Crowdsourced Rainfall Data

Cleaning, Validation, Visualisation, and Application

(WOW)

Tess O'Hara

A thesis submitted to Newcastle University in partial fulfilment of the requirements for the degree of Doctor of Philosophy within the Faculty of Science, Agriculture and Engineering.

September 2022

Acknowledgements

I would like to acknowledge the support of my supervisors, past and present. I stubbornly outlasted Dr Geoff Parkin, despite his best efforts to see this through. I wish him well in his retirement. Professor Hayley Fowler challenged the outputs at each stage to ensure the quality of this research. Dr Elizabeth Lewis has consistently been available for everything from technical support to encouragement during the tougher stretches. Dr Lewis contributed to running engagement activities with museums and schools, enthusiastically sharing her knowledge, experience, and supersized educational games. This work could not have been completed, or even contemplated, without the patient assistance of Dr Fergus McClean (AKA the python wizard), who painstakingly dismantled and rebuilt my amateur code, turning it into something pythonic and functional. Dr McClean taught me well, and I hope he takes pride in how far I travelled under his tutelage and beyond. Jake Brown of the Met Office is thanked for the provision of WOW data, and insight into the mechanics of the archive.

I would like to acknowledge the contribution of Dr Roberto Villalobos-Herrera, who provided a quality controlled Official rain gauge dataset for the UK, implemented the quality control algorithm for citizen science data, and patiently answered the myriad of questions I posed. I wish to thank Dr Susana Ochoa Rodriguez and Dr Lipen Wang for providing radar data and the software used in the generation of blended rain gauge and radar observations, and for their guidance in the interpretation of the outputs. Dr Ben Smith and Dr Steven Birkinshaw assisted with the set-up and parametrisation of SHETRAN, which was immensely helpful. I would also like to express my thanks to the numerous contributors to the Data, Risk, and Environmental Analytical Methods (DREAM) Centre for Doctoral Training, who devised and supervised activities to aid learning and development, and the Natural Environment Research Council for funding this research.

On a personal note, I would like to recognise the support of my family through not only the PhD but also the tough times of the pandemic. My partner Peter, son Ben and dog Clary have been there each step of the way, and I could not have done it without them. Finally, I would like to thank my parents for always letting me get on with it, never questioning one seemingly random endeavour after another.

Abstract

Rainfall observations from citizen scientists in Britain are shared from private automated weather stations (PAWS) to the Met Office Weather Observations Website (WOW). There is a reluctance by researchers and professionals to use the data which is ascribed to the relative difficulty in access, and an uncertainty in data quality. This research provides an assessment of the available data, metrics to assess data quality, examples of potential applications, and considers barriers and benefits to the provision of such data.

Over 2,700 PAWS records were subject to statistical quality control (SQC), which flagged 2.5% of the dataset for removal representing 95% of the recorded rainfall total, indicating the propensity for PAWS records to include implausibly high rainfall. The post-SQC data were assessed against comparable datasets, establishing that some PAWS (40%) were generating good quality data all the time (based on SQC performance) whilst other records included intermittent faults (50%), and some were unreliable observers (10%). A systematic manual SQC process is presented, allowing the selection of good quality PAWS records for use in post-event analysis of convective rainfall events. When post-SQC'd rain observations from WOW and Official sources were compared with radar there was no statistically significant difference. This research concludes if Official data are accepted for use, then WOW data selected using the presented SQC methods are also of acceptable quality.

The results of blending gauge observations with radar from four convective rainfall events are presented to show the value of WOW data in post-event analysis. Examples of improved delineation of the rainfall field and more accurate representations of rainfall than derived solely from radar are provided. This is further illustrated with an example of catchment modelling showing a rainfall dataset incorporating Official and WOW gauge data with radar improves model efficiency, as compared to using only gauge data or radar derived rainfall.

Finally, the challenges of sustainable citizens science rainfall observation are addressed. A series of approaches for the instigation and promotion of citizen science are presented. Conclusions on the most effective way to ensure a rewarding experience for participants include self-selection combined with easy access to data sharing, timely data validation with feedback, and peer/technical support. This approach supports willing observers, promoting the generation of good quality reliable rainfall data.

Contents

Acknowledgements iii

Abstract iv

Table of Acronyms ix

List of Figures x

List of Tables xvi

Chapter 1. Introduction 1

1.1 Preface 1
1.2 Rainfall Observations – Ground Based Rain Gauges and Radar 2
1.3 What is Citizen Science and Crowdsourcing? 4
1.4 Crowdsourcing, Citizen Science and WOW 5
1.5 Flooding in Britain 6
1.6 Thesis Structure/Outline 8
1.6.1 Aims and Research Questions 8

Chapter 2. Literature Review 11

2.1 Introduction 11
2.2 Rainfall Observations 11
2.2.1 Ground Based Rain Gauge Monitoring networks 11
2.2.2 Radar 12

2.3 Rain Gauge Data Quality 15
2.3.1 Fundamentals 15
2.3.2 Rain Gauge Installation 16
2.3.3 Rainfall Data Statistical Quality Control 16

2.4 Citizen Science 20
2.4.1 Citizen Science Motivation 21
2.4.2 Citizen Science/Crowdsourced Rainfall Data and WOW 21
2.4.3 Citizen Science Rainfall Data Quality 24
2.4.4 Citizen Science for Flood and Catchment Management 27

2.5 Conclusion 29

Chapter 3. What, Where, When, WOW? 30

3.1 Overview 30
3.2 Introduction 31
3.3 Data Access and Collation 31
3.3.1 Data Upload and Download Mechanisms 31
3.3.2 Available Parameters 32
3.3.3 Station Metadata 32

3.3.4 Observation Intervals....................34

3.4 Dataset Creation; Consolidation and Cleaning Methods....................36
3.4.1 Summary....................36
3.4.2 Multiple Parameter Records at Original Time Intervals....................37
3.4.3 Station Summary Statistics....................37
3.4.4 Selecting Stations....................38
3.4.5 Calculated Rainfall Observations at Original Time Intervals....................39
3.4.6 Rain Observations at Hourly Interval....................40

3.5 Dataset Description....................40
3.5.1 Selecting British Stations....................40
3.5.2 Filtered Raw Multi-Parameter Data from Britain....................41
3.5.3 Station Count and Longevity of Record....................41
3.5.4 Distribution of Weather Stations....................43
3.5.5 Intervals between Observations....................45
3.5.6 Calculating Rain Depth....................45
3.5.7 Resampling to Hourly....................46

3.6 Spatial Analysis....................47
3.6.1 Method....................47
3.6.2 Results....................48

3.7 Discussion....................52
3.7.1 Response to Research Questions....................54
3.7.2 Recommendations for Using WOW Rainfall Data....................56
3.7.3 Outputs....................57

Chapter 4. Describing Crowdsourced Rainfall Observations and Assessing Quality using a Statistical Quality Control Algorithm....................59

4.1 Overview....................59
4.2 Introduction....................60
4.2.1 Research Questions....................60
4.2.2 Data Sources....................60

4.3 Methods....................61
4.3.1 Determining WOW Rain Gauge Data Quality....................61
4.3.2 Appropriateness and Effectiveness of SQC for WOW Rain Gauge Data....................63
4.3.3 Exploring the potential for improvement in WOW data quality....................64

4.4 Results....................64
4.4.1 Data Availability Pre-SQC....................65
4.4.2 Impact of SQC....................70
4.4.3 Nearest Neighbour Comparison....................83
4.4.4 Identifying post-SQC Good Quality WOW Rainfall Data....................85
4.4.5 Efficacy of SQC for WOW Rain Observations....................89

4.5 Case Study – All Citizen Scientists are Equal, but Some are More Equal Than Others....................97
4.5.1 Introduction....................97

4.5.2 Results and Discussion 98

4.6 Discussion 101
4.6.1 Response to Research Questions..103
4.6.2 Recommendations for Using WOW Data 107
4.6.3 Outputs 107

Chapter 5. Event Based Quality Control with Data Validation and Visualisation 108

5.1 Overview..108
5.2 Introduction..109
5.3 Methods 109
5.3.1 Event selection methodology 109
5.3.2 Method for Acquiring Rainfall Data 110
5.3.3 Method for Determining WOW Rainfall Data Accuracy..111
5.3.4 Methods for Creating Interpolations and Assessing 'Value Add' of WOW Rain Gauge Data..113

5.4 Results 114
5.4.1 Events Selection 114
5.4.2 Determining Event Temporal and Spatial Extent and Data Availability 115
5.4.3 Manual Quality Control Results 124
5.4.4 WOW Data Quality Versus Official Data Quality. 127
5.4.5 Creating Interpolations – Determining Event Temporal Extent 140
5.4.6 Creating Interpolations – Determining Event Spatial Extent/Gauge Coverage 144
5.4.7 Assessing the 'Value Add' of Including WOW Data 147

5.5 Discussion 158
5.5.1 Response to Research Questions..160
5.5.2 Recommendations 161

Chapter 6. Case Study - Using WOW Rainfall Data in Hydrological Modelling 162

6.1 Overview..162
6.2 Introduction..163
6.2.1 Background 164
6.2.2 Catchment..165

6.3 Method 165
6.3.1 Interpolation and Blend Analysis Methods 167

6.4 Results 168
6.5 Discussion 173
6.5.1 Response to Research Question 175
6.5.2 Recommendations 175

Chapter 7. Observing the Observers – Motivations and Barriers to Sharing Data from Private Automated Weather Stations 176

7.1 Overview 176
7.2 Introduction 177
7.3 Method 177
7.3.1 Community Outreach 178
7.3.2 Schools and Special Interest Group Engagement 178
7.3.3 Weather Station Provision 179
7.3.4 Engaging with a Special Interest Group (COL) 180

7.4 Results 180
7.4.1 Community Outreach 180
7.4.2 Schools Engagement 182
7.4.3 Weather Stations Provided 184
7.4.4 The Climatological Observers Link 191

7.5 Discussion 194
7.5.1 Citizen Science Participation and Crowdsourcing Rainfall Data 194
7.5.2 Response to Research Questions 195

Chapter 8. Conclusions 198

8.1 Overview 198
8.2 Results Summary 199
8.3 Results Significance 201
8.4 Limitations and Further Work 204
8.4.1 Data Availability and Sources 204
8.4.2 Statistical Quality Control Developments 205
8.4.3 Follow-on Research 206

Appendix 1 207

Appendix 2 208

Appendix 3 213

Appendix 4 225

References 237

Table of Acronyms

Acronym	Explanation
AMS	American Meteorological Society
API	Application Programming Interface
AWS	Automated Weather Stations
BAY	Bayesian radar and point rainfall data merging technique
CDD	Consecutive Dry Day
CoCoRaHs	Community Collaborative Rain, Hail and Snow Network
COL	Climatological Observer's Link
CSO	Combined Sewer Overflow
DREAM	Data, Risk, and Environmental Analytical Methods
EA	Environment Agency
ETCCDI	Expert Team on Climate Change Detection and Indices
FAIR	Findability, Accessibility, Interoperability and Reuse
FCRM	Flood and Coastal Risk Management
GHCN	Global Historical Climatology Network
GMT	Greenwich Mean Time
ID	Identification
KED	Kriging with external drift radar and point rainfall data merging technique
LLFA	Lead Local Flood Authority
MFB	Mean Field Bias
MIDAS	Met Office Integrated Data Archive System
NOAA	National Oceanic and Atmospheric Administration
NRW	Natural Resources Wales
NWS	National Weather Service
PAWS	Private Automated Weather Station
PWS	Personal Weather Station
QA	Quality Assurance
QC	Quality Control
SEPA	Scottish Environmental Protection Agency
SQC	Statistical Quality Control
TBR	Tipping Bucket Rain gauge
UK	United Kingdom
UKEOF	United Kingdom Environmental Observation Framework
USA	United States of America
WGS	World Geodetic System
WMO	World Meteorological Organisation
WOW	Weather Observation Website
WU	Weather Underground

List of Figures

Figure 1.1 Photograph of a typical cumulonimbus cloud formation...........................2
Figure 1.2 Map of British Official ground based automated rain gauge and radar locations, 2011. 3
Figure 1.3 WOW screenshot showing the global distribution of registered citizen science weather stations (as of May 2022)....................................6
Figure 1.4 Photographs of different flood types from the UK - Top left: River Great Ouse in flood River Great Ouse, Pauline A Marsh / River Great Ouse in flood / CC BY-SA 2.0. Top right: 'Toon Monsoon, flooding in Newcastle Upon Tyne, 2012, chroniclelive.co.uk. Bottom left: Coastal flooding risk at Ilfracombe, JBA Consulting. Bottom right: Groundwater flooding, Compton, 2021, British Geological Society. 7
Figure 2.1 UK rainfall radar locations, after Harrison et al., (2015). 14
Figure 2.2 Illustrated examples of negative interpretations of commonly used names to describe people participating in citizen science, after Eitzel et al. Citizen Science Terminology Matters, 2017...........................20
Figure 3.1 Instructions for determining WOW site rating for a rain gauge................33
Figure 3.2 Photographs of example weather station types and set-up (from left to right: Davis Vantage Pro 2 combined parameter weather station, Davis Vue weather station, and a Netatmo rain gauge). 34
Figure 3.3 Flowchart of data processing in this thesis..36
Figure 3.4 Met Office raw data PAWS locations from Britain, the Isle of Man, Northern Ireland, Republic of Ireland and the Channel Islands. 41
Figure 3.5 Count of PAWS in Britain reporting to WOW each year (2011 – 2020). 42
Figure 3.6 Histogram of Reporting Durations (years) for WOW Weather Stations 42
Figure 3.7 Distribution of British WOW weather stations 2011 – 2020 (rural = blue, urban = grey)...44
Figure 3.8 Maps of British Official and WOW rain gauge coverage, 2018 (10km and 5km scale). 50
Figure 3.9 Maps of British Official and WOW rain gauge coverage of urban areas (5km scale) with West Midlands Insert...51
Figure 4.1 Map of Official rain gauge locations with duration of reporting (EA, SEPA, NRW) to 2018. 66
Figure 4.2 Map of WOW rain gauge locations with duration of reporting (2011 - 2020). 66
Figure 4.3 Histogram of record lengths per rain gauge (WOW and Official datasets). 67

Figure 4.4 Scatter plot of percent per gauge of data missing from Official gauges pre-SQC versus length of record. 68

Figure 4.5 Scatter plot of percent per gauge of data missing from WOW gauges pre-SQC versus length of record. 68

Figure 4.6 Histogram of data missing from Official gauges pre-QC. 69

Figure 4.7 Histogram of data missing from WOW gauges pre-QC. 69

Figure 4.8 Histogram showing data removed from Official gauges by SQC. 71

Figure 4.9 Histogram showing data removed from WOW gauges by SQC. 71

Figure 4.10 Scatter plot showing relationship between pre- and post-SCQ missing data for WOW gauges where >1% of observations were removed by SQC (colour shows duration in years). 72

Figure 4.11 Histogram showing rain removed from WOW gauges by SQC 73

Figure 4.12 Histogram showing rain removed from official gauges by SQC. 73

Figure 4.13 Combined scatter plot and histograms showing relative percentage of hours and rainfall removed from Official gauges by SQC. 74

Figure 4.14 Combined scatter plot and histograms showing relative percentage of hours and rainfall removed from WOW gauges by SQC. 74

Figure 4.15 Maps showing percentage of hours removed per rain gauge (left) and percentage of rain removed by rain gauge (right) by SQC for WOW dataset. 76

Figure 4.16 Maps showing percentage of hours removed per rain gauge (left) and percentage of rain removed by rain gauge (right) by SQC for Official dataset. 77

Figure 4.17 Comparison of percentage of WOW and Official gauges at which each rule was used to exclude suspicious hours. 78

Figure 4.18 Percentage of data per gauge flagged for removal by SQC rule - Official gauges (see numbers of gauges per rule in text). 80

Figure 4.19 Percentage of data per gauge flagged for removal by SQC rule – WOW gauges (see numbers of gauges per rule in text). 80

Figure 4.20 Plot showing pre-SQC timeseries for WOW gauge breaking R5 – Example Gauge 1. 82

Figure 4.21 Plot showing post-SQC timeseries for WOW gauge R5 - Example Gauge 1. 82

Figure 4.22 Plot showing pre-SQC timeseries for WOW gauge breaking R5- Example Gauge 2. 82

Figure 4.23 Histograms of nearest neighbour correlation plots (presence/absence, Spearman, and Pearson). 84

Figure 4.24 Box and Whisker plots showing nearest neighbour correlations between WOW and Official gauges. 85

Figure 4.25 Map of WOW gauges with no hours removed by SQC and Official Gauges (selected post-manual review). 86

Figure 4.26 Map of WOW Gauges with no hours removed by SQC, and good correlation with nearest Official neighbour (colour scale shows distance in km to nearest Official gauge) 87

Figure 4.27 Map of Very Good WOW gauges with no nearest neighbour check available (colour scale by duration of reporting). 88

Figure 4.28 Maps of British Official and WOW rain gauge coverage meeting Very Good quality control standard, 2018 (5km scale). 89

Figure 4.29 Example gauge 1 – complete timeseries of rainfall (mm) from WOW gauge (ID 41937481), and nearest Official neighbouring gauge. 91

Figure 4.30 Example gauge 1 - selected timeseries for WOW gauge (ID 41937481), highlighting apparent error. 91

Figure 4.31 Example gauge 1 – selected timeseries for WOW (ID 41937481), and nearest Official neighbour providing a comparison to explore an apparent error. 92

Figure 4.32 Example gauge 2 - high hourly rainfall at WOW gauge (ID:32265743) not exceeding 92mm hr^{-1} with corresponding timeseries from nearest official neighbour. 93

Figure 4.33 Example gauge 3 – complete timeseries at WOW rain gauge (ID:431606049) showing erroneous rainfall post-SQC. 94

Figure 4.34 Example of WOW observations with nearest neighbouring Official observations (full period of WOW gauge record) 95

Figure 4.35 Selected timeseries from WOW and neighbouring Official gauges showing discrepancy in observations 96

Figure 4.36 Duration of reporting from COL gauges. 98

Figure 4.37 Relative percentage of hours and rainfall removed from COL Gauges by SQC. 100

Figure 5.1 Satellite image showing the locations of the four case study event locations used in post-event reconstruction. 115

Figure 5.2 Satellite image of Canvey Island affected by flooding in July 2014. 116

Figure 5.3 Daily rain from Official rain gauge (ID: EARAYLER) close to Canvey Island for the period prior to flooding in July 2014. 116

Figure 5.4 Satellite image of Bar Hill most affected by flooding in August 2014. 117

Figure 5.5 Daily rainfall from Official rain gauge (ID: EA179635) close to Bar Hill for period prior to flooding in August 2014. 118

Figure 5.6 Map showing area of Birmingham affected by flooding, after Section 19 report, WSP, 2017. 119

Figure 5.7 Maps of archived radar data for the UK on 16/06/2016, showing the evolution of the storm (source: MetCheck). 119

Figure 5.8 Daily rain from Official rain gauge (ID: EA3176_RK) closest to Birmingham for the period prior to flooding in June 2016. 120

Figure 5.9 Satellite image of the area most affected by flooding in Milton Keynes on 25/05/2018....121

Figure 5.10 Archived UK radar images for the 27/05/18 (source: MetCheck)....122

Figure 5.11 Daily rainfall from Official rain gauge (ID: EA170823) closest to Milton Keynes for the period prior to flooding in July 2018. 123

Figure 5.12 Milton Keynes WOW rain gauges rejected following manual first pass checks....125

Figure 5.13 WOW gauge from Canvey event with poor SQC performance. 126

Figure 5.14 Pearson correlation matrix for all gauges within 40 km of Bar Hill with <10% data missing from start of 2014 to event date. 127

Figure 5.15 Pearson Correlation between hourly data for co-located radar and rain gauge data, with '---' indicating level below which correlation poor (0.368). 130

Figure 5.16 Line plot of selected Canvey Island event timeseries for WOW gauges W3 and W5 demonstrating discrepancy between radar and rain gauge (suffix 'rd' = radar, 'rg' = rain gauge) 131

Figure 5.17 Bar Hill selected event time series for Official gauges EA14 and MO4 demonstrating discrepancy between radar and rain gauge (suffix 'rd' = radar, 'rg' = rain gauge) 132

Figure 5.18 Milton Keynes selected event time series for WOW gauges W1 and W16 demonstrating discrepancy between radar and rain gauge (suffix 'rd' = radar, 'rg' = rain gauge) 133

Figure 5.19 Strip plots of difference between hourly rain observed at gauge and co-located radar square for all events (radar – rain gauge). 135

Figure 5.20 Histogram showing difference between co-located radar and gauge data for WOW and Official cohorts (normalised for sample size, grey = overlap) 136

Figure 5.21 Boxplot of RMSE between hourly rain gauge and co-located radar rainfall values for WOW and Official gauge data for selected rainfall events....137

Figure 5.22 Map of Canvey Island event gauges (26 WOW and 35 Official) 138

Figure 5.23 Map of Bar Hill event gauges (27 WOW and 31 Official) 138

Figure 5.24 Map of Birmingham event gauges (17 WOW and 25 Official) 139

Figure 5.25 Map of Milton Keynes event gauges (23 WOW and 25 Official) 139

Figure 5.26 Canvey Island radar peak rainfall hours on day of flooding (20/07/2014)....141

Figure 5.27 Bar Hill radar peak rainfall hours on day of flooding (08/08/2014). 142

Figure 5.28 Birmingham radar peak rainfall hours on day of flooding (16/06/2016)....143

Figure 5.29 Milton Keynes radar peak rainfall hours on day of flooding (27/05/2018)....144

Figure 5.30 Canvey Island total rainfall (radar from 14:00 - 16:00 20/07/2014), with Official and WOW rain gauge locations. 145

Figure 5.31 Bar Hill total rainfall (radar from 12:00 - 17:00 08/08/2014) with Official and WOW rain gauge locations. 145

Figure 5.32 Birmingham total rainfall (radar from 16:00 - 20:00 16/06/2016) with Official and WOW rain gauge locations. 146

Figure 5.33 Milton Keynes total rainfall (radar from 19:00 – 21:00 on 27/05/2018) with Official and WOW rain gauge locations 146

Figure 5.34 Radar and interpolations (Official and WOW block kriged, remainder block kriged with external drift) of event rainfall on Canvey Island 20/07/14, 14:00 - 16:00. 148

Figure 5.35 Difference plot between radar/Official interpolation - radar/WOW/Official interpolation for Canvey Island rain (14:00 - 16:00 20/07/2014). 150

Figure 5.36 Difference plots between radar and interpolations for Canvey Island Rain (14:00 - 16:00, 20/07/2014). 150

Figure 5.37 Interpolations Bar Hill, 08/08/2014 12:00 – 17:00 *R/O = Radar/Official **R/W/O = Radar/WOW/Official. 152

Figure 5.38 Difference plot between radar/Official interpolation - radar/WOW/Official interpolation for Bar Hill (12:00 - 17:00, 08/08/2014). 153

Figure 5.39 Interpolations Birmingham, 16/06/2016 17:00 – 20:00. 154

Figure 5.40 Difference plot between radar/Official interpolation - radar/WOW/Official interpolation for Birmingham (17:00 – 20:00, 16/06/2016). 155

Figure 5.41 Difference plots between radar and interpolations for Birmingham rain (17:00 - 20:00, 16/06/2016). 155

Figure 5.42 Interpolations Milton Keynes, 27/05/2018 19:00 – 21:00. 157

Figure 5.43 Difference plot between radar/WOW/Official - radar/Official interpolation for Milton Keynes (19:00 – 21:00, 27/05/2018). 158

Figure 6.1 Map of rain and river gauges relative to surface water catchment and areas affected by pluvial flooding around Birmingham. Source of affected area: "Birmingham City Council, June 2016 Flooding: Flood and Water Management Act, Section 19 Investigation”. 164

Figure 6.2 Hydrograph of simulated versus observed daily flow at Birmingham gauge 28003 for 2016. 168

Figure 6.3 Plots of sum of rainfall from radar/official/wow blend, radar/official blend, radar data and CEHGEAR-1hr with hydrograph of observed and simulated flow at river gauge 28003, Birmingham 06/06/2016 - 17/06/2016. 169

Figure 6.4 Plots of Rainfall from CEHGEAR-1hr, blended and radar data with hydrograph of observed and simulated flow with peaks shown as dashed lines (08/06/16 - 09/06/16). 170

Figure 6.5 Plots of Rainfall from radar/Official/WOW blend, radar/Official blend, radar and CEHGEAR-1hr data with hydrograph of observed and simulated flow with peaks shown as dashed lines (16/06/16 - 17/06/16) – note the observed and blended QMAX occurred at 23:00. 172

Figure 7.1 Daily rainfall from gauges at Corbridge Middle School and Maften. 183

Figure 7.2 Responses from volunteers hosting weather station immediately after receipt. 185

Figure 7.3 Netatmo rain and temperature modules in situ 186

Figure 7.4 Responses from volunteers 6 months after receipt of the weather station 187

Figure 7.5 Plot of daily rainfall from Netatmo weather stations and closest official rain gauge, March - May 2021. 188

Figure 7.6 Plot of daily rainfall from Netatmo weather stations, May 2021. 189

Figure 7.7 Responses from volunteers 18 months after receipt of the weather station 190

Figure 7.8 Visualisation comparing survey response per volunteer across three instances....190

List of Tables

Table 3.1 Number of PAWS reporting to WOW - total and urban (2011 - 2020). 43
Table 3.2 Count of modal observation intervals (in minutes) from WOW weather stations. 45
Table 3.3 Example Rain Record from WOW PAWS. 46
Table 3.4 British Official weather station and WOW PAWS coverage (5km scale). 49
Table 3.5 Urban British Official and WOW station coverage (5 km scale). 52
Table 4.1 Rules applied in SQC to remove suspicious observations (colour coded; blue - excessive rain, yellow - excessive dry, grey – either too wet or dry). 62
Table 4.2 Observations missing prior to SQC for WOW and Official rain gauges. 67
Table 4.3 Observations removed and missing post SQC for WOW and Official gauges. 70
Table 4.4 Rainfall rate (mm/hr) before and after SQC at WOW and Official gauges. 73
Table 4.5 Rainfall removed by SQC for WOW and Official gauges. 73
Table 4.6 Summary of pre- and post-SQC data and rainfall available from COL members. 99
Table 4.7 Descriptive statistics of nearest neighbour analysis on COL data. 100
Table 5.1 Case study pluvial flood events selected for analysis. 115
Table 5.2 Count of rain gauges in WOW, EA, and MIDAS within the event search areas. 123
Table 5.3 Number and percent of WOW gauges rejected by manual SQC Stage 1 per event. 124
Table 5.4 Gauges with co-located radar and rain gauge observations >0.368. 131
Table 5.5 Descriptive statistics for RMSE for WOW and Official rain gauge data and co-located radar during selected rainfall events. 137
Table 5.6 Event Overview, Duration, Intensity and Total Rain as recorded by radar 140
Table 5.7 Descriptive statistics from interpolations of Canvey Island rainfall datasets from different sources. 149
Table 5.8 Descriptive statistics from interpolations of Bar Hill rainfall data from different sources. 151
Table 5.9 Descriptive statistics from interpolations of Birmingham rainfall data from different sources. 154

Table 5.10 Descriptive statistics from interpolations of Milton Keynes rainfall data from different sources. 156

Table 6.1 Statistical comparison of simulations from CEHGEAR, Blended and Radar rainfall datasets, Birmingham 06/06/2016 - 17/06/2016. 169

Table 6.2 QMAX 8th/9th June, Observed and Simulated 171

Table 6.3 QMAX 16th/17th June, observed and simulated flows. 172

Table 7.1 Netatmo winner questionnaire responses (colour coded blue for positive and orange for negative). 181

Table 7.2 Free text responses to weather station winner's questionnaire. 182

Table 7.3 Pearson correlation between Corbridge Rain Gauge Daily Observations (blue = full period, green = March onwards after Netatmo gauge was stood upright). 183

Table 7.4 Locations of Netatmo weather stations hosted by volunteers. 184

Table 7.5 Pertinent points from COL members regarding PAWS and rainfall observations..193

Chapter 1. Introduction

1.1 Preface

This thesis was originally conceived as “follow-on” research undertaken by Starkey *et al.* (2017) centred around hydrology/meteorology participatory citizen science in a flood prone community in North East England. The concept was to use crowdsourced rain gauge data (in contrast to the prior intensive approach employed by Starkey *et al*.), assessing how the data could be used for flood mitigation. Rain gauge data were provided by the Met Office, as the industrial partner for this research, from their Weather Observations Website (WOW) database (Met Office, 2011). This research represents the first known attempt to use WOW rainfall data *en masse*.

After initial exploratory analysis of the data, it was apparent they were not fit for use. There were inconsistencies, extreme values and other issues that explained why data were not being used by the research community. Unlike Official rain gauge observations there was no systematic quality assurance/quality control (QA/QC) approach demonstrating WOW data were carefully collected from well positioned and managed weather stations, before being quality controlled and published. It was clear that something had to be done prior to the application of WOW data. This thesis constitutes that something, aiming to quality control and establish ways to validate WOW rainfall observations, then continuing to explore how such crowdsourced data could be used.

The methods developed and applied in this research may not be applicable to all circumstances, so outputs include easily accessible rainfall data from WOW stations in Britain from 2011 to 2020, at the original reported time interval, aggregated to hourly and following statistical quality control (SQC). This has been made available for other researchers in the hope that increased interest and scrutiny may lead to further insight and use of WOW rainfall data (see Appendix 1). The code used for the processing of data has been published open source for re-use if required (see Appendix 1). In support of Open Data and reusability this research adheres to the FAIR principles, namely: Findability, Accessibility, Interoperability and Reuse of digital assets (Go Fair, 2016; Wilkinson *et al.*, 2016).

1.2 Rainfall Observations – Ground Based Rain Gauges and Radar

The principle reason for operating a weather observation network in Britain is to provide the basis for reliable weather forecasts, which become critical when providing advance warning of extreme weather events, that may, for example, result in flooding caused by heavy and/or intense rainfall (Met Office, 2019). An introductory explanation of the essential characteristics of the three mechanisms that cause rainfall in the UK is available from the Royal Meterological Society (2020), with further detail on convective storms presented by the National Oceanic and Atmospheric Administration (2018). In summary:

Frontal rainfall is associated with the passing of low-pressure systems that cause the uplift of moisture laden air that cools as it rises resulting in rain.

Orographic rainfall is caused by the uplift of air masses over upland areas, cooling moisture laden air as it rises. The air cools at an average of 6.5^{o}C per 1000 meters, often resulting in cloud on the windward side of mountains that can result in rainfall. The lee side of the upland area is described as in the "rain shadow" and can be very dry in comparison to the windward side.

Convective rainfall is caused by localised heating of the ground leading to the formation of Cumulonimbus clouds (see Figure 1.1) with a flat base marking the height temperature cools sufficiently to cause the switch from evaporation to condensation. Cumulonimbus clouds are commonly referred to as thunder clouds in the UK, causing associated high winds and hail along with intense heavy rain. Storms may be single cell, multi-cell, or supercell, each larger and leading to longer periods of rain (from minutes to hours). Multi-cells may form a squall line that is narrow (10-20 km) but long (>100 km) and cause fast moving intense rainfall.

Figure 1.1 Photograph of a typical cumulonimbus cloud formation.

The Met Office maintains a network of ground-based weather stations throughout the UK to record key observations, most frequently including temperature, wind speed and direction, and rainfall. The stations are located at approximately 40 Km intervals, spaced predominately to capture information about the low-pressure frontal systems that dominate UK weather (Met Office, 2016b). In 2021 there were 256 ground-based weather stations with data shared via the Met Office Integrated Data Archive System (MIDAS) in the United Kingdom (UK) reporting hourly rainfall observations (Met Office, 2021b)[1].

The Flood Forecasting Centre (FFC) and responsible flood authorities including the Environment Agency (EA), Scottish Environmental Proception Agency (SEPA), National Resources Wales (NRW) and Lead Local Flood Authorities (LLFAs), rely on accurate weather forecasting to predict when and where flooding will occur. The authorities in the UK (referred to as 'the agencies' henceforth) supplement the Met office monitoring network with additional rain gauges for operational reasons relating to their remit relating to water quality and flood management. The agencies add approximately a further 1,200 rain gauges to the ground-based hourly or sub-hourly rain observation network in Britain, bringing the total to approximately 1,500 (Environment Agency, 2017; Natural Resources Wales, 2022; Scottish Environmental Protection Agency, 2022).

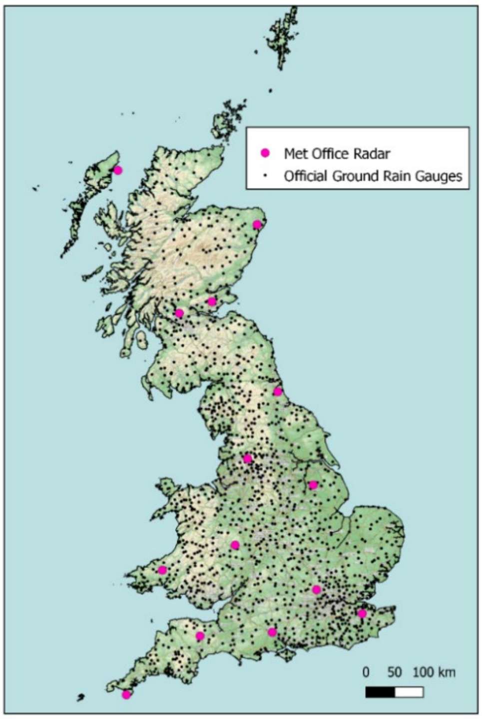


Figure 1.2 Map of British Official ground based automated rain gauge and radar locations, 2011.

[1] There are many more rain gauges reporting daily rainfall, operated by the Met Office, Climate Observers etc.

In addition to ground-based monitoring, weather observations are made using weather balloons, satellite, and weather radar (referred to simply as radar henceforth). Radar covers 99% of the UK: however, the accuracy of radar is affected by ground clutter, overshooting, "bright band" and drift (Wilson *et al.*, 1979b; Sauvageot, 1994; Joss *et al.*, 1995; Krajewski *et al.*, 2010). Radar is particularly useful for determining the temporal and spatial distribution of rain over a wide area (Sauvageot, 1994), and is considered most reliable when blended with rain gauge data to correct for errors in magnitude (Steiner *et al.*, 1999; Trapero *et al.*, 2009; Rabiei *et al.*, 2015).

The Met Office rain gauge network in Britain (see Figure 1.2), although comprehensive, may not always provide the density of rainfall observations required for all applications (Villarini *et al.*, 2008; McGregor *et al.*, 2009). One reason rain gauges networks may be inadequate in the future is that changing weather patterns and climate are projected to result in an increase in convective storms, especially in the summer months (Brooks, 2013; Kendon *et al.*, 2014; Miller *et al.*, 2017). Highly localised single cell convective storms that arise quickly and discharge relatively high volumes of rainfall may be small enough to occur between formal rain gauges. The discharge of the highest intensity rainfall from multi and supercell storms may not correspond with rain gauge locations (Schroeer *et al.*, 2018). Whilst the storms may be visible via radar and satellite, without merging with ground-based gauges there is no certainty in radar precipitation volume (Sauvageot, 1994; Berne *et al.*, 2013; Bárdossy *et al.*, 2018).

1.3 What is Citizen Science and Crowdsourcing?

Specific discussion of crowdsourced rainfall data is included in later sections; however, it is important to define some relevant terminology early on to avoid confusion. The terms "citizen science" and "crowdsource" are widely used (>15 million and 4.8 million search results respectively via Google) but the definitions are not always clear. In this thesis "citizen science" is used to describe the production of data that are useful to scientific investigation by those not professionally engaged to do so. "Crowdsourcing" is considered citizen science at its most basic, and is described as a way of obtaining data by soliciting content from a large group of people who may not understand or be aware of the research intent (Irwin, 1995; Haklay, 2012).

Eitzel *et al.* (2017) looked in detail at the terminology used to describe citizen science, emphasising the need to encourage participation via inclusive language (explored further in section 2.4). The paper points out that there can be differing applications for terms depending on the circumstances of the endeavour. In Chapters 3 – 5 of this thesis, data have been generated by citizen scientists: however, the utilisation of data is described as having been crowdsourced, as there was no active participation in the analysis or use of data. Chapter 6 includes elements of participatory citizen science, including the co-design of projects.

1.4 Crowdsourcing, Citizen Science and WOW

The English are known globally for our obsession with the weather, or at least talking about it (Fox, 2004). Although it might seem to just be making small talk, many people really are curious about what the weather is doing. It could be the changeable temperate maritime climate of the UK that makes us especially interested in weather observations, or the risk of being affected by flooding (Met Office, 2015; Flood Guidance, 2022). Businesses reliant on knowing what the weather is doing include racetrack operators needing to know when to water the grass and farmers keeping tabs on their crops (Future Farming, 2017; TurfTrax, 2019). There is certainly enough interest in the weather for people to have invested in private automated weather stations (PAWS) that record parameters typically including temperature, barometric pressure, wind speed and direction, and rainfall (Bell *et al.*, 2013).

Recorded weather observations can be shared online either automatically via Wi-Fi or mobile phone connections, or manually by uploading data files. This research is focused on rainfall data available via the Met Office WOW platform, as the Met Office is the industrial partner for the research. Many PAWS manufacturers provide data sharing via their own propriety platforms, e.g. Davis Weather Link (Davis Instruments, 2018) and Netatmo Weather Map (Netatmo Weather Map, 2022). Other providers facilitate data sharing for users of any type of PAWS, but may place restrictions/fees on data access, e.g. Weather Underground (WU) (Weather Underground, 2019).

The Met Office Weather Observation Website (WOW) is available online for sharing and viewing weather data (see https://wow.metoffice.gov.uk/). It was established in 2011 as a digital repository for manual weather observations collected by trained volunteer climate observers. WOW was expanded to accept observations from weather stations operated by any registered user.

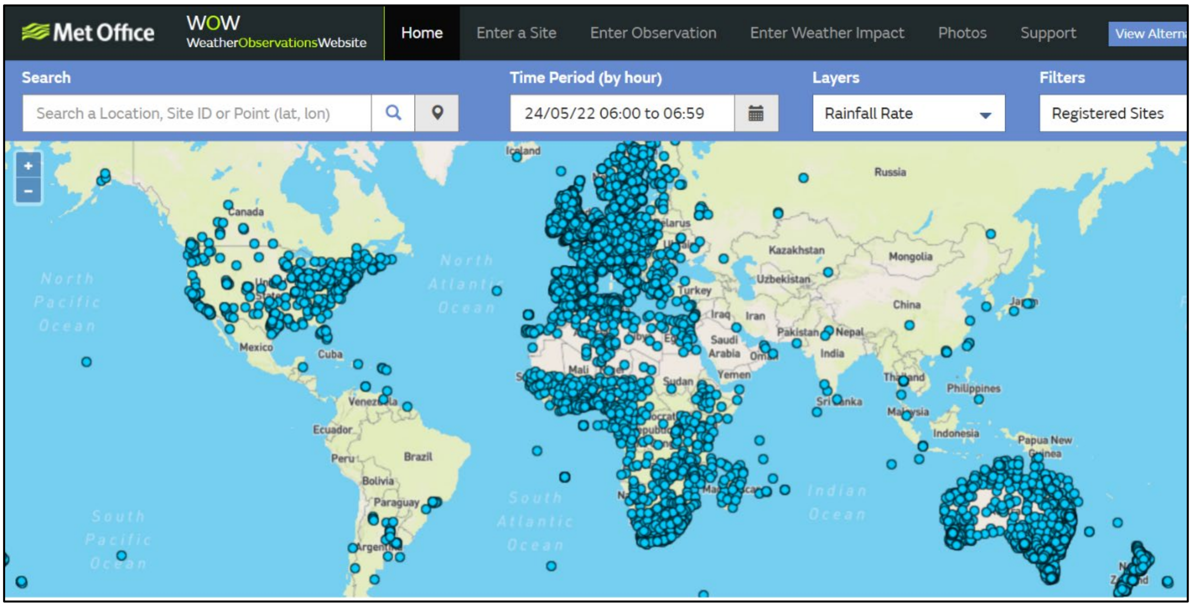


Figure 1.3 WOW screenshot showing the global distribution of registered citizen science weather stations (as of May 2022).

The platform accepts observations from manual and automatic equipment around the world. There are linkages with the Belgian, Dutch, Swedish, Irish, Australian and New Zealand meteorological services, resulting in relatively high numbers of weather stations reporting in the respective nations.

This research uses hourly (or shorter) interval observations from PAWS in Britain. The rationale for focusing on hourly/sub-hourly observations stems from the potential for citizen science to improve our understanding of spatially and temporally variable rainfall, that may not be captured in a sparse ground-based gauge observation network designed for monitoring large frontal weather systems. The spatial extent (Britain) was selected, as this is where WOW originated, and data were readily available from Official weather monitoring networks allowing for data quality comparisons.

1.5 Flooding in Britain

This thesis assesses the accuracy of rain gauge observations with a view to using the data in a range of hydrological applications, including flood forecasting and response. There are six types of flooding recognised in Britain, namely fluvial, pluvial, coastal, groundwater, sewer and reservoir flooding (Flood Guidance, 2022).

Figure 1.4 Photographs of different flood types from the UK - Top left: River Great Ouse in flood River Great Ouse, Pauline A Marsh / River Great Ouse in flood / CC BY-SA 2.0. Top right: 'Toon Monsoon, flooding in Newcastle Upon Tyne, 2012, chroniclelive.co.uk. Bottom left: Coastal flooding risk at Ilfracombe, JBA Consulting. Bottom right: Groundwater flooding, Compton, 2021, British Geological Society.

In northern Europe, convective storms are predicted to increase in intensity and number due to the warming effects of climate change, and extensive research has been undertaken to better predict and analyse flooding from convective rainfall (Parry *et al.*, 2007; Falconer *et al.*, 2009; Berg *et al.*, 2013; Flack *et al.*, 2019). As the research presented in this thesis primarily assesses rainfall data, it follows that the analysis and response to pluvial flooding is the most likely application; however, in locations where there is no/limited Official rain gauging crowdsourced data may prove a useful alternative, or addition to existing Official rainfall monitoring. In Chapter 5 the use of WOW rainfall data for post-event analysis of intense rainfall that resulted in pluvial flooding is considered.

1.6 Thesis Structure/Outline

This thesis is structed in the conventional manner, with this chapter providing the overview and introduction. Chapter 2 comprises a literature review, demonstrating why this research is required and how it fits into the existing published body of work on citizen science rainfall analysis.

Chapters 3 to 5 progressively demonstrate the stages of data acquisition, cleaning, validation, and application. Chapter 3 provides a detailed description of the raw data available from WOW and how theses rain gauge records were processed to create a useable dataset. Chapter 4 considers the quality of the WOW rain gauge data, using a comparable Official dataset for benchmarking, and incorporates a case study of an alternative approach to data collection. Chapter 5 presents a methodology for the selection of event based high quality rain gauge data, using radar data to compare data quality and for merging with gauge data. Case studies demonstrating the quality of citizen science data are included in this chapter, using gauge data blended with radar to improve the accuracy of the rainfall field during post event analysis of convective storms.

Chapter 6 examines how to get people involved in citizen science, reporting on several different approaches to promoting participation, and assessing the barriers that might limit participation or reduce data quality. Chapter 7 consolidates the conclusions and recommendations from Chapters 3 – 6.

1.6.1 Aims and Research Questions

The overarching aims of this research, and research questions addressed in this thesis are presented in each of the relevant chapters:

Chapter 3 – What, Where, When, WOW?

Aim: To examine if the spatial distribution of rainfall monitoring in Britain can be improved by including observations available from automated citizen science rain gauges reporting to WOW.

Q1. What observations are available from automated citizen science rain gauges reporting to WOW and how can they be wrangled into a useable format?

Q2. Where are PAWS rain gauges located within Britain?

Q3. To what extent would the inclusion of WOW stations improve the spatial distribution of rainfall monitoring in Britain?

Chapter 4 – Describing Crowdsourced Rainfall Observations and Assessing Quality using a Statistical Quality Control Algorithm

Aim: To describe crowdsourced rainfall data available via WOW and assess the quality as compared to rainfall data from Official ground-based gauge networks in Britain.

Q1. How good are WOW rain gauge data?

Q1a. How does the availability of WOW rain gauge data compare to that from Official rain gauges?

Q1b. What were the results of SQC on the WOW data and How does WOW data quality compare with Official data quality?

Q1c. Where are the post-SQC good quality WOW rain gauges located?

Q2. Is the method of SQC appropriate for WOW data?

Q3. How could the quality of WOW observations be improved?

Chapter 5 - Event Based Quality Control with Data Validation and Visualisation

Aim: To continue the statistical quality control process by applying manual quality control and cross checking with an external data source (radar) to validate the quality of WOW rainfall observations.

Q1. What manual checking can be used to improve the accuracy of WOW rain gauge data so it can be used in post-event analysis?

Q2. What impact does including WOW rainfall data have on the interpolation of rainfall during localised high intensity storm events?

Chapter 6 - *Using WOW Rainfall Data in Hydrological Modelling*

Aim: Assess the impact of the inclusion of WOW data in post-event hydrological analysis using a hydrological model.

Q1. What impact does including WOW rainfall data have on streamflow estimation in a hydrological model?

Chapter 7 - Observing the Observers – Motivations and Barriers to Sharing Data from Private Automated Weather Stations

Aim: To document the experience of citizen scientists in recording weather observations, considering the heterogeneity of participants regarding skills, knowledge and commitment levels to highlight motivations and barriers to participation.

Q1. What themes emerge that encourage or reduce participation in crowd sourcing weather observations, and are they common between groups of potential or active citizen scientists?

Q2. Can key elements that prevent or support the generation of good quality rainfall data via citizen scientists be identified?

Chapter 2. Literature Review

2.1 Introduction

The literature review provides context for this thesis, highlighting current research gaps, and establishing what research is required to fill these. The review considers the essentials of rainfall observations, with a focus on crowdsourced rainfall observation research carried out both in Britain and worldwide. The review also looks at associated topics including methodologies for the statistical quality control of rainfall observations and, research considering the drivers and barriers to citizen science participation.

2.2 Rainfall Observations

2.2.1 Ground Based Rain Gauge Monitoring networks

Officially operated ground-based rain gauge networks provide the essential rainfall observation data required for forecasting in Britain. The network evolved over time, primarily concerned with capturing observations of the large frontal weather systems that dominate British weather and cause much disruption, particularly in the winter (Met Office, 2019).

Although highly caveated due to access to records and definitions of Official monitoring, Kidd *et al.* (2017) calculated the rain gauges used for global precipitation monitoring cover roughly the area of a football pitch. Clearly then there is a need to interpolate between gauges to determine the rainfall in an area larger than, for example, the standard 5” (127 mm) opening of the classic Snowden MKII manual rain gauge. Various interpolation methods have been proposed and evaluated over many years (Shepard, 1968; Delfiner, 1975; Tabios III *et al.*, 1985; Dirks *et al.*, 1998). The selection of the most suitable method is highly dependent on several factors including the nature of the rainfall, the distribution/density of gauges, the computational capacity available for data processing, the speed with which data are required and the accuracy of the resulting dataset. As this research is focused on convective rainfall, further consideration is given here to the density of monitoring required to capture short-duration, high-intensity rainfall.

The significance of gauge density was explored by Schroeer *et al.* (2018) in an exceptionally densely gauged area of Austria, with 150 rain gauges in 300 km^2 forming the Wegener observation network. The researchers demonstrated that there can be >50% underestimation of short-duration high-intensity precipitation observed during convective storms when interpolating between gauges spaced >5km apart. A higher density of rain gauges is therefore required to accurately represent small-scale, high-intensity convective rainfall. Whilst it may not be practical for national monitoring agencies to install and operate gauges at the required density, the proliferation of citizen science rain gauges could present an alternative way of expanding the monitoring network.

2.2.2 Radar

Radar technology evolved over several decades after the physicist Heinrich Hertz first observed that electromagnetic waves could be focused into beams and reflected off objects in 1886 (Fabry, 2015). Radar was developed and widely employed during the second world war for the detection of ships and aircraft beyond visible range (Imperial War Museum, 2022). It was discovered that precipitation was visible via radar which made the detection of aircraft difficult, but equally provided valuable information about weather conditions. Radar is an indirect measure of reflection which is a function of particle size, type, and distribution. The reflection is converted to a rainfall rate using the equation:

$$Z = aR^b$$

Where, Z = reflectivity ($mm^6\ m^{-3}$), R = rainfall rate ($mm\ h^{-1}$), and a and b are parameters derived from raindrop size and distribution (Marshall, 1948; Collier, 1989).

Wilson *et al.* (1979b) described the fundamentals of using radar for precipitation detection, bemoaning *"…data are underutilized and both confusion and misunderstanding exist about the inherent ability of radar to measure rain, about factors that contribute to errors, and about the importance of careful calibration and signal processing.".* The complaint sounds familiar in the context of citizen science, indicating meteorologists have long been a cautious crowd. Wilson *et al.* (1979b) point out the importance of calibrating radar using point data from rain gauges, suggesting that one gauge per 1000 – 2000 km^2 is required, but stating that when high accuracy rainfall measurements are required with an average error of <10 – 20% then the number of gauges required for calibration is insufficient to provide the required accuracy.

Goudenhoofdt *et al.* (2009) note that merging methods are the final step in the processing of radar, when all else has been done to minimise error. They assessed radar/rain gauge merging divided into 'simple' and 'geostatistical' methods and conclude that the simple method of mean field bias (MFB) correction reduces the mean absolute error by 25%. A geostatistical method, kriging with external drift (KED), which primarily interpolates gauge data using radar to improve the spatial distribution of that interpolation, was found to be the most effective method, reducing mean absolute error by 40%. Sensitivity analysis undertaken as part of the assessment suggested that geostatistical merging methods were preferable regardless of the gauge density. The research considered data at the daily scale, which would not be appropriate for applications relating to rainfall during convective storms; however, the conclusions are supported by Jewell *et al.* (2015) in their research using UK data at the 15 to 60 minute scale looking at differing weather types ranging from large-scale slow-moving systems to fast localised heavy rainfall. The research highlights the limitation of merging radar and rain gauge at time intervals <15 minutes due to the mismatch in data accumulation and spatial representation.

A review by Ochoa-Rodriguez (2017), explores the pros and cons of radar merging techniques with particular focus on the temporal and urban scales that would support urban hydrological applications. The paper recommends the most accurate methods for merging rain gauge and radar are MFB, KED and Bayesian merging (BAY) and that further research is required into the sensitivities of each method. BAY was first proposed by Todini (2001), quantifying the estimation error covariance of radar and rain gauge observations before optimally combining the two. It can be particularly effective in areas with limited rain gauge coverage, or during convective events where the highest intensity rain may not have been observed in a gauge; however, it is computationally demanding (Ochoa-Rodriguez *et al.*, 2015).

An American Meteorological Society (AMS) conference paper details how radar evolved and how data processing is currently undertaken in the UK (Harrison *et al.*, 2015). In 2015 there were 14 C-band radars active in Britain (see Figure 2.1).

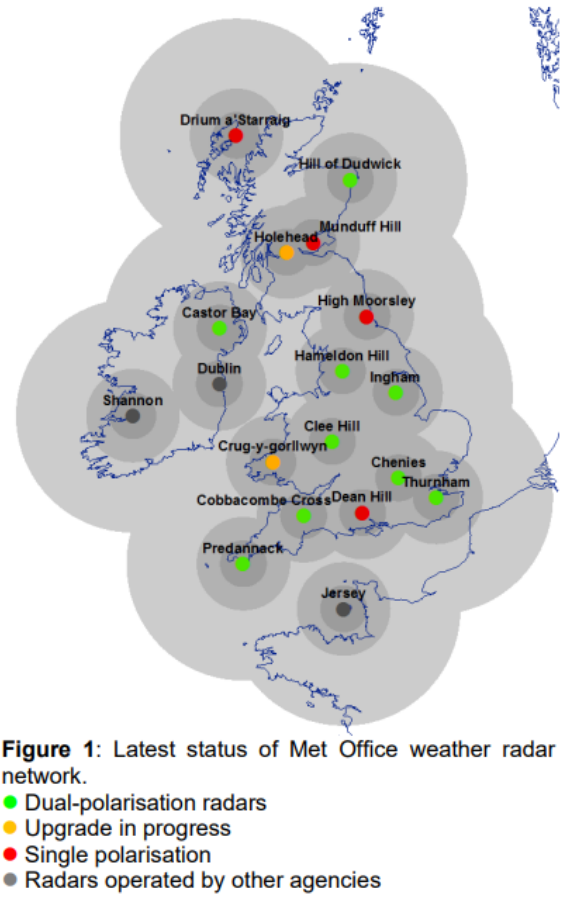


Figure 2.1 UK rainfall radar locations, after Harrison et al., (2015).

The stages of correction and quality control are outlined as:

- Clutter filtering
- Identification of partial beam blockage
- Correction of reflectivity for attenuation in precipitation

MFB adjustment is undertaken using rain gauges. Radar adjustment uses the criteria:

- the gauges must lie within 100 km of the radar;
- the radar must have detected precipitation within the collocated gauge pixel during most of a reporting hour;
- both radar and gauge must have measured >0.2 mm during the hour;

- The gauge must not be in an area where the radar suffers from the effects of ground clutter or anomalous propagation.
- Provided at least 50 rain gauge measurements meet the above criteria for a time window between 50 and 2000 hours, the gauge adjustment factor is calculated using the shortest possible time window.

In June 2021 there were 130 rain gauges being used in England for radar bias correction (pers. comm. Jake Brown, Met Office 2021). Whilst this may be a quick way to generate radar data for publication, as demonstrated by Goudenhoofdt *et al.* (2009) and summarised by Ochoa-Rodriguez *et al.* (2019) it is not the most effective way to achive accuracy during convective events.

2.3 Rain Gauge Data Quality

2.3.1 Fundamentals

McMillan *et al.* (2012) sought to benchmark observational rainfall uncertainties, stating that point measurements errors and spatial variability result in uncertainty in the accuracy of extrapolated rain observations. The principal issues impacting on rainfall data quality are; equipment accuracy; the suitability of the location and gauge set-up; and, the reliability and functionality of data transfer (Allerup *et al.*, 1980).

Automated rain gauges in the UK operated by official bodies are most commonly tipping bucket rain gauge' (TBR) (Met Office, 2006a; Environment Agency, 2021; Sepa, 2021) which tip once a known volume of rain has been collected in the gauge (commonly 0.125 - 0.5mm). TBRs can be linked to telemetry systems to allow for near real time data sharing. There are several issues with TBRs that may reduce the accuracy of data, including undercatch due to wind, wetting losses, and tipping errors (McMillan *et al.*, 2012; Pollock *et al.*, 2018).

The WMO compared the performance of a range of rain gauges in studies operating over several decades (1955 – 2008), with a particular focus on the impact of wind and evaporation on the accuracy of rainfall intensity (Sevruk *et al.*, 2009). In laboratory simulation of a consistent intensity of rainfall, TBRs were found to demonstrate errors greater than 5%. Subsequent field trials confirmed the conclusion that at high intensities, TBRs were not accurate (Lanza *et al.*, 2009). Given they are so common, it is most pragmatic to suppose that during the highest intensity rainfall events the observations they generate are likely to be an underrepresentation of the true rainfall.

2.3.2 Rain Gauge Installation

The WMO in their "Guide to Instrument and Methods of Observation" (WMO, 2018) emphasise the influence of wind on accurately capturing rainfall in a gauge. Pit gauges are recommended, as being more likely to capture rainfall, without wind-induced undercatch. The guide states that there should be drainage available, which may be problematic during a flood as the pit may not drain freely or could be inundated. Where pit gauges are not employed, the WMO recommend metadata be preserved for use during quality control including the height of the gauge above the ground, the height of any associated wind speed measuring instrumentation, and a detailed site description. Regarding exposure the guide states that there should be no objects closer to the gauge than twice the height of the gauge orifice i.e., a 30 cm tall gauge should be in a clear space >60 cm in diameter.

> *"The best sites are often found in clearings within forests or orchards, among trees, in scrub or shrub forests, or where other objects act as an effective wind-break for winds from all directions." (Ibid.)*

Similar conditions could well be encountered in domestic settings, where there are often boundary fences within a few meters of a potential gauge placement.

The Met Office provides a way to determine the site rating of a WOW rain gauge that can be shared via WOW metadata (see Figure 3.1).

2.3.3 Rainfall Data Statistical Quality Control

Statistical quality control (SQC) is required as part of the quality assurance process. There are multiple ways to apply SQC to rainfall data, and prior to using data it would be prudent to remove inaccurate observations. Approaches used in quality controlling official data and citizen science data are reviewed here. British Standard BS7483.2:2012 provides the basis (but not the specifics) of the procedures followed in the UK for the quality control of rainfall data from automated gauges at time steps from monthly to sub hourly. The Standard states;

"Quality control processes should be implemented to:

a) identify erroneous measurements of precipitation while distinguishing them from extreme precipitation events;

b) provide accurate estimates for missing or erroneous measurements; and

c) be applied in a consistent way over time."

The Standard notes the difficulty in applying SQC to short-duration convective systems, stating that the ideal accumulation period for effective SQC is 24 hours. It states that "*the principal method for the quality control of precipitation data should be based on the comparison of measurements from one rain gauge with those from neighbouring rain gauges."* Given the limited spatial extent of convective rainfall, or the wide distribution of gauges this is unlikely to be valid for all but the most persistent rainfall events. The fundamental variability of convective rain confounds SQC methods that rely on interpolation and nearest neighbour checks.

The World Meteorological Organisation (WMO, 2021) defines four types of errors from automated weather stations that can be found in weather station data, namely;

- Random error that can be an over or underestimation and are not dependent on the measured value.
- Systematic errors that tend to be biased either about or below the actual values, for example due to sensor drift.
- Large errors that are caused by equipment malfunctions or mistakes in data processing and are generally easily detected.
- Micrometeorological errors resulting from small-scale perturbations or weather systems affecting the observation that may give results that are discordant with surrounding observations.

Effective SQC methods need to be applied to data to capture and account for all these errors. The WMO suggests that once instantaneous checks have been completed the data can be flagged as:

- Good
- Inconsistent
- Suspect
- Erroneous
- Missing

Where data are good, inconsistent, suspect, or erroneous they should be passed to further checks (*Ibid*.).

Blenkinsop *et al.* (2017) lists 11 SQC measures applied in the creation of hourly gridded rainfall data for the UK. SQC steps included comparison against recorded maxima, noting multiple days reporting large accumulations, or regular reporting of high volumes at 9 am (traditionally the time that daily records are logged) and excessively long dry periods. The research emphasised the need for corroborative data to constrain results. An extension of the work by Blenkinsop was undertaken on British rain gauge data by Villalobos-Herrera *et al.* (2022).

In a paper published in the Journal of Hydrology Estevez *et al.* (2011) applied:

- Range/limit (fixed or dynamic) test
- Step test
- Internal consistency test
- Persistence test
- Spatial consistency test.

The assertion was that data not meeting the range test should be flagged as erroneous and that not meeting the other tests should be considered suspicious and be validated by manual inspection. In other research looking at data from PAWS (de Vos *et al.*, 2017), SQC steps applied were:

- Calculated the difference in cumulative daily rainfall with the previous time step
- Interpolated on a fixed timeline with constant steps where constant rainfall is assumed to remove erratic reporting time steps
- Discarded original intervals longer than 20 minutes

- Compared values of interpolated time series with median rainfall for all stations for each time interval and excluded values exceeding the median by more than 50mm h^{-1}
- Dry periods identified as >24h with zero rainfall – If a PWS reported continuous zero rainfall outside of the dry period it was assumed to be faulty.
- Determined inter-gauge correlations and discarded all results if average and median <0.21 was calculated between subject station and all other stations.
- Visual comparison with radar rainfall time series.

Qi et al. (2016) used an alternative method of SQC drawing together data from different sources to determine if gauge data can be considered accurate, part of the Multiple Radar/Multiple Sensor (MRMS) system utilised by the National Weather Service (NWS) of the National Oceanic and Atmospheric Administration (NOAA) of the United States of America (USA) (National Atmospheric and Oceanic Administration, 2015). The MRMS was specifically designed to support the forecasting of disasters, particularly flooding by combining data from radar, satellite, surface gauges, weather balloons, lightning reports, and predictive models. By drawing on a variety of data sources there is the opportunity to cross reference rainfall depths to derive a value that is most likely to be representative.

The literature makes it clear that the potential errors in rainfall collection and data transmission make it difficult to have certainty in the accuracy of rainfall observations. This is supported by Beven (2019) who asserts that hydrology is an 'inexact science' in part due to uncertainties in the accuracy of rainfall measurements. The inherent natural variability of rainfall means that interpolation can cause extrapolation errors. The highly variable nature of rainfall provides a strong case for the potential of citizen science data for filling gaps between Official observations; however, there is a need to ensure that such data are of a comparable accuracy to those obtained from Official sources if they are to be used in combination. Ultimately it is incredibly difficult to know whether any rainfall observations are accurate. SQC can remove wildly inaccurate observations; however, it is important to note that there can be value in retaining and considering extreme values in certain circumstances. Extreme rainfall is being observed globally (Parry *et al.*, 2007; Kendon *et al.*, 2014; Bevacqua *et al.*, 2019) so it is important not to be overzealous in the application of SQC, resulting in observations that may be of the most interest to researchers being disregarded. With climate change, new global and British weather records could be established in the coming years requiring reappraisal of the SQC measures of 'extreme'.

2.4 Citizen Science

As this thesis is primarily focused on citizen science it is important to have clarity on what that means. The definition of citizen science and the delineation of degrees of involvement has been researched at length (See *et al.*, 2016; Eitzel *et al.*, 2017). Haklay's widely acknowledged "*Levels of participation and engagement in citizen science projects"* starts with crowdsourcing with "*Citizens as Sensors"*, describing contributors who may not know they are participants in a citizen science project *per se* (Haklay, 2012). An pertinent example of this may be accessing temperature/rainfall data from weather stations (Muller *et al.*, 2015). Successive levels describe additional degrees of engagement and participation, up to Level 4 '*Extreme Citizen Science'* which describes collaboration including co-design of a project, data collection and analysis. Research into the importance of terminology in citizen science has shown how critical it can be, with some of the potential pitfalls highlighted in Figure 2.2.

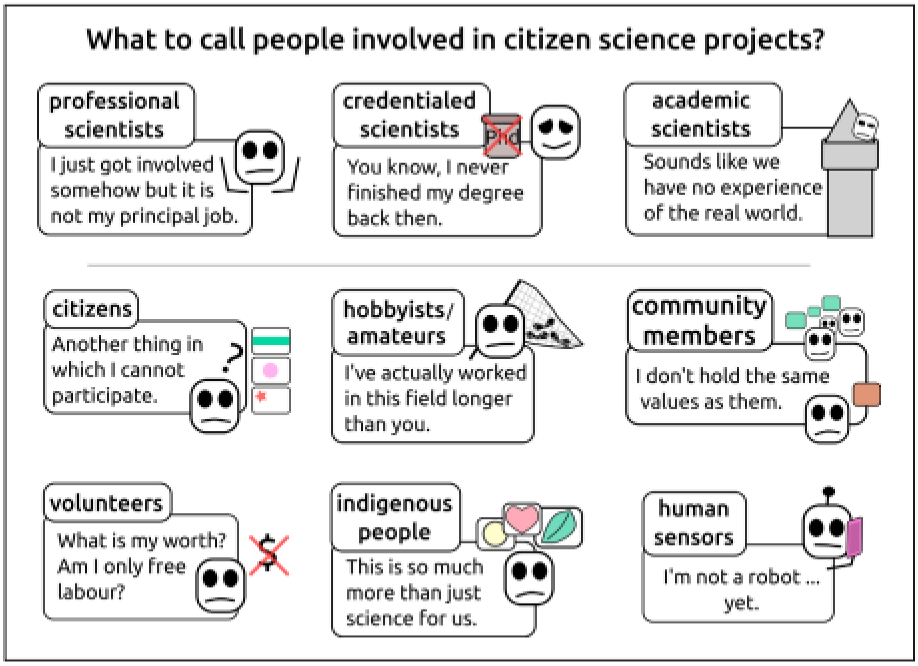


Figure 2.2 Illustrated examples of negative interpretations of commonly used names to describe people participating in citizen science, after Eitzel et al. Citizen Science Terminology Matters, 2017.

2.4.1 Citizen Science Motivation

The UK Environmental Observation Framework (UKEOF) looked at how citizen science works when gathering environmental data (Roy et al., 2012) and the motivations of participants (Geoghegan et al., 2016). The reports point out that citizen science participation is often related to enjoying knowing more about the natural world and making a valuable contribution to a scientific body of knowledge; however, motivations can vary widely, and it is difficult to design a participatory project that meets the needs of all contributors. Geoghegan et al. (2016) split participants into those gathering the data; the citizen scientists, and stakeholders in policy, practice and science fields; then assessed the motivations of the different groups. The report provides an excellent basis for any research into citizen science participation, highlighting the importance of communicating information, particularly via simplified depictions of data to communicate with non-specialists.

Goodchild *et al.* (2012) noted the advantages of volunteered geographic information, and issues with data quality. They presented ways to address data quality using approaches of participant cross validation, taking the consensus as acceptably accurate and/or using an understanding of the natural world to reality check contributions. Foody *et al.* (2015) also focused on crowdsourced contribution to mapping. An important conclusion was that the quantity of participants is not a critical concern, as they identified that a small group of core participants were able to generate more accurate contributions than a larger, possibly less dedicated group.

Research on users adding weather data to the WOW database concluded that participants felt an emotional connection to conducting citizen science, and that care should be taken in the aggregation of data to preserve that connection to sustain participation (Lin *et al.*, 2016). Reed (2008) talks about the need to emphasise 'empowerment, equity, trust and learning' when engaging volunteers in participatory environmental management, and that skilled facilitation is integral to participatory approaches.

2.4.2 Citizen Science/Crowdsourced Rainfall Data and WOW

There is limited published literature specifically relating to crowdsourced rainfall observations gathered in the UK, as mass digital collection of such data (generating a sizable body of widely accessible data) is relatively recent (e.g., the Met Office WOW platform started in 2011). Some researchers allude to WOW, but then focus on alternative citizen science/crowdsourced contributions such as tweets and photos (Lewis *et al.*, 2017).

The WOW platform is one of several ways data can be shared from PAWS. There are open platforms such as WOW and WU that accept data from a variety of weather stations. WU was developed by the University of Michigan but was sold to IBM in 2015, who subsequently commercialised the data. There are also proprietary platforms where data from a particular brand of weather station can uploaded (Davis Instruments; Netatmo Weather Map, 2022). The way data are accessed varies across all platforms, but all have limitations. It is possible to view the 'current' rainfall from PAWS reporting to Netatmo, Davis WeatherLink and WU, however, the reporting frequency is not divulged, any aggregation of observations is not noted, and it is assumed that hourly rates are depicted, based on an assumption as there is no clarity provided by the platforms. WOW offers more flexibility that the alternatives, by allowing the viewer to select a date and time interval to obtain a map of hourly rainfall for the given time, and a further facility allowing the viewer to select up to 30 days of data from a given station for download. WOW also provides comparison data with the closest Official station, which can be useful for making a rapid assessment of data reliability.

A review paper by Muller *et al.* (2015) looked at the uses and potential for crowdsourced data in atmospheric science, citing concern around QA/QC as the biggest challenge. They emphasised the increasing potential of crowdsourcing due to increases in connectivity and computing power. Given that the paper was drafted in 2014 it seems likely that these aspects have improved further in that time and the potential for such data still exists. The paper defines some key terms relating to the degree of participant engagement; a distinction is made between 'passive' and 'active' citizen science, with the prior being an automated process whilst the latter involves a degree of engagement. Additionally, the researchers define 'automated' as being without human intervention, 'semi-automated' as meaning data are collected automatically then uploaded manually, and 'manual' where both data collection and submission are done by a person. WOW would be defined as a passive automated approach if defined using these terms.

There can be a transience to crowdsourced/citizen science projects. Muller *et al.* (2015) provide an overview of 27 current related projects, of which 23 had web addresses. In June

2022, 16 were still operationally collating and sharing weather data (one was an archive) and the remainder appeared to have closed.

Research conducted by Bell *et al.* (2013), two years after WOW became available to the public, showed more citizen scientists uploading data to WU than WOW. Their paper included mapping of the spatial distribution of weather stations in the UK and Ireland including weather stations reporting to both platforms. The researchers considered the difficulties citizen science weather observers faced, emphasising concerns around data quality, and highlighting the challenge of achieving optimum siting of PAWS in domestic settings. Some comparative analysis of PAWS parameters was undertaken, but there was no assessment of rainfall data quality. The researchers advocated nearest neighbour checking for data validation.

Citizen Science rainfall data collection can be successful when done on a large scale. In the USA, Canada and the Bahamas the Community Collaborative Rain, Hail and Snow Network (CoCoRaHs) was established in 1998 to promote citizen science data collection (Community Collaborative Rain Hail and Snow Network, 1998; Bell *et al.*, 2015). The observations made by volunteers who upload data to a dedicated website are used by meteorologists, hydrologists, emergency managers and other organizations that rely on weather data. There may be a higher demand for citizen science data in geographically large countries like the USA, where there is an information demand but not the same density of formal monitoring as in the UK.

In places where such a comprehensive official monitoring network is available, data from PAWS could provide the detail required for capturing convective precipitation. Researchers in the Netherlands looked at data from 63 weather stations over a 4-month period and investigated whether the PAWS data could provide the high spatial and temporal density of observations needed for urban hydrological modelling (de Vos *et al.*, 2017). The distribution of Official weather stations in the Netherlands is relatively sparse (31 in the country) however, the addition of PAWS data improved spatial delineation and quantification of rainfall.

Wind speed and direction, temperature and pressure observations from WOW have been used by researchers to facilitate fine-scale analysis of a hailstorm that affected northern England in 2015. Clark *et al.* (2018) showed that PAWS observations provided detail of storm structure that was not visible from Met Office automated weather station (AWS) observations. A bias correction factor derived from Official data was applied to temperature, pressure, and wind speed observations prior to using the WOW data. It is notable that the researchers did not consider rainfall in their work, potentially indicative of the difficulty in applying bias correction to the highly variable rainfall parameter.

Holley *et al.* (2020), investigated a convective storm in Norfolk using a combination of data from EA gauges, gauge data from citizen scientists reporting to WU, and data from gauges located on farms (solicited directly from farmers). In total there were 115 citizen science gauges and 37 Official observations. They calculated the difference between the radar and the gauge data, interpolated the difference onto a regular grid then applied that as a bias correction to the radar. No quality control was applied to the gauge data prior to use. The research did not demonstrate the suitability of applying a correction factor. A discrepancy of 62 mm hr^{-1} between observations from gauges within 500 m was assumed to be either a true representation of rainfall variability, or potentially due to inaccuracies from rapid tipping of the TBR.

Research conducted on hydrometric and meteorological citizen science in a previously ungauged catchment in the North East of England (Haltwhistle Burn) demonstrated the potential of citizen science at the local scale (Starkey *et al.*, 2017). Quantitative and qualitative data were collected by members of the public to capture spatial and temporal river response to rainfall, then used in a physically based, spatially distributed hydrological model (SHETRAN). The research found that, during extreme events, citizen science data was more accurate than those available from Official networks, that citizen science data was valuable in filling gaps between Official monitoring, and that it could be applied to the validation of data. The research involved intensive interaction between the researcher and citizen science participants, which is not always practical.

2.4.3 Citizen Science Rainfall Data Quality

As noted in the abstract, whilst WOW has existed for over a decade there are barely any publications documenting data use. This is despite the recognition of the value of citizen science and international promotion efforts (Buytaert *et al.*, 2014; See, 2019; IHP, 2021). Data quality is frequently questioned in research using citizen science; therefore, this section considers key elements of rain gauge data quality.

Research focused on specific issues relating to the data quality from PAWS has been undertaken, highlighting that there are many ways in which data from PAWS may be compromised. Bell *et al.* (2015) conducted a year-long comparative study between 5 commercially available citizen science weather stations and professional Met Office equipment to determine whether the citizen science data was sufficiently accurate. Test parameters included temperature, humidity, dew point and rainfall, and all were found to exhibit substantial bias. The study aimed to parametrise bias and apply corrections; however, such an approach would not be possible were a suitable control weather station not available. The research concluded that the citizen science rain gauges tested were best suited to recording the intensity and timing of rainfall, whilst being less reliable for deriving accurate total rainfall volumes.

Contrary to the study undertaken by Bell *et al.* (2015), research focused on Amsterdam assessed the accuracy of Netatmo rain gauges when compared to the standard gauge used by the Royal Netherlands Meteorological Institute (de Vos *et al.*, 2017). The researchers demonstrated that there was little difference in the reported volumes of rainfall when the raw data from PAWS were compared to Official gauge data; however, there were errors introduced during data transfer due to issues with time step reporting. Error was introduced where the upload frequency or time steps differ from the reporting frequency from the rain gauge and can also complicate data processing when time steps vary widely.

Butler (2019) compared the different software available for the uploading of PAWS data to various online platforms. Opinions were also offered on the relative merits of the platforms available including Weather Underground and WOW. The conclusion was that there are several different options available to owners of PAWS: however, there was no critical assessment of the data available, other than a cursory look at the inbuilt presentation options e.g., graphical representations of results. The research highlighted the importance of usability, especially where the users are not commonly experts.

An SQC method was devised by de Vos *et al.* (2019), specifically for the real-time SQC of citizen science rain observations from Netatmo rain gauges. The method was developed using an urban data set from Amsterdam and a national dataset covering the Netherlands. The method relied on a reference dataset that extends for at least one month prior to the period of interest, derived from 2 c-band radars, validated using 356 rain gauges, at a 1Km and 5-minute resolution. The study documented sources of error in the data, including, gaps in time series; bias due to wind induced undercatch or calibration; under reported values due to blockages of the tipping mechanism; over reported records due to operator handling/cleaning etc.; and erroneous reports from a single station suggesting something fundamentally incorrect about the set-up or performance of the gauge.

The SCQ method comprises 4 stages, with errors identified by cross comparison with neighbouring rain gauges. There are three filters applied to identify false zeros, erroneously high rainfall and PAWS consistently reporting erroneous observations. The filters as presented rely on a high density of rain gauges, with the researchers indicating that the filters will be less effective where there are fewer than 5 rain gauges within 10km of the station under question. The method provided by de Vos *et al.* (2019) concludes with the application of bias correction based on known biases with Netatmo PAWS that can be revised during the SQC process to account for peculiarities unique to each individual rain gauge. This may be valid in study areas with consistent climatology where observations are generated by rain gauges of the same brand (hence similar biases can be expected), however the application of bias correction coefficients derived from rain gauges at potentially highly variable elevations across a wider area, and between differing brands of rain gauge has not been tested. The Netherlands consists of a land mass spanning 41540 Km^2, and is famously flat with elevation ranging from -7m to 322m above sea level (The World Bank *et al.*, 2020; Country Reports, 2023). For comparison the land mass of Britain is 230,147 km^2 with several mountain ranges in Wales, Scotland and England and a maximum elevation of 1,343m (Ben Nevis, Scotland) (*Ibid*.). The UK Hadley centre have divided Britain into 8 regions for the purposes of presenting regional variations in precipitation, further emphasising the variability in precipitation across Britain as compared to the Netherlands, which mean that the methods applied to Amsterdam is less likely to generate robust quality control for observations from rain gauges in Britain (Met Office, 2016a).

An alternative approach to SQC is presented by Bárdossy *et al.* (2021) who argue that PAWS observations are generally order-correct, if not absolutely accurate. Their research focused on an area of Germany approximately 36 000 km^2 in size, distributed throughout which were 111 weather stations operated by the German Meteorological Service (DWD), referred to as the Primary Network, and 3082 PAWS, referred to as the Secondary Network. The topography of the study area was more variable than the Netherlands, with a mountainous area and peak of

1,496m. The method developed required one year of PAWS observations, and due to the potential impact from freezing precipitation in the unheated PAWS gauges the analysis was restricted to observations made when the air temperature at the nearest Primary Network gauge was >5°C. There were three main steps in the SQC, comprising nearest neighbour consistency checks, the correction of individual station bias and the application of event-based outlier detection that predominantly removed false zeros. The correction of bias relies on those PAWS remaining after inconsistency filtering to generate consistently biased observations. The researchers recognise that this may not be true where wind occurs with rain, however, their results show a more detailed depiction of rainfall when data from both the Primary and Secondary Networks are interpolated. The research does not include cross validation with external sources, such as the radar used by de Vos, *et al* (2019). As with the previous method, it appears to be most suited when wishing to undertake SQC on relatively high-density PAWS networks, and also requires a robust primary network of high-quality rain gauges to facilitate bias correction. These two methods are therefore not ideal for application to WOW PAWS rainfall observations given the relatively low spatial density of PAWS reporting to WOW in many locations in Britain, and the potential for variation in biases introduced by the varying brands of weather station reporting to WOW.

Some PAWS parameters available from WOW undergo basic SQC checks prior to publication. There are validation checks on wind, whereby the direction must be expressed in degrees and therefore can only be 0 – 360, and temperature must fall within the range -50 to 50 degrees centigrade. When observations are accessed via the web platform there is an option to report a suspect station, triggering a manual review by a member of the Met Office team. There is currently no mechanism for automated SQC of rainfall data (*pers.* comm. Katherine O'Boyle, Met Office, March 2019).

2.4.4 Citizen Science for Flood and Catchment Management

Uncertainty in quantifying rainfall has an impact on accuracy further down the actual and metaphorical stream, leading to difficulties in flood forecasting, flood response and post-event analysis. Citizen science has been mooted to expand monitoring; however, a limitation of citizen science rainfall observations identified by Starkey *et al.* (2017) was that outside of extreme events participation rates drop markedly. The research was successful in achieving high rates of participation and engagement by developing a close relationship between the researcher and the community. Whilst participation rates may be high initially, or at times of extreme events the sustainability outside of these periods is questionable, suggesting that citizen science applied in this way is appropriate for supplementing, but not replacing, Official monitoring.

The EA has undertaken research into the best way to work with volunteers for effective Flood and Coastal Risk Management (FCRM) outcomes (O'Brien et al., 2014). The value of having volunteers provide rainfall and river level data with the benefit of local knowledge, which aids understanding of the local environment was a key finding of the research. The EA identified volunteer participation motivations for citizen science were primarily personal risk or experience of flooding, along with a technical interest in the data and a desire to support the community. Models of participation that followed a 'bottom up' approach that could be adapted to the specific community context and were multifaceted (i.e., may include other elements such as biodiversity promotion) were found to be the most successful, possibly because they fostered a sense of community spirit and broadened the appeal of citizen science.

There has been research into engaging citizen scientists in activities that would contribute to flood management, ranging from communities contributing local knowledge to the understanding of flood risk, to data gathering (McCallum *et al.*, 2016; Sy *et al.*, 2019; Puttinaovarat *et al.*, 2020). The motivations for citizen science participation in relation to water resources management are assessed in Buytaert *et al.* (2014), particularly in respect to remote areas that perhaps otherwise would not be monitored. There is discussion around the value of sensitizing participants to changes in their environment when they play an active role as observers, and how that can propagate action and behaviour change.

Research conducted by Kutija *et al.* (2014) looked at how citizen science data in the form of crowdsourced photographs and eyewitness reports could validate modelling of a significant surface water flood in Newcastle in 2012. The researchers used CityCAT, a hydrodynamic model designed for use in urban areas (Glenis *et al.*, 2018), to model the flood based on official reports, then used citizen science contributions for validation. The model output was considered a fair representation of flooding, except for road underpasses where accumulated water exceeded the modelled levels. The research highlights the potential for varied forms of citizen science/crowdsourced data, combined with quantitative rainfall measurements for flood model validation in urban areas.

2.5 Conclusion

The literature review has demonstrated that the Official British ground-based rain monitoring network is insufficient to adequately characterise convective storms and alternative methods such as radar are unreliable without gauge-based correction.

The WOW archive contains rainfall data from citizen scientists that may fill the gaps in the Official monitoring network. There is a need to assess and report on the quality of the WOW data. Establishing rainfall data quality is complicated due to the variety of inaccuracies that may exist; therefore, it is important to use an appropriate and comprehensive method of SQC. Research conducted to date has generated conflicting accounts on the accuracy of PAWS rainfall observations, in part due to the different methods employed and the degree to which observations have been scrutinised and validated. The lack of a reference dataset, and the inherent variability of rainfall observations, both spatially and temporally make it a particularly difficult parameter to assess.

This review has shown that citizen science is often presented as a way to engage communities with the environment and obtain scientific data. The potential for PAWS data means that encouraging people to participate could be desirable, but that can be a laborious process. Understanding the factors influencing citizen scientists will assist in developing future engagement and building confidence in the outputs.

This research is focused on Britain, but the circumstances are not unique, there are very few places globally where the Officially administered rain gauge network is sufficiently dense to accurately record rainfall during convective storms.

Chapter 3. What, Where, When, WOW?

10 years of Crowdsourced Weather Observations from the Met Office Weather Observation Website.

3.1 Overview

This chapter introduces crowdsourced rainfall data shared via the Met Office Weather Observation Website (WOW) over 10 years (2011 – 2020). The aim was to examine if the spatial distribution of rainfall monitoring in Britain could be improved by including observations available from automated citizen science rain gauges reporting to WOW. This was achieved by processing the available observations into usable datasets, mapping the locations of rain gauges with respect to official monitoring networks, and assessing the change in gauge density due to the inclusion of WOW observations.

The first step was to generate a data set of hourly rainfall depths that could be used for the research presented in this thesis, and by other interested parties. Three datasets were generated comprising; multi-parameter raw data in an easy-to-use format; rainfall data at the original reported time interval; and an hourly rainfall dataset. The process of compiling and preparing data is presented in this chapter detailing; how data were made more user friendly by reformatting from time wise (monthly files) to station wise (station files); how stations were selected for inclusion in this research; and, how rain depth from accumulation or rate was calculated. Issues identified during data processing are highlighted and discussed in this chapter.

The code used to wrangle WOW data and the resulting datasets are available via links presented in Appendix 1. The quality of data is assessed and discussed in Chapter 4 and potential applications are demonstrated in Chapters 5 and 6.

3.2 Introduction

The aim of this chapter is to examine if the spatial distribution of rainfall monitoring in Britain can be improved by including observations available from automated citizen science rain gauges reporting to WOW. It details the stages of generating usable datasets from a bulk download of data provided by the Met Office by consolidating and cleaning observations.

The research questions are to determine:

Q1. What observations are available from automated citizen science rain gauges reporting to WOW and how can they be wrangled into a useable format?

Q2. Where are PAWS rain gauges located within Britain?

Q3. To what extent would the inclusion of WOW stations improve the spatial distribution of rain monitoring in Britain?

3.3 Data Access and Collation

3.3.1 Data Upload and Download Mechanisms

Data can be uploaded to WOW from any PAWS. Connection instructions for different brands are available on the WOW help pages (https://wow.metoffice.gov.uk/support). PAWS data upload to WOW can be automated or manual upload is also possible.

WOW weather observations are available;

- from WOW via the homepage;
- using a registered developer account; or,
- by request from the Met Office.

It is possible to view observations via the website, and once registered as a user it is possible to download available parameters for a range of up to 28 days from a selected station. Due to these limitations, for this research WOW observations were provided via the Met Office using an in-house bulk download API. At the time of writing there was no public access to the bulk WOW archive. Observations from the UK and Ireland reporting to WOW were provided from when records began in January 2011 to the end of December 2020. The full dataset consisted of 101 files totaling 85 gigabytes, that were processed using python. Data from PAWS located in Britain were selected for this research.

3.3.2 Available Parameters

WOW records were provided as date and time stamped reports. A report comprised all parameter observations for a given station at a given time. A total of 55 parameters (including hazard warnings) could be contained within each record. The complete parameter list can be viewed in the "Data Formats and APIs" section of WOW support pages as a .csv template for data upload.

The parameters in each report are determined by the operator. In addition to weather observations, each report included a unique identifier, station ID, latitude, longitude, and a date and time stamp. No units were specified in the reports; however, there is a list of "Standard versus Alternative Units" on the WOW support webpage (see https://wow.metoffice.gov.uk/support/dataformats#dataFormats).

3.3.3 Station Metadata

PAWS operators can choose to add station metadata to WOW, including assessing location attributes to generate a site rating. There is a guide to finding the site rating, based on schemes used by the Climatological Observers Link (COL), the World Meteorological Organization (WMO) and the Met Office. The key requirements for the ideal positioning of a weather station are presented in Figure 3.1.

Measurements of rainfall

A: Standard ''five inch'' manually-read rain gauge or calibrated tipping-bucket rain gauge, at standard height above ground (30 cm), site exposure minimum = 3.
B: Standard ''five inch'' manually-read rain gauge or calibrated tipping-bucket rain gauge, the rim mounted at standard height above ground (30 cm), exposure = 2 or 3.
C: Standard ''five inch'' manually-read rain gauge or calibrated tipping-bucket rain gauge, the rim mounted at standard height above ground (30 cm), exposure 1 or less.
D: Non-standard rain gauge and/or tipping-bucket rain gauge, exposure 1 or less.
U: Instruments unknown or not stated.
0: No rainfall measurements made at this site.

STANDARD INSTRUMENTS in this context means: Standard-pattern (Snowdon or Met Office Mk II pattern) ''five-inch'' copper rain gauge, with deep funnel, the rim of the gauge level and mounted at 30 cm above ground level, meeting the minimum exposure requirement of being at least 'twice the height'' of the obstacle away from the obstacle.

Exposure

5: Very open exposure: no obstructions within 10h or more of temperature or rainfall instruments.
4: Open exposure: most obstructions/heated buildings 5h or from temperature or rainfall instruments, none within 2h.
3: Standard exposure: no significant obstructions or heated buildings within 2h of temperature or rainfall instruments.
2: Restricted exposure: most obstructions/heated buildings >2h from temperature or rainfall instruments, none within 1h.
1: Sheltered exposure: significant obstructions or heated buildings within 1h of temperature or rainfall instruments.
0: Very sheltered exposure: site obstructions or sensor exposure severely limit exposure to sunshine, wind, rainfall.
R: Rooftop site: Rooftop sites for temperature and rainfall sensors should be avoided where possible.
T: Traffic site: equipment sited adjacent to public highway.
U: Exposure unknown or not stated.
Exposure ratings relate to the site of the temperature and rainfall instruments only, which should ideally be at ground level. Sensors for sunshine, wind speed etc are best exposed as freely as possible, and rooftop or mast mountings are usually preferable. Exposure guidelines are based on a multiple of the height h of the obstruction above the sensor height; the standard is a minimum distance of twice the height (2h).

Figure 3.1 Instructions for determining WOW site rating for a rain gauge.

The built-up nature of many domestic settings means it can be difficult for operators to achieve a high site rating using the WOW schema. Many PAWS are sold as a single unit combining; rain gauge; wind gauges; and, temperature monitoring, which have different requirements for optimal siting (see Figure 3.2). The combined PAWS are convenient for home use, but potentially sacrifice achieving the most accurate observations.

Figure 3.2 Photographs of example weather station types and set-up (from left to right: Davis Vantage Pro 2 combined parameter weather station, Davis Vue weather station, and a Netatmo rain gauge).

Where shared by the operator, information on the station type and positioning can be accessed on a station-by-station basis via WOW. Station metadata are held separately from time series observations and were not available in bulk for use in this research[2].

3.3.4 Observation Intervals

A potential benefit of WOW data is the provision of hourly or shorter interval data. The interval between observations can be critical, for example when studying intense, short-duration convective storms that may be over within an hour (Eggimann *et al.*, 2017; Schroeer *et al.*, 2018). The reporting interval of WOW observations is not standardised. It is user determined and may range from 1 minute to >daily. After an initial period of data collection, the Met Office limited the accumulation of data, enforcing a minimum interval time between observations of 2 minutes to reduce the accumulation of extremely large datasets.

During exploratory analysis it was noted that the interval between reports can change for a given station. This variation appeared to be due to changes made by operators e.g., switching

[2] Metadata was accessed on an individual basis via WOW where additional insight was required.

from hourly to half-hourly reporting, where the reporting frequency was regular. There were also examples of irregular intervals, that appeared to be outside of the control of operators and may have occurred during data transfer. These inconsistencies made the data hard to handle, as it was not possible to make assumptions about the characteristics of a PAWS record e.g., that the station was reporting at daily intervals. In response to these inconsistencies a series of criteria for station selection were determined by trial and error and are presented in section 3.4.4 . The criteria were designed to select only stations with hourly or shorter interval observations, that were operational for at least one month.

3.4 Dataset Creation; Consolidation and Cleaning Methods

3.4.1 Summary

The focus of this research were PAWS within Britain reporting rainfall observations at an hourly or shorter interval. An overview of data processing is shown in Figure 3.3. Methods for data selection and transformation were devised as part of this research and are presented in this section.

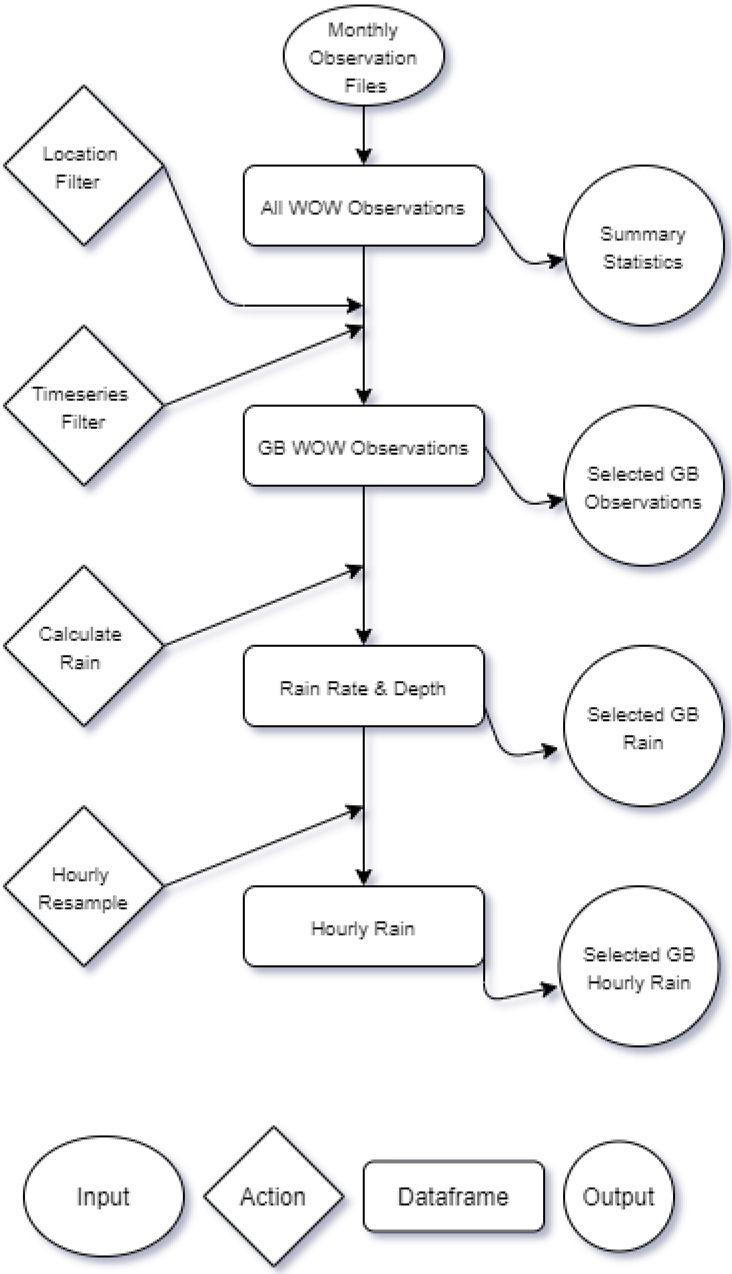


Figure 3.3 Flowchart of data processing in this thesis.

3.4.2 Multiple Parameter Records at Original Time Intervals

The first step towards creating an easily usable dataset required changing how data were stored and only selecting PAWS within Britain. Data were provided as all observations from specified time periods (varying from monthly to bi-annually), so were re-indexed to being stored as all time periods for a specific station. This was done to facilitate collecting station statistics and to allow for quicker and easier data selection at subsequent steps.

Parameters were chosen from the possible 55 provided based on how likely they were to be included in a record, and how useful they would be in subsequent analysis. The selected parameters were those most commonly shared, and consisted of air temperature (^{O}C), min. temp. (last 24hr), max. temp. (last 24hr), pressure at station (hPa), relative humidity (%), rainfall rate (mm/h), rainfall accumulation (mm), wind speed (knots), wind direction (degrees), wind gust (knots), wind gust direction (degrees). Stations within the mean high-water extent of Britain were selected for further processing.

3.4.3 Station Summary Statistics

PAWS summary statistics were generated to help understand what data were available and to facilitate the filtering process including:

- location (latitude and longitude – WGS 1984),
- count of the total reports,
- counts of reports per year,
- counts of observations for air temperature, dew point, pressure, relative humidity, rain rate, wind speed, wind direction and rain accumulation,
- start and end date of records,
- duration of reporting in days,
- mean, mode and minimum intervals between reports.

A summary table containing the statistics is available online (see Appendix 1).

3.4.4 Selecting Stations

When undertaking preliminary exploration of the data it was clear there were peculiarities. Anomalies were noted with the duration of reporting, with some stations having very short operational records, and the time intervals between observations could be variable or erratic.

The WOW instructions encourage users to set up a test station to ensure data are being shared as intended; however, it appeared that these test stations remain in the exported data. Without immediate access to the station metadata (where the station name may highlight the test status), a method of filtering stations was developed. Some assumptions were made that would limit the inclusion of test stations, or those unlikely to be providing ongoing hourly rainfall data.

The aim of filtering was to ensure only genuine stations providing a usable amount of hourly (or shorter) interval data were included in generating statistics and for further analysis. The filters were applied to the PAWS summary statistics. That list was used to select timeseries data from corresponding stations.

The following filters were applied to the station summary data table:

1. rainfall accumulation or rainfall rate > 600 records
2. reporting duration of >28 (days) and minimum interval between reports of < 65 (minutes) and maximum annual records > 365 (count) and mean interval of reporting < 65 (minutes).

Selecting Rainfall Data - The first filter selected stations with > 600 rain records. If there were less than 600 records the station was removed, on the basis that hourly reporting for a month (28 days) would generate 675 records. The reduction from 675 to 600 was to allow for lost hours in reporting[3].

Duration – If the duration of the record was less than 28 days it was assumed the data would not be particularly useful, and the station may have been set up as a test.

[3] The filter did not consider the observation at this point; therefore, there may have been stations where the depth of rain recorded by the station was zero.

Minimum Interval – The minimum interval between repots had to be <65 minutes to remove stations not reporting at hourly or sub-hourly intervals.

Modal Interval – a minimum modal interval of 65 minutes was used to remove stations reporting daily.

Annual Records – The maximum annual records from each station was calculated, if the maximum was less than 365 it indicated that the station was probably reporting daily data, or was not operational for long; therefore, these stations were removed from further analysis.

3.4.5 Calculated Rainfall Observations at Original Time Intervals

Individual PAWS files with the following variables were generated for the filtered stations:

- Station ID,
- Latitude,
- Longitude,
- Rain Rate (mm/h)
- Rain Accumulation (mm)
- Rain Amount (mm)

Rainfall depth was calculated from rain accumulation as the WOW instruction pages indicate that rain accumulation should constitute the reporting of rain. During exploratory analysis it was observed that rainfall was accumulated from one report to the next and reset daily. The time of the reset varied, with some operators using the British meteorological standard reporting day of 9 am GMT, whilst others used midnight (commonly the default for PAWS). Rain per report was calculated by deducting the rain at time t +1 from the rain at time t, with a correction where rain fell across the reset interval (see Appendix 1 for links to the code).

Prior to calculating the depth of rain at each interval it was necessary to remove duplicate reports discovered during exploratory data processing. There were many instances of weather stations sending simultaneous reports, potentially due to different sensors reporting, i.e., the wind vane generating a different report to the rain gauge. Where this occurred, it was noted that the rainfall accumulation field reported no accumulation of rain; therefore, records were sorted by date and time (ascending) and rainfall accumulation (descending) with the first record (containing any rain observation) being retained.

3.4.6 Rain Observations at Hourly Interval

Observations were available at a variety of time intervals ranging from 1 minute to 60 minutes. It was noted during data cleaning that the reporting interval could vary for any given station. For the purposes of applying SQC and to simplify the data rainfall observations were aggregated to hourly. There are applications where sub-hourly data are desirable, therefore the original data were retained and could be reviewed should more detailed analysis be required.

Data were resampled using the pandas library in python. The available observations were summed for the given hour, with the depth of rainfall assigned to the hour following the time at which it fell, in accordance with meteorological practice (e.g., all rain falling from 10:00 to 10:59:59 was assigned to 11am). The creation of this dataset included a review of the available rainfall depths, and the stations that did not report any rainfall within the whole record were removed i.e., all observations were 0 mm. Data in the accumulation field was primarily used to create an hourly dataset. If there were no records of rainfall accumulation, the rain rate was used as the alternative.

The Met Office recommends that rainfall accumulation be recorded, however it was noted during exploratory analysis that some PAWS only reported rain as a rate. When resampling from the original time interval to hourly the mean rate was taken, whilst the sum of the rainfall depth within the hour was used.

3.5 Dataset Description

3.5.1 Selecting British Stations

There were observations from 4,172 PAWS within the dataset provided by the Met Office. The locations are presented in Figure 3.4. The dataset included PAWS outside Britain (including Northern Ireland, the Republic of Ireland, and the Isle of Man).

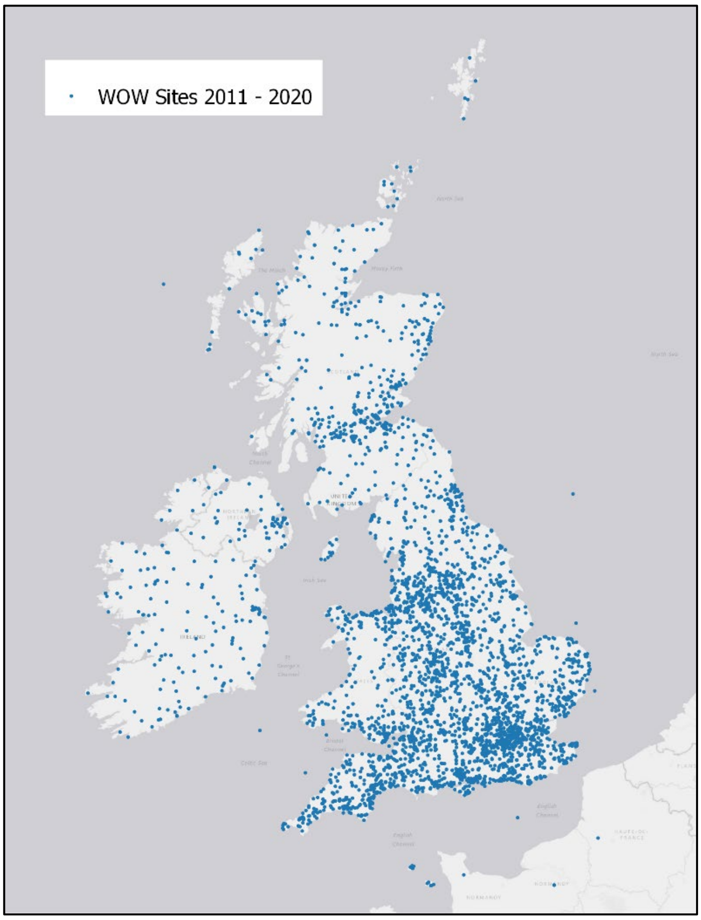


Figure 3.4 Met Office raw data PAWS locations from Britain, the Isle of Man, Northern Ireland, Republic of Ireland and the Channel Islands.

There were 3,920 PAWS within Britain. The distribution of British PAWS is shown in Figure 3.7.

3.5.2 Filtered Raw Multi-Parameter Data from Britain

PAWS within Britain were assessed based on the criteria described in section 3.4.4 with only those reporting hourly or sub-hourly for at least one month selected. The filtering process removed 1,203 (31%) of British PAWS, leaving 2,717 for further analysis. The summary statistics for selected stations have been published and are available via links in Appendix 1. The summary table does not include observation data but serves as a handy lookup to establish whether WOW data may be available.

3.5.3 Station Count and Longevity of Record

The number of selected PAWS reporting to WOW varied with time (Figure 3.5), peaking at 1,375 in 2016. The number of stations post-2014 was relatively stable, exceeding 1,200. This figure conceals a high degree of turnover, with new stations added and others no longer reporting.

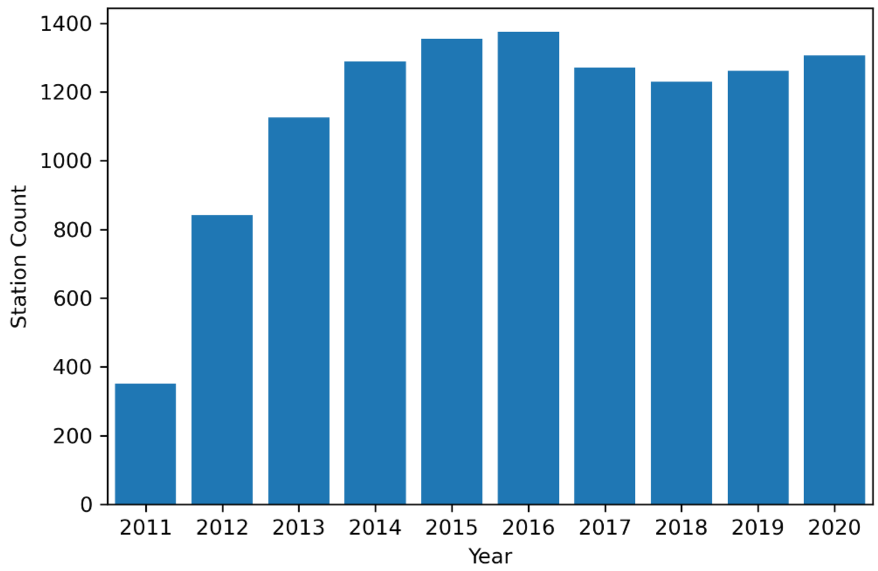


Figure 3.5   Count of PAWS in Britain reporting to WOW each year (2011 – 2020).

The mean duration of PAWS reporting to WOW was 3.6 years. There were 183 (7%) PAWS reporting for 9 or more years and 29% reported for 5 or more years. 643 (24%) of PAWS reported for less than 1 year.

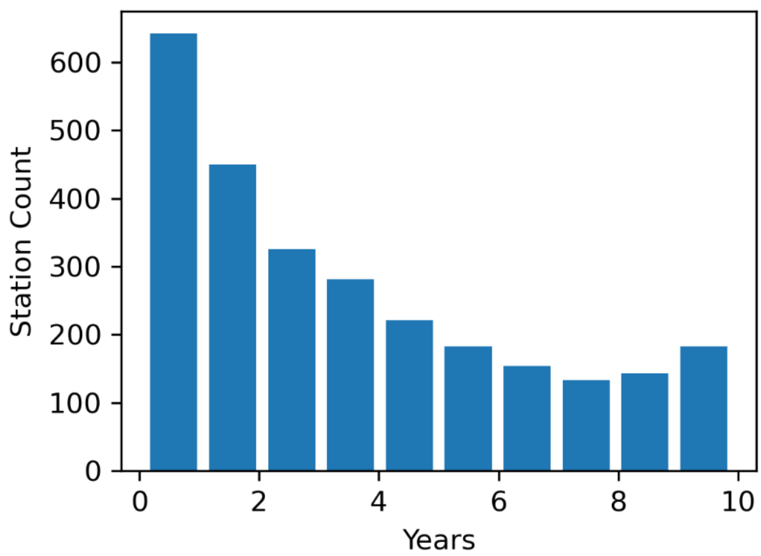


Figure 3.6   Histogram of Reporting Durations (years) for WOW Weather Stations

The duration of reporting is a function of when the station began uploading data, therefore as time passes it is expected that the number of longer records will increase. The relatively short station life span means WOW observations may not be ideal for long-term climate studies; however, this analysis does not include manual daily data provided by trained climate observers, which is also available via WOW and covers a longer period.

3.5.4 Distribution of Weather Stations

The distribution of WOW PAWS was not even across Britain, with a proliferation around London, the South, the centre of England and the central belt of Scotland. The distribution varied between years (see Figure 3.7). Although PAWS numbers reduced post 2016, their distribution was more widespread, particularly in Scotland.

Table 3.1 provides the count of PAWS reporting to WOW each year, with a breakdown of the number and percentage located in urban areas.

Year	Total Station Count	Urban Station Count	Urban (%)
2011	351	191	54
2012	842	439	52
2013	1126	561	50
2014	1288	644	50
2015	1354	689	51
2016	1375	714	52
2017	1271	658	52
2018	1230	633	51
2019	1262	627	50
2020	1306	623	48

Table 3.1 Number of PAWS reporting to WOW - total and urban (2011 - 2020).

The number of PAWS fluctuated from 351 in 2011 to 1,306 in 2020, with the percentage located in urban areas remaining reasonably consistent, ranging from 54% in 2011 to 48% in 2020. Their distribution was linked to population density in many areas, for example there was a high concentration of stations in London and the southeast of England, the central belt of Scotland and major cities including Newcastle, Sheffield, Birmingham, and Bristol (see Figure 3.7 and ArcStory Map via link in Appendix 1). There was a relatively high concentration of stations around Exeter, presumably due to the presence of the Met Office headquarters and people with a particular interest in the weather/WOW.

Figure 3.7 Distribution of British WOW weather stations 2011 – 2020 (rural = blue, urban = grey).

3.5.5 Intervals between Observations

The modal interval between observations was calculated for each PAWS. The number reporting at common intervals used in meteorological observations are shown in Table 3.2.

Modal Frequency of reporting (minutes)	Stations (Count)	Stations (%)
>2 - 2	21	0.8
2-5	190	7.0
5-10	1069	39.5
10-15	384	14.2
15-30	745	27.5
30-60	222	8.2
60	76	2.8

Table 3.2 Count of modal observation intervals (in minutes) from WOW weather stations.

The interval between reports is determined by the weather station operator. During this research, however, it was noted that data aggregation and transfer provided by a third-party application could take place at irregular intervals, varying from a few seconds difference to several hours. A comparable irregularity was noted by de Vos *et al.* (2019). The discrepancy between the time of the rain and the time of reporting is less significant when considering longer-duration events, but at the shortest time-scales desirable for urban hydrology the potential misrepresentation can be problematic. This a potentially a significant limitation for one of the most useful applications of citizen science and is discussed further in Chapter 7.

3.5.6 Calculating Rain Depth

Rainfall depth was calculated from accumulation. Although the Met Office recommends rainfall accumulation as the preferred field for rainfall measurements it was noted during the filtering process that 58 stations provided rainfall rate rather than accumulation (zero count in rainfall accumulation field). In fact, many PAWS had a higher record count for rate than accumulation, as shown in the example data from a WOW PAWS in Table 3.3.

Date and Time	Rainfall Rate (mm/h)	Rainfall Accumulation (mm)	Rain from Accumulation (mm)
04/10/2020 14:45	0.254	19.05	0
04/10/2020 15:45	0.254	19.304	0.254
04/10/2020 16:00	0.254	19.304	0
04/10/2020 16:30	0.254	19.304	0
04/10/2020 17:45	0	19.304	0

Table 3.3 Example Rain Record from WOW PAWS.

No intermittent rows were removed from Table 3.3, demonstrating the erratic timing of reports. The intervals between reports were 60 minutes, 15 minutes, 30 minutes, and 75 minutes. The overall modal interval of reporting (from the station summary table) was 15 minutes. Discrepancies in intervals of reporting undermine the quality of data from PAWS. Where reports are erratic the most practical use for the data is on aggregation to daily rainfall.

It was not clear why the rainfall rate was reported at more time steps than rain occurred (according to the rainfall accumulation field). This is something for further investigation with the Met Office, or to understand regarding how PAWS report. The anomaly was discounted where there were values available in the rainfall accumulation field, as this was assumed to be the correct value in accordance with WOW guidance.

3.5.7 Resampling to Hourly

There were 11 PAWS where no rain was reported that were removed from the dataset. There were 2,690 with data in the accumulation field, which was used as the primary source of rainfall data. There were 16 PAWS with no observations in the rainfall accumulation field; therefore, rainfall rate was used. These were reviewed individually and appeared to be reporting hourly, making the rainfall rate (mm/h) and the rainfall accumulation per hour the same. The hourly data were subject to SQC, presented in Chapter 4.

3.6 Spatial Analysis

3.6.1 Method

As an example, PAWS reporting to WOW during 2018 were used to assess gap filling[4] (see Figure 3.5). This analysis is a snapshot for illustrative purposes (given that PAWS report to WOW for a mean duration of 3.6 years). Official weather stations reporting hourly or sub-hourly observations in 2018 were obtained from the MIDAS (Met Office, 2006a) and from the EA, NRW and SEPA (from Villalobos-Herrera *et al.* (2022)). WOW PAWS were selected from the full record where:

- the sum of all rain observations was greater than 0mm
- there were at least 7 days of observations
- the modal interval was <= 61 minutes
- observations were made in 2018

To determine the ‘coverage’ of stations, both the WOW and Official weather stations were plotted in GIS, on a base map derived from the Mean High-Water Mark for Great Britain (ONS, 2021) with the urban areas as defined by Morton et al. (2020) for the Land Cover Map, 2015. The land area derived from these base maps is 230,147 km^2 for Britain, of which 15,593 km^2 is classified as Urban. The ‘coverage’ of Britain and Urban Britain was calculated at a series of extrapolation distances from station locations, using radii of 1, 2, 5, 10, 20, and 40 km. Any extrapolation extent beyond the British coastline has been excluded to avoid over representing the area of ‘coverage’. For Urban Britain, extrapolation extents were selected to coincide with urban land use and are within the coastline of Britain. From here, the area coverage (in km^2 and a percentage of total) was calculated for Britain and Urban Britain at each radial extent, for both the WOW stations and the Official stations independently. The ‘unique’ coverage at each radial extent was determined, and for Britain and Urban Britain, by deducting areas where there is overlap between the WOW and Official coverage maps. This allows the area (in km^2 and as a percentage of total) uniquely covered by WOW and Official stations to be calculated at each radial extent. Finally, the WOW and Official coverage at each radial extent is summed to determine the respective total cover for Britain and Urban Britain.

[4]2018 was the last complete year at the time of the research and was selected as the most current so as not to over-represent the potential by including any PAWS that ever reported to WOW.

3.6.2 Results

There were 1,230 PAWS of which 633 were in urban areas. There were 1,115 Official stations managed by the EA, SEPA, and NRW, and an additional 228 stations reporting via MIDAS, totalling 1,343 of which 178 were in urban areas. The rain gauge coverage of Britain is demonstrated in Figure 3.8, applying the 10 and 5 km radial extent as examples showing the distribution of Official and WOW rain gauges in Britain.

It is clear that there is overlap in many areas, however, in panel (c) (at far right) at both the 10 and 5 km extents, it is possible to see areas (in blue) where WOW rain gauges provide the only rain observations at the respective resolutions. The coverage for the varying extents for both WOW and Official rain gauges for Britain and urban areas are presented in Figure 3.9. For Britain, the area within 20 and 40 km of an Official rain gauge was 99.25%, meaning that there is likely to be little benefit in considering WOW gauges if this resolution is acceptable for use (except perhaps for cross-validation purposes, or in the event of an equipment malfunction). At the 10 km scale however, WOW gauges have more potential to aid in the delineation of rainfall, for example, 15% of Britain is within 5 km of WOW PAWS where there is no Official rain gauge to provide observations. The area of Britain covered by Official rain gauges at the 5 km extent is 36%; therefore, the addition of WOW stations represents an increase of coverage for Britain by the inclusion of WOW rain gauges of 41%.

In British urban areas, there a clear benefit in considering data from WOW PAWS at a 5 km scale, with 30% of urban areas being within 5 km of WOW gauges only. By combining WOW and Official gauges, the number of gauges in urban areas within 5 km of an observation point increases to 84% (from 54% for Official alone). At the 1 and 2 km extents, there are increases of 265 and 176%, respectively, when WOW locations are combined with Official. For high-resolution delineation of rainfall, the benefit of including WOW PAWS is therefore clear. As seen for Britain as a whole, the Official monitoring network has adequate coverage at the 20 and 40 km resolution for Urban Britain.

The higher density of WOW gauges in urban areas demonstrates the potential for providing high spatial resolution rainfall data in these locations. The radial area of 5 km has been used as an example for visualisation in Figure 3.9. There is merit in having data from rain gauges located at higher densities given the highly variable nature of rainfall, particularly during convective events, as overlap between Official and WOW extrapolations can be beneficial to improve rainfall spatial resolution. The temporal resolution of data is hourly or sub-hourly, meaning that WOW PAWS can increase the resolution of rainfall data in urban areas at a temporal interval that will support detailed hydrological modelling and post-event analysis.

<table>
<tr><th>Gauges</th><th>Coverage (km^2)</th><th>Coverage (%)</th><th>Unique cover (km^2)</th><th>Unique cover (%)</th><th>Combined cover: all official + unique WOW (km^2)</th><th>Combined cover (%)</th></tr>
<tr><td>WOW</td><td>58,755</td><td>25.4</td><td>32,637</td><td>14.2</td><td rowspan="2">116,030</td><td rowspan="2">50.4</td></tr>
<tr><td>Official</td><td>83,393</td><td>36.2</td><td>57,275</td><td>24.9</td></tr>
</table>

Table 3.4 British Official weather station and WOW PAWS coverage (5km scale).

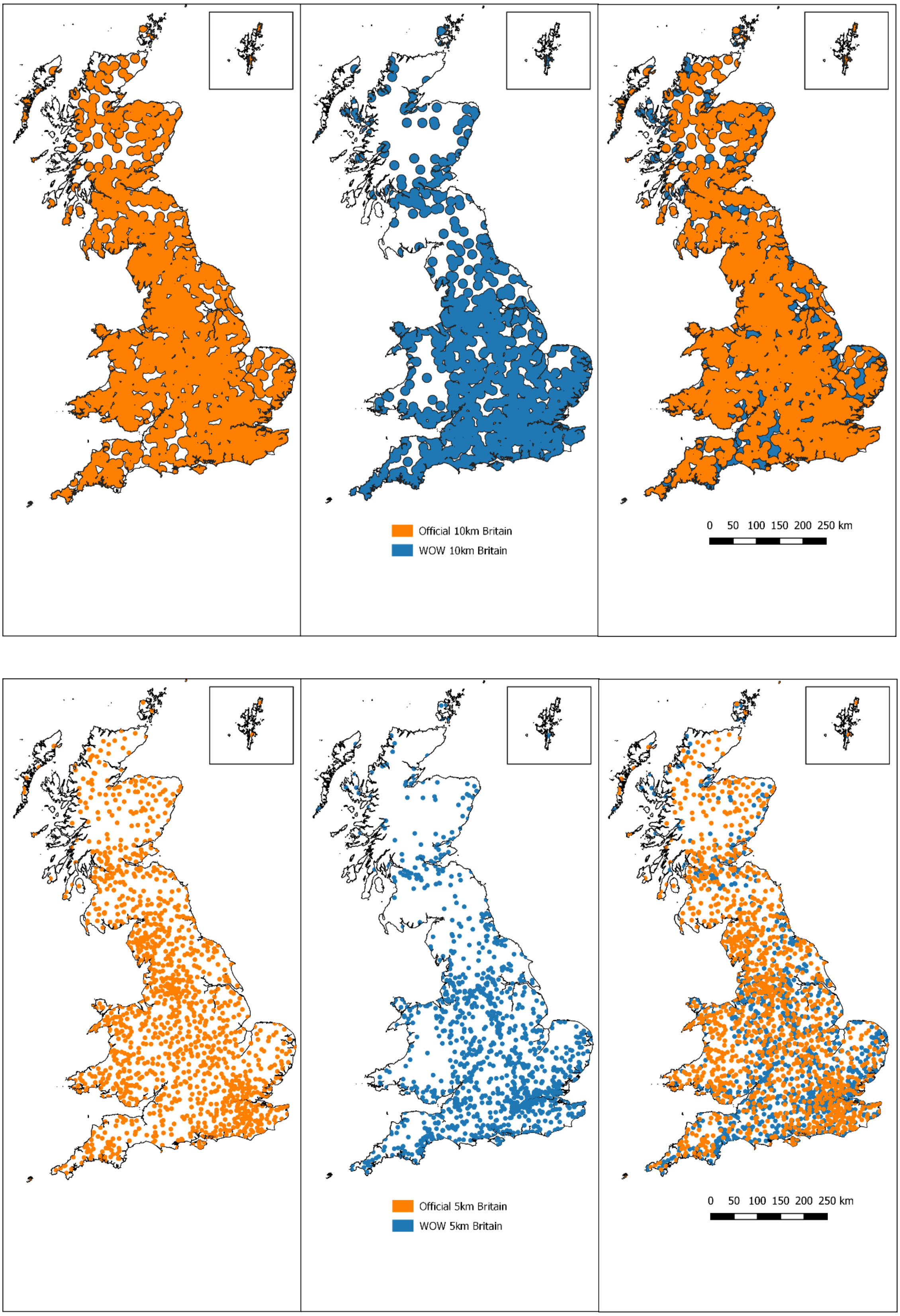


Figure 3.8 Maps of British Official and WOW rain gauge coverage, 2018 (10km and 5km scale).

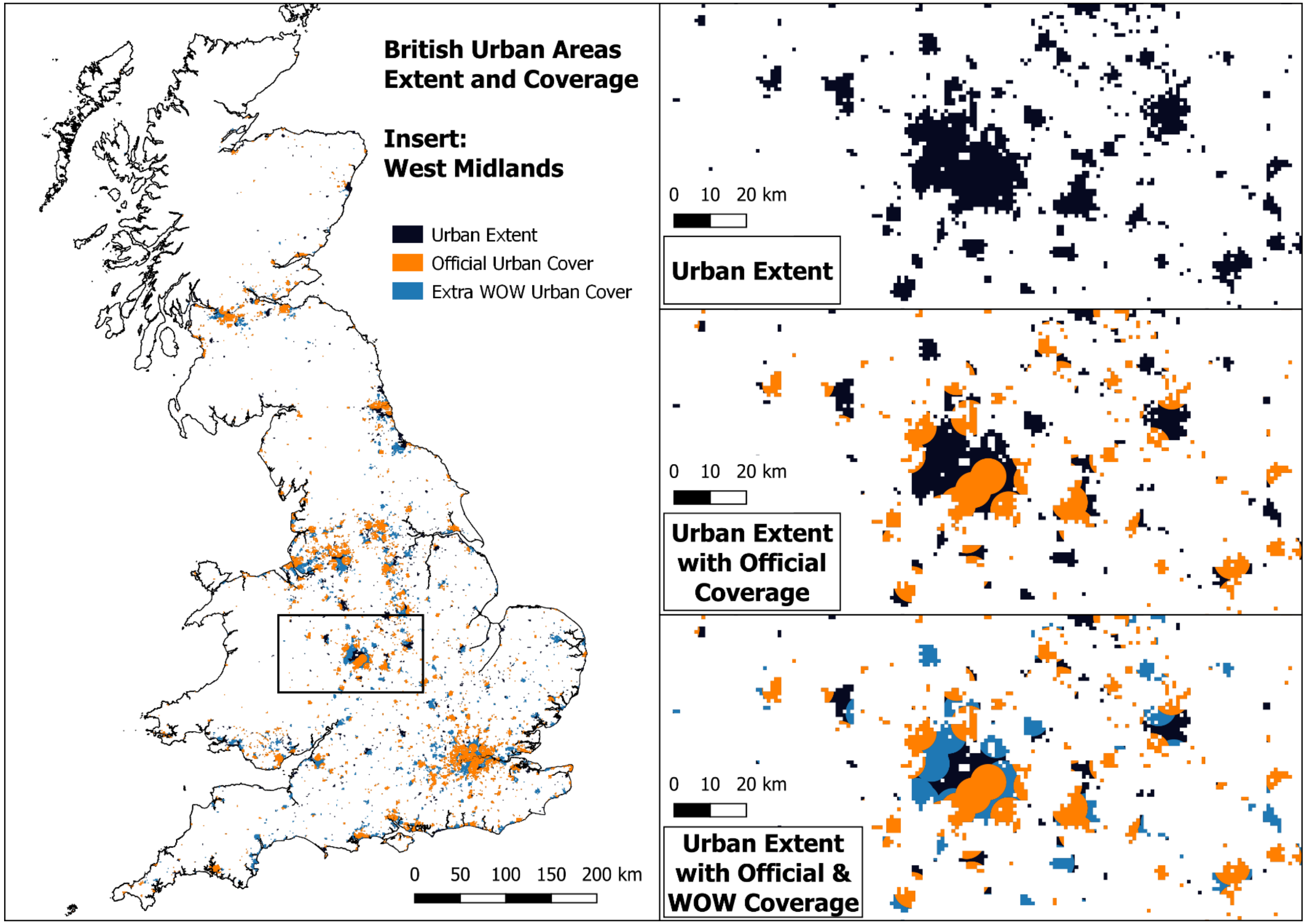


Figure 3.9 Maps of British Official and WOW rain gauge coverage of urban areas (5km scale) with West Midlands Insert.

Figure 3.9 and Table 3.5 demonstrate that 64% of British urban areas were within 5 km of WOW PAWS with rainfall data. 28% of urban areas were uniquely covered by WOW PAWS.

Gauges	Coverage (km^2)	Coverage (%)	Unique cover (km^2)	Unique cover (%)	Combined cover: all official + unique WOW (km^2)	Combined cover (%)
Urban WOW	9,925	64%	4,416	28%	12,643	81%
Urban Official	8,227	53%	2,720	17%		

Table 3.5 Urban British Official and WOW station coverage (5 km scale).

3.7 Discussion

The data processing undertaken as described in this chapter resulted in accessible and easy to use datasets of citizen science data from WOW spanning 10 years from 2011 to 2020 and provided metrics by which to make some initial judgements on the usefulness of the data. The resulting datasets and summary statistics are accessible online via links in Appendix 1.

The process for selecting data was lengthy and was made more difficult because of the ways the data are generated, shared, and stored. Some issues are beyond the control of WOW administrators, e.g., weather stations will be located where people wish to place them; however, there were some aspects that presented problems that could be addressed with relative ease:

- 'Ghost' test stations not closed on completion of testing,
- Non-standard and variable intervals of reporting,
- Inconsistent reporting of seemingly similar observations (rain rate and rain accumulation),
- Limited access to data (no bulk data access option)

It would be helpful if the WOW administrators periodically removed stations set up for testing purposes e.g., those named "Test". This would reduce the number of stations with very limited observations. Many weather stations have the capability to report at whatever interval the operator sets. The shortest available interval is generally the most desirable. If data capacity permits, operators should be encouraged to provide data at a 2-minute interval. Rainfall data usability would be improved if the "tip times" were available rather than rate or accumulation.

Tip times are recorded by Official rain gauge networks and are advantageous as data can be reliably aggregated to any preferred time interval, for example 1-minute for fine grain analysis of convective events, to daily for more regional climatic analysis. It is not known if citizen science rain gauges have this capacity, but requesting a particular time interval e.g., 2 minutes would be a quick interim fix.

The difficulties in data access are a clear demotivator for data use, additionally processing WOW data has highlighted issues that can undermine confidence, in particular the difficulty in accessing bulk station metadata. Researchers have long pointed out the necessity for good metadata (Muller *et al.*, 2013), going beyond what can be shared currently via WOW e.g., incorporating any QA/QC procedures applied to the data. Based on personal communication with the Met Office there is a reluctance to enforce the provision of metadata as users may be dissuaded; however, this ultimately denigrating confidence in data quality, rendering the database less attractive to potential users. If station metadata and a site rating were required parameters, additional analysis would be possible, e.g., assessing the reliability of different types of weather station, which has only been possible in relatively small scale trials (Bell *et al.*, 2015), or when working with data from a proprietary platform, e.g. Netatmo or Davis (de Vos *et al.*, 2017; Davis Instruments, 2018; de Vos *et al.*, 2019; Netatmo Weather Map, 2022). As with the weather observations, access to station metadata and site ratings would be more practical if they could be accessed in bulk.

The user experience is an important consideration that can be overlooked (Walker *et al.*, 2021). Support for citizen scientists that could make data sharing easier and reliable include.

- More simple data transfer mechanisms,
- Encouragement to continue sharing data,
- Mechanism for feedback where issues are identified and support to improve observations.

Simplicity for participation, interaction, and a sense of community have been documented as motivators for citizen science (Geoghegan *et al.*, 2016; Lin *et al.*, 2016; Starkey *et al.*, 2017). Based on some of the issues documented in this chapter, and via the intimate knowledge of WOW acquired during research, these elements are currently limited and could be improved.

Chapter 6 of this thesis looks in more detail at some of the motivations, barriers and benefits to participation operating a citizen science weather station.

3.7.1 Response to Research Questions

The research presented in this chapter demonstrates that the spatial distribution of rainfall monitoring in Britain is improved with the inclusion of automated citizen science rain gauges reporting to WOW, with the most benefit observed where rainfall is highly localised and interpolation beyond 5 km from the observation point would be considered unreliable. Responses to the research questions posed in section 3.2 are presented here. Conclusions, recommendations, limitations and suggested further work are presented in Chapter 7.

Q1. What observations are available from automated citizen science rain gauges reporting to WOW and how can they be wrangled into a useable format?

Between 2011 and 2020 there were 3,920 PAWS reporting to WOW. There were 2,117 PAWS with rain gauges located in Britain reporting hourly or shorter interval observations to WOW that were selected by filtering of those reporting for less than a month. After filtering observations were aggregated to hourly, as this was deemed to be a useful interval for using the data at later stages of research. It was also a pragmatic decision, as the interval of reporting could be erratic within a rain gauge record and between rain gauges. Some aspects of hydrological investigation are considered to be more robust when using shorter temporal interval data (Ochoa-Rodriguez *et al.*, 2015), hence the data were preserved in their raw form to allow observations to be processed for other uses at the most appropriate interval. This also supports the aim of open data for further research (Go Fair, 2016; Wilkinson *et al.*, 2016).

A total of 85,321,792 hours of observations from 2,717 WOW rain gauges were identified for statistical quality control in this research. This is significantly smaller than the number of observations in the Official data used in chapter 4, which totaled 241,161,095 hours of observations from 1,477 rain gauges, with records dating from 1962. The number of PAWS fluctuated each year, and at the last count (2020) there were 1,306. The mean duration of reporting for PAWS was 3.6 years, with 643 (24%) reporting for a year or less, and 183 (7%) reporting for 9 or more years.

The data are described in detail in section 3.5, which explains the filtering of raw data to create a dataset that would be fit for further analysis. There were some issues with very short record weather stations and the use of differing fields for reporting (rate or accumulation) that were addressed to provide a fair representation of available data.

The mean duration of reporting by PAWS of 3.6 years is far shorter than Official rain gauges meaning the data are unlikely for applications including climate studies or longer-term hydraulic assessments. As a comparison, the mean duration for reporting for the Official rain gauge data used in chapter 4 of this thesis was 3.6 years. As WOW only started in 2011 there may be some PAWS that do provide consistent longer-term records, which will become apparent over time.

Q2. Where are PAWS rain gauges located within Britain?

The PAWS reporting to WOW were in almost all corners of Britain, from the southwest (with a proliferation around the Met Office in Exeter), to the Shetland Islands in the northeast (see Figure 3.7). Initially, in the early years of WOW, there were very few gauges in the more mountainous parts of Scotland and Wales but by 2020 the numbers had increased. The split between urban and rural locations varied in absolute terms over the years but remained around 50%. This could be considered as an advantage as this research has shown that there is a high degree of overlap between WOW and Official rain gauges in urban areas (36%) which could be valuable during data validation, or when looking at applications requiring a high spatial density of rain gauges.

Given that most rain gauges reported to WOW for less than 4 years the PAWS network distribution is highly dynamic. PAWS are located wherever people wish to place them, unlike an Official monitoring network that is carefully planned with rain gauges situated to meet operational needs.

There are many more citizen science rain gauges operational in Britain, the data from which have not been included in this research. Whilst platforms offer free access to real-time data or mapping of observations, the historical record is not made available (Davis Instruments, 2018; Weather Underground, 2019; Netatmo Weather Map, 2022). It is assumed based on the experience of undertaking the research (speaking to weather station suppliers etc.) and published research on PAWS data obtained directly from famers (Holley *et al.*, 2020) that there are many PAWS operators who do not share their data. It is therefore difficult to determine how many more weather stations there are in Britain, but certainly the WOW archive only represents a portion of the data being generated. If the data from all PAWS could be collated the extent of the network could be significantly more extensive.

Question 3. To what extent would the inclusion of WOW stations improve the spatial distribution of rain monitoring in Britain?

PAWS reporting in 2018 were used as an example to calculate the distribution and 'coverage' each rain gauge provided, using 5km as the maximum extent to which rainfall data could reliably be interpolated. The 'coverage' of Britain using this method of calculation was 36% by Official gauges, and 25% by WOW gauges. Combining the Official and WOW rain gauges gave a national coverage of 50%; therefore, due to overlapping, WOW gauges provide an additional 14% of unique land coverage beyond that provided by Official gauges.

In urban areas the Official coverage was 53% and the WOW coverage was 64%, with 28% being unique to WOW. The combination of both datasets provided coverage of 81% in urban areas. The extent of overlap in urban areas between WOW and Official gauges is positive as the 5 km spacing may be insufficient during the highest intensity convective storms which often cause flash flooding in urban areas. Research has identified the value of including other types of crowd sourced rainfall data, derived from novel sensors and cameras etc. (Yang *et al.*, 2019), therefore the consideration of data from the more traditional rain gauge should be considered advantageous also.

3.7.2 Recommendations for Using WOW Rainfall Data

This section is provided to highlight some of the potential pitfalls when using WOW rainfall data to alert other researchers.

Processing WOW data is complex, because of several of the issues highlighted in this chapter. It is recommended that where possible, PAWS records are viewed or plotted for a 'sanity check' before proceeding to any detailed analysis or interpretation. The graphical representation of data was found to be a useful tool in identifying inconsistent or excessive rainfall depths. It was noted that issues can be intermittent, meaning that although there may be an error at one point in time, there may still be useful data from a given station.

The issue of standard and variable units was highlighted during data cleaning but not further investigated. In the case of rainfall data caution should be exercised when working with data from countries where imperial units are the norm, e.g., the USA. Davis Instruments manufacture one of the most popular PAWS with a 'bucket' size of 0.01 inches per tip for the imperial version of the weather station (0.254 mm) and 0.2 mm per tip for the metric version.

Due to variations in reporting intervals, rainfall rate can be misleading, as although the rate is given in mm h^{-1} it may be for a 2-minute interval for one record but then a 30-minute interval for the next. This makes rainfall accumulation an easier parameter to work with, and, as the rainfall rate can be back calculated from the rainfall accumulation, it is the recommended parameter for analysis.

Caution is advised when accessing data to ensure the number of reports and the interval of reporting are sufficient for the intended application. PAWS with more reports are likely to be more useful for a wider range of applications; however, with luck a station reporting for a short duration may have captured something of interest.

Some of the most popular PAWS are unheated. When temperatures are subzero it is therefore likely that precipitation may not be accurately recorded. They may also struggle to accurately represent hail, which can be associated with intense convective events.

3.7.3 Outputs

The code used in data processing and various iterations of the data have been made available (see Appendix 1). Publication ensures observations can be used in the way best suited to the required analysis, i.e., multi-parameter analysis, or consideration of the shortest available time interval, as opposed to hourly aggregations. The raw (as reported) observations for a series of common parameters (including air temperature, rainfall rate/accumulation and wind speed/direction) are included in the published data. The rainfall data were transformed from accumulation to depth and are available at the reported time interval and resampled to hourly. There is a useful look-up table that can be accessed online to help determine data availability, providing insight into the WOW archive based on this research (https://arcg.is/1zmui5). Links to all the digital outputs from this thesis can be found in Appendix 1:

I. Multiple parameter observations at original time intervals
II. Summary statistics for selected stations (a lookup table)
III. Calculated rainfall observations at original time intervals
IV. Rainfall accumulations at hourly intervals
V. Code used for processing the WOW bulk download.

Chapter 4. Describing Crowdsourced Rainfall Observations and Assessing Quality using a Statistical Quality Control Algorithm

4.1 Overview

Rain gauge observations from Private Automated Weather Stations (PAWS) reporting observations to the Met Office Weather Observation Website (WOW) (Met Office, 2011) from Britain from 2011 – 2020 formed the basis of this research. This chapter uses the WOW hourly dataset created in Chapter 3, describing the crowdsourced rainfall data available via WOW and assessing the quality as compared to rainfall data from Official ground-based gauge networks in Britain. It builds on the discussion of the spatial and temporal distribution of the WOW data by aiming to determine if data are of an acceptable quality for use in meteorological and hydrological applications. A method of statistical quality control (SQC) developed for global sub-daily rainfall data and contextualised for Britain (Lewis *et al.*, 2021; Villalobos-Herrera *et al.*, 2022) is applied to both the WOW and Official datasets for quality assessment.

The post-SQC WOW data have been published along with a summary table of SQC performance for WOW rain gauges (see Appendix 1).

4.2 Introduction

Some of the oldest British weather records available today were generated by people who could be described as citizen scientists in today's terms e.g., ships captains, who kept detailed weather logs whilst at sea (Hawkins *et al.*, 2022). Some of the heaviest downpours recorded in Britain pre-2000, and subsequently used in design, were made by non-professionals (Faulkner *et al.*, 1999). Despite this long and noble heritage there can be a reluctance to use contemporary crowdsourced weather data (Endfield *et al.*, 2012; Buytaert *et al.*, 2014). There is a persistent assumption that observations will be less accurate than from Official sources (Muller *et al.*, 2015; See, 2019). The previous chapter demonstrated that PAWS reporting to WOW had the potential to increase the ground-based rain gauge coverage. This chapter builds on the previous by considering WOW data quality.

4.2.1 Research Questions

The aim of the research presented in this chapter is to assess the quality of WOW rain gauge data, compared to the quality of rainfall data from Official ground-based gauge networks in Britain. The questions this research presented in this chapter seeks to answer are:

Q1. How good is WOW rain gauge data?

Q1a. How does the availability of WOW rain gauge data compare to that from Official rain gauges?

Q1b. What were the results of SQC on the WOW data and how does that compare with Official data?

Q1c. Where are the post-SQC good quality WOW rain gauges located?

Q2. Is the method of SQC appropriate for WOW data?

Q3. How could the quality of WOW observations be improved?

4.2.2 Data Sources

WOW data from 2,706 rain gauges with records spanning 2011 – 2020 were used for the analysis in this chapter. A full description of how the WOW dataset was created is presented in Chapter 3. To answer Q1b. and provide context for Q2. data were obtained from Official rain gauges in Britain operated by the EA, SEPA and NRW. There were two Official datasets made available for this research. The first included data from Official rain gauges reporting hourly or shorter interval observations (1,477) gauges. The second dataset consisted only of observations from Official gauges selected following manual validation, i.e., following the removal of gauges obviously generating poor quality data. There were 1,240 rain gauge records in the second dataset (Villalobos-Herrera *et al.*, 2022). All observations were aggregated to hourly prior to SQC. The longevity of Official observation records varied by rain gauge with the oldest observation being from 1962. Data up to July 2018 were available for use in this research; however, some records were shorter due to gauges ceasing to operate.

4.3 Methods

4.3.1 Determining WOW Rain Gauge Data Quality

The methods previously applied to crowdsourced rain gauge observations to assess and handle data quality are discussed in the literature review (Bell *et al.*, 2015; Davids *et al.*, 2019; de Vos *et al.*, 2019; Bárdossy *et al.*, 2021). Approaches included comparative studies between gauges and filtering with bias correction to create usable datasets (Bárdossy *et al.*, 2021). The application of bias correction in the two most recent SQC algorithms developed for crowdsourced rainfall data (de Vos *et al.*, 2019; Bárdossy; Seidel and Hachem, 2021) is less suited to data from WOW, due to additional variation introduced given data are obtained from differing brands of rain gauge. Given this limitation, rather than seeking to ‘correct’ WOW data a method of SQC was selected that would flag or eliminate suspect data from the time-series. The intention was to determine whether WOW rain gauge observations could be processed in the same way as data from Official monitoring networks, allowing intercomparisons of data quality between the two networks. After reviewing the available methods for SQC an algorithm created for the assessment of a global dataset was selected and applied to the WOW data (Lewis *et al.*, 2021). The choice of SQC method was influenced by the availability of recently completed research that developed and applied the same method to an Official rain gauge data set, allowing for direct comparison (Villalobos-Herrera *et al.*, 2022).

The WMO recommends comprehensive SQC should assess format, completeness, consistency, tolerance/range and spike/streak (WMO, 2021). All these elements (except for internal consistency, as only rainfall data were considered) are included in the SCQ method applied to the WOW and Official datasets respectively. A detailed breakdown of the checks undertaken in the SQC is provided in Appendix 2. The SQC process flags and/or removes observations (summarised in Table 4.1 and detailed in Appendix 2). The SCQ was localised to Britain by replacing the global hourly and daily maxima from Lewis *et al.* (2021) with applicable British maximum values (92mm and 341mm respectively) (Villalobos-Herrera *et al.*, 2022).

Rule	SQC Flags	SQC Measure Description
R1	K-largest (QC2) - k10 = 0	10 largest events in a year all equal zero – excludes full year
R2	Daily accumulations (QC13) = 1 (3) or 2 (4)	2 or more consecutive days of suspected daily accumulations
R3	Monthly accumulations (QC14) = 1 (2)	Suspected monthly accumulations
R4	Streaks (QC15) > 0	Streaks of repeated values (i.e., ≥2 identical consecutive high values or longer periods of repeated lower non-zero values)
R5	Hourly rainfall (QC10) > 92	Value exceeds UK record 1-hour rainfall (hour removed)
R6	24-hr rainfall (QC11) > 341	Daily value exceeds UK record daily rainfall (day removed)
R9	Hourly neighbours (dry) (QC19) = 3 & CDD (QC12) > 0	Excessive dry spells according to hourly neighbours and/or ETCCDI CDD[5]. In the absence of neighbours, the CDD flag used.
R10	Daily neighbours (dry) (QC18) = 3 & CDD (QC12) > 0	Excessive dry spells according to daily neighbours and/or ETCCDI CDD. In the absence of neighbours, the CDD flag used.

Table 4.1 Rules applied in SQC to remove suspicious observations (colour coded; blue - excessive rain, yellow - excessive dry, grey – either too wet or dry).

Rules performing nearest neighbour checks within the SCQ (R9 and R10) were suppressed when applied to WOW data as they were designed for cross-validation on the assumption that neighbouring gauges were fundamentally reliable and were generating accurate observations. For the purposes of this research, it was assumed that WOW data were unreliable and should not be used for inter-station validation. For rules R9 and R10 the SCQ applied to WOW data used the Consecutive Dry Day (CDD) index from the Climdex Global Historical Climatology Network (GHCN) Daily dataset defined by the Expert Team on Climate Change Detection and Indices (ETCCDI) (Donat et al., 2013).

As the quality controlled Official rain gauge data were available, that dataset was used instead for an alternative method of nearest neighbour analysis applied as a separate process outside

[5] CDD data are available for download from https://www.climdex.org/learn/datasets/

of the SQC algorithm. For the nearest neighbour comparisons, data were aggregated to daily intervals as in Lewis *et al.* (2021), as it was expected that hourly observations between gauges would be naturally variable and therefore correlation would be weak. A further variation in the nearest neighbour checking related to the requirement for 3 years of overlapping observations. Due to the relatively short lifespan of WOW gauges this could not be applied to 1,408 (51%) of gauges, limiting the usefulness of the check. As a compromise the following process was applied for the selection of nearest neighbours:

Initially, QGIS was used to identify the 10 nearest neighbours to a given WOW gauge, then any at a distance > 25 km, and with < 1 year of overlapping data were excluded. Once the nearest neighbour (within 25 km and with at least 1 year of overlapping data) was identified the following metrics were calculated:

- Number of overlapping days
- Pearson and Spearman's Rank correlation
- Number of matching days (matching presence/absence of rain)

Once quality-controlled data were available for the WOW and Official datasets, the availability of observations pre- and post-SQC was determined and compared between the datasets. The description of data availability involved calculating the duration of records and quantifying the extent of missing observations. The available observations were plotted geographically to determine the distribution of rain gauges and to assess whether there was any spatial correlation across Britain relating to SQC performance.

4.3.2 Appropriateness and Effectiveness of SQC for WOW Rain Gauge Data

To answer Q2: to determine whether the SQC was effective in identifying gauges with good or poor-quality data, a manual review of pre-QC and post-SQC observations was completed by plotting WOW gauge data alongside corresponding observations from the nearest neighbouring Official gauge. This was done for a selection of gauges and in conjunction with the analysis of gauge data within the search area of specific convective events (presented in Chapter 5). Plots of the full timeseries of data from >142 WOW gauges were reviewed, along with focus on specific event periods.

4.3.3 Exploring the potential for improvement in WOW data quality

To answer Q3 and consider whether WOW data quality could be improved, a case study was undertaken on rainfall data from members of the Climatological Observers Link (COL). The intention was to assess whether data from weather enthusiasts was of better quality than data available via WOW, and if so, to consider why that might be.

The steps for data collection and processing were:

- Initial request for data via the COL website.
- Request sent to members via the Secretary.
- Members willing to share data responded directly with a general offer or specific details.
- Members sent data via email or drop-off, many adding information about the dataset.
- Initial data processing of observations from original time interval to hourly and queries regarding any unusual data.

Observations were provided from COL members, with 10 supplying PAWS observations directly. A further 4 COL members provided their WOW station IDs for assessment.

Data from COL members were subject to the same SQC and nearest neighbour checks that were applied to WOW observations and were analysed using the same methods for comparison. The benefits and limitations of the COL approach to weather observations is explored in more detail in Chapter 6.

4.4 Results

A summary table detailing the performance of the WOW gauges was generated, as well as the SQC'd WOW gauge data (see Appendix 1).

4.4.1 Data Availability Pre-SQC

Data from 2,706 WOW gauges[6] and 1,477 Official gauges were included in this analysis. The Official rain gauge distribution and duration of reporting can be seen in Figure 4.1. There are regional patterns reflecting the different times hourly/sub-hourly gauging was adopted in Britain; for example, many gauges in the southwest had records of 20 years or less.

The rain gauges maintained by the respective agencies are located primarily to support fluvial flood forecasting, as that is where their responsibility lies. The locations of WOW gauges are shown in Figure 4.2. Locations are determined by the participation of citizen scientists who self-select to share data to WOW.

[6] The number of gauges in the WOW dataset was 2,717 however 11 stations were found not to have any rainfall in the record >0 therefore they were not included in the analysis.

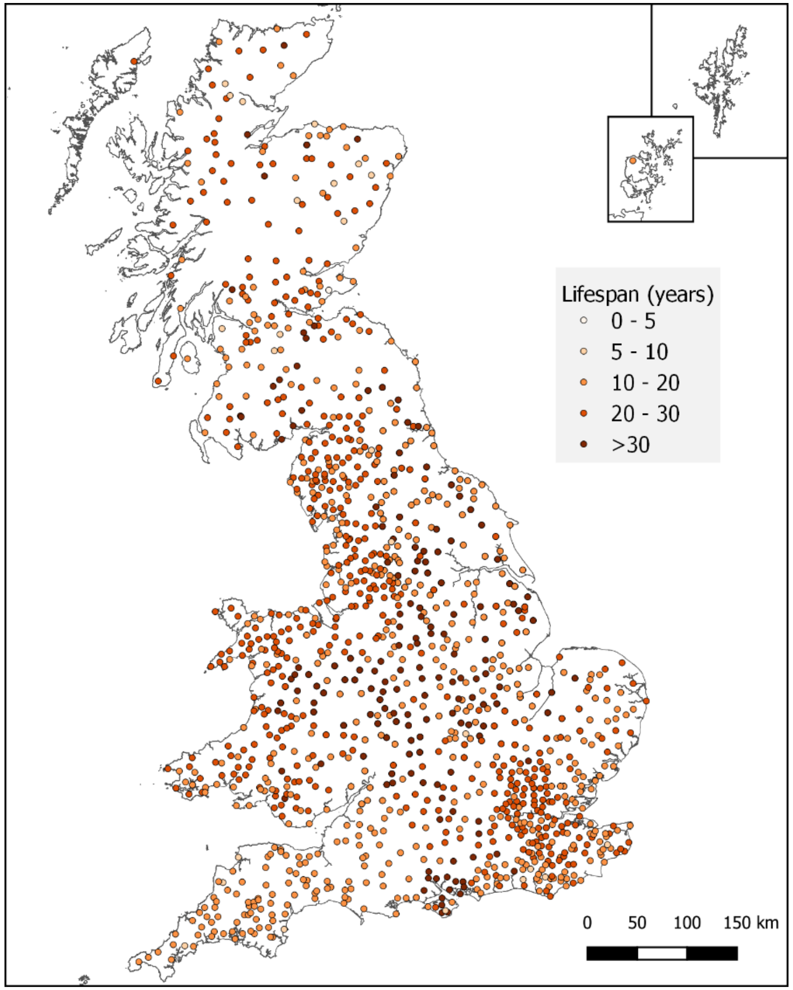


Figure 4.1 Map of Official rain gauge locations with duration of reporting (EA, SEPA, NRW) to 2018.

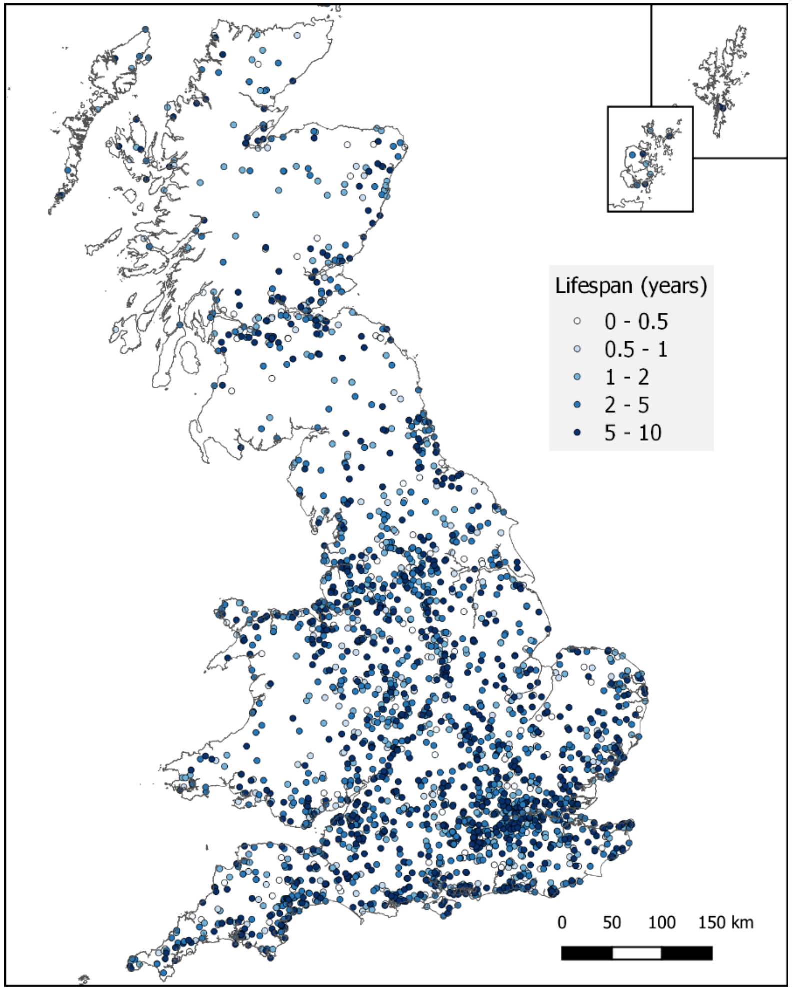


Figure 4.2 Map of WOW rain gauge locations with duration of reporting (2011 - 2020).

The number of observations available in each dataset (see Figure 4.3 and Table 4.2) varied due to the number of gauges reporting, the life span of those gauges (longer duration Official gauges) and the percentage of missing data.

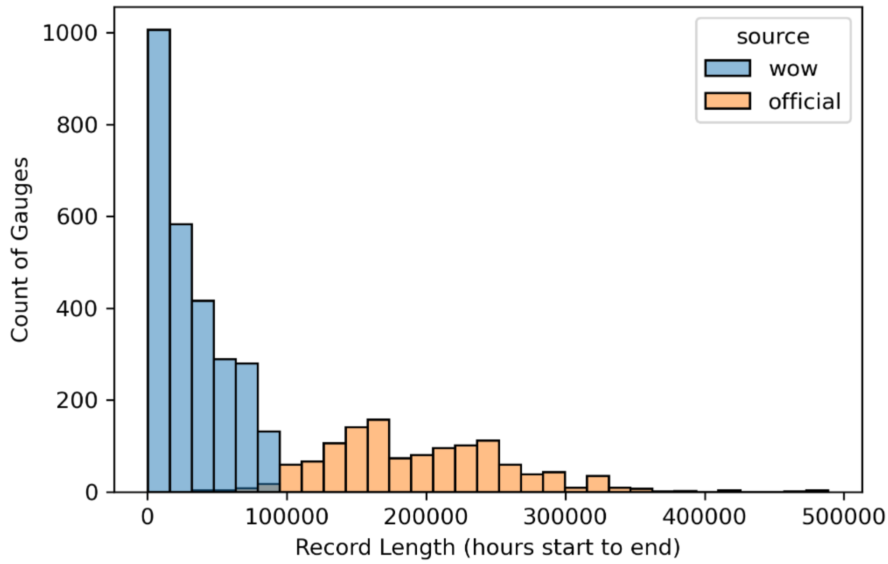


Figure 4.3 Histogram of record lengths per rain gauge (WOW and Official datasets).

The percentage of observations between the start and end point of reporting for all rain gauges within their respective dataset was compared between the WOW dataset and Official (see Table 4.2).

Hours	WOW (Hours)	WOW (%)	Official (Hours)	Official (%)
Total	85,321,792	100	241,161,095	100
Missing pre-SQC	17,732,746	20.78	11,901,320	4.94

Table 4.2 Observations missing prior to SQC for WOW and Official rain gauges.

The Official dataset was missing 4.94% of observations prior to SQC, whereas the WOW dataset was missing approximately one fifth (20.78%) of all observations. Hours missing were defined as any hour between the start and end point of reporting where no observation was reported, not including zero rainfall.

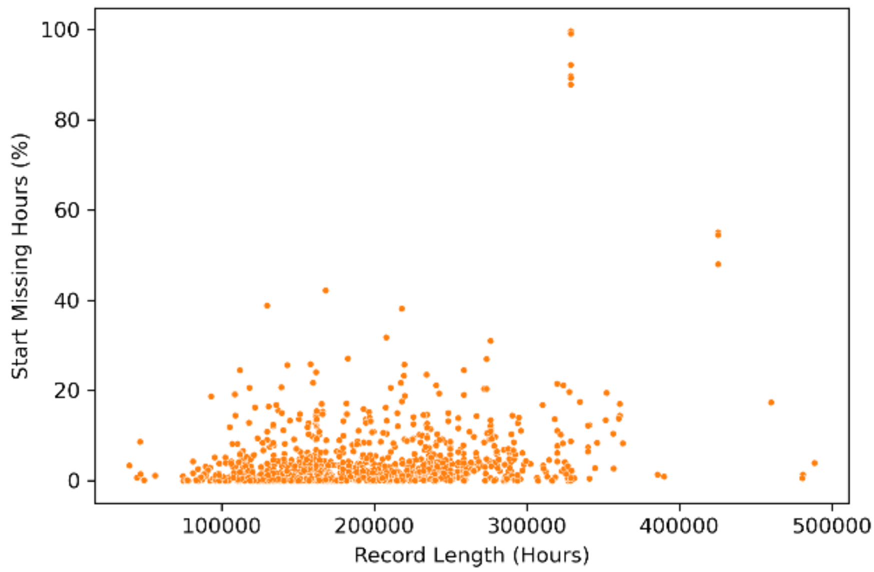


Figure 4.4 Scatter plot of percent per gauge of data missing from Official gauges pre-SQC versus length of record.

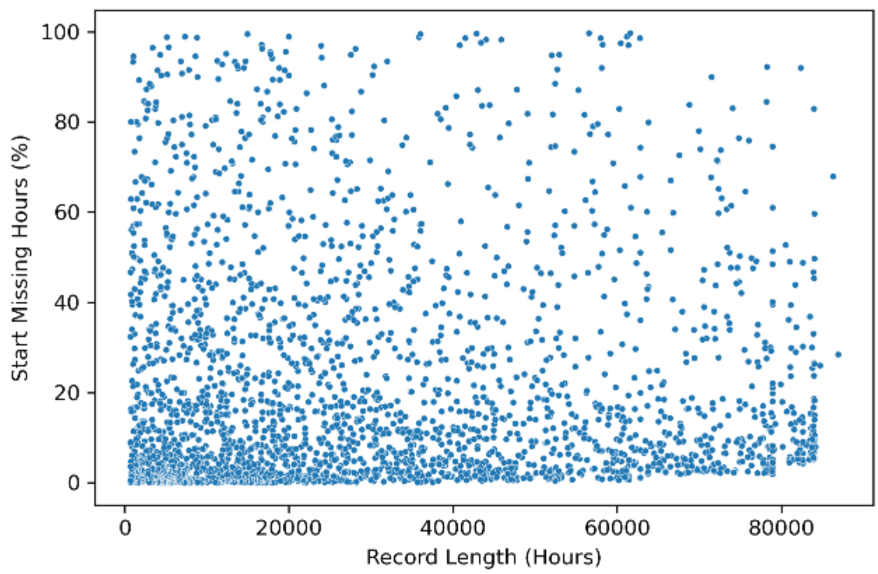


Figure 4.5 Scatter plot of percent per gauge of data missing from WOW gauges pre-SQC versus length of record.

Figure 4.4 and Figure 4.5 show the variation in missing data per gauge between the two datasets. The high percentages of missing data in WOW gauge reporting could be a function of short records with large gaps, however this was not the case, as seen in Figure 4.4. Conversely the Official data set is more complete with some outliers (Figure 4.5).

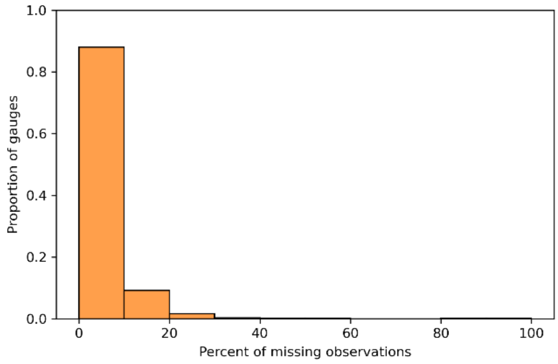


Figure 4.6 Histogram of data missing from Official gauges pre-QC.

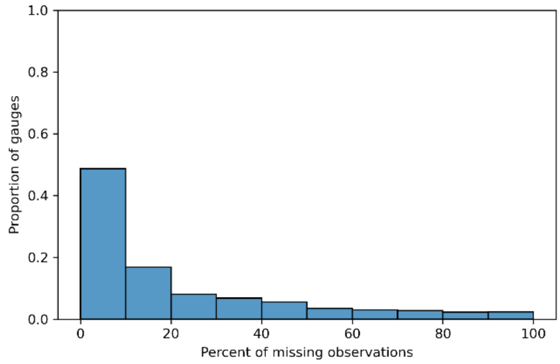


Figure 4.7 Histogram of data missing from WOW gauges pre-QC.

The extent of missing data is further demonstrated in Figure 4.6 and Figure 4.7 showing the proportion of gauges missing data prior to SQC being undertaken. There were 375 WOW gauges and only 8 Official gauges where the proportion of missing hours was greater than 50% (13.9% and 0.65% respectively). Of the Official gauges with >50% missing data, 6 were operated by SEPA with a record length >37 years and >87% missing data, a further 2 gauges were operated by the EA for 48.5 years and were missing 55% of data.

When looking at WOW rainfall observations there is no way of determining why there would be gaps in a gauge record. From a potential WOW data user perspective, it is important to be aware that although a WOW gauge may apparently be active at a particular point in time, there may not be any rainfall data. Missing data has been acknowledged as an issue with rain gauge data (Durre *et al.*, 2010) and can be problematic for analysis as important events may not be captured. It is often more advantageous to have a complete representation of rainfall, which is more readily available from the Official dataset (see Figure 4.4); however, there are applications where short-duration records in the right location are acceptable/desirable e.g., for the post-event analysis convective rainfall events causing flash floods.

Reasons for missing WOW rain observations may include power cuts (potentially more likely during high rainfall/flood events), data transfer issues, or gauges being non-operational (de Vos *et al.*, 2017). The Official dataset may be more complete as those gauges are likely to have a more robust method of locally storing observations in the event of power or transmission failures. PAWS are at risk of being moved or disconnected which would be unlikely for Official gauges. When manually reviewing PAWS records it was apparent that some were very long duration with relatively few observations. It appears that some WOW users stop sharing data from their weather stations without closing their WOW account, which artificially extends the longevity of the record. It was also noted during a manual record review that some stations report for several months (or years), followed by a break, and then resume reporting. There are also types of weather stations that operate opportunistically, in that they only report when it rains. It is possible that at least some of the WOW gauges operate this way, which would account for large gaps between observations. Similarly, Starkey *et al.* (2017) noted that participation is boosted at times of extreme events and tends to drop off during other periods. There could be a similar phenomenon with WOW gauges, or it is possible that people are less likely to notice issues with data reporting during dry periods.

During this research, a banner on the WOW homepage informed users of approximately 15 hours of data loss due to technical issues (June 2021). There has been no analysis of where/when the gaps exist to further explore this issue; however, research assessing Netatmo rain gauge data identified the issues highlighted in the previous paragraph (de Vos *et al.*, 2017). Whilst data analysis could provide some clues, it would seem most effective to approach people directly to ask why there are gaps in their specific record as looking at the data alone would always be a best guess. The large gaps in the WOW database are an example of where oversight of reporting and immediate/regular periodic SQC could help improve data availability, as interaction with a PAWS operator would provide an opportunity to fix any issues. The automated passive nature of WOW is potentially a benefit in encouraging participation, and a drawback, as it does not prompt users to pay attention to what their weather station is reporting (discussed further in Chapter 6).

4.4.2 Impact of SQC

Availability of Data

Applying the SQC resulted in 2.52% of available observations being removed from the WOW dataset and 2.79% of observations from the official dataset. The SQC therefore removed a slightly higher percentage (0.27%) of observations from the WOW dataset than the Official.

Hours	WOW (Hours)	WOW (%)	Official (Hours)	Official (%)
Total	85,321,792	100	241,161,095	100
Removed by QC	2,157,502	2.52	6,733,450	2.79
Missing after QC	19,890,248	23.31	18,634,770	7.73

Table 4.3 Observations removed and missing post SQC for WOW and Official gauges.

There was <=1 observation[7] removed from 1,242 WOW gauges (46.9%) and 249 official gauges (20.1%)[8]. However, there were 21 WOW gauges where on completion of the SQC there were no observations remaining. Whilst the majority of these were short lifespan gauges, some had been reporting for several years.

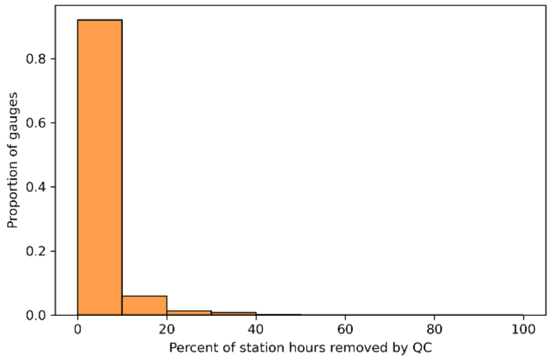


Figure 4.8 Histogram showing data removed from Official gauges by SQC.

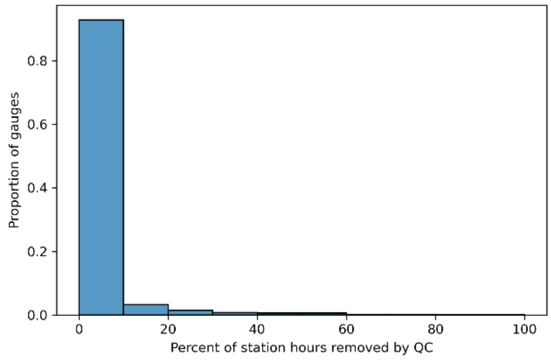


Figure 4.9 Histogram showing data removed from WOW gauges by SQC.

The histograms showing the percentages of data removed from each gague (Figure 4.8 and Figure 4.9). The majority of gauges (WOW: 2,512 or 93%, Official: 1,141 or 92%) had <10% of observations removed by the SQC. There were 1,274 (47%) WOW gauges with <10% observations removed by SQC in urban areas. The mean percentage of hours removed was 19.62% and 4.2% respectively for Offical and WOW gauges, the standard deviation was greater for the Official gauges than WOW gauges (32.28 and 8.45 respectively). The reason for this is explored further later in this section where the individual rules are considered.

The relationship between missing data pre- and post-SQC was assessed to determine if gaps in a gauge record were indicative of broader data quality issues in the WOW dataset, which was not a test in the SQC. The length of record was considered, to determine whether the percentage of missing data was higher in shorter records (as could be expected, as fewer missing observations would result in a higher percentage). Data were filtered to remove gauges where the percentage of missing data was <1%.

---

[7] An issue with SQC was identified that removed the final observation from WOW gauges with an end date and time of 1/1/21 (the last accepted observation for this research). Consequently, a minor modification was made for statistical analysis; rather than reporting on gauges with no observations removed, instead those with <=1 observation removed and an end date pre-2021 were considered.

[8] Caveat that nearest neighbour checks were not applied within the SQC for WOW data.

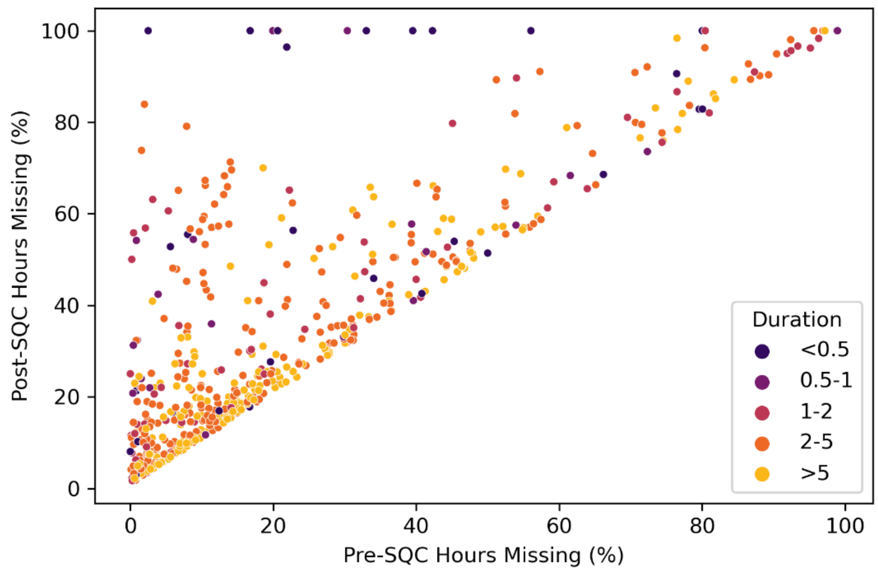


Figure 4.10 Scatter plot showing relationship between pre- and post-SCQ missing data for WOW gauges where >1% of observations were removed by SQC (colour shows duration in years).

Figure 4.10 shows the duration of reporting is unrelated to the percent of hours removed by SQC; however, 9 gauges with reporting frequencies under 1 year had all observations removed by SQC.

The availability of data can be seen to be highly variable between and within the datasets, and more so for WOW data than Official. No relationship was identified between missing data and data quality. The significance of missing WOW data means observations may not be available when expected. A lower percentage of WOW gauge data were removed by SQC than for Official gauges, suggesting that the WOW data are of a comparable if not superior quality, but only under the adage of "*lies, damned lies and statistics*", as the discrepancy in the duration of reporting between the two datasets makes such a comparison unreasonable. The superior quality of WOW data is further challenged when looking at the depth of rainfall reported by WOW gauges and removed by the SQC.

Depth of Rainfall Before and After SQC

The starting mean rainfall depth (in mm) was relatively high from WOW gauges when compared to Official gauges (see Table 4.4). The SQC clearly had a significant impact, reducing the mean rainfall rate from 2.46 mm $h^{-1}$to 0.12 mm $h^{-1}$for WOW gauges, which was comparable with the Official gauge mean rate (0.11 mm h^{-1}). The high original WOW mean rainfall was influenced by extreme values, there were 25,498 (0.04%) observations >1,000 mm at WOW gauges and 41,414 (0.06%) observations that exceeded the threshold value of 92 mm. There was comparatively little change post-SCQ in the Official mean rainfall, in part because the dataset was so much larger than the WOW dataset along with far fewer extreme values in the original Official data.

Mean rainfall per hour (mm)	WOW	Official
Before SQC	2.46	0.12
After SQC	0.12	0.11

Table 4.4 Rainfall rate (mm/hr) before and after SQC at WOW and Official gauges.

The SQC resulted in the removal of almost 95% of rainfall from the WOW dataset, as shown in Table 4.5. In comparison, ~9% of rainfall was removed by the SQC from the Official dataset.

Rainfall (mm)	WOW Rainfall (mm)	WOW Rainfall (%)	Official Rainfall (mm)	Official Rainfall (%)
Start Total	151,790,524	100	27,636,855	100
Removed by SQC	143,943,774	94.83	2,470,869	9.19
After SQC	7,846,751	5.17	25,165,986	90.81

Table 4.5 Rainfall removed by SQC for WOW and Official gauges.

There was a large variation between datasets as visible in the plots in Figure 4.11 and Figure 4.12, with more rainfall being removed from more WOW gauges than Official.

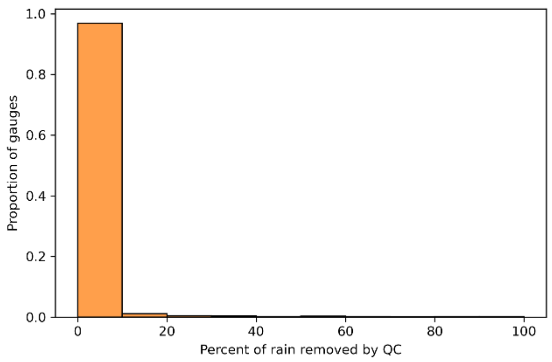


Figure 4.11 Histogram showing rain removed from WOW gauges by SQC

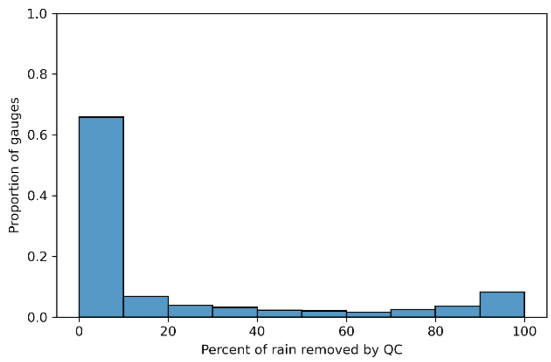


Figure 4.12 Histogram showing rain removed from official gauges by SQC.

The interquartile range was greater in the WOW data than the Official data; however, there were outlier gauges in the Official dataset. It is interesting to note the outliers in the Official gauges where >10% of the data were removed by the SCQ; as this is a closely managed network of gauges, the extent of poor-quality observations at multiple gauges is perhaps surprising.

There was a discrepancy between percentage hours and rainfall removed, with more rainfall being removed from WOW gauges, and from fewer hours of reporting (see Figure 4.13 and Figure 4.14). This is indicative of the excessive rainfall values reported at some WOW gauges, either regularly or intermittently.

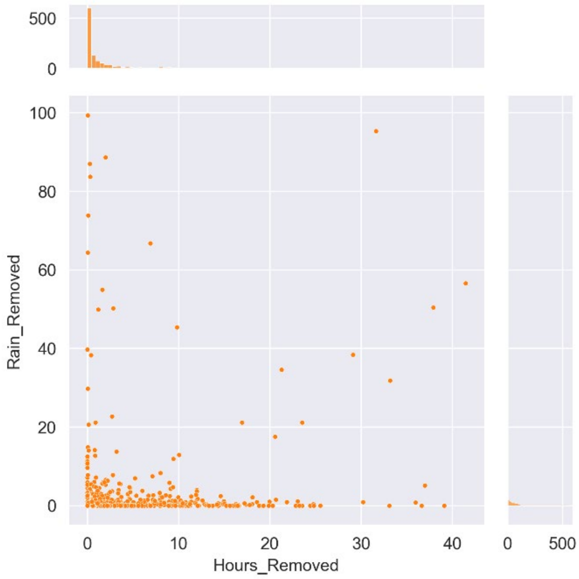


Figure 4.13 Combined scatter plot and histograms showing relative percentage of hours and rainfall removed from Official gauges by SQC.

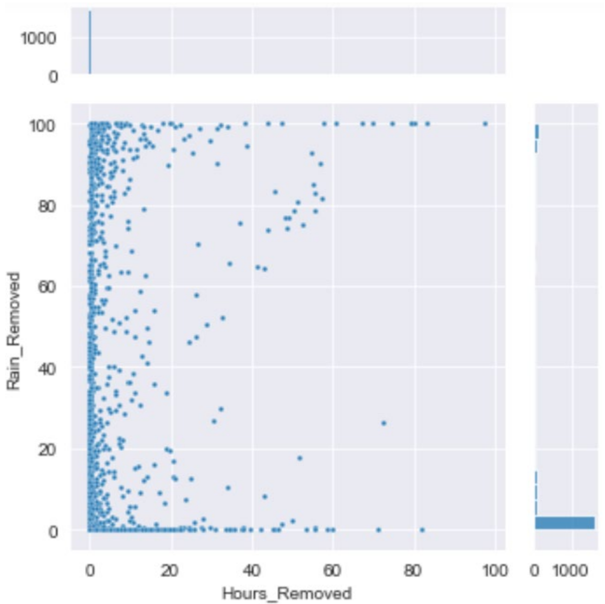


Figure 4.14 Combined scatter plot and histograms showing relative percentage of hours and rainfall removed from WOW gauges by SQC.

The comparison of rainfall depths pre and post SQC between the Official and WOW datasets shows that there are errors in both, but WOW gauges more frequently reported extremely high observations either consistently or occasionally. There were extreme values in the Official data, and as a percentage the SQC removed more Official observations than WOW.

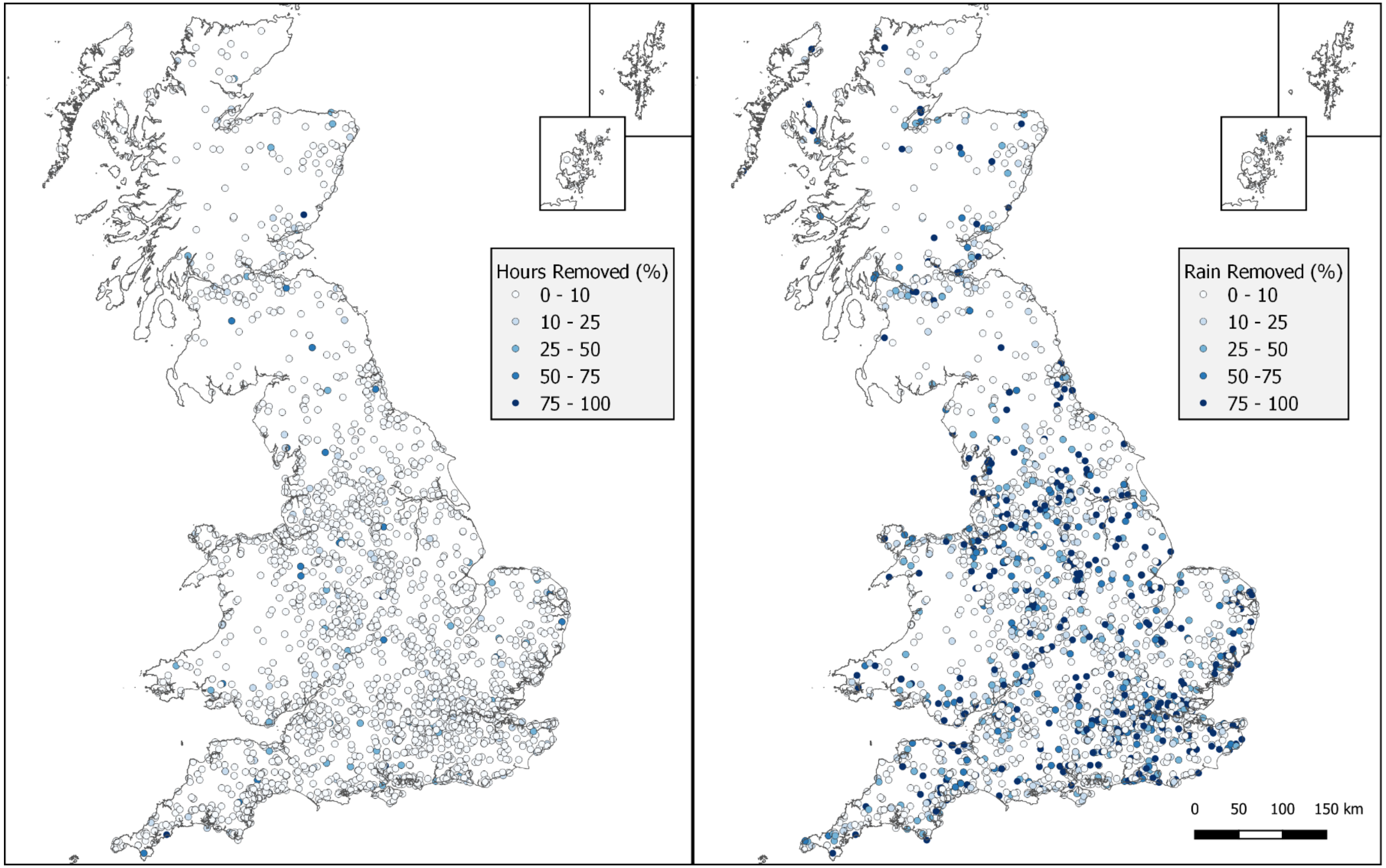


Figure 4.15 Maps showing percentage of hours removed per rain gauge (left) and percentage of rain removed by rain gauge (right) by SQC for WOW dataset.

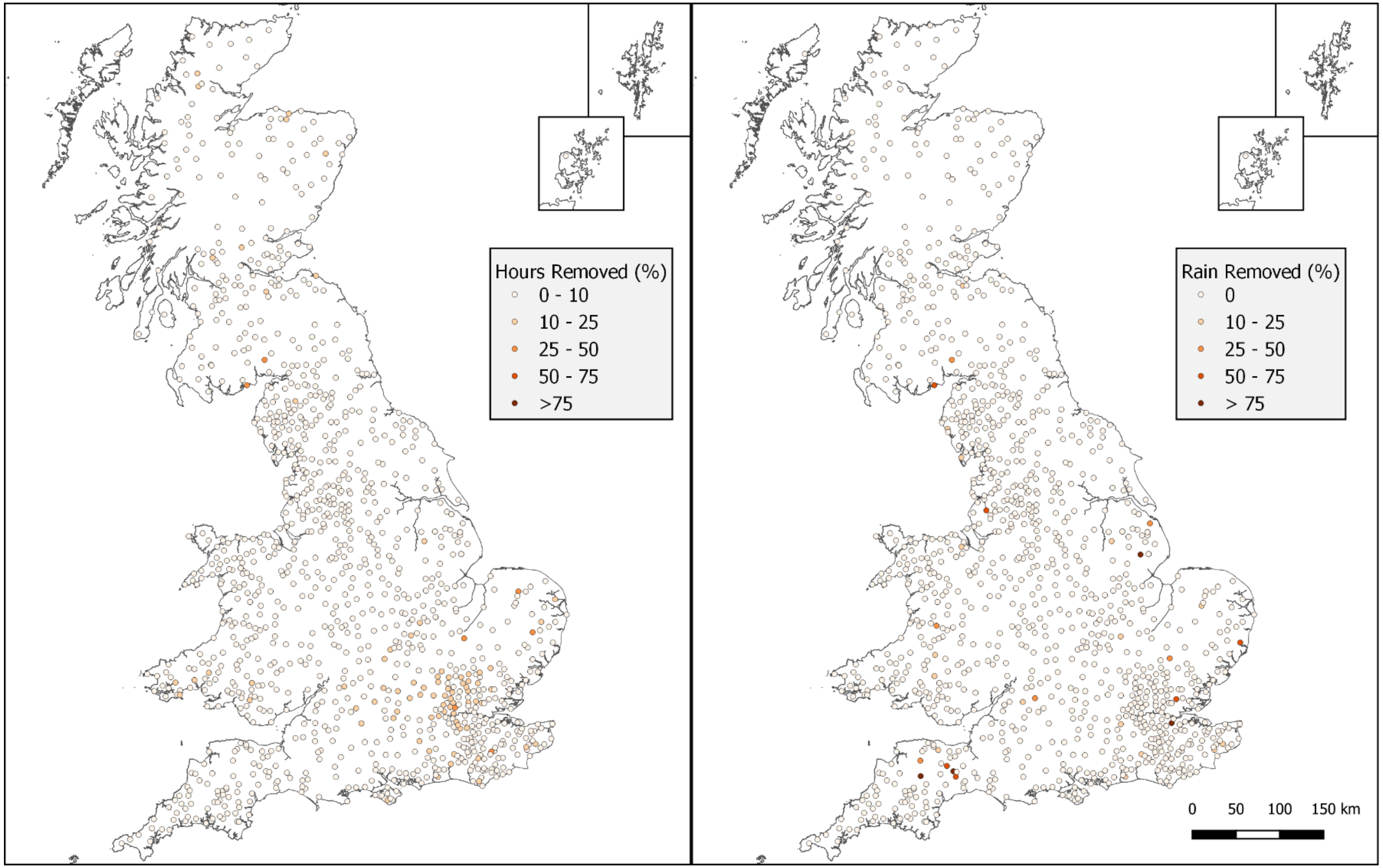


Figure 4.16 Maps showing percentage of hours removed per rain gauge (left) and percentage of rain removed by rain gauge (right) by SQC for Official dataset.

Figure 4.15 shows the discrepancy between the percentage of WOW observations and the rainfall depth removed by SQC per gauge. For comparison, the percentage of hours and rain removed per gauge for Official rain gauges are shown in Figure 4.17. The mapping shows no obvious pattern in the geographical location of gauges that over-report rain, e.g., it was not more common in wetter regions to the west of Britain, with errors appearing to be independent the rainfall amount.

SCQ Performance - Rule Breakdown

Analysis of the performance of gauges against individual rules was undertaken to determine if there were specific issues that were resulting in the removal of gauge data, and to establish whether the same data quality issues existed between the two datasets.

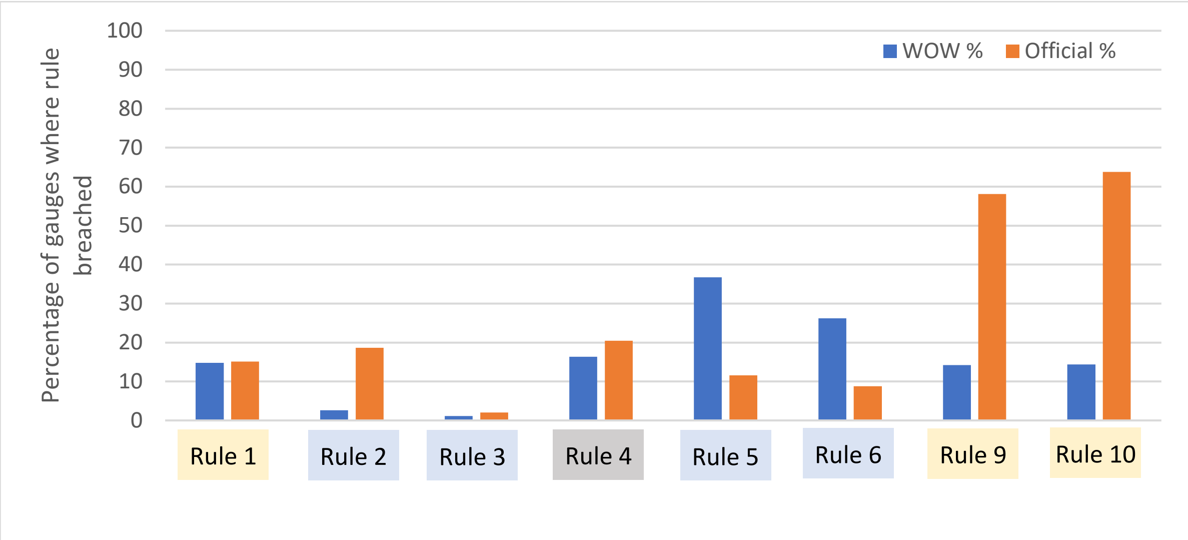


Figure 4.17 Comparison of percentage of WOW and Official gauges at which each rule was used to exclude suspicious hours.

Key to Rule colour coding:

Too much rain

Too little rain

Either too little or too much rain

Figure 4.17 shows the percentage of gauges at which each rule was triggered for each dataset. The figure aims to highlight where there is a discrepancy between the 2 datasets, to establish whether a systematic under- or over-reporting could be identified in either dataset. It is important to note that a rule may have been broken more than once at an individual gauge, and different rules may have been triggered by the same unreliable observation. For example, an hourly observation of 500 mm would trigger Rule 5 and Rule 6

Figure 4.17 indicates that a smaller percentage of WOW than Official gauges breach rules 1,2, 3, 4, 9 and 10, and vice-versa. A rule-by-rule break down is provided based on the percentage of Official and WOW gauges at which each rule was triggered on one or more instances:

Rule 1 was triggered where the 10 largest events in a calendar year were zero, i.e., there was no rain during the calendar year, resulting in all the observations for that year being removed. In instances where the gauge reported for only a short portion of a year, and there was no rain during the period the rule was triggered. This occurred at WOW gauges where the final observation was timestamped 01/01/21. Once the obvious issue in data processing was corrected this rule was triggered at 14.82% (401) of WOW gauges and 12.78% (178) of Official gauges.

Rule 2 was triggered when there were 2 or more consecutive days of suspected daily accumulations. The rule was rarely triggered in WOW gauges (2.62%, 71) compared to Official (16.38%, 220). It is assumed that much of the gauge data is uploaded to WOW automatically therefore it should be relatively unusual for data to be aggregated to daily, as there is no obvious mechanism for that to happen.

Rule 3 highlights where monthly accumulations are suspect, which was relatively rare in both datasets: 1.11% (30) and 2.04% (24) for WOW and Official respectively.

Rule 4 was triggered when there were streaks of repeated values, the count of stations where this was identified was very similar between the WOW and Official datasets; 16.37% (443) and 20.46% (241) respectively.

Rule 5 was triggered when the hourly rainfall was >92mm. This was encountered far more frequently in WOW gauges that Official; 36.7% (993) and 11.54% (136) respectively.

Rule 6 was triggered when the daily rainfall was >314mm which occurred more frequently in WOW gauges that Official; 26.2% (709) and 8.74% (103) respectively.

Rule 9 was triggered where there were excessive hourly dry spells, which occurred far less frequently in WOW gauges that Official; 14.23% (385) and 58.06% (684) respectively.

Rule 10 was triggered when there were excessive daily dry spells, which occurred far less frequently in WOW gauges that Official; 14.34% (388) and 63.75% (751) respectively.

Figure 4.17 shows the percentage of observations triggering each rule, allowing comparison between the datasets to distinguish if one type of error was more common in one than the other. The figure does not differentiate the frequency of occurrence of rule breaking at individual gauges, which is presented in

Figure 4.19 and Figure 4.18. Note that for the purposes of visualisation instances of zero were removed; therefore, the comparison is only for instances where the rule was breached.

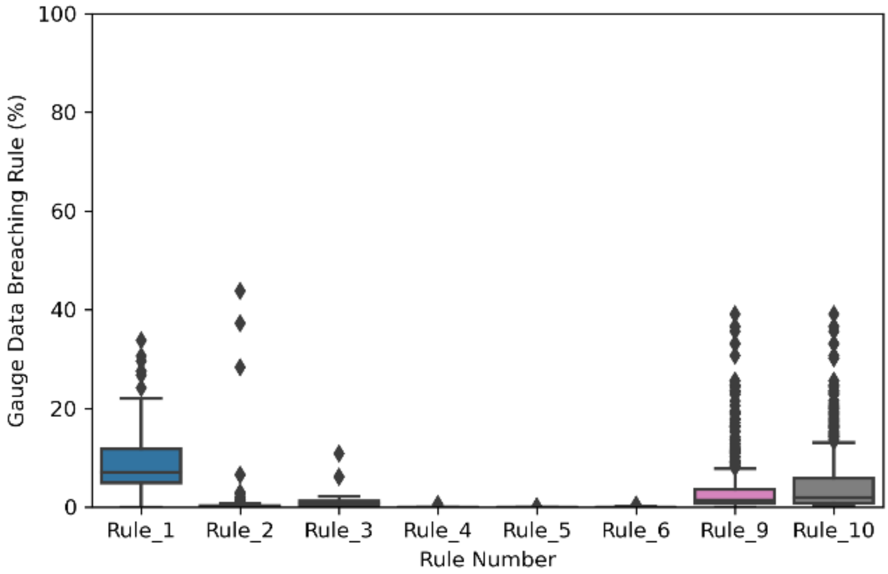


Figure 4.18 Percentage of data per gauge flagged for removal by SQC rule - Official gauges (see numbers of gauges per rule in text).

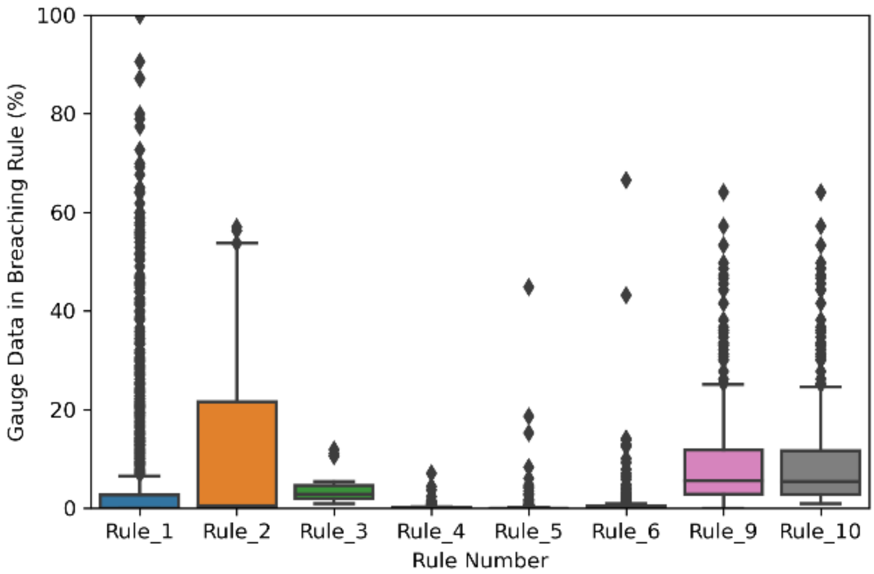


Figure 4.19 Percentage of data per gauge flagged for removal by SQC rule – WOW gauges (see numbers of gauges per rule in text).

Figure 4.17 shows there was a smaller percentage of WOW gauges than Official that breached all rules, bar 5 and 6; however,

Figure 4.19 and Figure 4.18 show that there were gauges where a higher percentage of data was removed by SQC. From this it can be inferred that there were a limited number of WOW gauges that generated poor-quality observations.

As this research is primarily concerned with extreme rainfall, Rule 5 relating to the over reporting of hourly rainfall was investigated further. There were 1,713 (63%) and 1,025 (69%) WOW and Official gauges respectively at which Rule 5 was not triggered. Breaching the rule resulted in 24 hours of observations being removed from the record, due to the low confidence in the observations. For the remaining 993 (37%) WOW and 452 (31%) Official gauges where Rule 5 was breached a selection of the time series were plotted; assess the nature of the error(s).

The first example gauge timeseries had a low percentage of data removed by SQC due to Rule 5, with only 1 hour exceeding 92mm at 1am on 01/01/2017 (see Figure 4.20). The hourly rainfall was reported as 822.96mm, which looked like an annual accumulation reported as an hourly total. The annual (2016-2017) observed rainfall from the gauge based on WOW data was 779.52mm therefore it is not possible to conclusively say this was the reason for the error. The error occurred for one year only (rather than a systematic error that reported annual aggregations each year), as the gauge was reporting from 2011 onwards. The SQC correctly identified and removed the offending observation, as can be seen in Figure 4.21.

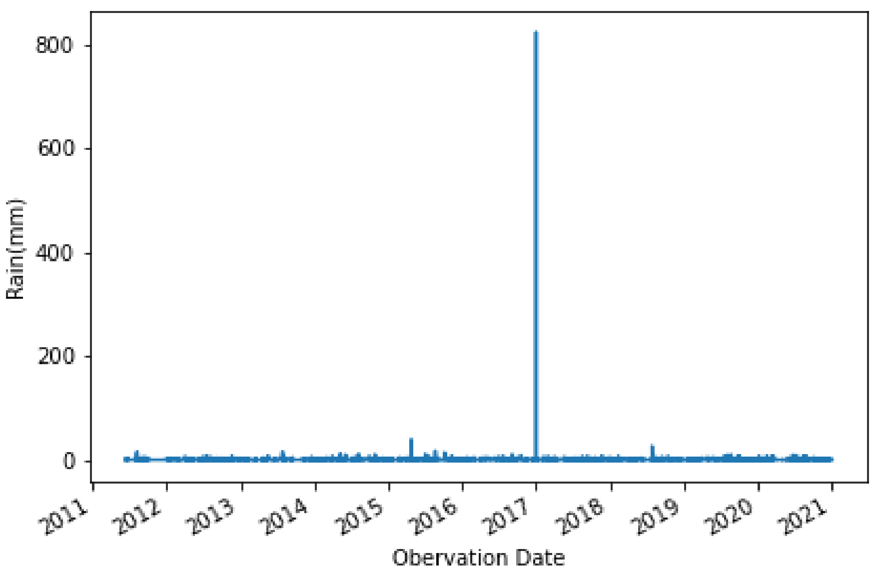


Figure 4.20 Plot showing pre-SQC timeseries for WOW gauge breaking R5 – Example Gauge 1.

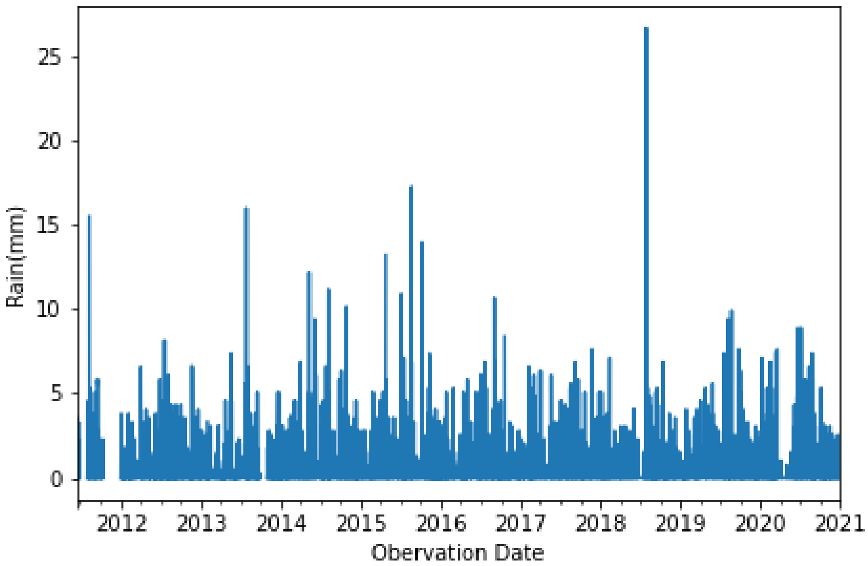


Figure 4.21 Plot showing post-SQC timeseries for WOW gauge R5 - Example Gauge 1.

At the other extreme presented is a WOW gauge where 45% of observations were flagged as exceeding Rule 5 (see Figure 4.22).

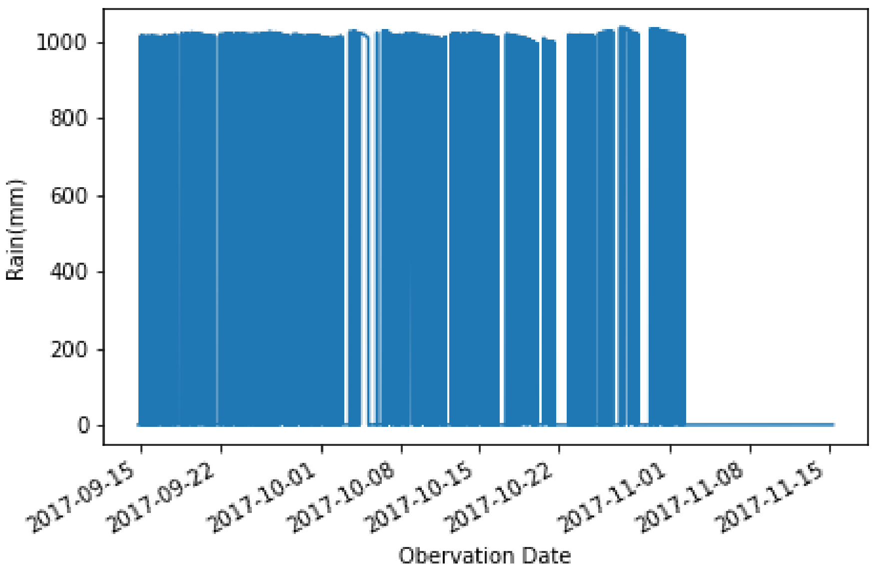


Figure 4.22 Plot showing pre-SQC timeseries for WOW gauge breaking R5-Example Gauge 2.

On inspection of the data there appeared to be an issue with data reporting. The rainfall rate (mm h^{-1}) was provided at 30-minute intervals. The rainfall accumulation was also reported, and was used for calculating the hourly total, but it fluctuated, increasing and/or decreasing between hours. This issue with observations from this gauge is therefore a data reporting/processing error. The inconsistency in data reporting by users risks the loss of valuable data (as in this instance) which could be minimised with enforcement of which data are shared and some periodic validation of gauge data to identify and rectify such errors.

4.4.3 Nearest Neighbour Comparison

The nearest neighbour comparison assessed the consistency between the daily total rainfall at a given WOW gauge and the nearest Official gauge. Official gauge data had been manually reviewed and 'poor' quality gauges removed (237), the remaining Official gauges were used in the nearest neighbour comparison (1,240).

Gauges were not considered to have a neighbour where there was <1 year of overlapping data[9] or where the distance between neighbours was >25 km. There were 1,727 WOW gauges remaining after filtering. Gauges not included in the nearest neighbour assessment were not necessarily reporting poor quality data, but the external validation of quality could not be undertaken. Overlapping days for neighbouring gauges were compared for presence/absence of rainfall on the day (rain >0 mm), Spearman's Rank correlation and the Pearson correlation. Figure 4.23 shows correlation results.

[9] Official observations were only available to June 2018

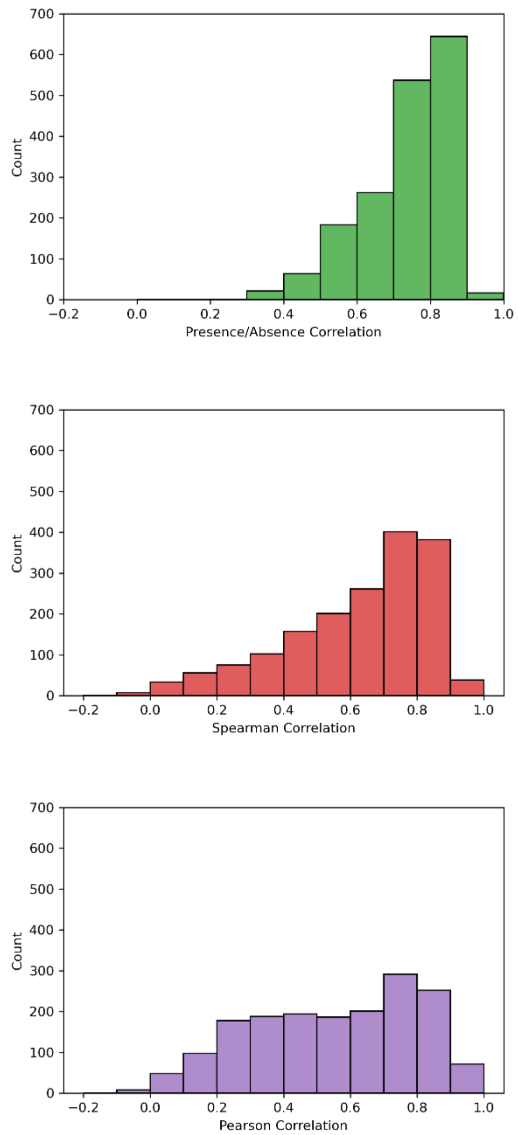


Figure 4.23 Histograms of nearest neighbour correlation plots (presence/absence, Spearman, and Pearson).

There were 1157 (67%) gauges where the correlation for the presence or absence of rain was >=0.7, 786 (46%) gauges where the Spearman correlation was >=0.7, and 592(34%) gauges where the Pearson correlation was >=0.7. The distribution of correlation scores can be seen in Figure 4.24, showing the variation in the interquartile range was greatest for the Pearson correlation.

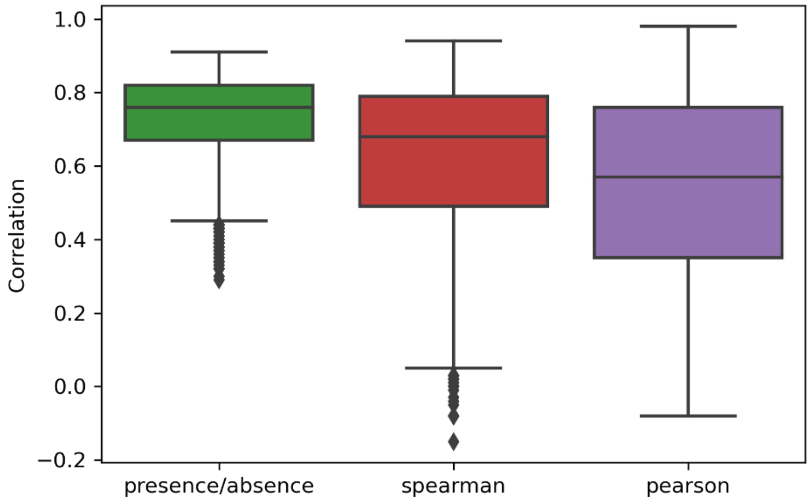


Figure 4.24 Box and Whisker plots showing nearest neighbour correlations between WOW and Official gauges.

There were 534 gauges (31%) where all correlation scores were >=0.7. The correlation scores highlight the variation in the WOW dataset, between gauges and within an individual gauge record. The difficulty was identifying reliable gauges. It is possible to cherry pick only those gauges where minimal observations have been removed by SQC, and where nearest neighbour correlation is strong, but this risks the exclusion of good data within gauge records with intermittent errors.

The nearest neighbour check can potentially be used as an indicator of the very best gauges, and a warning to 'proceed with caution' for stations performing less well in correlation and SQC. On review of individual gauge data, it was apparent that a poor correlation score did not necessarily mean that all observations were of poor quality. Some gauges reported erratically resulting in poor correlation scores, but there appeared to be good data within the record. The analysis of the nearest neighbour correlations suggests that WOW gauges can currently reliably be used as an indicator that rain is occurring, but for quantitative use further manual SQC is required.

4.4.4 Identifying post-SQC Good Quality WOW Rainfall Data

There were 1,242 (46.9%) WOW gauges from which no hours were removed by the SQC, i.e., every observation passed the checks and was found to be acceptable (see locations mapped in Figure 4.25). Of those there were 673 (24.87%) in urban areas. The figure shows both the Official and WOW gauges to demonstrate where WOW gauges may fill gaps in the Official monitoring network. In urban areas having high density of good quality data can be essential when assessing pluvial flooding (Schroeer *et al.*, 2018; Ochoa-Rodriguez *et al.*, 2019b). In some coastal areas and in the Highlands of Scotland, WOW gauges provide the only available rainfall data in the vicinity.

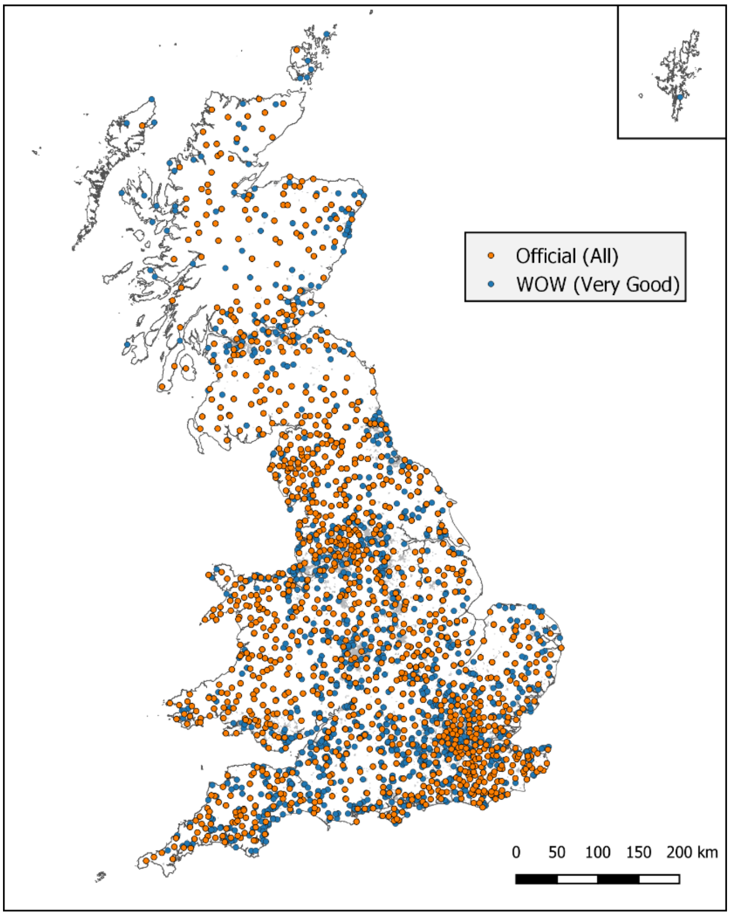


Figure 4.25 Map of WOW gauges with no hours removed by SQC and Official Gauges (selected post-manual review).

The next iteration of mapping constitutes the 'Gold Standard' (without manual validation of individual gauge records) defined as those WOW gauges with no hours removed by the SQC, and strong positive Pearson, Spearman's Rank and presence/absence correlation (all >=0.7). There were 300 (11%) WOW gauges meeting these criteria (see Figure 4.26). These gauges can be identified in the summary statistics for SQC via the link in Appendix 1. Figure 4.26 shows no obvious geographical grouping of the 'Gold' gauges, so the legend indicates the distance to the nearest Official gauge (from which the correlations were derived) to demonstrate where 'Gold Standard' WOW gauges fill gaps in the Official monitoring network. The timeseries of these 300 gauges were manually reviewed to assess the success of SQC and to verify if the data compared favourably with that from the nearest Official neighbour, as the correlation suggested. Examples of possible erroneous data remaining in the gauge records post SQC are presented in section 4.4.5.

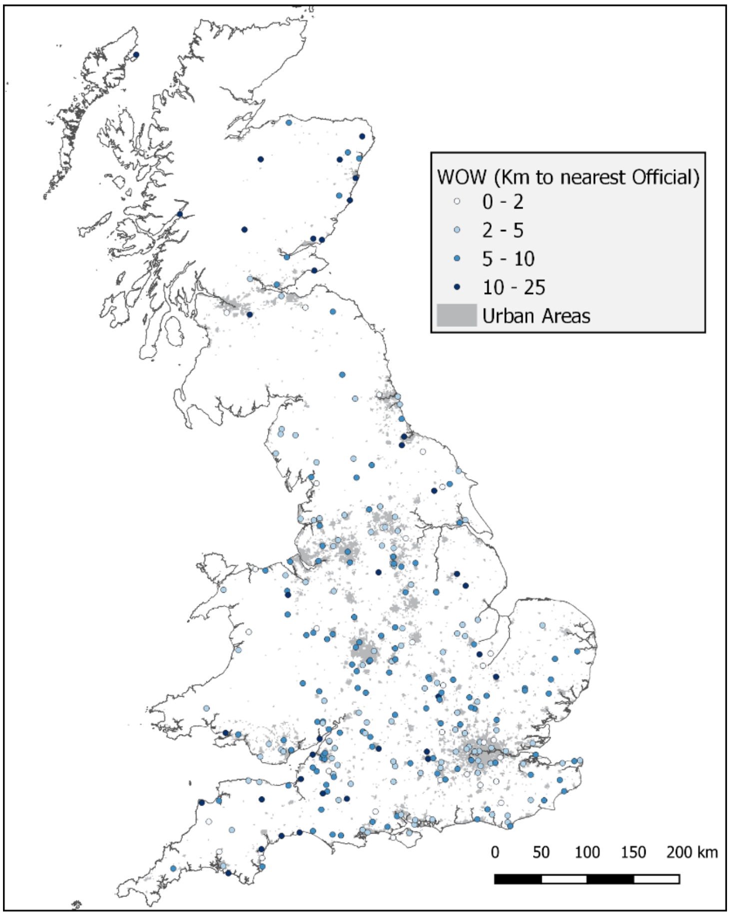


Figure 4.26 Map of WOW Gauges with no hours removed by SQC, and good correlation with nearest Official neighbour (colour scale shows distance in km to nearest Official gauge)

Due to proximity (>25 km to nearest neighbour), or lack of overlapping data it was not possible to calculate a nearest neighbour correlation for 671 (25%) of the WOW gauges with no data removed by the SQC (see Figure 4.27).

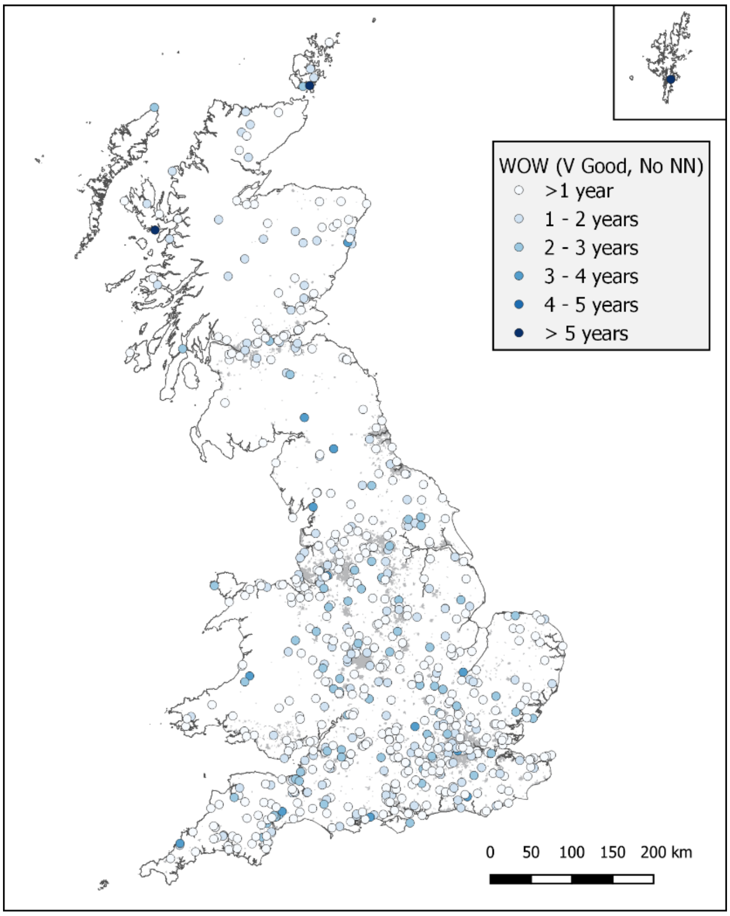


Figure 4.27 Map of Very Good WOW gauges with no nearest neighbour check available (colour scale by duration of reporting).

For these gauges an appraisal using an updated QC dataset (post-2018), manual review of the data, or an alternative measure of independent validation such as comparison with the published gridded rain or radar (Met Office, 2003; Lewis *et al.*, 2022) would help assess data accuracy.

The figures in this section are presented with the caveat that not all gauges were reporting at any one time. Temporal variation of gauge distribution is challenging to present on paper, showing all the iterations and at a scale that is compelling. The intention of this section is not to suggest that these are the only WOW gauges generating good quality data, rather it is to highlight a subset of gauges from which data could be used with minimal post-SQC validation. The criterion of high nearest neighbour correlations is intended to highlight gauges where the absolute rainfall values are most likely to be accurate.

As a follow on from Figure 3.8 presented in chapter three an assessment of the coverage provided by WOW rain gauges reporting in 2018, meeting the highest standard of SQC, with no observations removed and correlation with nearest neighbours being >=0.7 is pictured in Figure 4.28. There are 183 WOW gauges reporting in 2018 that met the 'very good' criteria.

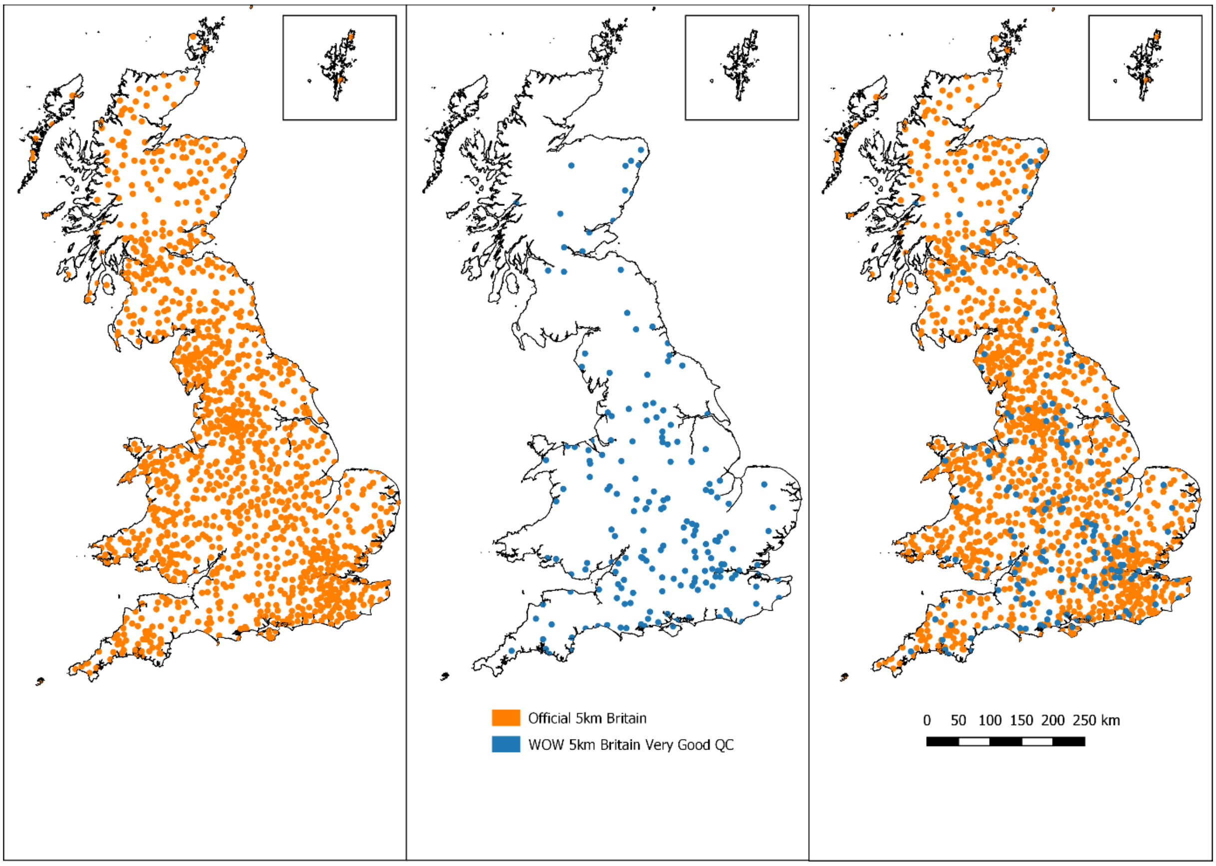


Figure 4.28 Maps of British Official and WOW rain gauge coverage meeting Very Good quality control standard, 2018 (5km scale).

The total coverage provided by the WOW gauges at an extrapolated distance of 5km^2 is 12,808km^2, and the 'unique' coverage, i.e., areas where there are WOW gauges without Official is 6,564km^2. Whilst these are relatively modest gains the distribution of the WOW gauges offers additional data points in urban areas, and rural locations where there may not be an Official gauge for several kilometres.

4.4.5 Efficacy of SQC for WOW Rain Observations

As the SQC was developed for rain gauges operated by national meteorological services there is a reasonable expectation that the data would have been gathered from gauges subject to quality assurance processes (e.g. sensor calibration), as per WMO guidelines (WMO, 2018; WMO, 2021). This section addresses the question of the suitability/reliability of the SQC when applied to WOW gauge data, given that WOW gauges are less likely than Official gauges to have been installed in accordance with WMO guidelines, and there is no formal process of quality assurance.

There is a broader consideration (that equally applies to Official gauge data) regarding the distinction between 'passing' SQC, versus the accuracy of an observation. Despite all the checks SQC makes, ultimately there is no guarantee that the resulting SQC'd output provides an accurate reflection of rainfall (Allerup *et al.* (1980) provide a comprehensive guide to rain gauge errors); rather, it is a confirmation that the checks have been passed. Determining accuracy is a far more difficult proposition, particularly for rainfall as a high degree of spatial variability is common (and indeed a driver for the consideration of crowdsourced data to fill gaps) (Ochoa-Rodriguez *et al.*, 2015; Schroeer *et al.*, 2018). The nearest neighbour checks provide some reassurance that there is a degree of consistency between rain gauges but does not guarantee absolute accuracy.

The analysis in the previous sections of this chapter was based on the premise that the SQC is appropriate, and functions as expected, resulting in datasets that are an accurate representation of rainfall after erroneous data have been removed as required. However, during manual review of gauge data, instances were noted where the SQC had not identified what looked like obvious errors. "Looked like" is a critical consideration, as it was the manual review of plotted timeseries that allowed the identification of apparent issues. This issue is explored further in this section.

SQC Efficacy

Some examples are presented of possible failures by the SQC to identify erroneous observations. The first example in Figure 4.28 is from a WOW gauge that had no observations removed by SQC, nearest neighbour correlations were all good (Pearson, Spearman's Rank and presence/absence of daily rain correlation between neighbours were >=0.7) and there was <10% missing data pre-SQC.

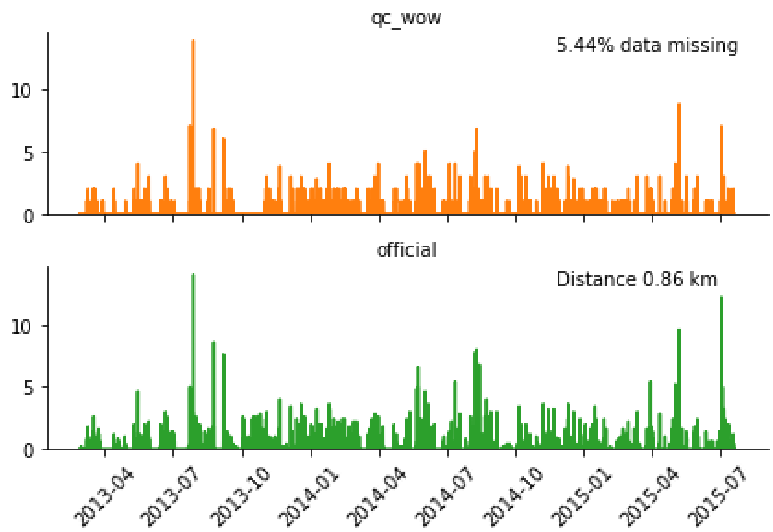


Figure 4.29 Example gauge 1 – complete timeseries of rainfall (mm) from WOW gauge (ID 41937481), and nearest Official neighbouring gauge.

The rainfall depths from the WOW gauge appear to frequently repeat, which may be correct as rainfall <5mm for most observations; however, there is more variation in rainfall depth at the nearest Official gauge (0.86 km away). One possible explanation could be a tipping-bucket size of 1mm (see Figure 4.29); however, in the figure there is one observation on 9th January 2014 that is not a factor of 1, and other data points outside the time series shown were found to vary. The original data were reviewed but there was no further insight as to whether an error existed.

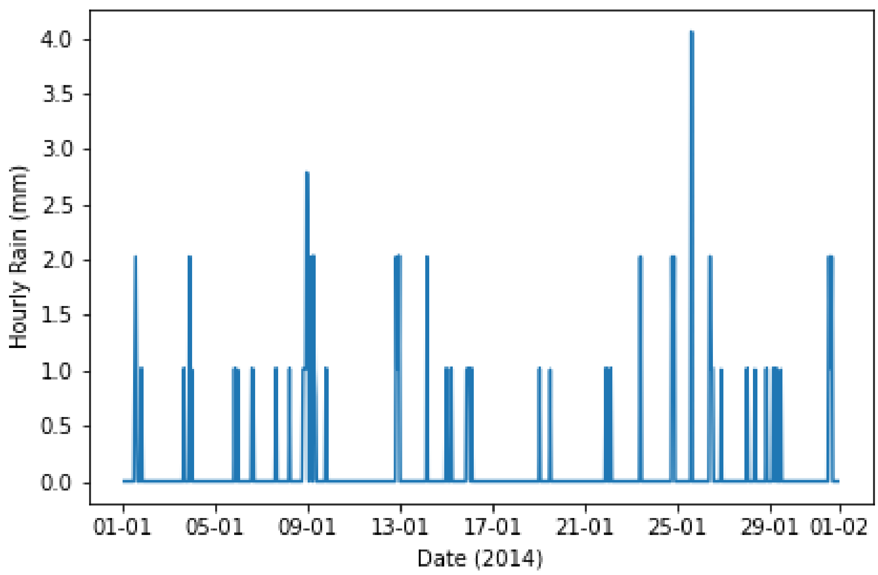


Figure 4.30 Example gauge 1 - selected timeseries for WOW gauge (ID 41937481), highlighting apparent error.

When plotted with data from the nearest official neighbour the strong correlation between the datasets was apparent (see Figure 4.30). The best approach to resolving the data reliability question is likely to be communicating with the PAWS operator.

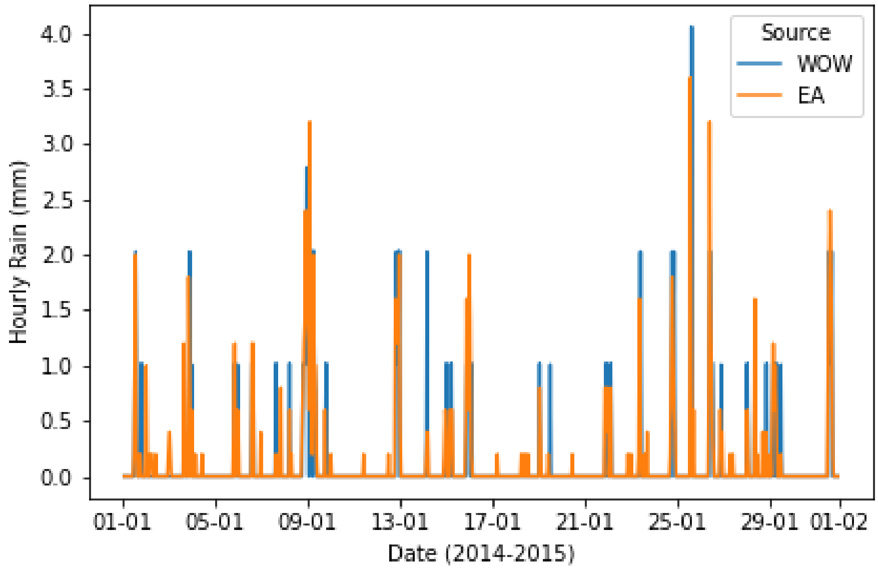


Figure 4.31 Example gauge 1 – selected timeseries for WOW (ID 41937481), and nearest Official neighbour providing a comparison to explore an apparent error.

The possible erroneous reporting highlights the need for timely, frequent data quality validation, as it may be possible to quickly resolve any issue or confirm the observation as accurate. This approach is endorsed by the WMO (WMO, 2021), and although it requires more active management of data it has the potential to enormously improve confidence in WOW data. The hourly observations in Example 1 are of relatively low consequence when considering the impact of extreme rain, as although the data appear erroneous, there is good correlation with a very close neighbouring gauge, and all the observations are low, with only one exceeding 10mm over the duration of reporting.

Other potential errors were identified from WOW gauges where no hours were removed by SQC, and the correlation was good. There were high values not exceeding the triggered value of 92 mm h^{-1}, that were therefore not highlighted by the SQC process. Example 2 is shown in Figure 4.31, where high rainfall was recorded at a WOW gauge with no corresponding rainfall at the nearest Official neighbour (2km away). There is no insight possible from the presented time series to determine whether the potential fault lies with the WOW gauge or observations from the Official monitoring network.

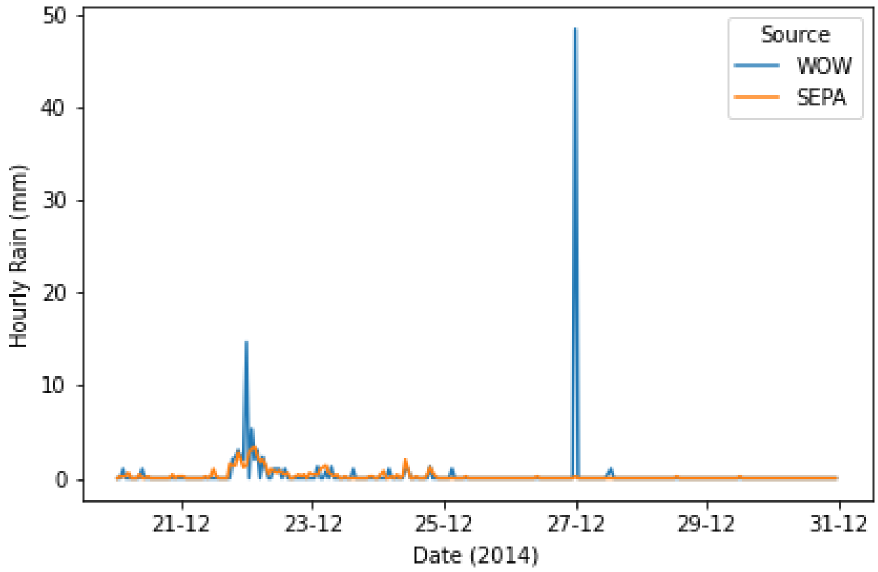


Figure 4.32 Example gauge 2 - high hourly rainfall at WOW gauge (ID:32265743) not exceeding 92mm hr^{-1} with corresponding timeseries from nearest official neighbour.

High rainfall observations were noted in two other 'Gold standard' WOW gauges where no observations were removed by SQC, and where there was strong correlation (>=0.7) with the nearest official neighbour. The erroneous observations may have been removed had the nearest neighbour check not been removed from the SQC.

In Example 3, the SQC has successfully removed erroneously high rainfall observations (see Figure 4.32); however, on removal of the extreme events, multiple high rainfall observations become apparent, along with persistent rainfall ~10 mm h^{-1}which are presumably erroneous results.

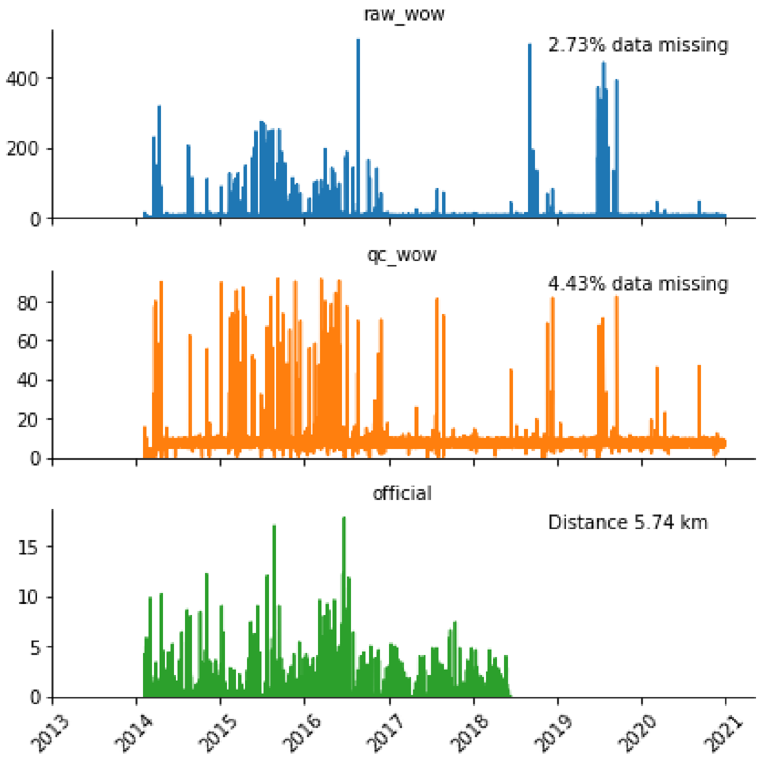


Figure 4.33 Example gauge 3 – complete timeseries at WOW rain gauge (ID:431606049) showing erroneous rainfall post-SQC.

A further example of a possible SQC failure is presented in Figure 4.33 displaying a timeseries from which zero hours were removed by SQC. The Spearman's rank and presence/absence correlations were both >0.8 and there were had 603 days of overlapping data with neighbouring official gauges. The Pearson correlation was lower (0.6) than the Spearman or presence/absence.

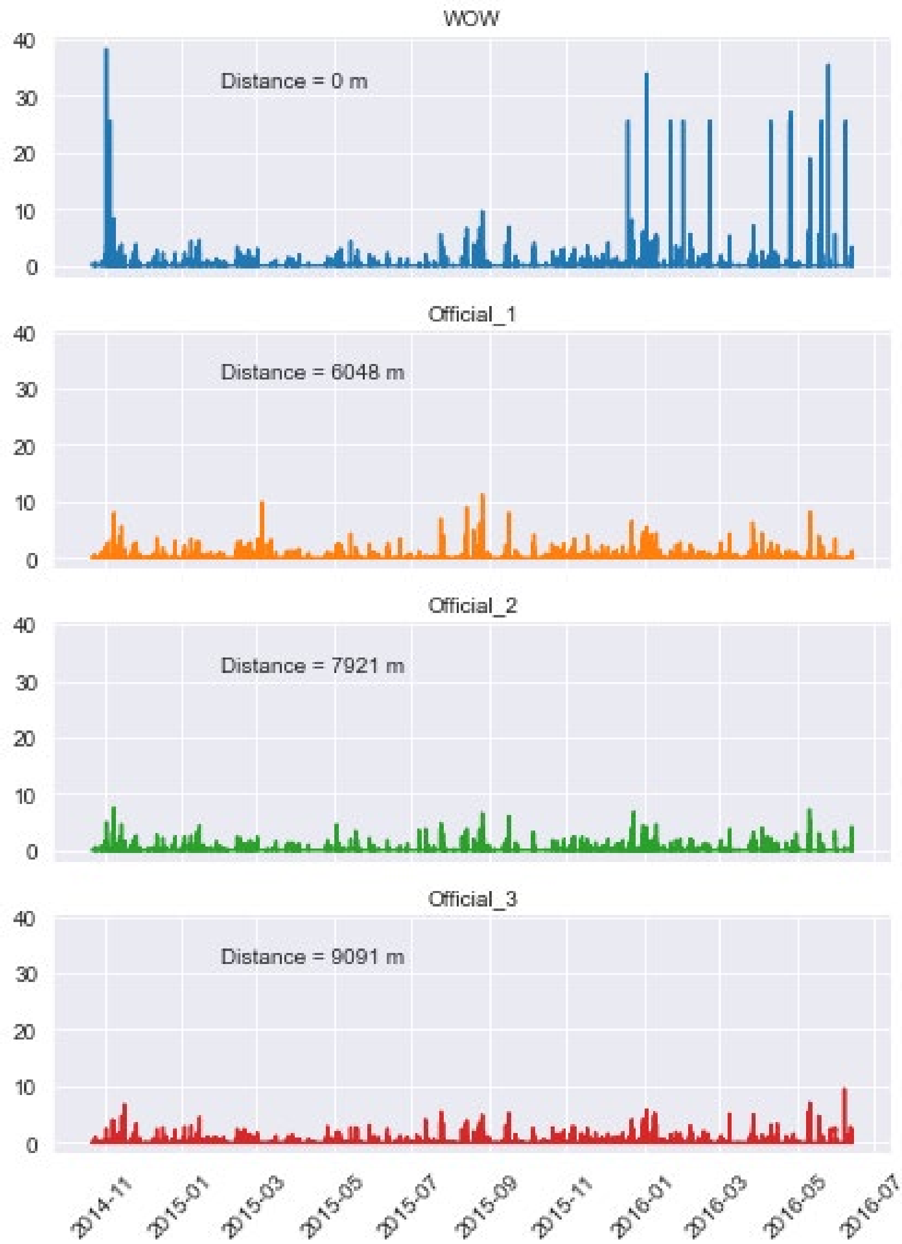


Figure 4.34 Example of WOW observations with nearest neighbouring Official observations (full period of WOW gauge record)

Figure 4.33 demonstrates where observations were not removed by SQC but were not likely be an accurate representation of rainfall (intermittent late 2015 onwards). When compared to rainfall in neighbouring Official gauges within 10km of the WOW rain gauge there are no corresponding elevated observations (>10mm h^{-1}).

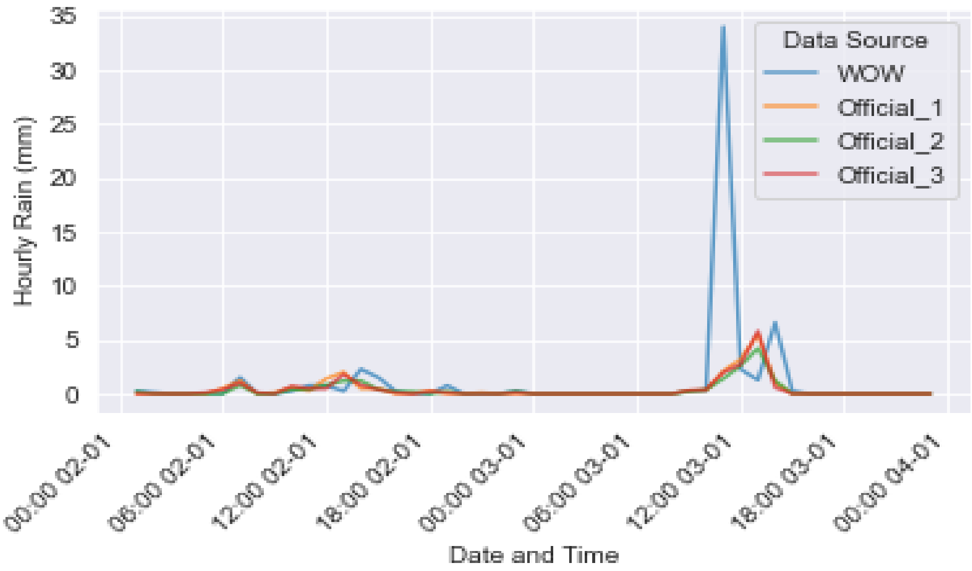


Figure 4.35 Selected timeseries from WOW and neighbouring Official gauges showing discrepancy in observations

Figure 4.34 shows a peak at midday on the 3rd of January. There were no missing observations in the hours preceding the peak that could result in an accumulation. In isolation it is possible the observation represented rainfall at the gauge; however, there were multiple observations inconsistent with surrounding gauges, so it appeared to be an intermittent error. This example demonstrates a gauge that is seemingly generating reasonable observations most of the time, but for some reason(s) occasionally loses accuracy. As the nearest official neighbour is 6km away this gauge could be useful if highly detailed localised observations were required for convective event analysis, but the potential for error means that it would most likely be excluded from analysis as there could be no confidence that an observation was a true reflection of rainfall, without the corroboration of a nearby official gauge.

The examples provided suggest that the SQC is not optimal for some of errors occurring at WOW gauges. Additionally, the use of the SQC in a bespoke manner (i.e., running nearest neighbour checks outside of the SQC) is sub-optimal. Further examples of SQC performance are provided in Chapter 5, including manual gauge review. It is impossible to determine the source/cause of any error from the data alone, and again this makes the case for a more proactive approach to quality control closer to the time of observation.

Observation Accuracy

Q2. related to the suitability of SQC to determine whether rain observations are an accurate representation of rainfall. One option for data validation is the comparison of observations against an independent dataset, in addition to the SQC applied here in this chapter. Given the extent and the size of the WOW database it was not practical to compare all gauge observations against the corresponding radar rainfall. In the subsequent chapter such an approach was adopted, with comparisons made between gauge observations and co-located radar. As radar is not without accuracy issues (Ciach *et al.*, 1999; Steiner *et al.*, 1999; Trapero *et al.*, 2009; Villarini *et al.*, 2010; Peleg *et al.*, 2013) such an approach is not a complete failsafe but adding in the additional layer of data scrutiny allows for further validation of the absolute rainfall.

4.5 Case Study – All Citizen Scientists are Equal, but Some are More Equal Than Others

4.5.1 Introduction

There are several factors that may diminish rainfall data quality, including the accuracy of the sensor/gauge; the suitability of the gauge location; the ability of the operator to recognise and address issues; the level of care and attention given to the gauge; and issues arising during the transfer of data (see Literature Review). PAWS operators may not employ a quality assurance process where gauges are calibrated, manual check gauges are referenced, and unusual results are investigated and rectified. This uncertainty regarding quality assurance, alongside the likely compromises in location and the quality of available equipment, mean there are multiple ways crowdsourced rainfall data suffer reduced accuracy. The opposing view is that PAWS operators can dedicate time and attention to their equipment and observations in a way that professionals maintaining an official network cannot. It could be several weeks between technician visits, during which time errors may have occurred and data quality diminished. In addition, a technician visiting multiple locations may not have an appreciation of the results to expect and may not be involved in the data processing/analysis and would therefore be one step removed from being able to identify and rectify issues.

The special interest group 'Climatological Observers Link' (COL) is comprised of members with a range of backgrounds, from professional meteorologists to schools. There are around 350 members, predominantly based in the UK, and many share their weather observations in a monthly bulletin circulated to members. There are limited data published online, with only an example bulletin available via the website (https://www.colweather.org.uk/). In January 2020 COL was approached via the Secretary and a request for rainfall data from PAWS was circulated amongst members. Data were requested that approximately overlapped with the WOW cohort reporting period (2011 – 2020).

4.5.2 Results and Discussion

Responses were received from 14 COL members, comprised of the direct provision of data and the sharing of WOW IDs. The longest record received spanned from 01/01/1998 – 21/12/2021 and the shortest spanned 01/01/2018 – 12/11/2021, which came from a volunteer who had moved house during the study period who provided records from both locations.

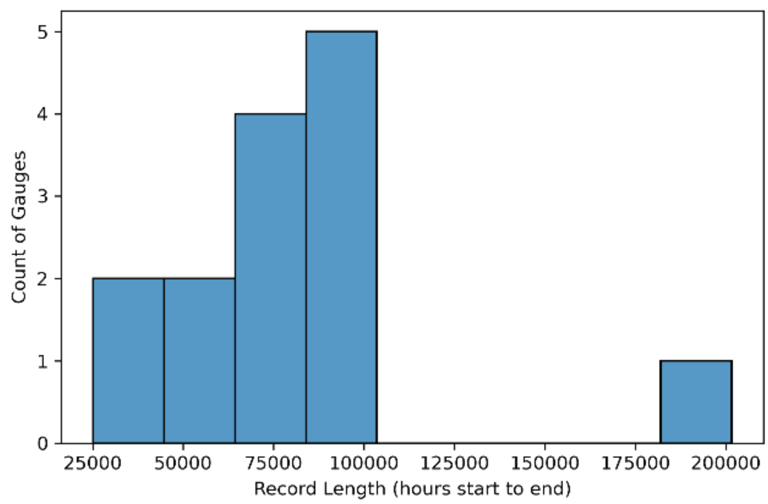


Figure 4.36 Duration of reporting from COL gauges.

A reason for the higher proportion of gaps in the WOW record became apparent during the analysis of COL data: the failure of rain gauges. It is presumed that in a packaged weather station like the Davis Vantage Vue 2 that results from the other sensors would continue to be sent if the rain gauge failed, which may erroneously appear as zero rain, rather than no rain recording.

Table 4.6 shows a summary of the pre- and post-SQC data availability provided by COL members. The difference between data sent directly or accessed via WOW is distinct, with a higher percentage of missing data pre-SQC from the gauges sharing data via WOW.

Observation Origin	Missing Pre-SQC (%)	Missing post-SQC (%)	Hours removed (%)	Rainfall removed (%)
Direct_1	0.00	0.00	0.00	0.00
Direct_2	0.00	5.95	5.95	0.00
Direct_3	0.96	1.67	0.71	0.00
Direct_4	10.72	0.00	0.00	0.00
Direct_5	0.00	0.00	0.00	0.00
Direct_6	0.00	0.00	0.00	0.00
Direct_7	0.01	0.01	0.00	0.00
Direct_8	0.00	0.00	0.00	0.00
Direct_9	0.01	0.02	0.01	15.92
Direct_10	0.00	1.70	1.70	0.00
WOW/COL_1	12.18	12.48	0.31	93.56
WOW/COL_2	6.25	6.31	0.06	1.60
WOW/COL_3	5.90	6.49	0.59	24.55
WOW/COL_4	10.99	11.05	0.05	5.46

Table 4.6 Summary of pre- and post-SQC data and rainfall available from COL members.

The high percentage of rainfall removed from WOW/COL_1 was due to erroneous extremes in the data, which were not encountered in the datasets provided directly. Extensive communication with several of the contributors and the time of data processing provided an insight into the diligence with which they collect and record weather data (see further discussion in section 6.4.4).

The combined hours of observations received from COL members was 1,093,380, with 49,020 hours missing (4.5%), with missing hours predominantly from COL records accessed via WOW (see Table 4.6). The records sent directly were missing a maximum of 10.72% of hours (from 'Direct_4'), and 5 gauges were missing zero observations. The operators of 'Direct_4' and 'Direct_8' provided a detailed explanation of why data were missing. From the 4 WOW/COL gauges the maximum missing data was 12%, with a mean of 8.9% missing. As a comparison the WOW dataset was missing 20.78% and the Official 4.94% hours.

Application of SQC resulted in 8,652 observations (0.80%) being removed from the COL dataset (WOW 2.52% and Official 2.79%). There were 6 gauges from which zero hours were removed. The mean rain depth from the COL data was 0.49mm hr^{-1} pre-SQC and 0.19mm hr^{-1} post-SQC, which is a higher post-QC mean than that of the WOW or Official datasets. The COL sample size was small, and the missing data were minimal which could have resulted in a higher mean than recorded in the Official dataset.

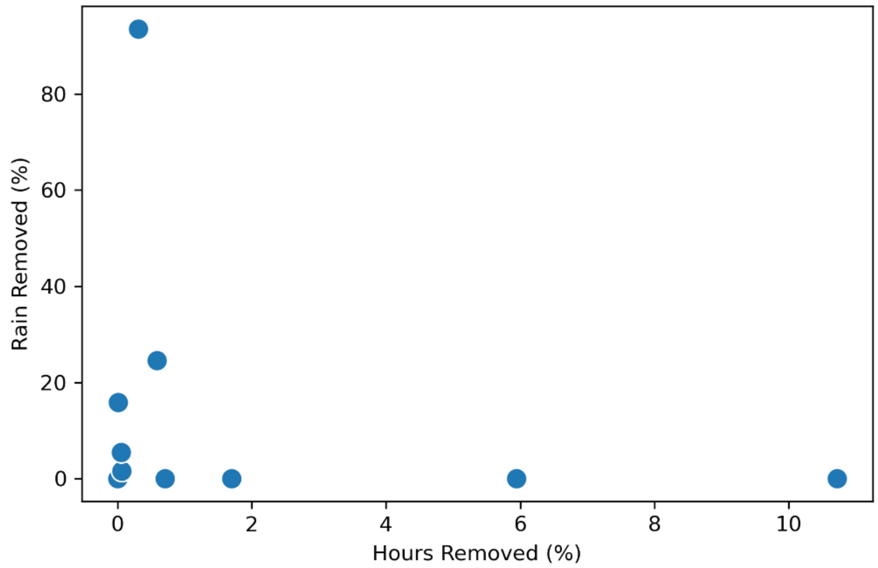


Figure 4.37 Relative percentage of hours and rainfall removed from COL Gauges by SQC.

As with the WOW dataset there was a higher percentage of rainfall removed by the SQC than hours (see Figure 4.36). The 10 rain gauges where COL data were submitted directly for research were subject to nearest neighbour comparison of daily data, descriptive statistics are provided in Table 4.7.

Statistic	Distance (m)	Pearson Correlation	Spearman's Rank Correlation	Presence/Absence Correlation
Minimum	1,779	0.33	0.37	0.61
Maximum	6,860	0.94	0.92	0.9
Mean	1,779	0.73	0.75	0.80
Standard deviation	1,812	0.21	0.18	0.08

Table 4.7 Descriptive statistics of nearest neighbour analysis on COL data.

The 50th centile for Pearson correlation was 0.8, indicating the high degree of agreement between COL rain gauge daily rainfall and their nearest Official neighbours (WOW for 0.57). It was clear on manual review that data provided directly rather than via WOW were less likely to have had data/rainfall removed by SQC and have a better correlation with the nearest Official neighbour than data submitted to WOW.

COL members have self-identified as having a specific enthusiasm for weather observing, therefore it is not surprising that their data are of a generally good quality. The variation between data submitted via WOW and that sent directly for research reflects the benefit of active participation in rainfall observation.

4.6 Discussion

The descriptive statistics and statistical quality control applied to the WOW and Official rainfall gauge datasets presented in this chapter provide insight into the quality of WOW data.

The automated SCQ of the WOW data using code developed for the SQC of a global hourly dataset (Lewis *et al.*, 2021) and applied to Official rain gauge data from the UK (Villalobos-Herrera *et al.*, 2022) allowed for data quality to be compared. There were a disproportionate number of hours and consequently the hours of reporting in the WOW dataset, which consisted of ~35% the number of observations as compared to the Official dataset. The COL dataset comprised of data from only 10 PAWS. As the WOW archive grows with time it would be valuable to reconsider the comparison of data quality.

The performance of Official data when subject to SQC may be a surprise for some users, in that the percentage of hours removed from the Official and WOW data were marginally more from the Official dataset than the WOW dataset (2.52% and 2.79% respectively). There were significant instances of over reporting of rainfall in the Official dataset, including 11.54% of Official gauges having >=1 observations >92mm. The SQC results serve to highlight that all data can be flawed, and there should be an equal degree of scrutiny of Official datasets prior to use to ensure reliability. Research by Villalobos-Herrera *et al.* (2022) has demonstrated the value of sub-hourly SQC, which allows for more careful assessment of data quality that can be lost during aggregation of data to hourly.

Rainfall is a difficult weather parameter to work with. The natural variability coupled with the potential for error in all measurement methods mean anyone using rainfall data must become comfortable with the concept of uncertainty. Beven (2019) argued that hydrology is an ‘inexact’ science due to unknowns, including the inaccuracies inherent in the measurement and extrapolation of rainfall. Placing too much emphasis on error can unnecessarily limit data use, and our desire to categorize data as ‘good’ or ‘bad’ runs the risk of missing opportunities to extract insight from the imperfect.

Determining whether WOW data were of ‘good’ quality proved to be complex. Previously researchers have used paired Official and PAWS rain gauges to demonstrate data quality (Bell *et al.*, 2015; de Vos *et al.*, 2017) with mixed results. Others have chosen not to look at the quality of individual rainfall data, either disregarding it (Holley *et al.*, 2020) or applying methods that limit the influence of any single rain gauge before using data in combination with Official data (Bárdossy *et al.*, 2021). Where data have been subject to quality control, they have originated from rain gauges from one manufacturer using a specific data transfer mechanism for upload, which allowed for errors to be more readily identified as there were fewer variables (de Vos *et al.*, 2019). Given the range of PAWS types and the different methods of data transfer it was assumed that an SQC method devised for weather stations globally would be most effective in identifying errors; however, it may not have captured all the errors, due to the issues like power cuts, gauges being moved, and operator apathy to calibration that are likely to be more prevalent in PAWS data than from and Official monitoring network. The inverse may be true of some PAWS operators; data sent directly from COL members was of excellent quality, with rich metadata complementing the observations. People with a passion for the weather may well pay more care and attention and be able to provide more insight that is available from an Official database, which is one of the key benefits of citizen science (Walker *et al.*, 2016; Liu *et al.*, 2020) not being well exploited via WOW currently. Ways to overcome these issues are discussed further in chapters 5 and 6.

This research demonstrated that crowdsourced citizen science rainfall data can be excellent quality, however, there is an issue with credibility. Accuracy is distinguished from credibility, with the latter being a function of the confidence users have in data (Flanagin *et al.*, 2008). Frustratingly once credibility has been achieved confidence can be misplaced, with users undeterred by fluctuations in accuracy. Official rain gauge data is a good example of this phenomenon, as despite the questions raised by Villalobos-Herrera *et al.* (2022) there is apparently a greater willingness to use Official data over citizens science data. It has been noted during this research that users of rainfall data tend to overestimate the validity of published Official rainfall data and are willing to accept errors that would not be tolerated in citizen science observations.

Given that Muller *et al.* (2015) were lauding the benefits of crowdsourced rainfall data, and seven years later it is still not routinely used it is critical that quality (whether that be improvements, or simply validation) is addressed if the potential for WOW rainfall data is to be fully realised. Given the low number of publications using WOW or other PAWS rainfall data it appears that there is very little will to take advantage of it for meteorological or hydrological applications. It is hoped that this research goes some way to promote the use of PAWS rainfall data.

4.6.1 Response to Research Questions

The aim of the research presented in this chapter was to assess the quality of WOW rain gauge data, compared to the quality of rainfall data from Official ground-based gauge networks in Britain. This has been achieved by addressing the research questions posed in section 4.2.1. as follows:

Question 1. How good is WOW rain gauge data?

2,706 WOW gauges were subject to SQC. The WOW rain gauge data was found to be of generally good quality, with 93% of WOW gauges having <10% of hours removed by SQC, as compared to 92% of Official gauges (see section 4.4.2). A further breakdown of the quality assessment is presented in the following research question responses.

Q1a. How does the availability of WOW rain gauge data compare to that from Official rain gauges?

Descriptive statistics and automated SQC were used to assess rainfall data quality in WOW. Prior to any SQC the WOW rain gauge records were found to be missing 21% of data as compared to Official gauges that were missing 5% (see section 4.4.1). The SQC removed 2.52% of WOW observations as compared to 2.79% of Official observations meaning there were almost 65.5 million hours of WOW observations post-SQC and 241 million hours of Official gauge data. The analysis may be skewed by the relatively short reporting lifespan of many WOW gauges, and the lower number of hours of observations in the WOW dataset as compared to the Official dataset (see section 4.4.2). The missing WOW data does mean that crucial observations may not be available as expected.

There was a discrepancy between the number of hours removed by the SQC and the depth of rain, due to some excessively high reporting of rain at some gauges (>10 m h^{-1}on occasion). Such obvious data errors were removed by the SQC, but further probing into the cause of the error closer to the time of reporting would help improve data quality. If that is not possible the application of sub-hourly SQC may be more effective in identifying and removing data errors (Villalobos-Herrera *et al.*, 2022).

Q1b. What were the results of SQC on the WOW data and how does that compare with Official data?

Contrary to the received wisdom that citizen science data are of poor quality, there were 1,242 (46.9%) of WOW rain gauges with no data removed by SQC, as compared to 249 (20.1%) of Official gauges. Most of the WOW and Official gauges had <10% of observations removed by SQC (WOW: 2512 or 93%, Official: 1141 or 92%) (see section 4.4.2).

Analysis of SQC demonstrated that WOW gauges disproportionately triggered rules relating to excessive rain, with 993 WOW gauges (26.2%) having observations greater than the UK hourly record of 92mm/hour, as opposed to 103 (8.74%) Official gauges. Conversely Official gauges were more likely to have data removed due to under reporting rainfall, as indicated by dry spells, with 751 Official gauges reporting excessive daily dry spells (63.75%) opposed to 385 (14.34%) WOW gauges.

It was not possible to subject all WOW gauges to nearest neighbour checks as some were at a distance >25Km, or the duration of reporting did not (sufficiently) overlap with their neighbours (see section 4.4.3). Of the 1,727 WOW gauges subject to nearest neighbour comparisons, 300 (11%) passed SQC with no data removed and with daily rain totals that correlated well with the nearest neighbour, indicating they were generating very good quality data.

Q1c. Where are the post-SQC good quality WOW rain gauges located?

WOW rain gauges generating good quality data were widely distributed around Britain (see 4.4.4). Of the 1,242 gauges with no observations removed by the SQC, there were 673 (24.87%) in urban areas. There were 1,274 (47%) WOW gauges with <10% observations removed by SQC in urban areas.

Q2. Is the method of SQC appropriate for WOW data?

The SQC removed 2.16 million erroneous observations from the WOW dataset, therefore it was effective in identifying errors. However, there is no assurance that the remaining observations are an accurate record of the actual rainfall and manual validation of the SQC highlighted instances where there appeared to be erroneous observations remaining (see section 4.4.5). Based on the analysis presented it appears that the SQC method is not optimised for WOW data, as there are several instances where errors have not been identified by the algorithm. It may be appropriate that the algorithm is conservative in nature, to avoid the unwarranted expulsion of observations from the time-series, provided the resultant datasets are subject to manual validation prior to application (see following chapter).

Improvements could be made, following the ontologies of SQC developed for Netatmo rain gauge observations in the Netherlands and Germany (de Vos *et al.*, 2019; Bárdossy *et al.*, 2021), whereby the SCQ algorithm is applied sequentially on successively filtered datasets. In the case of the SQC method presented in this chapter the SQC R1 – R4 could be implemented first, to filter PAWS with no rainfall for the year (R1), periods where accumulations of daily rain have been reported (R2), suspected instances of monthly accumulations (R3) and streaks of repeated values (R4). The next two rules to be applied could be removing values exceeding the hourly and daily maxima (R5 and R6). The implementation of rules R1 to R6 should remove much of the highly erroneous data and would allow for the application of the nearest neighbour checks of consistency of wet and dry hours and days between neighbouring stations (R7 – R10). If such changes were applied the SQC algorithm sensitivity to data gaps would require exploration, given they occur with a higher degree of frequency in PAWS data than in Official data sets, and may be an impediment to the determination of correlation. Caution is also required to ensure that nearest neighbour checks are not overzealous with the risk of excising observations where extreme rainfall is identified in a single gauge but not those neighbouring.

A further method for data validation is the comparison of observations against an independent dataset, in addition to the SQC applied here in this chapter. In the subsequent chapter such an approach is adopted, with comparisons between gauge observations and co-located radar. Given the extent and the size of the WOW database it was not practical to compare all gauge observations against the corresponding radar rainfall, therefore high intensity rainfall events that resulted in pluvial flooding were selected as test cases. As radar is not without accuracy issues (Ciach *et al.*, 1999; Steiner *et al.*, 1999; Trapero *et al.*, 2009; Villarini *et al.*, 2010; Peleg *et al.*, 2013) such an approach is not a complete failsafe but adding in the additional layer of data scrutiny allows for further validation of the absolute rainfall.

Q3. How could the quality of WOW observations be improved?

A case study of an alternative method of data gathering showed that where data were not shared automatically the observations perform better when subject to SQC (see section 4.5). Issues with the automated transfer of data have been highlighted in other research on weather observations, suggesting it is a systemic problem (de Vos *et al.*, 2019). There may also be an issue of 'out of sight, out of mind' where operators pay less attention to the observations being automatically uploaded to WOW, and therefore they are not aware of discrepancies in the dataset. Making the process more interactive to include an aspect of operator confirmation of observations may therefore be beneficial. The active promotion of data review could allow for errors to quickly be corrected, ghost stations to be removed, etc.

The data were provided by a specialist interest group whose members provide peer support including advice and scrutiny of results which leads to the prevention of errors occurring and means errors are addressed when identified. Creating a citizen science community has been demonstrated to be effective in improving the quality of data (Geoghegan *et al.*, 2016; Lin *et al.*, 2016), equally, it has been shown that often the best quality of data, and the most reliable contributions originate from a small number of highly dedicated participants (Haklay, 2012), so it may be the case that 'less is more'.

Finding a way to flag those participants that keep a close eye on their contribution may assist in identifying the most reliable data in WOW. Additionally, or alternatively, addressing data quality issues in a timely fashion would improve data quality, as problems could be explored and rectified so they would not persist in the dataset. Direct engagement with contributors would help foster a sense of community which could also contribute to improvements.

4.6.2 Recommendations for Using WOW Data

As with any dataset, WOW data should be reviewed before use. Data from gauges reporting for less than a year should be treated with caution and undergo manual SQC, an example of which is implemented in the following chapter.

The WMO advocate for a quality assurance process e.g. routine calibration, for good reason (WMO, 2021), and it may be that something similar would be beneficial for WOW PAWS. The function exists to provide such information in the metadata, but it is not readily available with the observations.

4.6.3 Outputs

The data outputs of this section comprise:

I. Quality controlled hourly rainfall data from WOW rain gauges in Britain from 2011 – 2020.
II. Summary statistics of SQC performance for PAWS reporting to WOW in Britain from 2011 – 2020

Links to the datasets are provided in Appendix 1.

Chapter 5. Event Based Quality Control with Data Validation and Visualisation

5.1 Overview

This thesis has so far considered the availability, distribution, and quality of WOW rain gauge data. This chapter further explores data quality by focusing on known convective events that resulted in urban pluvial flooding to determine whether relevant good quality WOW data can be identified and used in the post-event analysis of rainfall. The benefits and limitations of incorporating crowdsourced rain gauge data in the determination of rainfall distribution are presented.

During the exploratory analysis, WOW and Official data generated following Statistical Quality Control (SQC) were blended with radar data to produce spatial rainfall fields. This highlighted that some gauges (both Official and WOW) were reporting unreliable observations during the selected events, despite having passed the SQC. Consequently, this chapter details additional manual SQC validation comprising gauge by gauge data checking and comparison with radar for objective validation, demonstrating that after careful selection it is possible to identify WOW gauge data that are as reliable as those from Official sources.

This approach is the inverse to the more frequent application of using gauge observations to validate radar (Wilson *et al.*, 1979b; Collier, 1989).

5.2 Introduction

The previous chapter went some way to determining the quality of the WOW rain gauge data, but questions remained around the efficacy of the applied SQC method and ultimately the resulting accuracy of the dataset. This chapter aims to continue the statistical quality control process by applying manual quality control and cross checking with an external data source (radar) to validate the quality of WOW rainfall observations. This is undertaken by considering the following research questions:

Q1. What manual checking can be used to improve the accuracy of WOW rain gauge data so it can be used in post-event analysis?

Q2. What impact does including WOW rainfall data have on the interpolation of rainfall during localised high intensity storm events?

The objectives of this chapter are to:

I. Determine whether WOW data are available for specific intense rainfall events that resulted in pluvial flooding.

II. Assess the reliability of those data and generate a WOW rain gauge observations dataset to be used in further analysis.

III. Assess the 'value add' of WOW data by interpolating using Official data and data from WOW gauges to determine whether the pattern of rain is different/more closely matches available radar.

IV. Provide insight into the ways crowdsourced rainfall data can contribute to measuring rainfall in extreme events and highlight the limitations.

5.3 Methods

5.3.1 Event selection methodology

It has already been stated that WOW data would most likely contribute to better delineation of the spatial pattern of high intensity rainfall in urban areas i.e., where a high density of data are required across a highly variable rain field. On that basis, known pluvial flood events were reviewed with the aim of selecting four events for study (to accommodate the differences unique to each event/available dataset e.g., the intensity of the rain and the distribution/number of active WOW and Official gauges). No attempt was made in this research to determine the overall meteorological characteristics of events.

Lead Local Flood Authorities (LLFAs) in England and Wales are responsible for developing, maintaining, and strategizing for local flood risk management in their areas and must maintain a register of flood risk assets (UK Government, 2010). LLFAs are responsible for pluvial flooding, or when a river not designated as a 'main river' floods. One of the LLFA responsibilities under the Flood and Water Management Act (2010) is to investigate significant flooding following the event, to help determine if action can be taken to mitigate future risks. The requirement is detailed in Section 19 of the Act; therefore, the reports are often referred to as "Section 19" reports. The reports have a relatively standardised format; however, content of pertinent sections on rainfall volumes and recommended actions can vary in detail and quality due to availability and accessibility of data. These reports were used to provide information about significant flood events.

5.3.2 Method for Acquiring Rainfall Data

The process for identifying relevant rainfall data for a selected event was as follows:

a. Determine duration of event and location of interest using information from Section-19 report, radar, and contemporary news reporting etc.
b. Identify gauges with event observations within spatial and temporal boundaries appropriate to the extent of the storm, as determined in part a.

Three rain gauge data sources were accessed, comprising the post-SQC WOW data (as detailed in Chapter 4), the post-SQC Environment Agency (EA) data (as detailed in Chapter 4). Met Office rain gauge data were obtained via the CEDA Archive (Met Office, 2006a). The Met Office dataset is quality controlled prior to publication, however, the raw and QC'd iterations of a given observation were contained within the dataset along with rain accumulations at intervals greater than one hour, meaning the correct (QC'd, hourly) data had to be selected. No additional SQC was applied to these data[10].

[10] Code to extract hourly and daily rainfall data from MIDAS is provided via links in Appendix 1.

Radar datasets at 5 minute temporal and 1 km spatial resolution were extracted from the CEDA archive (Met Office, 2003) and provided by Rain++, a company specialising in the use of rain gauge data and radar (http://www.rainplusplus.com/)[11]. Radar is corrected using mean bias field bias (MFB) prior to publication (*Ibid.*). Radar was resampled to hourly for consistency with rain gauge data and due to issues with comparability between the datasets at shorter intervals (Bruno *et al.*, 2014; Jewell *et al.*, 2015).

The latitude and longitude of WOW and Official gauges were transformed from a geographic coordinate system (EPSG:4346) to a projected coordinate system appropriate for the UK (EPSG:27700). Gauges and radar were selected within a known bounding box of the central point of each storm event, as defined in the event selection.

5.3.3 Method for Determining WOW Rainfall Data Accuracy

There were two stages to determining WOW rain gauge accuracy. In the first stage WOW data were manually SQC'd. This was necessary a preliminary iteration of interpolations and blending (using the automatically SQC'd data) included erroneous observations and multi-hour/day data gaps. It was supposed that the Official datasets, having been SQC'd and published, were accurate and could be treated as reference data, however that was found not to be the case during manual SQC, therefore Official gauge data were assessed during this stage along with the WOW gauge data. The process was a 'first pass' to remove gauges that were obviously not reporting accurate observations. In the second stage the selected rain gauge data were compared with co-located radar rainfall as an independent check of veracity.

Stages of Manual SQC

The method considered the full time-series of observations to determine whether the automated SQC was able to identify potentially erroneous results, then focused on a known flood event for more detailed event specific appraisal.

Stage 1 – Manual SQC

The manual SQC process comprised:

Event Timeseries Review

1. Are there any data missing in the 24 hours prior to flooding?

[11] Rain++ provided blending software and data – see next section

2. Is there potential for accumulations after any gaps in the record?
3. Is there a gap due to QC?
4. Are there event pattern anomalies (erratic/low/high)?
5. Are there extreme highs during the event that are anomalous with other gauges?

Full Timeseries

6. Are there anomalies in the extended gauge data (e. g. repeated values)?
7. Are there pre-SQC extremes (>100mm)?
8. Are there post-SQC extremes (>45mm)?

Annual Daily Timeseries – From start of year to event occurrence

9. Is the daily Pearson correlation high (>0.7), medium (0.5 – 0.7), low (0.3 – 0.5) or very low (<0.3)?

In step 9, as inter-gauge correlation during the summer months can be unreliable due to the dominance of spatially isolated convective rainfall, data were assessed from the start of the year in which the flood occurred. Correlation was reviewed between all gauges within the search area as an indicator of gauges that may be inaccurate. Gauges with >10% missing data were not included in the appraisal as the correlation was considered invalid.

Based on the outcome of stage 1, a subjective judgement was made whether to include each WOW and Official gauge data in further analysis. Where gaps were present in post-SQC observations (indicating observations may have been removed by the SQC) a full review was conducted, starting with the raw data prior to aggregation, to establish if the cause of the SQC failure was apparent. This method was designed to systematically assess the gauge data; however, the ultimate decision to include or exclude a gauge was made on a case-by-case basis and adopted a precautionary approach (default was that suspect gauges were removed).

Stage 2 – Comparison with Radar

In the second phase of manual SQC, the collocated radar rain depths were extracted for each hour at each rain gauge. The focus of the comparison was 2-fold; 1. to compare the relative data quality between the WOW and Official datasets, and 2. to highlight outliers for further checks.

There are several methods for the collocation of radar and rain gauge data, that mitigate drift and mismatching between gridded products (Goudenhoofdt *et al.*, 2009; Jewell *et al.*, 2015; Rabiei *et al.*, 2015). In this instance the simplest method was adopted where the difference between the radar and gauge rain depth was calculated by subtracting the gauge rain observation from the radar rain observation for the corresponding/overlying 1 km grid square. For this analysis, instances of zero rain at the gauge and radar were removed.

Unlike traditional radar/rain gauge data validation, the pairing of collocated data in this research was used in a novel way to establish whether there was a difference between the paired datasets, i.e., the radar/Official gauge observations and the radar/WOW gauge observations, and not to establish which value most closely represented the true rainfall depth. The paired data were thus subject to a series of review measures:

- Strip plots showing rain gauge – radar rainfall between WOW and Official datasets to consider the pattern of distribution between the two cohorts.
- Box plots of the RMSE for both WOW and Official data.
- Review of Pearson correlation for both WOW and Official gauges
- Paired sample T-test
- Event time series plots including neighbouring gauges.

It was assumed that if deviations between gauge and radar were comparable to the gauge data datasets (Official and WOW) that WOW data were sufficiently reliable for use. This assumption allows for deviation between gauge and radar observations, as it is known that radar can be inaccurate despite having been SQC'd (Sauvageot, 1994) . On completion of these checks, and if there was no statistical difference between the cohorts it could be concluded that where Official data are considered good enough for use then WOW data should be too.

5.3.4 Methods for Creating Interpolations and Assessing 'Value Add' of WOW Rain Gauge Data

There are numerous methods to interpolate between rain gauges and to blend data with radar (see Literature Review, section 2.2.2). An automated tool for the creation of rain gauge interpolations and for blending with radar was provided for use in this research by Rain++. In this research block kriging (BK) and kriging with external drift (BKD) were used for interpolating between rain gauges and blending gauge data with radar, as they have been demonstrated to be the most effective methods available (Wang *et al.*, 2013; Jewell *et al.*, 2015; Ochoa-Rodriguez *et al.*, 2019a).

Rain gauge observations and radar for each selected pluvial flooding event were used to create interpolations and blended fields comprising:

- Block kriging of Official gauge data
- Block kriging of WOW gauge data
- Block kriging of Official and WOW gauge data
- Kriging with external drift with official gauge data and radar
- Kriging with external drift with official and WOW gauge data and radar

Radar was used for comparison with the created interpolations[12].

5.4 Results

5.4.1 Events Selection

A review was conducted of twenty known pluvial flood events that occurred prior to August 2018 (limit of available SQC'd Official rain gauge data) obtained from the EA Surface Water update to the Flood Resilience Review. An initial sift was made to discard those with limited WOW gauges in the vicinity of the flooding, (common in the early years of WOW - see Figure 3.5 in Chapter 3). Events were excluded where/when there were limited WOW gauges, e.g. Coverack (18/07/2017), which was also characterised as a rapid response fluvial flood (Flack *et al.*, 2019). Given the intermittent nature of some WOW gauges it was not until detailed gauge by gauge review was undertaken that the availability of WOW data was confirmed.

Table 5.1 presents the locations and the dates of the four selected flooding events, with the duration and spatial extent of the search area used to select WOW and Official gauges for each

---

[12] A blend of radar and WOW (without the official rain gauge data) was not made, as such an interpolation would not be generated where Official data were available.

event. The spatial extent of the study area and the duration of the analysis period was based on a subjective review of contemporaneous reporting (e.g., media, Section 19 reports, and weather bulletins), as described in section 5.4.2.

Name	Event Date	Timeseries Start	Timeseries End	Flood X (easting)	Flood Y (northing)	Search Extent (km)
Canvey Island	20/07/2014	10/07/2014	21/07/2014	578008	184090	30
Bar Hill	08/08/2014	04/08/2014	11/08/2014	538000	263000	40
Birmingham	16/06/2016	06/06/2016	17/06/2016	405843	282613	30
Milton Keynes	27/05/2018	21/05/2018	28/05/2018	485550	239024	30

Table 5.1 Case study pluvial flood events selected for analysis.

A satellite image showing the locations of the case study events can be seen in Figure 5.1.

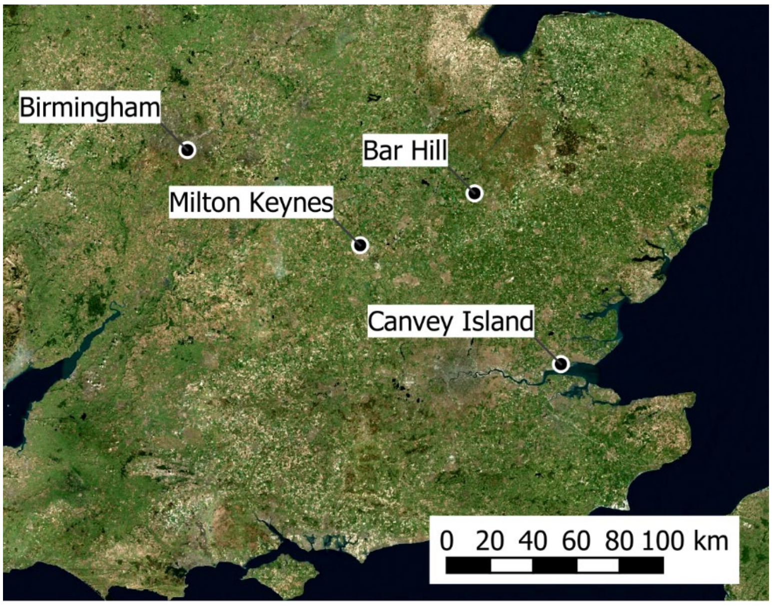


Figure 5.1 Satellite image showing the locations of the four case study event locations used in post-event reconstruction.

5.4.2 Determining Event Temporal and Spatial Extent and Data Availability

Event 1 – Canvey Island, 20/07/2014

Canvey Island is in Essex in southeast England. The island is approximately 3.5 km wide and 7.5 km long (see Figure 5.2). The Thames is south of the island, Holehaven Creek is southwest, East Haven Creek is northwest and Hadleigh Ray is north and east. These watercourses mean Canvey Island is a true island. Land has been reclaimed and lies below mean sea level, so pumps are used for drainage, particularly during the high tide (Essex County Council, 2014).

Figure 5.2 Satellite image of Canvey Island affected by flooding in July 2014.

The post-event analysis conducted on behalf of the LLFA states that flooding on Canvey Island was first reported at 17:45 on 20th July, 2014 (*Ibid*.). Heavy rainfall had been forecast and several teams were on standby to respond. The Met Office daily weather summary noted convective storms active across England on the day of the event, with the highest total daily rainfall being recorded in Essex (see daily weather summaries in Appendix 3) (Met Office, 2021a). The Section 19 post-event analysis report states that heavy, intense rain caused pluvial flooding and overwhelmed the drainage system with the magnitude of the rain beyond the design capacity of the surface water drainage system (Essex County Council, 2014).

The nearest Official rain gauge to Canvey Island active at the time of flooding was approximately 7 km to the northeast of the island (ID: EARAYLER). The daily data from the gauge were accessed to determine antecedent rainfall conditions and to help to determine the temporal extent of the event (see Figure 5.3)(Met Office, 2006b).

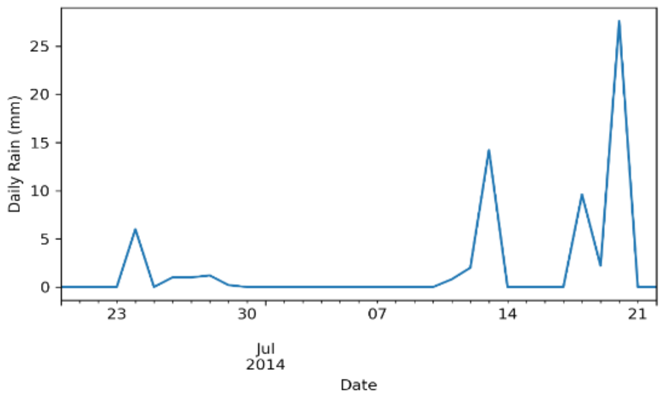


Figure 5.3 Daily rain from Official rain gauge (ID: EARAYLER) close to Canvey Island for the period prior to flooding in July 2014.

The plot of daily rainfall shows that rainfall was recorded on multiple days prior to the flooding on the 20th of July therefore WOW data were sought from the 10th of July to 21st July. The island is relatively small so to obtain sufficient gauge data a search polygon extending 30km from the approximate mid-point of flooding was applied.

Event 2 – Bar Hill, Cambridge, 08/08/2014

Bar Hill is a village to the north of Cambridge that was purpose built in the 1960's (Cambridge Science Park, 2020). According to local press reports the area has a history of flooding, with floods occurring in the 1960's and 1970's (Elliott, 2017). It is a distinct urbanised area covering no more than 2 km^2 to the northwest of Cambridge (see Figure 5.4).

Figure 5.4 Satellite image of Bar Hill most affected by flooding in August 2014.

Contemporary reporting of the 2014 flooding states that Cambridge Fire and Rescue began receiving calls for assistance from 12:45pm on the 08/08/2014 (BBC, 2014). Flooding was isolated but extensive, with properties being affected in several small towns and villages to the north of Cambridge and into Norfolk. The Section 19 report for the event is brief (5 pages) providing minimal rainfall analysis (Cambridgeshire County Council, 2014). The nearest official rain gauge to Bar Hill active at the time of flooding was EA179635, approximately 12Km to the northwest. The daily data for the gauge were reviewed to determine antecedent rainfall conditions and to help to determine the temporal extent of the event (see Figure 5.5) (Met Office, 2006b).

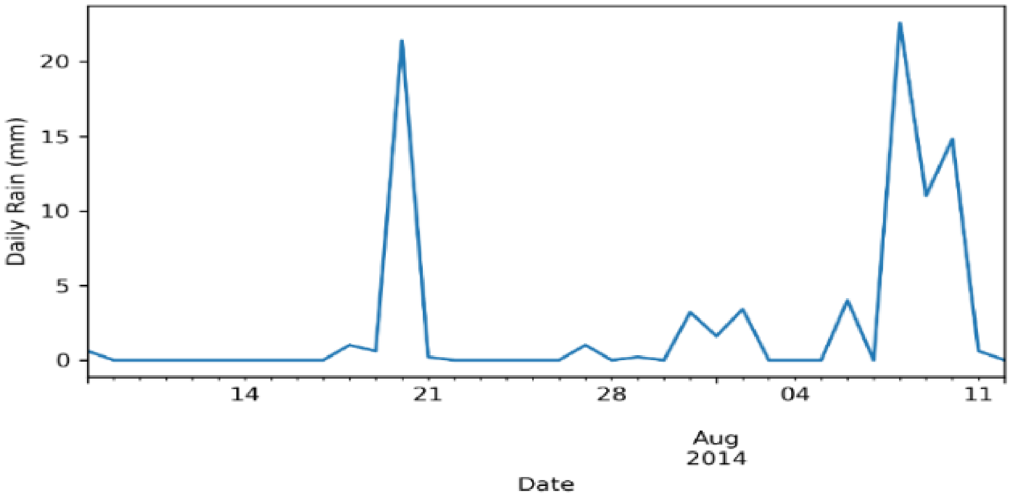


Figure 5.5 Daily rainfall from Official rain gauge (ID: EA179635) close to Bar Hill for period prior to flooding in August 2014.

The plot of daily rainfall shows there was only light rainfall within the week prior to the rainfall that caused flooding, following on from some rainy days earlier in July. WOW data were sought from to cover the period up to and beyond the flood occurring (4th August to 11th August). The flooding in Bar Hill was localised however post event reporting indicated a multi-cell convective system causing widespread rainfall (BBC, 2014; Cambridgeshire County Council, 2014). A polygon of 40 km was applied as a search area for WOW gauges acknowledging the spatial extent of the storm.

Event 3 – Birmingham, 16/06/2016

Birmingham is a large city in the West Midlands. It is a heavily modified urban area. There were a series of rainfall events resulting in multiple occurrences of pluvial flooding over several days. According to the Section 19 report, rainfall between the 8th and 10th June affected the north of the city, with rainfall on the 16th of June having more impact in the south (see Figure 5.6) (WSP, 2017). The rainfall on the 16th of June was of lower intensity and longer duration than the storms that caused flooding earlier in the month. News reports from the day state that West Midland Fire Service began to receive requests for assistance shortly after 18:00 on the 16th June (BBC, 2016). In total 435 incidents of flooding were reported following the storms (WSP, 2017).

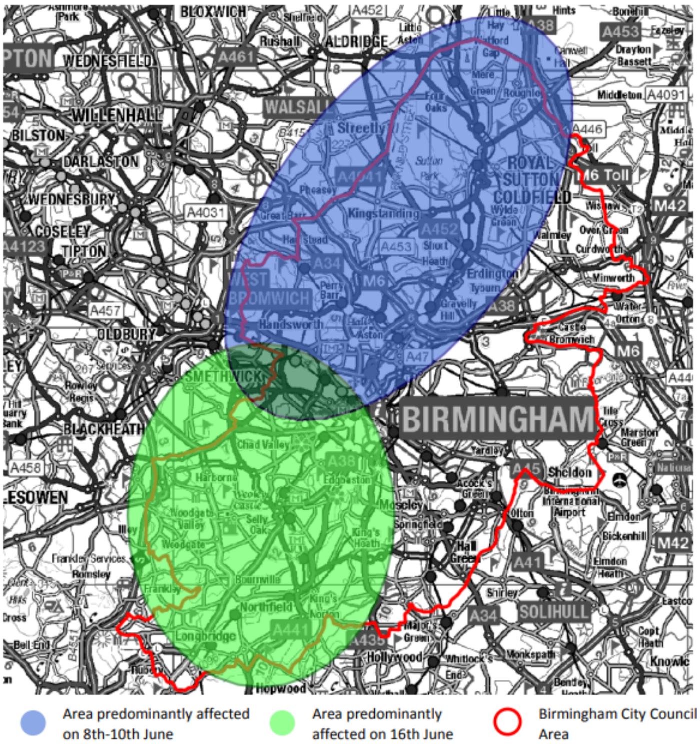
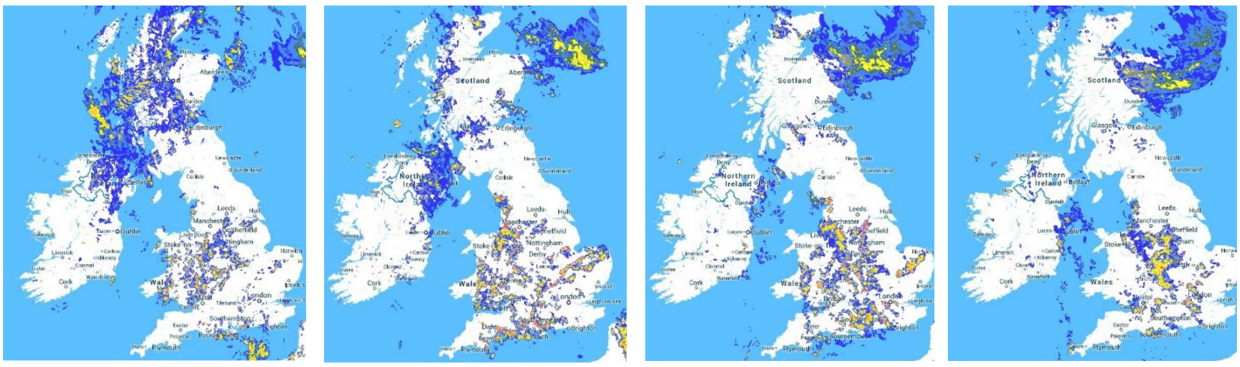


Figure 5.6 Map showing area of Birmingham affected by flooding, after Section 19 report, WSP, 2017.

Archived radar imaging at a 30 minute resolution for the UK indicates intermittent but persistent rainfall across the Midlands throughout the day on the 16th June, 2016 (see Figure 5.7). The daily weather summary issued by the Met Office states that there were scattered, slow-moving heavy showers and thunder storms across England and Wales (see Appendix 3)(Met Office, 2021a).



11:30 16:00 19:00 21:30

Figure 5.7 Maps of archived radar data for the UK on 16/06/2016, showing the evolution of the storm (source: MetCheck).

The nearest Official rain gauge to the flooding in Birmingham that was active at the time of flooding was ID: EA3176_RK, approximately 4 km to the southwest of the central point of flooding, and within an affected area. Daily data were accessed to determine the antecedent rainfall conditions and to help to determine the temporal extent of the event (see Figure 5.8)(Met Office, 2006b).

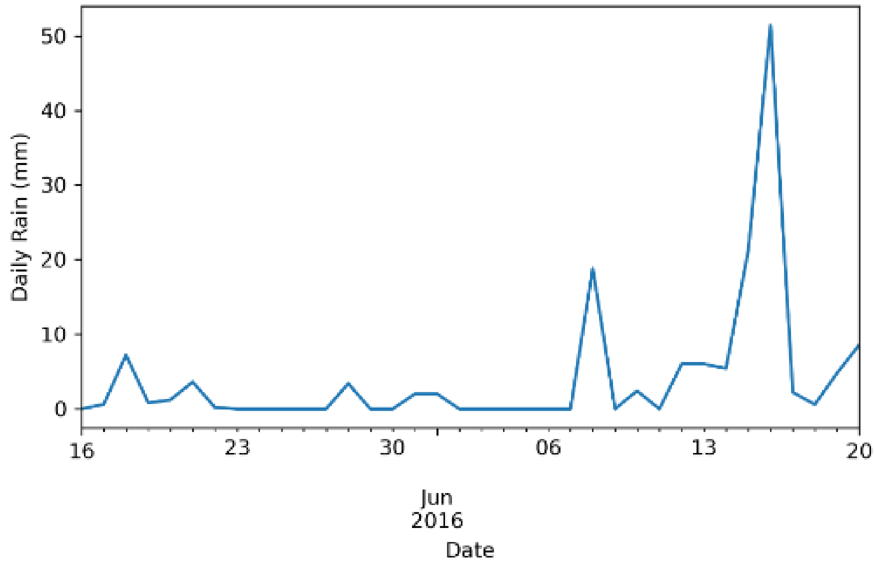


Figure 5.8 Daily rain from Official rain gauge (ID: EA3176_RK) closest to Birmingham for the period prior to flooding in June 2016.

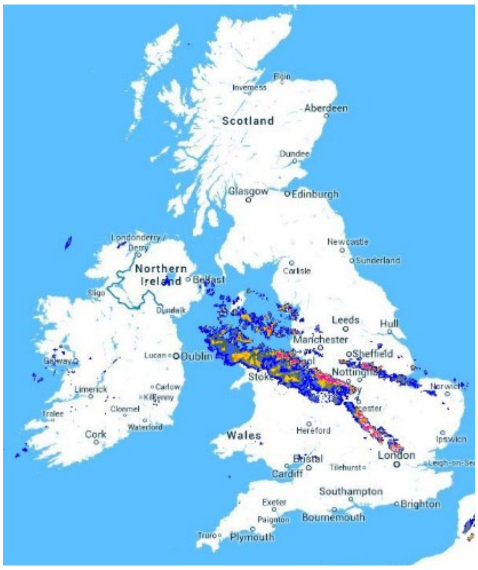
The plot of daily rainfall shows the rainfall that precipitated flooding earlier in May, in addition to the storms around the 16th when the most extensive flooding occurred. WOW data were sought from the 6th June to 17th June. The density of WOW gauges active at the time of flooding was relatively high, with several Official gauges also active, therefore a polygon of 30km was applied as a search area for WOW gauges.

Event 4 – Milton Keynes, 27/05/2018

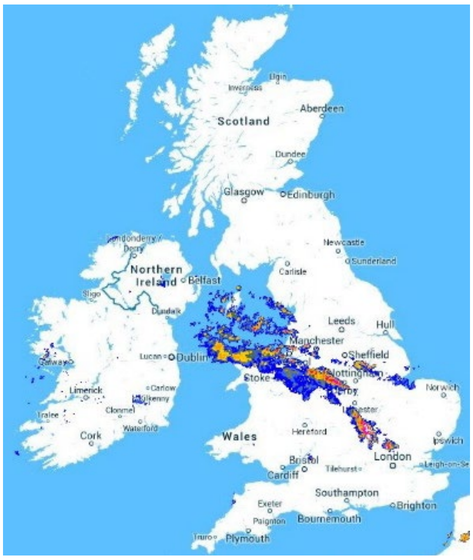
Milton Keynes is in Buckinghamshire, around 100 km north of London. It is a 'new town' purpose built in the 1960's (see Figure 5.9). Flooding occurred on 27/5/2018 at multiple locations identified broadly as Milton Keynes. Initial flood investigation reports were synthesised into an overarching report that addresses all the flooding under the LLFRs purview (AECOM, 2019; WSP, 2019). The Section 19 report indicated there was above average rainfall in May 2018, resulting in antecedent conditions conducive to flooding. Reporting from a local council meeting suggested have were several flood events in recent years, and actions to reduce the risk of flooding were insufficient (D Tooley, 2020).

Figure 5.9 Satellite image of the area most affected by flooding in Milton Keynes on 25/05/2018.

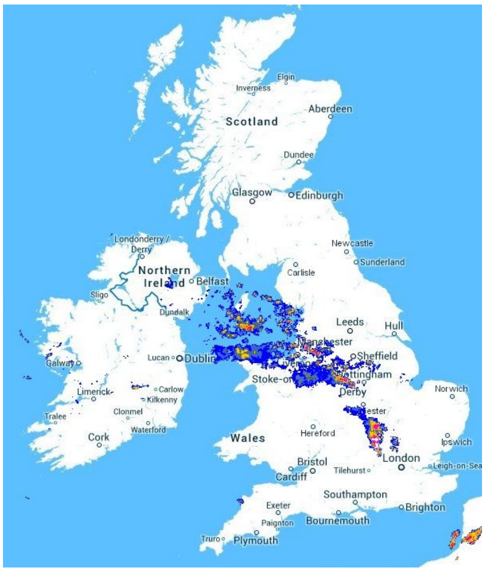
Archived radar imaging at a 30 minute resolution for the UK pictured an extensive multi-cell convective system trending NW to SE across central England on the 27/5/18 (see Figure 5.10). The daily weather summary issued by the Met Office confirms the presence of bands of heavy rain and thunderstorms crossing the UK on the day (see Appendix 3).



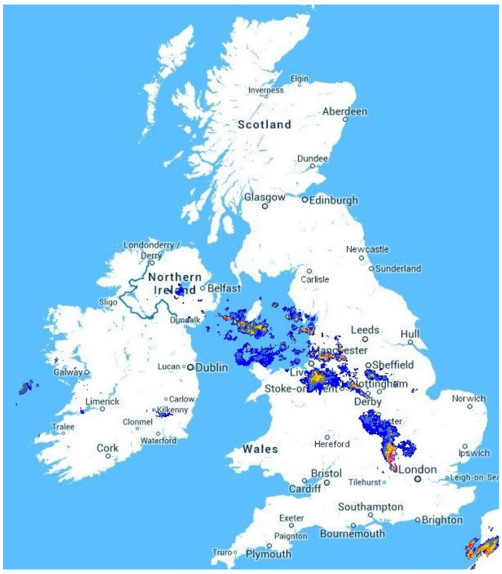
20:00



20:30



21:00



21:30

Figure 5.10 Archived UK radar images for the 27/05/18 (source: MetCheck).

There were several Official rain gauges in the vicinity of flooding in Milton Keynes. Daily rainfall data were sought from ID: EA170823, approximately 16km to the northwest of Milton Keynes to determine antecedent rainfall and to help to determine the temporal extent of the event (see Figure 5.11)(Met Office, 2006b).

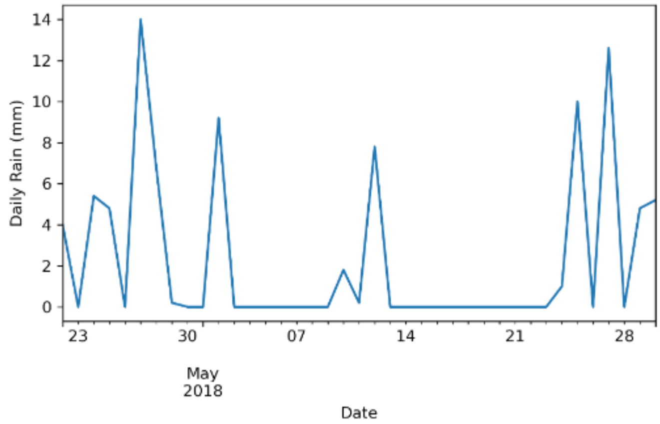


Figure 5.11 Daily rainfall from Official rain gauge (ID: EA170823) closest to Milton Keynes for the period prior to flooding in July 2018.

The plot of daily rainfall shows the gauge captured rainfall in late April exceeding the rainfall that precipitated flooding on the 27th of May. There was then a dry period of over a week during May therefore the data selected was limited to a relatively short duration, spanning from the 21st of May to 28th of May. The density of WOW gauges active at the time of flooding appeared to be relatively high, with several Official gauges also active; therefore, WOW gauges within a polygon extending 30 km from the approximate location of flooding in Milton Keynes were sought.

All Events

The number of gauges available from the WOW, EA and MIDAS datasets are shown in Table 5.2 for the four case-study events. Based on the archived radar dataset examination and background information, the search area was a 30 km polygon from the flood impacted location for Canvey Island, Birmingham, and Milton Keynes and 40 km around Bar Hill.

Event	WOW Gauges	EA Gauges	MIDAS Gauges
Canvey	40	32	3
Bar Hill	39	25	6
Birmingham	29	20	5
Milton Keynes	34	20	5

Table 5.2 Count of rain gauges in WOW, EA, and MIDAS within the event search areas.

5.4.3 Manual Quality Control Results

WOW data were subject to the tests outlined in section 5.3.3[13]. Tables with the manual SQC results are presented in Appendix 3 (including the original WOW to simplified ID to allow for independent data validation/use). On completion of the first stage of manual SQC data the number of gauges rejected for each event is presented in Table 5.3. The most common reason for rejection was missing data during the event.

Event	Starting Gauges (Count)	No data during event (Count)	Data gaps during event (Count)	Gauges with Poor Data (Count)	Rejected (%)
Canvey	40	4	4	14	35
Bar Hill	39	6	8	12	31
Birmingham	29	4	4	12	41
Milton Keynes	34	6	3	11	32

Table 5.3 Number and percent of WOW gauges rejected by manual SQC Stage 1 per event.

The errors identified by the manual validation process varied and included elements that could not realistically be identified by the automatic SQC developed by Lewis *et al.* (2021), and those that perhaps could be if the algorithm was modified. Based on a visual assessment there were 13 instances where the manual review suggested that the automated SCQ may have failed to detect an issue in the data; however, not all the erroneous observations occurred during the event period and therefore the data could be used. Examples of the errors identified by the manual SQC are presented here.

Example 1 – QC check on event time series: the timeseries for gauges rejected from the Milton Keynes event are shown in Figure 5.12. There were 6 gauges with no data during the event (W22, W24, W25, W31, W33, W34), one gauge (W26) where the depth of rain was inconsistently low for the duration of the event, three gauges (W21, W23, W28) had large gaps in the data suggesting reporting was erratic, and, there was one gauge (W9) that may have been accurate at the time of the event; however, there was poor correlation with neighbouring gauges and there were erratic highs in the full time series.

[13] The original WOW station IDs have been substituted with simplified IDs for ease of labelling. The original gauge IDs can be seen in the summary tables in Appendix 3.

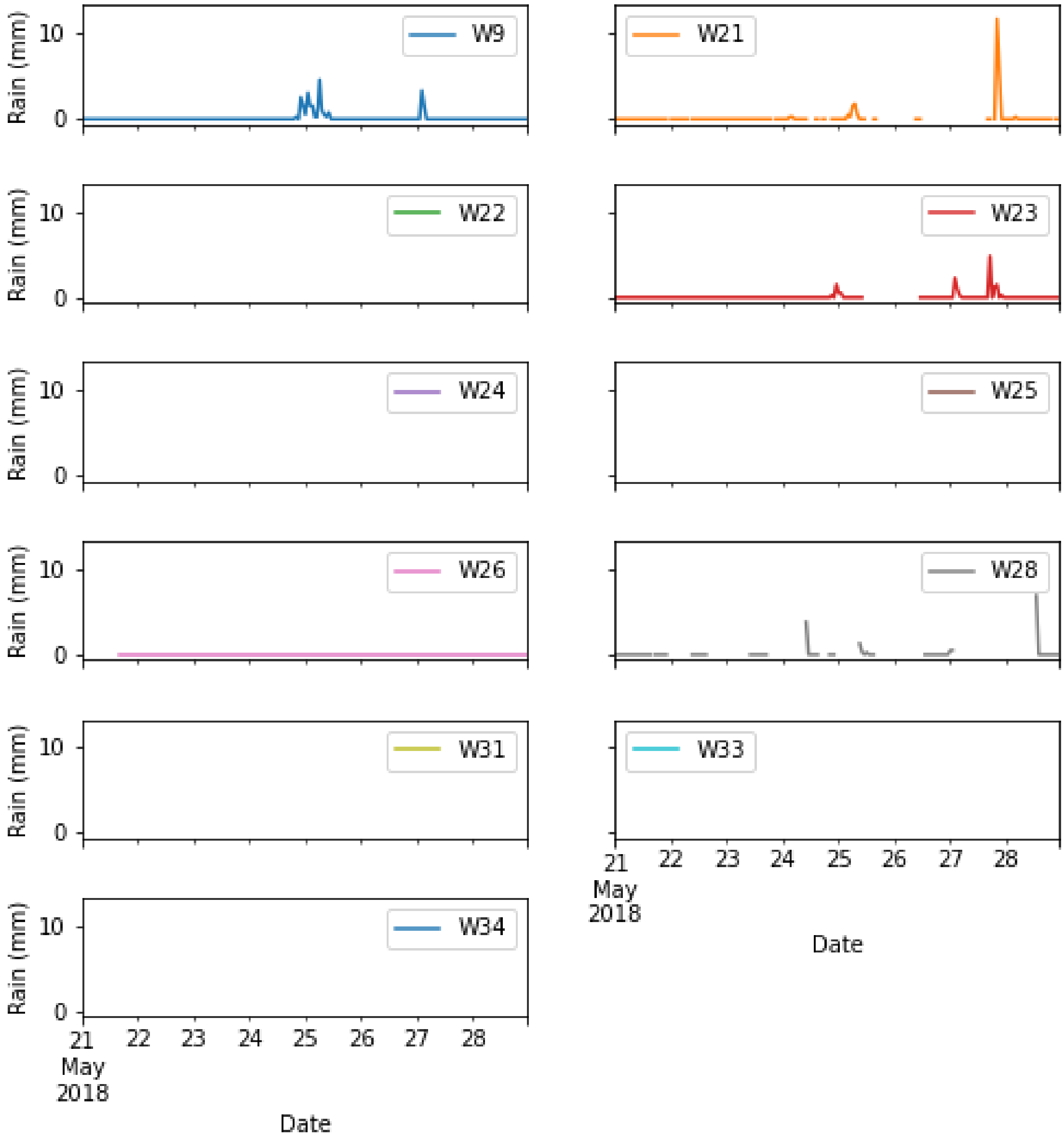


Figure 5.12   Milton Keynes WOW rain gauges rejected following manual first pass checks.

Example 2 – Check on full timeseries: An example of issues with the timeseries outside the event, and potentially issues with the automated SQC, can be seen during the Canvey event. There was one gauge (W10) where the automated SQC identified erroneously high rain depths, but also failed to identify erroneous data below the hourly threshold of 92mm/hr (see Figure 5.13).

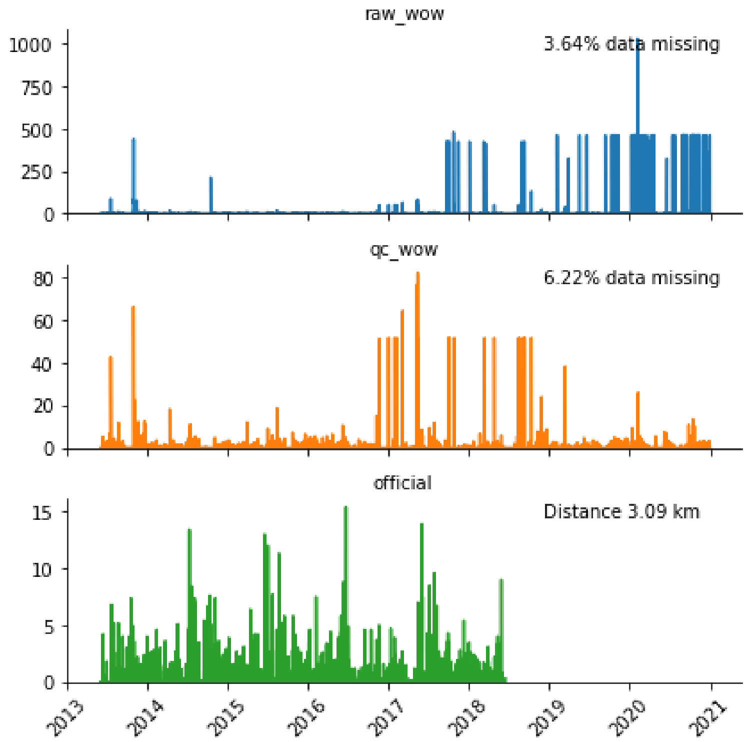


Figure 5.13 WOW gauge from Canvey event with poor SQC performance.

As the erroneous results did not regularly occur until after the event in question (June 2014), and the observations from the event period were consistent with other gauges, the gauge data was deemed to be fit for purpose and was included in subsequent analysis. This is an example of the judgement required when determining the suitability of gauge data for inclusion in analysis.

Example 3 – Daily correlation: Matrices created from the Pearson correlation between all gauges for a given event are presented in Appendix 3, and an example from Bar Hill is presented in Figure 5.14. As the gauge data originate from a 40 x 40 km area it is expected that data from the most distant gauges may not correlate well; however, the matrix makes it immediately obvious where a gauge correlates poorly with all the others in the search area.

Correlation alone was not used to reject a gauge but was found to be a good visual indicator of unreliable gauge data and provided a prompt to look more closely at the full gauge data timeseries and event specific rainfall observations. In this example gauges W1 and W19 were found to exhibit a lower correlation than others within the search area. The timeseries from gauge W1 (AKA: 172761) suggested that there was some underreporting at the start of 2014, which was resolved by the summer. W19 showed poor correlation throughout 2014 and there were large gaps in the event data, rendering that gauge unfit for inclusion in further analysis. Matrices for the remaining events are provided in Appendix 3.

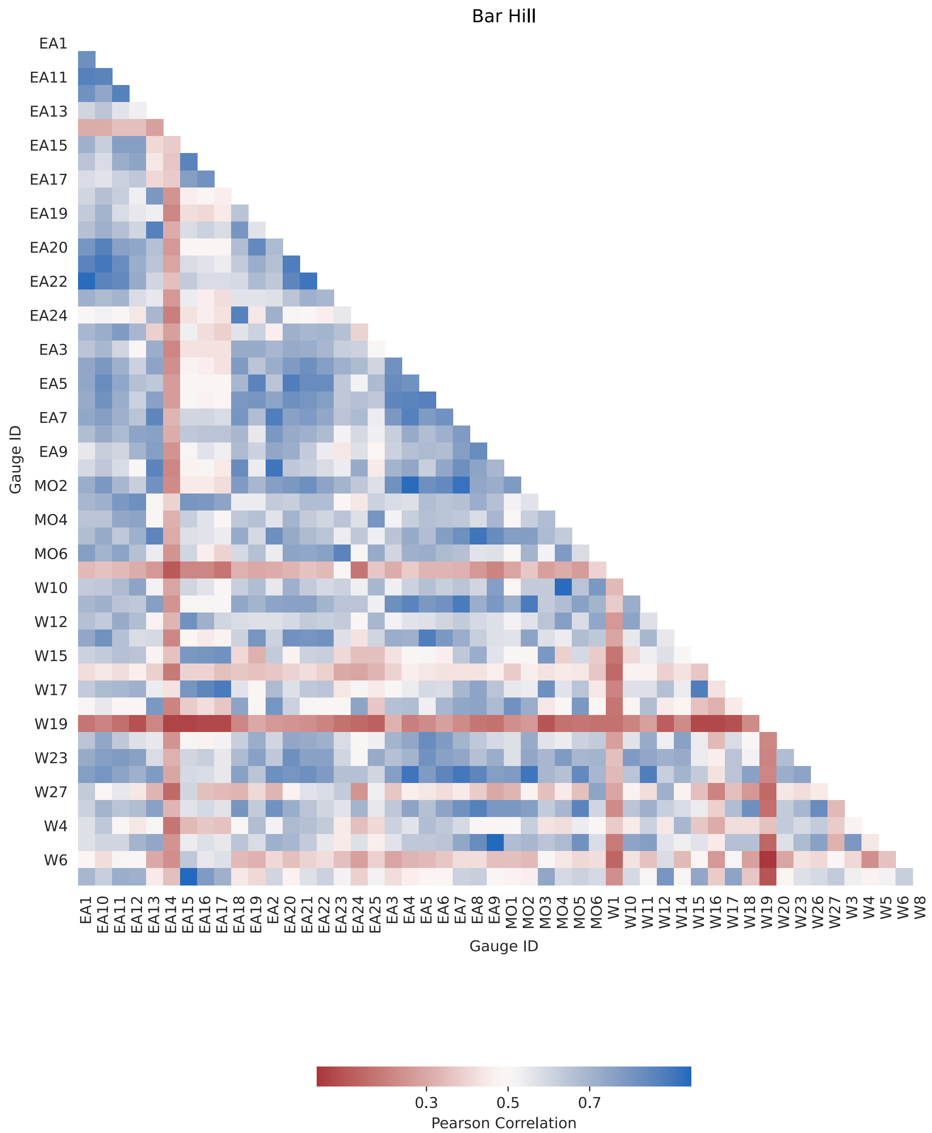


Figure 5.14 Pearson correlation matrix for all gauges within 40 km of Bar Hill with <10% data missing from start of 2014 to event date.

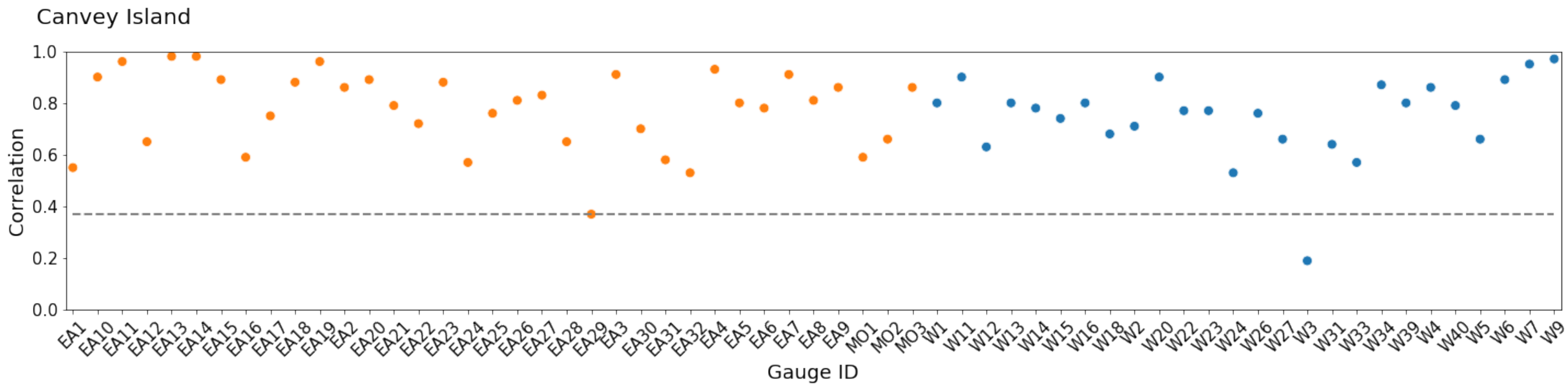
The correlation of data from Official gauge EA14 can be seen to be low in Figure 5.14, which is explored further in the next section (see Figure 5.17).

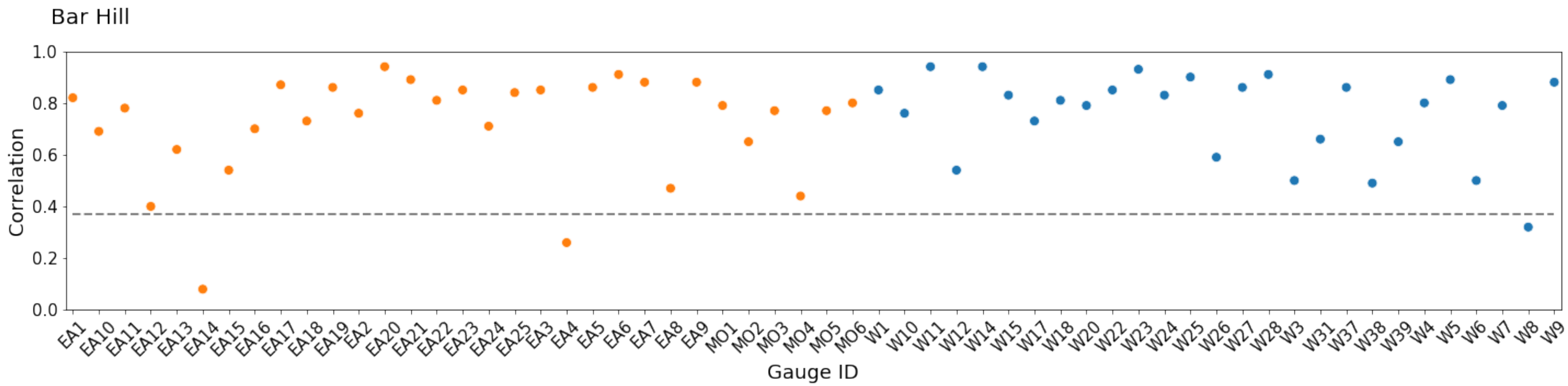
5.4.4 WOW Data Quality Versus Official Data Quality.

The second stage of manual SQC was to compare rain gauge observations against co-located gridded radar. Data were extracted from the radar square corresponding with each gauge for the duration of each event (as detailed in Table 5.1). The correlation between co-located radar and gauge data was assessed to determine any difference between WOW and Official rain gauge data, as compared to radar rainfall (and to potentially highlight any remaining dubious observations or gauges).

The absolute radar accuracy, even post MFB correction, can be questionable particularly during convective events such as those being studied, where there may be errors due to bright band effects, drift and variation within the 1 km grid square (Peleg *et al.*, 2013) Equally rain gauge observations exhibit errors that may be particularly acute during convective events, including wind-induced undercatch, tipping errors and inability to accurately measure precipitation in the form of hail (Habib *et al.*, 2001; Villarini *et al.*, 2008; Pollock *et al.*, 2018). Given the potential for errors in both radar and rain gauge data this aspect of the analysis does not attempt to establish the 'true' rainfall, but rather is used comparatively to establish if the extent of variation between the paired gauge and radar data is equal.

A test of variation between the datasets is presented in scatter plots showing the Pearson correlation between hourly data for the duration of the rainfall events that resulted in pluvial flooding (see Figure 5.15).


Canvey Island
Correlation
1.0
0.8
0.6
0.4
0.2
0.0
EA1 EA10 EA11 EA12 EA13 EA14 EA15 EA16 EA17 EA18 EA19 EA2 EA20 EA21 EA22 EA23 EA24 EA25 EA26 EA27 EA28 EA29 EA3 EA30 EA31 EA32 EA4 EA5 EA6 EA7 EA8 EA9 MO1 MO2 MO3 W1 W11 W12 W13 W14 W15 W16 W18 W2 W20 W22 W23 W24 W26 W27 W3 W31 W33 W34 W39 W4 W40 W5 W6 W7 W9
Gauge ID


Bar Hill
Correlation
1.0
0.8
0.6
0.4
0.2
0.0
EA1 EA10 EA11 EA12 EA13 EA14 EA15 EA16 EA17 EA18 EA19 EA2 EA20 EA21 EA22 EA23 EA24 EA25 EA3 EA4 EA5 EA6 EA7 EA8 EA9 MO1 MO2 MO3 MO4 MO5 MO6 W1 W10 W11 W12 W14 W15 W17 W18 W20 W22 W23 W24 W25 W26 W27 W28 W3 W31 W37 W38 W39 W4 W5 W6 W7 W8 W9
Gauge ID

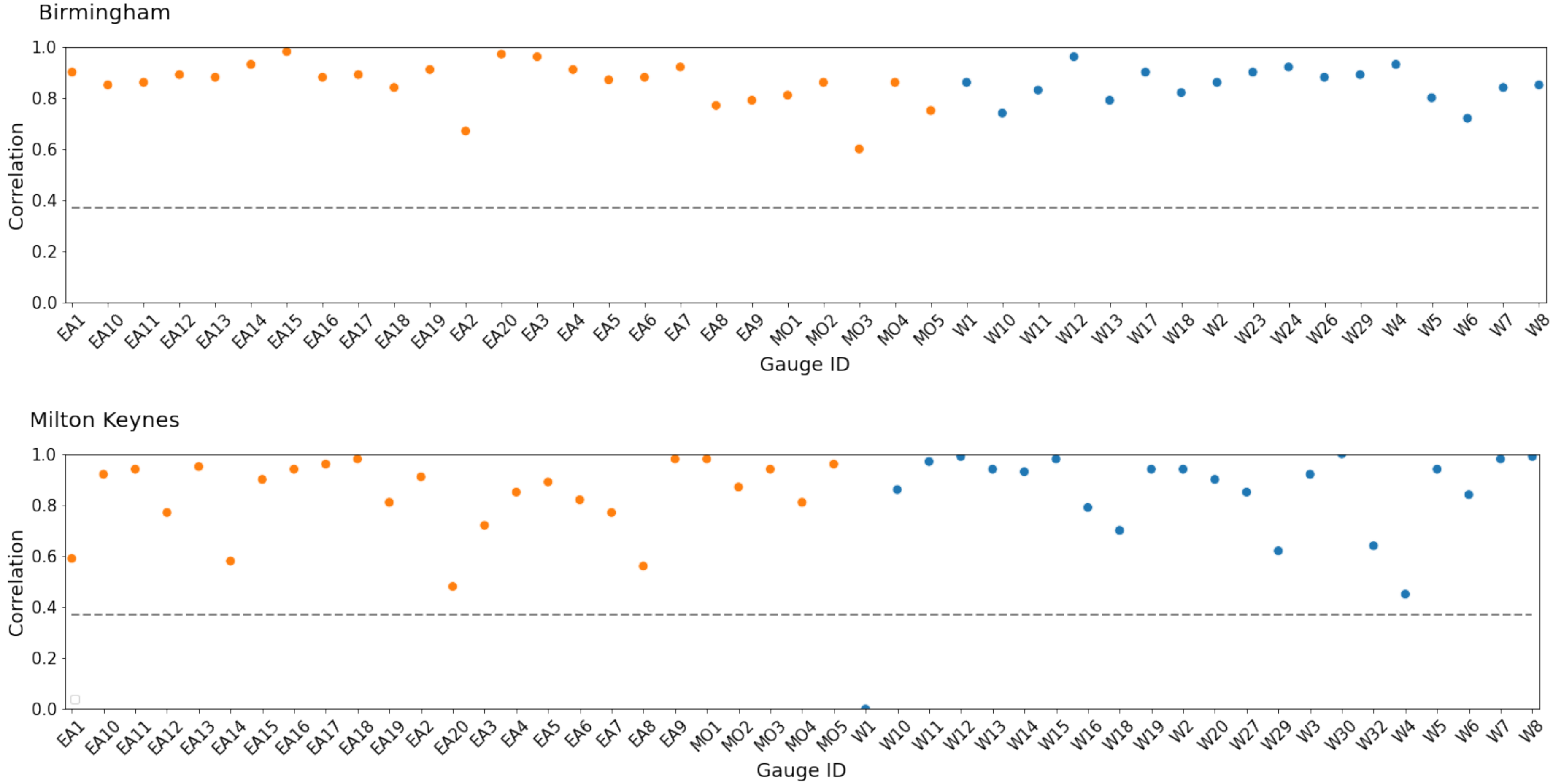


Figure 5.15 Pearson Correlation between hourly data for co-located radar and rain gauge data, with '---' indicating level below which correlation poor (0.368).

Research conducted by Ochoa-Rodriguez (2017) established that a Pearson correlation <0.368 was indicative of an error, therefore gauge observations were scrutinised to determine if the error could be identified . The potentially suspect gauges are detailed in Table 5.4.

Event	Gauge ID (Pearson Correlation)
Canvey	W3 (0.19)
Bar Hill	EA14 (0.08), EA4 (0.26), W8 (0.32)
Birmingham	EA7 (0.21)
Milton Keynes	W1 (0.0)

Table 5.4 Gauges with co-located radar and rain gauge observations >0.368.

For the gauges in Table 5.4, the timeseries was plotted with the radar against the rain gauge to determine if a persistent error could be identified, or whether certain hours were lowering the correlation. Additionally in some instances the timeseries of the closest gauges were plotted to further determine where the error may lie (gauge or radar). Discrepancies on the day of flooding were carefully considered (see Table 5.1 for dates), whereas errors occurring not occurring on the critical day were less of a concern (if isolated).

Canvey Island

The timeseries at W3 appeared to be advanced by 1 hour as compared to the radar and to the nearest gauge (W5) that was located at the same address (see Figure 5.16).

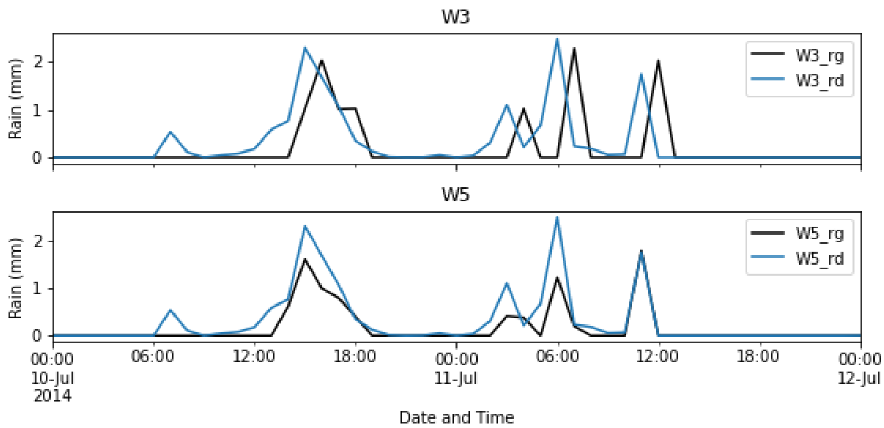


Figure 5.16 Line plot of selected Canvey Island event timeseries for WOW gauges W3 and W5 demonstrating discrepancy between radar and rain gauge (suffix 'rd' = radar, 'rg' = rain gauge)

There was a possibility data were being recorded in BST rather than the standard of GMT. The metadata for W3 available on WOW stated GMT was being used, with further information on the weather station available at the operator's website; https://cm2weather.co.uk/ . Despite the metadata there was a clear discrepancy therefore the gauge was removed from further analysis. Given the proximity to W5 there was no benefit in retaining potentially incorrect data, had that not been the case it may have been appropriate to amend by shifting the observations by 1 hour (if this was consistent across the event period), allowing the data to be used.

Bar Hill

There were three gauges in the Bar Hill event with poor correlation. The first (EA14) appeared to be reporting anomalous rain depths as compared to the radar and the nearest official gauge (MO4 at 10Km). Figure 5.17 shows that there were erroneous peaks reported by EA14 on the 7th and 11th of August that were not seen in the radar or MO4.

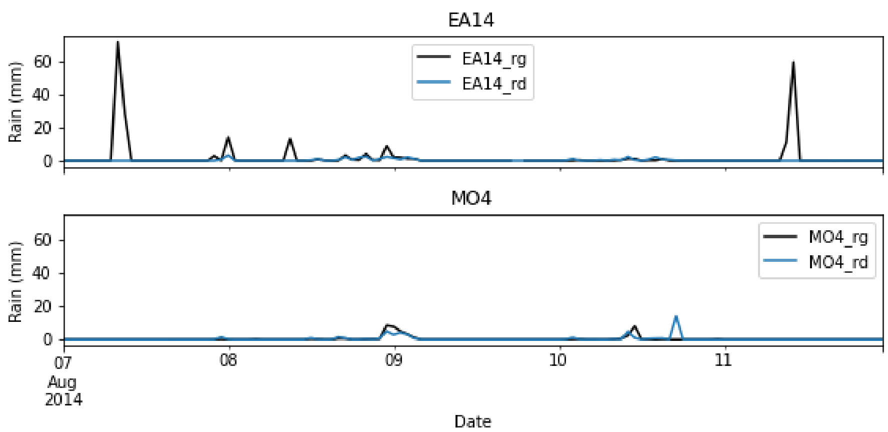


Figure 5.17 Bar Hill selected event time series for Official gauges EA14 and MO4 demonstrating discrepancy between radar and rain gauge (suffix 'rd' = radar, 'rg' = rain gauge)

Although the Pearson correlation displayed in Figure 5.14 was designed to visualise errors in the WOW gauge data, it showed that EA14 was inconsistent with other rain gauges in the locality. The mean Pearson correlation for E14 was 0.29. Observations from EA14 were subject to review by Villalobos-Herrera *et al.* (2022) as part of validation for sub-hourly SQC as the rain depth was flagged as extreme (>30mm/hour). The review concluded that data from EA14 were unreliable. Based on all the available evidence EA14 was removed from further analysis.

Birmingham

There were no gauges in the Birmingham area where the correlation was <0.368.

Milton Keynes

There was one gauge in the Milton Keynes area that had a Pearson correlation of 0. The timeseries suggests that W1 was underreporting rain (see Figure 5.18), having not recorded rain between the 24th and 28th May that was detected by co-located radar and at the closest gauge (W16, 3Km from W1).

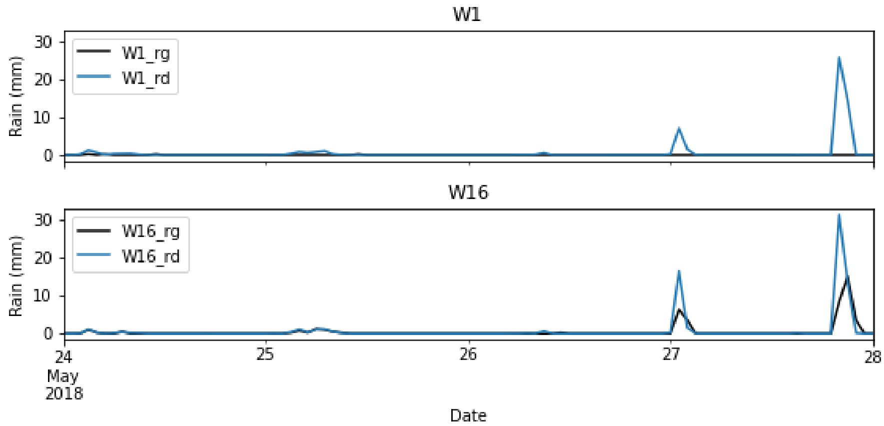
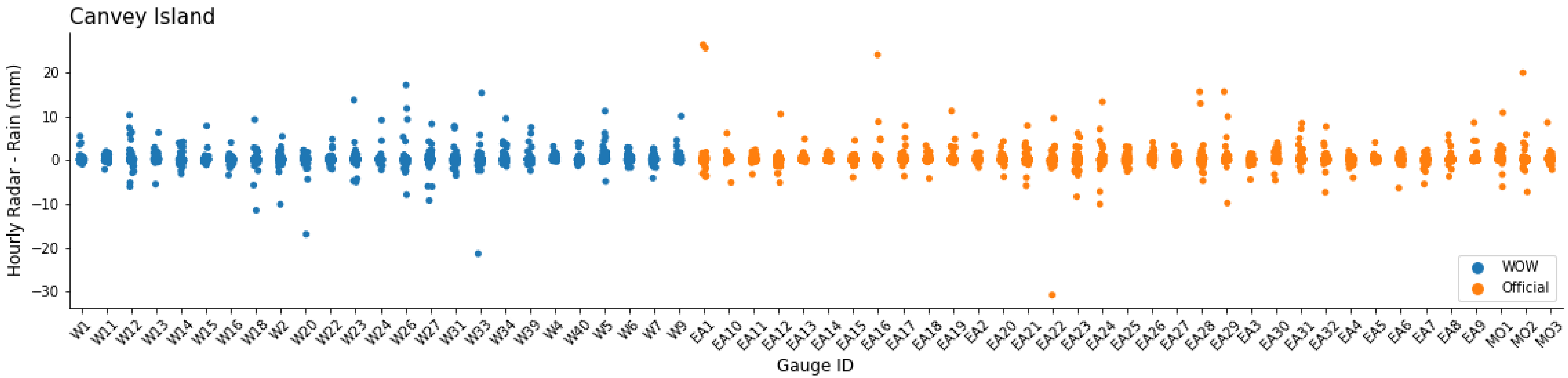
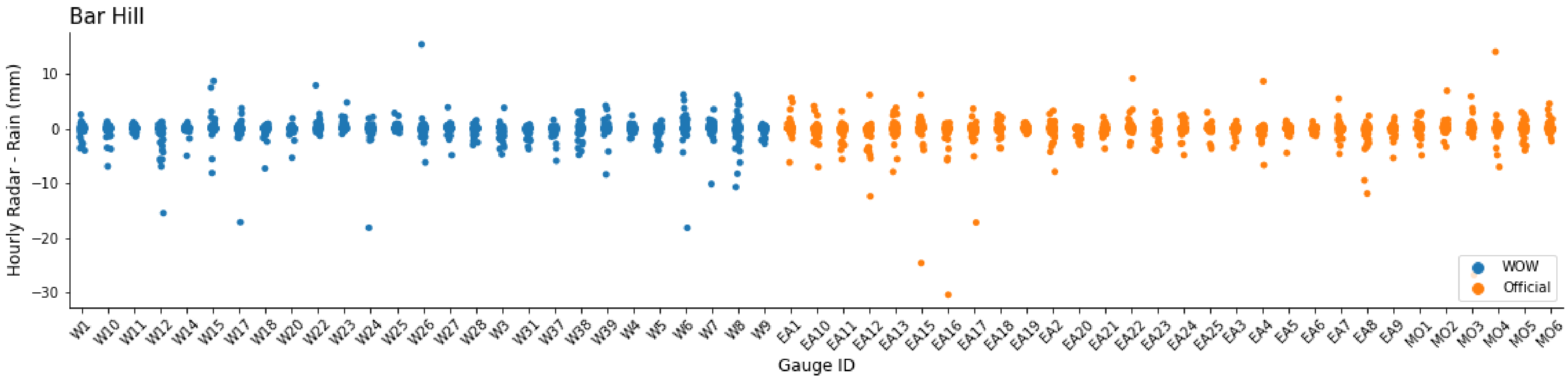


Figure 5.18 Milton Keynes selected event time series for WOW gauges W1 and W16 demonstrating discrepancy between radar and rain gauge (suffix 'rd' = radar, 'rg' = rain gauge)

Based on the evidence, W1 was removed from further analysis.

A comparison of the variance between the gauge datasets and radar can be seen in plots of the hourly difference between radar and rain gauge for the four selected events (see Figure 5.19).


Canvey Island
Hourly Radar - Rain (mm)
20
10
0
−10
−20
−30
WOW
Official
W1 W11 W12 W13 W14 W15 W16 W18 W2 W20 W22 W23 W24 W26 W27 W31 W33 W34 W39 W4 W40 W5 W6 W7 W9 EA1 EA10 EA11 EA12 EA13 EA14 EA15 EA16 EA17 EA18 EA19 EA2 EA20 EA21 EA22 EA23 EA24 EA25 EA26 EA27 EA28 EA29 EA3 EA30 EA31 EA32 EA4 EA5 EA6 EA7 EA8 EA9 MO1 MO2 MO3
Gauge ID


Bar Hill
Hourly Radar - Rain (mm)
10
0
−10
−20
−30
WOW
Official
W1 W10 W11 W12 W14 W15 W17 W18 W20 W22 W23 W24 W25 W26 W27 W28 W3 W31 W37 W38 W39 W4 W5 W6 W7 W8 W9 EA1 EA10 EA11 EA12 EA13 EA15 EA16 EA17 EA18 EA19 EA2 EA20 EA21 EA22 EA23 EA24 EA25 EA3 EA4 EA5 EA6 EA7 EA8 EA9 MO1 MO2 MO3 MO4 MO5 MO6
Gauge ID

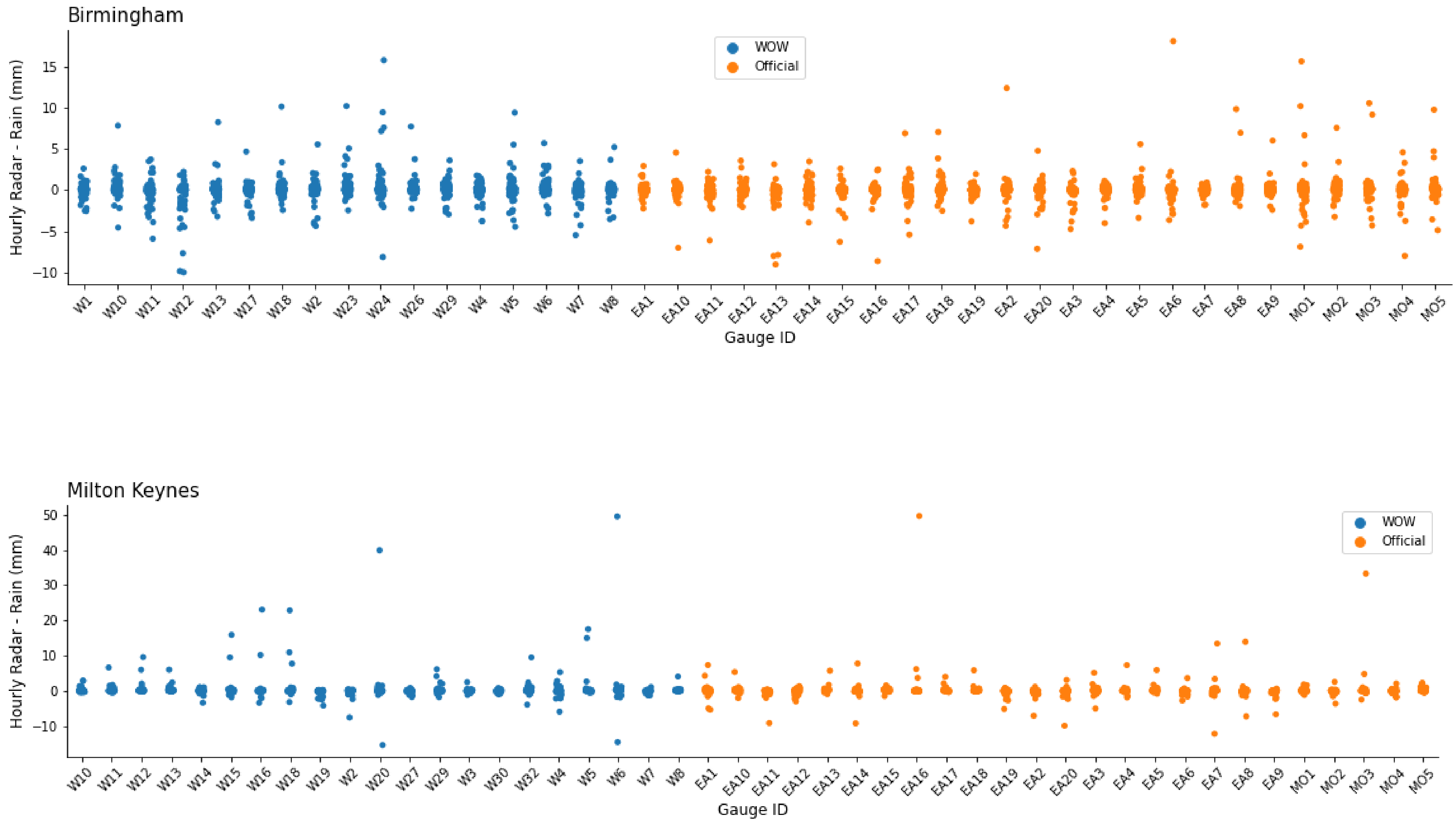


Figure 5.19 Strip plots of difference between hourly rain observed at gauge and co-located radar square for all events (radar – rain gauge).

Figure 5.19 demonstrates that whilst there are individual hours where there is a large discrepancy between the rain gauge and radar, there is no systematic difference between the WOW and Official rain gauges.

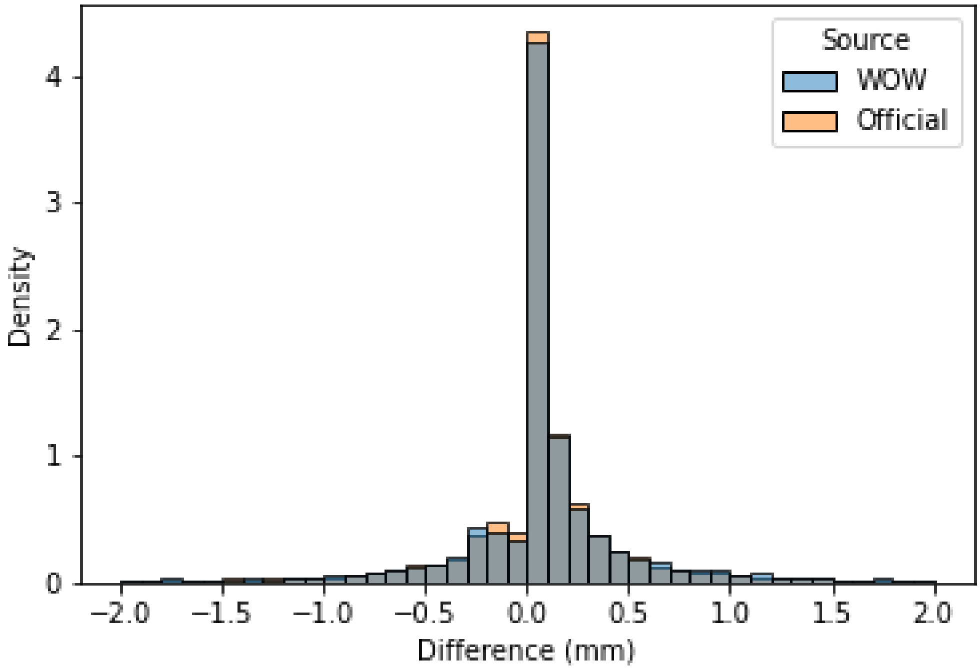


Figure 5.20 Histogram showing difference between co-located radar and gauge data for WOW and Official cohorts (normalised for sample size, grey = overlap)

Figure 5.20 shows the consistency between the manually-QC'd WOW and Official datasets when compared to radar observations, as recommended for rainfall validation by the WMO (see "2.4 Precipitation multi-source comparison test" {WMO, 2021 #418}). The Y axis has been normalised to account for the different sample size of the two datasets. The plot shows how little difference there is between the two datasets and is evidence indicating the WOW data are as accurate as the Official data for the majority of observations after manual SQC. The plot has been restricted on the X axis (-2,2), removing the extremes as they make it difficult to see the degree of similarity between the datasets. Extremes are of course very important, therefore, to test the similarity between the gauge datasets the RMSE was calculated between the rain gauge observation and co-located radar data for all gauges across all events and is presented in Figure 5.21.

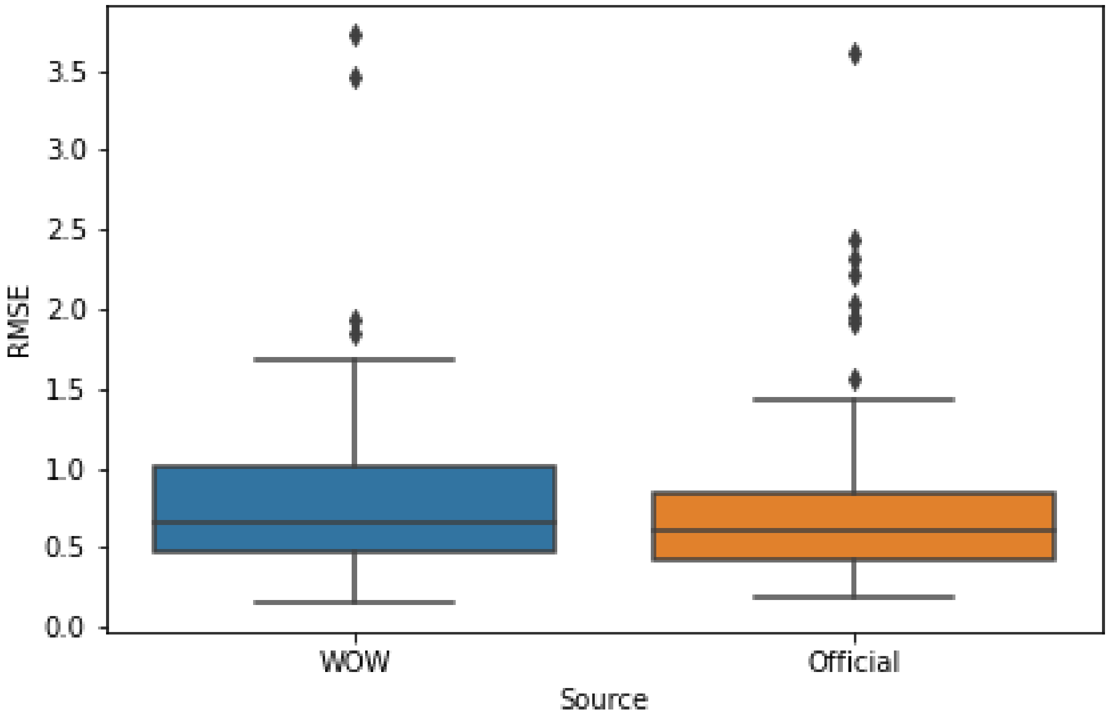


Figure 5.21 Boxplot of RMSE between hourly rain gauge and co-located radar rainfall values for WOW and Official gauge data for selected rainfall events.

The data presented in Figure 5.21 are also provided in Table 5.5. The mean, standard deviation and interquartile range was greater for WOW observations; however, there were more outliers in the Official observations.

Data Source	Count	Mean	Standard Deviation	Minimum	25%	50%	75%	Maximum
Official	115	0.75	0.53	0.19	0.43	0.61	0.84	3.62
WOW	90	0.81	0.59	0.14	0.47	0.66	1.01	3.73

Table 5.5 Descriptive statistics for RMSE for WOW and Official rain gauge data and co-located radar during selected rainfall events.

Finally, a t-test was conducted between the two datasets. The significance value was 0.43, far higher than the 0.05 required to demonstrate the similarity of the gauge datasets and confirming that there was no significant difference between the WOW and Official datasets. On the basis of results shown in this section it is concluded that the automated and manually SQC'd WOW data are as reliable as Official gauge data and can be used in the next stage (interpolation creation). The gauges with acceptable quality data for each event are presented in location maps in Figure 5.22 to Figure 5.25.

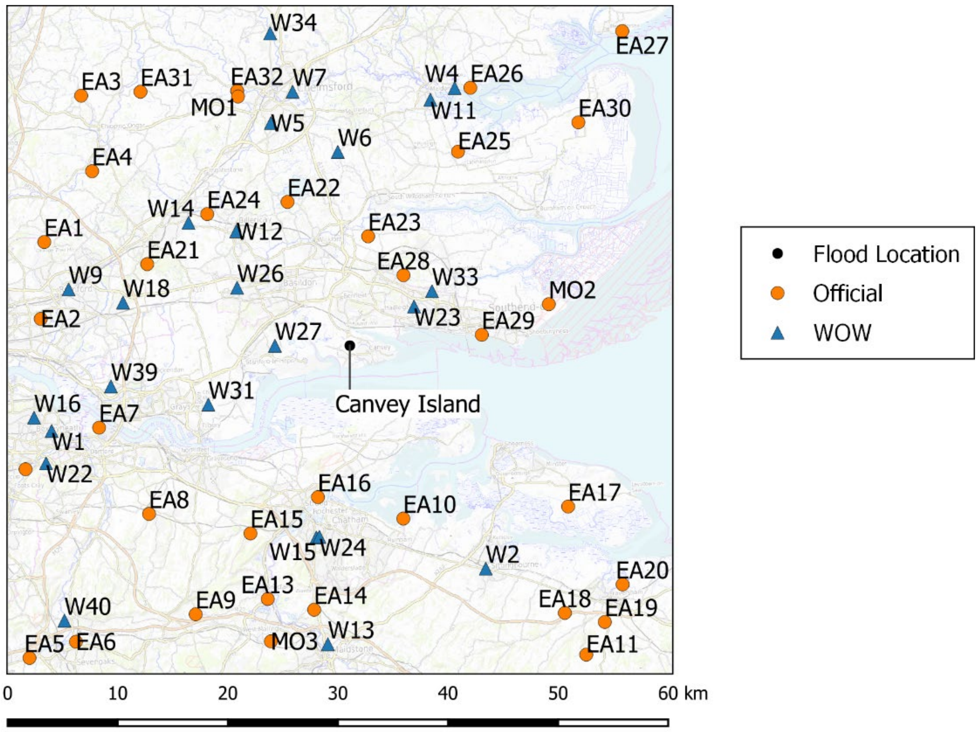


Figure 5.22 Map of Canvey Island event gauges (26 WOW and 35 Official)

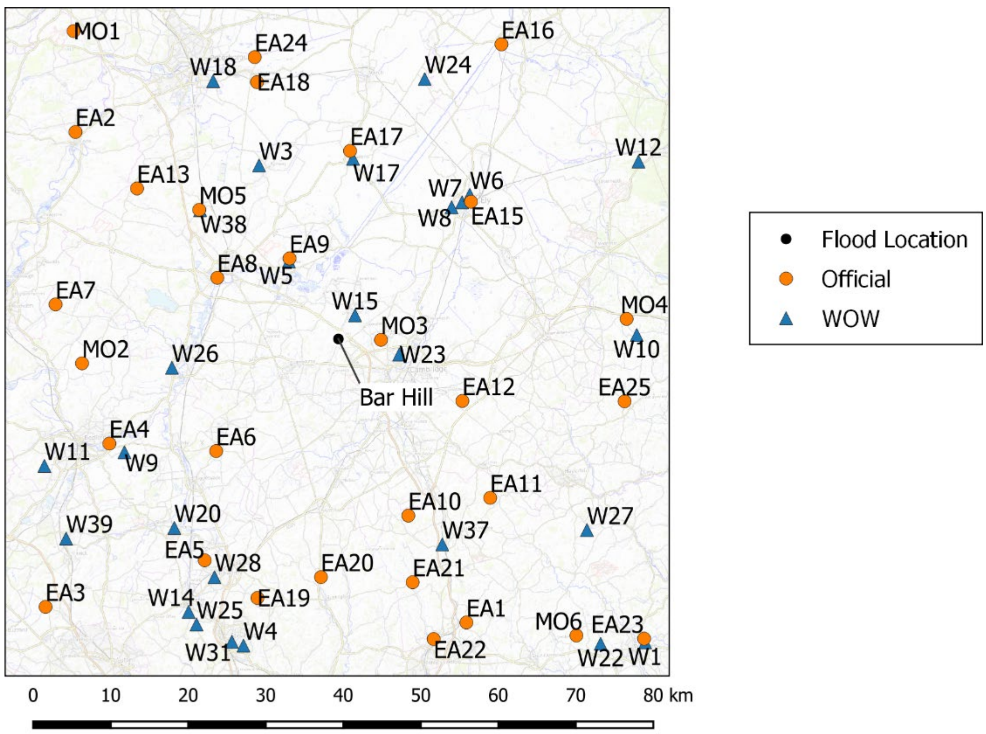


Figure 5.23 Map of Bar Hill event gauges (27 WOW and 31 Official)

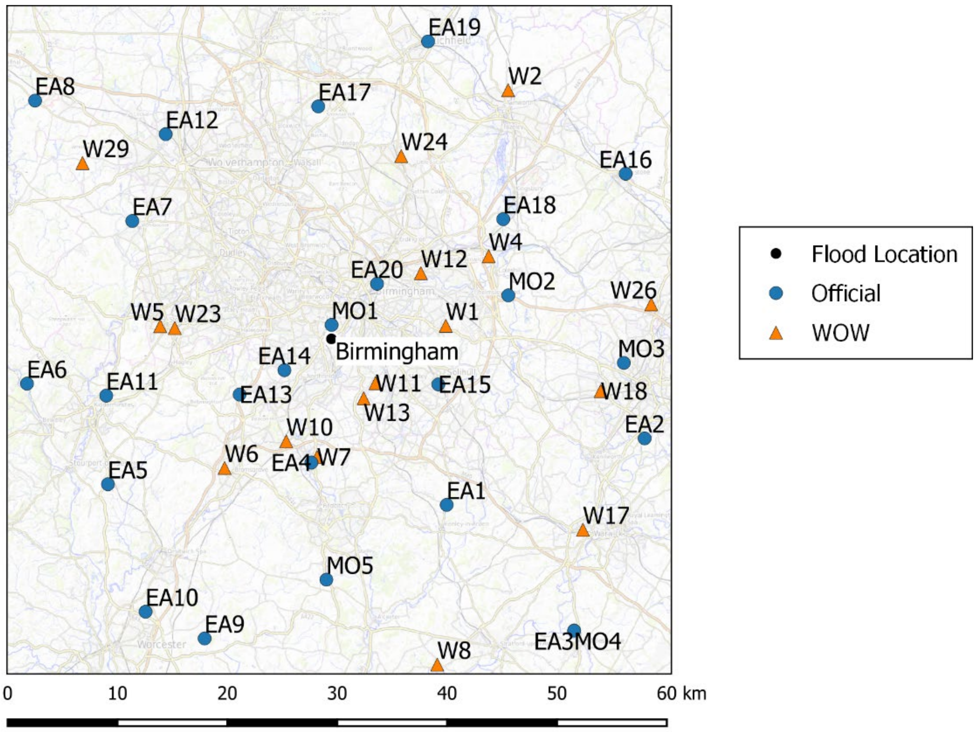


Figure 5.24 Map of Birmingham event gauges (17 WOW and 25 Official)

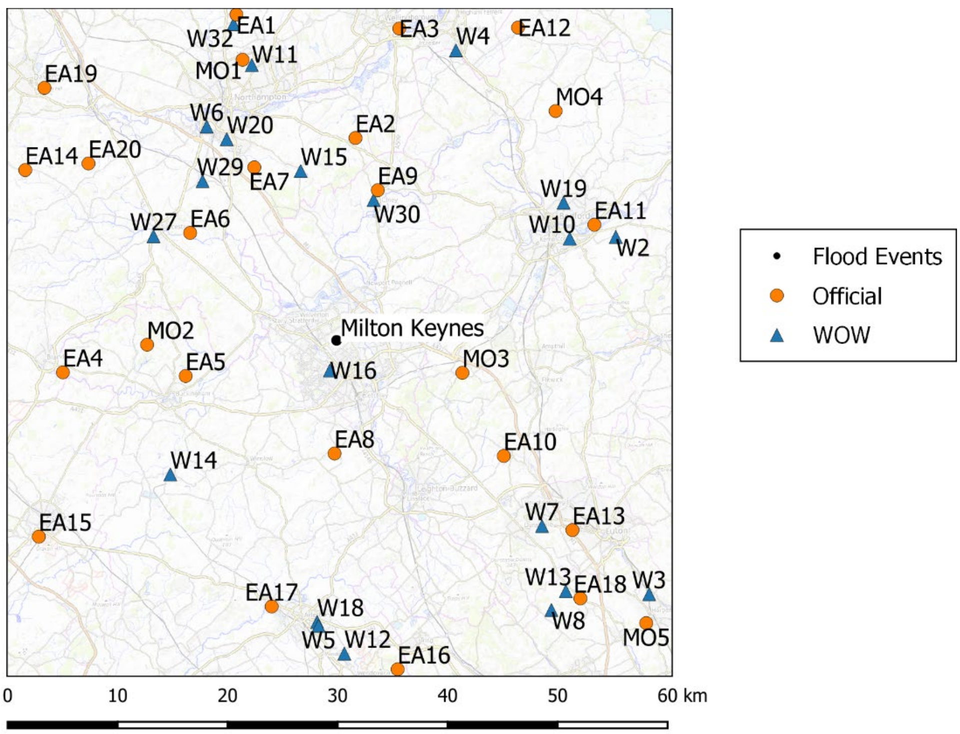


Figure 5.25 Map of Milton Keynes event gauges (23 WOW and 25 Official)

5.4.5 Creating Interpolations – Determining Event Temporal Extent

A short period around the time of flooding was selected for each event, to allow detailed comparison and analysis. The start and end point for the rain event was identified by looking at the radar and the timeseries for gauges closest to the flooded area and selecting the distinct rainfall period immediately before and after the flood event. Once the period of peak rainfall was identified comparisons were made between the interpolations on an hourly basis and then aggregated over the period.

Table 5.6 shows the relative duration and intensity of the rain (from radar) that resulted in flooding for each of the locations. The most intense rain was recorded during the Milton Keynes event (127 mm h^{-1}), whilst the lowest intensity was during the Birmingham event (49mm h^{-1}). A bigger difference between the maximum intensity and the maximum total rainfall indicatives lower intensity rain over a longer period. The intensity and duration of rain can impact the nature of flooding, high volume short events may result in highly localised 'flash flooding', whilst longer-duration lower-intensity rain may impact a larger area with longer-lasting effects and higher economic consequences due to damage (Zhai *et al.*, 2006; Flack *et al.*, 2019).

Event	Selected Duration (Hours)	Maximum Total Rainfall for duration (mm)	Maximum intensity rainfall (mm h^{-1})
Canvey Island	3	125	91
Bar Hill	6	97	72
Birmingham	4	94	49
Milton Keynes	3	134	127

Table 5.6 Event Overview, Duration, Intensity and Total Rain as recorded by radar

The radar data are corrected prior to publication using MFB and therefore should be accurate; however, the maximum hourly intensity for Milton Keynes (127mm h^{-1}) exceeds the UK maximum for gauged rainfall (92mm h^{-1}), and the Canvey Island maximum was close to the maximum (91mm). It is possible the radar was overestimating the rainfall for these two events, or that gauges have never captured the highest intensity rain due to the limited extent of convective storms overlying gauge locations (Lengfeld *et al.*, 2020).

Canvey Island

The period 14:00 – 16:00 on 20/07/14 was selected to represent the rainfall that resulted in flooding on Canvey Island as there was minimal rain outside of these hours (see Figure 5.26).

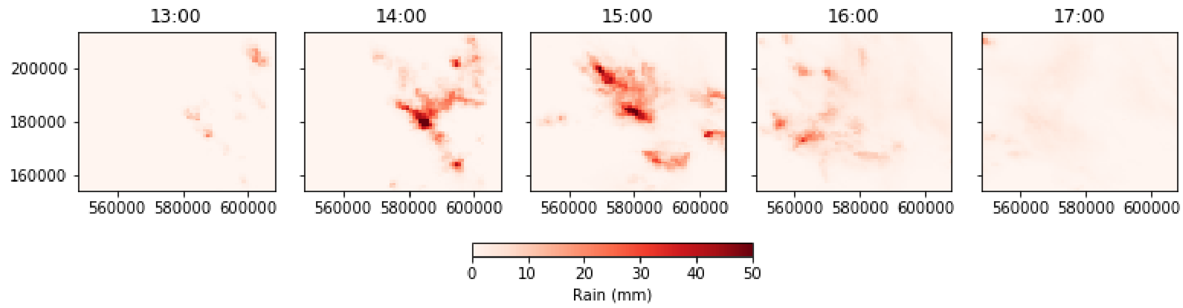


Figure 5.26 Canvey Island radar peak rainfall hours on day of flooding (20/07/2014).

The radar pictured in Figure 5.26 indicated that rainfall was localised, with a large part of the study area receiving minimal or no rain during the peak period. At 14:00 an area of high intensity rain was over Canvey Island, with the rain becoming more widespread up to 15:00. The maximum rainfall recorded via radar was 91 mm h^{-1}, close to the highest hourly observed rainfall in a gauge in the UK (92 mm h^{-1}). When aggregated from 14:00 – 16:00 the radar displays an area of intense rain centred on Canvey Island (Figure 5.30). The maximum radar rainfall over the 3-hour duration for a 1 km grid square was 125 mm.

Bar Hill

The radar showed small areas of intense rain at different times during the day on 08/08/2014, with the most intense rain falling on Bar Hill in the hour to 15:00. The period 12:00 – 17:00 was selected to represent the peak of the rainfall that resulted in flooding around Bar Hill (see Figure 5.27).

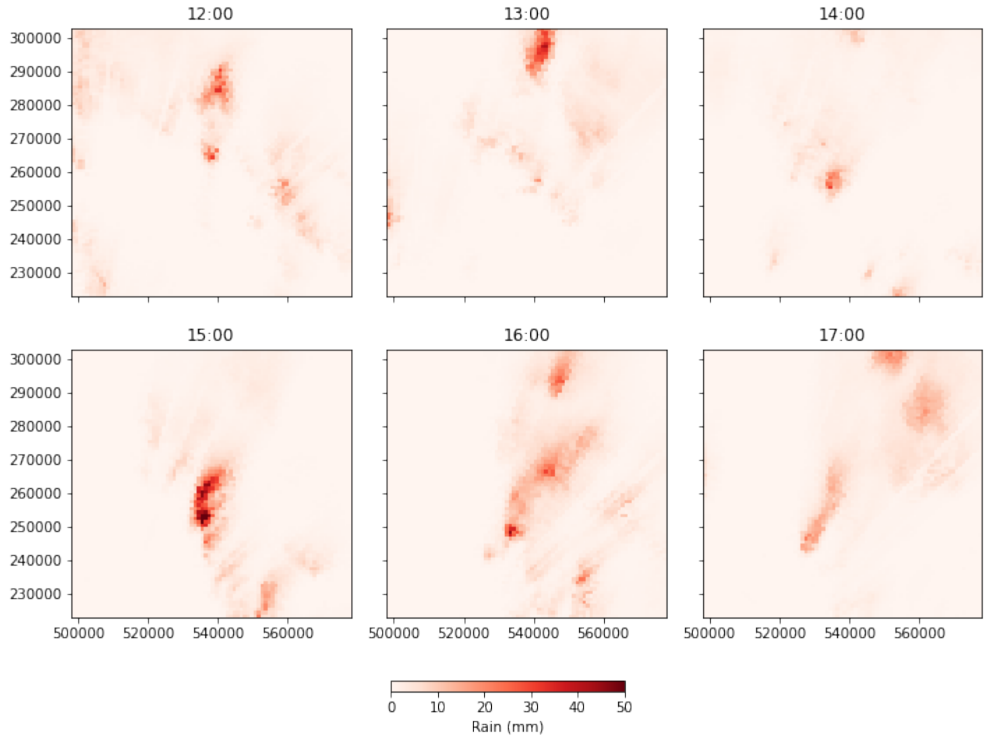


Figure 5.27 Bar Hill radar peak rainfall hours on day of flooding (08/08/2014).

The maximum rainfall for 1 km grid square recorded via radar was 72 mm h^{-1} (within the range of observed rainfall in UK rain gauges). The maximum total rainfall recorded by radar for the selected 6 hours for a 1 km grid square was 97 mm.

Birmingham

The period 17:00 to 20:00 on 16/06/2016 was selected to represent the rainfall that resulted in flooding in Birmingham, as there was minimal rainfall outside these hours (see Figure 5.28).

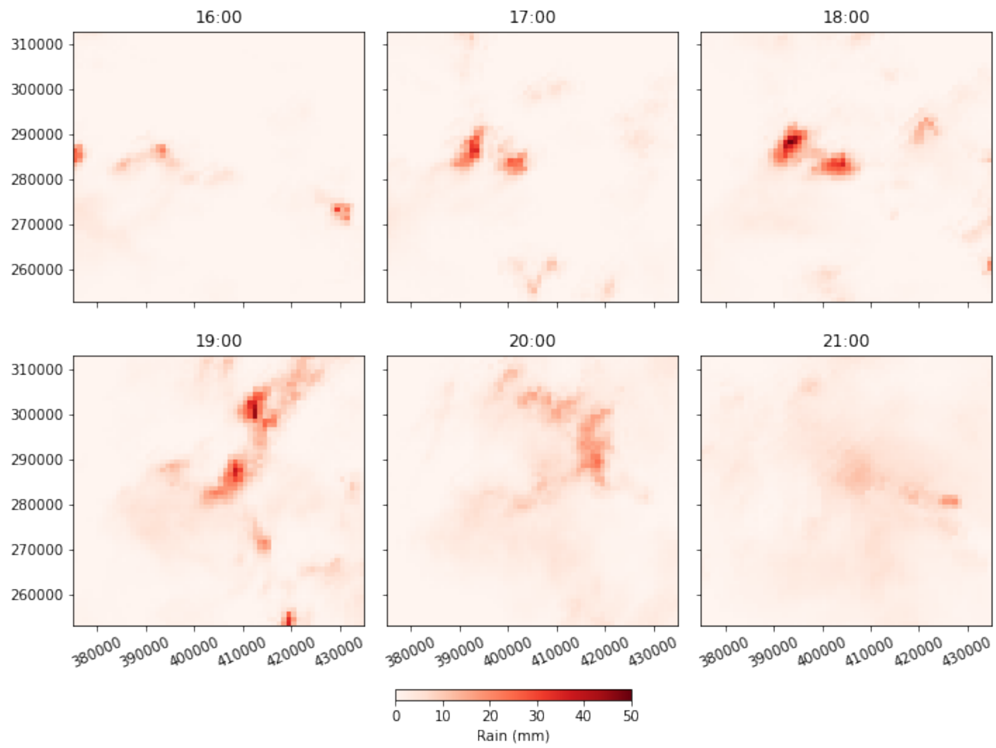


Figure 5.28 Birmingham radar peak rainfall hours on day of flooding (16/06/2016).

The radar showed an area of intense rainfall moving west to east across the city from around 17:00 – 20:00. The maximum total radar rainfall over 4 hours for a 1 km grid square was 94 mm. Rain was persistent over the period rather than short and intense as seen in Bar Hill. The maximum rainfall recorded via radar was 49 mm h^{-1}, which makes it the lowest intensity of the 4 events; however, as can been seen in Figure 5.28 the rain moved relatively slowly over the area. Birmingham is the most urbanised of the selected event locations, and there was heavy rainfall for several days prior to flooding which is likely to have contributed to the impact of the rain on the 16th of June (see Figure 5.8).

Milton Keynes

The period 19:00 to 21:00 on 27/05/2018 was selected to represent the rainfall that resulted in flooding around Milton Keynes, as there was minimal rainfall outside these hours.

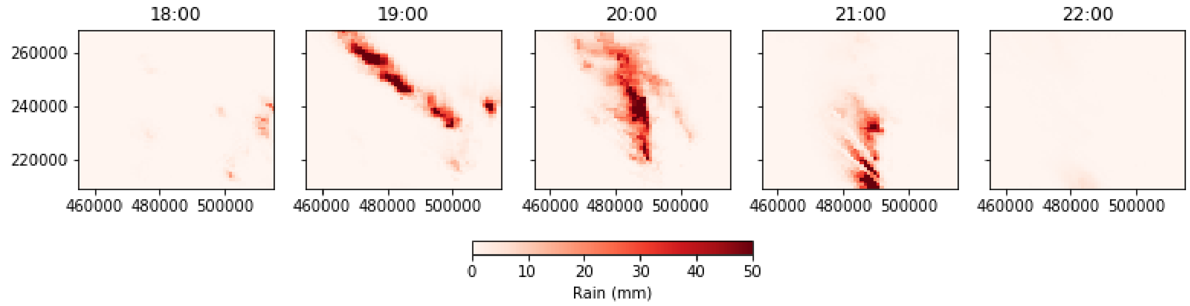


Figure 5.29 Milton Keynes radar peak rainfall hours on day of flooding (27/05/2018).

The radar shows that the rainfall that resulted in flooding in Milton Keynes was intense and widespread, the maximum was >105 mm $h^{-1}$from 19:00 – 21:00. The maximum rainfall over the period for a 1 km grid square was 134 mm. The maximum rain intensity was 127 mm h^{-1}, exceeding the gauged maximum hourly record for the UK.

5.4.6 Creating Interpolations – Determining Event Spatial Extent/Gauge Coverage

The spatial distribution of gauges was considered in relation to the flooded area and the area of most rain, as determined by the sum of the rainfall detected by radar over the rainfall event period. The rainfall was more widespread across Birmingham; hence it was captured in more gauges. Where the rainfall was highly localised in the other three events, there were fewer gauges in the area of highest intensity.

For the Canvey Island event there were no active gauges on the island at the time of the most intense rainfall, as several were removed during the manual SQC process due to gaps in the timeseries. There may have been sufficient flooding to prevent gauge operation, either inundation or electrical fault (see 5.4.2 for event details).

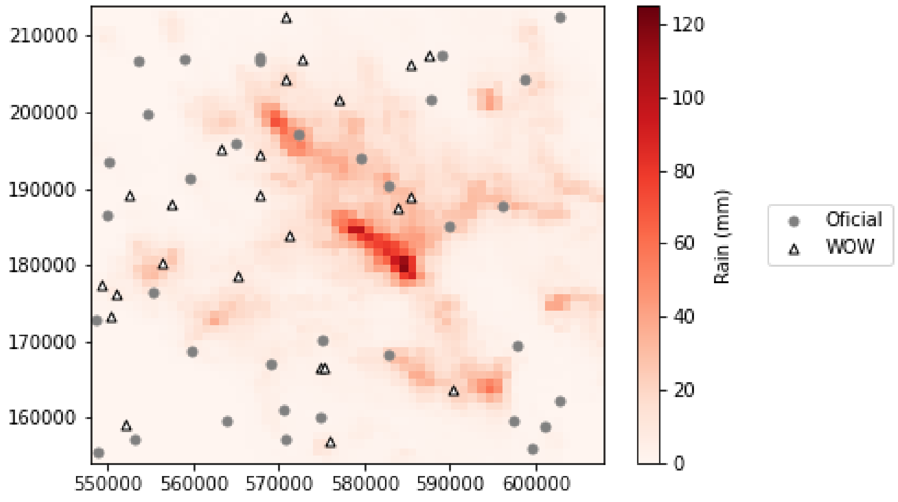


Figure 5.30 Canvey Island total rainfall (radar from 14:00 - 16:00 20/07/2014), with Official and WOW rain gauge locations.

The lack of gauges in the area of highest intensity rain means the selected block kriging may poorly represent peak rainfall. An alternative method such as Bayesian merging incorporating the numerical contribution of radar data and would likely be more effective here (Ochoa-Rodriguez *et al.*, 2019b). For the Bar Hill event there were similarly no gauges in the area of highest intensity rainfall and radar displays some errors, the effect of beam blockage can be seen in Figure 5.31 where data have been aggregated over 6 hours.

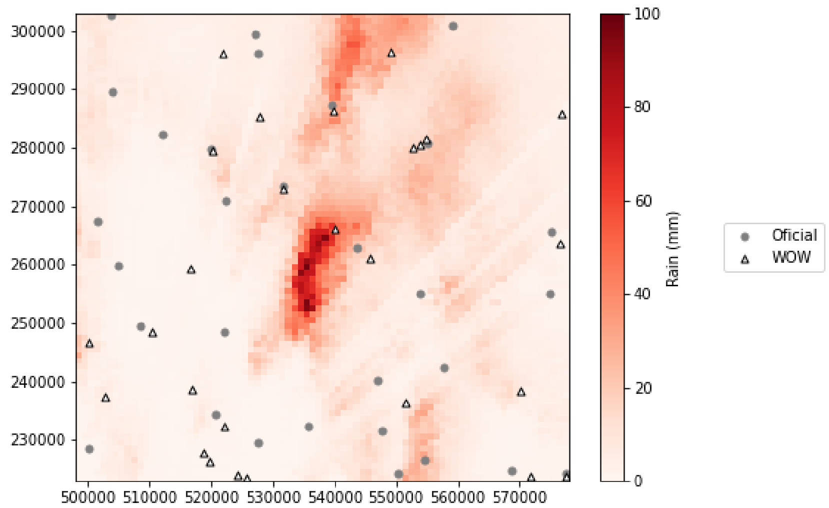


Figure 5.31 Bar Hill total rainfall (radar from 12:00 - 17:00 08/08/2014) with Official and WOW rain gauge locations.

For the Birmingham event there are multiple WOW and Official gauges located within the area of highest rainfall (see Figure 5.32).

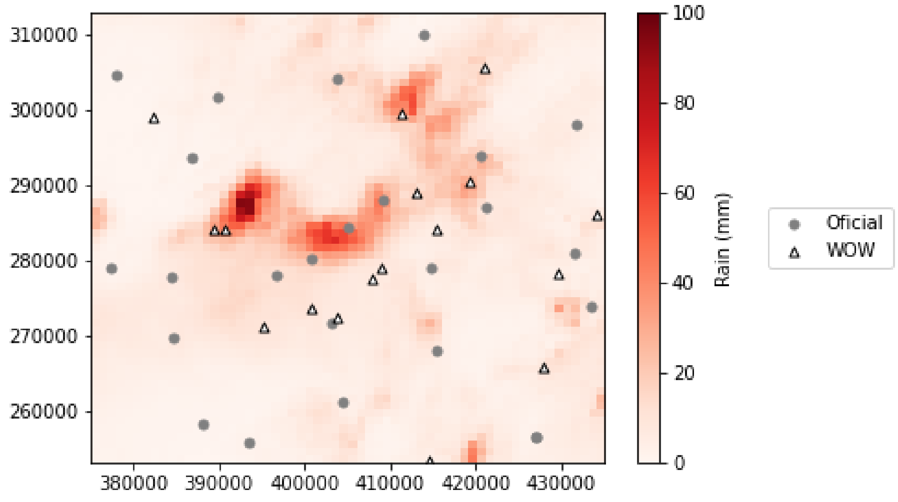


Figure 5.32 Birmingham total rainfall (radar from 16:00 - 20:00 16/06/2016) with Official and WOW rain gauge locations.

The more widespread nature of rainfall in the Birmingham event and the fortuitous position of gauges mean that the interpolations using gauge data were more likely (than the other events) to capture the spatial distribution of rainfall. For the Milton Keynes event, Figure 5.33 shows there were gauges located in an area of heavy rainfall to the northwest, but not in the central and southern area where heavy rainfall was also indicated by the radar.

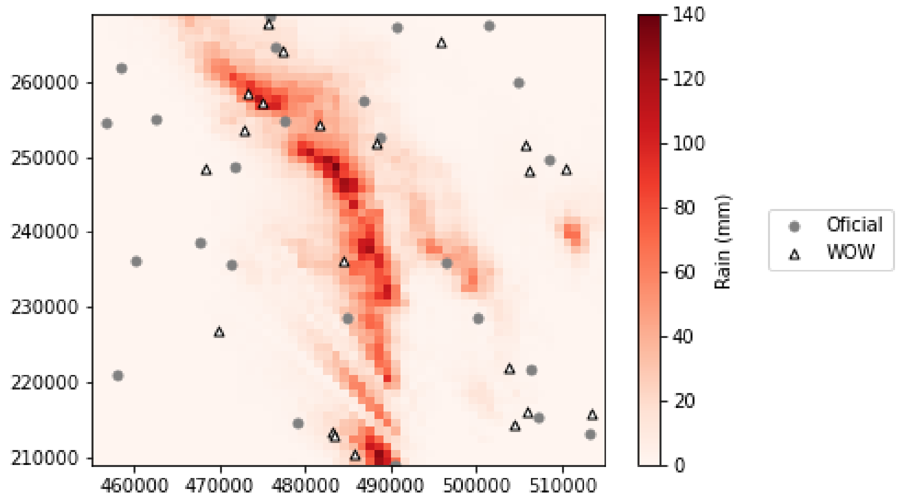


Figure 5.33 Milton Keynes total rainfall (radar from 19:00 – 21:00 on 27/05/2018) with Official and WOW rain gauge locations

The effect of radar beam blockage can be seen in Figure 5.33 highlighting a difficulty in rainfall analysis (as there are no perfect datasets).

5.4.7 Assessing the 'Value Add' of Including WOW Data

Multiple interpolations were created using the combinations of datasets detailed in the methodology (see section 5.3.4). The interpolations show the benefit of including all available data, from Official and WOW gauges and radar[14].

Canvey Island

Despite the lack of gauges in the region of highest rainfall (as indicated by the radar), Figure 5.34 shows the benefit of using both the Official and WOW data for interpolations, and the added value of blending radar with gauge observations. In the figure it possible to see the relative contribution to the interpolation by the different datasets, and the incremental improvement of including additional gauges (combined WOW/Official). It also shows the benefit of using radar to confine the gauge derived interpolation of rain depth for the event (Radar/WOW/Official).

As already mentioned in section 5.4.6, at a maximum of 91 mm h^{-1}it is possible that the radar overestimated the rainfall depth in the Canvey Island event but without gauges in the area of highest rainfall questions around radar accuracy remain. The interpolations using only gauge data show the potential to underestimate the depth of rain in 'peak' areas if gauges do not correspond to the locations of highest rainfall.

[14] In the following interpolation plots rain gauges are shown but are intentionally very light and small, to make it easier to focus on the interpolations whilst having an appreciation of how the gauge location may have influenced the output.

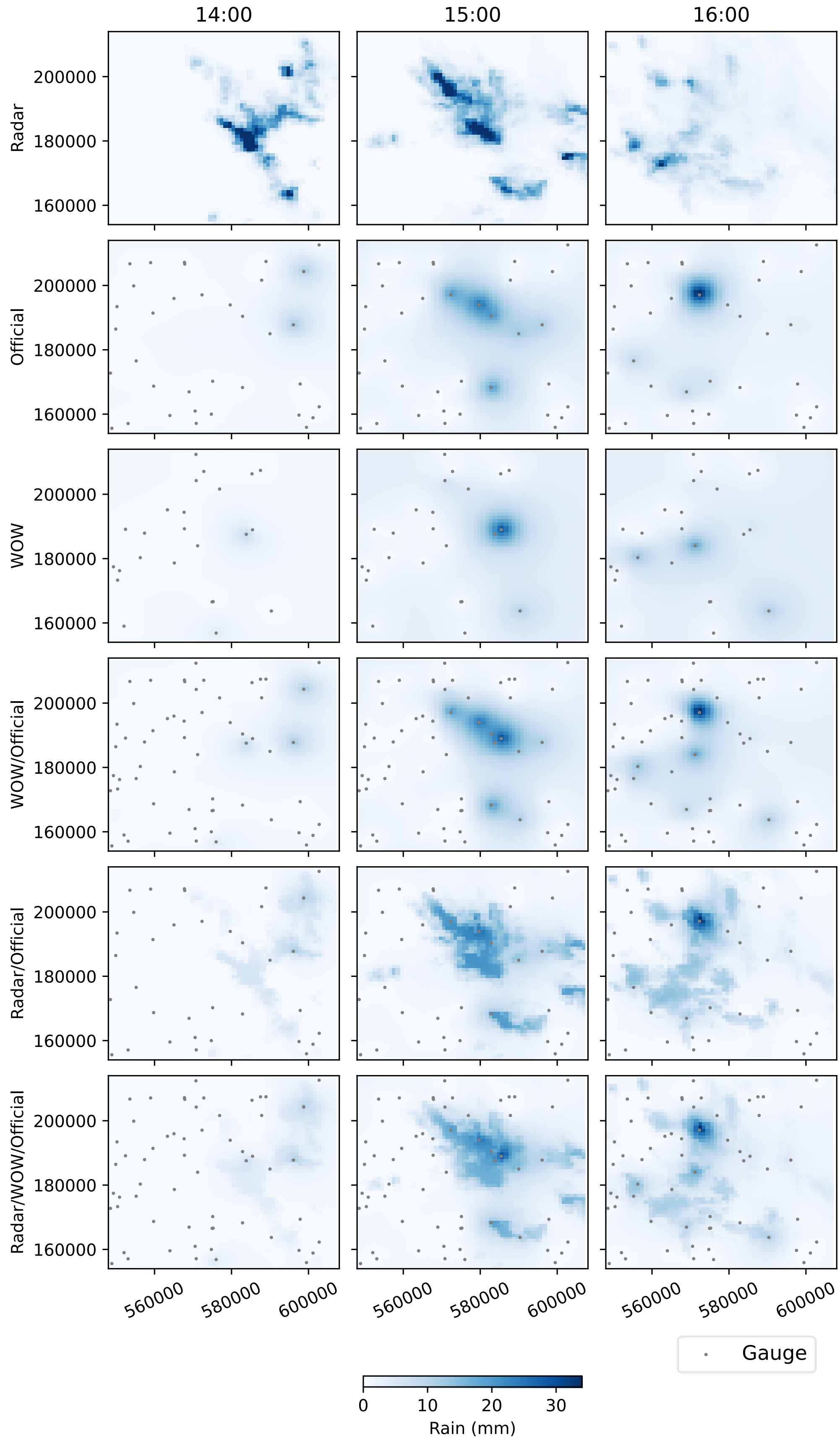


Figure 5.34 Radar and interpolations (Official and WOW block kriged, remainder block kriged with external drift) of event rainfall on Canvey Island 20/07/14, 14:00 - 16:00.

Variation between the interpolations is quantified in Table 5.7. For each interpolation the table shows the maximum hourly rainfall for a grid square, the maximum total rainfall over the storm event for a grid square, the number of grid squares where rainfall >0.125 mm was recorded/interpolated (approximately the smallest bucket size of a TBR gauge), the total rainfall across the study area and the mean rainfall for the grid squares where rain was observed.

Data Source	Maximum Event Total for 1km^2 (mm)	Hourly Maximum (mm h^{-1})	Extent of Rainfall (km^2)	Total Rain (mm)	Mean Rain (mm)
Radar	115.7	90.9	3236	27547.86	8.51
WOW	36.83	29.13	3481	27574.37	7.92
Official	55.33	36.46	3480	27951.68	8.03
WOW/Official	54.65	35.96	3474	28343.71	8.16
Radar/Official	55.75	35.36	3469	28143.64	8.11
Radar/WOW/Official	53.38	34.23	3459	27645.66	7.99

Table 5.7 Descriptive statistics from interpolations of Canvey Island rainfall datasets from different sources.

In Table 5.7 the hourly maximum and event maximum were higher for the radar than the other interpolations and blends, further confirming the gauges were not well positioned to capture the peak of the rainfall. In this instance there was a 117% increase between the maximum event total derived from the radar/WOW/Official blend versus the radar record. Conversely the total rainfall estimated by the radar was the lowest, as was the number of grid squares where rain was observed/interpolated, which highlights the importance of using the radar to determine the spatial distribution of rainfall. Point-to-point interpolation methods do not recognise breaks in rainfall between two points, hence the value of including radar to delineate the spatial distribution of rainfall.

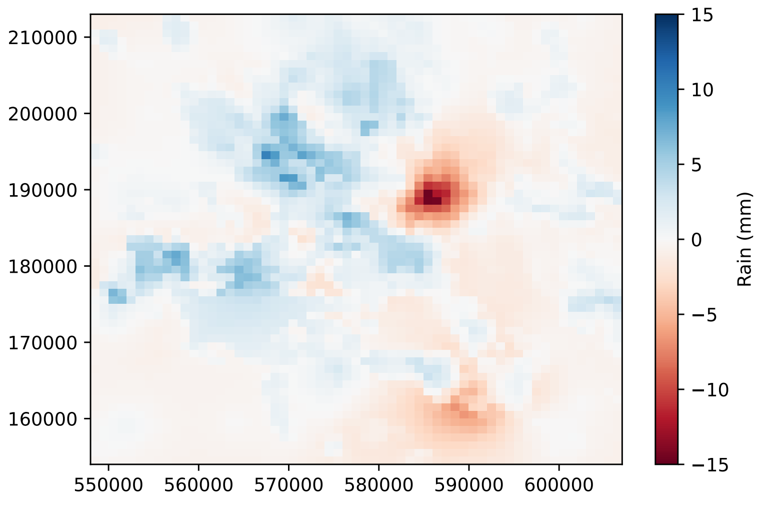


Figure 5.35 Difference plot between radar/Official interpolation - radar/WOW/Official interpolation for Canvey Island rain (14:00 - 16:00 20/07/2014).

In Figure 5.35 the difference between the interpolation generated using radar and Official data minus the interpolation created using radar, Official, and WOW data can be seen. Where the values are negative, rainfall was captured in a WOW gauge that was not recorded by an Official gauge. Where the values are positive, the additional WOW data has constrained the Official interpolation meaning that rainfall in these areas would not be overestimated.

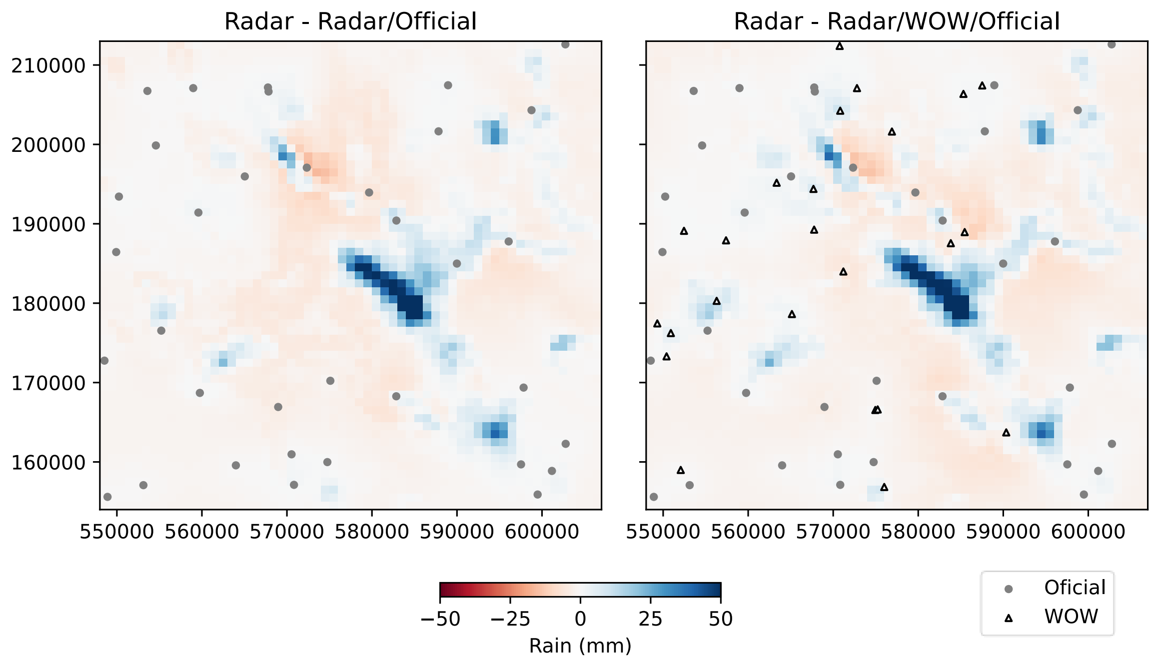


Figure 5.36 Difference plots between radar and interpolations for Canvey Island Rain (14:00 - 16:00, 20/07/2014).

In Figure 5.36 the left plot shows the difference between radar minus the interpolation created using radar and Official data. The right plot shows the difference between radar minus the interpolation created using both gauge data sets and blended with radar. It is hard to see where the differences lie in these plots in part due to the high intensity rainfall around Canvey Island which skews the scale. Around the south-central area of the plot an area of intense rainfall appears more confined and better defined when data from a WOW gauge are included in the interpolation, preventing the over extrapolation of radar.

Bar Hill

The interpolations in Figure 5.37 from the Bar Hill event suggest a discrepancy between the WOW and Official datasets at 16:00, not accounted for by the spatial distribution of rain as indicated by the radar. The rain at Official gauge MO3 was 38 mm h^{-1}, whilst the co-located radar recorded 11.2 mm h^{-1}, and the nearest WOW gauge (W15 4.6 km to the NW) recorded 21 mm h^{-1}. Given the disparity in the observations there is no way to be sure that the data from MO3 are incorrect; however, the benefit of incorporating additional data points in interpolations is minimising the reliance on observations from any given gauge. Additionally, where rainfall is highly localised, more data points improve the overall calculation of the rainfall field as the high value will not be over-extrapolated. In instances where observations are incorrect the same will be true and poor data will have a limited effect on the overall rainfall total. In Figure 5.37 it is possible to see the effect of beam blockage in the radar (Radar/WOW/Official row).

The variation between the interpolations is quantified in Table 5.8.

Data Source	Maximum Event Total for 1km^2 (mm)	Hourly Maximum (mm h^{-1})	Extent of Rainfall (km^2)	Total Rainfall (mm)	Mean Rainfall (mm)
Radar	97.13	72.09	6314	55682.93	8.82
WOW	44.09	24.06	6232	71291.84	11.44
Official	54.02	35.61	6240	68719.10	11.01
WOW/Official	54.19	34.03	6215	65445.48	10.53
Radar/Official	58.33	35.21	6238	76509.75	12.27
Radar/WOW/Official	61.29	33.76	6225	76705.91	12.32

Table 5.8 Descriptive statistics from interpolations of Bar Hill rainfall data from different sources.

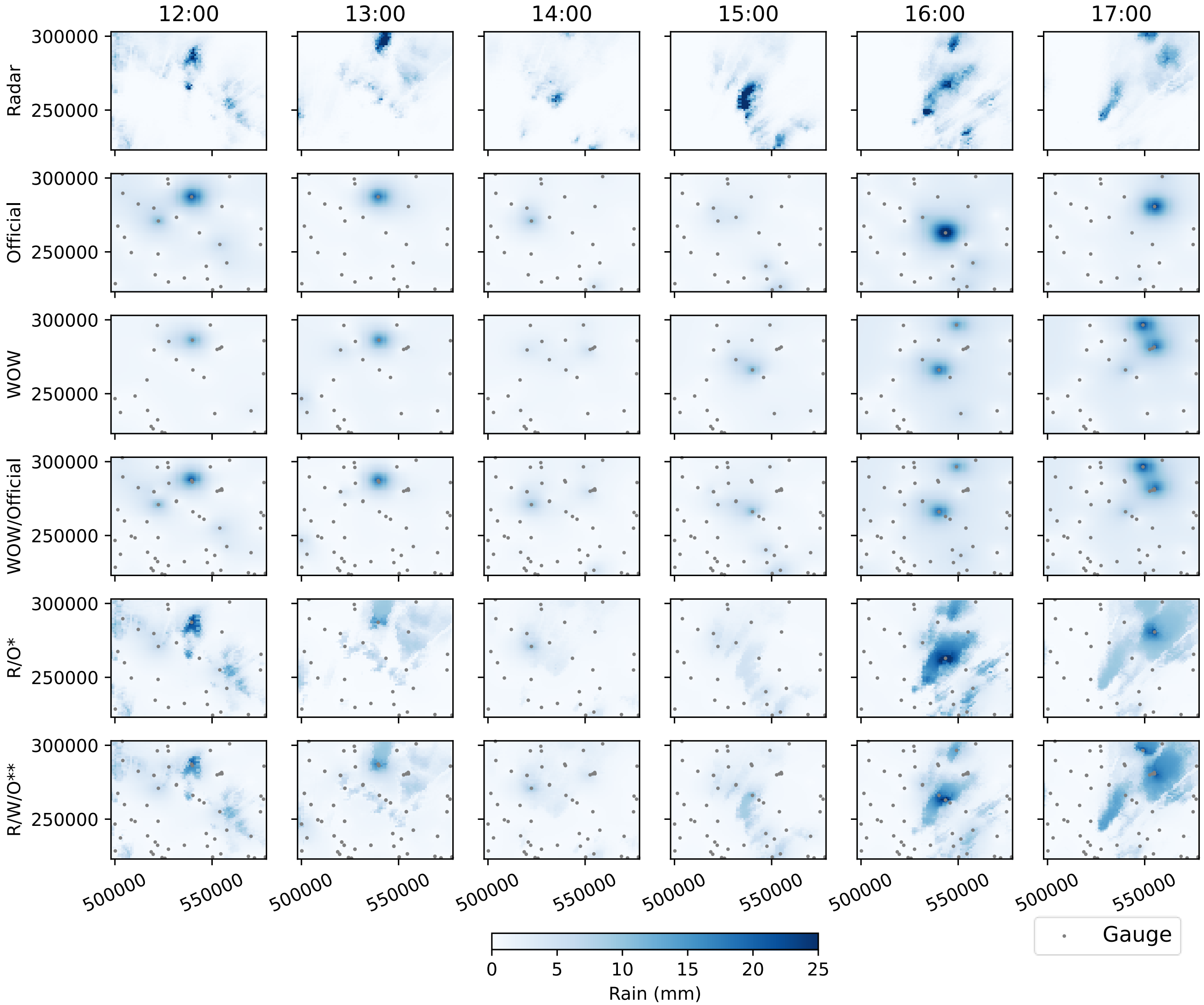


*Figure 5.37 Interpolations Bar Hill, 08/08/2014 12:00 – 17:00 *R/O = Radar/Official **R/W/O = Radar/WOW/Official.*

As with the Canvey Island event, the radar recorded an hourly maximum higher than observed at any of the gauges, leading to a discrepancy between the interpolations. Again, this appears to be due to a lack of gauges in the area of highest intensity rather than a disagreement between data sources. Interestingly, the total rain is much higher in the radar/WOW/Official interpolation than the radar. Looking at Figure 5.37 this appears to be at least partially due to an increase in the rain in the NE of the area at 17:00.

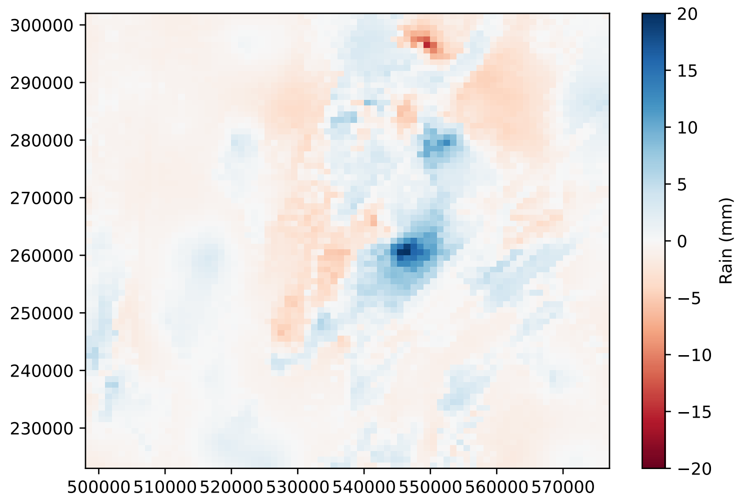


Figure 5.38 Difference plot between radar/Official interpolation - radar/WOW/Official interpolation for Bar Hill (12:00 - 17:00, 08/08/2014).

In Figure 5.38 it is possible to see the influence of additional gauges and the large difference between observations at two nearby gauges to the right of the centre of the plot as already mentioned in the discussion.

Birmingham

During the Birmingham event the gauges were better placed to detect areas of intense rainfall (see Figure 5.39). This may partially be due to the relatively low intensity, persistent rainfall that moved slowly across the area, and fewer gauges removed during manual SQC, as compared to the Canvey Island event. The variation between the interpolations is quantified in Table 5.9.

Data Source	Maximum Event Total for 1km^2 (mm)	Hourly Maximum (mm/hr)	Extent of Rainfall (km^2)	Total Rainfall (mm)	Mean Rainfall (mm)
Radar	80.54	48.84	3556	26217.62	7.37
WOW	39.31	19.3	3540	44774.48	12.65
Official	29.83	20.95	3540	26998.87	7.63
WOW/Official	39.47	21.28	3540	32913.57	9.3
Radar/Official	34.87	22.19	3540	28060.3	7.93
Radar/WOW/Official	39.42	21.18	3540	30683.54	8.67

Table 5.9 Descriptive statistics from interpolations of Birmingham rainfall data from different sources.

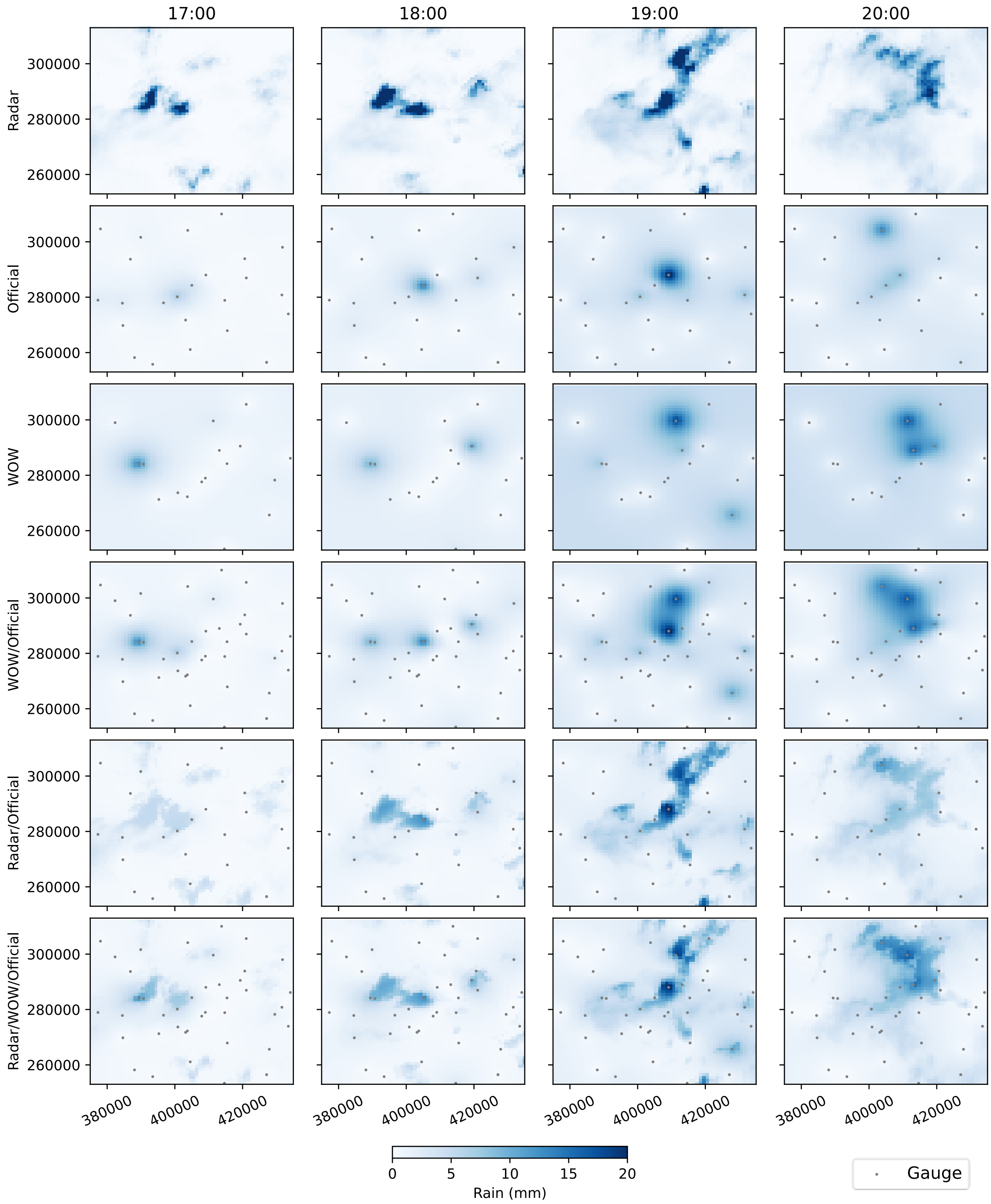


Figure 5.39 Interpolations Birmingham, 16/06/2016 17:00 – 20:00.

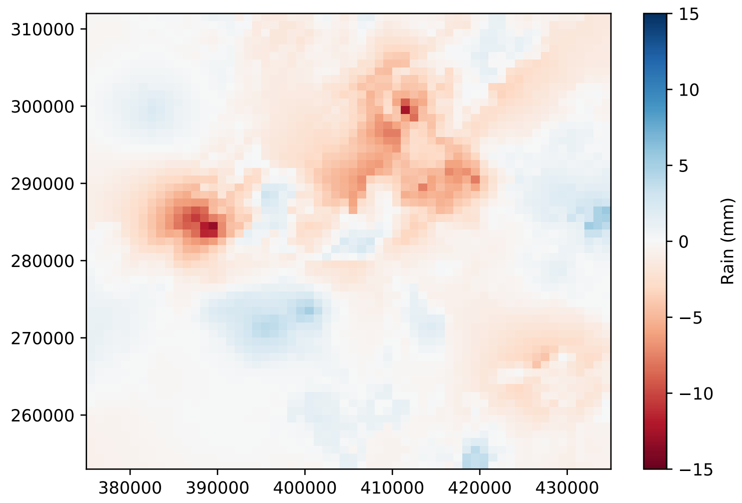


Figure 5.40 Difference plot between radar/Official interpolation - radar/WOW/Official interpolation for Birmingham (17:00 – 20:00, 16/06/2016).

In Figure 5.40 the contribution of WOW gauge data is very clear in places. The blend incorporating WOW data demonstrates areas to the north and west of the plot where rain was underestimated by almost 15 mm when only the radar and Official rain gauge data were used in interpolation.

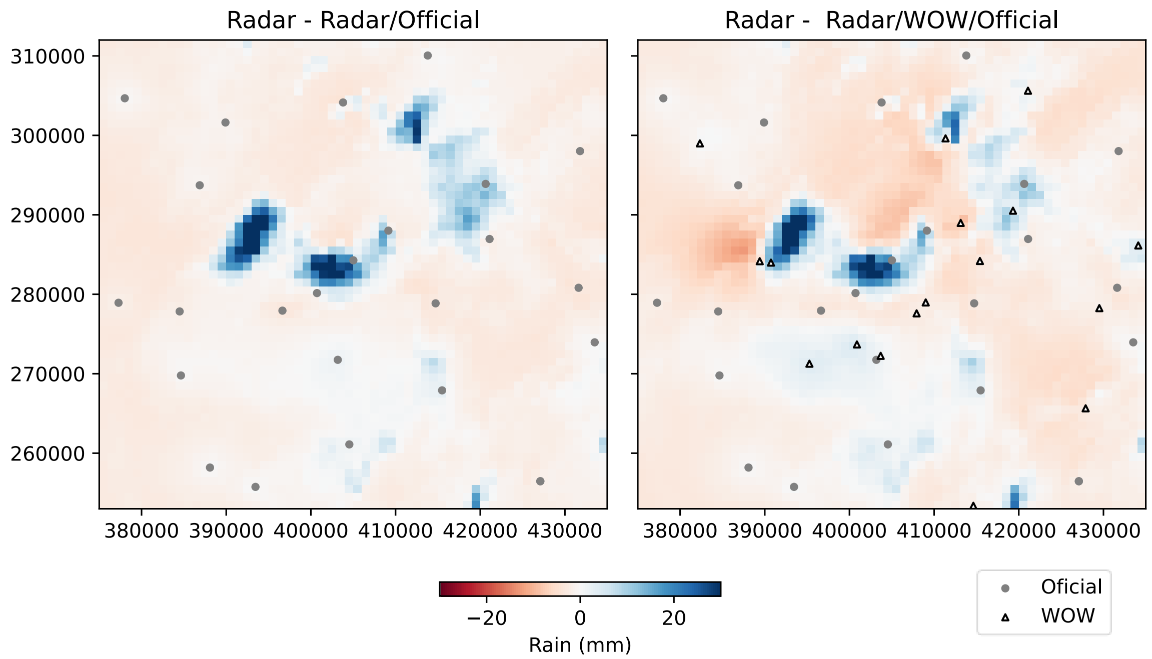


Figure 5.41 Difference plots between radar and interpolations for Birmingham rain (17:00 - 20:00, 16/06/2016).

In Figure 5.41 it is possible to see the benefit of including WOW data in the interpolation. The difference between the radar and the radar/WOW/Official data is greater in many areas then the difference between the radar and the radar/Official. This suggests the rainfall was more spatially confined than indicated by the Official data alone, as the WOW gauges prevent over-extrapolation from the Official rain gauges.

Milton Keynes

The rainfall that caused flooding in Milton Keynes was a part of a multi-cell convective system >100 km in length (see Figure 5.29) that caused flooding in other locations, including Birmingham (see Appendix 3, Daily Weather Summary). The intensity of the rainfall was high in the Milton Keynes area for three hours from 19:00 – 21:00 on the 27th May, 2018 (see Figure 5.42). Beam blockage is visible in the radar rainfall field at 21:00. The closest radar to Milton Keynes (and Bar Hill where a similar blockage is visible) is "Chenies" (see Figure 2.1), suggesting that at least up to 2018 there was an issue with the radar signal.

When comparing the interpolation including or excluding WOW gauges, two WOW gauges were well placed to observer high intensity rainfall at 19:00 and 20:00. WOW gauges W6 and W20 provide validation for the high rainfall intensities captured by the radar in the northwest of the study area, providing a more realistic representation of the rainfall than that observed by the Official gauges alone. Unfortunately, however, there were no gauges in the area of highest intensity rain in the south and central parts of the study area.

The variation between the interpolations is presented in Table 5.10.

Data Source	Maximum Event Total for 1km^2 (mm)	Hourly Maximum (mm/hr)	Extent of Rain (km^2)	Total Rain (mm)	Mean Rain (mm)
Radar	133.79	126.96	2605	35545.93	13.65
WOW	57.05	34.48	3588	27177.49	7.57
Official	33.24	22.18	3595	23348.61	6.49
WOW/Official	57.05	34.48	3588	25883.25	7.21
Radar/Official	32.69	22.44	3593	21747.96	6.05
Radar/WOW/Official	56.17	33.5	3584	24968.00	6.97

Table 5.10 Descriptive statistics from interpolations of Milton Keynes rainfall data from different sources.

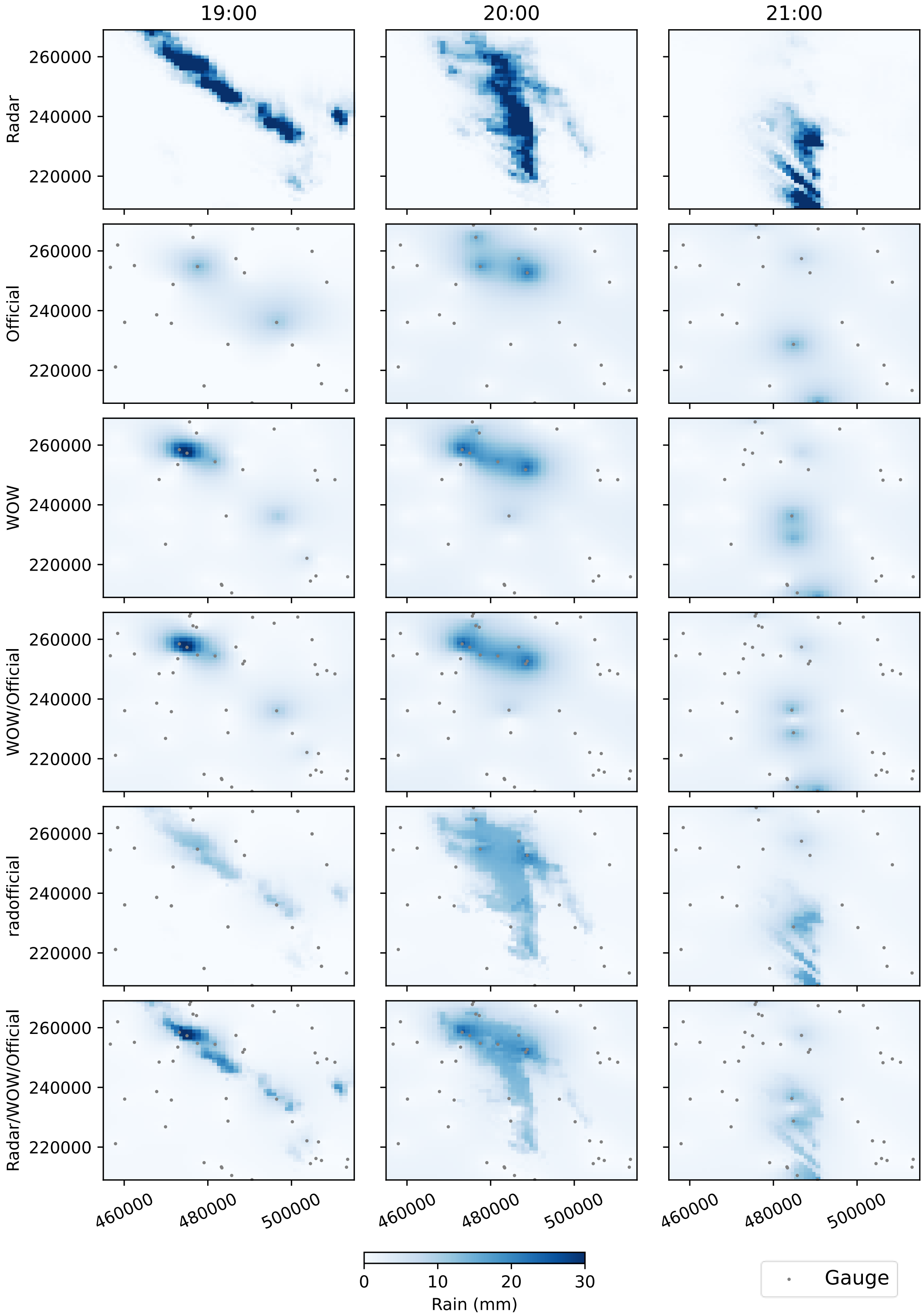


Figure 5.42 Interpolations Milton Keynes, 27/05/2018 19:00 – 21:00.

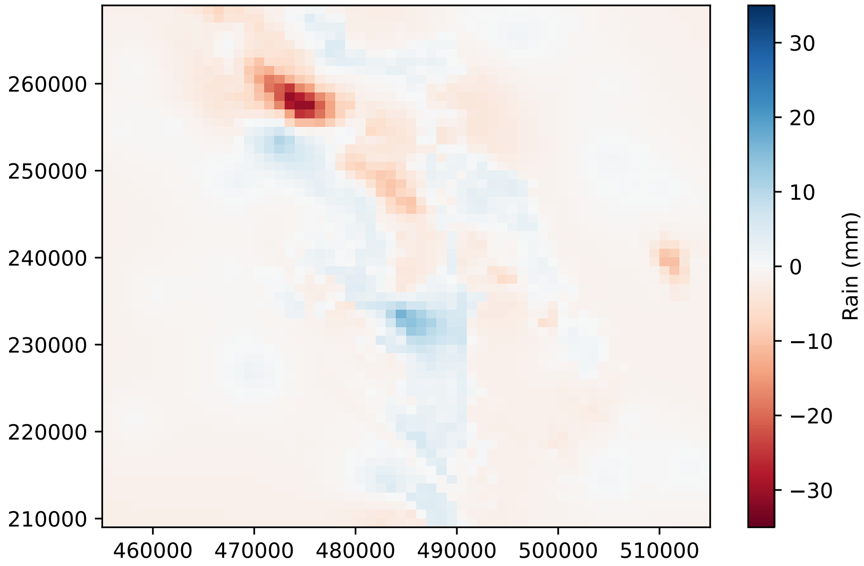


Figure 5.43 Difference plot between radar/WOW/Official - radar/Official interpolation for Milton Keynes (19:00 – 21:00, 27/05/2018).

Figure 5.43 provides the best example of the benefit of including WOW gauge data of the four events used in this study. The northwest area of the plot a WOW gauge captured rainfall that was not seen in Official gauges, resulting in an underestimation of 30 mm where WOW data were not included in the blending.

5.5 Discussion

The manual event-based rain gauge data checking process and cross validation with radar is significant as it demonstrates that the quality of WOW rain gauge data was comparable with data from Official rain gauges. It counters the often-encountered consensus that data from PAWS are of low quality. The demonstration that WOW data can be as good as Official may mean that some data users consider what degree of error is acceptable more broadly, but the methods applied in this research facilitate the validation of both Official and WOW rain gauge data. Whilst other researchers have demonstrated the good quality of PAWS rain gauge data, it has been on a small geographical scale (Starkey *et al.*, 2017) or focused on a particular type of weather station (de Vos *et al.*, 2017; de Vos *et al.*, 2019). This is the first research that has taken a crowdsourcing approach and has proven the quality of PAWS rain gauge data from a wide geographical area. The process of manual quality control was time consuming, and it may be that swifter data provision is desirable for application such as nowcasting, but aspects of the automated and manual SQC could be used in combination with techniques such as applied by Bárdossy *et al.* (2021), where the absolute observation of rainfall depth is less important than understanding the pattern of rainfall.

The selected events represented storms of differing characteristics. The Canvey Island and Milton Keynes events were relatively intense and of short duration with rainfall focused in a relatively small area, whereas the Birmingham and Bar Hill events were less intense. The results show that the spatial extent of most intense/highest rainfall can be small, and therefore there may not be a gauge well positioned to observe the rainfall. This is well documented, having been demonstrated by Schroeer *et al.* (2018) and the implications for hydrological modelling having been considered by Ochoa-Rodriguez *et al.* (2015). Incorporating data from all available sources, including WOW, means there is more chance of having a rainfall gauge in the critical location. The WOW dataset provides relatively easy access to citizen science data, particularly historical data. Other platforms are available that would further increase the number of available gauges for example from providers including Davis Instruments (2018); Weather Underground (2019); Netatmo Weather Map (2022).

For the Canvey Island event, gauges in the area of highest intensity rainfall and the area of flooding were removed from analysis due to suspect observations. It is possible that rain gauges could become inundated during flood events, with pit gauges (favoured for their accuracy in windy conditions) (Pollock *et al.*, 2018) particularly susceptible to flooding. Additionally, issues with data transmission arise as electrical supplies can be affected by flooding (Essex County Council, 2014).

The radar plots show that quality-controlled radar includes errors, despite techniques being available to mitigate poor representation of rainfall (Fornasiero *et al.*, 2006; Krajewski *et al.*, 2010; Libertino *et al.*, 2015). In this analysis beam blockage resulted in an underrepresentation of rainfall during the Milton Keynes and Bar Hill events. It is also likely that radar is overrepresenting rainfall in areas of high intensity, with rainfall depths close to or exceeding gauged hourly records in the Bar Hill and Milton Keynes event.

The analysis presented in this chapter is limited to four events, that may not be representative of all convective storms, additionally there was no consideration given to the meteorological characteristics of the events. Data from more events needs to be assessed to further test the usefulness of PAWS rain gauge contributions. It may also be beneficial to consider other parameters including wind (where available) to estimate undercatch. Given the difficulty in siting wind measuring equipment at the WMO recommend height of 10m (WMO, 2018), and the influence of buildings on wind direction and strength in domestic settings (Bell *et al.*, 2013) such approximations may be unreliable.

The selection of block kriging with external drift for interpolation was based on the availability of data for the selected events (Wang *et al.*, 2013; Ochoa-Rodriguez, 2017). It may be that in other circumstances where the data density is different an alternative method of interpolation may be more appropriate. Sensitivity testing to determine the most appropriate method should be incorporated, with the use of external data sources to validate interpolations, e.g., the relative performance of interpolations when used in hydrological modelling or using social media etc. to cross reference impacted areas. The use of multiple data sources has been the subject of several research studies and reviews (Kutija *et al.*, 2014; Muller *et al.*, 2015; Smith *et al.*, 2017; Yang *et al.*, 2017; See, 2019) and particularly where post event analysis is occurred shortly after the event it would be prudent to use corroborating data including images and social media reports to cross validate where possible.

5.5.1 Response to Research Questions

The aim of the research presented in this chapter was to continue the statistical quality control process by applying manual quality control and cross checking with an external data source (radar) to validate the quality of WOW rainfall observations. This has been achieved by answering the research questions, as summarised here;

Question 1. What manual checking can be used to ensure the accuracy of WOW rain gauge data so it can be used in post-event analysis?

A manual SQC process was devised to highlight common errors encountered in the automatically SQC'd data, including data missing at critical times, the accumulation of rainfall after gaps in the record and anomalous observations (extreme and repeated values). A comparison of the WOW and Official rainfall datasets with respective radar data demonstrated that on completion of the automated and manual SCQ there was no significant difference between the rain gauge datasets, indicating that WOW and Official data are equally valid in terms of quality (see section 5.4.4).

Q2. What impact does including WOW rainfall data have on the interpolation of rainfall during localised high intensity storm events?

A series of rainfall interpolations were created from 4 selected convective rainfall events that resulted in pluvial flooding, using combinations of ground-based rain gauge data from Official and WOW sources, along with radar and an official gridded rainfall data set. Comparisons of the outputs demonstrated that there is a benefit to including the SQC'd WOW observations in creating the rainfall interpolations as they additional data points improve the representation of the rainfall field. Where there are limited data points there is a risk that data from a given location is over-extrapolated resulting in either an under or over-estimate of rainfall. This is only the case for observations that have been subject to automated and manual SCQ checks to remove suspect observations, however this applies equally to both the Official and WOW datasets.

5.5.2 Recommendations

This research provides a method for the manual SQC of previously automatically QC'd rain gauge data. It is recommended that this process is followed to ensure WOW rain gauge data are of sufficiently good quality for use in post-event reconstruction and analysis (see section 5.3.3).

Chapter 6. Case Study - Using WOW Rainfall Data in Hydrological Modelling

6.1 Overview

This Chapter demonstrates the usefulness of WOW rainfall data by presenting a case study using the datasets created in the previous chapter. On completion of manual QC and validation with radar, rainfall generated from differing combinations of Official, WOW and radar data are used in a hydrogeological model to 'sense check' the results, and to ascertain the potential impact of incorporating WOW data into post-event flood analysis. The modelling is undertaken using SHETRAN, a physically based spatially distributed model with data drawn from the previous chapter for a storm that occurred over Birmingham in 2016. The results highlight the sensitivity of hydrological modelling to changes in the rainfall field, demonstrating the importance of using rainfall data that most closely represent reality.

6.2 Introduction

The research presented in the previous chapters has demonstrated there is good spatial coverage of WOW gauges, particularly in urban areas, and that good quality rainfall observations can be extracted from the WOW database. This case study aims to assess the impact of including WOW data in post-event hydrological analysis by using WOW observations in a hydrological model. The following research question is addressed;

Q1. What impact does including WOW rainfall data have on streamflow estimation in a hydrological model?

Rainfall observations are used to provide the starting conditions for rainfall forecasting (as discussed in section 2.2.1) but there are other applications, with flood forecasting and analysis being most relevant to this research. Beven (2012) noted that the accurate representation of precipitation is a fundamental requirement for the generation of numerical hydrological models. This can be challenging, due to the inherent variations in the rainfall field, and the significance of the spatial and temporal variation in rainfall throughout a catchment has been explore by many researchers (Wilson *et al.*, 1979a; Lopes, 1996; Singh, 1997). Outcomes of research commonly support Beven's assertion confirming that peak run off, time to peak and storm run-off are all highly sensitive to precipitation (Seibert *et al.*, 2018).

The selection of a hydrological model is highly dependent factors including the range of parameters to be represented, the data available for input, what outputs are required, and the computational power available. The most simplistic models include unit hydrographs and regression modelling which offer quick results, however for more detailed analysis that allows for sensitivity testing of variation in the rainfall field across a catchment a physically based, spatially distributed model is required. (Bárdossy *et al.*, 2022)demonstrated that for two spatially distributed hydrological models up to 50% or model error can be due to uncertainty in precipitation. The research used SHETRAN, a physically based model and HBV, a conceptual model, and found rainfall to be equally significant in both model types. They suggest that uncertainty in rainfall may be more significant in all types of hydrological model than has previously been acknowledged.

Schilling (1991) presents some key characteristics of rainfall data for the modelling of urban catchments, particularly with respect to urban drainage, stating that a time series spanning

>20 years, with a 1-minute temporal and 1km^2 spatial resolution is desirable. Modelling urban catchments can be challenging due to the presence of impermeable surfaces that reduce infiltration and sewer networks that rapidly convey water to watercourses which can reduce the time to peak, when compared with a rural catchment. {Birkinshaw, 2021 #444@@author-year} used SHETRAN to demonstrate a novel approach to hydrological modelling in urban catchments where there are insufficient records on the subsurface infrastructure and the connectivity of impervious areas that none the less would account for the highly urbanised nature of the catchment. Coincidentally one of the catchments presented in the paper was the Tame at Water Orton, therefore it was deemed appropriate to use the optimised SHETRAN catchment parameters for the hydrological modelling presented in this case study.

6.2.1 Background

There were two distinct extreme rainfall events resulting in pluvial flooding within the Birmingham City Council administrative area on the 8th - 9th June, and 16th June 2016 (see Figure 5.44).

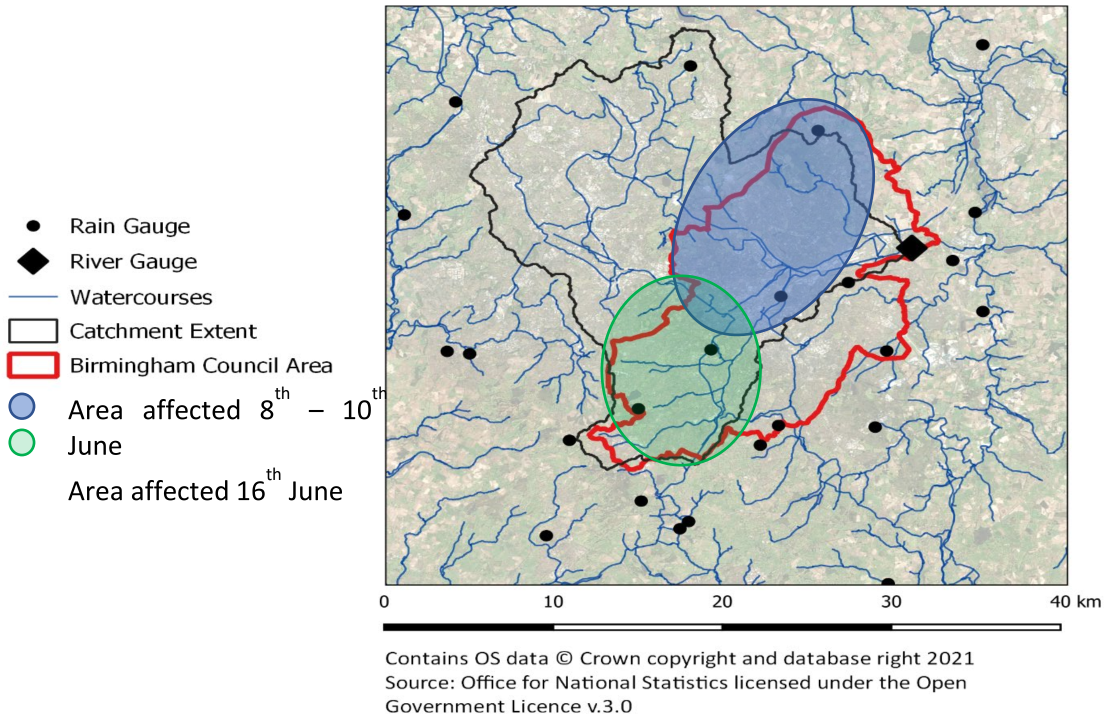


Figure 6.1 Map of rain and river gauges relative to surface water catchment and areas affected by pluvial flooding around Birmingham. Source of affected area: "Birmingham City Council, June 2016 Flooding: Flood and Water Management Act, Section 19 Investigation".

The catchment boundary and Birmingham City Council area are not aligned (see Figure 6.1). The catchment area is the primary focus for this analysis as the fluvial response to extreme rainfall as captured by flow gauging provides a simple metric for assessment, as opposed to the complexity of determining the extent and depth of pluvial flooding (Muthusamy *et al.*, 2019). The location of the river gauge is shown in Figure 6.1 (Centre for Ecology and Hydrology, 2022).

6.2.2 Catchment

The Birmingham catchment is almost fully urbanised and is underlain by superficial deposits that are highly variable in transmissivity from boulder clay to sands and gravels (see Appendix 3 – NFRA Land cover map) (*Ibid*.). Outflow from the catchment is gauged on the river Tame at Water Orton by the EA using a Simple Crump profile weir (ID:28003). The NRFA gauge information states there is high confidence in extrapolations for the gauge, and it is marked as suitable for pooling. The catchment is described as having a fast response. The 2016 annual hydrograph is shown in Appendix 3 - Annual hydrograph, Tame at Water Orton). On 16th June the annual maximum (AMAX) of 114.904 m^3/s was recorded at 23:00 (extrapolated as the flow exceeded the weir maximum capacity of 100.1 m^3/s). As the AMAX was an extrapolation there is a possibility it was inaccurate; however, for the purposes of this research it has been assumed to be an accurate representation of flow.

6.3 Method

SHETRAN was used for 'exploratory modelling' of a single flood event in Birmingham in 2016, in an attempt to demonstrate the value of including WOW data in the creation of rainfall fields and to cross-validate the rainfall fields (produced by interpolation/blending) with river flow (Beven, 2019). Flooding in Birmingham 2016 resulted from multiple days of convective rainfall moving relatively slowly across the catchment (see Table 5.6), providing the opportunity to assess a catchment level response to rain. The model set-up was based on predefined catchment characteristics as detailed in Lewis *et al.* (2018), with some modification of parameters to account for the highly urbanised nature of the selected catchment (presented in Appendix 3) {Birkinshaw, 2021 #444}. The use of a more sensitive urban model, e.g. CityCAT, was considered; however, given that a validated model was available, with modifications made to account for the highly urbanised nature of the catchment, it was considered most suitable. Had further parameters been available, including the drainage network is it possible that CityCAT would be more allow for a more nuanced representation of urban flow, however at given the uncertainties that a more detailed model would introduce the optimised SHETRAN model was deemed acceptable for use in this instance (Glenis *et al.*, 2018).

Simulated outflows generated from a series of rainfall datasets were compared with the observed to establish which rainfall was most accurate. Four rainfall datasets were used:

- Gridded estimates of hourly rainfall (CEH-GEAR1hr) (Lewis *et al.*, 2022) - published data available for the UK;
- Hourly radar (Met Office, 2003) - a widely available data source that is understood to capture the pattern of rainfall but may include inaccuracies in the depth of rain;
- Radar/official blend created during this research – a combination of rain gauge and radar data that may reduce the inaccuracies in the depth of rain as per radar; and
- Radar/official/wow blended dataset created during this research – a combination of all the available Official, radar, and WOW data to assess the relative merit of incorporating citizen science data.

Published CEHGEAR1hr gridded data were used as the primary source of rainfall data; starting in 2010 to provide a run-in period for the hydrological model prior to the period of interest in 2016. Rainfall data from radar and blended datasets for the period before and up to flooding occurring (06/06/16 – 17/06/16) were then substituted into the CEHGEAR1hr dataset to create alternate rainfall fields for the pluvial flooding event itself.

The model outputs were assessed by:

- Visual comparison of simulated and observed hydrographs for the full period and zoomed in onto the peak flow from the event;
- Comparison of the linear regression (r^2) value and Pearson correlation between the simulations and the observed flows (Moriasi *et al.*, 2007); and
- Calculation and comparison of the Nash-Sutcliff efficiency index for each iteration (Nash *et al.*, 1970).

Given the very short duration of the event, there are limitations on the efficacy of the available metrics, therefore time to peak and peak-runoff were considered to be most indicative of rainfall accuracy.

6.3.1 Interpolation and Blend Analysis Methods

There is no accurate independent check dataset for the validation of the blended rainfall data, as all rainfall interpolations include inaccuracies, making it difficult to determine the most accurate combination of radar and gauge data. A series of checks were implemented to enable comparison between the interpolations and blends.

A method to test the value and accuracy of WOW rainfall data was devised using SHETRAN; a physically based, spatially distributed hydrological model (https://research.ncl.ac.uk/shetran/index.htm). This is an alternative approach to that adopted by Glenis *et al* (2018), whereby the outputs of a CityCAT model were cross validated using citizen science (photos) to assess the accuracy of modelling water levels in urban areas.

Descriptive statistics for each blend/interpolation were generated comprised of the:

- Maximum rain total for any grid square (mm)
- Hourly maximum (mm/hr)
- Count of grid squares receiving rain during event
- Total rain across the area (mm)
- Mean rain across the area (mm)

These descriptive statistics provided a measure of how the rainfall depth and extent varied across the area.

The peak rainfall period, consisting of the hours of high rainfall immediately prior to flooding, was selected to analyse the validity of each interpolation and blend. Each interpolation/blend was plotted and visually inspected. Difference plots between interpolations were created for further visual analysis. The plots show the areas where the inclusion of additional gauges changes the interpolations.

6.4 Results

The model was run-in using CEHGEAR hourly gridded data from 2010. The NSE calculated on the daily mean flow excluding an initial run in period was 0.514, exceeding the 0.5 level that is considered a good simulation (Nash *et al.*, 1970; Moriasi *et al.*, 2007). A hydrograph of simulated and observed daily flow is shown in Figure 6.2. In low flow periods the simulated flow is consistently lower than the observed, possibly due to urban drainage not being well represented by the model (Coutu *et al.*, 2012; Birkinshaw *et al.*, 2021).

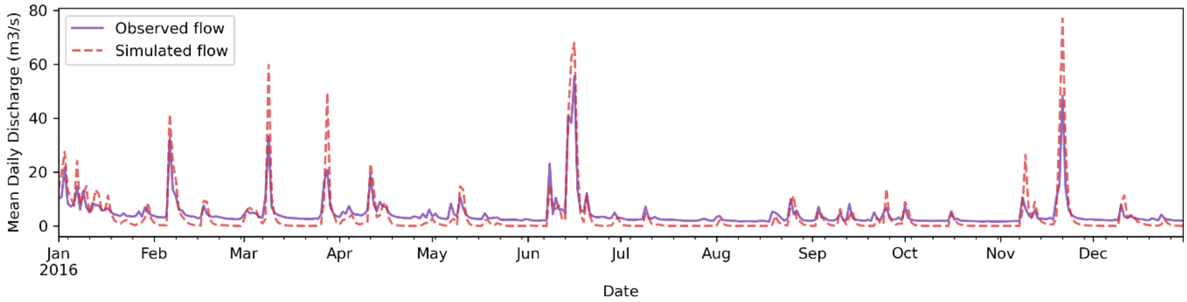


Figure 6.2 Hydrograph of simulated versus observed daily flow at Birmingham gauge 28003 for 2016.

The simulated outflows from the various rainfall data and the observed flow were plotted for visual comparison. The hyetograph and distribution of the total rainfall for the period is shown in Figure 6.3, along with the hydrograph.

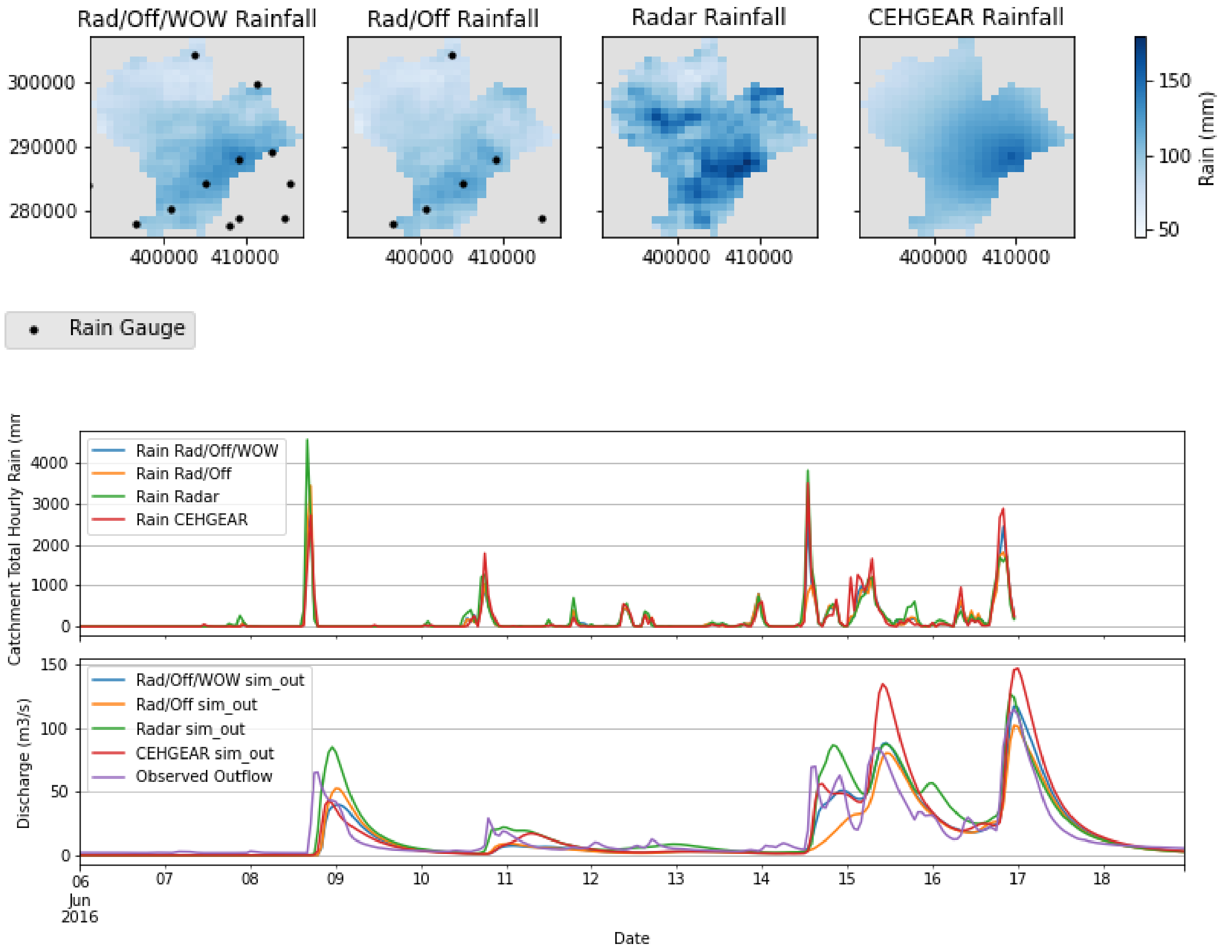


Figure 6.3 Plots of sum of rainfall from radar/official/wow blend, radar/official blend, radar data and CEHGEAR-1hr with hydrograph of observed and simulated flow at river gauge 28003, Birmingham 06/06/2016 - 17/06/2016.

A lag between the observed and simulated QMAX on the 8th, 14th and 15th of June is visible in Figure 5.46; however, the plot shows close correlation of the QMAX timing on the 16th of June. This is assumed to be due to the time it takes for the 'wetting' of the catchment during simulation, which could normally be addressed by amending model parameters, but is reported here to be faithful to the comparison between rainfall datasets. A series of statistical tests were performed to compare the simulated outflow to the observed and are presented in Table 6.1. The NSE calculated on the short time span of the event is an improvement over that of the model as a whole (0.51).

Measure	Radar/Official/WOW	Radar/Official	Radar	CEHGEAR
R^2	0.75	0.68	0.65	0.60
Pearson	0.89	0.84	0.89	0.90
NSE	0.75	0.68	0.65	0.60

Table 6.1 Statistical comparison of simulations from CEHGEAR, Blended and Radar rainfall datasets, Birmingham 06/06/2016 - 17/06/2016.

The improvement in the accuracy of the rainfall from CEHGEAR-1hr, to radar, then radar/official blend, and finally the radar/official/wow blend is demonstrated by the results in Table 6.1. The metrics show that overall, the highest model efficiency and linear regression were achieved using the radar/official/wow blend of data. The Pearson correlation was marginally higher for the simulation using the CEHGEAR-1hr dataset; however, the Pearson correlation does not account for differences in the magnitude between the simulated and observed flow.

Peak Analysis

According to the radar, rainfall resulting in flooding on 8th/9th June was highest to the west and central part of the catchment where no WOW or Official gauges were located. This resulted the simulated peak outflow for the blended data being lower than observed (Figure 6.4 and Table 6.2). Conversely the QMAX from the radar simulation exceeded the observed, indicating that the radar was likely overreporting rainfall.

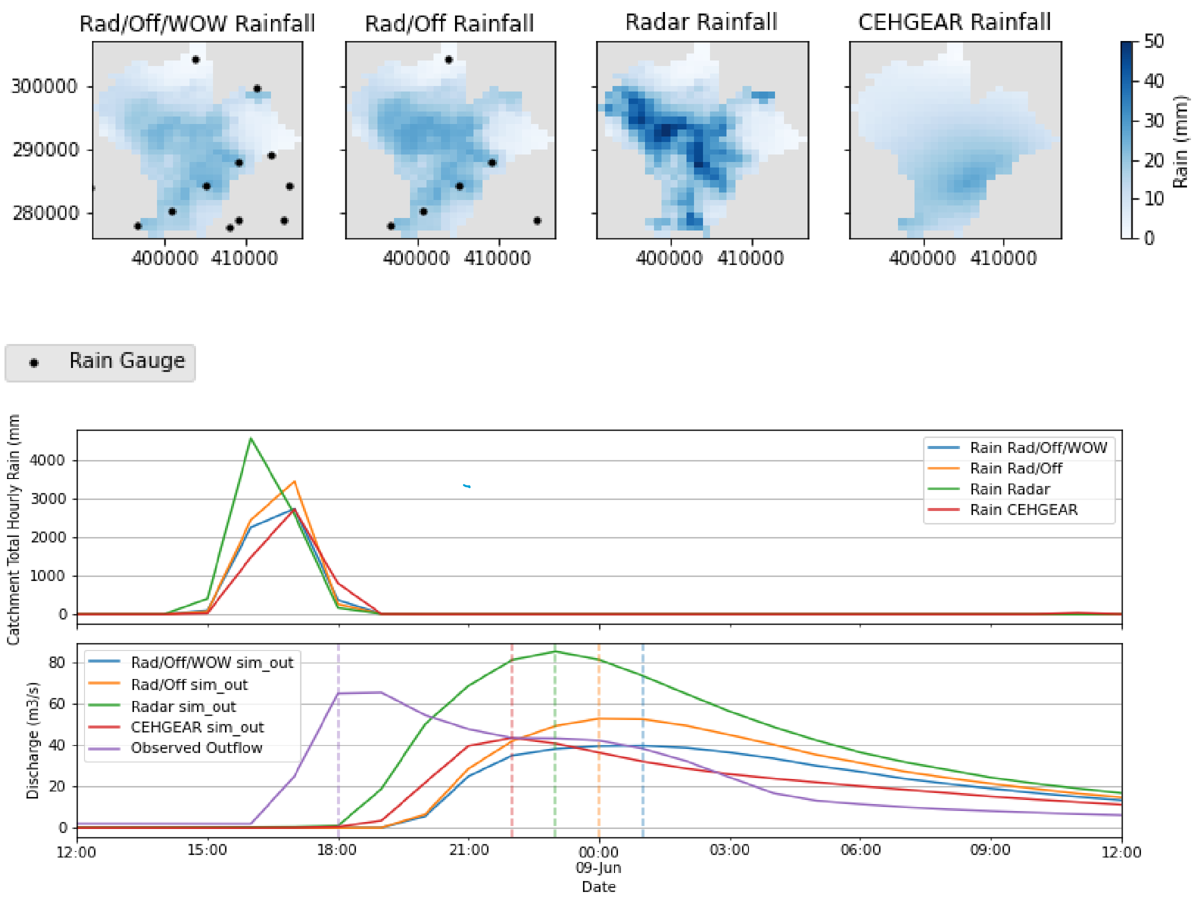


Figure 6.4 Plots of Rainfall from CEHGEAR-1hr, blended and radar data with hydrograph of observed and simulated flow with peaks shown as dashed lines (08/06/16 - 09/06/16).

The peak rainfall occurred at 16:00 in the radar, and at 17:00 in the remaining rainfall datasets. There was a lag between the observed QMAX (at 18:00) and the simulated QMAX, with the CEHGEAR-1hr dataset being least delayed (at 22:00) and the radar/Official/WOW blend data being the most delayed (01:00). Possible reasons for the delayed simulated response include:

- Poor representation of out of channel overland flow
- Hourly aggregation of rainfall amplifying slight mismatches in rainfall timing
- Radar accuracy in geographical location of rainfall

The magnitude and variation between the simulated QMAX and observed on 8th/9th June are presented in Table 6.2.

QMAX	Value	Variation simulation – observed (m^3/s)	Variation simulation – observed (%)
Observed	64.78	0	0
Radar/Official/WOW	39.60	-25.18	-38.87
Radar/Official	52.66	-12.12	-18.70
Radar	84.95	20.17	31.14
CEHGEAR	43.32	-21.46	-33.13

Table 6.2 QMAX 8th/9th June, Observed and Simulated

According to the radar and the blended datasets the rainfall resulting in flooding on the 16th/17th June was highest in the south of the catchment, close to both Official and WOW gauges. The rainfall simulation derived from the radar/Official/WOW blend most closely matches the observed QMAX magnitude and timing on the 16th of June (see Figure 6.5).

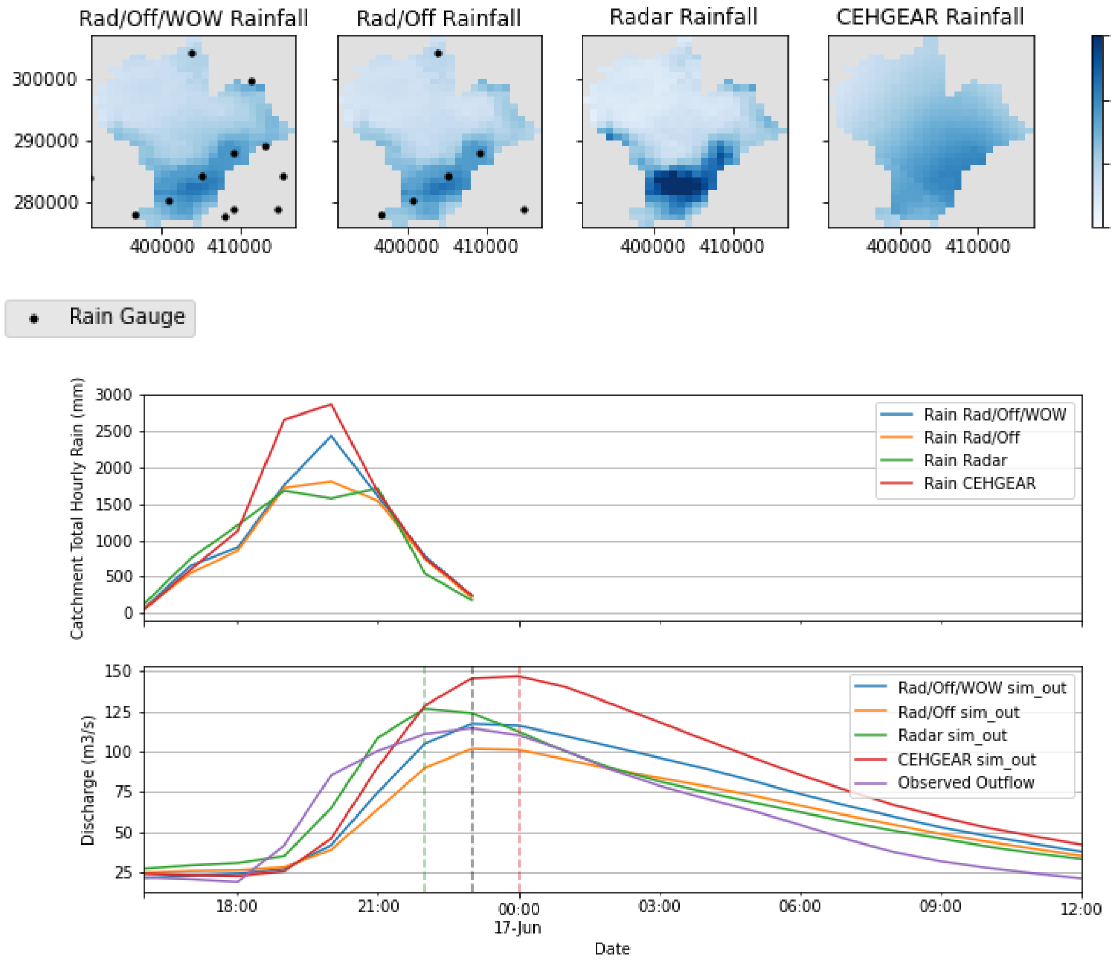


Figure 6.5 Plots of Rainfall from radar/Official/WOW blend, radar/Official blend, radar and CEHGEAR-1hr data with hydrograph of observed and simulated flow with peaks shown as dashed lines (16/06/16 - 17/06/16) – note the observed and blended QMAX occurred at 23:00.

The peak rainfall occurred at 21:00 in the radar, and at 20:00 in the remaining rainfall datasets. The observed QMAX and that simulated from the blended data occurred at 23:00 on the 16th of June, the radar simulation QMAX occurred at 22:00 and the CEHGEAR QMAX was at midnight on the 17th of June. The magnitude and variation of the QMAX from each simulation and as observed on 16th/17th June are presented in Figure 6.3.

QMAX	Value	Variation simulation – observed (m^3/s)	Variation simulation – observed (%)
Observed	114.5	0	0
Radar/Official/WOW	117.1	2.64	2.31
Radar/Official	101.8	-12.7	-11.1
Radar	126.6	12.14	10.60
CEHGEAR	146.8	32.27	28.18

Table 6.3 QMAX 16th/17th June, observed and simulated flows.

The QMAX from the radar/Official/WOW simulation was closest to the observed (2.31% variation) making it the most accurate on 16th June when timing and magnitude were considered.

6.5 Discussion

The importance of rainfall in hydrological modelling was evidenced in this chapter and has been demonstrated in other research. Bárdossy *et al.* (2022) established that up to 50% of hydrological model error could be attributed to uncertainty in rainfall representation. In this research a variation of 2% - 28% between simulated and observed flows confirms the importance of using data representative of rainfall, which is highly dependent on the location of gauges relative to rainfall, and the accuracy of radar data. This analysis reiterates that extrapolation of point data during convective events increases in accuracy as the density of rain gauges increases (Peleg *et al.*, 2013; Ochoa-Rodriguez *et al.*, 2015; Schroeer *et al.*, 2018). The modelling of the two peak rainfall events demonstrated that the simulated outflow using the blended datasets was more accurate when rainfall occurred in areas where it was captured by rain gauges. The WOW rain gauge data increased the accuracy of the rainfall field, by increasing the chance of capturing high intensity rainfall and by preventing the over extrapolation of rainfall.

The radar dataset was the most reliable in areas where there were no rain gauges. In this instance, when the peak rainfall occurred in an ungauged area, an alternative method of blending i.e., Bayesian may be better (Todini, 2001; Ochoa-Rodriguez *et al.*, 2015). This analysis emphasises the interpolation method needs to be selected on a case-by-case basis choosing the tools most appropriate to the resolution of data (temporal and spatial) and the application (Wang *et al.*, 2013).

The temporal scale of data used in the generation of interpolations was hourly, which is generally considered too coarse for the assessment of flooding in urban catchments (Ochoa-Rodriguez *et al.*, 2015). SQC can be undertaken on shorter time interval data (Villalobos-Herrera *et al.*, 2022), therefore an improvement to this analysis could be to use 15 minute observations, where available, to coincide with radar at a time interval that allows for some discrepancy in reporting between the two datasets (Krajewski *et al.*, 2010). As not all WOW rainfall data was available at sub-hourly intervals an alternative may be to downscale data to the desired time interval.

The analysis presented in this chapter is limited to four events, that may not be representative of all convective storms, additionally there was no consideration given to the meteorological characteristics of the events. Data from more events needs to be assessed to further test the usefulness of PAWS rain gauge contributions. It may also be beneficial to consider other parameters including wind (where available) to estimate undercatch. Given the difficulty in siting wind measuring equipment at the WMO recommend height of 10m (WMO, 2018), and the influence of buildings on wind direction and strength in domestic settings (Bell *et al.*, 2013) such approximations may be unreliable.

The selection of block kriging with external drift for interpolation was based on the availability of data for the selected events (Wang *et al.*, 2013; Ochoa-Rodriguez, 2017). It may be that in other circumstances where the data density is different an alternative method of interpolation may be more appropriate. Sensitivity testing to determine the most appropriate method should be incorporated, with the use of external data sources to validate interpolations, e.g., the relative performance of interpolations when used in hydrological modelling or using social media etc. to cross reference impacted areas. The use of multiple data sources has been the subject of several research studies and reviews (Kutija *et al.*, 2014; Muller *et al.*, 2015; Smith *et al.*, 2017; Yang *et al.*, 2017; See, 2019) and particularly where post event analysis is occurred shortly after the event it would be prudent to use corroborating data including images and social media reports to cross validate where possible.

The approach presented using SHETRAN and hourly data may not be the most appropriate for the analysis of convective storms (Birkinshaw *et al.*, 2021; Bárdossy *et al.*, 2022). For future analysis, a model optimised for flashy events in urban areas may be better suited than one designed for the analysis of multi-day events (Guerreiro *et al.*, 2017; Glenis *et al.*, 2018). This requires rainfall data at a higher temporal resolution (minimum. 15 minutes) than was available for this research and extensive detail on the subsurface infrastructure present in the catchment {Birkinshaw, 2021 #444}.

6.5.1 Response to Research Question

The aim of this case study was to assess the impact of the inclusion of WOW data in post-event hydrological analysis using a hydrological model. This was achieved by addressing the research question;

Q1. What impact does including WOW rainfall data have on streamflow estimation in a hydrological model?

The modelling case study demonstrated that WOW rain gauge data are valuable in the post-event analysis of rainfall. Including the WOW rain gauge data in the post-event analysis of convective storms improved the representation of the spatial distribution of rainfall when only gauge data were available. Where radar data was also available for inclusion in interpolations the WOW data added quantitative data to correct any bias in the radar as displayed in the interpolations (see section 5.4.7). The case study showed that in this example the interpolation incorporating the WOW, Official and radar rainfall data generated simulated flows most closely matching the observed, suggesting this combination was closest to representing the actual rainfall.

The importance of rainfall in accurate hydrological modelling has been identified (Bárdossy *et al.*, 2022), and was demonstrated during this research. Point rain gauge data from a limited number of Official gauges, or poorly corrected radar will not provide as accurate a representation of rainfall as could be generated when all rainfall data sources are incorporation into rainfall field interpolations.

6.5.2 Recommendations

On the basis of the comparison of interpolations and the improved simulation of river flow in hydrological modelling (see section 5.4.7) the use of quality controlled WOW data, where available, is recommended to improve the spatial and/or quantification of the rainfall field. Irrespective of whether WOW data are available; Official rain gauge or radar should be incorporated into interpolations for provide the best available representation of rainfall.

Chapter 7. Observing the Observers – Motivations and Barriers to Sharing Data from Private Automated Weather Stations

7.1 Overview

The research presented in this thesis has demonstrated potential benefits, albeit not yet fully realised, of crowdsourced data available via WOW (and other platforms). In this chapter research is presented providing insight into the motivations and barriers for citizen science participation. The aim was to document the experience of citizen scientists in recording weather observations, considering the heterogeneity of participants regarding skills, knowledge and commitment levels to highlight motivations and barriers to participation.

Three different engagement approaches were adopted, with participants including members of the public who had never previously encountered citizen science rainfall monitoring, a school group, and academics familiar with using Official rainfall data. Additionally, insight was sought from correspondence with members of the Climatological Observers Link (COL) special interest group. The analysis showed that self-selection, a peer support network, and timely data validation with feedback were important factors in assuring sustainable participation and the generation of good quality rainfall data.

7.2 Introduction

In Britain there are rural areas where there may not be an Official rain gauge for many kilometres, catchments with no rain gauge, or, where the density in urban areas falls far short of the high spatial density of observations that is required for hydrological applications (Ochoa-Rodriguez *et al.*, 2015). It follows therefore that where there are gaps in the Official network, it would be beneficial to fill those gaps. The expense of installing and maintaining an Official rain gauge must be borne by agencies and authorities; therefore, there is an attraction to 'free' rain observations from PAWS. As this and other research has demonstrated that such data can be of acceptable quality (de Vos *et al.*, 2019; Bárdossy *et al.*, 2021), promoting participation in collecting weather data seems an obvious solution to gap filling in monitoring networks. This chapter sees to document the experience of citizen scientists in recording weather observations, considering the heterogeneity of participants regarding skills, knowledge and commitment levels to highlight motivations and barriers to participation, and answer the research questions:

Q1. What themes emerge that encourage or reduce participation in crowd sourcing weather observations, and are they common between groups of potential or active citizen scientists?

Q2. Can key elements that prevent or support the generation of good quality rainfall data via citizen scientists be identified?

7.3 Method

Three approaches were made during this research to encourage/assess participation in weather observing:

1. Community outreach activities hosted by two family focused museums in Newcastle.
2. Engagement with local teachers from Tyne and Wear and Northumberland.
3. Provision of weather stations to researchers from Newcastle University.

In addition, data sought from COL members was also confirmed as high quality (see section 4.5), the modality of COL has also been assessed to see if lessons can be learned from their approach.

7.3.1 Community Outreach

Four workshops were hosted at the Great North Museum (https://greatnorthmuseum.org.uk) and the Centre for Life (www.life.org.uk/life/science-centre) in Newcastle over the school holidays to share information on the risks of flash flooding, natural flood attenuation, checking rainfall data online, operating a weather station and interpreting rain gauge data. Games illustrating the key concepts were made available, along with posters, demonstrations, and discussion with participating climate scientists from Newcastle University (see Appendix 4). Participants were directed to a website (https://research.ncl.ac.uk/cspaws/) and a support email address that we have developed for this project to support citizen scientists with any issues around installing a weather station and contributing their data to the Met Office Weather Observations Website (Met Office, 2011)

Two Netatmo weather stations with rain gauges were given out as prizes in a 'names from the hat' draw. The recipients were asked to complete a questionnaire shortly after receiving the weather station (see Appendix 4) and were offered support and advice via email. The responses to the questionnaire were assessed to determine how comfortable people were with the installation and set-up of the weather stations.

7.3.2 Schools and Special Interest Group Engagement

A presentation was given to 42 teachers in the Northeast to encourage schools to get involved with weather observations. There was a stipulation that equipment would not be provided by the researchers, but there was signposting to available funding (via the Royal Society) and an offer of support for writing a grant request and subsequent implementation of weather-related activities.

Contact was made with Environment Agency Community Liaison Officers for flooding for the Northeast Region. The liaison officers support local community groups, often formed in response to flooding or flood risk. A letter of invitation was sent to multiple groups via the liaison officer, offering support if there was any interest in the groups participating in weather observation activities (see Appendix 4). Offers of support/invites to participate were sent to other local and flood impacted groups via Tyne Rivers Trust, EA, NFF, Northumbria Water and the Water Group.

A talk was given to a citizen science forum group hosted by Newcastle University, with participants drawn from neighbouring academic institutions, the local councils and special interest groups. An activity was devised to demonstrate how unwieldy citizen science requirements could disincentivise potential participants (see Appendix 4).

The take-up rate was assessed, along with the degree of involvement with weather monitoring.

7.3.3 Weather Station Provision

Six volunteers from the School of Civil Engineering at Newcastle University living within 10 km of the city centre were provided with Netatmo weather stations with rain gauges. The study group was selected on the assumption that professionals working with meteorological data would be likely to have an interest in how those data were collected and may be more motivated than the general public to host a weather station.

The volunteers were provided with information to support the set-up of the provided equipment (see Appendix 4), and access to ongoing email support. They were asked to make data publicly available via the Netatmo weather website (Netatmo Weather Map, 2022), and connect to WOW via a third-party app recommended by the Met Office for that purpose. Volunteers were asked to complete the station meta data and were provided a guide to aid them (see Appendix 4).

Shortly after receiving the weather stations a questionnaire was shared for completion by participants, to gauge the ease of set-up and highlight any issues with the rain gauges (see Appendix 4). After a period of three months, data were requested via email, to determine whether gauges were active and generating 'sensible' results. Additionally, photographs were requested of the gauge set-up and a further questionnaire was deployed. Finally, after over one year of operation a questionnaire was sent to assess the success of the weather station deployment. The qualitative assessment of the questionnaires was conducted to determine the level of engagement, and the barriers/drivers for participation in using a PAWS.

7.3.4 Engaging with a Special Interest Group (COL)

Data were requested from COL members wishing to participate in the research. Members were asked to share hourly or shorter time-interval data from PAWS in a digital format, or to provide a WOW ID. The assessment of data quality is presented in detail in section 4.5. In addition to the data several COL members engaged in email correspondence with the researcher. The results of data analysis and discussion of the correspondence is considered to determine whether lessons can be learned from the COL approach regarding encouraging participation and the provision of good quality rainfall data.

7.4 Results

7.4.1 Community Outreach

Over 800 people interacted with the games and demonstrations over 4 days of engagement at Newcastle museums. The degree of interest and participation with the activities available on the day was high, with groups/individuals trying multiple activities and speaking with the three Newcastle University researchers on hand. The live Netatmo weather map was demonstrated, which was of great interest to many who sought out the weather near to their homes or places they had visited in the past. Participants were surprised at the availability of real-time reports and the number of weathers stations sharing their data.

The cost of weather stations was cited as a reason for not collecting rainfall data, along with not knowing it was possible. The workshops raised awareness amongst people who had not previously considered participating.

Two Netatmo weather stations were awarded as prizes out of 246 people who entered the draw. The weather stations were posted to the winners with a questionnaire comprising a 5-point Likert scale for a series of statements followed by 4 free text questions. A short interview was conducted with each winner to gather their responses (see Table 7.1). Winner 1 was a family with resident school-age children and winner 2 was a grandparent with non-resident school age grandchildren.

Statement	Winner 1	Winner 2
1. The weather station is really cool	Strongly agree	Agree
2. We have wanted a weather station for a while	Strongly disagree	Neutral
3. We just entered the competition; we hadn't thought about getting a weather station before!	Strongly agree	Disagree
4. The outdoor module was easy to set up	Agree	Agree
5. The rain gauge was easy to set up	Agree	Agree
6. We regularly check what the weather is doing on the app	Agree	Neutral
7. We compare our results to those of Netatmo units near by	Agree	Neutral
8. We pay more attention to the weather forecast now we can check what really happens	Agree	Neutral
9. We look at results for weather stations in other places	Agree	Agree
10. We would recommend getting a weather station to other people	Neutral	Neutral
11. At £175 the weather station is good value	Disagree	Disagree

Table 7.1 Netatmo winner questionnaire responses (colour coded blue for positive and orange for negative).

The responses to the statements were generally positive, with both winners agreeing the at the weather stations were interesting and easy to set up. Winner 2 was less enthusiastic than winner 1 about the use of the weather station and checking data, but as can be seen in Table 7.2, winner 2 did not have a home computer and was not comfortable using a smart phone, which was detrimental to accessing the weather map, etc. Neither winner indicated that they would recommend the weather station to others, possibly influenced by the price as they both disagreed that the weather station was good value. Three further questions were asked about the weather stations and the sessions the winners had attended at the museums (see Table 7.2). The winners said that they enjoyed the museum activities, and whilst winner 2 was aware of weather stations prior to the sessions, it was a new concept for winner 1. Winner 2 found out more about weather stations around the world so expanded their knowledge and understanding of how weather data are gathered.

Question	Winner 1	Winner 2
1. Do you have any comments on the weather station?	We didn't install the weather station for 2 months as we were moving house. The weather station looks nice and was easy to install. The Wi-Fi connection took a few goes. My son likes checking the other weather stations around our area, and when we are away from home.	My grandchildren set the weather station up and use it the most when they visit. They keep an eye on the weather here using the map online. We don't have a computer at home and I don't like using the phone so we look at it more when the grandkids visit. We had looked at some before when they did a school project but thought they were a bit expensive.
2. Was there anything at the museum demo session that you or your family really enjoyed?	My son enjoyed the virtual reality demo. The kerplunk game was good.	The kids enjoyed the games and the youngest did some colouring.
3. Did you learn anything new or interesting during the session?	We didn't really know much about weather stations so that was interesting to see the equipment.	We learned about all the weather stations in the world on the map, we had no idea there were so many.

Table 7.2 Free text responses to weather station winner's questionnaire.

The Netatmo weather map was checked in February 2022, and both rain gauges appeared to be still reporting and sharing data (although owner details are not visible the locations were consistent with the winners). The locations of the stations are not shared here for privacy.

7.4.2 Schools Engagement

Several schools expressed an interest in seeking funding to purchase weather stations, and support for grant applications was provided to three. Only one school was actively engaged in weather observations prior to lock-down due to Covid-19. There was not further communication from the two that were awaiting confirmation of funding, so it is presumed that either funding was not granted, or the priority of the activity was reduced.

To begin the engagement with Corbridge Middle School, a workshop was held with 25 children aged 9 – 13 years old to look in detail at how weather stations work, and the impacts of wild weather globally and locally. Three researchers from Newcastle University presented in an assembly for the whole school. On completion of the assembly, 2 staff members with responsibility for geography and science continued to work with pupils in the 'Weather and Climate Club'. The school was very active in promoting weather observation activities and were able to secure funding to purchase 2 weather stations: 1 Davis Vantage Pro and 1 Netatmo. Staff took responsibility for installing and operating the weather stations. Teachers provided feedback saying they had some difficulty finding a suitable location on the school grounds where the station would be secure whilst being sufficiently exposed. Technical assistance was provided in the analysis of data, which highlighted that the Netatmo station (which had been placed on a flat roof for safety) had fallen over, with no observations reported for >1 month. It was difficult on the school grounds to balance the competing need for a safe and secure location, with connectivity, exposure, and access for maintenance.

Technical support was provided to the school for the interpretation of data and a further remote meeting was arranged where pupils could ask questions about the weather stations. As an exercise to validate the data being generated by the weather stations at the school the daily rainfall was compared between the two stations and the nearest official monitoring station located approximately 1km away at Maften. Data were compared from 1/1/2020 to 31/1/2020 (Met Office, 2006b). The results were shared with the teachers and pupils at the school to demonstrate the accuracy of their monitoring.

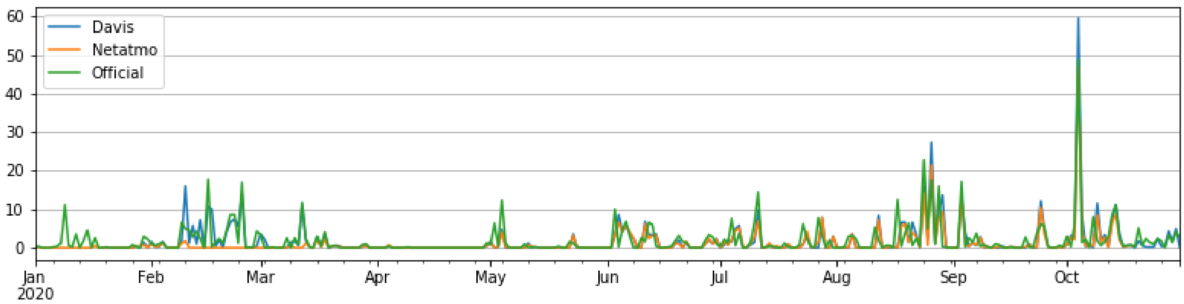


Figure 7.1 Daily rainfall from gauges at Corbridge Middle School and Maften.

Figure 7.1 shows that the observations from the Davis gauge were consistent with the Official rain gauge throughout the data analysis period.

Gauge	Davis	Netatmo	Official
Davis	1	0.92	0.82
Netatmo	0.98	1	0.76
Official	0.84	0.84	1

Table 7.3 Pearson correlation between Corbridge Rain Gauge Daily Observations (blue = full period, green = March onwards after Netatmo gauge was stood upright).

The Pearson correlation between gauges shown in Table 7.3 were all strong and were improved once the issue with the Netatmo gauge was rectified in March.

Observations continued throughout 2020-2021. Pupils reported their activities to the Royal Society, who provided the initial funding for the weather station, and participated in a Student Conference in February 2022. Communication with the school indicated that other activities were being pursued, and funding had been awarded to look at sustainable urban drainage. Although the focus of the Weather and Climate Club has moved on, the school continues to engage with Newcastle University on related activities. The success of the participation with Corbridge Middle School's involvement with weather observation really stemmed from the incredible efforts of the club leader. The enthusiasm and persistence were notable, and the staff member was active in the promotion of applied science in numerous ways.

7.4.3 Weather Stations Provided

There were 6 weather stations given to volunteers. To preserve the anonymity of the volunteers the locations are described rather than mapped (see Table 7.4).

Weather Station	Location	Relative position
1	Tynemouth	10km east of Newcastle University
2	Heaton	2km east of Newcastle University
3	Gosforth	3km north of Newcastle University
4	Jesmond	2km northeast of Newcastle University
5	Gosforth	3km north of Newcastle University
6	Gosforth	3km north of Newcastle University

Table 7.4 Locations of Netatmo weather stations hosted by volunteers.

The first survey was completed by the volunteers in March 2023, within 1 month of receiving their weather stations. The first section consisted of 8 statements with 3 options on a Likert scale (see Figure 7.2).

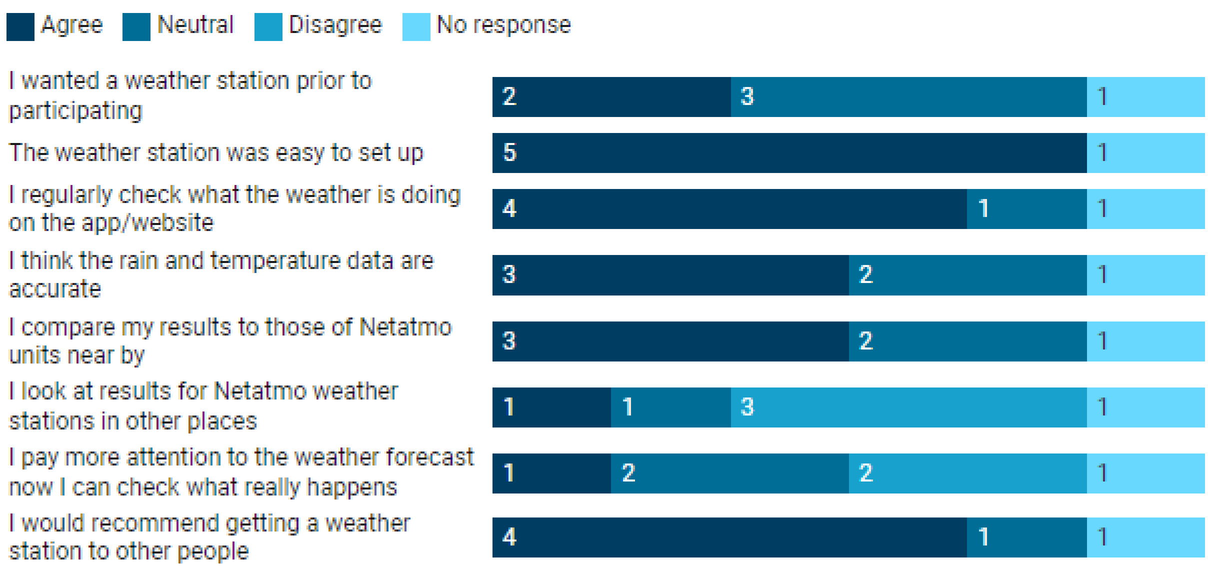


Figure 7.2 Responses from volunteers hosting weather station immediately after receipt.

Responses were received from 5 of the 6 (no response from volunteer 6). The results were generally positive, with all the respondents saying they found the unit easy to set up and 4/5 saying they would recommend a weather station to others.

In addition to the Likert scale, respondents were asked if there was anything they especially liked or disliked about the weather station and were invited to add further comments. The positives included:

- The simplicity of the set-up
- The easy to use and clear web interface and the unit appearance
- The additional indoor monitoring of CO_2 and noise
- Having local weather data

The negatives included:

- Not been given sufficient information on where to install the units for best performance
- Finding it hard to identify a suitable location for installation in a domestic property outdoor space
- No automated comparison facility with Official monitoring
- Not having any notification that the connection was lost or batteries low
- The rain gauge being unheated so not able to accurately record snowfall

- The limited guidance on connecting the data stream to additional platforms like WOW etc.

The positive comments suggested that the volunteers were happy with the ease of using the weather station. It was interesting to note that some of the additional features including CO_2 and noise monitoring were appealing to volunteers, perhaps broadening the appeal of the weather station. The negative comments were indicative of the knowledge of the volunteer cohort, given that they were all familiar with weather observations they were aware of the difficulties and limitations of siting the equipment in a domestic setting (see Figure 7.3) and the lack of heating element for snow observations. The ability to compare data with the nearest Official monitoring location is offered by WOW but is not a feature of the Netatmo weather map. Only one volunteer connected the weather station to WOW, as the others found the process too cumbersome, and they did not see the benefit.

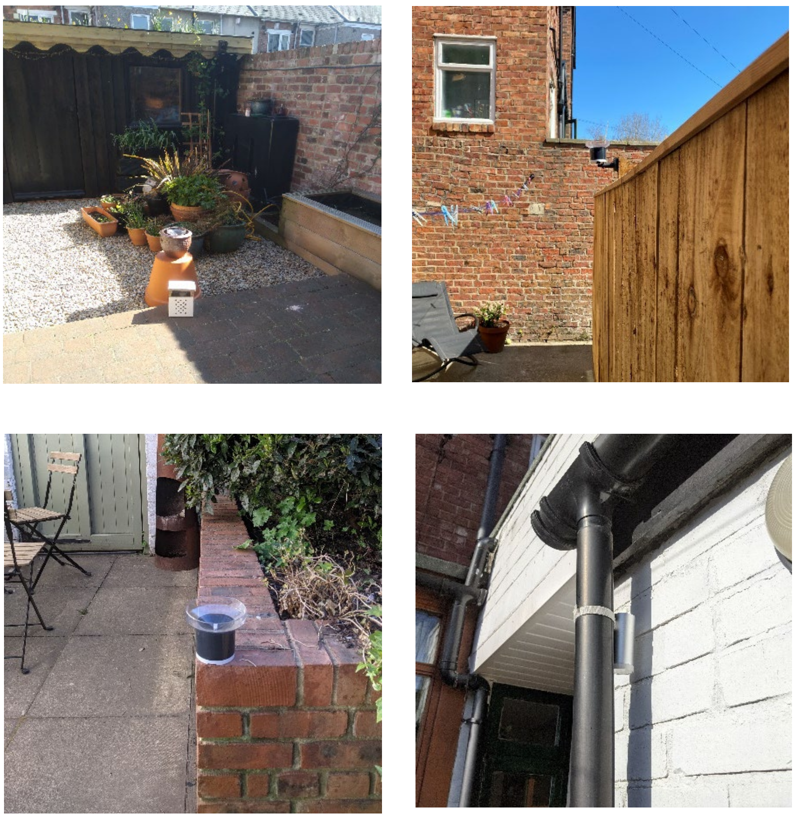


Figure 7.3 Netatmo rain and temperature modules in situ

In May 2021 a second survey was deployed. The first two questions were modified to capture the experience of operating the weather station, whilst the remaining questions were the same as the first survey (see Figure 7.4).

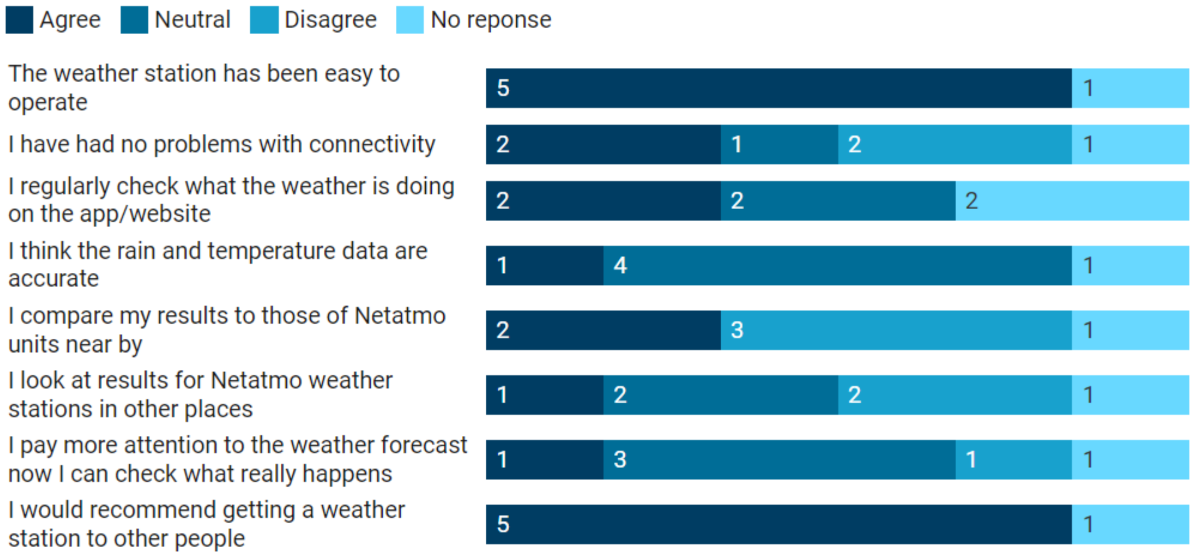


Figure 7.4 Responses from volunteers 6 months after receipt of the weather station

One person (volunteer 1) did not respond. Again, the results were generally positive, with all the respondents saying the weather station was easy to operate and that they would recommend it to others. In addition to the Likert scale respondents were asked some free text questions, which are presented below with summarised compile responses.

1. **Any issues operating the weather station or accessing data?**

Loss of connection, issues with syncing the rain gauge, limited download to previous 3 months, not able to download data from all modules at once, the need for monitoring blockages. One volunteer mentioned that later in the year the sun was catching the temperature module leading to temperature spikes.

2. **How have you used the weather information from the station?**

Comparing against forecasts, to quantify observed rainfall, to regulate home temperature, for fun – engaging kids in learning.

3. **Do you have any advice for other weather station operators?**

Consider where the sun will reach when selecting a position for the temperature module

4. **What would you like to see from the data your weather station has gathered?**

Summary metrics of monthly/annual rainfall, comparisons with seasonal and annual trends and with nearby weather stations, comparisons with official monitoring, how rainfall varies across the city during high intensity events.

Rainfall and temperature data were requested from the weather stations to allow for comparative analysis, and to establish if the weather stations were operational. Data were shared from 5 of the 6 weather stations, however one had only been operational for a few weeks and had not recorded any rain, therefore the results have been omitted from the following plots. The data were aggregated to daily and compared with the nearest Official Met Office rain gauge (Met Office, 2006b), located at Albermarle Barracks approximately 18 km west of Gosforth/Jesmond where the rain gauges with data were hosted.

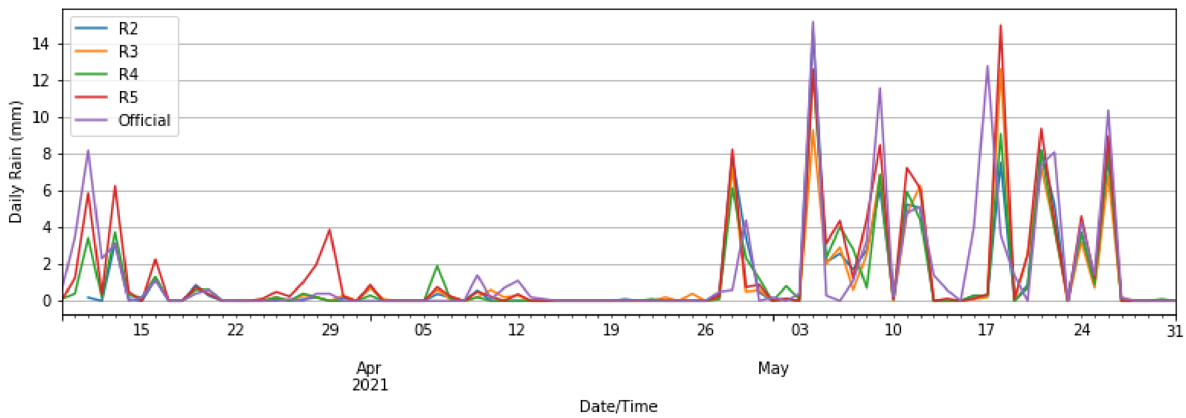


Figure 7.5 Plot of daily rainfall from Netatmo weather stations and closest official rain gauge, March - May 2021.

Figure 7.5shows the consistency in observed rainfall between the Netatmo rain gauges and the nearest official rain gauge. The Pearson correlations between the official gauge and the respective Netatmo rain gauges were 0.75, 0.67, 0.75 and 0.68 (R2, R3, R4, R5). There were a series of heavy rainfall events in May that were captured by the Netatmo rain gauges (see Figure 6.6).

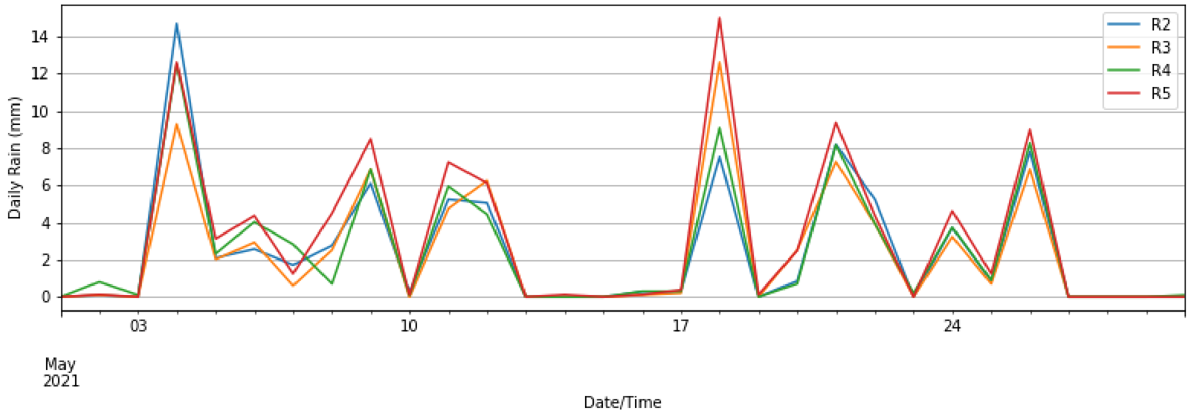


Figure 7.6 Plot of daily rainfall from Netatmo weather stations, May 2021.

Figure 7.6 shows that the magnitude of rain was inconsistent between gauges on several days, with a variation of up to 6mm. The Netatmo weather stations reported rain >5mm on 8 days in May. The highest rainfall was observed at R5 for all events except on 3rd May. Rain gauges R3 and R5 were located within a few hundred meters of each other, therefore the difference in the rain depth is pronounced. The results suggest that rain observations derived from Netatmo rain gauges may not be accurate. This could be a function of the gauge itself, although previous studies showed them to be accurate (de Vos *et al.*, 2017), or may be a function of the difficulty in siting a rain gauge in an appropriate position in a residential setting. Had data been available from R6, which was also within a few hundred meters of R3 and R5 it may have been possible to cross validate between the gauges.

In May 2022, approximately 18 months after the provision of the rain gauges, a final survey was deployed (see Figure 7.7). Responses were received from 5/6, with volunteer 1 not responding to surveys 2 or 3. The volunteers were asked if their weather stations were currently operational, with only volunteer 6 saying their weather station was. The same 8 questions were asked in survey 3 as in 1 and 2.

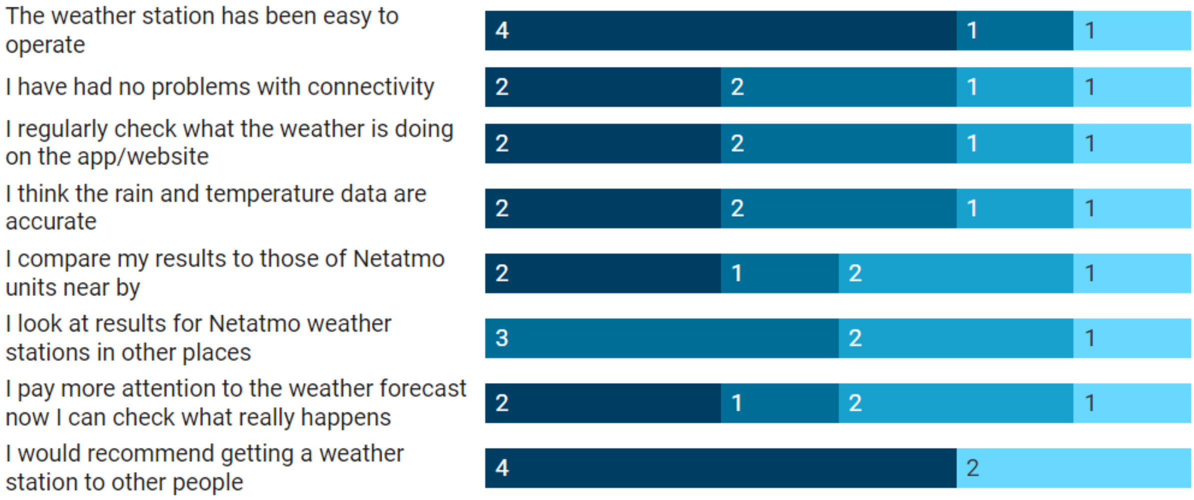


Figure 7.7 Responses from volunteers 18 months after receipt of the weather station

The survey responses were compared for each volunteer between instances to see if people had changed their minds about the weather stations over time (see Figure 7.8). Volunteer 1 has been omitted from the compiled results, as they only responded to survey 1. Each question has been simplified to a single word for ease of representation with the answers to each survey being presented sequentially across the x axis to visualise any change in opinion over time.

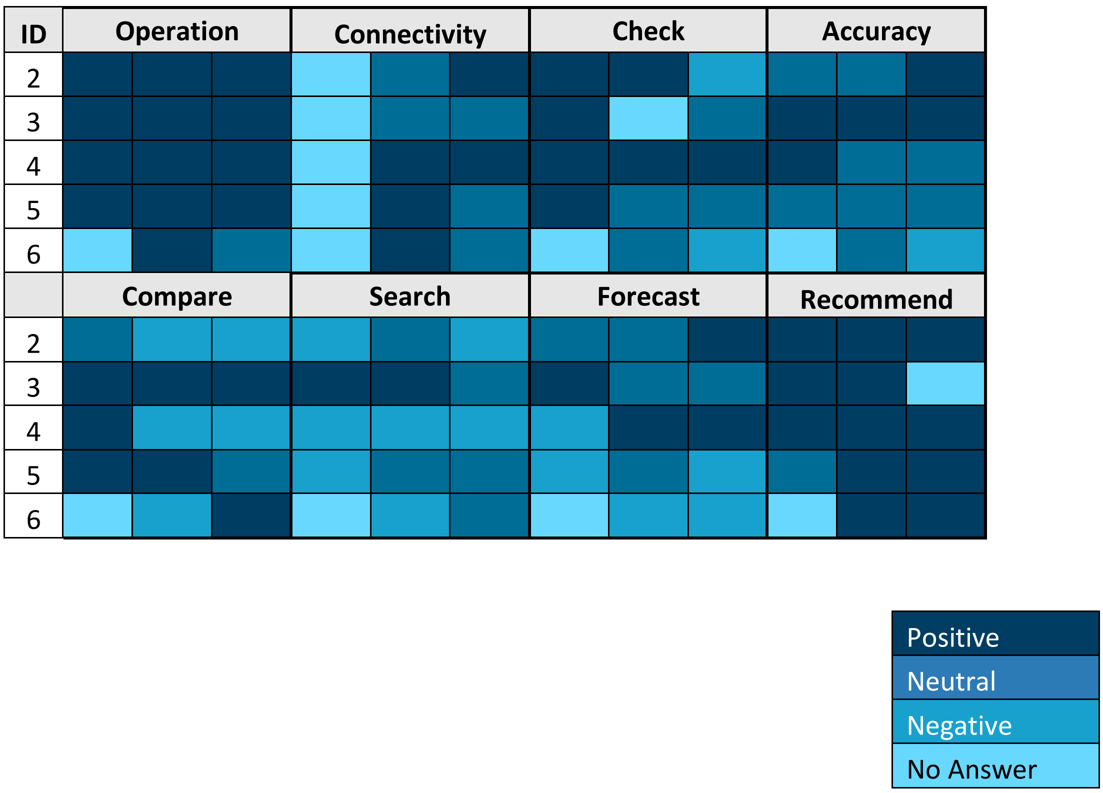


Figure 7.8 Visualisation comparing survey response per volunteer across three instances.

The results show that most volunteers found the weather stations easy to operate throughout except for volunteer 6, who was less positive at the time of the final survey. The only weather station operational at the time of the final survey belonged to respondent 6. Volunteer 2's perception of accuracy increased over the course of the study, whilst volunteers 4 and 6 felt the weather station was less accurate over time. Volunteer 6, who was less confident in the final survey expressed issues with blockages and the unit falling over, which seems likely to have influenced their response.

The volunteers in this study were less inclined to compare their results with other stations as time progressed. It could be that the novelty of this feature wore off, or they were reasonably comfortable that the weather station was providing good data and the comparison was not necessary. The volunteers were less engaged than those who won weather stations in looking at weather station data around the world. It is supposed that as all these volunteers worked with weather data and were familiar with alternative data sources that this feature was less appealing than it was to a non-specialist.

The volunteers said that they would recommend a weather station to others, which did not diminish over time. Volunteers were not asked to contribute financially to the weather stations they were given, which may have affected their decision to recommend them to others. Although the retention rate was low, the volunteers thought there would be a benefit to having a weather station.

7.4.4 The Climatological Observers Link

Data were provided directly by 10 COL members, and a further 4 provided their WOW ID. The analysis of COL data is presented in section 4.5. In summary, when subject to the same automated SQC as WOW and Official data, COL member data submitted directly to the researcher was better quality than that shared via WOW, and proportionally better than the WOW and Official datasets. The disparity between manually shared data and data accessed via WOW was primarily due to erroneous high rainfall observations. The automation of data collection and sharing has been noted as detrimental to quality if there is no human oversight (Muller *et al.*, 2015), and it was confirmed in this research.

There was a degree of interaction between the researcher and COL contributors via email. During initial discussion around the intent of research, several COL members provided information on their background and motivation for participating in weather watching. Eight of the ten contributors stated that they have worked professionally in meteorological services in some capacity from data collection to weather forecasting and 6 were actively sharing data with their communities via their own websites. Having professional experience in the field of study is expected to result in the generation of better-quality data than could be expected from someone with no experience.

The researcher contacted each contributor on completion of SQC to share the results and discuss the data quality. Several of the respondents were able for provide 'rich' metadata regarding the operation of their PAWS, and the information on processing observations. For example, on receipt of the COL dataset one operator relayed that his station had been offline for 2 years; therefore, there were no hourly data for that period, but daily rainfall was available and offered for use in analysis if required (he ran a daily manual gauge alongside the PAWS). Another operator explained the presence of missing data due to incidences of snow fall or sub-zero temperatures causing equipment of freeze. Again, daily totals were offered should a complete dataset be required. There was no such rich data available with the other datasets, making it impossible to know the reasons for missing data, or even if WOW contributors were aware of the gaps.

The combined experience of the COL members ran to many pages of correspondence, some of their insights are collated below regarding a range of issues, including the operation of PAWS, data quality and participation (see Table 7.5). The degree of dedication and depth of knowledge was palpable; the COL data providers were willing to assist the research and were approachable. The points raised by COL members demonstrated their very different approach to rainfall data collection, particularly the elements relating to the calibration and quality control of the data. Clearly having a background in meteorology would aid this; however, it seemed that a critical element resulting in higher quality data was the degree of scrutiny and collective validation of data between the members. As data are shared via the club bulletin the contributors take care (pride?) to ensure their figures are accurate, which are then subject to peer review. There is also an element of knowledge sharing, with talks and presentations on specific elements of weather observation, which improves the knowledge and skills of the group. Plus, there is the option of ready interaction with the highly skilled group.

<table>
<tr><th>General</th><th>Equipment</th><th>QA/QC</th></tr>
<tr><td>• Able to recount extreme weather events in detail.
• Have interest in associated features e.g., groundwater levels and river levels, water quality impacted by CSO's.
• Manual records dating back to the 1960's.
• Somewhat dismissive of 'have-a-go heroes' with cheap equipment dumped in gardens.
• Ways to measure snow – coring, heating with hairdryer.</td><td>• Provided detailed descriptions of the type of equipment, how long it was operational and where it was situated
• Packaging of sensors making it difficult to get good placement – serious observers will split element.
• Highlighted errors with rainfall rate to do with tip issue, vibrations on funnel cause early tip and high rain can lead to double tips etc. causing extreme rate.
• Calibrate equipment regularly, using jewel scales for accuracy.
• Noted issues with TBR blockages from leaves, pollen, grass, etc.
• Interactions with wildlife – squirrels eating through cables
• Damage of equipment due to wind – temperamental reed switches.</td><td>• Actively review their data regularly.
• Running manual check gauges alongside PAWS and use that to identify and chase errors.
• Noted breaks in data due to replacing PAWS.</td></tr>
<tr><th>Interaction/ Engagement</th><th>Data</th><th>WOW</th></tr>
<tr><td>• Actively compare their own data with that from nearby members and discuss why differences may arise.
• Share data to monthly bulletin.
• Receive requests for rainfall data via their website, usually relating to incidences of flooding.</td><td>• Use total from manual gauge and distribute using AWS – ignoring the recorded total.
• Davis imperial to metric conversion of 0.01 inch converts to 0.2mm, 0.5m, 0.7mm, 0.9mm – prefer to convert as one at end of record.
• Applied correction factors for periods where there was drift between check gauge and automated gauge.
• Downscale from check gauge to fill gaps.</td><td>• Aware there is no SQC of WOW data.
• Complained about difficulties interfacing with WOW.</td></tr>
</table>

Table 7.5 Pertinent points from COL members regarding PAWS and rainfall observations.

There was a major drawback in seeking data from COL, as the records were not available online or via any central repository. Data were supplied directly from each user to the researcher. Data formatting was unique to each user making it time consuming to consolidate records, plus there was a chance of transposing data or making other errors during processing. The lack of a COL database means that, in the region of 300 PAWS operated by people capable of generating some of the highest quality citizen science rainfall data is not widely available for operational and research purposes.

7.5 Discussion

The outputs of the research activities presented in this chapter demonstrate some of the difficulties in actively promoting sustainable citizen science participation that results in the generation of good quality rainfall data. Future campaigns to increase and/or improve citizen science/crowdsourced rainfall observations can hopefully be more effective where these challenges are recognised and addressed.

7.5.1 Citizen Science Participation and Crowdsourcing Rainfall Data

There has long been scepticism regarding crowdsourced/citizen science data quality (and specifically rainfall data) (Bell *et al.*, 2015; de Vos *et al.*, 2017; Guerrini *et al.*, 2018). Recently this may have been exacerbated by the ease with which people can now generate and share weather observations, for example, relatively low cost (<£200) Netatmo weather stations are widely available online, require no experience to use and can be set up within an hour (Netatmo). Prior to the wide availability of simple home weather stations, hardware was more costly (circa £1000), requiring purchase from specialist suppliers and some technical competency to set-up data sharing (Davis Instruments, 2018). As a result of this revolution making weather observations is no longer the preserve of the dedicated enthusiast but is now widely accessible to anyone with a Wi-Fi connection and a few hours to spare for set-up. The available technology that now allows anyone to recorded weather data means a potential shift from a true 'observation' to a mere data point.

There are ways to build credibility including ensuring metadata are complete (Eysenbach, 2008), reducing uncertainty around the data providers experience/capabilities (Callister Jr, 2000), and confirmation of data accuracy (Metzger *et al.*, 2003). Building user confidence in by showing data to be trustworthy by demonstrating the expertise of the providers aids establishing credibility (Hovland *et al.*, 1953). It may be possible to improve WOW observations by implementing peer monitoring in the same fashion as COL, or Wikipedia. Approaches that could be considered include rewards from peers for good data (similar to the voting enabled on peer support networks like stack exchange), training on the placement of weather stations/calibration etc similar to that provided to climate observers, and rapid error checking of observations to address issues in a timely manner.

One reason why the Netatmo weather map was so well received by participants in the engagement activities may be due to heuristics. Simple elements including appearance, ease of navigation and no obvious commercial intentions are used by our impressionable human minds as short cuts to assess credibility (Metzger *et al.*, 2010). It could be that the modern presentation of the Netatmo weather map is more enticing than the now more dated appearance of WOW, thus lowering user confidence in the data. Frankly books are judged by their covers as are websites by their colour scheme and operability. A line of thought could be "if the presentation is sub-par, so are the data".

It has been shown that familiarity and user reliance improve the perception of data quality, with an example being confidence in Wikipedia. Those that use the crowdsourced encyclopaedia regularly and are contributors who understand how it is created have more confidence in the entries than people who do not (Johnson *et al.*, 2000).

7.5.2 Response to Research Questions

The aim of the research presented in this chapter was to document the experience of citizen scientists in recording weather observations, considering the heterogeneity of participants regarding skills, knowledge and commitment levels to highlight motivations and barriers to participation. This has been achieved by addressing the research questions posed in section 7.2 and presented here.

Q1. What themes emerge that encourage or reduce participation in crowd sourcing weather observations, and are they common between groups of potential or active citizen scientists?

This research demonstrated that encouraging participation in citizen science weather observations is difficult, requiring skills that are not routinely taught to the engineers and physicists who go on to become meteorologists and hydrologists. Personal experience gained during this research suggests that interdisciplinary working, with social scientists could aid in the development of citizen science project development and support effective engagement.

Support for potential citizen scientists can encourage participation, but based on this research, the provision of weather stations is not recommended if the collection of data is to be sustainable. The aspects of research where weather stations were provided demonstrated that people without a burning desire to operate PAWS are unlikely to share their data or continue beyond the limited duration of intensive encouragement. This has been noted in other citizen science research (Starkey *et al.*, 2017).

Q1a. What actions promote or discourage participation?

The answer depends on the context, as it depends on the aim of the project. Two scenarios are considered, the first is the contribution of PAWS data to an online platform such as WOW. The second is a more involved citizen science project, where participants are involved in the design of the activities. The results of this research suggest that submitting observations could be promoted by making the process as easy as possible, with data transfer mechanisms that do not require a high level of skill and making data as accessible as practical once uploaded. It is assumed that people wishing to share data via WOW already have their own motivations for observing the weather, therefore the encouragement needs to be focused on where the data are shared.

Regarding establishing citizen science projects, this research has shown that encouraging prospective participants to raise their own funds to purchase equipment (as in the case with Corbridge Middle School) is more sustainable and secures a more engaged audience than providing volunteers with equipment (as with the volunteers from Newcastle University).

Distractions and lack of support were shown to be demotivators during engagement with schools and volunteer weather station hosts. Not having the technical skills to link the PAWS to WOW was also a frustration for some participants, to the point they decided it wasn't necessary.

Q2. Can key elements that prevent or support the generation of good quality rainfall data via citizen scientists be identified?

Beven et al. (2020) stated that a key element to improving observational data quality would be the reduction in uncertainties, as far as technologically possible. Regarding WOW this could include promoting calibration, enforcing the requirement for PAWS metadata, and removing conflicting data fields such as rain rate that have a lower accuracy than accumulation.

The interaction with COL members demonstrated that having decades of professional experience in meteorology and the collection of weather data can certainly assist in the generation of good quality rain observations, but that diminishes when the process of data collection and sharing is fully automated. If there could be more interaction with data providers, the quality of data would likely improve. There is an advantage of having near to real time data available via WOW, so a semi-manual process, including the manual upload of data would not be recommended; however, perhaps a hybrid format, where raw data are made immediately available and then SQC is applied prior to archiving would be beneficial. This is effectively equivalent to the process used for Official rainfall data sharing.

Chapter 8. Conclusions

8.1 Overview

This research has demonstrated that there are good quality hourly rainfall data available from thousands of automated gauges operated by citizen scientists in Britain. It has also shown that identifying Private Automated Weather Station(s) (PAWS) rain gauge observations that are good quality would be easier if there was less poor-quality data to filter out. To promote the use of the available WOW rainfall data a method of selecting event specific acceptable quality rainfall data has been presented in this thesis, and quality-controlled datasets have been published.

This chapter concludes the thesis, highlighting the outputs of research and considering what developments may be possible in the future.

8.2 Results Summary

Chapter 3 aimed to determine whether the spatial distribution of rainfall monitoring in Britain is improved with the inclusion of automated citizen science rain gauges reporting to WOW. This was found to be true, with the most benefit observed where rainfall is highly localised and interpolation beyond 5 km from the observation point would be considered unreliable. The chapter detailed the processing of data from private weather stations in the UK and Ireland reporting to the Met Office WOW archive from when it started in 2011 to the end of 2020 that were provided in bulk by the Met Office. Data were filtered to select PAWS within Britain that were reporting at an hourly or shorter time interval for at least 28 days. There were 2,717 weather stations identified; however, the number sharing data to WOW at any given time varied from 351 in 2011 to 1,306 in 2020. Stations had a mean lifespan of 3.6 years and 183 (7%) had records >9 years. The hourly rainfall depth was derived from rainfall accumulation and the potential coverage of Britain was determined at a range of extrapolation distances to assess whether PAWS could fill gaps in the Official rainfall monitoring networks operated by the Met Office, Environment Agency (EA), Scottish Environmental Protection Agency (SEPA), and Natural Resources Wales (NRW). The most benefit was identified at the 5km range where WOW PAWS increased rain gauge coverage of Britain by 14% as a whole, and by 28% in urban areas.

The research presented in Chapter 3 generated datasets of PAWS data and demonstrated the potential of WOW rainfall data to fill gaps in Official monitoring networks, but the quality of those data was not considered. In Chapter 4 data quality was addressed by applying an automated statistical quality control (SQC) algorithm that removed suspicious observations (Lewis *et al.*, 2021), to rainfall data resampled to an hourly interval with the aim of describing the crowdsourced rainfall data available via WOW and assessing the quality as compared to rainfall data from Official ground-based gauge networks in Britain. The results were compared to the SQC of data from 1,477 Official rain gauges (Villalobos-Herrera *et al.*, 2022). Some key differences pre-SQC were noted between the two datasets, namely, the longevity of reporting (some Official gauges had data spanning back >30 years), and the number of observations missing from the datasets (20.78% from WOW gauges, and 4.94% from Official gauges). The SQC removed 2.52% of WOW observations and 2.79% of Official observations, indicating WOW data were of marginally better quality than Official. However, the depth of rainfall removed by the SQC represented 94.83% and 9.19% of the total from the WOW and Official datasets respectively, highlighting the propensity for some WOW rain gauges to report extremely high rainfall depths. It was noted that Official rain gauges were more likely to under than over report rainfall, although observations exceeding the recorded maximum hourly rainfall for the UK (92mm) occurred at 11.54% of Official rain gauges. Manually validation of the performance of SQC showed that some errors, for example, repeated elevated rainfall values (not observed in neighbouring gauges) <92mm were not identified by the SQC. The manual checks highlighted that the SQC needs modifications to capture these errors. The analysis of SQC confirmed the findings of research in Amsterdam and Germany demonstrating the acceptable quality of rainfall observations from PAWS (de Vos *et al.*, 2019; Bárdossy *et al.*, 2021).

Chapter 4 included a case study, looking at data from a special interest group (Climatological Observers Link: COL) to assess whether the data from such a group would be any different from that shared via WOW. It was discovered that the COL member data were of excellent quality when provided manually (although there was some difficulty with inconsistent data formatting) but contained errors when provided automatically via WOW. The case study reiterated the value of creating peer support networks, the benefit of tapping into networks of highly dedicated/skilled citizen scientists, and the need to review automated observations to ensure errors are captured and addressed close to the time of data generation.

In Chapter 5 the quality of WOW and Official data was further scrutinized, using radar as the best available independent validation for hourly observations made during 4 convective rainfall events that resulted in urban pluvial flooding. A standardized approach to data selection was devised and applied to WOW data from PAWS within a defined search area ($30km^2$ – $40km^2$), to remove unreliable data based on inconsistencies during the event, or where the gauge timeseries appeared to include unreliable observations. The resulting WOW and Official observations were compared with rainfall derived from co-located radar values as a means of validation. There was no significant difference between the WOW and Official observations when compared to radar, confirming that the quality of WOW observations was on a par with Official data quality.

Having demonstrated the quality of WOW rainfall observations, the usefulness was assessed by blending combinations of Official, WOW and radar data to create 1km gridded rainfall fields. It was shown that including WOW data in interpolations improves the delineation of rainfall depth during convective events, as compared to using Official rain gauge data with radar, as there is an increased chance of a gauge being in the area of highest rainfall (hence capturing the peak rainfall), or WOW gauge observations outside the area of highest intensity rainfall confining extrapolation and preventing an over-representation of rainfall depth.

In Chapter 6 the importance of accurately representing the rainfall field was demonstrated using a hydrological model to compare observed flow with simulations created with and without WOW rainfall observations. An interpolation created using block kriging with external drift (considered to be the most appropriate for the density of available data (Wang *et al.*, 2013)), generated a maximum flow rate within 3 m^3/s of the observed, as compared to a difference of +/- 12 m^3/s when only radar or radar and Official rain gauge data were used. The result concurred with research arguing that rainfall is the most important parameter in hydrological modelling, and accurate representation of rainfall depth a distribution is essential (Bárdossy *et al.*, 2022).

Finally, in Chapter 7 the motivations and barriers for citizen scientist wishing to participate in weather observation were explored. Different modalities of engagement were trialled, including via museums and schools, and with staff from Newcastle University. The challenge of encouraging participation was demonstrated, even where participants were highly knowledgeable and familiar with the principles. It was concluded that the most effective way to encourage participation is by providing support to those who have already signalled an interest either by observing the weather or expressing an interest in doing so. A top-down approach including the supply of equipment does not guarantee enthusiastic volunteers. A similar phenomenon has been noted during the provision of toilets to promote sanitation, whereby receivers do not feel like they have ownership or responsibility for the equipment, whereas a modality promoting the benefits of sanitation has been found to be sustainable and effective (Kar *et al.*, 2008).

8.3 Results Significance

The value of crowdsourced rainfall observations in post-event reconstruction of rainfall has been demonstrated by this research. Climate change is resulting in an increasing intensity of convective storms, with subsequent increases in pluvial flooding (Fowler *et al.*, 2021), meaning

Section-19 reports on post-event analysis will become more common place. PAWS rainfall data have the potential to aid engineers in determining issues with urban drainage and make improvements based on the most accurate post-event reconstructions. The time required for SQC means that such post-event analysis is currently the most practical application for PAWS rainfall data. Using and further exploring the available data would further highlight both the possibilities and any issues in the data, and enlivened interest could be a catalyst for improvements.

The SQC method used during this research was selected as it had been demonstrated to be effective on automated data from around the world (Lewis *et al.*, 2021), and on Official observations from Britain, meaning a quality controlled dataset was available for comparison (Villalobos-Herrera *et al.*, 2022). In its current format the SQC requires multiple years of observations to run all the checks, meaning it may not be the most effective approach for the quality control of PAWS rainfall observations given the short records of many. Real-time quality control was attempted by de Vos *et al.* (2019) which relies on assumptions of a consistent PAWS reporting bias which may have been appropriate when looking at data from a single type of weather station, but may not be valid for the mixed bag of PAWS reporting to WOW. In short, this research has shown that none of the current SQC approaches are ideal for PAWS rainfall data validation, which is reflected in following section on Limitations and Further Work.

The SQC algorithm applied in this research takes time to implement, making it unsuitable for real-time evaluation of rainfall data. As the PAWS data increase the extent of rain gauge coverage in Britain, particularly in urban areas where pluvial flooding is expected to cause increasing damage in the coming years, the data at present could be used to indicate the potential pluvial flood risk. The data may not have been validated; however, if intense rainfall is detected in multiple PAWS gauges it can serve as an indicator of where and when pluvial flooding could occur, and prompt validation e.g., from Official monitoring, radar, and social media posts. This could be a useful tool for LLFA's as the responsibility for pluvial flooding lies with them.

The impact of discharges from Combined Sewer Overflows (CSOs) has recently been in the news in Britain (BBC, 2022), and has been researched in the context of heavy rainfall (Miller *et al.*, 2017). CSO's release sewerage into rivers and the sea when the flow into the drainage system threatens to overwhelm treatment plants causing sewerage to back-up. Attempts have been made to predict when CSO's may discharge using radar data (Mounce *et al.*, 2014), and water companies have sensors on some CSOs that share notifications of discharges (Surfers Against Sewerage, 2022), but not all CSOs have real-time data sharing. Organizations including Surfers Against Sewerage (*Ibid*.) and The Rivers Trust (The River's Trust, 2022) operate websites to inform the public of risk from CSOs that could be enhanced if real-time high intensity rainfall observations such as those derived from PAWS were incorporated into warnings. Such warnings help the public determine if it is safe to enter the water, and it would be encouraging to see citizen science derived data being used in this practical way to help others.

This research has focused on PAWS in Britain, principally for practical purposes for demonstrating data quality via comparison with alternative data sources (Official ground-based rain gauging and radar); however, the role of citizen science rainfall monitoring in countries with a less comprehensive rainfall observation network is implicit. Research on such potential has often been limited to small scale undertakings, presumably also for practical purposes, but limiting the ability to draw far reaching conclusions (Davids *et al.*, 2019; Fehri *et al.*, 2020; Shinbrot *et al.*, 2020). This research supports the assertion that crowdsourced rainfall data can be good quality, but also highlights the ways in which it can fail. The importance of QA procedures and close interaction with citizen scientists becomes critical where those data are the sole source of observations for an area. Research by Walker *et al.* (2021) is clear on the need to carefully implement citizen science projects for the benefit of all involved, which is consistent with the findings of this thesis on the importance of effective engagement with willing participants. However, it is also important to recognised the benefits of citizen science rainfall observations extend beyond the collection of data, promoting also the concept of 'experiential learning' whereby participants have a greater appreciation of their environment or have the opportunity to advocate for their communities and share their in depth knowledge with researchers (Haklay, 2012; Paul *et al.*, 2017).

8.4 Limitations and Further Work

8.4.1 Data Availability and Sources

This research only looked at WOW data. The improvements in spatial density demonstration in section 3.6.2 would be greater if data from other monitoring networks were readily available. Availability of real-time data is critical for certain applications, for example, real-time pluvial flood forecasting. A review of the speed in which data become available was not undertaken and would be beneficial. An entirely different approach to SQC is required if WOW data are to be used in real-time applications or now casting that require rapid data provision. This research has shown that there cannot always be a 'pre-judgement' of data quality, as quality can diminish at a gauge without warning. A method using rapid nearest neighbour checking and radar cross validation could be implemented, whereby the approach is more akin to using the WOW data to get a sense of the situation in real-time, using the observations of multiple gauges as indicative of rainfall (Goodchild *et al.*, 2012).

The time of reporting versus the time of rainfall was highlighted as an issue during the exploratory analysis that was undertaken for this research. It was noted that the method commonly employed by official monitoring organizations of recording 'tip-time' is a more robust way of recording rainfall. Each time the 'bucket' of the rain gauge empties, the time is recorded. This method gives a more precise indication of the time of rainfall, particularly during intense rain, and can be aggregated to any desired time interval. If citizen science rain gauge data was available as tip times, it would allow for more detailed analysis of rainfall during the events where such gauges have the potential to add the most useful data.

Station metadata are available via WOW but not provided with the observations. It is possible there could be systematic errors or, errors more likely to be noted from certain types of rain gauge, or due to data transfer from a given gauge or via and particular system (e.g., Netatmo weather station data can be aggregated via a third party prior to upload to WOW). Without easy access to metadata, it has not been possible to explore such issues. As with previous research (Bell *et al.*, 2015; de Vos *et al.*, 2019), if metadata were available it may be possible to identify systematic errors and work to correct them.

8.4.2 Statistical Quality Control Developments

The SQC code was developed for the assessment of hourly/sub-hourly rainfall data from Official monitoring networks around the world. There were 2 main limitations:

i. An assumption in nearest neighbour checks that neighbouring stations were reporting accurate observations that could be used in cross-validation.

ii. The application of all SCQ steps in one iteration, preventing the filtering of the 'worst' observations and thus limiting the ability of perform nearest neighbour checks withing the SQC process.

As highlighted in section 4.6.1 improvements could be made, following the ontologies of SQC developed for Netatmo rain gauge observations in the Netherlands and Germany (de Vos *et al.*, 2019; Bárdossy *et al.*, 2021), whereby the SCQ algorithm is applied sequentially on successively filtered datasets. This would require assessment of the impact of data gaps on the successful implementation of nearest neighbour checks and validation of the handling of high intensity rainfall in sparsely gauged areas, where rain may only be observed at a single location. Modifications could be made to better reflect some of the errors identified during the manual review of WOW gauge records. A rule could be developed to capture sporadic yet repeated extremes below the defined hourly maximum. Given the potential for the use of WOW data in the delineation of convective storms, it may be more productive to apply SQC to sub-hourly data, rather than hourly (aggregated) data. A method for sub-hourly SQC has been designed and implemented by Villalobos-Herrera *et al.* (2022).

Automation appears to have been the enemy of crowdsourced data accuracy. This could be addressed by instigating a semi-automated approach whereby data are manually uploaded to WOW, and encouraging participants to review prior to submission; however, this limits the potential for real- or near real-time reporting. Alternatively, a SQC process could be developed that is periodically implemented in a timely fashion on the available data. Where errors are discovered there could be direct interaction with the citizen scientist to address issues and if possible, prevent further errors. This would require more oversight of the WOW data archive but would be akin to the process applied to Official data. If the WOW database is to gain credibility and become a useful tool for research, etc., there needs to be more active management of data.

8.4.3 Follow-on Research

The case study presented in Chapter 6 using data in a hydrological model focuses on a single event, serving to demonstrate the key benefits and limitations of using the different rainfall data sources. There are instances where the outcomes may be different, for example when/where the radar is less reliable. The analysis should be repeated for other convective events where there are sufficient gauges available, with consideration of alternative blending methods. Bayesian merging may be more effective where convection is highly localised or where the highest intensity rainfall is captured by radar, but not rain gauges (Wang *et al.*, 2013).

In Chapter 7 the documentation of interactions with existing of aspiring (or ambivalent) citizen scientists consisted only of small-scale studies, that in many ways were subsidiary to the focus of WOW data validation and application. Working with practitioners skilled in engagement would aid in the development and assessment of activities promoting weather observations by citizen scientists. A significant limitation to some of the activities presented here relates to difficulties with engagement during the pandemic. The strength of COL appears to be the peer network, therefore where citizen science rainfall observation activities are being instigated, it could be more productive to meet in person as a group and to build a community. To better understand motivations a longer term, more comprehensive study would be useful comparing modalities or degrees of engagement.

Appendix 1

The datasets and code listed below and named in the thesis body text are available directly from T. O'Hara on request.

I. Multiple parameter observations at original time intervals
II. Summary statistics for selected stations (a lookup table)
III. Calculated rainfall observations at original time intervals
IV. Rainfall accumulations at hourly intervals
https://doi.org/10.25405/data.ncl.21724970.v1
V. Code used for processing the WOW bulk download. – see the Newcastle University Water Group GitHub; https://github.com/nclwater/WOW
VI. MIDAS data processing code, daily and hourly. – see the Newcastle University Water Group GitHub; https://github.com/nclwater/WOW

Arc StoryMap link with live map showing station statistics - https://arcg.is/1zmui5

Appendix 2

WMO test classification	Automated QC check (and rule if applicable)
Format	QC13 - Daily Accumulation (R2), QC14 - Monthly Accumulation (R3)
Completeness	QC1- Percentile, QC2 – K-largest, QC5 – Intermittency
Consistency - Internal	Unchecked – requires non-rain parameters
Consistency - Temporal	QC3 – Days of Week, QC4 – Hours of Day, QC6 – Breakpoints, QC13 – Daily Accumulations (R2), QC14 – Monthly Accumulations (R3)
Consistency - Spatial	QC16 – Daily Neighbour (wet) (R8), QC17 – Hourly Neighbour (wet) (R7), QC18 - Daily neighbours (dry) (R10), QC19 - Hourly neighbours (dry) (R9), QC20 – Monthly neighbours (R11).
Consistency - Summarization	QC8 - R99pTOT, QC9 - PRCPTOT
Tolerance/range	QC10 - World record* (R5), QC11 - Rx1day* (R6)
Spike and Streak	QC15 – Streaks (R4)

Automated SQC summary against WMO recommendations

WMO test classification	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Format													x	x						
Completeness	x	x			x															
Consistency - Temporal			x	x		x							x	x						
Consistency - Spatial																x	x	x	x	x
Consistency - Summarization								x	x											
Tolerance/range										x	x									
Spike and Streak															x					

SQC Matrix of Tests Against WMO Recommendations

The following tables describes each of the SQC checks, providing the short name, a description of the check purpose, the rationale for inclusion and any whether failing the check results in observations being removed. The first group of checks (QC1-QC7) seeks to identify gauge records in which substantial portions of the data (or even the whole record) could be suspicious. The text in these tables and providing further explanation is after Lewis *et al.* (2021) with minor modifications to update the applied maxima, as modified for Britain.

QC & Rule number	Short name	Description
QC1 (a)/(b)	Percentiles	Returns years where the 95th (a), and 99th (b) percentiles of the hourly rainfall distribution are zero. Corresponds to suspiciously low rainfall in many climate regions
QC2 (a)/(b)/(c) R1	K-largest	Hourly rainfall totals are put in descending order separately for each year. The check then returns the years where the largest 1st, 5th and 10th largest hourly rainfall totals are all equal to zero. The check helps to identify suspiciously low annual rainfall totals in all but the very driest climates. Very low annual totals may occur where missing values have been entered as zeros, for example.
QC3	Days of Week	The data are checked to look for biases towards certain days of the week. A two-sided t-test on the distribution of mean rainfall over the days of the week is performed. E.g., helps in identification of spuriously higher rainfall on Mondays, due to data being preferentially reported as a weekend accumulation.
QC4	Hours of the Day	As QC3 but for hours of the day. A two-sided t-test on the distribution of mean rainfall over the hours of the day to help indicate if daily accumulations of rainfall have been reported at 0900 or another standard time.
QC5	Intermittency	continuity and homogeneity are checked. Generally patchy records (i.e. numerous sequences of missing data) are identified using the check which returns years where more than five periods of missing data are bounded by zeros. The lack of continuity in the record highlighted by this check suggests that rainfall events may not have been properly recorded, for example due to dry periods being recorded but not rainfall events.
QC6 (a)/(b)	Breakpoints	A Pettit test is used, which may indicate changes in station location or measuring equipment. However, a breakpoint could also be detected due to a "real" change in rainfall, induced by climate change for example. As such the results of the Pettit test are tentative and supplied for the user's reference.
QC7	Minimum value change	Possible changes to the measuring equipment are reflected by the check which identifies how many times the minimum recorded value (greater than zero) changes from one year to the next.

Suspiciously High Values

QC & Rule number	Short name	Description and Rationale
QC8	R99pTOT	Compares the annual gauge statistics against the maximum ETCCDI indicating annual total precipitation on days where precipitation exceeds the 99th percentile of wet day precipitation.
QC9	PRCPTOT	Uses PRCPTOT climate indices to quantify annual total precipitation on wet days (≥1 mm).
QC10 (R5)[15]	World record*	The check compares each individual hourly rainfall value against the UK record hourly rainfall total.
QC11 (R6)	Rx1day*	Each individual hourly value is compared against the corresponding maximum ETCCDI Rx1day (maximum daily precipitation) value for the corresponding grid cell. Note that the check compares hourly values with Rx1day rather than daily values, again in order to account for the smoothing of spatial variation inherent in the gridded ETCCDI indices.

* These check thresholds have been amended for the UK , QC10 – UK value 92mm, QC11 – UK value 341.5 mm

Long Periods without Rainfall

QC & Rule number	Short name	Description and Rationale
QC12 (R9)	CDD	potentially suspicious long dry period are assessed using the ETCCDI CDD (consecutive dry days) index. This additional check is intended to provide more confidence in the identification of periods in which missing data may have been recorded as zeros.

Suspect Accumulations and Repeated Values

QC & Rule number	Short name	Description and Rationale
QC13 R2	Daily accumulations	The "daily accumulations" (QC13) and "monthly accumulations" (QC14) checks look for day- and month-long periods of zeros followed by a very wet hour and then another spell of 23 zeros (i.e. the rest of the day is dry). The purpose is to highlight cases where rainfall accumulated over a full day or month is mistakenly associated with a single hour. The checks use a threshold of double the ETCCDI SDII index for the grid cell underlying the gauge location to classify the wet hour as potentially spurious. The SDII is an index of the average daily precipitation total on wet days. Inspecting gauge time series showed that specifying that the wet hour needs to be followed by a period of 23 zeros strongly reduces the possibility of flagging genuinely large rainfall totals that happen to follow a legitimate dry period.

[15] Hourly value is massive compared to the next biggest in 'modern' data –it's about 65mm. So there may have been an issue with the potential citizen scientist who recorded that. See Flood Studies report VII, Meteorological studies, 1975 – p70 UK record events.

QC & Rule number	Short name	Description and Rationale
		Two additional elements are incorporated in the "daily accumulations" (QC13) check to help identify long periods of systematic daily accumulations of any magnitude. Firstly, periods of two or more days with any non-zero potential daily accumulations are identified. Secondly, periods of zeros in between potential daily accumulations are also flagged. This helps to identify continuous periods of potential daily accumulations, i.e., not only periods where a day receives rainfall.
QC14 R3	Monthly accumulations	
QC15 R4	Streaks	Repeated (identical) values based on several conditions in the "streaks" check (QC15). First, streaks of 2 or more consecutive hours of identical high rainfall values are flagged using double the ETCCDI SDII index as a threshold again. It is considered unlikely that 2 or more hours of identical large values would occur by chance. Second, streaks of 12 or more consecutive hours of identical values greater than the measurement resolution are also flagged. Third, streaks of 24 or more repeated hourly values greater than zero are flagged (i.e. very long streaks at the data resolution). The different thresholds in the second and third components of the check were chosen to reduce the chance of wrongly identifying spells of drizzle as potentially erroneous. Finally, the check flags periods of zeros bounded by streaks of ≥ 24 repeated values. This helps to find periods of uniform disaggregation or repetition of daily totals.

Neighbouring Gauge Checks on Large Values

These checks are performed after aggregation of the hourly series to the daily interval. Neighbour checks are not currently conducted on an hourly basis, due to the heightened intermittency and spatial variability at sub-daily time intervals.

QC & Rule number	Short name	Description and Rationale
QC16 R8	Daily neighbours (wet)	Uses the quality controlled GPCC daily precipitation database as reference data
QC17 R7	Hourly neighbours (wet)	Uses neighbouring gauges in the (GSDR) dataset being checked (after aggregation to the daily interval).

QC16 Not in UK version – no GPCC data.

The checks involve the following steps:

- Select the closest (up to 10) neighbouring gauges within 50km with sufficient overlap (>3 years) and close correspondence (i.e., in wet/dry matching statistics and correlation) with the target gauge. The check is not undertaken if there are insufficient gauges meeting these criteria
- Calculate a time series of differences between the target gauge and each neighbour (after normalisation of each record to the interval 0-1)
- For each neighbouring gauge, produce a time series of flags using the following approach:
- Subset the differences on wet days where the target gauge exceeds the (normalised) value of the neighbouring gauge
- Fit an exponential distribution to the subset of differences

- Identify threshold difference values associated with the 95th, 99th and 99.9th percentiles. Use these thresholds to flag exceedances (i.e,. notably large differences)
- Combine the flags (i.e., one series per neighbour) into a single series of flags, where the ultimate flag value is equal to the minimum flag value of all neighbours at each time step

Other Neighbouring Gauge Checks

QC number & Rule	Short name	Description and Rationale
QC18 R10	Daily neighbours (dry)	These tests examine the agreement between a target gauge and its neighbours on dry periods. The neighbour selection follows the same procedure as QC16&QC17. The checks are again differentiated by use of the GPCC daily precipitation database (QC18) and neighbouring gauges within the GSDR dataset being checked (QC19). For a moving 15-day window, if the target gauge shows no rainfall in the window, the number of wet days (>0 mm) in the same window for each neighbour are counted and the average taken. To receive the highest flag, all neighbours must have three or more wet days in a 15-day window that is dry in the record of the target gauge. Successively lower flags are used if all neighbours corroborate the occurrence of a lower number of wet days
QC19 R9	Hourly neighbours (dry)	
QC20 (R11)	Monthly neighbours	uses reference data from the quality-controlled GPCC monthly precipitation database. Monthly neighbouring gauges are selected based on proximity only. The check proceeds by calculating the percentage difference in monthly totals between each neighbouring GPCC monthly gauge and the target gauge (after aggregation to the monthly interval). No normalisation of monthly totals is currently undertaken. The percentage differences are grouped into classes and the flags associated with each of the neighbours are then combined into a single series of flags. Again, all neighbours need to signal notably high or low monthly totals for the most severe flags to be invoked. For example, if the target gauge monthly total is more than double the totals of each of the active neighbouring gauges then a high flag is assigned
QC21 - 24	Timing offset	identification of generally suspicious gauges by augmenting QC1-7. The check (QC21) computes the affinity index (AI) and correlation at the daily interval between the target gauge and its nearest neighbour in the GPCC database at lags of -1, 0 and +1 day. The AI is a matching statistic for wet and dry days (a dry day is defined here as having <0.1mm rain). The lag associated with the highest values of the statistics is returned as an indication of whether or not there could be potential timing issues. As part of the procedure we also calculate the “pre-QC affinity index” (QC22) and “pre-QC correlation” (QC23) relative to the nearest daily neighbour in the GPCC database as a means of identifying generally questionable records. Finally, the “daily factor” (QC24) and “monthly factor” (QC25) checks provide the mean difference (expressed as a factor) between the target gauge and its nearest daily or monthly GPCC neighbour, respectively, to help identify possible unit errors

Appendix 3

Met Office Daily Weather Summaries (Met Office, 2021a)

20/7/14 – Canvey Island

Summary of UK Weather for Sunday 20 July 2014. ***Thundery showers*** *across the north and east of East Anglia moved steadily north to affect Lincolnshire and Yorkshire during the early hours. These* ***thunderstorms were locally torrential****, with 44.8mm of rain recorded at Tealby, Lincolnshire in two hours. It was another very mild night in most places, with temperatures remaining in double figures Celsius nationwide. During the afternoon,* ***slow-moving locally intense and thundery showers*** *developed across north-east England and* ***eastern England****, and some of these gradually edged slightly west. There were reports of* ***localised flooding*** *in parts of Norfolk, Suffolk,* ***Essex*** *and Kent, with 45.8mm of rainfall falling in the hour between 1500 and 1600 at Norwich Airport. Other locations affected by the showers reported hourly rainfall accumulations of 20-40mm. It was another warm or locally very warm day, with maximum temperatures reaching the mid 20s Celsius fairly widely. EXTREMES Highest Maximum: 27.3 Solent, Most Rain: 66.4 Hanningfield, Essex*

8/8/14 – Bar Hill Cambridge

Summary of UK Weather for Friday 08 August 2014. Any mist and fog patches soon cleared through the morning, and the ***showery rain*** *across north-western parts moved eastwards, becoming locally heavy in places. Some* ***heavy showers developed across eastern parts of England*** *in the morning as well, these gradually moved northwards. Showers also started to develop across more central and northern parts of England into the afternoon, where they become* ***more intense and widespread, with some very heavy, thundery downpours.*** *Into the evening, another area of* ***showery rain spread*** *from the south across some central-southern areas, and then south-east England and East Anglia. By midnight, there was still an area of rain,* ***heavy at times across the south-east and east*** *of the country, and showery rain affecting coastal parts of north-east England and south-east Scotland, as well as Aberdeenshire. Otherwise, it was largely dry. Temperatures were generally close to or a little above average for the time of year. EXTREMES Highest Maximum:25.7 Wellesbourne, Warwickshire, Most Rain: 64.0 Bramham, North Yorkshire*

16/06/16 - Birmingham

Summary of UK Weather for Thursday 16 June 2016. Isolated showers continued throughout the early morning across central and southern England, whilst low cloud and sea fog edged on-shore along the north-east coast of England. The bulk of England and Wales saw sunny spells and ***scattered heavy, slow-moving and often thundery showers****. This led to some* ***localised flooding*** *in places. Despite the showers, Cornwall, Kent and Sussex saw the driest and brightest weather, and temperatures remained around average for the time of year across the country. EXTREMES Highest Maximum:22.4 Faversham, Kent, Most Rain:48.8 Monks Wood, Cambridgeshire*

27/05/18 – Milton Keynes

Summary of the UK Weather for Sunday 27 May 2018. ***Thunderstorms*** *continued moving northwards across southern parts of the UK during the early hours. There was very* ***heavy rain*** *across parts of London, and into the Midlands and by dawn, Wales. A band of heavy and thundery rain continued to move northwards across parts of Wales and the Midlands during the day. Late in the day some* ***intense thunderstorms developed across the Midlands****, particularly over the Birmingham area, causing some* ***localised flooding*** *in places. These moved northwards into parts of north Wales and into the evening.* ***Further thunderstorms*** *broke out across parts of* ***southeast England and East Anglia*** *late in the afternoon. EXTREMES 27.6°C London, St James's Park, Most Rain: 81.0mm Winterbourne No2*

Canvey Island Manual SQC results

ID	Time series gaps	Data missing within 24 hours	Potential accumulation after gap	Gap due to QC?	Event pattern anomalies	Event High Extreme	Time series anomalies	Pre QC extremes	Post QC Extremes	Annual Pearson High	Judgement	Comments
"192316" : "W1"	complete				No	No	No	Yes	Yes	High	Accept	
"10251822" : "W2"	complete				No	No	No	No	No	Med	Accept	
"23517106" : "W3"	complete				No	No	No	Yes	Yes	High	Accept	
"23688198" : "W4"	complete				No	No	No	No	No	High	Accept	
"30826477" : "W5"	complete				No	No	No	Yes	No	High	Accept	
"39116470" : "W6"	complete				No	No	No	No	No	High	Accept	
"39717800" : "W7"	complete				No	No	No	Yes	Yes	High	Accept	
"48804542" : "W8"	complete				No	No	Yes	Yes	Yes	Med	Reject	Correlation and timeseries concerns
"52369157" : "W9"	complete				No	No	No	Yes	No	High	Accept	
"53658882" : "W10"	complete				No	No	Yes	Yes	Yes	High	Reject	QC Fail?
"75188624" : "W11"	complete				No	No	No	Yes	Yes	High	Accept	
"405626113" : "W12"	complete				No	No	No	Yes	No	NA	Accept	
"414186106" : "W13"	complete				No	No	No	No	Yes	NA	Accept	
"436149203" : "W14"	complete				No	No	No	Yes	Yes	NA	Accept	
"692946068" : "W15"	complete				No	No	No	Yes	Yes	NA	Accept	
"147933" : "W16"	incomplete	No	No	No	No	No	No	No	No	High	Accept	
"760600" : "W17"	incomplete	Yes	Yes	No	Yes	Yes	No	No	Yes	High	Reject	Accumulation after gap
"786072" : "W18"	incomplete	No	No	No	No	No	No	Yes	Yes	High	Accept	
"10969640" : "W19"	incomplete	Yes	No	No	No	No	No	No	No	High	Reject	Data missing on 20th
"23170907" : "W20"	incomplete	No	No	No	No	No	No	Yes	No	NA	Accept	
"23342702" : "W21"	incomplete	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Reject	QC Fail - see timeseries? Large gaps during event
"23973496" : "W22"	incomplete	No	Yes	No	No	No	No	No	No	High	Accept	remove possible accumulation - 18/07/2014 08:00:00
"24037340" : "W23"	incomplete	No	No	No	No	No	No	No	No	High	Accept	
"24881929" : "W24"	incomplete	No	No	No	No	No	No	No	No	NA	Accept	Data missing to 18th
"25114645" : "W25"	incomplete	No	Yes	No	No	No	No	No	Yes	High	Reject	big data gap
"28836659" : "W26"	incomplete	No	No	No	No	No	No	Yes	No	NA	Accept	
"29197359" : "W27"	incomplete	No	No	No	No	No	No	Yes	Yes	High	Accept	
"33978011" : "W28"	incomplete	No	Yes	No	Yes	Yes	Yes	No	Yes	High	Reject	QC Fail? Minimal data during event
"37860072" : "W29"	incomplete	Yes					No	Yes	Yes	NA	Reject	No data for event duration
"39816399" : "W30"	incomplete	Yes					No	No	No	NA	Reject	No data for event duration
"40111807" : "W31"	incomplete	Yes	No	No	No	No	No	No	No	High	Accept	
"47109197" : "W32"	incomplete									NA	Reject	No data for event duration
"68140038" : "W33"	incomplete	Yes	No	No	No	No	No	No	No	High	Accept	
"76000035" : "W34"	incomplete	No	No	No	No	No	No	Yes	Yes	High	Accept	
"152266007" : "W35"	incomplete	Yes					No	Yes	Yes	NA	Reject	Big gap, no rain
"319466657" : "W36"	incomplete	No	No	No	No	No	Yes	Yes	No	High	Reject	QC Fail?
"344516003" : "W37"	incomplete						No	Yes	No	NA	Reject	No data for event duration
"368936192" : "W38"	incomplete	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Low	Reject	Multiple extremes in time series and event data inconsistent
"376516074" : "W39"	incomplete	Yes	No	No	No	No	No	No	No	High	Accept	
"876886001" : "W40"	incomplete	No	Yes	No	No	No	No	Yes	No	NA	Accept	remove possible accumulation - 13/07/2014 13:00:00

Bar Hill manual SQC results

ID	Time series gaps	Data missing within 24 hours	Potential accumulation after gap	Gap due to QC?	Event pattern anomalies	Event High Extreme	Time series anomalies	Pre QC extremes	Post QC Extremes	Annual Pearson High	Judgement	Comments
"172761" : "W1"	complete				No	No	No	No	No	Med	Accept	
"252684" : "W2"	incomplete	Yes									Reject	No data for duration of event
"804887" : "W3"	incomplete	Yes	No	No	No	No	No	No	No	High	Accept	
"1136422" : "W4"	complete				No	No	No	No	No	High	Accept	Some gaps in timeseries
"2859815" : "W5"	incomplete	No	No	No	No	No	No	No	No	High	Accept	
"5538222" : "W6"	complete				No	Yes	Yes	No	No	Med	Accept	Possible QC fail for 2016 - gap in reporting with high value
"5691150" : "W7"	complete				No	No	No	No	No	NA	Accept	Big data gaps pre/post QC
"22649470" : "W8"	incomplete	No	No	No	No	No	Yes	No	No	High	Accept	2019 no data
"23872180" : "W9"	complete				No	No	No	Yes	No	NA	Accept	Possible QC fail single high observation post Gap in reporting
"24072046" : "W10"	complete				No	No	Yes	Yes	No	High	Accept	Possible QC fail high post QC value in 2015, other extremes removed by QC
"30733919" : "W11"	complete				No	Yes	Yes	Yes	No	High	Accept	Possible QC fail, big difference in early 2013 data and after, multiple extremes
"33078600" : "W12"	incomplete	No	No	No	No	Yes	Yes	Yes	No	High	Accept	Possible QC fail repeated values
"37427472" : "W13"	incomplete	Yes									Reject	No data for duration of event
"39583046" : "W14"	complete				No	No	No	No	No	High	Accept	
"40058795" : "W15"	complete				No	Yes	Yes	Yes	No	High	Accept	Possible under-reporting across timeseries
"41355920" : "W16"	incomplete	No	No	No	No	No	Yes	Yes	No	Low	Reject	Several extreme highs pre QC
"48831203" : "W17"	incomplete	No	No	No	No	No	No	No	No	High	Accept	
"56672429" : "W18"	complete				No	No	No	No	No	High	Accept	
"72308033" : "W19"	incomplete	Yes									Reject	Large gaps including 8th
"72412105" : "W20"	complete				No	No	No	No	No	High	Accept	
"73519307" : "W21"	incomplete	Yes									Reject	Large gaps including 8th
"74636199" : "W22"	complete				No	No	No	No	No	NA	Accept	About 2 years missing data
"74961273" : "W23"	complete				No	Yes	No	No	No	High	Accept	Possible QC fail, very low rainfall, short lifespan, not sure what's going on
"77994178" : "W24"	complete				No	No	Yes	Yes	No	NA	Accept	2 year gap in reporting
"78500105" : "W25"	incomplete	No	No	No	No	No	Yes	Yes	No	NA	Accept	Possible under-reporting across timeseries
"142526008" : "W26"	complete				No	No	Yes	No	No	High	Accept	
"153666198" : "W27"	incomplete	Yes	No	No	No	No	Yes	Yes	No	High	Accept	Checked timeseries closely, no issues
"380416838" : "W28"	complete				No	No	No	No	No	NA	Accept	
"401996242" : "W29"	complete				Yes	No	No	No	No		Reject	Missed rain when compared to other gauges
"435016816" : "W30"	incomplete	Yes									Reject	No data for duration of event
"456876311" : "W31"	complete				No	No	Yes	Yes	No	NA	Accept	1 extreme pre-QC
"513876019" : "W32"	incomplete	Yes									Reject	No data for duration of event
"529146002" : "W33"	complete				Yes	No	No	Yes	No		Reject	Missed rain when compared to other gauges
"883166001" : "W34"	incomplete	Yes									Reject	No data for duration of event
"884006001" : "W35"	incomplete	Yes	No	No							Reject	2 day reporting gap prior to event
"886206001" : "W36"	complete				Yes	No	No	Yes	No		Reject	Decimated by QC
"887006001" : "W37"	complete				No	No	No	No	No	NA	Accept	
"888496001" : "W38"	complete				No	No	No	No	No	NA	Accept	
"891556001" : "W39"	incomplete	No	No	No	No	No	No	Yes	No	NA	Accept	

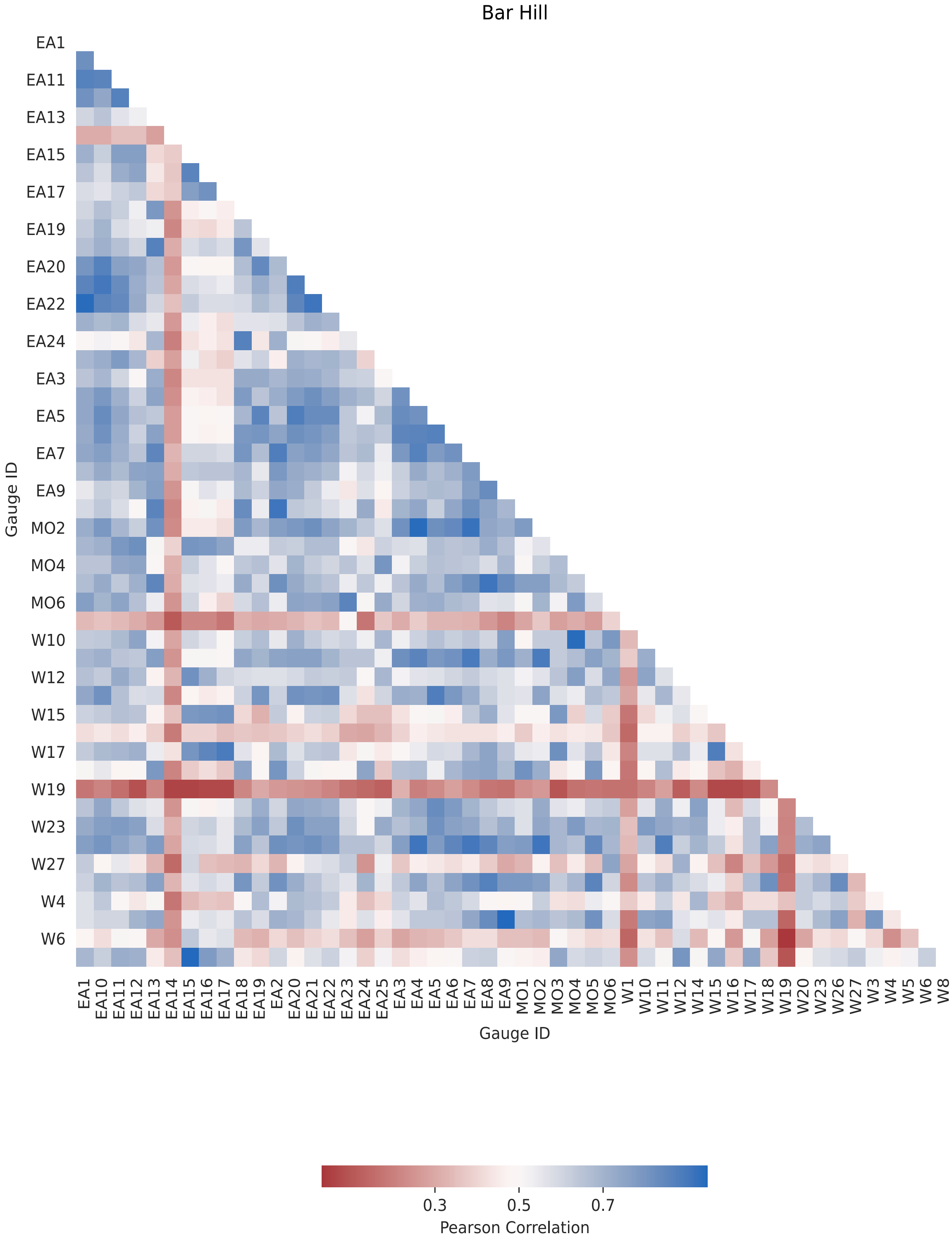
Birmingham manual SQC results

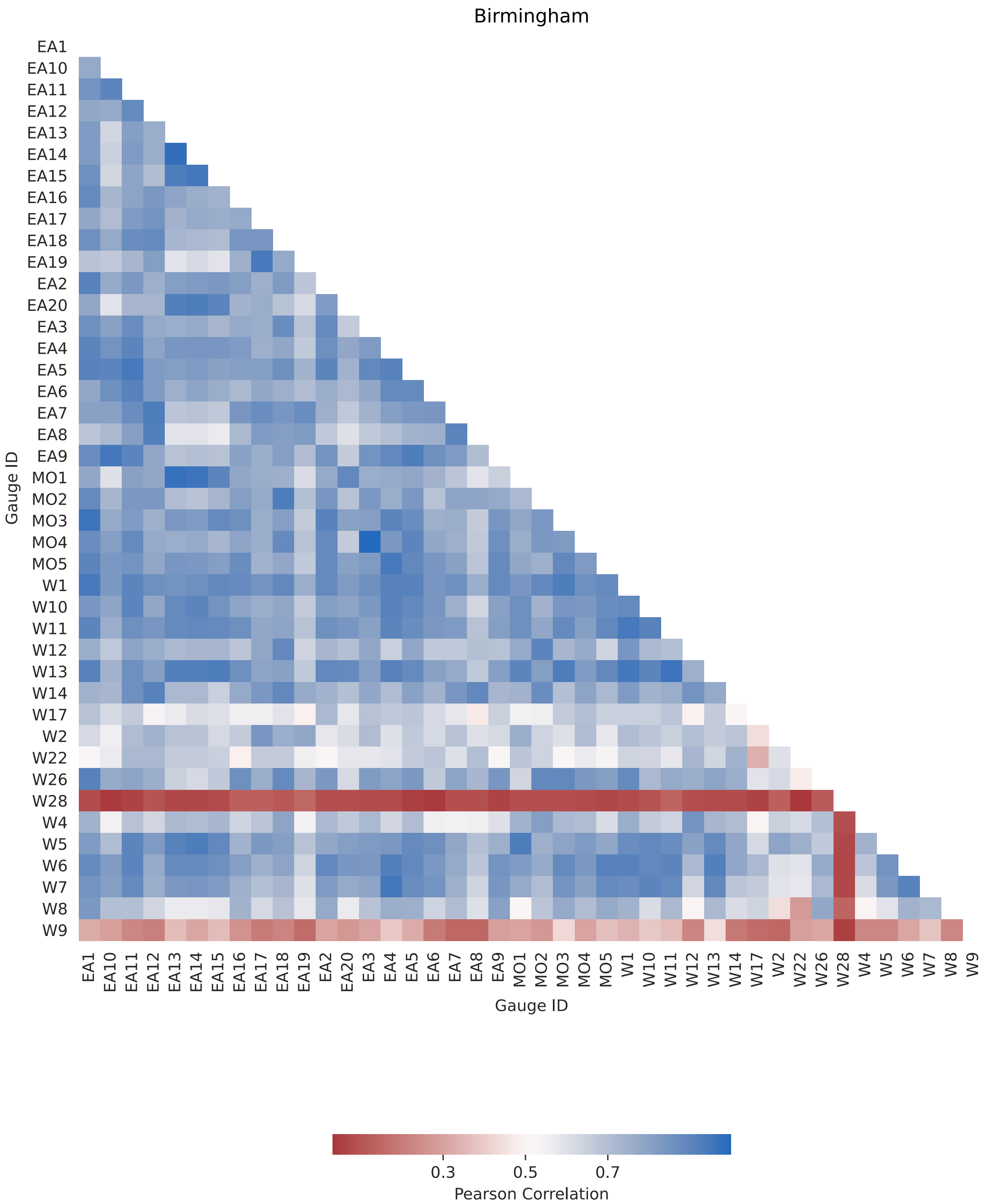
ID	Time series gaps	Data missing within 24 hours	Potential accumulation after gap	Gap due to QC?	Event pattern anomalies	Event High Extreme	Time series anomalies	Pre QC extremes	Post QC Extremes	Annual Pearson High	Judgement	Comments
"1229657" : "W1"	complete				No		No	No	No	High	Accept	
"1270610" : "W2"	complete				No		No	No	No	High	Accept	
"40138890" : "W3"	complete				No		No	Yes	Yes	NA	Reject	up to 2016 multiple gaps and erratic reporting
"40648046" : "W4"	complete				No		No	No	Yes	High	Accept	
"42644361" : "W5"	complete				No		No	Yes	Yes	High	Accept	
"53696116" : "W6"	complete				No		No	Yes	Yes	High	Accept	
"403336212" : "W7"	complete				No		No	Yes	No	High	Accept	
"551146108" : "W8"	complete				No		No	No	Yes	High	Accept	
"632296132" : "W9"	complete				No		No	Yes	Yes	Low	Reject	
"888466001" : "W10"	complete				No		No	No	No	High	Accept	
"892026001" : "W11"	complete				No		No	No	No	High	Accept	
"915006001" : "W12"	complete				Yes		No	No	No	High	Accept	40mm on 14th
"927156001" : "W13"	complete				No		No	No	Yes	High	Accept	
"222261" : "W14"	incomplete	Yes					No	No	No	High	Reject	Data missing on 16th
"23593182" : "W15"	incomplete						No	Yes	Yes	NA	Reject	No data for duration of event
"24697037" : "W16"	incomplete						No	No	No	NA	Reject	No data for duration of event
"27031057" : "W17"	incomplete	No	Yes	No	No		No	Yes	Yes	High	Accept	Remove data from 12/06/2016 11:00:00
"30826934" : "W18"	incomplete	No	No	No	No		No	Yes	Yes	NA	Accept	
"36916542" : "W19"	incomplete						No	No	Yes	NA	Reject	No data for duration of event
"38105363" : "W20"	incomplete						No	Yes	Yes	NA	Reject	No data for duration of event
"871176079" : "W21"	incomplete	No	No	Yes	Yes		No	Yes	No	NA	Reject	Large gaps in data
"876666007" : "W22"	incomplete	Yes	Yes	No	No		No	Yes	No	High	Reject	Large gap on 16th
"890726001" : "W23"	incomplete	No	No	No	No		No	Yes	No	High	Accept	
"928326002" : "W24"	incomplete	No	Yes	No	No		No	Yes	Yes	NA	Accept	Remove data from 14/06/2016 12:00:00
"931636001" : "W25"	incomplete	No	No	No	Yes		No	Yes	No	NA	Reject	Odd oscillation from 13th, QC fail
"933056001" : "W26"	incomplete	No	No	No	No		No	Yes	Yes	NA	Accept	
"939386001" : "W27"	incomplete	Yes	Yes				No	No	Yes	NA	Reject	Mostly gaps
"945636001" : "W28"	incomplete	No	No	No	No		No	Yes	Yes	Vlow	Reject	
"973326001" : "W29"	incomplete	No	Yes	No	No		No	No	No	NA	Accept	Remove data from 13/06/2016 15:00:00

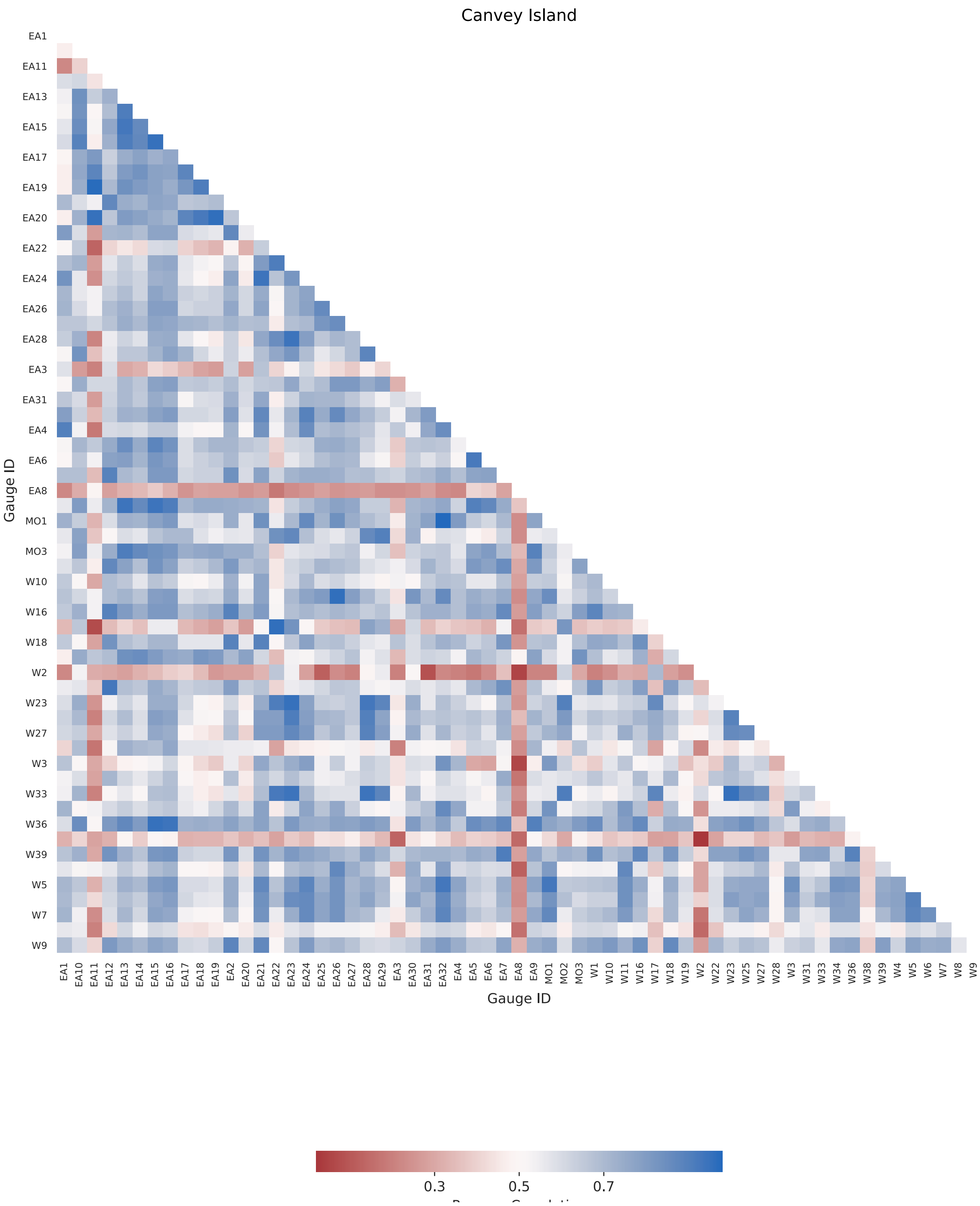
Milton Keynes manual SQC results

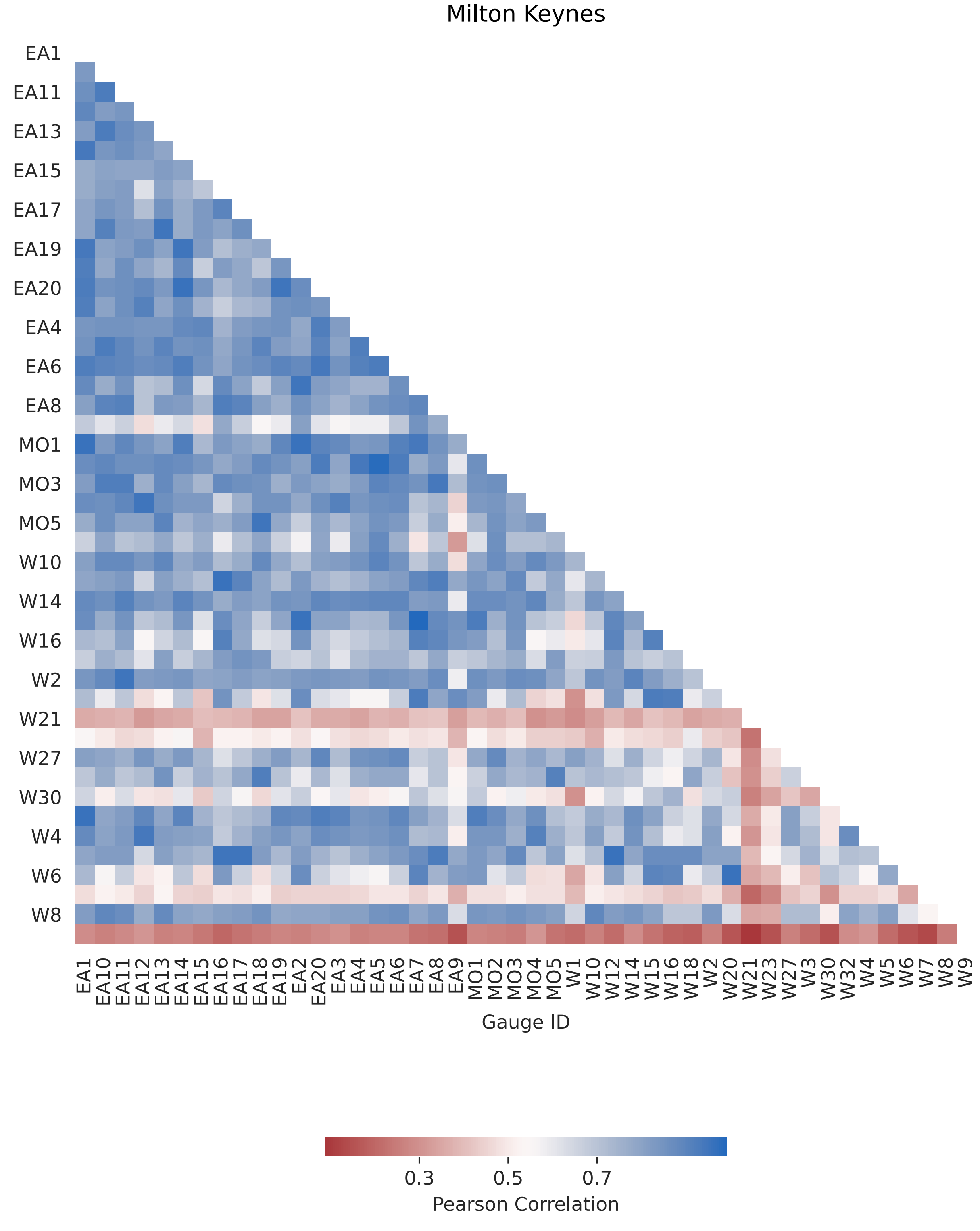
ID	Time series gaps	Data missing within 24 hours	Potential accumulation after gap	Gap due to QC?	Event pattern anomalies	Event High Extreme	Time series anomalies	Pre QC extremes	Post QC Extremes	Annual Pearson High	Judgement	Comments
"471104":"W1"	complete				Yes	No	No	Yes	Yes	High	Accept	OOOh, not much rain, is it right or wrong???
"23872180":"W2"	complete				No	No	No	No	Yes	High	Accept	
"24016212":"W3"	complete				No	No	No	Yes	Yes	High	Accept	
"24112255":"W4"	complete				No	No	No	Yes	No	High	Accept	
"27817197":"W5"	complete				No	No	No	Yes	Yes	High	Accept	
"31240469":"W6"	complete				Yes	No	No	No	No	High	Accept	High on 27/05/2018 19:00:00 30mm
"41509140":"W7"	complete				No	No	No	No	Yes	Med	Accept	
"41859025":"W8"	complete				No	No	No	No	Yes	High	Accept	
"68123610":"W9"	complete				No	No	No	Yes	Yes	Low	Reject	Poor correlation plus some odd highs
"529146002":"W10"	complete				No	No	No	No	Yes	High	Accept	
"531536045":"W11"	complete				No	No	No	Yes	Yes	NA	Accept	
"654116126":"W12"	complete				No	No	No	No	No	High	Accept	
"900796001":"W13"	complete				No	No	No	Yes	No	NA	Accept	
"915126001":"W14"	complete				No	No	No	No	No	High	Accept	
"940356001":"W15"	complete				No	No	No	No	Yes	High	Accept	
"942966001":"W16"	complete				No	No	No	No	No	High	Accept	
"963626001":"W17"	complete				No	No	No	No	No	High	Accept	
"9cf3666a-f9dc-e711-9402-0003ff5970c1":"W18"	complete				No	No				High	Accept	
"df552bf9-1958-e811-bd6d-0003ff5991d1":"W19"	complete				No	No				NA	Accept	
"234034":"W20"	incomplete	Yes	No	No	No	No	No	Yes	No	High	Accept	
"32867993":"W21"	incomplete	Yes	No	No	No	No	Yes	Yes	Yes	NA	Reject	Multiple erroneous highs in timeseries
"40138584":"W22"	incomplete	Yes					No	No	No	NA	Reject	No data for duration of event
"62594004":"W23"	incomplete	Yes	No	No	Yes	No	No	Yes	Yes	Med	Reject	Extreme value removed by QC, unsure abour remaing data
"73621774":"W24"	incomplete	Yes					No	No	No	NA	Reject	No data for duration of event
"380346051":"W25"	incomplete	Yes					No	Yes	Yes	NA	Reject	No data for duration of event
"905606001":"W26"	incomplete	Yes					No	No	No	NA	Reject	Multiple gaps in timeseries
"916386001":"W27"	incomplete	No	No	No	No	No	No	No	No	High	Accept	
"936256001":"W28"	incomplete	Yes	Yes	No	Yes		No	Yes	Yes	NA	Reject	Too many gaps during event
"940096001":"W29"	incomplete	Yes	Yes	No	No	No	No	No	No	NA	Accept	remove accumulation on 27/05/2018 12:00:00
"20432d26-a5ec-e611-93ff-000d3ab1ce94":"W30"	incomplete	Yes	No	No	No	No	No	No	Yes	High	Accept	
"8f6594e8-c3b4-e611-9408-0003ff59aed1":"W31"	incomplete						No	No	No	NA	Reject	No data for duration of event
"9e7320ef-5365-e611-9402-000d3ab1c008":"W32"	incomplete	Yes	No	No	No	No	No	Yes	Yes	High	Accept	
"dc612ee0-e78a-e611-9401-0003ff59aed0":"W33"	incomplete						No	Yes	yes	High	Reject	Multiple gaps in timeseries
"e23924c1-aad5-e611-9401-000d3ab1ce9d":"W34"	incomplete						No	Yes	Yes	NA	Reject	No data for duration of event

Pearson Correlation Matrices




Birmingham
EA1
EA10
EA11
EA12
EA13
EA14
EA15
EA16
EA17
EA18
EA19
EA2
EA20
EA3
EA4
EA5
EA6
EA7
EA8
EA9
MO1
MO2
MO3
MO4
MO5
W1
W10
W11
W12
W13
W14
W17
W2
W22
W26
W28
W4
W5
W6
W7
W8
W9
Gauge ID
Gauge ID
0.3
0.5
0.7
Pearson Correlation


Canvey Island
Gauge ID
Gauge ID
0.3
0.5
0.7



Milton Keynes
EA1
EA11
EA13
EA15
EA17
EA19
EA20
EA4
EA6
EA8
MO1
MO3
MO5
W10
W14
W16
W2
W21
W27
W30
W4
W6
W8
EA1 EA10 EA11 EA12 EA13 EA14 EA15 EA16 EA17 EA18 EA19 EA2 EA20 EA3 EA4 EA5 EA6 EA7 EA8 EA9 MO1 MO2 MO3 MO4 MO5 W1 W10 W12 W14 W15 W16 W18 W2 W20 W21 W23 W27 W3 W30 W32 W4 W5 W6 W7 W8 W9
Gauge ID
0.3
0.5
0.7
Pearson Correlation

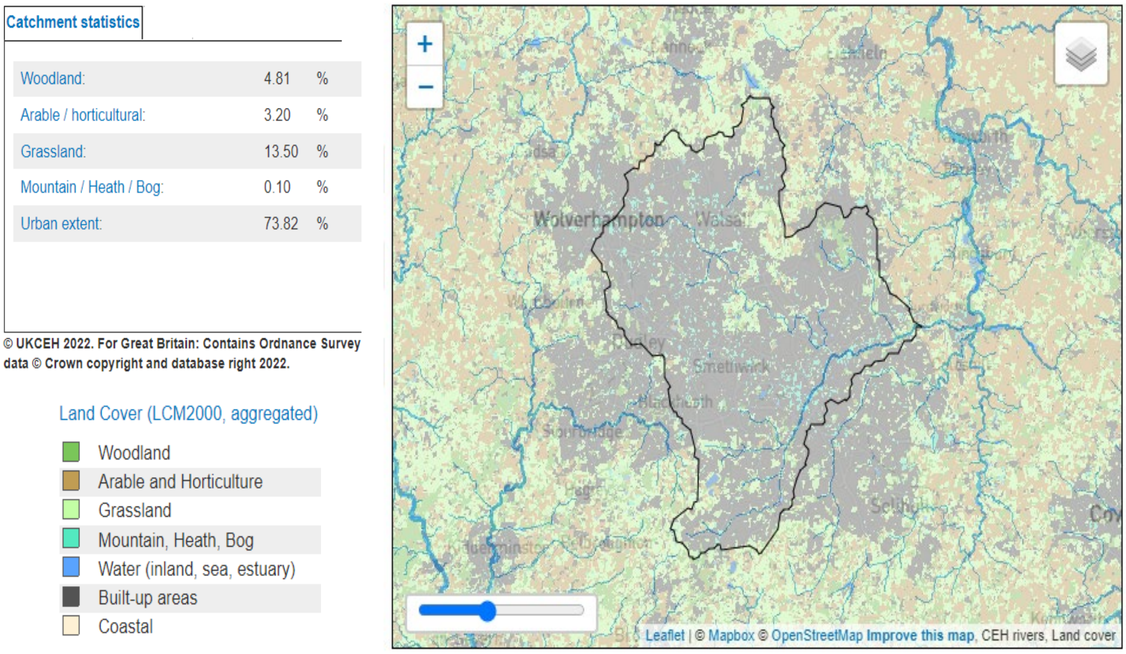
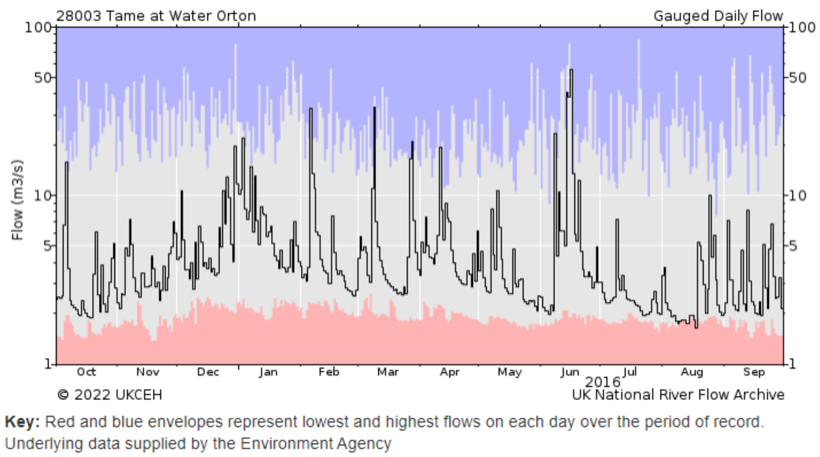


Figure showing the extent of catchment with land cover from NRFA (Centre for Ecology and Hydrology, 2022).



2016 Annual Hydrograph from Gauge 28003 Tame at Water Orton from NFRA (Centre for Ecology and Hydrology, 2022).

SHETRAN parameters modified from automated set-up

- Water flow strickler value = 100
- Overland flow strickler = 15
- Urban areas - impermeable

Date	CEHGEAR Total Rain (mm)	Blend Total Rain (mm)	Radar Total Rain (mm)
06/06/2016	0	0	0
07/06/2016	147	188	605
08/06/2016	5610	6101	8602
09/06/2016	46	0	0
10/06/2016	4508	3359	5186
11/06/2016	445	908	1435
12/06/2016	2165	2135	2544
13/06/2016	1166	1967	1781
14/06/2016	8936	8024	10333
15/06/2016	10803	8703	9698
16/06/2016	13820	12112	11117

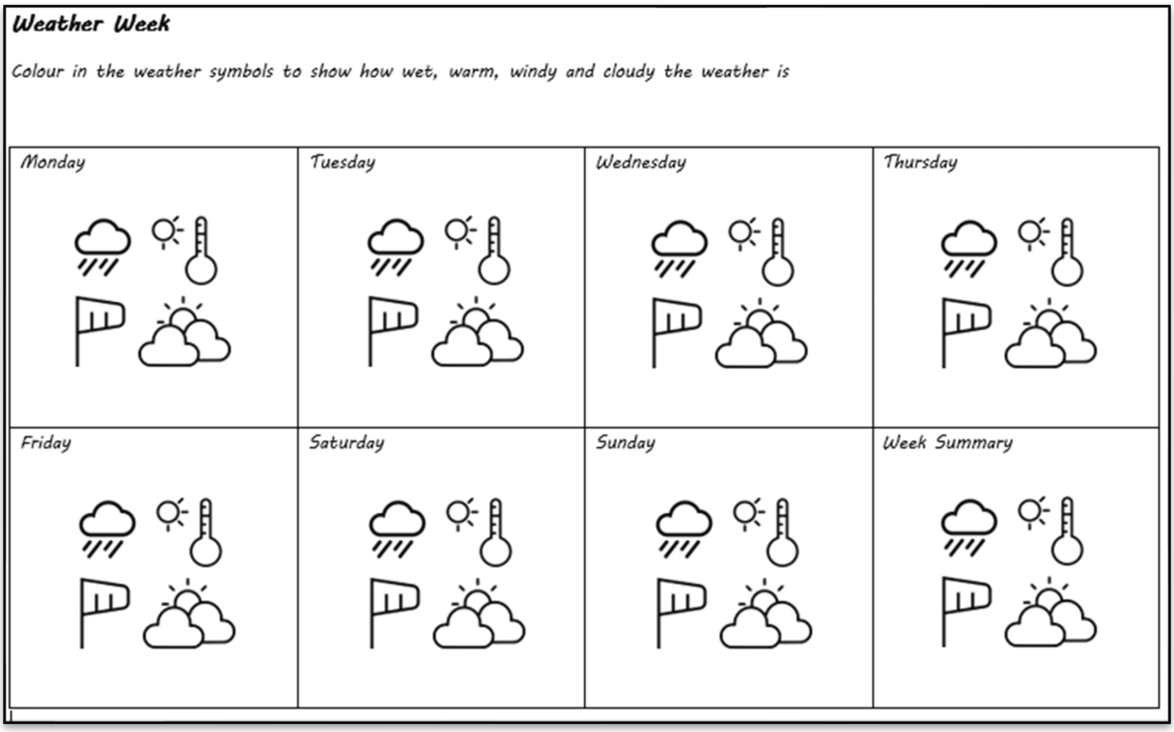
Table showing total daily rain for the catchment for CEHGEAR 1Hr, Blended and Radar datasets (highest daily total highlighted for each day)

The simulated and observed outflows around the peak on the 16th-17th June are shown in the following table:

Date/Time	Simulated CEHGEAR Outflow (m^3/s)	Simulated Blend Outflow (m^3/s)	Simulated Radar Outflow (m^3/s)	Observed Outflow (m^3/s)
16/06/2016 21:00	79	72	94	100
16/06/2016 22:00	130	110	147	111
16/06/2016 23:00	180	147	176	114
17/06/2016 00:00	209	170	180	110
17/06/2016 01:00	214	178	174	101
17/06/2016 02:00	207	177	163	89
17/06/2016 03:00	196	169	152	79

Table showing simulated and observed hourly outflows at gauge 28003 around time of AMAX, 17th June 2016

Appendix 4



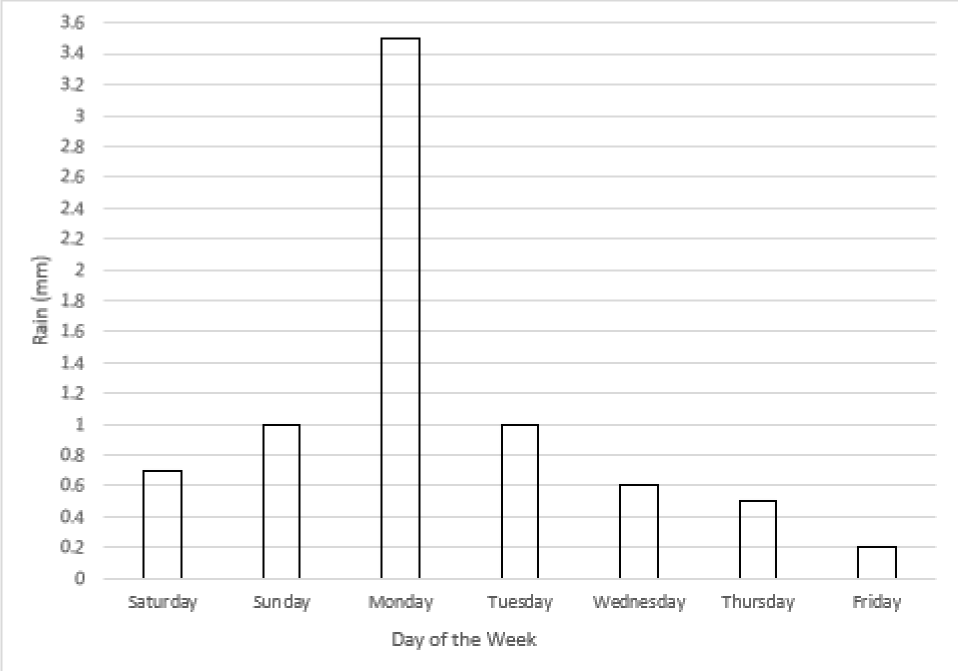
Activity sheet for weather recording by Middle School age children.

Rainy Days Exercise

The graph shows rainfall at Whitley Bay for one week in May this year.

Tasks

1. Colour the tallest bar blue
2. Colour the shortest bar red
3. Colour the Wednesday bar in green
4. Use a ruler to help mark the top of each bar on the rain axis on the left of the graph



Questions

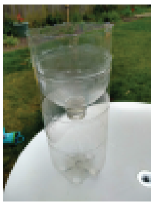
1. On which day did most rain fall?
2. On which day did least rain fall?
3. How many millimetres of rain fell on Wednesday?
4. How many millimetres of rain fell that week?

Activity sheet for weather data interpretation by Middle School age children

RECORD RAIN AT HOME

WHAT YOU NEED

- A plastic pop bottle
- Scissors
- A bucket or pot
- Some soil/sand/stone
- A measuring device eg a syringe from a medicine bottle or a 5ml teaspoon
- A ruler of tape measure



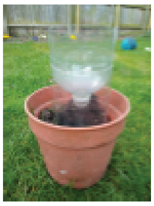
MAKE THE GAUGE & STABILISE

- Very carefully cut the top off the bottle with scissors
- Turn the top upside down to make a funnel
- Put the funnel in the bottle
- Measure the diameter of the gauge

Your rain gauge is very light and might fall over.

- Place the bottle in a pot, bucket, or dig a small hole in the ground to keep the bottle upright
- Put sand, soil or stones around the bottle to hold it

You have a rain gauge!

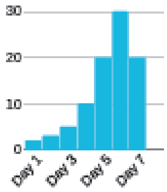


RECORD A STORM

- When rain is forecast put your rain gauge out
- Write down the start and end time of the rain and the date
- Measure the volume of rain in the gauge
- Share your results with us!

CREATE TIME SERIES DATA

- At 9 am everyday measure how much rain is in your gauge
- Use any measuring device you have to record the volume of rain
- Write the measurements in a table with the date and volume
- Share your results with us!



THINK LIKE A SCIENTIST

Make a graph;

- Put the date along the bottom (horizontal axis - x)
- Put the amount of rain going up (vertical axis - y)
- Which day had the most rain?
- Which day(s) had the least rain?

Share your results at https://forms.ncl.ac.uk/view.php?id=8315905

Newcastle University

Instructions for making a pop bottle rain gauge

Dear [WINNER],

Congratulations on winning! Apologies for the delay in getting this out to you, after all the difficulty in contacting you, I then had to wait for purchasing. Please find enclosed a Netatmo weather station including an indoor base unit, an outdoor module and a rain gauge. Hopefully setting it up is straight forward, but if you have any questions, please feel free to contact me at the number or email address above.

The outdoor module gives the most accurate readings when it is not in direct sunlight, but air can move around it. A bird table with a roof can be a good location, or a north facing wall/fence (shaded). The rain gauge works best when there is nothing above it to shield the rain – watch out for trees and tall fences etc. that could limit how much rain falls into the funnel. You might need to experiment with the best location for Wi-Fi signal, keeping 'line of sight' between the modules helps. Please share your data to the weather map and the Met Office WOW website so other people can benefit from your results!

As we are doing research on people recording the weather, it would be fantastic to hear about your experience with the Netatmo. I'd like to give you a call in a couple of weeks. If you could fill in the table on the back of this letter, it would be great to get your feedback.

Please could you confirm receipt by email?

All the best, Tess.

Weather station raffle winner letter.

Project Introduction - for groups in the NE, NW and Yorkshire interested in the weather and flooding

Have you thought about getting an automated weather station but would like some help understanding how it works and how to set it up? Or perhaps you already have a weather station and you want to know more about the readings you are getting? We would love to talk to you about the possibilities and see if we can work together.

We are researchers at Newcastle University working with rainfall data from home weather stations. We are looking for community groups in places at risk of flooding who are collecting (or would like to collect) weather data, including rainfall. The project aims to help communities understand how the weather effects them, and to see if the data can improve flood forecasting and resilience. We are working with the UK Met Office and flood forecasters, so the project could have a real impact on how floods are managed in England.

We can offer;

- An introductory workshop with activities and presentations on weather and climate change (suitable for ages 7+).
- Activities explaining natural flood management, the impact of monitoring and flash flooding.
- Demonstrations on how to set up a weather station and ensure the data is recorded and uploaded.
- Data analysis and discussion on what the recorded weather data means, to put your results into context.
- On-going technical support over the duration of the project.
-

There are many different types of automated weather station available ranging in price from around £100 to upwards of £1000. We can help you decide what type of system would best suit your needs, but at this stage we do not have funding to buy the weather station for you.

We ask that;

- You live or work in an area prone to flooding;
- You have a community group or a group of friends who want to start collecting weather data;
- You commit to buying a weather station and running it for at least a year; and,
- You can host an introductory workshop and participate in activities.

The project runs throughout 2019. We would like to start working with groups as soon as possible, so if you are at all interested, please get in touch. Please contact Tessa Gough and Elizabeth Lewis at t.gough2@newcastle.ac.uk by the 14th of February if you would like to know more.

Project Invitation Letter

The Importance of Meta Data and Citizen Science Engagement

Bringing the Crowd to the Cloud – Citizen Science and Private Automated Weather Stations

Making weather observations is tricky business, there are lots of ways data can be skewed. Some of the typical ways weather can be misrepresented are sheltered wind vanes, thermometers positioned in full sun and rain gauges located under trees. For official weather stations a lot of time and effort goes into getting the setting exactly right. For citizen science the possible locations may not be perfect. It's good to position the weather station in the best place possible and to provide details on the location so data users can make a judgement call on how observations might be affected.

The Met Office scores weather station locations based on exposure and the positions of the rainfall and temperature gauges. The station score is visible to all users of the online platform (WOW). The details of the site setting are not readily available.

Exercise

Look at the scenarios presented in the table and use the accompanying "Station Metadata" document to determine the exposure, temperature, and rainfall rating and therefore the star rating for each location.

Points for Discussion

- Was the rating easy to calculate?
- How would you feel about your star rating as a budding citizen scientist?
- As a data user would the star rating tell you what you needed to know?
- Any other points?

Tessa Gough, PhD researcher Newcastle University.

Please contact cspaws@newcastle.ac.uk if you would like to know more about "Bringing the Crowd to the Cloud" including participating in workshops and accessing support for recording weather observations.

Weather Station	Location	Exposure	Temperature	Rainfall	Site Rating
 £950	School playing field, 150m from 12m high school building. Pole mounted combined rain, temperature and wind modules at 1.5m	5	A	D	2*
 £700	Golf club, 10m from 5m high club house, pole mounted combined rain, temperature and wind modules at 2m	3	A	D	2*
 £170	Back garden, 6m from 8m high house. Ground mounted rain gauge at 30cm, temperature module on north facing window ledge	1	D	C	1*

Site rating exercise.

Setting up a Weather Station on the Met Office Weather Online Website

The Met Office have a website (WOW) where weather data can be shared (http://wow.metoffice.gov.uk/). The following steps will guide you through setting up an account and providing details about you weather station.

The way of uploading data to WOW changes depending on the type of weather system. There are some helpful websites linked on WOW (http://wow.metoffice.gov.uk/support/dataformats) to get your system connected. If you can't connect your weather station, please contact Citizen Science Weather (cspaws@newcaslte.ac.uk) and we can help you. You need to have a WOW account set up before contacting the research team about data upload.

Setting up a WOW Account

http://wow.metoffice.gov.uk/

If you click on "login" at the top right corner of the WOW homepage you will be given the choice of setting up a Met Office account or signing into WOW via your Facebook, Google or Twitter account. The Facebook and Twitter options don't work at the moment. You can sign in via Google. If you chose this option skip to Step 2. If you want to set-up a Met Office account, go to Step 1.

Step 1. Setting up a Met Office Account

Open the WOW home page and click "sign up" at the top right side of the page - http://wow.metoffice.gov.uk/

Click the box to open a new account - (https://register.metoffice.gov.uk/WaveRegistrationClient/public/newaccount.do?service=weatherobservations)

Enter your details making a note of your user name and password.

The password must be at least 9 characters and contain at least one digit, one uppercase letter, one lowercase letter and one special character. The Met Office take data security seriously!

Once you click "register" you will be sent an email that confirms your account and provides instructions on how to login. Click the link in the email to go back to the WOW site. For some reason there can be an error message saying the registration couldn't be completed. Ignore this, chances are your account is set-up.

Step 2. Login

Go to the home page (http://wow.metoffice.gov.uk/) and click "login" in the top right corner. Enter your user name and password or select the option to sign in via Google. The Facebook and Twitter links don't work.

Step 3. Enter a Site

This stage of the process takes up to 20 minutes. Only the latitude, longitude, site name and time zone are required information, however, it would be appreciated if you can complete as much of the form as possible as it can be really useful in understanding the data. If you are short on time you can put in the essential information only to create the site, then edit the details later to add further information.

From the WOW home page select "Enter a Site" from the black ribbon at the top of the page, this takes you to the site set up page.

1. Choose a location for this site

The easiest way is to click the "Search a Location" box and enter your place name. You can then pick up and move the place marker displayed in the map to your location. You can be as precise in your location as you chose to be. As you move the marker, you will see the latitude and longitude update. You must enter the height above sea level manually. If you don't know the elevation you can zoom into your location on https://www.freemaptools.com/elevation-finder.htm and the marker will display your elevation as "height above sea level in metres" for you to enter into WOW.

You can add site images at this point. Images can be useful for data users to understand the location of the weather station, but they are not compulsory. If you take photos, use a compass to find the north, south, east and west sides of the weather station. The photo is of the side of the weather station facing the given direction, meaning you will be facing the opposite way to take the picture, e.g., you face south to photograph the north facing side of the weather station. To remember which photo is which as you take it, it is recommended you do north, east, south, then west (so face south, west, north, then east).

2. Site Details

Chose a site name. You do not have to use your address or any identifying name, e.g., "North Town Weather Station" is fine. The name is not desperately important despite being a required piece of information as your weather station will be assigned a unique identifying ID. The name helps you pick out a weather station if you decide to add others to your account.

Select a time zone. "UTC+ 00:00 Dublin, Edinburgh, Lisbon, London" is for the UK and Ireland.

Create an Authentication Key – to automatically upload data you will need to create a 6-digit number here (it also gets referred to as a PIN and passcode). Pick any 6 digits and make a note of it.

You can choose if you want to add a site description and logo. It is helpful if you can state the type of weather station you are using (brand and model number) in the "Site Description" box.

3. Site Owner Details

All the boxes in this section are voluntary but it's helpful if you can make the "default access level" public as this allows data to be shared.

4. Site Data Preferences

Site Location Attributes
To complete this section is it helpful to refer to the information on the support page http://wow.metoffice.gov.uk/support/siteratings. Try to be as accurate as possible as it is very

helpful for data users to know the circumstances observations have been made in. The site rating is likely to be 1 or 2 stars for a private site as a lot depends on the space around the weather station. Don't be disheartened, all data is useful.

Site Data Preferences
Click the arrow on the right-hand side of the section for each group of parameters and select either "Not Captured" or the units relevant to your weather station.

5. Additional Information

If there is anything else you think could be of interest to data users you can add it here. If you have calibrated your weather station, you could add the details here or make note of any obstructions that might affect the observations.

Once you have filled in all the boxes click "Save Changes" and you site is now set up, congratulations!

Step 4. Add an Observation
Once you have your site set up you can start adding weather observation data. Click the "Add Observation" button, or if you need a break before continuing you can get to the same point later by clicking "Enter Observation" on the black ribbon at the top of the WOW home page.

You can choose to add a manual observation or do a bulk upload. Ideally if you have an automated weather system it is preferable to have an automated data upload. There are some instructions on the WOW support page http://wow.metoffice.gov.uk/support/dataformats however if you are unable to connect your weather station to WOW or need support with uploading a csv data file please contact the research team at cspaws@newcastle.ac.uk for assistance.

Congratulations, you should now be sharing weather data with WOW. Wow! Thanks.

Instructions for following the process for setting up a weather station on WOW.

Station Metadata

Location Attributes

WOW lets you provide details about a number of attributes that help other WOW users and the Met Office understand the surrounding environment. These attributes have been compiled based on site grading schemes used by the Climatological Observers Link (COL), the World Meteorological Organisation (WMO) and the Met Office.

http://wow.metoffice.gov.uk/support/siteratings

How site ratings are calculated

Each site is automatically allocated a 'site rating" based on the observing location attributes entries submitted on site registration. The system is based on the quality and exposure of the temperature and rainfall data:

5* = E5, T=A, R=A
4* = E >= 3, T=A, R=A
3* = E >= 3, T[=A,B or C], R[=A,B or C]
2* = E >= 1, T[=Any], R[=Any]
1* = E =0,1,R or U, T[=Any], R[=Any]

(Where E = Exposure, T = Temperature, and R = Rainfall).

*If temperature is measured at a site, but not rainfall, the site rating will be based on the quality and exposure of the temperature data alone. If rainfall is measured at a site, but not temperature, the site rating will be based on the quality and exposure of the rainfall data alone. If there is no temperature or rainfall data, the site will be classed as 1**

Exposure (E)

Exposure ratings relate to the site of the temperature and rainfall instruments only, which should ideally be at ground level. Sensors for sunshine, wind speed etc are best exposed as freely as possible, and rooftop or mast mountings are usually preferable.

Exposure guidelines are based on a multiple of the height h of the obstruction above the sensor height; the standard is a minimum distance of twice the height (2h). Thus, for a rain gauge at 30 cm above ground, a building 5 m high should be at least 9.4 m distant (5 m less 0.3 m, x 2), and a 10 m building should be at least 17 m from a thermometer screen (10 m less 1.5 m, x2)

5: Very open exposure: no obstructions within 10h or more of temperature or rainfall instruments.
4: Open exposure: most obstructions/heated buildings 5h or from temperature or rainfall instruments, none within 2h.
3: Standard exposure: no significant obstructions or heated buildings within 2h of temperature or rainfall instruments.
2: Restricted exposure: most obstructions/heated buildings >2h from temperature or rainfall instruments, none within 1h.
1: Sheltered exposure: significant obstructions or heated buildings within 1h of temperature or rainfall instruments.
0: Very sheltered exposure: site obstructions or sensor exposure severely limit exposure to sunshine, wind, rainfall.

1

R: Rooftop site: Rooftop sites for temperature and rainfall sensors should be avoided where possible.
T: Traffic site: equipment sited adjacent to public highway.
U: Exposure unknown or not stated.

Measurements of air temperature (T)
STANDARD INSTRUMENTS in this context means: Calibrated mercury-in-glass thermometers or calibrated electronic temperature sensors.

A: Standard instruments in Stevenson Screen, calibration within last 10 yr., site exposure minimum rating=3.
B: Standard instruments in Stevenson Screen or manufacturer supplied AWS radiation screen, calibration within last 10 yr., site exposure = 2 or 3.
C: Standard instruments in Stevenson Screen or manufacturer supplied AWS radiation screen, site exposure 1 or less.
D: Non-standard instruments and/or no or non-standard radiation screen and/or sheltered site, site exposure 1 or less.
U: Instruments unknown or not stated.
0: No air temperature measurements made at this site.

Measurements of rainfall (R)
STANDARD INSTRUMENTS in this context means: Standard-pattern (Snowdon or Met Office Mk II pattern) "five-inch" copper rain gauge, with deep funnel, the rim of the gauge level and mounted at 30 cm above ground level, meeting the minimum exposure requirement of being at least 'twice the height" of the obstacle away from the obstacle.

A: Standard "five inch" manually read rain gauge or calibrated tipping-bucket rain gauge, at standard height above ground (30 cm), site exposure minimum = 3.
B: Standard "five inch" manually read rain gauge or calibrated tipping-bucket rain gauge, the rim mounted at standard height above ground (30 cm), exposure = 2 or 3.
C: Standard "five inch" manually read rain gauge or calibrated tipping-bucket rain gauge, the rim mounted at standard height above ground (30 cm), exposure 1 or less.
D: Non-standard rain gauge and/or tipping-bucket rain gauge, exposure 1 or less.
U: Instruments unknown or not stated.
0: No rainfall measurements made at this site.

Measurements of wind
A: Wind sensors calibrated within last 10 years, mounted 10m above the ground on mast or pole, with no obstructions within 100m.
B: Wind sensors mounted above the ground on mast or pole, with no obstructions within 50m.
C: Wind sensors mounted on building or wall.
U: Instruments unknown or not stated.
0: No wind measurements made at this site.

2

Urban Climate Zone Index (UCZ)
1: Intensely developed urban zone with detached close-set high-rise buildings with cladding, e.g. downtown towers.
2: Intensely developed high density urban with 2 - 5 storey, attached or very close-set buildings often of brick or stone, e.g. old city core.
3: Highly developed, medium density urban with row or detached but close-set houses, stores & apartments e.g. urban housing
4: Highly developed, low density urban with large low buildings & paved parking, e.g. shopping mall, warehouses.
5: Medium development, low density suburban with 1 or 2 storey houses, e.g. suburban housing.
6: Mixed use with large buildings in open landscape, e.g. institutions such as a hospital, university, airport.
7: Semi-rural development with scattered houses in natural or agricultural area, e.g. farms, estates.
U: UCZ unknown or not stated.
UCZ descriptions as defined by the World Meteorological Organisation (WMO-No.8, 8th Edition)

Reporting hours
A: Will always aim to provide a weather report at 09:00 GMT. Daily temperature and rainfall values relate to standard 24 hour period morning to morning.
B: Will always aim to provide a weather report between 06:00 and 09:00 GMT. Daily temperature and rainfall values relate to standard 24 hour period morning to morning.
C: Daily temperature and rainfall values relate to the 24 hour period midnight to midnight. This is the default for most automatic weather stations.
D: Air temperature and rainfall terminal hour is other than A, B or C above, or extremes do not relate to 24 hour periods.
U: Reporting hours unknown or not stated.

The WOW explanation for determining site rating and provision of metadata, adapted from WOW support pages.

References

AECOM (2019) *Independent flood review, Milton Keynes, 27th May 2018 Flood Event*.
Allerup, P., et al. (1980) 'Accuracy of point precipitation measurements', *Nordic Hydrology*, 11(2).
Bárdossy, A., et al. (2022) 'Is Precipitation Responsible for the Most Hydrological Model Uncertainty?', *Frontiers in Water*, 4.
Bárdossy, A., et al. (2018) 'Intensity–duration–frequency curves exploiting neighbouring extreme precipitation data', *Hydrological Sciences Journal*, 63(11).
Bárdossy, A., et al. (2021) 'The use of personal weather station observations to improve precipitation estimation and interpolation', *Hydrol. Earth Syst. Sci*, 25, pp. 583-601.
BBC (2014) *Clean-up underway after flash flooding*. Available at: https://www.bbc.co.uk/news/uk-england-cambridgeshire-28721657.
BBC (2016) *Birmingham flash flooding causes more chaos*. Available at: https://www.bbc.co.uk/news/uk-england-birmingham-36556884
BBC (2022) 'Fresh wave of sewage pollution hits Britain's beaches' Marshall, C. *BBC* pp. https://www.bbc.co.uk/news/science-environment-62813573.
Bell, S., et al. (2013) 'The state of automated amateur weather observations', *Weather*, 68(2), pp. 36-41.
Bell, S., et al. (2015) 'How good are citizen weather stations? Addressing a biased opinion', *Weather*, 70(3), pp. 75-84.
Berg, P., et al. (2013) 'Strong increase in convective precipitation in response to higher temperatures', *Nature Geoscience*, 6(3), pp. 181-185.
Berne, A., et al. (2013) 'Radar for hydrology: Unfulfilled promise or unrecognized potential?', *Advances in Water Resources*, 51, pp. 357-366.
Bevacqua, E., et al. (2019) 'Higher probability of compound flooding from precipitation and storm surge in Europe under anthropogenic climate change', *Science advances*, 5(9), p. eaaw5531.
Beven, K. (2012) 'Down to Basics: Runoff Processes and the Modelling Process', in *Rainfall-Runoff Modelling*. pp. 1-23.
Beven, K. (2019) 'Towards a methodology for testing models as hypotheses in the inexact sciences', *Proceedings of the Royal Society A*, 475(2224), p. 20180862.
Beven, K., et al. (2020) 'Developing observational methods to drive future hydrological science: Can we make a start as a community?', *Hydrological Processes*, 34(3), pp. 868-873.
Birkinshaw, S.J., et al. (2021) 'Improved hydrological modelling of urban catchments using runoff coefficients', *Journal of Hydrology*, 594, p. 125884.
Blenkinsop, S., et al. (2017) 'Quality-control of an hourly rainfall dataset and climatology of extremes for the UK', *INTERNATIONAL JOURNAL OF CLIMATOLOGY Int. J. Climatol*, 37, pp. 722-740.
Brooks, H.E. (2013) 'Severe thunderstorms and climate change', *Atmospheric research*, 123, pp. 129-138.
Bruno, F., et al. (2014) 'Spatial reconstruction of rainfall fields from rain gauge and radar data', *Stochastic Environmental Research and Risk Assessment*, 28(5), pp. 1235-1245.
Butler, M.K. (2019) 'Personal weather stations and sharing weather data via the Internet', *Weather*, 74(1).

Buytaert, W., et al. 2 (2014) 'Citizen science in hydrology and water resources: Opportunities for knowledge generation, ecosystem service management, and sustainable development' *Frontiers in Earth Science*.
Callister Jr, T. (2000) 'Media literacy: On-ramp to the literacy of the 21st century or cul-de-sac on the information superhighway', *Advances in Reading/Language Research*, 7, pp. 403-420.
Cambridge Science Park (2020) *About*. Available at: https://www.cambridgesciencepark.co.uk/community/lifestyle/bar-hill/
Cambridgeshire County Council (2014) *Flood Investigation Report - Bar Hill*.
Centre for Ecology and Hydrology (2022) *National River Flow Archive Tame at Water Orton*. Available at: https://nrfa.ceh.ac.uk/data/station/info/28003.
Ciach, G.J., et al. (1999) 'Radar–Rain Gauge Comparisons under Observational Uncertainties', *Journal of Applied Meteorology*, 38(10), pp. 1519-1525.
Clark, M.R., et al. (2018) 'Fine-scale analysis of a severe hailstorm using crowd-sourced and conventional observations', *Meteorological Applications*, 25(3).
Collier, C.G. (1989) *Applications of weather radar systems*. Wiley.
Community Collaborative Rain Hail and Snow Network (1998) *Home Page*. Available at: https://www.cocorahs.org/.
Country Reports (2023) *Netherlands Geography*. Available at: https://www.countryreports.org/country/Netherlands/geography.htm.
Coutu, S., et al. (2012) 'Parsimonious hydrological modeling of urban sewer and river catchments', *Journal of Hydrology*, 464, pp. 477-484.
D Tooley (2020) *We must not forget the lessons learned from Milton Keynes flooding*. Available at: https://www.miltonkeynes.co.uk/news/politics/council/we-must-not-forget-lessons-learned-milton-keynes-flooding-2085259.
Davids, J.C., et al. (2019) 'Soda bottle science—citizen science monsoon precipitation monitoring in Nepal', *Frontiers in Earth Science*, 7.
Davis Instruments (2018) *Weather Link*. Available at: https://www.weatherlink.com/map (Accessed: November 2017).
de Vos, L., et al. (2017) 'The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam', *Hydrol. Earth Syst. Sci*, 21, pp. 765-777.
de Vos, L.W., et al. (2019) 'Quality Control for Crowdsourced Personal Weather Stations to Enable Operational Rainfall Monitoring', *Geophysical Research Letters*, 46(15).
Delfiner, P. (1975) 'Optimum interpolation by kriging', *Display and analysis of spatial data*, pp. 96-114.
Dirks, K., et al. (1998) 'High-resolution studies of rainfall on Norfolk Island: Part II: Interpolation of rainfall data', *Journal of Hydrology*, 208(3-4), pp. 187-193.
Donat, M.G., et al. (2013) 'Global land-based datasets for monitoring climatic extremes', *Bulletin of the American Meteorological Society*, 94(7), pp. 997-1006.
Durre, I., et al. (2010) 'Comprehensive automated quality assurance of daily surface observations', *Journal of Applied Meteorology and Climatology*, 49(8), pp. 1615-1633.
Eggimann, S., et al. 51 (2017) 'The Potential of Knowing More: A Review of Data-Driven Urban Water Management' *Environmental Science and Technology*.
Eitzel, M.V., et al. (2017) 'Citizen Science Terminology Matters: Exploring Key Terms', *Science: Theory and Practice,* , 2(1).
Elliott, C. (2017) 'Days when paradise was branded 'a disaster area'', *Cambridgshire Live*. Available at: https://www.cambridge-news.co.uk/news/history/days-paradise-branded-a-disaster-12518748
Endfield, G.H., et al. (2012) 'Exploring the role of the amateur in the production and circulation of meteorological knowledge', *Climatic Change*, 113(1), pp. 69-89.

Environment Agency (2017) 'Environment Agency Rainfall API'. Available at: https://environment.data.gov.uk/flood-monitoring/doc/rainfall.
Environment Agency (2021) 'Rainfall Data' *https://www.api.gov.uk/ea/rainfall/#rainfall*. August 2021. Available at: https://www.api.gov.uk/ea/rainfall/#rainfall.
Essex County Council (2014) *Flood Investigation Report, Canvey Island, Castle Point Borough*. [Online]. Available at: https://www.scribd.com/document/391823288/Flood-Investigation-Report-Can-Vey-Island.
Estevez, J., et al. (2011) 'Guidelines on validation procedures for meterological data from automatic weather stations', *Journal of Hydrology*, 402.
Eysenbach, G. (2008) *Credibility of health information and digital media: New perspectives and implications for youth*. MacArthur Foundation Digital Media and Learning Initiative.
Fabry, F.d.r. (2015) *Radar meteorology : principles and practice*. Cambridge : Cambridge University Press.
Falconer, R.H., et al. (2009) 'Pluvial flooding: new approaches in flood warning, mapping and risk management', *Journal of Flood Risk Management*, 2(3), pp. 198-208.
Faulkner, D., et al. (1999) *Flood estimation handbook*. Wallingford: Wallingford : Institute of Hydrology.
Fehri, R., et al. (2020) 'Testing a citizen science water monitoring approach in Tunisia', *Environmental Science and Policy*, 104.
Flack, D.L., et al. (2019) 'Recommendations for improving integration in national end-to-end flood forecasting systems: An overview of the FFIR (Flooding From Intense Rainfall) programme', *Water*, 11(4), p. 725.
Flanagin, A.J., et al. (2008) 'The credibility of volunteered geographic information', *GeoJournal*, 72(3), pp. 137-148.
Flood Guidance (2022) *Types of Flooding*. Available at: https://www.floodguidance.co.uk/what-is-resilience/types-flooding/.
Foody, G.M., et al. (2015) 'Accurate attribute mapping from volunteered geographic information: issues of volunteer quantity and quality', *The Cartographic Journal*, 52(4), pp. 336-344.
Fornasiero, A., et al. (2006) 'Enhanced radar precipitation estimates using a combined clutter and beam blockage correction technique', *Nat. Hazards Earth Syst. Sci.*, 6(5), pp. 697-710.
Fowler, H.J., et al. (2021) 'Anthropogenic intensification of short-duration rainfall extremes', *Nature Reviews Earth & Environment*, 2(2), pp. 107-122.
Fox, K. (2004) *Watching the English: The Hidden Rules of English Behaviour*.
Future Farming (2017) *6 tips to get started with on-farm weather stations*. Available at: https://www.futurefarming.com/Tools-data/Articles/2017/9/6-tips-to-get-started-with-on-farm-weather-stations-3076WP/ (Accessed: June 2019).
Geoghegan, H., et al. (2016) 'Understanding motivations for citizen science', *Ukeof.Org.Uk*, (May).
Glenis, V., et al. (2018) 'A fully hydrodynamic urban flood modelling system representing buildings, green space and interventions', *Environmental Modelling & Software*, 109, pp. 272-292.
Go Fair (2016) *FAIR Principles*. Available at: https://www.go-fair.org/fair-principles/.
Goodchild, M.F., et al. (2012) 'Assuring the quality of volunteered geographic information', *Spatial statistics*, 1, pp. 110-120.
Goudenhoofdt, E., et al. (2009) 'Evaluation of radar-gauge merging methods for quantitative precipitation estimates', *Hydrology and Earth System Sciences*, 13(2), pp. 195-203.
Guerreiro, S., et al. (2017) 'Pluvial Flooding in European Cities—A Continental Approach to Urban Flood Modelling', *Water*, 9(4).

Guerrini, C.J., et al. (2018) 'Citizen science, public policy', *Science*, 361(6398), pp. 134-136.
Habib, E., et al. (2001) 'Sampling errors of tipping-bucket rain gauge measurements', *Journal of Hydrologic Engineering*, 6(2), pp. 159-166.
Haklay, M. (2012) *Citizen Science and Volunteered Geographic Information – overview and typology of participation*. Berlin: Springer.
Harrison, D., et al. (2015) *AMS Conference on Radar Meteorology*.
Hawkins, E., et al. (2022) 'Millions of historical monthly rainfall observations taken in the UK and Ireland rescued by citizen scientists', *Geoscience Data Journal*, n/a(n/a).
Holley, D., et al. (2020) 'Thunderstorms and extreme rainfall in south Norfolk, 16 August 2020: meteorological analysis', *Weather*, n/a(n/a).
Hovland, C.I., et al. (1953) 'Communication and persuasion'.
IHP (2021) *Science for a Water Secure World in a Changing Environment*. UNESCO. [Online]. Available at: https://www.ncihp.si/wp-content/uploads/2021/09/IHP-IC-XXIV-Ref.1rev_IHP-IX-Strategic-Plan_in-session_fin.pdf.
Imperial War Museum (2022) *How Radar Changed the Second World War*. Available at: https://www.iwm.org.uk/history/how-radar-changed-the-second-world-war.
Irwin, A. (1995) *Citizen Science: A Study of People, Expertise and Sustainable Development* Routledge.
Jewell, S.A., et al. (2015) 'An assessment of kriging-based rain-gauge–radar merging techniques', *Quarterly Journal of the Royal Meteorological Society*, 141(691), pp. 2300-2313.
Johnson, T.J., et al. (2000) 'Using is believing: The influence of reliance on the credibility of online political information among politically interested Internet users', *Journalism & Mass Communication Quarterly*, 77(4), pp. 865-879.
Joss, J., et al. (1995) 'The application of radar gauge comparisons to operational precipitation profile corrections', *Journal of Applied Meteorology and Climatology*, 34(12), pp. 2612-2630.
Kar, K., et al. (2008) 'Handbook on community-led total sanitation'.
Kendon, E.J., et al. (2014) 'Heavier summer downpours with climate change revealed by weather forecast resolution model', *Nature Climate Change*, 4(7).
Kidd, C., et al. (2017) 'So, how much of the Earth's surface is covered by rain gauges?', *Bulletin of the American Meteorological Society*, 98(1).
Krajewski, W.F., et al. (2010) 'Radar-rainfall uncertainties: Where are we after thirty years of effort?', *Bulletin of the American Meteorological Society*, 91(1), pp. 87-94.
Kutija, V., et al. (2014) 'Model Validation Using Crowd-Sourced Data From a Large Pluvial Flood', *11th International Conference on Hydroinformatics*, (January 2017).
Lanza, L.G., et al. (2009) 'The WMO field intercomparison of rain intensity gauges', *Atmospheric Research*, 94(4), pp. 534-543.
Lengfeld, K., et al. (2020) 'Use of radar data for characterizing extreme precipitation at fine scales and short durations', *Environmental Research Letters*, 15(8), p. 085003.
Lewis, E., et al. (2018) 'Development of a system for automated setup of a physically-based, spatially-distributed hydrological model for catchments in Great Britain', *Environmental Modelling & Software*, 108, pp. 102-110.
Lewis, E., et al. (2021) 'Quality control of a global hourly rainfall dataset', *Environmental Modelling & Software*, 144, p. 105169.
Lewis, E., et al. (2022) 'Gridded estimates of hourly areal rainfall for Great Britain 1990-2016 [CEH-GEAR1hr] v2'. NERC EDS Environmental Information Data Centre. Available at: https://doi.org/10.5285/fc9423d6-3d54-467f-bb2b-fc7357a3941f.
Lewis, M.W., et al. (2017) 'Improvements in nowcasting capability: analysis of three structurally distinct severe thunderstorms across northern England on 1 July 2015', *Weather*, 72(4), pp. 91-98.

Libertino, A., et al. (2015) 'Radar Estimation of Intense Rainfall Rates through Adaptive Calibration of the Z-R Relation', *Atmosphere*, 6(10), pp. 1559-1577.
Lin, Y.-W., et al. (2016) 'Co-Observing the Weather, Co-Predicting the Climate: Human Factors in Building Infrastructures for Crowdsourced Data', *Science & Technology Studies*, 29(3), pp. 10-27.
Liu, B.F., et al. (2020) 'Eyes of the storm: How citizen scientists contribute to government forecasting and risk communication', *Weather, Climate, and Society*, 12(2).
Lopes, V.L. (1996) 'On the effect of uncertainty in spatial distribution of rainfall on catchment modelling', *CATENA*, 28(1), pp. 107-119.
Marshall, J.S. (1948) 'The distribution of raindrops with size', *J. meteor.*, 5, pp. 165-166.
McCallum, I., et al. (2016) 'Technologies to Support Community Flood Disaster Risk Reduction', *International Journal of Disaster Risk Science*, 7(2), pp. 198-204.
McGregor, P., et al. (2009) *Proceedings of the Institution of Civil Engineers-Water Management*. Thomas Telford Ltd.
McMillan, H., et al. (2012) 'Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality', *Hydrological Processes*, 26(26), pp. 4078-4111.
Met Office (2003) '1 km Resolution UK Composite Rainfall Data from the Met Office Nimrod System. '. CEDA Archive: NCAS British Atmospheric Data Centre, . Available at: https://catalogue.ceda.ac.uk/uuid/27dd6ffba67f667a18c62de5c3456350.
Met Office (2006a) 'MIDAS UK Hourly Rainfall Data'. NCAS British Atmospheric Data Centre,: September 2021. Available at: https://catalogue.ceda.ac.uk/uuid/bbd6916225e7475514e17fdbf11141c1.
Met Office (2006b) 'MIDAS: UK Daily Weather Observation Data. '. NCAS British Atmospheric Data Centre. Available at: https://catalogue.ceda.ac.uk/uuid/954d743d1c07d1dd034c131935db54e0.
Met Office (2011) *WOW online weather data platform*. Available at: https://wow.metoffice.gov.uk/ (Accessed: November 2015).
Met Office (2015) *Climate Zones*. Available at: https://www.metoffice.gov.uk/climate-guide/climate/zones (Accessed: 21/06/2019).
Met Office (2016a) *Met Office Hadley Centre observations datasets: HadUKP*. Available at: https://www.metoffice.gov.uk/hadobs/hadukp/.
Met Office (2016b) *Weather Stations*. Available at: https://www.metoffice.gov.uk/learning/making-a-forecast/first-steps/observations/weather-stations (Accessed: 3/11/18).
Met Office (2019) *The UK land observation network*. Available at: https://www.metoffice.gov.uk/learning/making-a-forecast/first-steps/observations/weather-stations (Accessed: 3/11/18).
Met Office (2021a) *Daily Weather Summary*. Available at: https://www.metoffice.gov.uk/research/library-and-archive/publications/daily-weather-summary.
Met Office (2021b) *MIDAS UK Hourly Rainfall Data.* . Available at: https://catalogue.ceda.ac.uk/uuid/bbd6916225e7475514e17fdbf11141c1 (Accessed: September 2021).
Metzger, M.J., et al. (2003) 'Bringing the concept of credibility into the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment', *Communication yearbook*, 27(1), pp. 293-335.
Metzger, M.J., et al. (2010) 'Social and heuristic approaches to credibility evaluation online', *Journal of communication*, 60(3), pp. 413-439.

Miller, J.D., et al. (2017) 'The impacts of urbanisation and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom', *Journal of Hydrology: Regional Studies*, 12, pp. 345-362.
Moriasi, D.N., et al. (2007) 'Model evaluation guidelines for systematic quantification of accuracy in watershed simulations', *Transactions of the ASABE*, 50(3), pp. 885-900.
Mounce, S., et al. (2014) 'Predicting combined sewer overflows chamber depth using artificial neural networks with rainfall radar data', *Water science and technology*, 69(6), pp. 1326-1333.
Muller, C.L., et al. (2013) 'Toward a standardized metadata protocol for urban meteorological networks', *Bulletin of the American Meteorological Society*, 94(8), pp. 1161-1185.
Muller, C.L., et al. 35 (2015) 'Crowdsourcing for climate and atmospheric sciences: Current status and future potential' *International Journal of Climatology*.
Muthusamy, M., et al. (2019) 'A remote sensing based integrated approach to quantify the impact of fluvial and pluvial flooding in an urban catchment', *Remote Sensing*, 11(5), p. 577.
Nash, J.E., et al. (1970) 'River flow forecasting through conceptual models part I—A discussion of principles', *Journal of hydrology*, 10(3), pp. 282-290.
National Atmospheric and Oceanic Administration (2015) *Multiple Radar / Multiple Sensor*. Available at: https://www.nssl.noaa.gov/news/factsheets/MRMS_2015.March.16.pdf.
National Oceanic and Atmospheric Administration (2018) *Severe Weather 101 - Thunderstorm Types*. Available at: https://www.nssl.noaa.gov/education/svrwx101/thunderstorms (Accessed: June 2022).
Natural Resources Wales (2022) 'River levels, rainfall and sea data'. Available at: https://rivers-and-seas.naturalresources.wales/?lang=en.
Netatmo *Smart Home Weather Station*. Available at: https://www.netatmo.com/en-gb/weather/weatherstation.
Netatmo Weather Map (2022) *Weather observations from Netatmo weather stations*. Available at: https://weathermap.netatmo.com/.
Ochoa-Rodriguez, S. (2017) *Rainfall estimates for urban drainage modelling: an investigation into resolution requirements and radar-rain gauge data merging at the required resolutions*. Imperial College London.
Ochoa-Rodriguez, S., et al. (2015) 'Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation', *Journal of Hydrology*, 531, pp. 389-407.
Ochoa-Rodriguez, S., et al. (2019a) 'A Review of Radar-Rain Gauge Data Merging Methods and Their Potential for Urban Hydrological Applications', *Water Resources Research*, 55(8), pp. 6356-6391.
Ochoa-Rodriguez, S., et al. 55 (2019b) 'A Review of Radar-Rain Gauge Data Merging Methods and Their Potential for Urban Hydrological Applications' *Water Resources Research*.
Parry, M.L., et al. (2007) *Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC*. Cambridge University Press.
Paul, J.D., et al. (2017) 'Citizen science for hydrological risk reduction and resilience building', *WIREs Water*, 5(1).
Peleg, N., et al. (2013) 'Radar subpixel-scale rainfall variability and uncertainty: lessons learned from observations of a dense rain-gauge network', *Hydrol. Earth Syst. Sci.*, 17(6), pp. 2195-2208.
Pollock, M.D., et al. (2018) 'Quantifying and Mitigating Wind-Induced Undercatch in Rainfall Measurements', *Water Resources Research*, 54(6).
Puttinaovarat, S., et al. (2020) 'Flood Forecasting System Based on Integrated Big and Crowdsource Data by Using Machine Learning Techniques', *IEEE Access*, 8.

Qi, Y., et al. (2016) 'A Real-Time Automated Quality Control of Hourly Rain Gauge Data Based on Multiple Sensors in MRMS System', *Journal of Hydrometeorology*, 17(6), pp. 1675-1691.
Rabiei, E., et al. (2015) 'Applying bias correction for merging rain gauge and radar data', *Journal of Hydrology*, 522, pp. 544-557.
Royal Meterological Society (2020) *In Depth - Rainfall and Relief*. Available at: https://www.metlink.org/resource/in-depth-rainfall-and-relief/ (Accessed: June 2022).
Sauvageot, H. (1994) 'Rainfall measurement by radar: a review', *Atmospheric Research*, 35(1), pp. 27-54.
Schilling, W. (1991) 'Rainfall data for urban hydrology: what do we need?', *Atmospheric Research*, 27(1), pp. 5-21.
Schroeer, K., et al. (2018) 'Strong Dependence of Extreme Convective Precipitation Intensities on Gauge Network Density', *Geophysical Research Letters*, 45(16).
Scottish Environmental Protection Agency (2022) 'Rainfall data for Scotland'. Available at: https://www2.sepa.org.uk/rainfall/.
See, L. (2019) 'A review of citizen science and crowdsourcing in applications of pluvial flooding', *Frontiers in Earth Science*, 7, p. 44.
See, L., et al. (2016) 'Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information', *ISPRS International Journal of Geo-Information*, 5(5), p. 55.
Seibert, J., et al. (2018) 'Upper and lower benchmarks in hydrological modelling', *Hydrological Processes*, 32(8), pp. 1120-1125.
Sepa (2021) 'Download Rain Data' *https://www2.sepa.org.uk/rainfall/DataDownload*. 2021/8//.
Sevruk, B., et al. (2009) 'The WMO precipitation measurement intercomparisons', *Atmospheric Research*, 92(3), pp. 376-380.
Shepard, D. (1968) *Proceedings of the 1968 23rd ACM national conference*.
Shinbrot, X.A., et al. (2020) 'Quiahua, the first citizen science rainfall monitoring network in Mexico: Filling critical gaps in rainfall data for evaluating a payment for hydrologic services program', *Citizen Science: Theory and Practice*, 5(1).
Singh, V. (1997) 'Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph', *Hydrological processes*, 11(12), pp. 1649-1669.
Smith, L., et al. (2017) 'Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework', *Journal of Flood Risk Management*, 10(3), pp. 370-380.
Starkey, E., et al. (2017) 'Demonstrating the value of community-based (‘citizen science’) observations for catchment modelling and characterisation', *Journal of Hydrology*, 548.
Steiner, M., et al. (1999) 'Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation', *Water Resources Research*, 35(8), pp. 2487-2503.
Surfers Against Sewerage (2022) *Safer Seas & Rivers Service*. Available at: https://www.sas.org.uk/safer-seas-service/.
Sy, B., et al. (2019) 'Flood hazard assessment and the role of citizen science', *Journal of Flood Risk Management*, 12.
Tabios III, G.Q., et al. (1985) 'A comparative analysis of techniques for spatial interpolation of precipitation 1', *JAWRA Journal of the American Water Resources Association*, 21(3), pp. 365-380.
The River's Trust (2022) *Sewage Map*. Available at: https://theriverstrust.org/key-issues/sewage-in-rivers.
The World Bank, et al. (2020) *World Bank Open Data*. Available at: https://data.worldbank.org/indicator/AG.SRF.TOTL.K2?locations=NL.

Thiessen, A.H. (1911) 'Precipitation averages for large areas', *Monthly weather review*, 39(7), pp. 1082-1089.
Todini, E. (2001) 'A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements', *Hydrol. Earth Syst. Sci.*, 5(2), pp. 187-199.
Trapero, L., et al. (2009) 'Uncertainty of precipitation estimates in convective events by the Meteorological Service of Catalonia radar network', *Atmospheric Research*, 93(1-3), pp. 408-418.
TurfTrax (2019) *The TurfTrax Weather System*. Available at: https://www.turftrax.co.uk/weather-systems.html (Accessed: June 2019).
Villalobos-Herrera, R.B., S; , et al. (2022) 'Sub-hourly resolution quality control of rain gauge data significantly improves regional sub-daily

return level estimates', *Quarterly Journal of the Royal Meteorological Society.* , (In review).
Villarini, G., et al. (2010) 'Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall', *Surveys in geophysics*, 31(1), pp. 107-129.
Villarini, G., et al. (2008) 'Rainfall and sampling uncertainties: A rain gauge perspective', *Journal of Geophysical Research: Atmospheres*, 113(D11).
Walker, D., et al. (2016) 'Filling the observational void: Scientific value and quantitative validation of hydrometeorological data from a community-based monitoring programme', *Journal of Hydrology*, 538.
Walker, D.W., et al. (2021) 'The benefits and negative impacts of citizen science applications to water as experienced by participants and communities', *Wiley Interdisciplinary Reviews: Water*, 8(1).
Wang, L.P., et al. (2013) 'Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology', *Water Sci Technol*, 68(4), pp. 737-47.
Weather Underground (2019) *Weather Observations from Private Automated Weather Stations*. Available at: https://www.wunderground.com/ (Accessed: November 2017).
Wilkinson, M.D., et al. (2016) 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific data*, 3(1), pp. 1-9.
Wilson, C.B., et al. (1979a) 'On the influence of the spatial distribution of rainfall on storm runoff', *Water Resources Research*, 15(2), pp. 321-328.
Wilson, J.W., et al. (1979b) 'Radar measurement of rainfall—A summary', *Bulletin of the American Meteorological Society*, 60(9), pp. 1048-1060.
WMO (2018) 'Guide to Instruments and Methods of Observation'. World Meteorological Organization Geneva.
WMO (2021) *Guidelines on Surface Station Data Quality Control and Quality Assurance for Climate Applications (WMO-No. 1269)*. WMO Geneva, Switzerland.
WSP (2017) *Birmingham City Council June 2016 Flooding: Flood and Water Management Act, Section 19 Investigation*. [Online]. Available at: https://www.birmingham.gov.uk/downloads/file/7167/flooding_section_19_investigation_-_june_2016.
WSP (2019) *Milton Keynes Borough Council, FWMA Section 29 Report Flood Event: May 2018*.
Yang, P., et al. (2017) 'Gauging Through the Crowd: A Crowd-Sourcing Approach to Urban Rainfall Measurement and Storm Water Modeling Implications', *Water Resources Research*, 53(11).
Yang, P., et al. (2019) 'Reward-Based Participant Management for Crowdsourcing Rainfall Monitoring: An Agent-Based Model Simulation', *Water Resources Research*, 55(10).
Zhai, G., et al. (2006) 'Flood risk acceptability and economic value of evacuation', *Risk Analysis*, 26(3).