



Automated quality-control filters for
undetected rainfall in citizen rain
gauge data

R.K. Hutten



Automated quality-control filters for undetected rainfall in citizen rain gauge data

by

R.K. Hutten

to obtain the degree of Master of Science at the Delft University of Technology, to be defended publicly on
Friday June 22, 2018 at 03:00 PM.

Student number:	4004078	
Project duration:	August 21, 2017 – June 22, 2018	
Thesis committee:	Dr. Ir. J. A. E. ten Veldhuis,	TU Delft, water resource management
	Dr. M. A. Schleiss,	TU Delft, Geoscience and remote sensing
	MSc. L. W. de Vos,	KNMI, RDWD
	Dr. Ir. H. Leijnse,	KNMI, RDWD
	Prof. Dr. Ir. N. C. van de Giesen	TU Delft, water resource management



Preface

This thesis is the last step to become a civil engineer specialised in water management. It has been a wonderful process in which I could combine my passion for the issue of rainfall in urban areas with solving data puzzles using python.

Water is a really old interest of me. In my words: 'I am a water rat'. From playing with water and mud in my childhood, to swimming competitions as teenager and rowing in the canals, kayaking on the rivers, rafting on the wild waters as a student. I like to live in an area below sea level with dikes and canals. And now, it is ending up with a real diploma in water management.

Problem solving with logic is my second nature. Already as a child I solved problems by systematic analysing, creating my own models and just practice. As my two sisters often experienced: don't disturb this habitat. Their creativity show up much more in languages and semantics.

Every master thesis has its ups and downs. But I am a lucky girl. Using my talents, I learnt a lot from the things that are difficult for me. The topic of my thesis is exactly what drives me to the best.

But you never do it alone. First, I am grateful to Marie-Claire ten Veldhuis for the shared passion and being my supervisor during my bachelor and master thesis. She knows how to cope with my personality. Marc Schleiss was a very helpful whiz kid to me. He challenges me to find the right answer by thinking critically backwards and systematically forwards. Lotte de Vos is the expert in citizen weather stations. Thanks for sharing your knowledge, for your patience and the many useful suggestions. If I really didn't know what to do I always went to Hidde Leijnse for a brainstorm session about a filter design, solving together the issues in the programming scripts or for just some original and unconventional advice. Nick van de Giesen really is a supervisor who gave me great input during the start of my thesis. Aart Overeem was my neighbour at KNMI. Like good neighbours do: he helped me through the difficult hours. I enjoyed his 'state of the Aart' jokes. Thanks to all of you for the learning environment, but also for your cooperation and the pleasant ambience at KNMI as well as at the TU Delft. It has been a great time for me.

A day has 24 hours and writing a master thesis does not stop outside the doors of the research building. My friends, especially Joreen, Sandra, Zoe, Marloes, Vera en Lotte, were not only helpful 'partners in crime', but also made me enjoy life. I also got great support from my room mates from room 4.93 at TU Delft. They always took time to listen to my stories and they offered me a cosy environment.

I also like to acknowledge my family. My parents for their emotional, but also financial support. I am lucky that they are engineers. Thanks for your involvement and your interest into my study in water management. When you emphasise how proud you are, I know my education has been successful. I am the oldest of three totally different sisters. Due to these differences we never had a dull moment at home, but we also share values and made a lot of fun together. It helps me through my single minded periods. Ela and Kris thanks!

Last, but not least I thank my boyfriend David! I know I have been insufferable when my focus is on my study, especially during writing my thesis. It must really be love if you can cope with that. Thank you so much for creating my home base and for having so much patience and for giving all the emotional things and love I need.

*R.K. Hutten
Delft, June 2018*

Contents

Preface	iii
Abstract	vii
Nomenclature	ix
1 Introduction	1
1.1 Introduction to Citizen Weather Stations and Platforms	2
1.2 Errors and Quality-Control of rainfall data from weather stations	3
1.3 Research Objective and Contributions	5
1.4 Guidelines for Reading	6
2 Data and Methods	7
2.1 Study Area	7
2.2 Data and Software	8
2.2.1 Data	8
2.2.2 Software	9
2.3 Methods	12
2.3.1 Research Methodology.	12
2.3.2 Data Collection of Netatmo Rainfall Timeseries in study area	12
2.3.3 Conversion of CWS data from Irregular to Regular Time Grid	14
2.3.4 Simulation of UR errors on Datasets	14
2.3.5 Confusion Matrices and Cost Function	15
3 Properties and Errors of CWSs	19
3.1 General Properties of CWSs in study area	19
3.2 Errors of CWSs in study area	20
4 Design of Filters for UR errors in CWSs	25
4.1 CWS Filter.	25
4.2 Radar Filter	28
5 Performance of designed Filters	31
5.1 CWS Filter.	32
5.2 Radar Filter	37
5.3 Comparison between the CWS and Radar filter based on Performance	40
6 Flags given to CWS Data by Filters	41
7 Discussion	47
7.1 Assessment of Data and Methods used in this study	47
7.2 Calibration of Tipping Bucket Volumes from CWSs	48
7.3 Limitations of CWS Filter	49
7.4 Limitations of Radar Filter	51
7.5 Limitations of CWS Data and UR Flags given to CWSs.	52
8 Conclusion and Recommendation	55
References	59
A Overview of CWS Platforms	63
B Determination of Threshold for the Simulation of UR Errors on datasets	67
C Determination and Verification of fixed Radius for CWS Filter	69

D	Determination of q_{rad} using Drizzle Check	71
E	Spatial Maps of Lacking Intervals and Flags given by the Filters	75

Abstract

Flooding in cities, known as Urban Pluvial Flooding (UPF), causes disruption of society, damage to cities and inconvenience for people (Douglas et al., 2010; Spekkers et al., 2015). Cities are expected to become more vulnerable to UPF due to more frequent and more intense extreme rainfall events with high spatial variability as result of climate change (Ashley et al., 2005; Hartmann et al., 2013). Additionally increasing urbanisation results in more impervious surfaces, inducing shorter response times of the drainage systems. A high spatial and temporal resolution of the urban rainfall network is required to forecast UPF. Professional rain gauge networks do not provide this necessary resolution. A strategy to increase this density is to use rainfall data obtained from Citizen Weather Stations (CWSs). The downside of CWS networks is the low quality of measurements compared to professional measurements. This is caused by the use of lower quality sensor, web-platform processes, set-up of the weather station, stability of data transfer and cleaning of the CWS (de Vos et al., 2017). Undetected Rainfall (UR) is one of the errors found in rain gauge data of a CWS network. An UR error is an incorrect zero rainfall value in the rainfall data, meaning that rainfall did occur at the gauge but it was not detected. A quality-control system that creates a more reliable CWS network is required to facilitate UPF forecasting. The aim of this research is to take a first step towards a quality-control system by developing and testing different automatic quality-control filters that flag for UR errors in rainfall data of a CWS network.

For this research 753 CWSs in the Rotterdam-The Hague region are used. They were obtained from Netatmo weathermap and give a density of 1 CWSs per 5.93 km². This is not a uniform density because more CWSs are located in densely populated areas. The availability of the CWSs grew from 369 to 668, measuring at least once per day between October 2015 and October 2017. Two filters are developed to identify UR errors in CWS rainfall data. The first filter is the CWS filter which uses neighbouring CWSs within a radius of 8 kilometre to find the UR errors. The second filter is the radar filter, which identifies UR errors using the overlying radar pixel. The performances of these filters are tested by artificially placing zeros at 10% of the rainfall sequences of 10-min rainfall data from four automatic KNMI gauge datasets within the study area. Both filters are applied on these datasets and evaluated using confusion matrices. Finally, the two filters are applied on the 753 CWS rainfall datasets for a period between October 2016 and October 2017.

Both filters are considered to be adequate in detecting and flagging of UR errors, however, the performance of the two filters decreases during low rainfall intensities. The CWS filter performs slightly better compared to the radar filter. The performance of the CWS filter decreases significantly in case of a low density of available neighbouring CWSs within an 8km radius or in the presence of physical barriers. In that case the radar filter is preferable due to the overall more stable performance as result of continuous data availability. A drawback of the radar filter is the higher number of incorrect flags compared to CWS filter. These appear due to the disagreement between radar and gauges on the occurrence of zero rainfall. These results demonstrate that both filters can automatically identify UR errors in rain gauge data, however the performance of the filters varies per region and study time. Future research should therefore focus on testing these filter in more detail to obtain more specifications on how they work under different circumstances as well as to establish a strategy for the selection of the correct filter in these different situations. Apart from this, future studies should focus on developing filters for other errors such as false rainfall intensities and cumulative values in data. Combined, this will eventually create one quality-system that can be easily applied on CWSs.

Nomenclature

List of abbreviations

Abbreviations	Full Meaning
AG	Automatic Gauge
API	Application Programming Interface
CWS(s)	Citizen Weather Station(s)
DM-curve	Double Mass curve
FDR	False Detection Rate
FN	False Negatives
FNR	False Negative Rate
FP	False Positives
FPR	False Positive Rate
KIS	Klimatological Information System
KNMI	Royal Dutch Meteorological Institute
QA	Quality-Assessment
QC	Quality-Control
QGIS	Quantum Geographical Information System
TB	Tipping Bucket
TN	True Negatives
TNR	True Negative Rate
TP	True Positives
TPR	True Positive Rate
UPF	Urban Pluvial Flooding
UR	Undetected Rainfall
URM	Urban Rainfall Monitoring
ZI	Zeros Implemented

List of symbols

Symbol	Unit	Definition
Δt	min	Temporal resolution of fixed grid
Δ_{\max}	min	Maximum time gap boundary
A_l	km ²	Land area of study area
A_s	km ²	Sea area of study area
A_t	km ²	Total area of study area
CWS_n	-	Regular timeline and rainfall data of CWS with ID
$CWS_n(t)$	-	Data of CWS at timestep t
CWS_{old}	-	Original CWS data retrieved with API's
$F_n(t)$	-	Output flag of filter at timestep t
$F_n(t-1)$	-	Output flag of filter at previous timestep t-1
LI_{cws}	%	Percentage of lacking intervals per CWS
$N_{med,n}(t)$	-	Number of neighbouring CWSs inside radius of CWS_n at timestep t
$P_{CWS,n}(t)$	mm	Precipitation depth of CWS_n at timestep t
$P_{med,n}(t)$	mm	Median precipitation depth of neighbouring CWSs inside radius of CWS_n at timestep t
$P_{rad,n}(t)$	mm	Precipitation depth of overlying radar pixel at timestep t
q_{ag}	mm	Threshold for artificially introduced zeros
q_{rad}	mm	Threshold for the occurrence of rainfall for radar
$R_{new,cumsum}$	mm	New cumulative rainfall timeseries after re-interpolation
$R_{old,cumsum}$	mm	Original cumulative rainfall timeseries
R_{old}	mm	Original rainfall timeseries corresponding to irregular timeline
T_{new}	-	Regular timeline after conversion
T_{old}	-	Original irregular timeline
x_{new}	-	Timesteps of T_{new} in which Δ_{\max} is exceeded
x_{old}	-	Timesteps of T_{old} in which Δ_{\max} is exceeded



Introduction

Urban Pluvial Flooding (UPF) causes disruption of society, damage to cities and inconvenience for people (Douglas et al., 2010; Spekkers et al., 2015). Just in the summer of 2017 alone, multiple extreme weather events occurred triggering UPF. A Dutch news provider stated on June 29th, 2017 that “*Berlin was shut down due to extreme rainfall*” (NOS, 2017). Another example is the quarter final of European football championship for women in Rotterdam on 29th of July 2017, which was cancelled because of flooded football fields due to heavy rainfall (AD, 2017). These news articles are some examples in which UPF causes problems for society. The effects of UPF have also been evaluated by many studies. For example, Spekkers et al. (2013) state that a relation exists between the number of insurance claims and excessive rainfall. ten Veldhuis (2011) shows that the cumulative damage of UPF over 10 years is equivalent to the damage of an event once per 125 years in Haarlem. In other words, UPF has significant impact on society and cities, which highlights the societal and scientific relevance of the topic.

In the future, cities are expected to become more vulnerable to UPF as result of climate change and urbanisation. Extreme rainfall events with high spatial variability are predicted to intensify and occur more frequently as result of climate change (Ashley et al., 2005; Hartmann et al., 2013). In addition, the amount of population residing in urban areas will increase from 54% in 2014 to 66% in 2050. For the Netherlands, this amount is expected to become 96.4% in 2050, compared to 90% in 2014 (United Nations, Department of Economic and Social Affairs, Population Division, 2015). This foreseen urbanisation will lead to more impervious surfaces inducing quicker response time of the drainage systems. Together these factors cause urban areas to be more prone to UPF and its impacts (Gaitan et al., 2016; Stone et al., 2011). Prevention, reduction and forecasting UPF is therefore a crucial issue for society, waterboards and municipalities.

As mentioned before, Rotterdam is an example of a city that experiences UPF. Adaptive strategies for reduction and prevention of UPF have been developed and implemented (De Urbanisten and Management team of Rotterdam Climate Proof, 2013; Municipality of Rotterdam, 2016). UPF forecasting requires an improvement of Urban Rainfall Monitoring (URM). Urban areas require a higher temporal and spatial resolution compared to rural areas for URM (Schilling, 1991). Specifically, a temporal resolution of 5 min and spatial resolution of 3 km for 1000 hectare urban catchment are recommended by Berne et al. (2004). These fine temporal and spatial scales can only be obtained by a high density gauge network (Kidd et al., 2017). Cristiano et al. (2017) emphasise as well that a dense network is required to capture the small-scale variability of rainfall. Such a network that measures at high temporal resolution, is still lacking in urban areas (Niemczynowicz, 1988). For example, the professional rain gauge network in Rotterdam, an area of 324.16 km² (32416 hectare), only consist of two automatic gauges measuring at Rotterdam airport and Hoek van Holland and one daily manual gauge at the northern border as shown in Figure 1.1. The automatic KNMI gauges measure every 10 minutes. In short, the density of the professional rain gauge network in Rotterdam falls short for URM to forecast UPF.

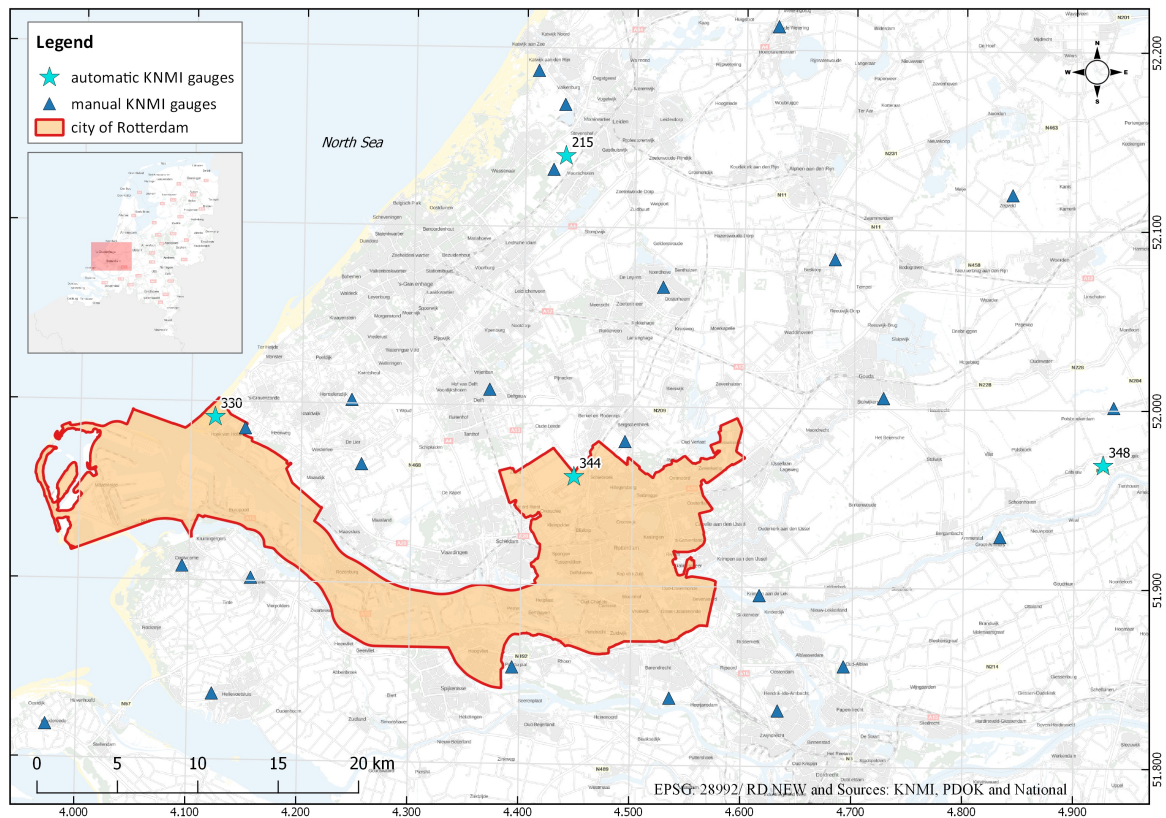


Figure 1.1: A map of the professional KNMI gauge network around the city of Rotterdam showing the shortage of rain gauges in the densely built areas.

1.1. Introduction to Citizen Weather Stations and Platforms

A strategy to increase the density of the rain gauge networks is the use of crowdsourced data. The implementation of crowdsourced data is enabled by the shift of hydrologic focus from the third paradigm 'computational simulation' to the fourth paradigm 'data intensive science' as discussed by Peters-Lidard et al. (2017). The acceptance of the fourth paradigm allows us to embrace crowdsourced data to improve our understanding of the hydrological cycle. In addition, Muller et al. (2015) state that crowdsourced meteorological data has the potential to solve the issues concerning observations in areas with low-density network and real-time monitoring, validation and quality control mechanisms that need to be in place to accomplish this. The benefits and implementation of crowdsourced Citizen Weather Stations (CWSs) for temperature monitoring have already been investigated in multiple studies (Bell et al., 2013, 2015; Meier et al., 2015), while only a few researches have been conducted to the potential and use of crowdsourced CWSs for rainfall monitoring. Bell et al. (2015) investigate the biases of CWSs between different types and brands compared to professional weather stations. de Vos et al. (2017) conclude that crowdsourced CWSs have the potential to be an additional tool to monitor urban rainfall without additional investment costs. In other words, crowdsourced CWSs can be an appropriate tool to the problem regarding URM.

Worldwide many crowdsourced weather networks are available using citizen science (Muller et al., 2015). Examples of these platforms are CoCoRaHS (CoCoRaHS, 1998), UCRaiN (Illingworth et al., 2014), CWOP (CWOP, 2000), WOW (Met Office and KNMI, 2011), Wunderground (Weather Underground, 1993) and Netatmo weathermap (Netatmo, 1993). Some of those platforms only show precipitation amounts obtained from CWSs, while other platforms as CWOP, WOW, Wunderground and Netatmo weathermap consist of more meteorological variables, for example temperature. Besides the number of meteorological variables measured by CWSs, more features vary between the platforms. These features are area of interest, temporal resolution, connection between station and platform (built-in or external software and so on), web-platform processes, presence of quality control, availability of meta-data, rain gauge types, weather station brands, density of networks, objectives of platform, involvement of contributors, feedback for contributors and sup-

port for contributors. Despite of the wide variety of crowdsourced CWS platforms only a few are suitable for URM because of the requirement of high temporal and spatial resolutions.

As mentioned above, all platforms have different properties. For some platforms these properties are shown in Table A.1 in appendix A. For example the temporal resolution varies between the different platforms. Both CoCoRaHS and UCRaiN upload daily new data, whereas the others do this on average every 5 to 10 minutes. The uploading of measurements on Netatmo and Wunderground platform happens with an irregular time interval of approximately every 5 and 10 minutes respectively. The temporal resolution of the platform is dependent on the type of gauge applied in the network. Manual gauges are used for the CoCoRaHS and UCRaiN. Netatmo weathermap, CWOP, WOW and Wunderground use automatic weather stations. In contrast to Netatmo weathermap, CWOP, WOW and Wunderground show gauge data of multiple brands that they include in their platform, e.g. Davis Vantage Pro2, Vaisala WXt510 and LaCrosse 2500. The density of a platform depends on locations that the platform covers and the number of stations available for that area.

Other differences between the platforms are in the communication link between station and platform. For the platforms with manual gauges, the communication takes place by email, telephone or by uploading the data on the webpage by contributor (CoCoRaHS). The Netatmo weathermap uses built-in software. The owners can decide if they want to share the data from the outdoor module to Netatmo weathermap. Other platforms use software to connect the stations with platform. Butler (2018) investigates several software packages available and applied by CWOP, WOW and Wunderground contributors to enable the connection between station and platform. The choice of software to upload to these platforms strongly depends on the owner's preferences. Advice on software is available on web pages for WOW and Wunderground contributors. The owners decide if they want to contribute by setting-up a link between their station and platform of interest.

Gharesifard and Wehn (2016) research the drivers and barriers of people to take part in sharing meteorological weather data of CWS to online platforms. People are more likely to contribute to weather platforms if a combination of high quality and good accuracy of CWS data is achieved, equitable cost is applicable for CWS and the web-platforms are easy to use. They also conclude that trust in the competence of citizen and in the quality of data is essential for people to share their data. The attachment of meta-data is a recommendation to create this trust. WOW is a platform that is front-runner in this. They give amateurs the possibility to classify their locations to inform data-users about the measurement set-up (Met Office and KNMI, 2011). CoCoRaHS supports its contributors by training them how to measure precipitation and teaches them more about how the gauge should be installed. CWOP recommends their members to request feedback about the performance of their station. Netatmo reduces the barriers by giving them an option during installation to automatically share the data, as well as, allowing people to retrieve details about rain gauge battery and quality of Wi-Fi signal.

Both the Netatmo and Wunderground are growing platforms that nowadays meet those requirements for URM closest based on the temporal resolution and density of the network in the Netherlands. The Wunderground platform has a higher CWS density compared to the Netatmo platform, as the platform includes all Netatmo devices automatically, as well as multiple CWSs of other brands. The temporal resolution of Wunderground, however, is lower than Netatmo, approximately 10 and 5 minutes respectively (de Vos et al., 2017). This means that temporal resolution of Wunderground is too low for URM as a minimum temporal resolution of 5 minutes was found in literature (Berne et al., 2004). Apart from the higher density of the growing crowdsourced CWS networks and no additional investment costs compared to professional weather networks, the CWS networks have a significant lower quality. This raises concerns about the potential of the data from the crowdsourced CWS networks, since they are not obtained according to the guidelines of World Meteorological Organisations and other professional weather organisations (Muller et al., 2015).

1.2. Errors and Quality-Control of rainfall data from weather stations

The downside of crowdsourced CWS networks is a lower quality of measurement accuracy due to the occurrence of many different errors. Errors often found for low-cost CWSs are larger compared to the ones of automatic rain gauges of authorities, but also new additional errors arise. They stem from the sensor quality, processes on web-platforms, lack of knowledge of owners on how to place of the gauge on their terrain, feed-

back from end-user (research, company or authority) to owner about performance of gauge, and the reality that many CWSs are situated in the backyards or upon roof terraces (Figure 1.2). The CWS sensors, for example, are cheaper than the ones used in professional networks and cannot measure solid precipitation resulting in larger measurement errors. Wrong placement of the gauge can for example lead to a gauge being located in wind shadows in addition to 2% to 10% systematic wind induced error above the gauge (Jarraud, 2008b). Sometimes CWSs are installed below or close to trees resulting in faulty measurements. Partial blockage of a gauge gives an underestimation of rainfall, whereas complete blockage results in zero rainfall measurement during an event. Jarraud (2008b) also states that there is an error due splashing in or out the gauges. This splashing error can become larger due to rain gauges close to (paved) surfaces such as a terrace. CWS owners are also advised to periodically clean their gauge from dirt e.g. insects and leaves (Butler, 2018) because not cleaning of the gauge will result in faulty measurement. The problem is that some owners clean their gauge with water resulting in incorrect measurements. Also, station outage can occur due to a non-operational sensor. Finally, the processes on the platforms can cause errors, such as rounding and conversion (from inch to millimetres) errors and connection loss with the device will lead to loss of measurement as found by de Vos et al. (2017).

All errors aforementioned are mainly present in CWS rainfall data through the following three types of erroneous behaviours in citizen rainfall data: (i) false rainfall intensities due to shielding errors, splashing errors and partial blockage of gauge, (ii) station outage due to non-operational sensors resulting in lacking intervals, or (iii) Undetected Rainfall (UR) due to complete blockage of the gauge. These errors are difficult to identify because a reliable ground-truth of CWSs is often lacking.



Figure 1.2: A picture of CWS on roof terrace in Essen, Germany. (Picture taken by Frans IJserinkhuijsen)

Despite that CWS networks have the potential for URM, they do not comply with quality rain gauges of professional networks. Quality-Assessment (QA) and Quality-Control (QC) of crowdsourced CWS rainfall data are therefore required to make it suitable for URM (Muller et al., 2015). The quality of automatic rain gauges of professional networks has extensively been assessed using the standards provided by WMO and other institutions (Estévez et al., 2011; Jarraud, 2008a; Zahumenský, 2004). Many studies have developed QC mechanisms based on those assessment to enhance the quality of the professional networks by implementing decision trees (Qi et al., 2016) or multiple procedures (Blenkinsop et al., 2017) flagging or filtering for the various er-

rors using statistical checks. The use of decision trees in combination with flagging is a principle that can easily be implemented in QC of CWS data. On the other hand, the use of many other variables for cross-checking as done by Qi et al. (2016) reduces the applicability of future QC filters substantially because of the requirement of other sources.

Upton and Rahimi (2003) present research that might contribute to development of QC for CWS rainfall data. They apply both single-gauge tests and checks using neighbouring stations for online detection of errors in tipping-bucket rain gauges. After analysis of the checks, only four tests are recommended for daily application. Two of the recommended checks are the computation of median number of tips and median inter-tip times from data obtained from neighbouring stations. So-called buddy checks and the computation of statistical median are both interesting for QC filters of citizen rainfall data. Data of neighbouring stations can be used for the spatial consistency component in the filters and makes the need of other sources redundant. By computing the median of neighbouring CWS data, the checks are less affected by the potential threat of outliers in those stations compared to the mean.

Some web-platforms include QC of CWSs as given in Table A.1. CoCoRaHS implements both consistency checks and data entry checks to determine the reliability of the data. The data entry checks are more preliminary QC in which errors are detected, such as incorrect station numbers, dates and times, or unreasonable precipitation values. CWOP executes QC on request of contributors. The QC includes a sub hourly validity check of accumulated rainfall. Wunderground has an algorithm consisting of range test, internal rate of change and spatial consistency check using neighbouring stations. But the exact performance of the filter is unknown (Butler, 2018). In other words, CoCoRaHS, CWOP and Wunderground have QC filters in place.

Lastly, de Vos et al. (2017) investigate the potential of crowdsourced CWSs for URM as mentioned before. They assess the quality of CWSs in Amsterdam mapping out several errors in the crowdsourced CWSs. The crowdsourced CWS data of both Netatmo and Wunderground platforms are processed from irregular time interval to a fixed time grid and filtered by basic QC mechanisms found by trial and error. A first attempt of QC is performed, however, an extended and automatic QC system for the crowdsourced CWSs of Netatmo platform is still missing and the operation of Wunderground QC algorithm remains an open question. The same problem is the case for Netatmo weathermap. A filter option can be selected during retrieval of rainfall data from platform, but it is unknown which checks are performed. Hence QC filters needs to be developed and tested for CWS rainfall data obtained from crowdsourcing platforms in order to use these networks for URM.

1.3. Research Objective and Contributions

The aim of this study is to develop and test different automated quality-control filters flagging for Undetected Rainfall (UR) errors in rain gauge data of a CWS network. By doing so, the first steps are taken for a roadmap towards a quality-control system that creates a more reliable CWS network for URM to facilitate forecasting of UPE. The CWS network of Netatmo weathermap is selected for several reasons: high temporal resolution of platform, the use of only one type of rain gauge (tipping bucket rain gauge of Netatmo brand), no web-platform processing, no quality-control applied on data (only on request), dense network in surrounding of Rotterdam and fewer data gaps compared to Wunderground. In order to accomplish the objective of this study, the following main research question is formulated:

How can undetected rainfall errors in citizen rain gauge data be identified by automated quality-control filters?

Different sub-questions are formulated to answer the main research question.

1. *What are the properties and errors of rainfall data in the CWS network in the study area?*

To enable the development of QC filters for citizen rain gauge data it is required that there is a better understanding of general features present in the data such as density of CWS network and availability of measurements. An analysis of the data on specific errors and properties forms the basis for subsequent steps. A general assessment of the rainfall data in the study area is therefore executed.

2. *What filters can be designed to automatically detect UR errors from citizen rain gauge data?*

Several filter designs are created including complete documentation to allow future users to understand the operations and the underlying assumptions to detect UR errors.

3. *What is the performance of the proposed filters?*

The development of a filter is an iterative process. By evaluating the performance of the filter, the effectiveness of the filter is measured. In addition to this, the performances give a baseline for the following sub-questions about how it flags and how filter can be used in future applications. This is essential, since the ground-truth of UR errors in CWSs are lacking.

4. *How do filters flag different CWSs in the study area over time and space?*

Analysing the flagging of CWSs by the filters is essential to obtain understanding about possible UR errors and reliability of the CWS network.

5. *What (dis)similarities are present between the filters based on neighbouring CWSs and radar?*

The comparison between two different filters shows how the filters differ from each other based on the performance of filters and the flagging of CWSs. The comparison should lead to recommendations with respect to application and limitations of the filters regarding their applicability and suitability.

6. *What general recommendations can be provided for using CWSs in the study area?*

The question will give information about reliability of stations in the study area and in which cases the filters are suitable.

1.4. Guidelines for Reading

In the next chapter, data and software used in the research are presented and the methods are explained used to achieve the objective. After this the properties and errors of CWS network in study area are given (sub-question 1). Chapter four gives the design of two filters to detect and flag the UR errors in CWS rainfall data (sub-question 2). After that the performances of the two filters are evaluated by applying the filters on automatic KNMI datasets with artificially implemented zeros (sub-question 3). In the same chapter also, a comparison between the performances of the filters is conducted (partial sub-question 5). The sixth chapter shows the flags given to CWSs in study area by the filters (sub-question 4) including a comparison between the flags given by the two filters (partial sub-question 5). After the results, the findings from the previous chapters are discussed and some first recommendations about the use of the filters is provided (sub-question 6). The recommendations are further addressed in the eighth chapter, as well as the conclusion of this study.

2

Data and Methods

The study area, data, software and methods that are applied in the research are discussed in this chapter. Firstly, a brief explanation of the study area is given including the demarcation. Secondly, the data required for the research are described followed by packages and specific functions from the software. Subsequently, the methods applied to achieve the objective of the research are given.

2.1. Study Area

The study area is selected in such a way that the area covers a radius of 35 kilometres around the Delftse Poort (51.924085°, 4.4718289°) in the city of Rotterdam. An X-band radar with a range of 30 kilometre is installed in September 2015 at the Delftse Poort, one of the highest buildings in Rotterdam. The construction of the rain radar was part of the EU-funded Rain Gain project that aims at improving the predictions of UPE. The University of Technology in Delft (TU Delft), specifically the department of Geoscience and Remote sensing and Water resource management of the faculty Civil Engineering, is also one of the partners involved in the project MUFFIN (Multiscale Urban Flood ForecastING). One of TU Delft's objectives is to discover the possibilities of rainfall data obtained from crowdsourced CWS networks of urban rainfall measuring in the city of Rotterdam. Both projects require a fine spatial and temporal resolution of rainfall measurements in urban areas. This is partially already accomplished by installing the X-band rain radar at Delftse poort, which measures at a 30×30 m² spatial resolution every minute. The X-band radar is however not further used in this study. The use of CWSs network has the potential to complement this as previously discussed. Hence the study area is chosen that the CWS network in the study area covers the range of radar.

The coordinates of the study area are given in Table 2.1. The study area given in Figure 2.1 includes two of the four biggest cities in the Netherlands, namely Rotterdam and The Hague. The total area (A_t) is 50545.51 km² in which land area (A_l) and sea area (A_s) are respectively ~4464.32 km² and ~581.19 km². The sea is located in the North-West part of the study area. Other significant land uses are the waterways, lakes, dunes and the harbour area including Maasvlakte 1 and 2 in Rotterdam. All these land classes are areas in which no people reside, which means that the expected availability of CWSs in those area is minimal. The area for CWSs (approximately land area) is thus smaller than the total area of the study.

Table 2.1: The latitudes and longitudes of the study area are given here.

	Coordinates
<i>Southern border</i>	51.60°N
<i>Northern border</i>	52.24°N
<i>Western border</i>	3.95°E
<i>Eastern border</i>	4.98°E

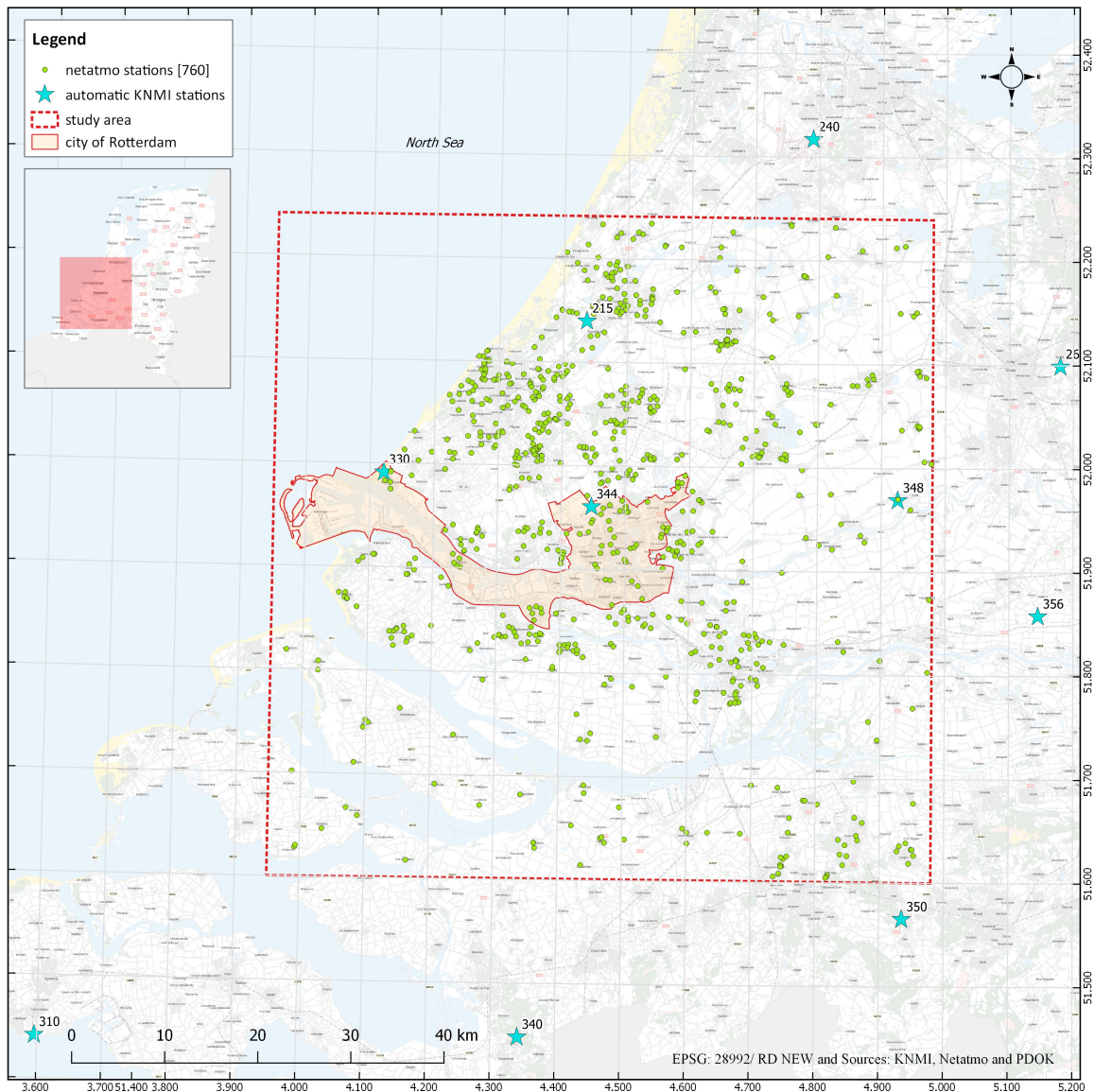


Figure 2.1: The study area of the research is selected that the range of the X-band radar located at Delftse Poort in Rotterdam is covered. In total, 760 Netatmo CWSs were found in the area, as well as, 4 automatic KNMI gauges measuring every 10 minutes. The Netatmo stations upload approximately every 5 minutes data to Netatmo weathermap.

2.2. Data and Software

This section discusses the data, software packages and specific functions from the software applied in the research. Data is briefly described to obtain more knowledge about the characteristics of the data. The packages and functions, required to complete the research, are termed here to enhance the reproducibility of the research.

2.2.1. Data

Several rainfall data sets are used in the research. Table 2.2 lists these data sets. The rainfall data obtained from Netatmo rain gauge¹ are input for the filters and are assessed. The rain gauges of Netatmo are tipping buckets with a default tipping bucket volume of 0.101 mm, shown in Figure 2.2. The diameter of the collecting funnel is 130 mm (de Vos et al., 2017). These tipping buckets have a detectable range between 0.2 mmh⁻¹ to 150 mmh⁻¹ with an accuracy of 1 mmh⁻¹ according to Netatmo company². The rainfall data is transferred

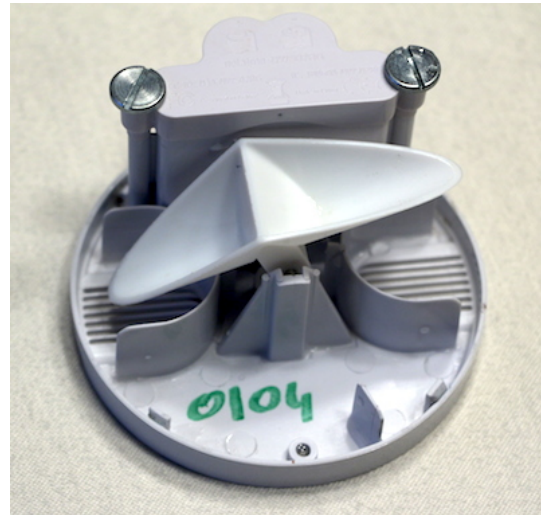
¹<https://weathermap.netatmo.com/>

²<https://www.netatmo.com/en-US/product/weather/weatherstation/specifications>

between rain gauge and indoor stations through wireless connection. The distance between the rain gauge and the indoor stations can therefore be up to a maximum of 100 meters. The automatic KNMI rainfall data are mainly used to evaluate the performance of the filters. The overlying radar pixel datasets are primarily implemented in a QC check of a filter. Table 2.3 shows the visualisation layers used in spatial images.



(a) Netatmo rain gauge



(b) Netatmo tipping bucket mechanism

Figure 2.2: (a) Picture of Netatmo rain gauge (courtesy of Netatmo). (b) Picture of tipping bucket mechanism inside of the Netatmo rain gauge (Picture is taken from <https://www.engadget.com/2014/04/24/netatmo-expands-ios-friendly-weather-station-with-rain-gauge/?guccounter=1>).

2.2.2. Software

This study uses Quantum GIS (QGIS) and programming language R and Python 2.7. QGIS is used to create spatial images and basic area calculations. Programming language R is applied to retrieve the CWSs data from the Netatmo weathermap. The retrieval of data with R requires the packages *httr* and *jsonlite*. The rest of the programming is performed with Python 2.7 using the packages: *pandas_ml*, *numpy*, *matplotlib.pyplot*, *os*, *datetime*, *time*, *scipy* (computation of euclidean distance), *pyproj*, *itertools*, *sklearn.metrics* (for confusion matrix), *csv*. Next to packages designed for Python, a function (source: https://rosettacode.org/wiki/Run-length_encoding#Python) is applied to perform running length code (equivalent to *rle* in R).

Table 2.2: An overview of the used rainfall data sets, including a description, the variables present in collected data, the retrieved time period, the application in the research and the source, is given here.

<i>Rainfall datasets</i>					
Name	Description	Variables	Year/Period	Application	Source
Netatmo CWSs	Netatmo rainfall data of the outdoor modules of CWSs is retrieved for a specific period.	Device code, module ID, coordinates irregular unix timestamps (~ 5-min data) and rainfall depths [mm]	1 October 2015 to 1 October 2017	The Netatmo rainfall data is input for the filters.	https://weathermap.netatmo.com/
Automatic KNMI gauges	From the synoptic measurement of the Netherlands database in the Climatological Information System (KIS), the unvalidated 10-min rainfall data of four automatic gauges in Voorschoten, Hoek van Holland, Rotterdam and Cabauw were collected.	Location of gauge, UTC datetime, station code, precipitation intensity electric rain gauge [mm/h]	1 October 2015 to 1 October 2016	The 10-min rainfall data from the automatic KNMI gauges is used to evaluate the performance of the filters. The same data is applied for the analysis of occurrence of rainfall between point measurement and overlying radar pixel.	Klimatological Information System (Only available for KNMI employees)
RAD_NL25_RAC_MFBS_5min (Radar)	The gauge-adjusted climatological radar datasets of rainfall depths from KNMI that have a 5-min temporal resolution and 1×1 km ² grid. The rainfall images originate from two C-band weather radars in de Bilt and Den Helder, adjusted with ground measurements from KNMI manual and automatic gauge networks (Overeem et al., 2009a,b).	Overlying radar pixel above the Netatmo CWSs: UTC datetime and 5-min precipitation accumulation [mm]	1 October 2015 to 1 October 2017	The 5-min precipitation accumulations are converted to 10-min precipitation accumulations. The adjusted data are then applied as check in a filter. The same data of the period 1 October 2015 to 1 October 2016 is applied for the analysis of occurrence of rainfall between point measurement and overlying radar pixel.	https://data.knmi.nl/datasets/rad_nl25_rac_mfbs_5min/2.0

Table 2.3: An overview of WMS and WFS layers used in visualisations including a description, the application in the research and the source, is given here.

<i>Visualisation layer</i>				
Name	Description	Year/Period	Application	Source
BRT Background map Grey (WMS)	BRT background map is created from TOP 10NL from the Basis Registration Topography (BRT) with streetnames of the Basis registration Adresses and Buildings (BAG) and updated every year.	-	BRT Background map Grey is applied as background for spatial visualisations in QGIS	https://www.pdok.nl/nl/service/wmts-brt-achtergrondkaart-grijs-0
CBS Neighbourhood 2017 (WFS)	CBS Neighbourhood 2017 consist of the geometries of municipalities and neighbourhoods. All the statistical core numbers are attributed to the geometries.	2017	The municipality layer form CBS Neighbourhood 2017 is extracted to create a shapefile of the city of Rotterdam in QGIS.	https://www.pdok.nl/nl/service/wfs-cbs-wijken-en-buurten-2017

2.3. Methods

2.3.1. Research Methodology

To accomplish the objective of the research, several steps subdivided in two parts, are completed. The first part consists of a literature review to gain more knowledge about CWSs and platforms, as well as to understand the errors found in (citizen) rain gauges and QC mechanisms created by previous studies to filter those errors. The second part consist of the research methodology to create the filter to detect and to flag UR errors in CWS rainfall data and answer the questions of the research.

Figure 2.3 shows the research methodology. In step A, the Netatmo data and other data are collected from the Netatmo platform (<https://weathermap.netatmo.com/>), data center of the Royal Dutch Meteorological Institute (KNMI) or the internal Klimatological Information System (KIS). Subsection 2.3.2 discusses the data collection of Netatmo rainfall data more in detail. The following stage (step B) is to pre-process all rainfall data to precipitation depth in with a temporal resolution of 10 minutes. The method applied to convert CWS data from an irregular time grid to a regular time grid is given in subsection 2.3.3. Other actions taken in step B are the identification of empty or NA-filled CWS rainfall timeseries and tipping bucket volumes found for the Netatmo's in the study area. In step C, CWS data in the study area is assessed to identify erroneous behaviour in CWS rainfall data and to derive properties of the data for the study area. This phase of the research has a bigger context than only UR errors.

The focus in step D to F is on the QC of UR in the Netatmo rainfall data. Step D is the design of the filter including the limitations, benefits, assumptions and computation relevant features required to run the filter. After this, the performances of the filters are assessed using the automatic KNMI rainfall timeseries with randomly implemented zeros to simulate UR intervals in step E, explained in subsection 2.3.4, as input for the filter. The UR flags given by the filter are then compared to the known locations of Zeros Implemented (ZI) using confusion matrices including a cost function. Subsection 2.3.5 gives more information about this method. If the filter performances appear to be sufficient, then the filter is considered to be working. A new phase (step F) is started or a new filter is developed. Otherwise, a closer and critical investigation of the output is conducted and an iterative loop back to the design step is taken. The evaluation of the filter performance is only carried out for the first year of the research period, namely 1 October 2015 to 1 October 2016.

Step F consist of UR flags given by the filter to Netatmo rainfall data for the period between 1 October 2016 to October 2017. This is evaluated over space and time. In order to obtain the (dis)similarities between the filters, comparisons are made at the end of steps E and F. Lastly, general recommendations are given for using CWSs to provide knowledge on how to use CWSs and the filters.

2.3.2. Data Collection of Netatmo Rainfall Timeseries in study area

As aforementioned, Netatmo owners can anonymously share data of their outdoor module to Netatmo weathermap. Others can retrieve the data within a specific area using Netatmo Weather Application Programming Interface (API). An image of the structure of Netatmo weather API is given in Figure 2.4. The retrieval of data from the Netatmo platform starts with creating a Netatmo account and herein an app to obtain access to the data. The app gives a Netatmo client ID and Netatmo secret ID to get an authorized token. Note that there is a maximum amount of retrieves before a new app needs to be generated.

The Getpublicdata API (<https://dev.netatmo.com/en-US/resources/technical/reference/weatherapi/getpublicdata>) enables that all device ID's (MAC adres), rainfall module ID's, locations and the most recent measurements are retrieved within the predefined study area (Table 2.1). Since the size of the study area is relatively large, the number of stations within the area exceeds the limit of returning stations. It is therefore subdivided into 100 tiles in order to return all ID's and locations of CWSs within the area. Meier et al. (2015) also encountered this problem. The ID's and locations obtained with Getpublicdata API are then applied to collect the rainfall timeseries for a specific CWS using the Getmeasure API (<https://dev.netatmo.com/resources/technical/reference/common/getmeasure>). Other entries required for this API to work are start and end time of the series in Unix timestamps. The Unix start and end timestamps are respectively 1443657600 seconds (1 October 2015 00:00:00 UTC) and 1506816000 seconds (1 October 2017 00:00:00 UTC). The use of the two API's gives two different type of files: a file with all CWSs and module ID's, coordinates and altitude, and rainfall timeseries for CWSs.

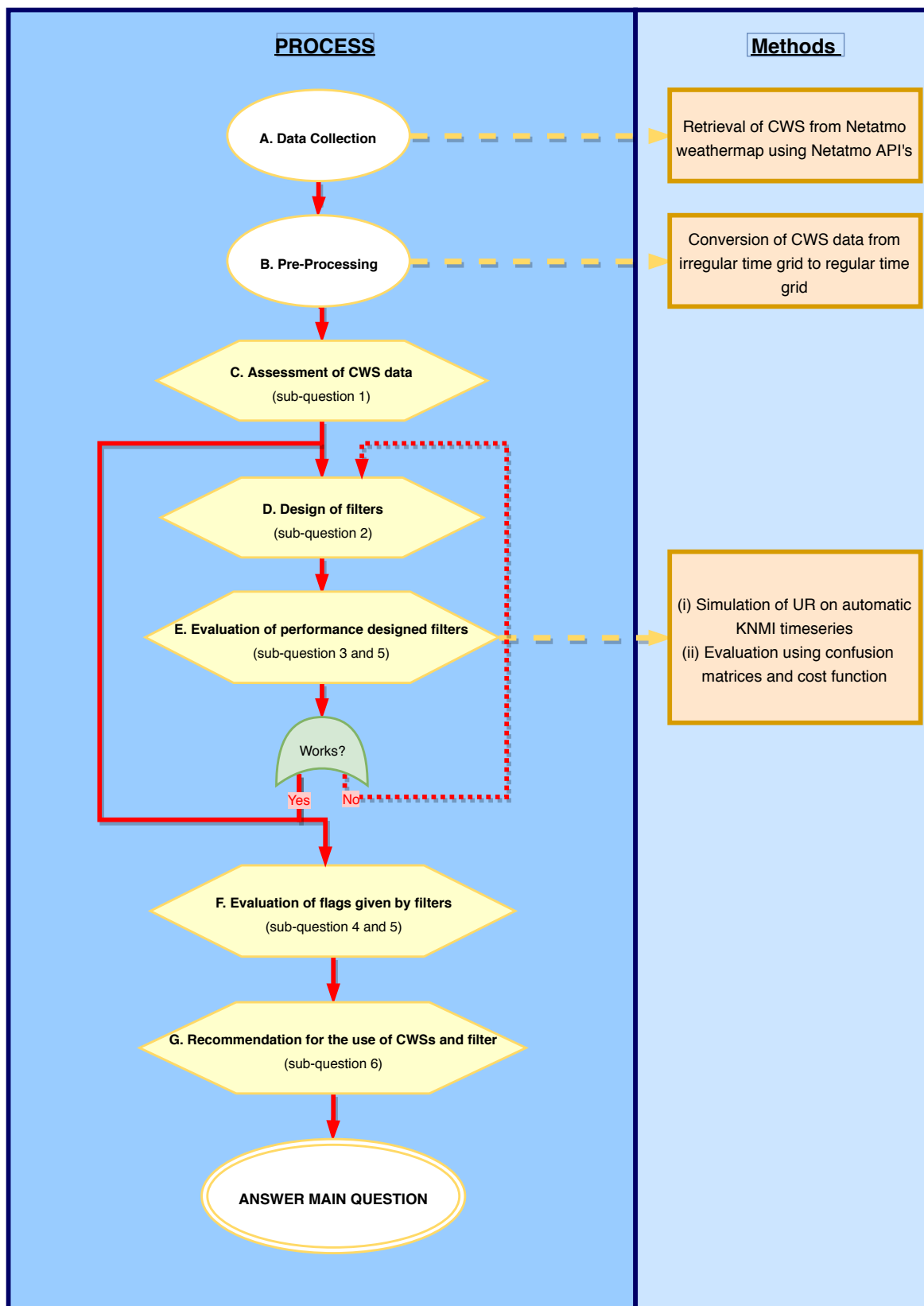


Figure 2.3: A diagram showing the research methodology carried out in this study to achieve the objectives. The process part gives the steps taken to answer the questions. The methods side shows methods that require more in-depth explanation. The red dashed line from performance check (green) to step D indicates that the performances of filter is insufficient (*an iterative process*). If the filter works, then there are two options. The first option is continuing to step F and the other option is to start a new filter.

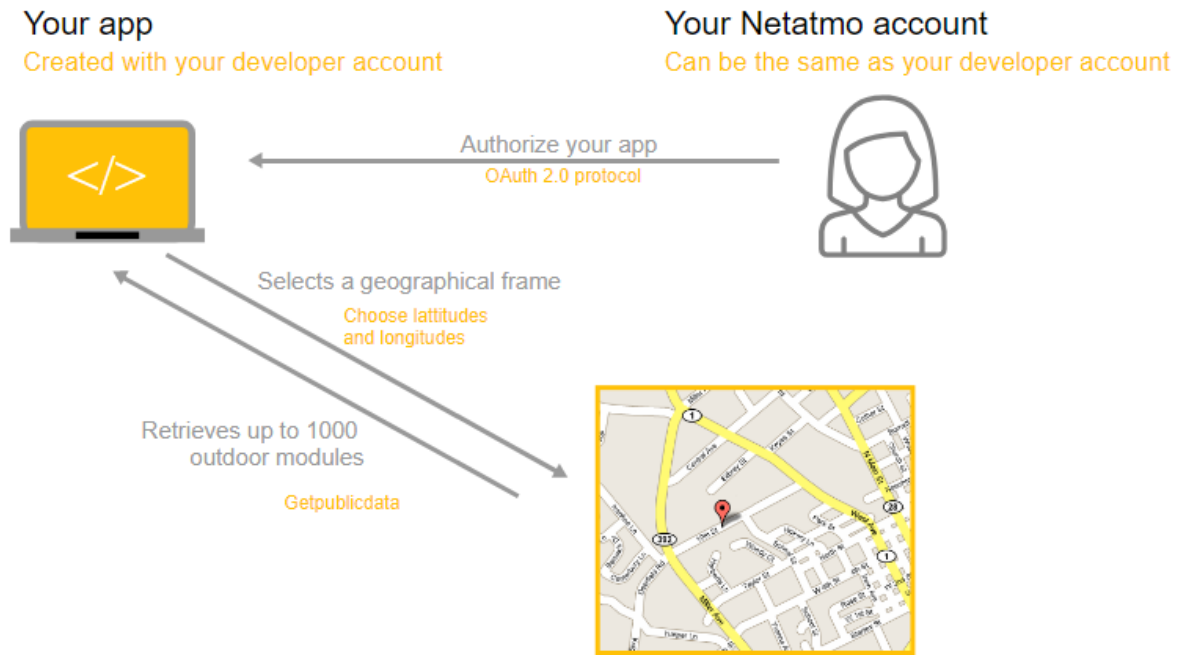


Figure 2.4: An visual representation of Netatmo Weather API structure to retrieve device ID's, rainfall module ID's, locations and most recent measurement using Getpublicdata (source: <https://dev.netatmo.com/en-US/resources/technical/reference/weatherapi>).

2.3.3. Conversion of CWS data from Irregular to Regular Time Grid

The rainfall timeseries of the Netatmo stations have an irregular time grid as already mentioned in section 1.1. About every ~ 5 minutes the number of tips since the last update is reported to the indoor module from the gauge and immediately communicated to Netatmo weathermap (de Vos et al., 2017). Due to this irregularity in time measurements, it is necessary to convert the rainfall timeseries to a regular time grid. A regular time interval of 10 minutes³ is chosen for this study. The transformation of the irregular time grid to the regular time grid is executed including the re-interpolation of rainfall data based on the assumption that constant rainfall intensities occur over the time interval as proposed by the supervisors of this study.

The principle of transformation from an irregular timeline to a regular timeline with re-interpolation of rainfall data works as follows. A regular timeline T_{new} is created. All NA-values in the irregular CWS timeseries CWS_{old} are removed, after which the time series (T_{old}) and rainfall series (R_{old}) are selected from CWS_{old} . The following step is summing R_{old} cumulatively over time resulting in $R_{old,cumsum}$ series. These series $R_{old,cumsum}$ of an irregular timeline T_{old} is then re-interpolated to a regular timeline T_{new} using a linear interpolation function of the package *numpy*. This yields the new variable $R_{new,cumsum}$. A maximum data gap boundary condition is applied, because larger time intervals than 10 minutes in the original timeseries can occur. The maximum time gap Δ_{max} is set on 20 minutes for this research. The timesteps x_{old} in which an interval between two measurement times t and $t-1$ in T_{old} exceeds the maximum time gap boundary, are determined. With these timesteps x_{old} , the timesteps in the regular timeline x_{new} are then retrieved. The values of cumulative rainfall series $R_{new,cumsum}$ at timesteps x_{new} are replaced by NA-values. Lastly, the rainfall series R_{new} corresponding to the regular timeline T_{new} are calculated by subtracting the cumulative sum of the previous timestep from the cumulative sum of the actual timestep. Note that the first step is always a NA-value as result of not being able to quantify over which period the first measurement of original timeseries is representative.

2.3.4. Simulation of UR errors on Datasets

The exact timesteps in which UR errors occur in CWS rainfall data are unknown. Datasets are therefore developed in which the exact timesteps of UR errors are known. The automatic KNMI gauges within the study area are artificially implemented with zeros to simulate UR errors on the datasets. This is done based on the

³Remark that any time interval can be chosen here.

assumption that no UR errors occur in the raw rainfall data of the automatic KNMI gauges. These simulated UR errors are called Zeros Implemented (ZI). An important requirement for the simulation of UR errors on the data is that the zeros are randomly implemented and consist of different lengths of consecutive ZI intervals. Hence a technique is developed to randomly simulate UR intervals on 10 % of the rainfall sequences datasets as shown in Figure 2.5.

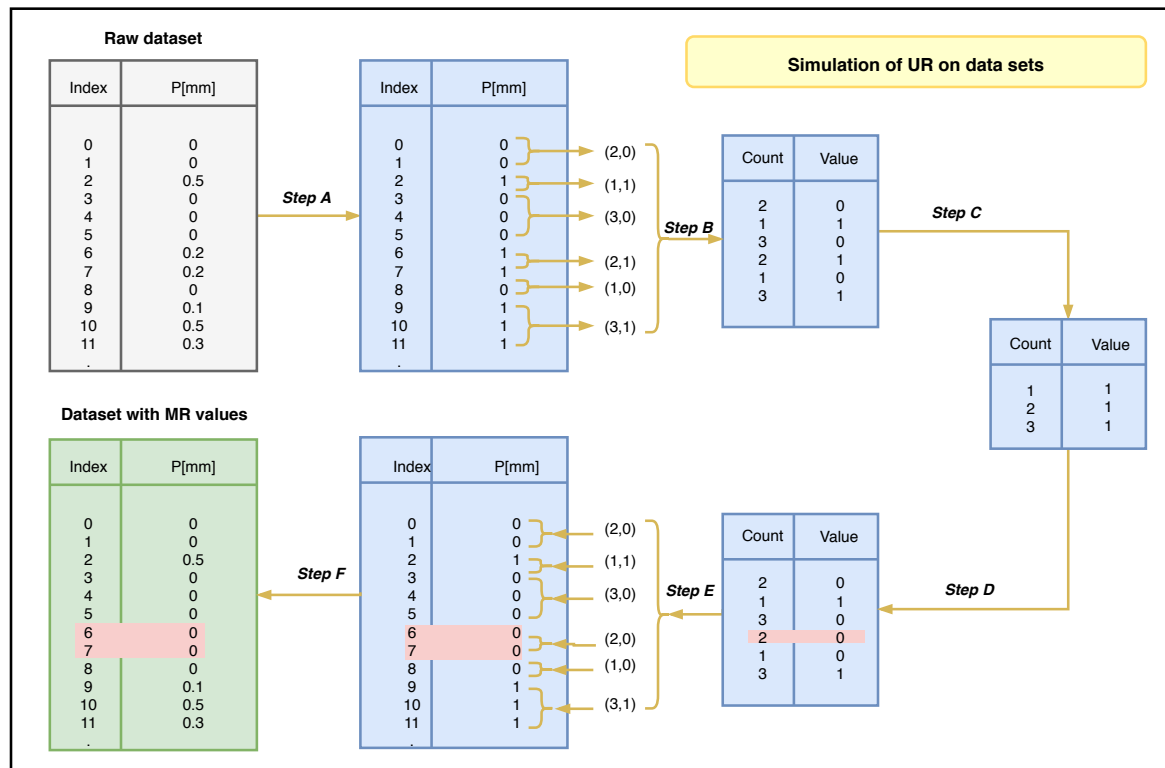


Figure 2.5: A visualisation of the method for simulation UR errors on data sets. The simulation of UR errors is executed by randomly implementing zeros on 10% of the rainfall sequences. UR errors can only be found for the intervals with zeros in which the original rainfall depths were higher than 0.05. Zeros Implemented (ZI) are shown in red and represent simulated UR errors.

The steps of Figure 2.5 are described below.

- A The raw rainfall timeseries is converted to timeseries in which rainfall values, zero values and NA-values become respectively 1-values, 0-values or 2-values.
- B These new type of timeseries are then encoded using the running length encoding function described in subsection 2.2.2. The encoded series consist of the variable 'counts' and 'values'. The 'count' variable give the frequency that a certain value type consecutively occurs. The 'value' variable gives information about the type of value: rainfall, no rain or NA.
- C Rainfall sequences are the rows in encoded series in which the value variable is 1. These sequences are extracted from the encoded series.
- D For 10 % of the rainfall sequences, the values variable are randomly converted from 1-value (rainfall) to 0-value (ZI). These 10 % of altered rainfall sequences replace the original rows in the encoded series.
- E Subsequently, the encoded series with ZI intervals are decoded back to timeseries. Four type values are given in these timeseries, namely rainfall values (1), original zeros and ZI-values (0) and NA-values (2).
- F The type 1-values are replaced by the correct rainfall amount as given in the raw dataset. ZI's, which are implemented on time points that the rainfall amount was lower than the threshold q_{ag} of 0.05 s, are reversed. This means that a simulated UR interval gets overwritten by its original rainfall depth. Appendix B shows the determination of the threshold value q_{ag} .

2.3.5. Confusion Matrices and Cost Function

A confusion matrix is a technique for visualising the performance of a classification algorithm. The QC filters applied in this study are classification algorithms that use the principle of decision trees to classify the rainfall

data on the occurrence of UR errors. Each timestep automatically continues the filter to obtain no flag or UR flag that states if it is or is not an UR interval. To determine if the filters perform well, a confusion matrix is applied. Confusion matrices summarise the possible outcomes of the truth and the predictions. Predictions are here the outcomes of 'no flag' or 'UR flag' given by the filters. In other words, a confusion matrix shows the ways in which the QC filters are confused when it is flagging.

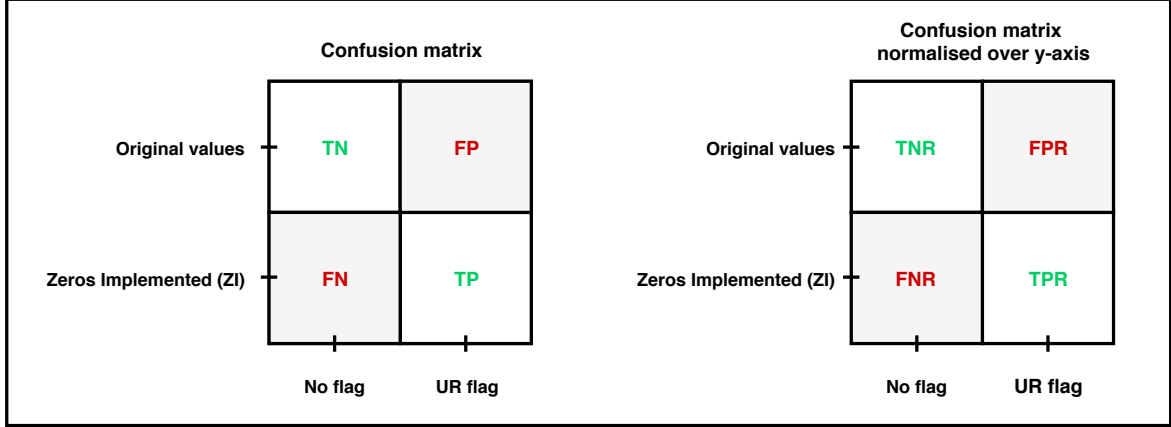


Figure 2.6: A concept of a confusion matrix to evaluate the performance of the filters in the detection of UR errors. On the left side is the normal confusion matrix given with its characteristics. A confusion matrix normalised over y-axis is shown on the right side. Correct and incorrect possibilities are shown by green and red colour, respectively. An explanation of the abbreviations is shown in Tables 2.4 and 2.5.

Figure 2.6 shows the concept of the confusion matrix applied to evaluate the performance of the filters in the detection of UR errors. The left side of the figure shows the normal confusion matrix. Filters flag or do not flag a specific timestep as UR. This option is given by x-axis. The y-axis is the truth axis. Original values implies that no UR errors are implemented on the specific timestep, so these intervals are original values. ZI means that an UR error is simulated at that interval. Within the confusion matrix four options are given, namely True Negatives (TN), False Positives (FP), False Negatives (FN) and True Positives (TP). Table 2.4 gives the terminology for the different options as applied in this study.

Table 2.4: Basic terminology of the confusion matrix for this research.

Term	Definitions
True Negatives (TN)	The filter correctly flagged that no UR error occurred
False Positives (FP)	The filter incorrectly flagged that UR error occurred
False Negative (FN)	The filter incorrectly flagged that no UR error occurred
True Positives (TP)	The filter correctly flagged that UR error occurred

Several metrics can be computed from the confusion matrix. Table 2.5 gives the most important metrics for this study. Equation 2.1 to 2.5 are the mathematical formulas for the computation of True Negative Rate (TNR), False Positive Rate (FPR), False Negative Rate (FNR), True Positive Rate (TPR) and False Detection Rate (FDR), respectively. These terms can also be calculated by normalising the confusion matrix over the y-axis as is displayed on the right side of Figure 2.6. For FDR, the confusion matrix should be normalised over the x-axis. In this research, the focus is mainly on TPR, FNR and FDR.

$$TNR = \frac{TN}{TN + FP} \quad (2.1)$$

$$FPR = \frac{FP}{TN + FP} \quad (2.2)$$

$$FNR = \frac{FN}{TP + FN} \quad (2.3)$$

$$TPR = \frac{TP}{TP + FN} \quad (2.4)$$

$$FDR = \frac{FP}{TP + FP} \quad (2.5)$$

Table 2.5: Basic terminology of metrics computed from the confusion matrix for this study

Term	Definitions
True Negatives Rate (TNR)	Percentage of original values that are correctly flagged.
False Positives Rate (FPR)	Percentage of original values that are incorrectly flagged.
False Negative Rate (FNR)	Percentage of ZI's (simulated UR errors) that are incorrectly flagged.
True Positives Rate (TPR)	Percentage of ZI's (simulated UR errors) that are correctly flagged.
False Discovery Rate (FDR)	Percentage of UR flags that are incorrectly given by filter.

Since the objective of the filters is to detect all UR errors in data, the performance of the filters is mainly focussed on getting FNR as close as possible to 0 without having unreasonable give-and-takes. This means that the frequency of the filter incorrectly flagging ZI's, has to be low. The frequency of the filter correctly flagging ZI's, is required to be high. The filter performance is thus optimised based on FNR. The lower FNR becomes as a result of strict filters, the higher the number of FP are. The strict filters prefer to detect and flag all UR values. By doing so good data is sacrificed (original values with no UR errors). Finally, FDR also becomes higher as result of this.

An additional performance indicator is investigated to better evaluate the performance of the filter, namely the relative error or relative cost. This indicator is similar to FNR. Instead of using the counts of occurrence, as done by FNR, the true rainfall amounts of FN, the so-called costs, are used. The total cost is the sum over all true rainfall amounts of FN. The relative cost can be computed as the total cost dividing by the sum of the total cost and total benefit. The total benefit here is the total rainfall amount of ZI's that are correctly flagged by the filters (TP). A comparison between FNR and relative cost is expected to show that the filters perform better during higher rainfall intensities.

3

Properties and Errors of CWSs

In this chapter, the properties and errors of rainfall data in the CWSs network within the study area are described in order to answer the first sub-question: *What are the properties and errors of rainfall data in the CWS network in the study area?* First, the properties of rainfall measurements from CWSs in the study area are discussed. In section 3.2, the errors found in the rainfall data are explained.

3.1. General Properties of CWSs in study area

Figure 2.1 shows the 760 operational Netatmo CWS linked to Netatmo platform within the study area. These CWSs are obtained with Getpublicdata API (subsection 2.3.2). The Getmeasure API is used to collect the time-series of those stations for research period (1 October 2015 until 1 October 2017). Of the 760 CWSs linked to the Netatmo platform, 7 stations only contain measurements before or after the selected period for this study. The remaining 753 CWSs are used in our analysis.

The 753 CWSs generally are in close proximity of each other in populated areas as shown in Figure 2.1. This is also stated by Kidd et al. (2017). The CWS network is thus denser in cities, e.g. Leiden, The Hague and Rotterdam. Based on the assumption that CWSs are uniformly spread over land area in the study area, a density of 1 CWSs per 5.93 km² is found. For the city of Rotterdam a density of 1 CWS per 5.49 km² is computed as result of 59 CWSs within the municipality of Rotterdam. The densities differ significantly less from each other than expected. A reason for this could be that the municipality of Rotterdam not only consists of populated areas, such as the Europoort harbour.

Not all 753 CWS are continuously measuring during the total research period as shown in Figure 3.1. However the amount of available CWSs is growing from 369 CWSs to 668 CWSs during the period 1 October 2015 to 1 October 2017. A CWS is assumed to be available when it measures at least once per day. The growth of available CWSs increases gradually with here and there some sudden growths and halts in the increase of CWSs. The month December in 2015, for example shows a rapid growth of available Netatmo stations. This rapid growth can be a result of people getting Netatmo CWSs due to the holiday seasons. Between April and June 2016, a halt occurs in the increase of available CWSs. A specific reason cannot be given for these alternations. In contrast to the growth of number of available CWSs, the steepness of this growth becomes less over time.

Owners of Netatmo CWSs can calibrate their Netatmo station, however it is not required. Figure 3.2 shows that 70 owners did calibrate their station. The calibration of these 70 owners are based on deviating tipping bucket volumes from default value of 0.101 millimetre. Most (67) of the owners, who decided to calibrate the CWS, found higher tipping bucket volumes than the default value. These result can be explained by two reasons: the actual tipping bucket volume being higher than stated by the manufacturer or people pouring water too quickly in the bucket during calibration resulting in a higher calibrated tipping volume. For three CWSs (id's: 59,490 and 738) determination of the calibrated tipping bucket volume is not possible, since the timeseries have only zero-values. In short, most owners do not calibrate the tipping bucket volume. But when the tipping bucket volume is calibrated, often a higher volume is found.

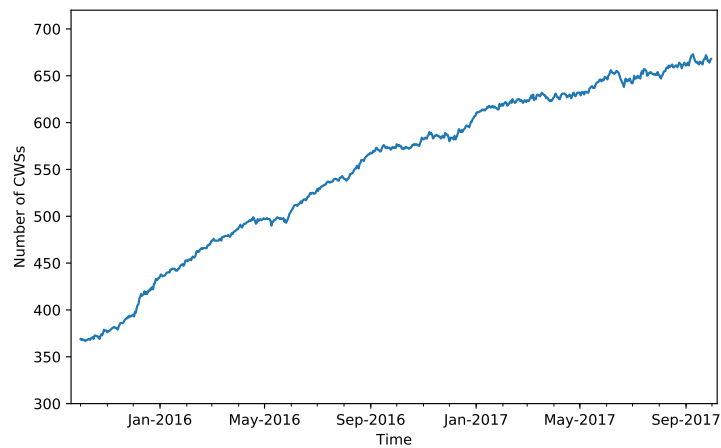


Figure 3.1: Number of available Netatmo CWSs measuring at least once per day over period from 1 October 2015 to 1 October 2017. The number of available CWSs per day is gradually growing from 369 CWSs to 668 CWSs.

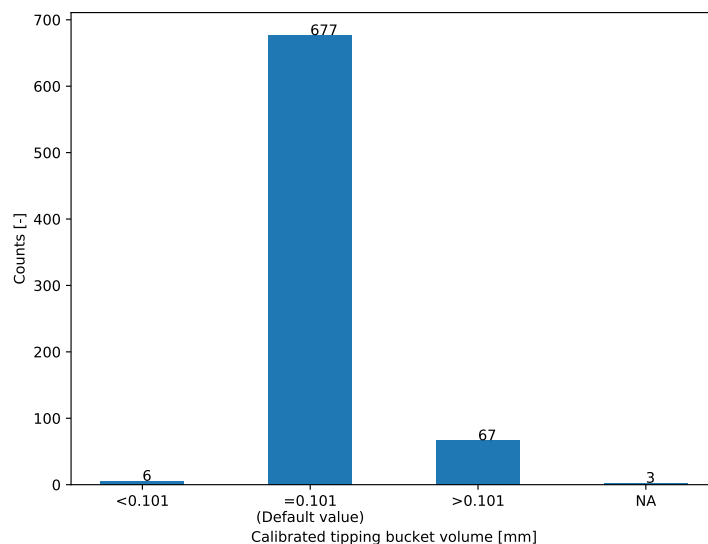


Figure 3.2: Calibrated tipping bucket volumes found for 750 stations. For three CWSs (59, 490 and 738) no calibrated tipping bucket volumes can be determined, since these stations contain only zero-values.

In short, the CWSs within the study area are mainly located at densely populated area. Most of the CWSs are not available the entire research period, however, the number of CWSs is growing over the two years. In total, 7 of the 760 stations are not further used due to measurements taken before or after the selected period of this study. Lastly, 70 owners calibrated the Netatmo rain gauges. Most calibrated device have tipping bucket volumes larger than the default value of 0.101 millimetre.

3.2. Errors of CWSs in study area

Three types of erroneous behaviours are expected to be encountered for in citizen rainfall data, as discussed in section 1.2. These behaviours are the occurrence of lacking intervals (NA-filled intervals), the identification of false rainfall intensities and UR errors. In this section, the citizen rainfall data is investigated to check if these behaviours occur in the data.

One of the errors expected in the data are lacking intervals. Over the selected period for this study, ~5% of all 10-min intervals are lacking (data not available). The monthly variations in lacking intervals are given in grey in Figure 6.3. The minimum and maximum percentages of monthly lacking intervals are 3% and 6% for respectively October 2015 and March 2017. Besides this, the figure shows that the monthly lacking interval percentage increases slightly over time on average. The percentage of lacking intervals per CWSs (LI_{CWS}) is shown in Figure 3.3¹. The lacking intervals per CWS is computed by dividing the number of intervals in which lacking intervals occurred by the total number of intervals that a station is operational. The figure shows that the percentage of lacking intervals per CWS deviates between 0% and 74%. Many CWSs (711) have a station outage between 0% and 20%, 29 CWSs are not giving data between 20% and 40% of their measurement period and 13 CWSs are disrupted more than 40% of measurement time. In other words, the percentages of lacking intervals deviate significantly more over the different CWSs compared to percentages of lacking intervals over time.

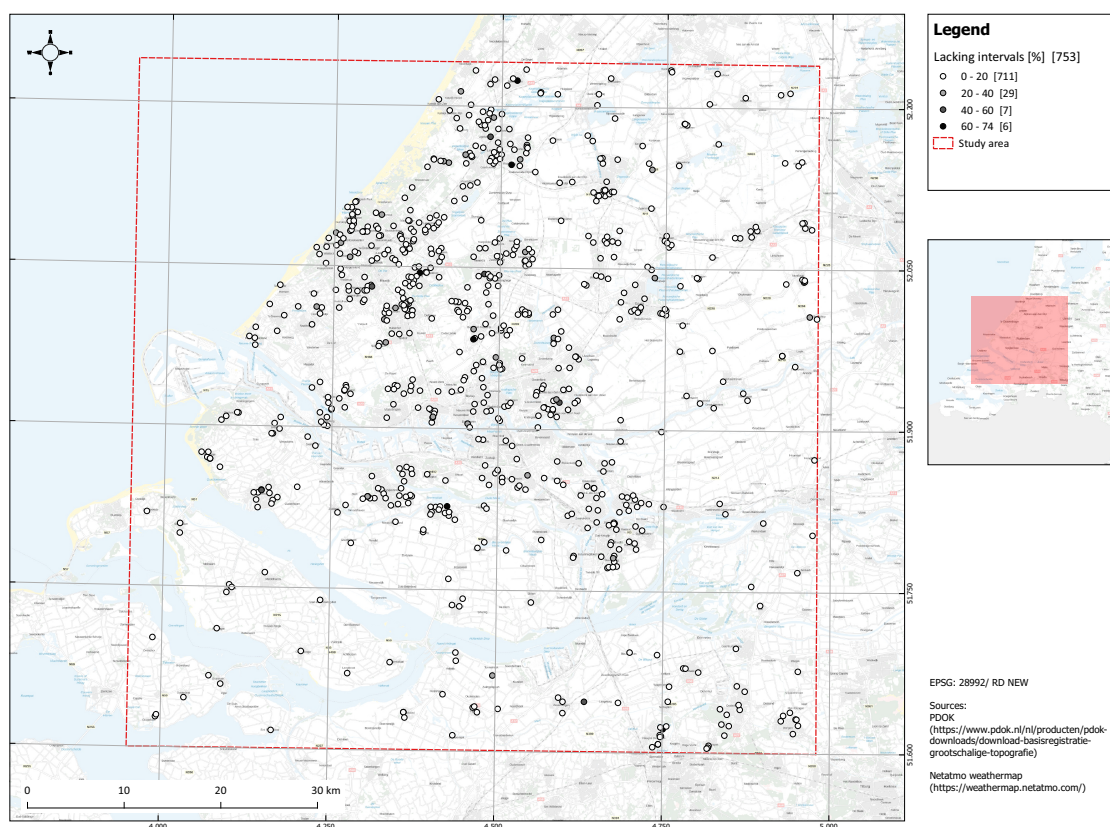


Figure 3.3: Percentages of lacking intervals per CWS in study areas. Appendix E gives the figure on actual size.

The analysis below includes measurements of 293 CWSs that measure the total research period (October 2015 to October 2017). Additionally, only the first year (October 2015 to October 2016) of the data is applied.

Figure 3.4 shows that yearly accumulated rainfall amounts of 293 CWS deviate strongly per CWS. 212 CWSs have a yearly rainfall sum between 600 and 1000 millimetres. 28 CWSs measured more than 1000 millimetres in first year, and 53 CWSs have a yearly rainfall sum lower than 600 millimetres. In the 28 CWSs that measure more than 1000 millimetres for the first year, six CWSs measure more than 2000 millimetres of rainfall in one year. Five out of six of those CWSs have a default tipping bucket volume. One station has a tipping bucket volume of 0.119 millimetre. The percentages of lacking intervals for these stations is between 0.16% and 5.5% meaning that the fractions of lacking intervals are less or at most 1% higher than the total percentage of lacking intervals (5%). Three CWSs (with ID's 25,173 & 645) have unlikely high yearly rainfall amounts of more than 30000 millimetres. After visual inspection, these CWSs appear to function partially normal and partially

¹Appendix E gives the figure on actual size

biased due to cumulative values. However, the cumulative parts are occasionally reset. This behaviour is also shown in the Double-Mass (DM) curve of CWS ID173 given in Figure 3.5. CWS 173 has cumulative values in the beginning after which it is reset and follows approximately the same slope as the other CWSs in the figure. The other three CWSs (with ID's 215, 483 & 706) do not have cumulative values, therefore, it is assumed that those CWS have false rainfall intensities in the data.

In the 53 CWS, which yearly sums are below 600 millimetres, 3 CWSs have a yearly sum below 10 millimetres. The tipping bucket volume of these CWSs are all set on the default value. One CWS (ID104) only tipped once in the entire year. However, the percentage of lacking intervals is only 3.4%. This could indicate that CWS is situated in an area in which no rainfall can be captured. Two CWSs (with ID's 246 & 264) have a yearly sum of 9.7 and 2.8 millimetres. Beside those three CWSs, the other 50 CWSs also give yearly sums, which are outside of the lower limit of a reasonable yearly rainfall sum for the study area. For some CWSs, this is a result of high percentage of lacking intervals, others have false rainfall intensities in the data.

Two DM-curves² are given in Figures 3.6 and 3.7. On the x-axis the cumulative precipitation depths at the automatic KNMI gauges are given and, on the y-axis the cumulative rainfall depths of CWSs are shown. The CWSs are only taken into account if they measure for first year and are within 5 kilometres radius of the automatic KNMI gauge. Figures 3.6 and 3.7 show respectively the data of 14 and 12 CWSs. Subsequently, 9 and 11 CWSs are underestimating against 5 and 1 CWSs overestimating for Voorschoten and Rotterdam. This implies that most CWSs are underestimating the rainfall based on the assumption that the automatic KNMI gauges give the real amount. Bell et al. (2013) also found that CWSs of other brands underestimate compared to an official rain gauge maintained by Met Office. Because of this, it can be stated that CWSs do not agree with the references, however, the magnitude of this dissimilarity strongly varies per CWSs.

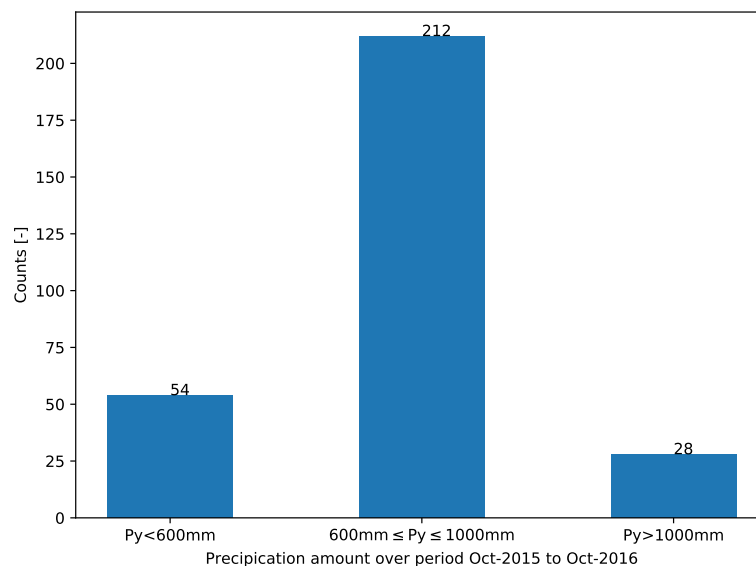


Figure 3.4: Yearly accumulated rainfall sums of CWSs of first year (Oct-2015 to Oct-2016), which measured the entire research period.

²All timesteps with NA-values are simultaneously removed for both CWSs, as well as, automatic KNMI gauges

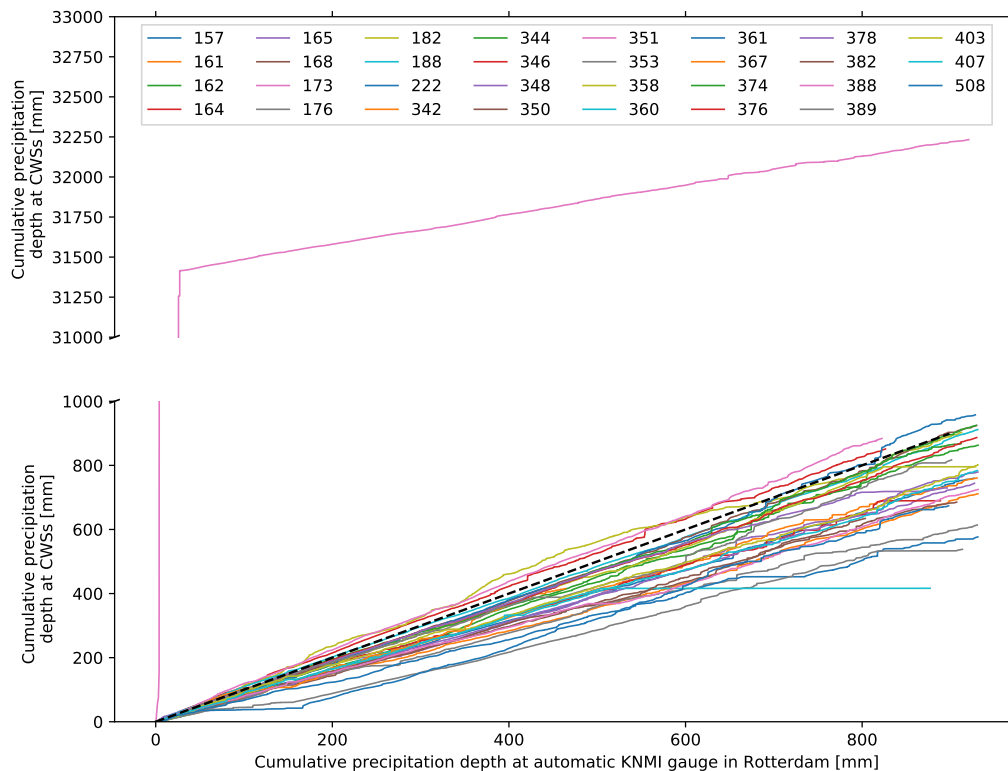


Figure 3.5: Double mass curve between the automatic KNMI gauge in Rotterdam on x-axis and CWSs in a radius of 8 kilometre on y-axis. CWS ID173 appears to function partially normal and partially biased due to cumulative values. These cumulative parts are reset once in a while.

Multiple reasons can be addressed for deviation of slope of CWSs compared to slope of automatic KNMI gauges. CWS ID461 in Figure 3.6 overestimates significantly due to a tipping bucket volume of 0.236 millimetre. CWS ID454 in Figure 3.6 is a relatively stable stations that slightly overestimates. A cause for the overestimation is most likely the higher tipping bucket volume of 0.12 millimetre compared to the default volume. CWSs that are underestimating and close to the slope in Figure 3.7, have a tipping bucket volume of 0.101 millimetre, the default tipping bucket volume. All these results together could imply that the effective tipping bucket volume is higher than 0.101 millimetre but lower than 0.12 millimetre.

Besides the deviations of slopes between CWSs and the automatic KNMI gauge, these figures also show that some CWSs are significantly fluctuating and have no straight slope (CWS ID318 and CWS ID453). Both station have a yearly sum higher than 1000 millimetre but tipping bucket volumes of 0.101 millimetre. False rainfall intensities therefore occur in these CWSs. Another type of erroneous behaviour is UR errors. An UR errors occurs when a CWS gives zero rainfall, but it was actually raining at that time. So, the cumulative precipitation depth of CWS does not increase, but the cumulative precipitation depth of automatic KNMI increases. This gives horizontal lines in DM-curves. UR errors are, for example, found in CWSs ID274 and ID449 of Figure 3.7. CWS ID274 has a horizontal line near 470 millimetres on x-axis, and in CWS 449, a relatively long period of UR intervals is shown.

Otherwise stated, the three types of erroneous behaviours mentioned in section 1.2 are occurring in the rainfall data of CWSs in the study area. Additional errors are identified on top of that, namely the occurrence of cumulative values in the data and one station with only one tip the entire year.

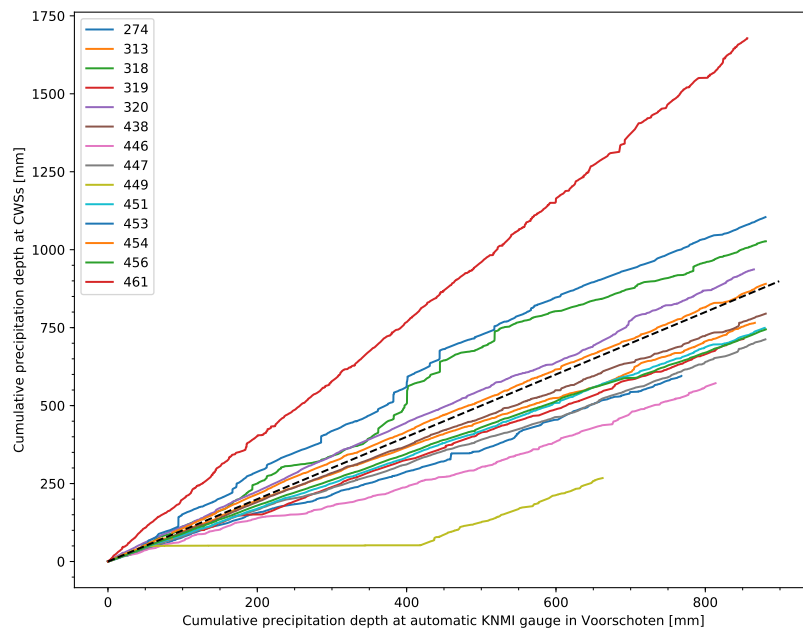


Figure 3.6: Double mass curve between the automatic KNMI gauge in Voorschoten on x-axis and CWSs in a radius of 5 kilometre on y-axis.

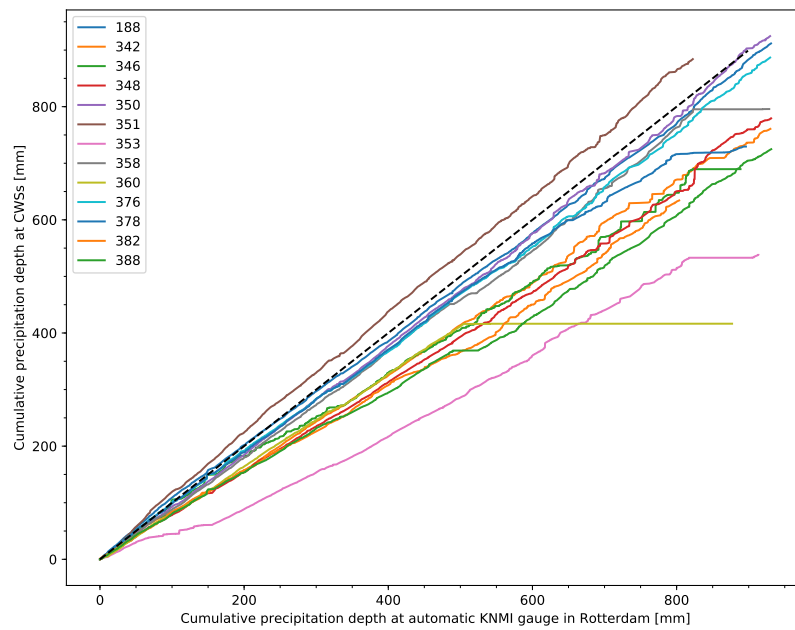


Figure 3.7: Double mass curve between the automatic KNMI gauge in Rotterdam on x-axis and CWSs in a radius of 5 kilometre on y-axis.

4

Design of Filters for UR errors in CWSs

In this chapter, the designs of filters for identification of UR errors are discussed including their limitations, benefits, assumptions and computations of relevant features required to run the filters. Two types of filters are developed. The first filter uses neighbouring CWSs to detect UR intervals, therefore it is called CWS filter. The second filter applies the overlying radar pixel to find UR intervals, therefore it is called Radar filter.

The task of the filters is to automatically detect and flag UR's in CWS_n at every timestep. Table 4.1 shows the output options (flags) that a filter can give to the CWS_n at every timestep. Note that an UR interval is not similar to lacking intervals. Lacking intervals are Not Available (NA) values.

Table 4.1: Different flags assigned by the filters to $CWS_n(t)$.

Tag number	Colour	Description
$F_n(t) = 0$	green	$CWS_n(t)$ passes filter resulting in no UR flag
$F_n(t) = 1$	red	$CWS_n(t)$ flagged as UR

4.1. CWS Filter

Figure 4.1 shows the design of the CWS filter. Table 4.2 explains the mathematical expressions applied in the steps of CWS filter. Each CWS of interest follows the different steps of the CWS filter to determine the most probable output flag $F_n(t)$ for $CWS_n(t)$. An output flag marks if an interval is identified as UR or if it passes the filter. Two different flags can be assigned to $CWS_n(t)$ as shown in Table 4.1. The CWS filter stops after the last time step available for CWS_n is processed. Input of the CWS filter is the CWS_n retrieved from your CWS database.

Table 4.2: Description of variables and symbols implemented in CWS filter.

Variable	Unit	Description
CWS_n	-	Regular timeline and rainfall data of CWS with ID
$CWS_n(t)$	-	Data of CWS attimestept
$F_n(t)$	-	Output flag of filter attimestept
$F_n(t-1)$	-	Output flag of filter at previous timestep t-1
$P_{CWS,n}(t)$	mm	Precipitation depth of CWS_n attimestept
$N_{med,n}(t)$	-	Number of neighbouring CWSs inside radius of CWS_n attimestept
$P_{med,n}(t)$	mm	Median precipitation depth of neighbouring CWSs inside radius of CWS_n attimestept

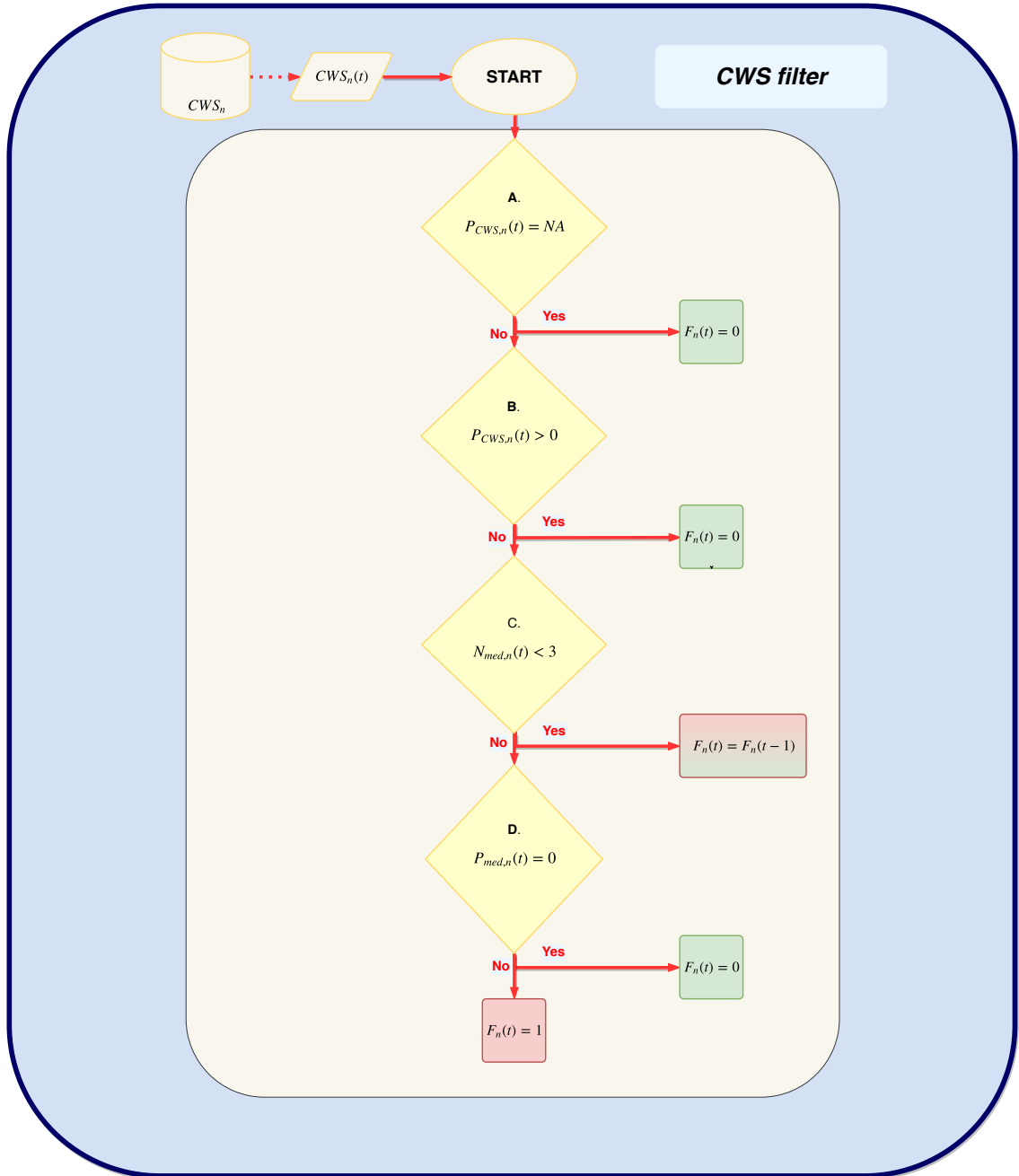


Figure 4.1: A QC decision tree to identify UR errors in CWS rainfall data using neighbouring CWSs within a radius of 8 kilometre. The QC filter mechanisms consist of decisions¹ (◇), processes (□), data (▭), database (⊞)

A *Is the CWS available at timestep t?*

Statement: Precipitation depth of $CWS_n(t)$ is not available. (Equation 4.1)

Explanation: Station outage is one reason why rainfall data is lacking (not available). Since lacking intervals are not the same as UR, further continuation of the filter is not required for a specific timestep t . Equation 4.1 gives therefore the decision: if precipitation depth is not available at t , then CWS_n passes the CWS filter at t . If the answer is no to the decision A, then decision B becomes applicable.

Required actions: Determine $P_{CWS,n}(t)$.

$$P_{CWS,n}(t) = NA \quad (4.1)$$

B *Is precipitation occurring at timestep t?*

Statement: Precipitation occurs at timestep t. (Equation 4.2)

Explanation: The second step checks whether precipitation occurs at CWS_n for timestep t. If precipitation occurs at CWS_n(t), then an UR error is precluded. The flag F_n(t) is zero meaning that the CWS passes the filter at this timestep and gets no UR flag. A 'No' answer on equation 4.2 means that CWS_n(t) shows no precipitation. Two different events are then possible: (i) no precipitation is occurring, or (ii) CWS_n(t) is an UR error. To find out which of the two is happening, more additional checks are required. Hence the CWS_n(t) continues to decision C.

Required actions: The same computation as executed by step A.

$$P_{CWS,n}(t) > 0 \quad (4.2)$$

CWSs that give a 'No' answer to step B at timestep t, can be an UR error as mentioned above. These CWSs require therefore additional checks, namely step C and D.

C *Is the number of neighbouring CWSs within radius of CWS_n less than 3 at timestep t?*

Statement: Number of neighbouring CWSs within radius of CWS_n is less than 3 at timestep t.

Explanation: The third step is implemented as preparation for the next step to determine if the computation in step D is allowed. If the number of neighbouring CWSs within the radius is less than 3, the CWS_n(t) gets the flag obtained in the previous timestep. If the number of neighbouring CWSs is at least 3, then the CWS_n(t) continues to the last step D.

Required actions: Determine the number of CWSs within the radius approved of CWS_n.

$$N_{med,n}(t) < 3 \quad (4.3)$$

D *Are the neighbouring CWSs within the radius also measuring zero precipitation at timestep t?*

Statement: The median precipitation depth of the neighbouring CWSs within the radius of CWS_n is zero at timestep t. (Equation 4.4)

Explanation: If the median precipitation depth P_{med,n}(t) of the neighbouring CWSs within the radius of CWS_n is zero, then it is more likely that no rainfall is occurring at timestep t. This results in a flag F_n(t) is zero resulting in no UR flag. If the statement 4.4 is false, then CWS_n gets the flag F_n(t) is one meaning that it is a possible UR error.

Required actions: (i) Retrieve the rainfall depths of the neighbouring CWSs within radius of CWS_n at t; (ii) compute the median from the rainfall depths;

$$P_{med,n}(t) = 0 \quad (4.4)$$

Determination of radius

CWS rainfall data of each timestep passes the filter (F_n(t)=0) or receives an UR flag (F_n(t)=1) as result of quality-control checks of the CWS filter. The most important checks are steps C and D. These steps determine if an interval of CWS rainfall data is a possible UR error by using neighbouring CWSs within a fixed radius of 8 kilometre. This fixed radius is determined by applying the CWS filter on raw rainfall data from automatic KNMI gauges (Table 4.3) within the borders of the study area. By doing so, the assumption is made that the CWS filter should not give UR flags for the rainfall data obtained from automatic KNMI gauges for radii between 5 km and 15 km. Based on the lowest percentages of UR flags an optimum radius of 8 kilometre is selected with focus on the gauges in Voorschoten and Rotterdam. Therefore, the radius of 8 kilometre is verified by the percentage of CWSs within the study area that have at least 3 neighbouring CWSs within this radius. This verification shows that 99.5% of all CWSs have at least 3 neighbouring CWSs within a radius of 8 kilometre. The radius of 8 kilometre is therefore accepted. Appendix C demonstrates the determination and verification of the fixed radius more in detail.

Benefit

The CWS filter does not use other sources to detect and flag UR values in CWS data such as radar and automatic KNMI gauges.

Table 4.3: ID, location and distance to Delftse Poort of automatic KNMI gauges within the study area.

ID	Location	Distance to Delftse Poort
AG215	Voorschoten	24.14 km
AG330	Hoek van Holland	25.21 km
AG344	Rotterdam	4.43 km
AG348	Cabauw	31.61 km

Limitations and assumptions

- Due to the implementation of decision C, the filter is not able to start with flagging before the moment that the CWS is available. This results in shorter processed datasets. This delay in flagging is determined: (i) by the moment it starts raining or that intervals are lacking; (ii) or the moment that at least three neighbouring CWSs are available within the radius. In short, check C can determine the moment that the CWS filter starts flagging and is therefore a possible temporal constraint for CWS filter.
- Decision D is a spatial consistency check in which is assumed that the majority of the neighbouring CWSs within radius of CWS of interest function properly.
- Decision D has another constraint, namely the requirement of at least three neighbouring CWSs to compute the median. To avoid median precipitation depths that are not available in decision D, decision C is therefore implemented.

4.2. Radar Filter

Figure 4.2 shows the design of the radar filter. Table 4.4 explains the mathematical expressions applied in the steps of the radar filter. Each CWS of interest follows the different steps of the radar filter to determine the most probable UR flag $F_n(t)$ for $CWS_n(t)$. The same options of flags as for the CWS filter can be assigned to $CWS_n(t)$. These are given in Table 4.1. Instead of using neighbouring CWSs as done by the CWS filter, the radar filter uses the precipitation depths of the overlying radar pixel.

Table 4.4: Description of variables and symbols implemented in CWS filter.

Variable	Unit	Description
CWS_n	-	Regular timeline and rainfall data of CWS with ID
$CWS_n(t)$	-	Data of CWS attimestept
$F_n(t)$	-	Output flag of filter attimestept
$P_{CWS,n}(t)$	mm	Precipitation depth of CWS_n at timestep t
$P_{rad,n}(t)$	mm	Precipitation depth of overlying radar pixel attimestept t
q_{rad}	mm	Threshold value

A Is zero precipitation occurring at timestep t?

Statement: Precipitation depth of $CWS_n(t)$ is zero. (Equation 4.5)

Explanation: The first step checks whether zero precipitation occurs in CWS_n at timestep t. If precipitation occurs or the interval is lacking at $CWS_n(t)$, then CWS_n passes the radar filter at timestep t and gets a green flag. A 'Yes' answer on equation 4.5 means that $CWS_n(t)$ gives zero precipitation. Two different events are then possible: (i) zero precipitation is occurring, or (ii) $CWS_n(t)$ is a possible UR error. To find out which of the two is right, $CWS_n(t)$ continues to step B.

Required actions: Determine $P_{CWS,n}(t)$.

$$P_{CWS,n}(t) = 0 \quad (4.5)$$

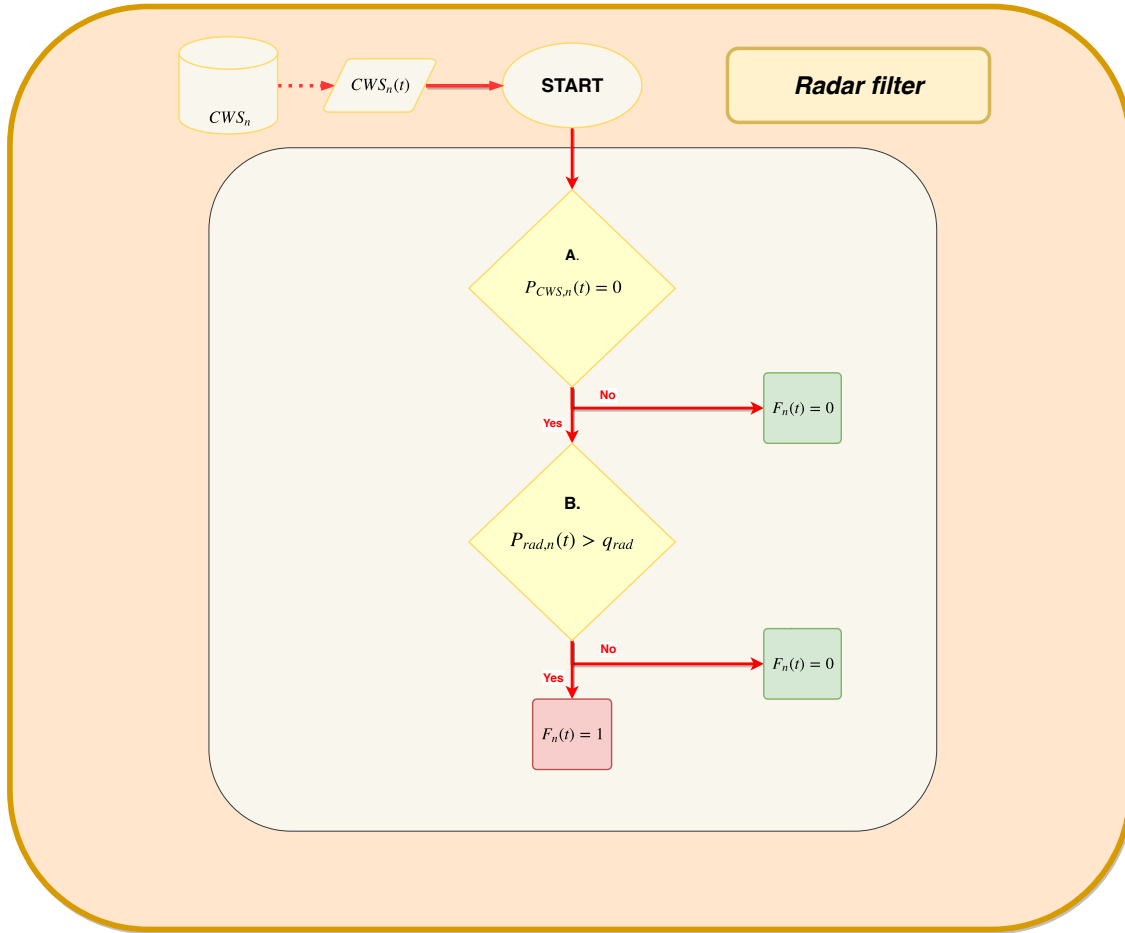


Figure 4.2: A QC decision tree to detect and flag UR errors in CWS rainfall data using the overlying radar pixel. The QC filter mechanisms consist of decisions² (◇), processes (□), data (▭), database (⊞)

B Does the overlying radar pixel give precipitation depth higher than 0.06 millimetre (q_{rad})?

Statement: The precipitation depth of the overlying radar pixel is greater than 0.06 millimetre. (Equation 4.6)

Explanation: The last check is to find out whether the overlying radar pixel estimates a precipitation depth higher than q_{rad} . If equation 4.6 is false, then $CWS_n(t)$ gets a green flag ($F_n=0$). A 'Yes' answer to the statement results in an UR flag ($F_n=1$)

Required actions: Determine $P_{rad,n}(t)$.

$$P_{rad,n}(t) > q_{rad} \quad (4.6)$$

Determination of q_{rad}

The time interval of CWS does or does not pass the filter dependent on the flag they get during the two checks. An UR flag awarded to $CWS_n(t)$ implies for the radar filter that the rainfall depth measured by CWS at timestep is zero but the precipitation depth of the overlying radar pixel exceeds the radar threshold q_{rad} of 0.06 millimetre. A comparison on the occurrence of rainfall between the automatic KNMI gauges and radar showed the need of a threshold. Appendix D demonstrates that from this comparison can be found that the overlying radar pixel estimates rainfall values at certain timesteps, whereas the gauge measures zero. The radar filter would automatically flag these timesteps as UR values without being an actual UR value, therefore a more in detail investigation of these intervals is done, called drizzle check. The drizzle check investigates whether these timesteps are low rainfall values. The drizzle check, demonstrated in appendix D, shows that the 75th percentile of these timesteps is approximately 0.06 millimetre. This means that 75% of timesteps are low rainfall values below 0.06 millimetre in which the overlying radar pixel gives rain and the gauge measures no rainfall. It is therefore decided to implement a threshold value. An additional analysis is executed to determine the best threshold value as last in appendix D. A threshold of 0.06 millimetre is considered to be the

optimum threshold value. With this threshold is assumed that values of the overlying radar pixel above the threshold are rainfall values, whereas values below or equal to the threshold are zero rainfall values.

Benefit

The radar filter has no delays in detection and flagging of UR errors in CWSs, since it only requires the precipitation depth of the overlying radar pixel. The radar filter is therefore more suitable to regions, which have a sparse CWS network.

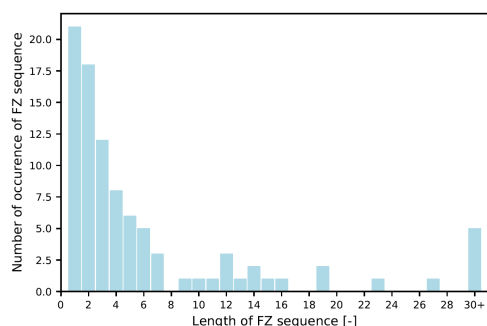
Limitations

- Due to the use of radar, the filter is no longer an independent filter.
- The performance of the radar filter depends on the rainfall intermittency that corresponds to the use of radar and rain gauges. Rainfall intermittency is the alternation between rainfall and zero rainfall periods.
- Disagreement about the occurrence of rainfall may cause the radar filter to perform subpar. These disagreements result from differences in measurement techniques and in space-times resolutions between radar and gauges. Radar artefacts exacerbate these disagreements.

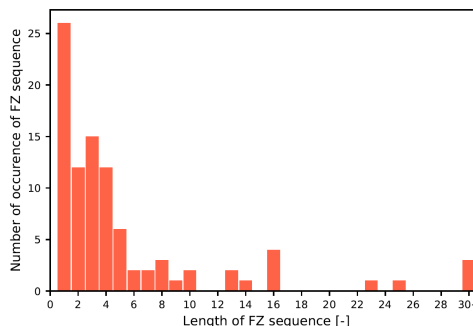
5

Performance of designed Filters

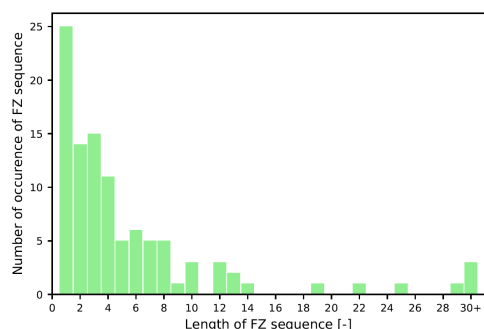
In this chapter the performances of the designed filters are evaluated. In doing so, UR errors are artificially introduced by implementing zeros on the automatic KNMI data sets of stations given in Table 4.3. An example of the lengths of consecutive ZI intervals are given in Figures 5.1a to 5.1d. In the same figures the frequencies of these lengths that are implemented on the automatic KNMI rainfall data are also given. The figures show that ZI sequences have mainly consecutive lengths between 1 and 4 UR values. This means that the rainfall sequences selected in step B of simulation method, as described in subsection 2.3.4, have mainly sequences between 1 and 4 UR values. The graphs also shows that some longer ZI sequences are simulated on the datasets.



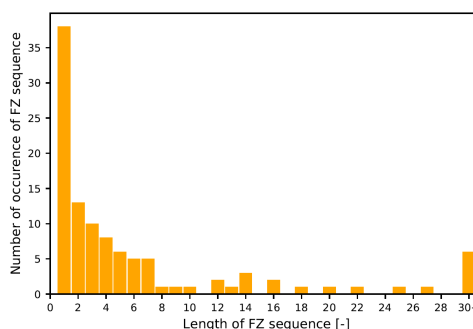
(a) Automatic KNMI gauge 215 in Voorschoten



(b) Automatic KNMI gauge 330 in Hoek van Holland



(c) Automatic KNMI gauge 344 in Rotterdam



(d) Automatic KNMI gauge 348 in Cabauw

Figure 5.1: Lengths of consecutive ZI's (=FZ in this graph) including the frequency of these lengths for the automatic KNMI rainfall data implemented to test the UR filters

The performances of the filters are discussed below using confusion matrices and the data set of automatic KNMI gauges whereupon UR errors are simulated. The latter is from now on called data sets. It is important to keep in mind is that the performances of the filters are considered to be good the closer the percentages of ZI's (simulated UR errors) that are incorrectly flagged (FNR), are to zero.

5.1. CWS Filter

Before starting with the performances found for the CWS filter, it is essential to get a better understanding of the spatial distribution of the CWSs within the radius of the automatic KNMI gauges. Figure 5.2 shows the number of neighbouring CWSs for increasing radius. The number of CWSs are significantly higher and have a steeper slope for the gauges in Voorschoten and Rotterdam. This is a result of Rotterdam and Voorschoten being situated in a more densely populated area. However, the number of stations is overall lower for the station in Voorschoten compared to the gauge in Rotterdam. By comparing this with the fact that the North Sea and dunes are located West of the gauge, shown in Figure 5.4, it can be seen that the fraction of sea area inside the radius thus becomes larger with increasing radius. This results in less available area for CWSs. The same principle can also be seen at the station in Cabauw. The larger the radius becomes, the larger the fraction of radius area outside of the scope of this research. On the other hand, the gauge in Hoek van Holland is close to the sea and Europoort Harbour, but the slope of the number of CWSs is steeper compared to the one of Cabauw due to a higher coverage of a more densely populated area with increasing radius.

Figure 5.3 presents the number of neighbouring CWSs available per day within 8-kilometre radius from the automatic KNMI gauges. The number of neighbouring CWSs strongly fluctuates over time for the stations in Rotterdam and Voorschoten, but an increasing trend can be seen for the gauges. For the gauge in Hoek van Holland, less than 3 CWSs are available between October 2015 and the beginning of December 2015. Additionally, the number of neighbouring CWSs within radius of the station in Hoek van Holland is lower compared to the station in Cabauw till the beginning of August 2016. This is despite the fact that the station in Hoek van Holland has more neighbouring CWSs based on locations. In other words, the temporal behaviour of the neighbouring CWSs can fluctuate significantly over time. The effect of this temporal behaviour in the availability of neighbouring CWSs, however, mainly gives uncertainties for the CWS filter tested on the data of the station in Hoek van Holland.

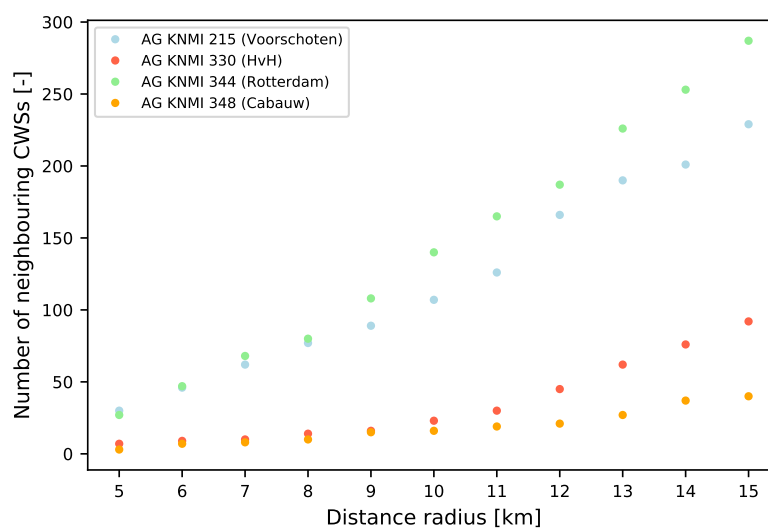


Figure 5.2: Number of neighbouring CWSs within different radius of the automatic KNMI gauges in the study area for a period from 10/1/2015 to 10/1/2016.

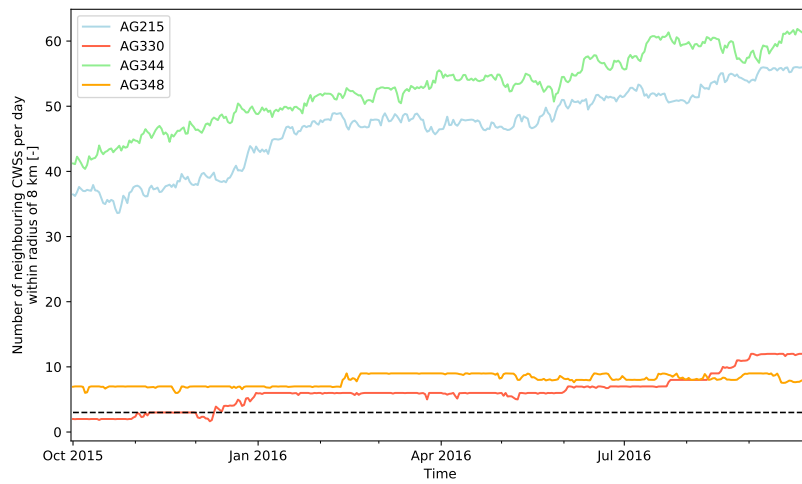


Figure 5.3: Number of neighbouring CWSs available per day over period between 10/1/2015 to 10/1/2016 within a radius of 8 kilometres from automatic KNMI gauges in the study region.

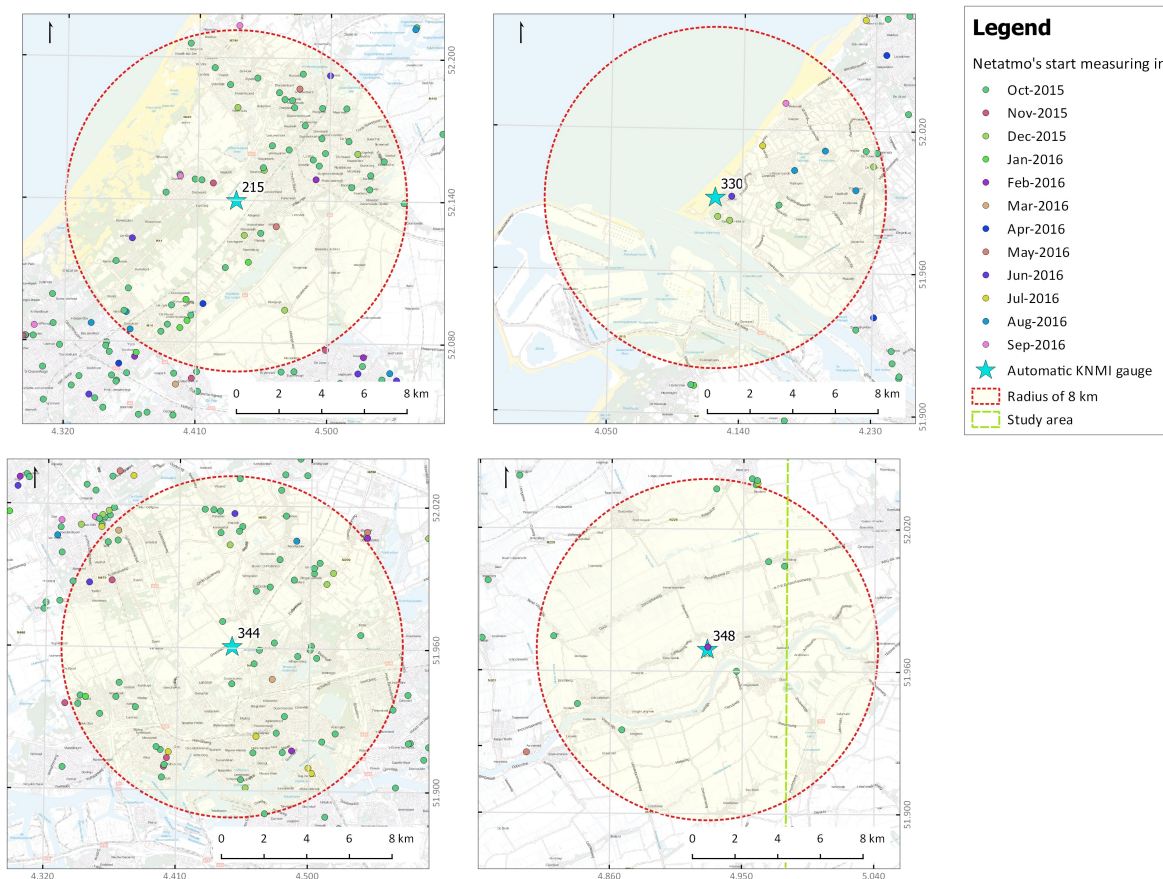


Figure 5.4: Spatial maps of the CWSs that measure during the first year of the research period within an 8-kilometre radius of the automatic KNMI gauges. Voorshoten (215), Hoek van Holland (330), Rotterdam (344) and Cabauw (348). Coloured symbols indicate month and year in which CWS started measuring.

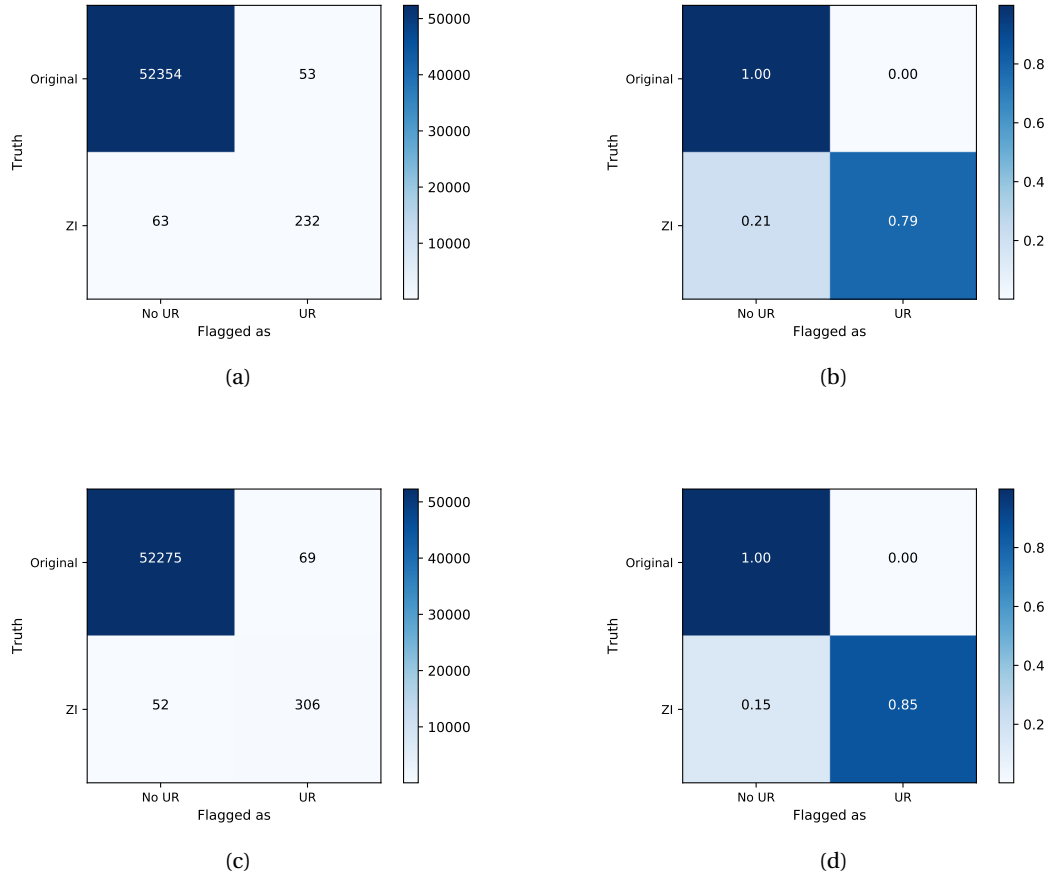


Figure 5.5: Confusion matrices for the automatic KNMI gauges AG215 (Voorschoten) and AG344 (Rotterdam) showing the performance of the CWS filter. Figures (a) and (c) show the counts, figures (b) and (d) give the rates after normalising the confusion matrices over the total number of original values or ZI's (sum of counts on horizontal axes in (a) and (c)) for respectively AG215 and AG344.

In order to get more understanding of these characteristics, as well as deriving the statistics of the performance rates of CWS filter, the following steps are executed 100 times: (i) Simulation of UR errors on the automatic KNMI gauge datasets within the study area to create datasets with ZI intervals; (ii) running the CWS filter on these datasets.

The 100 flagged timeseries created by the CWS filter are analysed using confusion matrices and its characteristics including the relative cost as explained in subsection 2.3.5.

Figure 5.6a shows that the median True Positive Rates (TPR) of the performances of the CWS filter are larger than 75% for all stations. This means that 50 out of 100 runs of each station, the CWS filter correctly detects and flags at least 75% of the ZI's. In the same figure also violin plots are shown in lightblue. These plots give the density distribution of the percentages. Figure 5.6a also shows that the TPR's of the station in Hoek van Holland (AG330) have the largest range, as well as the lowest median and mean in correctly flagging of ZI's. A large range in TPR's implies that the performance of CWS filter deviates more significantly per repetition indicating that the station of Hoek van Holland is more dependent on the timesteps in which the UR errors are artificially introduced.

The lower overall performance of the CWS filter for the station in Hoek van Holland can be attributed to the fact that parts of the dataset have only three neighbouring CWSs within radius of 8 kilometre. These periods are also interrupted by periods with less than three neighbouring CWSs. This causes the decision C in Figure 4.1 to become relevant and the flag of previous time step is taken over. By applying the flag of the previous time step an additional error in correctly flagging can be present. The occurrence and magnitude of the mistake in flagging depends on the locations of the ZI's in combination with the availability of neighbouring CWSs. The second lowest performance of the CWS filter is for the automatic KNMI gauge in Rotterdam

(AG344) as shown in Figure 5.6a. The TPR also has a larger range than the other two stations (AG215 and AG348).

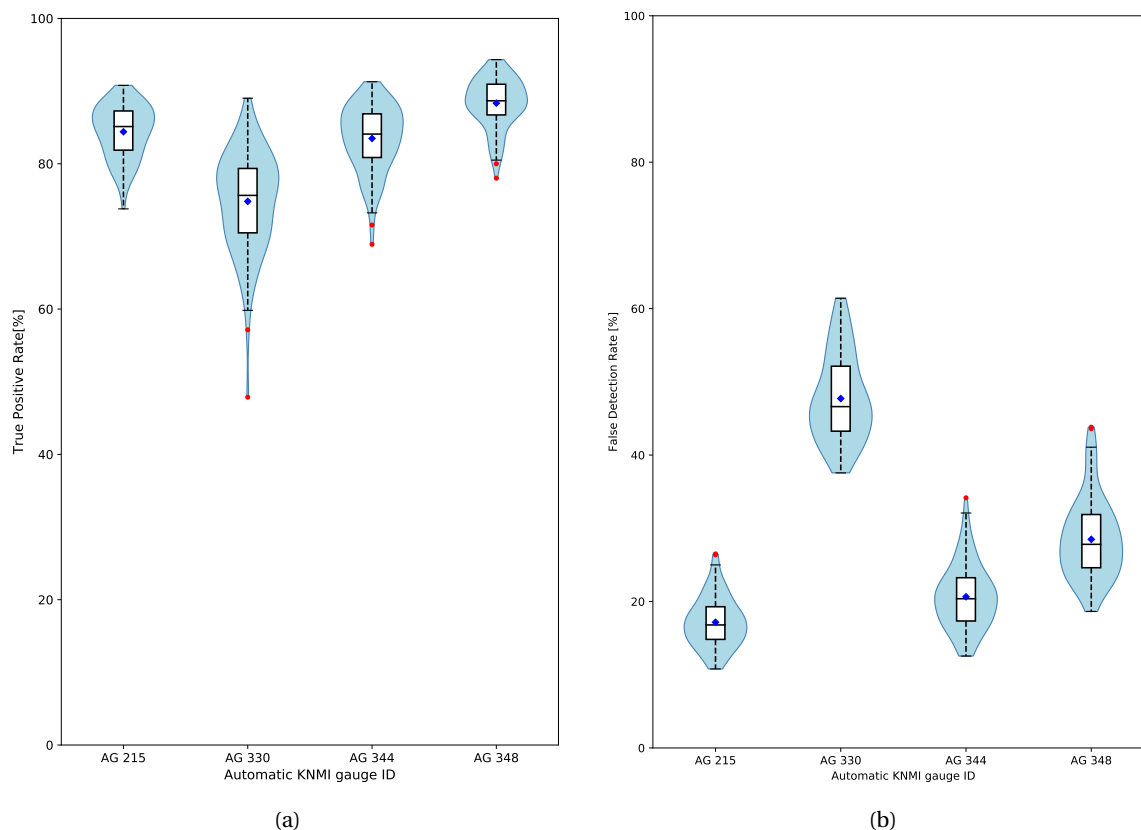


Figure 5.6: True Positive Rates (a) and False Detection Rates (b) after application of the CWS-based filter for 4 automatic KNMI weather stations. The boxplot in combination with a violin plot shows the statistics of the rates based on 100 simulated timeseries with 10% artificially introduced zeros. The density of the rates is given by violin, indicated by the lightblue shape behind the boxplot. The blue diamond gives the mean rate. The black line in the middle of the box is the median. The horizontal borders of the box indicate 25 to 75 percentile range.

In contrast to the station in Hoek van Holland, the automatic KNMI gauge in Rotterdam has at least 38 neighbouring CWSs within the radius of 8 kilometre to compute the median precipitation depth. This indirectly implies that there is another reason why the TPR's for the station in Rotterdam have a larger range for the TPR's than the stations in Voorschoten (AG215) and Cabauw (AG348). A reason could be that the neighbouring CWSs within the radius also have themselves UR errors in their rainfall timeseries. If this is the case, then it influences the applicability of the assumption that the majority of the neighbouring CWSs within the radius of the CWS of interest need to function properly. Despite this, Table 5.1 still shows that the mean TPR's are above 75% for all the stations. Additionally, the 25 to 75 percentile range is at most 9%. TPR's of 100 percent indicate the highest sensitivity possible, which creates a more strict filter that indirectly leads to more False Positives (FP), also called false alarms. The performance of the CWS filter is therefore accepted to be good regardless of not correctly flagging all the simulated UR errors.

False alarms occurs in the datasets at time steps where the automatic KNMI gauge measures zero and the neighbouring CWSs have a median higher than zero. The number of false alarms are 53, 207, 69 and 99 for gauges AG215, AG330, AG344 and AG348 respectively, between 1 October 2015 and 1 October 2016. These counts remain constant over the 100 computations. For the stations in Voorschoten, Hoek van Holland and Rotterdam, most (>70 %) of the false alarms are found in the first half year. From the constant FP's and varying TP's from the 100 computations, the False Detection Rates (FDR) is computed. Figure 5.6b gives the FDR for the four automatic KNMI gauges. The FDR of the CWS filter is on average 48% for the station in Hoek van

Holland, which has the highest FDR of all stations. This means that on average 48% of UR flags given by the CWS filter are incorrect (false alarms). The mean FDR's are 17%, 21% and 28% for AG215, AG344 and AG348 respectively.

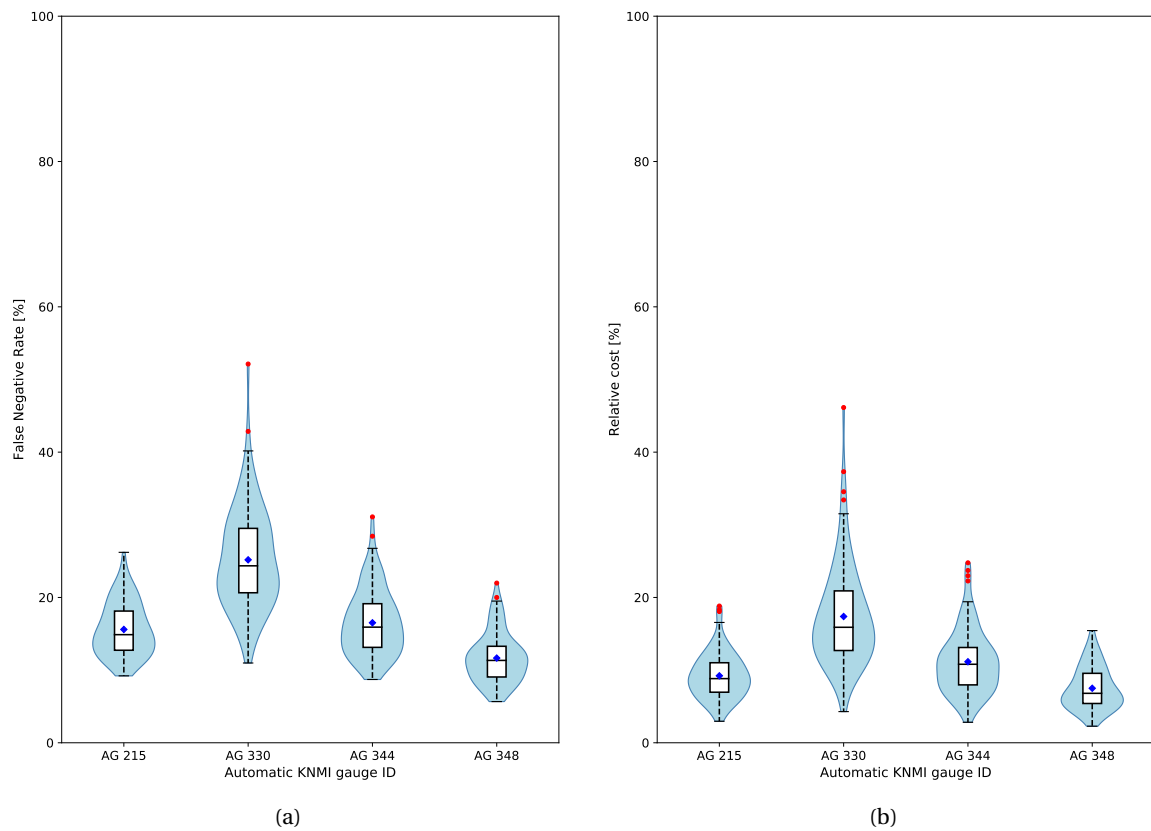


Figure 5.7: False Negative Rates (a) and Relative Cost (b) after application of the CWS-based filter for 4 automatic KNMI weather stations. The boxplot in combination with a violin plot shows the statistics of the rates based on 100 simulated timeseries with 10% artificially introduced zeros. The density of the rates is given by violin, indicated by the lightblue shape behind the boxplot. The blue diamond gives the mean rate. The black line in the middle of the box is the median. The horizontal borders of the box indicate 25 to 75 percentile range.

The False Negative Rate (FNR), which is the opposite of the TPR, is together with relative cost given in Figure 5.7. On average 16%, 25%, 17% and 12% of the ZI's are not detected and flagged as UR by the CWS filter for the stations in Voorschoten, Hoek van Holland, Rotterdam and Cabauw respectively. The relative costs of all stations (Figure 5.7b and Table 5.1) are lower compared to FNR's (Figure 5.7a and Table 5.1) meaning that the number of ZI's that are incorrectly flagged mainly occur during lower rainfall values. This confirms that the CWS filter performs better during the occurrence of higher rainfall intensities.

Table 5.1: Mean, 25th percentile (p25) and 75th percentile (p75) of True Positive Rate (TPR), False Negative Rate (FNR) and Relative Cost found for the CWS filter.

ID	TPR			FNR			Relative Cost		
	mean	p25	p75	mean	p25	p75	mean	p25	p75
AG215	84%	82%	87%	16%	13%	18%	9%	7%	11%
AG330	75%	70%	79%	25%	21%	30%	17%	13%	21%
AG344	83%	81%	87%	17%	13%	19%	11%	8%	13%
AG348	88%	87%	91%	12%	9%	13%	8%	5%	10%

5.2. Radar Filter

In this section, the performance of the proposed radar filter is discussed. A limitation of the radar filter is that rain gauges and the overlying radar pixel do not always agree when it is raining at a certain timestep. An analysis about the disagreement between the automatic KNMI gauges and the overlying radar pixel on the occurrence of rainfall is therefore executed using confusion matrices. Remark that the threshold of 0.06 millimetre is included in this analysis, since the disagreements of radar estimating rainfall and the gauges measuring zero rainfall occur at lower rainfall amounts as shown in appendix D. Radar values above this threshold are seen as rainfall values, and values below or equal to this threshold are zero values. The result of this analysis are shown first followed by the performances of the radar filter.

Figures 5.8a to 5.8d give confusion matrices (normalized over the total number of rainfall or zero rainfall values of the overlying radar pixels), which correspond to the amount of agreement between the automatic KNMI gauges and the overlying radar pixel with respect to zero rainfall. The FNR's denote the rates in which rainfall are estimated by the overlying radar pixel and zero-rainfall is measured by the gauge. The FNR's, shown in the figures, are for the radar filter of great importance, since the radar filter considers these timesteps (False Negatives) in which this ($P_{rad} > 0.06 \text{ mm} \ \& \ P_{ag} = 0$) occurs on forehand as UR errors. These timesteps are thus flagged as UR without being an UR error. With the assumption that the raw datasets of the automatic KNMI gauges do not contain UR errors as given in subsection 2.3.4.

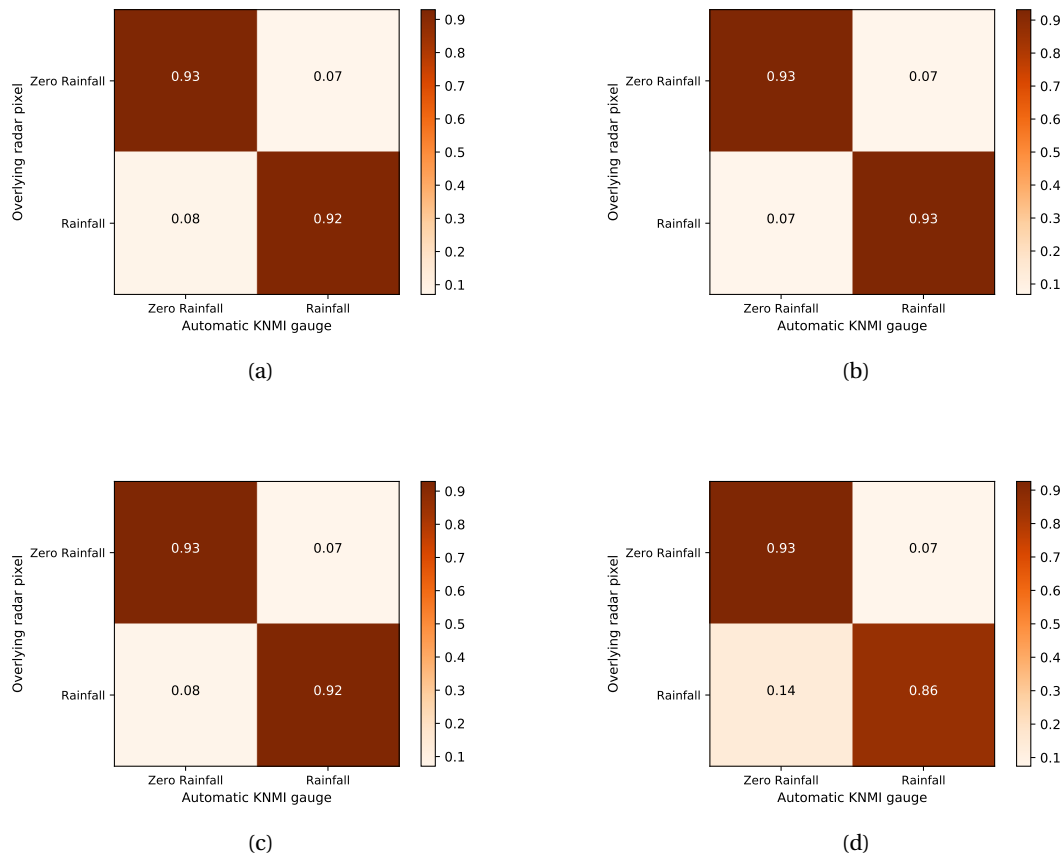


Figure 5.8: Confusion matrices for the degree of agreement between the automatic KNMI gauges and the overlying radar pixels with respect to zero rainfall. (a), (b), (c) and (d) are respectively the stations in Voorschoten (AG215), Hoek van Holland (AG330), Rotterdam (AG344) and Cabauw (AG348). The confusion matrices are normalised over the total number of rainfall values or zero rainfall values of overlying radar pixel in the period between 10/1/2015 to 10/1/2016.

These figures show that the automatic KNMI gauges in Voorschoten, Hoek van Holland and Rotterdam have similar FNR's, namely 8%, 7% and 8%. However, the gauge in Cabauw has a FNR of 14%. This means that

14% of the time intervals no rain is measured by the gauge, when the overlying radar pixel does show rain. The occurrence of these false negatives for the gauge of Cabauw is also shown in the hexbin plot (scatter density plot) in Figure 5.9 because the figure gives a horizontal line of hexbins on the zero rainfall depth of the gauge. The same problem is also encountered by de Vos et al. (2017). The reason for this phenomenon are radar artefacts. The radar measures most probably the reflections of the mast in Cabauw. For this reason the FNR found for the station in Cabauw is neglected. In short, 8% of the intervals of gauge data is automatically flagged by the radar filter as UR without being an UR error. This 8% is prior to the application of radar filter already a false alarm (False Positive). For the application of the radar filter on the CWSs, this means that on average 8% of the flags that are given by the filter are not UR errors, but timesteps in which the overlying radar pixel and the CWS did not agree.

The FDR gives in this analysis the rate in which rainfall is measured by the gauge, but zero rainfall is estimated by the overlying radar pixel. The FDR's found are 54%, 55%, 54% and 57% for the gauges in Voorschoten, Hoek van Holland, Rotterdam and Cabauw respectively. This means that on average 55% of the timesteps that the gauge measures rainfall, no rain is estimated by the overlying radar pixel. UR errors, which occur on those intervals, can therefore not be identified by the radar filter. In other words, there are some drawbacks for the application of the radar filter due to the significant amount of disagreement about the occurrence of rainfall between gauges and the overlying radar pixels.

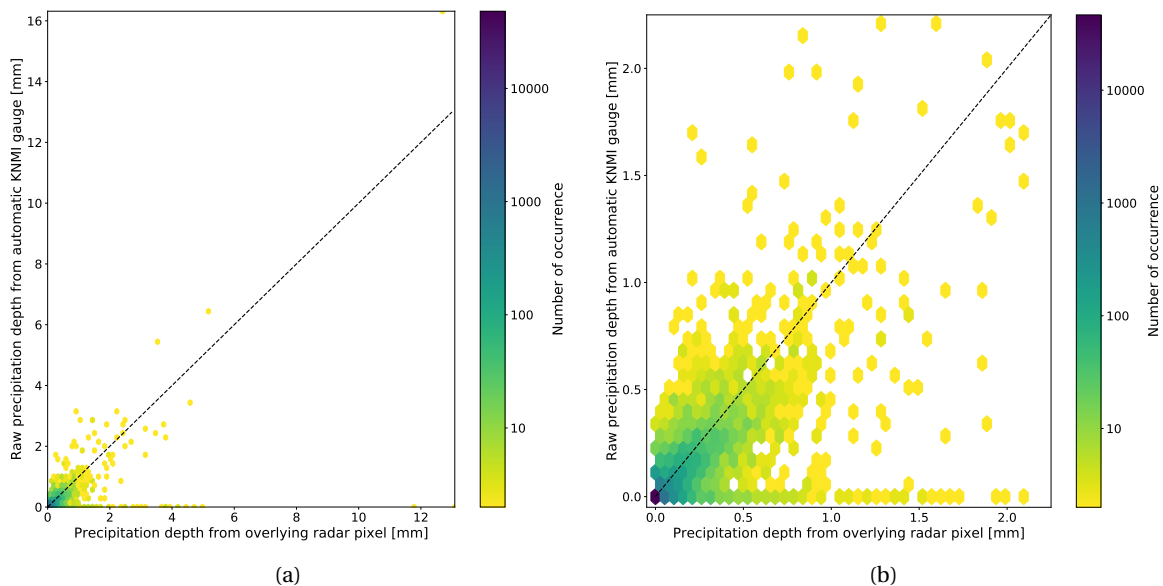


Figure 5.9: Hexbin plots (scatter density plots) of precipitation depth of the overlying radar pixel on x-axis and precipitation depth of the automatic KNMI gauge at Cabauw (AG348) on y-axis. (a) all values, (b) values in 0 to 2.2 mm range.

Figure 5.10a gives that the median TPR's of the gauges found for the radar filter are larger than 76%. This means that 50 out of 100 computations of each station detects and correctly flags at least 76% of the ZI's. The same figure gives that the TPR's of the stations in Voorschoten and Hoek van Holland both have a difference of 32% between minimum and maximum TPR, whereas the other two stations only differ 27% and 23% for Rotterdam and Cabauw respectively. The difference in minimum and maximum TPR is thus larger for the gauges close to North Sea. The larger difference in TPR's implies that radar filter performs less consistent for those stations. Between 74% to 78% of ZI's are on average correctly flagged by the radar filter, as given in Table 5.2. Despite of not having a TPR of 100%, the performance of the filter is considered to be adequate.

False alarms are accurate measurements (original values) that are incorrectly flagged as UR. So the FDR's give the percentages of incorrect UR flags. Figure 5.11b shows the FDR percentages for the radar filter. The differences between minimum and maximum FDR of the four stations are between 26% and 29% for the 100

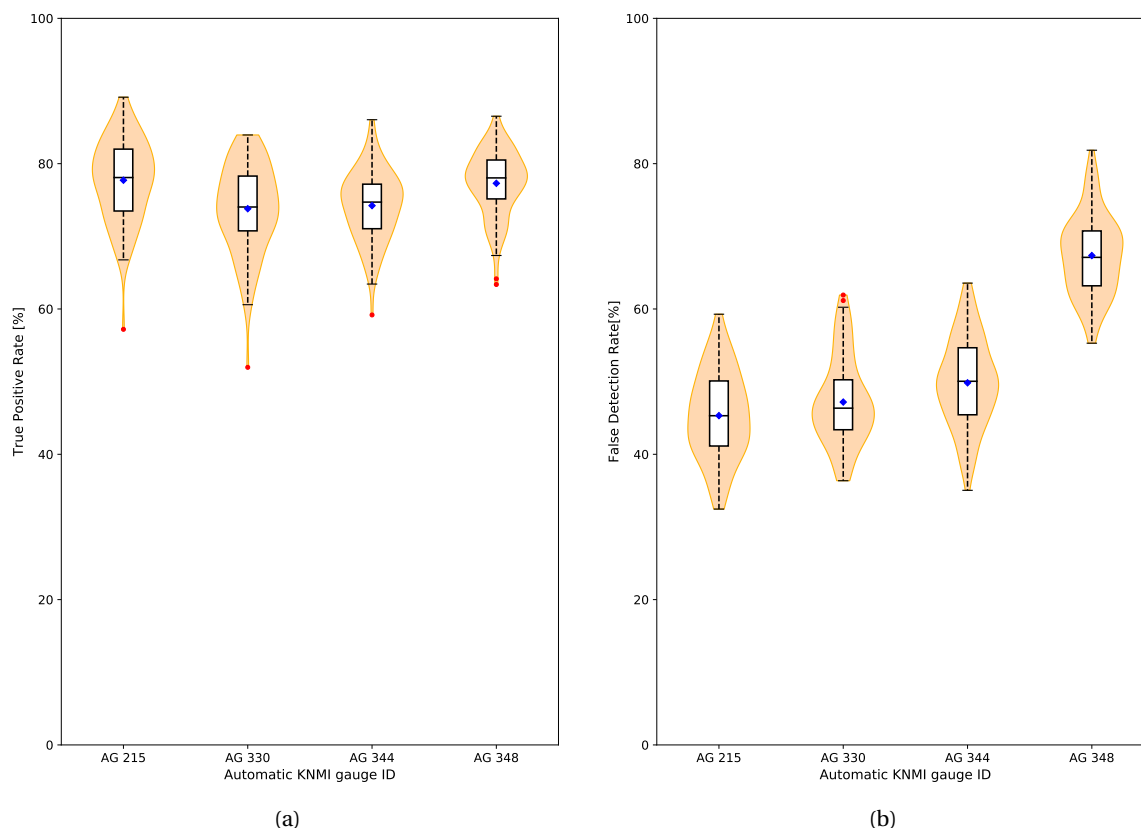


Figure 5.10: True Positive Rates (a) and False Detection Rates (b) after application of the radar-based filter for 4 automatic KNMI weather stations. The boxplot in combination with a violin plot shows the statistics of the rates based on 100 simulated timeseries including 10% artificially introduced zeros. The density of the rates is given by violin, indicated by the lightblue shape behind the boxplot. The blue diamond gives the mean rate. The black line in the middle of the box is the median. The horizontal borders of the box indicate 25 to 75 percentile range.

computations. The differences between minimum and maximum FDR's are thus relatively uniform for all the automatic KNMI gauges. The mean FDR's are 45%, 47%, 50% and 67% for the stations in Voorschoten, Hoek van Holland, Rotterdam and Cabauw respectively. The higher mean FDR percentage of the station in Cabauw compared to others results from the occurrence of the radar artefacts. Due to this the mean FDR of the station in Cabauw is left out in the next result. On average 47% of the UR flags given by the radar filter are incorrect. Together almost half of the UR flags given by the filter are false alarms. This results partially from the disagreement between gauges and overlying radar pixel on the occurrence of zero rainfall.

Table 5.2: Mean, 25th percentile (p25) and 75th percentile (p75) of True Positive Rate (TPR), False Negative Rate (FNR) and Relative Cost found for the radar filter.

ID	TPR			FNR			Relative Cost		
	mean	p25	p75	mean	p25	p75	mean	p25	p75
AG215	78%	73%	82%	22%	18%	27%	9%	7%	11%
AG330	74%	71%	78%	26%	22%	29%	12%	9%	15%
AG344	74%	71%	77%	26%	23%	29%	11%	9%	12%
AG348	77%	75%	81%	23%	19%	25%	10%	8%	12%

Figure 5.11 shows the percentages found for FNR and relative cost. The mean FNR percentages are respectively 22%, 26%, 26% and 23% for the gauges in Voorschoten, Hoek van Holland, Rotterdam and Cabauw, as shown in Table 5.2. This table also gives the mean relative cost to range between 9% and 12%. These percent-

ages are significant lower than the mean FNR percentages. This means that the radar filter has difficulties with flagging of UR errors that appear during drizzle rainfall events.

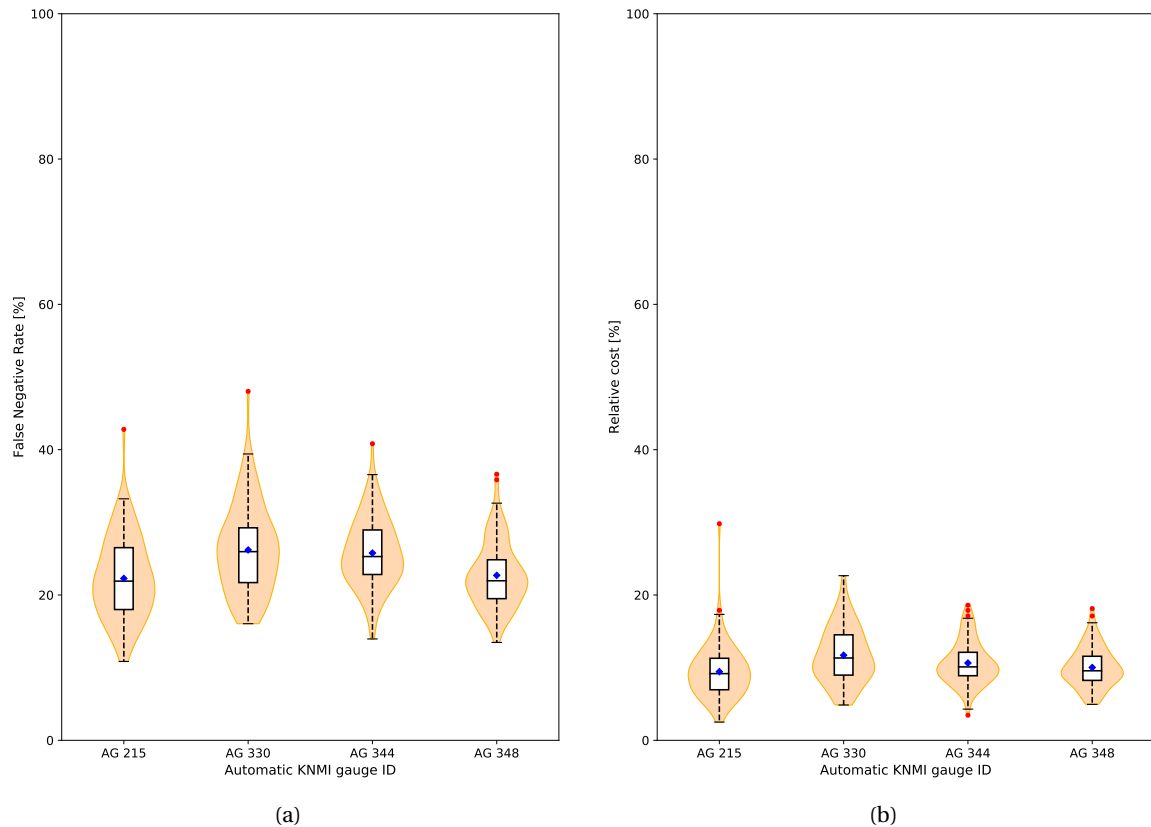


Figure 5.11: False Negative Rates (a) and Relative Cost (b) after application of the radar-based filter for 4 automatic KNMI weather stations. The boxplot in combination with a violin plot shows the statistics of the rates based on 100 simulated timeseries including 10% artificially introduced zeros. The density of the rates is given by violin, indicated by the lightblue shape behind the boxplot. The blue diamond gives the mean rate. The black line in the middle of the box is the median. The horizontal borders of the box indicate 25 to 75 percentile range.

5.3. Comparison between the CWS and Radar filter based on Performance

Both filters are considered to be adequate in detecting and flagging of UR errors, however, the performance of the radar filter is overall more stable due to continuous availability of required information to perform filter. For the CWS filter, it is found that the performance decreases as result of having less than 3 neighbouring stations as shown for the automatic KNMI gauge in Hoek van Holland. The presence of a physical barrier, such as the sea and harbour, in combination with a normative southwestern wind direction for the coastal area of the Netherlands and neighbouring CWSs that are all located in north east can also be a reason for the CWS filter to perform less well for the gauge in Hoek van Holland.

On the other hand has the radar filter a significant higher FDR than the CWS filter. The disagreement between the gauges and the overlying radar pixel on the occurrence of rainfall causes the filter to give more false alarms. The occurrence of radar artefacts in the data, as seen for the station in Cabauw, results in an additional increase in false detection rate. Another problem for the radar filter is that on average 7% of the intervals that the gauge measures rainfall, no rain is given by the overlying radar pixel. This causes that the radar filter to be unable to detect UR errors that occur in the 7%.

Both filters perform better for higher rainfall intensities as shown by the lower percentages of relative cost compared to the FNR. They both experience a decrease in performance during drizzle rainfall events.

6

Flags given to CWS Data by Filters

?? In this chapter the two filters, CWS-based and radar, are run for all CWS in the study area that measure over the period October 2016 to October 2017. The aim is to evaluate the flags given to CWSs by the filters. The results of the evaluation of flags given to CWSs are presented in Figure 6.1. The statistics corresponding to this figure are given in Table 6.1. On average 1.17% and 1.38% of all the zeros and nonzeros timesteps are UR flags given to CWSs by the CWS filter and radar filter respectively. For the CWS filter the minimum and maximum percentages of UR flags given to CWSs are 0% and 7.26%. The percentages of UR flags given to CWSs by the radar filter deviates between 0.02% and 6.42%. Comparing the results of the two filters gives that the UR flags given to CWSs by CWS filter have a larger range. The median percentages of UR flags given by both filters is slightly lower than the mean, namely 0.89% and 1.12% for respectively the CWS filter and radar filter. The lower median and the lower 25 and 75 percentile of UR flags given by the CWS filter compared to radar filter indicates that the number of flags are generally lower for the CWS filter. However, the larger difference between 25 and 75 percentile also implies that the number of UR flags vary more between the CWSs in the study area for the CWS filter.

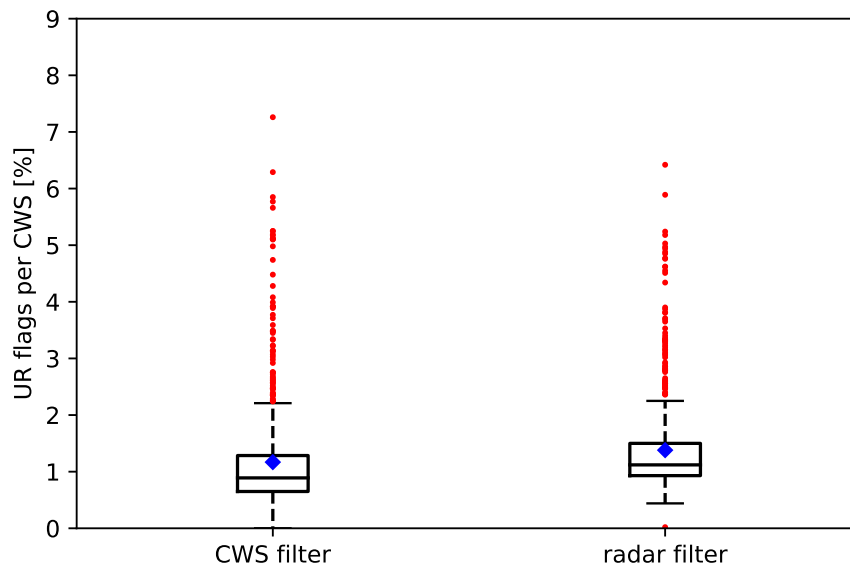


Figure 6.1: Percentage of UR flags given to 747 CWSs over total number of (non)zeros timesteps by each of the two filters separately for a period from October 2016 to October 2017. The blue diamond gives the mean percentage of UR flags per CWS. The black line in the middle of the box is the median. The horizontal borders of the box indicate 25 to 75 percentile range.

Table 6.1: Statistics of percentage of UR flags per CWSs given by the two filters separately. These statistics correspond to Figure 6.1. p25, p50, p75 are 25th, 50th and 75th percentile.

	CWS filter	Radar filter
<i>mean</i>	1.17%	1.38%
<i>std</i>	0.92%	0.80%
<i>min</i>	0.00%	0.02%
<i>p25</i>	0.65%	0.93%
<i>p50</i>	0.89%	1.12%
<i>p75</i>	1.29%	1.50%
<i>max</i>	7.26%	6.42%

A CWS¹ in the southern part of the study area has the highest percentage of UR flags for both filters as shown in maps 2 and 3 of Appendix E. The map with UR flags percentages given by the CWS filter shows that the neighbouring CWSs within a radius mostly have percentages of UR flags of at most 3%. The same map also shows that most (665) CWSs have percentages of UR flags between 0% and 2%, 64 CWSs got flagged as UR between 2% and 4% of their total number of zeros and nonzeros, and 18 CWSs have a percentage of UR flags larger than 4%. For the radar filter, shown in map 3 in appendix E, the number of CWSs for similar percentages are 658 (0.02% to 2%), 73 (2% to 4%) and 16 (4% to 6.42%). No clear spatial trend of UR flags given to CWSs is found for both filters.

Table 6.2: UR Flags given over all the intervals including the number of UR flags given for the same timesteps by both filters. Intervals are the total of zero and nonzero measurements given by the 747 CWSs that measure over the period from 1 October 2016 to 1 October 2017.

Variable	CWS filter	Radar filter
<i>Intervals [-]</i>	32348103	32348103
<i>UR flags [-]</i>	370047	441314
<i>UR flags given to the same intervals by both filters [-]</i>	177817	177817
<i>UR flags [%]</i>	1.14%	1.36%
<i>UR flags given to the same intervals by both filters [%]</i>	48.05%	40.29%

Over the period between 1 October 2016 and 1 October 2017, 1.14% and 1.36% of all zero and nonzero measurements from the 747 CWSs are flagged as UR by respectively CWS filter and radar filter, as given in Table 6.2. The radar filter gives 19% more UR flags compared to the CWS filter. This is most likely a result of the disagreement between rain gauge measurement and overlying radar pixel on the occurrence of rainfall. A percentage of 8% difference is found between automatic KNMI gauges and overlying radar pixel. Moreover, the filters are not always flagging the same timestep as UR. The CWS filter gives 48% of the UR flags to the same timesteps as the radar filter for the period between 1 October 2016 to 1 October 2017. This percentage is 40% for the radar filter. Overall the CWS filter is flagging more at the same time with the radar filter, then the other way around. This difference is expected because the radar filter gives 19% more flags due to the disagreement about the occurrence of rainfall. The same principle is also shown in Figure 6.2. This figure gives the percentage of UR flags given by both filter for the same time interval over the total number of UR flags awarded to CWSs.

Figures 6.3 and 6.4 show the percentages of UR flags over the total number of measurements per month for the CWS filter and the radar filter respectively. On average 1.09% of the monthly intervals are flagged as UR by the CWS filter. This fraction is 1.30% for the radar filter. The maximum monthly fraction of UR flags over all months occurred for the CWS filter in November 2016. This percentage is 2.05%. November 2016 also has the second highest monthly number of nonzero 10-min timesteps. September 2017 has the highest monthly fraction of UR flags for the radar filter. This again corresponds with a higher monthly number of nonzero timesteps. The minimum monthly fractions of UR flags occur in the same month for both filters, namely December 2016. For December 2016 the lowest monthly number of nonzero timesteps is measured by the

¹indicated by darkest blue for CWS filter and darkest red for radar filter

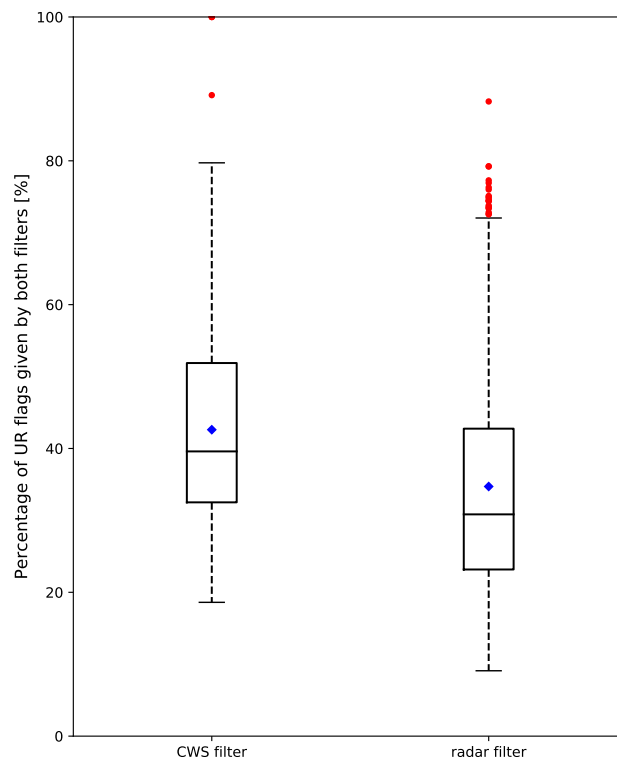


Figure 6.2: Number of UR flags given to CWSs for the same timesteps by both filters. The blue diamond gives the mean percentage of UR flags per CWS given for the same time interval by both filters. The black line in the middle of the box is the median. The horizontal borders of the box indicate 25 to 75 percentile range.

CWSs. Figure 6.5 shows the monthly fractions of UR flags given by the CWS filter and the radar filter overlap each other in October 2016 and November 2016. December 2016 is the only month in which the monthly fraction of UR flags is higher for the CWS filter compared to the radar filter. In other words, a higher monthly fraction of UR flags is generally found for months with higher number of nonzero timesteps for both filters. Besides the monthly fractions of UR flags given by radar filter are higher than the ones of the CWS filter, except December 2016.

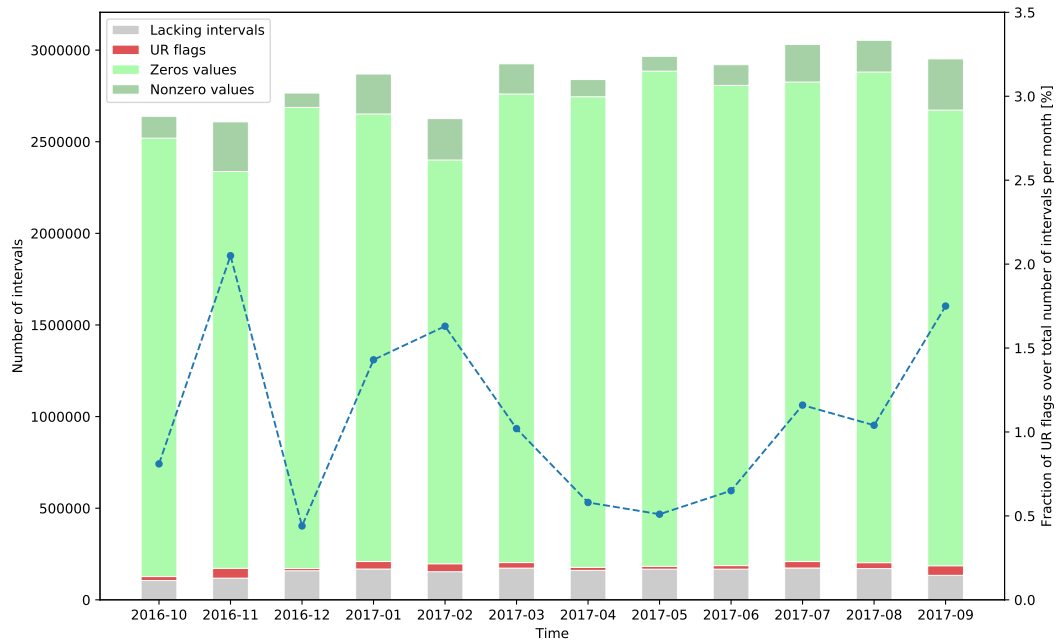


Figure 6.3: Number of time intervals per month with lacking intervals, UR timesteps flagged by CWS filter, zero and nonzero values occur. The blue line is the fraction of monthly UR timesteps including lacking intervals flagged by the CWS filter.

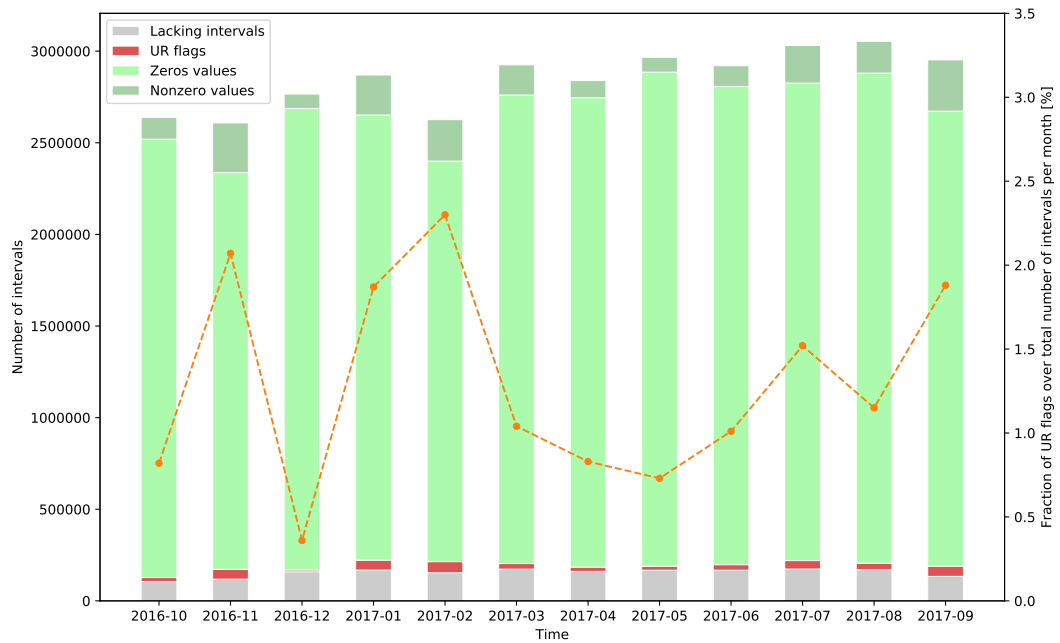


Figure 6.4: Number of time intervals per month with lacking intervals, UR timesteps flagged by radar filter, zero and nonzero values occur. The blue line is the fraction of monthly UR timesteps including lacking intervals flagged by the radar filter.

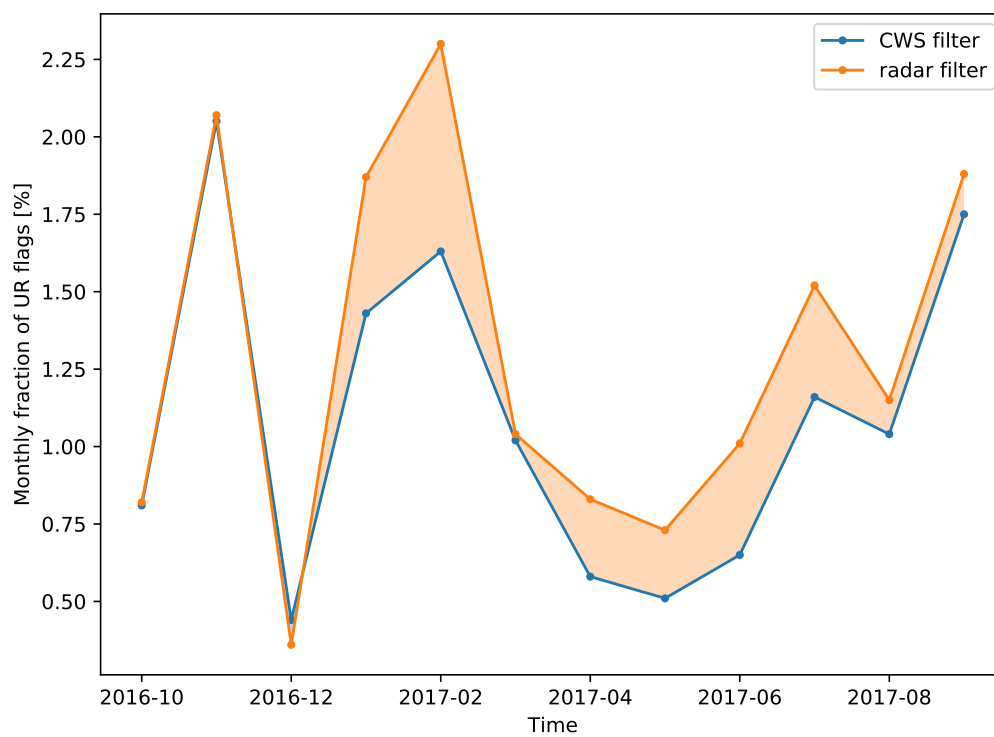


Figure 6.5: Monthly fraction of UR flags for both filters. An orange fill indicates that the monthly fraction of UR flags given by radar filter is larger. For blue fill, it is the other way around.

7

Discussion

In this chapter, the results obtained are discussed and put into wider perspective on how to use CWSs and the filters. Firstly, the data and methods are discussed. Secondly, an analysis of the calibration of tipping bucket volumes of CWSs within the study area is given. After this, the limitations of the two filters are discussed in more detail. Lastly, the limitation of CWS data and UR flags given to CWSs are outlined.

7.1. Assessment of Data and Methods used in this study

For this study, the Netatmo CWS data are retrieved from Netatmo weathermap using the Getpublicdata API and Getmeasure API provided by the manufacturer and a R-script provided by Lotte de Vos, both for a period of 1 October 2015 to 1 October 2017. The retrieval of the device ID's, module ID's, locations and the most recent measurement of 760 CWSs is done on October 23, 2017 using the Getpublicdata API. All the stations known of by Netatmo weathermap are returned independent of the period that they are measuring. For the retrieval of this data the area is divided into 100 tiles to ensure that all CWSs in the region are returned. The need of splitting of the study area into subareas is also stated by Meier et al. (2015). Because the retrieval of station ID's and locations happens in stages, duplicates in data did occur. Also stations outside of the sub areas are returned due to the 30-min cache system. The stations outside the sub-tiles and duplicates of stations are therefore identified and removed after retrieval with Getpublic API and before the Getmeasure API is applied to obtain the timeseries of CWSs.

As mentioned in section 3.1, seven out of the 760 CWSs only contain measurements before or after the selected period. Five of those stations are empty timeseries implying that no measurements are available for the desired period. These stations stopped measuring before 1 October 2015. The other 2 CWSs are stations in which the timeseries are totally NA-filled. These stations started with measuring after 1 October 2017.

An option in the Getpublic API to apply quality-control on the timeseries of CWSs is available. No explanation of this quality-control options is however provided. This in combination with the preference of raw data resulted in the decision not to use the quality-control option.

Netamo weathermap is not the only platform for CWS data. The Netatmo weathermap platform is chosen based on high temporal resolution of ~ 5 minutes, only having one type of rain gauge (a tipping bucket rain gauge of Netatmo brand), no web-platform processing, dense network in the Rotterdam-The Hague region and less data gaps compared to Wunderground. Wunderground has an even higher density of CWS network for this region because Netatmo CWSs are automatically linked from Netatmo weathermap to Wunderground and other CWS brands are uploaded by owners themselves. This higher density of Wunderground can be preferred in regions where the CWS network of Netatmo weathermap is to sparse. Disadvantages of Wunderground are the lower temporal resolution (approximately ~ 10 minutes) than Netatmo weathermap and the occurrence of conversion and rounding errors on platform. Besides the different types of rain gauges on Wunderground all give their own uncertainties as result of for example different measurement techniques. This is seen as a disadvantage for this research. For this reason, the retrieval of CWS data from Netatmo

weathermap is preferred above the CWS data from Wunderground.

Netatmo CWS are tipping bucket rain gauges with a default volume of 0.101 millimetre. The sampling of small-scale rainfall temporal variability of convective storm by tipping buckets has some drawbacks as stated by Habib et al. (2001). The performances of the tipping buckets are lower during drizzle rain and tipping buckets have difficulties with the determination of the start and end time of rainfall events. Tipping buckets of at most 0.1 millimetre should be used combined with sampling intervals with a magnitude of seconds in order to reduce sampling errors of tipping bucket rain gauge regarding small-scale rainfall temporal variability (Habib et al., 2001). For the Netatmo tipping buckets applied in this study this means that sampling errors mainly become larger as result of the sampling timescale of ~ 5 minutes.

The rainfall timeseries of Netatmo stations have an irregular time grid of 5 minutes, therefore it is required to convert these irregular timeseries to a regular grid. A temporal resolution of 10 minutes is selected for this study but any temporal resolution can be chosen. This conversion from irregular rainfall timeseries to regular rainfall timeseries is executed based on the assumption that constant rainfall intensities occur over the time interval. This is to a great extent true for small temporal resolutions. For higher temporal resolutions, this assumption is less justified due to small-scale rainfall variability. Application of this method at higher temporal resolutions may cause the small-scale rainfall variability to average out. This contradicts the reason to use a CWS network, that is to make use of a higher temporal and spatial resolution to capture variability of rainfall in urban areas in order to forecast UPF.

Knowledge about the occurrence of UR errors in CWS rainfall data is lacking. Zeros are therefore artificially introduced on 10% of rainfall sequences of automatic KNMI gauges in which the original rainfall amounts are higher than 0.05 millimetre (q_{ag}). The implementation of q_{ag} in this method results from a previous evaluation of the performance of the CWS filter. From this evaluation emerges that the CWS filter has trouble with identifying UR errors at low rainfall quantities. The motive of forecasting UPF made that flagging of UR errors at timesteps with lower rainfall quantities is not crucial. The implementation of this thresholds indirectly implies that the filters perform poorly during drizzle rain. As mentioned above, the performances of the tipping buckets are also poor during low rainfall intensities. The application of Netatmo tipping bucket rain gauges in order to measure drizzle rainfall in urban areas already shows some significant problems. For this reason it can be argued that Netatmo CWSs should not be applied to monitor drizzle rainfall. This makes the implementation of q_{ag} well-founded.

Confusion matrices are applied to evaluate the performances of the two filters. In these matrices a position is adopted to detect all UR errors in CWS rainfall data without having unreasonable give-and-takes. By doing so, strict filters are created in which incorrectly UR flags (false alarms) are more acceptable than missed UR errors. The higher the true positive rates are to 1, the higher the false detection rates. An optimal performance of approximately 75% is considered to be sufficient.

The radar filter uses the overlying radar pixel to identify UR flags. The gauge-adjusted climatological radar data sets are used for this. Application of the radar filter in near-real time disables the use of this radar product. Instead, the unadjusted raw radar product should be used in these situations. The use of the unadjusted raw radar product is expected to give a larger disagreement between radar and gauges on the occurrence of rainfall. The performances of the radar filter in flagging are therefore expected to change. For this reason the performance of the radar filter should also be investigated with the use of the unadjusted raw radar product.

7.2. Calibration of Tipping Bucket Volumes from CWSs

This study found that 73 owners calibrated the tipping bucket volumes. Most (67) owners got higher tipping bucket volumes than the default tipping bucket volume of 0.101 millimetre resulting in overestimation of rainfall depth compared to professional gauges as shown in Figures 3.6 and 3.7. The over- and underestimation in these figures can be an indication that the actual tipping bucket is not exactly 0.101 millimetre. CWS ID461 gives an overestimation of rainfall compared to the automatic KNMI gauge in Voorschoten and has a tipping bucket volume of 0.236 millimetre. This raises the question whether the overestimation is due to miscalibration.

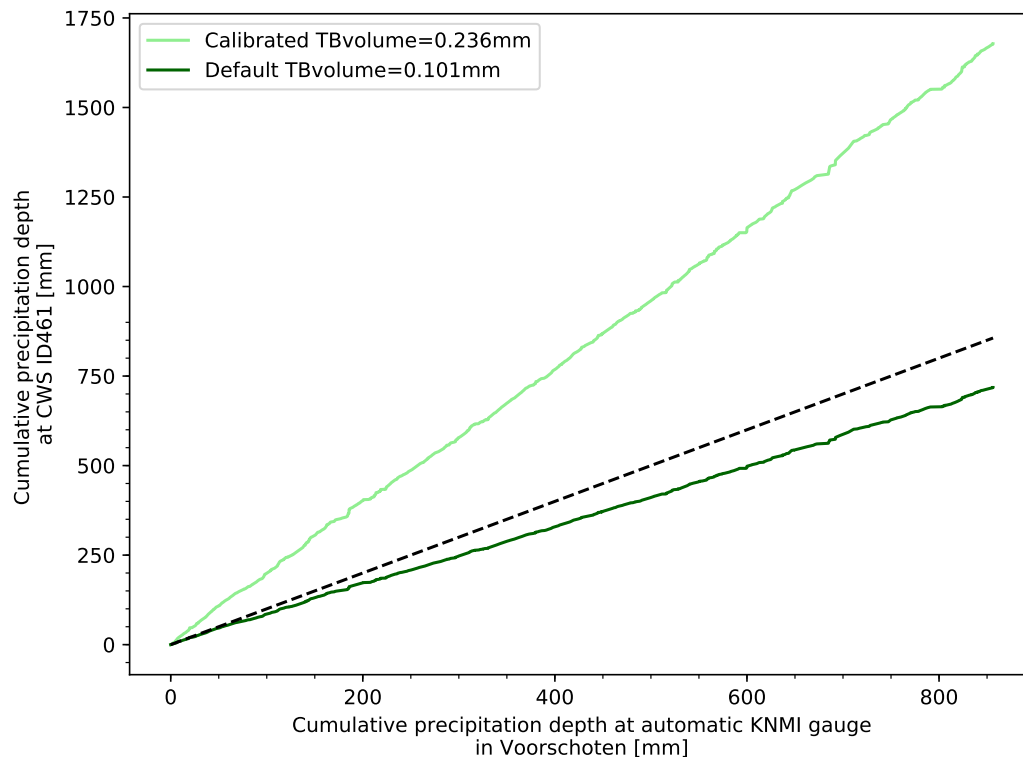


Figure 7.1: Double mass curve between the automatic KNMI gauge in Voorschoten on x-axis and CWS ID461 on y-axis. CWS ID461 is at a distance of 4.73 kilometre from the professional gauge. A tipping bucket (TB) volume of 0.236 was found by the owner through calibration. The DM-curve corresponding to this TB volume is shown in light green. If the owner did not decide to calibrate the tipping bucket volume, then the darkgreen line would have been found.

Figure 7.1 shows two DM-curves: a lightgreen line for a calibrated tipping bucket volume of 0.236 millimetre and a darkgreen line for a default tipping bucket volume. On the x-axis it shows the cumulative precipitation depth at automatic KNMI gauge in Voorschoten and on the y-axis the cumulative rainfall depths of CWS ID461. This figure shows that the DM-curve for the default tipping bucket volume is closer to 1:1 ratio (black dashed line) than the calibrated tipping bucket volume. This indicates that CWS ID461 is miscalibrated. Miscalibration can result from the owner pouring water into the gauge too quickly during calibration leading to higher tipping bucket volumes after calibration. This problem should be investigated in future research.

In Figure 7.1 is also shown that the DM-curve for the default tipping bucket volume is underestimated compared to the professional gauge. This gives the question whether the default tipping bucket volumes underestimate rainfall to begin with. Table 7.1 gives the tipping bucket required to derive the exact yearly sum of the professional gauges at close distance for four CWSs. The table shows that the tipping buckets required a range between 0.102 millimetre for the CWS ID188 and 0.120 millimetre for the CWS ID 461 in which the CWS ID188 is the closest to the automatic KNMI gauge and CWS ID461 the furthest. Additionally, this the original tipping buckets with default tipping bucket volumes give significantly lower tipping bucket volumes to estimate the exact yearly sum of professional gauge than the ones in which the owner calibrated the device. All this together can imply default tipping bucket volumes slightly underestimate rainfall. Calibration, if done correctly, can improve the rainfall measurements.

7.3. Limitations of CWS Filter

For the CWS filter the assumption is made that the majority of neighbouring CWSs should function properly in order to identify UR errors in CWS of interest. This means that the majority of the neighbouring CWSs

Table 7.1: Tipping bucket volumes required (TB_{new} volume) to derive the exact yearly sum of the professional gauge at close distance for four CWSs. The distance to AG including the ID of the AG, as well as, the original tipping bucket volume (TB_{original} volume) are given here.

CWS	Distance to AG	TB_{original} volume	TB_{required} volume
ID188	1.27 km to AG344	0.101 mm	0.102 mm
ID454	3.08 km to AG215	0.12 mm	0.110 mm
ID350	4.00 km to AG344	0.101 mm	0.103 mm
ID461	4.73 km to AG215	0.236 mm	0.120 mm

should not have UR errors in their data. 75% of the CWSs are flagged at most 1.29% of the zero and nonzero timesteps as UR flags. This implies that the majority of CWSs function relatively good. From this 75th percentile of UR flags it cannot be determined if these flags occurred at the same timesteps for all CWSs. The latter should be investigated in order to accept or reject this assumption made to check whether the computation of median precipitation depth of neighbouring CWSs is substantiated.

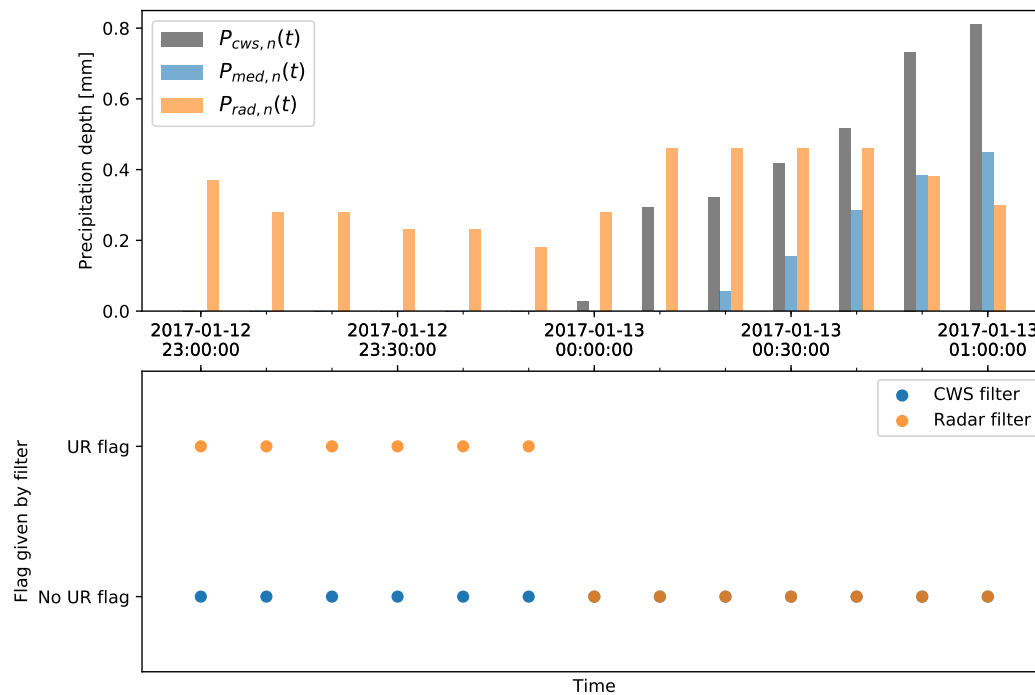


Figure 7.2: An example of how the CWS filter and the radar filter flag during a snow event at CWS ID188. $P_{\text{cws},188}(t)$ is the precipitation depth measures at the CWS of interest. $P_{\text{med},188}(t)$ is the median precipitation depth of neighbouring CWSs used for the CWS filter. $P_{\text{rad},188}(t)$ is the precipitation depth in the overlying radar pixel.

A snow event can cause all CWSs to be blocked causing the CWS filter not to work while the radar filter flags these timesteps of the snow event as UR. An example of this is given in Figure 7.2. The occurrence of a 2-hour snow event is found in the hourly data of the automatic KNMI gauges in Voorschoten and Rotterdam from 12 January 2017 at 11 pm to 13 January 2017 at 1 am. The rainfall depths, median precipitation depth of neighbouring CWSs, the precipitation depth of overlying radar pixels as well as the flags given by the two filters are selected at CWS ID188. This figure shows that both CWS ID188 and neighbouring CWS do not measure precipitation during the first hour, whereas the radar does measure nonzero values as expected. This results

in the CWS filter giving no UR flags and the radar filter flagging the first hour as UR. In other words, the CWS filter does not work during a snow event. Remark that the CWS filter and the radar filter are created to identify UR and not for snowfall. But the fact that the snowfall is flagged as an UR timestep by radar filters implies the need of an additional check to identify snowfall.

Another limitation of the CWS filter is that some CWSs have less than three neighbouring CWSs within the 8 kilometre radius. Step C in Figure 4.1 is implemented because the computation of the median precipitation depth requires at least three neighbouring CWSs. A drawback of this step is that the filter may start later with flagging than the moment the CWS starts with measuring. This results in shorter processed datasets. The delay in flagging is determined by the moment it starts raining or that lacking intervals occur; (ii) or the moment that at least three neighbouring CWSs are available within the 8 kilometre radius. The use of a fixed 8 kilometre radius also shows that CWSs cannot be flagged at all due to continuously having less than three neighbouring CWSs. Figure 7.3 shows that 4 CWSs cannot be flagged, 15 CWSs have shorter processed datasets and the other 576 CWSs are normally flagged by the CWS filter for a period between 1 October 2015 and 1 October 2016.

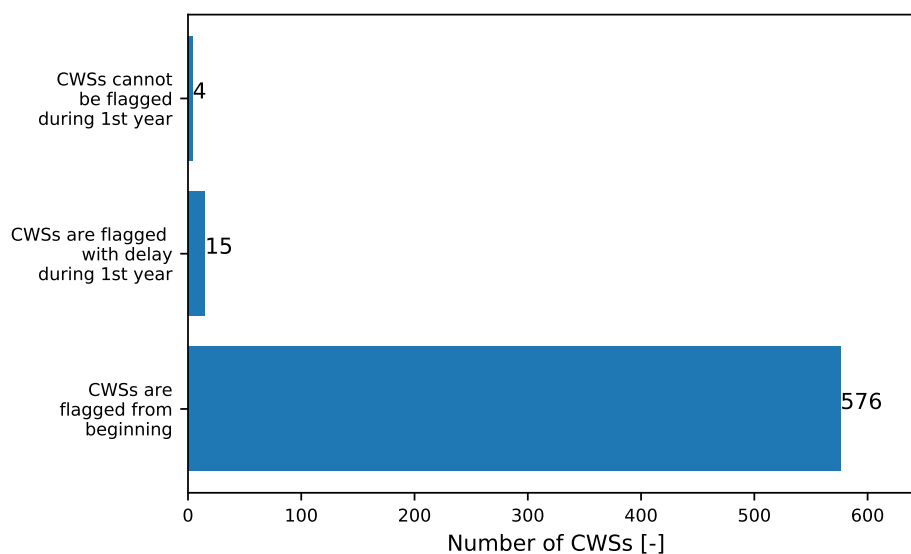


Figure 7.3: Effect of 8 kilometre in relation with the number of neighbouring CWSs for the CWS filter in a period from 1 October 2015 till 1 October 2016 (first year).

The rainfall data with ZI's from the automatic KNMI gauge in Cabauw is used to evaluate the performance of the CWS filter. Figure 5.4 shows that this station is located close to Eastern border of the study area (indicated by the green dashed line). A fraction of the 8 kilometre radius is therefore outside the study area causing that potential neighbouring CWSs can be missed during data retrieval. Retrieval of all CWSs within the 8 kilometre radius around the CWS of interest is advised if the CWS filter is applied. By doing so, the best possible representation of neighbouring CWSs within the fixed radius can be achieved.

7.4. Limitations of Radar Filter

The radar filter generally has a higher number of incorrect UR flag false alarms and a lower overall performance compared to the CWS filter (FNR is less close to 1), this is due to the limitations of the radar filter. The differences in measurement techniques and space-time resolutions can cause the radar and gauges to disagree about the occurrence of rainfall. Also the alternation between rain and dry periods differs between the two different types of devices. In this section the effects of these limitations on the radar filter are discussed.

Radar and gauges are different measurement techniques. The CWSs used in this research are tipping bucket rain gauges. Tipping buckets measure the number of tips over time. The number of tips is converted to rain-

fall accumulation by multiplying it with the tipping bucket volume. "Radar transmits pulses of microwaves signals and measure the power of signal reflected back by raindrops, snowflakes and hail stones. Rainfall is estimated using the reflectivity" according to Cristiano et al. (2017). It integrates the reflectivity over volume. This volume increases with increasing distance from the radar giving that rainfall data is smoothed out over space. This means that if the radar filter is applied on a region far away from the radar, the possibility exist that the overlying radar pixel cannot flag UR errors. The smoothing over space caused that the precipitation depth in the overlying radar pixel is below the threshold q_{rad} . Additionally, radar can also overshoot a rainfall event at large distance from the radar due to the curvature of the Earth. This results in a radar estimating no rainfall for a certain pixel, however, it is raining at CWSs in the pixel causing that possible UR errors are missed by the radar filter. Finally, radar measures snowfall where the CWSs only measure rain. The radar filter can flag the timesteps in which snowfall occurs as UR errors, however, they are not UR.

The gauge-adjusted climatological radar product used in this research to evaluate the performances of the radar filter, has a 5-min temporal resolution and $1 \times 1 \text{ km}^2$ grid. The CWSs have a temporal resolution of approximately 5 minutes and are point measurement with a collecting funnel of 13 centimetres. The temporal resolution of both devices are approximately equal. However, the spatial scales differ significantly between radar and gauges. The larger space resolution makes that the radars has difficulties with capturing highly localized rainfall events (Gires et al., 2014). Suppose that a very localized rainfall event causes rain at a completely blocked gauge and that the radar is not able to detect this small-scale rain event. This UR error cannot be identified by the radar filter. Also the position of the gauge can cause a disagreement between radar and gauges on the occurrence of rainfall. For example a storm has not yet arrived at the gauge but the precipitation depth estimated by the radar is already higher than q_{rad} . The timestep in which this occurs will be flagged as UR. In other words, the disagreements on the occurrence of rainfall between radar and gauges as result of a difference in space, can negatively affect the performance of the radar filter.

The difference in spatial scales between radar and gauges has also an effect on the differences in rainfall intermittency between radar and gauges. According to Schleiss et al. (2011) "small areas are more likely to be dry than large ones" implying that the gauge is more likely to give a dry period than radar. The difference in rainfall intermittency between the two can therefore affect the performances of the radar filter. For example when an overlying radar pixel has a rainfall depth higher than q_{rad} and the CWS measures zero rainfall due to rainfall intermittency, then this timestep is flagged as UR without being an actual UR error. The differences in rainfall intermittency is also magnified due to the use of tipping buckets. Their binary behaviour can cause dry periods within a rainfall event as result of not tipping of the bucket during a certain interval (Molini et al., 2001). The lacking of tips mainly occurs during low rainfall intensities. Consequently, the intermittent behaviour of the CWSs is expected to be more significant for these values, thereby it emphasises the need of the threshold q_{rad} .

7.5. Limitations of CWS Data and UR Flags given to CWSs

All the errors expected on forehand are found in the CWS data in this study area. In addition, CWSs are found to function partial normal and partial cumulative in which the cumulative parts are reset once in a while. The CWSs with cumulative datasets are also flagged by the two filters in order to identify UR errors. UR errors cannot be identified by the two filters since both filters think that it is raining at those timesteps. Besides the fact that no algorithm is developed to automatically detect these CWSs with cumulative parts. They are discovered by manual investigation of CWSs. The exact number of CWSs in which this happens is therefore unknown. Future research should develop a method to automatically identify these datasets followed by exclusion of the CWSs or conversion of cumulative parts to normal values.

UR flags are given to CWS data by both filters for the period from 1 October 2016 to 1 October 2017. Less than 50% of UR flags are given by the CWS filter and the radar filter for the same interval. Suppose that only the CWS filter is selected to detect UR errors in the rainfall data of CWSs, then a large percentage of UR flags are missed according to the radar filter. This makes the use of the filters separately less appealing for the detection of all UR errors. A solution can be to apply both filters on the datasets in order to identify all timesteps with UR errors. A problem of this solution is that the application of the CWS filter is not always possible. Another effect due to the application of both filters is more incorrect UR flags. For this reason, the why and when of UR flags given by only one of the two filters should be investigated in order to obtain more understanding

about what is actually happening.

A remark has to be made concerning the percentages of UR flags given to CWSs by the two filters. The total number of zero and nonzero timesteps used to obtain these percentages differs between CWSs. UR flags given to CWSs with lower number of zeros and nonzeros compared to other CWSs result in higher percentages. This should be taken into account when using those percentages.

The number of UR flags given to the rainfall data with ZI of the automatic KNMI gauges is on average higher for the radar filter, except the station in Hoek van Holland. For this station the number of UR flags given by the CWS filter is higher. The same station also has a lower performance for the CWS filter compared to the radar filter due to issues with availability of neighbouring CWSs. These CWSs have almost no spreading of over total area within an 8 kilometre radius due to presence of a physical boundary. All this together could indicate that the radar filter is preferred above the CWS filter for situations in which the number of UR flags given by the CWS filter is higher than the radar filter.

The percentages of UR flags given to CWSs are generally higher for the radar filter. This phenomena is shown in Figure 7.4. This figure shows that more counts are found above the black line implying that the percentage of UR flags given to CWSs are higher for the radar filter. However, a larger difference between minimum and maximum is found for the CWS filter. At the higher percentages of UR flags, these percentage are mainly beneath black line meaning that the percentage of UR flags given by CWS filter is higher for these CWSs. Linking this to the fact that the number of UR flags given by the CWS filter is higher than the radar in situations in which the CWS filter has issues, could indicate that these CWSs experience the same problems.

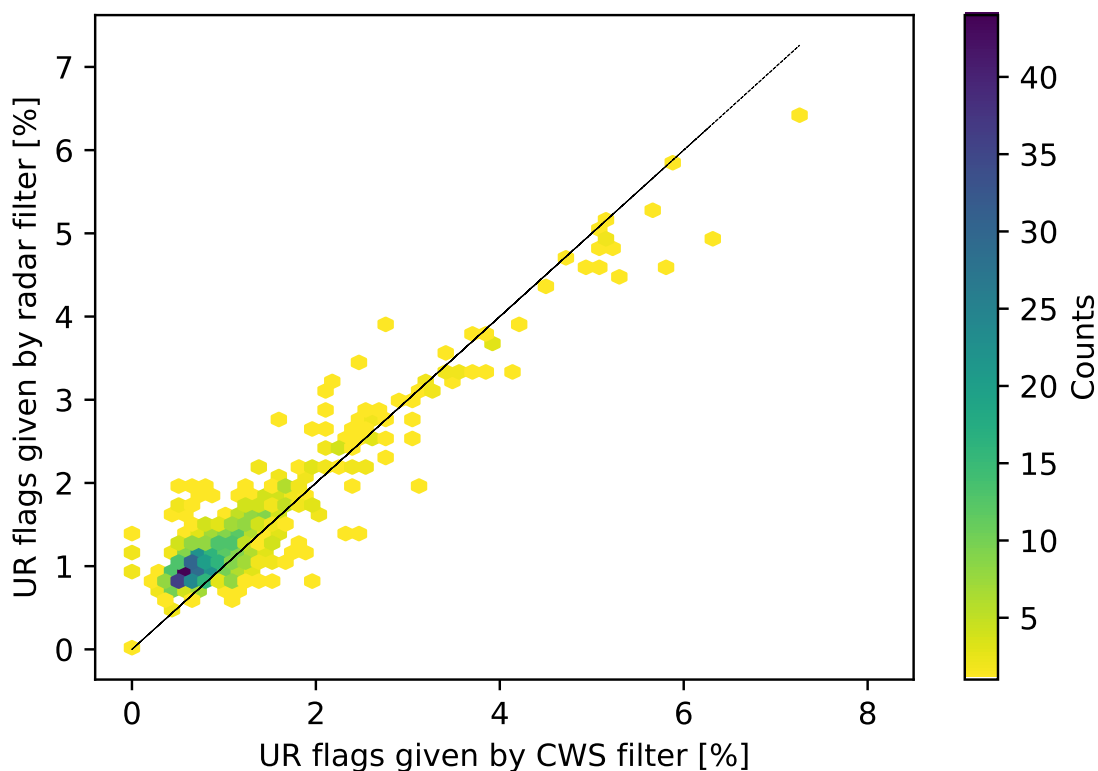
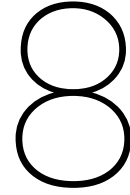


Figure 7.4: Hexbin plots (scatter density plots) of percentage of UR flags given to CWS by CWS filter on x-axis and percentage of UR flags given to CWSs by radar filter on y-axis.



Conclusion and Recommendation

Professional rain gauge networks fall short to monitor urban rainfall in order to forecast urban pluvial flooding. Citizen rain gauge data have the potential to fill this gap, however the quality of CWSs is significantly lower due to errors. These errors are mainly present in CWS rainfall data through three types of erroneous behaviours, namely lacking intervals, false rainfall intensity values and UR errors. The objective of this research was to develop and test different automated quality-control filters flagging for undetected rainfall errors in rain gauge data of CWS network by answering the main research question. To recapitulate, the main research question is:

How can undetected rainfall errors in citizen rain gauge data be identified by automated quality-control filters?

UR errors in citizen rain gauge data can be automatically identified by the CWS filter using neighbouring CWSs or by the radar filter using precipitation depth of overlying radar pixel. It is found that both filters adequately flag undetected rainfall errors in citizen rain gauge data. The CWS filter performs slightly better than the radar filter, however both filters have their benefits and drawbacks. A shared drawback is the lower performance during low rainfall intensities, therefore the application of filters during drizzle events should be minimized. In case of sparse CWS network density or presence of physical boundary, the radar filter is preferred above the CWS filter. Due to its higher performances in these situations, the CWS filter should be applied for regions with high CWS density and adequate spreading of neighbouring CWSs around the CWSs of interest. Overall the radar filter is a more consistent choice as a quality-control mechanism due to the continuous availability of data for filtering. In contrast, users should take into account that the radar filter gives a higher number of incorrect flags compared to the CWS filter. All this together concludes that both filters can automatically identify UR errors in citizen rain gauge data, however the choice of a filter should be a consideration based on region characteristics, available neighbouring CWSs over time and space, spreading of these neighbouring CWSs and higher number of incorrect flags by radar filter.

To attain the answer to the main question, several sub-questions presented in section 1.3 are answered. These answers are:

- *What are the properties and errors of rainfall data in CWS network in the study area?*

The density and availability of CWSs are properties of the CWS network addressed in this study. A higher CWS network density (5.93 km² per CWS) is found compared to density of professional gauge network (1116.08 km² per automatic KNMI gauge) in the study area. Nonetheless, a CWS network is more dense in populated areas (Kidd et al., 2017) and not all CWSs are continuously available over time. The number of available CWSs roughly doubled from 369 CWSs measuring at the start to ~ 668 CWSs measuring at the end of the selected period from October 2015 to October 2017. The density of CWS network and number of available CWSs differ in other regions from the spatial and temporal properties found here. Other more densely populated regions are also expected to have a higher CWS network density compared to professional rain gauge network. However, the number of available CWSs depends on the

period of interest and CWS of interest. On the other hand, this study shows an increase in number of available CWSs and the author expects more growth in the future resulting in a higher CWS network density.

Multiple types of erroneous behaviours were encountered in rainfall data of CWSs. The first error identified is the lacking 10-min intervals. $\sim 5\%$ of all 10-min intervals are lacking. The number of lacking 10-min intervals per CWS deviates between 0% and 74%. Secondly, some CWSs appear to function partially normal and partially biased due to cumulative values, which are reset once in a while. Thirdly, false rainfall intensities in the rainfall data of CWSs are found. Lastly, UR errors are identified as result of a comparison between CWSs and nearby automatic KNMI gauge (in 5-kilometre radius). An UR error is an incorrect zero in the CWS data, when it is actually raining at that moment. The focus in this study is on the identification and flagging of these UR errors in the CWS rainfall data.

- *What filters can be designed to automatically detect UR errors from citizen rain gauge data?*

Two types of filters are developed to detect and flag UR errors from citizen rain gauge data at every timestep, namely the CWS filter and the radar filter. The CWS filter uses the median precipitation depth computed from neighbouring CWSs within a radius of 8 kilometre. In order to compute the median precipitation depth at least 3 neighbouring CWSs are required and the assumption is made that the majority of neighbouring CWSs within the radius of the CWSs of interest function properly. If less than 3 CWSs are available within the required radius, then the flag of the previous timestep is taken over. An advantage of the CWS filter is that it does not require other rainfall sources in order to work. The radar filter uses the overlying radar pixel to detect and flag undetected rainfall values. A precipitation depth estimated by the overlying radar pixels larger than 0.06 millimetre (q_{rad}), while there is zero rainfall measured by CWS, gives that the radar filter hands out an UR flag. A benefit is that the radar filter has continuous availability of required information to perform the filter for all CWSs.

- *What is the performance of the proposed filters? What (dis)similarities are present (filter performance wise) between the filters based on neighbouring CWSs and radar?*

Both the CWS filter and radar filter are considered to be adequate in detecting and flagging of UR errors. On average 83%, 84%, 75% and 88% of the ZI's (simulated UR errors) are flagged by the CWS filter, whereas the radar filter flags on average 74%, 78%, 74% and 77% of these intervals for the automatic KNMI gauges in Rotterdam, Voorschoten, Hoek van Holland and Cabauw. The performances of the CWS filter to detect and flag these undetected rainfall values is thus higher compared to the radar filter for stations in Rotterdam, Voorschoten and Cabauw. Except for the station in Hoek van Holland, the CWS filter has a similar performance to radar filter. The lower performance of CWS filter for the station in Hoek van Holland results from having less than 3 neighbouring stations combined with the presence of a physical barrier. This mitigates the suitability of neighbouring CWSs to identify UR values. Despite the lower overall performance, the radar filter is more stable due to the continuous availability of required information to perform this filter.

The radar filter, on the other hand, gives a higher number of false alarms (incorrect flags) compared to CWS filter due to the amount of disagreement between radar and gauges on the occurrence of zero rainfall. This false detection rate increases even more as result of radar artefacts. The disagreements between radar and gauges about the occurrence of rainfall also cause that the radar filter is not always able to detect all UR errors. Between 22% and 26% of ZI's are missed by the radar filter (false negative rates). These percentages are lower for the CWS filter, namely between 12% and 25% of ZI's. The relative costs compared to the false negative rates demonstrate that mainly ZI's on lower rainfall values are missed by both filters. Both filters experience therefore a performance decrease during drizzle rainfall events. In other words, the CWS filter and radar filter perform better during higher rainfall intensities.

- *How do filters flag different CWSs in study area over time and space? What (dis)similarities are present (flagging of CWSs wise) between the filters based on neighbouring CWSs and radar?*

Over a period from 1 October 2016 to 1 October 2017, 1.14% and 1.36% of all zero and nonzero measurements are flagged as UR by respectively the CWS filter and the radar filter. Less than 50% of these intervals are flagged by both the CWS filter and the radar filter. In general the radar filter is flagging less for the 10-min timesteps compared to the CWS filter due to the disagreement between gauge measurements and radar on the occurrence of rainfall. The percentages of UR flags over the total number of zeros and nonzeros timesteps per CWSs are on average lower for the CWS filter compared to the radar filter. However, the difference between minimum and maximum percentage of UR flags implies that the number of UR flags given to 747 CWSs vary more between the CWSs for the CWS filter than for the radar filter. In addition to this, the percentages of UR flags per CWSs were spatially mapped. From these spatial maps no clear trend in flagging of UR errors between CWSs is uncovered. Lastly, a higher monthly percentage of UR flags for both filters is generally found for months in which a higher number of time intervals with rainfall occurs. The radar filter has a higher monthly percentage of UR flags for all months, except December 2016. In this month higher monthly percentage of UR flags is found for CWS filter. December 2016 is also the month with the lowest percentage of UR flags and lowest percentage of Nonzero timesteps compared to the other months.

Future research should focus on testing these filters more in detail in order to get more specifications on how these filters work for different circumstances and to establish a strategy to select the appropriate filter in each situation. A suggestion on how to use the CWS filter was already discussed in chapter 7, namely data retrieval of all neighbouring CWSs within a radius of kilometres. The discussion also suggested that the CWS filter does not work properly during snow events. The exact performance of the filter during this type of events is still unknown, therefore data of historical snow events should be selected and run through CWS filter. Discussed in chapter 7 was also the application of raw radar product instead of gauge adjusted radar product for the radar filter. The use of raw radar product can amplify the disagreement on the occurrence of rainfall between gauges and radar. This induces that the performance of the radar filter decreases during near-real time application. In order to verify to what extent this happens, a comparison of performance of the radar filter should be carried out using both radar products.

Besides specific testing per filter, both the CWS filter and the radar filter should be validated for various circumstances. A future research step would be to evaluate the performance of filters in other climatic regions. A region with dominant convective storms might cause the CWS filter to no be applicable due to the local character of these storms. The radar filter could experience problems due to this as result of averaging over radar pixels. Besides different regions identify physical characteristics such as mountain ranges for which the suitability of filters is not analysed. Alongside these unknown circumstances, this study showed only the performance of the filters for 10-minutes resolution, while a smaller or higher temporal resolution of data can be required. Because of this step B and C from figure 2.3 should be repeated for a wide-variety of temporal resolutions followed by an intercomparison. Another constraint of this research was the application of only Netatmo CWSs, however the Netatmo platform is not always sufficient. For this reason, it is advised to repeat the same steps in this research for Wundermap, since this platform contains multiple CWS brands with different measurement techniques. Note that the application of Wundermap gives a lowest temporal resolution of approximately 10-min and that the platform includes processes which lead to additional errors.

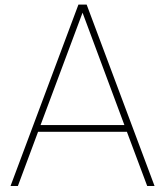
The focus of the research was on the identification and flagging of undetected rainfall values in CWS rainfall data. In order to create a reliable network for URM to facilitate forecasting of UPE, a QC system is required that includes filters for undetected rainfall intervals, as well as the other errors given above. Future research should therefore be directed towards developing quality-control mechanisms that detect and flag lacking intervals, cumulative behaviour in data and false rainfall intensities. Together these filters should be implemented in one QC system that is easy to apply on CWSs for practitioners.

References

- AD. EK-kwartfinale in Rotterdam afgelast wegens noodweer, July 2017. URL <https://www.ad.nl/buitenlands-voetbal/ek-kwartfinale-in-rotterdam-afgelast-wegens-noodweer~aa30b94a/>. [Accessed: 2017-10-07].
- R.M. Ashley, D.J. Balmfort, A.J. Saul, and J.D. Blanskby. Flooding in the future - predicting climate change, risks and responses in urban areas. *Water Science and Technology*, 52(5):265–273, 2005.
- S. Bell, D. Cornford, and L. Bastin. The state of automated amateur weather observations. *Weather*, 68(2): 36–41, 2013.
- S. Bell, D. Cornford, and L. Bastin. How good are citizen weather stations? addressing a biased opinion. *Weather*, 70(3):75–84, 2015.
- A. Berne, G. Delrieu, J. Creutin, and C. Obled. Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*, 299(3):166 – 179, 2004.
- S. Blenkinsop, E. Lewis, S. C Chan, and H. J Fowler. Quality-control of an hourly rainfall dataset and climatology of extremes for the uk. *International Journal of Climatology*, 37(2):722–740, 2017.
- M. K. Butler. Personal weather stations and sharing weather data via the internet. *Weather*, 2018.
- CoCoRaHS. Community Collaborative Rain, Hail and Snow Network, 1998. URL <https://cocorahs.org/>. [Accessed: 2017-11-10].
- E. Cristiano, J.A.E. ten Veldhuis, and N.C. van de Giesen. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review. *Hydrology and Earth System Sciences*, 21(7): 3859–3878, 2017.
- CWOP. Citizen Weather Observer Program, 2000. URL <http://www.wxqa.com/>. [Accessed: 2017-11-10].
- De Urbanisten and Management team of Rotterdam Climate Proof. Rotterdam: Climate change adaptation strategy. Technical report, Municipality of Rotterdam, 2013.
- L.W. de Vos, H. Leijnse, A. Overeem, and R. Uijlenhoet. The potential of urban rainfall monitoring with crowd-sourced automatic weather stations in amsterdam. *Hydrology and Earth System Sciences*, 21(2):765–777, 2017.
- I. Douglas, S. Garvin, N. Lawson, J. Richards, J. Tippet, and I. White. Urban pluvial flooding: a qualitative case study of cause, effect and nonstructural mitigation. *Journal of Flood Risk Management*, 3(2):112–125, 2010.
- J. Estévez, P. Gavilán, and J.V. Giraldez. Guidelines on validation procedures for meteorological data from automatic weather stations. *Journal of Hydrology*, 402(1):144–154, 2011.
- S. Gaitan, N.C. van de Giesen, and J.A.E. ten Veldhuis. Can urban pluvial flooding be predicted by open spatial data and weather data? *Environmental Modelling and Software*, 85:156–171, 2016.
- M. Gharesifard and U. Wehn. To share or not to share: Drivers and barriers for sharing data via online amateur weather networks. *Journal of Hydrology*, 535:181 – 190, 2016.
- A. Gires, I. Tchiguirinskaia, D. Schertzer, A. Schellart, A. Berne, and S. Lovejoy. Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data. *Atmospheric Research*, 138: 125 – 138, 2014.
- E. Habib, W. F. Krajewski, and A. Kruger. Sampling errors of tipping-bucket rain gauge measurements. *Journal of Hydrologic Engineering*, 6(2):159–166, 2001.

- D. L. Hartmann, A. M. G. Klein Tank, M. Rusticucci, L. V. Alexander, S. Brönnimann, Y. A. R. Charabi, F. J. Dentener, E. J. Dlugokencky, D. R. Easterling, A. Kaplan, B. J. Soden, P. W. Thorne, M. Wild, and P. Zhai. Observations: Atmosphere and surface. *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 159–254, 2013.
- S.M. Illingworth, C.L. Muller, R. Graves, and L. Chapman. Uk citizen rainfall network: a pilot study. *Weather*, 69(8):203–207, 2014.
- M. Jarraud. Guide to meteorological instruments and methods of observation (wmo-no. 8). Technical report, World Meteorological Organisation, 2008a.
- M Jarraud. Guide to meteorological instruments and methods of observation (wmo-no. 8). Technical report, 2008b.
- C. Kidd, A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, and D. B. Kirschbaum. So, how much of the earth’s surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, 98(1):69–78, 2017.
- F. Meier, D. Fenner, B. Grassmann, T. and Jänicke, M. Otto, and D. Scherer. Challenges and benefits from crowd sourced atmospheric data for urban climate research using berlin, germany, as testbed. In *ICUC9–9th International Conference on Urban Climate jointly with 12th Symposium on the Urban Environment*, 2015.
- F. Meier, D. Fenner, T. Grassmann, M. Otto, and D. Scherer. Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Climate*, 19:170 – 191, 2017.
- Met Office and KNMI. Weather Observation Website in Netherlands. <https://wow.knmi.nl/>, 2011. [Accessed: 2017-11-16].
- A. Molini, P. La Barbera, L.G. Lanza, and L. Stagi. Rainfall intermittency and the sampling error of tipping-bucket rain gauges. *Physics and Chemistry of the Earth, Part C Solar, Terrestrial & Planetary Science*, 26(10):737 – 742, 2001.
- C.L. Muller, L. Chapman, S. Johnston, C. Kidd, S. Illingworth, G. Foody, A. Overeem, and R.R. Leigh. Crowdsourcing for climate and atmospheric sciences: current status and future potential. *International Journal of Climatology*, 35(11):3185–3203, 2015.
- Municipality of Rotterdam. Duurzaam disther bij de rotterdammer. Technical report, Municipality of Rotterdam, 2016.
- Netatmo. Netatmo, 1993. URL <https://www.netatmo.com/nl-NL/site/>. [Accessed: 2017-10-15].
- J. Niemczynowicz. The rainfall movement — a valuable complement to short-term rainfall data. *Journal of Hydrology*, 104(1):311 – 326, 1988.
- NOS. Zware buien leggen Berlijn plat, June 2017. URL <https://nos.nl/artikel/2180640-zware-regenbuien-leggen-berlijn-plat.html>. [Accessed: 2017-10-07].
- A. Overeem, T. A. Buishand, and I. Holleman. Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar. *Water Resources Research*, 45(10), 2009a.
- A. Overeem, I. Holleman, and T. A. Buishand. Derivation of a 10-year radar-based climatology of rainfall. *Journal of Applied Meteorology and Climatology*, 48(7):1448–1463, 2009b.
- C. D. Peters-Lidard, M. Clark, L. Samaniego, N. E.C. Verhoest, T. van Emmerik, R. Uijlenhoet, K. Achieng, T. E. Franz, and R. Woods. Scaling, similarity, and the fourth paradigm for hydrology. *Hydrology and Earth System Sciences*, 2017.
- Y. Qi, S. Martinaitis, J. Zhang, and S. Cocks. A real-time automated quality control of hourly rain gauge data based on multiple sensors in mrms system. *Journal of Hydrometeorology*, 17(6):1675–1691, 2016.

- W. Schilling. Rainfall data for urban hydrology: what do we need? *Atmospheric Research*, 27(1):5 – 21, 1991. ISSN 0169-8095.
- M. Schleiss, J. Jaffrain, and A. Berne. Statistical analysis of rainfall intermittency at small spatial and temporal scales. *Geophysical Research Letters*, 38(18), 2011.
- M. H. Spekkers, M. Kok, F. H. L. R. Clemens, and J. A. E. ten Veldhuis. A statistical analysis of insurance damage claims related to rainfall extremes. *Hydrology and Earth System Sciences*, 17(3):913–922, 2013.
- M. H. Spekkers, F. H. L. R. Clemens, and J. A. E. ten Veldhuis. On the occurrence of rainstorm damage based on home insurance and weather data. *Natural Hazards and Earth System Sciences*, 15(2):261–272, 2015.
- K Stone, R van Duinen, W van Veerbeek, and S Dopp. Sensitivity and vulnerability of urban systems—assessment of climate change impact to urban systems. Technical report, Deltares, Netherlands, 2011.
- J.A.E. ten Veldhuis. How the choice of flood damage metrics influences urban flood risk assessment. *Journal of Flood Risk Management*, 4(4):281–287, 2011.
- United Nations, Department of Economic and Social Affairs, Population Division. World urbanization prospects: the 2014 revision. Technical report, 2015.
- G.J.G. Upton and A.R. Rahimi. On-line detection of errors in tipping-bucket raingauges. *Journal of Hydrology*, 278(1):197 – 212, 2003.
- Weather Underground. Weather Underground, 1993. URL <https://www.wunderground.com/>. [Accessed: 2017-10-15].
- I. Zahumenský. Guidelines on quality control procedures for data from automatic weather stations. Technical report, World Meteorological Organization, 2004.



Overview of CWS Platforms

Table A.1: Overview of properties from different Citizen Weather Stations initiatives based on the following sources: CoCoRaHS (<https://www.cocorahs.org/>), UCrain (Illingworth et al., 2014), CWOP (<http://wxqa.com/>), WOW (<http://wow.metoffice.gov.uk/> & <https://wow.knmi.nl/>), Wundermap (<https://www.wunderground.com/wundermap>) and Netatmo weathermap (<https://weathermap.netatmo.com/>)

	CoCoRaHS	UCrain	CWOP	WOW	Wundermap	Netatmo Weathermap
Complete name	Collaborative Rain, Hail, Snow Network	UK Citizen Rain-fall Network (pilot study)	Citizen Weather Observer Program	Weather Observation Website	-	-
Operated by	Colorado State University	Departments from University of Manchester, University of Birmingham and University of Leicester	National Oceanic and Atmospheric Administration (NOAA)	UK Met Office (& KNMI)	Weather ground	Under- Netatmo
Number of stations	-	13	> 22 000 sites (Butler, 2018)	>10 000 stations	>250 000 station	Meier et al. (2017) found 870 stations in Basel Switzerland, 547 stations London, United Kingdom, 190 stations in New York City, United states of America, and 4742 station in Paris, France on March 1, 2016.
Areas	United States of America, Canada and Bahamas	Manchester, Birmingham and Leicester in United Kingdom	Worldwide, but large portion in USA	220 countries	worldwide	worldwide
Type of stations	Manual gauges	Self-built manual gauges	Different types of stations are used	Variety of stations	Variety of stations	Netatmo weather stations

Table A.1 – continued from previous page

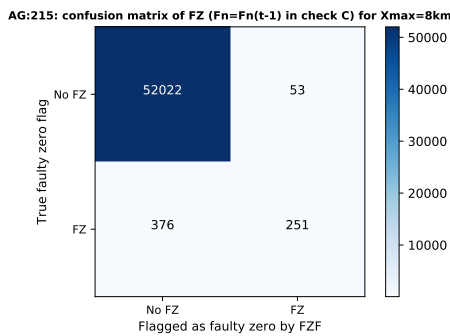
	CoCoRaHS	UCRaIN	CWOP	WOW	Wunderground	Netatmo Weathermap
Connection between station and platform	Reporting data per phone or by online access to your data to upload the data	Reporting of data per email, twitter or contact with one of university contact persons	Different types of software are used for the connection	WOW-nl advice software for connection between station and platform	Wunderground advice software for connection and query data from Netatmo weathermap	Netatmo has built-in software in the stations to upload the data from station to platform. The owners themselves have to select if they want to share outside module
Measured variables	Rainfall, hail and snow	Rainfall	Temperature, relative humidity, wind, rainfall and barometric pressure	Temperature, relative humidity, wind, rainfall and barometric pressure	Temperature, relative humidity, wind, rainfall and barometric pressure	Temperature, wind and rainfall (outside module)
Temporal Resolution	daily at ~ 7am	daily on weekdays between 3 June 2013 and 12 July 2013	5 min	~ 10 min	~ 10 min	~ 5 min
Meta data	Name, station number (include position), station name and time of measurement	-	Station code, latitude, longitude, elevation, location, county, data provider, station type, software for connection and picture of top view of site	Position of station and picture, as well as, the contributors get the possibility to classify their station based on several characteristics: situation of site, lay-out of urban area, time of measurement, measurement of air temperature, rainfall and wind.	Station ID, latitude, longitude, altitude, elevation and brand of device if known	MACaddress, moduleID, timestamp, latitude, longitude, altitude, rain gauge battery status and signal-quality of wifi.

Table A.1 – continued from previous page

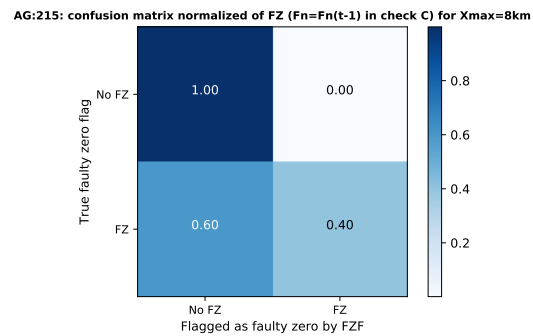
	CoCoRaHS	UCRaiN	CWOP	WOW	Wunderground	Netatmo Weathermap
Quality-Control	Yes, both data entry Quality-Control (QC) and consistency checks	No	Yes, there is option to request quality-controlled data from stations. QC checks are sub-hourly validity check for rainfall. Other variables also have internal, temporal and spatial consistency checks every (sub-)hour. The spatial checks are buddy checks that uses nearby stations that have already passed the QC checks.	No, QC but more knowledge about the situation on how the measurement were taken	Yes, Wunderground has several QC checks in place. These checks are validity, internal rate of change and spatial checks. The spatial checks uses nearby stations. However the exact algorithm applied is still unknown (Butler, 2018)	Yes, a filter is available, but the specifications of filter are unknown.

B

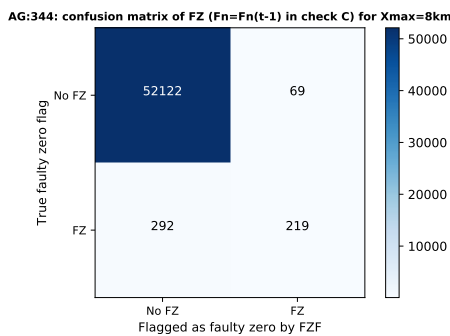
Determination of Threshold for the Simulation of UR Errors on datasets



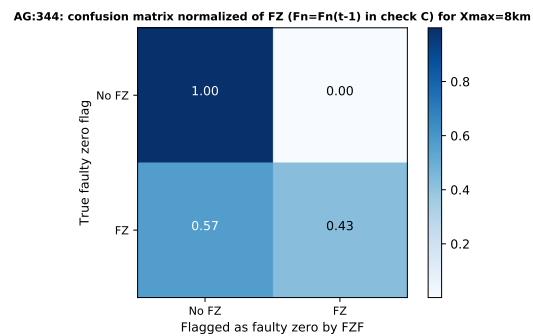
(a) Counts of TN, FP, FN and TP of ID 215



(b) TNR, FNR, FPR and TPR of ID 215



(c) Counts of TN, FP, FN and TP of ID 344



(d) TNR, FNR, FPR and TPR of ID 344

Figure B.1: Confusion matrices of the automatic KNMI gauges 215 (Voorschoten) and 344 (Rotterdam) developed for the evaluation of the performance of CWS filter. Figures a and c show the counts, whereas figures b and d give the rates after normalising the confusion matrices over the column axis. Remark that these performances are for the automatic KNMI datasets with simulated UR errors with threshold $q_{ag} = 0$. Also remark that FZ are UR flags for x-axis and ZI's for y-axis. No FZ are no UR flags for x-axis and original values for y-axis.

The automatic KNMI datasets with simulated UR errors are used to obtain the performances of the two filters. After that the first design of the CWS filter was created, it was run on these datasets with simulated UR errors with threshold $q_{ag} = 0$. Performances of the CWS filter for the stations in Voorschoten and Rotterdam are shown in Figures B.1a to B.1d. Analysis of B.1b and B.1d gives that 40% and 43% of ZI's are correctly flagged

as UR intervals by the CWS filter for respectively Voorschoten and Rotterdam. In other words, the CWS filter identifies less than 50% of the simulated UR errors giving an insufficient performance.

Because the performances of the CWS filter are rather low for these datasets, manual inspection of UR flags given by CWS filter for the station in Rotterdam is carried out. The enumeration below gives the results of the manual inspection.

- Simulated UR sequences with a length of 1 or 2 ZI('s) are not detected and flagged by the CWS filter.
- Only once, a simulated UR sequence with a length of 4 ZI's is totally flagged.
- Longer simulated UR sequences have often higher percentages of ZI's that are flagged by the CWS filter.
- From the 292 missed ZI's (false negatives in Figure B.1c) only 23 have originally higher rainfall values than the default tipping bucket volume. This means that 269 missed ZI's have originally lower rainfall depths than default tipping bucket volume of Netatmo CWS.

From the more in-depth investigation emerges that the CWS filter has difficulty with detecting the occurrence of simulated UR errors that are originally low rainfall amounts. A reason for the lower performance of CWS filter for ZI's at lower rainfall amounts is the use of neighbouring CWSs to detect and flag UR errors. As mentioned in section 3.1, CWSs applied in this research are tipping buckets with volumes of 0.101 millimetres or higher manually set as calibrated values. Lower rainfall intensities or drizzle rainfall events can cause that the tipping of buckets of the neighbouring CWSs did not take place at the specific time step. Simulated UR error in a station of interest can therefore be missed.

As discussed in section 1.3, the objective of this research is to take the first step towards a QC system that enhances the quality of CWSs and improves the suitability of crowdsourced CWS network for URM to facilitate forecasting of UPE. The lower performance of CWS filter during low rainfall intensities is therefore not of great importance for the research, since the occurrence of UR errors during low rainfall intensities are not normative to trigger urban pluvial flooding. Together with the fact that the Netatmo CWSs can only measure rainfall values equal to the tipping bucket volume, the decision is made to reverse the implementation of simulated UR errors back to the original rainfall values on the automatic KNMI gauge data for time steps that have original rainfall amounts below a threshold q_{ag} of 0.0505 millimetres.

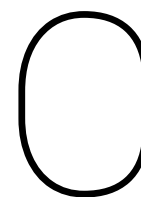
Table B.1: Total number and total rainfall amounts of the reversed UR errors due to implementation threshold q_{ag} of 0.0505 millimetres compared to threshold q_{ag} of 0 between 1 October 2015 and 1 October 2016 for automatic KNMI gauges within study area.

ID	number of reversed ZI's	P_{total} [mm]
215	346	5.3157
330	276	4.3687
344	305	4.8879
348	371	5.4177

Table B.1 gives the total number and total rainfall amounts of ZI's that are reversed per automatic KNMI gauges in the study area due to the implementation of the new threshold over the period between 1 October 2015 and 1 October 2016. Approximately less than 6 millimetres of rainfall is missed by filters as result of the new threshold. These low missed rainfall amounts are considered to have negligible effect for UPE. Table B.2 shows the effect of the new threshold ($q_{ag}=0.0505$ mm) on the performances of CWS filter. The performances of the CWS filter to correctly flag ZI's increase substantial for the new datasets with $q_{ag}=0.0505$ mm compared to old datasets $q_{ag}=0$, therefore the threshold of $q_{ag}=0.0505$ mm is accepted.

Table B.2: Performances of CWS filter for datasets with simulated UR errors with respectively $q_{ag} = 0$ mm (old) and $q_{ag} = 0.0505$ mm (new). The performance increase as result of the implementation of the new threshold.

ID	TPR_{old}	TPR_{new}	ΔTPR
215	44%	79%	+35%
330	52%	66%	+14%
344	43%	85%	+42%
348	50%	95%	+45%



Determination and Verification of fixed Radius for CWS Filter

As mentioned in section 4.1, the radius applied in decision C and D is chosen to be fixed. The determination and verification of the fixed radius are shown below.

Determination of fixed radius

Table C.1 shows the percentages of UR flags is less than 0.5 percent for all the radius and automatic KNMI gauges. The automatic KNMI rain gauges in Voorschoten and Rotterdam have the lowest percentage of UR flags ($F_n(t) = 1$) for radii of 9 or 10 km and 7 km respectively. A radius of 8 km is therefore an optimum, since the radius has the second lowest percentage of UR flags for both stations. The same is also the case for the automatic KNMI gauge in Cabauw. Hence a fixed radius of 8 km is selected.

Table C.1: Percentages of UR flags on raw rainfall data obtained from automatic KNMI gauges.

Radius [km]	5	6	7	8	9	10	11	12	13	14	15	
215	$F_n(t) = 0$ [%]	99.87	99.87	99.88	99.90	99.91	99.91	99.89	99.86	99.84	99.84	99.84
	$F_n(t) = 1$ [%]	0.12	0.12	0.11	0.10	0.09	0.09	0.10	0.13	0.15	0.16	0.16
	No flags [%]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
330	$F_n(t) = 0$ [%]	98.98	98.98	98.98	98.71	99.68	99.78	99.84	0.10	99.81	99.83	99.83
	$F_n(t) = 1$ [%]	0.13	0.12	0.12	0.39	0.32	0.21	0.16	0.17	0.18	0.17	0.17
	No flags [%]	0.89	0.89	0.89	0.89	0.00	0.00	0.00	0.00	0.00	0.00	0.00
344	$F_n(t) = 0$ [%]	99.81	99.84	99.90	99.87	99.84	99.83	99.86	99.86	99.85	99.86	99.84
	$F_n(t) = 1$ [%]	0.18	0.16	0.10	0.13	0.15	0.16	0.14	0.14	0.14	0.14	0.15
	$F_n(t) = 1$ [%]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
348	$F_n(t) = 0$ [%]	99.48	99.68	99.80	99.81	99.67	99.69	99.74	99.78	99.78	99.73	99.72
	$F_n(t) = 1$ [%]	0.08	0.32	0.20	0.19	0.33	0.31	0.26	0.22	0.22	0.27	0.27
	No flags [%]	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Verification of fixed radius

As mentioned in subsection 4.1, three neighbouring CWSs are required to compute the median precipitation depth in step D. Figure C.1 shows that 75 percent of the CWSs has three neighbouring CWSs inside a radius of approximately 2 km. In fact, only four CWSs of the 755 CWSs (760 CWSs including the empty CWSs) or 0.53 percent require a larger radius than 8 km for three neighbouring stations inside the radius. The radii of these stations ranges between 8.3 km and 9.3 km. Three of those CWSs are located in the province Zeeland and one is located close to East border. The automatic KNMI gauges have 3 CWSs within a radius of 1.262 km to 1.698 km. In other words, 99.5 percent of the CWSs and all automatic KNMI gauges have three neighbouring CWSs within a radius of 8 km based on their locations.

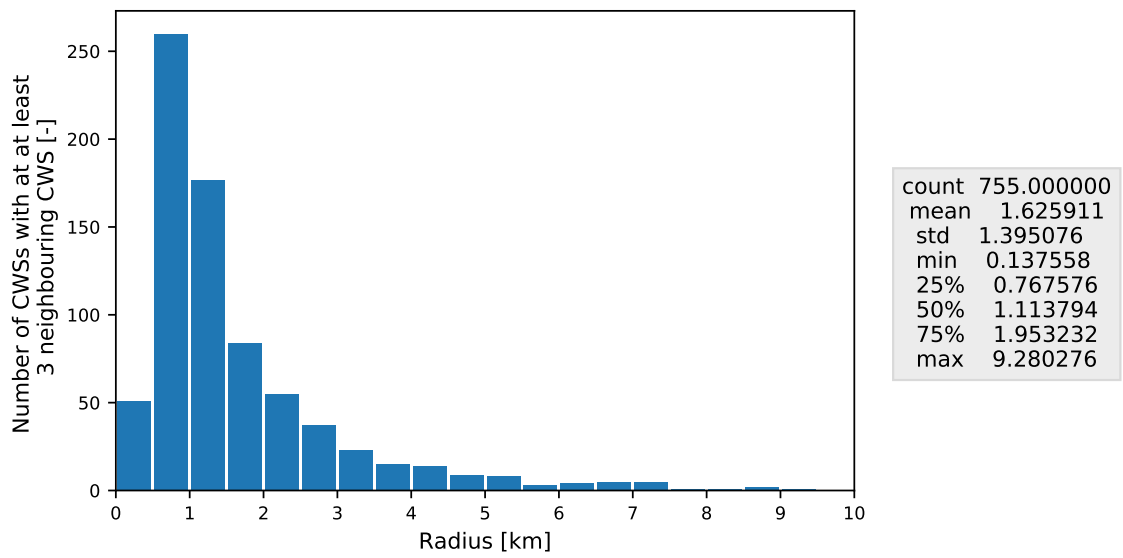


Figure C.1: A histogram of number of CWSs with at least 2 neighbouring CWSs within circle with radius varying between 0 and 10 km.



Determination of q_{rad} using Drizzle Check

As mentioned in section 4.2, weather radars and gauges can disagree on the occurrence of rainfall for certain timestep. This appendix therefore starts with presenting the (dis)agreements on the occurrence of rainfall data between automatic KNMI gauges and the overlying radar pixels without q_{rad} .

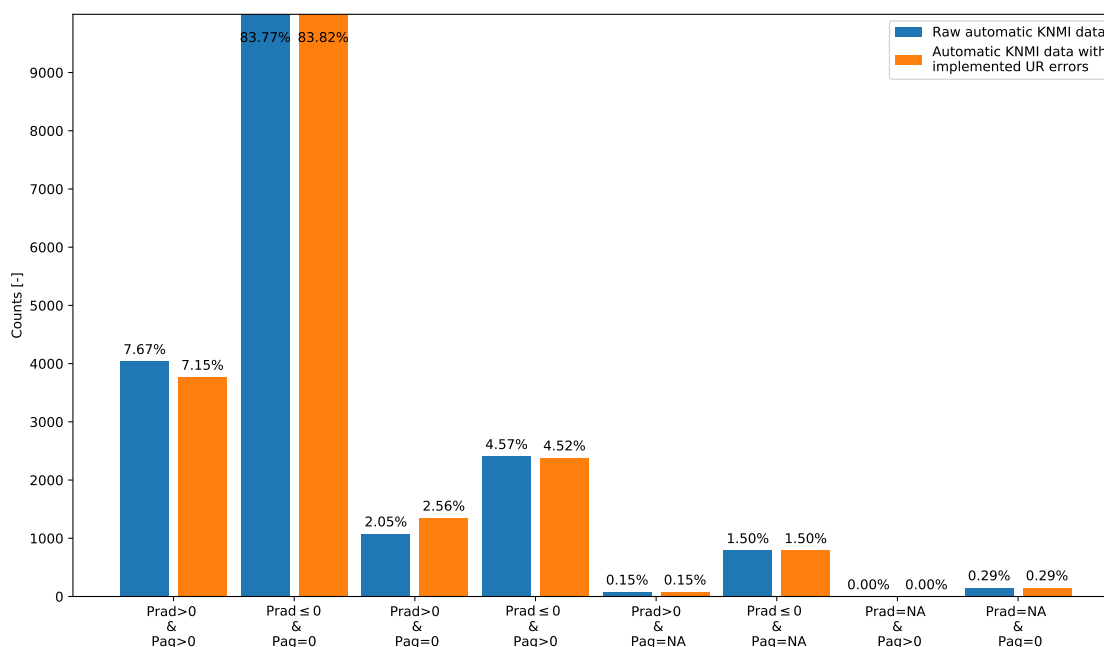


Figure D.1: A barplot with the occurrence of combinations of value types between radar P_{rad} and automatic KNMI gauge P_{ag} in Voorschoten (AG215). The possible combinations of values types are given for overlying radar pixel of this stations in relation to raw automatic KNMI dataset (blue) and automatic KNMI dataset with implemented UR errors (orange). $P_{\text{rad}} \leq 0$ means no rainfall. The percentages shown in the bars represent the fraction of the specific combination type divided by the total dataset.

Figures D.1 and D.2 show the occurrence of combinations value types for the overlying radar pixel in relation to raw automatic KNMI data and automatic KNMI data with implemented UR errors, which are used in chapter 5 to determine the performances of the two filters. This figure shows that zero rainfall is estimated by radar and gauges approximately 84% at the same intervals in the datasets from the gauges in Voorschoten and Rotterdam. Assuming that all lacking intervals (NA) are also zero values brings these percentages on almost 86%. Both radar and gauges measure rainfall on average 7% of the intervals. This means that approximately 91% of the intervals an agreement on the occurrence of rainfall or no rainfall is achieved between radar and gauges. The other 7% disagreement between radar and gauge however causes the limitations of the radar filter.

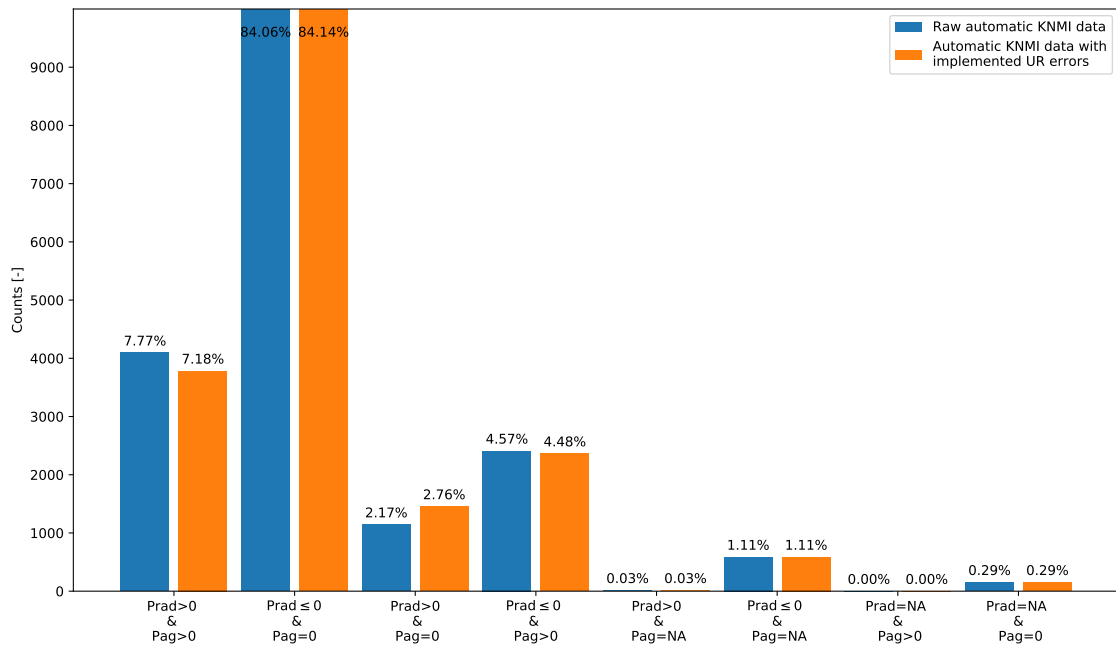
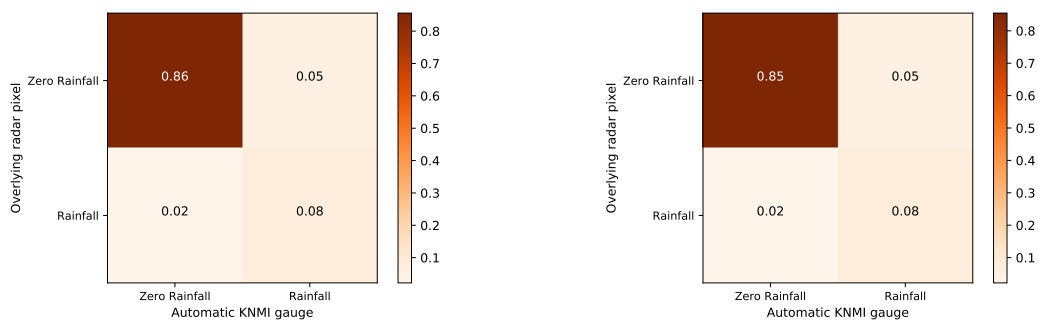


Figure D.2: A barplot with the occurrence of combinations of value types between radar P_{rad} and automatic KNMI gauge P_{ag} in Rotterdam (AG344). The possible combinations of values types are given for overlying radar pixel in relation to raw automatic KNMI dataset (blue) and automatic KNMI dataset with implemented UR errors (orange). $P_{\text{rad}} \leq 0$ means no rainfall. The percentages shown in the bars represent the fraction of the specific combination type divided by the total dataset.

The disagreements on the occurrence of rainfall between overlying radar pixel and automatic KNMI data with implemented UR errors are shown in the top right and bottom left square of Figures D.3a and D.3b. For 2% of all the intervals, radar estimates rainfall, whereas the automatic KNMI gauge gives no rainfall. These intervals are no UR errors but they will automatically be flagged as UR. This is the reason for the decision to implement a threshold q_{rad} . Only when the rainfall estimates from the radar exceeds this threshold the interval is considered to be a rainfall timestep.



(a) AG215 in Voorschoten

(b) AG344 in Rotterdam

Figure D.3: Normalised confusion matrices showing the (dis)agreement on the occurrence of rain or no rain between automatic KNMI gauges AG215 and AG344 and their overlying radar pixel.

An analysis is executed to obtain more knowledge about the rainfall amounts that occur for the phenomena described above and to be able to determine the value of the threshold. This analysis is called the drizzle check because these rainfall amounts are expected to be low rainfall amounts. The 75 percentile of these rainfall amounts is 0.06 millimetre for the gauges in Voorschoten, Hoek van Holland and Rotterdam as shown in Table D.1. For the station in Cabauw the 75 percentile is 0.07 millimetre. In other words, most of the intervals

in which the phenomena occurred are drizzle rains. This emphasises again the need of the threshold q_{rad} .

Table D.1: 75 percentile of radar rainfall amounts for the intervals that $P_{\text{rad}} > 0$ and $P_{\text{ag}} = 0$ for the automatic KNMI gauges in Voorschoten (AG215), Hoek van Holland (AG330), Rotterdam (AG344) and Cabauw (AG348).

ID	75 percentile of P_{rad}
AG215	0.06 mm
AG330	0.06 mm
AG344	0.06 mm
AG348	0.07 mm

The (dis)agreements on the occurrence of rainfall between the overlying radar pixel and automatic KNMI data with ZI's are investigated in the same manner, as is done in Figure D.3 for the application of three threshold for q_{rad} , namely 0.05 mm (half default tipping bucket volume), 0.06 mm (obtained from drizzle check) and 0.101 mm (default tipping bucket volume). Figures D.4a to D.4f give the normalised confusion matrices showing the (dis)agreement between radar and gauges. These figures show that the (dis)agreements are similar for the two gauges in Voorschoten and Hoek van Holland. These results compared with Figure D.3 show the disagreements about radar giving rainfall and gauge measuring no rainfall, decrease with 1% or 2% for the thresholds. The implementation of the thresholds results in approximately 1% (q_{rad} of 0.05 mm) or 0% (other thresholds) of all intervals being flagged as UR on forehand. It should be taken into account however that increasing the threshold causes an increasing number of intervals to disagree on the occurrence of rainfall at the gauge but no rainfall estimated by radar. These disagreements are 6%, 7% and 8% for q_{rad} being respectively 0.05 mm, 0.06 mm and 0.101 mm. The higher these disagreements becomes the higher the possibilities are that UR errors in rainfall data are missed by the radar filter. This is the reason the optimum q_{rad} of 0.06 mm is chosen to remove 75% of the incorrect UR flags given by the radar filter a priori without compromising on the ability to flag UR errors.

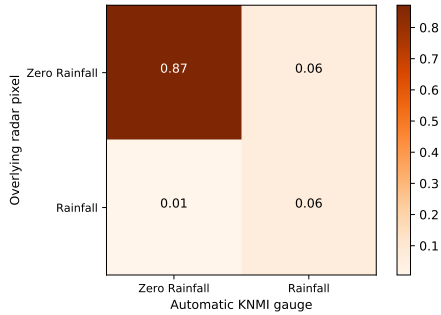
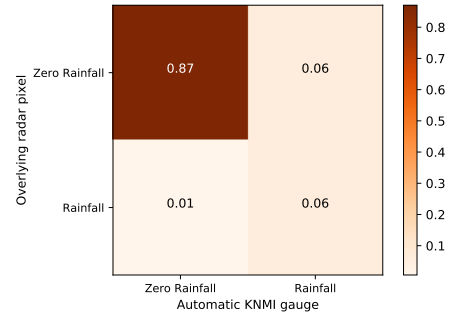
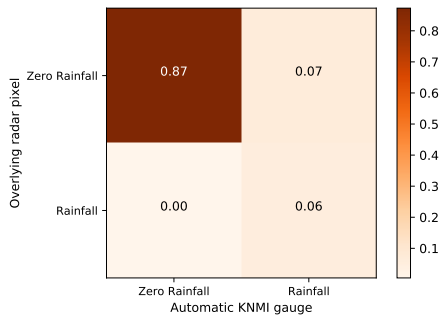
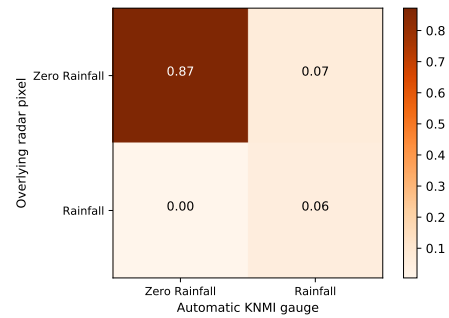
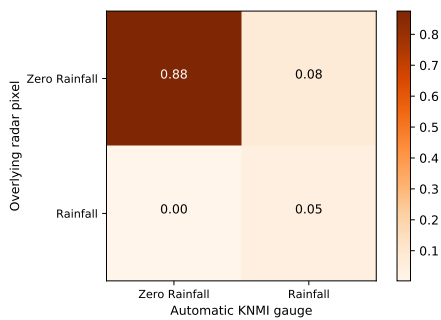
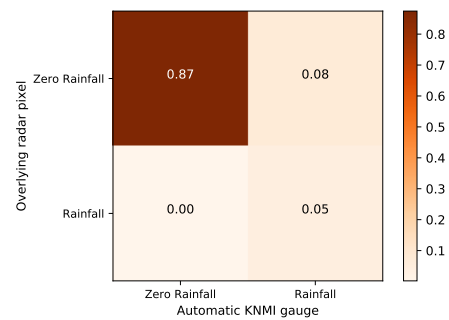
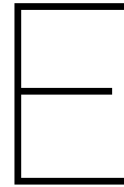
(a) AG215 in Voorschoten with $q_{\text{rad}}=0.05$ mm(b) AG344 in Rotterdam with $q_{\text{rad}}=0.0505$ mm(c) AG215 in Voorschoten with $q_{\text{rad}}=0.06$ mm(d) AG344 in Rotterdam with $q_{\text{rad}}=0.06$ mm(e) AG215 in Voorschoten with $q_{\text{rad}}=0.101$ mm(f) AG344 in Rotterdam with $q_{\text{rad}}=0.101$ mm

Figure D.4: Normalised confusion matrices showing the (dis)agreement on the occurrence of rain or no rain between automatic KNMI gauges AG215 and AG344 and their overlying radar pixel with the implementation of three thresholds: 0.0505 mm, 0.06 mm and 0.101 mm

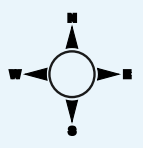
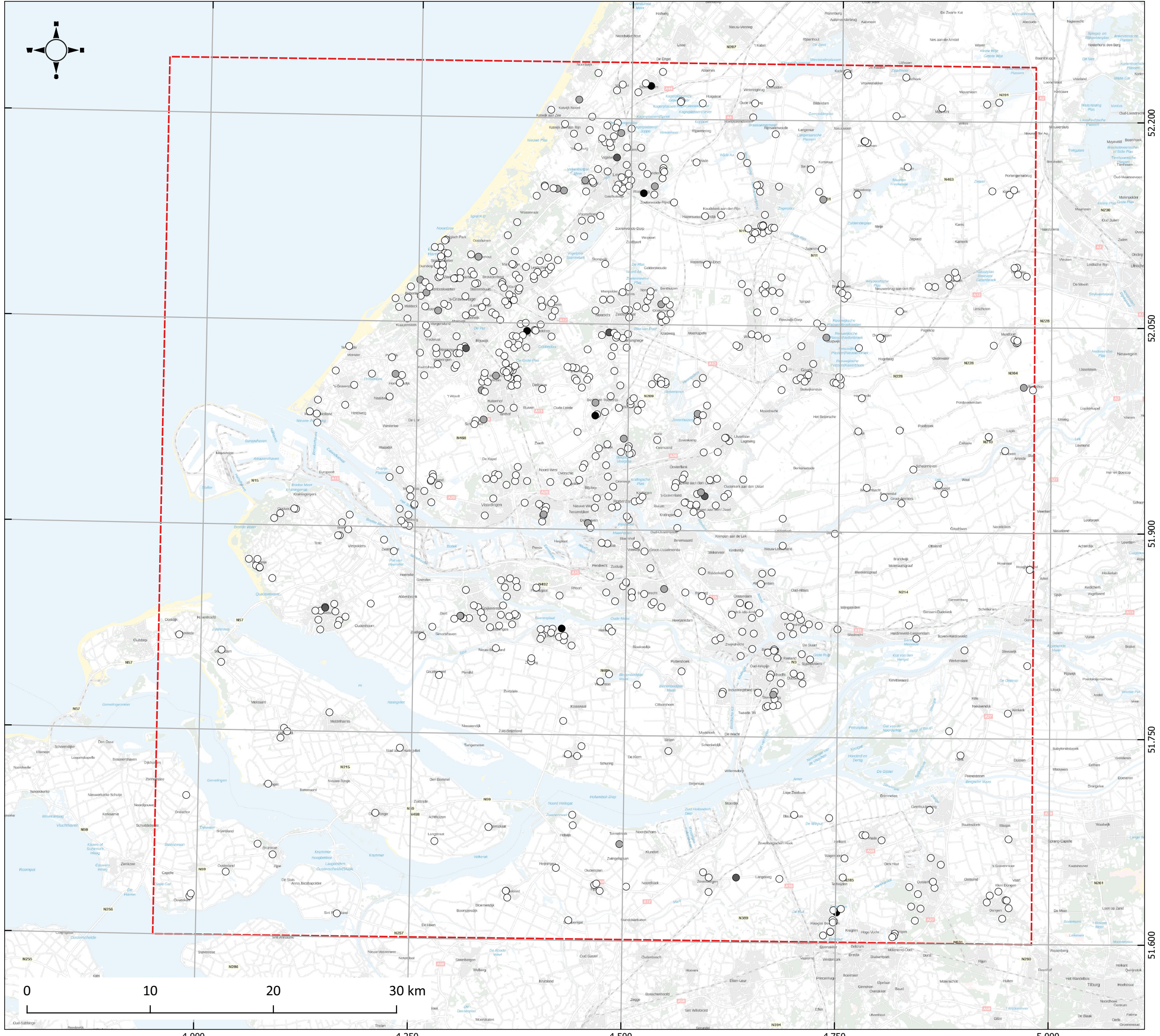


Spatial Maps of Lacking Intervals and Flags given by the Filters

The spatial maps are given here. The maps are shown on the next pages in the following order:

1. *Percentage of lacking intervals (figure 3.3 in chapter 3):* The start and end time for a period from 1 October 2015 to 1 October 2017 are determined for each CWSs. The start time and end time are the first and last timesteps that a station gives zero or nonzero values. The percentage of lacking intervals is the number of timesteps that no measurement is available divided by the total number of timesteps between start and end time.
2. *Percentage of UR flags per CWS given by CWS filter:* This percentage is computed from the number of UR flags given by CWS filter divided by the total number of zero and nonzero values for a period between 1 October 2016 and 1 October 2017.
3. *Percentage of UR flags per CWS given by radar filter:* This percentage is derived from the number of UR flags given by CWS filter divided by the total number of zero and nonzero values for a period between 1 October 2016 and 1 October 2017.

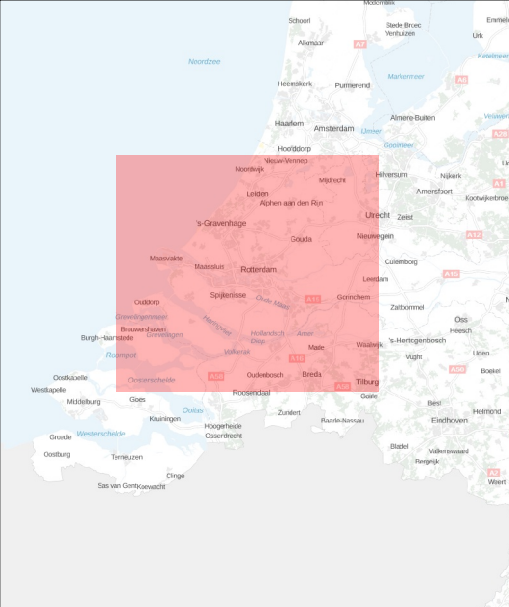
This page is intentionally left blank.



Legend

Lacking intervals [%] [753]

- 0 - 20 [711]
- 20 - 40 [29]
- 40 - 60 [7]
- 60 - 74 [6]
- ▭ Study area

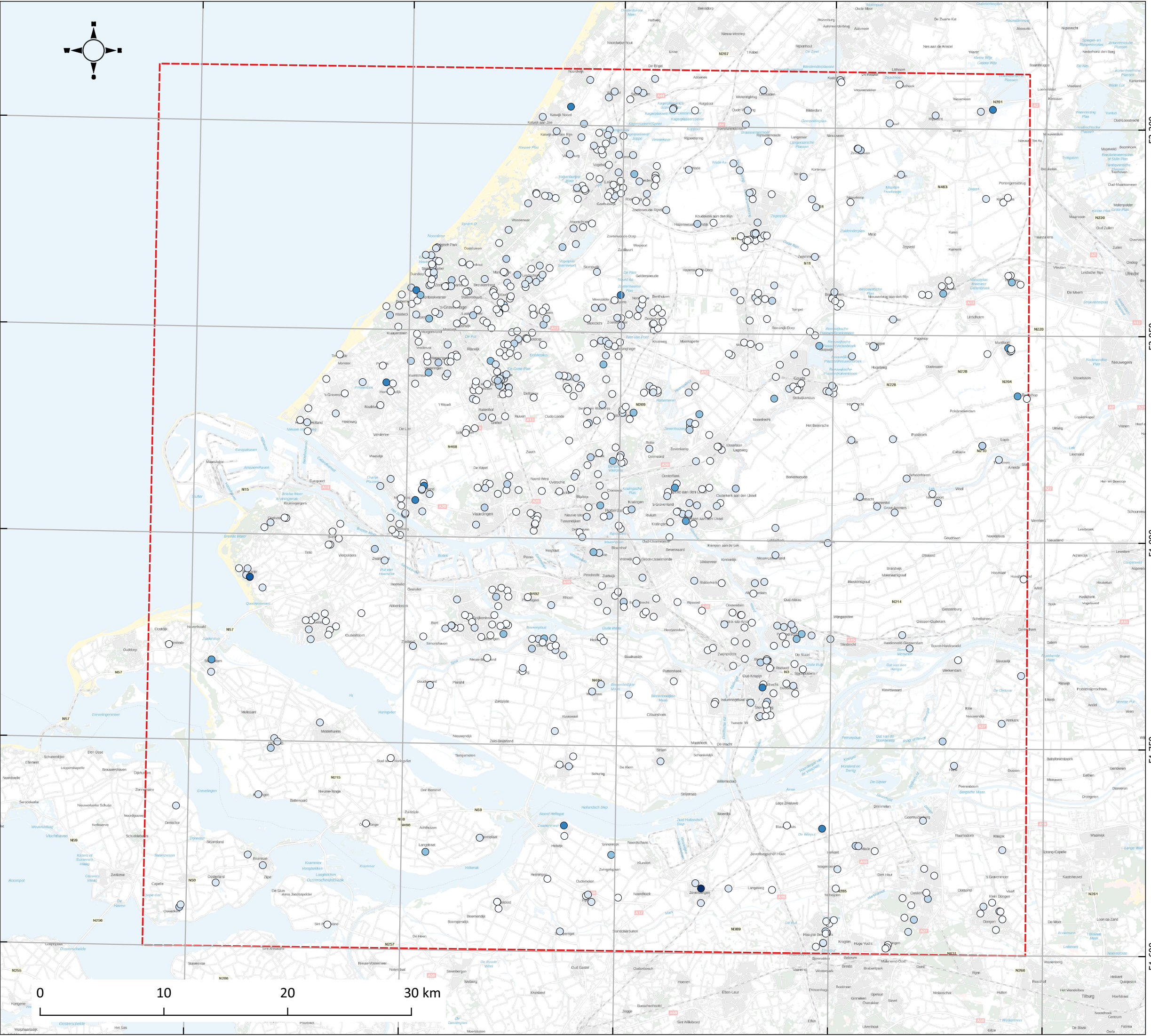
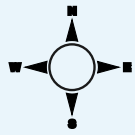


EPSG: 28992/ RD NEW

Sources:
 PDOK
 (<https://www.pdok.nl/nl/producten/pdok-downloads/download-basisregistratie-grootschalige-topografie>)

Netatmo weathermap
 (<https://weathermap.netatmo.com/>)

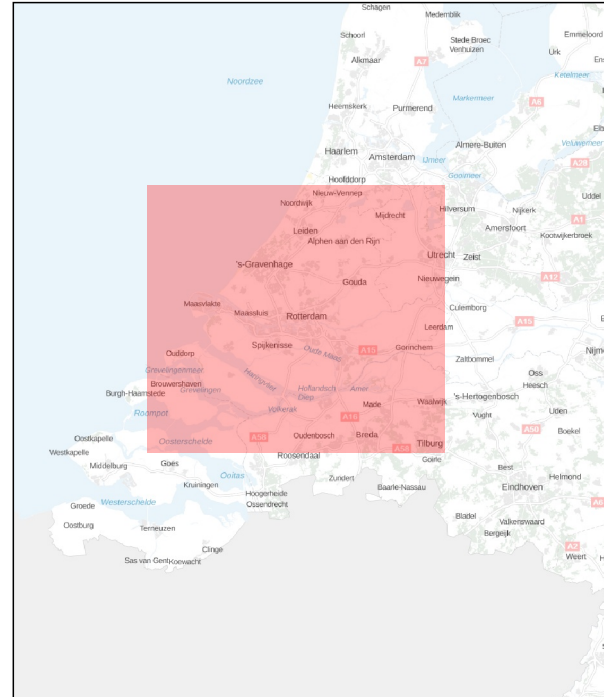
This page is intentionally left blank.



Legend

UR flags given by CWS filter [%] [747]

- 0.00 - 1.00 [434]
 - 1.00 - 2.00 [231]
 - 2.00 - 3.00 [43]
 - 3.00 - 4.00 [21]
 - 4.00 - 5.00 [5]
 - 5.00 - 6.00 [11]
 - 6.00 - 7.00 [1]
 - 7.00 - 7.26 [1]
- Study area



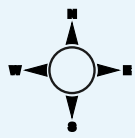
EPSG: 28992/ RD NEW

Sources:
PDOK
(<https://www.pdok.nl/nl/producten/pdok-downloads/download-basisregistratie-grootchalige-topografie>)

Netatmo weathermap
(<https://weathermap.netatmo.com/>)



This page is intentionally left blank.

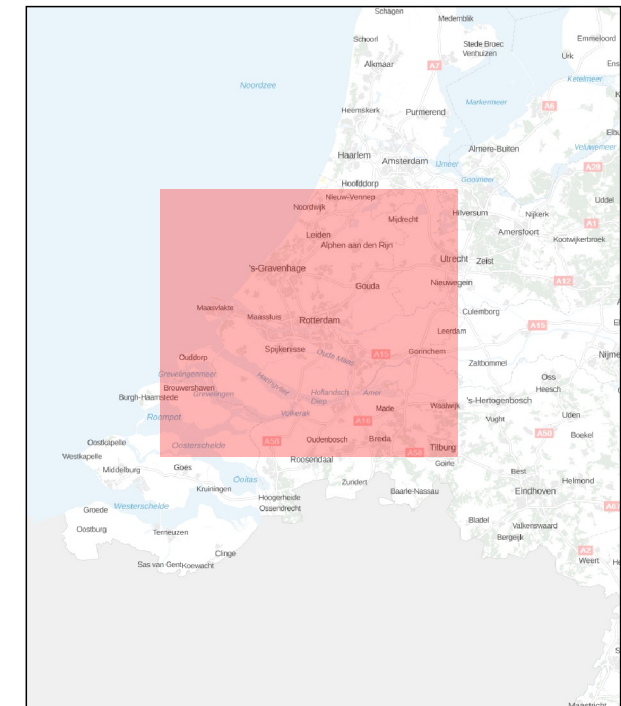


Legend

UR flags given by radar filter [%] [747]

- 0.02 - 1.00 [279]
- 1.00 - 2.00 [379]
- 2.00 - 3.00 [48]
- 3.00 - 4.00 [25]
- 4.00 - 5.00 [11]
- 5.00 - 6.00 [4]
- 6.00 - 6.42 [1]

Study area



EPSG: 28992/ RD NEW

Sources:
PDOK
(<https://www.pdok.nl/nl/producten/pdok-downloads/download-basisregistratie-grootschalige-topografie>)

Netatmo weathermap
(<https://weathermap.netatmo.com/>)

