical polarization.

4.4.2 Validation of the rainfall retrieval algorithm

Out of the 72 rainfall events identified for this study, three unique rainfall events were presented in detail, demonstrating the performance of the rainfall retrieval without applying

any correction to the other phenomenon. These three events were selected such that they represent different rainfall types demonstrating the various phenomena of the signal attenuation. The raw RAL data were resampled to 15-minutes when the results were compared with the commercial microwave link data.

a) Convective event

Figure 4-2 shows an example of one of the convective events that occurred during the 3rd and 4th May 2018. This event lasted for 237 minutes (equivalent to 3.95 hours). It was selected as it occurred after a prolonged dry period, such that there would be no other phenomenon contributing to the dynamics of the signal other than rain. The received power level started to drop once the rainfall began; however, there were some fluctuations observed (about 2 dB variation from -21.6 dB to -23.5 dB) in the observed signal during the dry period. Note that for all this analysis, the baseline or reference signal was determined based on a 24-hour moving window median of the received power level during the dry period. Also, to have an idea of the variation of the baseline signal, the 5th and 95th percentiles have been shown along with the median value. Initially, this event started with a low-intensity rainfall of about 0.3 mm h⁻¹ before reaching two peaks of about 7 mm h⁻¹ and 22 mm h⁻¹ each. This is also reflected in the specific attenuation plot, with approximately 1.8 dB km⁻¹ and 7.2 dB km⁻¹ corresponding to the two peaks. Specific attenuation based on the RAL showed a strong correlation (with site T = 0.91 and site R = 0.71) with measured attenuation based on the disdrometers; however, the magnitude was overestimated from the beginning of the event. The two parameters (*a* and *b*) derived for the RAL frequency using the data from the two disdrometers showed little variability in the drop size distribution at the two ends of the link path. The rainfall derived from the RAL was also higher than the observed from the beginning of the event, and for most of the timesteps, the differences increased with the magnitude of the rainfall intensity. However, there was a strong correlation of 0.95 between the RAL-derived rainfall and the path-averaged ground observations. As there was no other phenomenon involved, this overestimation was explained as being caused by the wet antenna attenuation. The magnitude of this overestimation can also be seen on the accumulation plot shown in Figure 4-3. The total accumulated rainfall estimated from the RAL was almost 100% higher than the path-averaged ground observations. For this event, the relative humidity started to increase once the rainfall started and stayed at the same level for the whole duration of the event; however, the temperature dropped from 20°C to 13°C during the event. Similarly, the average wind speed gradually decreased from 4 m s⁻¹ to 1 m s⁻¹.

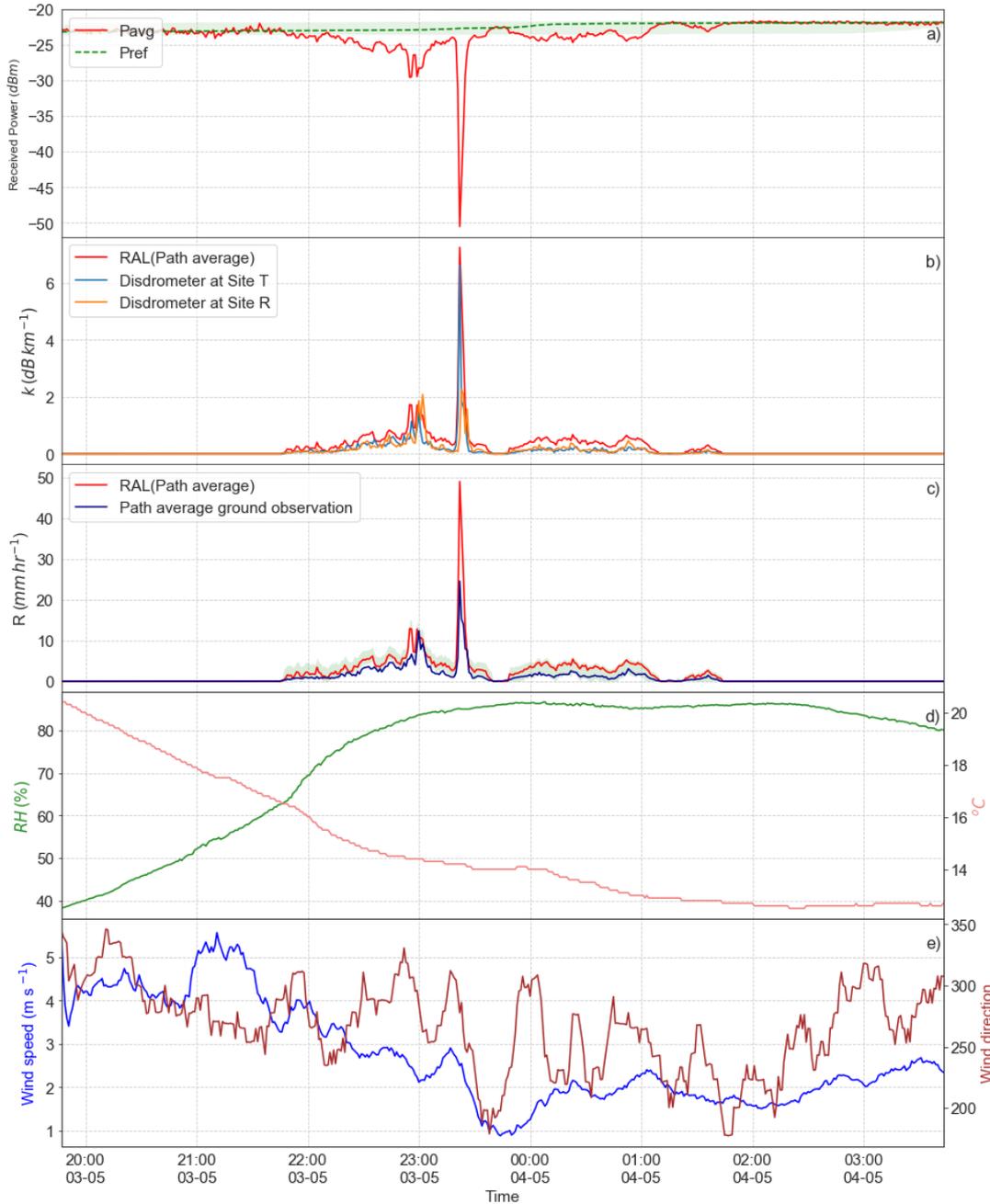


Figure 4-2: Time series of an event on the 3rd and 4th May 2018 of (a) received signal level and the reference signal (green dashed line); the shaded colour shows the 5th and 95th percentile of the reference signal during the dry period; (b) specific attenuation derived based on the RAL data and theoretical specific attenuation derived for the RAL frequency using the processed disdrometer data; (c) rainfall intensity derived from the RAL data using the R - k power law and path-averaged rainfall intensity using three tipping bucket rain gauges and two disdrometers; (d) variation of relative humidity (green line) and average temperature (red line) for the same event; (e) average wind speed (blue) and wind direction (brown) at the transmitter site.

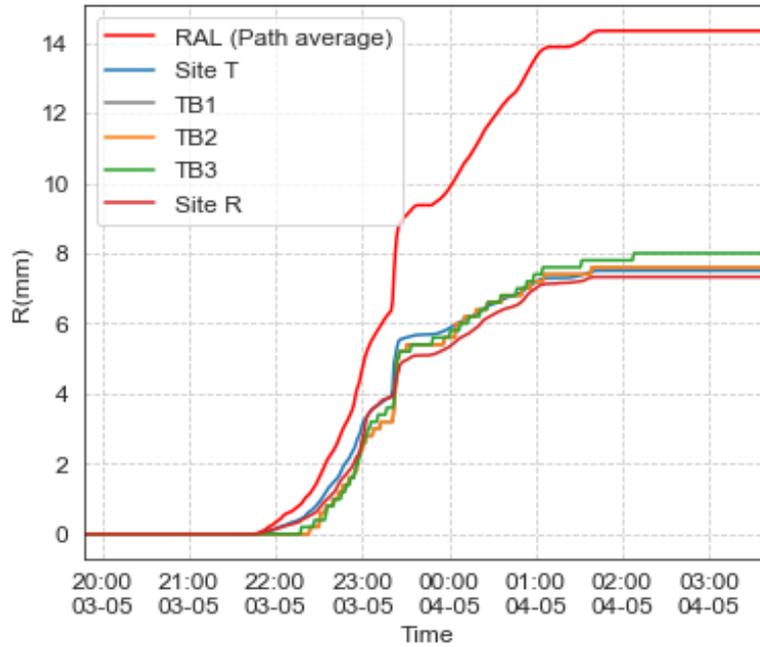


Figure 4-3: Accumulated rainfall plot for an event on 3rd and 4th May 2018.

To understand and compare the performance of the rainfall retrieval using the operational microwave link for the same event, the RAL data were resampled at the same interval of 15-minute similar to the received signal level data from the telecommunication operator.

Figure 4-4 shows the time series plot of the RAL and the CML data for the same event. There is an offset of about 26 dB between the RAL and CML received signal level data, which was mainly due to the differences in the transmitting power of these links. The rainfall intensity derived from both the RAL and the CML showed a strong correlation with the path averaged ground observations; however, both of these estimates were overestimated. The overestimation was observed to be of a larger magnitude on the CML when compared with the RAL. This was attributed to the larger decrease of the signal level on the CML at the start of the event, and after the peak when the path-averaged ground observations were almost close to zero. The signal level of the CML took a longer time to return to the level of the baseline signal, which could be due to different antenna covers used on the two links, becoming wet due to the rain and subsequently drying at variable paces after the event (Minda and Nakamura, 2005; Leijnse et al., 2008).

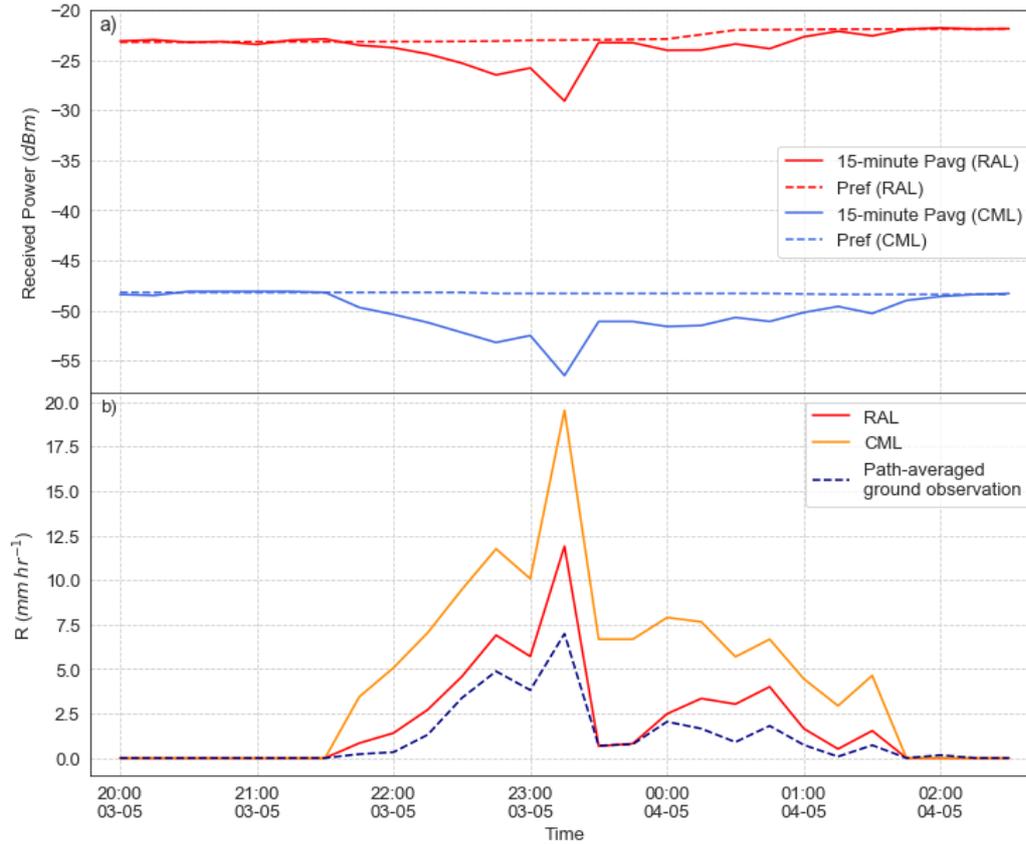


Figure 4-4: Time series of an event on the 3rd and 4th May 2018 using the 15-minute dataset of (a) received power levels (solid lines) from the CML, the RAL and the reference signal (dotted line); (b) rainfall intensities derived from the link attenuation using the R - k power law.

b) Stratiform event

Figure 4-5 shows an example of one of the stratiform rainfall events, which occurred on 18th August 2018. This event lasted for 84 minutes (equivalent to 1.4 h). During this event, the signal level fluctuated around 2 dB during the dry period. The specific attenuation obtained from the RAL showed a strong correlation with the observed attenuation based on the data from the two disdrometers (Site T = 0.83 and Site R = 0.82). However, the magnitude of the specific attenuation and the rainfall rate was overestimated by the RAL for the whole event. This overestimation was observed to be consistent throughout the entire event, but a different magnitude compared to the convective event and almost double in magnitude compared with the path-averaged ground observation estimates from the disdrometers.

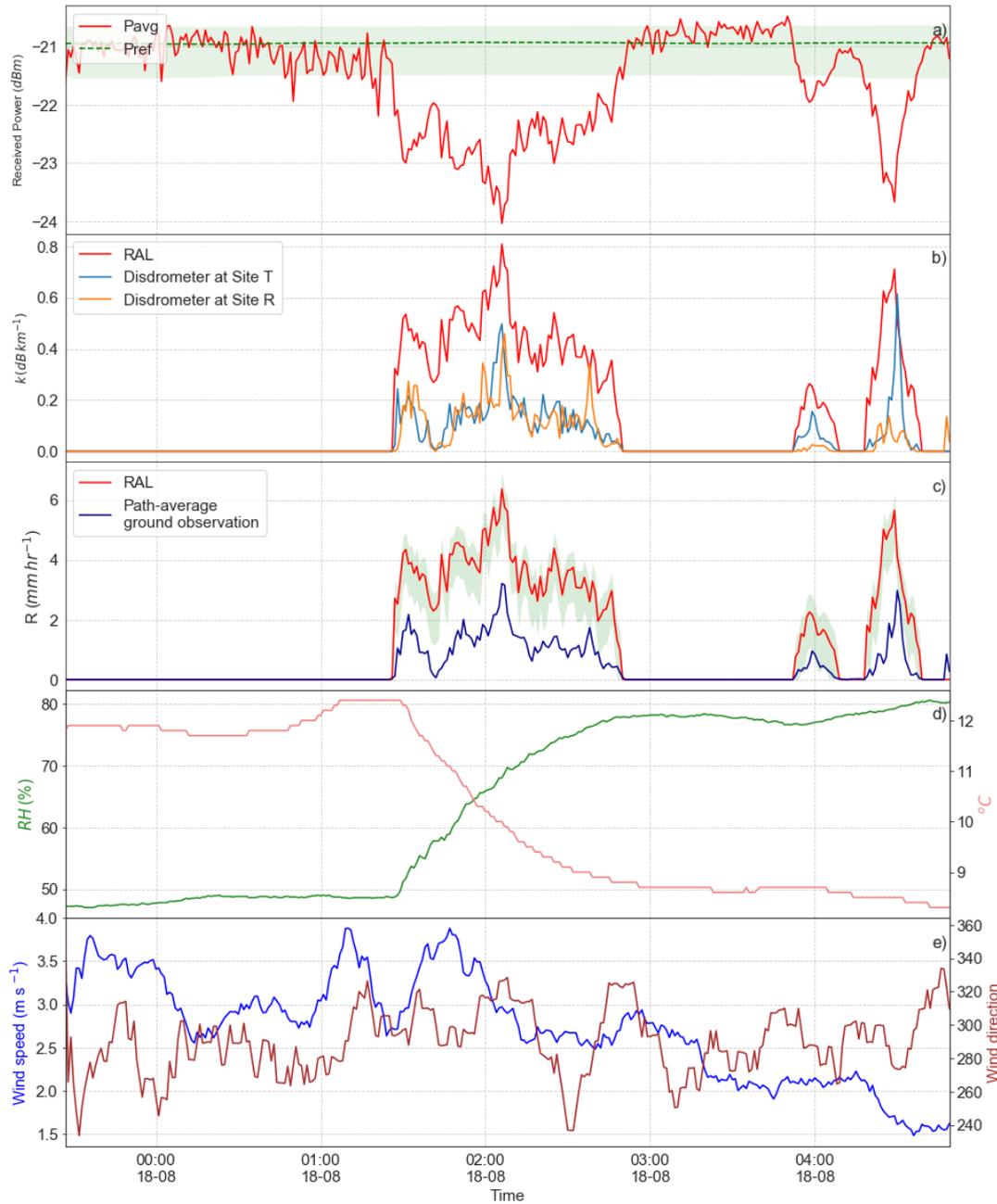


Figure 4-5: Same as Figure 4-2 but for an event on 18th August 2018.

Figure 4-6 shows the time series plot of the RAL and CML data for the same 18th August event. The variation of the power level from the RAL and the CML was very similar to the convective event, showing a correlation of 0.92. The signal behaviour of the RAL and CML were slightly different once the rainfall stopped; however, the CML signal took more time to return to the baseline compared with the RAL. This was also reflected in the rainfall estimation based on the CML and the RAL. Even though the rainfall estimates based on both the CML and RAL were overestimated, the magnitude of the overestimation was higher on

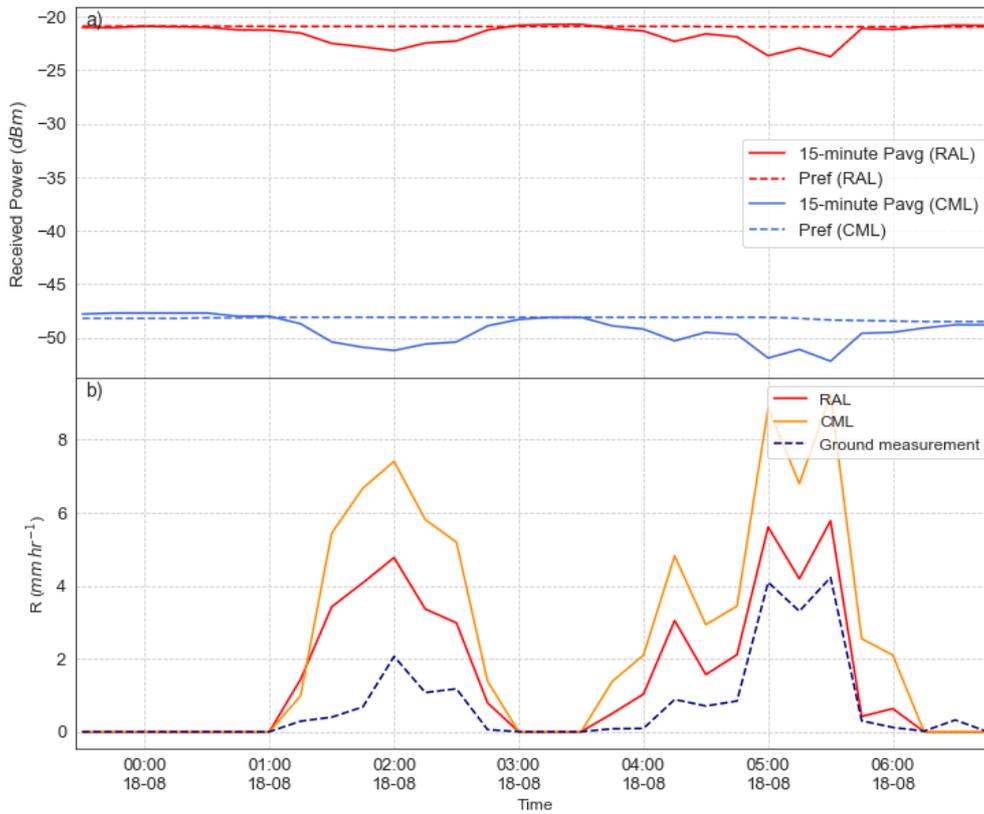


Figure 4-6: Same as Figure 4-4 but for an event on the 18th August 2018.

the CML estimates throughout the event. This difference was more likely related to the antenna film material used on the link, which may have different characteristics reflected in the wet antenna attenuation. For the RAL link, a hydrophobic material was used; however, for the CML, a standard polycarbonate was used as the antenna film.

c) Longest rainfall event

Figure 4-7 shows the longest rainfall event out of the 72 rainfall events observed. It occurred from the 16th till the 17th June 2018. This event lasted for 663 minutes (equivalent to almost continuous 11 hours), including both low and high-intensity rainfall. It started with low-intensity; however, there were several peaks of high-intensity rainfall, as well as some short intermittent dry periods during the rainfall event. Similar to the other two rainfall events discussed, both the RAL attenuation and the RAL-derived rainfall intensity showed a strong correlation with the path-averaged ground observations with correlation coefficients greater than 0.8. Again, the rainfall intensity was overestimated from the beginning of the event; however, the magnitude of the overestimation was lower compared with the two other events.

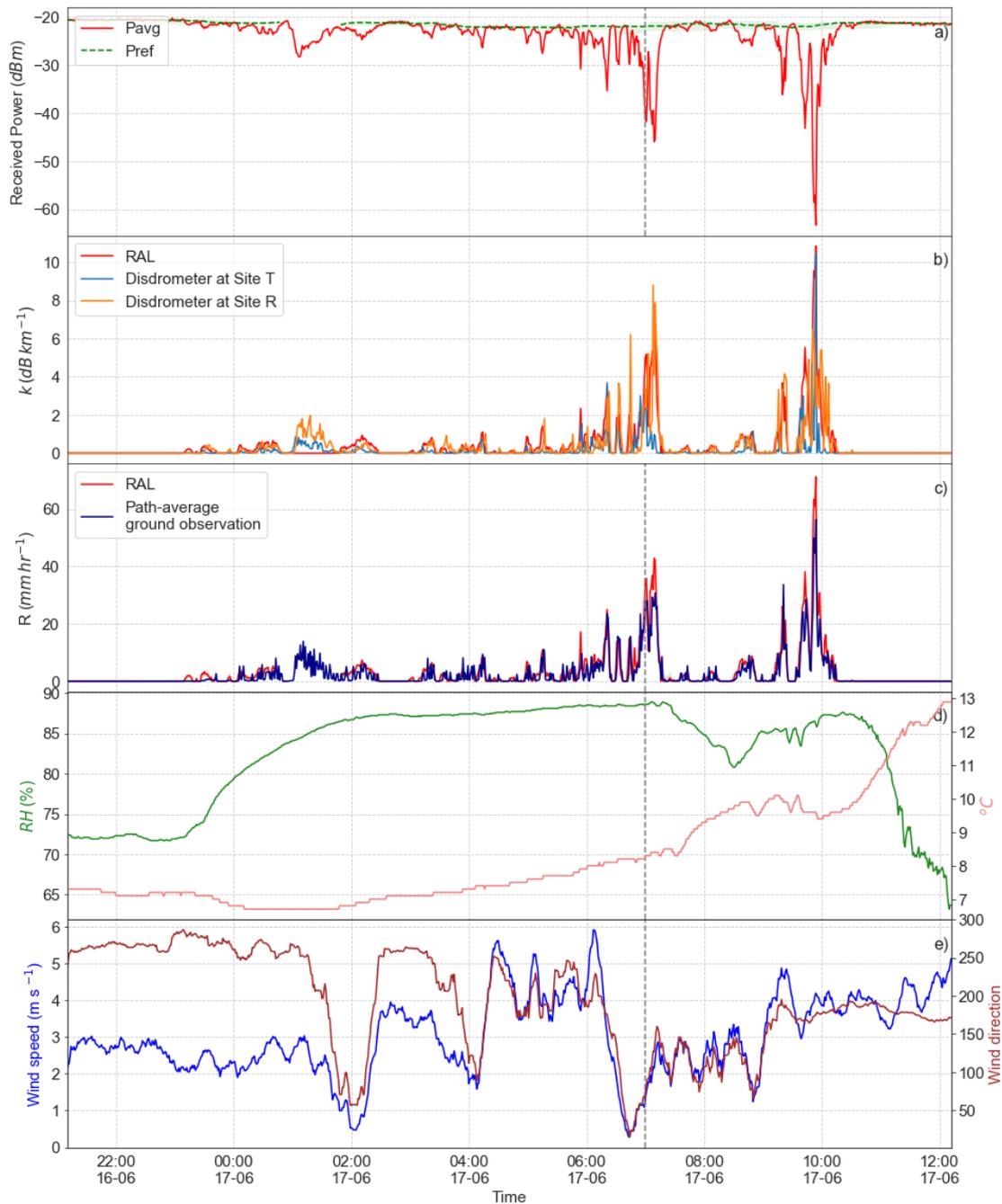


Figure 4-7: Same as Figure 4-2 but for an event on the 16th and 17th June 2018.

The main differences were observed during the peaks when the wind speed was increasing; the magnitude of the overestimation at such peaks was lower compared with when the wind speeds were decreasing. Table 4-4 shows the variability of a and b parameters at the two sites compared with the parameter derived based on the whole experimental dataset.

Table 4-4: Comparison of the a and b parameters of the rainfall-attenuation relationship for three rainfall events.

		a	b
3 rd and 4 th May	Site T	8.17	0.90
	Site R	7.95	0.89
18 th August	Site T	7.63	0.89
	Site R	6.28	0.82
16 th and 17 th June	Site T	7.29	0.98
	Site R	7.65	0.93
Whole experimental data		7.79	0.93

Figure 4-8 shows the time series plot of the RAL and CML data for this long rainfall event. The variation of the power level for the RAL and CML were very similar, showing a correlation of 0.95. The derived rainfall based on both the RAL and CML were overestimated as in the other two cases; however, the major difference observed was during the peaks of the rainfall event; both the RAL and CML estimated a similar magnitude of rainfall. This was mainly due to lower excess attenuation estimated by the CML compared with the two previous events. Even though the difference between the baseline and received signal level was higher for the CML most of the time, during the peak of the rainfall event, the excess attenuation was almost equal to the RAL. This is likely to be due to the increasing wind speed during the rainfall peak, as also observed on the 1-minute dataset, resulting in the reduction of the excess attenuation caused due to the wet antenna. After the peak rainfall event, the CML signal level showed similar behaviour as in the previous two examples showing an overestimation.

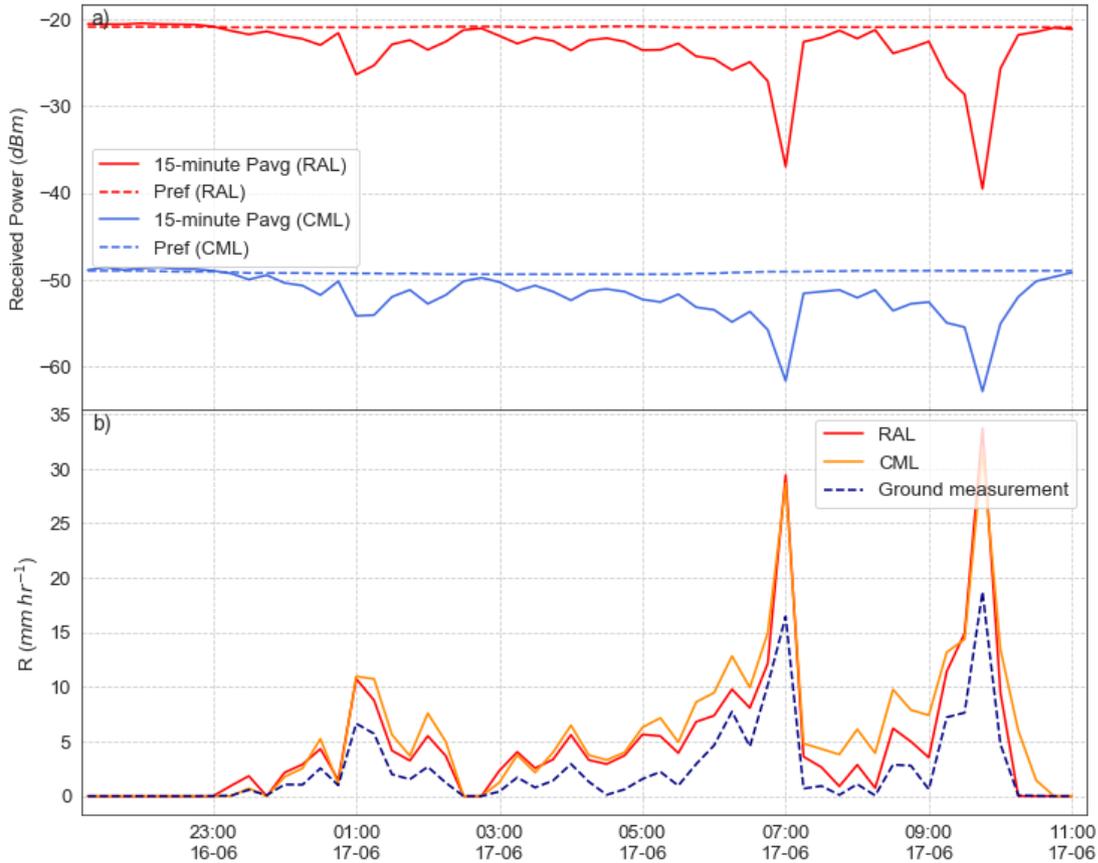


Figure 4-8: Same as Fig 4-4 but for an event on the 16th and 17th June 2018.

d) Summary of all 72 rainfall events

Figure 4-9 shows a summary of the rainfall retrieval for all 72 rainfall events analysed. The RAL-derived rainfall estimates showed a strong correlation with the path-averaged ground observations with a $\rho = 0.87$; however, there was an overestimation of about 90% with a multiplicative factor (regression slope of 1.24). This overestimation was also observed in the double mass curve plot for RAL and ground observations in Figure 4-9(b); a plot commonly used for the continuous evaluation of rainfall. This plot suggested that the total rainfall amount estimated based on RAL was almost double that of the path-averaged ground observation. This overestimation could be attributed mainly to phenomena other than rain such as wet antenna, dew formation, which was erroneously processed as rain in the retrieval algorithm. This was also due to the fact that for this analysis, no correction was applied to the rainfall retrieval process. Also, when comparing the specific attenuation in Figure 4-9 (c), similar results were obtained with slightly higher bias and CV compared with the rainfall intensity, with the multiplicative factor (regression slope) being 1.29. This indeed indicated

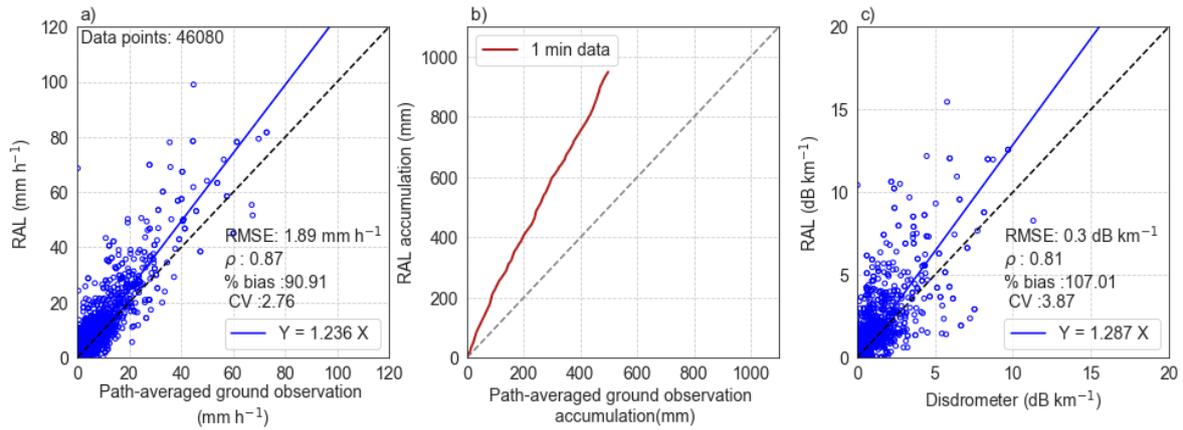


Figure 4-9: (a) Scatterplot of rainfall rate estimated for the RAL versus path-averaged ground observation based on the 1-minute dataset; (b) double mass curve derived based on the RAL and path-averaged ground observation; and, (c) scatter plot of the specific attenuation measured from the RAL versus based on two disdrometers at two ends.

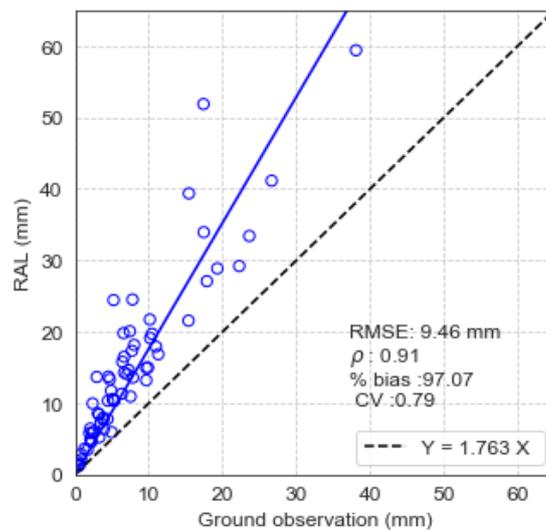


Figure 4-10: Scatter plot of total rainfall depth from the RAL and ground observation for all 72 rainfall events; each data points on the scatter plot represent one rainfall event.

that the uncertainties in the $R-k$ relation did not statistically explain uncertainties in the rainfall estimates.

Further, to understand the performance of all 72 rainfall events, total rainfall depth derived based on the RAL and path-average ground observation for those events were compared in Figure 4-10. Rainfall depths for all 72 rainfall events were overestimated by about 97% with

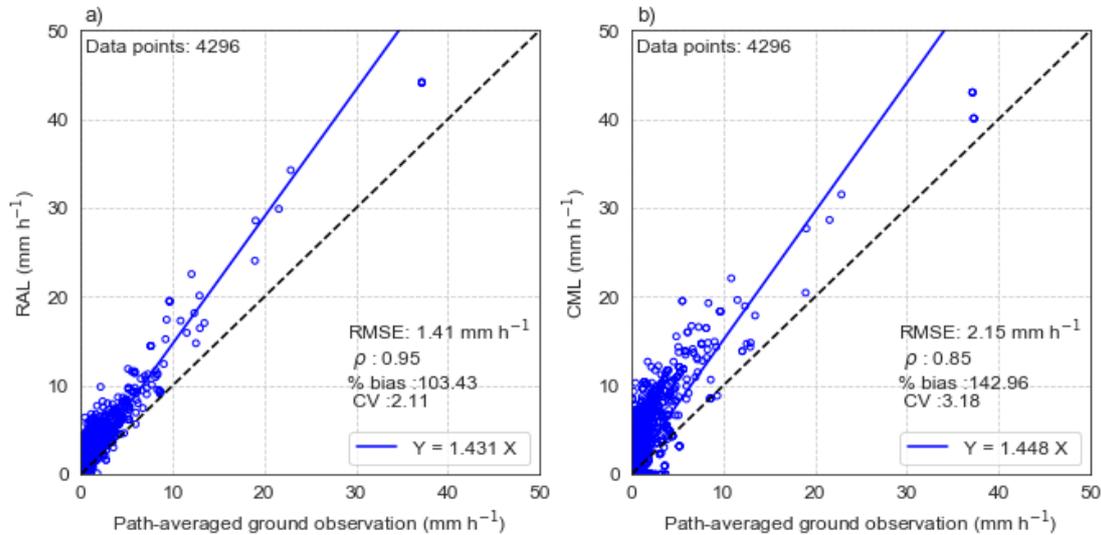


Figure 4-11: Scatterplots of the link derived rainfall intensities versus path-averaged ground observed rainfall as obtained from two disdrometers and three tipping bucket rain gauges for (a) 1-minute, and (b) 15-minute.

a multiplicative regression scale of 1.76, however with a high correlation coefficient of 0.91 indicating the best match between the RAL estimation and path-averaged ground observation.

Further analysis was made based on the operational network. To make a fair comparison of the CML and the RAL link performance, raw RAL data were resampled to 15-minute average. Figure 4-11 shows the comparison of the scatter plot with the ground observation based on the RAL and CML. The RAL showed a bias of 103%, whereas the CML showed a higher magnitude bias of 142%. Also, the correlation coefficient was higher for the RAL link compared with the operational microwave link data. In addition, there was more noise in the CML derived rainfall estimates (CV value of 3.18) compared with the RAL link (CV value of 2.11).

Further comparison based on the RAL and the CML derived rainfall estimates is provided as a complementary cumulative distribution function (CCDF) of the estimated rainfall rate in Figure 4-12. Both the RAL and the CML derived rainfall estimates were higher compared to the path-average ground observation; however for the higher probabilities of occurrence, the RAL estimates were closer to the path-average ground observation compared with the CML. At probabilities of 0.01, both showed similar estimates, and the RAL even showed higher estimates at the lower probabilities, which represent higher intensity rainfall.

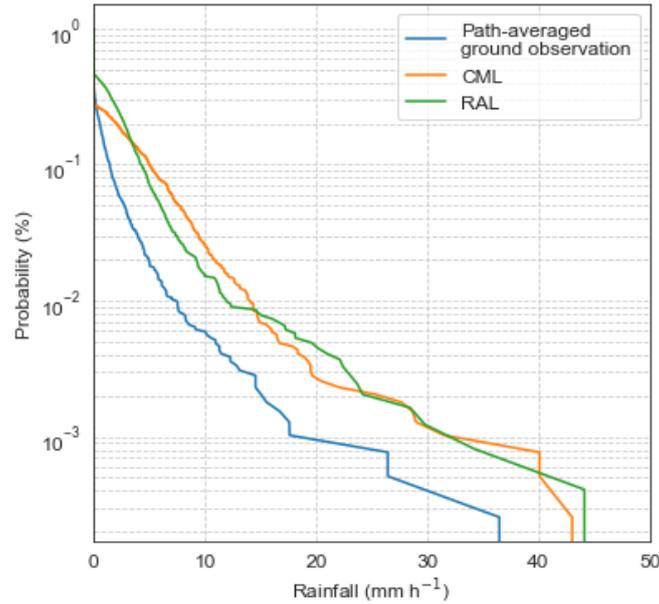


Figure 4-12: Complementary cumulative distribution function of path-average ground observation and derived based on the RAL and CML.

4.4.3 Impact of sampling strategies

Figure 4-13 shows the scatter plot for 5-minute sampling data for three different strategies (Average, MinMax and Instantaneous). All three strategies showed a strong correlation with the path-averaged ground observation, with all being greater than 0.80 and the Average strategy showing the highest correlation being 0.92. The Average strategy also showed the least bias of 112%, with the highest overestimation of 185% shown by the Instantaneous strategy. However, comparing the RMSE and the CV for the three strategies, the MinMax strategy showed the lowest values, indicating the lower spread of the results compared with the other two strategies. Similarly, looking at the data based on one sampling strategy versus the others, there was a strong correlation among them for all three cases, with correlations greater than 0.8. However, there was a larger spread among the data from the Instantaneous when compared with the Average and Instantaneous strategies as indicated by the higher CV. The lowest spread was observed between the Average and MinMax data.

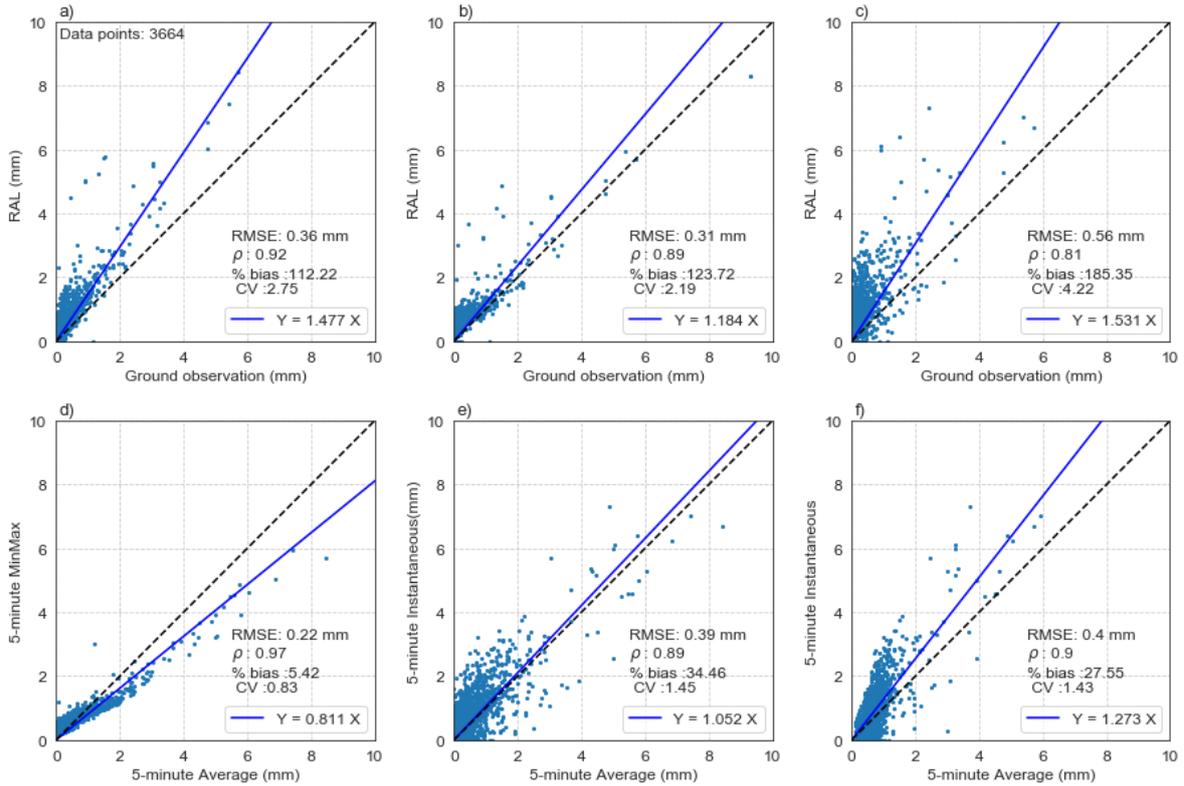


Figure 4-13: (a)-(c) Scatter plot for 5-minute sampling derived based on the RAL versus path-averaged ground observation; (d) Average versus MinMax; (e) Average versus Instantaneous; and (f) MinMax versus Instantaneous.

Figure 4-14 shows the scatter plot for 15-minute sampling data for three sampling strategies. For this case, the Average and the MinMax strategies showed a strong correlation, with the ground observation greater than 0.8. However, the Instantaneous strategy showed a correlation of only 0.52. The percent bias for the MinMax and Instantaneous strategies were higher compared to the Average strategy indicating that the RAL-derived rainfall estimates for these two sampling strategies significantly overestimate rainfall rates. Also, the spread in the results was higher, indicating more noise in the results compared with the Average strategy. Considering all statistics, RMSE, ρ , percent bias and CV, the 15-minute Average strategy showed the best performance of the three strategies tested. In order to evaluate the performance of one sampling versus another, scatter plots among the three strategies are shown in Figure 4.16 (d)-(e). In this case, the Average and the MinMax strategies showed a strong correlation of 0.88 and a low CV of 1.0. However, comparing the Instantaneous with the Average and MinMax strategies showed a lower correlation and more spread with higher CV values. This was mainly due to the fact that there was only one snapshot of data during the period of 15-minute as opposed to the other two sampling data.

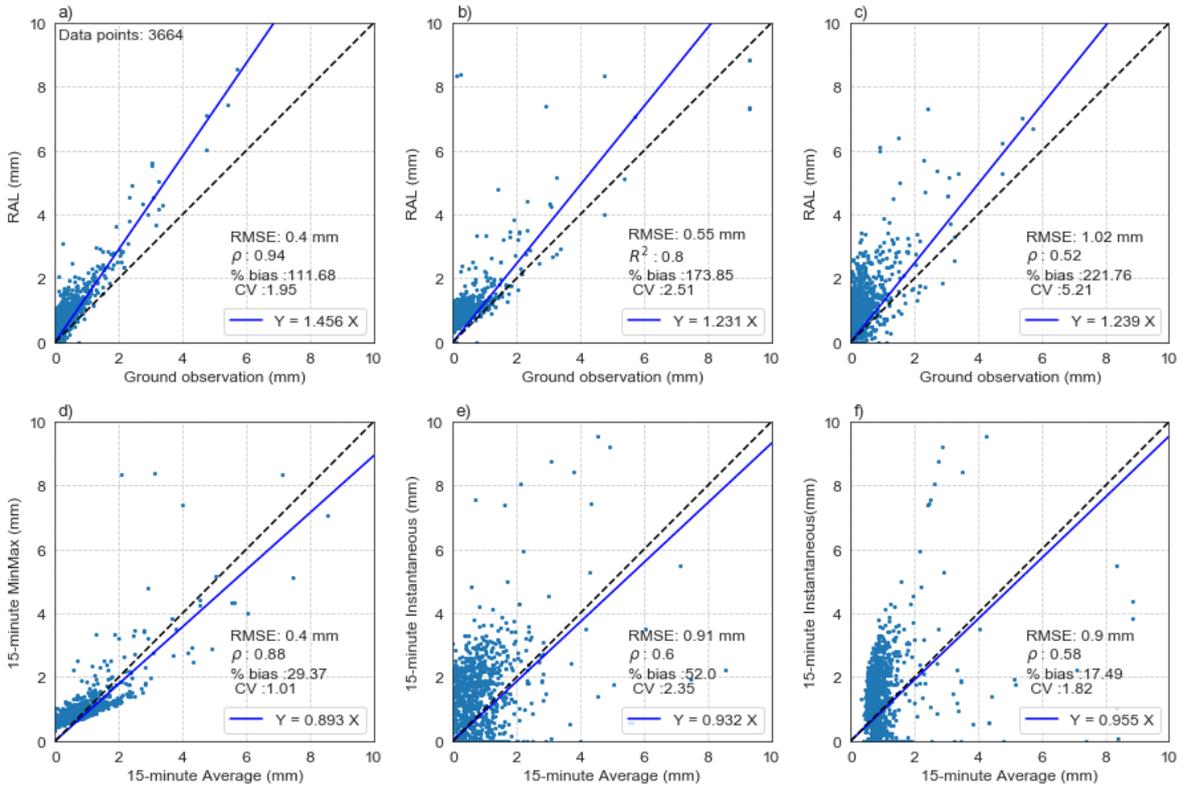


Figure 4-14: Same as Figure 4-13 but for 15-minute sampling.

To further analyse the performance of rainfall retrieval based on different sampling strategies, the complementary cumulative distribution function (CCDF) of rainfall depth was plotted, as shown in Figure 4-15. As there was no correction applied for the RAL derived rainfall, 1-minute RAL derived rain was considered as the benchmark estimate along with the path-averaged ground observation. For 5-minute sampling data, derived rainfall based on the Average strategy was found to be quite similar for all the probabilities; however, some differences were observed for the MinMax and Instantaneous strategies. For the MinMax data, the probabilities of occurrence of rainfall above 1 mm were observed to be lower compared with the Average sampling strategy. Similarly, there were also differences observed based on the Instantaneous sampling strategy.

Similarly, for the 15-minute sampling data, rainfall derived based on the Average sampling strategy was closer to the benchmark 1-minute data for all the probabilities. For the 15-minute MinMax strategy, the probability of occurrence of rainfall was higher for rainfall depth up to 9 mm; however, for the higher rainfall depths, the probability was lower compared with the other two sampling strategies.

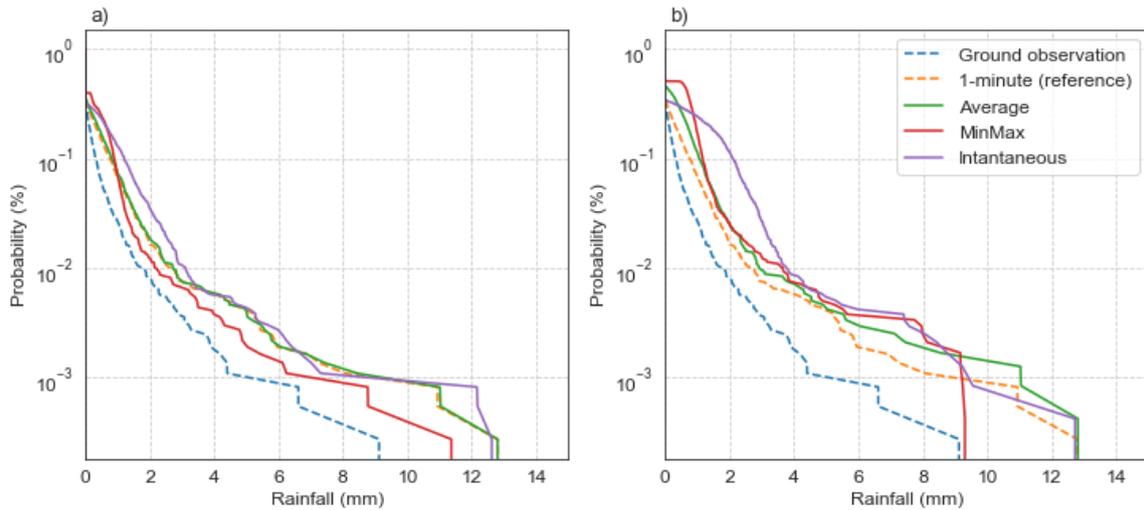


Figure 4-15: Complementary cumulative distribution plot of rainfall depth for (a) 5 and (b) 15-minute sampling data.

Figure 4-16 shows the scatter plot of rainfall depth derived based on 5 and 15-minute sampling strategies compared with the 1-minute benchmark. Figure 4-16 (a-c) shows the scatter plot of 5-minute rainfall depth based on the Average, MinMax and Instantaneous strategies compared with the 1-minute Average strategy (considered as the benchmark). For all three cases, the correlation coefficient was higher than 0.9; however, the percent bias and CV was higher for the MinMax and Instantaneous strategies compared with the Average strategy. Similarly, as in Figure 4-16 (d-e) all statistics (RMSE, correlation coefficient, percent bias and CV) showed better performance of the 15-minute Average strategy compared with the MinMax and Instantaneous strategies for the same time interval. This demonstrated that the uncertainty in the rainfall estimates from both the 5 and 15-minute Average strategy was lower compared with the MinMax and Instantaneous strategies.

Comparing the 15-minute and 5-minute sampling intervals, Figure 4-16 (g-i) showed that the Average strategy had the best performance with a correlation of 0.99, percent bias of only 3.6% and CV of 2.67. This performance deteriorated for the MinMax strategy with the percentage bias and CV increased and showing more scatter in the rainfall depths.

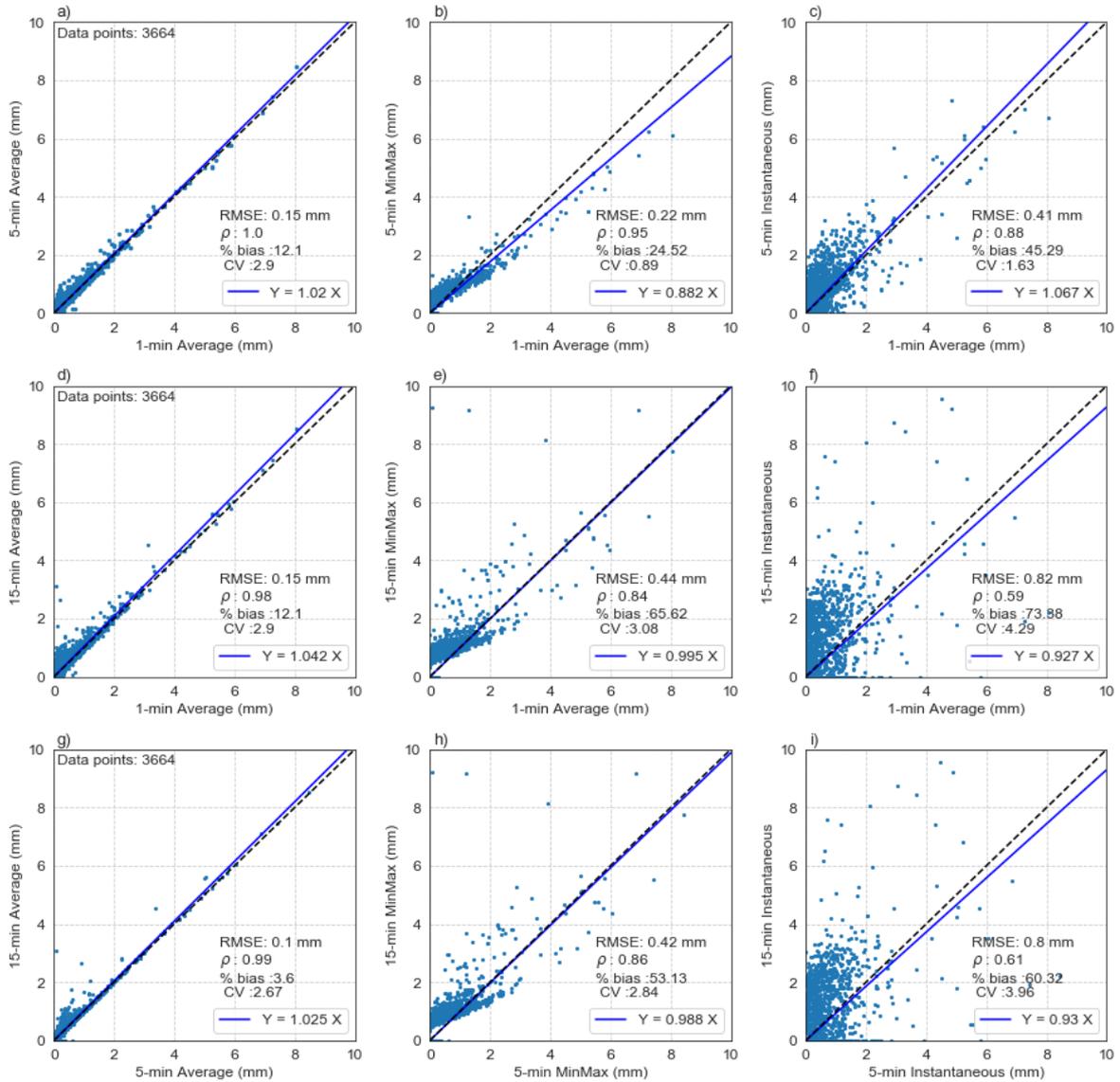


Figure 4-16: Scatter plot of (a) 5-min Average vs 1-min Average; (b) 5-min MinMax vs 1-min Average; (c) 5-min Instantaneous vs 1-min Average; (d) 15-min Average vs 1-min Average; (e) 15-min MinMax vs 1-min Average; (f) 15-min Instantaneous vs 1-min Average; (g) 15-min Average vs 5-min Average; (h) 15-min MinMax vs 5-min MinMax; and (i) 15-min Instantaneous vs 5-min Instantaneous

The distribution of the rainfall based on the three different sampling strategies across all 72 rainfall events are presented in Figure 4-17. Here, the Average sampling strategy again showed better performance compared with the MinMax and Instantaneous strategies for both 5-minute and 15-minute intervals, with the mean rainfall depth being closer to the reference (1-minute data) and the spread being lower compared with the other two strategies, indicating

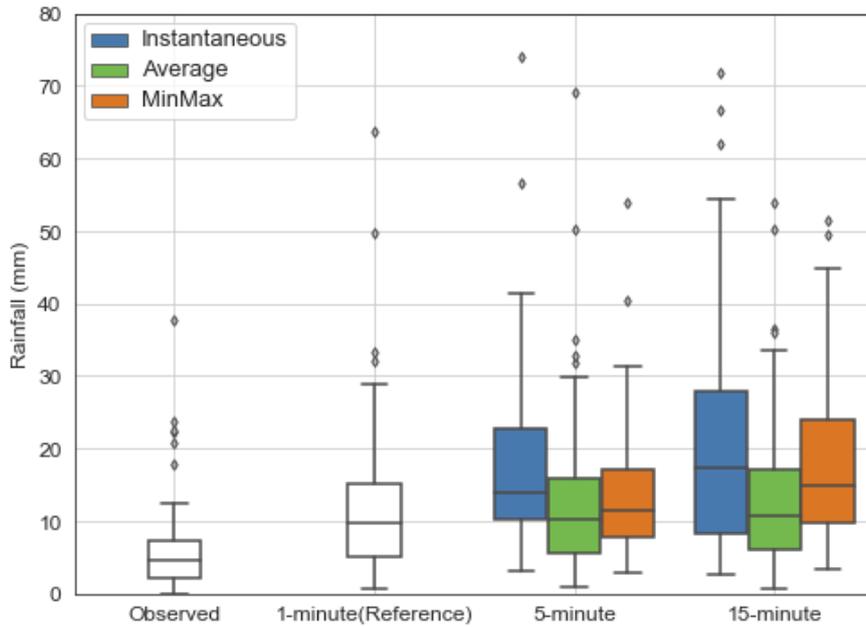


Figure 4-17: Distribution of rainfall based on three sampling strategies for 5 and 15-minute compared with observed and reference (1-minute).

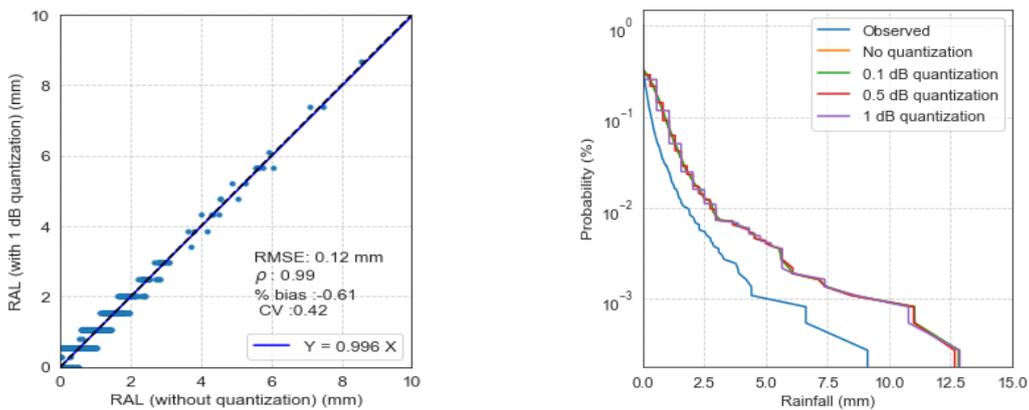


Figure 4-18: (a) Scatter plot of rainfall depth based on the RAL data with no quantization vs 1 dB quantized data; and (b) Complementary cumulative distribution function for various level of quantized data.

that the Average sampling strategy provides more reliable rainfall estimation compared with the MinMax and Instantaneous strategies.

4.4.4 Impact of quantization errors

To understand the impact of data quantization on the overall rainfall retrieval process, the raw RAL data was artificially modified with different quantization levels from 0.1 to 1dB. The rainfall retrieval algorithm used was exactly the same as the one used previously for 1-



Figure 4-19: Photograph of the receiver antenna just after receiver antenna sprayed with water during the wet antenna experiment. Water droplets formation was observed on the antenna film.

minute data. Figure 4-18 shows a summary of the results for the quantization. The scatterplot with the 1 dB (highest quantization) versus no quantization in Figure 4-18 (a) showed only minor variation in the rainfall results with a correlation coefficient of 0.99 and with -0.6% bias. Similarly, the cumulative distribution function plot with three different levels of quantization of data showed a minor variation in the result, indicating that the impact of quantization up to 0.5 dB for the RAL link (24 GHz) was negligible compared with other uncertainties like data sampling strategy. This is similar to the results obtained by Leijnse et al. (2008) using simulated microwave link data from radar when comparing with 0 dB and 1 dB quantization level.

4.4.5 Wet antenna attenuation experiment

A simple experiment was conducted to assess the impact of wet antenna attenuation on rainfall retrieval. For this, during a dry, sunny day (23rd August 2019) with an ambient temperature of 24°C, the receiver antenna of the RAL link was artificially wetted with a water spray bottle. As the transmitter antenna was installed on a 30m tower, a similar experiment could not be conducted on the transmitter antenna; however, as the antenna material was the same, it was assumed to have a similar effect. Figure 4-19 shows photographs taken during this experimental setup, with water droplets observed on the antenna film after the water spray was applied on the antenna surface.

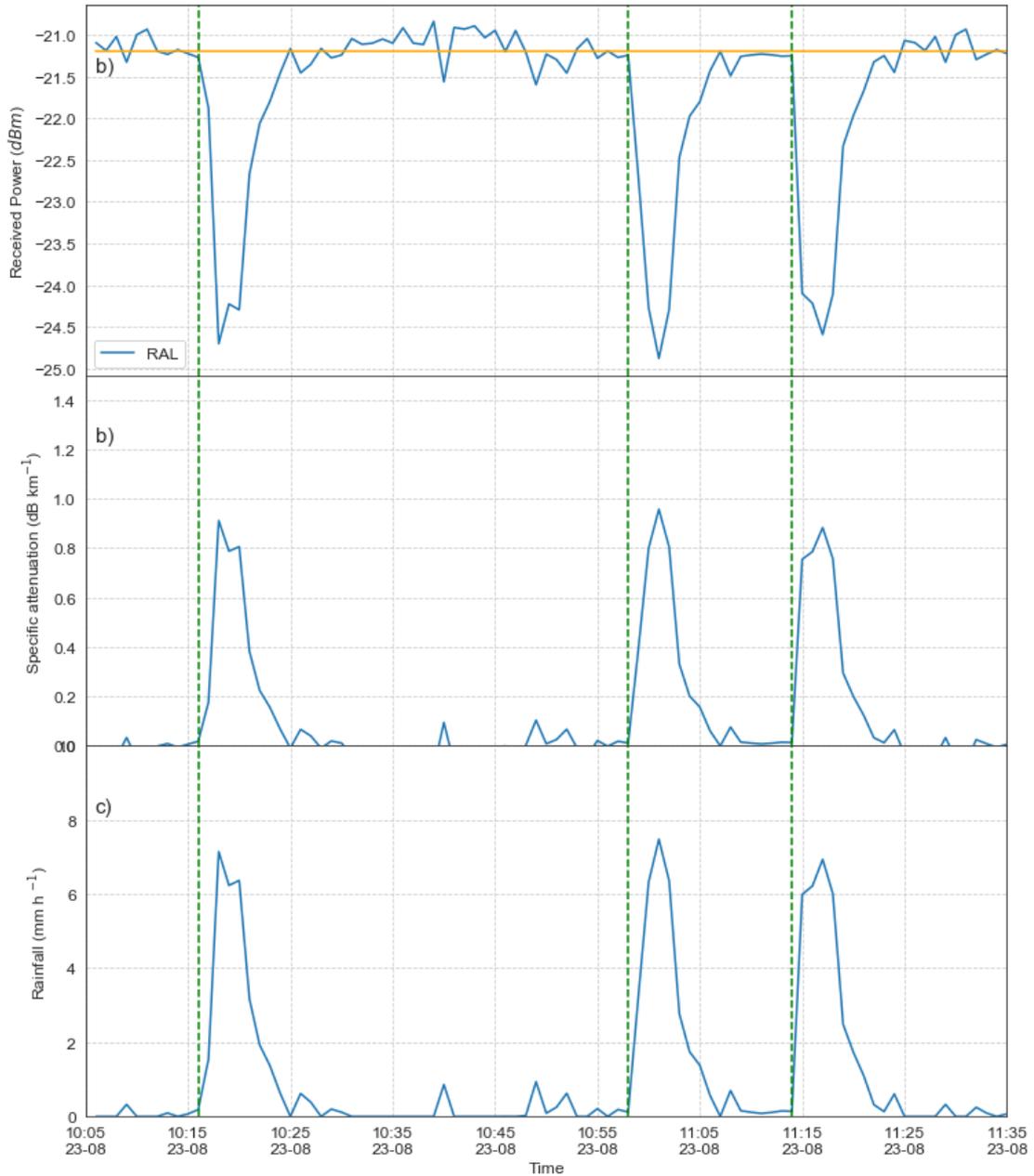


Figure 4-20: Time series of the wet antenna simulation experiment conducted on 23rd August 2019. Three experiments were conducted at different times (dotted line on the plot indicate the start time of the water spray [I :10:16 AM, II :10:58 AM and III: 11:14 AM]); (a) received signal level with the baseline signal plotted on orange colour; (b) specific attenuation of the RAL link; and (c) rainfall equivalent using the R-k relationship.

Figure 4-20 shows the resulting impact on the received signal level data. Once the antenna got wet, it resulted in an attenuation of up to 3.5 dB. This corresponds to a rainfall intensity of 7.5 mm h⁻¹ at 24 GHz using the power relationship derived in equation 4.6. The drop in

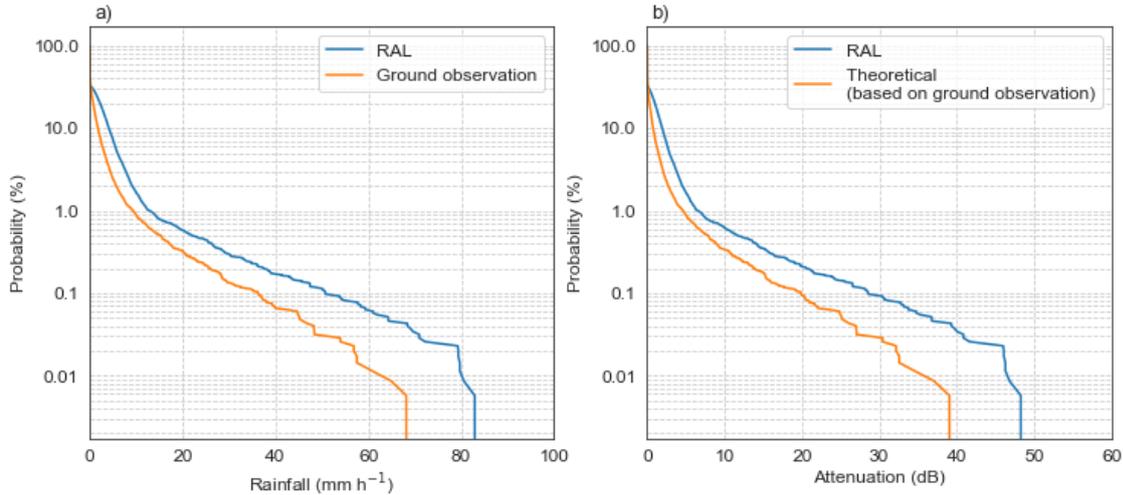


Figure 4-21: Complementary cumulative distribution function of (a) Rainfall measured by the RAL and ground observation, and (b) total path attenuation. For attenuation, the ground observation was obtained based on the theoretical raindrop attenuation.

the signal level was immediate but took almost 2 minutes to reach the maximum attenuation for all three cases. It took another 10 to 12 minutes to return to the original signal level. For all three experiments, the drop in the signal level was not the distinct decay function observed by Minda and Nakamura (2005) for their wet antenna experiment. This result was more similar to the experiment conducted by van Leth et al. (2018), where no distinct distribution was observed for their experimental link at 38 GHz. However, due to the different drying rate, the signal level in this case returned to the original level more quickly than what was observed by van Leth et al. (2018). Even though the formation of a few large beads was observed on the corner of the antenna surface during the start of the experiment, this was slowly evaporated, thus showing the gradual reduction in the attenuation level.

4.4.6 Modelling wet antenna attenuation

Figure 4-21 (a) shows the CCDFs of measured rainfall intensity based on the RAL and path-averaged ground observation. For all probabilities, the RAL-derived rainfall intensities were higher than the ground measurements. As an example, for 10% probability, the RAL-derived rainfall intensities were about 6 mm h⁻¹ as opposed to the ground measurement that was only 2 mm h⁻¹. Similar trends were observed on the measured attenuation based on the RAL and the ground observations, as shown in Figure 4-21 (b). For the total attenuation, theoretical attenuation was derived based on the power relationship using the path-averaged ground-

based rainfall measurement. This difference was mainly attributed to the wet antenna attenuation. Besides the wet antenna attenuation, this difference could also be due to variability in the drop size distribution along the path as the RAL link was quite long (~ 4 km). However, the theory predicts that the drop size distribution at this frequency (24 GHz) would have a negligible impact on the $R-k$ relationship and subsequent rainfall retrieval.

Figure 4-22 shows a direct comparison between the measured total attenuation and path-averaged ground observation. Pearson's correlation coefficient and Spearman's rank correlations were calculated to identify any relationship between these two parameters. The Pearson's correlation was observed to be 0.83, and the Spearman's rank coefficient was 0.86, indicating that the total attenuation from RAL is a function of rainfall rate. Also, the rank correlation between the rainfall intensity and wet antenna attenuation (WAA) estimated from the difference between the theoretical attenuation and measured attenuation from the RAL was 0.97, providing evidence that wet antenna attenuation and rainfall intensity were correlated.

a) Model parameterization

The two wet antenna correction models are compared in Figure 4-23. The results obtained were in line with the result of WAA at 27 GHz reported by Leijnse et al. (2008). The total WAA observed at both antennas combined to a total of about 2 dB for light rainfall with a rate lower than 4 mm h^{-1} , 4 dB for moderate rainfall with a rain rate of 30 mm h^{-1} and 5 dB for rain rates of 100 mm h^{-1} . The difference between measured and predicted attenuation, without considering the wet antenna effect, increased with decreasing probability (increasing rain rate). With the wet antenna effect included, the difference between measurements and predictions was almost constant for all levels of probability. In absolute terms, the difference between the measured and predicted attenuation for 0.1% probability level was reduced from 29 dB to 23 dB. The disagreement between the measurement and predictions can be seen more prominently for lower percentages of time i.e for high rainfall intensity. This is more likely due to the variation of drop size distribution during high-intensity rainfall.

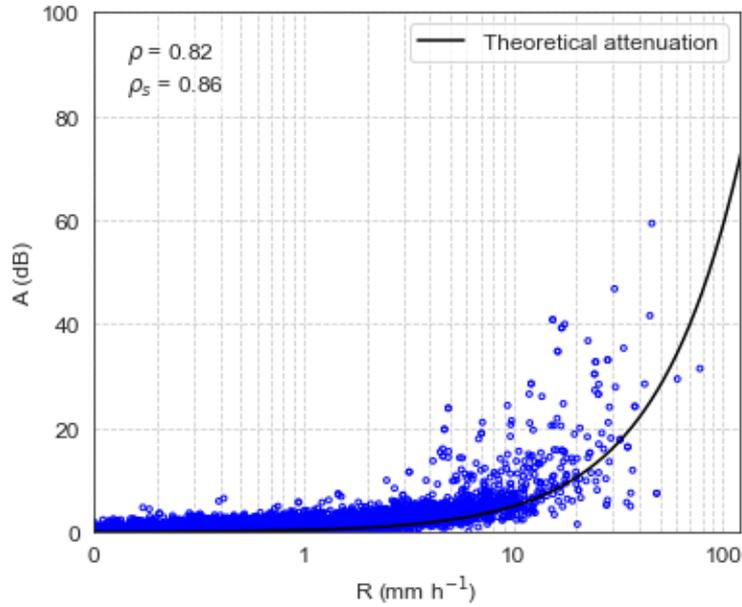


Figure 4-22: Measured attenuation related to the path-averaged rainfall intensities shown at log-scale. Pearson’s correlation (ρ) and Spearman rank correlations (ρ_s) are displayed on the plot.

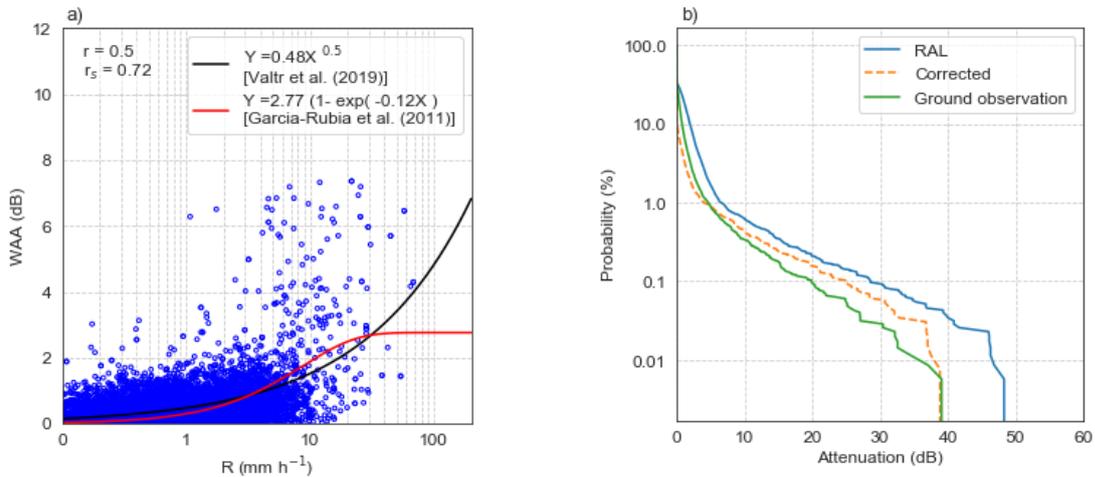


Figure 4-23: (a) Fitting of WAA of antenna pairs on rate for RAL using the least squares method, rain rate plotted on a log scale and; (b) CCDFs of attenuation based on the RAL and ground observation based on the theoretical raindrop attenuation with the correction applied.

b) Applying WAA on RAL link

Figure 4-24 shows the scatter plot after applying wet antenna correction based on three different models. The percent bias, RMSE and CV after applying the wet antenna correction

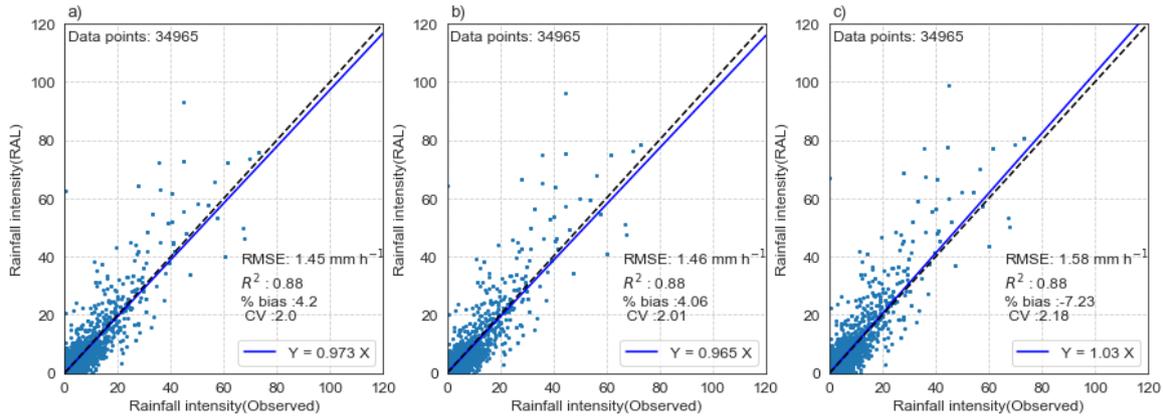


Figure 4-24: Scatter plot of rainfall intensity derived from the RAL and path-averaged ground observation after applying the wet antenna correction based on (a) the power-law model of Valtr et al. (2019); (b) the exponential decay model of Garcia-Rubia et al. (2011); and (c) a constant wet antenna attenuation of 1.1 dB.

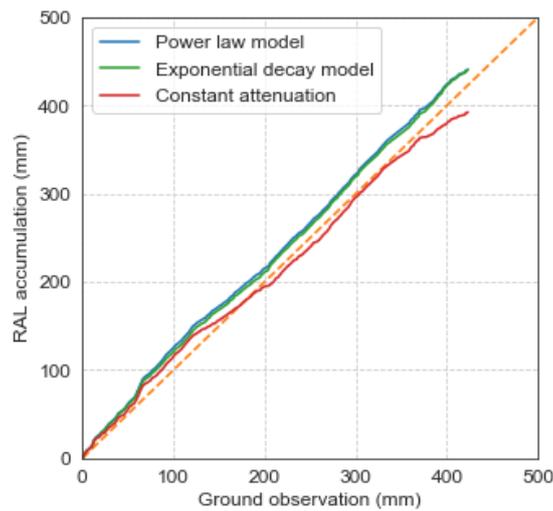


Figure 4-25: Double mass curve of the RAL derived rainfall versus path-averaged ground observation after applying WAA

reduced for all three models. The percent bias reduced from $\sim 90\%$ to 4.2% using the model proposed by Valtr et al. (2019), to 4.06% using Garcia-Rubia et al. (2011), and to -7.23% using a constant attenuation model. Other statistics (RMSE and CV) were also reduced, with the overall performance for the models based on Valtr et al. (2019) and Garcia-Rubia et al. (2011) found to be very similar. Even though the overall performance of a constant attenuation model was lower compared with the other two, it showed a significant improvement of percent bias in rainfall retrieval.

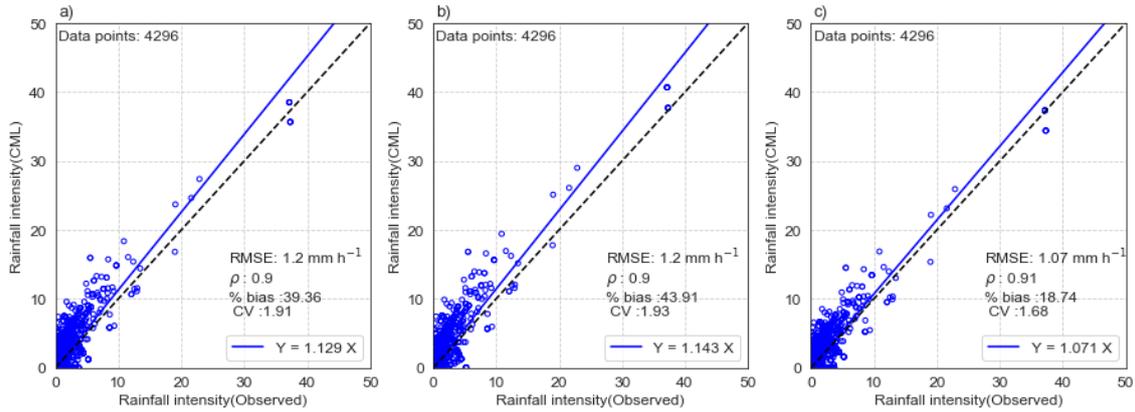


Figure 4-26: Scatter plot of the CML data after applying wet antenna correction using (a) the power law model; (b) the exponential decay model; and (c) the constant attenuation model.

This result has been further plotted as a double mass curve plot in Figure 4-25. Both the power-law and exponential model this was slightly overestimated (above the 1:1 line), with the magnitude similar to the path-averaged ground observation. Similarly, the constant attenuation model was slightly underestimated but was very close to the path-averaged ground observation.

c) Applying WAA on CML data

The WAA parameters derived based on the RAL were applied to the CML data, as shown in Figure 4-26. All three models showed substantial improvement for the rainfall retrieval after applying the WAA correction. However, considering the statistics (RMSE, ρ and percent bias), the simple constant attenuation model was found to be performing better compared with the other two (power-law and exponential decay) models. The main reason behind this is that the material used on the RAL antenna was different from the CML antenna. The RAL antenna used hydrophobic material as an antenna cover; however, the CML antenna used a standard polycarbonate material. This impacts the parameters of both the power-law and exponential decay models.

4.4.7 Comparison with the analytical result

Figure 4-27 shows the distribution of the accumulated rainfall amount for all 72 events based on the empirical and analytical approaches. These results are based on the 15-minute minimum and maximum RSL data, with no wet antenna correction applied. For the empirical

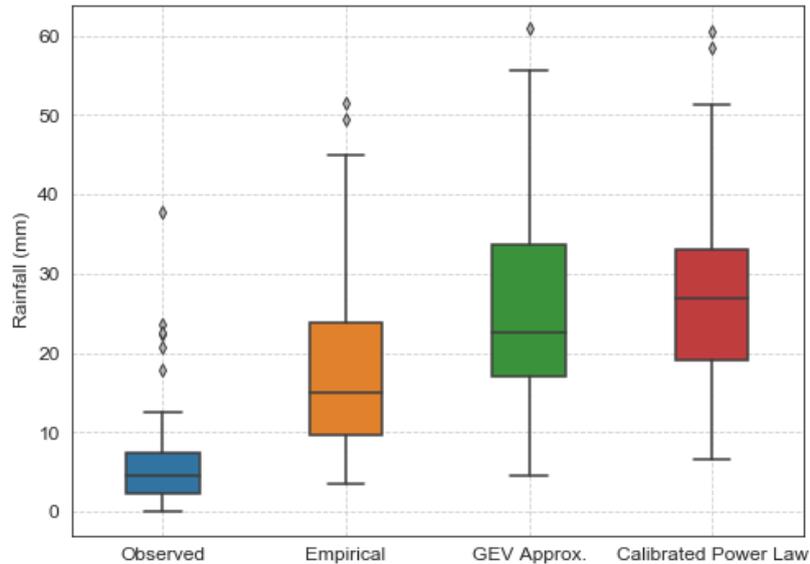


Figure 4-27: Distribution of accumulated rainfall for 72 rainfall events based on analytical method compared with the empirical approach and observed data.

method, the weightage factor (α) was considered as 0.31. The rainfall amounts for all 72 rainfall events were overestimated for both analytical methods (GEV approximation and Calibrated power-law), with a higher mean rainfall amount compared to the empirical and observed values. The spread of the GEV approximation was found to be the largest compared with the empirical and calibrated power-law. This indicated that the empirical method used in this study provided rainfall estimates closer to the ground observations even without applying any further correction.

4.5 Chapter summary

This study validated the rainfall retrieval algorithm based on a dedicated experimental microwave link. Additionally, this synthesis study recommended a new parameter set for the power-law model which will be used for the rainfall retrieval from the commercial microwave link in Chapters 5 and 6. Although this chapter provided insights on the impact of data sampling strategy on rainfall retrieval accuracy, there is a need to understand the impact on the whole rainfall retrieval process, including the dry/wet classification which will be conducted in Chapter 5.

Chapter 5

Validation of rainfall retrieval using commercial microwave links

This chapter provides an evaluation of rainfall retrieval using commercial microwave link signal extinction data for the Greater Metropolitan area of Melbourne. An open-source algorithm named RAINLINK was used for this analysis. For this study, the parameters of RAINLINK were calibrated using the commercial microwave link dataset for Melbourne. Additionally, this chapter provides a comparison of the rainfall retrieval performance using two common data sampling approaches for the same period, with data based on minimum-maximum and average sampling from the same link paths compared. This chapter is the basis of an article submitted to the Journal of Hydrology.

5.1 Background

Backhaul towers operated by Mobile Network Operators (MNO) have proven to be a complementary source of rainfall information (Messer, 2018; Uijlenhoet et al., 2018; Chwala and Kunstmann, 2019). This technique of rainfall measurement gained popularity with feasibility and validation studies undertaken for a variety of locations around the world including: Brazil (Rios Gaona et al., 2015), Burkina Faso (Doumounia et al., 2014), Czech Republic (Fencl et al., 2013; Fencl et al., 2017), Germany (Chwala et al., 2012; Chwala et al., 2016; Smiatek et al., 2017; Graf et al., 2020), Israel (Messer et al., 2006; Goldshtein et al., 2009), Italy (Roversi et al., 2020), The Netherlands (Leijnse et al., 2007; Overeem et al., 2011, 2013, 2016b; de Vos et al., 2019), Pakistan (Sohail Afzal et al., 2018) and Switzerland (Bianchi et al., 2013). These validation studies have been conducted based on a few links to a couple of thousand links covering an entire country such as The Netherlands (Overeem et al., 2013) and Germany (Graf et al., 2020). The temporal resolution of such CML rainfall estimates typically varies from a few seconds to 15 minutes, with most telecommunication operators sampling the received signal level (RSL) at 10 Hz but storing it at a much coarser temporal resolution. In most studies, 15-minute minimum and maximum RSL data, as stored operationally by the MNO's network management systems, were used for rainfall retrieval (Leijnse et al., 2007; Goldshtein et al., 2009; Overeem et al., 2011, 2016b; Rios et al., 2017). There have been a few studies using 1-min and even higher temporal resolution, up to a

second, instantaneous RSL data for rainfall estimation (Doumounia et al., 2014; Chwala et al., 2016). Overeem et al. (2016b) evaluated 2.5 years of microwave link rainfall estimates for the Netherlands with more than 3000 microwave links (using 15-minute minimum-maximum sampling) against gauge-adjusted radar rainfall data, showing a relative underestimation of 9% for 15-min interpolated rainfall maps with a 74 km² resolution. However, the interpolated hourly rainfall map using CMLs outperformed automatic rain gauges compared with gauge adjusted radar data. Similarly, Chwala et al. (2012) used 1-minute averaged RSL data recorded with the data loggers for five microwave links, showing a good correlation between link and radar-derived rainfall.

Some of the Mobile Network Operators (MNO) also provide instantaneous RSL (periodic snapshots) data over the 15-minutes: de Vos et al. (2019) compared the performance of instantaneous versus minimum and maximum RSL data for The Netherlands. Even though this comparison was based on data from two different periods, each having a different network, the use of minimum and maximum sampled data outperformed the instantaneous 15-minute data. Similarly, average sampling of the received signal level over the 15-minute interval is also common for telecommunication operators in some parts of the world, but this has not been evaluated against the widely used minimum and maximum RSL strategy. Accordingly, this study tests this alternative strategy while demonstrating for the first time the capability of rainfall retrieval using CML signal extinction data in the Australian continent.

To date, there has not been a study evaluating the errors introduced by the minimum-maximum sampling as opposed to average sampling. This study explores the capability of rainfall retrieval using CML signal extinction data for the Greater Metropolitan area of Melbourne, the second-largest city in Australia, with a population of 4.48 million (ABS, 2016). This study compares the performance of rainfall retrieval using two commonly sampled datasets for the same period, where data based on minimum-maximum and average sampling from the same link paths are compared.

5.2 Study area and data used

5.2.1 Study area

The study area covers the greater Melbourne region in the Australian state of Victoria. This region has a temperate oceanic climate (Cfb, Köppen-Geiger classification), with an annual average rainfall (based on 29 years of rainfall data from 1990 until 2018 for 73 stations) varying from 500 mm in the west of Melbourne to 1400 mm in the Dandenong ranges towards the eastern part of the city, with a standard deviation of 175 mm. Most of the rainfall occurs during the southern hemisphere winter (June, July and August) and spring (September, October and November). On average, there are 110 days each year with at least 1 mm of rainfall. The average temperature (based on the same period with 21 stations) of the study area varies between 18°C to 24°C for summer and varies between 6°C to 12°C for the winter season. The elevation of the study area ranges from sea level to 1803 m.

5.2.2 Data used

This study was based on the data collected from one of the telecommunication operators. The detail of the dataset is explained in Chapter 3. A total of 135 microwave links were used, covering approximately 2 years of data. These CML data were sampled with a frequency of 10 Hz and stored as the minimum, maximum and average over 15-min intervals with constant transmitted power.

5.3 Methodology

5.3.1 Preliminary data processing and quality check

The CML dataset was delivered by the operator in two separate files: one with 15-min RSL data for all the links stored daily and the other with the corresponding metadata. These files were received on a monthly basis at the end of each month. Metadata included the location of transmitter and receiver nodes, the elevation of the antennas, the assigned microwave frequencies (including minimum and maximum frequency range), the polarization of the signal, path lengths and the IP addresses of each transmitter and receiver. Using the IP address as a unique identifier in the CML data and metadata, merged RSL data with necessary fields (frequency, latitude, longitude, and polarization) were prepared for further processing.

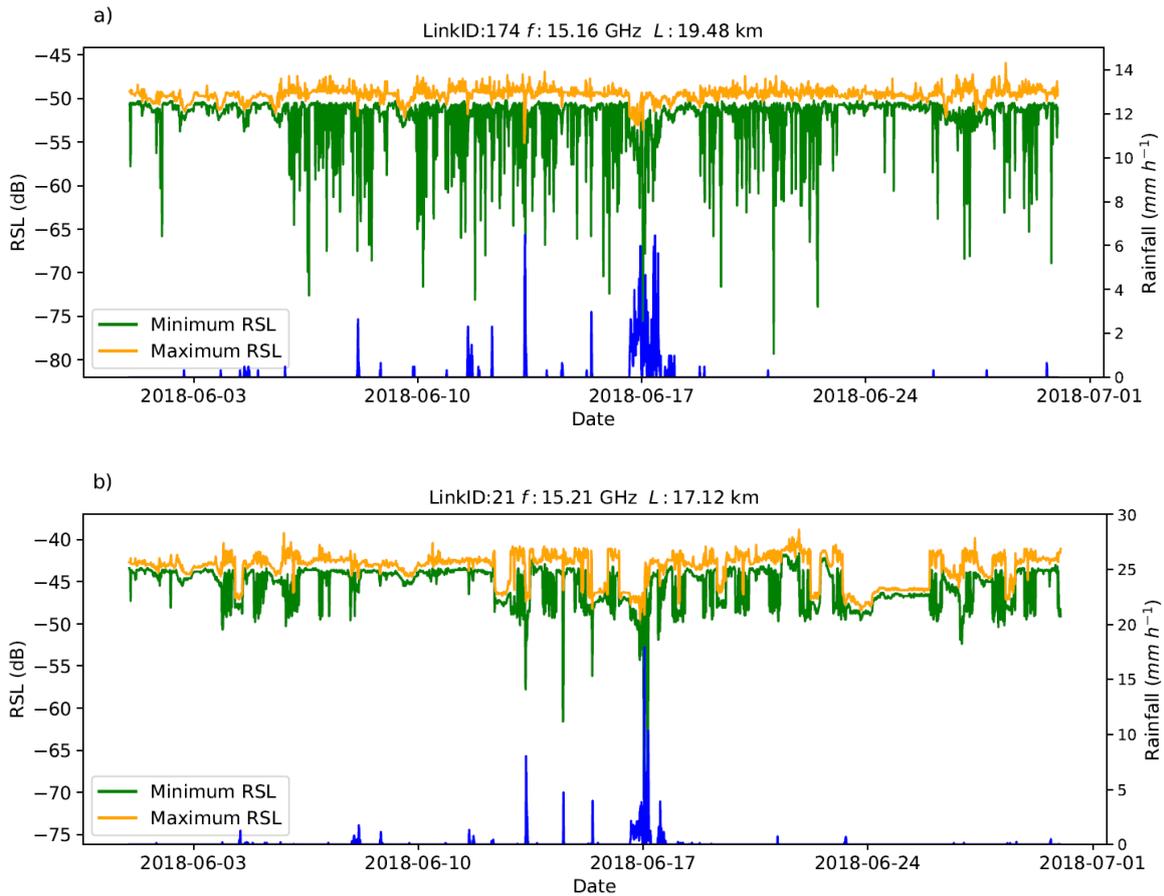


Figure 5-1: Example of some suspicious microwave link data: (a) There is noise observed in the minimum (green) and maximum (yellow) RSL; (b) There is a sudden drop in the received power level even if the transmitted power remains constant. Blue lines indicate observed rainfall rates from *Rainfields*.

Among all 144 links, there were nine links that showed some suspicious behaviour in the data. Three distinct behaviours were identified in these links: (a) presence of noise in the dataset; (b) a sudden drop in the signal level during dry periods; and (c) a gradual increase/decrease in the signal level. Examples are shown in Figure 5-1. These suspicious links were excluded from further analysis.

5.3.2 Use of RAINLINK

After pre-processing the data, the freely available RAINLINK package developed by Overeem et al. (2016a) was used for retrieving rainfall rates. Originally, RAINLINK was designed for handling minimum and maximum RSL data with a constant transmitting power. de Vos et al. (2019) describe the pre-processing necessary to handle instantaneous received signal level data, where the transmitted power is allowed to vary. Studies employing average

RSL data have not been published to date. This study is, therefore the first to use average RSL data in RAINLINK for retrieving rainfall. Here average and minimum/maximum RSL data were processed separately. Further details of the RAINLINK package can be found in Overeem et al. (2016a).

The RAINLINK package includes the following processing steps:

- 1) **Pre-processing of the data.** Duplicated link identifiers, identifiers with inconsistent metadata, and links with frequencies outside the range 10–40 GHz are excluded from the analysis.
- 2) **Dry/wet classification.** Rainy periods are identified based on spatial correlation. When at least half the nearby links (default radius of 15 km) experienced a drop in the minimum or average signal level, the time interval was considered as “wet”. This radius was increased to 20 km (in contrast to the default value of 15 km) based on the spatial distribution of the microwave links in the Melbourne metropolitan area. This drop in the signal level was calculated based on the difference between the RSL data as compared with the maximum value of the link over the previous 24-hour period, both as a difference and a difference divided by the path length. If the median of all nearby links was less than the predefined threshold median $Q_{mP}(\text{dB})$ for the difference and $Q_{mPL}(\text{dB km}^{-1})$ for difference per kilometre link path, the link was labelled as “wet” for that interval.
- 3) **Reference signal level.** Based on the moving median of the signal level during the previous 24-hour dry period, the reference signal level (P_{ref}) was determined. The difference between this reference level signal and the actual signal level provided the attenuation due to rainfall.
- 4) **Outlier removal.** Based on a filter that relies on the principle of the rainfall distribution over space (Overeem et al., 2016a), outliers were removed. This filter discards the time interval of a link for which the cumulative difference between its specific attenuation and that of the surrounding links (default radius of 15 km, although for this study increased to 20 km) over the 24 hour becomes lower than the outlier filter threshold value.
- 5) **Attenuation level.** The corrected minimum ($P_{Cor/min}$), maximum ($P_{Cor/max}$) and average ($P_{Cor/avg}$) attenuation level was calculated for each time interval as:

$$P_{\text{Cor}/\text{min}} = \begin{cases} P_{\text{min}} & \text{if wet AND } P_{\text{min}} < P_{\text{ref}} \\ P_{\text{ref}} & \text{otherwise} \end{cases} \quad (5.1)$$

$$P_{\text{Cor}/\text{max}} = \begin{cases} P_{\text{max}} & \text{if } P_{\text{Cor}/\text{min}} < P_{\text{ref}} \text{ AND } P_{\text{max}} < P_{\text{ref}} \\ P_{\text{ref}} & \text{otherwise} \end{cases} \quad (5.2)$$

$$P_{\text{Cor}/\text{avg}} = \begin{cases} P_{\text{avg}} & \text{if wet AND } P_{\text{avg}} < P_{\text{ref}} \\ P_{\text{ref}} & \text{otherwise} \end{cases}, \quad (5.3)$$

where P_{min} , P_{max} and P_{avg} are the raw minimum, maximum and average attenuation, respectively.

- 6) **Wet antenna attenuation.** A constant wet antenna attenuation (A_a) was deducted from the corrected total attenuation. This attenuation was later divided by the path length to obtain the specific attenuation (k).
- 7) **Rainfall estimation.** The rainfall rate R was calculated from k using the power law equation proposed by Olsen et al. (1978):

$$R = ak^b, \quad (5.4)$$

where the values of the parameters a and b in Eq. 4 were derived for Melbourne using data obtained from an OTT PARSIVEL¹ optical disdrometer and shown in Figure 5-2. Extinction cross sections were estimated based on the T-matrix method developed by Mishchenko and Travis (1994) using a python interface developed by Leinonen (2014) being the most comprehensive and computationally efficient method for the calculation of electromagnetic scattering of particles of arbitrary shape. This was later used to derive the specific attenuation, which were related to rainfall rates using the power-law model based on the least-squares method. For the case of using MinMax, the weighing factor α was used to calculate the average rainfall from R_{min} and R_{max} .

$$R = \alpha \cdot R_{\text{max}} + (1 - \alpha) \cdot R_{\text{min}} \quad (5.5)$$

However, for the case of using average sampling, this α parameter was obviously not required.

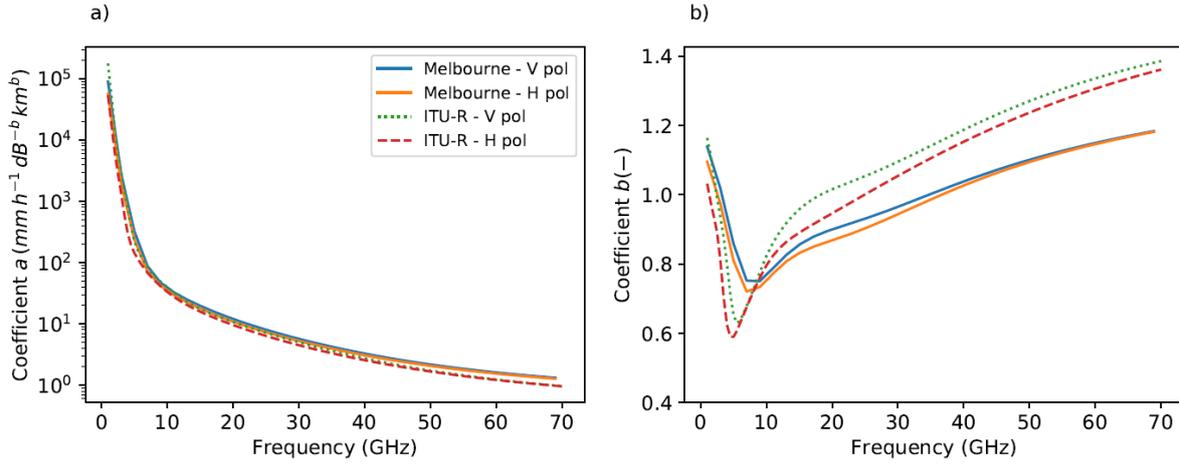


Figure 5-2: (a) Coefficients a and (b) exponent b of the power-law relation between R and k for both horizontally and vertically polarized signals for frequencies ranging from 1 to 70 GHz. The values recommended by the International Telecommunication Union, Radio communication (ITU-R, 2005) for computing specific attenuation for given rain rates for world wide application are shown in dotted and dashed lines.

5.3.3. Calibration of the RAINLINK parameters

RAINLINK has 14 parameters for rainfall retrieval, including those related to wet-dry classification, reference signal determination, outlier filter, wet antenna attenuation and rainfall retrieval using the power law. The optimal values of these parameters are likely to differ between different climatic conditions and microwave link networks (*e.g.* regarding sampling strategy, spatial link density, and resolution of the RSL data); thus it is recommended to calibrate the most important parameters on a subset of the data (de Vos et al., 2019). For this study, based on a sensitivity analysis using one month of a dataset, the three parameters Q_{mP} , Q_{mPL} and A_a for both the Average and MinMax dataset were identified as most important for the overall rainfall retrieval. Additionally, α parameter was also identified as the most important for the MinMax RSL. For the sensitivity analysis, the cost function proposed by de Vos et al. (2019) was used, which includes the POD, FAR, CV, percentage bias and correlation coefficient. Besides these, two additional parameters a and b are required; based here on local drop size distribution data from Guyot et al. (2020). More information on these parameters is provided in Table 5-1.

Table 5-1: List of variables used for calibration of RAINLINK. Here, the alpha coefficient provides the weightage between the minimum and maximum rainfall. The parameters a and b used in this relationship are based on a local disdrometer data.

Variable description	Symbol	Unit
Threshold median	Q_{mP}	dB
Threshold median per unit length	Q_{mPL}	dB km ⁻¹
Wet antenna attenuation	A_a	dB
Alpha coefficient	α	-
Prefactor of rainfall-attenuation relationship	a	mm h ⁻¹ dB ^{-b} km ^b
Exponent of rainfall-attenuation relationship	b	-

A subset of 30 rainy days spread over the 2-year period was selected for the calibration of the RAINLINK parameters. This subset was selected such that there is high availability of RSL data from microwave links during the selected periods, also ensuring that data from different seasons were included. This subset represents a total rainfall amount of 390 mm, ranging from 3 mm to 33 mm per day. The details of the calibration dataset are shown in Table A1 in the Appendix.

a) Parameters for the average RSL

The optimised values of the three most important parameters Q_{mP} , Q_{mPL} and A_a for the average RSL data have been identified using the optimization procedure in de Vos et al. (2019). The detail of the calibration procedure is described later in this section. The radius for nearby links was increased from 15 km to 20 km, as the density of the links was lower for the present case, as compared to The Netherlands. Other parameters besides these were kept at their default values in the RAINLINK package. Based on the calibration dataset, hourly rainfall estimates were calculated for the various combinations of values of Q_{mP} , Q_{mPL} and A_a . Here, hourly rainfall estimates were considered to minimize the sampling error caused by the measurement lag in the radar due to its providing measurements aloft that usually take a couple of minutes for the rainfall to reach the earth's surface. Accordingly, Q_{mP} was varied from -2.5 to -0.1 dB, Q_{mPL} was varied from -2.0 to -0.1 dB km⁻¹ and A_a was

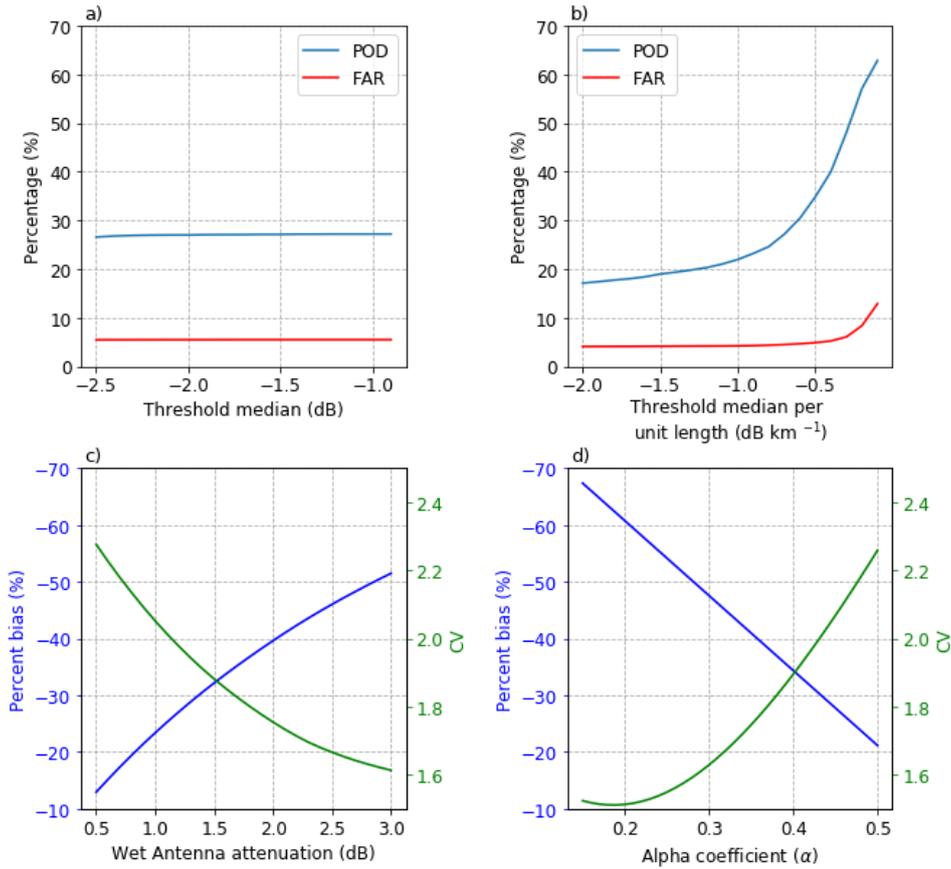


Figure 5-3: Sensitivity analysis using MinMax RSL data for: (a) Threshold median; (b) Threshold median per unit length; (c) Wet antenna attenuation; and (d) Alpha coefficient.

varied from 0.5 to 3 dB (steps of 0.1 for all parameters). The sensitivity of each of these parameters based on the MinMax data is shown in Figure 5-3. The results obtained for each of the combinations were evaluated against the gauge-adjusted radar product, also accumulated to path-averaged hourly values. The path-averages were calculated based on the weights of the intersecting CML paths for each radar pixel.

b) Parameters for minimum/ maximum RSL

For the minimum and maximum RSL data, an additional parameter α is required in the optimization process of the rainfall retrieval (Overeem et al., 2016a). The value of α was varied between 0.10 to 0.50 (with steps of 0.01). Other three parameters, including the wet antenna attenuation A_a were obtained in a similar manner as for the average RSL.

Table 5-2: Confusion matrix based on gauge adjusted radar and CML

		<i>Gauge adjusted radar (Rainfields).</i>	
		Wet (0)	Dry (1)
<i>CML</i>	Wet (0)	True Positive (TP)	False Positive (FP)
	Dry (1)	False Negative (FN)	True Negative (TN)

TP: True positive (both R_{link} and R_{radar} detect rainfall),

TN: True negative (both R_{link} and R_{radar} show no rainfall),

FP: False positive (R_{link} detects rainfall but R_{radar} shows no rainfall),

FN: False negative (R_{link} shows no rainfall but R_{radar} detects rainfall).

5.3.4 Performance metrics

The overall performance of the path-average rainfall retrieval was assessed based on a series of evaluation criteria, covering the two main steps in the rainfall retrieval: (a) Wet-dry classification and; (b) Rainfall retrieval.

(a) Wet-dry classification

This provides a measure of how well the link observations correctly estimate the occurrence of rainfall. The following criteria (de Vos et al., 2019; Graf et al., 2020) were used to assess the performance of the classification based on the confusion matrix, as shown in Table 5-2.

1) The probability of detection (POD) provides a measure of the proportion of actual wet periods that are identified by both the CML and the radar. In this case, POD is defined as the percentage of wet periods identified using the nearby link approach when both R_{link} and R_{radar} detect rainfall. The POD is given as:

$$POD = \frac{TP}{TP + FN} \times 100 \% . \quad (5.6)$$

The POD value ranges from 0 to 100%, with 100% being a perfect score and 0% being the worst.

- 2) **The false alarm ratio (FAR)** provides a measure of the proportion of the actual wet periods that are incorrectly identified. In this case the R_{link} detects rainfall but R_{radar} shows no rainfall. The FAR is given as:

$$FAR = \frac{FP}{FP + TP} \times 100 \% . \quad (5.7)$$

Similarly, the FAR value also ranges from 0 to 100 %, but with 100% being the worst score and 0% being the best score.

- 3) **The Matthews correlation coefficient (MCC)** provides a measure of the quality of the binary classification (wet-dry in the case of classification for CMLs) (Matthews (1975)). This is considered as one of the best ways to report the result of the confusion matrix:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} . \quad (5.8)$$

The MCC value ranges from 0 to 1, with 1 being the best score and 0 being the worst.

- 4) **The Error rate (ERR) or misclassification rate** provides the performance measure of binary classification based on the miss-classification from both positive and negative classes and is calculated as:

$$ERR = \frac{(FP + FN)}{(TP + TN + FP + FN)} . \quad (5.9)$$

Similarly, ERR ranges from 0 to 1, with 0 being the best score and 1 being the worst score.

(b) Rainfall retrieval

This set of evaluation parameters provides a measure of how well the CML-derived rainfall relates to the reference rainfall depths (in this case, the gauge-adjusted radar).

- 1) The **Pearson correlation coefficient (ρ)** provides the correlation between rainfall depths measured by the link R_{link} and the gauge-adjusted radar (R_{radar}). It is given as:

$$\rho = \frac{\text{cov}(R_{\text{link}}, R_{\text{radar}})}{\text{std}(R_{\text{link}}) \text{std}(R_{\text{radar}})}, \quad (5.10)$$

where $\text{cov}(x, y)$ is the covariance between x and y , and $\text{std}(x)$ is the standard deviation of x . ρ values range from 0 to 1, with 1 being the best and 0 the worst performance.

- 2) **The coefficient of variation (CV)** provides a measure of the dispersion of data points between the rainfall intensity derived by the link (R_{link}) and the gauge-adjusted radar (R_{radar}). It is given as:

$$\text{CV} = \frac{\text{std}(R_{\text{res}})}{\bar{R}_{\text{radar}}}, \quad (5.11)$$

where $R_{\text{res}} = R_{\text{link}} - R_{\text{radar}}$ and \bar{R}_{radar} is the mean of the gauge adjusted radar data. The smaller the CV the better the performance.

- 3) **The relative bias** provides the average error between the rainfall intensity measured by the link R_{link} and the gauge-adjusted radar (R_{radar}). It is given as:

$$\frac{\bar{R}_{\text{res}}}{\bar{R}_{\text{radar}}} \times 100\%, \quad (5.12)$$

where, \bar{R}_{res} is the mean of the residual. Similarly, values closer to 0 are better, however, positive values indicate overestimation and negative values indicate underestimation compared with the reference.

5.4. Results

5.4.1 Calibration

Table 5-3 shows the calibration results for the four most important RAINLINK parameters. Two parameters, threshold median (Q_{mP}) and threshold median per unit length (Q_{mPL}), are related to the wet-dry classification while the remaining two, wet antenna attenuation (A_a) and Alpha (α), are related to rainfall retrievals. For Average, both Q_{mP} and Q_{mPL} are less negative compared to the default values of RAINLINK, while the A_a value is also lower compared to MinMax, but higher compared to instantaneously sampled data. For MinMax, the threshold median is slightly higher, but the threshold median per unit length is less

Table 5-3: Calibration results for a selection of four of the RAINLINK parameters and comparison with values obtained for The Netherlands.

Dry/ wet classification parameters		Rainfall retrieval parameter	
Threshold Median, Q_{mP} (dB)	Threshold Median L, Q_{mPL} (dB km ⁻¹)	Wet antenna attenuation, A_a (dB)	Alpha, α
<i>Average data</i>			
-0.7	-0.2	1.6	-
<i>MinMax data</i>			
-1.50 (-1.4)	-0.40 (-0.7)	1.4 (2.30)	0.29 (0.33)
<i>Instantaneous RSL data (de Vos et al., 2019)</i>			
-0.6	-0.4	1.4	-

negative than the default value. This indicates that the time interval is more likely to be classified as wet, with the threshold median per unit length being closer to 0 than the default value, which is also the case for the threshold median for Average.

Table 5-4 shows the performance for the calibration dataset for both the Average and MinMax data presented at an hourly timescale. For dry/wet classification, MinMax had a better performance when compared to Average data, although its POD value was lower. For the rainfall retrieval, both datasets showed similar performance for ρ . MinMax had a small negative bias compared to the almost unbiased Average data, and had a much lower value for CV. Hence, MinMax resulted in the best overall performance.

5.4.2 Validation result

After processing the dataset (excluding the data used for calibration), the number of data points containing results for the two sampling strategies (MinMax and Average) were of different sizes, because the outlier filter used in processing the raw data removed different time intervals for specific links from MinMax and Average data. Thus, in order to make a fair comparison, the time intervals with available data for both strategies were retained. Also, to note that all the performance evaluation was based on the path-average rainfall depths against the reference.

Table 5-4: Performance criteria for the calibration period for Average and MinMax.

Dataset	Dry/ Wet classification			Rainfall retrieval		
	POD	FAR	MCC	Relative Bias (%)	CV	ρ
Average	68.25	10.25	0.47	0.20	1.41	0.67
MinMax	64.62	7.38	0.55	-0.68	1.12	0.68

a) Performance of wet-dry classification

Figure 5-4 shows the time series of RSL with corresponding rainfall intensities for a selected event, together with the wet-dry classification for one of the microwave links. For both the Average and MinMax data, most of the wet periods with higher rainfall intensities were identified correctly, as shown by the true positives. However, some time-steps with low rainfall intensities during a wet period were classified as dry, *e.g.* false negatives. There were instances where the time-steps are incorrectly classified as wet even though there was no rain observed on the ground indicated by the false alarm. These false alarms were observed during the time-steps where the RSL dropped below the baseline signal level due to reasons other than rain. For this event, Average data correctly identified wet periods with a POD of 90 % and a slightly lower POD of 84 % for the MinMax data. Also, for the Average data 18 % and for the MinMax only 9 % of the dry periods were incorrectly classified as wet. Similarly, 10 % and 16 % of wet periods were missed by the Average and MinMax, respectively.

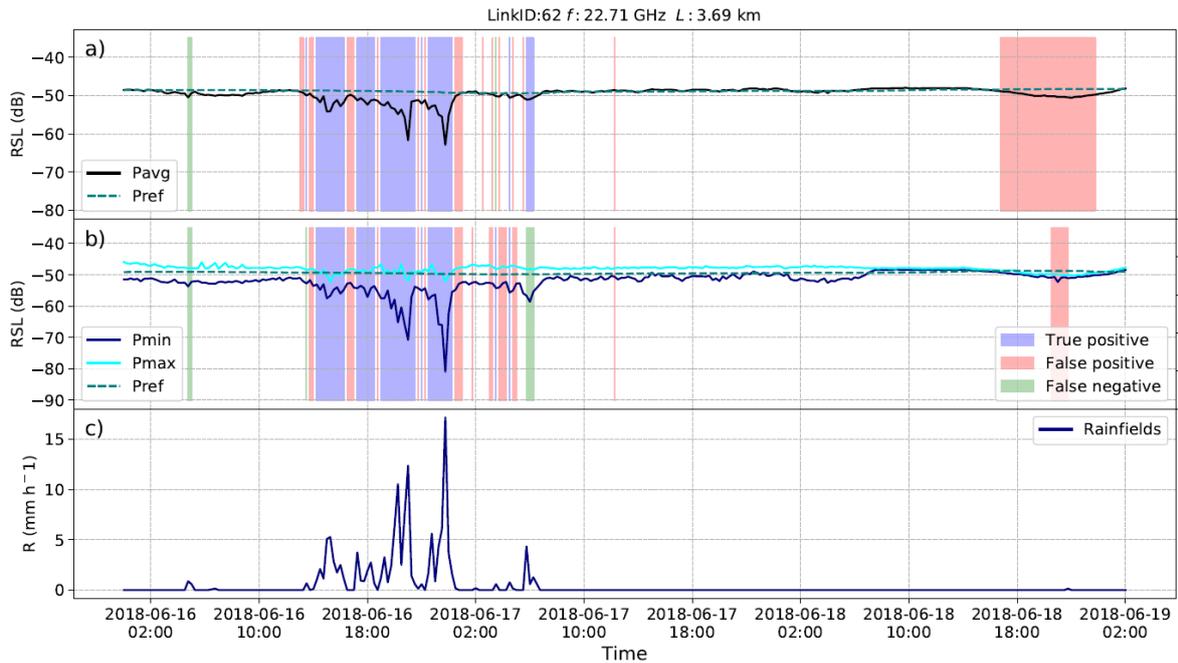


Figure 5-4: Time-series of (a) average RSL; (b) minimum and maximum RSL; and (c) gauge adjusted radar rainfall intensities for a selected rainfall event for LinkID 62. The wet-dry classification using the calibrated parameters is shown as shaded colours.

Table 5-5: Performance of wet-dry classification for 15-min average and 15-min minimum/maximum RSL data.

Data set	Threshold for time interval to be wet								
	0 mm h ⁻¹			0.1 mm h ⁻¹			0.5 mm h ⁻¹		
	POD	FAR	MCC	POD	FAR	MCC	POD	FAR	MCC
Average	64.34	8.76	0.38	68.05	8.94	0.38	77.91	9.57	0.35
MinMax	53.58	3.70	0.42	57.34	3.84	0.46	68.41	4.35	0.45

Table 5-5 shows the performance summary of the wet-dry classification for three threshold values used to distinguish each of the time intervals between wet and dry periods based on the gauge-adjusted radar data for all the dataset. Average data showed a higher probability of detection (POD) value of 64%, whereas MinMax showed a slightly lower POD of 54%. Also, 9% and 4% of the dry periods were incorrectly classified as wet based on the Average and MinMax data, respectively. Similarly, 36% and 46% of wet periods were missed by Average

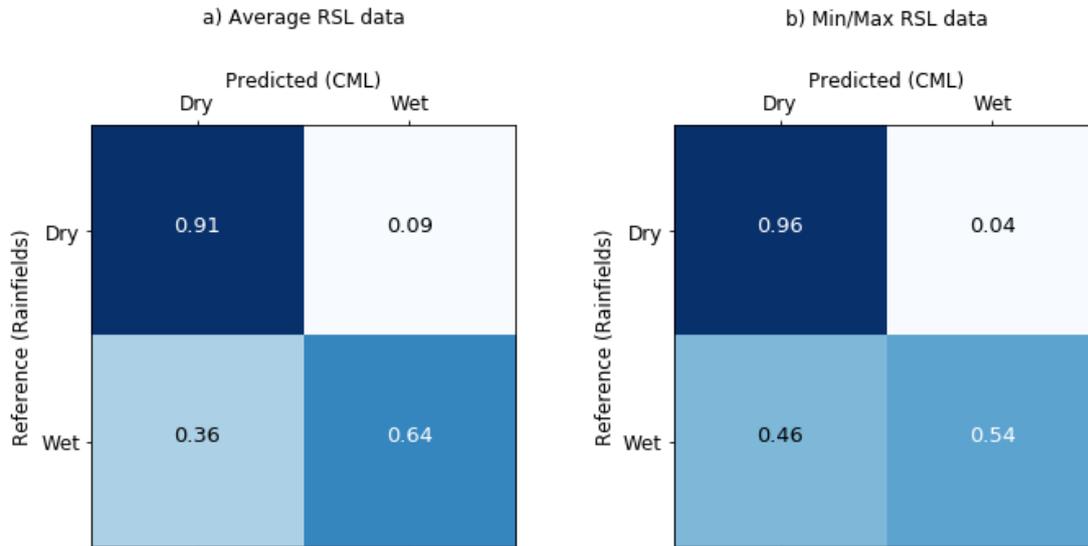


Figure 5-5: Normalized confusion matrix for wet-dry classification for (a) Average data; and (b) MinMax data

and MinMax data, respectively (see Figure 5-5). For all threshold, Average data showed a higher probability of detection (POD) when compared with the MinMax data. However, at the same time, it also showed a much higher false alarm ratio (FAR), meaning there were a larger number of time intervals that were misclassified as wet as compared with MinMax. Based on the MCC values, for all the threshold values, MinMax data outperformed the Average data.

As a further investigation of the performance for wet-dry classification, Figure 5-6 shows the box plot for four different statistics (POD, FAR, MCC and ERR with the wet-dry threshold of 0.1 mm h^{-1}). Considering all four statistics, three of the values showed better performance of MinMax compared with Average data. Although the POD value was higher for Average, there were a higher number of both positive and negative misclassifications, which is reflected in the higher value of ERR. The wider range of all four statistics suggests that some of the links were performing poorly. Graf et al. (2020) obtained similar statistics in their study for Germany, where they found the median MCC value of 0.47, which is slightly higher than that obtained for MinMax data here. In their case, they used data with a temporal resolution of 1 min for the wet-dry classification.

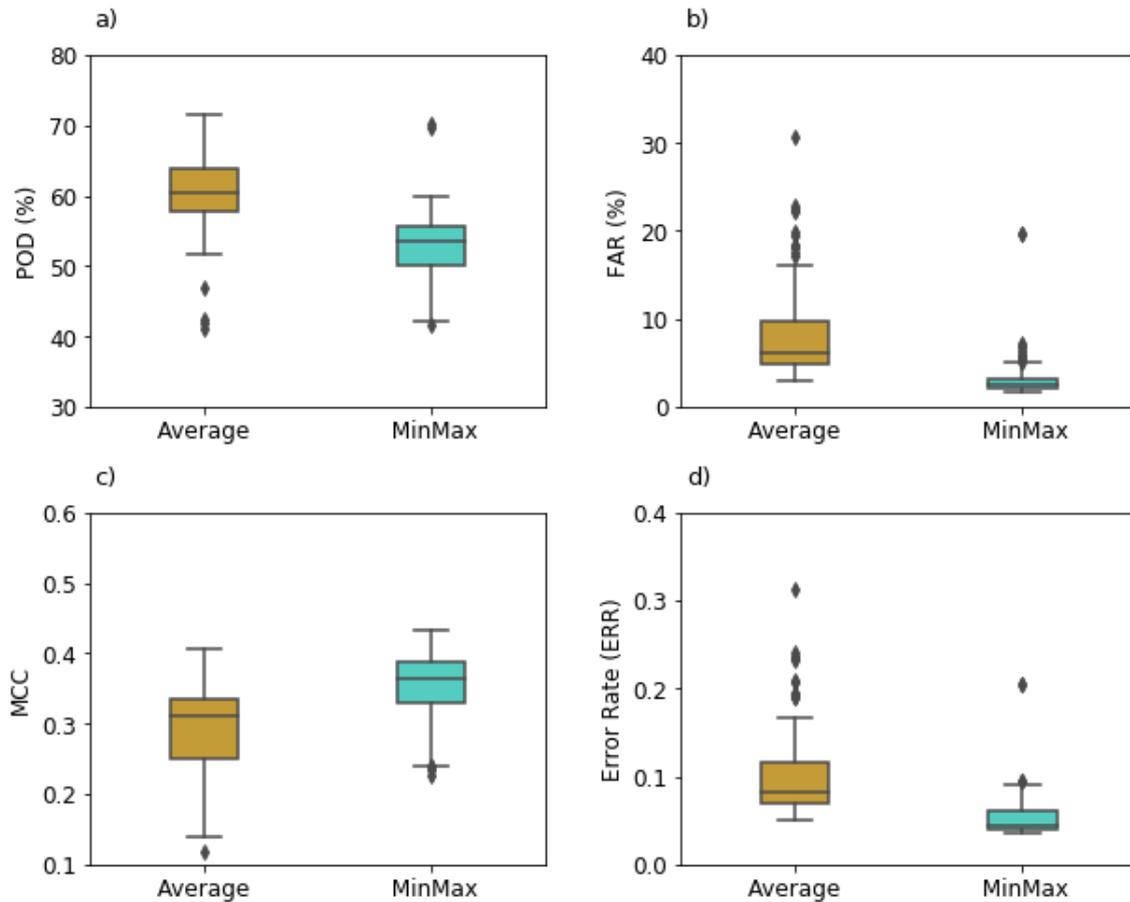


Figure 5-6: Box plots showing the performance criteria for the wet-dry classification using the nearby links methods with: (a) Probability of detection (POD); (b) False alarm ratio (FAR); (c) Matthew correlation coefficient (MCC); and (d) Error rate (ERR) for Average and MinMax data.

In order to further investigate the performance of the wet-dry classification for different rain rates, the POD was calculated for each of the different classes of data exceeding certain threshold rain rates from the reference data as shown in Figure 5-7. The probability of detecting rain increases with higher rain rates, reaching more than 90 % when 5 mm h^{-1} was used as a threshold rain rate rather than a lower rainfall rate. This suggests that the correct detection rate of wet periods is better under more intense rainfall conditions.

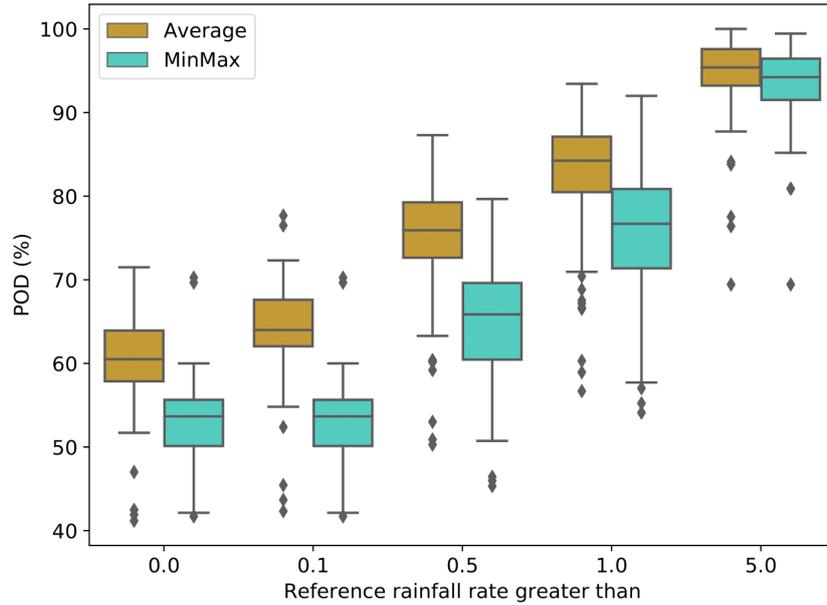


Figure 5-7: Probability of detection (POD) based on a reference rainfall rate. Here five thresholds (0, 0.1, 0.5, 1.0 and 5.0-mm h⁻¹) for the gauge adjusted radar data are used to filter the data; thus this result provides POD result for only wet intervals.

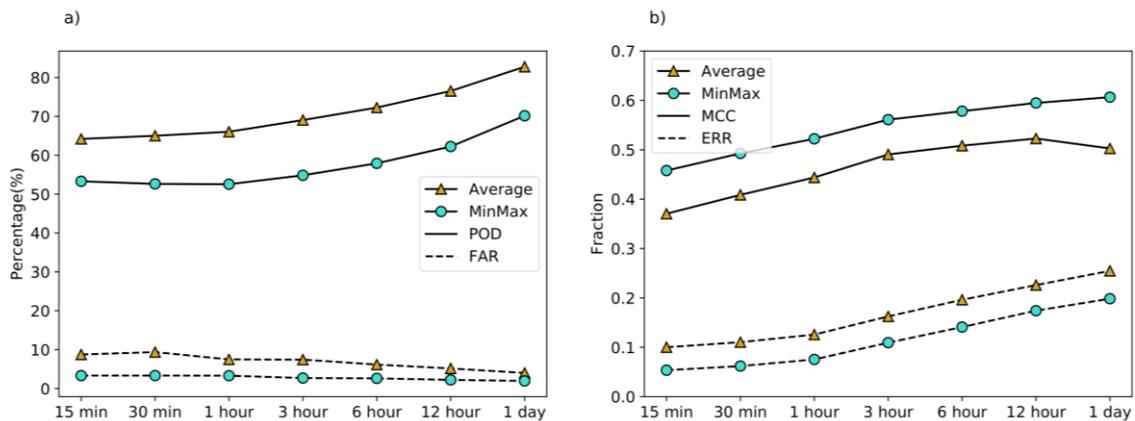


Figure 5-8: Validation criteria showing the ability of CMLs to detect rainy periods against reference data; (a) POD and FAR; and (b) Matthew correlation coefficient (MCC) and Error rate (ERR). All four parameters were calculated for Average and MinMax for various accumulation intervals, using a threshold of 0.1 mm h⁻¹ to detect rain occurrence.

Similarly, Figure 5-8 shows the accuracy of the links in determining rainfall occurrence for various accumulation intervals. Here, a threshold of 0.1 mm h⁻¹ was used to distinguish wet

from dry periods using the gauge-adjusted radar data. Although the POD increased and FAR decreased as the considered intervals became longer, the error rate simultaneously increased. This is due to a decrease in the relative proportion of true negative (TN) values compared with the lower accumulation interval. The increase in MCC values suggests that the performance increases for longer accumulation intervals for both Average and MinMax data.

b) Performance for the rainfall retrievals

Figure 5-9 shows the comparison of link-derived rainfall with gauge-adjusted radar data for 15-min, 1-hour, 3-hour and 1-day accumulations for both MinMax and Average. The accuracies of the link-derived rainfall increase for longer durations for both sampling strategies. This can be seen by the decrease in the value of CV (for Average RSL 2.84 to 1.49 and for MinMax 2.70 to 1.22), and the increase in the value of the correlation coefficient (for Average 0.29 to 0.75 and MinMax 0.30 to 0.79). A systematic overestimation in link-derived rainfall estimates with respect to gauge-adjusted radar data was found for all accumulation intervals for both sampling strategies. The lower values of RMSE and CV for MinMax, the smaller overestimation and generally similar values for the correlation coefficient indicate that it outperforms Average.

In addition, Table 5-6 shows the results for the performance of the rainfall retrieval on a seasonal basis. For both sampling strategies (Average and MinMax), link-derived rainfall is overestimated for all four seasons, with the largest overestimation occurring during the Summer and Autumn. This larger magnitude of overestimation is mainly attributed to precipitation events with a higher intensity of rainfall during these two seasons. The performance in terms of bias and CV was better for winter and spring for both sampling strategies compared with the other two seasons. This result differs from the result presented by Graf et al. (2020) for the winter season in Germany. They obtained the lowest performance during the winter month (with the highest overestimation and higher CV values of 16.04) due to the presence of both mixed and solid precipitation, which is not the case for this study. Further, this overestimation during other seasons was most likely due to dew formation on the antenna covers as solid/melting snow does not occur in the study area. Similarly, for all seasons, link-derived rainfall corresponded well to the validation data for hourly accumulations. The best performance in terms of correlation coefficient and CV was found in the Spring, followed by the Summer. Based on the two sampling strategies, MinMax had a lower overestimation and a better CV value compared with Average data.

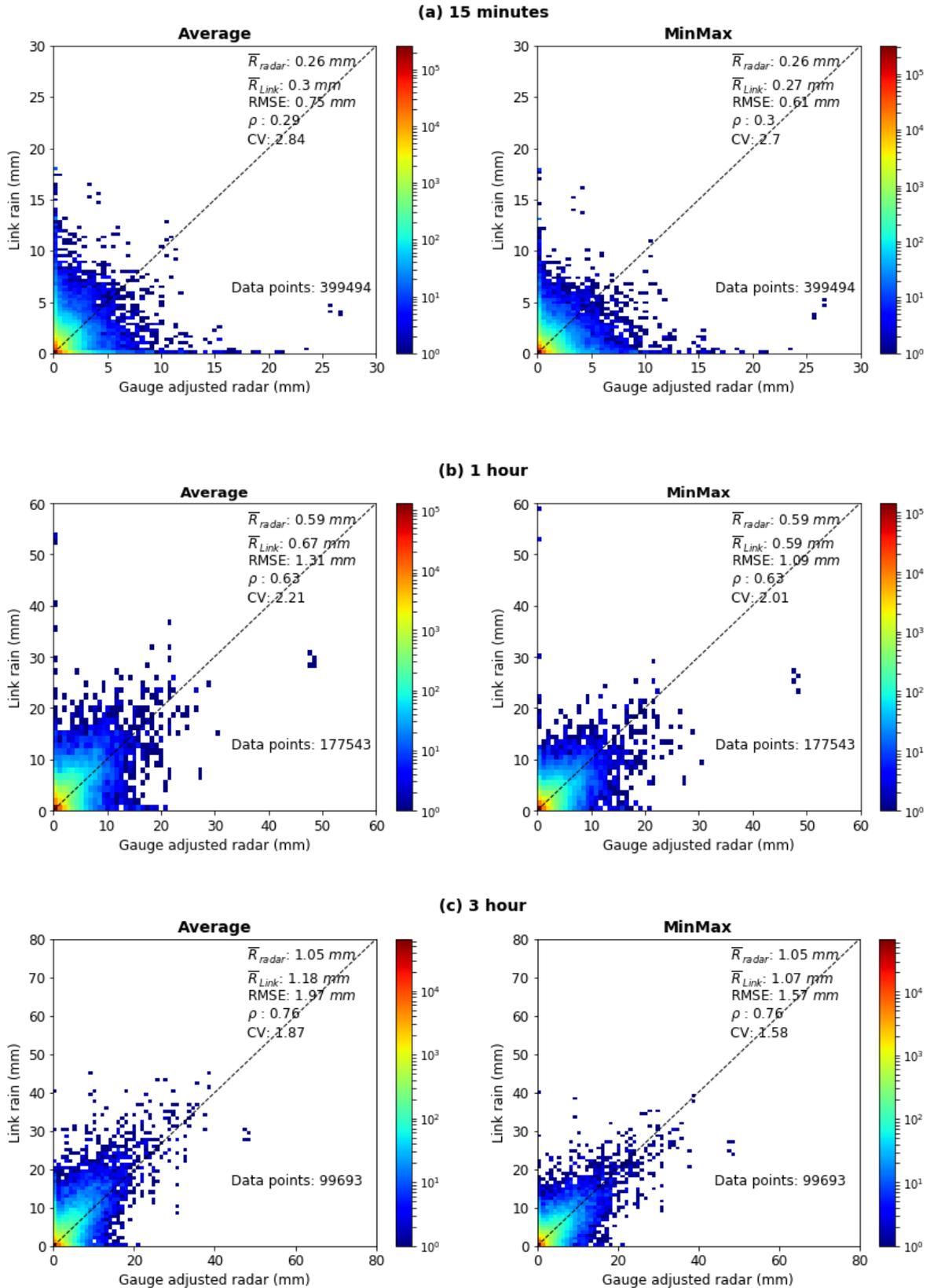


Figure 5-9: Validation of path-average CML-rainfall against gauge-adjusted radar rainfall. Scatter density plot of link-derived rainfall with radar over intervals of (a) 15 minutes; (b) 1 hour; and (c) 3 hour. Only gauge-adjusted radar rainfall depths greater than zero were selected.

Table 5-6: Validation of 15-min and 1-hour accumulation link-derived rainfall against gauge-adjusted radar rainfall (reference) on a seasonal basis.

Dataset	15-minute			1-hour		
	Relative bias (%)	CV	ρ	Relative bias (%)	CV	ρ
<i>Summer (Dec, Jan, Feb)</i>						
Average	15.75	2.87	0.29	15.80	1.92	0.69
MinMax	9.05	2.85	0.32	9.05	1.90	0.69
<i>Autumn (Mar, Apr, May)</i>						
Average	12.73	2.81	0.30	12.71	2.60	0.52
MinMax	13.68	2.65	0.29	13.70	2.32	0.51
<i>Winter (Jun, Jul, Aug)</i>						
Average	9.56	2.52	0.22	8.78	2.21	0.53
MinMax	3.60	1.38	0.23	3.62	1.90	0.56
<i>Spring (Sep, Oct, Nov)</i>						
Average	7.92	2.71	0.32	8.12	1.90	0.71
MinMax	7.26	2.58	0.33	7.51	1.72	0.72

To have a better understanding of the performance of CML rainfall retrievals, event-based results were also analysed for CML-derived and gauge-adjusted radar rainfall data as shown in Figure 5-10. Here, a rainfall event was defined as a rain period separated by a 1-hour or longer rain-free period and having each 15-min time interval with a minimum rainfall rate of 0.1 mm h^{-1} . There were altogether 342 such rainfall events with periods lasting from 45 minutes to 29.25 hours. Compared to the results presented in Figure 5-9, correlation coefficients were significantly higher, with values of 0.86 for both Average and MinMax data. In terms of relative bias, the Average was closer to the reference gauge-adjusted radar dataset compared with MinMax. However, other statistics (RMSE and CV) were lower for MinMax, showing better performance.

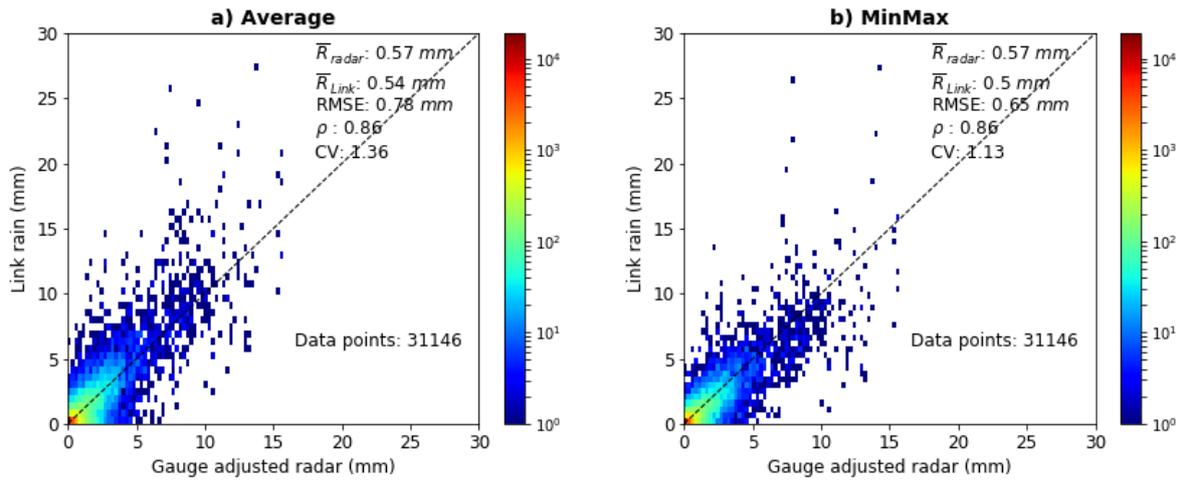


Figure 5-10: Scatter density plots of path-average CML-rainfall against gauge-adjusted radar rainfall for 342 rainfall events.

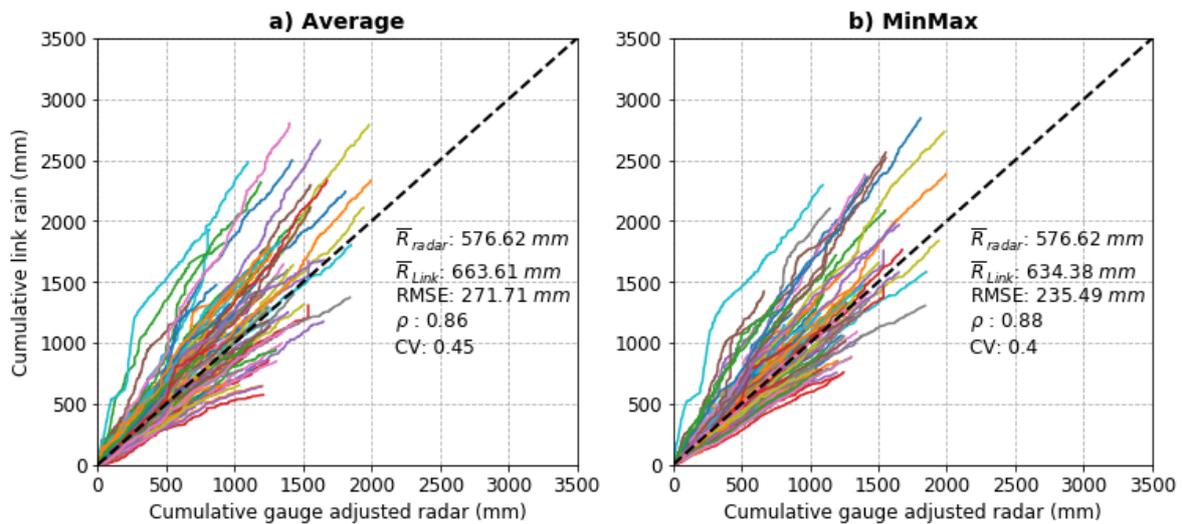


Figure 5-11: Double mass curves per event for all links: (a) Link-derived cumulative rainfall using average data versus cumulative gauge-adjusted radar rainfall; and (b) Link-derived cumulative rainfall using min/max RSL data versus cumulative gauge-adjusted rainfall data.

In order to further investigate the continuous performance of link-derived rainfall estimation, double mass curves between link-derived and gauge-adjusted radar rainfall are shown as accumulation plots in Figure 5-11. Intervals, where either link or radar had missing data were excluded. Most of the links rainfall were passing through the 1:1 black line indicates a good agreement between the link-derived rainfall and the gauge-adjusted radar reference data. In the overall comparison, both Average and MinMax showed a positive bias of about 15 %.

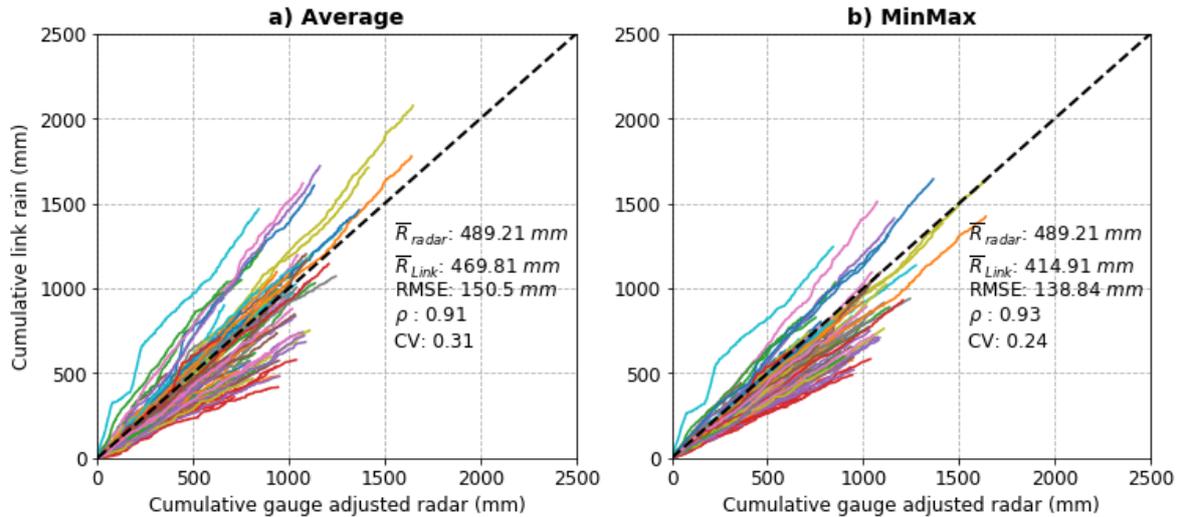


Figure 5-12: Density plot of the double mass curve for all links during only the wet period (no false alarm included) for: (a) Average, and (b) MinMax data.

However, 37 links for Average and 40 links for MinMax showed a mean negative bias of 19.30% and 17.34%, respectively. There were almost similar numbers of links (11 links for Average and 10 links for MinMax) for both RSL data types showing overestimation above 50% and up to 135%.

The results presented in Figure 5-11 for the rainfall retrieval include both timesteps with the false-positives and false-negatives. As there was a significant difference in the false alarm ratio between the two datasets (Average and MinMax), to further investigate the performance for only wet periods, the density plot of double mass curves for all links are plotted in Figure 5-12 by including only time intervals which have non-zero rainfall for both the links and the radar. In terms of bias, Average data showed better performance with a negative bias of 3.96% compared with MinMax data showing a stronger negative bias of 15.18%. For a higher number of links (35 links), there was an overestimation based on Average data compared with only 18 links for MinMax data. However, other statistics (RMSE, ρ , CV) suggest that MinMax performed slightly better compared with the Average RSL.

Figure 5-13 provides more insights regarding the performance of rainfall retrieval by the two sampling strategies. The average link-derived rainfall using the MinMax data was lower than the Average RSL (Figure 5-14 (a)) for time intervals where it was wet according to both sampling datasets (not involving the reference). Furthermore, almost all higher rainfall depths were lower for the MinMax data. In addition, Figure 5-14(b) shows the density scatter plot

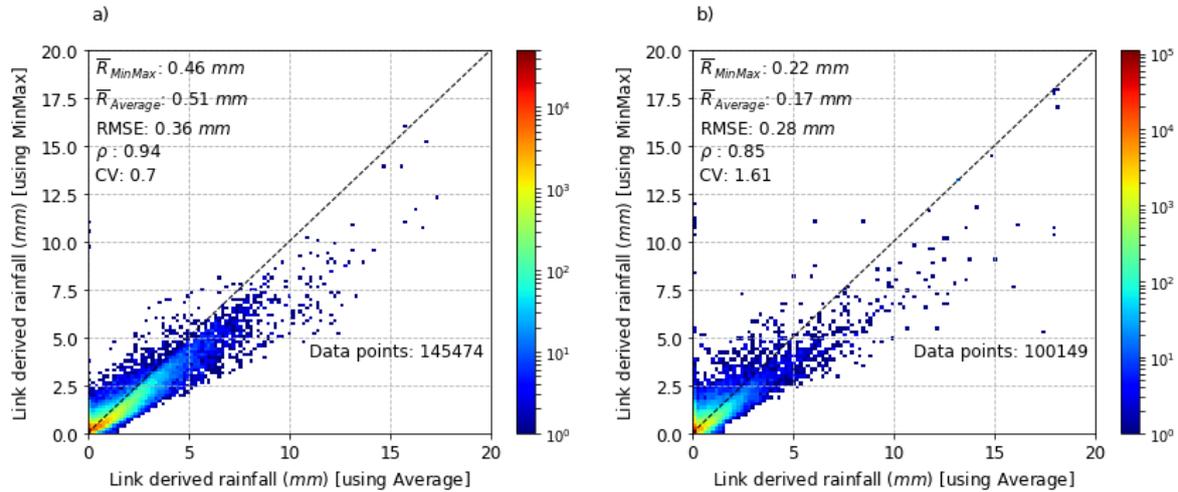


Figure 5-13: Comparison of link-derived rainfall retrievals using 15-min Average vs MinMax RSL data for: (a) wet periods only; and (b) all time intervals with false alarm only.

of link-derived rainfall during all the periods when there was a false alarm for both sampling strategies when compared to the reference. This result suggests that the mean rain rate obtained from MinMax data was higher compared with the Average data even though the higher rainfall depths were lower. So, on average, overestimations during false alarms were higher for MinMax but higher for Average in the case of larger rainfall depths.

In order to explore the spatial variability of the performance of the rainfall retrieval based on the two-sampling data, the correlation coefficient and RMSE based on an hourly accumulation over the study area are shown in Figures 5-15 and 5-16. The correlation coefficients were found to be homogeneously distributed over the area for both Average and MinMax, with some of the outliers in the south-west and south of the Melbourne CBD, where the values were the lowest, and for some individual links with the lower frequencies. The RMSE distribution over the study area was homogenous over the Melbourne CBD and closer to the radar, with a couple of links over the study area showing a higher RMSE for both sampling datasets.

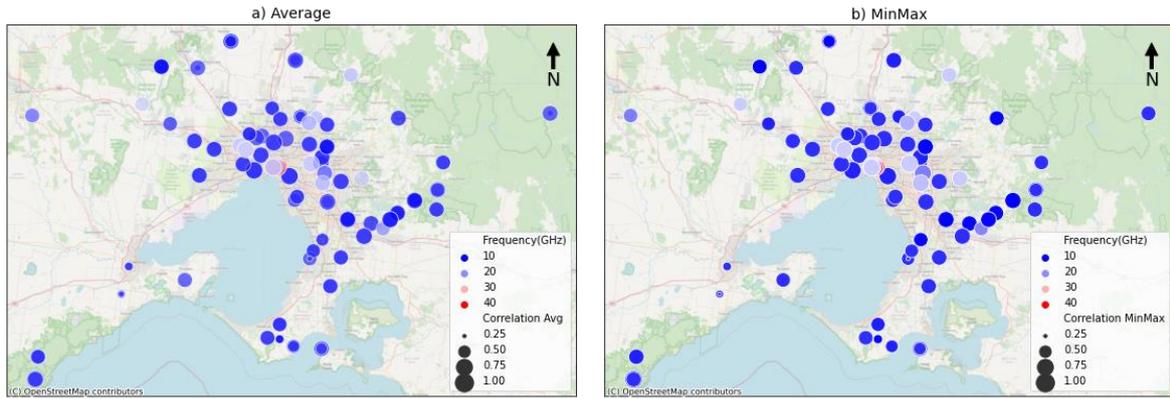


Figure 5-14: Spatial distribution of the correlation coefficient (circle size) between the reference and CML-derived hourly rainfall accumulation at different frequencies (circle colour) for (a) Average; and (b) MinMax data.

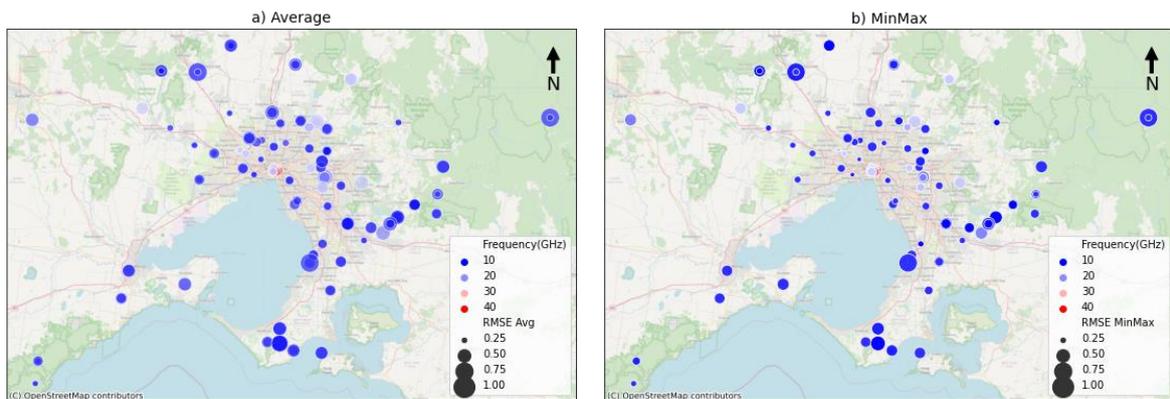


Figure 5-15: As for Figure 5-14 but for room mean square error (RMSE).

5.5 Discussion

5.5.1 Optimized parameters for RAINLINK

Among the parameters used in RAINLINK, only the four most sensitive parameters were optimized for this Melbourne dataset, specifically the threshold median (Q_{mP}), the threshold median per unit length (Q_{mPL}), wet antenna attenuation (A_a) for Average RSL data and the alpha (α) for MinMax RSL data. In addition, two parameters, a and b , were obtained based on local disdrometer data from the study of Guyot et al. (2020), and are the most critical for rainfall retrieval of all the parameters. In the absence of these parameters, one needs to use the generalized values from the ITU recommendations (ITU-R, 2016) or, when available, those from other studies from a similar climate. Q_{mP} and Q_{mPL} , which are related to dry/wet

classification, were obtained separately for the Average and MinMax datasets. For Average RSL data, Q_{mP} and Q_{mPL} were found to be -0.7 dB and -0.2 dB km⁻¹ (no reference is available for comparison using the Average RSL data). For the MinMax RSL data, Q_{mP} and Q_{mPL} were found to be -1.50 dB and -0.40 dB km⁻¹, which is less negative compared with a similar dataset for The Netherlands. This means using the new parameter values enabled corresponding time intervals to be classified as wet with a lower deviation in median and median per unit length values. The average POD and FAR for the MinMax reported in this study were in a similar range to the values for RAINLINK reported by de Vos et al. (2019).

The optimized value of the parameter A_a was found to be 1.6 dB and 1.4 dB for the Average and MinMax data, respectively, and was within the range of values suggested by Overeem et al. (2011) (1.2-1.9 dB) for MinMax data. However, for the MinMax data, Overeem et al. (2013) used $A_a = 2.3$ dB, which is higher than the value obtained for the Melbourne dataset. In their case, A_a and α was optimized for the rainfall after determining the parameters for the wet-dry classification separately and is therefore different than the case presented herein as both wet-dry and rainfall were optimized together. Also, the weighting factor ($\alpha = 0.29$) for the minimum and maximum attenuation for the Melbourne dataset was slightly lower than obtained by Overeem et al. (2013).

There are other parameters in RAINLINK (like the minimum number of available links, the period over which the reference level has to be determined, the minimum number of hours that should be dry in the preceding period, the outlier filter threshold, and the radius for finding nearby links) which have not been altered from the default values of RAINLINK. These parameters are more likely to remain constant and have less dependency on the dataset and climatology of the study area. However, the radius for finding the nearby links was increased from 15 km to 20 km, as the density of links in the employed dataset for the Melbourne metropolitan area was low compared to other studies in The Netherlands.

5.5.2. Effect of sampling type on overall rainfall retrievals

For the wet-dry classification, the Average RSL showed a higher POD when compared with the MinMax RSL data, but at the same time, the FAR value was also higher for the Average RSL data. This means that the Average RSL missed less rain, but the higher FAR shows that often false alarms are provided during dry periods, whereas the MinMax RSL strategy indeed has a lower POD but also a lower FAR. This is mainly because for each of the time intervals, the MinMax RSL contains additional information of 15-minute data characteristics as opposed to the Average RSL data. A similar result has been reported by de Vos et al. (2019) for the MinMax data from the Netherlands. They obtained a POD of 38% and a FAR of 35% for a threshold of 0 mm, a POD of 50% and a FAR of 40% for a threshold of 0.1 mm. Also, taking into account the MCC and ERR, the MinMax data outperformed the Average RSL strategy for the wet-dry classification.

For the rainfall retrieval, the Average data performed very similarly to the MinMax data. In a few cases, considering relative bias, Average data performed slightly better. However, other statistics, such as the CV and RMSE, are in favour of the MinMax sampling data. Thus, in the overall comparison, the MinMax outperformed the Average data. The main reason behind this is that the MinMax data had better wet-dry classification compared with the Average data. Even though for some cases, results favour Average data, for the overall rainfall retrieval process, both wet-dry classification and rainfall retrieval are necessary steps. There are no other studies to date that compare the performance of Average versus MinMax data, but the MinMax, performance is quite similar to the results of Overeem et al. (2016b).

5.6 Chapter summary

This chapter presented an evaluation of rainfall retrieval based on commercial microwave link data for Melbourne, Australia. An open-source algorithm named RAINLINK was used, and the most important parameters for the rainfall retrieval were calibrated using the local commercial microwave links dataset. The main purpose of this study was to compare the overall performance for rainfall retrieval using the two commonly available data sampling strategies.

Chapter 6

Deep Learning model for improving rainfall estimates

This chapter presents a novel approach of using deep learning to improve the rainfall estimation from commercial microwave links when there is only limited information available from the telecommunication operators. This includes the development a special type of recurrent neural network known as the long short-term memory model. The data collected from one disdrometer was used for the training and evaluation. This chapter is the basis of a published article in Water Resources Research.

6.1 Background

The CML-derived rainfall retrieval accuracy and temporal resolution is dependent on a number of factors, one of which is the way the received signal level (RSL) is stored by the telecommunication companies or operators (Leijnse et al., 2008). Operators use a Network Management System (NMS) to collect and store RSL data in their cellular network for quality control monitoring of their network. In most cases, the minimum and maximum values are stored for a 15-minute window (Messer, 2018; Uijlenhoet et al., 2018). Usually, this 15-minute time window and parameters (minimum, maximum RSL and transmit power) stored by the NMS are hard-coded by the hardware provider, which is usually sufficient for network quality monitoring purposes. However, parameters for the rainfall retrieval using the power relationship, derived from (typically 30-s) drop size distribution (DSD) data and applied to 15-min minimum and maximum or average attenuation could lead to uncertainty. Also, within this 15-minute interval, the exact time when the minimum and maximum RSL reading occurred is unknown (Ostrometzky and Messer, 2014), which could lead to uncertainties in the rainfall retrieval. Several studies tried to make use of such available datasets to retrieve rainfall (Overeem et al., 2011, 2013; Ostrometzky and Messer, 2014; Overeem et al., 2016a; Ostrometzky and Messer, 2017). Using only minimum and maximum RSL values, Overeem et al. (2011) employed an extra coefficient to provide the estimate of time-averaged rain rate over the 15-minute sampling period. This coefficient provides the relative contribution of minimum and maximum attenuation (α and $1 - \alpha$ for minimum and maximum, respectively) to obtain the weighted average attenuation. They suggested α as 0.33 based on a set of data for The Netherlands with the same value used for all time steps. Thus, this α remains

uncertain as the distribution of the rainfall or attenuation is usually not consistent across any 15-minute period. Similarly, Ostrometzky and Messer (2017) also used various extreme value distribution functions for estimating time-average rain rates from minimum and maximum signal levels.

Theoretically, there is no physical limit of a NMS on recording interval, meaning that it is possible to record the average RSL, poll the data at a higher frequency or even record in real-time, except for the limitation of storage capacity or data transfer (Messer, 2018; Uijlenhoet et al., 2018; Chwala and Kunstmann, 2019). However, as these cellular networks have been designed and optimized for providing efficient telecommunication rather than measuring rainfall, the opportunistic use for rainfall monitoring needs to deal with the existing scheme of data sampling and storage. Thus, there is a need for a robust and more accurate methodology of estimating accurate rainfall using only the minimum and maximum RSLs that are currently recorded.

Recent developments in the field of deep learning, especially in the branch of deep neural networks, offer a great opportunity to model various physical processes from data, especially if large quantities of data are available (Reichstein et al., 2019). As rainfall occurrence and RSL are time-dependent phenomena, and as a prediction is required at each time step, recurrent neural networks (RNNs) are well-suited: they can learn time dependency and are applicable to series of varying lengths (Shi et al., 2014). Such networks have been successfully used in other fields of computer science, such as speech recognition, language translation, and video and motion prediction (Graves and Schmidhuber, 2009; Mathieu et al., 2015). Closer to the application presented here, Mishra et al. (2018) have implemented a deep learning framework to distinguish dry and wet periods from communication satellite data to improve rainfall retrievals. Recently, Habi and Messer (2018) and Polz et al. (2019) have also used machine learning techniques for the wet-dry classification of commercial microwave links. Similarly, there have been a few other studies using deep learning for rainfall-runoff modelling (Hu et al., 2018; Kratzert et al., 2018), but according to an extensive search, this is the first work that employs a recurrent neural network for improving the rainfall estimation from commercial microwave link data. The primary objective of this study was to design and apply a deep learning model for improving rainfall estimation using CML data. The specific objectives of this study were: (1) to train and validate a deep learning model using a disdrometer dataset; and (2) to use this model to predict rainfall using limited data (only

minimum and maximum RSL data) from a CML. To achieve these objectives, laser-based disdrometer data were collected for more than one year, and then used for training, validation and testing. Subsequently, the disdrometer trained model was used to retrieve rainfall from a CML situated in the proximity of the disdrometer. A range of deep learning model architectures was evaluated and compared with the existing approach (in this case, the result obtained using the weighted average method). In the overall rainfall retrieval process, all other steps are based on the steps followed by Overeem et al. (2016a), including baseline estimation, wet antenna attenuation and attenuation-rainfall relationship.

6.2 Study area and data

For this study, disdrometer data from an OTT¹ PASIVEL installed at Mount View Reservoir; Glen Waverly [37°53'24"S, 145°10'23"E] was used. For the detail description of the processing chain involved in deriving the specific attenuation, please refer to the Chapter 4.3 section. Besides this, the same commercial microwave link along the dedicated experimental microwave link path was also used for the independent testing of the model.

6.3 Methodology

6.3.1 Long short-term memory network

Long Short-Term Memory (LSTM) architecture is a specific type of recurrent neural network (RNN) that was originally designed to capture long-term dependencies in time series data. It has the capability of overcoming issues of vanishing and exploding gradients (Hochreiter and Schmidhuber, 1997). This architecture preserves the states over a longer period of time without losing temporal dependencies (Hochreiter and Schmidhuber, 1997). For the problem related to the temporal distribution with non-linearity in data, such as natural language processing, image classification and sound translation, this method has proved the most useful when compared with other statistical and conventional feed-forward models (Shen, 2018; Shen et al., 2018).

In this study, among various architectures within the LSTM, the sequence-to-sequence LSTM network has been used herein (Figure 1a). The output $y = y_1, y_2, \dots, \dots, y_n$ is the average attenuation from an input $x = x_1, x_2, \dots, \dots, x_n$ consisting of n consecutive time steps. The

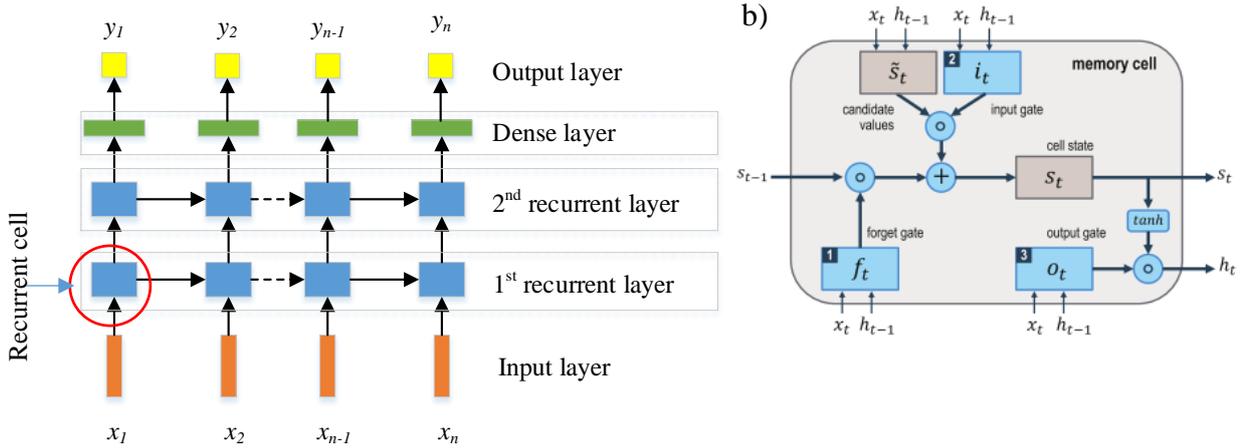


Figure 6-1: (a) General architecture of a two-layer sequence-to-sequence recurrent neural network. The output of the second layer for each of the time steps is fed into a dense layer to calculate the prediction ($y_1, y_2 \dots y_n$). (b) The internal architecture of the LSTM recurrent cell [Adapted from: Hu et al. (2018)].

input variables are the minimum and maximum attenuation and the output is the average attenuation.

a) Each time step t ($1 \leq t \leq n$), the current input x_t is processed in the LSTM recurrent cells of each layer in a network, as in Figure 1(b). The LSTM cell is composed of an input layer, one or more memory cells and an output layer. The major feature of LSTM networks is that they contain hidden layers, which are referred to as memory cells. Each of these memory cells is composed of three gates for adjusting the internal cell state (s_t): the forget gate (f_t), an input gate (i_t) and an output gate (o_t). Details of the LSTM algorithm are explained by Hochreiter and Schmidhuber (1997), and can be summarized as

$$g_t = \tanh(W_{gx}x_t + W_{gh}h_{(t-1)} + b_g) \quad (6.1)$$

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{(t-1)} + b_i) \quad (6.2)$$

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{(t-1)} + b_f) \quad (6.3)$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{(t-1)} + b_o) \quad (6.4)$$

$$s_t = g_t \odot i_t + s_{t-1} \odot o_t \quad (6.5)$$

$$h_t = \tanh(s_t) \odot o_t \quad (6.6)$$

$$y_t = (W_{hy}h_t + b_y), \quad (6.7)$$

where σ is the logistic sigmoidal function, \odot is an element wise multiplication, g_t is input node, x_t is the input forcing, W are the network weights, b are bias parameters and y is the

output for the time step t . Among these parameters, W and b are learnable (adaptable) parameters which are updated at each time step based on a given loss function.

6.3.2 Modelling approach

For this study, a two-layer LSTM model was designed with fully connected hidden layers, 50 LSTM neurons and a single dense layer for predicting output was implemented to form the base architecture of the model. This architecture was adopted based on the prior sensitivity analysis using different numbers of layers and neurons. The Mean Squared Error (MSE) was used as a loss function (quantified and minimized for average attenuation) and the Adaptive Moment Estimation, in short, known as Adam, was used to optimize the model. Adam is an adaptive learning rate optimization algorithm proposed by Diederik and Ba (2014), which performs better than the conventional stochastic gradient descent algorithms.

This model was run for 200 epochs (i.e. 200 passes through the data), with variations in the number of layers and hidden neurons tested to understand the sensitivity and optimize model performance. A newly developed architecture of the RNN called Grated Recurrent Unit (GRU), developed by Cho et al. (2014), was also tested. The GRU has a similar goal of tracking the long-term dependencies in the time series data. This GRU contains two gates called a reset gate and an update gate as opposed to the LSTM, which contains three gates. Similarly, a conventional artificial neural network (which is a memoryless model that does not capture the temporal trend in the data) using dense layers was tested and compared to the performance of the LSTM model. In order to prevent the model from overfitting, a dropout rate of 50% was adopted for each of the layers.

The input dataset (15-minute maximum and minimum attenuation) for the model was produced from the 30-second disdrometer derived attenuation. This whole dataset was split into sets of sub-sequences of size N_w , which helps in the backpropagation of the LSTM model over time. This is helpful in back-propagation through time, which unfolds this LSTM into a feed-forward neural network with a number of recurrent steps, N_w . This N_w was determined based on the most prevalent duration of rain events (based on the Poisson-type distribution). Time steps of 10, 15 and 20 minutes were considered for this. Samples with all dry time steps were discarded. For the rain events longer than the considered time-steps, these events were split into overlapping sequences. Here we define a rainfall event as a rain period separated by a 1-hour or longer rain-free period and having a minimum rainfall rate of 0.1mm/hr.

The chosen dataset was split into two groups of sequences; (i) the training group corresponding to 80% of the whole sequence, and (ii) the testing group corresponding to the remaining 20% of the whole sequence (typical independent test amount range between 10 and 20%). And within a training group, 80% of the data was used to optimize the model and the remaining 20 % was used as a validation set to monitor the learning process. This validation was conducted to make sure that there was no overfitting of the model parameters during the training phase.

6.3.3 Rainfall retrieval from CML data

The algorithm introduced by Overeem et al. (2016a) was used for the rainfall retrieval from commercial microwave link data. A brief description of the algorithm is given below:

a) Wet/dry classification: The algorithm proposed by Overeem et al. (2016a) uses a spatial correlation by looking at the nearby links to identify the given time steps as either wet or dry. In the case presented herein, there was only one link, and so this classification was based solely on the disdrometer data. Time steps for which the rainfall rate observed by the disdrometer was greater than or equal to 0.1 mm h⁻¹ were classified as “wet”, and all remaining time steps were classified as “dry”.

b) Identification of the reference/baseline signal: This was calculated as the moving median of the signal level during the previous 24-hour dry period.

c) Wet antenna attenuation correction: A constant wet antenna attenuation was obtained based on the optimization a microwave link using the algorithm suggested by de Vos et al. (2019). In this case, average RSL data was used to obtain the wet antenna attenuation of 1.2 dB. The resultant (total – wet antenna) attenuation was later divided by the path length to obtain the rain-induced specific attenuation (k).

d) Computation of the rainfall rate: This was done using the power-law relationship between rainfall intensity (R) and specific attenuation (k), (Olsen et al., 1978) as

$$R = ak^b, \quad (6.8)$$

where, R is the rainfall intensity (mm h⁻¹), k is the specific attenuation of the signal (dB km⁻¹), and a and b are parameters depending on the frequency, polarization, drop size distribution, drop shape and canting angle. The values for parameters a and b ($a = 9.563$ and

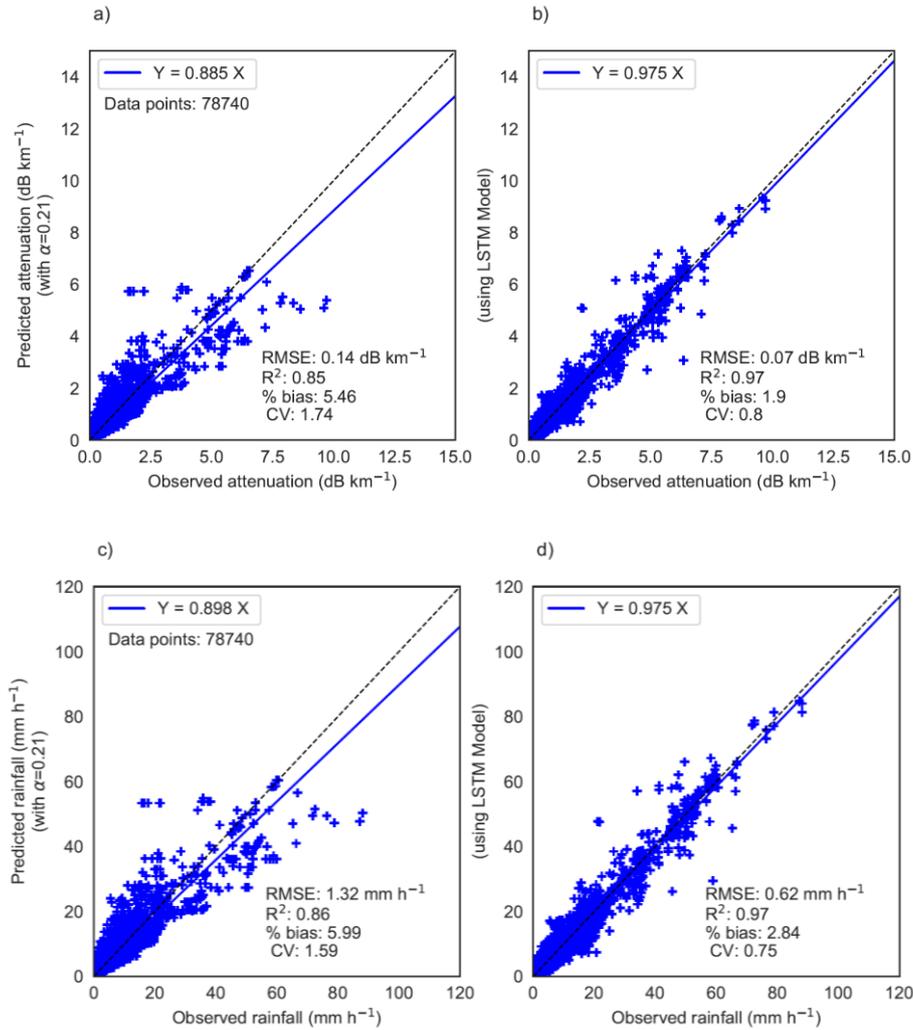


Figure 6-2: Comparison of the predicted to the observed (a, b) average attenuation based on disdrometer derived attenuation and (c,d) rainfall rate based on disdrometer derived attenuation estimates for the weighted average method with $\alpha=0.21$; and for the LSTM Model.

$b = 0.956$) were derived for Melbourne for an equivalent 22.715 GHz microwave link using data obtained from an OTT PARSIVEL¹ optical disdrometer over a three-year period.

6.4 Results

6.4.1 Verification result

Figure 6-2 shows an illustrative comparison of the rainfall estimation (obtained from the attenuation estimate using equation (10) with the same a and b parameters) using the weighted average method with parameter $\alpha = 0.21$ and the LSTM model. The factor $\alpha = 0.21$

was obtained by optimizing the minimum and maximum attenuation values. It shows how the LSTM model outperformed the constant weighted average method when compared with observed attenuation and rainfall directly derived from the 30-second disdrometer data. Using the weights 21% and 79% for the maximum and minimum, respectively, resulted in a larger relative bias in both average attenuation and rainfall rate compared to using the LSTM model. The Root Mean Square Error (RMSE) between the attenuation measured by the disdrometer and that obtained from LSTM dropped from 0.14 dB km^{-1} to 0.07 dB km^{-1} , while the relative bias reduced from 5.6% to 1.9% and the coefficient of determination (R^2) increased from 0.85 to 0.97. Similar improvements were observed for rainfall intensity. To further examine the performance of the LSTM model over a constant weighted average method, other α values, ranging from 0 to 1 with a step of 0.1, were also tested. It was found that the statistics (RMSE, R^2 , CV and percentage bias) of the LSTM model were better compared to those of the constant weighted average method for all values of α considered.

Figure 6-3 shows the time series of the rainfall intensity observations and model estimates for the longest rainfall event that occurred on 2 and 3 December 2017 and lasted for 31 hours. The LSTM model is able to accurately capture the entire range of rainfall intensities, including higher rainfall intensities. The RMSE is reduced from 4.29 mm h^{-1} to 1.04 mm h^{-1} . Similarly, the relative bias in the mean is reduced from 55.95% to -2.01% for the LSTM model compared with other conventional approaches.

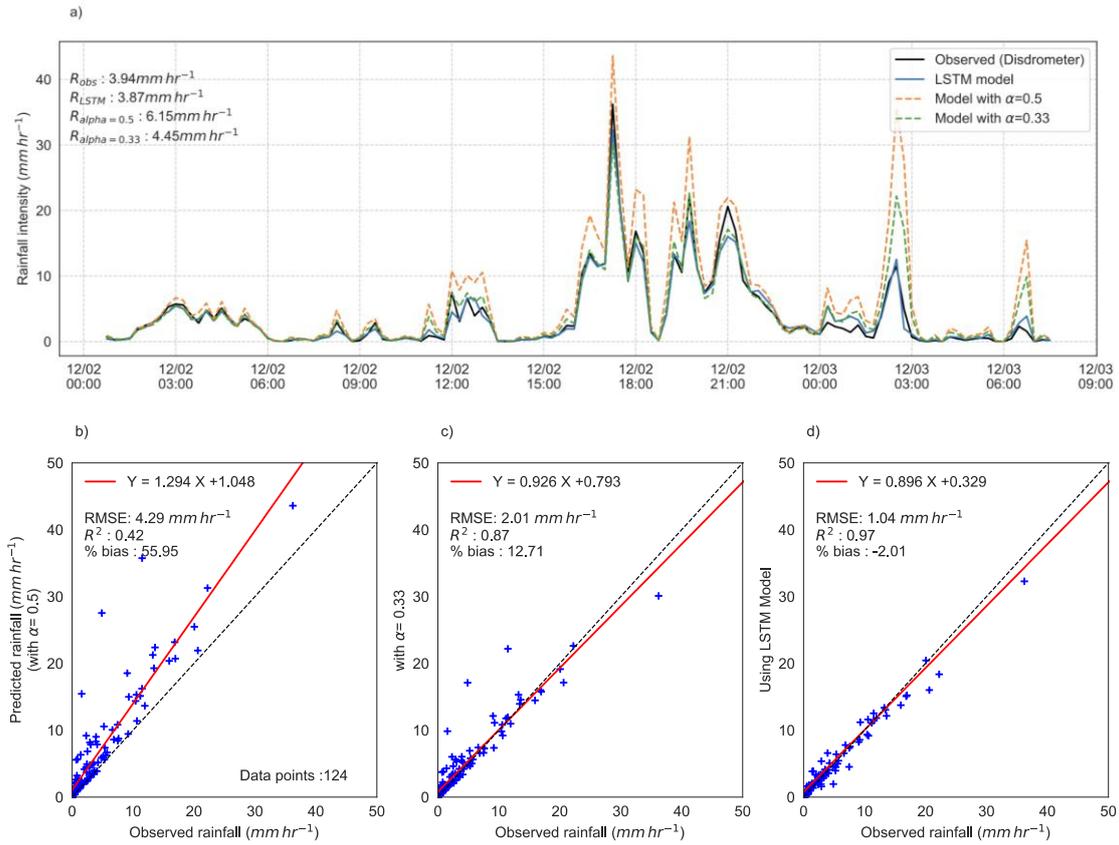


Figure 6-3: (a) Time series of predicted and observed rainfall intensity based on disdrometer data from conventional models and LSTM for 2 and 3 December 2017. Bottom panels: Comparison of the predicted to the observed rainfall for the same days for (b) the conventional approach with $\alpha=0.31$; (c) the conventional approach with $\alpha=0.21$; (d) the LSTM model.

6.4.2 Performance of various architectures of the deep learning model

Table 6-1 shows five statistical measures for three different architectures of the deep learning model. Among these, the dense layer was a traditional artificial neural network with no capacity to model time dependencies and the other two (GRU and LSTM) were different types of RNN architectures. For both the training and test datasets, the performance of the LSTM and GRU were very similar, while the GRU performed better in terms of relative bias. The remaining four statistical parameters were similar. The RMSE, MAE, R^2 and Coefficient of Variation (CV) values for the LSTM model indicate that it performed slightly better than the GRU and much better than the dense layer.

Table 6-1: Comparison of the performance of various model architectures for rainfall prediction for the training and test datasets based on the disdrometer.

Data sets	Architecture	R ²	RMSE (mm h ⁻¹)	MAE (mm h ⁻¹)	Relative bias (%)	CV
Training datasets						
31495 sets	Dense Layer (ANN)	0.85	1.12	0.21	-6.23	1.48
	GRU	0.96	0.61	0.17	-0.17	0.81
	LSTM	0.97	0.62	0.17	2.87	0.76
Test datasets						
7874 sets	Dense layer (ANN)	0.86	1.11	0.20	-0.13	1.47
	GRU	0.95	0.71	0.20	-0.45	0.84

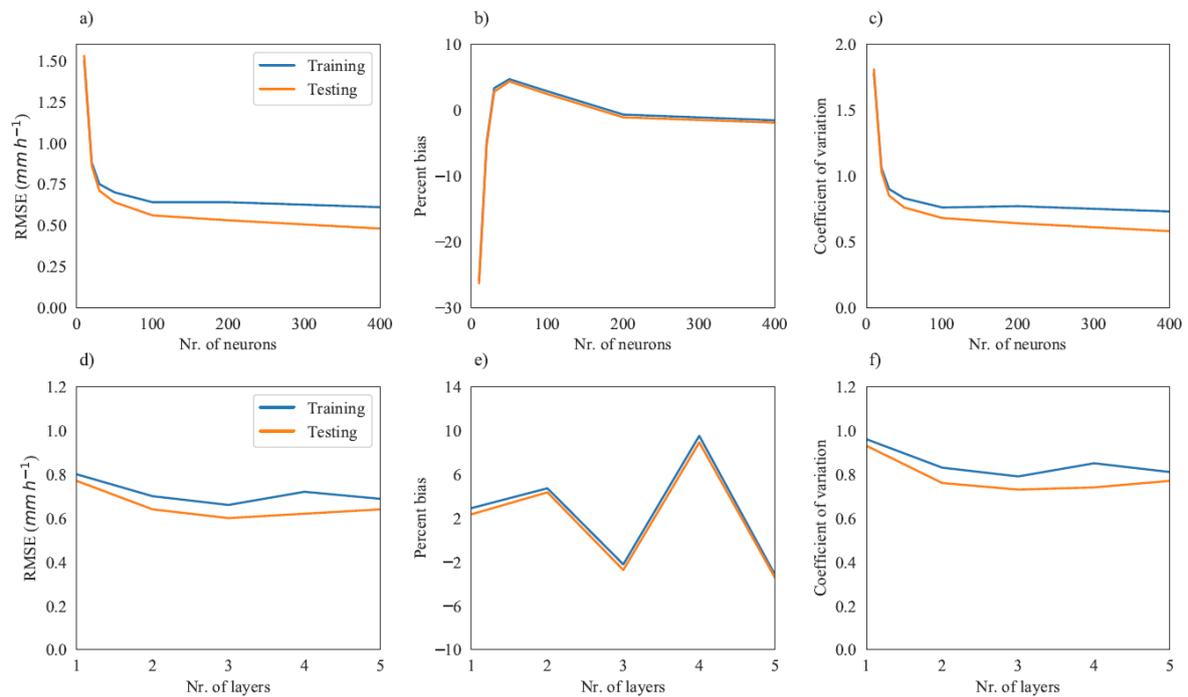


Figure 6-4: LSTM model performance for (a) RMSE; (b) Relative percentage bias; (c) Coefficient of variation vs No. of hidden neurons; (d) RMSE; (e) Relative percentage bias; and (f) Coefficient of variation vs no. of hidden layers.

6.4.3 Sensitivity analysis

a) Sensitivity of the LSTM model

Figure 6-4 shows the LSTM model sensitivity for different numbers of neurons and hidden layers for the training and test datasets. The RMSE decreased significantly with an increasing amount of neurons up to 100 and remained constant after increasing the number of neurons further. Similar characteristics were observed for the relative bias and the coefficient of variation. For the hidden layers, the LSTM model performance increased from one to three layers and started to deteriorate beyond that value. The performance of the model showed the fluctuation based on percent bias but its best performance was observed with three hidden layers.

b) Sensitivity on input data

To understand the impact of input data on the overall performance of LSTM model, the following two cases were considered:

- 1) Case I: adding a constant bias of ± 0.4 dB km⁻¹ in both minimum and maximum attenuation, which is almost equal to the wet antenna attenuation applied to the commercial microwave link data (refer to Figure 6-5 and 6-6).
- 2) Case II: adding a random noise of -0.3 to +0.3 dB km⁻¹, which is equivalent to the error that may be caused due to quantization and other errors (refer to Figure 6-7).

Figure 6-5 and Figure 6-6 showed that the performance of the LSTM model deteriorates when compared with clean attenuation data. However, overall performance of the LSTM model itself was still better for both cases compared with the constant weighted average method with parameter α , although there was a significant bias in the rainfall estimation. Thus, the outcome of the LSTM model is linked to the nature of the bias in the input data (in our case, in the attenuation obtained from the min and max RSL data from CML).

Figure 6-7 showed the performance of The LSTM model with added random noise data. For this case also, the overall performance of the model was still better compared with the constant weighting factor method. Also, when comparing the two cases, the LSTM model with added random noise input showed lower bias in the rainfall estimates compared with constant bias.

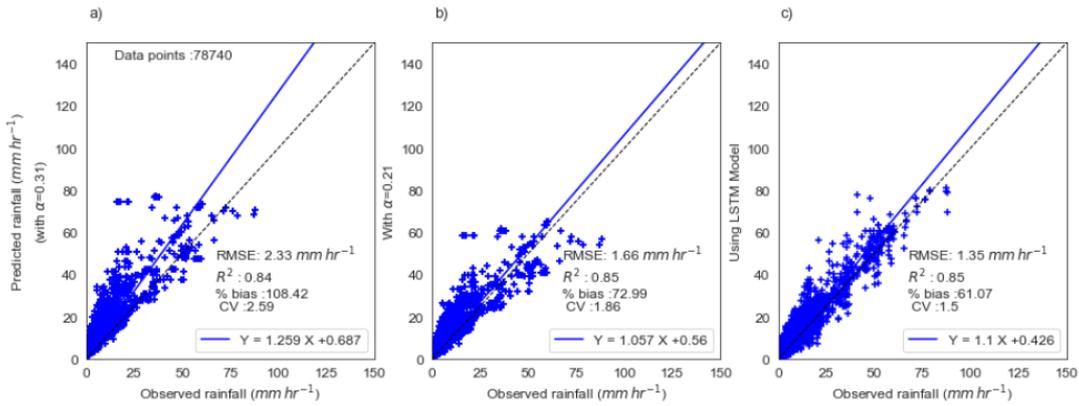


Figure 6-5: Scatter plot of the predicted *versus* observed rainfall intensity with +0.4 dB km⁻¹ constant bias for (a) with $\alpha=0.31$; (b) with $\alpha=0.21$; and (c) LSTM model

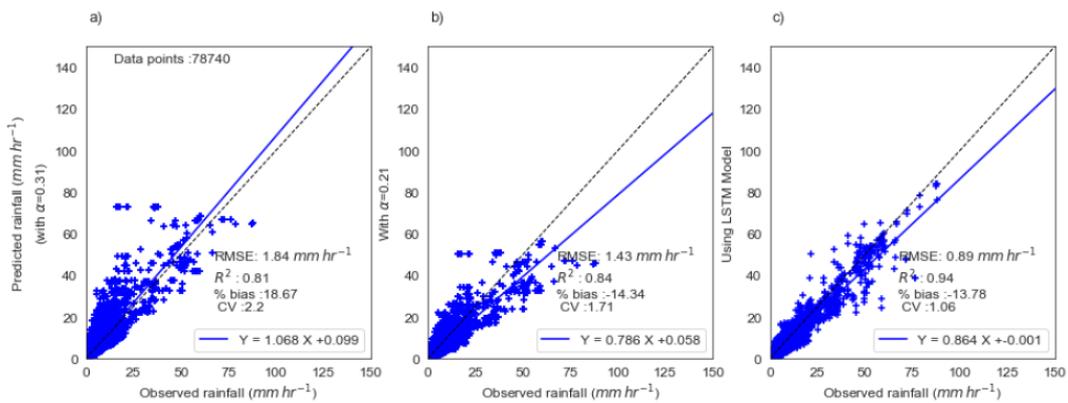


Figure 6-6: Scatter plot of the predicted *versus* observed rainfall intensity with -0.4 dB km⁻¹ constant bias (a) with $\alpha=0.31$; (b) with $\alpha=0.21$; and (c) LSTM model

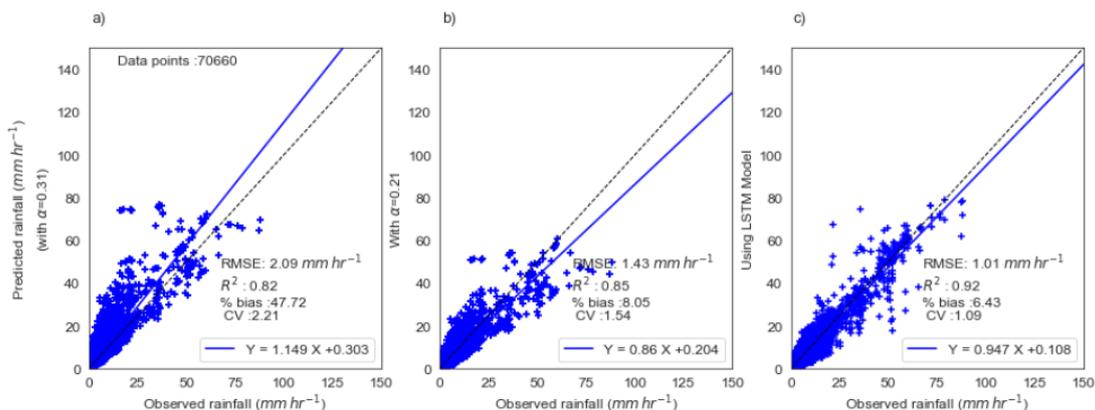


Figure 6-7: Scatter plot of the predicted *versus* observed rainfall intensity with added random noise from -0.3 to +0.3 dB km⁻¹ for (a) a weighting factor with $\alpha=0.31$; (b) a weighting factor with $\alpha=0.21$; and (c) LSTM model.

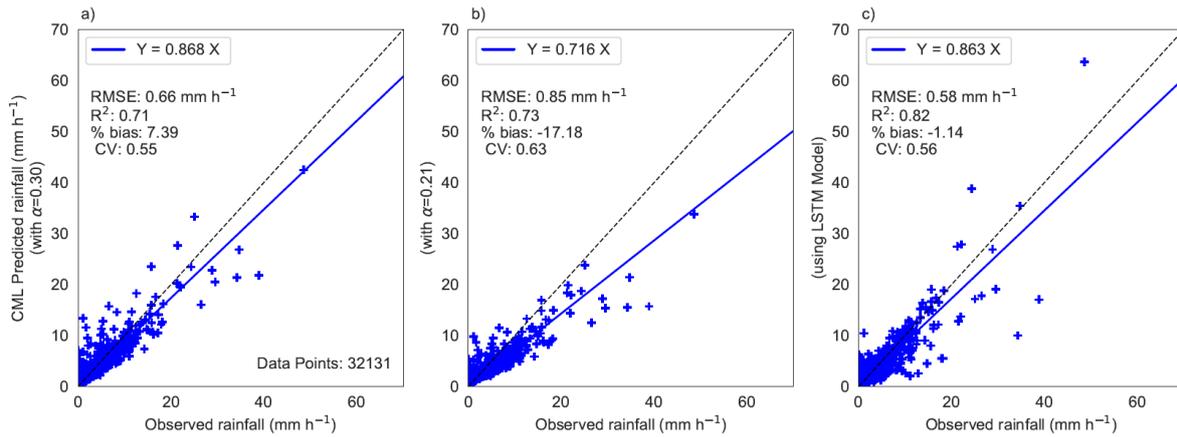


Figure 6-8: Scatter plots of the predicted *versus* observed rainfall intensity using commercial microwave link data for: (a) the weighted average method with $\alpha=0.30$; (b) the weighted average method with $\alpha=0.21$; and (c) the disdrometer-trained LSTM Model.

6.4.4 Application of the disdrometer-trained LSTM model to the CML dataset

Figure 6-8 shows the comparison of the predicted rainfall using the CML observations for the 22 GHz CML link closest to the disdrometer site. The factor $\alpha = 0.21$ and wet antenna attenuation of 1.2 dB were obtained through optimisation of average rainfall obtained from minimum and maximum rainfall based on the gauge-adjusted radar data obtained from the Bureau of Meteorology. The LSTM model outperformed the two weightage average method results ($\alpha = 0.30$ and 0.21) for all but a few events. In particular, shorter duration convective rainfall events were under-estimated by the LSTM model compared with the weighted average method.

6.5 Summary

This chapter proposed a new data-driven model which uses deep learning techniques for more accurate rainfall estimation. The developed model was independently verified with a dataset that had not been used for training the model. These results have demonstrated the ability to use such data-driven models for rainfall retrieval. Generally speaking, this technique is the first study to propose such an approach based on building the model using disdrometer data and applying it to real commercial microwave link data.

Chapter 7

Conclusions and future work

7.1 Conclusions

Backhaul towers from telecommunication companies provide a valuable source of rainfall information close to the ground surface, complementing in-situ measurement devices like rain gauges and weather radar. The accuracy of such rainfall estimates depends on a number of factors. Researchers have made significant advances in developing algorithms and models for improving the rainfall estimates from such opportunistic data sources. However, real-world applications have been minimal due to a lack of widespread testing. Therefore, the main contribution of this research was validating the existing rainfall retrieval algorithms for the Australian climate, developing an understanding of uncertainties in rainfall retrieval using different sampling strategies, and developing a new data-driven rainfall methodology. Such techniques have never been tested in the Australian weather conditions, providing the basis to further develop and improve rainfall estimation.

7.1.1 Uncertainties associated with microwave link rainfall estimates

Key parameters used in the power-law model, which is the basis for the commonly used rainfall retrieval algorithms, were assessed using data collected from the two OTT disdrometers. These parameters were later used to validate a total of 72 rainfall events collected between February 2018 and December 2019. Later, this data was resampled according to three commonly used strategies; namely average, minimum/maximum, and instantaneous. Moreover, this was undertaken for two different time intervals i.e. 5 and 15-minutes, to understand the impact on the rainfall retrieval accuracy and compared with the reference 1-minute dataset. This work also intercompared the performance of three wet antenna attenuation models. Analyses led to several conclusions:

- The path-averaged rainfall estimation from the experimental link showed a strong correlation of 0.95 with ground observations. However, a systematic overestimation of about 103% was observed. When applied to commercial microwave links the rainfall estimation showed a larger bias of 142%. In addition, there was more noise

observed in the CML derived rainfall estimates (CV of 3.18) compared with the experimental link (CV of 2.11).

- Based on a comparison of three commonly used sampling strategies, for both 5 and 15-minute time periods, it was found that the average sampling strategy showed better rainfall estimates compared with the minimum/maximum and instantaneous sampling strategies.
- For the quantization errors, for the 24 GHz link it was found that there was no such significant impact of data quantization.
- The signal attenuates up to 3.5 dB when the receiver antenna got wet, corresponding to 7.5 mm/h rainfall at 24 GHz. It was not possible to test the effect on the transmitting antenna.
- All three wet antenna attenuation models tested showed a substantial improvement in rainfall retrieval. Based on all performance statistics, both the exponential decay and power-law models showed very similar performance and outperformed the constant wet antenna model. However, when applied to the commercial microwave link data, the constant wet antenna model performed best.

7.1.2 Rainfall retrieval from commercial microwave links

This study presented rainfall retrievals over the greater Melbourne Metropolis using 135 commercial microwave links operating at frequencies ranging from 10 to 40 GHz with path lengths of 0.2 km to 25 km. This study was the very first to conduct a comparison of rainfall retrieval using CML data with two different sampling strategies over the same link paths (Average and MinMax RSL data over 15 minutes). For this study, the RAINLINK package was used, and a new set of parameters derived for rainfall retrieval for Melbourne. The analyses led to the following conclusions:

- For the wet-dry classification using the nearby link approach, MinMax data was found to perform best, with a lower FAR (3.7%) and similar POD compared with the Average data. Other statistics, including Matthew's correlation coefficient (0.45 for MinMax and 0.35 for Average) and error rate, suggested that the use of MinMax data achieved fewer false alarms.
- The Average data sampling approach had a similar rainfall retrieval performance to using the MinMax data sampling approach. When compared for different

accumulation intervals, the MinMax data sampling approach provided the best performance based on statistics (relative bias, CV, RMSE and ρ) and double mass curve.

7.1.3 Deep learning for improved rainfall estimates

A novel approach for improving the rainfall estimation from commercial microwave links (when only limited information such as minimum and maximum RSL data are available) by using a deep learning model was proposed. This used the data collected from a disdrometer for training and testing prior to applying the model to the commercial microwave link data. Analyses led to the following conclusions:

- Results showed a substantial improvement based on simulated microwave link attenuation data from a disdrometer and for real commercial microwave link data, compared with two weighted average methods with parameters of $\alpha = 0.31$ or $\alpha = 0.21$.
- Although the performance of the deep learning model was lower when used on commercial microwave link data rather than the simulated data, there was still a good improvement in bias and R^2 of the rainfall estimation compared to the weightage average method.

7.2 Limitations and Future work

There are a couple of limitations in this research, and so the corresponding future needs are as listed below:

- 1) For the experimental study, only two disdrometers were installed, with one at each end of the experimental microwave link. As this microwave link was about 4 km in length, the drop size distribution measured at two ends only would not fully capture the spatial distribution of the drop size distribution (DSD) along the path. It is therefore recommended to install more disdrometers along the path in future work in order to investigate the variability of DSDs at a small spatial scale. This would provide more insights on the distribution of the received signal data sampled at the coarser time resolutions of 15-minutes. Additionally, it is also recommended for future studies to include a high-resolution time-lapse camera on the transmitter and receiver antenna so that other environmental phenomena such as dew formation and its impact on the rainfall retrieval could be investigated.

- 2) This study has evaluated the use of CML data for the Melbourne metropolitan area. Here, gauge-adjusted radar data were used as the reference data even though there are limitations of such radar-derived products. Thus, it is recommended to compare with nearby gauge data for the link performance. Furthermore, it is recommended that CML rainfall estimates be used as a complementary source of information where there are no radars or as an alternative rainfall measurement instrument. At the city scale, CML-derived rainfall estimates can help overcome data gaps due to radar clutter from high-rise buildings and the absence of traditional rain gauges. Accordingly, there is a great opportunity to combine the three datasets into a “merged product”. Moreover, studying the performance of different sampling strategies in detail with a high-resolution dataset is expected to provide greater insight to the optimal sampling strategy. This may also lead to an improved rainfall retrieval algorithm.
- 3) The proposed deep learning model was only verified on a single CML link close to the disdrometer, but a similar methodology could be adopted for other links having different frequencies within the same climate. Usually, links having frequencies ranging from 20 to 40 GHz are most suitable as these tend to be associated with a close to linear specific attenuation-rain rate relationship (Berne and Uijlenhoet, 2007). Following the approach developed in this thesis and based on a time series of attenuation and rainfall rates from a single disdrometer data set, various deep learning models for each given frequency can be designed and implemented. The limitations of this data-driven approach reside first in the size and representativeness of the collected disdrometer data set for a given location. The applicability of this trained model to retrieve rainfall for a wider area still has to be demonstrated, and likely the size (duration, diversity and quantity of the recorded rainfall events) of the disdrometer data will also be an important factor. Second, this disdrometer-trained model replaces only one of the steps among numerous successive steps in the rainfall retrieval, such as dry/wet classification, baseline estimation, and wet antenna attenuation. Such data-driven deep learning approaches for dry/wet classification have also been recently explored by Polz et al. (2019) and Habi et al. (2018). Similarly, one could think of a similar approach for baseline estimation and wet antenna attenuation. Also, adaptive learning and/or transfer learning techniques with

such deep learning models could be implemented and applied in places where such disdrometer data are unavailable.

Appendix

Table A1: Dataset used for calibration of RAINLINK parameters.

Day	Time steps with rain	Total rain (mm)	Maximum rainfall intensity (mm/hr)
2017-08-06	34	10.64	4.40
2017-08-07	32	7.87	3.34
2017-08-15	33	8.87	6.17
2017-08-18	51	9.53	2.86
2017-08-26	28	13.65	8.47
2017-09-05	33	12.61	7.12
2017-09-07	38	10.35	3.76
2017-09-08	36	5.85	2.28
2017-09-13	33	8.21	3.92
2017-09-15	51	30.13	19.64
2018-06-08	35	12.64	5.25
2018-06-15	22	6.74	6.07
2018-06-16	48	30.19	7.44
2018-07-07	57	19.55	5.03
2018-07-08	70	15.42	3.66
2018-08-07	31	9.80	4.08
2018-08-12	39	6.06	2.27
2018-08-18	51	18.52	6.57
2018-08-19	38	7.59	3.99
2018-08-21	38	4.29	2.35
2018-10-09	37	7.25	4.11
2018-10-16	24	8.85	6.99
2018-10-17	21	4.20	3.72
2018-10-19	27	13.35	6.87
2018-11-07	32	9.66	5.17
2018-11-20	31	32.12	17.69
2018-11-21	24	13.14	8.97
2018-11-22	64	33.66	9.22
2018-11-23	58	17.93	4.07
2018-12-20	26	3.18	0.94

References

Australian Bureau of Statistics:

https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/quickstat/2GMEL?opendocument, access: 16/04/2020, 2016.

Alemayehu, T., Kilonzo, F., van Griensven, A., and Bauwens, W.: Evaluation and application of alternative rainfall data sources for forcing hydrologic models in the Mara Basin, *Hydrology Research*, 49, 1271-1282, <https://doi.org/10.2166/nh.2017.081>, 2017.

Atlas, D., Srivastava, R. C., and Sekhon, R. S.: Doppler radar characteristics of precipitation at vertical incidence, *Reviews of Geophysics*, 11, 1-35, 10.1029/RG011i001p00001, 1973.

Atlas, D., and Ulbrich, C. W.: Path-and area integrated rainfall measurement by microwave attenuation in the 1-3 cm band, *J. Appl. Meteorol.*, 16, 10, 1977.

Berne, A., and Krajewski, W. F.: Radar for hydrology: Unfulfilled promise or unrecognized potential?, *Adv Water Resour*, 51, 357-366, <https://doi.org/10.1016/j.advwatres.2012.05.005>, 2013.

Berne, A., and Uijlenhoet, R.: Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability, *Geophysical Research Letters*, 34, <https://doi.org/10.1029/2007gl029409>, 2007.

Bianchi, B., Rieckermann, J., and Berne, A.: Quality control of rain gauge measurements using telecommunication microwave links, *Journal of Hydrology*, 492, 15-23, <https://doi.org/10.1016/j.jhydrol.2013.03.042>, 2013.

Blazquez, R. R., Vivier, E., and Godara, B.: Meteorology Using microwaves links: A comparative study, 2009 6th International Symposium on Wireless Communication Systems, 2009, 328-332,

Blevins, B.: Losses due to rain on radomes and antenna reflecting surfaces, *IEEE Transactions on Antennas and Propagation*, 13, 175-176, <https://doi.org/10.1109/TAP.1965.1138384>, 1965.

- Brandes, E. A., Zhang, G., and Vivekanandan, J.: An Evaluation of a Drop Distribution–Based Polarimetric Radar Rainfall Estimator, *Journal of Applied Meteorology*, 42, 652-660, [https://doi.org/10.1175/1520-0450\(2003\)042<0652:Aeoadd>2.0.Co;2](https://doi.org/10.1175/1520-0450(2003)042<0652:Aeoadd>2.0.Co;2), 2003.
- Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., and Levizzani, V.: Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data, *Journal of Geophysical Research: Atmospheres*, 119, 5128-5141, <https://doi.org/10.1002/2014JD021489>, 2014.
- Cherkassky, D., Ostrometzky, J., and Messer, H.: Precipitation Classification Using Measurements From Commercial Microwave Links, *Ieee T Geosci Remote*, 52, 2350-2356, 10.1109/TGRS.2013.2259832, 2014.
- Christopher, S. R., Kultegin, A., Savyasachee, M., and Justin, P. B.: 35-GHz Dual-Polarization Propagation Link for Rain-Rate Estimation, *Journal of Atmospheric and Oceanic Technology*, 13, 419-425, 1996.
- Chwala, C., Gmeiner, A., Qiu, W., Hipp, S., Nienaber, D., Siart, U., Eibert, T., Pohl, M., Seltmann, J., Fritz, J., and Kunstmann, H.: Precipitation observation using microwave backhaul links in the alpine and pre-alpine region of Southern Germany, *Hydrology and Earth System Sciences*, 16, 2647-2661, <https://doi.org/10.5194/hess-16-2647-2012>, 2012.
- Chwala, C., Keis, F., and Kunstmann, H.: Real-time data acquisition of commercial microwave link networks for hydrometeorological applications, *Atmospheric Measurement Techniques*, 9, 991-999, <https://doi.org/10.5194/amt-9-991-2016>, 2016.
- Chwala, C., and Kunstmann, H.: Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges, *Wiley Interdisciplinary Reviews: Water*, 6, e1337, <https://doi.org/10.1002/wat2.1337>, 2019.
- David, N., Alpert, P., and Messer, H.: The potential of cellular network infrastructures for sudden rainfall monitoring in dry climate regions, *Atmos Res*, 131, 13-21, <https://doi.org/10.1016/j.atmosres.2013.01.004>, 2013.
- de Vos, L. W., Overeem, A., Leijnse, H., and Uijlenhoet, R.: Rainfall Estimation Accuracy of a Nationwide Instantaneously Sampling Commercial Microwave Link Network: Error Dependency on Known Characteristics, *Journal of Atmospheric and Oceanic Technology*, 36, 1267-1283, <https://doi.org/10.1175/jtech-d-18-0197.1>, 2019.

- Doumounia, A., Gosset, M., Cazenave, F., Kacou, M., and Zougmore, F.: Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a West African test bed, *Geophysical Research Letters*, 41, 6016-6022, <https://doi.org/10.1002/2014GL060724>, 2014.
- Doumounia, A. G., M.; Cazenave, F.; Kacou, M.; Zougmore, F.: Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a West African test bed, *Geophysical Research Letters*, 41, 6016-6022, 10.1002/2014GL060724, 2014.
- Emad, H., Ehab, A., and Ananda, V.: Effect of Local Errors of Tipping-Bucket Rain Gauges on Rainfall-Runoff Simulations, *Journal of Hydrologic Engineering*, 13, 488-496, [https://doi.org/10.1061/\(ASCE\)1084-0699\(2008\)13:6\(488\)](https://doi.org/10.1061/(ASCE)1084-0699(2008)13:6(488)), 2008.
- Ericsson: Ericsson microwave outlook: trends and needs in the microwave industry, 2017.
- Fencl, M., Dohnal, M., Rieckermann, J., and Bareš, V.: Gauge-adjusted rainfall estimates from commercial microwave links, *Hydrology and Earth System Sciences*, 21, 617-634, <https://doi.org/10.5194/hess-21-617-2017>, 2017.
- Fencl, M., Rieckermann, J., Schleiss, M., Stransky, D., and Bares, V.: Assessing the potential of using telecommunication microwave links in urban drainage modelling, *Water science and technology : a journal of the International Association on Water Pollution Research*, 68, 1810-1818, <https://doi.org/10.2166/wst.2013.429>, 2013.
- Fencl, M., Rieckermann, J., Sýkora, P., Stránský, D., and Bareš, V.: Commercial microwave links instead of rain gauges: fiction or reality?, *Water science and technology : a journal of the International Association on Water Pollution Research*, 71, 31, <https://doi.org/10.2166/wst.2014.466>, 2015.
- Fencl, M., Valtr, P., Kvicera, M., and Bares, V.: Quantifying Wet Antenna Attenuation in 38 GHz Commercial Microwave Links of Cellular Backhaul, *IEEE Geoscience and Remote Sensing Letters*, 1-5, <https://doi.org/10.1109/LGRS.2018.2876696>, 2018.
- Fencl, M., Valtr, P., Kvičera, M., and Bareš, V.: Quantifying Wet Antenna Attenuation in 38-GHz Commercial Microwave Links of Cellular Backhaul, *IEEE Geoscience and Remote Sensing Letters*, 16, 514-518, <https://doi.org/10.1109/LGRS.2018.2876696>, 2019.

- Fenicia, F., Pfister, L., Kavetski, D., Matgen, P., Iffly, J., Hoffmann, L., and Uijlenhoet, R.: Microwave links for rainfall estimation in an urban environment: Insights from an experimental setup in Luxembourg-City, *Journal of Hydrology*, 464-465, 69-78, <https://doi.org/10.1016/j.jhydrol.2012.06.047>, 2012.
- Garcia-Rubia, J. M., Riera, J. M., Benarroch, A., and Garcia-Del-Pino, P.: Estimation of Rain Attenuation From Experimental Drop Size Distributions, *IEEE Antennas and Wireless Propagation Letters*, 10, 839-842, <https://doi.org/10.1109/LAWP.2011.2163609>, 2011.
- Germann, U., Galli, G., Boscacci, M., and Bolliger, M.: Radar precipitation measurement in a mountainous region, *Quarterly Journal of the Royal Meteorological Society*, 132, 1669-1692, <https://doi.org/10.1256/qj.05.190>, 2006.
- Giuli, D., Toccafondi, A., Gentili, G. B., and Forni, A.: Tomographic reconstruction of rainfall fields through microwave attenuation measurements, *J. Appl. Meteorol.*, 30, 17, 1991.
- Goldshtein, O., Messer, H., and Zinevich, A.: Rain Rate Estimation Using Measurements From Commercial Telecommunications Links, *IEEE Transactions on Signal Processing*, 57, 1616-1625, <https://doi.org/10.1109/TSP.2009.2012554>, 2009.
- Goldshtein, O., Messer, H., and Zinevich, A.: Rain Rate Estimation Using Measurements From Commercial Telecommunications Links, *IEEE Transactions on Signal Processing*, 57, 1616-1625, <https://doi.org/10.1109/TSP.2009.2012554>, 2009.
- Graf, M., Chwala, C., Polz, J., and Kunstmann, H.: Rainfall estimation from a German-wide commercial microwave link network : Optimized processing and validation for 1 year of data, *Hydrol. Earth Syst. Sci.*, 24, 2931-2950, <https://doi.org/10.5194/hess-24-2931-2020>, 2020.
- Graves, A., and Schmidhuber, J.: Offline handwriting recognition with multidimensional recurrent neural network, *Advances in neural information processing systems*, 545-552, 2009.
- Guyot, A., Pudashine, J., Protat, A., Uijlenhoet, R., Pauwels, V. R. N., Seed, A., and Walker, J. P.: Effect of disdrometer type on rain drop size distribution characterisation: a new dataset for south-eastern Australia, *Hydrology and Earth System Sciences*, 23, 4737-4761, <https://doi.org/10.5194/hess-23-4737-2019>, 2020.

- Habi, H. V., and Messer, H.: Wet-Dry Classification using LSTM and Commercial Microwave Links, IEEE 10th Sensory Array and Multichannel Signal Processing Workshop, 2018.
- Hochreiter, S., and Schmidhuber, J.: Long Short-Term Memory, *Neural Computation*, 9, 1735-1780, 1997.
- Hoedjes, J., Kooiman, A., Maathuis, B., Said, M., Becht, R., Limo, A., Mumo, M., Nduhiu, J., Shaka, A., and Su, B.: A Conceptual Flash Flood Early Warning System for Africa, Based on Terrestrial Microwave Links and Flash Flood Guidance, *ISPRS International Journal of Geo-Information*, 3, 584-598, <https://doi.org/10.3390/ijgi3020584>, 2014.
- Holt, A. R. K., G. G.; Rahimi, A. R.: Comparison of the use of dual-frequency and single-frequency attenuation for the measurement of path-averaged rainfall along a microwave link, *Iee P-Microw Anten P*, 150, 315-320, 10.1049/ip-map:20030616, 2003.
- Hu, C., Wu, Q., Li, H., Jian, S., Li, N., and Lou, Z.: Deep Learning with a Long Short-Term Memory Networks Approach for Rainfall-Runoff Simulation, *Water*, 10, 1543, <https://doi.org/10.3390/w10111543>, 2018.
- Islam, M. R., and Tharek, A. R.: Measurement of wet antenna effect on microwave propagation at 23, 26 and 38 GHz, *IEEE Antennas and Propagation Society International Symposium. Transmitting Waves of Progress to the Next Millennium. 2000 Digest. Held in conjunction with: USNC/URSI National Radio Science Meeting (C, 2000, 2094-2098 vol.2094*,
- ITU-R: Recommendation ITU-R P.838-3, 2005.
- ITU-R: Attenuation by atmospheric gases, Geneva, 276-211, 2016.
- Jaffrain, J., and Berne, A.: Experimental Quantification of the Sampling Uncertainty Associated with Measurements from PARSIVEL Disdrometers, *J Hydrometeorol*, 12, 352-370, <https://doi.org/10.1175/2010jhm1244.1>, 2011.
- Jameson, A. R.: A comparison of microwave techniques for measuring rainfall, *J.Appl. Meteorol*, 30, 25, 1991.
- Joss, J., Waldvogel, A., and Collier, C. G.: Precipitation Measurement and Hydrology, in: *Radar in Meteorology: Battan Memorial and 40th Anniversary Radar Meteorology*

- Conference, edited by: Atlas, D., American Meteorological Society, Boston, MA, 577-606, 1990.
- Kharadly, M. M. Z., and Ross, R.: Effect of wet antenna attenuation on propagation data statistics, *IEEE Transactions on Antennas and Propagation*, 49, 1183-1191, 10.1109/8.943313, 2001.
- Kidd, C., and Levizzani, V.: Status of satellite precipitation retrievals, *Hydrology and Earth System Sciences*, 15, 1109-1116, <https://doi.org/10.5194/hess-15-1109-2011>, 2011.
- Kramer, S. V., H. R.; Redder, A.: Improvement of X-band radar rainfall estimates using a microwave link, *Atmos Res*, 77, 278-299, <https://doi.org/10.1016/j.atmosres.2004.10.028>, 2005.
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M.: Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks, *Hydrology and Earth System Sciences*, 22, 6005-6022, <https://doi.org/10.5194/hess-22-6005-2018>, 2018.
- Kucera, P. A., Ebert, E. E., Turk, F. J., Levizzani, V., Kirschbaum, D., Tapiador, F. J., Loew, A., and Borsche, M.: Precipitation from Space: Advancing Earth System Science, *Bulletin of the American Meteorological Society*, 94, 365-375, <https://doi.org/10.1175/bams-d-11-00171.1>, 2013.
- La Barbera, P., Lanza, L. G., and Stagi, L.: Tipping bucket mechanical error and their influence on rainfall statistics an extremes., *Water Science and Technology*, 45, 4, 2002.
- Leijnse, H., Uijlenhoet, R., and Stricker, J. N. M.: Rainfall measurement using radio links from cellular communication networks, *Water Resources Research*, 43, <https://doi.org/10.1029/2006wr005631>, 2007.
- Leijnse, H., Uijlenhoet, R., and Stricker, J. N. M.: Rainfall measurement using radio links from cellular communication networks, *Water Resources Research*, 43, n/a-n/a, <https://doi.org/10.1029/2006wr005631>, 2007.
- Leijnse, H., Uijlenhoet, R., and Stricker, J. N. M.: Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation, *Adv Water Resour*, 31, 1481-1493, <https://doi.org/10.1016/j.advwatres.2008.03.004>, 2008.

- Leinonen, J.: High-level interface to T-matrix scattering calculations: architecture, capabilities and limitations, *Opt. Express*, 22, 1655-1660, <https://doi.org/10.1364/OE.22.001655>, 2014.
- Liberman, Y., Samuels, R., Alpert, P., and Messer, H.: New algorithm for integration between wireless microwave sensor network and radar for improved rainfall measurement and mapping, *Atmospheric Measurement Techniques*, 7, 3549-3563, <https://doi.org/10.5194/amt-7-3549-2014>, 2014.
- Liebe, H. J.: Modelling attenuation and phase of radio waves in air at frequencies below 1000 GHz, *Radio Science*, 16, 1183-1199, 1981.
- Matthews, B. W.: Comparison of the predicted and observed secondary structure of T4 phase lysozyme, *Biochimica et Biophysica Acta*, 405, 442-451, 1975.
- Messer, H.: Capitalizing on Cellular Technology—Opportunities and Challenges for Near Ground Weather Monitoring, *Environments*, 5, 73, <https://doi.org/10.3390/environments5070073>, 2018.
- Messer, H., Zinevich, A., and Alpert, P.: Environmental Monitoring by Wireless Communication Networks, *Science*, 312, 713-713, <https://doi.org/10.1126/science.1120034>, 2006.
- Messer, H. Z., Artem; Alpert, Pinhas: Environmental Monitoring by Wireless Communication Networks, *Science*, 312, 713-713, 10.1126/science.1120034, 2006.
- Michaelides, S.: *Precipitation: Advances in Measurement, Estimation and Prediction*, 1. Aufl. ed., Berlin, Heidelberg: Springer-Verlag, Berlin, Heidelberg, 2008.
- Michaelides, S., Levizzani, V., Anagnostou, E., Bauer, P., Kasparis, T., and Lane, J. E.: Precipitation: Measurement, remote sensing, climatology and modeling, *Atmos Res*, 94, 512-533, <https://doi.org/10.1016/j.atmosres.2009.08.017>, 2009.
- Minda, H., and Nakamura, K.: High Temporal Resolution Path-Average Rain Gauge with 50-GHz Band Microwave, *Journal of Atmospheric and Oceanic Technology*, 22, 165-179, <https://doi.org/10.1175/jtech-1683.1>, 2005.
- Mishchenko, M. I., and Travis, L. D.: T-matrix computations of light scattering by large spheroidal particles, *Optics Communications*, 109, 16-21, [http://dx.doi.org/10.1016/0030-4018\(94\)90731-5](http://dx.doi.org/10.1016/0030-4018(94)90731-5), 1994.

- Mishra, V. K., Gharanjik, A., Bhavani, M. R., and Bjorn, O.: Deep Learning Framework for Precipitation Retrievals from Communication Satellites, 10th European Conference on Radar in Meteorology and Hydrology, Wageningen, Netherlands, 2018.
- Morin, E., Krajewski, W. F., Goodrich, D. C., Gao, X., and Sorooshian, S.: Estimating Rainfall Intensities from Weather Radar Data
The Scale-Dependency Problem, *J Hydrometeorol*, 4, 782-797, 2003.
- Mueller, G.: Propagation of 6-millimeter waves, *P. IRE*, 183, <https://doi.org/10.1109/JRPROC.1946.234240>, 1946.
- Olsen, R., Rogers, D. V., and Hodge, D. B.: The αR^b relation in the calculation of rain attenuation, *IEEE Transactions on Antennas and Propagation*, 26, 318-329, <https://doi.org/10.1109/TAP.1978.1141845>, 1978.
- Ostrometzky, J., Cherkassky, D., and Messer, H.: Accumulated Mixed Precipitation Estimation Using Measurements from Multiple Microwave Links, *Advances in Meteorology*, 2015, 1-9, <https://doi.org/10.1155/2015/707646>, 2015.
- Ostrometzky, J., and Messer, H.: Accumulated Rainfall Estimation Using Maximum Attenuation of Microwave Radio Signal, *IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, A Coruna, Spain, 2014, 193-196,
- Ostrometzky, J., and Messer, H.: Comparison of Different Methodologies of Parameter-Estimation From Extreme Values, *IEEE Signal Processing Letters*, 24, 1293-1297, <https://doi.org/10.1109/LSP.2017.2723544>, 2017.
- Ostrometzky, J., Raich, R., Eshel, A., and Messer, H.: Calibration of the Attenuation-Rain Rate Power-Law Parameters Using Measurements from Commercial Microwave Networks, *41st IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, 2016,
- Overeem, A., Leijnse, H., and Uijlenhoet, R.: Measuring urban rainfall using microwave links from commercial cellular communication networks, *Water Resources Research*, 47, <https://doi.org/10.1029/2010WR010350>, 2011.
- Overeem, A., Leijnse, H., and Uijlenhoet, R.: Country-wide rainfall maps from cellular communication networks, *Proceedings of the National Academy of Sciences of the*

- United States of America, 110, 2741-2745, <https://doi.org/10.1073/pnas.1217961110>, 2013.
- Overeem, A., Leijnse, H., and Uijlenhoet, R.: Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network, *Atmospheric Measurement Techniques*, 9, 2425-2444, <https://doi.org/10.5194/amt-9-2425-2016>, 2016a.
- Overeem, A., Leijnse, H., and Uijlenhoet, R.: Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks, *Water Resources Research*, n/a-n/a, <https://doi.org/10.1002/2016WR019412>, 2016b.
- Overeem, A. L., H.; Uijlenhoet, R.: Measuring urban rainfall using microwave links from commercial cellular communication networks, *Water Resources Research*, 47, Artn W12505; 10.1029/2010wr010350, 2011.
- Overeem, A. L., H.; Uijlenhoet, R.: Country-wide rainfall maps from cellular communication networks, *Proceedings of the National Academy of Sciences of the United States of America*, 110, 2741-2745, 10.1073/pnas.1217961110, 2013.
- Overeem, A. L., H.; Uijlenhoet, R.: Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network, *Atmospheric Measurement Techniques*, 9, 2425-2444, 10.5194/amt-9-2425-2016, 2016a.
- Overeem, A. L., H.; Uijlenhoet, R.: Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks, *Water Resources Research*, n/a-n/a, 10.1002/2016WR019412, 2016b.
- Polz, J., Chwala, C., Graf, M., and Kunstmann, H.: Rain event detection in commercial microwave link attenuation data using convolutional neural networks, *Atmos. Meas. Tech. Discuss.*, 2019, 1-22, 10.5194/amt-2019-412, 2019.
- Polz, J., Chwala, C., Graf, M., and Kunstmann, H.: Rain event detection in commercial microwave link attenuation data using convolutional neural networks, *Atmos. Meas. Tech.*, 13, 3835-3853, 10.5194/amt-13-3835-2020, 2020.
- Rahimi, A. R., Holt, A. R., Upton, G. J. G., and Cummings, R. J.: Use of dual-frequency microwave links for measuring path-averaged rainfall, *J Geophys Res-Atmos*, 108, <https://doi.org/10.1029/2002jd003202>, 2003.

- Rayitsfeld, A., Samuels, R., Zinevich, A., Hadar, U., and Alpert, P.: Comparison of two methodologies for long term rainfall monitoring using a commercial microwave communication system, *Atmos Res*, 104-105, 119-127, <https://doi.org/10.1016/j.atmosres.2011.08.011>, 2012.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, *Nature*, 566, 195-204, <https://doi.org/10.1038/s41586-019-0912-1>, 2019.
- Rincon, R. F., Bidwell, S. W., Jameson, A. R., and Thiele, O. W.: A multi-frequency, dual-polarization, microwave link for rainfall estimation, *Int Geosci Remote Se*, 402-404, 1996.
- Rincon, R. F., and Lang, R. H.: Microwave link dual-wavelength measurements of path-average attenuation for the estimation of drop size distributions and rainfall, *Ieee T Geosci Remote*, 40, 760-770, Pii S0196-2892(02)04595-3; Doi 10.1109/Tgrs.2002.1006324, 2002.
- Rincon, R. F. L., R. H.: Microwave link dual-wavelength measurements of path-average attenuation for the estimation of drop size distributions and rainfall, *Ieee T Geosci Remote*, 40, 760-770, Pii S0196-2892(02)04595-3; Doi 10.1109/Tgrs.2002.1006324, 2002.
- Rios Gaona, M. F., Overeem, A., Leijnse, H., and Uijlenhoet, R.: Measurement and interpolation uncertainties in rainfall maps from cellular communication networks, *Hydrol. Earth Syst. Sci.*, 19, 3571-3584, <https://doi.org/10.5194/hess-19-3571-2015>, 2015.
- Rios Gaona, M. F., Overeem, A., Raupach, T. H., Leijnse, H., and Uijlenhoet, R.: Rainfall retrieval with commercial microwave links in São Paulo, Brazil, *Atmospheric Measurement Techniques*, 11, 4465-4476, <https://doi.org/10.5194/amt-11-4465-2018>, 2018.
- Rios, G. M. F., Overeem, A., Raupach, T. H., Leijnse, H., and Uijlenhoet, R.: Rainfall retrieval with commercial microwave links in São Paulo, Brazil, *Atmospheric Measurement Techniques*, 1-21, <https://doi.org/10.5194/amt-2017-287>, 2017.

- Roversi, G., Alberoni, P. P., Fornasiero, A., and Porcù, F.: Commercial microwave links as a tool for operational rainfall monitoring in Northern Italy, *Atmos. Meas. Tech.*, 13, 5779-5797, 10.5194/amt-13-5779-2020, 2020.
- Schip, V. h., Overeem, A., Leijnse, H., Uijlenhoet, R., Meirink, J. F., and van Delden, A. J.: Rainfall measurement using cell phone links: classification of wet and dry periods using geostationary satellites, *Hydrological Sciences Journal*, 62, 1343-1353, <https://doi.org/10.1080/02626667.2017.1329588>, 2017.
- Schleiss, M., and Berne, A.: Identification of Dry and Rainy Periods Using Telecommunication Microwave Links, *IEEE Geoscience and Remote Sensing Letters*, 7, 611-615, <https://doi.org/10.1109/LGRS.2010.2043052>, 2010.
- Schleiss, M., Rieckermann, J., and Berne, A.: Quantification and Modeling of Wet-Antenna Attenuation for Commercial Microwave Links, *IEEE Geoscience and Remote Sensing Letters*, 10, 1195-1199, <https://doi.org/10.1109/LGRS.2012.2236074>, 2013.
- Seed, A., Leahy, C., Duthie, E., and Chumchean, S.: Rainfields : The Australian Bureau of Meteorology System for Quantitative Precipitation Estimation, and It's Use in Hydrological Modelling, *Proceedings of Water Down Under 2008*, Modbury, SA, 2008, 661-670,
- Sene, K.: *Meteorological Forecasting*, Dordrecht: Springer Netherlands, Dordrecht, 67-99 pp., 2010.
- Sene, K.: *Precipitation Measurement*, 33-70 pp., 2013.
- Sevruk, B.: WMO questionnaire on recording precipitation gauge measurements, *Water Sci. Technol.*, 42, 10, 2002.
- Seybold, J. S.: *Introduction to RF propagation*, John Wiley distributor, 2005.
- Shen, C.: A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists, *Water Resources Research*, 54, 8558-8593, <https://doi.org/10.1029/2018WR022643>, 2018.
- Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F. J., Ganguly, S., Hsu, K., Kifer, D., Fang, Z., Fang, K., Li, D., Li, X., and Tsai, W. P.: HESS Opinions: Incubating deep-learning-powered hydrologic science advances as a community, *Hydrology and*

- Earth System Sciences, 22, 5639-5656, <https://doi.org/10.5194/hess-22-5639-2018>, 2018.
- Shi, X., Gao, Z., Lausen, L., Wang, H., and Yeung, D. Y.: Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model, *Adv. Neural. Inf. Process. Syst*, 30, 5617-5627, 2014.
- Skoulikaris, C., Anagnostopoulou, C., and Lazoglou, G.: Hydrological Modeling Response to Climate Model Spatial Analysis of a South Eastern Europe International Basin, *Climate*, 8, 1, 2020.
- Smiatek, G., Keis, F., Chwala, C., Fersch, B., and Kunstmann, H.: Potential of commercial microwave link network derived rainfall for river runoff simulations, *Environmental Research Letters*, 12, 034026, <https://doi.org/10.1088/1748-9326/aa5f46>, 2017.
- Sohail Afzal, M., Shah, S. H. M., Cheema, M. J. M., and Ahmad, R.: Real time rainfall estimation using microwave signals of cellular communication networks: a case study of Faisalabad, Pakistan, *Hydrology and Earth System Sciences Discussions*, 1-20, <https://doi.org/10.5194/hess-2017-740>, 2018.
- Stratton, J.: The effect of rain and fog on the propagation of very short waves, *P. IRE*, 18, 10, <https://doi.org/10.1109/JRPROC.1930.222101>, 1930.
- Testik, F. Y., and Gebremichael, M.: *Rainfall: state of the science*, Washington: WILEY, Washington, 2013.
- Tokay, A., Petersen, W. A., Gatlin, P., and Wingo, M.: Comparison of Raindrop Size Distribution Measurements by Collocated Disdrometers, *Journal of Atmospheric and Oceanic Technology*, 30, 1672-1690, <https://doi.org/10.1175/jtech-d-12-00163.1>, 2013.
- Tromel, S., Ziegert, M., Ryzhkov, A. V., Chwala, C., and Simmer, C.: Using Microwave Backhaul Links to Optimize the Performance of Algorithms for Rainfall Estimation and Attenuation Correction, *Journal of Atmospheric and Oceanic Technology*, 31, 1748-1760, <https://doi.org/10.1175/Jtech-D-14-00016.1>, 2014.
- Uijlenhoet, R., Overeem, A., and Leijnse, H.: Opportunistic remote sensing of rainfall using microwave links from cellular communication networks, *Wiley Interdisciplinary Reviews: Water*, 5, e1289, <https://doi.org/10.1002/wat2.1289>, 2018.

- Upton, G. J. G., Holt, A. R., Cummings, R. J., Rahimi, A. R., and Goddard, J. W. F.: Microwave links: The future for urban rainfall measurement?, *Atmos Res*, 77, 300-312, <https://doi.org/10.1016/j.atmosres.2004.10.009>, 2005.
- Valtr, P., Fencel, M., and Bareš, V.: Excess Attenuation Caused by Antenna Wetting of Terrestrial Microwave Links at 32 GHz, *IEEE Antennas and Wireless Propagation Letters*, 18, 1636-1640, <https://doi.org/10.1109/LAWP.2019.2925455>, 2019.
- van Leth, T. C., Overeem, A., Leijnse, H., and Uijlenhoet, R.: A measurement campaign to assess sources of error in microwave link rainfall estimation, *Atmospheric Measurement Techniques*, 11, 4645-4669, <https://doi.org/10.5194/amt-11-4645-2018>, 2018.
- Vande, H. C.: *Light Scattering by Small Particles*, Dover, Mineola, N.Y, 470, 1981.
- vanhet, S. T. I., Overeem, A., Leijnse, H., Uijlenhoet, R., Meirink, J. F., and van Delden, A. J.: Rainfall measurement using cell phone links: classification of wet and dry periods using geostationary satellites, *Hydrological Sciences Journal*, 62, 1343-1353, 10.1080/02626667.2017.1329588, 2017.
- Wang, Z. S., M.; Jaffrain, J.; Berne, A.; Rieckermann, J.: Using Markov switching models to infer dry and rainy periods from telecommunication microwave link signals, *Atmospheric Measurement Techniques*, 5, 1847-1859, 10.5194/amt-5-1847-2012, 2012.
- Watson, J., and Challinor, A.: The relative importance of rainfall, temperature and yield data for a regional-scale crop model, *Agricultural and Forest Meteorology*, 170, 47-57, <https://doi.org/10.1016/j.agrformet.2012.08.001>, 2013.
- Watson, R. J., and Hodges, D. D.: Estimation of Rainfall Rate from Terrestrial Microwave Link Measurements, 2009 Ieee International Geoscience and Remote Sensing Symposium, Vols 1-5, 1565-1568, <https://doi.org/10.1109/Igarss.2009.5417770>, 2009.
- Wood, S. J., Jones, D. A., and Moore, R. J.: Accuracy of rainfall measurement for scales of hydrological interest, *Hydrol. Earth Syst. Sci.*, 4, 523-530, 2000.
- Zinevich, A., Alpert, P., and Messer, H.: Estimation of rainfall fields using commercial microwave communication networks of variable density, *Adv Water Resour*, 31, 1470-1480, <https://doi.org/10.1016/j.advwatres.2008.03.003>, 2008.

Zinevich, A., Messer, H., and Alpert, P.: Prediction of rainfall intensity measurement errors using commercial microwave communication links, *Atmos. Meas. Tech.*, 3, 1385-1402, 10.5194/amt-3-1385-2010, 2010.