Improving radar rainfall forecasting for hydrological applications

Mayra Codo de Oliveira

Department of Civil Engineering

University of Bristol

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Engineering

April 2020

Para minha mãe, Yara

Abstract

Rainfall forecasting plays an essential role to forecast extreme precipitation events for real-time flood forecasting. Uncertainties from both radar rainfall estimations and forecasts propagate into runoff modelling and impact the ability of an event to be correctly forecasted. Weather radars provide measurements with high temporal and spatial resolutions necessary for hydrological applications; however, radar rainfall is subject to different sources of uncertainties. Short-term radar-based rainfall forecasts (known as nowcasting) are subjected to uncertainties deriving from radar rainfall estimations, uncertainties related to the nowcasting model and uncertainties related to the temporal evolution of the precipitation field. This thesis proposes new methods to quantify and account for rainfall forecast uncertainties and assesses how some of these uncertainties propagate into hydrological modelling in small urban areas and large river catchments. Uncertainties related to both radar rainfall estimations and the temporal evolution of velocity fields were studied.

The study area focused on the north of England, where data from three weather radars and more than 200 rain gauge stations were available. A radar rainfall estimation ensemble generator was implemented to model the uncertainties in radar rainfall. The radar rainfall estimation ensembles were computed based on comparing historical weather radar rainfall estimations and rain gauge measurements. The radar rainfall estimation ensembles were used to drive a radar-based forecasting model to produce ensemble rainfall forecasts. These radar estimation ensemble forecasts were compared against the forecasts produced with a stochastic ensemble generator. The results showed an improvement in the forecasting ability of the radar rainfall estimation ensembles during the first hour of the forecasting time. For flow forecasting applications, the radar rainfall estimation ensembles overperformed the stochastic ones in the first forecasted hour and could reproduce flow peaks more accurately. To assess uncertainties related to the temporal evolution of velocity fields, a new methodology to generate ensembles using rainfall advection fields from a time window that goes from 10 min up to 2 hr before the forecast's start was proposed. The results show that the extra information provided by the rainfall velocity fields from the previous hours can improve the rainfall forecast skill up to 3h ahead. The forecasts were used to predict sewer flows in an urban area, and the results showed that these forecasts provide improvement compared to deterministic forecasts. Merging radar rainfall with rain gauge measurements was studied to improve rainfall estimations and tests to assess if sub-hourly temporal resolutions could be used to merge radar and rain gauge data. A radar – rain gauge

merging method combines the spatial distribution of precipitation from weather radars with the accuracy of rain gauge measurements to produce a product with the best from both information sources. However, using radar-rain gauge merging techniques to produce rainfall forecasts is a challenge because the temporal correlation of the radar rainfall advection field is lost. A new rainfall forecasting method that merges radar and rain gauge rainfall using kriging with external drift (KED) and using advection velocity fields from original radar data was developed. The results showed that this method produces a better rainfall forecast than using KED rainfall or radar rainfall.

The methods used to account for uncertainties in radar estimations had a more substantial influence in improving the forecasting skill up to 1 hour lead time; during this period, radar estimations are the main errors sources in nowcasting. Ensembles produced by varying the rainfall velocity fields showed improved estimations compared to a stochastic ensemble generator at longer lead times.

Acknowledgements

Firstly, I would like to my deepest gratitude to my supervisor, Dr Miguel A. Rico-Ramirez, for the invaluable expertise, guidance, motivation, and comprehension at all times during my PhD. I could not have done this without his help, understanding and support.

Thank you to my friend in Woodland Road that helped me and motivated me, particularly during the PhD's early times. I would like to thank N. Nanding for the assistance when needed. I would especially like to thank Francesca Cecinati for her inspiration, for all the help, for being a friend and for making Conference trips more enjoyable.

I would like to thank Jacob Taekke for always being there, helping me through all these years and making me smile during the most challenging moments. I would like to thank Leon, that arrived in the middle of my PhD. You added an extra degree of challenge to this research, but you are also capable of making all my days bright, even the cloudiest ones.

Thank you to my father, Gilberto Maringoni, for the constant incentive and great drawings and guidance about research life. To my best friend and sister, Natalia Codo, that even from across the ocean is a constant presence in my life. I would also like to thank all my family and friends for continuous motivation and so needed distractions.

The research and development of this thesis were only made possible by the support of the "Ciência Sem Fronteiras" program, an initiative of the Brazilian government whose targets were "to promote the consolidation, expansion and internationalisation of science and technology, innovation and Brazilian competitiveness through exchange and international mobility", as established by its guidelines. Several State bodies have made joint efforts in this regard, in particular the Ministries of Science, Technology and Innovation (MCTI) and the Ministry of Education (MEC), through their development institutions: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes). The program was developed from 2011 to 2017. I wish to acknowledge my profound gratitude to all who made it possible to accomplish my PhD studies in the United Kingdom.

Finally, I would like to acknowledge the UK MetOffice, Environment Agency, Yorkshire Water Services Ltd and the British Atmospheric Data Centre for providing the hydrological and meteorological datasets and models used for this research.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award.

Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others is indicated as such.

Any views expressed in the dissertation are those of the author.

SIGNED: .....

DATE:.....

Publications

Codo, M. and Rico-Ramirez, M. A. (2018) 'Ensemble Radar-Based Rainfall Forecasts for Urban Hydrological Applications', Geosciences, 8(8), p. 297. doi: 10.3390/geosciences8080297

Codo, M. and Rico-Ramírez, M. A. (2015) 'Assessing the uncertainties related to the temporal evolution of the rainfall velocity field in radar nowcasting', in *37th Conference on Radar Meteorology Meeting*. Norman: American Meteorological Society.

Codo, M. and Rico-Ramírez, M. A. (2016) 'Short-term ensemble radar rainfall forecasts for hydrological applications', in *American Geophysical Union*. San Francisco.

Contents

Chapter 1. Introduction 1

1.1 Background and motivation 1

1.2 Research questions and objectives 3

1.3 Thesis layout 4

Chapter 2. Rainfall observation and forecast 6

2.1 Rain gauges 8

2.1.1 Tipping Bucket Rain Gauge 12

2.1.2 Weighing Rain Gauges 15

2.2 Weather radar 17

2.2.1 Radar uncertainties 21

2.3 Radar and rain gauge merging 29

2.4 Short-term rainfall forecasts (radar nowcasts) 33

2.5 Concluding comments 36

Chapter 3. Datasets and study areas 38

3.1 Introduction 38

3.2 Rainfall data 38

3.2.1 Rain gauge data 38

3.2.2 Weather radar data 40

3.3 Nowcasting model 42

3.3.1 Nowcasting component of the STEPS model 43

3.4 Study areas 45

3.4.1 Urban catchment - Ilkley 46

3.4.2 River catchments 47

Chapter 4. Radar rainfall ensembles to represent uncertainties in nowcasts 54

4.1 Introduction 54

4.2 Methodology 57

4.2.1 Analysis using radar rainfall estimation ensembles to generate short-term forecast ... 58
4.3 Rainfall Forecasting ... 59
4.4 Hydrological application in flow predictions ... 65
4.4.1 Urban catchment ... 65
4.4.2 River Catchments ... 71
4.5 Conclusion ... 71
Chapter 5. Ensemble forecasts based on the temporal variation of the velocity field ... 82
5.1 Introduction ... 82
5.2 Methodology ... 83
5.3 Results ... 87
5.3.1 Analysis using perturbations on the temporal evolution of velocity fields to generate short-term forecast ... 87
5.3.2 Hydrological application in flow predictions in Ilkley ... 92
5.4 Conclusion ... 98
Chapter 6. Radar and rain gauge merging ... 100
6.1.1 Introduction ... 100
6.2 Methodology ... 102
6.2.1 Ordinary kriging (OK) ... 103
6.2.2 Kriging with external drift (KED) ... 104
6.2.3 Variogram estimation ... 105
6.3 Results ... 106
6.4 Conclusion ... 120
Chapter 7. Radar and rain gauge kriging with forced velocity fields to generate ensemble forecasts 122
7.1 Introduction ... 122
7.2 Methodology ... 123
7.3 Results ... 126
7.3.1 Velocity fields for deterministic and KED forecasts ... 126

7.4 Conclusion 147

Chapter 8. Conclusion 149

Chapter 9. References 155

List of Tables

Table 2-1. Coefficient a and b for some typical Z-R relationships (Ulbrich and Lee, 1999). 20
Table 2-2. Order of magnitude of hydrometeors uncertainties due to attenuation at 5cm wavelength (Collier, 1996). 29
Table 3-1. Official recent national rain gauge network (Nanding, 2016). 39
Table 3-2. River catchment's area 48
Table 3-3. River catchment's mean and peak flows 48
Table 3-4 River catchment´s land-use 49
Table 3-5. River catchment´s permeability 49
Table 3-6. PDM model parameters (Moore, 2007) 53
Table 4-1. Event start dates, duration, peak measured flow, maximum average rainfall and storm types (S—stratiform and C—convective) for the Ilkley urban catchment. 58
Table 4-2. Contingency table for ROC curve 60
Table 4-3. RMSE mean for RE and STEPS ensembles for 0–1 hr after the forecast's start. 68
Table 4-4. NSE Coefficient for calibration and validation of the PDM for the river catchments studied using radar and gauge 71
Table 4-5. Events dates, duration peak measured flow, maximum average rainfall, average GOF scores for RE and STEPS ensembles and storm types (S – stratiform and C – convective) for the Arnford catchment. 76
Table 4-6. Events dates, duration peak measured flow, maximum average rainfall, average GOF scores and storm types (S – stratiform and C – convective) for the Brigflatts catchment. 76
Table 4-7. Events dates, duration peak measured flow, maximum average rainfall, average GOF scores and storm types (S – stratiform and C – convective) for the Henthorn catchment. 77
Table 4-8. Events dates, duration peak measured flow, maximum average rainfall, average GOF scores and storm types (S – stratiform and C – convevtive) for the Killington catchment. 77
Table 5-1. Event start dates, duration and measured flow peaks. 87
Table 6-1. Event dates, maximum average rainfall and storm types (S—stratiform and C—convective). 102
Table 6-2 KED correlation between observed and predicted rainfall and MAE for each event for temporal accumulation of 15 min, 30 min and 60 min. 117
Table 7-1. Contingency table (Jolliffe and Stephenson, 2011) 125

List of Figures

Figure 2-1 'Snowdon' rain gauge and graduated cylinder (Strangeways, 2010). 10
Figure 2-2 (a) Nipher Shield. (b) Alter Shield. (c) Tretakov screen (Strangeways, 2006) 12
Figure 2-3 Pit rain gauge (Strangeways, 2010). 12
Figure 2-4 Tipping bucket (Strangeways, 2010). 13
Figure 2-5 Schematic drawing of tipping bucket rain gauge operating (Netatmo Rain Gauge | Weather Station Tools, 2013). 13
Figure 2-6 Example of a blocked TBR (Source: Personal collection). 15
Figure 2-7 Schematic drawing of a Weighing Rain Gauge (Strangeways, 2006) 16
Figure 2-8 Weighing rain gauge 17
Figure 2-9 Schematic of how weather radars work (Source: Personal collection). 18
Figure 2-10. Illustration of uncertainties related to radar rainfall estimation. Where: 1. Overshooting of precipitation, 2. Sea clutter, 3. Wind turbines, 4. Bio-scatter, 5. Radar attenuation, 6. Ground clutter, 7. Variation of the vertical profile of reflectivity and 8. Occultation (Source: Personal collection). 22
Figure 2-11. Idealized vertical reflectivity profile (Harrison, Driscoll and Kitchen, 2000) 25
Figure 2-12. VPR during stratiform precipitation. The melting of snowflakes takes place around 1.8 km (bright band). Rainfall is present below the bright band. Reflectivity above the bright band is due to the presence of snowflakes, and below it due to raindrops (Rico-Ramirez et al., 2007). 26
Figure 2-13. VPR during convective precipitation. The BB does not occur at a specific height. It can occur over several kilometres (Rico-Ramirez et al., 2007). 26
Figure 2-14. Schematic representation of loss of information content in relation to lead time. The dashed, doted, and solid lines represent, respectively, NWP models, nowcasting models and theoretical limit of predictability (Golding et al., 1998). 36
Figure 3-1 Map of the study domain with the rain gauge locations (red circles). The region located in the North of England covers 256 x 256 $km^{2,}$ and the region is also under the coverage of three MetOffice weather radars. 40
Figure 3-2 Location of weather radars across the British Isles(Scottish Flood Forecasting Service, 2014). 41
Figure 3-3. Location of the urban catchment (7), the river catchments (1-6) and rain gauges' positions (blue circles) in the study region. 45
Figure 3-4 – Location of flow monitors (1 to 7) and rain gauges (RG02 to RG05) in the sewer system (Liguori et al., 2012). 46

Figure 3-5 – Monitoring network (Liguori et al., 2012)..... 47
Figure 3-6 The PDM rainfall-runoff model (Moore, 2007) 52
Figure 4-1. RE ensemble forecasts at time t = 30 min (**a**,**d**), t = 60 min (**b**,**e**), t = 90 min (**c**,**f**) starting on 1 August 2008 at 02:00. Each probabilistic forecast is formed by 25 ensemble members that are valid at the simultaneously. 59
Figure 4-2.ROC curves for event on 7 July 2008 starting at 15:00 for RE and STEPS ensembles, respectively and thresholds (th) equal to 0.1 mm/hr(a, d), 1.0 mm/hr(b, e) and 3.0 mm/hr(c, f). 61
Figure 4-3. ROC curves for RE and STEPS ensembles for event on 1 August 2008 starting at 01:30 for RE and STEPS ensembles, respectively and thresholds (th) equal to 0.1 mm/hr(a, d), 1.0 mm/hr(b, e) and 3.0 mm/hr(c, f). 61
Figure 4-4. ROC curves for RE and STEPS ensembles for event on 4 October 2008 starting at 15:00 for RE and STEPS ensembles, respectively and thresholds (th) equal to 0.1 mm/hr(a, d), 1.0 mm/hr(b, e) and 3.0 mm/hr(c, f). 62
Figure 4-5. ROC curves for RE and STEPS ensembles for event on 14 October 2008 starting at 15:00 for RE and STEPS ensembles, respectively and thresholds (th) equal to 0.1 mm/hr(a, d), 1.0 mm/hr(b, e) and 3.0 mm/hr(c, f). 62
Figure 4-6. Area beneath ROC curves for all the events using RE ensembles and STEPS ensembles and threshold 0.1 mm/hr. 63
Figure 4-7. Area beneath ROC curves for all the events using RE ensembles and STEPS ensembles and threshold 1 mm/hr..... 63
Figure 4-8. Area beneath ROC curves for all the events using RE ensembles and STEPS ensembles and threshold 3 mm/hr..... 64
Figure 4-9. Measured flows (Qm) and radar simulated flows (Q_R) for 04 and 05 of October 2008..... 66
Figure 4-10. Measured flows (Qm) and gauge simulated flows (Q_G) for 04 and 05 of October 2008. .. 66
Figure 4-11. Flows (a) and peak flows (b) for events on 07 July 2008 starting at 15:00. $Q_{F\ Ens\ Radar}$, $Q_{F\ Ens\ STEPS}$, $Q_{F\ Det}$, Q_m, Q_G and Q_R are, respectively, forecasted RE ensemble flows, forecasted STEPS ensemble flows, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow. 67
Figure 4-12. Flows (a) and peak flows (b) for events on 09 November 2008 starting at 14:00. $Q_{F\ Ens\ Radar}$, $Q_{F\ Ens\ STEPS}$, $Q_{F\ Det}$, Q_m, Q_G and Q_R are, respectively, forecasted RE ensemble flows, forecasted STEPS ensemble flows, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow. 67
Figure 4-13. Flows (a) and peak flows (b) peak flows for events on 1 August 2008 starting at 02:30 and flows (c) and peak flows (d) for events on 1 August 2008 at 03:00. $Q_{F\ Ens\ Radar}$, $Q_{F\ Ens\ STEPS}$, $Q_{F\ Det}$, Q_m, Q_G

and Q_R are, respectively, forecasted STEPS ensemble flows, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow. 68
Figure 4-14. RMSE for RE and STEPS ensembles and 0–1 hr after the start of the forecast. 69
Figure 4-15. Calibration and Validation of Arnford catchment, where Qobs represents the observed flow and Qsim represents the simulated flow using radar or gauge data. 72
Figure 4-16. Calibration and Validation of Henthorn catchment, where Qobs represents the observed flow and Qsim represents the simulated flow using radar or gauge data. 73
Figure 4-17. Calibration and Validation of Killington catchment, where Qobs represents the observed flow and Qsim represents the simulated flow using radar or gauge data. 74
Figure 4-18. Flows for events on 14/10/2008 starting at 16:00 at Arnford catchment, 04/10/2008 starting at 14:00 at Henthorn catchment, and 06/09/2008 starting at 03:00 at Killington catchment. Q_{ens}, Q_{obs}, $Q_{det,}$ Q_{rad} and Q_{gau} are, respectively, forecasted RE ensemble flows, measured flow, forecasted deterministic flow, estimated radar flow and estimated gauge flow. 78
Figure 4-19. Flows for events on 14/10/2008 starting at 16:00 at Arnford catchment, 04/10/2008 starting at 14:00 at Henthorn catchment, and 06/09/2008 starting at 03:00 at Killington catchment. Q_{ens}, Q_{obs}, Q_{det}, Q_{rad} and Q_{gau} are, respectively, forecasted STEPS ensemble flows, measured flow, forecasted deterministic flow, estimated radar flow and estimated gauge flow. 79
Figure 4-20. RMSE for RE ensembles at Arnford, Brigflatts, Henthorn and Killington catchments. $Radar_{ens}$, Det, $Radar_{data,}$ and $Gauge_{data}$ are, respectively, RE ensemble forecast, deterministic forecast, estimated radar flow and estimated gauge flow. 80
Figure 4-21. RMSE for STEPS ensembles at Arnford, Brigflatts, Henthorn and Killington catchments. $Radar_{ens}$, Det, $Radar_{data,}$ and $Gauge_{data}$ are, respectively, STEPS ensemble forecast, deterministic forecast, estimated radar flow and estimated gauge flow. 80
Figure 5-1. Average advection velocity fields and their standard deviation calculated over a 3 hr window. The blue area shows the regions where there was rainfall during this period. 84
Figure 5-2. Example of velocity fields used to generate ensemble forecasts with a time window of 30 min and starting time at t= 00. 86
Figure 5-3. Radar image (top left) at 21:00 on 04/10/2008, deterministic forecast (top right) and two ensembles with 60 min lead time (bottom). 88
Figure 5-4. Radar image(top left)) at 21:00 on 04/10/2008, deterministic forecast (top right) and two ensembles (bottom) with 180 min lead time. 89
Figure 5-5. ROC curves for the events on 02/09/2008 and 14/10/2008 with 30 min, 1-, 2- and 3 hr lead time and time window of 30, 60, 90 and 120 min. (FAR – False Alarm Rate; HR – Hit Rate) 90

Figure 5-6. – Area beneath the ROC curve for a rainfall threshold of 0.1 mm/hr, time window of 30-, 60-, 90- and 120 min and forecasting lead times of 30-, 60-, 120- and 180 min. (tb refers to the time window used to produce the advection fields) 91
Figure 5-7. – Area beneath the ROC curve for a rainfall threshold of 1.0 mm/hr, time window of 30-, 60-, 90- and 120 min and forecasting lead times of 30-, 60-, 120- and 180 min. 91
Figure 5-8. – Area beneath the ROC curve for a rainfall threshold of 3.0 mm/hr, time window of 30-, 60-, 90- and 120 min and forecasting lead times of 30-, 60-, 120- and 180 min. 92
Figure 5-9. Flows for event on 20 August 2008 starting at 20:30. Q_{EnsF}, Q_{DetF}, Q_m, $Q_{G,}$ and Q_R are, respectively, forecasted ensemble flows (with time-back window of 30-, 60, 90 and 120 min), forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow. 94
Figure 5-10. Flows for events on 04 October 2008 starting at 23:30. Q_{EnsF}, Q_{DetF}, Q_m, $Q_{G,}$ and Q_R are, respectively, forecasted ensemble flows (with time-back window of 30-, 60, 90 and 120 min), forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow. 95
Figure 5-11. Flows for events on 20 August 2008 starting at 20:30 and 04 October 2008 starting at 23:30. Q_{EnsF}, Q_{DetF}, Q_m, Q_G and Q_R are, respectively, forecasted ensemble flows (STEPS ensembles), forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow. 96
Figure 5-12. Peak flow boxplot for rainfall advection field forecasts with a time-back window of 30-, 60, 90 and 120 min, on 20 August 2008 starting at 20:30. Q_m, Q_G and Q_R, Q_{DetF} are, respectively, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow. 97
Figure 5-13. Peak flow boxplot for rainfall advection field forecasts with a time-back window of 30-, 60, 90 and 120 min, on 04 October 2008 starting at 23:30. Q_m, Q_G and Q_R, Q_{DetF} are, respectively, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow. 98
Figure 6-1. Variogram example with representation of range, sill and nugget. 106
Figure 6-2. Variograms calculated using rain gauge (left) data and radar (right) data for event on 15/01/2008 at 19:00 for accumulation periods of 15 min, 30 min and 60 min. 107
Figure 6-3. Rainfall for hourly accumulated radar data in mm/hr, KED and OK on 15/01/2008 at 19:00 and 15-, 30- and 60 min accumulation. 108
Figure 6-4. Variograms calculated using rain gauge (left) data and radar (right) data for event on 30/04/2008 at 10:00 for accumulation periods of 15 min, 30 min and 60 min. 109
Figure 6-5. Rainfall for hourly accumulated radar data in mm/hr, KED and OK on 30/04/2008 at 10:00 and 15-, 30- and 60 min accumulation. 110
Figure 6-6. Variograms calculated using rain gauge data and radar data for event on 04/10/2008 at 16:00 for accumulation periods of 15 min, 30 min and 60 min. 111

Figure 6-7. Rainfall for hourly accumulated radar data in mm/hr, KED and OK on 04/10/2008 at 16:00 and 15-, 30- and 60 min accumulation. 112
Figure 6-8. Variograms calculated using rain gauge data and radar data for event on 14/10/2008 at 18:00 for accumulation periods of 15 min, 30 min and 60 min. 113
Figure 6-9. Rainfall for hourly accumulated radar data in mm/hr, KED and OK on 14/10/2008 at 18:00 and 15-, 30- and 60 min accumulation. 114
Figure 6-10. KED cross-validation for events on 15/01/2008 and 30/04/2008 with 15-, 30- and 60-min accumulation. 115
Figure 6-11. KED cross-validation for events on 04/10/2008 and 14/10/2008 with 15-, 30- and 60-min accumulation. 116
Figure 6-12 Correlation between observed and predicted rainfall for all the events with 15-, 30- and 60-min accumulation. 118
Figure 6-13 Root mean square error and mean absolute error for all the events with 15-, 30- and 60-min accumulation. 118
Figure 6-14 Correlation between observed and predicted rainfall for all the events with 15-, 30- and 60-min accumulation and rainfall thresholds of 0.1-, 1.0-, and 3.0 mm/h. 119
Figure 6-15 Root mean square error for all the events with 15-, 30- and 60-min accumulation and rainfall thresholds of 0.1-, 1.0-, and 3.0 mm/h 119
Figure 7-1. KED rainfall estimation, deterministic forecast, KED forecast and velocity fields for the event on 15/01/2008 with the forecast initial time at 09:15 for 15 min accumulation. 128
Figure 7-2. KED rainfall estimation, deterministic forecast, KED forecast and velocity fields for the event on 14/10/2008 with the forecast initial time at 13:15 for 15 min accumulation. 129
Figure 7-3. KED rainfall estimation, deterministic forecast and KED forecast for the event on 30/04/2008 with the forecast initial time at 00:00 for 15 min accumulation. 131
Figure 7-4. KED rainfall estimation, deterministic forecast and KED forecast for the event on 30/04/2008 with the forecast initial time at 00:00 for 30min accumulation. 132
Figure 7-5. KED rainfall estimation, deterministic forecast and KED forecast for the event on 30/04/2008 with the forecast initial time at 00:00 for 60 min accumulation. 133
Figure 7-6. RMSE and Bias for the event on 30/04/2008 at 15 min accumulation 135
Figure 7-7. RMSE and Bias for the event on 30/04/2008 at 30 min accumulation 135
Figure 7-8. RMSE and Bias for the event on 30/04/2008 at 60 min accumulation 135
Figure 7-9. KED rainfall estimation, deterministic forecast and KED forecast for the event on 04/10/2008 with the forecast initial time at 17:00 for 15 min. 136

Figure 7-10. KED rainfall estimation, deterministic forecast and KED forecast for the event on 04/10/2008 with the forecast initial time at 17:00 for 30min. 137
Figure 7-11. KED rainfall estimation, deterministic forecast and KED forecast for the event on 04/10/2008 with the forecast initial time at 17:00 for 60 min. 138
Figure 7-12. RMSE and Bias for the event on 04/10/2008 at 15 min accumulation 140
Figure 7-13. RMSE and Bias for the event on 04/10/2008 at 30 min accumulation 140
Figure 7-14. RMSE and Bias for the event on 04/10/2008 at 60 min accumulation 140
Figure 7-15. RMSE and bias boxplot for threshold of 0.1 mm/hr and lead time of 1-, 2- and 3 hr..... 141
Figure 7-16. RMSE and bias boxplot for a 3.0 mm/hr threshold and lead time of 1-, 2- and 3 hr. 142
Figure 7-17. Critical Success Index (CSI) boxplot for threshold of 0.1 mm/hr and 3.0 mm/hr, and lead time of 1-, 2- and 3 hr. 144
Figure 7-18. Hit rate and false alarm ratio boxplot for threshold of 0.1 mm/hr and lead time of 1-, 2- and 3 hr. 145
Figure 7-19. Hit rate and false alarm ratio boxplot for threshold of 3.0 mm/hr and lead time of 1-, 2- and 3 hr. 146

List of Acronyms

AOD	Above Ordnance Datum
AP	Anomalous propagation
AR-2	Second-order autoregressive process
BADC	British atmospheric data centre
BB	Bright band
CED	Co-kriging with external drift
CEH	Centre for Ecology and Hydrology
CK	Co-kriging
COTREC	Continuity of tracking radar echoes by correlation vectors
CSI	Critical success index
DSD	Drop size distribution
EA	Environmental agency
FAR	False alarm rate
FLS	Fuzzy logic system
FSR	Frequency shift reflector
GANDOLF	Generating Advanced Nowcasts for Deployment in Operational Land-based Forecasts
GOF	Goodness of fit
GPM	Global precipitation measurement
HR	Hit rate
JAXA	Japan aerospace exploration agency
KED	Kriging with external drift
KEDFV	KED with forced velocities fields
MAE	Mean absolute error
MFB	Mean field bias correction
MIDAS	Met office integrated archive system
NASA	National Aeronautics and Space Administration
NN	Neural network
NORA	Nowcasting of orographic rainfall by means of analogues
NRFA	National river flow archive
NSE	Nash-Sutcliffe efficiency
NWP	Numerical weather prediction
OK	Ordinary kriging
PDM	Probability distributed model
PET	Potential evapotranspiration
POD	Probability of Detection
QPE	Quantitative precipitation estimations
QPF	Quantitative precipitation forecasts
RE ensembles	Radar rainfall estimation ensembles
RG	Rain gauge

RMSE	Root mean square error
ROC	Receiving operating characteristic
SAMR	Statistical advective method radar
SCE-UA	Shuffled complex evolution developed at the University of Arizona
SCIT	Storm cell identification and tracking
SLA	Service level agreement
STEPS	Short-term ensemble prediction system
TBM	Turning band method
TBR	Tipping bucket rain gauge
TITAN	Thunderstorm Identification, Tracking, Analysis, and Nowcasting
TREC	Tracking radar echo with correlation
TRMM	Tropical rainfall measuring mission
VPR	Vertical profile of reflectivity
WFD	Water framework directive
WR	Weighing rain gauges
Z_{DR}	Differential reflectivity

Chapter 1. Introduction

1.1 Background and motivation

Flooding is a recurrent natural disaster around the world, and climate simulations for the next 100 years suggest changes in the rainfall regime, causing more floods and droughts problems all over the world due to the increase of CO_2 emissions (Haines *et al.*, 2006; Lau *et al.*, 2010). The increase of CO_2 levels in the atmosphere and global climate changes are escalating flooding incidences, sewer overflows, and droughts. Along with climate change, variations in the land-use also contribute to the fact that floods events have become more frequent and intense over the past century in many regions worldwide (Haines et al., 2006; Lau et al. Adopting measures to adapt to a changing climate, alongside policies to mitigate CO_2 emissions are more important than it has ever been (Haines *et al.*, 2006). Large floods events are life-threatening phenomena, have a significant impact on the economy and infrastructure, with damages lasting much longer than the event's duration. As floods are becoming more frequent and intense, there is a need to increase protection measures to minimise the losses. There are two possible strategies to manage floods hazards. The first of them is non-structural measures, based on issuing floods warnings using real-time flood forecasting systems. The second one, a much more expensive approach, is structural measures to protect from an estimated flood level (Few, 2003; Brocca, Melone and Moramarco, 2011). Nevertheless, more intense floods that can overcome the protective structures will eventually occur. Alternatives measures using floods forecasting and warning systems should also be in place, as forecasting floods hours ahead can allow a timely emergency response to take place (Golding, 2009) and help to deploy flood barriers in critical locations.

Rainfall-runoff models have a vital function when issuing flood warnings (Brocca, Melone and Moramarco, 2011). Over the last couple of decades, much research has been done to increase forecasting rainfall reliability (Pierce *et al.*, 2000; Bowler, Pierce and Seed, 2006; Cloke and Pappenberger, 2009; Foresti *et al.*, 2013). The need for high-quality rainfall data with high temporal (e.g., 5 min or lower) and spatial (e.g., 1 km or lower) resolutions to forecast flows are often emphasised (Berne and Krajewski, 2013; Thorndahl *et al.*, 2016). Weather radars are able to estimate rainfall based on back-scattered radiation. It has been used for decades in meteorological applications and has the advantage of producing rainfall estimations with high temporal and spatial resolution. Radar resolution has a particular impact on localised events with high rainfall intensity (Zhu, Xuan and Cluckie, 2014). Weather radars are capable of estimating rainfall in real-time and provide high-resolution data that is required. However, radar estimations have shown to be subjected to a number of uncertainties that affect the reliability of flow predictions. Probabilistic radar estimations can be used to quantify residual

uncertainties in weather radars, as proposed by Germann *et al.* (2009). This method uses rain gauge historical data set as a reference to produce ensembles. Radar and rain gauge merging have been successfully used for years to improve rainfall estimations. However, it needs a dense rain gauge network to produce skilful estimations (Jewell and Gaussiat, 2015). Kriging with external drift is a well-performing merging method but using it in short-term forecasts is challenging as there is little research on KED products with a high temporal resolution, and it lacks a temporal correlation between time-steps.

Nowcasts are short-term forecasts (up to 6 hours lead time) and have been used for years in real-time applications. Nowcasting models can produce high resolution, very short-term forecasts; however, they quickly lose their skill and are highly dependent on the quality of radar data available (Foresti *et al.*, 2013). Besides uncertainties from rainfall estimation, rainfall forecast is also subjected to uncertainties inherent to the model (Bowler, Pierce and Seed, 2004; Foresti *et al.*, 2013). NWP models can extend the forecasting lead-time and produce forecasts with 1.5 km resolution (Simonin, Ballard and Li, 2014; Dance *et al.*, 2019). Combining nowcast and NWP forecasts can produce skilful forecasts for a more extended period: the nowcast provides accuracy at the start of the forecast, while NWP main advantages are only seen after at least 2h from the start of the forecast. Therefore, when the accuracy of rainfall forecasts start to drop rapidly, it is also the moment that the NWP forecasts start to gain more skill (Bowler, Pierce and Seed, 2006; Seed, Pierce and Norman, 2013). Both nowcast and NWP models are subject to a number of different error sources. Although nowcasts have limitations due to different sources of uncertainties, it is still more accurate than NWP forecasts at short-term and remains an important tool to produce flow forecasts. (Golding *et al.*, 1998; Seed, Pierce and Norman, 2013; Foresti *et al.*, 2016). Using nowcasting instead of quantitative precipitation estimations (QPE) from radar or rain gauge can increase the lead-time in flow forecasting (Foresti *et al.*, 2016).

Rainfall estimation is one of the primary sources of uncertainties for the first-hour lead-time (Foresti *et al.*, 2013). Advances in radars hardware and correction algorithms continuously improve radar data quality; nevertheless, residual errors will always be present and affect rainfall forecast (Germann *et al.*, 2009). As nowcasts are based on advected rainfall fields to produce forecasts, without accounting for information on the atmospheric state, the temporal evolution of rainfall is not considered. Using probabilistic forecasts can assess some of these uncertainties by perturbing the initial forecast. Probabilistic forecasts are a number of forecasts (ensembles) generated by adding noise to the deterministic forecast, and they are valid at the same time. Probabilistic forecasts are commonly used to quantify the forecast uncertainties and give further information about possible scenarios (Araghinejad and Burn, 2005; Laio and Tamea, 2007; Liguori and Rico-Ramirez, 2012a, 2013a).

Understanding and improving the estimation of these uncertainties is therefore essential to increase the ability to produce rainfall forecasts and subsequently enhance flow predictions.

This thesis proposes different approaches to account for uncertainties related to radar rainfall estimation and the temporal evolution of rainfall fields and assess how these uncertainties propagate into flow forecasts in urban and rural catchments.

1.2 Research questions and objectives

This study aims to develop methods to improve radar-based rainfall forecasting by addressing limitations related to the radar estimation of precipitation and the temporal evolution of rainfall fields and analysing how the forecasts produced can be useful to produce flow forecasts. With this in mind, the thesis addresses the following research questions:

Q1. Weather radars have the advantage of high temporal and spatial resolution. However, it is subjected to a number of measurements and estimation errors. Can probabilistic radar estimations be used successfully to account for radar uncertainties for nowcast applications? How rainfall estimation uncertainties propagate into hydrological models?

Q2. Is it possible to produce ensembles to address uncertainties due to the temporal evolution of rainfall in a more realistic way?

Q3. How temporal resolution affects radar and rain gauge merged products? Do high-resolution radar and rain gauge merged products have the potential to be used in nowcast models?

Q4. How to produce nowcasts using radar and rain gauge merged products? How the temporal resolution of radar and rain gauge merged products affects the forecasting skill? How to produce nowcasts using radar and rain gauge merged products when they lack temporal correlation?

In order to answer the research questions above, the objectives of this thesis are:

O1. To assess and quantify ensemble radar rainfall estimations in radar-based forecasting (nowcasting), sewer flow forecasting and river flow forecasting.

O2. To propose a new method to produce ensemble forecasts by adding noise to probabilistic forecasts in a more realistic way by using the information on how the precipitation evolved before the forecast starts. To assess the use of this new technique in producing ensemble rainfall forecasts and their application in sewer flow forecasting.

O3. To assess rainfall estimation accuracy using radar and rain gauge merged, KED (kriging with external drift) will be used in this thesis due to its robustness and performance. Different

temporal resolutions will be used to assess how the KED products' accuracy varies with the accumulation periods.

O4. To develop a new technique that can improve radar-based forecasts by using KED products to overcome the lack of temporal correlation between KED products' time steps.

1.3 Thesis layout

Considering some of the limitations regarding nowcasting uncertainties and their propagations, three methods to address nowcasting uncertainties were discussed. Two methods target radar rainfall estimation uncertainties and one method that assess the uncertainties related to the temporal evolution of rainfall fields.

The thesis layout includes a review of previous findings in the literature and chapters presenting novel research. In Chapter 4 and Chapter 5, the methods proposed are based on using probabilistic forecasts to quantify uncertainties. Chapter 6 tests different KED temporal resolutions that are an essential step to allow KED use in nowcasting and hydrological applications. Chapter 7 presents an original method to use KED in nowcasting. Unlike the techniques in Chapter 4 and Chapter 5, the use of KED in nowcasting aims to reduce forecasting uncertainties. Finally, Chapter 8 gives an overall conclusion of this thesis findings.

Chapter 2 outlines the theory behind rainfall measurement and describes the main uncertainties affecting rain gauges measurements and radar rainfall estimations. In Chapter 1, an overview of nowcasting techniques and uncertainties related to the nowcasting models is presented.

Chapter 3 defines the rain gauge and radar dataset used in this thesis and the nowcast and hydrological models used in the thesis.

Chapter 4 addresses the research question Q1. It describes an ensemble generator to produce ensemble radar rainfall estimations proposed by (Germann *et al.*, 2009) that uses a historical data set to estimate residual errors in radar rainfall. The method has been tested for rainfall and flow estimations, but this method's applicability in rainfall forecast for hydrological applications has not been evaluated. Firstly, with this as the chapter's final aim, ensemble radar rainfall forecasts were produced and compared with stochastic ensemble rainfall forecasts. In the next step, the rainfall forecast was used to forecast flows in an urban area and river catchments.

Chapter 5 proposes a new method to generate ensembles to account for uncertainties related to the temporal evolution of the velocity field to address the issues raised on the research question Q2. Utilising ensemble rainfall forecasts to account for these uncertainties is commonly used. However,

ensembles are usually produced by perturbing the rainfall fields with stochastic noise. The method proposed in this chapter uses radar images that are not usually used to produce the forecast, but they provide useful information on how the rainfall developed prior to the beginning of the forecast. In this way, different velocity fields are related to the rainfall event and can be used to produce the forecasts. The results are used to forecast flows in an urban catchment and to assess if the rainfall forecasts can be used for this application.

Chapter 6 gives an overview of rain gauge – radar rainfall merging techniques, specifically kriging with external drift (KED). In order to allow the use of KED in rainfall forecasting applications and to address research question Q3, different accumulation periods used in the merging are tested. KED is often applied using hourly accumulations. A few studies use high temporal resolution; however, further tests were needed to assess the prospect of using KED in nowcasting models.

Chapter 7 addresses the research question Q4 by presenting a novel technique to produce rainfall forecasts using KED rainfall and consider the temporal correlation of precipitation. KED products do not maintain the temporal correlation of rainfall. This is a major issue when generating nowcasts as the rainfall advection field calculated can be opposite from the reality, advecting the rainfall in different directions. The method proposed uses the high temporal correlation obtained from radar images to compute the advection field. The KED merged product is then advected using the radar-based advection field. The chapter assesses this novel technique for radar-based precipitation forecasting.

Chapter 8 summarises the thesis's conclusions and the limitations of the work; it also provides recommendations for future research.

Chapter 2. Rainfall observation and forecast

Accurate precipitation estimation is very important in hydrological modelling. It is a crucial input in rainfall nowcasting and can be used in rainfall-runoff modelling, flood forecasting applications, such as issuing warnings for extreme events, hydrological applications, and climate modelling. Flood forecasting applications in large rural catchments require rainfall with good spatial and temporal resolutions. For instance, in the UK, weather radar rainfall with spatial and temporal resolutions of 1 km and 15 min respectively is used for flood forecasting in large catchments (Price et al., 2012). However, for flood forecasting applications in small urban areas, rainfall data should have high temporal (e.g., 5 min or lower) and spatial (e.g., 1 km or lower) resolutions, as the small scale of the catchment combined in changes in the land use can lead to rapid changes in the catchment flow. A high resolution also allows the hydrological process dynamics to be better captured (Berne and Krajewski, 2013; Thorndahl *et al.*, 2016). Over the last couple of decades, much research has been done to increase forecasting rainfall reliability (Pierce *et al.*, 2000; Bowler, Pierce and Seed, 2006; Cloke and Pappenberger, 2009; Foresti *et al.*, 2013). Hydrodynamic models for real-time flow predictions, utilising radar rainfall and radar-based rainfall forecasts, can be used for the real-time control of drainage systems in urban areas (Liguori et al., 2012). However, it is essential to know the uncertainties related to the radar rainfall measurements to produce reliable hydrological simulations and forecasts.

Several sensors are available to measure precipitation, such as rain gauges, disdrometers, weather radar, microwave links and satellite. Each has its own advantages and disadvantages in terms of measurement, accuracy and spatial and temporal resolutions. This thesis uses rain gauge and weather radar data to improve precipitation measurement and forecasting for flood forecasting applications in large river catchments and urban areas. For this reason, this chapter will focus mainly on these two types of precipitation measurements.

Disdrometers are used to measure the raindrop size distribution (DSD) of precipitation events at a given point location, but these instruments do not measure rainfall rates directly. They measure the number of drops of different sizes, which give the DSD, which then can be used to the rainfall rate (Rico-Ramirez, 2019). The rainfall rate is approximately the 3.67^{th} moment of the DSD, providing the raindrops have reached terminal fall velocities (Rico-Ramirez, 2019). Disdrometers are often used to calibrate rainfall data from weather radars (Tokay *et al.*, 2003), given the fact that the radars measure the radar reflectivity from precipitation particles that are related to the sixth moment of the DSD (Rico-Ramirez, 2018). As with any other precipitation measurement method, disdrometer measurements are also prone to uncertainties. For instance, downdrafts (or updrafts) can cause an increase (or decrease) of the raindrops fall velocities and consequently affecting the rainfall rates. Since disdrometers provide

point measurements with a relatively small sampling area (around 50 cm^2), the variability of precipitation in space is not accounted for, in particular during convective rainfall events. Studies from Jaffrain and Berne (2011) and Tokay et al. (2003) estimate that sampling uncertainties in disdrometers are around 10-20%.

Satellite measurements are particularly helpful in remote areas and oceans where other types of rainfall measurements are not available or have poor quality. Satellites providing global rainfall estimates are handy in hydrological applications due to their spatial coverage (Sorooshian *et al.*, 2000), and the quality of estimations based on satellites have been improving over the years (Rico-Ramirez, 2019). Nonetheless, using satellites for operation applications still encounters limitations due to their uncertainties (Aghakouchak *et al.*, 2012). The Tropical Rainfall Measuring Mission (TRMM) from the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA) orbits in a range that covers tropical and sub-tropical ranges, providing rainfall, water vapours and clouds measurements (Kummerow *et al.*, 2000; Prigent, 2010). After its end, the Global Precipitation Measurement (GPM) mission was created as an extension to the TRMM. The GPM provides measurements from across the globe and includes an infrared sensor, microwave imager, dual-frequency Ku/Ka-band precipitation radar and infrared sensors (Smith *et al.*, 2007) for measuring microphysical properties, water fluxes and latent heat releases (Hou *et al.*, 2014). Rainfall measurements from GPM have a 0.1 degree spatial resolution and 3 hr temporal resolution (Smith *et al.*, 2007) with a latency time of about 6hr, which opens up new opportunities for real-time hydrological applications. The GPM mission was also designed to fill the gap of the TRMM regarding estimation of light-intensity precipitation and snow (Hou *et al.*, 2014). The GPM mission can provide valuable information for monitoring and predicting hurricanes, tropical cyclones, and precipitation events. Satellite data is also valuable for numeric weather (NWP) predictions as it gives further information about the atmospheric state (Hou *et al.*, 2014) and can potentially be used for data assimilation or to validate NWP forecasts.

Rain gauges and weather radars are vastly used to measure rainfall (Cecinati, 2017). Rain gauges are instruments that measure precipitation at ground level with high accuracy (Wilson, Brandes and Noaa, 1979); however, it lacks the spatial resolution needed for flood forecasting (Jarraud, 2010; Berne and Krajewski, 2013). Weather radars estimate rainfall in real-time and provide high-resolution data, and have great potential for being used in flood forecasting applications. The spatial resolution usually obtained with weather radar (e.g. 1km in the UK) would be difficult to achieve, and even dense rain gauge networks are unable to produce such resolution.

Weather radars can be used for short-term precipitation forecasting, which is also known as nowcasting. Nevertheless, weather radar rainfall estimations are not as accurate as rain gauge measurements, and often this can limit its application in hydrology (Wilson, Brandes and Noaa, 1979; Borga, 2002; Villarini and Krajewski, 2010). Accurate rainfall forecast inputs are one of the main factors that influence accurate flow predictions. Increasing the predictability of these forecasts has proven to be a challenge given the small spatial scales involved, and therefore reducing radar rainfall-related uncertainties can improve radar-based rainfall forecasts (Bowler, Pierce and Seed, 2004; Foresti *et al.*, 2013). Both rain gauges and weather radar rainfall are affected by different sources of errors that propagate into the rainfall forecasts produced by nowcasting models (Germann *et al.*, 2009; Liguori and Rico-Ramirez, 2012c). Geostatistical and non-geostatistical spatial interpolating methods, such as kriging, co-kriging, inverse distance weighting and regression models, have been developed to reduce uncertainties related to weather radar. (Seed and Austin, 1990; Kitchen and Blackall, 1992; Ciach and Krajewski, 1999; Habib, Krajewski and Kruger, 2001; Germann et al., 2006; Villarini et al., 2008). Methods to blend radar and rain gauge measurements will be further discussed in Chapter 6.

The wide availability of weather and its high temporal and spatial resolution make weather radars a good choice for rainfall and hydrological forecasting, particularly for hydrological application in urban catchments. Although there are limitations due to its accuracy, there are many methods to address the primary sources of inaccuracies in weather radars, and new researches are continuously being carried out. Apart from that, forecasting models can also include steps to account for measurements errors and improve the forecast outcome.

In this thesis, methods to account for rainfall estimation errors in nowcasting are proposed. Understanding how rainfall is measured, estimated and the limitations of the different instruments available are essential when deciding how to address these uncertainties and proposing ways to overcome them. This chapter provides an overview of the general theory to measure rainfall. It discusses limitations and uncertainties related to rain gauges and weather radars and provides an overview of methods used to generate short-term rainfall forecasts. This chapter only gives a general description of rainfall estimation and forecasting. This is a vast subject, and more detailed information can be found in the cited references.

2.1 Rain gauges

Rain gauges are instruments used to measure rain at ground-level. Rain gauges are used as a reference and are known to be reliable precipitation instruments; however, its use in forecasting rainfall for hydrological applications would be desirable a dense network of rain gauges, in order to capture the

process dynamics, and therefore this is often not practical (Wilson, Brandes and Noaa, 1979). Rain gauges are point measurements; being so, they only represent measurements over a limited area. However, using rain gauges data in combination with weather radar rainfall can provide better estimations of precipitation over a large area (Jarraud, 2010) with better accuracy and spatial/temporal resolutions.

Rain gauges can be classified according to how measurements are made: manual gauges, mechanical recording gauges, and electrical rain gauges (Strangeways, 2010). As the name says, manual gauges need an observer to record the amount of rain that accumulated over a period of time. Measurements can be made daily, weekly or even monthly, depending on the measuring cylinder's volume. It is estimated that there are over 50 different kinds of manual rain gauges in the world (Strangeways, 2010), such as the 'Snowdon' copper rain gauge (Figure 2-1). The 'Snowdon' rain gauge is a 5-inches gauge that has been used as a standard in the UK and Ireland for over a century and can provide reasonable rainfall estimations due to its large funnel and accurate rim. It stores rainwater on a bottle situated below ground level. (Strangeways, 2010; Burt, 2013). Readings are daily and made using a measuring cylinder. Mechanical recording gauges can record the measurements. Gauges such as siphon gauges are able to record the measurements, but an observer is needed on-site to collect the data. These gauges are usually used to give precise totals and are used as reference rain gauges. Electrical rain gauges, such as tipping bucket rain gauge (TBR), electronic weighing rain gauges (WR), capacitance gauges and drop-counting gauges, can record measurements remotely with the help of a data logger and have the advantage of better temporal resolution (Strangeways, 2006, 2010). As TBRs and WRs are vastly used, a more detailed description of these gauges is found in section 2.1.2 and 2.1.1.

Although rain gauges are usually used as a reference, they are also subject to measurement errors. Uncertainties are related to site and location errors, instrumental errors and random errors (Krajewski *et al.*, 2003; Rico-Ramirez, Liguori and Schellart, 2015). Precipitation measured by rain gauges will be usually underestimated with errors of up to 30% (Jarraud, 2010). Systematic errors in the measurement are usually more minor for liquid than solid precipitation. Precipitation can be estimated by correcting some of the systematic errors (Jarraud, 2010). The following uncertainties described in this section are a general description that are, up to some extent, common to most of the rain gauges. Uncertainties specific related to WR and TBR are going to be described in sections 2.1.2 and 2.1.1.

Figure 2-1 'Snowdon' rain gauge and graduated cylinder (Strangeways, 2010).

Uncertainties in rain gauge measurements can be related to wetting losses on the walls of the collecting funnel strainers or any debris collected, and losses in the emptying of the container can account for up to 15% of errors during warm weather and up to 8% in winter periods (Jarraud, 2010). These uncertainties are accumulative and are dependent on the number of times the container is emptied and on the type of gauges (e.g. WR are not subject to losses related to wetting losses in the container) (Strangeways, 2006; Jarraud, 2010).

Evaporation losses are mainly important during the summer or in hot locations, leading to uncertainties of up to 4% (Jarraud, 2010). In the Snowdon rain gauge, an average of 0.2 mm is lost in each rain event (Strangeways, 2006). In heavy rainfall events, evaporation is not as significant in short intermittent events of light rainfall. The water inside the gauges is less prone to evaporation by implementing measures such as housing the container underground and minimising the ventilation. Evaporation losses also are dependent on the gauge type and is an issue especially significant in gauges without a funnel to collect rainwater as it allows a more extensive water surface to be exposed. The former two sources of uncertainties can be minimised when electrical rain gauges are used (Jarraud, 2010).

Raindrops can splash when falling into the gauges, and vertical sides have proven to be useful to avoid losses. On the other hand, it can increase evaporation losses due to an increased wetting area, and it also increases wind-effect errors. Out- and in splashing losses can also cause errors of around 1-2% of rainfall (Kurtyka, 1953; Strangeways, 2006; Jarraud, 2010).

The collector has to be large enough to minimise uncertainties (with a diameter of at least 4 inches) (Strangeways, 2006). Gauges with less than 3 inches in diameter can present under-catch of up to 6%.

This error is reduced to up to 2% when the diameter is larger than 4 inches (Kurtyka, 1953). In manual gauges, a larger collector also means an increase in the size of the container. The UK Met Office uses a collector of 5 in on their manual gauges. For mechanical gauges, it is important to have a large collector so the water collected can overcome levers friction. For a TBR, the collection should be higher than 10 ml per tip. (Strangeways, 2006).

Locations and installation of rain gauges are also essential to reduce uncertainties. The gauges should be levelled as 1-degree tilt leads to a 1% error, and maintenance should be made to assure that it has not moved (Kurtyka, 1953; Strangeways, 2006). Also, the installation should be at a distance from other objects. As a general rule, it is acceptable for the distance to be twice longer than the object's height, although four times the height is often recommended (Strangeways, 2006).

Wind effects are one of the largest sources of errors in rain gauges, leading to uncertainties of 2-10% for rain and up to 80 % for snow (Kurtyka, 1953; Jarraud, 2010). As a general rule, the rain gauges should not be installed in windy locations. Wind effects are related to the interactions between the wind flow, the body of the rain gauge and the particles falling through the atmosphere and mainly affect lighter particles that are carried away from the gauge. The trajectories of these particles are distorted by the wind speed, displacement and acceleration close to the rain gauge. The use of windshield devices plays an essential role in reducing turbulence around rain gauges. The shape of the gauge and the orifice and data corrections can also improve rainfall estimations (Sevruk and Nespor, 1994; Duchon and Essenberg, 2001; Jarraud, 2010; Dai *et al.*, 2013). As it is a great source of uncertainties, different types of windshields have been developed throughout the years. One of the first attempts to shield rain gauges was made by Nipher at the end of the 19th century. The Nipher shield is shaped like an inverted trumpet. The design is not ideal for cold weather as snow can accumulate in the shield, but it is a widely used method (Kurtyka, 1953; Strangeways, 2006). The Alter shield is designed with hinged metal strips that are loosely hanging. In windy conditions, the strips swing inwards and form a conical shape similar to the Nipher shield. Its hanging strips avoid snow accumulation in the shield (Kurtyka, 1953; Strangeways, 2006). Tretyakov made a windshield that combines both the Nipher and Alter design effect and has fixed plates (see Figure 2-2). Other ways of avoiding wind effects include the pit gauge (see Figure 2-3) and turf wall, that although they are effective, can get filled with snow and sand (Strangeways, 2006). Aerodynamic gauges were developed in order to reduce the gauge profile to cause less interference in the wind flow. However, they present increased out-splashing errors. Although, in real applications, the airflow is turbulent, and its direction and speed can rapidly change. The designs were made through experiments in wind tunnels that use laminar airflow (Strangeways, 2004).

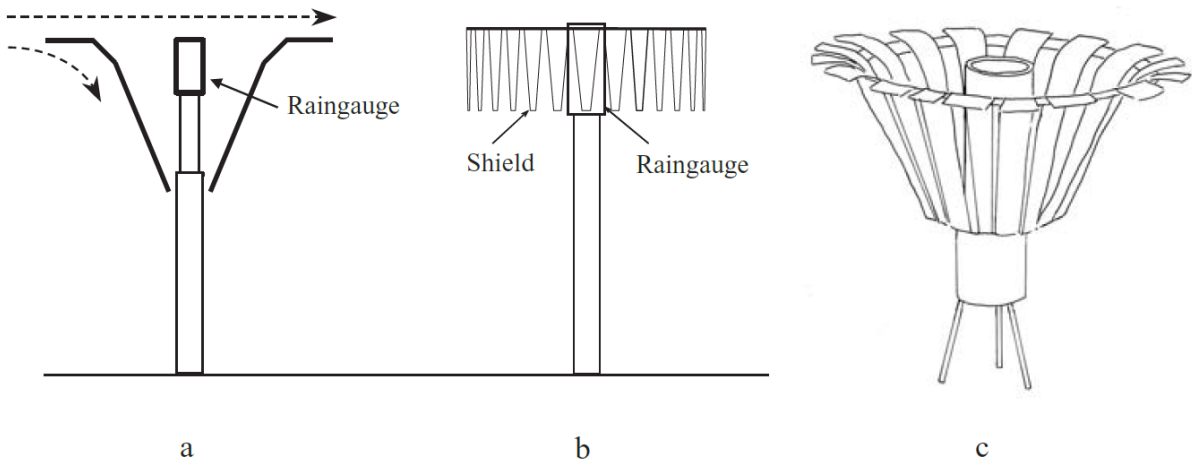


Figure 2-2 (a) Nipher Shield. (b) Alter Shield. (c) Tretakov screen (Strangeways, 2006)

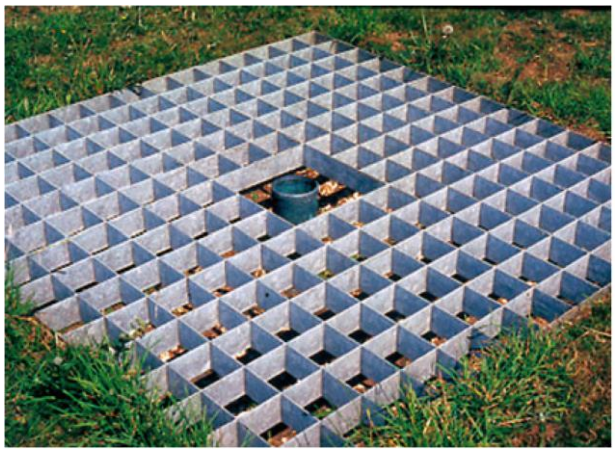


Figure 2-3 Pit rain gauge (Strangeways, 2010).

Spatial uncertainties occur as rain gauges are point measurements and are used to estimate areal rainfall. Rainfall intensity, rain gauge network density and domain size play an important role in spatial errors. Nonetheless, even in a highly dense network, sampling uncertainties will still be present. In this context, using spatial interpolating techniques can improve rainfall estimation.

2.1.1 Tipping Bucket Rain Gauge

Tipping bucket rain gauges (TBR) are the most common type of electrical rain gauges and play an essential role in rainfall-runoff modelling and urban hydrology. They are also used to calibrate or validate other sources of rainfall measurements (e.g. radar, satellite, microwave links). TBRs are widely used in the UK as they rely on a simple mechanism, are reliable and do not consume energy (Habib, Krajewski and Kruger, 2001; Strangeways, 2006). A TBR is based on a container divided into two triangular sections (Figure 2-4). The container is balanced over an axis and rests against one of two stops. Figure 2-5 illustrates the operation of a TBR during rainfall. The collector directs the water for the upper compartment, and when filled, the compartment becomes unbalanced and tips over. While

it empties, the other compartment is filled. Records are made every time the compartment tilts through a pulse-generating switch connected to a receiver. Every pulse or tip is related to the resolution of the TBR. Once a record is only done when the bucket switches, TBR only measure accumulated rainfall between tips. Rainfall estimates are done by counting the number of tips and the sections' capacity (Habib, Krajewski and Kruger, 2001; Jarraud, 2010). TBR can be designed with different compartment sizes. Very small buckets are more prone to uncertainties as the bucket would tip with a very small water amount. TBR are usually designed to tilt at intervals of 0.1, 0.2, 0.25, 0.5 and 1 mm of rain (Strangeways, 2006). In the UK, 0.1 and 0.2 mm intervals are often used, while in the US the most used interval is 0.25 mm and 1 mm.

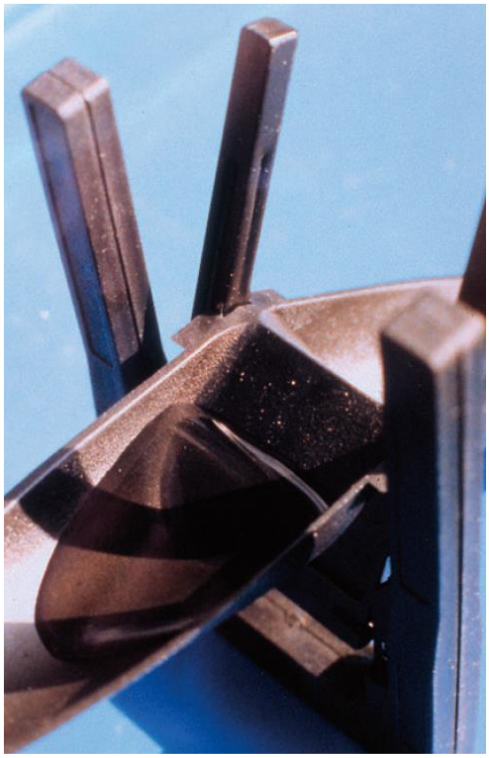


Figure 2-4 Tipping bucket (Strangeways, 2010).

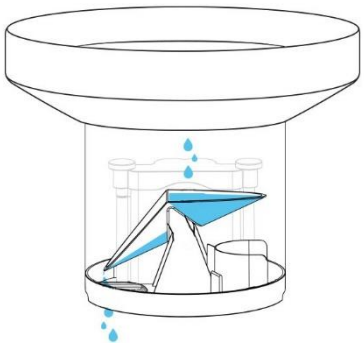


*Figure 2-5 Schematic drawing of tipping bucket rain gauge operating (*Netatmo Rain Gauge | Weather Station Tools*, 2013).*

Some of the uncertainties of TBRs were already discussed in section 2.1. However, TBRs are subjected to some specific errors. Systematic errors in TBR include wind effects and losses due to over-splashing, wetting and evaporation and are responsible for most of the measurements uncertainties (Humphrey *et al.*, 1997; Habib, Krajewski and Kruger, 2001). Due to the high surface area, TBRs are subject to evaporation losses. In short events of light rainfall, when the accumulated rain in the container is not enough to tip it, the remaining water can evaporate and therefore is not measured. TBRs are not recommended to measure solid precipitation. Heating is often necessary to melt the solid precipitation, leading to higher losses due to evaporation. Evaporation due to heating was found to be comparable or more significant than uncertainties due to wind effect (Strangeways, 2006, 2010; Jarraud, 2010; Savina *et al.*, 2012). Water can also adhere to the gauge and change the water balance of the compartment or leads to evaporation losses in cases when the bucket does not tip again. It might also occur that the compartment is not completely emptied when it tilts. In both events, more water is necessary to tip the bucket resulting in underestimation of short events. In longer events, after the third tip, this error is reduced. In very light rain, temporal errors can occur. It happens because the rainfall amount takes a longer time to tip the bucket, and the start and end of the event might not be accurately determined (Jarraud, 2010; Strangeways, 2010). During the time that it takes to tip over, undercatch happens when rainfall might still fall in the compartment that is being empty and cause underestimation of rainfall. During heavy rainfall events (over 250 mm/hr), these errors increase as the time it takes to reposition the buckets is not fast enough. Devices, such as a precise programmable pump, control the rate that the water falls into the compartments or that accelerate the tipping action can be used to reduce this uncertainty. Although this source of error can be reduced, it cannot be eliminated. The nozzle shape, position, and size can also lead to overreading according to the stream of water falling into the compartment (Jarraud, 2010; Strangeways, 2010). Calibration with different rainfall rates can also be made and taken into account in the data processing. Dynamic calibration is recommended to reduce measurements uncertainties and is an effective way to minimise the underestimation of rainfall in high-intensity events (Niemczynowicz, 1986; Humphrey *et al.*, 1997; Jarraud, 2010). Other error sources in TBRs are blockages due to leaves or dirt accumulated in the bucket (see Figure 2-7). It can be detected when looking at precipitation measurements from a nearby rain gauge, but regular maintenance of the TBR can help reduce blockages.

Figure 2-6 Example of a blocked TBR (Source: Personal collection).

2.1.2 Weighing Rain Gauges

Weighing Rain Gauges (WR) measure the weight of a container continuously, and consequently, the precipitation. The rainfall rate is calculated based on the volume of water collected during a given time interval (Nystuen *et al.*, 1996; Jarraud, 2010). In order to weigh the water, the gauge measures the rainfall amount with a spring mechanism or with a system of balanced weights. The electronic version of WR uses a straingauge load cell to weigh the water, and it produces more accurate measurements than the mechanical version. Figure 2-7 illustrates a mechanical weighing rain gauge, and a photo of a typical weighing rain gauge can be found in Figure 2-8. For continuous operation, a siphon is necessary to empty the bucket regularly; however, this is not shown in Figure 2-7. In order to reduce evaporation losses, evaporation suppressants, such as oil, are added to the gauge container (Strangeways, 2006; Jarraud, 2010). WRs are designed to measure all types of precipitation; however, they cannot identify the type of precipitation measured. In places where snow or hail are present, the gauges need to have a heating system, or anti-freezing solutions should be added to melt the snow. Besides that, snow measurements are more prone to be affected by wind as snowflakes are easier to be deflected by it (Strangeways, 2006; Jarraud, 2010).

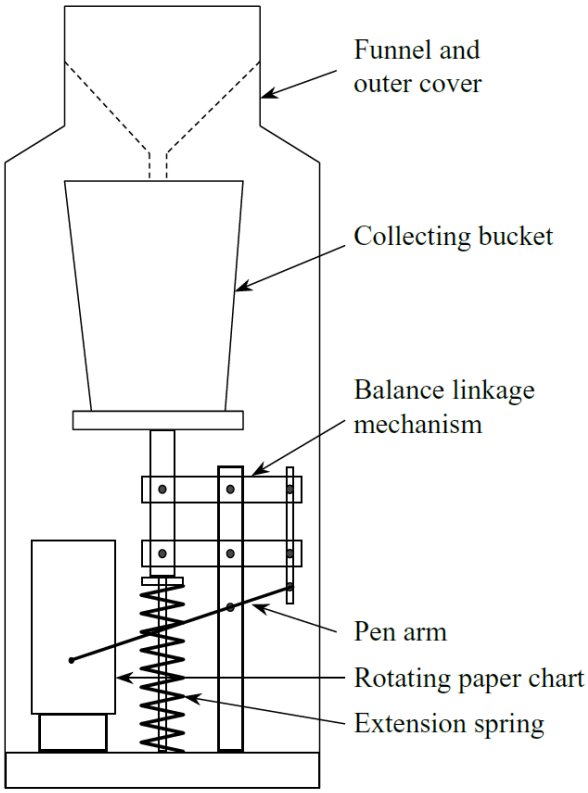


Figure 2-7 Schematic drawing of a Weighing Rain Gauge (Strangeways, 2006)

WRs are subject to some uncertainties related to the gauge type and how measurements are done, and most of it was discussed in section 2.1. An issue related to solids' measurements is that they can stick to the funnel orifice and not fall into the container until it melts, leading to temporal uncertainty (Jarraud, 2010). Calibration is an essential part of gauge maintenance and should be done 3 to 4 times per year. Maintenance also includes inspection of all components, cleaning, and recharging of the anti-freezing solution. WRs do not use mechanical moving parts to measure rainfall, so maintenance is simplified as there is no mechanical degradation (Vuerich *et al.*, 2009; Jarraud, 2010). WR can be sensitive to temperature, wind flow over the gauge, and raindrops impact. Data processing is mandatory to deal with these noises (Vuerich *et al.*, 2009).

Advantages of WR include the fact that they are not subject to wetting losses when the container is emptied, and the continuous weighing of rainfall produces a better temporal measurement than TBR (Jarraud, 2010; Strangeways, 2010). As the measurements are constant, WR evaporation losses are also smaller than in TBR and can be quantified. As WR offer advantages when compared to TBR, currently, the UK Environment Agency (EA) is in the process of replacing their TBR rain gauges for WRs (M. A. Rico-Ramirez, personal communication, 6th of June, 2019).

Figure 2-8 Weighing rain gauge

2.2 Weather radar

In World War II, the United Kingdom and the United States of America worked together in developing radar technology to locate enemy aircraft. The invention of the magnetron allowed radars to operate with shorter wavelengths, and as a result, the antenna's size could be reduced and increased the accuracy of the radar in detecting target aircraft (Probert - Jones, 1962; Rinehart, 1997; Rico-Ramirez, 2019). During the war, an issue with radars was that back-scattered energy from meteorological phenomena also appeared in the radar display, overshadowing aircraft signals (Probert-Jones, 1962). After the war, meteorologists started researching the use of radar to study the weather. Since then, improvements in all parts of radars have been made using computers to process and correct radar data and improve the quality of radar-based rainfall estimations (Rinehart, 1997).

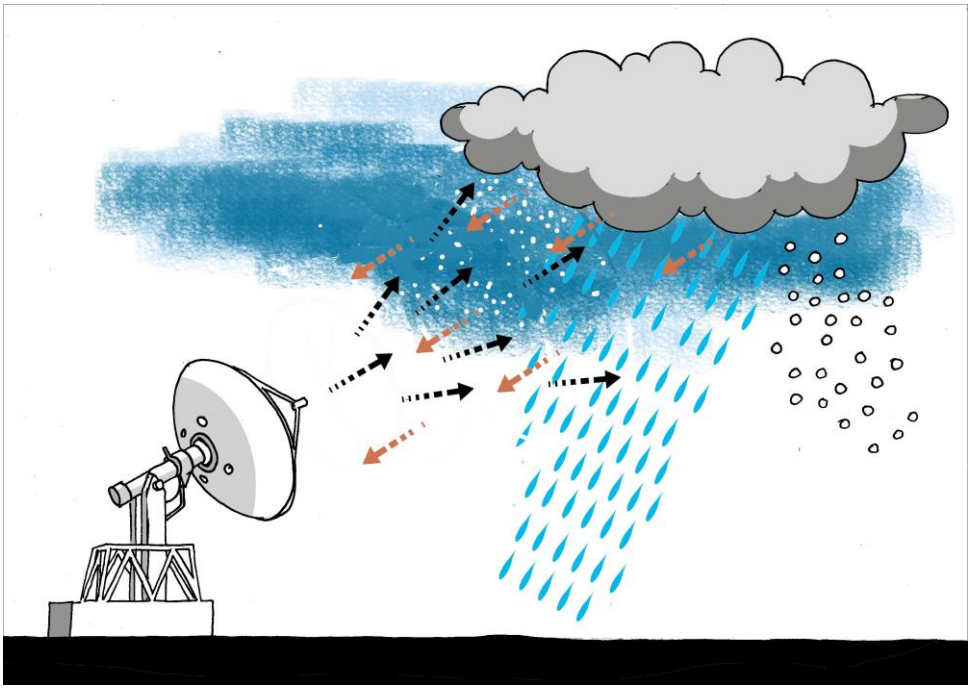


Figure 2-9 Schematic of how weather radars work (Source: Personal collection).

Radars are based on four main components: transmitter (produces electromagnetic radiation), antenna (concentrates the radiation and receives part of the back-scattered radiation), receiver (receives the back-scattered radiation and converts it to a low-frequency signal) and a display system (Collier, 1996; Rinehart, 1997). The microwave frequency from radars depends on the radar type (5 GHz for C-band, 3 GHz for S-band and 10 GHz for X-band) (Rico-Ramirez, 2019). When an electromagnetic wave hits a target, part of it is reflected to the radar. Rainfall measurements are made based on the back-scattered radiation, also called echo or radar reflectivity. It also incorporates the difference between horizontal and vertical polarised radiation in the case of dual-polarisation radars (Collier, 1996; Germann *et al.*, 2006; Met Office, 2009). The echoes produced constantly vary according to the movement of hydrometers. Therefore it is necessary to integrate a number of uncorrelated power reflected measurements in order to reduce the noise in the radar reflectivity. Figure 2-9 illustrates the principle used by weather radars, where the radar sends electromagnetic waves, and the backscattered radiation is used to measures rainfall. The back-scattered energy varies according to the number of particles measured within the volume measured by the beam, physical state, orientation, composition, state, size and shape of the hydrometer. Probert-Jones (1962) developed a radar equation to calculate the average power received.

$$P_r = \frac{\pi^3}{1024\ln(2)}\frac{P_0 h}{\lambda^2}G^2\theta_1\frac{1}{r^2}|K|^2\sum D^6 = \frac{C|K|^2 Z}{r^2} \qquad (2\text{-}1)$$

Where P_r is the received power, P_0 is the transmitted power, h is the pulse length, G is the gain of the antenna, θ_1 beamwidth, $|\mathrm{K}|^2$ is the dielectric constant (0.93 for water and 0.2 for ice particles), D is the

drop diameter, λ is the radar wavelength, and r is the range from the target to the radar location. Note that $Z = \sum D^6$ (see below). Adequate calibration of the radar constant (C) can help to improve the overall calibration of the radar system.

For average raindrop diameters small when compared to the radar wavelength ($D \ll \lambda$), it is possible to apply the approximation of the Rayleigh scattering (Rinehart, 1997; Rico-Ramirez, 2019). In this condition, the reflectivity can be defined as a function of the raindrop diameter and the drop size distribution (Collier, 1996):

$$Z = \int D^6 N(D) \mathrm{d}D; \ \ \mathrm{mm^6\ m^{-3}} \qquad (2\text{-}2)$$

Where $N(D)$ is the drop-size distribution. The radar reflectivity can be affected by the size of hydrometers considered, being so express it in terms of dBZ units is able to express this large variation in an appropriate way (Collier, 1996):

$$dBZ = 10\log_{10}(Z) \qquad (2\text{-}3)$$

The most common drop-size distribution relation is the two-parameter exponential form defined by Marshall and Palmer (1948):

$$N(D) = N_0 . e^{-\Lambda D} \qquad (2\text{-}4)$$

$$\Lambda = 4.1 . R^{-0.21}\ m^{-1} \qquad (2\text{-}5)$$

Where N_0 is $N(D = 0) = 8000\ m^{-4}$, Λ is the slope factor, and R is the rainfall rate.

The rainfall rate can also be expressed in function of the diameter, the rainfall drop-size distribution and raindrop velocity (Atlas and Ulbrich, 1977):

$$R = \frac{\pi}{6} \int v(D) D^3 N(D) \mathrm{d}D\ ;\ \mathrm{mm\ hr^{-1}} \qquad (2\text{-}6)$$

Where $v(D)$ is the terminal raindrop velocity (raindrop velocity when air friction or air resistance and the gravity force are balanced) and can be expressed by (Atlas and Ulbrich, 1977; Rico-Ramirez, 2004):

$$v(D) = 3.87 D^{0.67}; m\ s^{-1} \qquad (2\text{-}7)$$

From equations (2-6) and (2-7), we have that:

$$R = \frac{\pi}{6} (3.87) \int D^{3.67} N(D) \mathrm{dD}\ ;\ \mathrm{mm\ hr^{-1}} \qquad (2\text{-}8)$$

It is important to notice from equation (2-8) that R is related to the 3.67th moment of the drop size distribution, while Z is related to the 6th moment of the drop size distribution. Therefore, Z is more sensitive to larger drops than R.

A semi-empirical relation for calculating rainfall rates can be used (Marshall, Langille and Palmer, 1947; Collier, 1996; Harrison, Driscoll and Kitchen, 2000).

$$Z = aR^b \qquad (2\text{-}9)$$

Where a and b are constants that depend on the type of precipitation and can be estimated by using disdrometer observations where both *Z* and *R* are available or by using radar and rain gauge measurements. Using dual-polarization radars in conjunction with disdrometers can provide measurements of the raindrop size distribution. The N(D) can be calculated using the equation (2-2). The *Z-R* relationship depends on the calibration of the radar. However, in the literature, several relationships are dependent on the climatology of the area or the nature of precipitation (Marshall and Palmer, 1948; Marshall, Hitschfeld and Gunn, 1955; Atlas and Ulbrich, 1977; Fulton *et al.*, 1998). Marshall, Hitschfeld and Gunn (1955) defined a relationship that is more suitable for stratiform events, and it is often used in the UK:

$$Z = 200R^{1.6} \qquad (2\text{-}10)$$

In the US, the commonly used equation is more suitable for convective storms (Fulton *et al.*, 1998):

$$Z = 300R^{1.4} \qquad (2\text{-}11)$$

Apart from the ones presented on equations (2-10) and (2-11), many other *Z-R* relationships can be found in the literature. Ulbrich and Lee (1999) presented a list of some typical values. Ulbrich and Lee (1999) highlighted that the Z-R relationships for thunderstorm represent the core of the precipitation, and the values for *a* and *b* can change according to the development stage of precipitation. Given the fact that precipitation types change in space and time, the use of a single climatological *Z-R* equation can lead to errors in the estimation of precipitation with radar when looking at particular events.

Table 2-1. Coefficient a and b for some typical Z-R relationships (Ulbrich and Lee, 1999).

Source	Rainfall type	*a*	*b*
Joss and Waldvogel (1970)	Average of several types	300	1.5
Marshall and Palmer (1948)	Stratiform	200	1.6
Atlas and Chmela (1957)	Stratiform	255	1.41
Fujiwara (1965)	Stratiform	205	1.48
Jones (1956)	Thunderstorm	486	1.37
Fujiwara (1965)	Thunderstorm	450	1.46

For forecasting applications, rainfall data should have a high temporal and spatial resolution. Weather radars can scan large areas measure rainfall in real-time, with a temporal resolution of 5 min or lower, and provide high-resolution data (1 km or lower) at real-time (Wilson, Brandes and Noaa, 1979; Sun *et al.*, 2000; Sokol, 2003; Germann *et al.*, 2006). However, weather radar rainfall data are affected by different sources of errors (Rico-Ramirez, Liguori and Schellart, 2015). Radar errors can propagate in short-term forecasts; generating inaccurate forecasts and understanding these errors is essential for rainfall forecasting applications.

2.2.1 Radar uncertainties

Despite advances in the weather radar technology in the past few decades, measurements errors such as radar calibration, ground clutter, occultation (beam blockage), radar beam overshooting in shallow precipitation, variation in the vertical profile of radar reflectivity (VPR), reflectivity-rainfall (Z-R) relationships, radar signal attenuation, and anomalous propagation (AP) still affect rainfall measurement and limits its application in hydrology (see Figure 2-10) (Pellarin *et al.*, 2002; Germann *et al.*, 2006; Rico-Ramirez *et al.*, 2007; Rico-Ramirez and Cluckie, 2008). Understanding, identifying and correcting rainfall estimation uncertainties produce better quantitative precipitation estimations (QPE). According to Gorgucci, Scarchill and Chandrasekar (1996), precipitation estimates can have a large uncertainty (higher than 100%) when the reflectivity error is only a few decibels at C-band radars. Dual-polarisation radars also permit the hydrometeor characteristics measurements (Rico-Ramirez, Cluckie, and Han, 2005), which can improve the quality of the radar rainfall measurements. The following sections describe in more detail the error sources in radar rainfall estimates.

2.2.1.1 Radar Calibration

For accurate measurements, the radar's hardware components must be well maintained and up to date to guarantee an accurate calibration of the constants from the radar equation (2-1). Calibration of radar should also be made regularly to avoid inaccurate readings (Harrison, Driscoll and Kitchen, 2000). Lack of accurate determination of the calibration constant can lead to systematic errors in the observations. The calibrated constants should be verified before any longer processing is applied to the data. There are different methods available to calibrate weather radars. However, Ryzhkov *et al.*, (2005), Gourley et al. (2003) and Atlas, (2002) highlight that not having a universal method for calibrating radars and independently calibrating radars of an entire network can lead to discrepancies in rainfall estimation.

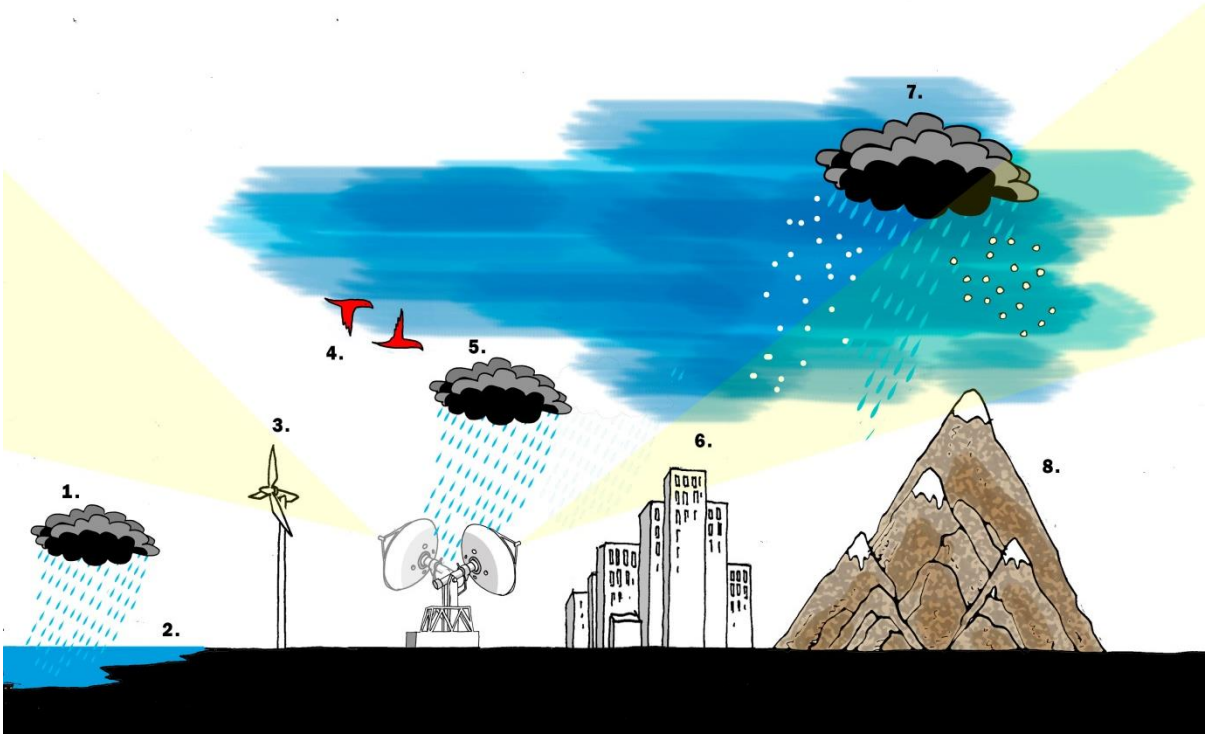


Figure 2-10. Illustration of uncertainties related to radar rainfall estimation. Where: 1. Overshooting of precipitation, 2. Sea clutter, 3. Wind turbines, 4. Bio-scatter, 5. Radar attenuation, 6. Ground clutter, 7. Variation of the vertical profile of reflectivity and 8. Occultation (Source: Personal collection).

The determination of the radar constant can be realized through field tests that use a metal sphere (Anagnostou, Morales and Dinku, 2001) that can be dropped at constant intervals or carried by aircraft, balloons or drones (Atlas, 2002; Duthoit *et al.*, 2017; Suh *et al.*, 2017). The location of each ball can be compared with the radar data (Atlas, 2002). Issues related to the tracking of the target, mainly in a narrow beam, could be overcome by using GPS (Atlas, 2002). A recent paper by Yin *et al.* (2019) showed how an unmanned aerial vehicle (also known as a drone) could be used to carry the metal sphere to calibrate weather radar systems. Because the backscattering cross-section of a metal sphere is known, then the reflectivity measurements can be calibrated. The frequency shift reflector (FSR) is another well-known method to calibrate radars and uses a parabolic reflector. The reflector possess a horn at the focus that is shorted by a diode and can be used in any weather conditions (Atlas, 2002). A method to calibrate ground radars using a space-based radar was proposed by Anagnostou et al. (2001) and the methodology demonstrated a significant reduction in uncertainties in the rainfall estimations related to miscalibration and would allow calibrations to be more frequent.

2.2.1.2 Ground clutter, occultation and anomalous propagation

The choice of the radar's position and the characteristics of the landscape around the radar can also produce errors in radar QPE (Quantitative Precipitation Estimation). Ideally, the radar beam should not encounter any kind of obstacles. However, the reality is that ground targets are more frequently than

desired, and higher interferences occur in orographic regions. In order to obtain rainfall measurements close to the ground, scans must also be carried at low angles (Rico-Ramirez and Cluckie, 2008). Ground targets, such as building, mountains, and vegetation, cause interference when they are in the way of the radar beam. Furthermore, unless they are appropriately corrected, they can lead to overestimating rainfall (Collier, 1996; Dufton and Collier, 2015). Ground interference can also cause occultation of the radar beam when the main part of the beam is blocked, and it can only be corrected if it does not obstruct more than 60% of the beam (Collier, 1996). It is also possible that the targets, usually hills or mountains, partially block the radar beam and can have a more significant effect in overestimating rainfall at low elevation angles. The problem of the data at these angles being so important is that often they give very useful information about rainfall at ground levels (Bech *et al.*, 2003; Cluckie and Rico-Ramirez, 2004).

Measurements closer to the ground are also affected by anomalous propagation (AP). Anomalous propagation occurs when the radar beam is deflected to the ground due to variations in atmospheric temperature, pressure, and humidity (e.g. a layer of warm and dry air over moist and cold air) (Rico-Ramirez and Cluckie, 2008). AP is most usual in nocturnal inversions on clear-nights or after thunderstorms when a severe difference in the atmospheric moisture is present. This condition leads to unwanted returned signals, extending ground echoes, that is challenging to accurately identify due to less predictable location (Collier, 1996; Harrison, Driscoll and Kitchen, 2000; Cho *et al.*, 2006; Rico-Ramirez and Cluckie, 2008; Dufton and Collier, 2015).

AP and ground clutter interference can be dealt with in the radar installation (related to the location and hardware used) and processing the radar data. When the radar is installed, possible ground clutters should be considered, and the radar needs to be positioned in a place where ground clutter is minimal. However, it is impossible to eliminate all of it. Measuring rainfall directing the radar beam at higher altitudes can reduce part of the ground clutter problem. However, rainfall measurements higher in the atmosphere are not necessarily the same as those rainfall observations measured at ground level. Therefore, rainfall measurements closer to the ground are still necessary, and therefore ground clutter echoes have to be taken into account (Collier, 1996). Comparing reflectivity data from different elevation angles is advised to reduce AP. Its conditions often influence lower elevation angles, and the reflectivity presents a higher spatial variability than precipitation. However, differently than ground clutter, AP also presents higher temporal evolution that can be similar to precipitation, including the existence of growth and decay comparable to storms (Steiner and Smith, 2002).

As ground clutter echoes are usually permanent, having reference radar data permits that a ground clutter map can be used to identify them. However, it has a high false alarm rate and fails to detect

anomalous propagation accurately. Giuli *et al.* (1991) proposed using the spatial variability of the differential reflectivity (Z_{DR}) to identify ground clutter, as it presents a larger Z_{DR} than rainfall alone. Rico-Ramirez & Cluckie (2008) used fuzzy and Bayes classifiers to successfully identify ground clutter and anomalous propagation (AP). Although the systems were training using dual-polarization radar data, the classifiers can also be used when only single-polarization data is available by combining Z, the radial velocity V, which is around zero for ground clutter, and the clutter map. Their results also corroborate with Giuli et al. (1991) findings about the importance of the Z_{DR} spatial variability in identifying ground clutter. A remaining challenge in the correction methods for clutter elimination rests on the need to assure that precipitation echoes are not removed, resulting in underestimating rainfall (Harrison, Driscoll and Kitchen, 2000; Cho *et al.*, 2006). Steiner and Smith (2002) proposed an algorithm based on the vertical gradient of the reflectivity and Z_{DR} spatial variability in polar coordinates. A decision tree is used pixel by pixel and to assess the occurrence of rainfall. Neural networks (NN) can be used to identify AP. However, they need to use long data sets for training. NN has the potential to be used for operational identification, but NNs are black-box systems and do not provide information about the classification process (Grecu and Krajewski, 2000; Steiner and Smith, 2002). Grecu and Krajewski (2000) used neural networks to identify clutter based on volume scan reflectivity observations. Using fuzzy logic is also valuable to identify ground clutter and AP. Many researchers have studied this approach (Kessinger, Ellis and Van Andel, 1999; Berenguer *et al.*, 2006; Cho *et al.*, 2006; Gourley, Tabary and Parent du Chatelet, 2007; Dufton and Collier, 2015). Kessinger, Ellis and Van Andel (1999) proposed algorithms using reflectivity, radial velocity and spectrum width as inputs to detect ground clutter and anomalous propagation. Cho *et al.* (2006) proposed a fuzzy logic function where its memberships and proper weights are derived from statistics of the reflectivity standard deviation, the absolute value of radial velocity and the vertical gradient of reflectivity. It provides better results in identifying ground clutter and AP compared with reflectivity-independent approaches that can remove precipitation echoes in heavy rainfall events.

2.2.1.3 Variation in the vertical profile of radar reflectivity

Variation in the vertical profile of reflectivity (VPR) is related to changes in size and shape distribution of hydrometeors (Krajewski *et al.*, 2011) and can be manifested in different forms of errors. VPR uncertainties account, along with beam-propagation, is one of the most challenging error sources in radar rainfall estimation (Bech, Gjertsen and Haase, 2007). As the distance from the radar increases, there is also an enlargement of the measurements' sampling volume. The velocity of hail, snow, and water are also different (due to differences in the density and shape of these hydrometeors and the difference in the density of the atmosphere due to altitude variations) (Rinehart, 1997). High reflectivity can be a result of the presence of melting snow, hailstones and very large raindrops (Rico-Ramirez *et*

al., 2007). At a higher altitude, the distribution of hydrometeors changes, leading to a difference between measured rainfall and rainfall that falls in the ground. This variation is caused by factors such as evaporation of raindrops below the radar beam, growth of precipitation, melting of snow and hail, orographic enhancement (growth of precipitation over hills at low levels that can trigger convective events) and wind effects (Harrison, Driscoll and Kitchen, 2000; Rico-Ramirez and Cluckie, 2007; Foresti and Pozdnoukhov, 2012; Schellart, Shepherd and Saul, 2012). Figure 2-11 shows an idealized VPR of the atmosphere in stratiform precipitation (Harrison, Driscoll and Kitchen, 2000).

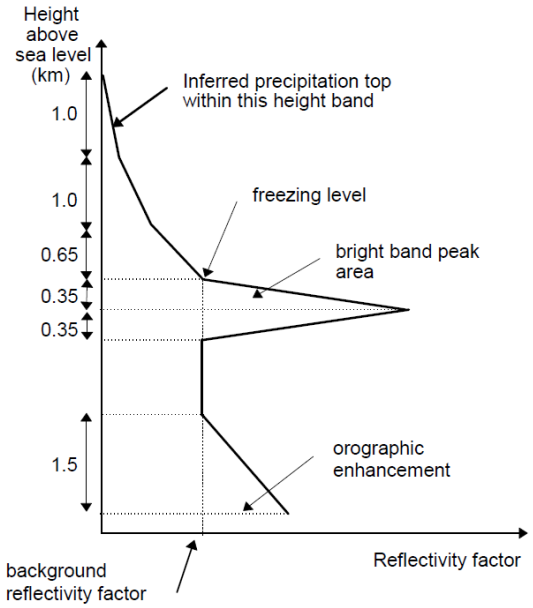


Figure 2-11. Idealized vertical reflectivity profile (Harrison, Driscoll and Kitchen, 2000)

At higher altitudes, snow falls quite slowly until it reaches the freezing level, where it starts to melt slowly. The bright band (BB) area is the region where the melting of snowflakes occurs. In this region, the outer part of the hydrometeor is formed by water, but the inside remains as snow, and it is seen by the radar as a large raindrop with reduced velocity. The scattering of microwaves enhances the reflectivity to the radar and can cause overestimation of rainfall. As the hydrometeors keep falling, their size becomes smaller, and the velocity increases, leading to a decrease in reflectivity. At a longer distance from the radar, part of the sampling volume could be in the rain in the lower part as the upper part of the sampling volume could be in the bright band area or even be in a region without precipitation, affecting the reflectivity measurements and, consequently, the rainfall estimation (Rico-Ramirez and Cluckie, 2007). Errors due to the radar beam intercepting the BB can lead to overestimating precipitation up to 5 times the actual rainfall. In cases where the radar beam is above the BB, there is an underestimation of rainfall, and uncertainties can be up to 4 times less than actual precipitation per kilometre (Rico-Ramirez and Cluckie, 2007). In Figure 2-12, an example from a typical VPR pattern

during an event with stratiform precipitation is shown. The bright band, where the melting of snow takes place, in this case, is centred at 1.8 km (it can be identified by the high reflectivity values). In the region below the bright band, the reflectivity is due to raindrops, and above the bright band, the reflectivity measured is due to snowflakes (Rico-Ramirez and Cluckie, 2007).

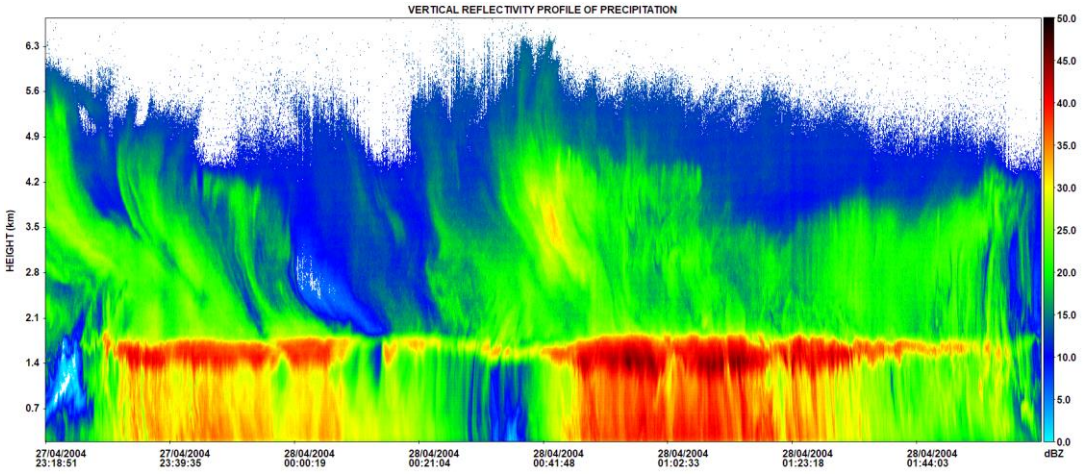


Figure 2-12. VPR during stratiform precipitation. The melting of snowflakes takes place around 1.8 km (bright band). Rainfall is present below the bright band. Reflectivity above the bright band is due to the presence of snowflakes, and below it due to raindrops (Rico-Ramirez *et al.*, 2007).

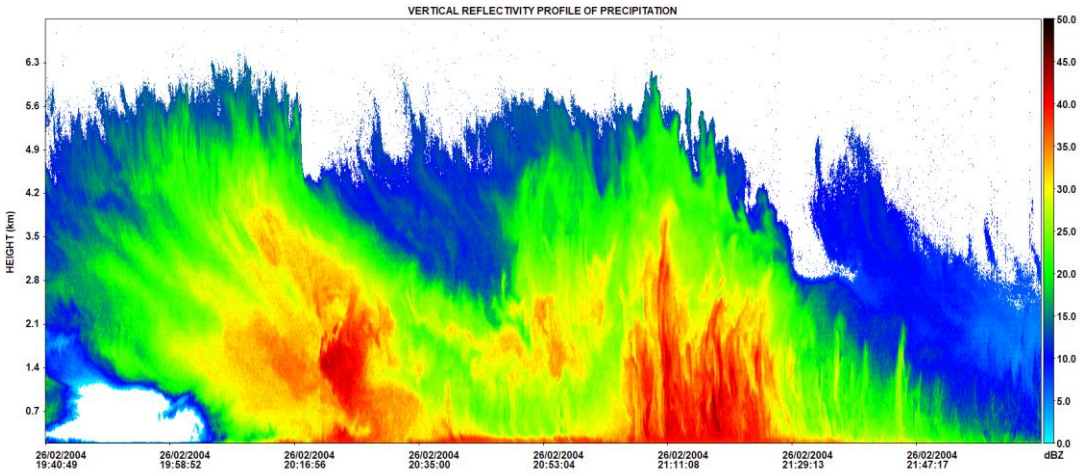


Figure 2-13. VPR during convective precipitation. The BB does not occur at a specific height. It can occur over several kilometres (Rico-Ramirez *et al.*, 2007).

In Figure 2-13, an example of the VPR on a convective event can be seen. It differs from a stratiform event as it does not present a typical BB form, and the reflectivity tends to be higher from the ground to several kilometres above. During convective events, its often difficult to identify the bright band using reflectivity measurements (Rico-Ramirez *et al.*, 2007). However, dual-polarisation measurements can help to identify precipitation particles, including the bright band. For instance, Gray et al. (2002) found a set of typical VPR profiles and used them to estimate rainfall reflectivity at ground level. Rico-

Ramirez et al. (2005) proposed using a fuzzy logic system (FLS) to classify rain, snow and melting snow. Krajewski et al. (2011) developed a method to describe the structure of the VPR concerning the altitude-related to the elevation angle of the radar.

2.2.1.4 Range degradation

Range degradations are significant at long ranges, leading to underestimation of rainfall. It includes partial beam filling, rainfall overshooting and vertical gradient of reflectivity. Partial beam filling (when the radar beam is partially blocked) can reduce estimates at far ranges, and this is due to the radar beam geometry. At far ranges, intense small features are averaged, and as a result, the reflectivity measured by the radar is biased (Villarini and Krajewski, 2010). However, this is a less significant issue when compared to the other range degradation uncertainties (Fulton *et al.*, 1998; Harrison, Driscoll and Kitchen, 2000; Villarini and Krajewski, 2010). A higher sampled volume means that readings of smaller and intense features are averaged and also that readings include different hydrometeors, consequently producing biased products (Villarini and Krajewski, 2010). It is common for the radar beam to come across the bright band at far ranges and during summer months. Although corrections can usually remove errors at these ranges, sometimes overcorrections can occur (Harrison, Driscoll and Kitchen, 2000).

During radar operation, in order to make measurements at higher distances from the radar, the radar beam height also increases. This is due to the earth's surface curvature and a higher elevation angle in the scan (Hunter, 1996). In situations of shallow rain (more often in cold seasons) or in locations that beam blockages have a more significant interference in low angles, the precipitation rate can be underestimated and, in some cases, cannot be detected by the radar (Hunter, 1996; Fulton *et al.*, 1998; Harrison, Driscoll and Kitchen, 2000). When there are signals detected from lower angles, it is possible to use extrapolation techniques to measure rainfall (Hunter, 1996; Fulton *et al.*, 1998). In situations where precipitation cannot be detected, the correction is more challenging (Fulton *et al.*, 1998).

2.2.1.5 Reflectivity-rainfall (Z-R) relationships

Z-R relationships depend on the drop size distribution of rainfall. As discussed in section 2.2, Z is more sensitive to large drops than R is. Experimentally measuring the distribution of drop sizes and their statistical correlation to the reflectivity is essential to estimate rainfall. Hundreds of Z-R relationships were determined for distinct locations, types of storms and time. As the difference between them are relatively small, only a few different Z-R relationships need to be used (Rinehart, 1997). Although the variation between storms, situations, and locations still contribute to rainfall estimation uncertainties. Melting snowflakes produce higher reflectivity when compared to convective rain, and the difference is even higher if compared to drizzle from shallow clouds, even when the rain intensity is the same. It

is common to use a constant Z-R relationship to estimate rainfall, leading to increased uncertainties on radar rainfall estimations.

Nevertheless, even with promising results, there are still progress to be made regarding raindrop size distribution and shape and estimations for light rain (Ryzhkov *et al.*, 2005). Hasan et al. (2014) developed an error model to assess how rain gauge uncertainties impact the *Z-R* equation. Although the model was shown to improve rainfall estimation, the difference is less than 5%. Dual-polarisation radars have enabled the use of different rainfall algorithms that combine not only reflectivity measurements but also phase measurements (which are immune to radar signal attenuation) in order to obtain better rainfall rates than using a single Z-R equation (Bringi, Rico-Ramirez and Thurai, 2011).

2.2.1.6 Signal attenuation

Attenuation is the loss of electromagnetic waves when they pass through any kind of material. As can be expected, some materials can reduce the radiation more than others (Rinehart, 1997). In rainfall estimation applications, attenuation is more significant during heavy rainfall events. Attenuation is not significant for radar using longer wavelengths (S-band radars) and can be compensated (Collier, 1996; Islam *et al.*, 2014). Therefore, using longer wavelengths can reduce attenuation (Steiner and Smith, 2002). Nonetheless, this approach has the downside of deteriorating the relationship between rainfall and ground clutter and worsening the resolution (Steiner and Smith, 2002). However, at smaller wavelengths (e.g. X-band or C-band radars), hydrometeors such as raindrops produce higher attenuation. They should be taken into account, as they can significantly alter the rain areas in the presence of attenuation. With a wavelength of 5 cm, C-band radars are operated in Europe in many different locations and are well known to produce rainfall estimations with uncertainties related to attenuation (Smyth and Illingworth, 1998). Table 2-2 summarizes the uncertainties due to the attenuation at wavelengths of 5 cm (Collier, 1996). Although snowflakes produce a lower attenuation than raindrops in light precipitation events, it is difficult to relate the size of those hydrometeors to attenuation (Collier, 1996). In cases where attenuation is excessive, it can be identified in reflectivity maps, but in the case of attenuation of only a few decibels, this task becomes difficult (Gorgucci, Scarchill and Chandrasekar, 1996).

Using dual-polarised radars offer advantages in identifying regions affected by attenuation. Smyth & Illingworth (1998) used the difference between horizontal and vertical differential phase shift to correct for attenuation, which is numerically stable and with a low sensitivity to hail. Bringi, Keenan and Chandrasekar (2001) also proposed an algorithm to correct attenuation using dual-polarisation radars using a constraint-based algorithm that is useful in stabilising attenuation corrections based on differential propagation phase. An interpolation method using data from S-band and X-band radars was

proposed by Lengfeld *et al.* (2016) to address uncertainties on X-band radars. The method showed that the method could be particularly useful in urban areas in nowcasting applications. Delrieu, Caoudal and Creutin (1997) study the use of mountain returns to reduce attenuation uncertainties and found the method to provide satisfactory results. Islam et al. (2014) calculated the differential propagation phase constraint varying the bright band's location, calculated by an NWP model. The constraint is then used in a correction algorithm that is valid also to correct attenuation above the bright band. The radar data processing algorithm used by the Met Office includes attenuation correction. However, it can be unstable in cases of severe attenuation, and further uncertainties can occur with attenuation correction in cases where the weather radars are not correctly calibrated (Harrison, Driscoll and Kitchen, 2000). Rico-Ramirez (2012) summarised different algorithms used in the literature to correct attenuation. They can be primarily classified into algorithms that use reflectivity measurements only and algorithms that use both reflectivity and phase measurements. Their results showed that algorithms using phase measurements could successfully correct *Z* for rain attenuation when comparing attenuation-corrected reflectivity measurements with reflectivity measurements from a nearby radar not affected by attenuation.

Table 2-2. Order of magnitude of hydrometeors uncertainties due to attenuation at 5cm wavelength (Collier, 1996).

Hydrometer type	Rainfall type	One-way attenuation effect
Clouds (non-precipitating)	Liquid	0.03 dB km^{-1} (g m^{-3})$^{-1}$
	Ice	0.0015 dB km^{-1} (g m^{-3})$^{-1}$
Snow	Dry	0.0012 dB km^{-1} at maximum rate 3 mm h^{-1} (i.e. ≈30 mm snow h^{-1})
	Wet	5 – 50 times dry snow
Rain	Wet	0.003 dB km^{-1} (mm h^{-1})$^{-1}$
Hail	Dry	0.025 – 0.6 dB km^{-1}
	Wet	0.25 – 3.5 dB km^{-1}
Fog	Visibility, 30 m	0.013 dB km^{-1}
	Visibility, 300 m	0.004 dB km^{-1}
(Gaseous atmosphere)		0.008 dB km^{-1}

2.3 Radar and rain gauge merging

Using RG (rain gauge) data to improve radar estimations have been used for years, and a variety of methods have been developed. The most traditional methods rely on a bias correction approach. Bias

correction methods adjust the radar field to agree on average with the gauges in terms of bias and can improve accuracy on significant QPE deviations (Harrison, Driscoll and Kitchen, 2000; Overeem, Leijnse and Uijlenhoet, 2013). Methods for bias correction can range from simple scaling to complex distribution mapping, which can improve low and high extreme rainfall estimations (Chen *et al.*, 2013; Teegavarapu, 2014; Lee, Kim and Suk, 2015). Brandes (1975) developed one of the first techniques to combine radar and rain gauges measurements. This was done using a factor based on the ratio of rain gauges and radar measurements, and subsequently, it was applied to the radar domain to correct rainfall measurements. Issues related to isolated rainfall events, attenuation and beam blockage were not adequately addressed by this method. Since then, a number of different methods for merging radar and rain gauges have been developed. Geospatial interpolation methods use weighting factors to produce a merged QPE product that considers the spatial variability of precipitation. It is important to highlight that the computational cost used by geospatial interpolation methods is much higher, and the time taken to run those methods has to be taken into account in real-time applications (Jewell and Gaussiat, 2015).

There are several geostatistical interpolation methods available in the literature and applied to different types of data sets. Ordinary kriging (OK) is an interpolation technique that can be used to interpolate a set of point measurements in order to estimate the values at unknown locations. These point measurements can come from a rain gauge network to estimate the rainfall at unsampled locations. The advantage of this technique is that it takes into account the spatial correlation of the measurements to perform the interpolation. In this way, the rainfall estimation at unknown locations will be correlated to the nearest available measurement based on the spatial correlation among the available measurements. Although OK does not merge rain gauge with radar data, its principles are used in many kriging-based methods, and a short description of the method will be discussed in section 6.2.1. Radar-gauge based merging methods that use kriging take advantage of the spatial interpolation in several ways. There are two main categories of kriging merging methods. In the first category, two interpolation fields are produced (one with rain gauges data and one with radar data), and the bias between them is used to correct the radar field (Jewell and Gaussiat, 2015). Examples of this approach are conditional merging (Pegram, 2004; Sinclair and Pegram, 2005) and radar-based error correction (Jewell and Gaussiat, 2015). The second category merges the radar and rain gauge datasets first and then generates interpolated rainfall fields by using weights to adjust the radar-based rainfall estimations at ungauged locations (Jewell and Gaussiat, 2015). This approach was used by Krajewski (1987) and Creutin, Delrieu and Lebel (1988), who proposed the use of co-kriging (CK) methods to merge radar and rain gauge data. CK only uses radar data only at rain gauge locations, so the advantage of having a high spatial resolution is not fully explored, and the spatial structure of rainfall is not taken into account. In the past two

decades, kriging with external drift (KED) has been developed in radar rainfall applications. KED combines rain gauge data with other additional information such as radar rainfall, satellite rainfall, terrain elevation, etc., to generate the interpolated rainfall fields, and the technique has been successfully applied by the radar research community (Grimes, Pardo-Igúzquiza and Bonifacio, 1999; Velasco-Forero *et al.*, 2004, 2009; Haberlandt, 2007; Berndt, Rabiei and Haberlandt, 2014; Jewell and Gaussiat, 2015). In KED radar rainfall applications, rain gauge data is merged with radar data. In KED, the spatial distribution of precipitation is primarily based on radar rainfall estimations, while the accuracy of the rainfall estimates relies mainly on the rain gauge measurements. KED takes into account the spatial correlation of the rainfall field but does not account for temporal correlations. Sideris *et al.* (2014) proposed the use of co-kriging with external drift (CED) to overcome this problem. CED is more advanced than KED because it takes into account not only spatial information, but also temporal information as secondary co-kriged variables. They found that this method produced better results than KED, especially at shorter accumulation periods of 10min, where the temporal correlation of the rainfall field is significant. However, at longer accumulation periods of 60min, the errors in CED are very similar to the errors in KED. CED, however, is computationally more expensive than KED, which is important to consider for real-time applications. Another method that uses block-kriging and Kalman filtering to merge radar and rain gauge measurements was proposed by Todini (2001) using a Bayesian approach. This technique was an improvement in rainfall estimations, and it also was able to reduce noise variations.

The rain gauge network density has a major influence when using kriging-based interpolation methods to estimate the rainfall field. The gauge network density is usually a balance between the installation and maintenance costs to maintain them and the number of gauges required to appropriately represent the spatial distribution of precipitation within the catchment area. Besides the network density, it is also important to take into account that rain gauges can present problems and be unavailable for some time. Jewell and Gaussiat (2015) carried a comparative study for merging techniques and ordinary kriging for different rain gauge network densities, concluding that all of them are highly dependent on the number of rain gauges available. The effect of the gauge network density is even more noticeable at high thresholds but merging radar and rain gauge data still produces superior results than any of these data sets alone. Goudenhoofdt and Delobbe, (2009) found that the geostatistical radar-gauge merging methods are more sensitive to the gauge network density than simple methods such as mean field bias correction. Similar findings were reported by Nanding et al. (2015). They also found that simple mean field bias correction methods can outperform geostatistical methods for a low-density rain gauge network.

Rain gauge and radar merging is a powerful tool to improve rainfall estimation as it has the advantage of combining the strength of both measurement techniques: higher accuracy from rain gauges and high spatial distribution of precipitation from weather radars. Chapter 6 presented a method to merge radar and rain gauge measurements using kriging with external drift (KED) for different accumulation periods. KED is a robust method to merge rainfall data and has been widely used to improve radar rainfall estimation accuracy (Jewell and Gaussiat, 2015; Cecinati, 2017; Ochoa-Rodríguez *et al.*, 2019). The previous chapter's results agree with the literature regarding that even at short accumulation periods, using KED still produces better rainfall estimations than using rain gauge or radar data alone. The radar-gauge merging product's quality suggests that KED merging could be used instead of radar rainfall data alone as input to nowcasting models. However, the need for high temporal resolution of nowcast models imposes a challenge in using KED for short-term rainfall forecasting. In addition to this, the temporal correlation of rainfall is not taken into account during the KED merging, and this can produce distorted rainfall advection fields that might not be correlated in time, causing problems to estimate the advection field by the nowcasting model (Keller, 2013; Ochoa-Rodríguez *et al.*, 2013).

The importance of temporal correlation of precipitation fields has been studied, and some methods to overcome this issue have been proposed. Sideris *et al.* (2014) developed a method called co-kriging with external drift (CED) to merge radar rainfall and rain gauge data, including information not only from the current time step *t* but also from previous time steps (e.g. *t*-1) to improve the rainfall product and indirectly taking into account the temporal correlation of the precipitation field. The model showed to perform better for short accumulation periods. However, for hourly accumulation periods, both KED and CED have similar performances. CED shows a significant benefit over KED in events with a high temporal correlation between time steps. As expected, the model is more complex than KED and demand more computing power and time to solve the equations. A study carried out by Keller, (2013) showed that in events where small rain cells are not detected by rain gauges and are estimated only by radar, CED could not replicate the rainfall advection fields satisfactory even when information from the previous time step is available. The analysis of CED and KED showed that both methods performed similarly to reproduce the rainfall advection fields. This study used hourly accumulations and did not find an advantage in using CED over KED, meaning that the method's extra complexity could not justify its use for this accumulation period. Keller (2013) assessed that the temporal evolution issue in both KED and CED is a limitation for the use of these methods in nowcasting.

2.4 Short-term rainfall forecasts (radar nowcasts)

Once the radar data have been corrected for the different error sources and adjusted with rain gauge measurements, a radar-based rainfall forecasting model can be applied. Quantitative Precipitation Forecasts (QPFs) can be produced with Numerical Weather Prediction (NWP) Models (with a forecasting lead time of a few hours to days) and with radar-based extrapolation, also known as nowcasting (1-6 hr forecasting lead time). Rainfall forecasts can refer to both short-term and long-term rainfall forecast, while nowcast is a short-term rainfall forecast, with only a few hours lead-time. NWP forecasts have a low performance at the beginning of the forecast, but the performance remains constant with forecasting lead time. However, the performance of radar-based forecasts (nowcasts) is high at the beginning of the forecast but decrease with forecasting lead time. Nowcasting models are based on the extrapolation of Lagrangian trajectories that can produce rainfall forecasts with a few hours lead-time (with increased loss of performance after 2–3 hr ahead). They have a significant role in enhancing rainfall warning systems, especially when predicting extreme events or flash floods (Zahraei *et al.*, 2012; Liguori and Rico-Ramirez, 2013b). Nowcasting models use a sequence of weather radar scans in order to produce an advection field that can be used to extrapolate the latest radar rainfall scan into the future. The generated precipitation forecasts produced with this technique have the same spatial and temporal resolutions as the original radar rainfall fields. The performance of the forecasts is high at the beginning of the forecast (Bowler, Pierce and Seed, 2006; Liguori and Rico-Ramirez, 2013b). However, uncertainties intrinsic to the nowcasting model result in an increasing forecasting skill loss after 1 hr lead-time. Uncertainties in radar nowcasts are caused mainly by (Seed, 2003; Liguori and Rico-Ramirez, 2013b; Foresti and Seed, 2014): uncertainties inherent to the specific nowcasting model used, errors in radar rainfall estimation, uncertainties due to the temporal development of the velocity field and uncertainties caused by precipitation processes such as growth and decay not being taken into account. The uncertainties related to the rainfall estimation account for a loss of accuracy in nowcasting up to 1 hr lead time (Foresti *et al.*, 2013). A description of the main errors in radar rainfall estimation was given in section 2.2.1.

Assuming that rainfall advection fields are homogeneous over the forecasted domain and that the rainfall fields are advected instead of evolving during the forecast are significant limitations in nowcasting models. Nowcasting models use a recent sequence of radar rainfall images to calculate how the rainfall evolves over time. However, they are not able to calculate how the rainfall velocities and changes in intensity happen in the future (Bowler, Pierce and Seed, 2006; Berenguer, Sempere-Torres and Pegram, 2011; Atencia and Zawadzki, 2014). The radar images are used to extrapolate recent rainfall measurement in order to compute the rainfall forecasts. That means that both growth and

decays of rainfall and temporal evolution of the rainfall advection fields are not accounted for by the nowcasting model. The non-stationary in space and time characteristic of rainfall make these factors even more critical in cases of convective precipitation and in large domains (Bowler, Pierce and Seed, 2006; Liguori and Rico-Ramirez, 2013b). Lagrangian nowcast models assume that the velocity fields do not change during the forecast length (Bowler, Pierce and Seed, 2006; Berenguer, Sempere-Torres and Pegram, 2011; Atencia and Zawadzki, 2014; Rossi *et al.*, 2015). Accounting for these uncertainties is still challenging, even with extensive research that has been done to assess how to reduce these errors (Seed, Pierce and Norman, 2013) and how to account for precipitation growth and decay.

There are several nowcasting techniques available in the literature:

- Tracking radar echoes by correlation – TREC (Rinehart and Garvey, 1978) and COTREC (Li, Schmid and Joss, 1995).
- Tracking of rain cell centroids (Johnson *et al.*, 1998).
- VET, Optical flow techniques (Pierce *et al.*, 2000; Bowler, Pierce and Seed, 2004)
- Use of NWP advection techniques (Toth and Kalnay, 1997; Golding *et al.*, 1998).
- Blending techniques (Nowcasting + NWP forecasts): STEPS (Short-term ensemble prediction system) (Bowler, Pierce and Seed, 2006).

Rinehart and Garvey (1978) developed a pattern recognition model based on a correlation coefficient to calculate motion vectors in storms known as TREC (tracking radar echo with correlation). The TREC algorithm has been modified, and new models are based on this approach. Li, Schmid and Joss (1995) proposed improvements for the TREC model by correcting some wrong vectors that are caused by shielding, clutter and some random errors in the radar rainfall estimation and investigating growth and decay process based on smoothed fields of echoes motion. These changes allowed a better forecasting skill of orographic rainfall with up to 20 min lead time. Sokol *et al.* (2013) compared forecasts produced using two different models. The COTREC model is based only on extrapolation from radar images and assumes that rainfall trajectories do not change with time. The SAMR model (Statistical Advective Method Radar) utilizes the same technique and factors in a statistical model to correct precipitation estimations and was designed for being used in the warm period of the year. Results showed that SAMR provides slightly better results, but it cannot predict new storms and is unable to forecast any significant changes in existing storms accurately (Sokol *et al.*, 2013). Berenguer, Sempere-Torres and Pegram, (2011) also use TREC and COTREC algorithms to generate forecasts and produce probabilistic forecasts. SBMcast uses the String of Beads model to maintain space and time structures. Algorithms that

determine the movement of storms based on cross-correlation or use storm-cell centroid techniques have been developed over decades. The first is efficient in providing information over a larger area, while the second is better tracking isolated storms. The Storm Cell Identification and Tracking (SCIT) algorithm proposed by Johnson *et al.* (1998) is based on tracking techniques and identifying rain cell centroids. The SCIT algorithm is able to identify, track, characterize and forecast storm cells movements. Although the algorithm can forecast efficiently isolated and well-organized storms, it is unable to detect small, shallow cells or cells with a maximum reflectivity of 30 dBZ (Johnson *et al.*, 1998). The GANDOLF scheme (Generating Advanced Nowcasts for Deployment in Operational Land-based Forecasts) was developed to increase convective rainfall predictability. An object-oriented model of convection is used that incorporates a model of the life cycle of convective clouds (Pierce *et al.*, 2000). This advection scheme showed deficiencies in severe rainfall events. An attempt to address this issue was made by dividing the rain analysis into blocks and forcing adjacent blocks to have a smooth transitional variation of the velocity. NWP models are useful for forecasting precipitation at longer lead-times than nowcasts as they are based on modelling the changes in the atmospheric state using meteorological variables such as temperature, pressure, humidity, winds, etc. NWP models determine large and slow evolving scales and use the local details in parameterisation or statistics. NWP models have limited resolution, and as a consequence, they have a lower initial skill at the beginning of the forecast than nowcasting models. However, NWP does not suffer from the same loss of skill as nowcasting models, and the skill remains more or less constant with a longer forecasting time (Golding *et al.*, 1998; Sokol *et al.*, 2013) as it is able to model precipitation growth and decay. NWP models require high computing power, and although the forecast resolutions are increasing (Sokol *et al.*, 2013; Simonin *et al.*, 2017), they are still more computationally demanding than nowcasting models. Figure 2-14 illustrates the loss of predictability with lead time for both NWP and nowcasting models. Nowcasting models start with high initial skill or performance. However, their predictability is lost rapidly with lead time as the temporal evolution of rainfall and growth and decays processes are not resolved. Bray *et al.* (2011) studied uncertainties sources in NWP models. They showed that the domain size and the buffer zone distance have are sources of uncertainty that can be easily overlooked as they are not subject to particular guidelines. They showed that these uncertainties are particularly important for lead-times shorter than 24 hours and state the need for long warm-up periods for NWP forecasts. Bowler, Pierce and Seed (2004) derived a new optical flow algorithm, enhancing the GANDOLF system's capability to calculate the advection field. This algorithm was used to further develop the STEPS model (Short-Term Ensemble Prediction System) (Bowler, Pierce and Seed, 2006). In the STEPS model, ensemble radar nowcasts are blended with Numerical weather prediction (NWP) forecasts. NWP models have been shown to have better forecasting skill after several hours of lead-time. They can improve the ability to forecast growth

and decay of precipitation when blended with a nowcasting model (Bowler, Pierce and Seed, 2006). Blending nowcasting with NWP forecasts have been shown to improve the forecasting skill (Liguori *et al.*, 2012), and it has been successfully used in real-time applications (Bowler, Pierce, and Seed, 2006; Seed, Pierce and Norman, 2013). The uncertainties in radar rainfall analysis and temporal evolution of precipitation were accounted for using a stochastic perturbation system, in which probabilistic forecasts are produced by adding spatially correlated stochastic noise to the deterministic forecast. Although the STEPS performance is higher than GANDOLF's, moderate and, heavy rain results still do not match up accurately with the observed precipitation. A newer version of STEPS has been developed to take into account radar errors (Seed, Pierce and Norman, 2013), using a statistical model to generate ensembles proposed by Germann *et al.* (2009).

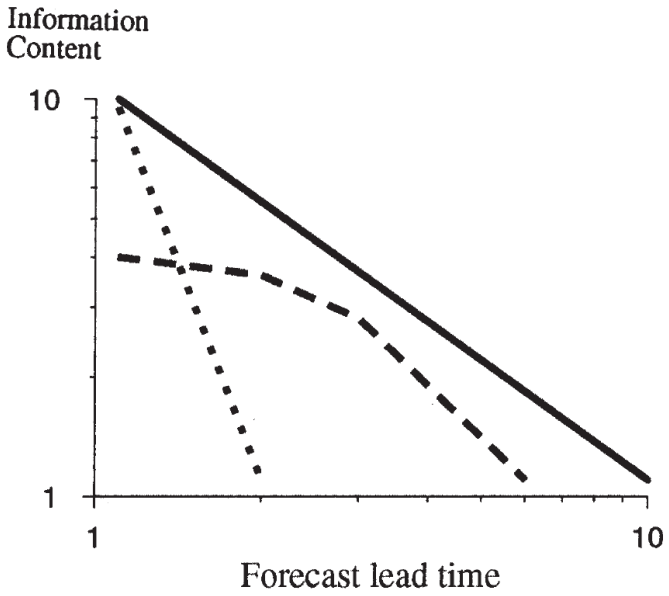


Figure 2-14. Schematic representation of loss of information content in relation to lead time. The dashed, doted, and solid lines represent, respectively, NWP models, nowcasting models and theoretical limit of predictability (Golding *et al.*, 1998).

2.5 Concluding comments

This chapter described the different techniques to measure precipitation. The advantages in the measurement of precipitation with weather radar are in terms of high spatial and temporal resolutions, whereas rain gauge can provide point measurements with reasonable accuracy. Radar rainfall can be affected by different error sources. Advances in weather radar systems and algorithm development have been able to minimise some of these error sources in order to improve precipitation estimation. However, even with the advance in technology to measure rainfall and merging techniques that take into account the strength of different methods, residual errors often remain and have to be taken into

account when producing short-term forecasts. Polarimetric radars potentially can improve the quality of estimations. However, its use is still limited as the relationship between polarimetric radar variables and estate variables are not linear (Zhang *et al.*, 2019). Measurement errors not only imply rainfall estimation, but they also propagate into the forecast models. It is important to take into account these errors in a probabilistic way to account for the uncertainty in the measurements (e.g. chapter 4) and also to try to reduce this uncertainty (e.g. chapter 6).

Nowcasting models can produce short-term radar-based forecasts that have a good performance at the beginning of the forecast, but their performance decrease with forecasting lead time. This is due to the fact that precipitation growth and decay is not accounted for by nowcasting models. The use of ensemble forecasting techniques can help account for the uncertainties in radar-based precipitation forecasting, as shown in the following chapters. However, there are still many challenges to improve radar precipitation measurement and forecasting. Blending nowcast with NWP is helpful to increase lead-time in forecasts as the atmospheric state is considered in the forecast. However, NWP also is subject to its sources of uncertainties, and its forecast skill is better hours after the start of the forecast.

Chapter 3. Datasets and study areas

3.1 Introduction

This chapter aims to describe the models and data sets used in this thesis. As described in Chapter 1, this thesis assesses different sources of uncertainties in short-term rainfall forecasts and how it propagates into river/sewer flow prediction models. The meteorological data comprises point rain gauge data and weather radar data for different places in the UK. Section 3.2 will describe the data used for rainfall estimation, including weather radar and rain gauge data. A brief description of the nowcast model used in this research is in section 3.3, and the urban and hydrological models used in different parts of the thesis are introduced in section 3.4.

3.2 Rainfall data

3.2.1 Rain gauge data

The UK has one of the densest rain gauge networks in the world, as a consequence of its dense drainage river network and diversity in climate, geology, water and land-use (Marsh, 2002). The UK Environment Agency (EA), partnered with other organizations, such as the UK Met Office, are able to provide rainfall data from a dense network of rain gauges. Rain gauge data is freely available upon request by the UK Environmental Agency under the Open Government License, whereas the UK Met Office Integrated Archive System (MIDAS) is freely available through the British Atmospheric Data Centre. The EA rain gauge data have a temporal resolution of 15min, whereas the Met Office has a temporal resolution of 1h. A recent study by Lewis *et al.* (2018) developed a methodology to quality control and combine rain gauge data from both networks to produce a continuous data set in space and time at 1km/1h resolutions over Great Britain. Nanding (2016) looked at the number of gauges in different countries and the average area covered for each gauge (Table 3-1), and their results show how well rain gauges cover the British territory. In this thesis, a rain gauge network of 229 tipping bucket rain gauges (TBRs) with 15 min temporal resolution was made available for the study area in the north of England. The rain gauge data was provided by the EA. The location of the rain gauges is available in Figure 3-1.

Table 3-1. Official recent national rain gauge network (Nanding, 2016).

Country	Gauge number	Area per gauge (km^2)
UK	4400	49
France	4300	116
Netherlands	357	99
Switzerland	450	92
Spain	8135	62
Belgium	260	117
Italy	3000	100
Germany	3500	102
South Africa	1500	814
India	2140	1535
Japan	1300	291
China	14000	686
South Korea	520	193

As mentioned in Chapter 1, rain gauges are used as ground-truth; nonetheless, they are subject to different sources of errors. In order to use only reliable rain gauge data, measurements from different rain gauges were compared, and only reliable rain gauge data was used. Rain gauge data that presented significant deviation compared to the surrounding rain gauges or showed an anomalous behaviour (e.g., blockages) were discarded (22 rain gauges in our case). It is also necessary to deal with sampling errors when comparing weather radar and rain gauges data. Rain gauges measurements are point measurements at ground level, whereas weather radars measurements occur at a higher altitude and with a more extensive sampling volume in space (Nanding, Rico-Ramirez, and Han, 2015). Therefore, part of both measurements' discrepancies is due to differences in the sampling volumes (Kitchen and Blackall, 1992). Providing the rain gauge data is quality-controlled, they can be considered ground-true measurements to validate the radar rainfall observations even though the rain gauge measurements represent only a particular point in space.

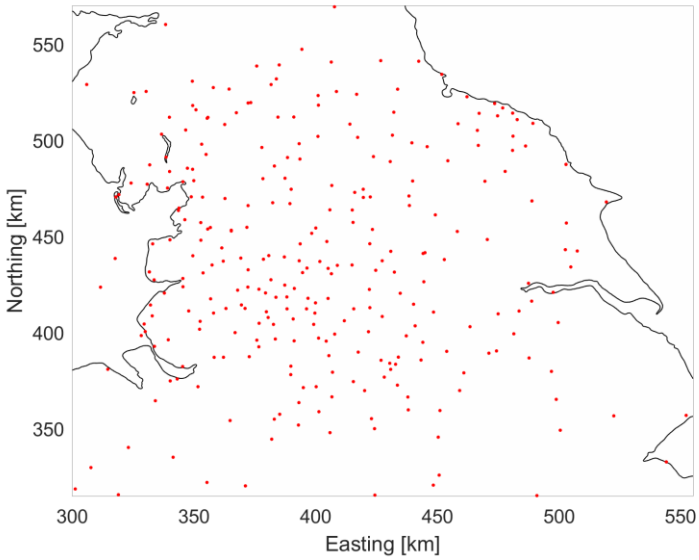


Figure 3-1 Map of the study domain with the rain gauge locations (red circles). The region located in the North of England covers 256 x 256 km^{2,} and the region is also under the coverage of three MetOffice weather radars.

3.2.2 Weather radar data

Composite radar data from the Nimrod system from the UK Met Office, available through the British Atmospheric Data Centre (BADC), were used in this study. The MetOffice operates 15 out of the 18 C-band weather radars existent in the British Isles (Met Office, 2012). The UK Met Office radar network can produce high-resolution precipitation data over the UK with a spatial and temporal resolution of 1km/5min, respectively (Met Office, 2003, 2009). The location of the radars, including the radars in Ireland, is shown in Figure 3-2. The Met Office is also responsible for data processing to reduce uncertainties in the rainfall estimations. A series of volume scans, usually between 0.5 and 4.0 degrees in elevation, are made by each radar every 5 min. The number of scans varies from four to eight and the angle is dependent on the height of the hills around the radar. In distances up to 75 km of the radar, scans have 1 and 2 km resolution and covers over 85% of UK area. From 75 km up to 255 km, the resolution is 5 km. The radars data used in this study has been already processed by the Met Office and have a temporal resolution of 5 min and a spatial resolution of 1 km x 1km. The data set used in this study is from 2007 and 2008 in order to agree with the available flow observations.

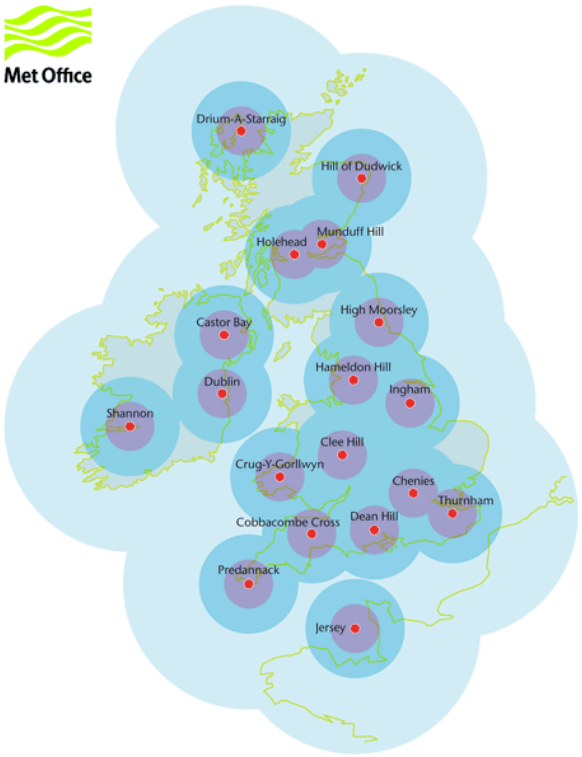


Figure 3-2 Location of weather radars across the British Isles(Scottish Flood Forecasting Service, 2014).

The UK Met Office applies different quality control and correction algorithms to the raw radar data. These corrections include noise filtering, clutter identification, beam blockage and occultation correction, residual spurious echo identification, attenuation correction, correction for vertical variations of the radar reflectivity, conversion from reflectivity to rainfall rate, adjustment for antenna pointing and a mean field bias adjustment based on rain gauge data (see Table 1 in Harrison, Kitchen and Scovell, 2009). Ground clutter is usually observed when the radar beam intercepts high ground, and therefore a small percentage of energy is reflected back to the radar. Because the radar performs the same elevation scans every 5min, then the ground clutter map for each elevation can be computed. Ground clutter frequency maps can be computed during sunny days for each elevation. The ground clutter echoes can then be identified and replaced with reflectivity data from higher elevations not contaminated with ground clutter. The *Z-R* relationship is constant to all rainfall kinds ($a = 200$ and $b = 1.6$). The radar data processing includes an algorithm to correct for rain attenuation, which can be significant at C-band frequencies. The algorithm can be unstable in cases of severe attenuation and further uncertainties can occur with attenuation correction in cases where the weather radars are not properly calibrated (Harrison, Driscoll and Kitchen, 2000). Currently, there are 18 C-band weather radars in the UK (Rico-Ramirez, Liguori and Schellart, 2015). These radars have been upgraded with dual-polarisation technology in the last few years. This will result in significant improvements in terms of data quality (e.g., better identification of non-meteorological echoes), attenuation correction and

rainfall estimation, as demonstrated by several studies (Bringi and Chandrasekar, 2001; Rico-Ramirez and Cluckie, 2008; Bringi, Rico-Ramirez and Thurai, 2011; Rico-Ramirez, 2012). Images from adjacent radars and previous images from the same radar are analysed to discard corrupted images. Anomalous propagation is removed from the radar data by assessing the probability of precipitation using a combination of Meteosat images and reports about the type and amount of clouds and present weather condition (Harrison, Driscoll and Kitchen, 2000). In order to take into account the variation in the vertical reflective profile, an idealized vertical profile, as described in section 2.2.1.3, is identified at each radar pixel. It is defined by the background reflectivity factor and incorporates simple parameterizations. A radar horizon is used to correct the radar beam's occultation (Harrison, Driscoll and Kitchen, 2000). The Met Office radar data processing system also includes correction algorithms for uncertainties due to noise filtering, antenna pointing, mean field bias and conversion from Cartesian to polar coordinates (Harrison et al. 2009)

3.3 Nowcasting model

As discussed in the introduction of this thesis, there are different nowcasting models available in the literature. Some of these models track the centroids of individual storms (e.g. TITAN - Thunderstorm Identification, Tracking, Analysis, and Nowcasting) of several square kilometres in size, whereas the most advanced models are able to track the movement of every rainfall pixel in the rainfall advection field (e.g. STEPS) (Dixon and Wiener, 1993; Bowler, Pierce and Seed, 2006). In this thesis, the STEPS model was used. The STEPS model was provided by the UK Met Office. There is also a freely available version for the research community (Pulkkinen *et al.*, 2019). STEPS is a rainfall-forecasting model that blends rainfall extrapolation nowcasts with NWP rainfall forecasts. The nowcasting module isolates small characteristics (estimation of the advection field, the temporal evolution of rainfall and spectral decomposition) into multiplicative cascades. This ensures that features that cannot be accurately predicted by the model are substituted by stochastic noise. The model assumes that the rate of the temporal evolution of rainfall and temporal development of velocity fields remain stationary during the forecast. Even with the uncertainties inherent to the model, radar nowcasting produces more skilful forecasts than NWP for at least 2 hr., and STEPS uses a multi-cascade approach to blend the two forecasting components (i.e. radar nowcasts with NWP forecasts) (Bowler, Pierce and Seed, 2006; Seed, Pierce and Norman, 2013). For this study, the STEPS model's nowcasting component was used to produce the forecasts up to 3 hr ahead. To account for the uncertainties in the radar rainfall analysis and the temporal evolution of the rainfall advection field, nowcasting models often provide probabilistic or ensemble forecasts. These forecasts are produced by adding spatially-correlated noise to the deterministic forecast. Liguori and Rico-Ramirez (2012) produced probabilistic nowcasts using

the STEPS model. They concluded that a number of ensembles larger than 10–20 members do not effectively increase the forecast's accuracy. In different research papers, a number of ensembles between 20 and 30 are commonly used (Kharin and Zwiers, 2003; Zappa *et al.*, 2010; Panziera *et al.*, 2011; Liguori and Rico-Ramirez, 2013a; Dai *et al.*, 2015; Foresti and Seed, 2015; Lewis *et al.*, 2015) On this research, each probabilistic forecast is formed by 25 ensemble members.

3.3.1 Nowcasting component of the STEPS model

The STEPS model's nowcasting component is based on the Lagrangian extrapolation of radar images to produce deterministic forecasts. Probabilistic forecasts are made by adding spatially-correlated stochastic noise to the deterministic forecast. Uncertainties related to the nowcast model are related to the determination of the advection velocity fields and the evolution of the rainfall advection field. Nowcast models are based on extrapolation, and it assumes that the rainfall advection fields are only advected and does not evolve with time (Seed, 2003). STEPS uses the optical flow algorithm described by Bowler et al. (2004) to produce the advection field. The algorithm uses a sequence of radar images to determine the advection velocity field. Two images are compared by partitioning each image into a series of blocks and determines the motion of each block, and a smoothing filter for each block is applied to improve the forecasting skill (Bowler, Pierce and Seed, 2004). One single advection vector is applied in the entire domain before the optical flow algorithm to avoid fast-moving rain issues. By maximizing the correlation between the newest radar image and the previous one, it is possible to determine the advection vector (Bowler, Pierce and Seed, 2006). Unlike the approach made by Bowler, Pierce and Seed (2004), STEPS uses a back-in-time advection system. The velocity fields are then subjected to an exponential smoothing in time filter (Bowler, Pierce and Seed, 2006):

$$v_{smooth}(t) = \alpha v_{smooth}(t - \Delta t) + (1 - \alpha) v_{estimate}(t) \qquad (3\text{-}1)$$

where v_{smooth} is the temporary smoothed velocity field, $v_{estimate}$ is the original velocity field and α is a parameter that gives the smoothing time-scale. Best performance is achieved when $\alpha = 0.85$; for this value, the time-scale for the smoothing is around 90 min.

Analyzing how the velocity fields vary with time is then used to account for advection uncertainties. After removing the bias from the forecasted velocity, the total diagnosed velocity can be compared to the forecasted velocity (Bowler, Pierce and Seed, 2006):

$$v_{total}(t + t_1) = v_{smooth}(t)\left\{1 - \frac{(t_1 + 1)}{120} f(|v_{smooth}(t)|)\right\} - \frac{1}{t - 1}\sum_{s=1}^{t_1} v_{smooth}(t + s) \qquad (3\text{-}2)$$

where $t_1$is the lead time in minutes.

The uncertainties arising from growth and decays are taken into account by adding stochastic noise to the advection vectors. The Lagrangian temporal evolution of the field is modelled using a second-order autoregressive process (AR-2) (Bowler, Pierce and Seed, 2006; Foresti *et al.*, 2016):

$$\mathrm{Y}^{e}_{k,i,j}(t+t_1) = \emptyset_{k,1}(t)\mathrm{Y}^{e}_{k,i,j}(t+t_1-\Delta t) + \emptyset_{k,2}(t)\mathrm{Y}^{e}_{k,i,j}(t+t_1-2\Delta t) \quad (3\text{-}3)$$

Where $\mathrm{Y}^{e}_{k,i,j}(t+t_1)$ and $\mathrm{Y}^{e}_{k,i,j}(t+t_1-\Delta t)$ are the lagrangian temporal evolutions that have been forwarded in time, $\emptyset_k$ are weights that control the Lagrangian evolution rate at each scale(3-4), Δt is the forecast time-step and t_1 is the lead time.

$$\emptyset_{k,0}(t) = \sqrt{\frac{1+\emptyset_{k,2}(t)}{1-\emptyset_{k,1}(t)}}\left[\left\{1-\emptyset_{k,2}(t)\right\}^2 - \left\{\emptyset_{k,1}(t)\right\}^2\right] \quad (3\text{-}4)$$

The STEPS forecasts combine the nowcasting model's outputs, the NWP forecast, and a stochastic noise component. At the beginning of the forecast, the nowcast will be the main component of the forecast. As the forecasting lead time increases, more weight is given to the NWP forecast than the radar nowcast. Although STEPS also comprises a downscaling NWP model, it is not used in this research, and therefore it will not be further discussed. A more detailed description of the STEPS model can be found in Bowler et al. (2006) and Foresti *et al.* (2016).

The choice of not using the NWP component is due to this research focusing on producing very-short term forecast in this research. Although STEPS blends both nowcast and NWP, using only the nowcast component is also used in real-time applications: the Australian Bureau of Meteorology uses STEPS in a range of applications, including a 30 member nowcast with a temporal resolution of 1 km and lead-time of up to 90 min (Seed, Pierce and Norman, 2013). For short-term forecasts, particularly for up to 2 hours lead time, advection nowcasting models can produce more reliable forecasts than NWP. Also, nowcasts are not as costly or time-consuming as NWP. That is why nowcast is still the optimal solution for producing a very-short term forecast (Seed, Pierce and Norman, 2013).

A sequence of three consecutive radar images with a time-step of 5 min was used to produce forecasts every 5 min with a forecasting lead-time between 3-6 h at 1 km spatial resolution. The forecasts produced in Chapters 4 and 5 cover a domain of 600 x 600 km^2 in size located in North England, as shown in Figure 3-1. For Chapters 6 and 7, a smaller domain with an area of 256 x 256 km^2 is used.

3.4 Study areas

The catchments used in this study are situated in North England. The region has been subject to several flood events over the years, causing economic and property losses and put thousands of people at risk (Convery and Bailey, 2008; Wilby, Beven and Reynard, 2008; Roberts *et al.*, 2009).

Figure 3-3 shows the river and urban catchments location and the rain gauges and radars in the region. The dense gauge network and the availability of three weather radars coupled with available hydrological data were the main reasons for choosing this region to develop this work. In order to assess how the uncertainties propagate into hydrological models, it is also essential to verify how the size of the catchments influences the flow forecast. Small catchments are subjected to higher uncertainties. Compared to observed rainfall, spatial displacement of the forecasted rainfall of only a couple of kilometres could be enough to generate inaccurate forecasts. For this reason, analyzing how the uncertainties propagate in catchments of different sizes are essential to gain further information on the error propagation. For this reason, this thesis explores the use of small urban catchments where sewer flow is one of the main drivers of the flood response, as well as larger catchments where river flows are the primary source of flooding.

This study utilizes the geographic grid reference of the Ordnance National Grid for the radar rainfall data (Ordnance Survey., 2018).

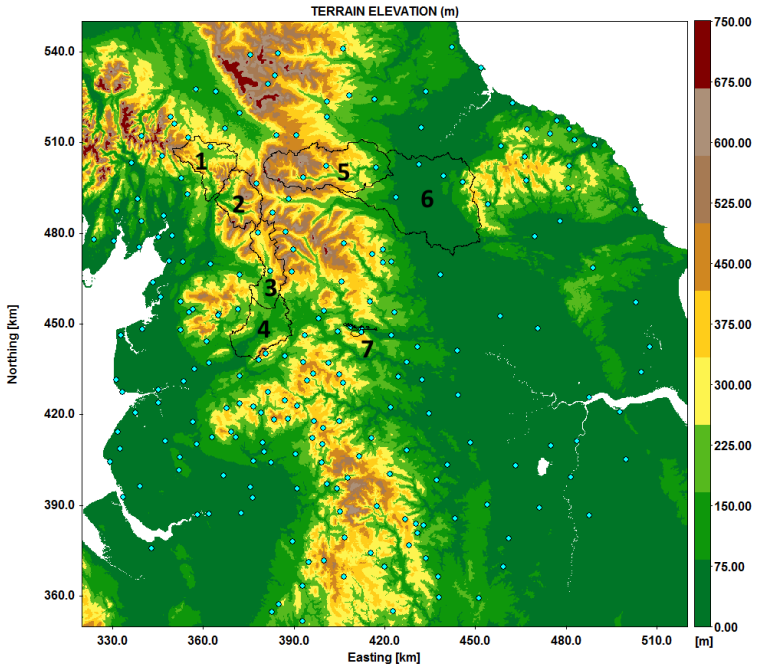


Figure 3-3. Location of the urban catchment (7), the river catchments (1-6) and rain gauges' positions (blue circles) in the study region.

3.4.1 Urban catchment - Ilkley

The rainfall forecast output was used as an input in a hydrodynamic sewer network model built in the Infoworks CS software package to simulate the flows in the sewer network. Rainfall-runoff processes and the flow through the sewer network conduits were modelled utilizing the Infoworks model provided and calibrated by Yorkshire Water for research purposes. The sewer system is mainly combined, being used to carry both wastewater and rainfall-runoff and is approximately 60 km long. The urban area, Ilkley, is located in the Pennine hills and has an area of 11.06 $km^{2,}$ with the majority of this area being permeable land (Figure 3-4). The town area falls into 25 different radar pixels with $1km^2$ resolution, and the model counts with 432 nodes (pipe connections or inlet points), 444 links (conduits), 13 pumps and 134 sub-catchments. The sewer system is shown in Figure 3-4, and this also shows the location of the rain gauges and flow monitors (Liguori *et al.*, 2012). A schematic drawing of the monitoring network is represented in Figure 3-5, giving more details about the pumping network and the flow monitor's location in relation to the pumping stations. Most of the flow monitors are situated upstream from the pump stations (Liguori *et al.*, 2012). Infoworks CS uses both rainfall-runoff volume and runoff routing models to simulate flows in the catchment, and in this study, the New UK percentage Runoff model, the Wallingford model and the Double Linear Reservoir model were used. The full St Venant equations are used in the model to calculate the flows in the sewer conduits (Liguori *et al.*, 2012). Liguori *et al.*, (2012) and Schellart, Shepherd and Saul (2012) provide further information about the Infoworks CS model used in this study. The calibration of the urban hydrological model was previously performed using current industrial standards (WaPUG, 2002) and data from three storm days and one dry day from events that happened between March and April 2000 were used. Data from 5 tipping bucket rain gauges and flow monitors within the urban area were used for calibration.

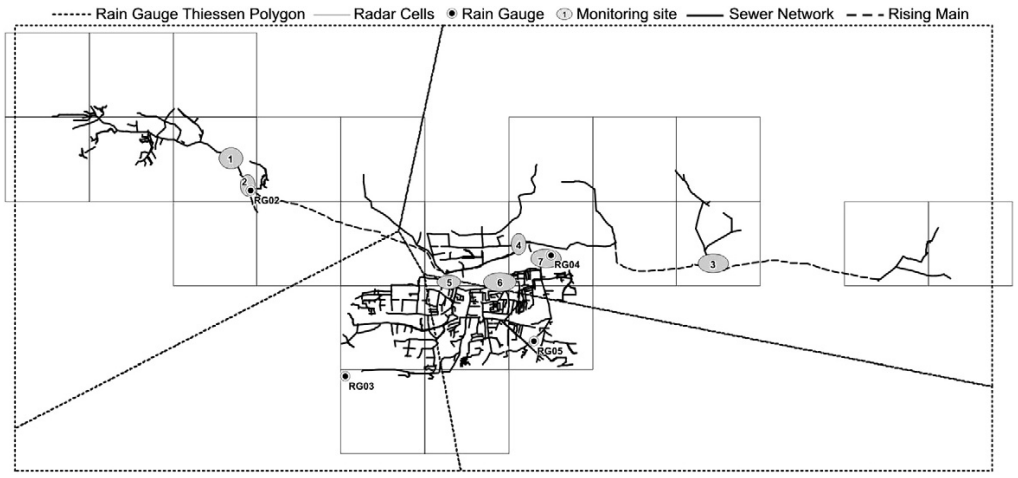


Figure 3-4 – Location of flow monitors (1 to 7) and rain gauges (RG02 to RG05) in the sewer system (Liguori *et al.*, 2012).

Data from 7 depths monitors, 4 additional rain gauges and 16 flow monitors are available from 2007 until 2009 within the urban area (Liguori *et al.*, 2012). In this study, data from 2008 was used for validation. The hydrodynamic model's capacity in simulating flows was assessed using radar and gauge data from 15 April 2008 until 31 December 2008. The period was chosen to include a wide range of events. The calibration and validation of the model was not part of the scope of this research. As the data used was from 2008, data from the second semester of 2007 until the first event was used to simulate the initial conditions. Between each event, flow simulations were carried out to update the initial conditions for the specific event. Radar and gauge data and rainfall forecasts with 5 min temporal resolution were used to simulate and forecast flows.

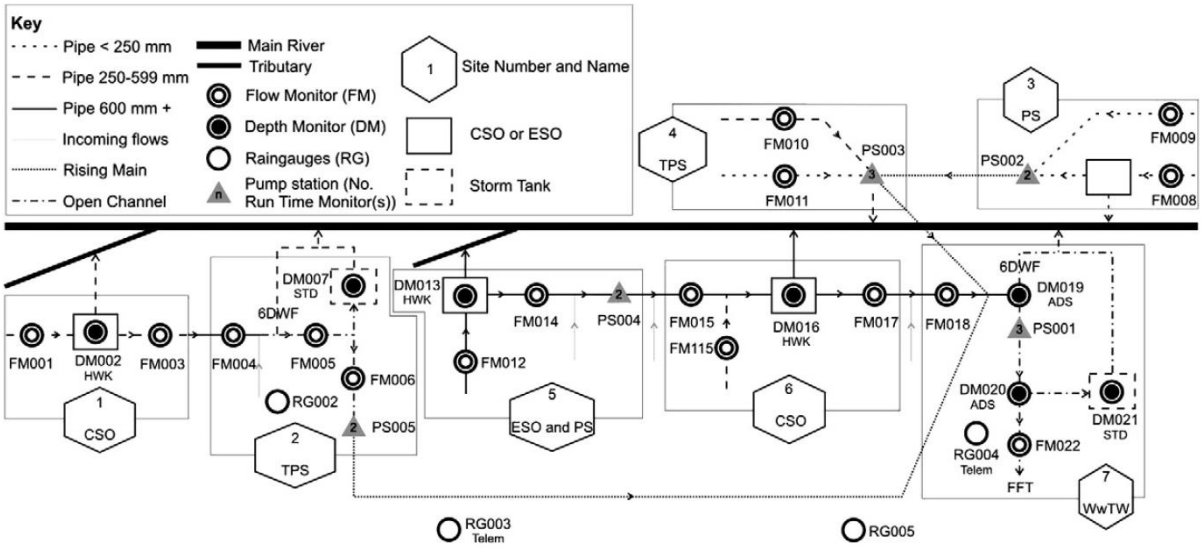


Figure 3-5 – Monitoring network (Liguori *et al.*, 2012).

3.4.2 River catchments

Rainfall forecast output was also used to forecast river flows in six different river catchments using the PDM (probability distributed hydrological model) to represent the rainfall-runoff processes in the catchments. The catchments chosen for this study followed the following criteria: the orographic impacts and snow accumulation should not be significant, and the catchments chosen are near-natural catchments. Near-natural catchments do not directly or significantly impact human activities, such as industrial, agricultural and groundwater abstraction, hydroelectric power plants, and reservoirs that can potentially change the river flow regime (Marsh and Lees, 2003). These criteria ensure the catchments have good river flow data quality to carry this study. The six river catchments studied are located in Cumbria and Lancashire districts.

Table 3-2 shows the name and area of the river catchments, and the position of the catchments and the urban area are presented in Figure 3-3. The mean flow of each catchment can be found in Table

3-3. The table also presents information about the peak flow in the water years of 2007-2008 and 2008-2009. Each water year start in October and ends in September of the following one. Crakehill is the largest catchment studied, and it also has the largest mean flow. However, the catchment has one of the lowest peak flows among the other catchments studied. Results from Table 3-3 show that the catchment's size does not necessarily correspond to the flow intensity in each catchment.

Table 3-2. River catchment's area

Location	Killington	Brigflatts	Arnford	Henthorn	Catterick	Crakehill
Rivers	Lune	Rawthey	Ribble	Ribble	Swale	Swale
Area (km^2)	219	200	204	456	499.4	1363
Number (Figure 1)	1	2	3	3+4	5	5+6

Table 3-3. River catchment's mean and peak flows

Location		Water year 2007-2008		Water year 2008-2009	
	Mean Flow (m3/s)	Peak flow (m3/s)	Stage (m)	Peak flow (m3/s)	Stage (m)
Killington	10.20	250.76	2.41	320.48	2.73
Brigflatts	9.34	234.40	2.96	318.70	3.30
Arnford	7.40	118.87	2.03	118.74	2.02
Henthorn	13.69	286.71	3.02	240.18	2.74
Catterick	12.90	388.64	3.14	327.44	2.91
Crakehill	20.63	196.27	5.29	154.07	4.87

Most of the land of the catchments are used as grassland, with the exception of River Swale, where horticultural uses take place. In all the catchments studied, the urban areas occupy less than 5% of each catchment.

Table 3-4 details how land is used in each catchment studied. The permeability of each catchment is described in Table 3-5. Apart from the River Luna catchment that has a low permeability bedrock, the rest of the catchments have bedrocks moderate permeable. The catchments also have low superficial permeability.

Table 3-4 River catchment´s land-use

Location	Landuse (%)				
	Woodland	Arable / horticultural	Grassland	Mountain / Heath / Bog	Urban
Killington	4.43	1.13	83.48	8.21	0.90
Brigflatts	5.25	0.24	85.19	7.69	0.60
Arnford	4.21	0.87	79.97	11.30	1.90
Henthorn	6.41	2.90	80.43	6.09	3.28
Catterick	5.49	6.71	54.24	30.27	1.95
Crakehill	6.37	35.25	41.42	12.48	3.40

Table 3-5. River catchment´s permeability

Location	Bedrock permeability (%)			Superficial deposits permeability (%)		
	High	Moderate	Low	High	Moderate	Low
Killington	0.00	29.81	70.19	1.06	9.29	46.06
Brigflatts	0.00	74.04	25.96	1.02	11.52	37.46
Arnford	0.00	92.36	7.64	2.46	17.16	46.23
Henthorn	0.00	96.52	3.48	3.95	7.81	61.86
Catterick	0.04	99.96	0.00	1.41	23.18	42.69
Crakehill	32.32	48.55	11.63	15.43	11.04	57.51

3.4.2.1 River Ribble

The Ribble basin was selected by the European Commission and officials of the Member States of EU, along with 13 other river catchments in Europe, as a pilot to develop and test the EU Water Framework Directive (WFD) (Watson and Howe, 2006). The River Ribble estuary is located on the northwest coast of England. At Henthorn station, the river drains an area of 456 km2 through North Yorkshire and Lancashire. The Ribble´s northern half in Yorkshire Dales and Langstrothdale Chase and drains at Arnford station an area of 204 km^2. The River Ribble estuary is incised into Permo-Triassic bedrock (Van Der Wal, Pye and Neal, 2002; Watson and Howe, 2006; NRFA, 2014). The Ribble´s geology is predominantly from carboniferous limestone and some millstone grit with post-glacial deposits and Boulder Clay in the valleys. Besides the River Ribble, the basin is also formed by the Crossens drainage Network and the Calder, Darwin, Yarrow, Douglas and Hodder river systems. Due to its variable surface, the basin has a natural runoff pattern with ´flashy´ river flows. The catchment maximum altitude is 691.4 m above Ordnance Datum (AOD) and the station level altitude is 38.8 m AOD. The catchment flow has been recorded in these sites since 1968 and the maximum gauge level and peak flow registered at Heathorn was, respectively, 3.625 m and 403.9 m3 in 31/10/2010. Along the catchment there is Moorland in the upper catchment, and downstream from Arnford there is mixed farming and several small towns however, only minor effluent is discharged in the basin. The basin is home for a vast

number of protected species that has been suffering with damage to its habitats. The reason the Ribble basin to be chosen as a pilot basin to the WFD by the EA is due to the variety of land and water uses within the basin, having both natural and modified landscapes and the variety of stakeholders in the basin (Marsh and Lees, 2003; Watson and Howe, 2006; NRFA, 2014) .

3.4.2.2 River Lune

The River Lune, located in Cumbria and Lancashire, plays an essential role in water resources due to its habitat and conservation status, being particularly important for the Atlantic salmon. The area is often home to conflicts as a result of the agricultural interests in the region. The river drains an area of 219 km^2 of the eastern Lake District fells and parts of the Yorkshire Dales National Park and has excellent water quality. The river regime is flashy and is prone to inundation during wintertime. The topography of the region presented a challenge for urban development and intensive agricultural use; for this reason, land-use is limited mainly to Moorland, arable farming and grass. The catchment drains Silurian slates to the West, Carboniferous conglomerate and Limestone in the North and East, Peat moss on high moors to the Northwest, heather moss in the North. Lower valleys are covered with Boulder Clay. The maximum catchment altitude is 675.6 m AOD, and the station altitude is 82.8 m AOD. The catchment flow has been recorded in these sites since 1968, and the maximum gauge level and peak flow registered was, respectively, 3.369 m and 471.813 m^3 on 08/01/2005 (Orr and Carling, 2006; NRFA, 2014).

3.4.2.3 River Swale

The river Swale is located in Yorkshire Ouse, and it drains water from North Yorkshire Dales and flows to the southeast to converge with the River Ure, draining a 1363 km^2 area at Crakehill station. The upper-catchment area is 499.4 km^2 and drains at Catterick station. The catchment is formed by Carboniferous limestone and millstone grit, Triassic mudstone and sandstone. The river starts in a steep-sided hill with a gradient of 14.8 m km^{-1} and as it matures into a slow-flow river downstream from Catterick with a gradient of 0.8 m km^{-1}. The river sediments also change as the river develops. In the upland region, there are boulders and cobbles with sand sediments, and in Crakehill, there is an accumulation of sand and silt in addition to the Boulder Clay (Bowes, House and Hodgkinson, 2003; Bowes *et al.*, 2005; NRFA, 2014). The maximum catchment altitude is 714.3 m AOD, and the Crakehill station altitude is 12 m AOD. The catchment flow has been recorded in these sites since 1969, and the maximum gauge level and peak flow registered in Crakehill station was, respectively, 5.68 m and 238.70 m^3 on 27/09/2012. The catchment is responsive and has a fairly natural regime, and has primarily moderate permeability. There is mainly Moorland and grassland in the headwater, having a considerable area of arable cover (NRFA, 2014).

3.4.2.4 River Rawthey

The River Rawthey is located in Cumbria and Lancashire and drains an area of 200 km^2 of Carboniferous Limestone and Millstone Grit. The catchment has peat on the highest moors and Boulder Clay in the valleys and on lower slopes. The river has a natural flow regime. The maximum catchment altitude is 734 m AOD, and the station altitude is 84.1 m AOD. The River Rawthey, being a steep river, has very responsive flows. The catchment flow has been recorded in these sites since 1969, and the maximum gauge level and peak flow registered in Brigflatts station was, respectively, 3.843 m and 460.4 m^3 on 31/01/1995 (NRFA, 2014).

3.4.2.5 River catchments data

Hydrometric data was obtained from the National River Flow Archive (NRFA), hosted by the Centre for Ecology and Hydrology (CEH). The hydrometric data has a 15 min temporal resolution and is subject to automated and manual validation by the NRFA. A Service Level Agreement (SLA) was made in the UK to improve the quality and continuity of hydrometric data and comprises the transfer and validation of river flow data. Data is transmitted from the gauging stations to regional operational databases, where it receives a primary quality control before being sent to the NRFA. The data is not added to the national archive until its properly validated by NRFA own software applications and qualified regional representatives. Visual and manual validation, including the comparison between different near-neighbour sites and assessments of time-series statistics, by representatives that are familiar with the rivers, flow patterns are the most efficient way to quality-control river flow data (Marsh, 2002; Dixon, Hannaford and Fry, 2013). Further details of the data sensing, recording and validation in the UK river basins by NRFA are presented in Dixon, Hannaford and Fry, (2013).

3.4.2.6 PDM model

The Probability Distributed Model (PDM) is a flexible lumped conceptual rainfall-runoff model which uses as input rainfall and probable evaporation data into fast and estimates the soil moisture storage and routes the surface and subsurface flows (Moore, 1985; Moore and Bell, 2002; Li *et al.*, 2011). PDM has been extensively used in literature (Arnell, 1999, 2011; Pierce *et al.*, 2005; Cabus, 2008; Cole and Moore, 2008; Liguori, Rico-Ramirez and Cluckie, 2009; Kay *et al.*, 2009; Ferket, Samain and Pauwels, 2010; Li *et al.*, 2011; Bringi and Thurai, 2012; Srivastava *et al.*, 2014; Arnell, Charlton and Lowe, 2014; Cecinati *et al.*, 2017). The model is currently used by the Environment Agency for real-time flood forecasting across England and Wales. The basic structure of a PDM model is outlined in Figure 3-6.

The soil's capacity to take up water, including canopy and surface retention, controls the runoff production at any given point within the catchment. The storage´s (S_1) inlet is rainfall (P), and the outlets are evaporation losses (E) and recharge, and when it is full, it generates runoff. The only difference in

each point of the catchment is the capacity; then, its spatial variation can be described by a probability density function. Integrating the point runoff is then used to model the runoff over the catchment. The model gives as output the outflow from subsurface and surface storages with established flows (e.g. constant abstractions or releases). The model is based on mass balance principles and takes into account averaged rainfall and evaporation over the catchment to compute recharge and runoff production. The model represents the subsurface and surface storages by a transfer function discretely coincident with a cascade of two linear reservoirs or by solving the Horton-Izzard equation (Moore, 2007). Further information about the model and the equations used can be found in Moore (1985, 2007).

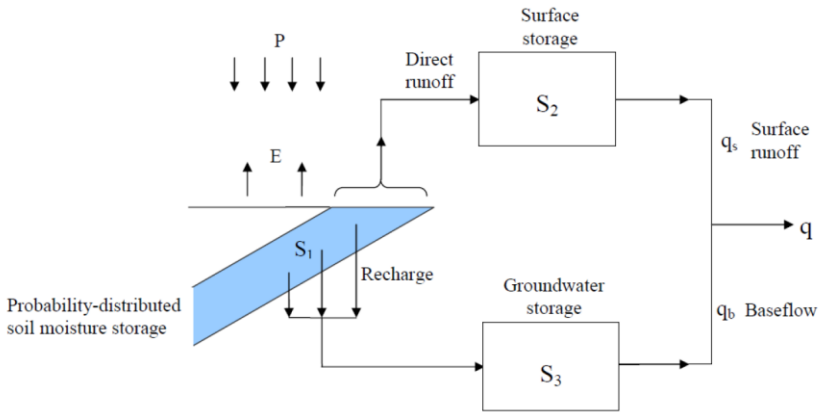


Figure 3-6 The PDM rainfall-runoff model (Moore, 2007)

The PDM uses 14 parameters that are listed in Table 3-6 and uses potential evapotranspiration (PET), catchment area and precipitation (measured or forecasted) as inputs. The meteorological data used to compute the potential evapotranspiration was provided by the UK Met Office through the Met Office Integrated Data Archive System (MIDAS). The Land surface and marine surface observation data from the UK Met Office comprises wind parameters, air temperatures, dew point temperature, air pressure and solar radiation measurements and rain measurements (Met Office (2012), 2014). The potential evapotranspiration was computed using the FAO Penman-Monteith algorithm (Allen *et al.*, 1998). The rain gauge data had to be averaged over each catchment to be used as input in the PDM model. The Thiessen method was used to compute the catchment averaged rainfall.

The simulations were made using hourly data. For this reason, both meteorological and hydrological data were accumulated (rainfall) or averaged (flows). The PDM model used was initially set up by Nanding (2016). In this research, data from the 1st of January until the 31st of December 2007 was used for the PDM calibration. Calibration was made optimising the parameters using a global optimization method.

Table 3-6. PDM model parameters (Moore, 2007)

Parameters	Description Unit	Unit
f_c	Rain factor	-
τ_d	Time delay	hr
c_{min}	Minimum storage capacity	mm
c_{max}	Maximum storage capacity	mm
b	Exponent of the soil moisture distribution	-
b_e	Exponent in the actual evaporation function	-
k_g	Groundwater recharge function	$hr.mm^{bg-1}$
b_g	Exponent of recharge function	-
s_t	Soil tension storage capacity	mm
k_b	Time constant of the groundwater routing	$hr.mm^2$
q_c	Constant flow representing returns/abstractions	m^3s^{-1}
k_1	Time constant of the surface routing	hr
k_2	Time constant of the surface routing	hr

The Shuffled Complex Evolution developed at the University of Arizona (SCE-UA) (Duan, Sorooshian and Gupta, 1992, 1994; Duan, Gupta and Sorooshian, 1993). Optimization was also tested using root mean square error (RMSE) and mean absolute error (MAE), but SCE-UA provided the best results overall with the NSE. The Nash-Sutcliffe efficiency (NSE) coefficient was used to assess the model's ability to estimate the catchment flow. The NSE ranges from -∞ up to 1, where 1 indicates that the model can perfectly reproduce the observed flows and negative values indicate that the model ability to reproduce the flow is worse than the observed flow mean (Nash and Sutcliffe, 1970). It is worth mentioning that the SCE-UA method optimises the model parameters in order to achieve the best NSE performance. However, it is well-known that different combination of parameter values might also produce similar model performance (known as equifinality), and therefore parameter uncertainty in hydrological modelling is an active area of research. In this work, however, the best-calibrated model obtained through SCE-UA was used in the analysis in order to test how the uncertainties in the rainfall and forecasts are propagated in the forecasting of river flows (or flows in the sewer system for the urban area). The calibration was undertaken in order to maximize the NSE score. The validation was undertaken using data from the 1st of January until the 31st of December 2008. The forecasted flows are compared to measured flows and estimated flow using radar data and gauge data. The gauge estimated flows are produced using the PDM calibrated using rain gauge data. For the radar estimated flow and forecasted flows, radar data was used to calibrate the PDM. This was done in order to account for any bias that might be present in the radar rainfall measurements that affect the hydrological simulations. In this way, the best-calibrated models using radar and rain gauge measurements were identified and used in the forecasting analysis.

Chapter 4. Radar rainfall ensembles to represent uncertainties in nowcasts

4.1 Introduction

The quantification of uncertainty in rainfall measurements is vital in flood forecasting applications. In Chapter 2, rain gauge and weather radar uncertainties were discussed, along with a review of some of the existing techniques to assess these uncertainties. In Chapter 3, a short explanation about the UK's data processing used to account for these uncertainties was also presented. It is important to note that even with efforts to minimize them, residual uncertainties (both in radar rainfall measurements and forecasts) are still present and have to be considered, in particular for hydrological applications (Germann *et al.*, 2009). Radars errors are particularly significant in the first hours of the forecasts (Foresti *et al.*, 2013) and they propagate and produce uncertainties in the flood forecasts. Therefore, it is important to model or quantify the uncertainties in the rainfall observations so they can be accounted for and propagated into rainfall and river flow forecasts. At present, it is unclear how the uncertainties in radar rainfall affect the rainfall forecasts as well as the flow forecasts in sewer systems and river catchments. This chapter investigates how we can use existing models to quantify the uncertainties in radar rainfall to provide more significant rainfall and flow forecasts compared to existing models that only add spatially-correlated stochastic noise to the rainfall advection field.

There are currently different models available in the literature to model the uncertainties in radar rainfall measurements. For instance, Cecinati et al. (2017) proposed a geostatistical method to generate random error fields with the correct error characteristics when comparing radar with rain gauge observations. Other methods proposed using multiplicative error models to account for biases in the rainfall observations (McMillan *et al.*, 2011). In contrast, other researchers argue that the errors between radar and rain gauge measurements can be modelled by a multiplicative function that accounts for biases and random components. Approaches that are based on the modelling of individual sources of errors can be challenging as radar estimations can have large variations due to meteorological conditions and are complexly correlated (Seed, Pierce and Norman, 2013; Rico-Ramirez, Liguori and Schellart, 2015). There are also models based on the difference between radar and a reference, usually a dense rain gauge network (Germann *et al.*, 2009; Seed, Pierce and Norman, 2013). These models are able to account for the fact that the radar errors change in space and time and that radar errors depend upon a number of different factors, as explained in Chapter 2. Once they compare radar data with a reference, they are able to provide a direct rainfall estimation that includes all uncertainties sources together (Rico-Ramirez, Liguori and Schellart, 2015). One of the models that take

into account the space-time error structure was proposed by German et al. (2019). Germann *et al.* (2009) use a series of stochastic ensembles to assess uncertainties in weather radar rainfall estimation.

In comparison with models that individually account for the different sources of errors, these models consider all the residual errors combined and decreases the complexity of the modelling process. Moreover, this model has been successfully used to model rainfall uncertainties for nowcast and also to simulate hydrological flows (Seed, Pierce and Norman, 2013; Rico-Ramirez, Liguori and Schellart, 2015), indicating its suitability to be used in real-time rainfall and flow forecast. The ensembles are generated by means of the difference between the radar measurements and a reference (e.g., rain gauge observations) obtained from a large historical data set. The model proposed by Germann *et al.* (2009) can represent the spatial and temporal correlations of the error mean and covariances. The ensembles are formed by adding a stochastic perturbation to the deterministic radar precipitation field (original unperturbed field), as shown in equation (4-1) (Germann *et al.*, 2009).

$$\underbrace{\Phi_{t,i}}_{probabilistic} = \underbrace{R_t}_{deterministic} + \underbrace{\delta_{t,i}}_{stochastic} \quad (4\text{-}1)$$

Where $\Phi_{t,i}$ is the resulting precipitation field for ensemble member i at time t, $\delta_{t,i}$is the perturbation field for ensemble member i, and R_t is the original radar precipitation field. $\delta_{t,i}$ is modelled using the 'LU decomposition algorithm', and the result is a Gaussian correlated field with covariances and means that are pre-defined. R_t is composed by the original radar estimates at time t and location x^T (Germann *et al.*, 2009).

$$R_t^T = (R_{t,x_1}, R_{t,x_2}, R_{t,x_3}, \dots, R_{t,x_M}) \quad (4\text{-}2)$$

Where M is the number of radar pixels in the study area Ω. In the same way, $\Phi_{t,i}$ and $\delta_{t,i}$ are also formed by $\Phi_{t,x_k,i}$ and $\delta_{t,x_k,i}$ at the M locations. Hereafter, R and δ will be described as 'deterministic' and 'stochastic' components, respectively. The resulting $\Phi_{t,i}$ is then a result of the estimated radar precipitation field and the N different realizations of $\delta_{t,i}$ time-series. It is crucial to have a large enough N for the resulting precipitation field to be able to capture the range of uncertainties in the radar estimates (Germann *et al.*, 2009).

Due to the multiplicative nature of radar uncertainties, the residual error ϵ is defined in dB as (Germann *et al.*, 2009):

$$\epsilon = 10\log\left(\frac{S_t}{R_t}\right) \quad (4\text{-}3)$$

Where S_t is the real or true precipitation field that is unknown. Perturbations are hence, added logarithmically, and Equation (4-1) can be written as (Germann *et al.*, 2009):

$$10log\left[\Phi_{t,i}\right] = 10\log[R_t] + \delta_{t,i} \quad (4\text{-}4)$$

Where the perturbation $\delta_{t,i}$ is in dB. Using 'the LU decomposition algorithm' is used to estimate $\delta_{t,i}$. The algorithm core multiplies a random Gaussian vector with zero mean by the square root of the covariance matrix (C); the result is then added to the mean error μ. C and μ dimensions are now, respectively, $MQxMQ$ and MQ, where M is the number of radar pixels and Q is the number of time steps. Values with no rainfall are excluded from the calculations. Temporal correlation of the errors is imposed using an autoregressive model AR(2), allowing $\delta_{t,i}$ to be calculated step-by-step. In order to use AR(2), it is necessary to calculate the lag-1 and lag-2 errors correlations in time (r_1 and r_2) (Germann *et al.*, 2009).

Initially, an assumption that the ϵ is independent in space and time is added using AR(2) and that μ, C, r_1 and r_2 do not vary in time. μ and C are then defined as (Germann *et al.*, 2009):

$$\begin{aligned} \mu &\equiv E\{\epsilon_{x_k}\} \\ C_{kk} &\equiv Var\{\epsilon_{x_k}\} \\ C_{kl} &\equiv Cov\{\epsilon_{x_k}, \epsilon_{x_1}\} \end{aligned} \quad (4\text{-}5)$$

Where, ϵ_{x_k} is the residual error at location x_k, E is the expectation, Var is the variance, and Cov is the covariance. Note that the covariance matrix will account for the variation of the errors in space, whereas the autoregressive model will account for the variation of the errors in time.

The ensemble generator consists of two steps (Germann *et al.*, 2009):

1. Errors space-time structure estimation - Estimate C and μ for the locations M. At this point, the dimensions of C and μ are, MxM and M. Decompose C and estimate r_1 and r_2.
2. Generate perturbation fields and produce ensembles – Generate $\delta_{t,i}$using Gaussian random white noise, the LU composition algorithm and the autoregressive filter AR(2). Each ensemble member is formed by adding $\delta_{t,i}$ to R_t.

Using the LU decomposition algorithm, the covariance matrix is decomposed into an upper- and a lower- triangular matrix (4-7). Equation (4-7) describe the LU decomposition algorithm (Germann *et al.*, 2009).

$$C = L.L^T \quad (4\text{-}6)$$

$$\delta_{t,i} = \mu.Ly_{t,i} \quad (4\text{-}7)$$

Where $y_{t,i} = N_M(0, I)$ is a random Gaussian vector with zero mean and unit-variance. Equation(4-7) results in a multivariate vector, $\delta_{t,i}$ distributed in N_M (Germann *et al.*, 2009).

An interactive procedure calculates δ_t by linking $Ly_{t,i}$ with $\delta'_{t-1,i}$ and $\delta'_{t-2,i}$ (Equation (4-8). At the end of the interaction, the mean error is added (Germann *et al.*, 2009).

$$\delta_{t,i}^{'} = Ly_{t,i} - a_1\delta_{t-1,i}^{'} - a_2\delta_{t-2,i}^{'}, \quad (4\text{-}8)$$

$$\delta_t = \mu - \upsilon\delta_{t,i}^{'} \quad (4\text{-}9)$$

A detailed description of the method can be found in Germann *et al.* (2009).

4.2 Methodology

The radar rainfall errors were modelled using the model proposed by Germann *et al.* (2009) described above. The radar rainfall estimation ensembles (RE ensembles) attempt to assess the residual uncertainties that remain even after the correction algorithms are applied to the radar data in a realistic way. These radar rainfall ensembles are used instead of the original radar images to generate forecasts using the nowcasting model from STEPS. A deterministic forecast is produced using each of the 25 radar estimation ensemble members. As a result, each of these 25 forecasts become the ensemble forecasts. Ensemble forecasts based on this method will be referred to as radar rainfall estimation ensembles (RE ensembles) in this chapter, while ensemble forecasts generated using the stochastic ensemble generator by STEPS will be referred to as STEPS ensembles. Note that the STEPS ensembles are generated by adding spatially correlated noise to the deterministic radar forecasts without taking into account how the radar errors change in space.

The rainfall forecast output was used as an input in the hydrodynamic sewer network model from Ilkley, a town in northern England. A description of the sewer system can be found in Section 3.4.1. The rainfall forecast output was also used to forecast flow in river catchments using a PDM model. In Section 3.4.2, there is a more detailed description of the river catchments and the PDM model setup.

An overview of the events with the start date and time, duration, peak flow, maximum average rainfall, and storm type is shown in Table 1. Events with high peak flow were selected, and flow forecasts were performed every 30 min, including forecasts of low, medium and high flows. In order to classify the storms as convective or stratiform, each pixel of the radar scan was classified as either stratiform or convective using the algorithm proposed by Steiner et al. (Steiner, Jr, and Yuter, 1995). A storm was classified as convective if more than 3% of the pixels in the study area are convective for a period longer than 3 h. The events that did not fulfil these requirements are classified as stratiform (Rico-Ramirez,

Liguori and Schellart, 2015). Although the percentage of convective pixels seems to vary from storm to storm and depends upon the size of the rainfall domain, it seems that most of the stratiform or widespread rainfall systems have a small number of convective pixels (i.e. pixels associated with a high reflectivity or rainfall intensity for several hours) in comparison to convective events. Therefore, the thresholds of 3% of convective pixels over 3 hr to classify convective storms were selected by trial and error by looking at different storms that clearly showed convective or stratiform behaviour.

Table 4-1. Event start dates, duration, peak measured flow, maximum average rainfall and storm types (S—stratiform and C—convective) for the Ilkley urban catchment.

Event Date	Starting Time	Duration (h)	Peak Flow (m^3/s)	Max Average Rainfall (mm/hr)	Storm Type
29 April 2008	23:30	6.5	0.3974	5.344	S
30 April 2008	15:30	3.3	0.3477	8.511	S
12 May 2008	04:00	6	0.5582	4.117	S
26 June 2008	13:00	9	0.2869	11.426	S
7 July 2008	15:00	4	0.5024	7.144	S
29 July 2008	05:00	4	0.5826	6.444	S
1 August 2008	01:30	4.5	0.9904	8.073	C
20 August 2008	20:00	5	0.5800	10.564	S
4 October 2008	15:00	13	0.3378	8.489	S
14 October 2008	15:00	5	0.2754	7.358	C
8 November 2008	20:00	5	0.2592	5.412	S
9 November 2008	13:30	7	0.5630	18.410	S
13 December 2008	00:00	9	0.4551	1.813	S

4.2.1 Analysis using radar rainfall estimation ensembles to generate short-term forecast

Radar rainfall estimation ensembles were generated using both the method described by Germann et al. (Germann *et al.*, 2009) and the STEPS ensembles (Bowler, Pierce and Seed, 2006). Figure 4-1 illustrates an example of radar scans and two different RE ensemble members at different time steps. In these figures, it is possible to see how the rainfall is developing according to the radar measurements. In Figure 4-1, the forecast started at 02:00, so the results show forecasts with a lead-time of 30 min, 60 min, and 90 min. There are apparent differences between the radar rainfall and the rainfall forecasts, and the results show how these differences increase with lead-time. Using different ensembles to

estimate and forecast rainfall can give more information about how the storm will develop according to the initial uncertainty of the radar rainfall measurement.

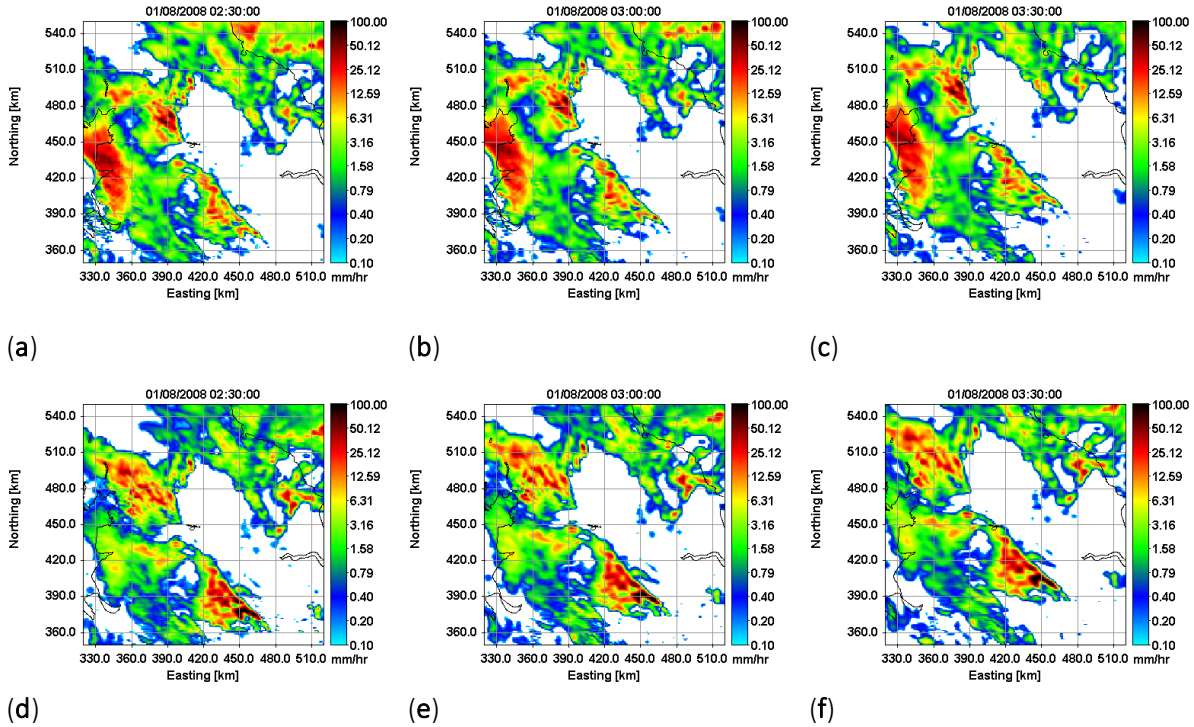


*Figure 4-1. RE ensemble forecasts at time t = 30 min (**a**,**d**), t = 60 min (**b**,**e**), t = 90 min (**c**,**f**) starting on 1 August 2008 at 02:00. Each probabilistic forecast is formed by 25 ensemble members that are valid at the simultaneously.*

4.3 Rainfall Forecasting

The Receiving Operating Characteristic (ROC) curves were calculated to assess the ensembles rainfall forecasts' predictability. ROC curves have been widely used to analyze uncertainties in probabilistic forecasting systems and measure the ability of a model to correctly identify the occurrence of an event. The method is based on a binary system, where yes/no forecasts and yes/no observations are computed (Liguori *et al.*, 2012). For a sequence of threshold, a 'hit-rate' (HR) (proportion of events correctly forecasted) and a 'false-alarm rate' (FAR) (proportion of events that were not forecasted) are computed and used to define the ROC curve (Mason and Graham, 2002). The better the forecast, the higher the HR (and lower the FAR). The area beneath the ROC curve should be above 0.5 (random forecast) and is equal to 1 when the model can correctly forecast an event that occurs or not (Liguori *et al.*, 2012).

Table 4-2. Contingency table for ROC curve

	Event observed	
Event forecast	Yes	No
Yes	True positive (TP)	False positive (FP)
No	False negative (FN)	True Negative (TP)

For a deterministic forecast, HR and FAR can be calculated according to equations (4-10) and (4-11) (Kharin and Zwiers, 2003):

$$HR = \frac{TP}{TP + FP} \tag{4-10}$$

$$FAR = \frac{FN}{FN + TP} \tag{4-11}$$

For probabilistic forecasts, the HR and FAR can be expressed in function of the probability of the forecast exceeding a critical threshold (P_{cr}). In this case, HR and FAR can be calculated according to equations (4-12) and (4-13) (Kharin and Zwiers, 2003):

$$HR = \int_{\Omega_P} f(P|E = 1)dP \tag{4-12}$$

$$FAR = \int_{\Omega_P} f(P|E = 0)dP \tag{4-13}$$

Where, P is the probability of the forecast to exceed a threshold, Ω_P is the forecast probabilities $P > P_{cr}$ and $f(P|E)$ is the conditional probability density function. E is a dichotomous predictant defined as $E = 1$ in the occurrence of an event and $E = 0$ when the event does not occur.

Some of the results obtained for a selection of events with rainfall intensities equal to or higher than 0.1 mm/hr, 1.0 mm/hr and 3.0 mm/hr are presented in Figure 4-2 until Figure 4-5 at different forecasting lead-times.

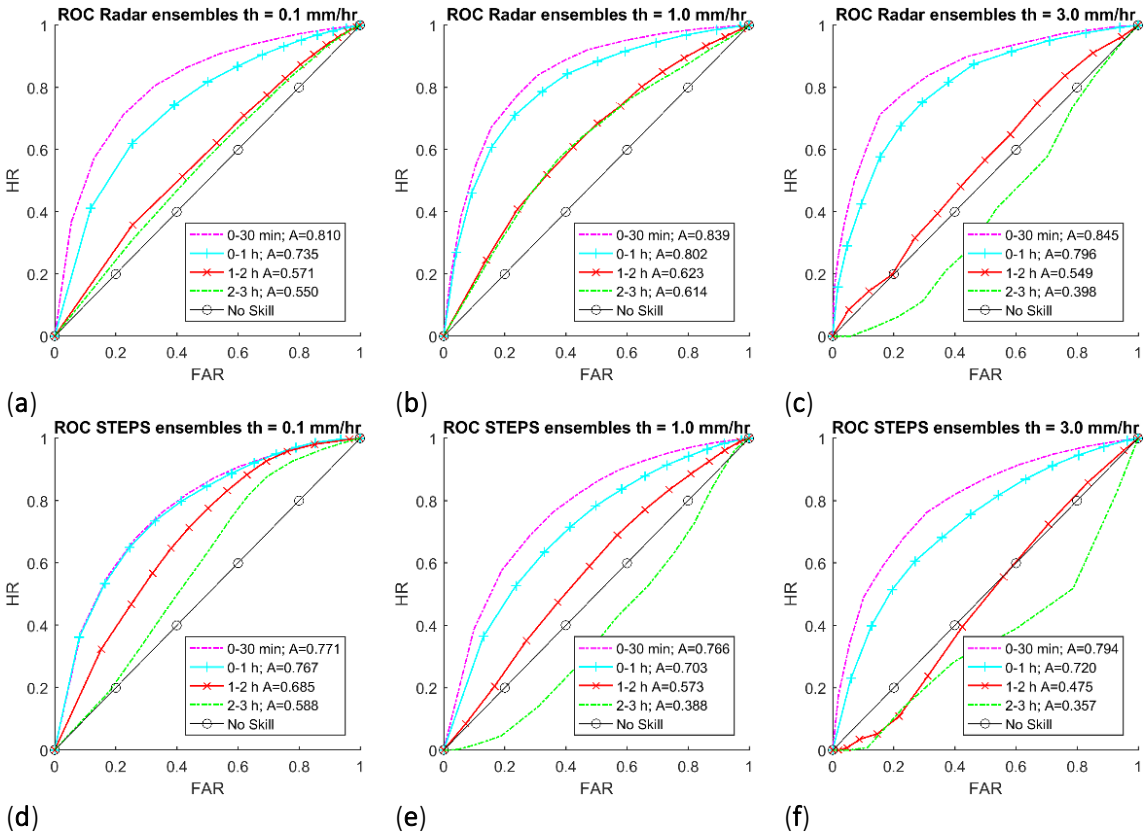


Figure 4-2.ROC curves for event on 7 July 2008 starting at 15:00 for RE and STEPS ensembles, respectively and thresholds (th) equal to 0.1 mm/hr(a, d), 1.0 mm/hr(b, e) and 3.0 mm/hr(c, f).

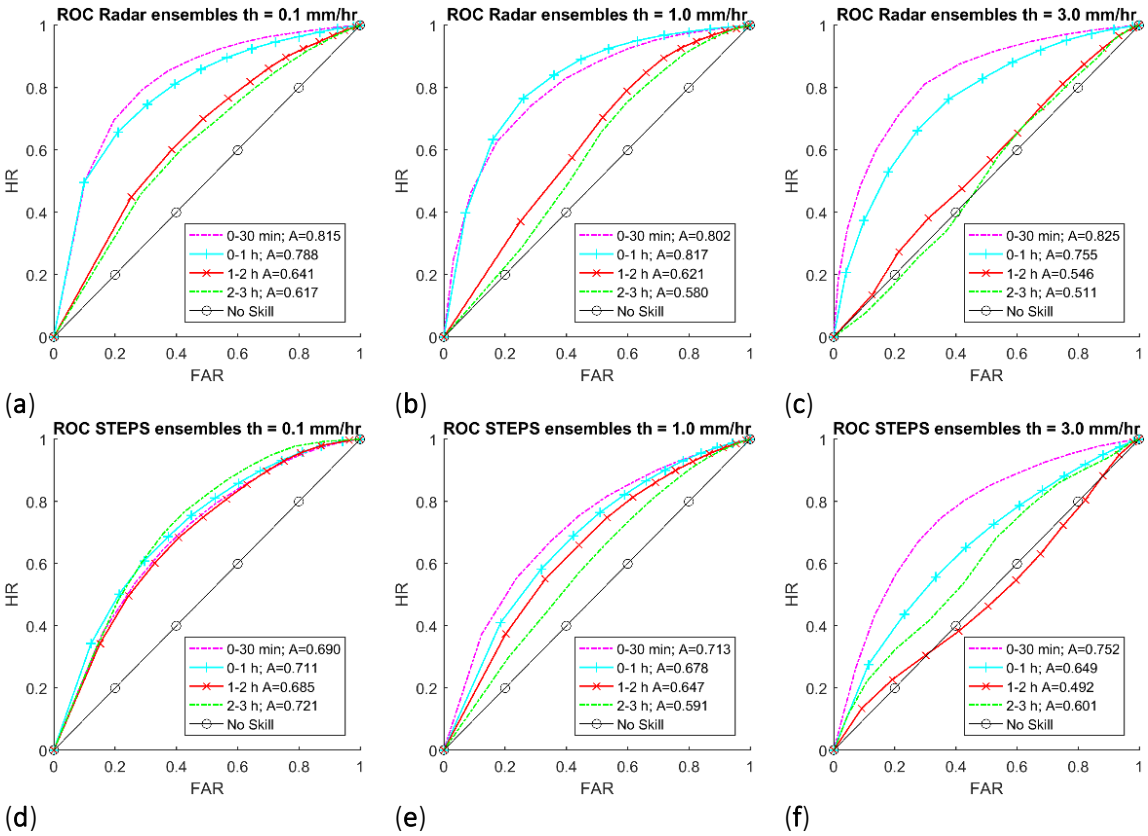


Figure 4-3. ROC curves for RE and STEPS ensembles for event on 1 August 2008 starting at 01:30 for RE and STEPS ensembles, respectively and thresholds (th) equal to 0.1 mm/hr(a, d), 1.0 mm/hr(b, e) and 3.0 mm/hr(c, f).

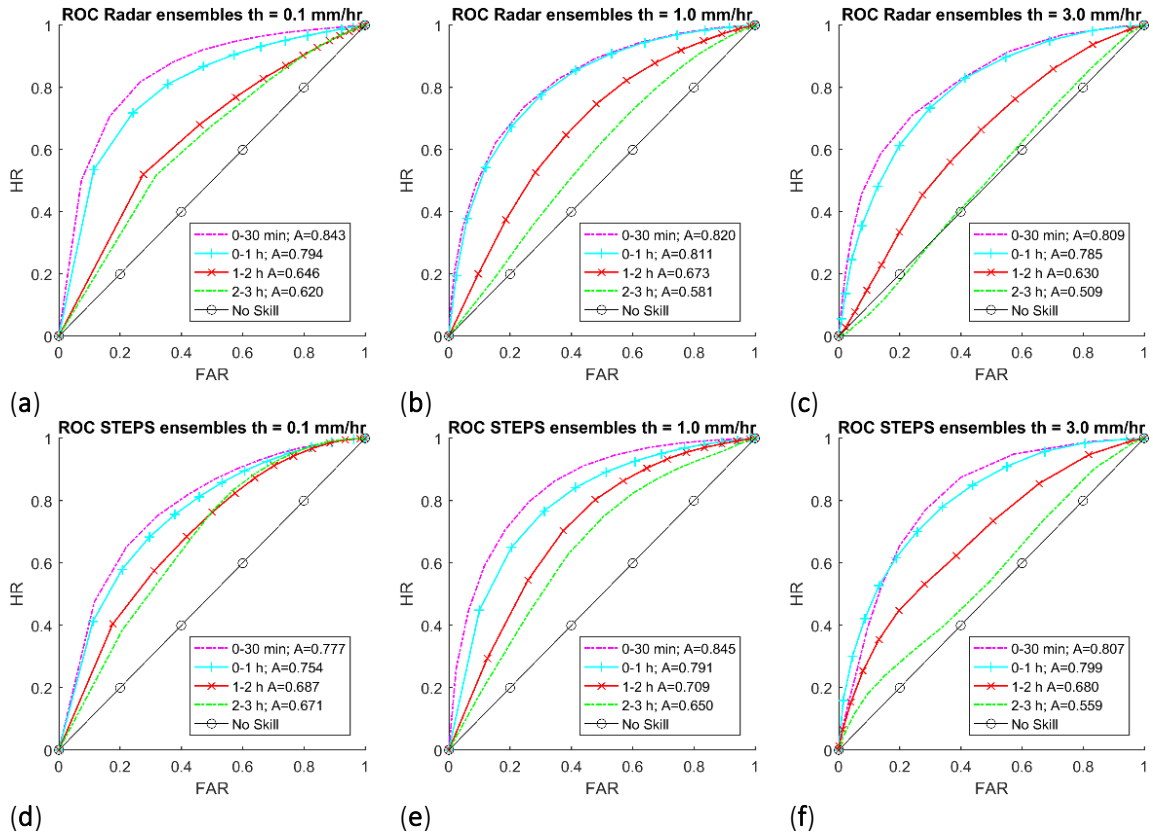


Figure 4-4. ROC curves for RE and STEPS ensembles for event on 4 October 2008 starting at 15:00 for RE and STEPS ensembles, respectively and thresholds (th) equal to 0.1 mm/hr(a, d), 1.0 mm/hr(b, e) and 3.0 mm/hr(c, f).

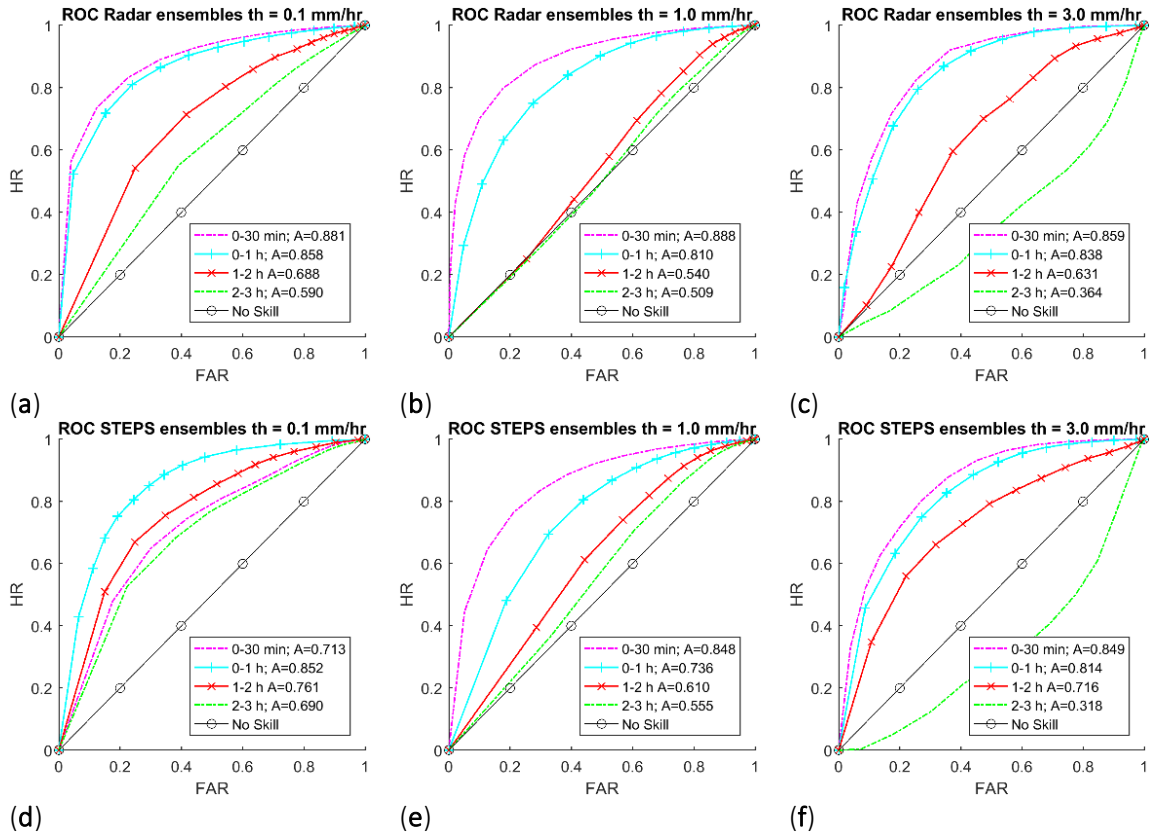


Figure 4-5. ROC curves for RE and STEPS ensembles for event on 14 October 2008 starting at 15:00 for RE and STEPS ensembles, respectively and thresholds (th) equal to 0.1 mm/hr(a, d), 1.0 mm/hr(b, e) and 3.0 mm/hr(c, f).

To summarise the results from the events analyzed, the area beneath the ROC curves for different rainfall thresholds and forecasting lead-times are plotted in Figure 4-6 – Figure 4-8. The results show that the forecast skill decreases with both forecasting lead-time and higher rainfall intensities.

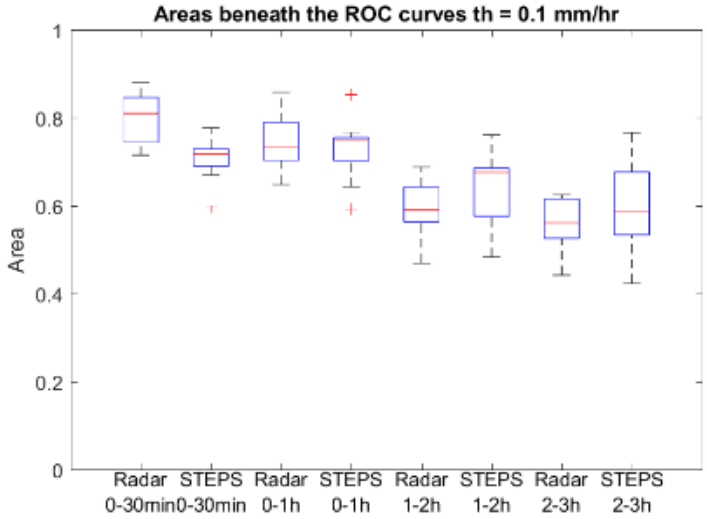


Figure 4-6. Area beneath ROC curves for all the events using RE ensembles and STEPS ensembles and threshold 0.1 mm/hr.

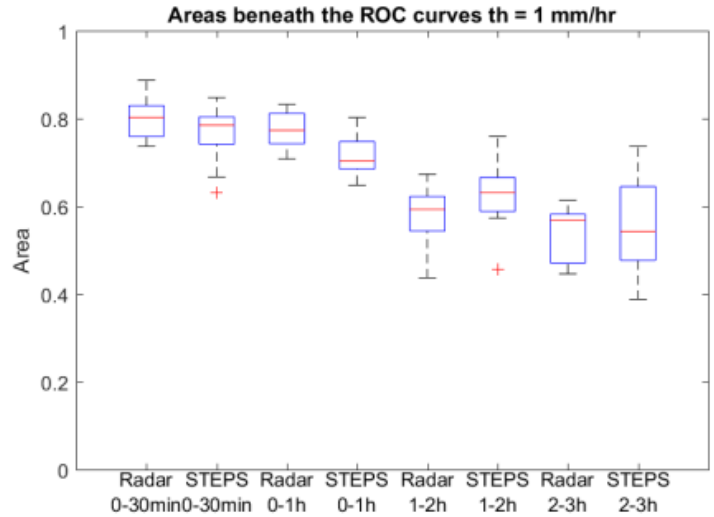


Figure 4-7. Area beneath ROC curves for all the events using RE ensembles and STEPS ensembles and threshold 1 mm/hr.

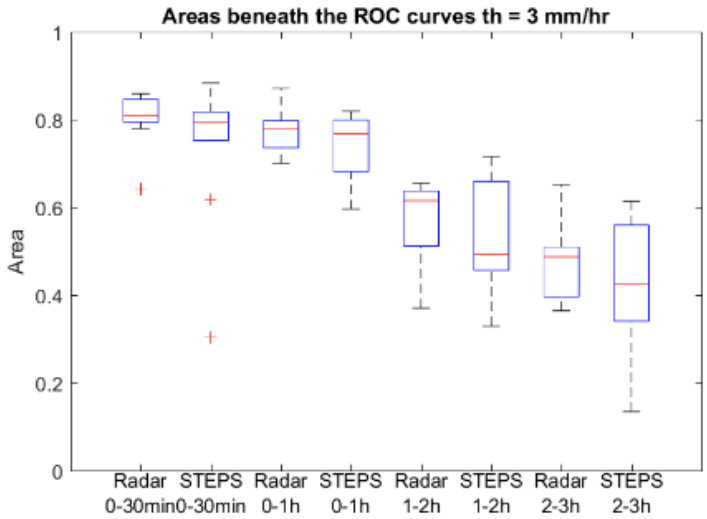


Figure 4-8. Area beneath ROC curves for all the events using RE ensembles and STEPS ensembles and threshold 3 mm/hr.

The tendency observed is that the forecast accuracy decreases with longer lead-time and higher rainfall intensities, as shown in Figure 4-6. The first forecasted hour shows a high forecast ability; however, with the increase of lead-time, the ability is rapidly reduced. The forecast skill also decreases with higher rainfall thresholds, indicating that high rainfall intensities are more challenging to forecast and are subject to more errors. In most cases, the ROC curve is below the random forecast line for a threshold of 3 mm/hr, demonstrating that in these situations, the model fails to forecast higher rainfall intensities efficiently. Figure 4-2 to Figure 4-5 demonstrate how the forecast efficiency decline after the first hour forecasted for a threshold of 3 mm/hr. For a rainfall threshold of 0.1 mm/hr, all the events for both probabilistic forecasts produce skilful predictions up to 3 hr lead-time. With an increased lead-time, however, the forecast ability also decreases.

The events analyzed produced good forecasts up to 1 hr lead-time for all rainfall thresholds, but overall the RE ensembles performed slightly better than the STEPS ensembles during this period. The STEPS ensembles lose forecasting skill less rapidly and more constantly between the time-steps used when the thresholds are between 0.1 mm/hr and 1.0 mm/hr. Significant forecast accuracy is lost after the first hour for the RE ensembles. The area beneath the ROC curve is reduced by approximately 20% for rainfall intensities higher than 0.1 mm/hr. For a threshold of 3.0 mm/hr. The area beneath the ROC curve is reduced by around 24% after the first forecasted hour. For lead-times longer than 1 h, the STEPS ensembles perform better than the RE ensembles in most cases. Seeing as radar errors are the predominant source of uncertainties in the first hour forecasted, an ensemble generator based on the modelling of the radar residual errors is expected to produce more accurate results at the beginning of the forecast (Germann *et al.*, 2009; Foresti and Seed, 2015).

The decrease in forecasting accuracy can be up to 27% per hour after 1 hr lead-time for higher rainfall intensities. In most cases with a 3.0 mm/hr threshold and lead-time longer than 2 hr, the forecast is unable to predict the rainfall intensities accurately at these small spatial scales and presents areas beneath the ROC curve equal or lesser than 0.5. The loss of efficiency in the forecast is consistent with other studies (Liguori *et al.*, 2012; Foresti *et al.*, 2016). Due to regions with high-intensity rainfall are smaller, there is a decline in the performance of the ensemble rainfall forecasts for higher thresholds. The area beneath the ROC curves for the RE ensembles are, on average, 10% higher than the STEPS up to 30 min lead-time. The difference falls to 6% when the forecasting lead-time is up to 1 hour.

4.4 Hydrological application in flow predictions

4.4.1 Urban catchment

The hydrodynamic model was verified using radar rainfall and the additional rain gauge data within the urban area as input. The root mean square error (RMSE) was calculated by comparing simulated flow with the measured flow. The RMSE was only computed for measured flow higher than 0.1 m^3/s to exclude flows measured in dry periods and to minimize uncertainties related to the flow measurement. The RMSE for the radar flow simulations is 0.0956 m^3/s and for the gauge flow simulation is 0.0838 m^3/s using a period from 15 April 2008 to 31 December 2008. Figure 4-9 presents the results comparing the measured flow with flow simulated using radar data, and Figure 4-10 shows the results comparing measured flows and flows simulated using gauge data. The results shown are from 04/10/2008 and 05/10/2008. In this specific case, the radar data was able to capture better the high peak flows; however, it also predicted peaks that did not occur. The rain gauge data was not able to estimate the first high-intensity peak. In the second peak, the rain gauge data could not estimate the flow pattern and estimated high intensities for a shorter period. The performance of radar and rain gauge was strongly dependant on the event, and in many cases, the rain gauge data produced the best results.

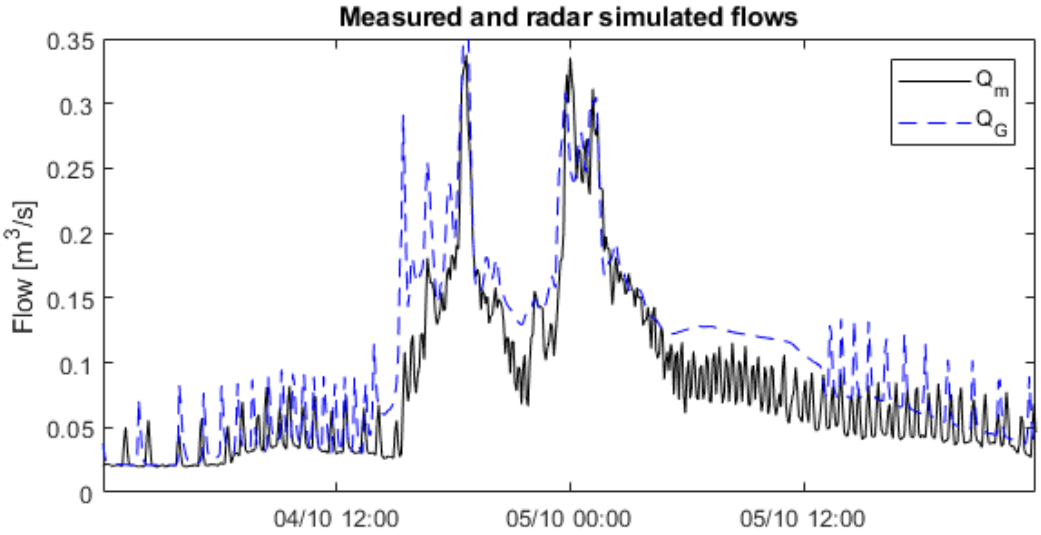


Figure 4-9. Measured flows (Qm) and radar simulated flows (Q_R) for 04 and 05 of October 2008.

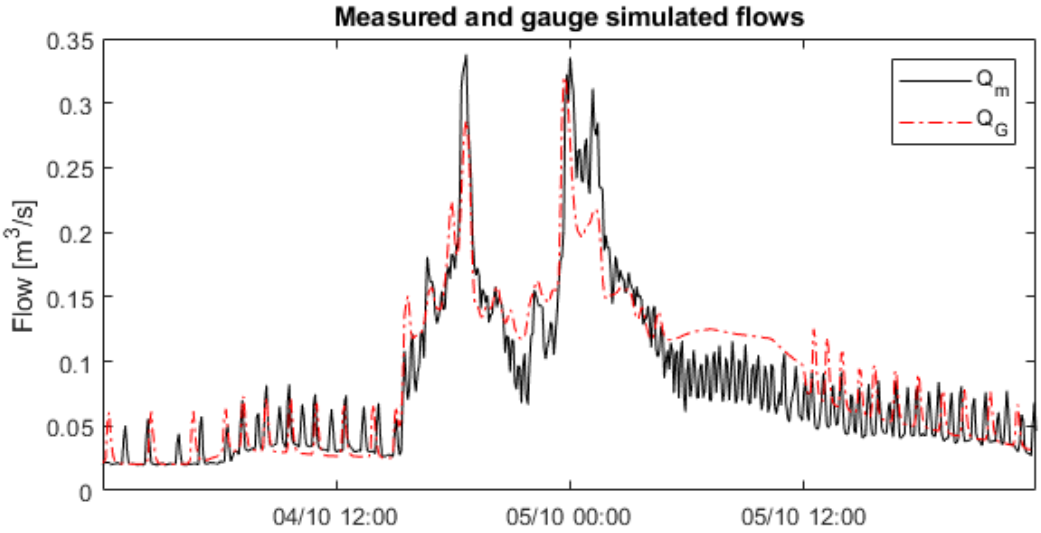


Figure 4-10. Measured flows (Qm) and gauge simulated flows (Q_G) for 04 and 05 of October 2008.

The sewer flow simulation results obtained using radar rainfall, rain gauge measurements and rainfall forecasts were compared against the sewer-measured flows. Since the forecasted peak flow sometimes appeared a few minutes later (or earlier) compared to the flow observations, it was decided to compare the peak flows within a particular time window (e.g., every 30 min or 1 h). This was decided as for real-time applications; it is essential to accurately forecast the flow peaks intensity. For real-life applications, having the flow forecasted within a few minutes from the measured flow still allows protective measures to be taken. The results are presented in Figure 4-11 until Figure 4-13.

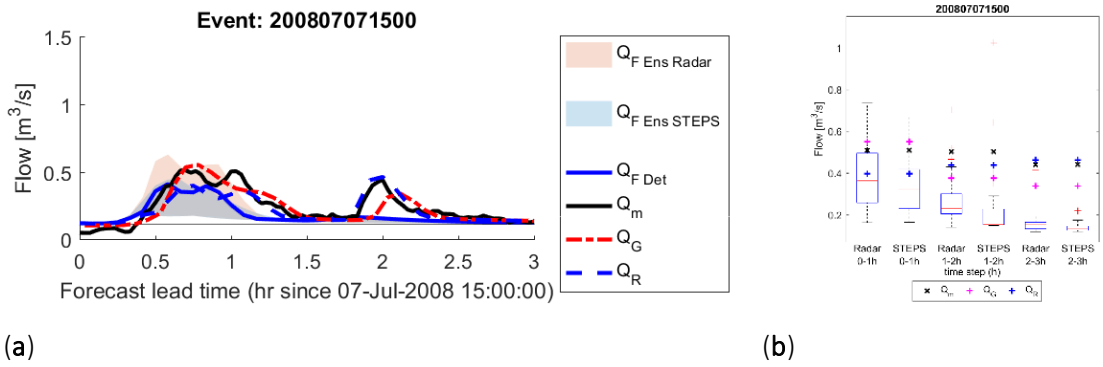


Figure 4-11. Flows (a) and peak flows (b) for events on 07 July 2008 starting at 15:00. $Q_{F\ Ens\ Radar}$, $Q_{F\ Ens\ STEPS}$, $Q_{F\ Det}$, Q_m, Q_G and Q_R are, respectively, forecasted RE ensemble flows, forecasted STEPS ensemble flows, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow.

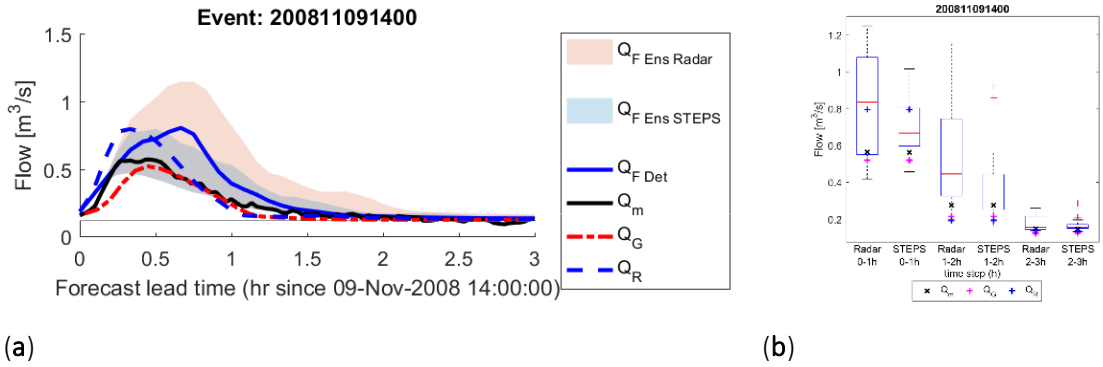


Figure 4-12. Flows (a) and peak flows (b) for events on 09 November 2008 starting at 14:00. $Q_{F\ Ens\ Radar}$, $Q_{F\ Ens\ STEPS}$, $Q_{F\ Det}$, Q_m, Q_G and Q_R are, respectively, forecasted RE ensemble flows, forecasted STEPS ensemble flows, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow.

RMSE calculations were carried out to assess the performance of the ensembles. Due to the time lag present between the measured peak flow and the ensemble-simulated peak flows, a cross-correlation correction was performed between the measured flows and ensemble flows. As the lags are not consistent, the cross-correlation was adjusted for each case. RMSE for each ensemble member was calculated for measured flows higher than 0.1 m^3/s (Figure 4-14). Table 2 presents the mean of the RMSE for both STEPS and RE ensembles. Given the fact that there is a significant loss of forecast efficiency after 1 hr of lead-time, the RMSE is only shown for this period. In most events, the time lags between measured flow and forecasts are less than 15 min.

The forecasted flow in Ilkley only produced reliable forecasts with lead-times up to 30 min. In some cases, the predictability maintained up to 1 hr ahead depending upon the nature of the rainfall event. For low flow situations, the model can accurately estimate flows using radar and gauge data. However, for flows over 0.5 m^3/s, the gauge estimations underestimate the flow peaks in most cases. Underestimation can be related to the model calibration. Using more events with high peak flows might be able to improve flow prediction in similar conditions. Results using radar data produce better

estimates for higher flows and can predict peak intensities more accurately for flows around 0.5 m^3/s. For higher intensity flows, the model also underestimates the flow in most cases. For low-intensity flows, both the gauge and radar estimations produce accurate results and mimic the flow pattern. Flow simulation with rain gauges has a lower RMSE (0.084 m^3/s) than using radar rainfall (0.096 m^3/s). These results are expected, given that the rain gauges used to simulate the flow are located within the urban area.

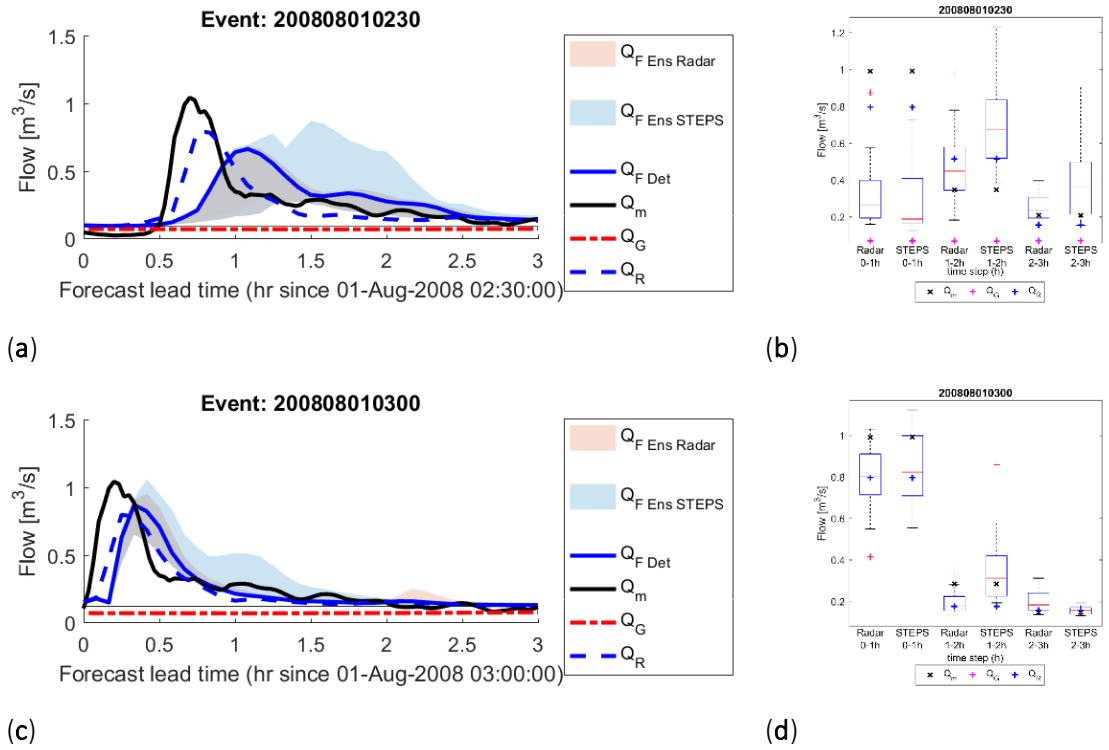


Figure 4-13. Flows (a) and peak flows (b) peak flows for events on 1 August 2008 starting at 02:30 and flows (c) and peak flows (d) for events on 1 August 2008 at 03:00. $Q_{F\ Ens\ Radar}$, $Q_{F\ Ens\ STEPS}$, $Q_{F\ Det}$, Q_m, Q_G and Q_R are, respectively, forecasted STEPS ensemble flows, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow.

Table 4-3. RMSE mean for RE and STEPS ensembles for 0–1 hr after the forecast's start.

RMSE (m^3/s)			
Event Date	Starting Time	RE Ensembles	STEPS Ensembles
7 July 2008	15:00	0.170	0.200
1 August 2008	02:30	0.405	0.430
1 August 2008	03:00	0.140	0.166
9 November 2008	14:00	0.287	0.169

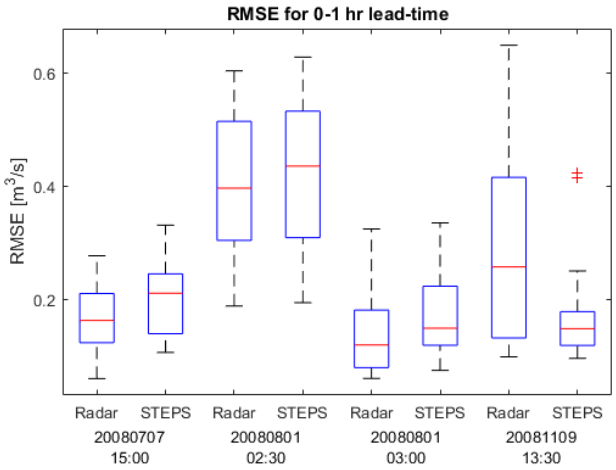


Figure 4-14. RMSE for RE and STEPS ensembles and 0–1 hr after the start of the forecast.

In the event shown in Figure 4-11 the radar probabilistic forecast is able to simulate the peaks in the first hour of the forecast and accurately simulate the second flow peak. The STEPS probabilistic forecasts underestimated both flow peaks. For this event, it is clear that using rain gauge data to produce the radar rainfall ensembles adds valuable information to improve the forecasts, as the RE ensembles predicted peaks are closer to the measured flow, while the STEPS ensembles underestimate the flow peaks. This allows the flow peak and flows patterns to be forecasted better using the RE ensembles. For RE ensembles, the measured flow (0.509 m^3/s) is very close to the 75th percentile flow (0.418 m^3/s). For the STEPS ensembles, the measured flow is only captured by the more extreme ensembles, as can be seen on the boxplot. The flow simulations using radar data and the deterministic forecast show similar results to the STEPS ensembles peak flows, both underestimating the flow peaks. The simulation using rain gauge data replicate the first flow peak more accurately but underestimate the second one. The RE ensembles have the advantage of combining information from both the radar and the rain gauges and are able to predict both flow peaks. The second large flow peak occurred around 2 hr lead-time and cannot be forecasted by any of the forecasts. This indicates that at lead-time longer than 1 hr, the flow forecast loses its forecasting skill. For this event, the RE ensembles perform better in analysing the RMSE mean. For the first hour, the mean is approximately 15% smaller for RE ensembles when compared to the STEPS ensembles.

The simulated flow peak in Figure 4-12 is overestimated using radar rainfall. Flow simulations using rain gauge data are much closer to the measured peak flow, but there is a delay in time of a few minutes. In this example, both RE and STEPS forecasts can capture the flow peak. However, the RE ensembles produce a flow forecast with a higher spread than the STEPS forecast. The measured peak flow (0.563 m^3/s) falls between the 25th and 75th percentile for the RE ensembles (Figure 4-12b) in the first forecasted hour. In the second hour of the forecast, the STEPS ensembles forecast the flow intensity more accurately than the RE ensembles. The RE ensembles produce a larger spread, and therefore

there is an overestimation of the flow by a large part of the ensembles. This leads to an increased RMSE (0.287 m^3/s) for the first hour of the forecasts and a higher value than using the STEPS ensembles (0.0169 m^3/s). The STEPS ensembles produce a smaller RMSE mean for the whole duration of the event.

In Figure 4-13, the nowcasting model's efficiency is rapidly lost with increased lead-time. This figure shows the results of forecasting the same rainfall event but at different starting times (30 min apart). The forecasts for both ensembles initiated at 02:30 on 1 August 2008 (Figure 4-13a) fail to predict the flow peak correctly, and the time lag between the forecasted flow and the measured flow is higher than for shorter lead-times. In the forecasts initiated at 03:00 (Figure 4-13b), the peak flow falls into the first forecasted hour, and the forecast produced replicates the peak flow better. In this case, both ensembles were able to capture the peak flow. However, only some of the RE ensembles can reproduce the peak flow correctly. Because the STEPS model produces a higher spread, more ensemble members can forecast the peak flow under these circumstances. The event presented in Figure 4-13 presents the highest measured flow among all the selected events. Accurately forecasting high rainfall intensities has proved to be more challenging, and both forecasts were able to predict the peaks at a short forecasting lead-time. The RMSE mean for the forecast initiated at 02:30 is higher (0.405 m^3/s for RE ensembles and 0.430 m^3/s for STEPS ensembles). In contrast to the other events, the flows were forecasted with a delay of around 20 min. Starting the forecast 30 min later improved the ability to predict flow peaks, and the RMSE for RE ensembles was nearly a third of the previous forecast. In the event represented in Figure 4-13, there were no rain gauge data available. So, the advantage of using radar rainfall ensembles is more evident. The RMSE indicates that the RE ensembles overperform the STEPS ensembles during the first hour forecasted.

Analysing the 13 events for longer forecasting lead-times, the performance is case dependent, but both RE ensembles and STEPS ensembles tend to lose their accuracy as forecasting lead-time increases.

In the majority of the events, there is a time-lag between the measured and forecasted flows. However, this time-lag is not consistent. In some events, the flow peaks are predicted in advance of the actual flow and in other cases, the forecasted flow has a delay when compared to the measured flow. The simulated flows for both radar and gauge do not present the same time lag, confirming that these are uncertainties inherited in the rainfall forecast. The fact that the area studied is an urban area of small dimensions with both permeable and impermeable surfaces means that the catchment response time is minimal, and any uncertainties related to the rainfall forecast have an almost immediate effect on the flow forecast. Tests performed with the ensemble forecasts, where high-intensity rainfall pixels were only displaced a few kilometres, could have a high impact on the flow simulation. This highlights the importance of improving the accuracy of rainfall forecasts for applications in urban areas.

4.4.2 River Catchments

As mentioned in Chapter 3, the PDM for the river catchments described were calibrated using the shuffled complex evolution method (SCE-UA) (Duan, Sorooshian and Gupta, 1992, 1994) in order to obtain a maximum Nash-Sutcliff efficiency (NSE) coefficient. The calibration was carried using one-year radar data from 2007. Data from 2008 was used for validation. The calibrations of all the catchments were done using both radar and rain gauge data, and therefore two models for each catchment were produced.

Table 4-4 contains the NSE coefficient for the six river catchments studied for both calibration and validation data periods. The results obtained for the calibration and validation for Arnford, Henthorn and Killington catchments are also presented below (Figure 4-15 until Figure 4-17). For all the catchments, rain gauge validation results were better than radar. Catterick and Crakehill validations showed that the model for these catchments could not reproduce large peaks, hence the poor NSE performance.

Table 4-4. NSE Coefficient for calibration and validation of the PDM for the river catchments studied using radar and gauge

Event Date	Arnford	Brigflatts	Catterick	Crakehill	Henthorn	Killington
Calibration (radar)	0.82	0.76	0.83	0.90	0.85	0.77
Calibration (gauge)	0.88	0.84	0.88	0.90	0.90	0.92
Validation (radar)	0.76	0.68	0.58	0.42	0.79	0.62
Validation (gauge)	0.89	0.84	0.84	0.78	0.89	0.90

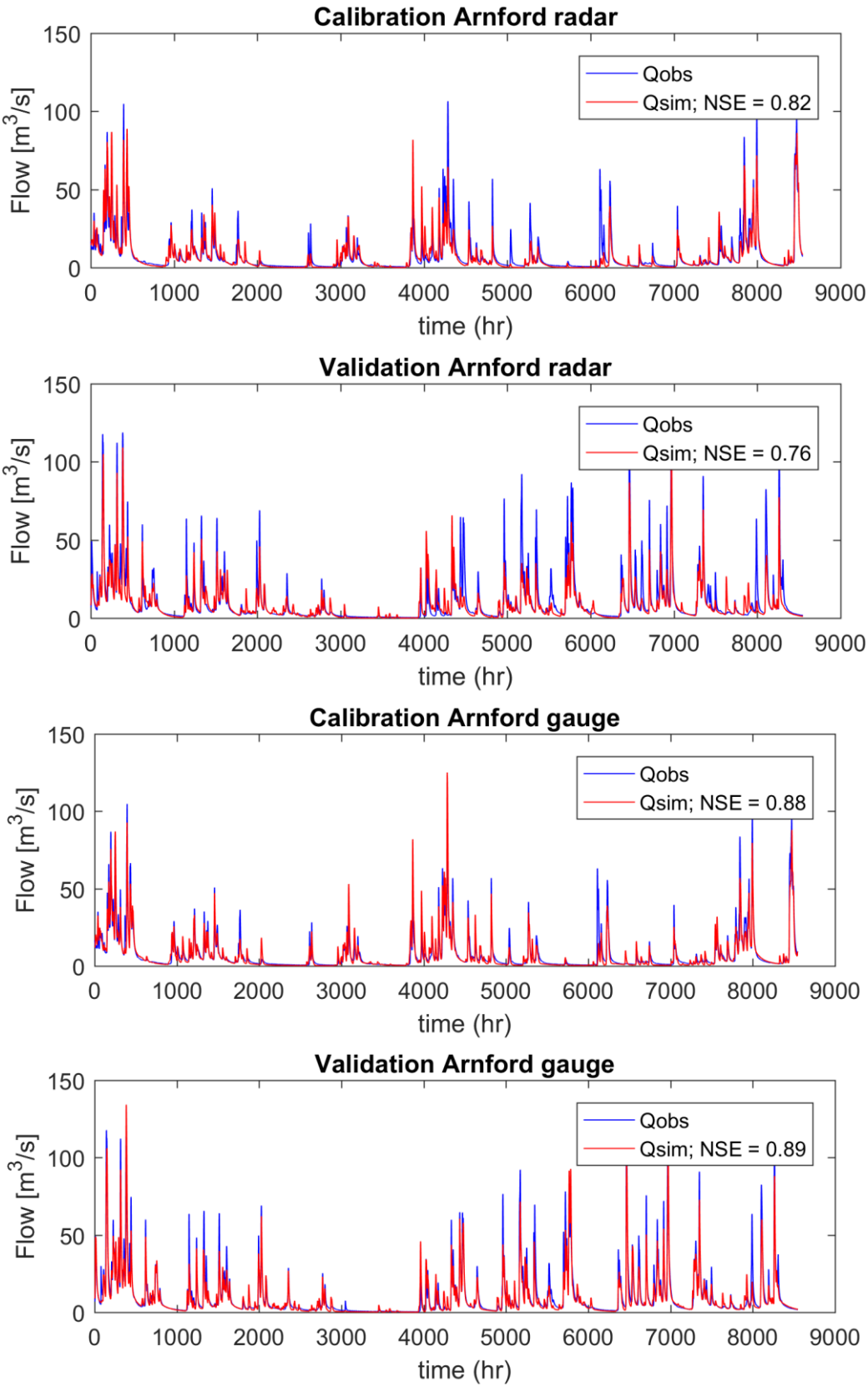


Figure 4-15. Calibration and Validation of Arnford catchment, where Qobs represents the observed flow and Qsim represents the simulated flow using radar or gauge data.

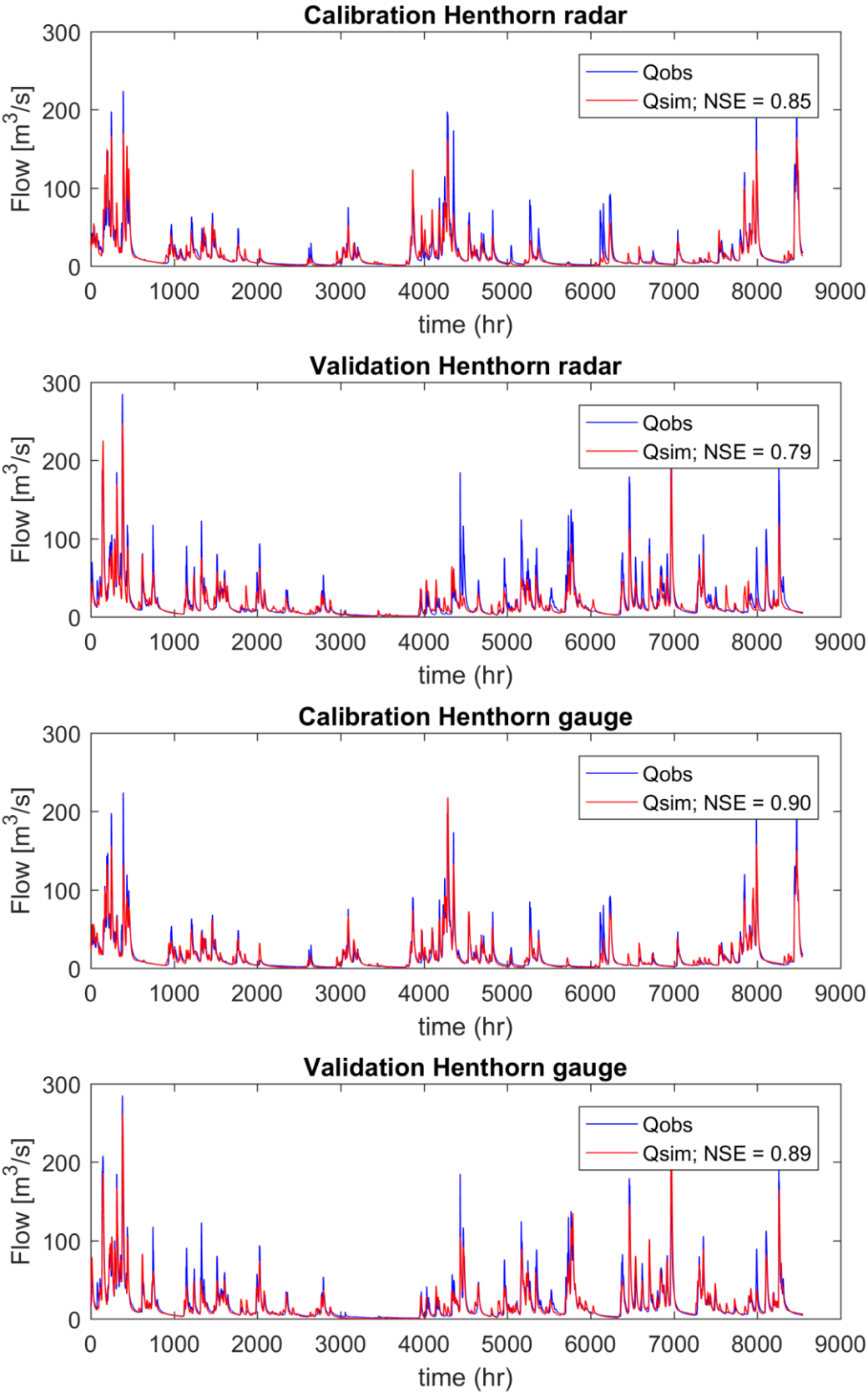


Figure 4-16. Calibration and Validation of Henthorn catchment, where Qobs represents the observed flow and Qsim represents the simulated flow using radar or gauge data.

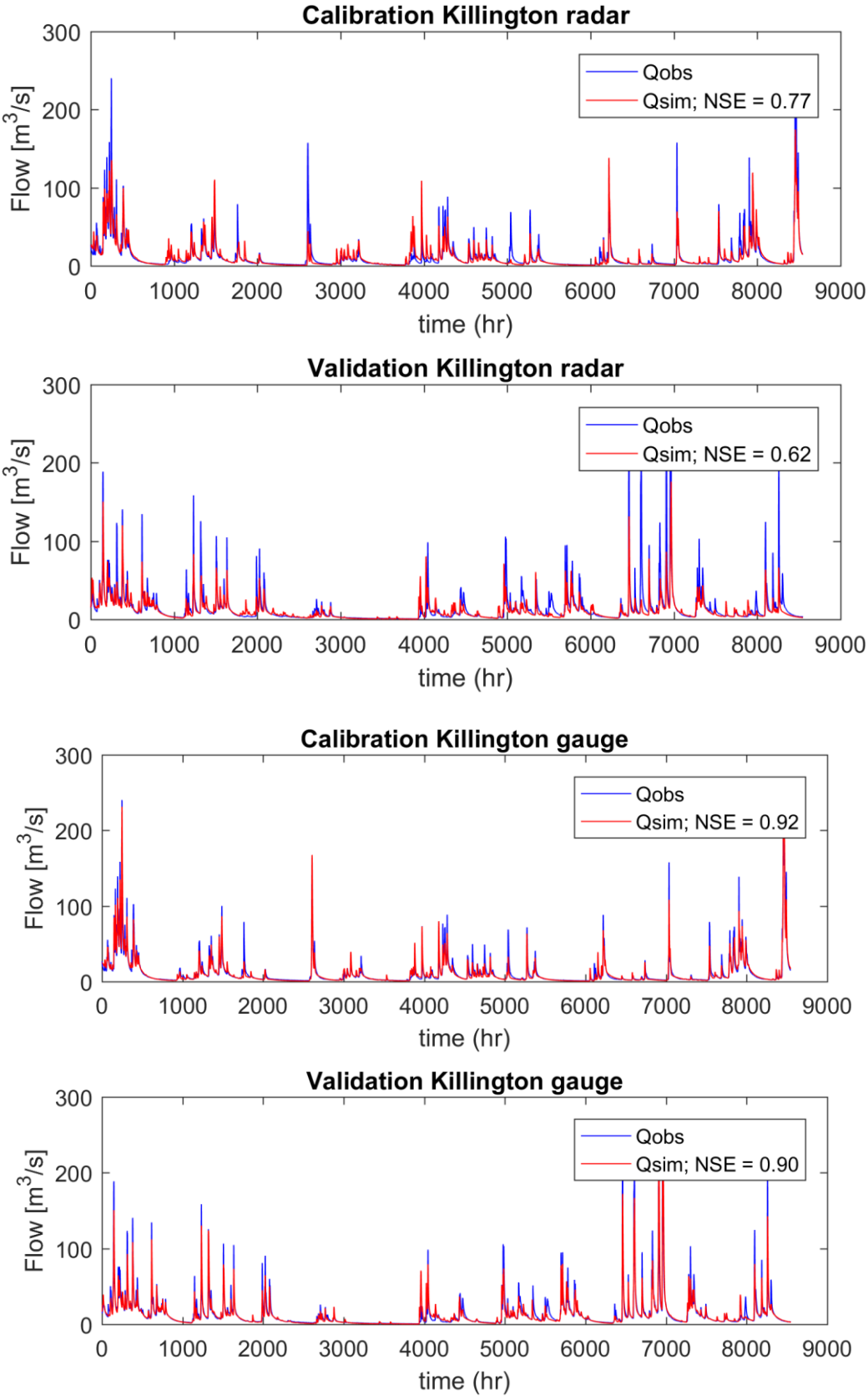


Figure 4-17. Calibration and Validation of Killington catchment, where Qobs represents the observed flow and Qsim represents the simulated flow using radar or gauge data.

The RE ensembles produced better forecasts in the vast majority of the cases, generating more accurate results when the radar data could not accurately estimate the flows. Taking radar uncertainties into account in the forecasts enabled flow peaks to be forecasted even when the rain gauges cannot help to estimate the simulated flows correctly. As some of the ensembles produced are closer in value to the deterministic forecast and some are able to predict the flow peaks, the ensemble spread was higher for the RE ensembles for a majority of the cases. In Figure 4-18 and Figure 4-19, the forecasted flow that falls between the 15 and 85 percentiles, measured flow, estimated flow using radar and gauge data for three different rural catchments are shown for an event starting on 04/10/2008 at 14:00 for RE and STEPS ensembles respectively. The RE ensembles for the Arnford catchment produce forecasts with a higher spread, but it can also predict the peak more accurately, although both probabilistic forecasts underestimate the flow. For the Henthorn catchment, there is a significant difference between the forecasts produced. The STEPS ensembles are close to the deterministic forecast, and it completely fails to reproduce the measured and estimated flows pattern. On the other hand, the RE ensembles present a very large spread, but both estimated and measured flows are between 15 and 85 percentiles. The estimated flows using radar and gauge data present pretty different results, with the estimated flow based on gauge data being closer to the measured flow, implicating that in this case, the use of RE ensembles can clearly improve the outcome of the forecast. For the Killington catchment, however, the estimated flow using radar and gauge are pretty similar. The fact of using a historical data set for the radar rainfall errors means that results presented by the RE ensembles are again more accurate than the STEPS ensembles and are capable of predicting the measured flow peak and pattern.

The Root Mean Square Error (RMSE) and the goodness of fit were calculated for all events to assess the results. The goodness of fitness (GOF) estimator used was proposed by Cecinati et al. (2017) and is calculated as shown on the equation below.

$$GOF = 1 - \frac{\sum_{i=1}^{n_{tot}} q_1 + \sum_{i=1}^{n_{tot}} q_{100}}{n_{tot}} \quad (4\text{-}14)$$

Where q_1 is the number of time steps in the first quartile, q_{100} is the number of time steps in the last quartile, and n_{tot} is the number of time-steps. The GOF measures how well the ensembles can match the measured flow by calculating the number of time steps where flow falls between the first and last quartile. The GOF ranges from 0 to 1. A GOF equal to 1 means that the measured flow in all time steps falls within the ensemble spread.

Table 4-5, Table 4-6,

Table 4-7 and Table 4-8 show an overview of the events studied for the four different catchments, including event start date, duration, peak flow, maximum average rainfall, GOF score and storm types.

Note that these events were selected because they produced a significant peak flow, and therefore these events are more interesting for forecasting purposes.

Table 4-5. Events dates, duration peak measured flow, maximum average rainfall, average GOF scores for RE and STEPS ensembles and storm types (S – stratiform and C – convective) for the Arnford catchment.

Arnford						
Event	Duration (h)	Peak flow (m3/h)	Max average rainfall (mm/hr)	GOF_{RE}	GOF_{STEPS}	Storm type
200801131400	25	57.74	1.02	0.28	0.16	S
200801170700	12	64.30	3.90	0.38	0.23	S
200804300000	5	104.88	5.38	0.49	0.37	S
200806261300	4	16.31	0.76	0.33	0.20	S
200809050800	19	281.38	3.80	0.27	0.29	C
200810040800	18	145.44	3.55	0.40	0.20	S
200810141500	3	99.50	2.72	0.76	0.51	C
200811082000	5	61.94	2.69	0.35	0.17	S
200811091500	3	115.13	2.03	0.60	0.33	S

Table 4-6. Events dates, duration peak measured flow, maximum average rainfall, average GOF scores and storm types (S – stratiform and C – convective) for the Brigflatts catchment.

Brigflatts						
Event	Duration (h)	Peak flow (m3/h)	Max average rainfall (mm/hr)	GOF_{RE}	GOF_{STEPS}	Storm type
200801131400	32	66.16	1.51	0.24	0.21	S
200801170100	16	87.74	3.22	0.26	0.33	S
200804300000	5	38.80	4.52	0.51	0.44	S
200806261300	5	10.36	1.75	0.60	0.56	S
200808010300	3	16.52	3.20	0.50	0.59	C
200809050800	22	134.85	6.71	0.45	0.50	C
200810040800	18	77.30	2.80	0.37	0.24	S
200810141500	3	34.68	2.21	0.48	0.58	C
200811082000	5	44.15	2.69	0.43	0.30	S
200811091700	2	85.09	1.37	0.55	0.64	S

Table 4-7. Events dates, duration peak measured flow, maximum average rainfall, average GOF scores and storm types (S – stratiform and C – convective) for the Henthorn catchment.

Heathorn						
Event	Duration (h)	Peak flow (m3/h)	Max average rainfall (mm/hr)	GOF_{RE}	GOF_{STEPS}	Storm type
200801131400	32	76.05	3.91	0.31	0.20	S
200801170100	22	71.98	2.60	0.21	0.11	S
200804300000	5	22.88	0.77	0.61	0.51	S
200806261300	4	22.10	1.10	0.24	0.14	S
200808010300	3	52.23	10.75	0.06	0.06	C
200809050800	22	74.90	2.54	0.30	0.31	C
200810040800	20	204.75	8.05	0.45	0.19	S
200810141500	4	95.10	4.39	0.78	0.65	C
200811082000	5	66.28	6.07	0.34	0.18	S
200811091500	3	103.00	1.31	0.76	0.65	S

Table 4-8. Events dates, duration peak measured flow, maximum average rainfall, average GOF scores and storm types (S – stratiform and C – convevtive) for the Killington catchment.

Killington						
Event	Duration (h)	Peak flow (m3/h)	Max average rainfall (mm/hr)	GOF_{RE}	GOF_{STEPS}	Storm type
200801131400	32	122.00	4.57	0.33	0.18	S
200801170100	16	55.63	3.59	0.45	0.42	S
200804300000	5	21.88	4.37	0.40	0.27	S
200806261500	3	58.63	0.15	0.37	0.35	S
200809050800	22	101.75	5.72	0.39	0.42	C
200810040800	16	347.50	9.35	0.56	0.31	S
200810141500	3	123.25	4.11	0.70	0.61	C
200811082000	5	41.63	4.33	0.55	0.30	S
200811091700	2	86.85	2.56	0.82	0.27	S

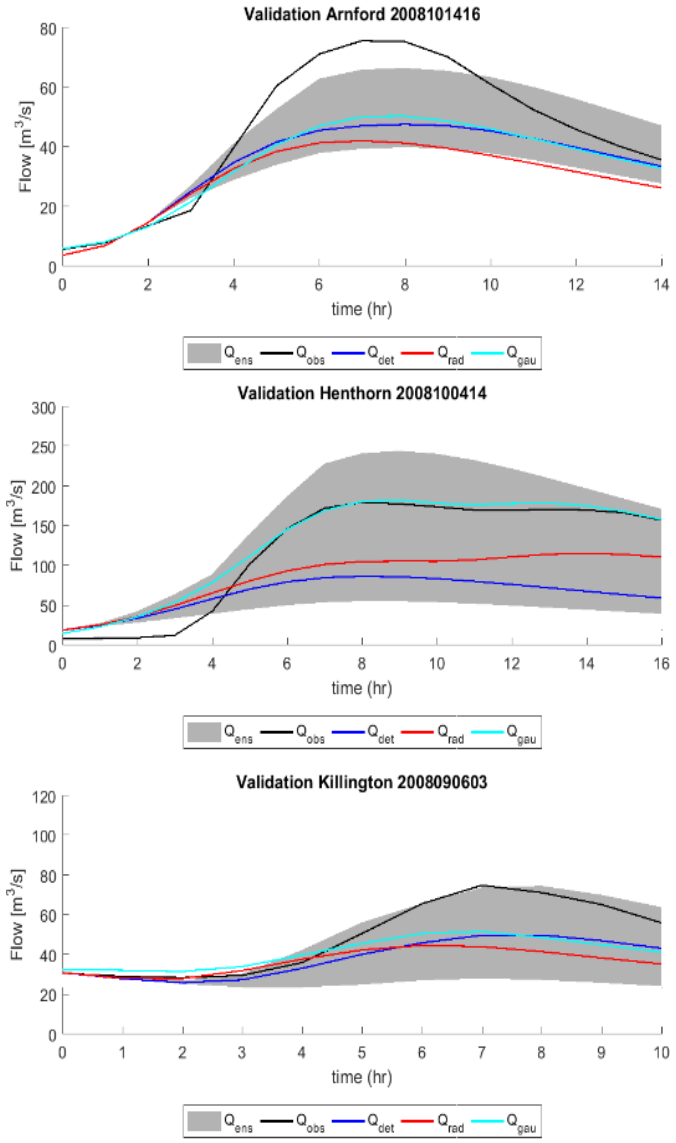


Figure 4-18. Flows for events on 14/10/2008 starting at 16:00 at Arnford catchment, 04/10/2008 starting at 14:00 at Henthorn catchment, and 06/09/2008 starting at 03:00 at Killington catchment. Q_{ens}, Q_{obs}, Q_{det}, Q_{rad} and Q_{gau} are, respectively, forecasted RE ensemble flows, measured flow, forecasted deterministic flow, estimated radar flow and estimated gauge flow.

To summarize the results, the mean RMSE for the RE and STEPS ensembles are shown in Figure 4-20 and Figure 4-21; the GOF for different events are presented in Table 4-5,

Table 4-6,

Table 4-7 and Table 4-8. The RMSEs for the RE ensembles are lower than for the STEPS ensembles (as shown in Figure 4-20 and Figure 4-21) for most of the events. However, in some events, these differences are very small for the two types of ensembles. Although the RMSE still provides better results for the RE ensembles, the slight difference in values would not justify the use of RE ensembles if only this method were used to compare the forecasts. RE ensembles take a longer time and require

more computer power to be produced than the STEPS ensembles. For most cases, the GOF for RE ensembles was higher than for the STEPS ensembles, with the GOF_{RE} being more than double of GOF_{STEPS}. This indicates that RE ensembles are able to capture more often the measured flows.

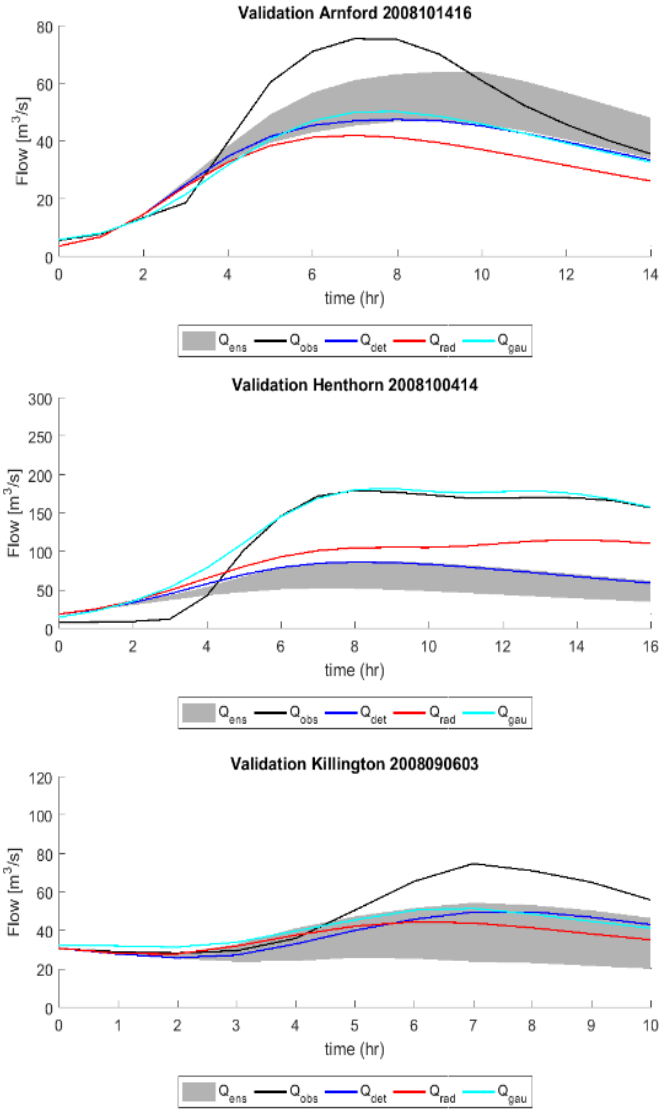


Figure 4-19. Flows for events on 14/10/2008 starting at 16:00 at Arnford catchment, 04/10/2008 starting at 14:00 at Henthorn catchment, and 06/09/2008 starting at 03:00 at Killington catchment. Q_{ens}, Q_{obs}, Q_{det}, Q_{rad} and Q_{gau} are, respectively, forecasted STEPS ensemble flows, measured flow, forecasted deterministic flow, estimated radar flow and estimated gauge flow.

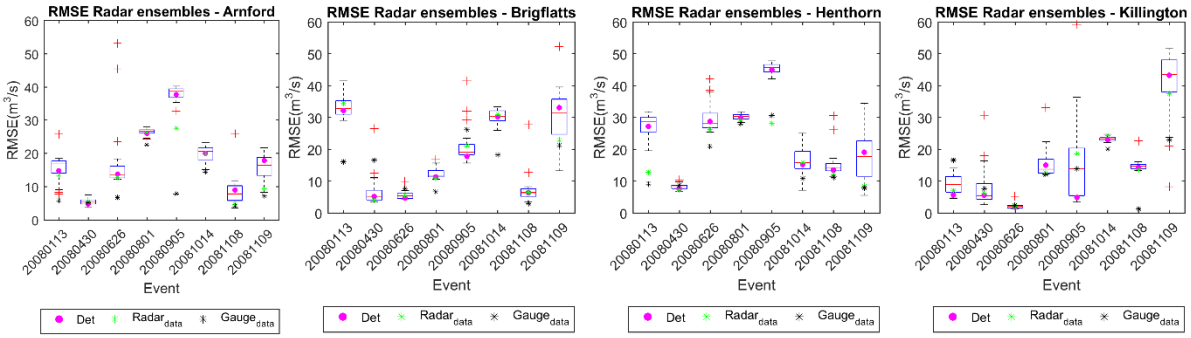


Figure 4-20. RMSE for RE ensembles at Arnford, Brigflatts, Henthorn and Killington catchments. $Radar_{ens}$, Det, $Radar_{data}$, and $Gauge_{data}$ are, respectively, RE ensemble forecast, deterministic forecast, estimated radar flow and estimated gauge flow.

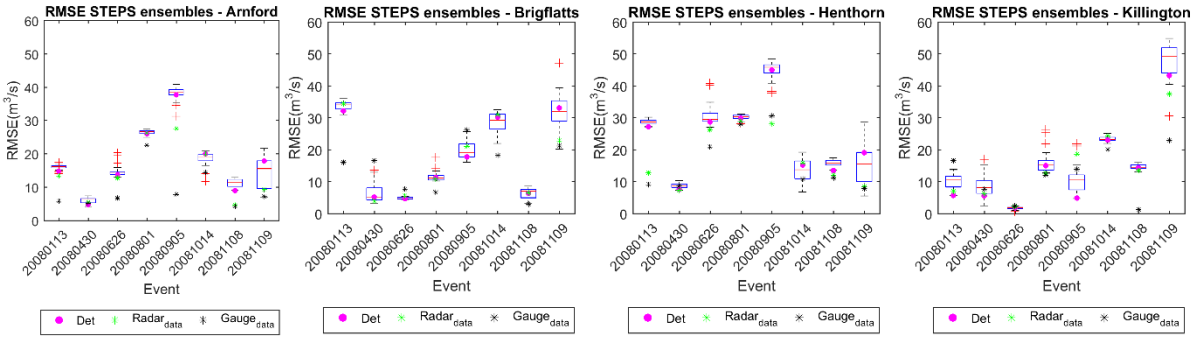


Figure 4-21. RMSE for STEPS ensembles at Arnford, Brigflatts, Henthorn and Killington catchments. $Radar_{ens}$, Det, $Radar_{data}$, and $Gauge_{data}$ are, respectively, STEPS ensemble forecast, deterministic forecast, estimated radar flow and estimated gauge flow.

The GOF results show that in almost 85% of the cases, the RE ensembles produce a better fit than the STEPS ensembles. In most cases where the STEPS produces better forecasts, the difference between GOF_{STEPS} and GOF_{RE} is less than 10%. Results also show that the storm type does not necessarily lead to more uncertainties in flow predictions, as the convective events do not always have the highest RMSE and lower GOF values.

4.5 Conclusion

This research assessed how radar rainfall uncertainties propagate from the radar rainfall measurements into radar rainfall forecasts and, further on, into urban sewer forecasting. The work also compared the accuracy of flow forecast prediction using two different rainfall ensemble generators. A stochastic ensemble generator, which adds spatially correlated noise to the deterministic forecast, was used as a reference (STEPS ensembles). An ensemble generator that adds spatially correlated noise based on the residual radar error between radar rainfall and rain gauge measurements was used to assess the radar rainfall uncertainties (RE ensembles) and how they propagate into the simulated sewer flows of an urban area.

Results from the rainfall forecasts show that both ensembles can produce skilful forecasts for lead-times up to 3 hr for all rainfall intensities larger or equal to 0.1 mm/hr. With intensities larger than 1

mm/hr, the results vary depending on the event. Skilful forecasts could be produced up to 1 hr lead-time in most cases, and in some events, it was possible to produce a skilled forecast even when the lead-time was 3 h. For a high rainfall threshold (larger than 3 mm/h), reliable forecasts were produced for at least 1-h lead-time. Foresti *et al.* (2016) used a modified version of STEPS to produce ensembles and concluded that for high-intensity events (5 mm/hr), the forecast is only reliable for 30 min. For lower intensities, the ensembles produced by Foresti *et al.* (2016) could forecast rainfall up to 90 min ahead. The RE ensembles produced slightly better results than the STEPS ensembles in the first hour of the forecasts. After this, the RE ensembles lost accuracy more rapidly than the STEPS ensembles.

The flow forecasts in the urban areas were generated using an Infoworks CS model. The RE ensembles produced better results than the STEPS ensembles in the first hour of the forecasts and were able to reproduce the flow peaks better. The RE ensembles RMSE was lower than the STEPS ensembles RMSE for most of the events in the first hour forecasted. This was true even in cases where the simulated flow using the radar rainfall is overestimated or underestimated, thus being able to reproduce the flow hydrograph better. With lead-times longer than 1 h, all the forecasts lose predictability independently of which ensemble generator is used.

The results show a time-lag of a few minutes between the measured and forecasted peak flows for the urban catchment. However, even with this limitation, it is possible to improve the forecast of peak flows in urban areas using the method proposed. This method can be used in real-time to enhance existing warning systems in urban areas with up to one-hour lead-time. The nowcast skill can be potentially improved by blending radar nowcasts with NWP forecasts, especially as the forecasting lead time increases. With up to one hour lead-time, the nowcast has a significant impact on the forecast, and this improvement enhances the flow forecasts during this period. Future works can incorporate uncertainties caused by the growth and decay of precipitation using, for instance, the method described by Foresti *et al.* (2018).

For river catchments, the RE ensembles produced forecasts with higher accuracy at predicting the flow peaks when compared to the STEPS ensembles for most cases. While the RE ensembles produced forecasts with a higher spread, large peaks that the STEPS ensembles would miss could be forecasted by the RE ensembles. RE ensembles' advantage was more evident in situations when radar data could not estimate the peak. However, it also struggled in events where both radar and rain gauge simulated flows were very different from the observed flow. For the events analysed, the convective events did not have the highest RMSE and lowest GOF, meaning that the storm type in these cases did not show a tendency of more difficult prediction. In both river and urban catchments, the underestimation in high-intensity peaks is challenging to predict, also highlighted by Zhu, Xuan and Cluckie, (2014).

Chapter 5. Ensemble forecasts based on the temporal variation of the velocity field

5.1 Introduction

Nowcast models based on Lagrangian extrapolation of rainfall are important tools to predict precipitation at short lead-times. However, they can quickly lose their predictability power due to uncertainties inherent to the nowcasting model, uncertainties in the initial radar rainfall analysis and uncertainties in precipitation evolution (e.g. precipitation growth and decay). While radar errors count for most of the nowcast uncertainties at the first hour forecasted, changes in the rainfall pattern, such as growths and decay and temporal evolution of rainfall, count for most of the uncertainties as the lead-time increases (Foresti *et al.*, 2013).

Rainfall events present high spatial and temporal variability (Berne and Krajewski, 2013), and using high temporal resolution data from weather radar can help capture the most recent precipitation distribution. The forecast's temporal resolution depends not only on the data available but also on the forecast application. Temporal resolutions ranging from 5 -15 min are more commonly used, but time-steps of 30 min are also found in the literature (Pierce *et al.*, 2000, 2004; Tilford, Sene and Collier, 2003; Ebert *et al.*, 2004; Turner, Zawadzki and Germann, 2004; Wilson *et al.*, 2004; Schroeder *et al.*, 2006; Liguori and Rico-Ramirez, 2012a; Schellart *et al.*, 2012; Caseri *et al.*, 2016; Foresti *et al.*, 2016; Kato *et al.*, 2017; Simonin *et al.*, 2017). Therefore, rainfall information from up to 1 hr before the start of the event can be used to generate the rainfall forecasts. As it is expected, the radar precipitation distribution measured during this period can have a high variation. Nonetheless, the radar rainfall scans still are able to produce meaningful forecasts.

In order to account for the uncertainties due to the temporal evolution of rainfall advection fields and growth and decays, it is common to use probabilistic forecasts by adding spatially correlated stochastic noise to the deterministic forecast (Seed, 2003; Bowler, Pierce and Seed, 2006; Berenguer, Sempere-Torres and Pegram, 2011). STEPS uses random velocity fields to account for uncertainties related to the temporal evolution of rainfall. Atencia and Zawadzki (2014) proposed a stochastic ensemble-generation technique by producing a reflective field that keeps the temporal and spatial structures of the rainfall advection fields, being able to represent some of the forecast uncertainties. The proposed ensemble generator has some conceptual similarities with SBMcast and STEPS related to the conservation of some properties (e.g. temporal correlation and power-spectrum slope) to generate the ensemble members. The main difference is that while SBMcast and STEPS only maintain these properties in the

first hour forecasted, the ensemble generator is able to add changes to these properties, being able to keep them in the forecast for up to 10 hours. Panziera *et al.*, 2011 proposed a method called NORA (Nowcasting of Orographic Rainfall by means of Analogues), an analogue approach to produce ensembles for orographic regions by comparing the present situation to an analogue one from extensive historical data. The analogues are compared using mesoscale flows, air mass stability, average rainfall rate and proportion of wet pixels in the radar image. Foresti *et al.* (2013) extended the NORA model to also consider the analysis of rainfall advection fields when retrieving analogue events. Ensembles then can be produced by taking into account the temporal evolution of the rainfall present in analogue events. Although the method could represent the forecast uncertainties in a number of situations, it fails in cases of extreme events that have not been observed before. Atencia and Zawadzki, 2015 used a 15-yr dataset to produce forecasts using the analogue approach. The study concluded that this is due to Lagrangian extrapolation based nowcasts do not take into account the temporal evolution of rainfall advection fields and growth and decay of rainfall. Analogue based seems promising to replicate the temporal variability of rainfall, but the main issue remains in retrieving the analogues and the requirement of having a long historical dataset. Further research in the application of analogues is still necessary.

The idea of using ensemble forecasts for accounting for the uncertainties due to the temporal evolution of the rainfall advection fields is not new. However, changes in the temporal evolution of precipitation are challenging to predict, and therefore the methods that use historical data are unable to capture this evolution. In this chapter, a new method to produce ensembles is proposed by using the information on how the storm developed during the last two hours prior to the forecast. In this way, any changes in the temporal evolution of precipitation are event-specific that can potentially produce more meaningful ensemble forecasts.

5.2 Methodology

It is well known that using a constant rainfall advection field during the duration of the forecast is known to produce uncertainties because the temporal evolution of precipitation is not accounted for. Figure 5-1 shows the average advection velocity fields and their standard deviation. The velocity fields were calculated for a particular precipitation event during a 3-hr window, and this illustrates how the rainfall advection fields can actually vary during the forecast. In some regions, the rainfall fields are not subjected to significant changes (e.g. south of the region shown on the right of Figure 5-1). Consequently, the uncertainties related to the temporal evolution of rainfall advection fields remain small. However, there are regions where the rainfall advection of precipitation can be more uncertain,

making it clear that this assumption can produce uncertain forecasts (e.g. east of the region shown on the right of Figure 5-1).

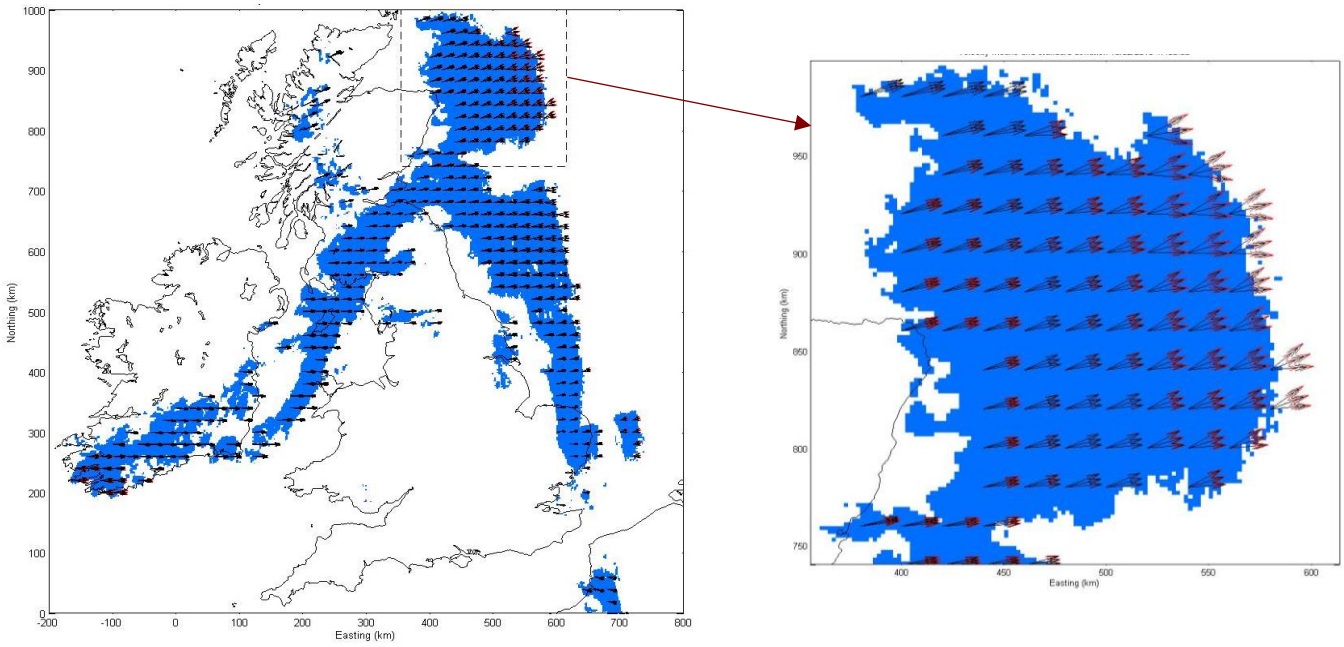


Figure 5-1. Average advection velocity fields and their standard deviation calculated over a 3 hr window. The blue area shows the regions where there was rainfall during this period.

The nowcasting component from the STEPS model extrapolates the rainfall advection field by solving the optical flow constraint over an area using the least-square approach. To avoid issues with the derivative estimation, a single advection vector, found by maximizing the correlation between the last two radar images, is used on the entire domain before solving the optical flow constraints. After solving the optical flow constraints, a smoothness constraint is applied to the advection field. The STEPS model uses a backward in time advection to avoid stripes that can appear when the rainfall advection fields are diverging. The rainfall advection field is perturbed using a Gaussian distributed field with white noise to produce an ensemble member. In this way, the perturbed rainfall advection field keeps the proper correlation structure. Further details of how the velocity fields are calculated can be found in section 3.3.1. However, although the precipitation field is perturbed, the advection field remains constant, which is clearly not the case, as shown in Figure 5-1.

Therefore, a new method to account for the uncertainties in the advection fields to generate ensemble forecasts is proposed. As different time-steps can be used to generate the advection field and the forecasts, the idea is to use rainfall advection fields at different time-steps. It means that not only the most recent radar images can provide relevant information on how the rain develops with time, but also precipitation fields from previous time steps. Accounting for how the rainfall evolved during a set period can give insights into how it will change in the near future. The ensemble generator proposed

uses radar rainfall advection fields measured up to two hours before the start of the forecast. Different velocity fields are computed using radar images that are between 5 and 30 min apart, which can be used to advect the most recent radar images. In order to analyse how far back in time, the velocity fields still influence and add skill to the forecasts, tests were made going back in time 30-, 60-, 90- and 120 min. The choice of these time windows is linked to the number of ensembles produced. Using a shorter time window would mean that not enough ensembles would be part of the probabilistic forecast. Simultaneously, producing 63 ensembles is already time demanding, so using a more extended period would only be advisable if, after analysing the results, the difference between having 45 and 63 ensembles was high, suggesting a much better performance at longer time-windows.

Figure 5-2 shows the number of velocity fields that can be generated to produce a probabilistic forecast just by doing combinations of precipitation fields during the previous 30 min before the start of the forecast. The rainfall advection fields calculated using radar estimations 5, 10 and 15 min apart are used. These time-steps between radar images were chosen based on the 5 min temporal resolution of the radar measurements. By using a time window of 30 min, 9 ensembles are produced. In an analogue way, using a time window of 60 min, produce 27 ensembles; using a 90 min time window produces 45 ensembles and having a 120 min time window is able to generate 63 ensembles. The choice of not going back further in time is that this would require a very large number of ensembles to be produced. This would be very demanding, requiring more computational power and a longer time to produce the forecasts. Every velocity field will produce a forecast that is moving faster or slower or in a slightly different direction (see, e.g. Figure 5-1) and therefore accounting for the temporal evolution of precipitation. Generating ensembles in this way can potentially produce more meaningful ensemble forecasts, as information from a particular precipitation event is used instead of adding random white noise to generate the ensembles.

To generate the ensemble rainfall forecasts, advection velocity fields were produced every 5 min with 5-, 10- and 15 min time-steps between them. As the STEPS model calculates the velocity fields for every forecast produced, deterministic forecasts were produced for each time window (forecasts every 5, 10 and 15 min). These velocity fields will be used to produce the forecasts. A modified version of the STEPS model was developed that allows a selection of the velocity field that will be used to advect the most recent radar rainfall advection field. The probabilistic forecasts were produced every 5 min, using the relevant velocity fields, for the length of the event using the different velocity fields.

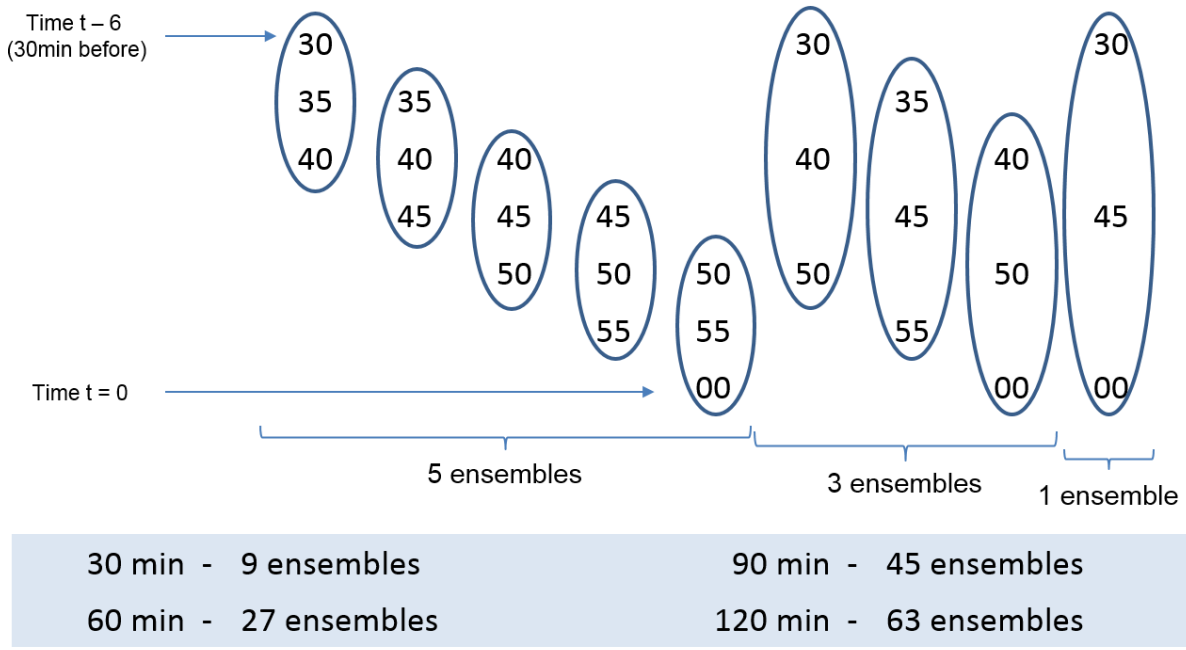


Figure 5-2. Example of velocity fields used to generate ensemble forecasts with a time window of 30 min and starting time at t= 00.

The results were assessed using the ROC curves, as described in Chapter 4. The ROC curves were produced using thresholds of 0.1 mm/hr. 1.0 mm/hr and 3.0 mm/hr. The ensembles were used as input to the urban model in the Ilkley catchment to assess how the rainfall ensemble forecasts are able to forecast floods in an urban area.

Forecasts were produced for 14 events, as described in Table 5-1. Although the flow was forecasted for the whole duration of the events, only the periods with high flow in Ilkley were taken into account to analyse the results in order to concentrate on how the model can forecast flow peaks. Initially, longer periods were accounted for, but as there were no problems with forecasting low and constant flow, it was decided to focus on the flow peaks. Besides being more challenging to be forecasted, forecasting flow peaks plays an essential role in real-time warning systems.

Table 5-1. Event start dates, duration and measured flow peaks.

Event Date	Starting Time	Duration (h)	Peak Flow (m^3/s)
29 April 2008	23:00	5.5	0.407
30 April 2008	15:30	0.5	0.348
01 May 2008	04:00	4	0.558
26 June 2008	13:00	6	0.192
7 July 2008	15:00	2	0.511
29 July 2008	03:00	4	0.588
1 August 2008	03:00	1	1.041
20 August 2008	22:00	2	0.482
2 November 2008	13:00	2	0.462
4 October 2008	17:00	11	0.359
14 October 2008	15:00	2	0.279
8 November 2008	20:00	4	0.298
9 November 2008	03:30	13	0.199
13 December 2008	00:00	6	0.317

5.3 Results

5.3.1 Analysis using perturbations on the temporal evolution of velocity fields to generate short-term forecast

The results were produced for all the events. Figure 5-3 and Figure 5-4 show examples of the radar image, the deterministic forecast and two ensemble members for the event on 04/10/2008 for 60 and 180 min lead-time. In Figure 5-3, although the ensembles are different from each other, they present a similar shape of the overall precipitation and are more agreeable with the radar image than for a longer lead-time. However, in Figure 5-4, it is easier to notice the difference between the ensembles and the impact that the velocity fields have in the forecast after 3 hr. In this particular example, the deterministic forecast and the ensembles fail to accurately forecast areas of higher rainfall intensity and wrongly forecast rainfall over large regions. This is mainly due to the changes in precipitation growth and decay.

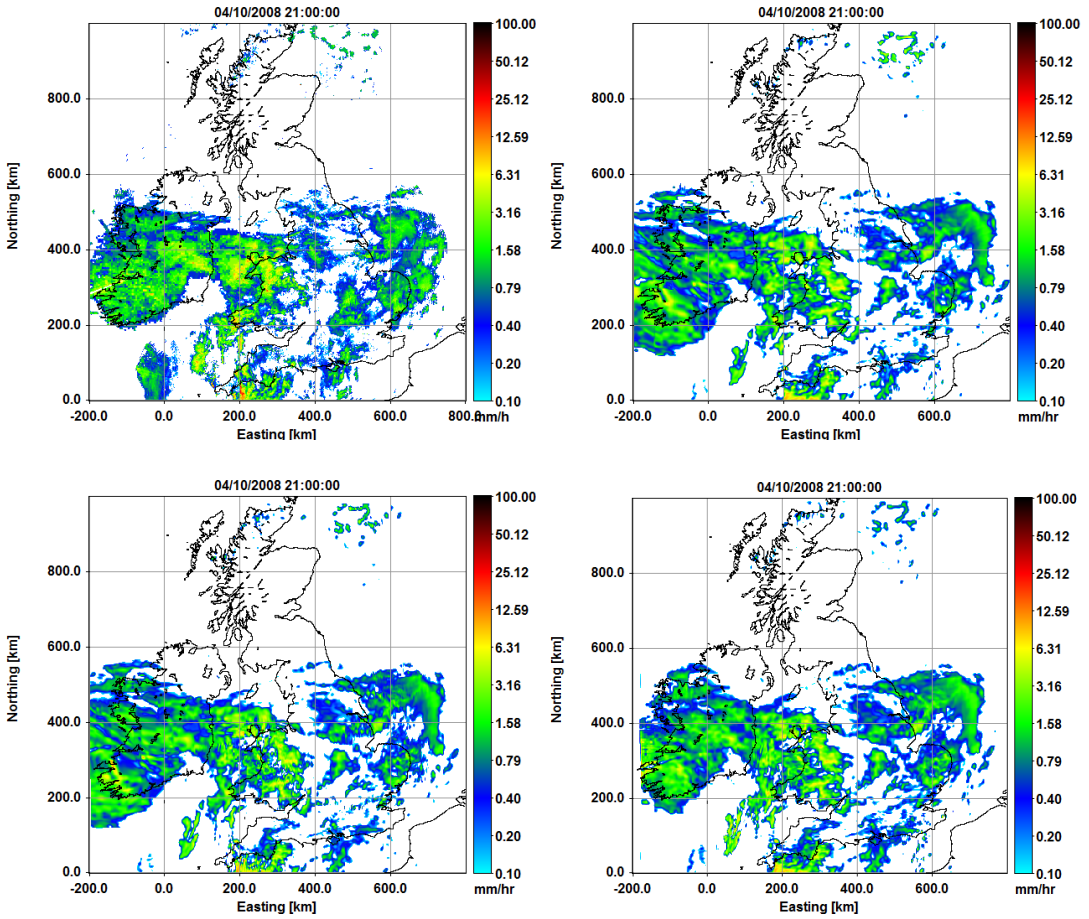


Figure 5-3. Radar image (top left) at 21:00 on 04/10/2008, deterministic forecast (top right) and two ensembles with 60 min lead time (bottom).

In order to assess the quality of the ensembles, ROC curves for each event were produced. Figure 5-5 shows the results for the ROC curves for events on 02/09/2008 and 14/10/2008 for a threshold higher than 0.1 mm/hr and for time-back windows of 30-, 60-,90- and 120 min. The results show that all the rainfall forecasts produce ROC curves with high areas underneath them. The area beneath the ROC curve increases with the number of ensembles in the forecast and decreases with longer lead time. The same pattern is seen across almost all the ensembles. Although, in some cases, the ROC areas are kept constant even when increasing the number of ensemble members. The ROC curves were also calculated for thresholds equal to 1.0 mm/hr and 3.0 mm/hr.

To summarize the results, Figure 5-6 until Figure 5-8 present the results for thresholds 0.1-, 1.0- and 3.0 mm/ hr, respectively. The figures compare the area underneath the ROC curves for time-back windows of 30-, 60-,90- and 120 min and lead time of 30-, 60-, 120- and 180 min. In Figure 5-6, it is possible to observe a tendency of the area beneath the ROC curves increasing slightly with the number of ensembles. However, the difference is very small, and that can also be verified when looking on a case by case basis. Figure 5-7 shows that having more ensembles have more impact with higher thresholds, and this can also be verified in Figure 5-8, where this difference can be noted. The results

also show that there is more gain in using more ensembles rainfall forecast members for longer lead-times. For all the cases, having longer lead-time and higher thresholds lead to a smaller area beneath the ROC curve, and this pattern is observed independent of the number of ensembles in the forecast.

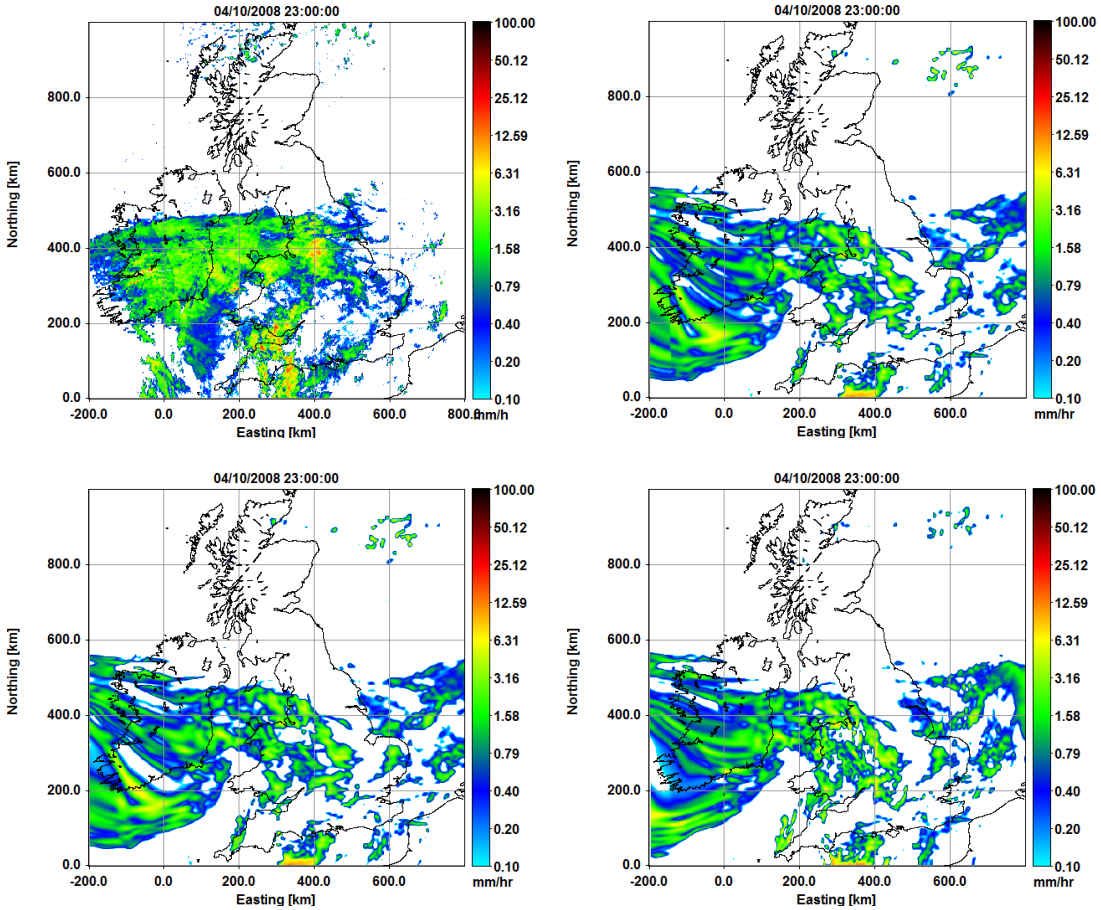


Figure 5-4. Radar image(top left)) at 21:00 on 04/10/2008, deterministic forecast (top right) and two ensembles (bottom) with 180 min lead time.

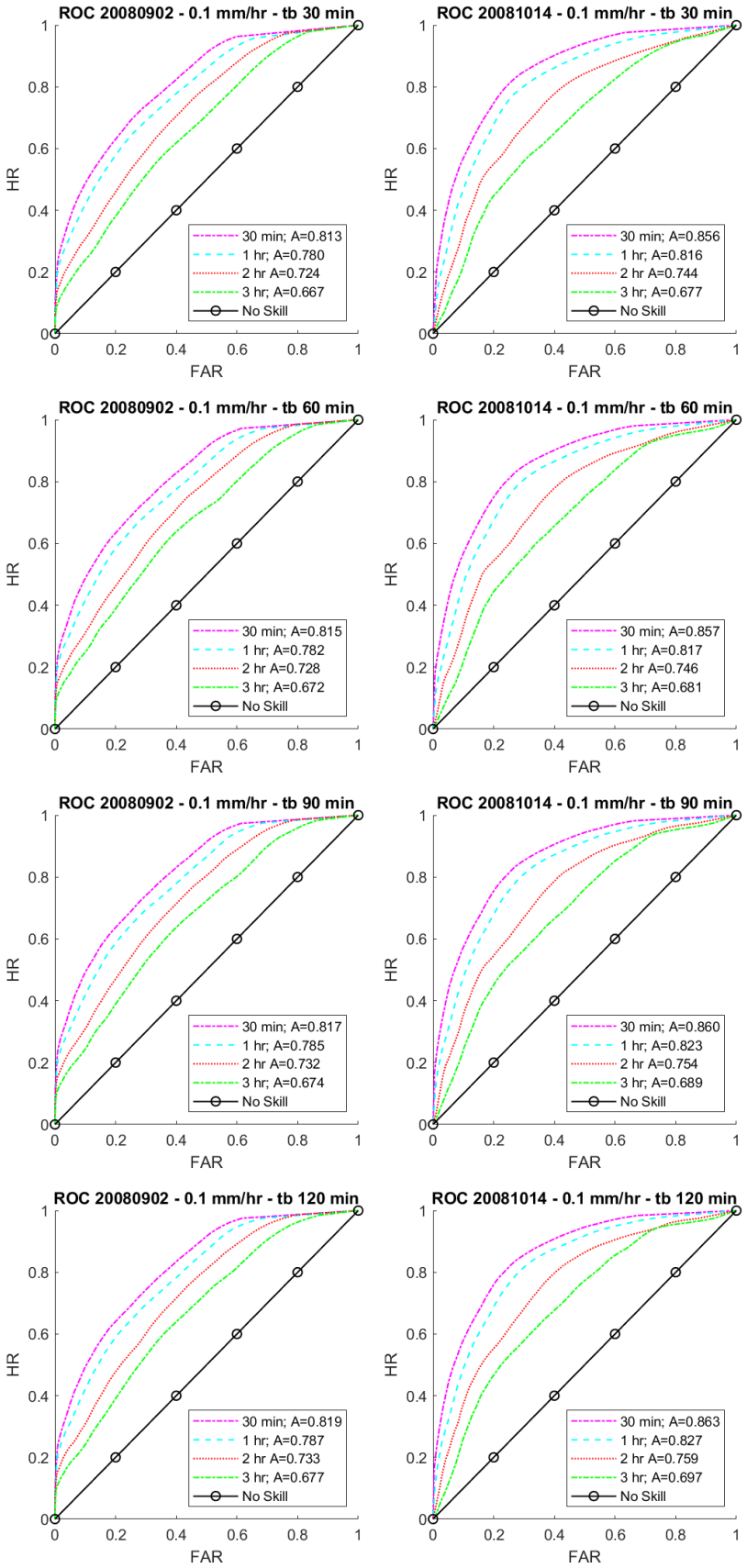


Figure 5-5. ROC curves for the events on 02/09/2008 and 14/10/2008 with 30 min, 1-, 2- and 3 hr lead time and time window of 30, 60, 90 and 120 min. (FAR – False Alarm Rate; HR – Hit Rate)

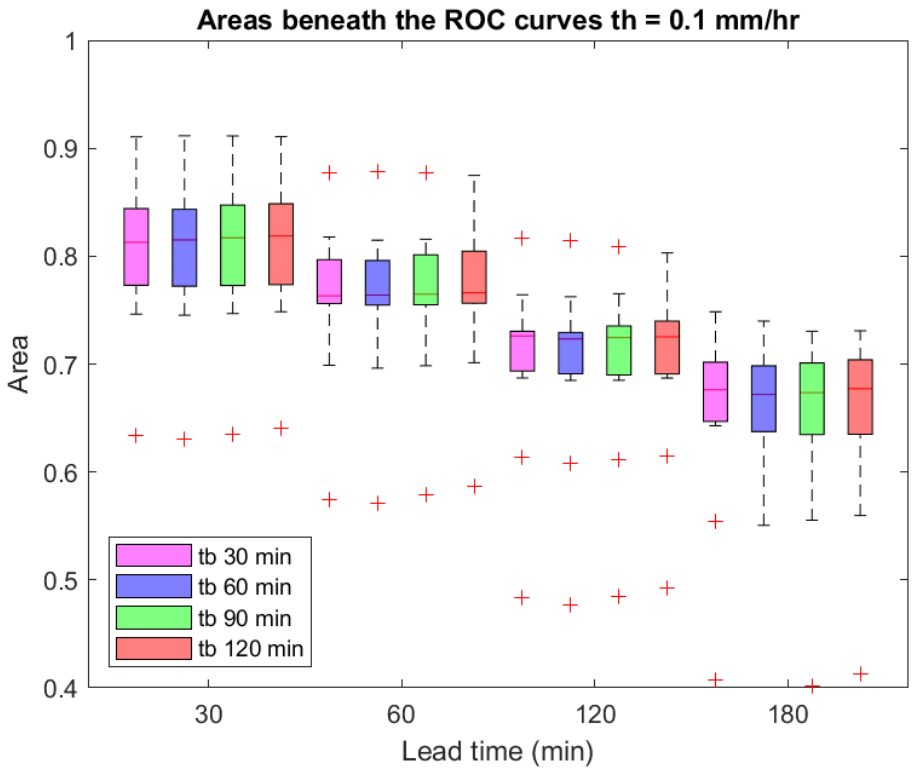


Figure 5-6. – Area beneath the ROC curve for a rainfall threshold of 0.1 mm/hr, time window of 30-, 60-, 90- and 120 min and forecasting lead times of 30-, 60-, 120- and 180 min. (tb refers to the time window used to produce the advection fields)

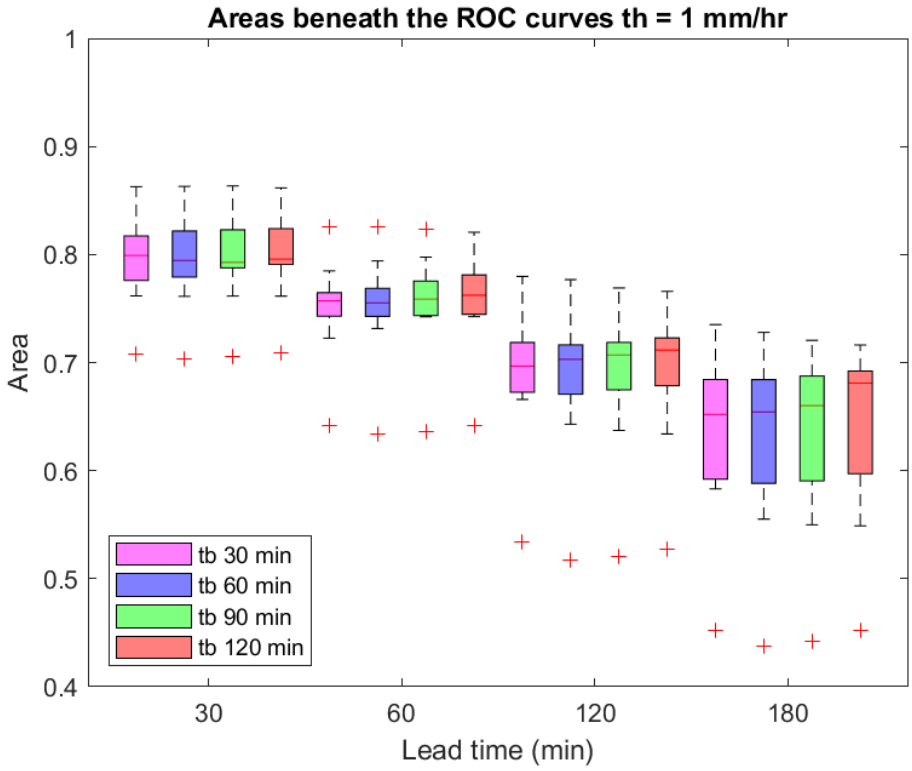


Figure 5-7. – Area beneath the ROC curve for a rainfall threshold of 1.0 mm/hr, time window of 30-, 60-, 90- and 120 min and forecasting lead times of 30-, 60-, 120- and 180 min.

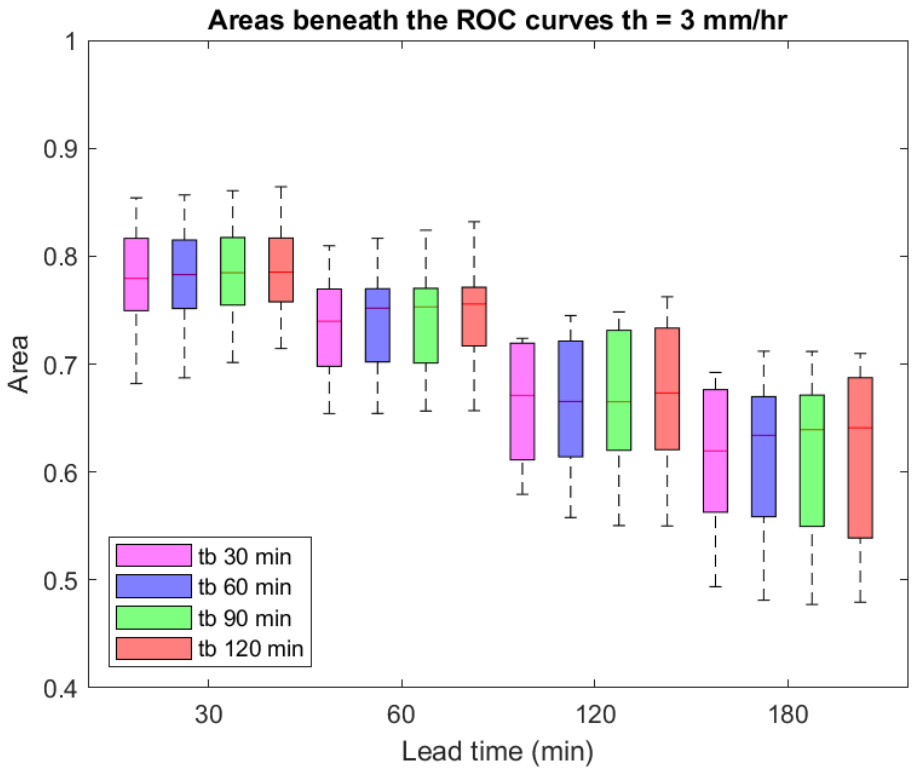


Figure 5-8. – Area beneath the ROC curve for a rainfall threshold of 3.0 mm/hr, time window of 30-, 60-, 90- and 120 min and forecasting lead times of 30-, 60-, 120- and 180 min.

5.3.2 Hydrological application in flow predictions in Ilkley

The deterministic and ensemble rainfall forecasts were used as input to an urban model for sewer flow simulation. The results were compared with the measured flow, simulated flow using radar rainfall data and simulated flow using rain gauge data.

Figure 5-9 and Figure 5-10 show results for events on 20/08/2008 at 20:30 and 04/10/2008, respectively, with a time-back window of 30-, 60, 90-, and 120 min. Figure 5-11 show results for the same events from Figure 5-9 and Figure 5-10 using STEPS ensembles to produce forecasts. The forecasts were produced using 25 ensembles. In Figure 5-9, the first peak could not be forecasted. As both forecasts rely on radar data, if the radar data fails in estimating the flow, it is challenging for the flow to be correctly forecasted. The first peak is in the first hour of the forecast, and at this period, radar uncertainties play a major role in the forecast outcome. Only the rain gauge data was able to predict this peak, but still, the flow is underestimated. Rain gauge flow also underestimates the following peak. In this case, radar data is able to estimate the peak better than rain gauges, and consequently, the forecasts are also more reliable. The radar estimations slightly overestimate the peaks, while the deterministic forecasts underestimate them. This is due to changes in the rainfall pattern that are not

reproduced by the deterministic forecasts. The rainfall advection field forecast was able to estimate the peak for all time windows. However, for the peak that occurs around 1.5 hr after the start of the forecast, the deterministic forecast was able to predict the flow peak, however with a slight time lag and all the ensemble rainfall forecasts were able to capture the peak flow even with the existing time lag. The forecasts produced using different time windows all produced the same behaviour of presenting a higher spread with an increased lead time. For the deterministic forecast and ensemble rainfall forecasts, results are practically the same for lead time up to 30 min, showing that for this period, the differences in the velocity fields do not have a significant impact on the forecast. Figure 5-10 presents more constant flows during the length of the forecast that can be easily predicted. The forecasts are not able to predict the increased flow at the beginning of the forecast; however, from 30 min until 1 hr, it provided better results than the radar estimations that underestimated the flow. The rainfall advection field forecast was able to estimate the flows during most of the forecasting period. However, it was still able to predict the increased flow between 1.5 – 2 hr lead time. As is happened in all the events, differences between the deterministic forecast and the ensemble rainfall forecasts were only visible after 30 min lead time and increasing as the forecast advances.

Comparing the results obtained using the proposed method with probabilistic forecasts using STEPS ensemble generator (results from Chapter 4), it is possible to conclude that although the forecasts can be useful in flow forecasting, they are not able to predict peaks that radar data only fails to estimate. Being so, the first flow peak of the event on 20/08/2008 cannot be forecasted using any of those methods.

In the same way that flow simulations in Chapter 4, the flow peaks were slightly displaced in time when compared to the measured flow. For this reason, the flow peaks within a time window were used to assess the forecasts.

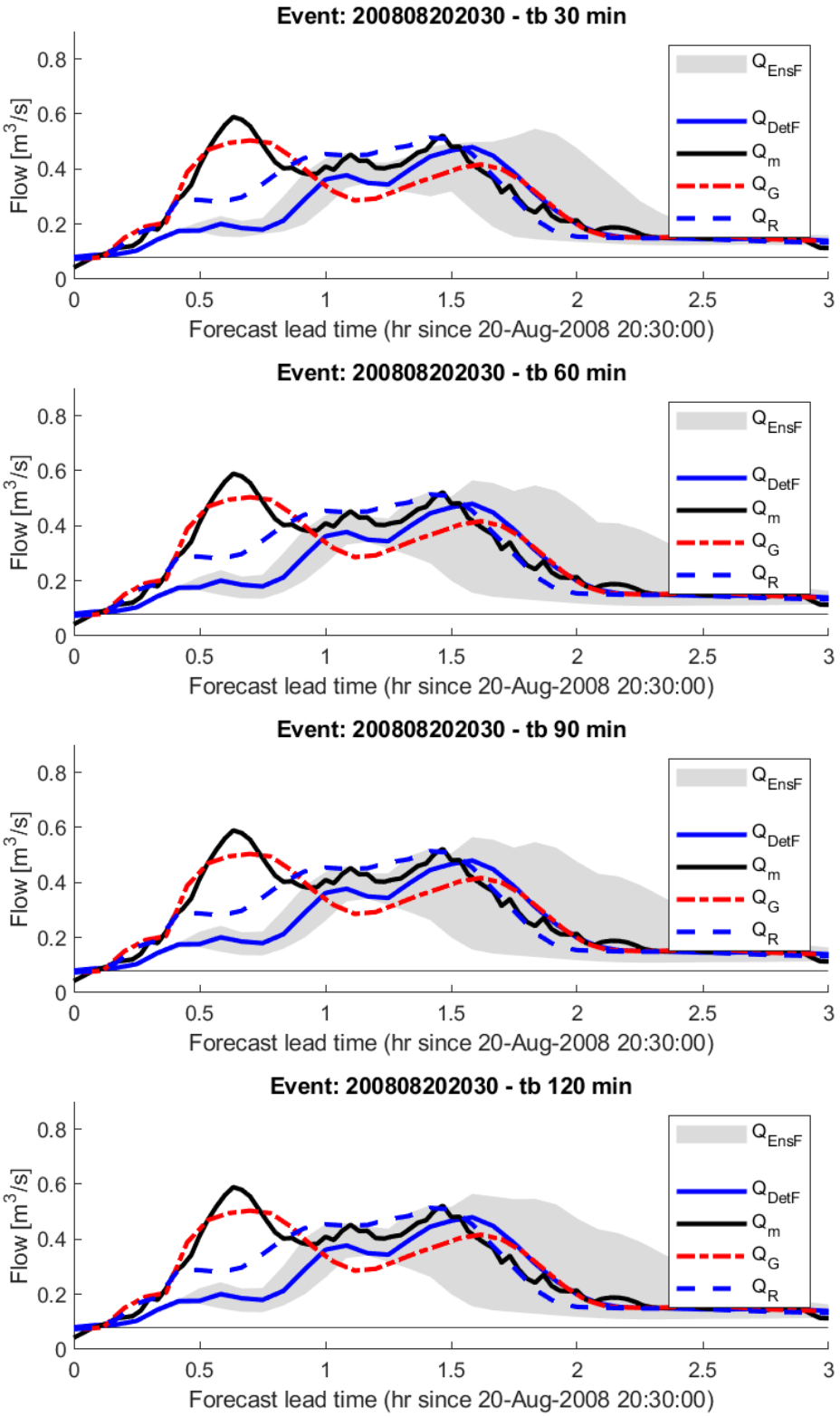


Figure 5-9. Flows for event on 20 August 2008 starting at 20:30. Q_{EnsF}, Q_{DetF}, Q_m, Q_G, and Q_R are, respectively, forecasted ensemble flows (with time-back window of 30-, 60, 90 and 120 min), forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow.

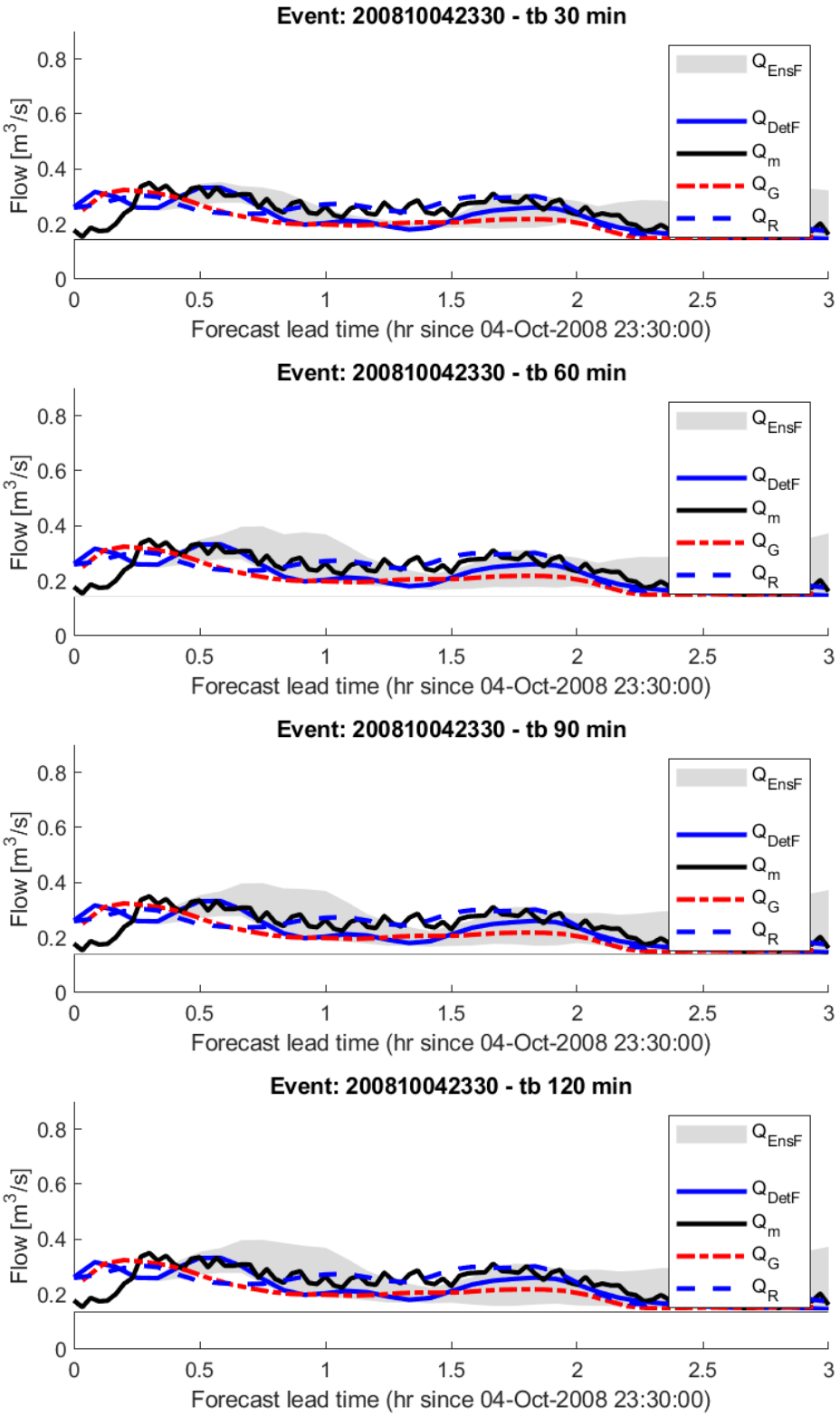


Figure 5-10. Flows for events on 04 October 2008 starting at 23:30. Q_{EnsF}, Q_{DetF}, Q_m, Q_G, and Q_R are, respectively, forecasted ensemble flows (with time-back window of 30-, 60, 90 and 120 min), forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow.

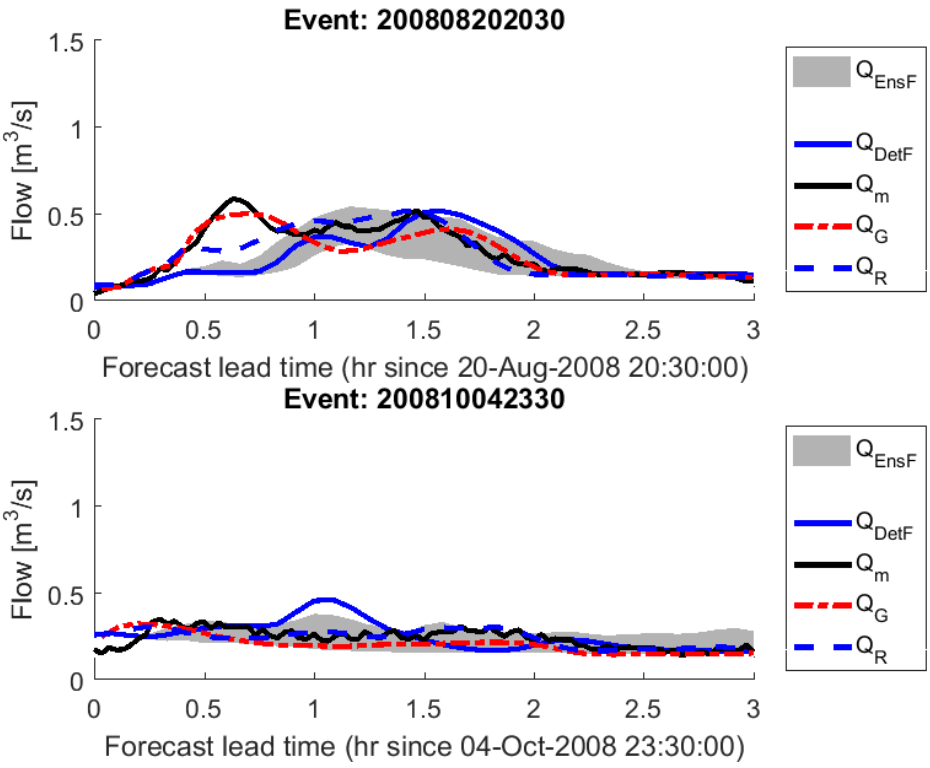


Figure 5-11. Flows for events on 20 August 2008 starting at 20:30 and 04 October 2008 starting at 23:30. Q_{EnsF}, Q_{DetF}, Q_m, Q_G and Q_R are, respectively, forecasted ensemble flows (STEPS ensembles), forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow.

Figure 5-12 and Figure 5-13 show the peak flows for the events on 20/08/2008 at 20:30 and 04/10/2008, respectively with a time-back window of 30-, 60, 90-, and 120 min compared with the peak measured flows, peak estimated flow using radar and rain gauge data and forecasted flow using deterministic forecast. In Figure 5-12, the peak flow was underestimated for all cases, with the rain gauge data being able to estimate it slightly better. For 1-2 hr after the forecast started, the measured flow peak was captured by some ensembles for all the cases. At longer lead times, as the measured flow was low, all forecasts were able to predict it accurately. The forecasts were not able to forecast the first flow peak on the event on 04/10/2008. However, the following peak that occurred soon after the first one produced flows of the same magnitude, leading to the peak flow for the first hour forecasted being predicted by the ensembles forecasts. For a time-back window of 30 min, most of the ensemble members underestimated the peak flow, but even though the forecast could still predict the peak flow. The peak flow coincided with the ensemble members ' median for forecasts with more ensembles (longer time-back window). For lead-time from 1-2 hours, the forecasts with a time-back window of at least 1 hour could predict the peak flows, being only slightly underestimated using the 30 min time-back window. For this period, the deterministic forecast underestimated the peak flow. For 2-3 hours ahead, as the flows were very low, all the probabilistic forecasts were able to predict it, although the deterministic forecast underestimated the peak flow.

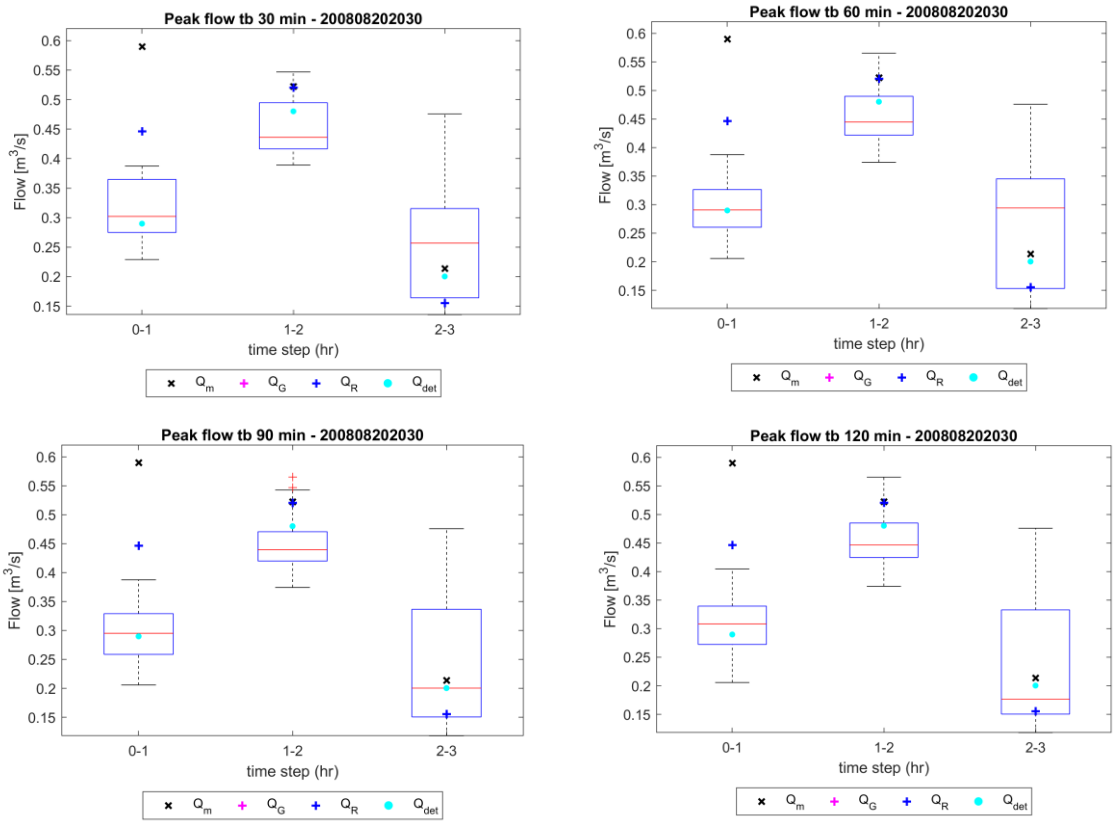


Figure 5-12. Peak flow boxplot for rainfall advection field forecasts with a time-back window of 30-, 60, 90 and 120 min, on 20 August 2008 starting at 20:30. Q_m, Q_G and Q_R, Q_{DetF} are, respectively, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow.

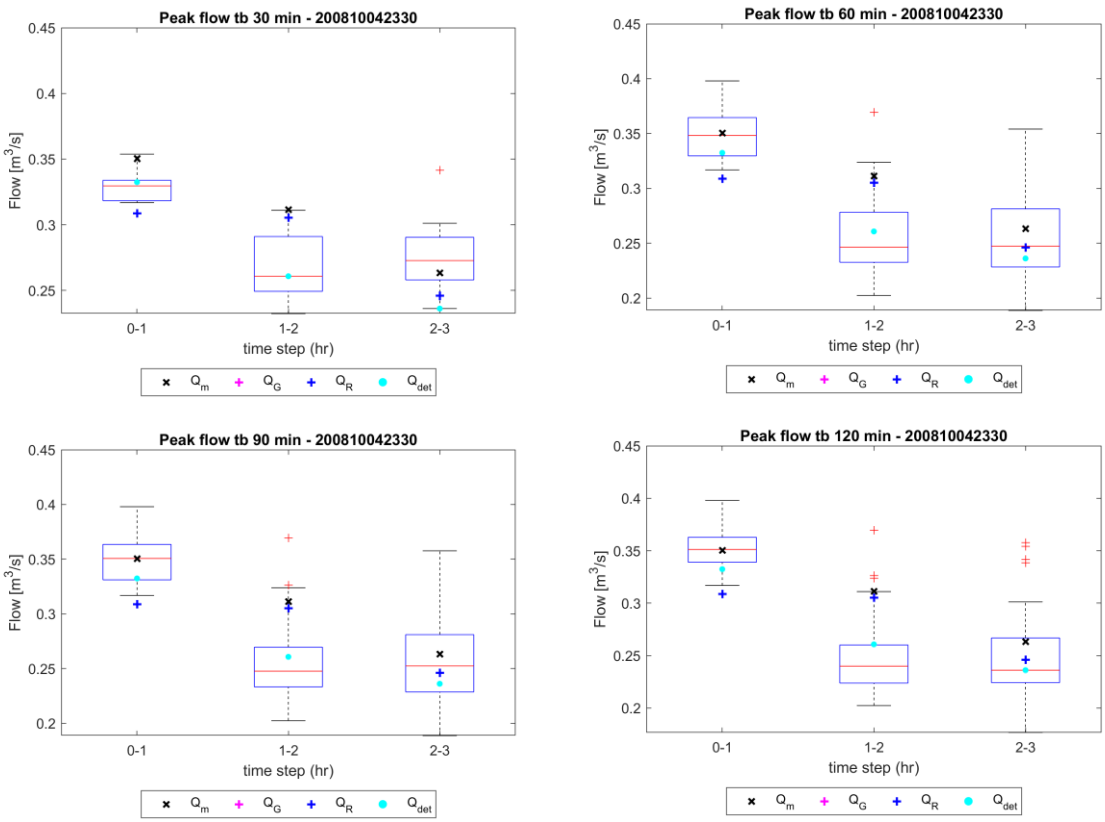


Figure 5-13. Peak flow boxplot for *rainfall advection field* forecasts with a time-back window of 30-, 60, 90 and 120 min, on 04 October 2008 starting at 23:30. Q_m, Q_G and Q_R, Q_{DetF} are, respectively, forecasted deterministic flow, measured flow, estimated gauge flow and estimated radar flow.

5.4 Conclusion

In this chapter, an ensemble generator was proposed using velocity fields calculated in different time steps before the start of the forecast. Each ensemble member was produced by advecting the most recent radar image using one of the computed advection fields. Rainfall advection fields calculated using 5-, 10- and 15 min apart from up to two hours before the start of the forecast. This study aimed to assess how these ensembles perform to produce forecasts and how much time-back should be looked into for the forecast. Options using 9, 27, 45 and 63 ensembles were used (time-back window of 30-, 60-, 90- and 120 minutes). Results were also used as an input in a hydrological model to simulate flows in an urban catchment, to assess how uncertainties propagate in the hydrological model.

ROC curve analysis showed that there is an increase in the forecasting ability with a higher ensemble number. The Area beneath the ROC curves seems to take advantage of a higher number of ensembles with longer lead times and higher thresholds. According to Foresti *et al.*, 2013, nowcasts are subjected to mainly radar errors up to 1 hr lead time. After this period, the primary sources of uncertainties are

due to the temporal evolution of rainfall advection fields and the growth and decay of rainfall. So, having a higher difference between the rainfall advection field forecasts at longer lead-times agree with the literature as the ensemble generator proposed is to account for uncertainties related to the temporal evolution of the rainfall advection fields. Even with a higher number of ensemble members, the forecast still loses predictability as the forecast advances in time. Comparing the results of the proposed method with some of the results from Chapter 4, it is possible to realise that the ensemble rainfall forecasts produce ROC curves with larger areas beneath them than both STEPS and RE ensembles for lead times longer than 1 hr. This means that these ensemble forecasts produce better performance and account for uncertainties due to the temporal evolution of the rainfall advection field.

The results were also used as an input for a hydrological model in Ilkley. Flow forecasts were only able to reproduce flow peaks that could also be estimated using radar data. For rainfall advection field forecasts with a time-back window longer than 60 min, the flow forecasting capability still improves, although forecasts produced using time windows of 90 min and 120 min produce similar results. Comparing the forecasted flows with flows forecasted using STEPS ensembles (Chapter 4), it was concluded that using the proposed ensemble generator does not provide higher accuracy than using the STEPS ensembles. This might be because the urban area taken into account is very small, and having rainfall displaced just a couple of pixels in a slightly different direction can give completely different results. For this reason, it is important to assess the progression of uncertainties of these ensembles in bigger catchments, both urban and rural, in order to obtain more information about the efficiency of the forecast in hydrological applications.

Due to the computer power and time necessary to produce forecasts with a large number of ensembles (it takes more than 60% more time to run the 90 min forecasts and over double of the time for the 120 min one when compared to the 60 min time window), and considering that having 90 min or 120 time-back windows to produce forecasts do not cause a significant improvement in the forecasts. Using a time window of 60 min should be enough for most applications.

Chapter 6. Radar and rain gauge merging

6.1.1 Introduction

Chapter 1 highlighted the importance of how high accuracy and temporal and spatial resolutions are important in rainfall nowcasting and, consequently, in flood predictions. It also described how radars measurements are subjected to different error sources that affect the accuracy of QPE (Quantitative Precipitation estimation). On the other hand, rain gauges are better in estimating precipitation at ground levels at specific point locations, and even with uncertainties inherited to rain gauges. Combining these independent measurements from radar and rain gauges allow the generation of a rainfall merged product that has high spatial resolution representing the spatial distribution of precipitation and better accuracy than the original measurements separately (Krajewski, 1987; Goudenhoofdt and Delobbe, 2009; Berndt, Rabiei and Haberlandt, 2014; Jewell and Gaussiat, 2015).

Chapter 4 considered the uncertainties related to radar QPE by producing probabilistic estimations based on the bias between weather radar and rain gauge measurements using historical time series. Although the radar rainfall ensembles were able to reproduce these uncertainties, they are not designed to reduce them. The use of radar-gauge merging techniques comes into place to reduce the uncertainties in radar QPE and produce better rainfall forecasts.

The rainfall accumulation period plays an important role in the performance of the radar-gauge rainfall merging methods. For hydrological applications, it is important to have both accuracy and high temporal resolution QPE due to the precipitation field's fast-changing nature. However, it is usual to have merged products at hourly or daily accumulations, at least due to the disparities in representativity between radar and rain gauge estimations. At shorter accumulation periods, the radar rainfall accumulation is the result of fewer scans. Rain gauge measurements also suffer from shorter accumulation periods, especially when TBR (Tipping bucket rain gauges) are used. When the bucket takes longer to tilt for low rainfall intensity events, there will be time-steps with no rainfall followed by a period of overestimation of precipitation. Issues related to the difference in spatial resolution between radar and rain gauge measurements, radars can detect very localized showers that nearby rain gauges can miss. However, for longer accumulation periods, there are higher chances for it to be detected by the rain gauges. Jewell and Gaussiat (2015) also compared the use of 15 and 60 min accumulation periods in radar-gauge merging. They concluded that although the performance is considerably better at hourly accumulations, merged products with 15 min accumulation are still better than using rain gauges or radars data alone. Berndt, Rabiei and Haberlandt (2014) compared the performance of a wide range of temporal accumulation periods (from 10 min to 6h accumulation) and

network densities for OK, KED and conditional merging. Their results showed that the correlation between radar and rain gauge increases and RMSE decreases as with higher accumulation periods, and even at high temporal resolutions, there is an improvement in the performance of merged products when compared to OK.

Most of the studies carried out to compare the vast number of radar-gauge merging techniques concluded that the KED rainfall merging approach is a robust method and has shown to produce reliable results for both stratiform and convective events (Goudenhoofdt and Delobbe, 2009; Li and Heap, 2011; Jewell and Gaussiat, 2015; Nanding, Rico-Ramirez and Han, 2015). Ochoa-Rodriguez *et al.* (2019) presented a summary of the outcome of different merging techniques comparison, and KED has an overall best performance than the other methods considered. Although the benefit of KED increases with rain gauge network density, sensitivity analysis showed that KED, along with other geostatistical methods, are is less sensitive to the median separation of rain gauges (Jewell and Gaussiat, 2015). An advantage of KED when compared to other methods with similar performance is the lower computational cost needed. For these reasons, KED was selected to perform radar-gauge merging in this study.

For radar-based rainfall forecasting applications (nowcasting), short temporal resolutions are required (e.g. 15min or lower). Even with a vast number of studies about merging techniques, there is still much work to be done on how temporal resolution affects hydrological forecasts. For nowcasting and flow prediction applications, the data used must have a high temporal resolution. Most of the studies carried on radar-gauge merged products use an hourly resolution or higher. Jewell and Gaussiat (2015) and Berndt, Rabiei and Haberlandt (2014) explored the influence of the temporal resolution on the merged product, and both studies concluded that a higher temporal resolution leads to a loss in the accuracy of rainfall estimations. Shehu and Haberlandt (2021) compared the performance of five merging methods at high temporal resolution. They concluded that conditional merging outperforms the other methods, as its products maintain high temporal and spatial correlation. It was also concluded that, regarding the RMSE, KED performed better than conditional merging. Using sub-hourly accumulation periods in KED has been subjected only to a limited number of research papers, but it is an essential step to allow the application of KED in nowcasting. In this thesis, the KED product will be used to generate rainfall forecasts to drive the nowcasting model (Chapter 7), and the error propagation will be assessed for the different temporal resolutions. With that in mind, in this chapter, a comparison between OK, KED and radar-only QPE products are carried out for different temporal resolutions suitable for hydrological applications. This will allow a comparison of the different accumulation periods effect in the nowcasting model when KED is used as an input. The use of KED in nowcasting has not

been widely studied, and there was no further research on the effect of temporal accumulation in it. In this chapter and Chapter 7, this research gap will be studied. Section 6.2 describes the methodology, followed by the results presented in section 6.3. Section 6.4 will present a review of the findings.

6.2 Methodology

Radar and rain gauge data from 2008 were used to produced radar-rain gauge merged rainfall estimates. For merging radar with rain gauges, KED was chosen. OK is not a merging method. However, it is used in this chapter as a reference only. The results were compared against OK and against rainfall data. In order to assess the accuracy of the merged product, three different temporal resolutions: 15 min, 30 min and 60 min were explored. For this, rain gauge measurements and radar estimations were accumulated accordingly to each temporal resolution. Note that most of the events were observed for 24h, starting from 00:00 until 23:59. The dates, average rainfall and storm types are presented in Table 6-1. The storm classification was carried in the same way as described in Chapter 4, where a storm is classified as convective if it has more than 3% of convective pixels throughout at least 3 hours.

Table 6-1. Event dates, maximum average rainfall and storm types (S—stratiform and C—convective).

Event Date	Max Average Rainfall (mm/h)	Storm Type
14/01/2008	4.24	S
15/01/2008	13.41	S
16/01/2008	5.87	S
17/01/2008	7.65	S
18/01/2008	6.86	S
29/04/2008	5.11	S
30/04/2008	9.88	S
07/07/2008	6.78	S
31/07/2008	8.83	S
01/08/2008	7.35	C
02/09/2008	8.12	C
04/09/2008	3.70	S
05/09/2008	22.84	C
06/09/2008	12.09	C
04/10/2008	12.57	S
14/10/2008	10.53	S
08/11/2008	7.43	S
09/11/2008	11.92	S
12/12/2008	4.57	S
13/12/2008	17.40	S

The OK and KED methods are described in sections 6.2.1 and 6.2.2. The variogram calculation is described in section 6.2.3.

6.2.1 Ordinary kriging (OK)

Krige (1952) developed a geostatistical interpolation method using distance weighted average, taking into account distance from adjacent rain gauges and the distance between them. Rain gauges within a fixed distance of each other are assumed to have some spatial correlation, while far away gauges are assumed to be independent (Krige, 1951; Jewell and Gaussiat, 2015). As stated previously, OK is only an interpolation method and does not merge radar and rainfall data. However, it is generally used as a benchmark to compare merged products against (Jewell and Gaussiat, 2015; Nanding, Rico-Ramirez and Han, 2015).

In OK, the value at each unknown location is calculated as the weighted average of the surrounding gauges and given by a linear combination of the available measurements, that is (Jewell and Gaussiat, 2015; Nanding, Rico-Ramirez and Han, 2015; Cecinati, 2017):

$$Z(x_0) = \sum_{\alpha=1}^{n} w_\alpha\,(x_0).Z(x_\alpha) \qquad (6\text{-}1)$$

Where $Z(x_0)$ is the estimated rainfall at a location x_0, $Z(x_\alpha)$ are the measured rainfall at the rain gauge location x_α , n is the number of available rain gauges and w_α are the kriging weights

With the condition that the estimation is unbiased and its variance must be minimized, the resulting OK equations can be written as:

$$\begin{cases} \sum_{\alpha=1}^{n} w_\alpha = 1 \\ \sum_{\beta=1}^{n} w_\alpha\,\gamma(x_\alpha - x_\beta) + \mu_1 = \gamma(x_\alpha - x_0) \quad \alpha = 1, \dots, n \end{cases} \qquad (6\text{-}2)$$

Where x_α and x_β are generic rain gauge locations, γ is the parametric variogram generated from the rainfall readings and μ_1 represents the Lagrange multiplier (Cressie, 1990). The variogram represents the spatial correlation of the measurements, and it varies for each time step. However, it is important to have sufficient measurements available to compute a reliable variogram, and therefore several time steps might be required. It is also important to highlight that the variogram is isotropic (parametric) and does not change with direction. However, Velasco-Forero et al. (2009) considered anisotropy due to the rainfall field's spatial variation and showed that a 2D non-parametric correlogram could improve the results. When calculating the variogram, it is assumed that the spatial correlation is a function only

of the distance between the rain gauges. It does not take into account the location of the rain gauges and their impact on the spatial correlation.

The equation system (6-2) can also be written as a matrix as follow:

$$\begin{bmatrix} \gamma(x_1 - x_1) & \dots & \gamma(x_1 - x_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(x_n - x_1) & \cdots & \gamma(x_n - x_n) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma(x_1 - x_0) \\ \vdots \\ \gamma(x_n - x_0) \\ 1 \end{bmatrix} \quad (6\text{-}3)$$

Solving the matrix allows the weighting factors to be computed and then used in equation (6-1) to estimate the rainfall rate at the unknown location x_0.

6.2.2 Kriging with external drift (KED)

Kriging with external drift generates a merged product using both radar and rain gauge data by taking into account additional constraints in the computation of the weights. It differs from OK by assuming that the global mean and the covariance of rainfall fields are non-stationary in space and vary across the field (Jewell and Gaussiat, 2015). The local mean field is extrapolated from radar data, and a residual noise component from rain gauge data is incorporated by interpolation. KED is determined by three constraints:

$$\begin{cases} \sum_{\alpha=1}^{n} w_\alpha = 1 \\ \sum_{\alpha=1}^{n} w_\alpha Z_R(x_\alpha) = Z_R(x_0) \\ \sum_{\beta=1}^{n} w_\alpha \gamma(x_\alpha - x_\beta) + \mu_1 + \mu_2 . Z_R(x_\alpha) = \gamma(x_\alpha - x_0) \quad \alpha = 1, \dots, n \end{cases} \quad (6\text{-}4)$$

Where Z_R is the radar rainfall estimation at location x and μ_2 is the second Lagrange multiplier required. Using the local mean field instead of the global one can lead to an inaccurate representation of rainfall in regions with sparse rain gauge coverage (Jewell and Gaussiat, 2015).

The equation system described in equation (6-4) can also be written as a matrix:

$$\begin{bmatrix} \gamma(x_1 - x_1) & \dots & \gamma(x_1 - x_n) & 1 & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \gamma(x_n - x_1) & \cdots & \gamma(x_n - x_n) & 1 & 1 \\ 1 & \cdots & 1 & 0 & 0 \\ Z_R(x_1) & \cdots & Z_R(x_n) & 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \gamma(x_1 - x_0) \\ \vdots \\ \gamma(x_n - x_0) \\ 1 \\ Z_R(x_0) \end{bmatrix} \quad (6\text{-}5)$$

The calculated weights are used in equation (6-1) to calculate the merged rainfall at the unknown locations.

Both OK and KED skills depend on the number of available rain gauges and their measurements' accuracy. When a low number of gauges are available, or for short accumulation periods (that can limit the number of rain gauges in the raining region), calculating a variogram based on rain gauges measurements can be challenging (Cecinati, 2017)

6.2.3 Variogram estimation

The rain gauge network spatial correlation can be represented using the variogram. As discussed earlier, a dense network is desirable in order to produce high-quality spatial interpolated products. Having a variogram that can adequately represent the spatial correlation of the measurements is therefore needed (Schiemann et al., 2011). The variogram for OK and KED in radar-gauge rainfall applications is usually calculated based on rain gauges measurements. It provides essential information, such as the nugget, range and sill, which are the three parameters of the variogram (see Figure 6-1) (Burrough, McDonnell and Lloyd, 2015). The nugget is a discontinuity at the origin of the variogram and is related to the variation of spatial sampling errors or measurement errors. The sill is the region where the variogram flats out and represents the variance. The range is the distance where the decorrelation occurs, meaning that if the spacing between two rain gauges is bigger than the range, they are not correlated. (Berne *et al.*, 2004; Burrough, McDonnell and Lloyd, 2015). Figure 6-1 shows a variogram where the range, nugget and sill are all marked in the figure.

The variogram is usually experimentally using rainfall data, usually from rain gauges. The variogram uses the separation distance, d, to calculate the variance, according to (6-6). Empirically calculating the variogram using rain gauges has already limitations depending on the density of the network (Cecinati, 2017). In this thesis, as the KED results will be used for forecasting applications, the rainfall rate for the following time-steps is unknown. The results for variograms calculated with both rain gauge and radar data sets will be compared in the next session. In cases where a variogram could not be calculated due to the lack of enough data points with rainfall, the previous variogram was used.

$$\gamma(d) = \frac{1}{2} E\left\{\left(\delta(x) - \delta(x+d)\right)^2\right\} \qquad (6\text{-}6)$$

Where δ is the residual error.

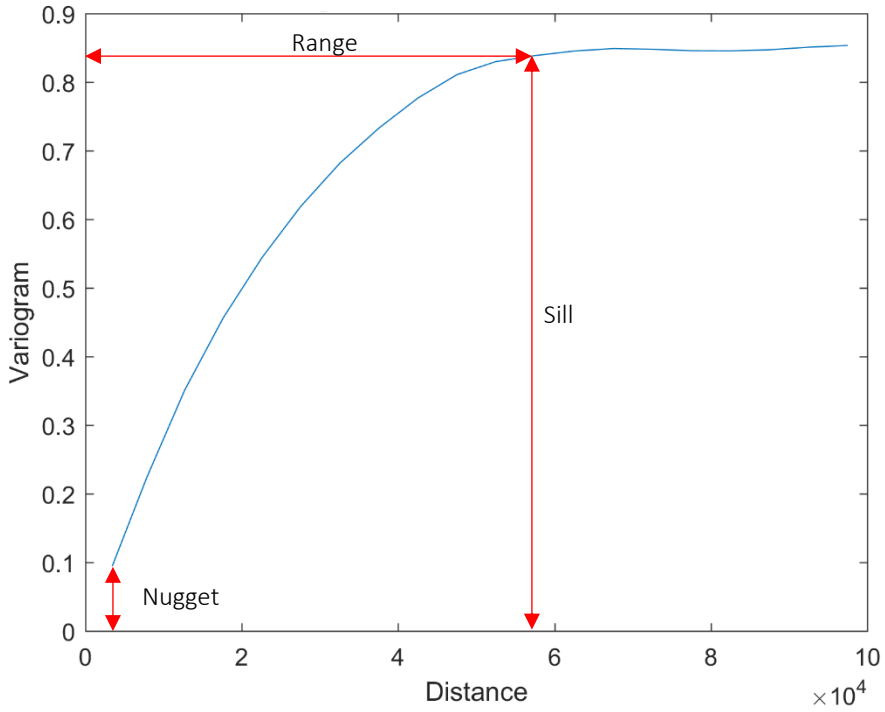


Figure 6-1. Variogram example with representation of range, sill and nugget.

6.3 Results

Using rain gauge data to produce the variograms was tested, but it was impossible to generate reliable variograms once there are not enough data points. Being so, the variograms were derived from the rainfall data of each time-steps. Due to the high spatial resolution, the number of data points available is much higher. Even being less accurate than rain gauge data, radar data is still accurate enough to produce variograms, and they have the advantage of being able to reproduce the spatial variability of the rainfall fields more accurately than rain gauges. The results were produced for 15-, 30- and 60-min accumulation periods. Figure 6-2, Figure 6-4, Figure 6-6 and Figure 6-8 show the variograms calculated for different events and time steps (15/01/2008 at 19:00, 30/04/2008 at 10:00, 10/04/2008 at 16:00 and 14/10/2008 at 18:00) to illustrate the different variograms. Figure 6-3, Figure 6-5, Figure 6-7 and Figure 6-9 shows the accumulated rainfall using OK interpolation and KED estimations for different events and for rain rates accumulated at the time scales of 15-, 30-, and 60 min. For each event, the variograms were calculated for each time step. Figure 6-2 shows that the variograms computed using radar rainfall are much smoother than the variograms computed using rain gauge data. This is due to the fact that for a single time step, there are more radar pixels available than rain gauge measurements that can help to compute a more reliable variogram. Therefore, for real-time applications, it is better to compute the variogram using radar rainfall rather than rain gauge measurements unless there is a

very dense rain gauge network or the time scale used to compute the variogram includes several time steps. This figure also shows that the variograms computed at 15min time scales show a higher variance (high sill) compared to the variograms computed at 60min time scales. This is expected, since the variability of precipitation is higher at shorter time scales. This result is consistent when looking at different rainfall events and time steps. The implication of this is that the interpolated OK/KED rainfall field will show a higher variance for shorter time scales. In Figure 6-3, the radar data present light rainfall for a much larger area than the one measured by rain gauges. As a result, the merged product also has a smaller area with rainfall, as in places where rain gauges did not detect any rainfall, the precipitation intensities were smoothed down. At 60 min accumulation, rain gauges were able to measure precipitation over a larger area, and as a result, rainfall was estimated in more locations of the domain.

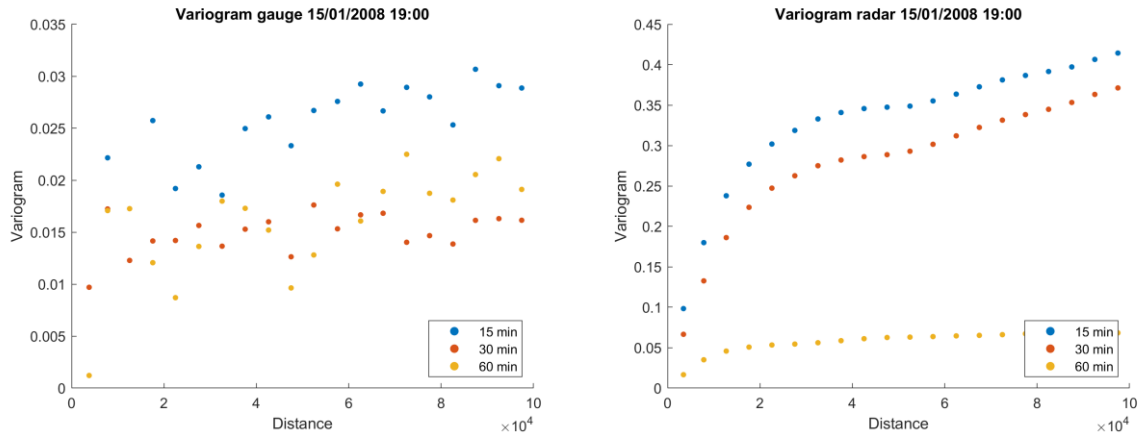


Figure 6-2. Variograms calculated using rain gauge (left) data and radar (right) data for event on 15/01/2008 at 19:00 for accumulation periods of 15 min, 30 min and 60 min.

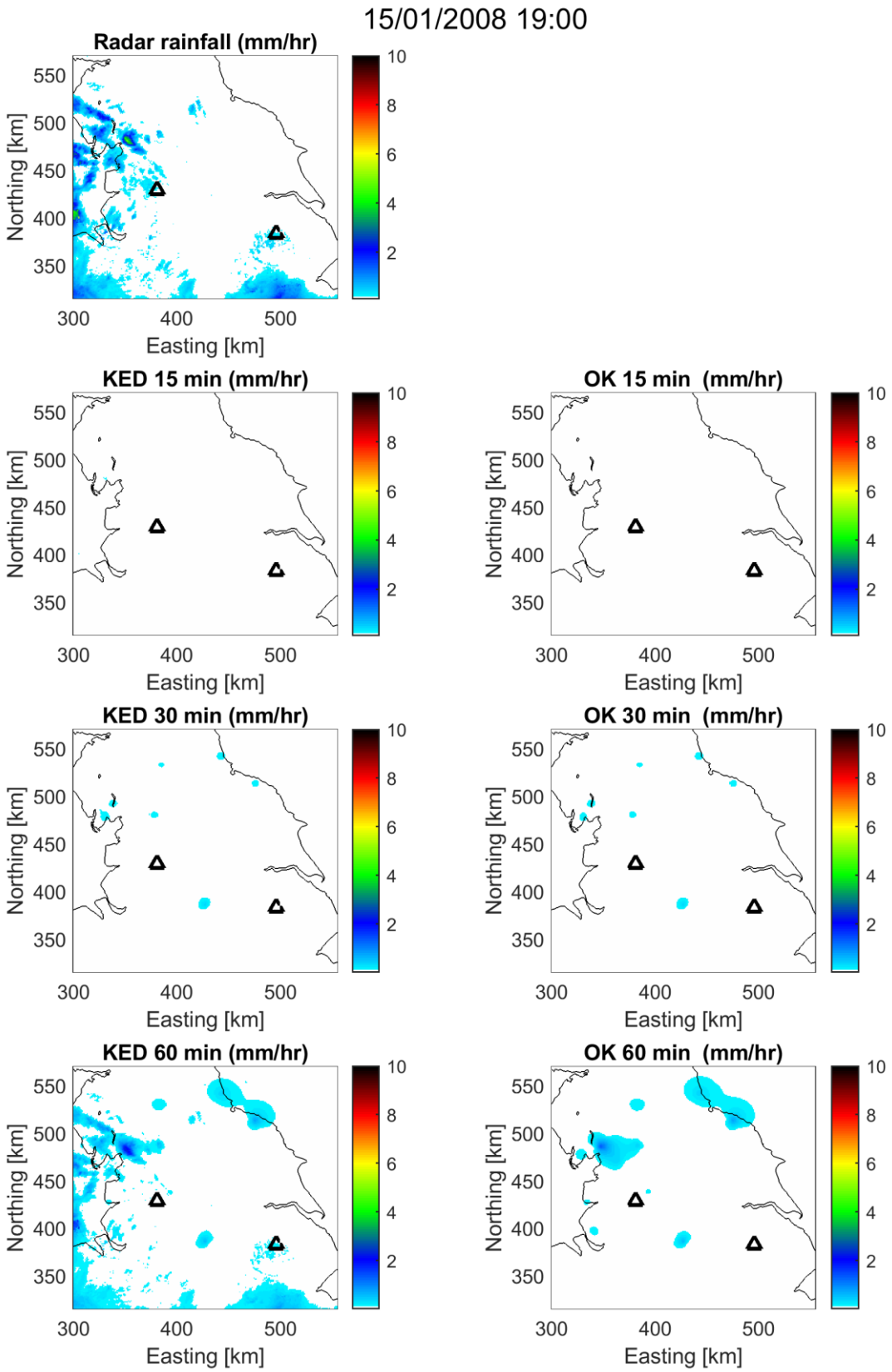


Figure 6-3. Rainfall for hourly accumulated radar data in mm/hr, KED and OK on 15/01/2008 at 19:00 and 15-, 30- and 60 min accumulation.

In Figure 6-4, it was only possible to calculate the variogram for 60 min accumulation when rain gauge data is used. At longer accumulation periods, more data points are likely to be available. This is also due to the fact that most of the precipitation for this particular time steps occurred outside the rain gauge network domain. In cases where no variogram is computed, the previous variogram is used. However, when using radar data, there are more rainfall points to be considered, and the variogram could be computed, even for shorter accumulation periods. Therefore, calculating the variogram with radar data is important when single time-steps are used. The variogram computed using radar data also showed that short accumulation periods produce a high spatial variance. Considering the radar image and the OK interpolation in Figure 6-5, it can be noted that the light rainfall occurs in small locations over the domain rather than over a large area. Thus, it is more difficult for rain gauges to accurately represent the amount of rainfall in the ground as they are point measurements. Over a 60 min accumulation period, the probability of precipitation falling at the rain gauge location is higher, and as a result, there are more points available to calculate the variogram. The variograms calculated in Figure 6-6 using the radar data are almost straight lines for 30- and 60 min, meaning that there is still a spatial correlation between the data points even at longer distances. This behaviour was not very common, but it happened in other events and is driven by the large precipitation cell shown on the top left corner of the domain, where most of the precipitation occurs in this time step. At 60 min accumulation, the variograms computed with radar or rain gauges data are very similar. This is due to the fact that most of the precipitation occurs over the rain gauge network, and therefore both variograms show similar results. The radar variogram for 30 min and 60 min are very similar to each other, meaning that the spatial correlation, in this case, does not change with more extended accumulation periods for this particular time step.

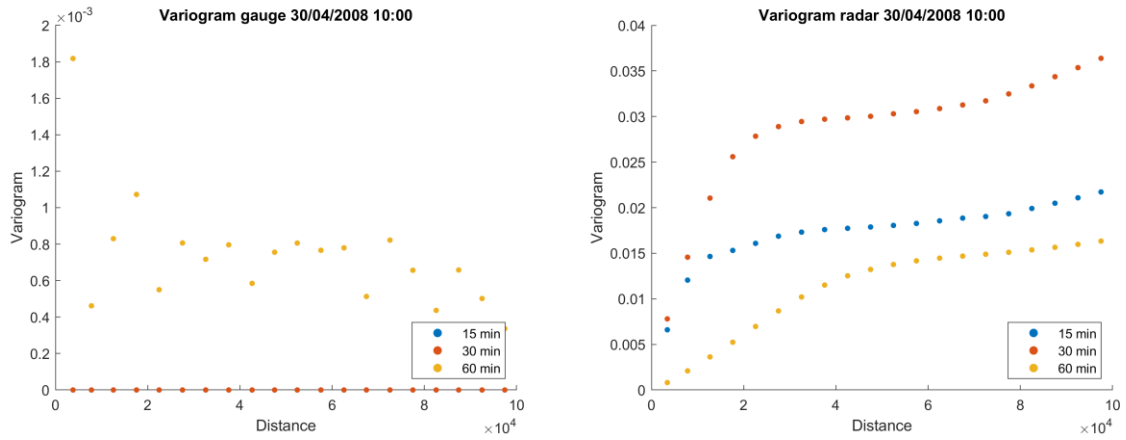


Figure 6-4. Variograms calculated using rain gauge (left) data and radar (right) data for event on 30/04/2008 at 10:00 for accumulation periods of 15 min, 30 min and 60 min.

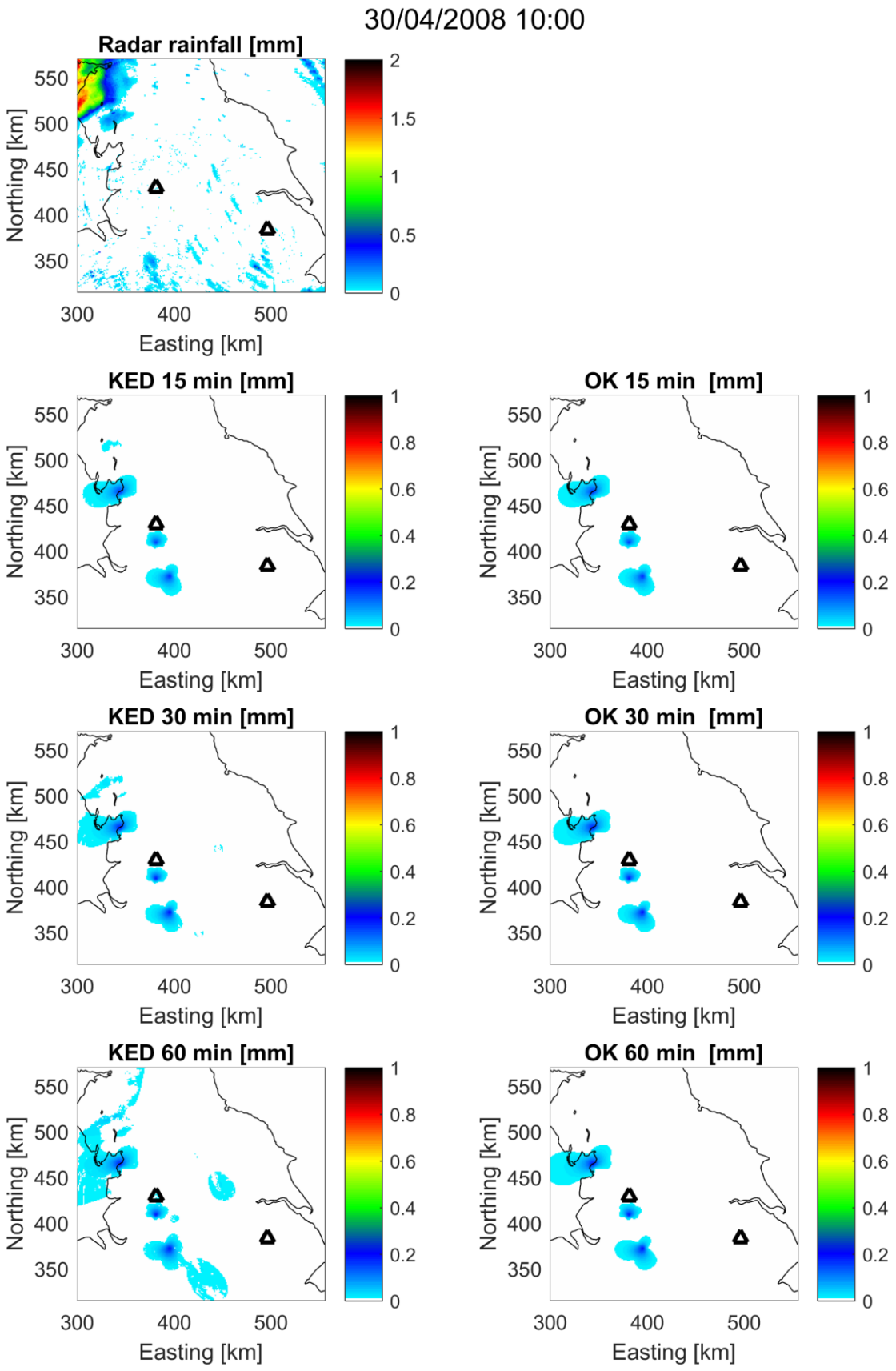


Figure 6-5. Rainfall for hourly accumulated radar data in mm/hr, KED and OK on 30/04/2008 at 10:00 and 15-, 30- and 60 min accumulation.

In Figure 6-7, the radar seems to overestimate rainfall at all accumulation periods when compared to the rain gauge. As a result, the KED predictions adjust the radar rainfall field to agree with the rainfall intensities measured by the rain gauge network. However, the overall spatial distribution of precipitation observed by radar is preserved in the KED estimates. The KED rainfall field is affected by both radar and rain gauge observations, and if they are inaccurate or show artefacts, the KED estimations will reproduce this. For instance, the KED estimations reproduce some of the radar rainfall artefacts (e.g. the beam blocking shown on the radar scan towards the north and north-west from the Hameldon Hill radar location), especially at longer accumulation periods.

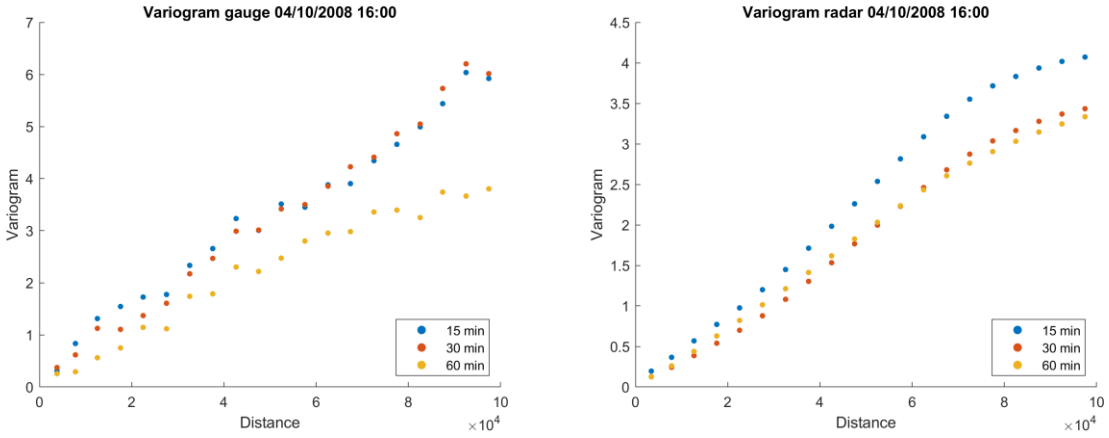


Figure 6-6. Variograms calculated using rain gauge data and radar data for event on 04/10/2008 at 16:00 for accumulation periods of 15 min, 30 min and 60 min.

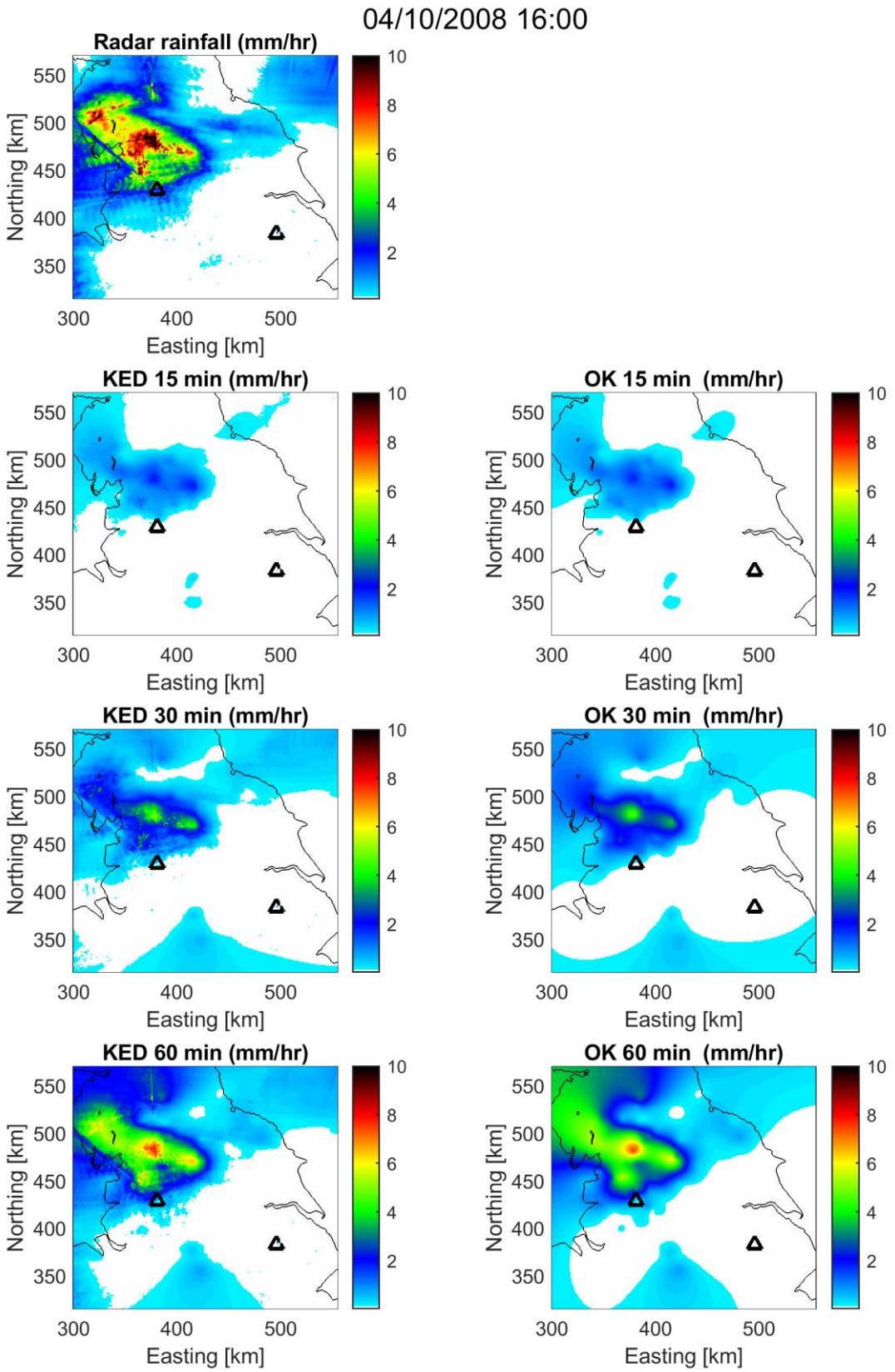


Figure 6-7. Rainfall for hourly accumulated radar data in mm/hr, KED and OK on 04/10/2008 at 16:00 and 15-, 30- and 60 min accumulation.

The variograms from Figure 6-8 also show that the variance increases with the temporal resolution. Although the radar variogram for 60 min accumulation was smoother than for the gauge data, at 60 min accumulation, both variograms have a similar shape and variance values. Looking into Figure 6-9, as the rainfall is spread over a large area of the domain, more rain gauge measurements are available for the variogram calculations. In this event, the rain gauge measurements show a higher intensity rainfall than estimated by radar in the region located in Northwest England. The radar overestimates the rainfall in other locations when compared to the rain gauge. As KED merging is highly based on the rain gauges measurements, the KED intensities are closer to the measured by rain gauges.

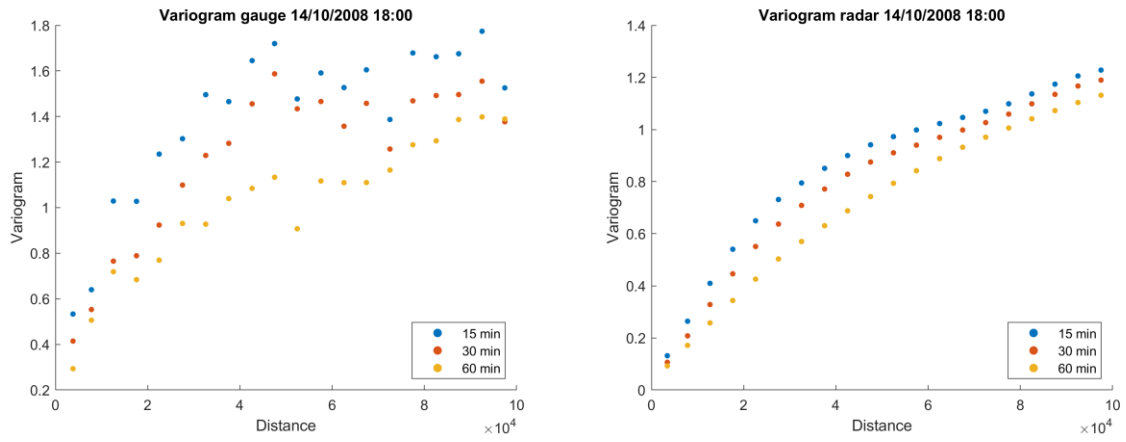


Figure 6-8. Variograms calculated using rain gauge data and radar data for event on 14/10/2008 at 18:00 for accumulation periods of 15 min, 30 min and 60 min.

In order to assess the results, cross-validation was used, and performance indicators such as RMSE and correlation were calculated for each event. Figure 6-10 and

Figure 6-11 show the results for KED cross-validation for the events on 15/01/2008, 30/04/2008, 04/10/2008 and 14/10/2008. Ideally, the correlation should be close to 1, and the RMSE should be close to zero. The cross-validation was carried out using the leave-one-out technique, and the figures show the estimated rainfall at the rain gauge location left out, compared with the rainfall measured on that location. The results show that with shorter temporal resolutions (e.g. 15min), the discrepancy between the two values is considerably higher than when using longer accumulation intervals (e.g. hourly), and therefore the RMSE error decreases (and correlation increases) with hourly accumulation intervals. These results were expected once there are fewer data points available to calculate the variograms, resulting in more uncertainties in the spatial correlation of the KED product. All results agree between them and show a pattern of having the data closer to the control line as the accumulation interval increases.

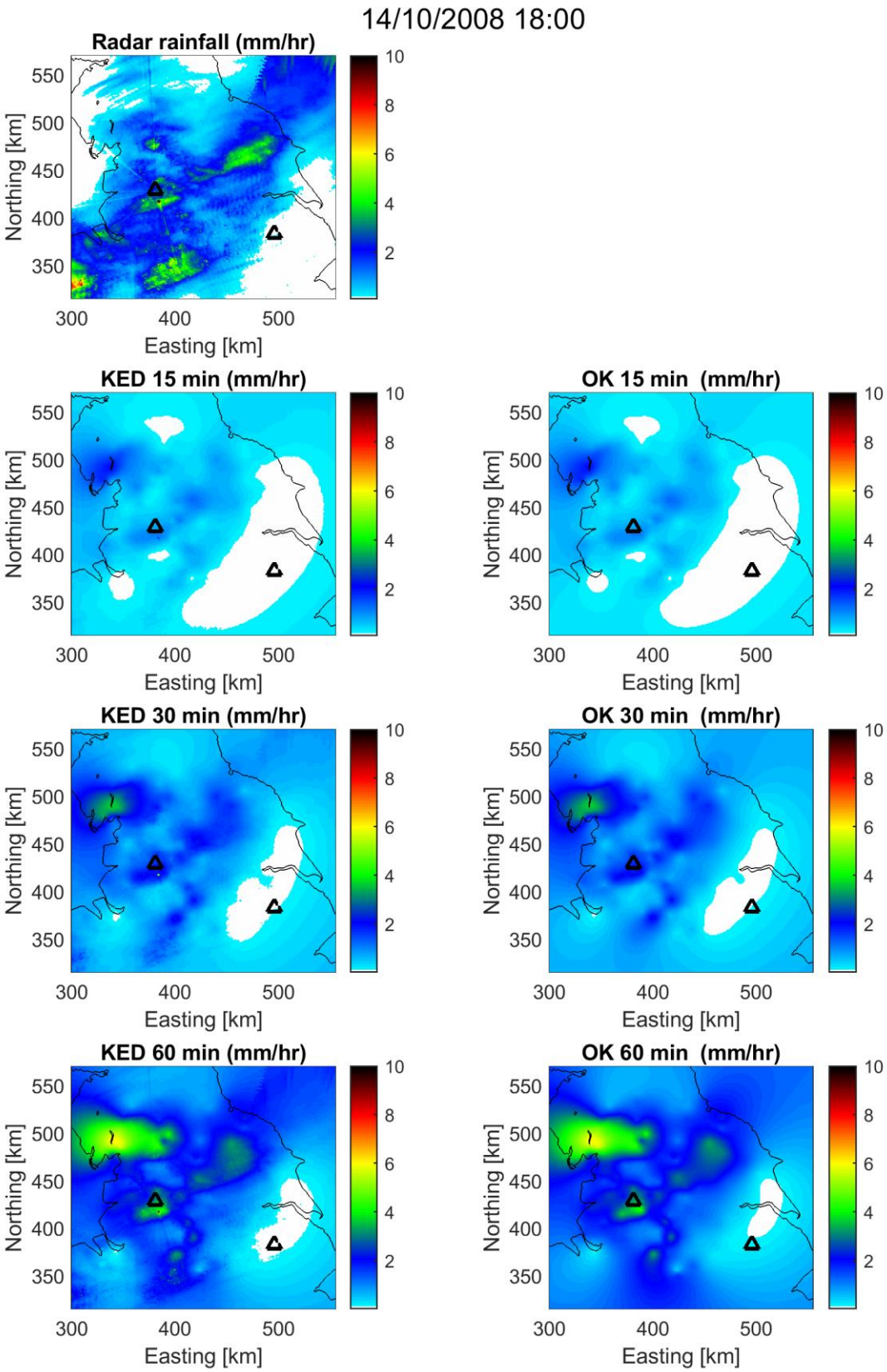


Figure 6-9. Rainfall for hourly accumulated radar data in mm/hr, KED and OK on 14/10/2008 at 18:00 and 15-, 30- and 60 min accumulation.

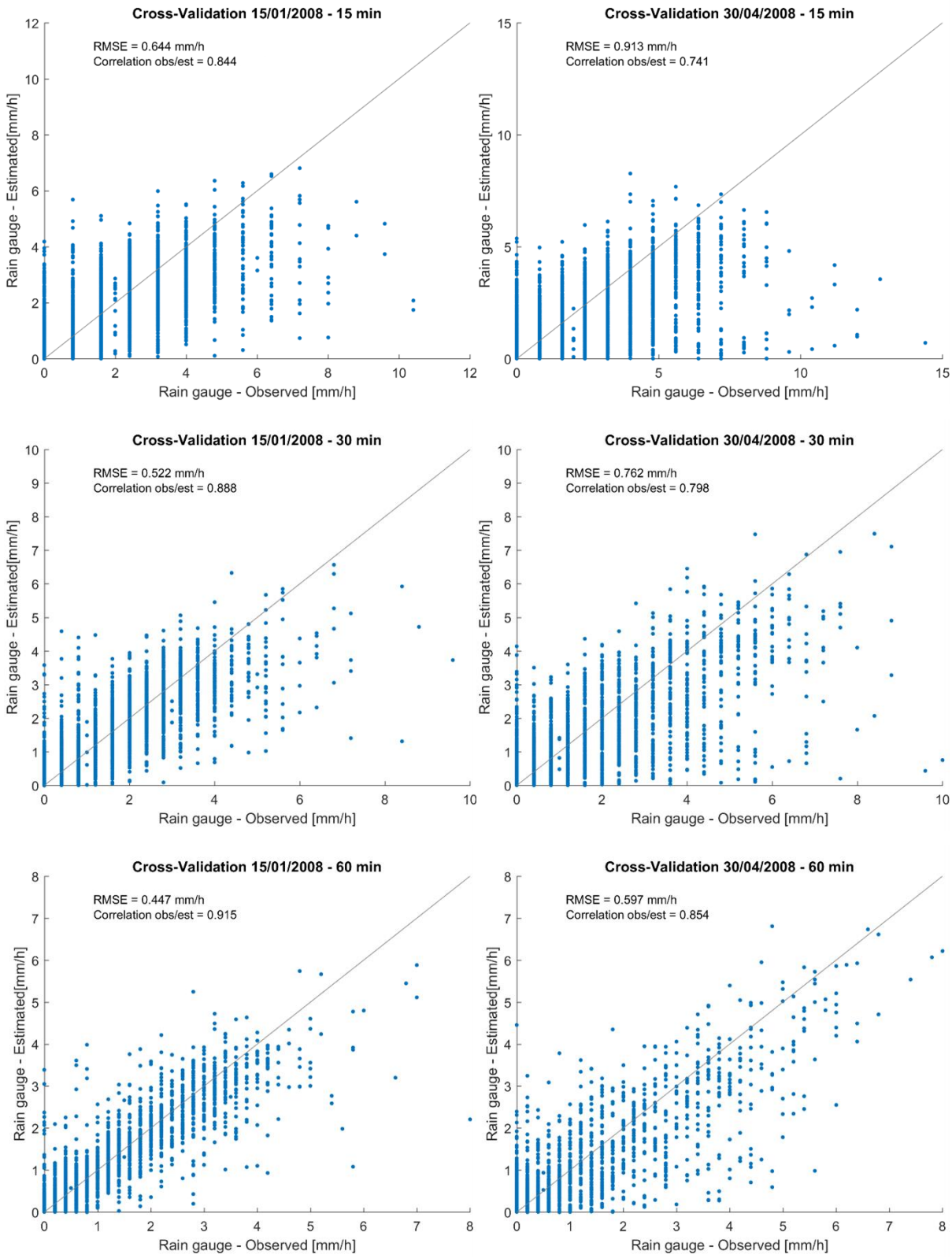


Figure 6-10. KED cross-validation for events on 15/01/2008 and 30/04/2008 with 15-, 30- and 60-min accumulation.

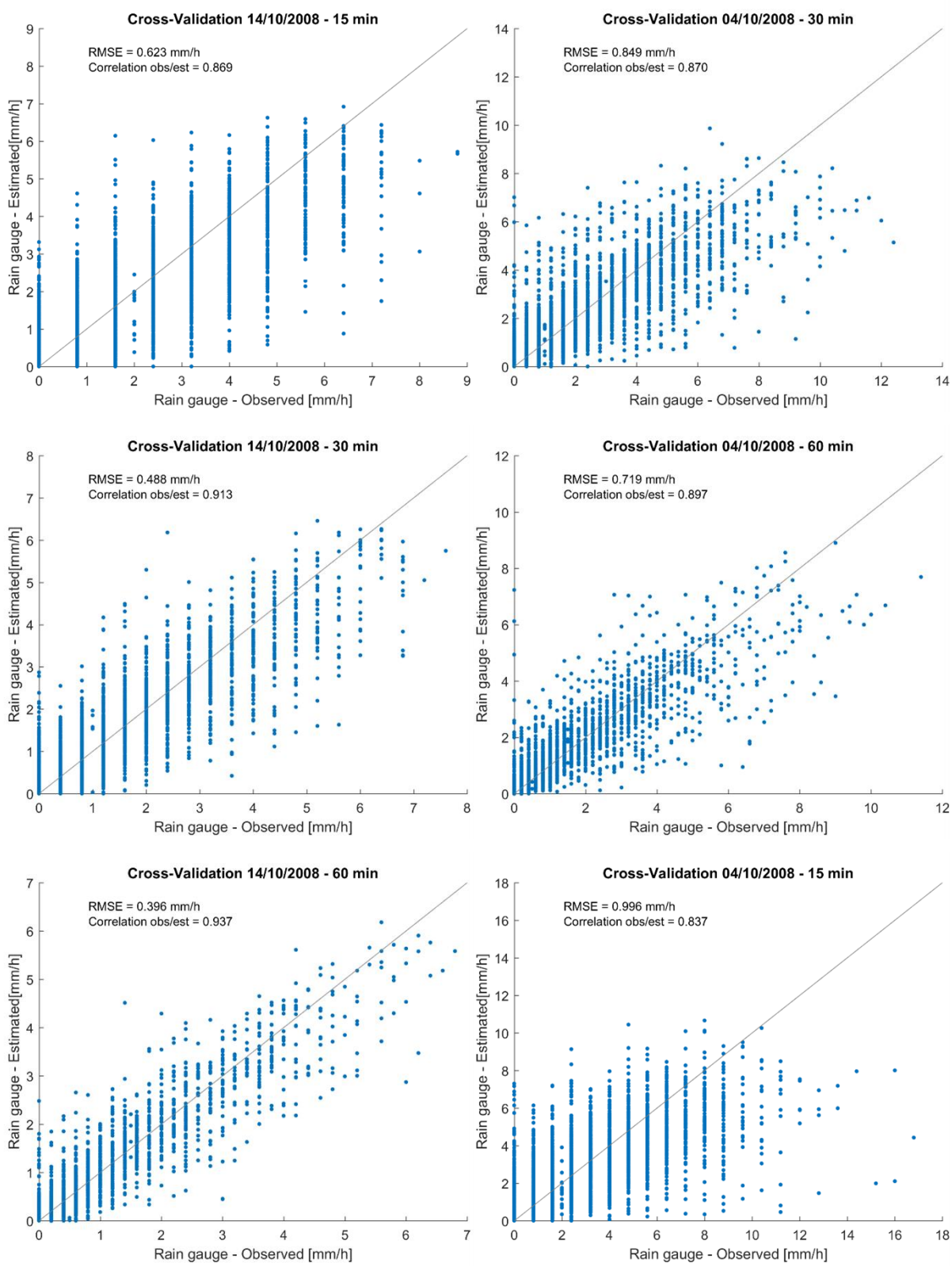


Figure 6-11. KED cross-validation for events on 04/10/2008 and 14/10/2008 with 15-, 30- and 60-min accumulation.

The summary of the MAE and correlation between observed rain gauge measurements and predicted rainfall at left-out rain gauge locations for KED for all the events can be seen in Table 6-2. For all the events, the correlation increases, and the MAE decreases as the accumulation time increases.

Table 6-2 KED correlation between observed and predicted rainfall and MAE for each event for temporal accumulation of 15 min, 30 min and 60 min.

Event Date	15 min		30 min		60 min	
	Correlation	MAE (mm/h)	Correlation	MAE (mm/h)	Correlation	MAE (mm/h)
14/01/2008	0.59	0.34	0.71	0.27	0.78	0.21
15/01/2008	0.84	0.36	0.89	0.28	0.92	0.24
16/01/2008	0.72	0.34	0.79	0.26	0.84	0.21
17/01/2008	0.73	0.36	0.80	0.28	0.85	0.22
18/01/2008	0.81	0.32	0.85	0.26	0.88	0.22
29/04/2008	0.84	0.16	0.90	0.13	0.93	0.11
30/04/2008	0.74	0.44	0.80	0.38	0.85	0.31
07/07/2008	0.35	0.42	0.42	0.36	0.53	0.29
31/07/2008	0.58	0.53	0.66	0.42	0.74	0.34
01/08/2008	0.62	0.56	0.68	0.47	0.74	0.39
02/09/2008	0.43	0.44	0.54	0.36	0.66	0.29
04/09/2008	0.43	0.34	0.50	0.29	0.67	0.24
05/09/2008	0.73	0.63	0.79	0.53	0.84	0.44
06/09/2008	0.59	0.75	0.66	0.65	0.75	0.54
04/10/2008	0.84	0.49	0.87	0.42	0.90	0.35
14/10/2008	0.87	0.34	0.91	0.26	0.94	0.22
08/11/2008	0.73	0.36	0.82	0.28	0.84	0.23
09/11/2008	0.74	0.43	0.81	0.35	0.85	0.29
12/12/2008	0.56	0.24	0.64	0.20	0.73	0.19
13/12/2008	0.66	0.49	0.74	0.39	0.81	0.32

The results presented in Table 6-2 are presented as box-plots in Figure 6-12 and Figure 6-13, where it is clearer to realise the tendency of the correlation to increase with the accumulation time. Longer accumulation times also produced merged products with lower RMSE.

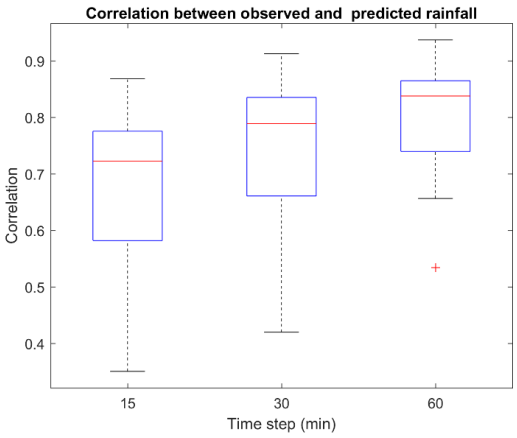


Figure 6-12 Correlation between observed and predicted rainfall for all the events with 15-, 30- and 60-min accumulation.

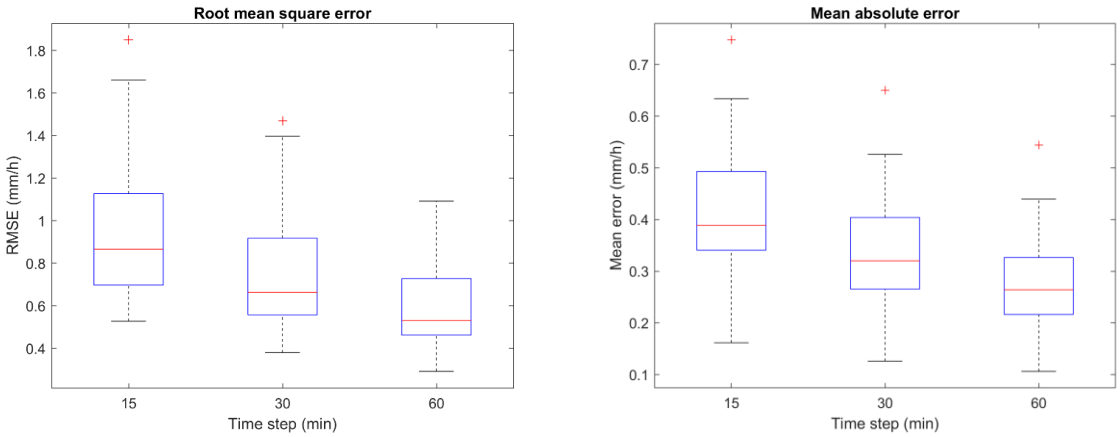


Figure 6-13 Root mean square error and mean absolute error for all the events with 15-, 30- and 60-min accumulation.

RMSE and correlation were also calculated for three different thresholds to assess how the model is able to estimate rainfall at higher intensities. The results agree with the expectations that the RMSE increases at higher thresholds. In an analogue way, the correlation between measured rainfall and KED rainfall decreases at higher thresholds.

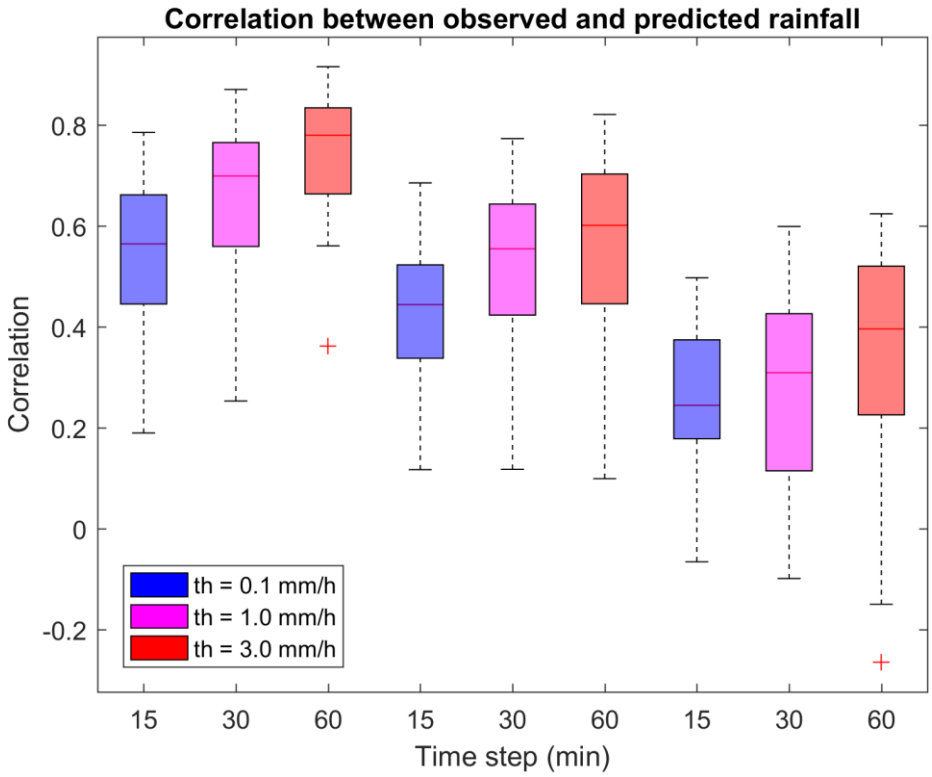


Figure 6-14 Correlation between observed and predicted rainfall for all the events with 15-, 30- and 60-min accumulation and rainfall thresholds of 0.1-, 1.0-, and 3.0 mm/h.

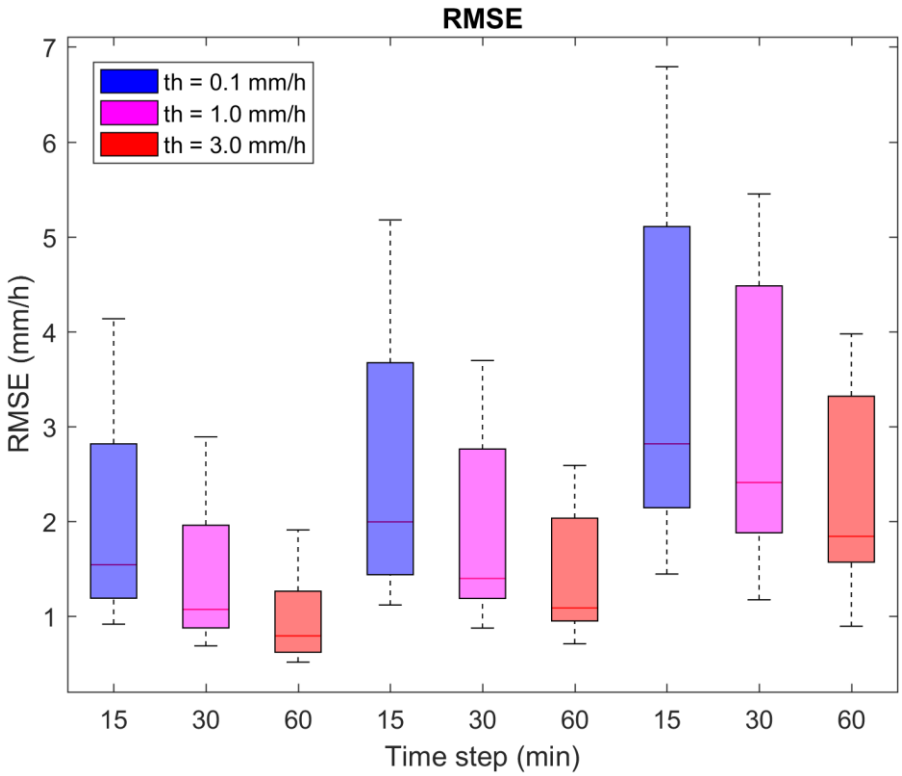


Figure 6-15 Root mean square error for all the events with 15-, 30- and 60-min accumulation and rainfall thresholds of 0.1-, 1.0-, and 3.0 mm/h.

6.4 Conclusion

Radar and rain gauge merging was carried for 20 events in 2008 using the KED rainfall merging technique. Although some studies have been carried out in the past to merging rain gauge and radar measurements using KED, there is only a few of them that analyse how the temporal resolution of the data sets influences the results. A high temporal resolution is particularly important for hydrological applications; in this chapter, 15 min, 30 min, and 60 min accumulation were tested for three rainfall intensity thresholds.

It was assessed that the variograms calculated using rain gauge data were not reliable enough as the number of data points with rainfall might be low for single time steps for a number of cases, and so the calculation of the variograms was not even possible in these cases. Jewell and Gaussiat (2015) also compared 15- and 60 min accumulation periods. They stated that computing a variogram with rain gauge data for 15 min accumulation during stratiform low-intensity events can be challenging as not enough data points are available. Nonetheless, the radar data allowed to compute reliable variograms even at shorter time steps. Therefore, radar data was chosen over rain gauge data to calculate the variograms for each time-step. Berndt, Rabiei and Haberlandt, (2014), in their study, compared KED results for 10-, 60- and 360 min accumulation periods and also used radar data to compute variograms as it provides sufficient data points for the variogram computation. The variance for 15 min accumulation was high and decreased as the accumulation period increased.

For all the events, the accuracy of the KED rainfall estimations increased with an increment of the accumulation period. However, even for accumulations of 15 min, KED rainfall estimates had fewer uncertainties than radar data. During a longer accumulation period, gauge data measurements tend to be more accurate, and there are higher chances for localized showers to be detected. Previous studies (Berndt, Rabiei and Haberlandt, 2014; Jewell and Gaussiat, 2015) also indicated that KED takes great advantage of accumulating rainfall for longer periods.

As the goal of merging rain gauge and radar data is to use the product in a new way to generate nowcasts, assessing the uncertainties for different accumulation periods is essential to the next step of this research. For nowcast applications, it is advisable to have short time-steps (15 min or less). At the same time, the quality of the merged product decreases with shorter accumulation periods. However, using merged products have the potential to be used in rainfall forecasts and hydrological applications in river catchments. Even shorter accumulations periods are needed for urban hydrology, and consequently, KED products will have higher uncertainties. Further studies are necessary to assess if using KED can be beneficial for urban hydrology. Choosing the best balance between lower accuracy in

KED for shorter accumulations and appropriate temporal resolution for the nowcasting models can result in more accurate forecasts. If time-steps of 60 min are used, it is more likely that the rainfall advection fields will be more unreliable due to changes in the precipitation pattern as they are calculated using 3 time-steps (i.e. over a 2-hour window if 60 min time-step is used). The challenge remains in how to use the KED rainfall products in nowcasting to produce more reliable forecasts.

Chapter 7. Radar and rain gauge kriging with forced velocity fields to generate ensemble forecasts

7.1 Introduction

Accurate rainfall measurement with high temporal and spatial resolutions is essential to produce accurate forecasts for hydrological applications. As discussed in Chapter 1, despite recent advances in weather radar technology and radar rainfall algorithms, uncertainties are still present even with high-quality hardware and a vast number of corrections techniques available. Rain gauges are able to represent rainfall at ground level with higher accuracy, although rain gauge measurements are still subjected to different sources of uncertainties. However, even when using dense rain gauge networks, the spatial resolution is not enough to capture the distribution of precipitation, having important implications in flow forecasting, especially in small catchments or small urban areas. Radar rainfall estimations are used as input to nowcasting models (i.e. radar-based forecasting models), and improving radar rainfall estimations will reduce errors in rainfall forecasts. Rainfall estimation errors account for the main source of uncertainties in the first hour forecasted; after this, uncertainties related to the model itself increase and have a higher impact on the forecast (Foresti *et al.*, 2013).

The higher accuracy of merged products in terms of rainfall estimation could improve the nowcast accuracy as the initial radar estimation errors are reduced. Therefore it has the potential to improve the forecasts. Caseri *et al.* (2016) used a non-conditional Turning Band Method (TBM) of random fields in time and space. The TBM used in this study is a Gaussian random field generator, adapted by Leblois and Creutin (2013) to be able to simulate intermittent rainfall advection fields. The TBM uses a space-time variogram, and the velocity incorporated by the advection technique is assumed to be constant and uniform during the length of the event. This is then followed by a conditioning step at rain gauge locations using residual substitution kriging and Markov Chain Monte Carlo sampling. The method uses the space-time variogram, direction and velocity of rain cells, wind velocity and direction of the event, the average percentage of zero rain, the mean and standard deviation of precipitation data and the rain gauge data. In order to perform the geostatistical analysis, it is necessary 4 hours of data after the start of the event. For consecutive time steps, data from the whole event is used. The results indicate a better performance when compared to a persistence-based method (where the last observed rainfall advection field using radar is assumed as the rainfall nowcast during the forecasted period). Ochoa-Rodríguez *et al.* (2013) used KED and Mean Field Bias Correction (MFB) to adjust radar data using rain gauge data. The forecasts produced using radar data and MFB present similar behaviour, while the KED forecasts show a different pattern. This is due to the fact that MFB adjusts the radar estimations by

using a multiplicative factor and therefore do not change the temporal or spatial structure of the rainfall advection field, in contrast to the merged KED product. While the radar rainfall advection field can present errors in the measurements compared with rain gauge observations, weather radars are able to capture the spatial distribution of rainfall with more detail than rain gauges, with higher spatial and temporal correlations. After merging radar and rain gauge rainfall measurements, these correlations can be distorted and cause inconsistencies in the nowcast due to the fact that KED does not take into account the temporal evolution of rainfall. Ochoa-Rodríguez *et al.* (2013) suggest that KED might not be appropriate for nowcasting, and a method that preserves the spatial and temporal correlation of rainfall might be more suitable for nowcasting applications. Shehu and Haberlandt (2021) used conditional merging as input in nowcast models and concluded that improving the rainfall estimation accuracy could extend the nowcast model predictability from 20 to 60 minutes.

Although there is a promising indication that KED would improve nowcast performance, further research is needed to assess how the nowcast performance improves compared to radar only forecasts. Previous studies suggest that the lack of temporal correlation of rainfall advection fields in KED radar gauge merging remains a challenge for nowcasting applications. The research question that this chapter attempts to address is how can KED rainfall be used to improve short-term rainfall forecasts and to take into account the temporal correlation of the rainfall advection field?

7.2 Methodology

KED is a rain gauge – radar merging technique that interpolates rain gauge data and uses radar rainfall as additional information in the interpolation process. The rainfall intensity from the merged rainfall product strongly relies on the rain gauge measurements. At the same time, the spatial distribution of the merged precipitation field is driven by the radar rainfall information. This is done by incorporating additional constraints in the kriging interpolation weights. The KED estimations produced in Chapter 6 for different accumulation periods will also be tested in order to analyse how the different accumulation periods affect the nowcasting performance. Nowcasting models benefit from a high temporal and spatial resolution. A shorter accumulation period is expected to have a higher impact in forecasting convective events where the behaviour of the rainfall pattern and intensity can change very quickly in space and time. Comparing 15-, 30- and 60 min accumulation periods can give an insight about the optimal accumulation period to be considered for nowcasting, balancing the loss of accuracy in the KED rainfall product as the accumulation time is reduced and the differences in the rainfall patterns at longer accumulations.

One of the problems of using KED rainfall in radar nowcasting is that the temporal correlation of the KED precipitation field is not taken into account in the merging process. Although the KED rainfall shows better performance than ground rainfall observations, there is a detrimental impact of the merging technique in the temporal correlation of the rainfall advection field that can produce significant effects when computing the advection field for the nowcasting extrapolation. Therefore, a new approach is proposed in this chapter that uses the temporal evolution of the rainfall advection field from the original radar rainfall advection field while at the same time taking into account the accuracy of the improved KED rainfall product. This is made in two steps: radar rainfall advection fields are calculated using radar data, followed by generating forecasts that use KED products, but instead of calculating the advection fields for this situation, they are forced to use the radar only advection field.

The nowcasting component of the STEPS model was used to generate deterministic forecasts. The model uses three consecutive rainfall estimation images to calculate the advection velocity field. This is computed and further applied to the most recent rainfall estimation to extrapolate the rainfall advection field and produce the forecast. The original radar rainfall scans were used to produce the radar advection field. The nowcasting model was modified to allow the radar advection field to be applied in the forecast (see analysis in Chapter 5) rather than the KED rainfall advection field. In this method, instead of using the KED products alone to generate the forecasts, the radar rainfall field is used as an extra input to the nowcast model. Therefore, the nowcasting model was forced to use the original radar data advection fields to generate the forecasts instead of the KED rainfall advection fields. The rainfall forecasts are produced by advecting the latest KED rainfall estimation instead of the original radar rainfall. In this way, the nowcast still benefits from a higher accuracy from the KED rainfall product. However, it is not affected by the lack of consistent temporal evolution of the rainfall advection fields of the merged KED rainfall product. This new approach is more computationally demanding than using any of the methods alone as it requires both the KED merged rainfall product and the radar advection field computed in advance as inputs for the nowcasting model. However, the time needed to generate the nowcast still short enough to allow its use for real-time applications. The forecasts were produced using accumulation periods of 15-, 30- and 60 min, as described in Chapter 6. The temporal resolution of the weather radar is 5 min, so it was necessary to accumulate the rainfall before calculating the rainfall fields. The forecasts produced by using this method will be referred to as KED with forced velocities fields (KEDFV) forecasts. At this moment, the KEDFV forecasts are produced according to the following steps:

1. Radar rainfall accumulation;
2. Radar and rain gauge merging using KED;

3. Deterministic forecasts using three consecutive radar rainfall images to calculate the radar rainfall field;
4. Deterministic forecast using KED product and radar rainfall field.

The results were assessed by comparing the forecasts against the KED rainfall product for different forecasting lead times. Different forecasting products were produced, such as the KEDFV forecasts (produced by extrapolating KED rainfall with radar advection fields), the KED forecasts (produced by extrapolating KED rainfall with KED rainfall advection fields) and the deterministic radar forecasts (produced by extrapolating radar rainfall with radar advection fields). Different performance indicators such as the RMSE, multiplicative bias, Hit Rate (HR), False alarm ratio (FAR), and Critical Success Index (CSI) were used to assess the forecast performance. HR, FAR and CSI are based on yes/no forecasts by using a contingency table. The contingency table that summarizes the outcomes of an event is shown in Table 7-1. Hit rate measures the proportion of occurred events that were correctly forecasted. FAR gives an estimation of how often events that did not occur are forecasted as rain by the model; for a perfect forecast, FAR should be 0. CSI is an index that measures the forecast's ability to simultaneously produce high POD (Probability of Detection) and low FAR, and a value of 100% means that the forecast can correctly predict rainfall (Jolliffe and Stephenson, 2011). The methods used to calculate RMSE, Bias, HR, FAR and CSI are described below on equations (7-1) until (7-5).

Table 7-1. Contingency table (Jolliffe and Stephenson, 2011)

	Event observed		
Event forecast	Yes	No	Total
Yes	a (hits)	b (false alarms)	$a+b$
No	c (misses)	d (correct rejections)	$c+d$
Total	$a+c$	$b+d$	$a+b+c+d=n$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2} \quad (7\text{-}1)$$

$$Bias = \frac{\frac{1}{N}\sum_{i=1}^{N} F_i}{\frac{1}{N}\sum_{i=1}^{N} O_i} \quad (7\text{-}2)$$

$$HR = \frac{a}{a+c} \quad (7\text{-}3)$$

$$FAR = \frac{b}{a+b} \quad (7\text{-}4)$$

$$CSI = \frac{a}{a+b+c} \quad (7\text{-}5)$$

Where F is the forecasted rainfall and O is the observed rainfall.

The radar and rain gauge data available as described in Chapter 1 were used, and a domain of 256x256 km^2 in North England was used (see Chapter 3). For assessing the forecasts' performance, 20 events between January and December 2008 were used to consider seasonality differences. Although most of the events consisted of heavy rainfall during the whole day, a 24-hour period was used to assess all events and take into account situations of low rainfall intensity. The events chosen were classified between stratiform and convective rainfall, being 17 stratiform events and three convective events. The domain and events used in this chapter are the same as those used in Chapter 6. As the results obtained in Chapter 6 are used as input in this one.

The forecasts were generated using a domain of 256x256 km^2. However, due to light precipitation issues appearing at the boundary of the domain, only a domain size of 200x200 km^2 was used to assess the forecasts' performance to avoid contouring issues.

7.3 Results

7.3.1 Velocity fields for deterministic and KED forecasts

Rainfall forecasts produced using radar rainfall and KED rainfall were used to assess the difference in the velocity fields computed for each rainfall product. Figure 7-1 shows an example of a forecast produced for the event on 15/01/2008, starting at 09:15. The first and second column in this figure are the Radar and KED rainfall advection fields; respectively, at different time steps or lead times (LTs), whereas the third and fourth columns represent the radar-based deterministic forecast ($Radar_f$) and KED forecast (KED_f) respectively, both produced using their corresponding advection fields. The deterministic radar forecast ($Radar_f$) indicates that the precipitation moves south, with a small eastern component. In contrast, the KED forecast (KED_f) produces velocities fields with an overall west direction with regions where the rainfall would be displaced north and regions with a north direction. As a result, the region with higher rainfall intensities ends up in the wrong location if the KED forecasts are used compared to the radar deterministic forecasts. Both rainfall forecasts with 3 hr lead time also show how the forecast loses its predictability at higher lead times as the nowcasting model does not take into account the temporal evolution of the rainfall advection fields nor growth and decay of precipitation. The event on 14/10/2008 with the initial forecast time at 13:15 is represented in Figure 7-2. In this case,

the radar rainfall is overestimated compared to the KED rainfall and therefore, the use of KED rainfall to drive the nowcasting model could potentially produce reasonable forecasts. However, the KED rainfall advection fields indicate that the movement of precipitation goes entirely in the wrong direction, resulting in the higher intensity rainfall regions being advected outside the domain.

Even though there are cases where both forecasted rainfall advection fields show similar patterns, there is a high number of forecasts where the differences in the advection fields produce entirely different results, for instance, where the direction of the rainfall varies greatly from one forecast to the next one. This can be seen for all accumulation periods, but it seems to have a higher impact on shorter accumulations. This initial analysis of the KED forecasts agrees with the literature. Ochoa-Rodríguez et al. (2013) discussed that the inability of the KED to capture the storm patterns is a limitation to its use in nowcasting.

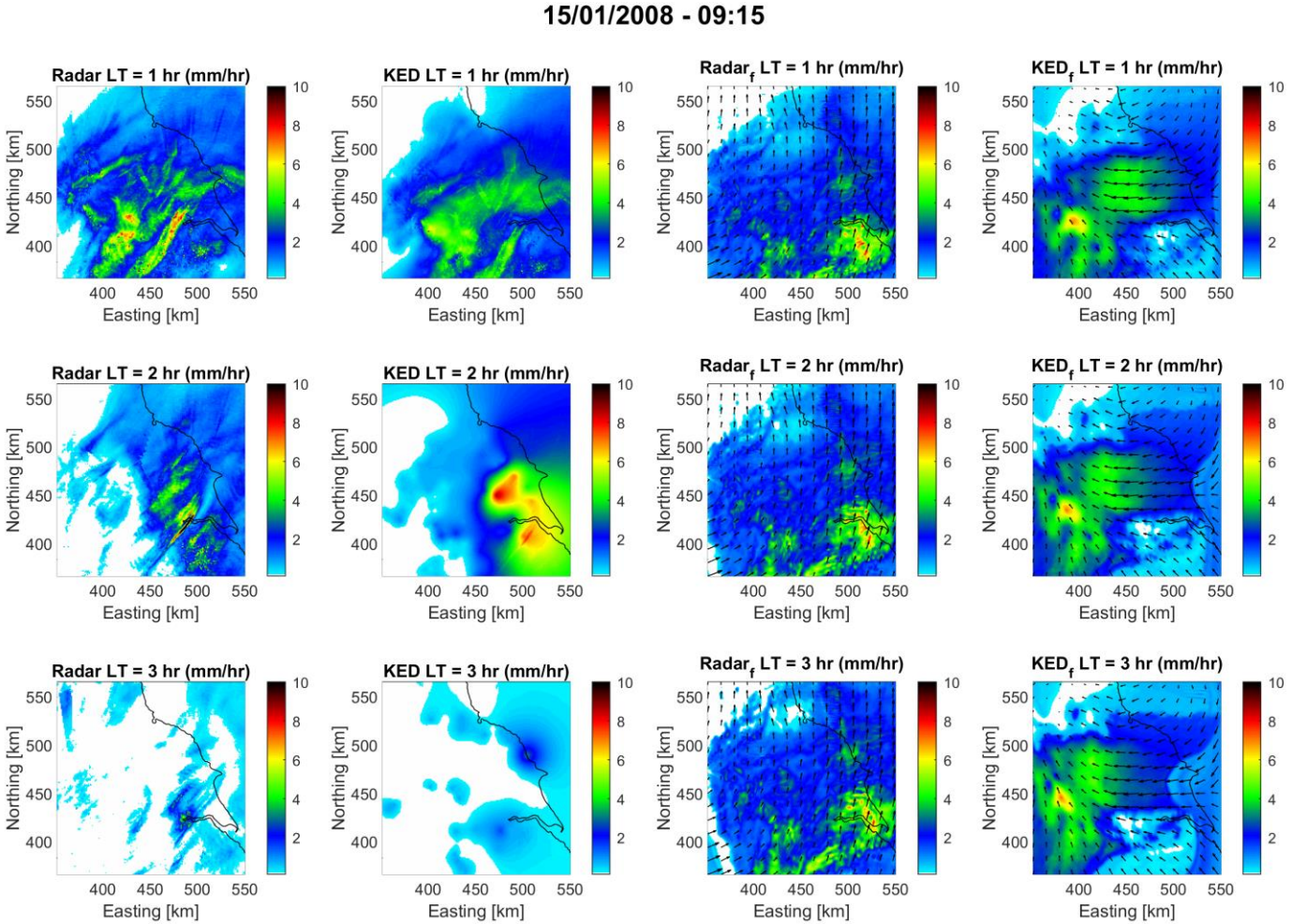


Figure 7-1. KED rainfall estimation, deterministic forecast, KED forecast and velocity fields for the event on 15/01/2008 with the forecast initial time at 09:15 for 15 min accumulation.

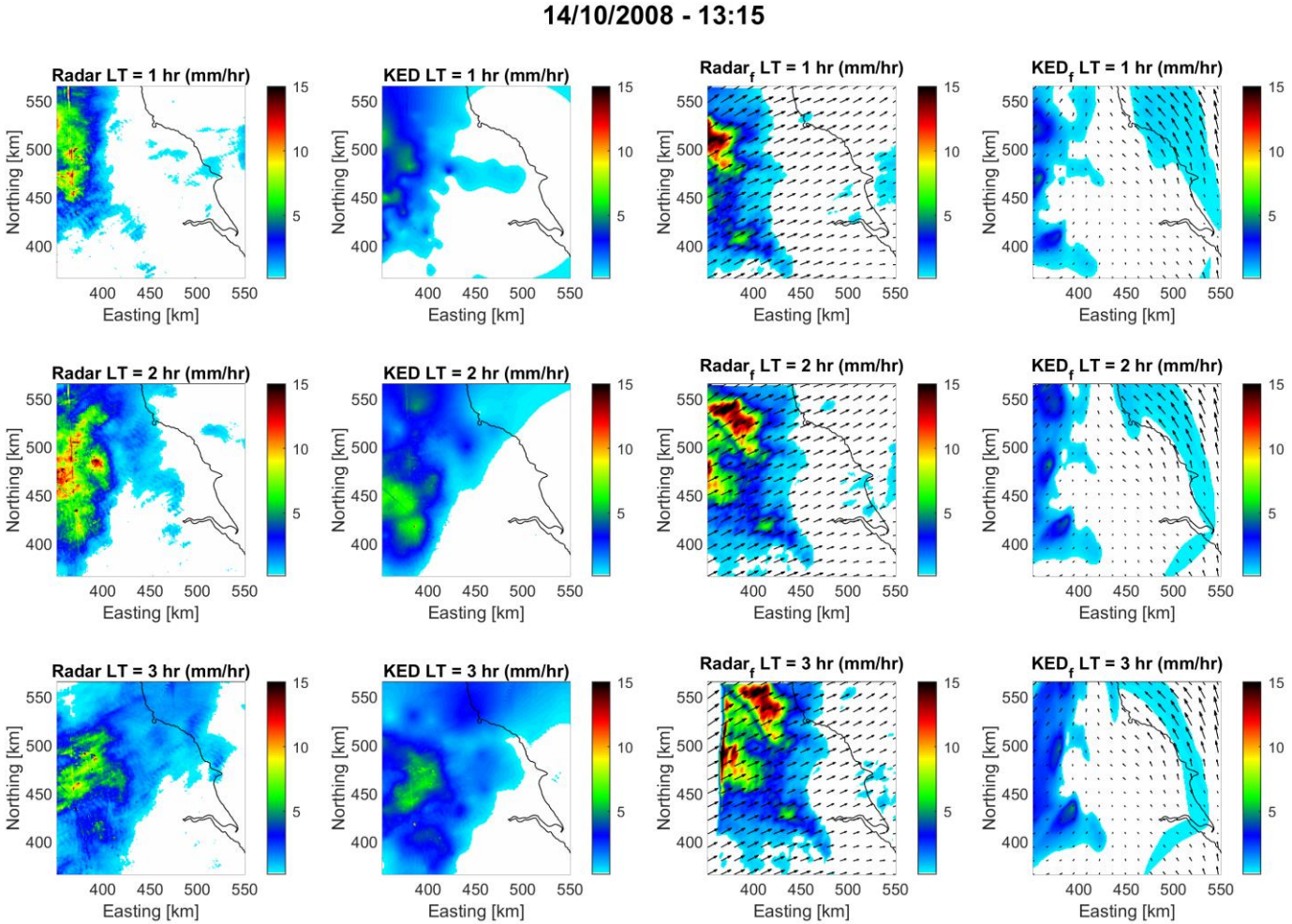


Figure 7-2. KED rainfall estimation, deterministic forecast, KED forecast and velocity fields for the event on 14/10/2008 with the forecast initial time at 13:15 for 15 min accumulation.

7.3.1.1 Forecasting KED estimates using radar-based velocity fields

To overcome the anomalies in the velocity (i.e. advection) fields derived from the KED rainfall advection fields, velocity fields derived from radar data were used, meaning that both the radar data deterministic forecast and the KEDFV forecasts have been produced with the same radar-based velocity field. However, the initial rainfall analysis differs from the two cases. Figure 7-3 until Figure 7-5 present the forecast results for the event on 30/04/2008, with the forecast starting at 00:00 and using different accumulation periods of 15-, 30-, and 60 min. For all the accumulation periods, the deterministic forecast overestimates the actual precipitation.

Figure 7-3 shows the radar rainfall (Radar, first column), KED rainfall (KED, second column), radar forecast produced with radar-based advection fields ($Radar_f$, third column), KED forecasts with KED-based advection fields (KED_f, fourth column) and KED forecasts with radar-based advection fields ($KEDFV_f$, fifth column). The radar-based advection field (arrows are shown in the third column) clearly captures the movement of the storm observed by radar. In this particular case, the storm is moving north from the south. However, the KED-based advection field (arrows in the fourth column) is unable to capture the storm movement, and the advection field is more or less stationary within the storm. The results for KEDFV show that the use of the radar-based advection field improves the forecasts, and, in this case, the KED rainfall forecasts move towards the north, which is in agreement with the movement of the storm. Although the changes in intensity with forecasting lead time are not captured by the nowcasting model, the KEDFV can still forecast the rainfall inside the intense precipitation region. At forecasting lead times longer than 2 hr, the KED forecasts miss most of the high-intensity regions and produce false alarms for high-intensity rainfall in regions of none or little rainfall.

In Figure 7-4, the KED-based and radar-based advection fields are similar, and both can capture the direction of the storm, which moves from the south towards the north. However, the KED-based advection field is slightly slower than the radar-based advection field. The radar-based velocity field of the KEDFV forecasts displaces the rainfall with a slightly stronger northern component. Hence, the

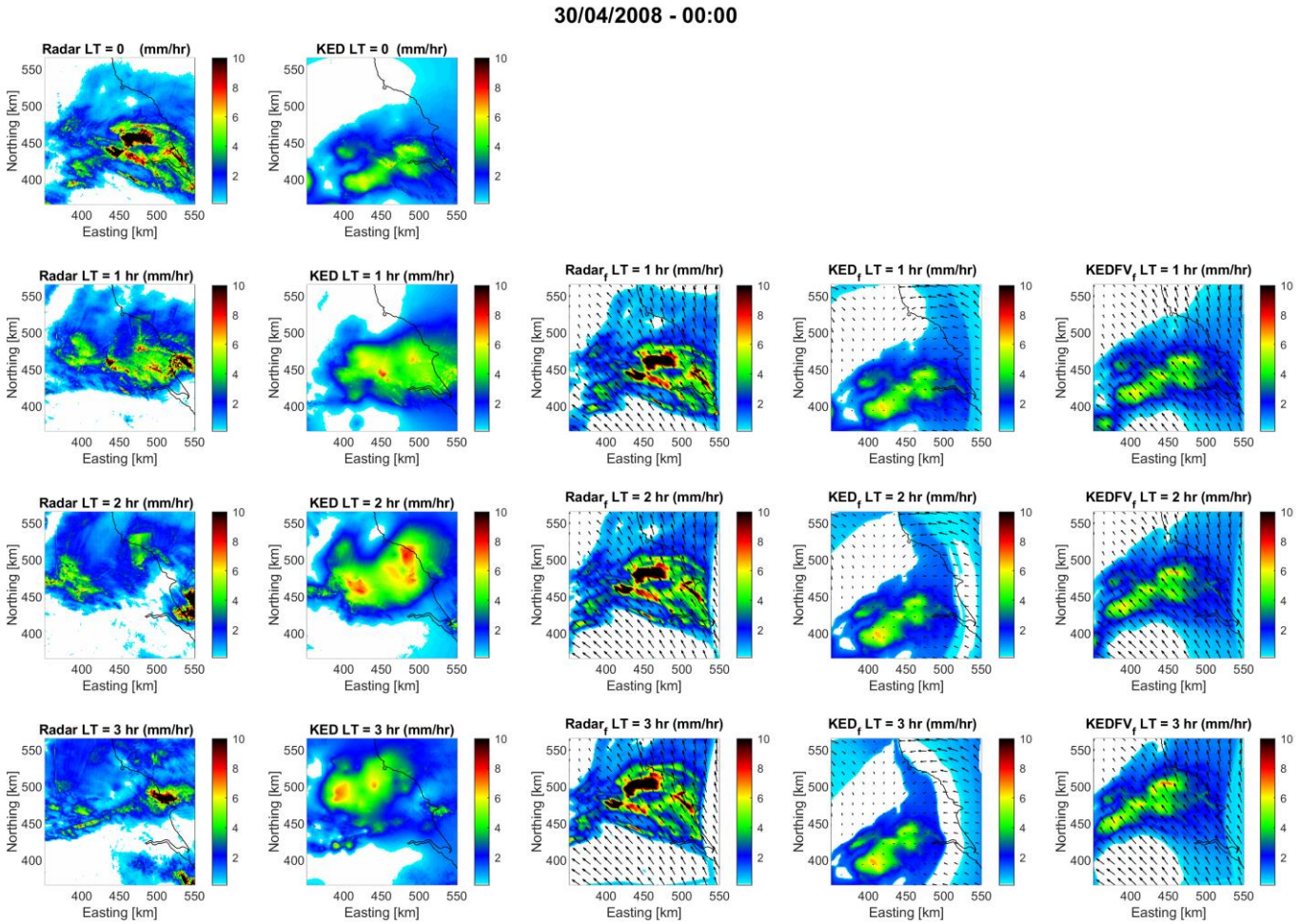


Figure 7-3. KED rainfall estimation, deterministic forecast and KED forecast for the event on 30/04/2008 with the forecast initial time at 00:00 for 15 min *accumulation.*

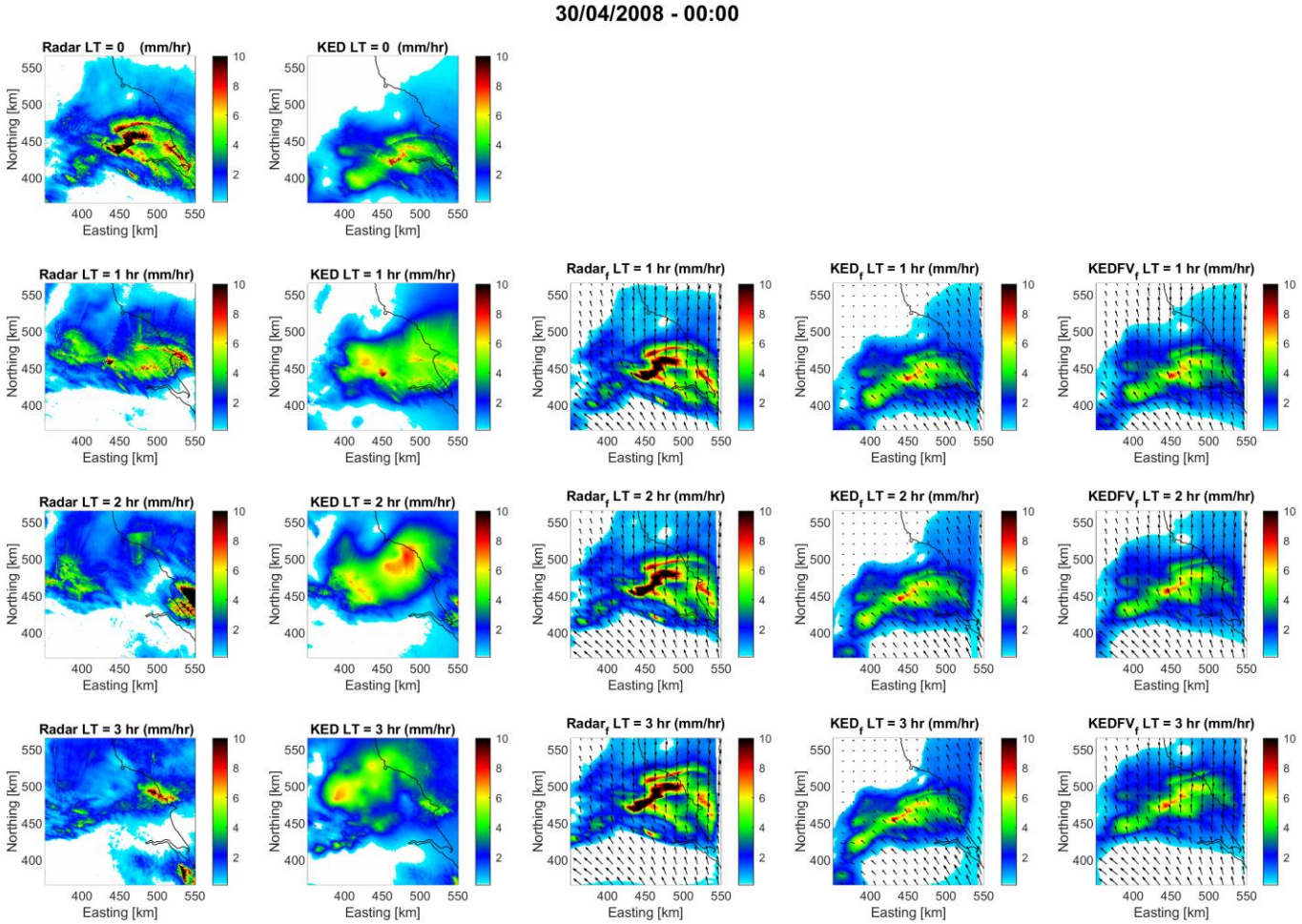


Figure 7-4. KED rainfall estimation, deterministic forecast and KED forecast for the event on 30/04/2008 with the forecast initial time at 00:00 for 30min accumulation.

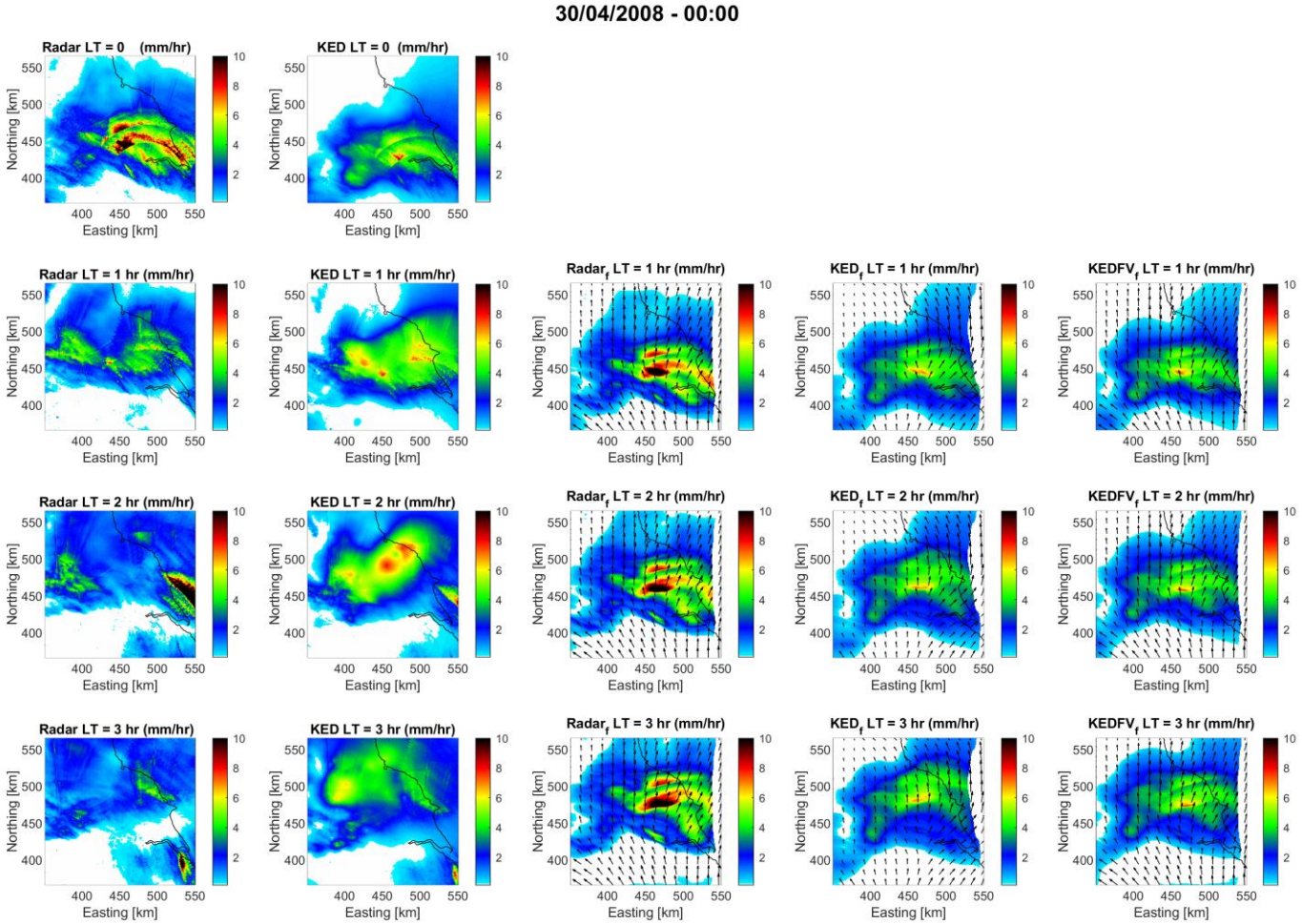


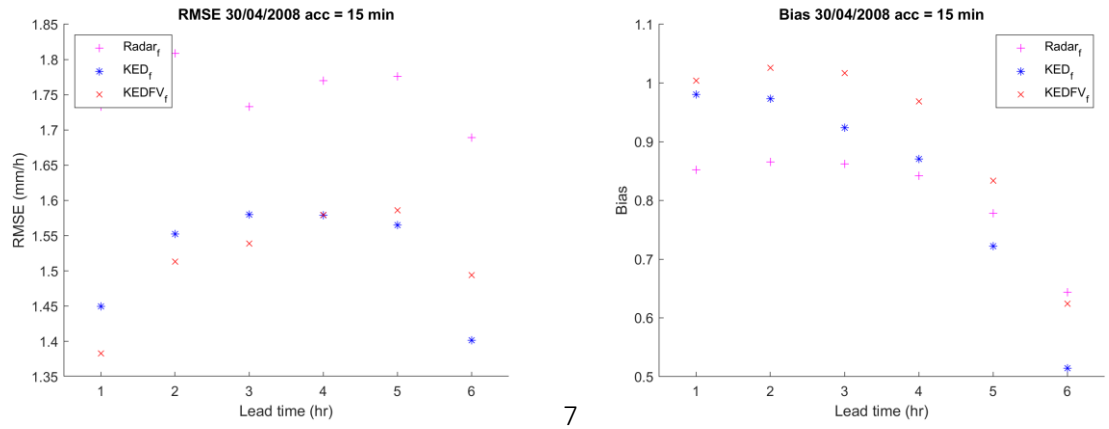
Figure 7-5. KED rainfall estimation, deterministic forecast and KED forecast for the event on 30/04/2008 with the forecast initial time at 00:00 for 60 min accumulation.

differences in the KED and KEDFV forecasts are smaller, although the KEDFV forecast can better reproduce the direction and location of the storm and can also reproduce better the low-intensity rainfall regions. These results also show that the rainfall intensity is also overestimated by the radar forecasts.

In Figure 7-5, the radar-based and KED-based velocity fields are again able to capture the storm movement. In this case, the radar-based velocity field is slightly faster than the KED-based velocity field. However, the KED and KEDFV forecasts are much more similar. However, both fail to reproduce the higher intensity regions accurately at longer lead times, showing that the nowcast loses predictability quickly for this particular event. Figures 7-3 – 7-5 also demonstrate that shorter accumulation times of 15min produce unreliable KED-based advection fields when compared with radar-based advection fields with the same accumulation time. This is due in part to the variability of the KED merged product between time steps. However, as the accumulation time increases (e.g. 30min or 60min), the KED-based advection field looks more similar to the radar-based advection field. The differences in the advection fields are present on all events studied, even if, in some cases, these differences are smaller or more noticeable in just part of the domains. By visually analysing the radar images and the forecasts, it is possible to see the differences between the radar rainfall and the forecasted KED direction. In addition to this, when KEDFV forecasts are produced, these issues are corrected, and the forecast quality improves.

Figure 7-6 until Figure 7-8 show the RMSE and bias for this event at 15-, 30- and 60 min accumulation rainfall products. The RMSE and bias shown in these figures were computed on a pixel basis and were computed using all the forecasts produced for this event during 24 hours. As expected from the forecasting results shown in the previous figures, there is a higher difference between the forecast performance at 15 min accumulation than at longer accumulation times; that is, the RMSE and the bias tend to be worse at shorter accumulation times. At 60-min accumulations, the KEDFV forecasts produce results with smaller RMSE and better bias (a bias equal to one indicates an unbiased forecast) up to 3 hrs lead time. After this period, the KED forecast and KEDFV forecast produce similar RMSE, and after 5 hrs lead time, the KED forecast overperforms the KEDFV forecast in terms of RMSE. During the whole length of the forecast, the radar forecast produces higher RMSE and unbiased forecasts than the KED-based forecasts. It is also interesting to note that the radar forecasts are the worst in terms of RMSE and bias compared to either the KED or KEDFV forecasts. For 30-min accumulation scans (Figure 7-7), the KEDFV forecast produces a forecast with lower RMSE only during the first hour ahead, and after this period, the KED/KEDFV forecasts produce similar RMSE. The bias is better for the KEDFV forecast

than the KED forecast during most of the forecasting period apart from the last hour of the forecast. During the whole forecasting period, the radar-based forecasts produce



7

Figure 7-6. RMSE and Bias for the event on 30/04/2008 at 15 min accumulation

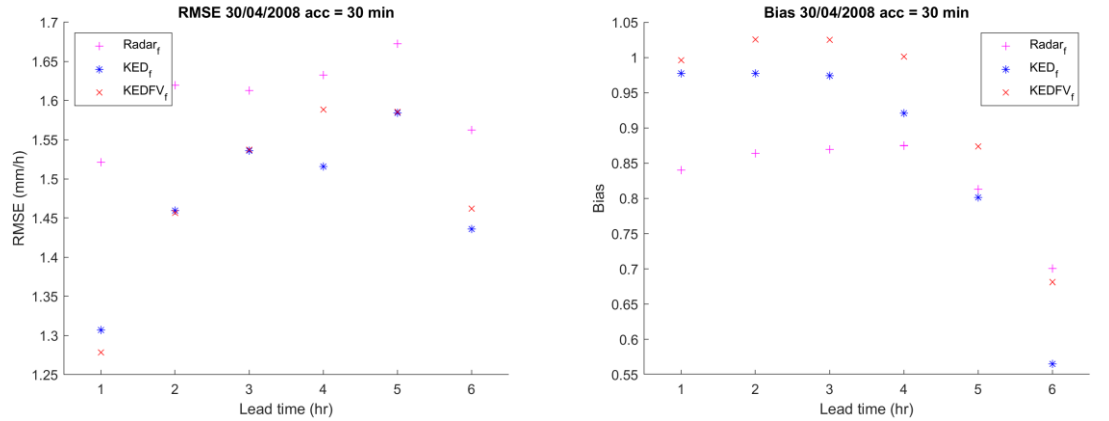


Figure 7-7. RMSE and Bias for the event on 30/04/2008 at 30 min accumulation

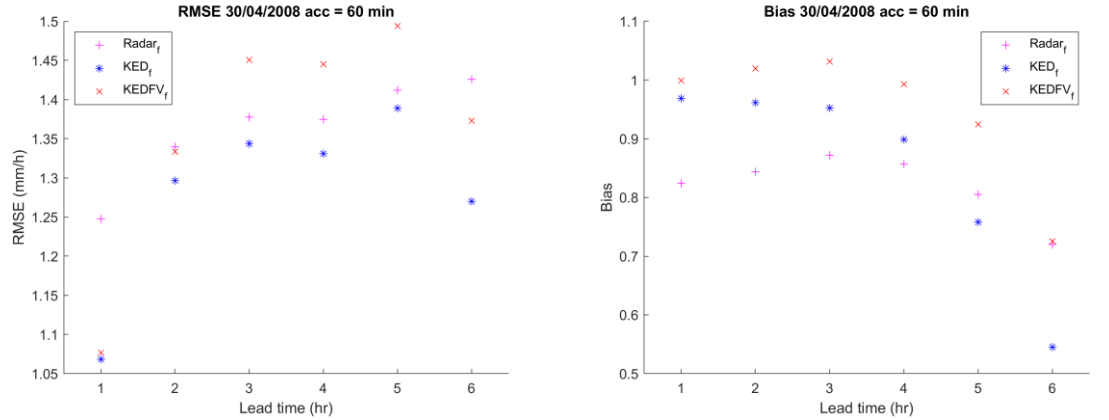


Figure 7-8. RMSE and Bias for the event on 30/04/2008 at 60 min accumulation

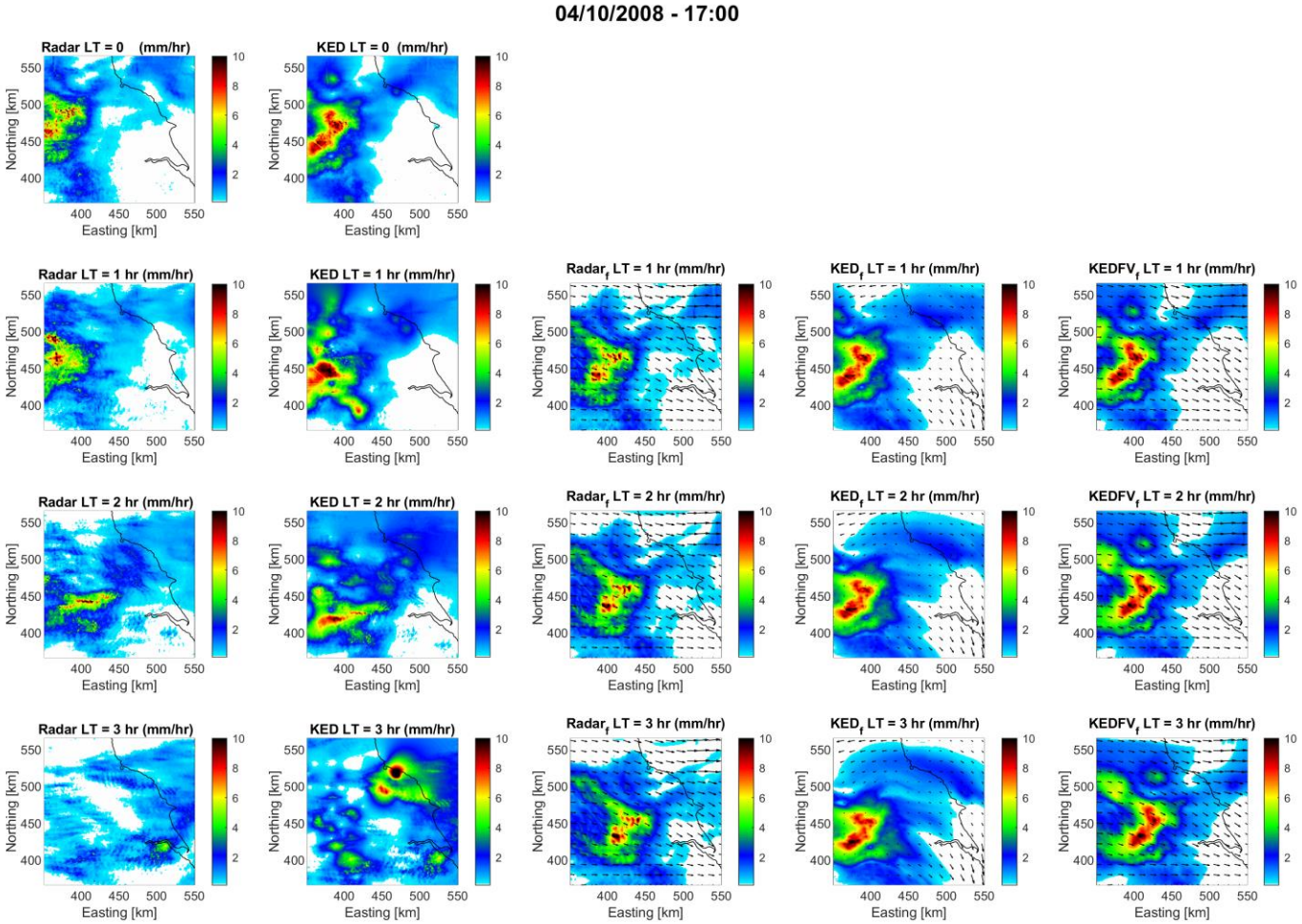


Figure 7-9. KED rainfall estimation, deterministic forecast and KED forecast for the event on 04/10/2008 with the forecast initial time at 17:00 for 15 min.

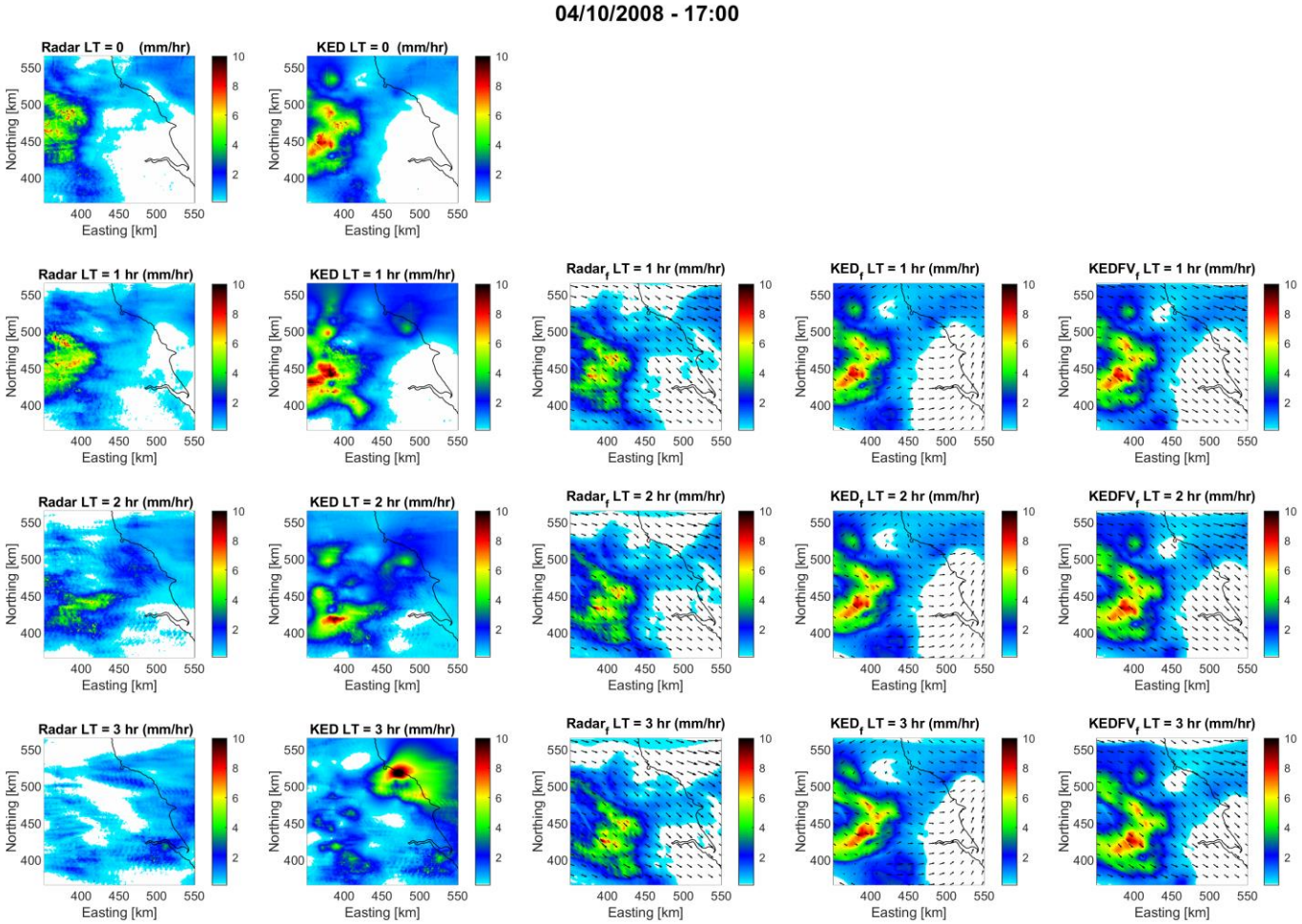


Figure 7-10. KED rainfall estimation, deterministic forecast and KED forecast for the event on 04/10/2008 with the forecast initial time at 17:00 for 30min.

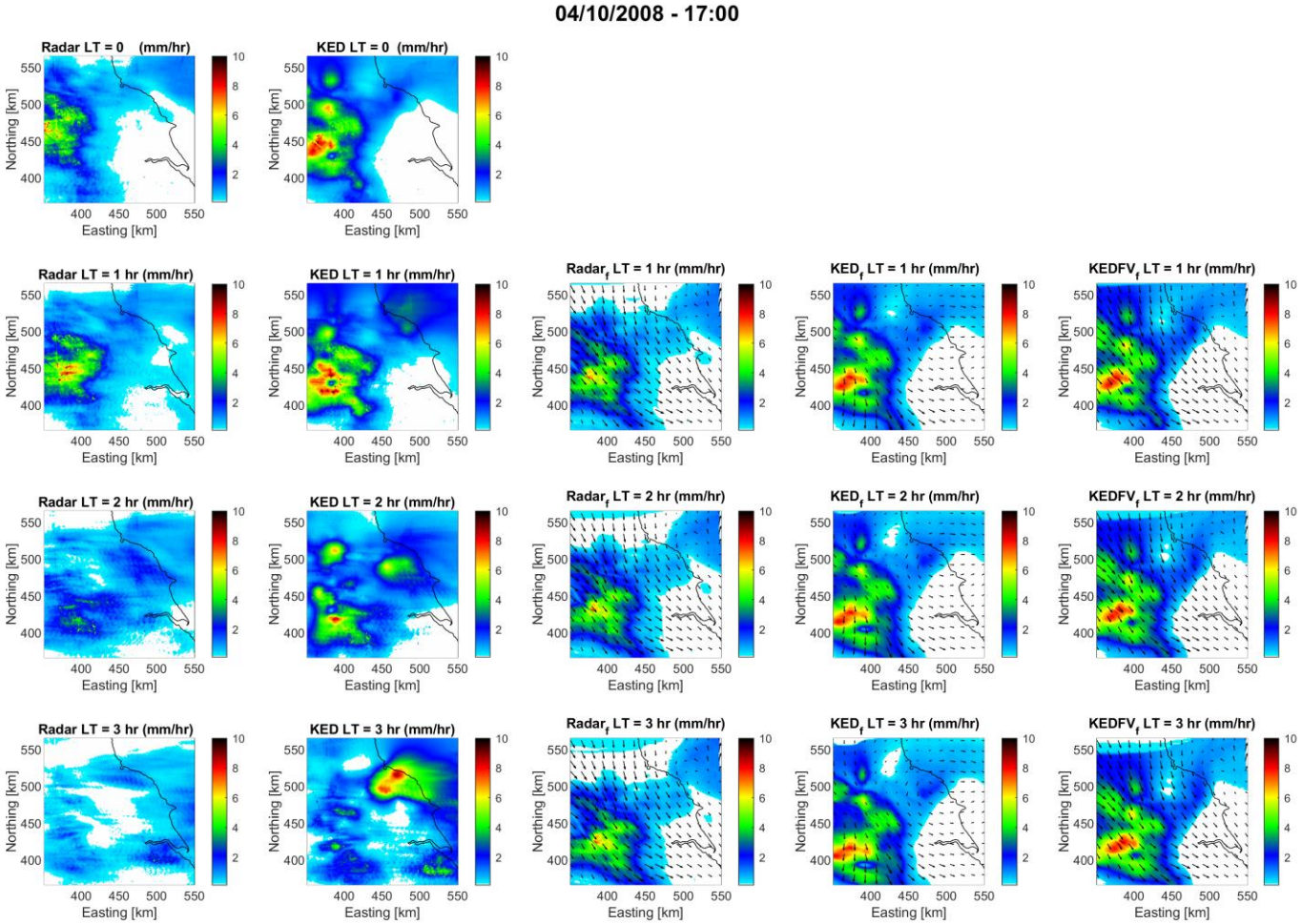


Figure 7-11. KED rainfall estimation, deterministic forecast and KED forecast for the event on 04/10/2008 with the forecast initial time at 17:00 for 60 min.

the worst results in terms of RMSE and bias. Figure 7-8 shows that the KED forecasts have a lower RMSE than the radar or KEDVF forecasts. However, the KEDFV forecast bias is better than the bias from the radar and KED forecasts. Figures 7-6 to 7-8 show that the KEDFV forecasts show the best results in terms of bias for all accumulation periods. Being able to particularly improve the forecasts at shorter accumulation periods, e.g. 15 min, shows that KEDFV forecasts have the potential to be used in hydrological applications as well.

Figure 7-9 until Figure 7-11 show the forecast results for the event on 04/10/2008, with the forecast starting at 17:00 and accumulation periods of 15-, 30-, and 60 min. In Figure 7-9, the radar-based velocity field for KEDFV moves faster than the KED-based velocity field. As a result, the rainfall displacement is more similar to the rainfall estimation. However, for three hours lead-time, it is noticeable that there is a lot of precipitation growth and decay and therefore increasing the uncertainties in all the forecasts. The KED and KEDFV forecasts shown in Figure 7-10 are closer to each other. However, all the forecasts fail to predict the high intensity and location of rainfall at a 3 hr lead time. Figure 7-11 show similar forecasts for both KED and KEDFV. Nevertheless, even with longer accumulation periods, the forecasts are not able to predict rainfall at longer lead times.

Figure 7-12 until Figure 7-14 show the RMSE and bias for the event on 04/10/2008 at 17:00. In the same way that it was done with the previous event, the RMSE and bias were calculated for the event's whole duration. The KEDFV forecast has a lower RMSE in the first hour forecasted for all the accumulation periods; however, the radar forecasts have slightly smaller errors after this period. The significant differences in the KED-based velocity field compared to the radar-based velocity field produce more significant errors in the KED forecasts compared to either the radar forecasts or the KEDVF forecasts. Taking into account the bias, the KEDFV forecast performs better than the other forecasts for all accumulation periods and during the whole length of the forecasts.

Comparing the nowcasting images for the different events, it is possible to see that the method proposed adds value to the KED forecast by correcting the rainfall advection in the KED advection fields. In the following figures, the results for all the events will be assessed by the performance indicators listed in section 7.2.

The RMSE and bias were calculated for all the events considered and rainfall thresholds (Th) of 0.1-, 1.0- and 3.0 mm/hr. The results for 0.1- and 3.0 mm/hr are shown in Figure 7-15Error! Reference

source not found. and Figure 7-16. The figures show a tendency of the RMSE to decrease as the accumulation period increases, with more

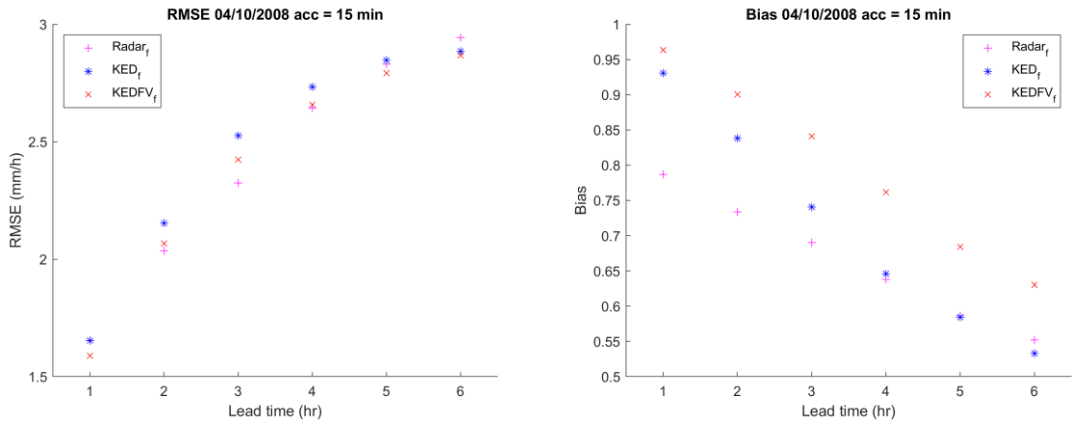


Figure 7-12. RMSE and Bias for the event on 04/10/2008 at 15 min accumulation

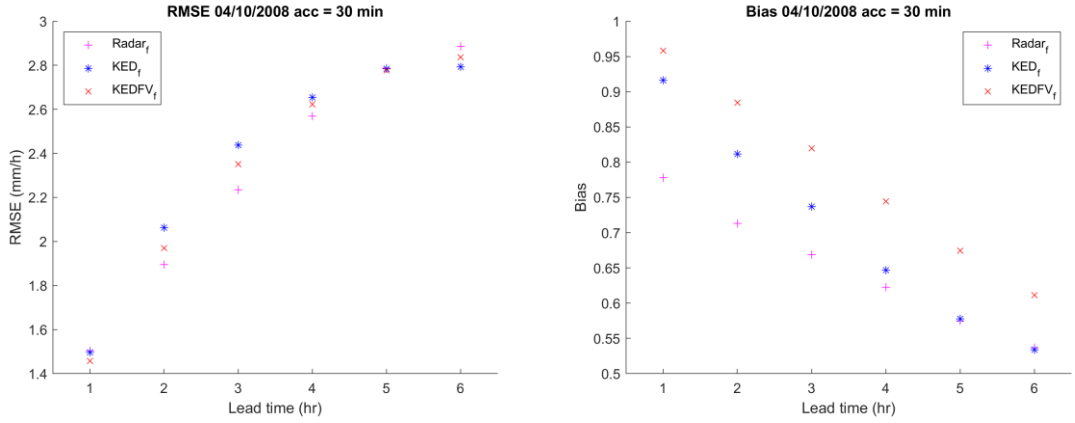


Figure 7-13. RMSE and Bias for the event on 04/10/2008 at 30 min accumulation

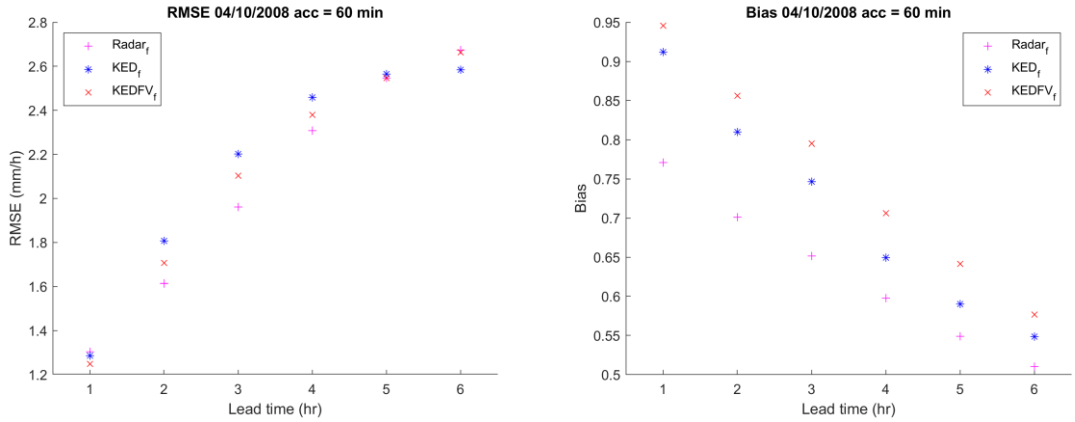


Figure 7-14. RMSE and Bias for the event on 04/10/2008 at 60 min accumulation

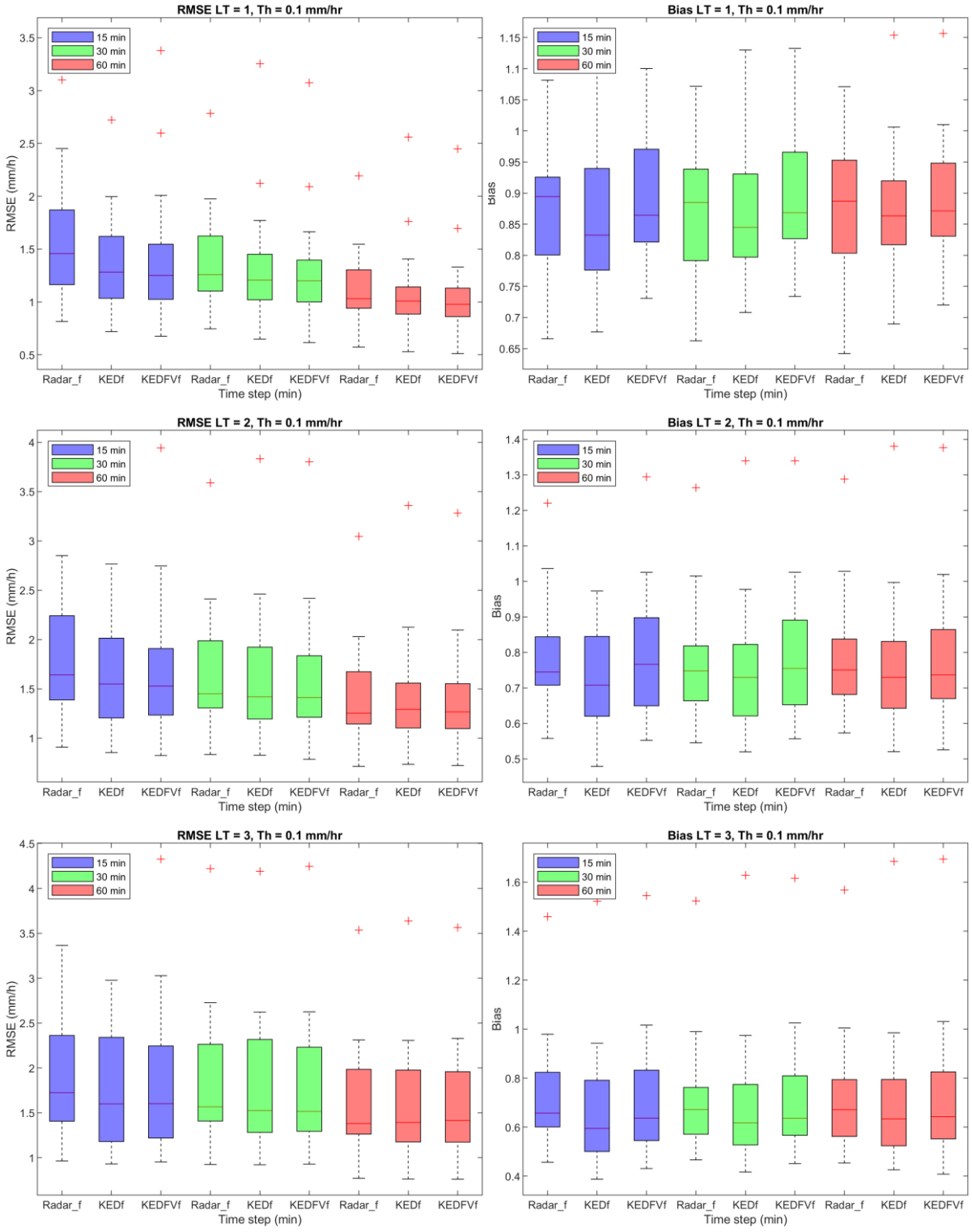


Figure 7-15. RMSE and bias boxplot for threshold of 0.1 mm/hr and lead time of 1-, 2- and 3 hr.

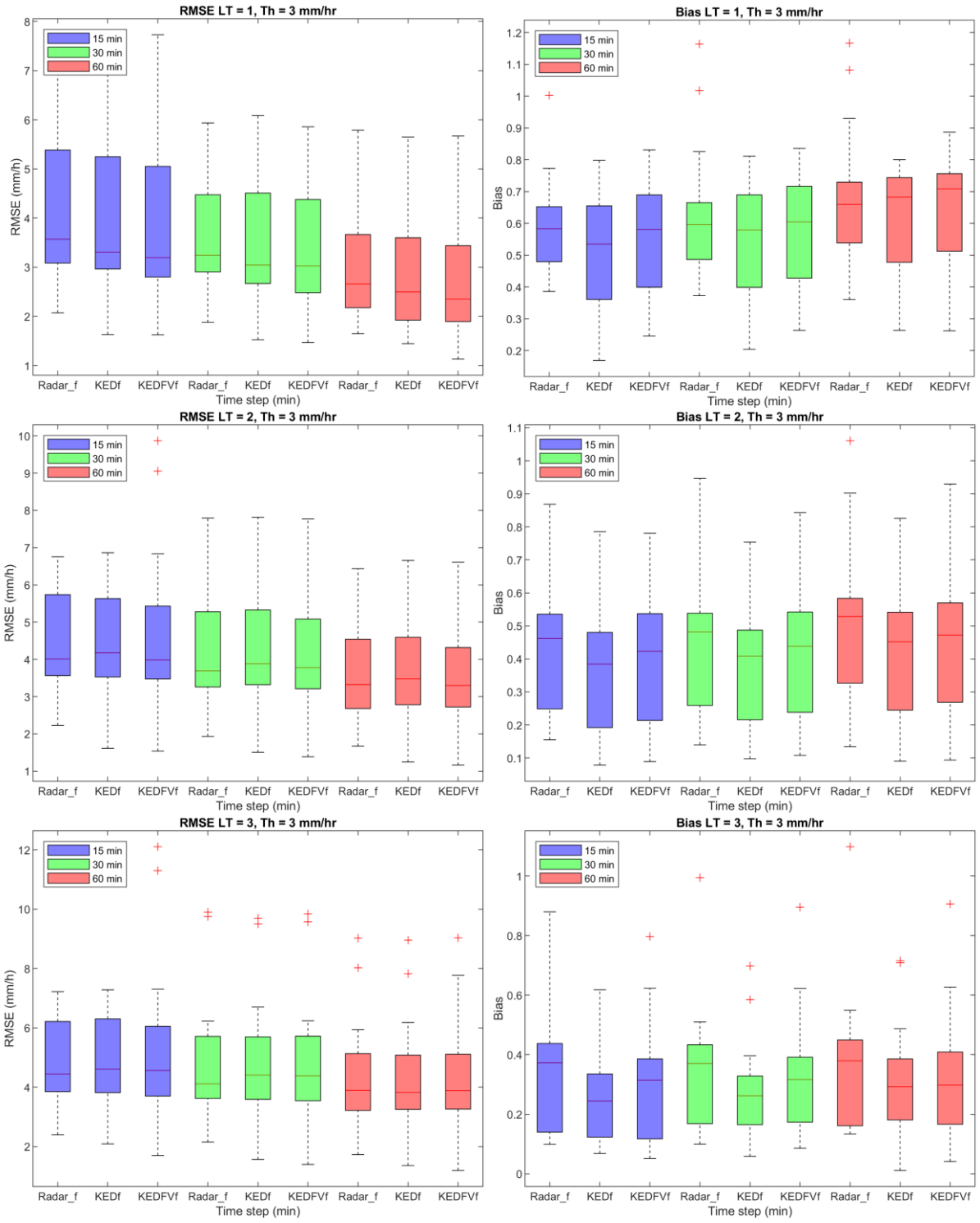


Figure 7-16. RMSE and bias boxplot for a 3.0 mm/hr threshold and lead time of 1-, 2- and 3 hr.

significant differences in the first-hour forecast. This is an expected result, as the uncertainties related to the rainfall estimation also reduces at longer accumulation times. Although the KEDFV seems to have a lower RMSE in a significant part of the events, the difference between KEDFV and KED regarding RMSE is very small. The bias of the KEDFV forecast is usually better than the bias of the KED forecast. However, the radar forecasts are able to produce bias closer to 1 in many cases. At longer lead times, the difference between the three forecasting methods gets smaller, probably as a result that all the forecasts lose predictability at longer lead times. As expected, the forecasts also perform worse when the rainfall thresholds increase.

The CSI, HR and FAR were also calculated for all events to give further insights into the nowcasts. The CSI measures the accuracy of the forecast when correct negatives are not taken into account. The CSI boxplot can be seen in Figure 7-17Error! Reference source not found. for the 1-, 2-, and 3 hr lead and thresholds of 0.1- and 3.0 mm/hr. The CSI shows that the KEDFV produce better results than the KED and radar forecasts. At 60 min accumulation, although the KEDFV performs slightly better than the KED forecast, both KED and KEDFV produce similar CSI. As previously discussed, the KEDFV can have a more significant impact on the forecast accuracy for shorter accumulation periods. The CSI values decrease with lead time for the three kinds of forecasts. The CSI values increase with a higher accumulation period. The CSI gets relatively small for higher thresholds, and at 3 hr lead time and 3 mm/hr thresholds, there is almost no difference between the different forecasts. Figure 7-18 and Figure 7-19 show the Hit rate and false alarm ratio for the events at 1-, 2- and 3 hr forecast. The hit rate for the KEDFV forecast is better than for the radar and KED forecasts for most cases. However, for longer accumulation times and lead times, in a similar manner to the CSI, the values for KED and KEDFV get closer. The false alarm ratio for the KEDFV forecast is usually similar to the KED forecast, however for 15- and 30 min accumulation and 1 hr lead time, the KEDFV overperforms the other forecasts. For most cases, the radar forecast produces a lower FAR than the KED based forecasts. The hit rate tends to decrease with longer lead times and higher thresholds. Even with the KEDFV performing better than the KED and radar forecasts in general, when taking into account HR, FAR, and CSI, it is important to note that for lead times longer than 2 hr and a threshold higher than 1 mm/hr, any of the forecasts produce reliable predictions.

All the performance indicators show that the KEDFV overperform both radar only and KED forecasts. KEDFV forecasts have the potential to be successfully used in hydrological applications, as it mainly improves forecasts at shorter accumulation periods.

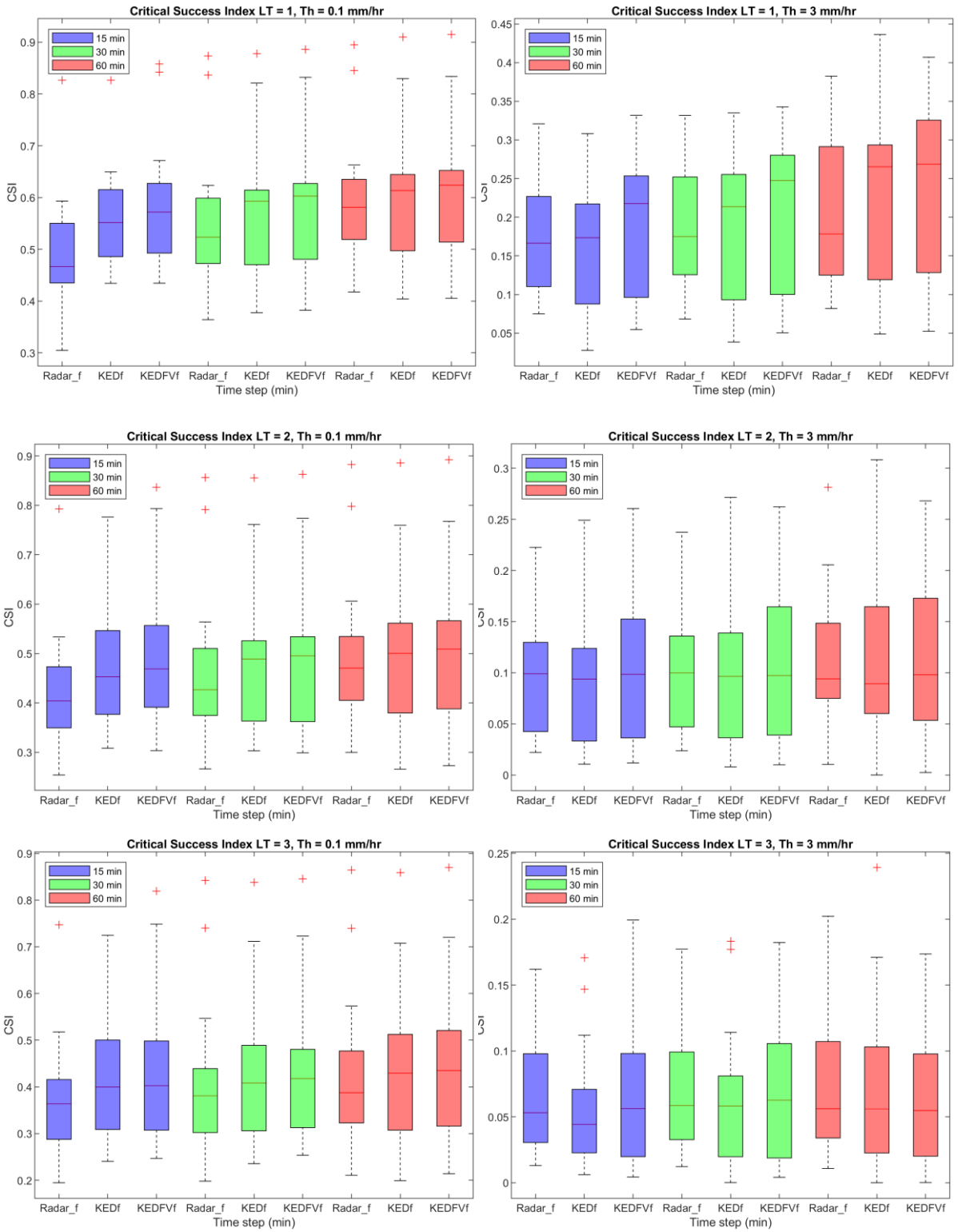


Figure 7-17. Critical Success Index (CSI) boxplot for threshold of 0.1 mm/hr and 3.0 mm/hr, and lead time of 1-, 2- and 3 hr.

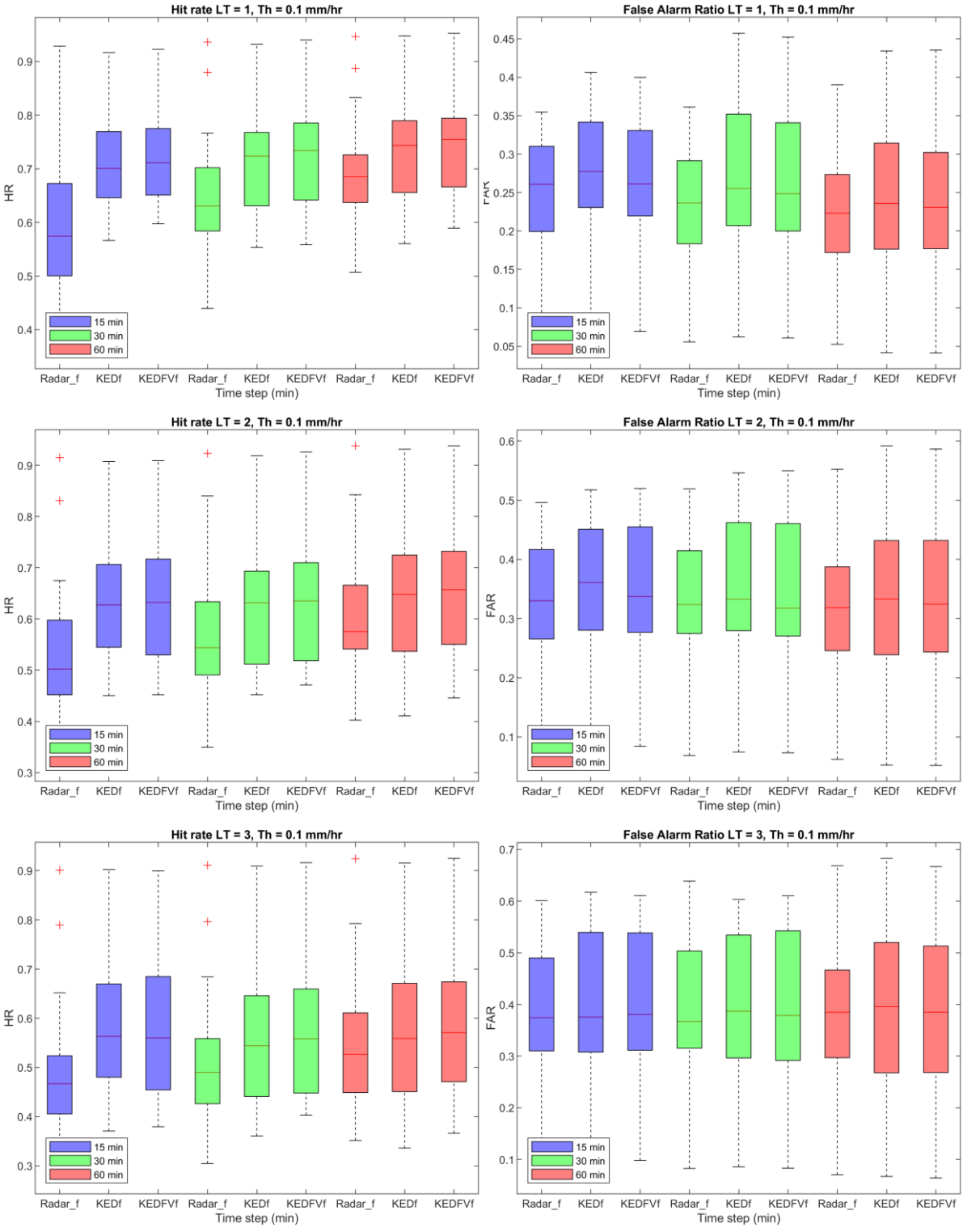


Figure 7-18. Hit rate and false alarm ratio boxplot for threshold of 0.1 mm/hr and lead time of 1-, 2- and 3 hr.

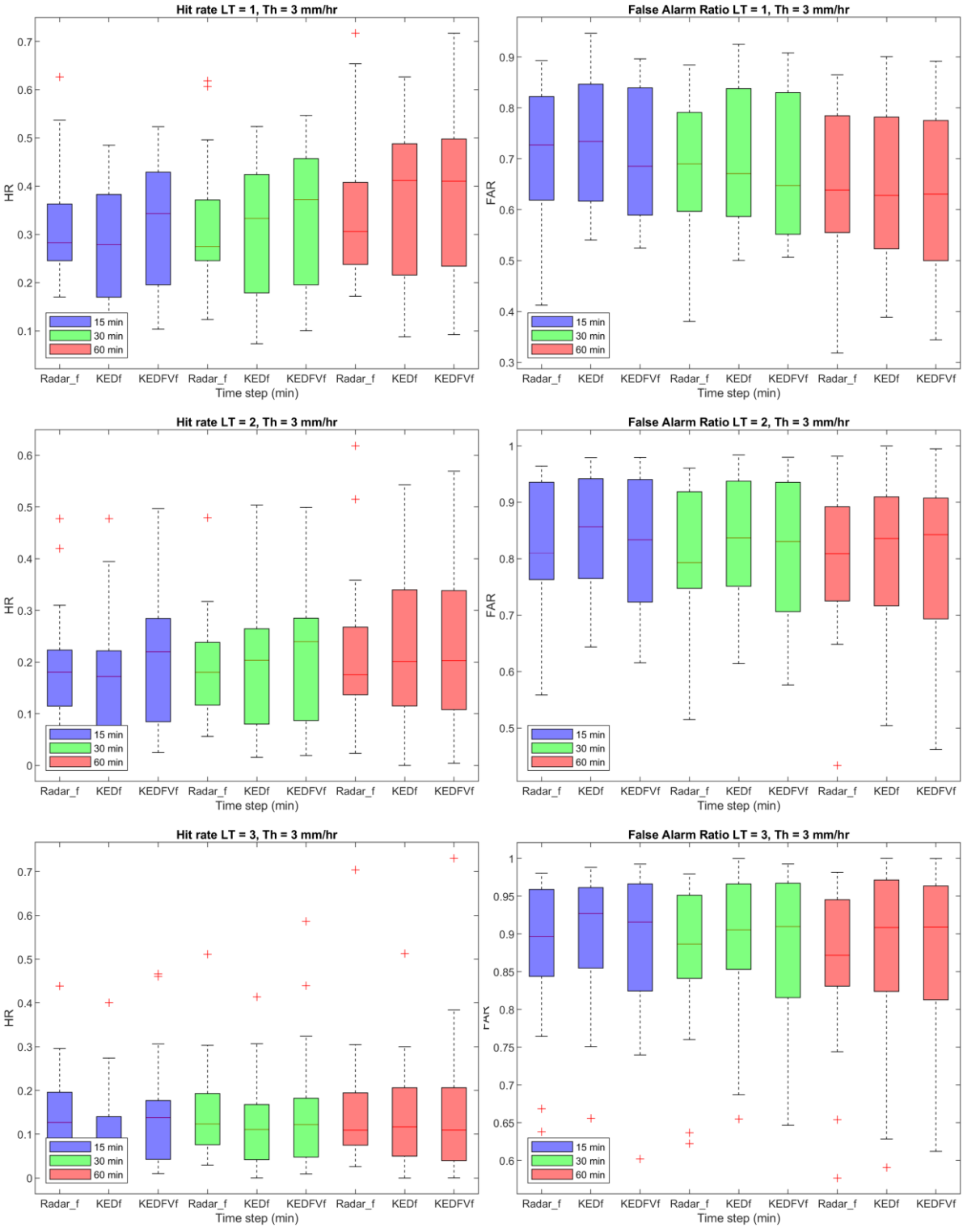


Figure 7-19. Hit rate and false alarm ratio boxplot for threshold of 3.0 mm/hr and lead time of 1-, 2- and 3 hr.

7.4 Conclusion

KED rainfall estimation is more reliable than radar only estimations. However, limitations due to the lack of temporal correlation between time-steps impose a limitation to its use in nowcasting. For this reason, there is a lack of research papers addressing nowcasting using KED. A new method to produce short-term forecasts using KED rainfall products was proposed to address this limitation. Results comparing the velocity fields of radar and KED forecasts support previous findings stating that the precipitation distribution for the KED forecasts had an erratic behaviour due to the lack of temporal correlation of precipitation between consecutive time steps (Ochoa-Rodríguez *et al.*, 2013). The KED rainfall product ensures that the spatial correlation of rainfall is kept at each time-step. However, there is no continuity between different time steps. This means that between time-steps, the calculated direction of the rainfall can be very different from the real one; as a result, the forecasted rainfall can be extrapolated for a completely different location. Ultimately, there is an increase in the forecast uncertainty in cases with large differences between the radar advection fields and the KED rainfall advection fields. The method proposed in this chapter takes into account the fact that the advection field produced by the original radar rainfall is better than the advection field produced by KED and can capture the rainfall pattern of rainfall. However, because the KED rainfall estimates are better in estimating rainfall intensities than radar data alone, the method extrapolates the KED rainfall advection field with the advection field from radar. In this way, the method takes into account the temporal evolution of precipitation, which is often lost in the KED rainfall products. The forecasts were assessed by comparing them with a deterministic radar-based forecast and a deterministic forecast based on KED only estimations.

Results indicate that there is a tendency for RMSE to decrease and the bias to increase as the accumulation period increases. RMSE for KEDFV forecast is slightly smaller than KED forecast RMSE. However, the KEDFV forecast usually produces better bias. At longer lead-times, the forecasts' bias and RMSE values get smaller and closer to each other due to all of them losing predictability at longer lead times.

In most cases, HR is better for KEDFV. The radar-based forecasts usually produce the lowest far than the KED based forecasts. The CSI, which compare HR and FAR, are better for KEDFV forecasts. HR, FAR and CSI and tend to present worse performance as lead time and thresholds increase. Lead-times longer than 2hr and threshold higher than 1 mm/hr, HR, FAR and CSI indicate that any of the forecasts produce

reliable predictions. As the accumulation periods increase, KED and KEDFV performance indicators values get closer, indicating that KED rainfall advection fields are closer to radar fields and suffer less with loss of temporal correlation.

The very similar results for the KED and KEDFV for 60 min accumulation mean that there is little gain in using a more sophisticated method as proposed above for longer accumulation periods. However, for shorter accumulations, the KEDFV does produce more accurate forecasts because it takes advantage of the radar-based advection field. It was expected that 15 min accumulation forecast would not perform well as its subject to more uncertainties from the KED merging. However, the results produced showed a better performance than using radar data alone. Results show that the nowcasting model benefits from the method proposed, with a more significant impact in shorter accumulation periods. This finding is exciting as it indicates that the KEDFV forecasts can be used in hydrological applications that need higher temporal resolution and should be the subject of further research. Using the KEDFV method showed that it is possible to use KED merging to produce a forecast with a high temporal resolution with higher accuracy than using radar data alone.

Chapter 8. Conclusion

The research carried in this thesis has as main objective to improve the reliability of radar-based rainfall forecasting and flow forecasting by addressing uncertainties related to rainfall estimation and its propagation into nowcasting and hydrological models. Two primary sources of uncertainties in nowcasting models were studied (uncertainties related to radar rainfall estimations and the temporal evolution of rainfall), and methods to account for them were proposed

In order to achieve this, an understanding of rainfall measurements and their uncertainties was necessary. Reviewing the current nowcasting techniques and their limitations was important to be able to propose ways to account for the uncertainties inherent to the model.

Chapter 2 discussed a review of some of the current techniques to measure precipitation, their importance and main limitations. Hardware updates and algorithm development are important to reduce uncertainties, but residual errors still propagate into rainfall and flow forecasting models. In this chapter, a description of nowcasts models and their main uncertainties were highlighted.

In Chapter 4, an ensemble generator that adds spatially correlated noise to radar images was used to assess the propagation of weather radar uncertainties into rainfall forecasting and in flow forecasts. The chapter focused on answering the research question Q1, and a summary of the results are shown below.

- Ensembles were generated by comparing historical rain gauge and weather radar data sets to model residual errors in radar estimations and were used in a nowcasting model to produce rainfall forecasts. Flow forecasts for an urban area and river catchments were generated.
- Rainfall forecast results showed that both ensemble systems performed better in low-intensity events
- For high-intensity rainfall, forecasts are only reliable up to 1-hour lead-time and lost predictability rapidly after this period.
- RE ensembles produced more accurate forecasts than STEPS ensembles in the first hour; this coincides with the period where radar estimations have a higher impact on the nowcast accuracy.

- For lead-times longer than 1-hour, RE ensembles could still produce better forecasts than STEPS ensembles in some events. However, in general, its forecasting ability decreased quicker than for the STEPS ensembles.
- Urban catchment results:
    - RE ensembles were better than STEPS ensembles regarding flow estimation, being able to better reproduce the flow peaks intensities, even in cases that the radar data could not correctly estimate the flow.
    - Any of the probabilistic forecasts could forecast flow peaks efficiently after 1-hour lead-time in the urban catchment.
    - In events where both radar and rain gauges failed to estimate Ilkley's flow peaks, any of the ensembles could predict them either.
    - Results for urban flow forecasting showed a time-lag of a few minutes; however, analysing the results using cross-validation and comparing the peak flow forecasted over a period of time show that the flow forecasts can be used in real-time applications for up to 1-hour lead-time.
- River catchments results
    - RE ensembles were able to predict flow peaks for most events better.
    - The fact that RE ensembles had a higher spread also meant that it was able to predict large peaks that were missed by the STEPS ensembles due to the fact that the peaks were also not estimated by radar data.
    - In cases where radar and rain gauge simulated flows were very different from the observed flow, the peaks were more difficult to predict.

Chapter 4 results showed that rainfall estimation ensembles are able to produce better ensembles than the STEPS probabilistic forecasts, mainly with up to 1 hour lead time. The benefits of using RE ensembles can also be seen in flow forecasts for both urban and river catchments. However, the forecasts lose predictability fast after the first hour.

In Chapter 5, a new method to generate ensembles forecasts was proposed in order to address uncertainties related to the temporal evolution of rainfall. Ensembles are generated using rainfall velocity fields up to two hours before the forecast initialization. Using past rainfall advection fields that were also calculated using different time-steps gives valuable information on how the rainfall pattern

develops with time. Depending on the time-back considered, probabilistic forecasts with a different number of ensemble members were produced. Following there is a list of the main findings in the chapter:

- Results showed that the forecasting ability increased with the number of ensembles members.
- For longer lead-times, having more ensemble members in the probabilistic forecast showed to have more impact than for up to 1-hour lead time. This is probably due to the temporal evolution of rainfall advection fields uncertainties being more significant after 1-hour lead-time. However, the forecast keeps losing ability with longer lead times.
- Comparing the area beneath the ROC curve of the rainfall advection fields ensembles produced better forecasts after 1-hour lead time than the probabilistic forecasts generated in Chapter 4. This means that the forecasts are able to produce better improve the performance of nowcast regarding uncertainties due to the temporal evolution of rainfall advection fields.
- Flow peaks could only be forecasted in cases when they were also estimated by radar data.
- Using a 60-min time window (27 ensembles) was enough to improve the forecast. The use of more ensembles would not be justified when considering the time necessary to produce them and the gain from having more ensembles.

The assumption that past information about the temporal evolution of rainfall could be used to improve probabilistic forecasts has been proved right. Increasing the time-window used does improve the forecast even when data for up to two hours before the forecast is used and have a more significant effect at longer lead times. However, this also leads to the need for more computer power and time to produce the forecasts, with little gain to its efficiency. Probabilistic nowcasts with 27 ensemble members are much faster to be produced than those with 63 ensembles, and they can still improve the forecasts.

In Chapter 6, the merging of rain gauge and radar data was discussed to address the research question Q3. Merging using KED was carried out for three different accumulation periods, and the impact of the temporal resolution on rainfall estimation for different intensities were analysed. As shorter accumulation periods were also used, using rain gauge data to calculate the variograms was not possible as there were not enough data points. Radar data was used to calculate variograms. In this chapter, it was found that:

- Variograms for a higher temporal resolution presented high variance, but variance decreased as temporal resolution got sparse.
- Using shorter accumulation periods for KED merging is particularly important for forecasting applications, and results indicated that even for a 15 min KED merging produced better rainfall estimations than radar data.

Although it was shown that it is possible to improve radar rainfall estimations by merging it with rain gauges using KED even at short temporal resolutions, it is important to highlight that the uncertainties increase with the temporal resolution. Nonetheless, the merged product still offers advantages compared to radar data alone, even at 15 min temporal resolution, showing that it could be advantageous to use KED merged products in nowcasting.

A new approach to produce short-term forecasts using the KED product was proposed in Chapter 7. It addresses the research question Q4 by proposing a way to use KED in nowcasting and investigating how temporal resolution affects the nowcast outcome. The forecast produced takes advantage of the high temporal correlation from weather radar images, coupled with higher accuracy in estimating rainfall obtained from KED merging. The deterministic forecasts are produced by applying radar only motion fields to KED estimations. In this way, the lack of temporal correlation of KED products is not part of the forecast. Following, there is a summary of the main findings of this chapter:

- Bias, HR and CSI indicate that, overall, KEDFV was the forecast with better performance.
- Regarding false alarm rates, KED forecasts had the lowest values when compared to both radar and KEDFV. Nonetheless, the CSI rate indicates that even then, KEDFV produces better forecasts.
- Radar forecasts results have lower RMSE than KEDFV forecasts. The RMSE between KED and KEDFV forecasts were comparable.
- For 15 min accumulation, the difference in performance between KED and KEDFV was higher for short accumulation periods.
- As the accumulation time increased, the performance of the forecasts got closer, and the two methods produced similar results for hourly accumulations, indicating that for 60 min accumulations, KED is less subject to time correlation errors.
- All the forecasts skills decreased quickly after a 2-hour lead time.

- In applications that hourly temporal resolution is sufficient, KED forecasts are as reliable as KEDFV.

KEDFV is a unique method as it overcomes one of the main limitations of using KED for nowcasting, having a more significant impact at shorter accumulation periods. Being able to produce reliable forecasts even at 15 min temporal resolution means that there is space for extending this research to use KEDFV for hydrological application.

Two different methods were used to address the issues related to rainfall estimation in forecasting and hydrological applications. The first one is based on RE ensembles' production using historical data for accounting for residual errors. This method has been developed by Germann *et al.* (2009) and further implemented into the STEPS model (Seed, Pierce and Norman, 2013). In this thesis, the method was tested for hydrological applications. The ensembles are generated before the forecast, and for each ensemble member, a deterministic nowcast is carried out. The second method is an entirely new technique to use KED to generate nowcasts. Compared with the radar estimation ensemble generator, this method uses rain gauges measurements contemporary to the radar data to adjust the radar estimations. Both methods have a higher impact in the first hour forecasted.

The KEDFV method was used to generate only deterministic forecasts. Probabilistic forecasts can also be generated. However, as in this thesis, only a deterministic forecast was produced; it is not possible to assess how the two methods compare. To further expand this research, it would be interesting to generate forecasts using both the ensemble generator from STEPS and the rainfall advection fields ensemble generator proposed in Chapter 5. A probabilistic forecast that combines KED rainfall estimations and rainfall advection field ensembles has the potential to produce more reliable forecasts for a longer lead time. In an analogue way, the RE ensembles could also be combined with the temporal evolution ensembles. Combining the two methods would also have the potential to improve the forecasts for a longer time. However, the research should also include an analysis of the number of ensemble members to be used for each method. In this thesis, it was produced 25 RE ensembles, and it is not realistic to keep the same number of ensembles in addition to the temporal evolution ensembles.

The merging was done using KED due to its robustness and performance compared with other techniques (Ochoa-Rodríguez *et al.*, 2019). Nonetheless, it is highly dependable on the rain gauge network density (Goudenhoofdt and Delobbe, 2009; Jewell and Gaussiat, 2015). However, in many

places rain gauge network is not dense enough to produce accurate merged results, limiting the application of this technique. Satellite technology advances mean that temporal and spatial resolution are increasing, and data is available even for ungauged catchments. Using satellite and radar data, or satellite, rain gauge, and radar data is an alternative that should be studied due to its particular importance in places where there is a lack of ground-based rainfall measurements. Regarding the KEDFV, research in assessing and quantifying the uncertainties propagation into hydrological models was not carried out. For urban areas, a higher temporal resolution is usually required, but for river catchments rainfall forecasts show that it has the potential to improve flow forecasts.

The study case for urban application in Ilkley was able to provide a good insight into the error propagation of rainfall advection field and RE ensembles; however, the size of the catchment is a plausible explanation for extra uncertainties in the flow prediction. If forecasted rainfall is only a couple of pixels displaced, the amount of precipitation that reaches the ground can be very different from the forecast. Further research using a larger urban area would be important to assess how the catchment size affects flow forecasts.

Blending nowcast with NWP has been helpful to produce skilful forecasts for a longer period. Using these ensemble generators with the NWP model has the potential to increase the forecast skill from the start of the forecast and for longer lead times.

The dataset used in this thesis is from 2007 and 2008. Since then, the MetOffice weather radar network has been renewed, and the single polarization radars have been upgraded to dual-polarization radars. Therefore, the accuracy of radar-based rainfall estimations has improved. During the same period, little improvement has been carried in the rain gauge network, which is still heavily based in TBR. It is essential to point out that as rainfall intensities in KED estimations are under a more substantial influence from rain gauges than from radars, and it is difficult to know how much KED estimations will benefit from more accurate radar data before further research is carried. However, as the rainfall advection field ensembles use only radar data, the quality of the probabilistic forecasts using this method should increase, mainly in up to 1 hour lead time.

Chapter 9. References

Aghakouchak, A. *et al.* (2012) 'Systematic and random error components in satellite precipitation data sets', *Geophysical Research Letters*, 39, pp. 3–6. doi: 10.1029/2012GL051592.

Allen, R. G. *et al.* (1998) 'Crop evapotranspiration : Guidelines for computing crop water requirements', *FAO irrigation and drainage paper*, 56, pp. 1–15. doi: 10.1016/j.eja.2010.12.001.

Anagnostou, E. N., Morales, C. A. and Dinku, T. (2001) 'The use of TRMM precipitation radar observations in determining ground radar calibration biases', *Journal of Atmospheric and Oceanic Technology*, 18(4), pp. 616–628. doi: 10.1175/1520-0426(2001)018<0616:TUOTPR>2.0.CO;2.

Araghinejad, S. and Burn, D. H. (2005) 'Probabilistic forecasting of hydrological events using geostatistical analysis', *Hydrological Sciences Journal*, 50(5), pp. 837–856. doi: 10.1623/hysj.2005.50.5.837.

Arnell, N. W. (1999) 'A simple water balance model for the simulation of streamflow over a large geographic domain', *Journal of Hydrology*, 217(3–4), pp. 314–335. doi: 10.1016/S0022-1694(99)00023-2.

Arnell, N. W. (2011) 'Uncertainty in the relationship between climate forcing and hydrological response in UK catchments', *Hydrology and Earth System Sciences*, 15(3), pp. 897–912. doi: 10.5194/hess-15-897-2011.

Arnell, N. W., Charlton, M. B. and Lowe, J. A. (2014) 'The effect of climate policy on the impacts of climate change on river flows in the UK', *Journal of Hydrology*. Elsevier B.V., 510(2014), pp. 424–435. doi: 10.1016/j.jhydrol.2013.12.046.

Atencia, A. and Zawadzki, I. (2014) 'A comparison of two techniques for generating nowcasting ensembles. Part I: Lagrangian ensemble technique', *Monthly Weather Review*, 142(11), pp. 4036–4052. doi: 10.1175/MWR-D-13-00117.1.

Atencia, A. and Zawadzki, I. (2015) 'A comparison of two techniques for generating nowcasting ensembles. Part II: Analogs selection and comparison of techniques', *Monthly Weather Review*, 143(7), pp. 2890–2908. doi: 10.1175/MWR-D-14-00342.1.

Atlas, D. (2002) 'Radar Calibration: Some Simple Approaches', *Bulletin of the American Meteorological*

Society, 83(September), pp. 1313–1316. doi: 10.1175/1520-0477-83.9.1313.

Atlas, D. and Chmela, A. C. (1957) 'Physical-synoptic variations of raindrop size parameters', in *Sixth Weather Radar Conference*. Cambridge, MA: American Meteorological Society, pp. 21–30.

Atlas, D. and Ulbrich, C. W. (1977) 'Path- and Area-Integrated Rainfall Measurement by Microwave Attenuation in the 1–3 cm Band', *Journal of Applied Meteorology*, 16, pp. 1322–1331. doi: 10.1175/1520-0450(1977)016<1322:PAAIRM>2.0.CO;2.

Bech, J. *et al.* (2003) 'The sensitivity of single polarization weather radar beam blockage correction to variability in the vertical refractivity gradient', *Journal of Atmospheric and Oceanic Technology*, 20(6), pp. 845–855. doi: 10.1175/1520-0426(2003)020<0845:TSOSPW>2.0.CO;2.

Bech, J., Gjertsen, U. and Haase, G. (2007) 'Modelling weather radar beam propagation and topographical blockage at northern high latitudes', *Quarterly Journal of the Royal Meteorological Society*, 133(626), pp. 1191–1204.

Berenguer, M. *et al.* (2006) 'A Fuzzy Logic Technique for Identifying Nonprecipitating Echoes in Radar Scans', *Journal of Atmospheric and Oceanic Technology*, 23, pp. 1157–1180.

Berenguer, M., Sempere-Torres, D. and Pegram, G. G. S. (2011) 'SBMcast - An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation', *Journal of Hydrology*. Elsevier B.V., 404(3–4), pp. 226–240. doi: 10.1016/j.jhydrol.2011.04.033.

Berndt, C., Rabiei, E. and Haberlandt, U. (2014) 'Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios', *Journal of Hydrology*. Elsevier B.V., 508, pp. 88–101. doi: 10.1016/j.jhydrol.2013.10.028.

Berne, A. *et al.* (2004) 'Temporal and spatial resolution of rainfall measurements required for urban hydrology', *Journal of Hydrology*, 299(3–4), pp. 166–179. doi: 10.1016/S0022-1694(04)00363-4.

Berne, A. and Krajewski, W. F. (2013) 'Radar for hydrology: Unfulfilled promise or unrecognized potential?', *Advances in Water Resources*. Elsevier Ltd, 51, pp. 357–366. doi: 10.1016/j.advwatres.2012.05.005.

Borga, M. (2002) 'Accuracy of radar rainfall estimates for streamflow simulation', *Journal of Hydrology*, 267, pp. 26–39.

Bowes, M. J. *et al.* (2005) 'Phosphorus-discharge hysteresis during storm events along a river catchment: The River Swale, UK', *Water Research*, 39(5), pp. 751–762. doi: 10.1016/j.watres.2004.11.027.

Bowes, M. J., House, W. A. and Hodgkinson, R. A. (2003) 'Phosphorus dynamics along a river continuum', *Science of the Total Environment*, 313(1–3), pp. 199–212. doi: 10.1016/S0048-9697(03)00260-2.

Bowler, N., Pierce, C. E. and Seed, A. W. (2004) 'Development of a precipitation nowcasting algorithm based upon optical flow techniques', *Journal of Hydrology*, 288(1–2), pp. 74–91. doi: 10.1016/j.jhydrol.2003.11.011.

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

Brandes, E. A. (1975) 'Optimizing Rainfall Estimates with the Aid of Radar', *Journal of Applied Meteorology*, pp. 1339–1345. doi: 10.1175/1520-0450(1975)014<1339:orewta>2.0.co;2.

Bray, M. *et al.* (2011) 'Rainfall uncertainty for extreme events in NWP downscaling model', *Hydrological Processes*, 25(9), pp. 1397–1406. doi: 10.1002/hyp.7905.

Bringi, V. N. and Chandrasekar, V. (2001) *Polarimetric Doppler Weather Radar - Principles and Applications*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511541094.

Bringi, V. N., Keenan, T. D. and Chandrasekar, V. (2001) 'Correcting C-band radar reflectivity and differential reflectivity data for rain attenuation: A self-consistent method with constraints', *IEEE Transactions on Geoscience and Remote Sensing*, 39(9), pp. 1906–1915. doi: 10.1109/36.951081.

Bringi, V. N., Rico-Ramirez, M. A. and Thurai, M. (2011) 'Rainfall Estimation with an Operational Polarimetric C-Band Radar in the United Kingdom: Comparison with a Gauge Network and Error Analysis', *Journal of Hydrometeorology*, 12(5), pp. 935–954. doi: 10.1175/JHM-D-10-05013.1.

Bringi, V. N. and Thurai, M. (2012) 'River flow simulations with polarimetric weather radar', in *Weather Radar and Hydrology (Proceedings of a symposium held in Exeter, UK, April 2011)*. IAHS, p. 2012.

Brocca, L., Melone, F. and Moramarco, T. (2011) 'Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting', *Hydrological Processes*, 25, pp. 2801–2813.

Burrough, P. A., McDonnell, R. A. and Lloyd, C. D. (2015) *Principles of geographical information systems*. Oxford: Oxford University Press.

Burt, S. (2013) 'An unsung hero in meteorology: Charles Higman Griffith (1830-1896)', *Weather*, 68(5), pp. 135–138. doi: 10.1002/wea.2059.

Cabus, P. (2008) 'River flow prediction through rainfall-runoff modelling with a probability-distributed model (PDM) in Flanders, Belgium', *Agricultural Water Management*, 95(7), pp. 859–868. doi: 10.1016/j.agwat.2008.02.013.

Caseri, A. *et al.* (2016) 'A space-time geostatistical approach for ensemble rainfall nowcasting', *E3S Web of Conferences*, 7, pp. 2–6. doi: 10.1051/e3sconf/20160718001.

Cecinati, F. *et al.* (2017) 'Representing radar rainfall uncertainty with ensembles based on a time-variant geostatistical error modelling approach', *Journal of Hydrology*. The Authors, 548, pp. 391–405. doi: 10.1016/j.jhydrol.2017.02.053.

Cecinati, F. (2017) *Uncertainty estimation and propagation in radar-rain gauge rainfall merging using kriging-based techniques*. University of Bristol.

Chen, J. *et al.* (2013) 'Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America', *Water Resources Research*, 49(7), pp. 4187–4205. doi: 10.1002/wrcr.20331.

Cho, Y. H. *et al.* (2006) 'Identification and removal of ground echoes and anomalous propagation using the characteristics of radar echoes', *Journal of Atmospheric and Oceanic Technology*, 23(9), pp. 1206–1222. doi: 10.1175/JTECH1913.1.

Ciach, G. J. and Krajewski, W. F. (1999) 'On the estimation of radar rainfall error variance', *Advances in Water Resources*, 22(6), pp. 585–595. doi: 10.1016/S0309-1708(98)00043-8.

Cloke, H. L. and Pappenberger, F. (2009) 'Ensemble flood forecasting: A review', *Journal of Hydrology*. Elsevier B.V., 375(3–4), pp. 613–626. doi: 10.1016/j.jhydrol.2009.06.005.

Cluckie, I. D. and Rico-Ramirez, M. A. (2004) 'Weather radar technology and future developments', in

GIS and Remote Sensing in Hydrology, Water Resources and Environment t (Proceedings of ICGRHWE held at the Three Gorges Dam, China, September 2003). Three Gorges Dam, China: IAHS, pp. 11–20.

Cole, S. J. and Moore, R. J. (2008) 'Hydrological modelling using raingauge- and radar-based estimators of areal rainfall', *Journal of Hydrology*, 358(3–4), pp. 159–181. doi: 10.1016/j.jhydrol.2008.05.025.

Collier, C. G. (1996) *Applications of weather radar systems: a guide to uses of radar data in meteorology and hydrology*. 2nd edn, *Wiley-Praxis series in atmospheric physics*. 2nd edn. Chichester: John Wiley & Sons, Incorporated & Praxis.

Convery, I. and Bailey, C. (2008) 'After the flood: the health and social consequences of the 2005 Carlisle flood event', *Journal of Flood Risk Management*, 1(2), pp. 100–109. doi: 10.1111/j.1753-318x.2008.00012.x.

Cressie, N. (1990) 'The origins of kriging', *Mathematical Geology*, 22(3), pp. 239–252. doi: 10.1007/BF00889887.

Creutin, J. D., Delrieu, G. and Lebel, T. (1988) 'Rain measurement by raingage-radar combination: A geostatistical approach', *Journal of Atmospheric and Oceanic Technology*, pp. 102–115.

Dai, Q. *et al.* (2013) 'The impact of raindrop drift in a three-dimensional wind field on a radar-gauge rainfall comparison', *International Journal of Remote Sensing*. Taylor & Francis, 34(21), pp. 7739–7760. doi: 10.1080/01431161.2013.826838.

Dai, Q. *et al.* (2015) 'Probabilistic radar rainfall nowcasts using empirical and theoretical uncertainty models', *Hydrological Processes*, 29(1), pp. 66–79. doi: 10.1002/hyp.10133.

Dance, S. L. *et al.* (2019) 'Improvements in forecasting intense rainfall: Results from the FRANC (Forecasting Rainfall exploiting new data Assimilation techniques and Novel observations of Convection) project', *Atmosphere*, 10(3). doi: 10.3390/atmos10030125.

Delrieu, G., Caoudal, S. and Creutin, J. D. (1997) 'Feasibility of using mountain return for the correction of ground-based X-band weather radar data', *Journal of Atmospheric and Oceanic Technology*, 14(3), pp. 368–385. doi: 10.1175/1520-0426(1997)014<0368:FOUMRF>2.0.CO;2.

Dixon, H., Hannaford, J. and Fry, M. J. (2013) 'The effective management of national hydrometric

data: experiences from the United Kingdom', *Hydrological Sciences Journal*. Taylor & Francis, 58(7), pp. 1383–1399. doi: 10.1080/02626667.2013.787486.

Dixon, M. and Wiener, G. (1993) 'TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology', *Journal of Atmospheric and Oceanic Technology*, 10(6), pp. 785–79.

Duan, Q., Sorooshian, S. and Gupta, V. (1992) 'Effective and efficient global optimization for conceptual rainfall-runoff models', *Water Resources Research*, 28(4), pp. 1015–1031. doi: 10.1029/91WR02985.

Duan, Q., Sorooshian, S. and Gupta, V. (1994) 'Optimal use of the SCE-UA global optimization method for calibrating watershed models', *Journal of Hydrology*, 158(3–4), pp. 265–284. doi: 10.1016/0022-1694(94)90057-4.

Duan, Q. Y., Gupta, V. K. and Sorooshian, S. (1993) 'Shuffled complex evolution approach for effective and efficient global minimization', *Journal of Optimization Theory and Applications*, 76(3), pp. 501–521. doi: 10.1007/BF00939380.

Duchon, C. E. and Essenberg, G. R. (2001) 'Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields', *Water Resources Research*, 37(12), pp. 3253–3263. doi: 10.1029/2001WR000541.

Dufton, D. R. L. and Collier, C. G. (2015) 'Fuzzy logic filtering of radar reflectivity to remove non-meteorological echoes using dual polarization radar moments', *Atmospheric Measurement Techniques*, 8(10), pp. 3985–4000. doi: 10.5194/amt-8-3985-2015.

Duthoit, S. *et al.* (2017) 'A new approach for in-situ antenna characterization, radome inspection and radar calibration, using an Unmanned Aircraft System (UAS)', *2017 IEEE Radar Conference, RadarConf 2017*, pp. 0669–0674. doi: 10.1109/RADAR.2017.7944287.

Ebert, E. E. *et al.* (2004) 'Verification of nowcasts from the WWRP Sydney 2000 Forecast Demonstration Project', *Weather and Forecasting*, 19(1), pp. 73–96. doi: 10.1175/1520-0434(2004)019<0073:VONFTW>2.0.CO;2.

Ferket, B. V. A., Samain, B. and Pauwels, V. R. N. (2010) 'Internal validation of conceptual rainfall-runoff models using baseflow separation', *Journal of Hydrology*. Elsevier B.V., 381(1–2), pp. 158–173.

doi: 10.1016/j.jhydrol.2009.11.038.

Few, R. (2003) 'Flooding, vulnerability and coping strategies: local responses to a global threat', *Progress in Development Studies*, 3(1), pp. 43–58. doi: 10.1191/1464993403ps049ra.

Foresti, L. *et al.* (2013) 'Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall', *Meteorological Applications*, 22(2), pp. 141–155. doi: 10.1002/met.1416.

Foresti, L. *et al.* (2016) 'Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium', *Hydrology and Earth System Sciences*, 20(1), pp. 505–527. doi: 10.5194/hess-20-505-2016.

Foresti, L. *et al.* (2018) 'Non-stationary radar precipitation ensembles: a stochastic nested generator and orographic precipitation growth and decay trends', *Geophysical Research Abstracts*, 20, p. 6739.

Foresti, L. and Pozdnoukhov, A. (2012) 'Exploration of alpine orographic precipitation patterns with radar image processing and clustering techniques', *Meteorological Applications*, 19(4), pp. 407–419. doi: 10.1002/met.272.

Foresti, L. and Seed, A. W. (2014) 'The effect of flow and orography on the spatial distribution of the very short-term predictability of rainfall from composite radar images', *Hydrology and Earth System Sciences*, 18(11), pp. 4671–4686. doi: 10.5194/hess-18-4671-2014.

Foresti, L. and Seed, A. W. (2015) 'On the spatial distribution of rainfall nowcasting errors due to orographic forcing', *Meteorological Applications*, 22(1), pp. 60–74. doi: 10.1002/met.1440.

Fujiwara, M. (1965) 'Raindrop-size distribution from individual storms', *Journal of Atmospheric Sciences*, 22, pp. 585–591.

Fulton, R. A. *et al.* (1998) 'The WSR-88D Rainfall Algorithm', *Weather and Forecasting*, 13(2), pp. 377–395. doi: 10.1175/1520-0434(1998)013<0377:TWRA>2.0.CO;2.

Germann, U. *et al.* (2006) 'Radar precipitation measurement in a mountainous region', *Quarterly Journal of the Royal Meteorological Society*, 132(618), pp. 1669–1692. doi: 10.1256/qj.05.190.

Germann, U. *et al.* (2009) 'REAL - Ensemble radar precipitation estimation for hydrology in a mountainous region', *Quarterly Journal of the Royal Meteorological Society*, 135(639), pp. 445–456.

Giuli, D. *et al.* (1991) 'Rainfall and clutter discrimination by means of dual-linear polarization radar

measurements', *Journal of Atmospheric and Oceanic Technology*, pp. 777–789. doi: 10.1175/1520-0426(1991)008.

Golding, B. W. *et al.* (1998) 'Nimrod : A system for generating automated very short range forecasts', 16, pp. 1–16.

Golding, B. W. (2009) 'Long lead time flood warnings: reality or fantasy?', *Meteorological Applications*, 16, pp. 3–12.

Gorgucci, E., Scarchill, G. and Chandrasekar, V. (1996) 'Error structure of radar rainfall measurement at C-band frequencies with dual polarization algorithm for attenuation correction', *Journal of Geophysical Research-Atmospheres*, 101(D21), pp. 26461–26471. doi: 10.1016/B978-0-12-397933-9.00002-3.

Goudenhoofdt, E. and Delobbe, L. (2009) 'Evaluation of radar-gauge merging methods for quantitative precipitation estimates', *Hydrology and Earth System Sciences*, 13(2), pp. 195–203. doi: 10.5194/hess-13-195-2009.

Gourley, J. J., Kaney, B. and Maddox, R. A. (2003) 'Evaluating the calibration of radars: a software approach', in *31st International Conference on Radar Meteorology*, pp. 459–462. doi: 10.1029/2002JD002184.Woo.

Gourley, J. J., Tabary, P. and Parent du Chatelet, J. (2007) 'A fuzzy logic algorithm for the separation of precipitating from nonprecipitating echoes using polarimetric radar observations', *Journal of Atmospheric and Oceanic Technology*, 24(8), pp. 1439–1451. doi: 10.1175/JTECH2035.1.

Gray, W. R., Uddstrom, M. J. and Larsen, H. R. (2002) 'Radar surface rainfall estimates using a typical shape function approach to correct for the variations in the vertical profile of reflectivity', *International Journal of Remote Sensing*, 23(12), pp. 2489–2504. doi: 10.1080/01431160110070834.

Grecu, M. and Krajewski, W. F. (2000) 'An efficient methodology for detection of anomalous propagation echoes in radar reflectivity data using neural networks', *Journal of Atmospheric and Oceanic Technology*, 17(2), pp. 121–129. doi: 10.1175/1520-0426(2000)017<0121:AEMFDO>2.0.CO;2.

Grimes, D. I. F., Pardo-Igúzquiza, E. and Bonifacio, R. (1999) 'Optimal areal rainfall estimation using raingauges and satellite data', *Journal of Hydrology*, 222(1–4), pp. 93–108. doi: 10.1016/S0022-162

1694(99)00092-X.

Haberlandt, U. (2007) 'Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event', *Journal of Hydrology*, 332(1–2), pp. 144–157. doi: 10.1016/j.jhydrol.2006.06.028.

Habib, E., Krajewski, W. F. and Kruger, A. (2001) 'Sampling Errors of Tipping-Bucket Rain Gauge Measurements', *Journal of Hydrologic Engineering*, 6(2), pp. 159–166. doi: 10.1061/(asce)1084-0699(2001)6:2(159).

Haines, A. *et al.* (2006) 'Climate change and human health: Impacts, vulnerability and mitigation', *Public Health*, 120(7), pp. 2101–2109. doi: 10.1016/j.puhe.2006.01.002.

Harrison, D. L., Driscoll, S. J. and Kitchen, M. (2000) 'Improving precipitation estimates from weather radar using quality control and correction techniques', *Meteorological Applications*, 7(2), pp. 135–144.

Harrison, D. L., Kitchen, M. and Scovell, R. W. (2009) 'High-resolution precipitation estimates for hydrological uses', *Proceedings of the ICE - Water Management*, 162(2), pp. 125–135. doi: 10.1680/wama.2009.162.2.125.

Hasan, M. M. *et al.* (2014) 'Correcting bias in radar Z-R relationships due to uncertainty in point rain gauge networks', *Journal of Hydrology*. Elsevier B.V., 519(PB), pp. 1668–1676. doi: 10.1016/j.jhydrol.2014.09.060.

Hou, A. Y. *et al.* (2014) 'The global precipitation measurement mission', *Bulletin of the American Meteorological Society*, 95, pp. 701–722. doi: 10.1175/BAMS-D-13-00164.1.

Humphrey, M. D. *et al.* (1997) 'A New Method for Automated Dynamic Calibration of Tipping-Bucket Rain Gauges', *Journal of Atmospheric and Oceanic Technology*, 14(6), pp. 1513–1519. doi: 10.1175/1520-0426(1997)014<1513:anmfad>2.0.co;2.

Hunter, S. M. (1996) 'WSR-88D Radar Rainfall Estimation: Capabilities, Limitations and Potential Improvements', *National Weather Service Office*, 20, pp. 26–36. Available at: http://www.srh.noaa.gov/mrx/research/precip/precip.php.

Islam, T. *et al.* (2014) 'Sensitivity associated with bright band/melting layer location on radar

reflectivity correction for attenuation at C-band using differential propagation phase measurements', *Atmospheric Research*. Elsevier B.V., 135–136, pp. 143–158. doi: 10.1016/j.atmosres.2013.09.003.

Jaffrain, J. and Berne, A. (2011) 'Experimental Quantification of the Sampling Uncertainty Associated with Measurements from PARSIVEL Disdrometers', *Journal of Hydrometeorology*, 12, pp. 352–370. doi: 10.1175/2010JHM1244.1.

Jarraud, M. (2010) *Guide to meteorological instruments and methods of observation*. 2nd edn, *World Meteorological Organisation: Geneva, Switzerland*. 2nd edn. Geneva: World Meteorological Organization. doi: Guide to meteorological instrument and observing practices.

Jewell, S. A. and Gaussiat, N. (2015) 'An assessment of kriging-based rain-gauge-radar merging techniques', *Quarterly Journal of the Royal Meteorological Society*, 141(691), pp. 2300–2313. doi: 10.1002/qj.2522.

Johnson, J. T. *et al.* (1998) 'The Storm Cell Identification and Tracking Algorithm : An Enhanced WSR-88D Algorithm', *Weather and Forecasting*, 13, pp. 263–276.

Jolliffe, I. T. and Stephenson, D. B. (2011) *Forecast Verification : A Practitioner's Guide in Atmospheric Science*. 2nd edn. Edited by D. B. Stephenson. Chichester: John Wiley & Sons, Incorporated.

Jones, D. M. A. (1956) *Rainfall drop-size distribution and radar reflectivity*, *Illinois State Water Survey - Laboratory*. Urbana, IL.

Joss, J. and Waldvogel, A. (1970) 'A method to improve the accuracy of radar measured amounts of precipitation', in *14th Conference on Radar Meteorology*. Tucson, AZ,: American Meteorological Society, pp. 237–238.

Kato, R. *et al.* (2017) 'Predictability of meso-γ-scale, localized, extreme heavy rainfall during the warm season in Japan using high-resolution precipitation nowcasts', *Quarterly Journal of the Royal Meteorological Society*, 143(704), pp. 1406–1420. doi: 10.1002/qj.3013.

Kay, A. L. *et al.* (2009) 'Comparison of uncertainty sources for climate change impacts: Flood frequency in England', *Climatic Change*, 92(1–2), pp. 41–63. doi: 10.1007/s10584-008-9471-4.

Keller, D. (2013) 'Evaluation and comparison of radar-rain gauge combination methods', *Scientific Report MeteoSwiss*, 94(94), p. 84 pp.

Kessinger, C., Ellis, S. and Van Andel, J. (1999) 'A fuzzy logic, radar echo classification scheme for the WSR-88D', in *29th International Conference on Radar Meteorology*. Montreal,QB: American Meteorological Society, pp. 576–579.

Kharin, V. V. and Zwiers, F. W. (2003) 'On the ROC score of probability forecasts', *Journal of Climate*, 16(24), pp. 4145–4150. doi: 10.1175/1520-0442(2003)016<4145:OTRSOP>2.0.CO;2.

Kitchen, M. and Blackall, R. M. (1992) 'Representativeness errors in comparisons between radar and gauge measurements of rainfall', *Journal of Hydrology*, 134(1–4), pp. 13–33. doi: 10.1016/0022-1694(92)90026-R.

Krajewski, W. F. (1987) 'Cokriging radar- rainfall and rain gage data', *Journal of Geophysical Research*, 92(D8), pp. 9571–9580. doi: 10.1029/JD092iD08p09571.

Krajewski, W. F. *et al.* (2003) *Towards probabilistic quantitative precipitation WSR-88D algorithms: Preliminary studies and problem formulation*, *Final Report, National Oceanic and Atmospheric Administration - National Weather Service*. Iowa City.

Krajewski, W. F. *et al.* (2011) 'Statistical model of the range-dependent error in radar-rainfall estimates due to the vertical profile of reflectivity', *Journal of Hydrology*. Elsevier B.V., 402(3–4), pp. 306–316. doi: 10.1016/j.jhydrol.2011.03.024.

Krige, D. G. (1951) 'A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand', *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52(6), pp. 201–215. doi: 10.2307/3006914.

Kummerow, C. *et al.* (2000) 'The Status of the Tropical Rainfall Measuring Mission (TRMM) after Two Years in Orbit', *Journal of Applied Meteorology*, 39, pp. 1965–1982.

Kurtyka, J. C. (1953) *Precipitation Measurements Study: Methods of Measuring Precipitation for Use with the Automatic Weather Station*. Urbana, IL.

Laio, F. and Tamea, S. (2007) 'Verification tools for probabilistic forecasts of continuous hydrological variables', *Hydrology and Earth System Sciences*, 11(4), pp. 1267–1277. doi: 10.5194/hessd-3-2145-2006.

Lau, C. L. *et al.* (2010) 'Climate change, flooding, urbanisation and leptospirosis: Fuelling the fire?',

Transactions of the Royal Society of Tropical Medicine and Hygiene. Royal Society of Tropical Medicine and Hygiene, 104(10), pp. 631–638. doi: 10.1016/j.trstmh.2010.07.002.

Leblois, E. and Creutin, J. D. (2013) 'Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method', *Water Resources Research*, 49(6), pp. 3375–3387. doi: 10.1002/wrcr.20190.

Lee, J.-K., Kim, J.-H. and Suk, M.-K. (2015) 'Application of bias correction methods to improve the accuracy of quantitative radar rainfall in Korea', *Atmospheric Measurement Techniques Discussions*, 8(4), pp. 4011–4047. doi: 10.5194/amtd-8-4011-2015.

Lengfeld, K. *et al.* (2016) 'A simple method for attenuation correction in local X-band radar measurements using C-band radar data', *Journal of Atmospheric and Oceanic Technology*, 33(11), pp. 2315–2329. doi: 10.1175/JTECH-D-15-0091.1.

Lewis, E. *et al.* (2018) 'A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr', *Journal of Hydrology*. Elsevier, 564, pp. 930–943. doi: 10.1016/j.jhydrol.2018.07.034.

Lewis, H. *et al.* (2015) 'From months to minutes - exploring the value of high-resolution rainfall observation and prediction during the UK winter storms of 2013/2014', *Meteorological Applications*, 22(1), pp. 90–104. doi: 10.1002/met.1493.

Li, J. and Heap, A. D. (2011) 'A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors', *Ecological Informatics*. Elsevier B.V., 6(3–4), pp. 228–241. doi: 10.1016/j.ecoinf.2010.12.003.

Li, L., Schmid, W. and Joss, J. (1995) 'Nowcasting of Motion and Growth of Precipitation with Radar over a Complex Orography', *Journal of Applied Meteorology*, 34, pp. 1286–1300.

Li, Y. *et al.* (2011) 'Assimilation of streamflow discharge into a continuous flood forecasting model', in *Risk in Water Resources Management (Proceedings of Symposium H03 held during IUGG2011 in Melbourne, Australia, July 2011)*. Melbourne: IAHS Publication, pp. 107–113.

Liguori, S. *et al.* (2012) 'Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments', *Atmospheric Research*. Elsevier B.V., 103, pp. 80–95. doi: 10.1016/j.atmosres.2011.05.004.

Liguori, S. and Rico-Ramirez, M. A. (2012a) 'Probabilistic forecasting of rainfall from radar nowcasting and hybrid systems', in *Weather Radar and Hydrology (Proceedings of a symposium held in Exeter, UK, April 2011)*. Exeter, UK: IAHS Publication.

Liguori, S. and Rico-Ramirez, M. A. (2012b) 'Quantitative assessment of short-term rainfall forecasts from radar nowcasts and MM5 forecasts', *Hydrological Processes*, 26(25), pp. 3842–3857. doi: 10.1002/hyp.8415.

Liguori, S. and Rico-Ramirez, M. A. (2012c) 'Using the radar rainfall error for probabilistic forecasting', in *The Seventh European Conference on Radar in Meteorology and Hydrology*. Toulouse, France.

Liguori, S. and Rico-Ramirez, M. A. (2013a) 'A practical approach to the assessment of probabilistic flow predictions', *Hydrological Processes*, 27(1), pp. 18–32. doi: 10.1002/hyp.9468.

Liguori, S. and Rico-Ramirez, M. A. (2013b) 'A review of current approaches to radar-based quantitative precipitation forecasts', *International Journal of River Basin Management*, 12(4), pp. 391–402. doi: 10.1080/15715124.2013.848872.

Liguori, S., Rico-Ramirez, M. A. and Cluckie, I. D. (2009) 'Uncertainty propagation in hydrological forecasting using ensemble rainfall forecasts', in *Hydroinformatics in Hydrology, Hydrogeology and Water Resources (Proc. of Symposium JS.4 at the Joint IAHS & IAH Convention, September 2009)*. Hyderabad, India: IAHS Publication, pp. 30–40.

Marsh, T. J. (2002) 'Capitalising on river flow data to meet changing national needs - A UK perspective', *Flow Measurement and Instrumentation*, 13(5–6), pp. 291–298. doi: 10.1016/S0955-5986(02)00056-0.

Marsh, T. J. and Lees, M. L. L. (2003) *Hydrometric Register and Statistics 1996-2000*. Wallingford.

Marshall, J. S., Hitschfeld, W. and Gunn, K. L. S. (1955) 'Advances in radar weather', *Advances in Geophysics*, 2, pp. 1–56. doi: 10.1016/S0065-2687(08)60310-6.

Marshall, J. S., Langille, R. C. and Palmer, W. M. (1947) 'Measurement of Rainfall by Radar', *Journal of Meteorology*, 4, pp. 186–192.

Marshall, J. S. and Palmer, W. M. K. (1948) 'The Distribution of Raindrops with Size', *Journal of Meteorology*, 5, pp. 165–166. doi: 10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2.

Mason, S. J. and Graham, N. E. (2002) 'Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation', *Quarterly Journal of the Royal Meteorological Society*, 128(584), pp. 2145–2166. doi: 10.1256/003590002320603584.

McMillan, H. *et al.* (2011) 'Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models', *Journal of Hydrology*. Elsevier B.V., 400(1–2), pp. 83–94. doi: 10.1016/j.jhydrol.2011.01.026.

Met Office (2003) *1 km Resolution UK Composite Rainfall Data from the Met Office Nimrod System [WWW Document]. NCAS British Atmospheric Data Centre*. Available at: http://catalogue.ceda.ac.uk/uuid/27dd6ffba67f667a18c62de5c3456350 (Accessed: 3 April 2017).

Met Office (2009) *National meteorological library and archive: Fact sheet 15 - radar*. Available at: https://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/j/h/fact_sheet_no._15.pdf (Accessed: 20 January 2019).

Met Office (2012) *Weather radar network renewal - Met Office*. Available at: https://www.metoffice.gov.uk/services/business-industry/water/radar-improvements (Accessed: 12 February 2019).

Met Office (2012) (2014) *Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current)*, *NCAS British Atmospheric Data Centre*. Available at: http://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0 (Accessed: 16 March 2019).

Moore, R. J. (1985) 'The probability-distributed principle and runoff production at point and basin scales', *Hydrological Sciences Journal*, 30(2), pp. 273–297. doi: 10.1080/02626668509490989.

Moore, R. J. (2007) 'The PDM rainfall-runoff model', *Hydrology and Earth System Sciences*, 11(1), pp. 483–499. doi: 10.5194/hess-11-483-2007.

Moore, R. J. and Bell, V. A. (2002) 'Incorporation of groundwater losses and well level data in rainfall-runoff models illustrated using the PDM', *Hydrology and Earth System Sciences*, 6(1), pp. 25–38. doi: 10.5194/hess-6-25-2002.

Nanding, N. (2016) *Hydrological applications of radar-raingauge rainfall merging techniques*. University of Bristol, Bristol.

Nanding, N., Rico-Ramirez, M. A. and Han, D. (2015) 'Comparison of different radar-raingauge rainfall merging techniques', *Journal of Hydroinformatics*, 17(3), p. 422. doi: 10.2166/hydro.2015.001.

Nash, J. E. and Sutcliffe, J. V (1970) 'River flow Forecasting through Conceptual Models. Part I - A Discussion of Principles', *Journal of Hydrology*, 10, pp. 282–290. doi: 10.1016/0022-1694(70)90255-6.

Netatmo Rain Gauge | Weather Station Tools (2013). Available at: https://www.scientificsales.com/Netatmo-Rain-Gauge-for-Netatmo-Weather-Station-p/nrg01-ww.htm (Accessed: 26 May 2019).

Niemczynowicz, J. (1986) 'The Dynamic Calibration of Tipping-Bucket Raingauges', *Hydrology Research*, 17(3), pp. 203–214. doi: 10.2166/nh.1986.0013.

NRFA (2014) *UK wide flood daily and peak data*. Available at: https://nrfa.ceh.ac.uk/data/search (Accessed: 7 March 2019).

Nystuen, J. A. *et al.* (1996) 'A Comparison of Automatic Rain Gauges', *Journal of Atmospheric and Oceanic Technology*, 13, pp. 62–73. doi: 10.1175/1520-0426(1996)013<0062:ACOARG>2.0.CO;2.

Ochoa-Rodríguez, S. *et al.* (2013) 'Improving rainfall nowcasting and urban runoff forecasting through dynamic radar-raingauge rainfall adjustment', in *7th International Conference on Sewer Processes & Networks*.

Ochoa-Rodríguez, S. *et al.* (2019) 'A Review of Radar-Rain Gauge Data Merging Methods and Their Potential for Urban Hydrological Applications', *Water Resources Research*, 55(8), pp. 6356–6391. doi: 10.1029/2018WR023332.

Ordnance Survey. (2018) *A Guide to Coordinate Systems in Great Britain V3.2*. Available at: https://www.ordnancesurvey.co.uk/docs/support/guide-coordinate-systems-great-britain.pdf (Accessed: 5 July 2018).

Orr, H. G. and Carling, P. A. (2006) 'Hydro-climatic and land use changes in the river lune catchment, North West England, implications for catchment management', *River Research and Applications*, 22(2), pp. 239–255. doi: 10.1002/rra.908.

Overeem, A., Leijnse, H. and Uijlenhoet, R. (2013) 'Country-wide rainfall maps from cellular communication networks', *Proceedings of the National Academy of Sciences of the United States of*

America, 110(8), pp. 2741–2745. doi: 10.1073/pnas.1217961110.

Panziera, L. *et al.* (2011) 'NORA-Nowcasting of Orographic Rainfall by means of analogues', *Quarterly Journal of the Royal Meteorological Society*, 137(661), pp. 2106–2123. doi: 10.1002/qj.878.

Pegram, G. G. S. (2004) *Spatial Interpolation and Mapping of Rainfall (SIMAR), vol. 3: Data Merging for Rainfall Map Production. WRC Report no. 1153/1/04*. Pretoria, RSA.

Pellarin, T. *et al.* (2002) 'Hydrologic visibility of weather radar systems operating in mountainous regions: Case study for the Ardeche Catchment (France)', *Journal of Hydrometeorology*, 3(5), pp. 539–555. doi: 10.1175/1525-7541(2002)003<0539:hvowrs>2.0.co;2.

Pierce, C. E. *et al.* (2000) 'GANDOLF: a system for generating automated nowcasts of convective precipitation', *Meteorological Applications*, 7(4), pp. 341–360. doi: Doi 10.1017/S135048270000164x.

Pierce, C. E. *et al.* (2004) 'The Nowcasting of Precipitation during Sydney 2000: An Appraisal of the QPF Algorithms', *Weather and Forecasting*, 19(1), pp. 7–21. doi: 10.1175/1520-0434(2004)019<0007:TNOPDS>2.0.CO;2.

Pierce, C. E. *et al.* (2005) 'Use of a stochastic precipitation nowcast scheme for fluvial flood forecasting and warning', *Atmospheric Science Letters*, 6(1), pp. 78–83. doi: 10.1002/asl.102.

Price, D. *et al.* (2012) 'Operational use of a grid-based model for flood forecasting', in *Proceedings of the Institution of Civil Engineers: Water Management*, pp. 65–77. doi: 10.1680/wama.2012.165.2.65.

Prigent, C. (2010) 'Comptes Rendus Geoscience Precipitation retrieval from space: An overview', *Comptes rendus - Geoscience*. Academie des sciences, 342(4–5), pp. 380–389. doi: 10.1016/j.crte.2010.01.004.

Probert-Jones, J. R. (1962) 'The radar equation in meteorology', *Quarterly Journal of the Royal Meteorological Society*, 88(378), pp. 485–495.

Pulkkinen, S. *et al.* (2019) 'Pysteps: An open-source Python library for probabilistic precipitation nowcasting (v1.0)', *Geoscientific Model Development*, 12(10), pp. 4185–4219. doi: 10.5194/gmd-12-4185-2019.

Rico-Ramirez, M. A. (2004) *Quantitative weather radar and the effects of the vertical reflectivity profile*. University of Bristol, Bristol.

Rico-Ramirez, M. A. *et al.* (2007) 'A high-resolution radar experiment on the island of Jersey', *Meteorological Applications*, 14(2), pp. 117–129.

Rico-Ramirez, M. A. (2012) 'Adaptive attenuation correction techniques for C-band polarimetric weather radars', *IEEE Transactions on Geoscience and Remote Sensing*, 50(12), pp. 5061–5071. doi: 10.1109/TGRS.2012.2195228.

Rico-Ramirez, M. A. (2019) 'Advances in the Measurement and Forecasting of Precipitation with Weather Radar for Flood Risk Management', in *Technology, Science, and Culture: A Global Vision 2018*. Universidad de las Américas Puebla: IntechOpen, pp. 40–51. doi: doi.org/10.5772/intechopen.83691.

Rico-Ramirez, M. A. and Cluckie, I. D. (2007) 'Bright-band detection from radar vertical reflectivity profiles', *International Journal of Remote Sensing*, 28(18), pp. 4013–4025. doi: 10.1080/01431160601047797.

Rico-Ramirez, M. A. and Cluckie, I. D. (2008) 'Classification of ground clutter and anomalous propagation using dual-polarization weather radar', *IEEE Transactions on Geoscience and Remote Sensing*, 46(7), pp. 1892–1904. doi: 10.1109/TGRS.2008.916979.

Rico-Ramirez, M. A., Cluckie, I. D. and Han, D. (2005) 'Correction of the bright band using dual-polarisation radar', *Atmospheric Science Letters*, 6(1), pp. 40–46. doi: 10.1002/asl.89.

Rico-Ramirez, M. A., Liguori, S. and Schellart, A. N. A. (2015) 'Quantifying radar-rainfall uncertainties in urban drainage flow modelling', *Journal of Hydrology*. Elsevier B.V., 528, pp. 17–28. doi: 10.1016/j.jhydrol.2015.05.057.

Rinehart, R. E. (1997) *Radar for Meteorologists*. 3rd edn. Grand Forks, ND: Rinehart Publications.

Rinehart, R. E. and Garvey, E. T. (1978) 'Three-dimensional storm motion detection by conventional weather radar', *Nature*, 273(5660), pp. 287–289. doi: 10.1038/273287a0.

Roberts, N. M. *et al.* (2009) 'Use of high-resolution NWP rainfall and river flow forecasts for advance warning of the Carlisle flood, north-west England', 16, pp. 23–34. doi: DOI: 10.1002/met.94.

Rossi, P. J. *et al.* (2015) 'Kalman filtering-based probabilistic nowcasting of object-oriented tracked convective storms', *Journal of Atmospheric and Oceanic Technology*, 32(3), pp. 461–477. doi:

10.1175/JTECH-D-14-00184.1.

Ryzhkov, A. V. *et al.* (2005) 'Calibration issues of dual-polarization radar measurements', *Journal of Atmospheric and Oceanic Technology*, 22(8), pp. 1138–1155. doi: 10.1175/JTECH1772.1.

Savina, M. *et al.* (2012) 'Comparison of a tipping-bucket and electronic weighing precipitation gage for snowfall', *Rainfall in the Urban Context: Forecasting, Risk and Climate Change*. Elsevier B.V., 103, pp. 45–51. doi: 10.1016/j.atmosres.2011.06.010.

Schellart, A. N. A. *et al.* (2012) 'Analysis of different quantitative precipitation forecast methods for runoff and flow prediction in a small urban area', in *Weather Radar and Hydrology (Proceedings of a symposium held in Exeter, UK, April 2011)*. Exeter, UK, pp. 614–619. doi: 10.1080/02626667.2014.920505.

Schellart, A. N. A., Shepherd, W. J. and Saul, A. J. (2012) 'Influence of rainfall estimation error and spatial variability on sewer flow prediction at a small urban scale', *Advances in Water Resources*. Elsevier Ltd, 45, pp. 65–75. doi: 10.1016/j.advwatres.2011.10.012.

Schiemann, R. *et al.* (2011) 'Geostatistical radar-raingauge combination with nonparametric correlograms: Methodological considerations and application in Switzerland', *Hydrology and Earth System Sciences*, 15(5), pp. 1515–1536. doi: 10.5194/hess-15-1515-2011.

Schroeder, A. J. *et al.* (2006) 'An automated high-resolution, rapidly relocatable meteorological nowcasting and prediction system', *Monthly Weather Review*, 134(4), pp. 1237–1265. doi: 10.1175/MWR3118.1.

Scottish Flood Forecasting Service (2014) *Radar coverage and suitability for flood forecasting in Scotland*. Available at: https://floodforecastingservice.net/2014/11/03/radar-coverage-and-suitability-for-flood-forecasting-in-scotland/ (Accessed: 12 April 2020).

Seed, A. W. (2003) 'A Dynamic and Spatial Scaling Approach to Advection Forecasting', *Journal of Applied Meteorology*, 42(3), pp. 381–388. doi: 10.1175/1520-0450(2003)042<0381:ADASSA>2.0.CO;2.

Seed, A. W. and Austin, G. L. (1990) 'Variability of summer Florida rainfall and its significance for the estimation of rainfall by gages, radar, and satellite', *Journal of Geophysical Research*, 95(D3), pp. 2207–2215. doi: 10.1029/JD095iD03p02207.

Seed, A. W., Pierce, C. E. and Norman, K. (2013) 'Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme', *Water Resources Research*, 49(10), pp. 6624–6641. doi: 10.1002/wrcr.20536.

Sevruk, B. and Nespor, V. (1994) 'The Effect of Dimensions and Shape of Precipitation Gauges on the Wind-Induced Error', in Desbois, M. and Désalmand, F. (eds) *Global Precipitations and Climate Change*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 231–246. doi: 10.1007/978-3-642-79268-7_14.

Shehu, B. and Haberlandt, U. (2021) 'Relevance of merging radar and rainfall gauge data for rainfall nowcasting in urban hydrology', *Journal of Hydrology*. Elsevier B.V., 594(May 2020), p. 125931. doi: 10.1016/j.jhydrol.2020.125931.

Sideris, I. V. *et al.* (2014) 'Real-time radar-rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland', *Quarterly Journal of the Royal Meteorological Society*, 140(680), pp. 1097–1111. doi: 10.1002/qj.2188.

Simonin, D. *et al.* (2017) 'Performance of Met Office hourly cycling NWP-based nowcasting for precipitation forecasts', *Quarterly Journal of the Royal Meteorological Society*, 143(708), pp. 2862–2873. doi: 10.1002/qj.3136.

Simonin, D., Ballard, S. P. and Li, Z. (2014) 'Doppler radar radial wind assimilation using an hourly cycling 3D-Var with a 1.5 km resolution version of the Met Office Unified Model for nowcasting', *Quarterly Journal of the Royal Meteorological Society*, 140(684), pp. 2298–2314. doi: 10.1002/qj.2298.

Sinclair, S. and Pegram, G. G. S. (2005) 'Combining radar and rain gauge rainfall estimates using conditional merging', *Atmospheric Science Letters*, 6(1), pp. 19–22. doi: 10.1002/asl.85.

Smith, E. A. *et al.* (2007) 'International Global Precipitation Measurement (GPM) Program and Mission: An Overview', in *Measuring Precipitation from Space*. Springers, pp. 611–653.

Smyth, T. J. and Illingworth, A. J. (1998) 'Correction for attenuation of radar reflectivity using polarizatoin data', *Quarterly Journal of the Royal Meteorological Society*. 2006th edn, 124(551), pp. 2393–2415. doi: 10.1002/qj.49712455111.

Sokol, Z. (2003) 'Utilization of regression models for rainfall estimates using radar-derived rainfall data

and rain gauge data', *Journal of Hydrology*, 278, pp. 144–152.

Sokol, Z. Z. *et al.* (2013) 'Comparison of precipitation nowcasting by extrapolation and statistical-advection methods', *Atmospheric Research*. Elsevier B.V., 123, pp. 17–30. doi: 10.1016/j.atmosres.2012.10.013.

Sorooshian, S. *et al.* (2000) 'Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall', *Bulletin of the American Meteorological Society*, 81(9), pp. 2035–2046.

Srivastava, P. K. *et al.* (2014) 'Sensitivity and uncertainty analysis of mesoscale model downscaled hydro-meteorological variables for discharge prediction', *Hydrological Processes*, 28(15), pp. 4419–4432. doi: 10.1002/hyp.9946.

Steiner, M., Jr, R. A. H. and Yuter, S. E. (1995) 'Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data', *Journal of Applied Meteorology and Climatology*, 34, pp. 1978–2007.

Steiner, M. and Smith, J. A. (2002) 'Use of three-dimensional reflectivity structure for automated detection and removal of nonprecipitating echoes in radar data', *Journal of Atmospheric and Oceanic Technology*, 19(5), pp. 673–686. doi: 10.1175/1520-0426(2002)019<0673:UOTDRS>2.0.CO;2.

Strangeways, I. (2004) 'Improving precipitation measurement', *International Journal of Climatology*, 24(11), pp. 1443–1460. doi: 10.1002/joc.1075.

Strangeways, I. (2006) *Precipitation: Theory, Measurement and Distribution*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511535772.

Strangeways, I. (2010) 'A history of rain gauges', *Weather*, 65(5), pp. 133–138.

Suh, J. S. *et al.* (2017) 'Drone-Based External Calibration of a Fully Synchronized Ku-Band Heterodyne FMCW Radar', *IEEE Transactions on Instrumentation and Measurement*. IEEE, 66(8), pp. 2189–2197. doi: 10.1109/TIM.2017.2687518.

Sun, X. *et al.* (2000) 'Flood estimation using radar and raingauge data', *Journal of Hydrology*, 239, pp. 4–18.

Teegavarapu, R. S. V. (2014) 'Statistical corrections of spatially interpolated missing precipitation data estimates', *Hydrological Processes*, 28(11), pp. 3789–3808. doi: 10.1002/hyp.9906.

Thorndahl, S. *et al.* (2016) 'Weather radar rainfall data in urban hydrology', *Hydrology and Earth System Sciences Discussions*, 21, pp. 1–37. doi: 10.5194/hess-2016-517.

Tilford, K. A., Sene, K. and Collier, C. G. (2003) *Flood Forecasting - Rainfall Measurement and Forecasting*. Bristol.

Todini, E. (2001) 'A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements', *Hydrology and Earth System SAciBeanyceessia*, 5(2), pp. 187–199.

Tokay, A. *et al.* (2003) 'Rain Gauge and Disdrometer Measurements during the Keys Area Microphysics Project (KAMP)', *Journal of Atmospheric and Oceanic Technology*, 20, pp. 1460–1477.

Toth, Z. and Kalnay, E. (1997) 'Ensemble Forecasting at NCEP and the Breeding Method', *Monthly Weather Review*, 127, pp. 3297–3319.

Turner, B., Zawadzki, I. and Germann, U. (2004) 'Predictability of Precipitation from Continental Radar Images. Part III: Operational Nowcasting Implementation (MAPLE)', *Journal of Applied Meteorology*, 43(2), pp. 231–248. doi: 10.1175/1520-0450(2004)043<0231:POPFCR>2.0.CO;2.

Ulbrich, C. W. and Lee, L. G. (1999) 'Rainfall measurement error by WSR-88D radars due to variations in Z-R law parameters and the radar constant', *Journal of Atmospheric and Oceanic Technology*, 16(8), pp. 1017–1024. doi: 10.1175/1520-0426(1999)016<1017:RMEBWR>2.0.CO;2.

Velasco-Forero, C. A. *et al.* (2004) 'A non-parametric methodology to merge raingauges and radar by kriging: to errors in radar measuresensitivityments', in *ERAD 2004: Third European Conference on Radar in Meteorology and Hydrology*, pp. 21–24.

Velasco-Forero, C. A. *et al.* (2009) 'A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data', *Advances in Water Resources*. Elsevier Ltd, 32(7), pp. 986–1002. doi: 10.1016/j.advwatres.2008.10.004.

Villarini, G. *et al.* (2008) 'Rainfall and sampling uncertainties: A rain gauge perspective', *Journal of Geophysical Research Atmospheres*, 113(11), pp. 1–12. doi: 10.1029/2007JD009214.

Villarini, G. and Krajewski, W. F. (2010) 'Review of the Different Sources of Uncertainty in Single Polarization Radar-Based Estimates of Rainfall', *Surveys in Geophysics*, 31(1), pp. 107–129.

Vuerich, E. *et al.* (2009) *WMO field intercomparison of rainfall intensity gauges. Instruments and*

Observing Methods. WMO/TD no. 1504. Geneva, Switzerland.

Van Der Wal, D., Pye, K. and Neal, A. (2002) 'Long-term morphological change in the Ribble Estuary, northwest England', *Marine Geology*, 189(3–4), pp. 249–266. doi: 10.1016/S0025-3227(02)00476-0.

WaPUG (2002) *Code of practice for the hydraulic modelling of sewer systems*, *Wastewater Planning Users Group*. Available at: http://www.ciwem.org/knowledge-networks/groups/urban-drainage/publications/code-of-practice.aspx (Accessed: 19 January 2017).

Watson, N. and Howe, J. (2006) 'Implementing the EU Water Framework Directive: Experiences of participatory planning in the Ribble basin, North West England', *Water International*, 31(4), pp. 472–487. doi: 10.1080/02508060608691951.

Wilby, R. L., Beven, K. J. and Reynard, N. S. (2008) 'Climate change and fluvial flood risk in the UK: more of the same?', *Hydrological Processes*, 22, pp. 2511–2523. doi: 10.1002/hyp.6847.

Wilson, J. W. *et al.* (2004) 'Sydney 2000 Forecast Demonstration Project: Convective Storm Nowcasting', *Weather and Forecasting*, 19(1), pp. 131–150. doi: 10.1175/1520-0434(2004)019<0131:SFDPCS>2.0.CO;2.

Wilson, J. W., Brandes, E. A. and Noaa, E. R. L. (1979) 'Radar Measurement of Rainfall — A Summary', *Bulletin of the American Meteorological Society*, 60(9), pp. 1048–1060. doi: 10.1016/S0304-4238(00)00152-7.

Yin, J. *et al.* (2019) 'UAV-Aided Weather Radar Calibration', *IEEE Transactions on Geoscience and Remote Sensing*. IEEE, 57(12), pp. 1–14. doi: 10.1109/tgrs.2019.2933912.

Zahraei, A. *et al.* (2012) 'Quantitative Precipitation Nowcasting: A Lagrangian Pixel-Based Approach', *Atmospheric Research*. Elsevier B.V., 118, pp. 418–434. doi: 10.1016/j.atmosres.2012.07.001.

Zappa, M. *et al.* (2010) 'Propagation of uncertainty from observing systems and NWP into hydrological models: COST-731 Working Group 2', *Atmospheric Science Letters*, 11(2), pp. 83–91. doi: 10.1002/asl.274.

Zhang, G. *et al.* (2019) 'Current Status and Future Challenges of Weather Radar Polarimetry: Bridging the Gap between Radar Meteorology/Hydrology/Engineering and Numerical Weather Prediction', *Advances in Atmospheric Sciences*, 36(6), pp. 571–588. doi: 10.1007/s00376-019-8172-4.

Zhu, D., Xuan, Y. and Cluckie, I. (2014) 'Hydrological appraisal of operational weather radar rainfall estimates in the context of different modelling structures', *Hydrology and Earth System Sciences*, 18(1), pp. 257–272. doi: 10.5194/hess-18-257-2014.