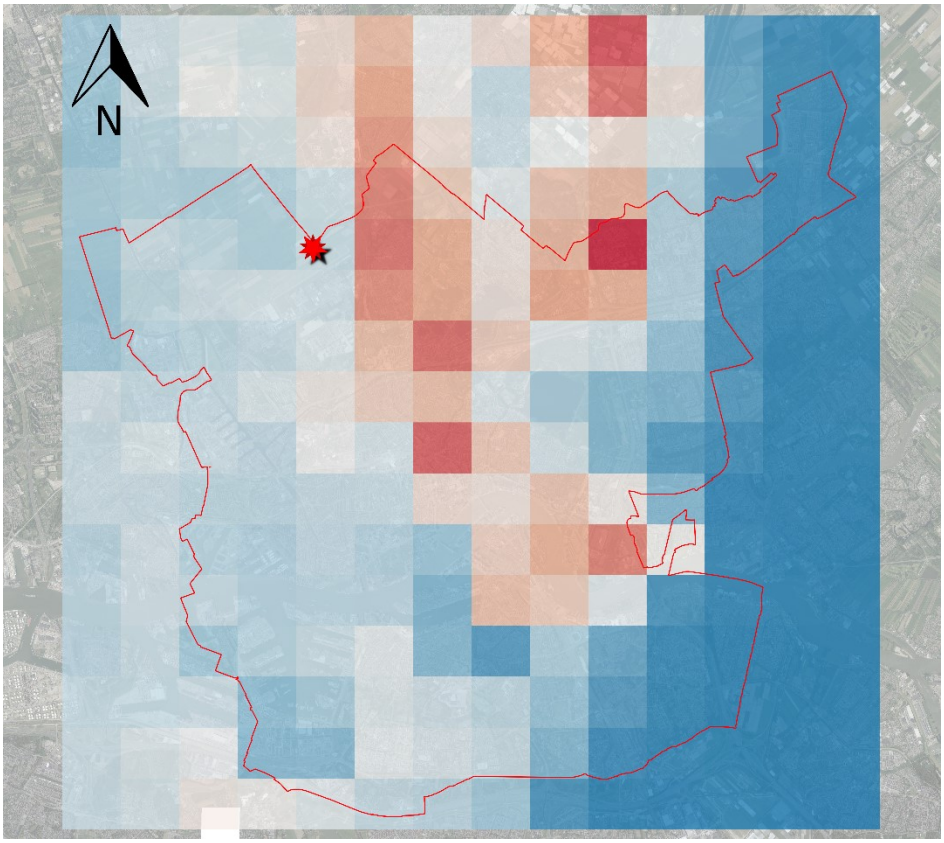
The potential of citizen observatories for improving spatial measurements of rainfall in cities



By J. Schoester

To obtain the degree of Master of Science at the Delft University of Technology

Student number	4166035	
Thesis committee	Dr. Ir. J. A. E. ten Veldhuis	TU Delft, water resources department
	Prof. Dr. Ir. N. C. van de Giesen	TU Delft, water resources department
	Dr. M. A. Schleiss	TU Delft, geoscience and remote sensing
Advisors	Dr. X. Tian	TU Delft, water resources department
	Ir. E.P.C. Bes	Gemeente Rotterdam, urban water

Contents

Abstract

Urban pluvial flooding occurs when the run-off converted rainfall exceeds the capacity of the sewer or stormwater system (Houston et al., 2011). This can result in damage to ecology, infrastructure, disruption to human activities, injuries and in the worst-case scenario, loss of life (Biniyam et al., 2017). With the expected increase in the world's population living in urban areas to 68% by 2050 (UN, 2018) and a more frequent occurrence of extreme weather events with a longer duration and higher intensity in the future (IPCC, 2012), there is a need to improve early warning system and disaster management (Restrepo-Estrada et al., 2017). Lopez et al. (2005) mentions that one of the key factors in hydrological models to determine accurate flood estimates is to have accurate rainfall input. However, many urban areas might lack this information because sensors are not available, or the number of sensors is too few to cover the entire region with an acceptable resolution (Restrepo-Estrada et al., 2017). Kidd et al. (2017) mentions that the density of rain gauges varies per region, with Europe and Eastern-Asia (including Japan) having a decent coverage by rain gauges, while the rest of the world has a sparse coverage of rain gauges. Citizen science or community-based monitoring may offer a solution to fill in some of the data gaps (Paul et al., 2018). Citizen science is often defined as the participation of the general public in the research design, data collection and interpretation process together with scientists (Buytaert et al., 2014; Paul et al., 2018). In a study by Eilander et al. (2016), where social media like twitter and Facebook were used to follow real-time flood events, the limiting factor was the number of social-media users. With the increase of population in urban regions (UN, 2018), these might offer good testing grounds for pilot projects involving citizen science.

This study assesses the potential of citizen observatories for improving the spatial measurements of rainfall in comparison to a single professional station. Two types of citizen observatories are used, semi-professional stations (TU Delft) and citizen weather stations (Netatmo), that are located near the city of Rotterdam. In total there are 9 TU Delft stations and 73 Netatmo stations in a region of 256.7 km^2. Based on a spatial variance analysis between these two types of citizen observatories and a professional station (KNMI) in the region, weighing factors are determined in order to merge and interpolate the data from the citizen observatories into a single rainfall map. This interpolated map is then compared to the radar rainfall maps, which are bias-corrected based on a network of professional ground stations (Overeem et al., 2016), and provide rainfall data with a 1x1 km spatial resolution. The baseline for improvement comes from the assumption that the rainfall measured by the KNMI station is uniform over the entire study area. By comparing the differences between KNMI and radar and citizen observatories and radar, it can be assessed whether the citizen observatories are an improvement.

The results show that the interpolated rainfall maps are better at capturing the spatial structure of the rainfall event, both on the small-scale (10-minute timestep) and event-scale (total accumulation). When it comes to the actual rainfall values, it became clear that rainfall intensities within a 4 km radius from the KNMI station were better represented by the KNMI station while after a distance of 8 km, the citizen stations clearly helped give a better representation.

Nomenclature

List of symbols

Symbol	Unit	Definition
$R_i(X_{KNMI})$	mm	the 10-minute rainfall value at time t_i from the KNMI station
$\mathrm{R_i}(X_h)$	mm	the 10-minute rainfall value at time t_i from a citizen station.
X_h	x, y	location of a citizen station
X_{KNMI}	x, y	location of the KNMI station
γ	mm²	semivariance of a citizen station
N	-	number of datapoints per pair
$R(x_0)$	mm	the 10-minute rainfall value to be determined in a pixel
$N, M, \dots, O$	-	number of stations per type of citizen station
$a, b, \dots, z$	-	type of citizen station
$d_{i,j,\dots,k_{a,b,\dots z}}$	m	distance of citizen station to a pixel
$R(d_{i,j,\dots,k_{a,b,\dots,z}})$	mm	the 10-minute rainfall value of a citizen station
$\lambda_{i,j,\dots k_{a,b,\dots,z}}$	mm^{-2}	weight based on distance from pixel
n+α	mm²	sill
α	mm²	partial sill
n	mm²	nugget
β	m	pseudo-range
d	m	distance to KNMI station
RMSD	mm²	Root Mean Squared Difference
$\widehat{y_i}$	mm	mean-value of the difference per pixel
y_i	mm	difference per pixel
n	-	number of observations
p	-	Pearson-correlation
$cov(x, y)$	mm²	covariance between variables
$\sigma_{x,y}$	mm²	standard deviation of variables
O_i	mm²	observed semivariance
M_i	mm²	modelled semivariance
N	-	number of citizen stations

List of figures

List of tables

Table C6	Median of the absolute difference when radar pixels either measure rainfall or they don't, using the faulty zero filter by Hutten et al. (2018) and the manual quality control.
Table C7	Median of the RMSD and Pearson-correlation for the four events, using only the manual quality control
Table C8	Median of the RMSD and Pearson-correlation for the four events, using the faulty zero filter by Hutten et al. (2018) and the manual quality control

1. Introduction

In 2018, 55% of the world's population lived in urban areas. This is expected to increase to 68% by 2050 (UN, 2018). Because of high levels of imperviousness compared to natural systems, response times tend to be shorter with higher run-off ratios and consequently higher peak flows (Veldhuis et al., 2017). Pluvial flooding occurs when the runoff converted rainfall exceeds the capacity of the sewer or stormwater system (Houston et al., 2011). If the system is a combined sewer system, this can lead to polluted sewer water to flood the streets, which poses a health hazard. In the case there is a combined sewer overflow (CSO), polluted water will enter the surface water system, which can lead to death of the marine life. During the heavy rain event of May 31st, 2018 in Rotterdam, the rainfall rate exceeded the capacity of the combined sewer system, leading to CSO and reports were made of sightings of massive loss of fish in the canals (AD, 2018). In addition, extreme rainfall events which lead to flooding can result in damage to ecology, infrastructure, disruption to human activities, injuries and in the worst-case scenario, loss of life (Biniyam et al., 2017). According to the IPCC (2012), extreme weather events are expected to increase in frequency, intensity and duration for many regions in the future. To prevent or mitigate future risks, there is a need for dense, high-resolution rainfall measurements in highly populated regions (Muller et al., 2015), that can improve early warning systems and disaster management (Restrepo-Estrada et al., 2017) as well as flood and urban drainage management (Muller et al., 2015). This means that it is important to have accurate and timely surface precipitation data (Brouwer et al., 2017). Lopez et al. (2005) mentions that one of the key factors in hydrological models to determine accurate flood estimates is to have accurate rainfall input. However, many urban areas might lack this information because sensors are not available, or the number of sensors is too few to cover the entire region with an acceptable resolution (Restrepo-Estrada et al., 2017). Kidd et al. (2017) mentions that the density of rain gauges varies per region, with Europe and Eastern-Asia (including Japan) having a decent coverage by rain gauges, while the rest of the world has a sparse coverage of rain gauges. Citizen science or community-based monitoring may offer a solution to fill in some of the gaps in data (Paul et al., 2018). Citizen science is often defined as the participation of the general public in the research design, data collection and interpretation process together with scientists (Buytaert et al., 2014; Paul et al., 2018). A study by Eilander et al. (2016), where social media like twitter and Facebook were used to follow real-time flood events, the limiting factor was the number of social-media users. With the increase of population in urban regions (UN, 2018), these might offer good testing grounds for pilot projects involving citizen science.

This research focusses on various crowdsourced rainfall data, and how their combined use can give better insight on the spatial variability of rainfall in urban areas compared to having just a single high-quality weather station. This rainfall product will then be compared to radar images, since radars can obtain data with a high temporal and spatial resolution (Lopez et al., 2005). The main research question therefore is: Can citizen observatories be used to improve information on the spatial variability of rainfall? To answer this question, a few sub-questions need to be answered. Which citizen observatories stations should be used? This question is related to the availability of the data from the stations, in addition to a manual quality control on the data. How do we quantify uncertainty of the citizen observatories? Section 2 will explain how radar data will be used to assess the data from the citizen observatories. Since the rainfall product will be compared to radar, it needs to be known how citizen observatories directly compare the radar. How can various data sources be merged into a single rainfall product? Each

type of rainfall station has its own quality and uncertainty. This needs to be accounted for when creating merged products. And lastly, how does the structure of the rainfall event relate to the added value of a network of citizen observatories?

1.1 Context of the study: FloodCitiSense

This research is done as part of the FloodCitiSense project. FloodCitiSense is a project involving many research organisations, a few of which are the Vrije Universiteit Brussel (VUB), Delft University of Technology (TUD), Imperial College London and the International Institute for Applied Systems Analysis, Ecosystems Services and Management Program (IIASA). The aim of the project is the development of an urban pluvial flood early warning service for and by citizens and city authorities. Citizens will no longer be seen as potential "victims" of floods but are actively engaged in the monitoring and mapping of urban pluvial floods. Figure 1.1 shows the concept of the FloodCitiSense project. In this project, citizens are involved in measuring rainfall data and making reports of pluvial flooding because of rainfall. By using the citizen data, a model can be made to determine a threshold for pluvial floods and a more location-precise early warning system can be created.

The FloodCitiSense project is in its pilot phase and will be tested in three cities, Birmingham, Brussels and Rotterdam. This research focuses on the city of Rotterdam.

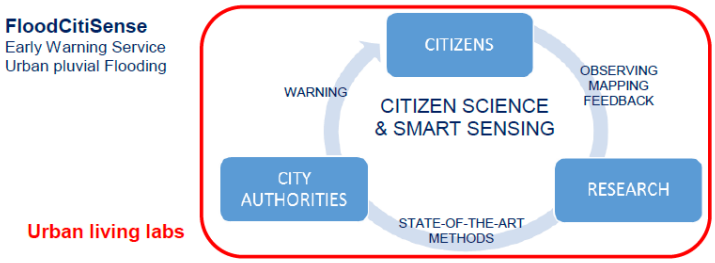


Figure 1.1: Concept of FloodCitiSense. (source:ENSUF-full-proposal-form_FloodCitiSense)

1.1.1 Engagement strategies

One of the challenges in citizens science is the number of citizens participating (Eilander et al., 2016). What is the correct way to engage these citizens? According to Wehn et al. (2015), three aspects need to be considered. Firstly, the goals of the involvement. Secondly, the participants themselves. Who are they (Fung, 2006)? What might be their reasons for joining (Goodchild, 2007)? Lastly, the means through which they are invited. They should be considered active coordinators during an incident instead of victims. Buytaert et al. (2014) gives even more insight in reasons for participating, some of which are self-driven environmental concern, scientific curiosity or local stakeholders. Critically, the participants should feel valued and not patronized. Starkey et al. (2017) explains the means with which they reached the people. A workshop was given for the local community and key partners (i.e. land owners and residents). Reaching participants from a wider range was done by social media, newspapers and leafleting.

In the case of FloodCitiSense, multiple workshops were organized to explain the project, the goals and the value of the citizens. In the latest workshop, participants were able to build their own sensor, to show the ease with which rainfall sensors can be created. It will also act as a way for the participants to feel actively involved and hopefully feel more responsible for the sensor with regards to maintenance. Secondly, the workshop will explain the mobile app that has been created, which is a way in which the participants can give active feedback and reports. The app will also be one of the ways for the participants to receive updates from FloodCitiSense. Furthermore, a questionnaire was created for the citizens to give feedback to help and improve future workshops and to have a better understanding on how to maintain the interests of the participants for a long-term relation. Participants were initially gained by posting on the main employee website of the municipality of Rotterdam, which is one of the partners of the FloodCitiSense project. Employees of the urban management department were mailed additionally. In addition, flyers were posted on the Water Management department of the TUD.

1.2 Guidelines for reading

In section 2, the study area will be introduced as well as the data and methodology. Section 3 will show results and findings. Section 4 will discuss the limitations of the study and possible future improvements. Section 5 will conclude with a short summary of the study and the most important findings.

2. Data & Methods

This section will introduce the study area, the data used to conduct the study, the software and the methods.

2.1 Study Area

The study area is influenced the FloodCitiSense project, focussing on the city of Rotterdam. The boundaries of the study area are chosen such, that the radar image covers the entire city. The coordinates of the study area are given in table 1.

	Coordinates (EPSG 28992: Amersfoort)
Northern Boundary	446500
Southern Boundary	430500
Eastern Boundary	85400
Western Boundary	101500

Table 1: Boundary coordinates of study area

The total surface area of the study region is 257.6 km². The total area of the city of Rotterdam is 125.3 km². The municipality of Rotterdam is working hard to make the city climate adaptive (RCI 2008). In order to reach this goal, they have 6 separate subjects, one of which is '*Flooding due to extreme weather events*.' The city of Rotterdam is located in the delta area where the rivers Meuse, Rhine and Waal reach the North-Sea. During extreme rainfall intensities the city experiences street floods, flooding of the basements and combined sewer overflow. Whenever flooding occurs, citizens can make a report to the municipality. A clear relation can be seen between reports and the imperviousness of an area in the city of Rotterdam. If an area has an imperviousness of 70% or more, the number of reports increases significantly (Bouwens et al., 2018). Having a better understanding of the spatial variability of rainfall events can help predict whether a certain area or district will be flooded, and an early-warning system can be created to increase preparedness as well as help post-flood analysis. In the case of property damage, it can also help the citizens with their insurance claims.

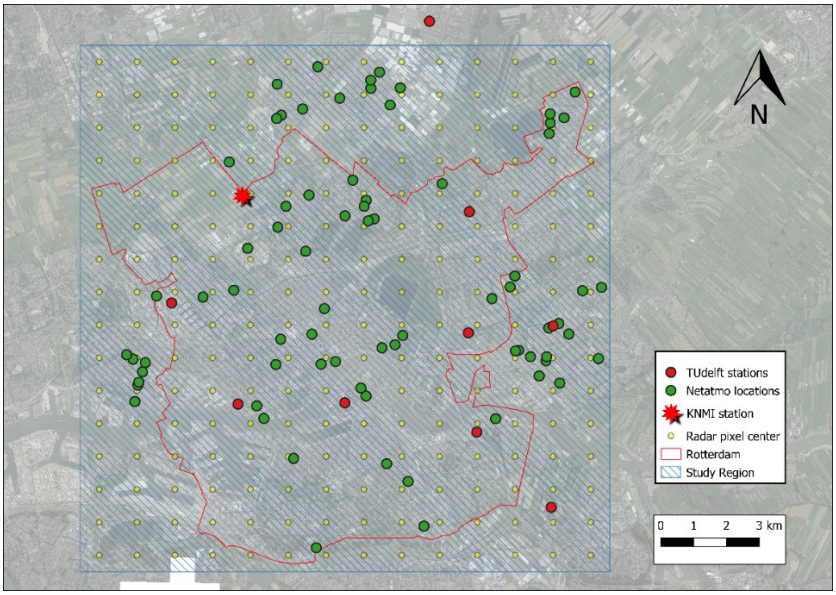


Figure 2. 1: Map of study area with locations of the citizen observatories used in this research

2.2 Data

In this section the different data types are mentioned. The characteristics of the data and why the data is used. Table 3 gives an overview of the data.

2.2.1 Netatmo weather stations

Netatmo manufactures low-cost citizens weather stations (source: https://www.netatmo.com/en-gb/weather/weatherstation/accessories#raingauge) . According to the product specifications provided by Netatmo, the rain gauges have a measuring range of 0.2 to 150 mmh^{-1} with an accuracy of 1 mmh^{-1}. The rainfall gauge is a tipping bucket with a volume of 0.101 mm. The funnel has a diameter of 130 millimetre. The rain gauge sends the data wirelessly to the indoor station that can be up to 100 meters away from the station. (source: https://www.netatmo.com/en-US/product/weather/weatherstation/specifications, *specifications smart rain gauges → Sensors and measurements*)

The Netatmo rainfall data is used as part of the citizen observatory data to verify if these stations can be used to gain better insight on the information of the spatial variability of rainfall.

Figure 2.2: Netatmo rain gauge (source: https://res.cloudinary.com/dxyci1a78/image/upload/c_scale,dpr_1.25,f_auto,q_auto/v1/weather/accessories/landing)

2.2.2 TU Delft weather stations

TU Delft is involved with the RainGain project (which is over at this point in time) and the climate institute. From this project, there are 15 stations located in and around Rotterdam. These stations are maintained about once every 2 months, or if it becomes clear there are big discrepancies between them. Unfortunately, six of the stations are no longer active but their historical data is still available. The rainfall gauge is a tipping bucket with a volume of 0.2 mm. The rain gauge sends the data wirelessly to the database every minute. (source: http://www.raingain.eu/sites/default/files/fs1_tech_spaansepolder_0.pdf)

According to the manufacturer (Campbell Scientific), the gauges have a measuring range of 0 to 700 mmh^{-1} with an accuracy of 2% if the intensity is less than 250 mmh^{-1} and an accuracy of 3% if the intensity if larger than 250 mmh^{-1}. The funnel as a diameter of 200 millimetres. (source: https://www.campbellsci.com/tb4mm)

The TU Delft rainfall data is used as part of the citizen observatory data to verify if these stations can be used to gain better insight on the information of the spatial variability of rainfall. The TU Delft stations are assumed to be of intermediate quality, between amateur and professional quality. However, they will be referred to as citizen stations, as they are part of the citizen observatories.

Figure 2. 3: TU Delft rain gauge. (source: https://weather.tudelft.nl/photos/Oost/2017-04-10_05_oost.jpg)

2.2.3 FloodCitiSense Intervalometers

From the FloodCitiSense project, intervalometers provided by Disdrometrics are given to the participants. During the workshop, 16 intervalometers were distributed. These rainfall sensors are placed and maintained by the citizens. Once a month, the citizens will upload the data to the database. The intervalometers measures the time between raindrops that fall on the sensor to estimate the rainfall intensity (Giesen et al. 2017). This means that that when it rains, the intervalometer measures the rainfall intensity with a temporal resolution in the order of milliseconds. The temporal resolution of the intervalometer can be set by the user (ie. minute, hourly, daily, etc.).

The intervalometer rainfall data is used as part of the citizen observatory data to verify if these stations can be used to gain better insight on the information of the spatial variability of rainfall. Unfortunately, in the first month after the distribution of the sensors, the Netherlands experienced a drought combined with high temperatures. During the rainfall events that followed, reports from the participants indicated that the glue had melted due to the heat and caused the sensor to no longer be waterproof, as well as the internal battery to expand in some cases. This caused sensors to stop functioning. Therefore, not enough data could be collected and unfortunately the intervalometers could not be used in this research.

Figure 2. 4: FloodCitiSense intervalometer

2.2.4 KNMI weather station

The KNMI weather station is located near the Rotterdam-The Hague airport. The KNMI station is an automated gauge and is maintained every day. It measures rainfall by means of a heighted weighing bucket and can therefore measure both liquid and solid precipitation. The KNMI has 4 guidelines when installing their weather stations. 1) Within a radius of 25 meters, no crops or vegetation higher than 0.5 meters is allowed, 2) within a radius of 50 meters, no crops or vegetation is allowed higher than 1.5 meters, 3) within a radius of 100 meters, no obstacles such as trees or bushes are allowed and 4) within a radius of 400 meters, no obstacles such as barns, buildings and forests are allowed (source: http://projects.knmi.nl/hawa/pdf/Handboek_H01.pdf). The open source rainfall data has an hourly temporal resolution. However, employees have access to unvalidated 10-minute resolution data. Special thanks to Hutten et al. (2018) for providing this data.

The KNMI weather station will be used as the baseline for this research, which will be explained in section 2.3, the methodology.

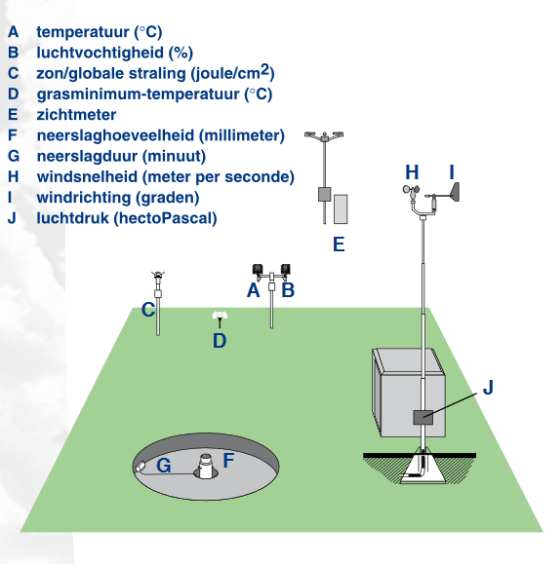


Figure 2. 5: Sketch of an automated weather station. F) measures rainfall depth and G) measures rainfall duration. (source: http://projects.knmi.nl/hawa/pdf/Handboek_H01.pdf)

2.2.5 Radar

Meteorological radar has the advantage of being able to cover a wide area with a high temporal and spatial resolution (Lopez et al., 2005). Radar samples precipitation by transmitting an electromagnetic wave, which interacts with various objects in the atmosphere (usually raindrops or other forms of precipitation) and measuring the backscatter. Unfortunately, the rainfall estimated by radar can be affected by various reasons: ground clutters caused by topographic interference (Espinosa et al., 2005) or errors in reflectivity-rainfall (Z-R) relationships (AghaKouchak et al., 2010). Also, systematic errors due to calibration can occur (Piccolo et al., 2005).

The radar data will be retrieved from the 'Nationale Regenradar'. It is a combined radar product using Dutch, German and Belgium radars. It is corrected for bias using all known reliable ground stations to give rainfall data with a 5-minute temporal resolution and a 1x1 km spatial resolution (Nationale Regenradar, 2013).

The radar data will be used as validation for the spatial variability of the citizen stations.

2.2.6 Software

This study makes use of the programming language Python 3.6.5 and Quantum GIS (QGIS). QGIS is mainly used for visualization of the rainfall maps and any subsequent maps used for analysis as well as some basic distance calculations. Python 3.6.5 is used to download the radar data from the Nationale Regenradar using an API provided by Nelen&Schuurmans. The API downloads the radar with a 1x1.15 km spatial resolution instead of 1x1 km, which does not influence this research. Furthermore, Python 3.6.5 is used for programming, making use of the following packages: *pandas, numpy, matplotlib.pyplot, os, scipy, pyproj*.

2.2.7 Selected rainfall events

Four rainfall events are manually selected to test the main research question. The events are selected based on high rainfall intensities. Table 2 gives and overview of the events. Because the KNMI data and the Netatmo data range from October 2015 to October 2017, only events from this period will be taken. In appendix B, for each event, the data from the stations and radar is plotted. The intensities of the radar, TU Delft stations and Netatmo stations are aggregated to a 10-minute resolution, as this is the resolution of the KNMI station.

Event Number	Start event (yyyy/mm/dd hh:mm)	End event (yyyy/mm/dd hh:mm)	Duration event
1	2015-11-29 18:00	2015-11-29 20:50	3 hours
2	2016-06-22 23:20	2016-06-23 01:50	2.5 hours
3	2016-11-17 15:00	2016-11-17 22:50	8 hours
4	2017-07-25 05:00	2017-07-25 07:50	3 hours

Table 2: Duration and dates of the four chosen events to be used to test the merging method

Furthermore, the radar data retrieved from the Nationale Regenradar presented some issues. One of the issues resulted in some of the radar pixels having an infinite rainfall depth, as can be seen in figure 2.6. Before the Nationale Regenradar data is made public, some filters are applied. It is unclear what the purpose of some of these filters is, as there is no documentation on them, but it caused the radar pixels to have impossibly high values. In finding the four rainfall events, it was important that these pixels were not present.

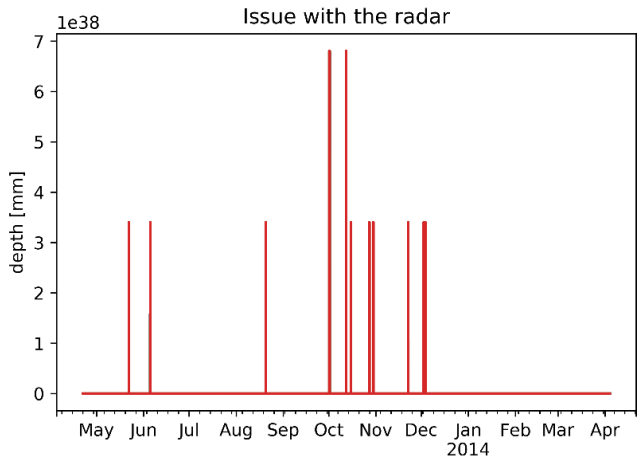


Figure 2.6: Nationale Regenradar showcasing the issue with some radar pixels giving impossibly high values

Another issue with the radar, can be seen in Appendix B, figure B4d. There is one pixel that does not follow the same trend as the other pixels, giving strange rainfall values. Looking at the location of this pixel, it is right above the Waalhaven, which is a location containing radio-towers which can disrupt the radar signal. This might therefore happen to all radar images and not just the Nationale Regenradar.

Rainfall datasets				
Name	**Description**	**Period**	**Application**	**Source**
Netatmo stations	Citizen weather stations from Netatmo, providing rainfall data with a 5-minute resolution. A total of 73 stations.	1 October 2015 to 1 October 2017	Input for the merged rainfall product.	Special thanks to R. Hutten, for providing the raw Netatmo data, which was used in her own research
TU Delft stations	Weather stations set up as part of the RainGain project in the city of Rotterdam, providing rainfall data with a 5-minute resolution. A total of 9 stations.	23 April 2013 to 30 April 2018	Input for the merged rainfall product.	http://weather.tudelft.nl/csv/
KNMI weather station	Automated weather station from the KNMI in the city of Rotterdam, giving validated rainfall data with a 10-minute resolution.	1 October 2015 to 1 October 2017	Baseline to evaluate if the merged rainfall product improves information on the spatial variability as well as 'truth' to calculate the cross-variogram.	Special thanks to R. Hutten, for providing the validated KNMI data, which was used in her own research.
Nationale Regenradar	Radar image created by a joint effort from the Dutch, German and Belgium radar and corrected for bias using known reliable ground stations. Spatial and temporal resolution of 1x1.15 km and 5-minute respectively.	23 April 2013 to 30 April 2018	The radar is used to assess whether the merged rainfall product improved information on spatial variability.	https://demo.lizard.net/api/v3/rasters/730d6675-35dd-4a35-aa9b-bfb8155f9ca7/data/?\

Table 3: Overview of data sources

2.3 Methods

In order to find out whether citizen stations can improve spatial rainfall estimates, it is important to set a baseline against which the citizen stations can be compared. This baseline is the KNMI weather station in the north-west of the study area. The rainfall measured by the KNMI station is assumed uniform over the study area. Using the radar, the difference per radar pixel can be determined. The next step is to use the citizen stations to create a merged rainfall map and compare this with the radar. Comparing both difference-maps allows to assess the value of having a dense network of citizen stations as opposed to having a single high-quality station.

Usually, before the data from the citizen stations can be used, it is important to apply some basic form of quality control. In this study, the data will undergo a manual check, to remove malfunctioning stations and obviously wrong data. They will not be subjected to any automatic filters or calibration, as the objective is to verify whether they can be used as is.

The criteria for removal of stations or data points during the manual check will be large data gaps, long periods with zero-measurements while close by stations do measure rainfall and impossibly high rainfall values.

2.3.1 Cross-variogram

The next step is to merge the data together to a single rainfall product. Since this study is making use of different types of citizen stations, each with their own quality and uncertainty, this uncertainty needs to be quantified before the data can be merged. A cross-variogram (eq.1) between the KNMI station and citizen stations is chosen as the means to quantify the uncertainty and quality. The main idea behind a variogram is that nearby values tend to be more similar than values that are further apart. Therefore, there should be a decreasing correlation as a function of distance. By calculating the cross-variogram, a relation between distance and dissimilarity can be determined for each of the network of citizen stations. This assumes that rainfall data from the KNMI weather station is the 'truth' and free of measurement errors.

$$\gamma(\|X_h - X_{KNMI}\|) = \frac{1}{2N} * \sum_{i=1}^{N}[R_i(X_h) - R_i(X_{KNMI})]^2 \quad [\text{mm}^2] \qquad (1)$$

$R_i(X_{KNMI})$	[mm]	the 10-minute rainfall value at time t_i from the KNMI station
$R_i(X_h)$	[mm]	the 10-minute rainfall value at time t_i from a citizen station.
X_h	[x, y]	location of a citizen station
X_{KNMI}	[x, y]	location of the KNMI station
γ	[mm²]	semivariance of a citizen station
N	[-]	number of datapoints per pair

Note that this is not a true cross-variogram, in the strict sense that it does not have two moving variables. The location of the KNMI station is constant, therefore only the location of the citizen stations is variable.

The variogram can be characterised by three variables: sill, range and nugget. After a certain distance, a variogram levels out. This distance is called the range. After this distance, points are no longer spatially correlated. The value at which the variogram decorrelates is called the sill.

The nugget is the discontinuity at distance zero. If two identical rain gauges are located at the same location, they should have similar measurements, giving a nugget of zero. In the case of this research, a professional weather station (KNMI) is compared to a citizen station (Netatmo & TU Delft), so a nugget is to be expected.

Having quantified the uncertainty and quality of each type of citizen station, the method to merge the data to a single rainfall product is shown in equation 2. The intensity in the pixel is calculated by equation 2.

$$R(x_0) = \frac{\sum_{i=1}^{N} \lambda_{i_a} * R_a(d_{i_a}) + \sum_{j=1}^{M} \lambda_{j_b} * R_b\left(d_{j_b}\right) + \cdots + \sum_{k=1}^{O} \lambda_{k_z} * R_z(d_{k_z})}{\sum(\lambda_a) + \sum(\lambda_b) + \cdots + \sum(\lambda_z)} \qquad [\text{mm}] \qquad (2)$$

$R(x_0)$	[mm]	the 10-minute rainfall value to be determined in a pixel.
$N, M, \ldots, O$	[-]	number of stations per type of citizen station.
$a, b, \ldots, z$	[-]	type of citizen station.
$d_{i,j,\ldots,k_{a,b,\ldots z}}$	[m]	distance of citizen station to a pixel
$R(d_{i,j,\ldots,k_{a,b,\ldots,z}})$	[mm]	the 10-minute rainfall value of a citizen station.
$\lambda_{i,j,\ldots k_{a,b,\ldots,z}}$	[mm^{-2}]	weight based on distance from pixel.

The weights $\lambda_{i,j,\ldots k_{a,b,\ldots,z}}$ are based on the cross-variogram and will be approximated by $\lambda \propto \frac{1}{\gamma}$. Where γ is dependent on the distance of the pixel to the citizen station. Being further away from the station results in a higher γ and therefore a lower weighing factor. In this manner, stations closer to the pixel have a higher contribution in calculating the rainfall. The cross-variogram is calculated as a function of distance between the KNMI station and the citizen stations, which means that in order to calculate the weight for a citizen station to a pixel, a model is required that describes the cross-variograms. In order to find this model, first the variogram needs to be calculated and second a model needs to be fitted.

Possible models, among others, that could fit a variogram are the Gaussian model, the Spherical model, the Linear model and the Exponential model. For simplicity, an Exponential model (eq. 3) will be used to fit the cross-variograms.

$$\gamma = n + \alpha * (1 - e^{-3 * \frac{d}{\beta}}) \qquad [\text{mm}^2] \qquad (3)$$

n+α	[mm²]	sill
α	[mm²]	partial sill
n	[mm²]	nugget
β	[m]	pseudo-range
d	[m]	distance to KNMI station

The model will be fitted to the cross-variograms using the non-linear, weighted least-square method. However, each pair of citizen-KNMI station has a different number of datapoints (fig. 2.7 and fig. 2.8) when calculating the semivariance. This difference in number of datapoints is added as a weight to the fit in the form of $N_i^{\frac{1}{2}}$, in which N_i is the number of datapoints per pair.

This will mean that semivariance values that are calculated with more datapoints, will have a higher impact on the fit.

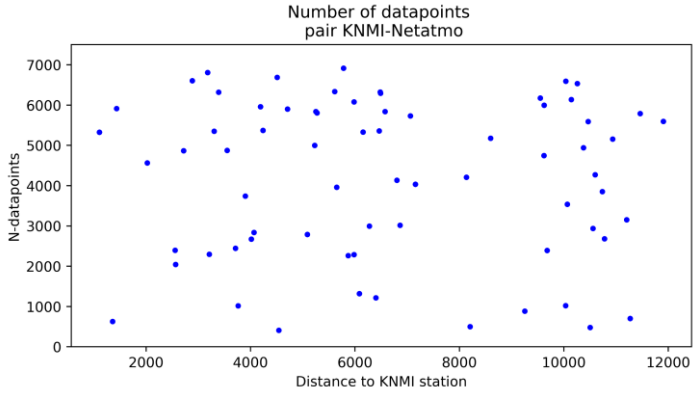


Figure 2.7: N-datapoints > 0, pair KNMI-Netatmo

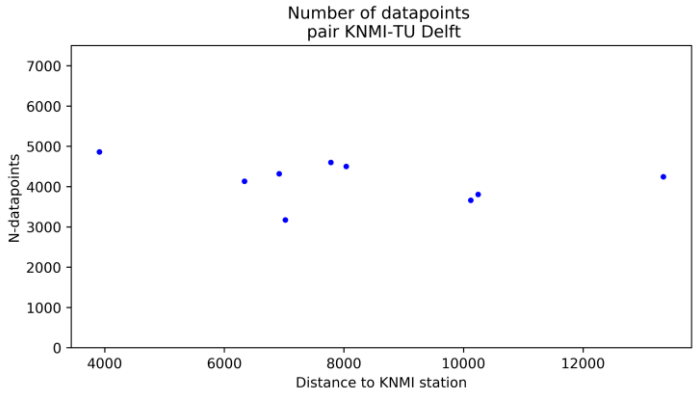


Figure 2.8: N-datapoints > 0, pair KNMI-TU Delft

2.3.2 Assessment of rainfall maps

The rainfall map of the study area based on eq. 2 will be assessed using two methods. Firstly, the RMSD-method (eq. 4). Secondly, the Pearson-correlation (eq. 5).

2.3.2.1 RMSD-method

The RMSD is defined as 'the square root of the mean of the squared differences between corresponding elements of the forecasts and observations' (Barnston, 1992). In this research, the radar rainfall data are the observations and the rainfall maps based on the citizen stations are the forecasts.

$$RMSD = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \quad [\text{mm}] \qquad (4)$$

$\hat{y}_i$	[mm]	mean-value of the difference per pixel
y_i	[mm]	difference per pixel
n	[-]	number of observations

The RMSD-method will be applied in two ways. Firstly, the RMSD will be calculated over an entire rainfall event. These rainfall events are selected manually, see section 2.2.7. By looking at an entire event, a RMSD can be determined for each pixel in the study area. Secondly, the same events will be used to determine the RMSD per timestep within the event. This will not give value per pixel but rather give a single value for the entire map. Per timestep it will become clear which rainfall map produced a better image of the rainfall.

2.3.2.2 Pearson-correlation

The Pearson-correlation p is a measure of the strength of the linear relationship between two variables. The value of p ranges from -1 to +1, with -1 meaning a perfect negative linear correlation and +1 meaning a perfect positive linear correlation. If the correlation is zero, it means there is no linear correlation.

$$p = \frac{cov(x,y)}{\sigma_x \sigma_y} \quad [-] \qquad (5)$$

$cov(x,y)$	[mm²]	covariance between variables
$\sigma_{x,y}$	[mm²]	standard deviation of variables

We expect there to be a positive correlation, since the radar data and the rainfall maps based on the ground stations are estimates of the same process. Table 4 shows the guideline to determine whether two datasets have a strong correlation.

	Coefficient, p	
Strength of Association	Positive	Negative
Small	0.1 to 0.3	-0.1 to -0.3
Medium	0.3 to 0.6	-0.3 to -0.6
Large	0.6 to 1.0	-0.6 to -1.0

Table 4: Guideline to determine strength of correlation

Like the RMSD-method, the Pearson-correlation will be taken over an entire rainfall event duration, to show the correlation of each pixel, and the individual timesteps to get a correlation for the entire map.

3. Results

3.1 Manual Quality Control

This research will try to use the raw data as much as possible to verify whether they can be used as is, without having to go through many filters. The data will undergo a visual check, to remove faulty stations or faulty data as explained in section 2.3.

Based on the visual check, we can identify some stations that have obviously wrong values. (example. fig 3.1 and 3.2). It seems that the Netatmo station is not measuring any rainfall, while the TU Delft station has a few outliers. The outliers from the TU Delft station are set to not a number, while the faulty Netatmo station is removed completely.

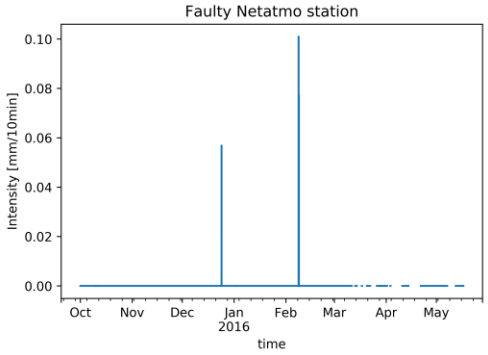


Figure 3.1: Faulty Netatmo station, no rainfall measured

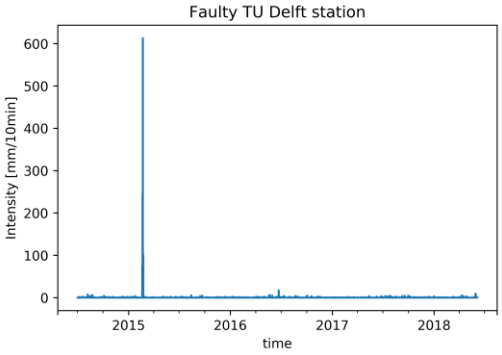


Figure 3.2: Faulty TUdelft station, outliers in rainfall data

In total there were two TU Delft stations (Ridderkerk and Oost) that were adjusted and three Netatmo stations (id354, id385 and id485) that were removed. See appendix A, table A1, for the overview and reasons for the adjustment of the stations.

3.2 Empirical variograms

To merge the data sources to a single output, the difference in the quality of stations needs to be quantified. The cross-variogram is created by comparing the citizen stations to the KNMI station. There are nine TU Delft stations, while there are seventy Netatmo stations after quality control. Research by Hutten et al. (2018) concluded that citizen stations contain faulty zero measurements. The research created filters to flag the faulty zeros, to show the reliability of the stations. Therefore, in comparing the citizen stations to the KNMI station, only non-zero values are considered, to remove the possibility of faulty zeros.

Figure 3.3 shows the Netatmo:KNMI cross-variogram. Each point represents the semivariance between a single Netatmo station and the KNMI station, based on the distance to the KNMI station. Immediately it is clear that there is a lot of noise. This can be explained by the number of datapoints between the pair. Most of the noise comes from pairs that have a low amount of datapoints. Furthermore, the Netatmo station that is closest to the KNMI station, does not have

the lowest semivariance. This is unexpected, as stations closer to each other are assumed to correlate better. However, this again stresses the importance of the quality of the dataset.

Additionally, there do seem to be two points with high semivariance values compared to the rest. No clear reason can be found as to why these two stations give such high semivariance values, as their location and rainfall values don't show anything out of the ordinary.

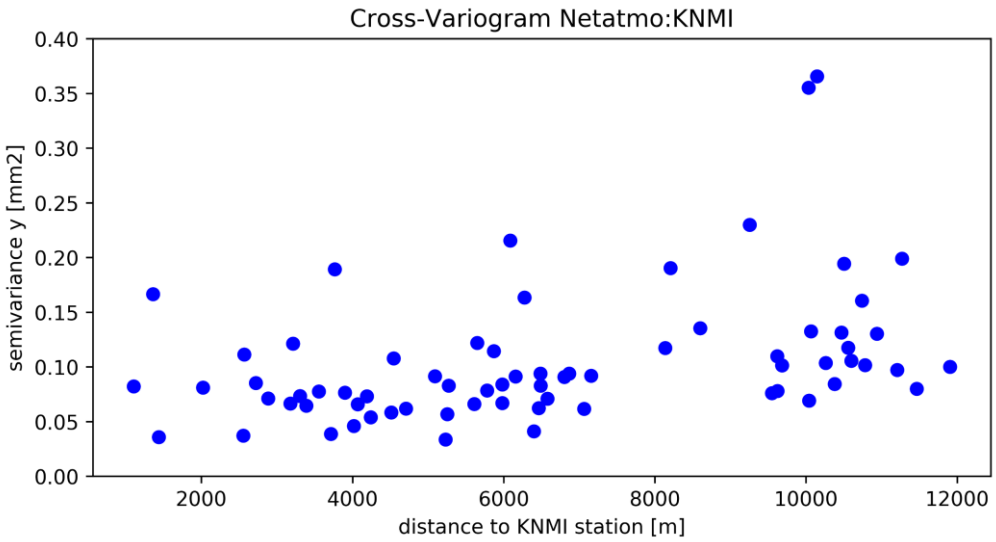


Figure 3.3: Cross-variogram Netatmo to KNMI. 10-minute temporal resolution.

Figure 3.4 shows the TU Delft:KNMI cross-variogram. In addition to having much fewer points, the stations are also far away from the KNMI station. Therefore, it seems that these points are already in the levelled-out part of the variogram. Additionally, this variogram also has the issue that the lowest semivariance value does not come from the TU Delft station that is closest to the KNMI station.

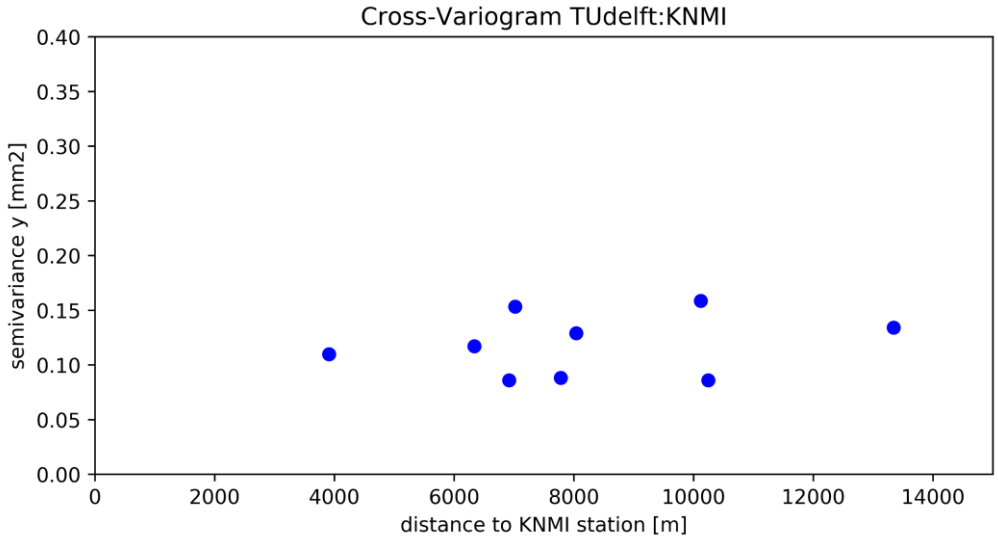


Figure 3.4: Cross-variogram TUdelft to KNMI. 10-minute temporal resolution

3.3 Fitting the model to the variograms

In order to fit the model (eq. 3) to the variograms, the three variables (sill, range and nugget) need to be considered. By setting boundaries for the variables, the fit will become easier. Without setting boundaries, the fit is poor, as shown for the Netatmo stations in figure 3.5.

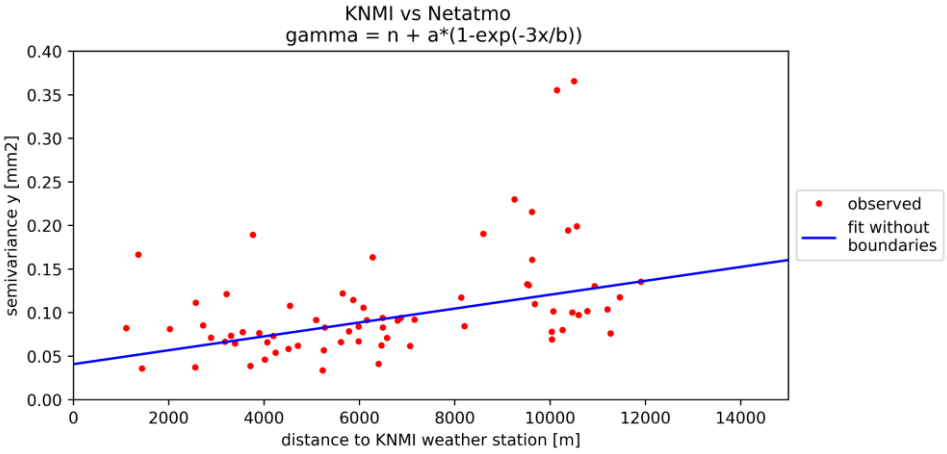


Figure 3.5: Netatmo-KNMI cross-variogram, fit without boundaries giving a bad fit

The sill is in the order of magnitude of 600 mm² with a range of 250,000 km and the nugget has a value that is larger than the minimum semivariance of the cross-variogram. Considering the cross-variogram, there are issues with all three of the parameters. The sill and range are much too high in comparison to the variogram, while the nugget is too large. Per definition, the nugget cannot be higher than the minimum semivariance value, which would mean that a station at distance x > 0, has a higher spatial correlation than a station at distance x = 0. It should be clear that at distance x = 0, the spatial correlation should be the highest and therefore the semivariance should be the lowest.

By imposing boundaries on the variables, specifically on the nugget and sill, the fit can be improved. First, the Netatmo:KNMI cross-variogram will be fitted as there are much more stations with a better spread over distance.

3.3.1 Netatmo:KNMI cross-variogram, imposing boundaries on the nugget

Initially, a boundary was imposed on the nugget, since the nugget has clear boundary conditions, namely, {0 < nugget < minimum semivariance}. To test the fit based on the nuggets, a goodness of fit test was done by minimizing the median of the absolute difference between the model and the observed data (eq. 6).

$$GoF = \min[median\{|O_i - M_i|\}]\, with\ i = [1, 2, \ldots, N] \quad [\text{mm}^2] \qquad (6)$$

O_i	[mm²]	Observed semivariance
M_i	[mm²]	Modelled semivariance
N	[-]	Number of citizen stations

The goodness of fit (eq. 6) resulted in the nugget going towards the minimum semivariance value, resulting in a value of 94 mm² for the sill and 32,000 km for the range. These values were

once again unrealistically large. Zimmerman et al. (2008) also dealt with the issue of identifying the proper value for the nugget for semivariogram (so not a cross-variogram as is the case in this research). Their solution was to impose the nugget as the semivariance value at the shortest distance, which for the case of a semivariogram will be the lowest value for the semivariance. Unfortunately, the semivariance at the shortest distance with the cross-variograms in this research are not the lowest values, which means they cannot be used as the nugget.

3.3.2 Netatmo:KNMI cross-variogram, imposing boundaries on the sill

Seeing that finding a fit by imposing the nugget is not giving the proper results, the sill will be used to fit the variogram. To find the boundary for the sill, it needs to be assumed that the variogram levels out within the boundaries of the cross-variogram. By looking at the variogram, even though there is quite some noise, a trend can be seen. The sill can be approached by taking the median of the levelled-out part of the variogram. By looking at the variogram (fig. 3.3), this is estimated by taking all the points at a distance x > 9 km. The fit is calculated by rewriting equation 3 (see eq. 7).

$$y = (sill - \alpha) + \alpha * (1 - e^{-3\frac{d}{\beta}}) \qquad [\text{mm}^2] \qquad (7)$$

, in which the sill is imposed as the median of the levelled-out part of the variogram.

This approach resulted in a sill of 0.12, a nugget of 0.02 mm² and a range of 15.9 km. The minimum semivariance of the Netatmo:KNMI variogram is 0.03 mm², so the nugget satisfies the boundary conditions. The range is also reasonable, as it stays within the range of the study area. The fitted model is shown in figure 3.6.

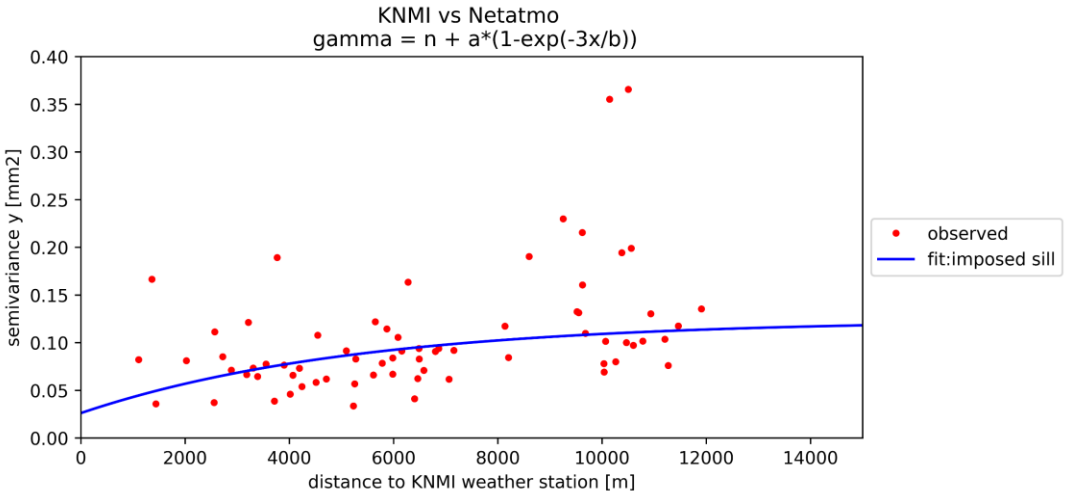


Figure 3.6: Exponential fit to the Netatmo:KNMI cross-variogram by imposing the sill

3.3.3 TU Delft:KNMI cross-variogram, imposing boundaries on the sill

The same method (eq. 7) is applied for the TU Delft:KNMI cross-variogram (fig 3.4). However, the variogram presents three issues: 1) the number of datapoints, 2) the spread over the distance and 3) the semivariance values. Considering the first issue, less datapoints means a lower precision fit. The second and third issue are related to each other, because there are no TU Delft

stations near the KNMI station and the variogram shows that the points are already in the almost levelled-out section of the variogram. The expectation is therefore that the best fit would be a straight line. The fitted model by imposing the sill is shown in figure 3.7.

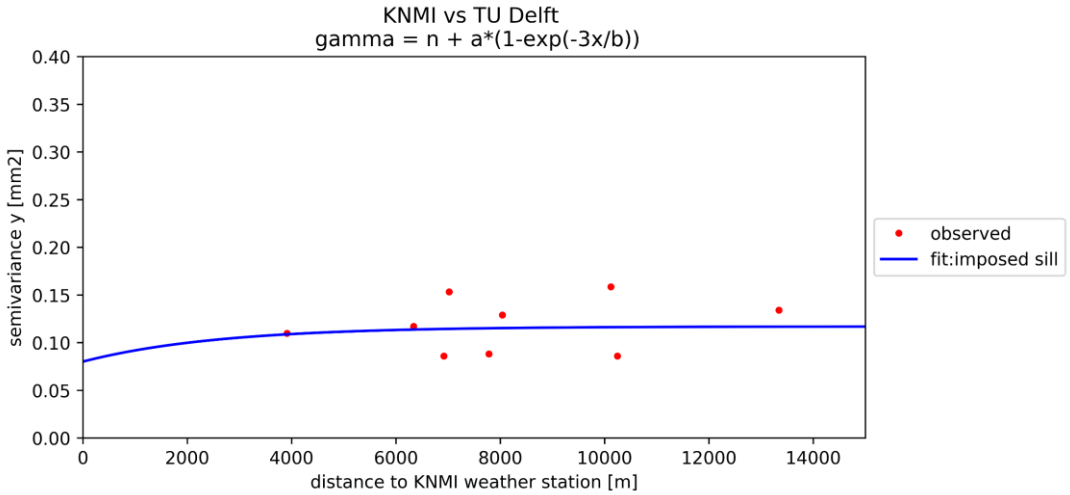


Figure 3.7: Exponential fit to the TU Delft:KNMI cross-variogram by imposing the sill

The value for the sill is 0.12 mm², the range is 7.8 km and the nugget 0.08 mm². This nugget value, which satisfies its boundary conditions of being lower than the minimum semivariance of the variogram (0.086 mm²), is four times larger than the nugget estimated in the Netatmo:KNMI cross-variogram. This means that if a Netatmo station and a TU Delft station would be placed next to each other, the rainfall data from the Netatmo station would be more reliable. The Netatmo stations are placed by citizens with little maintenance, while the TU Delft stations are maintained every two months or if there is a big discrepancy between the stations. A valid expectation is that the TU Delft stations are of a better or at least similar quality than the Netatmo stations. This is confirmed by comparing the stations directly to the radar data. Only pixels containing both TU Delft and Netatmo stations are used for comparison (fig. 3.8 and fig. 3.9). In this comparison, all zero-values are removed.

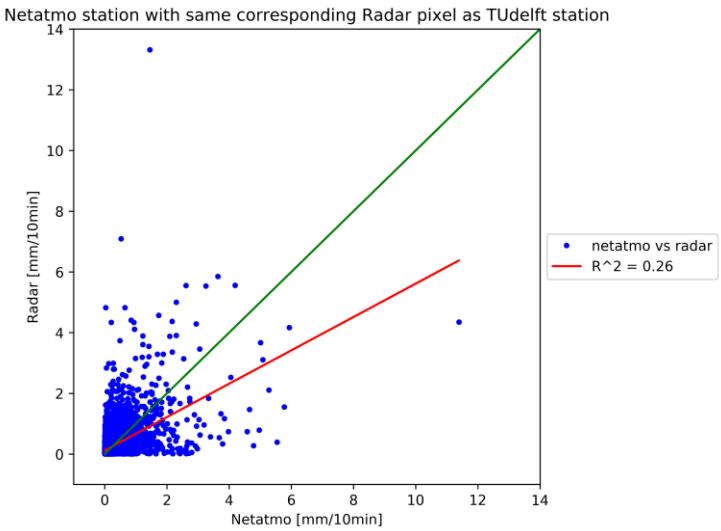


Figure 3.8: Netatmo data vs radar, only pixels containing both types of citizen stations

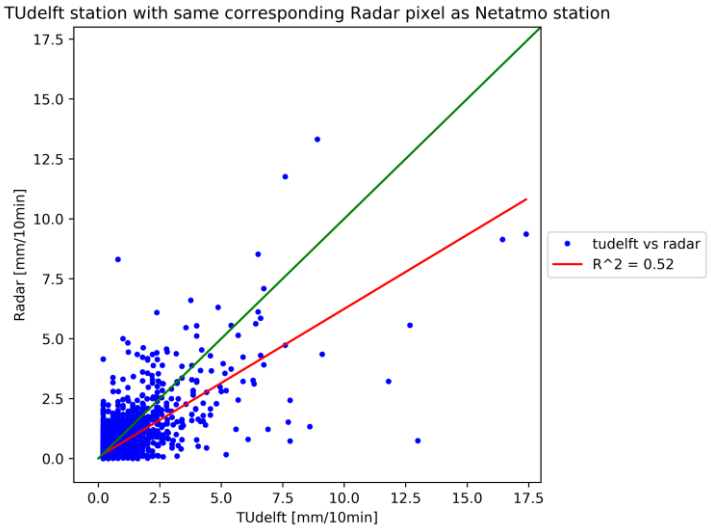


Figure 3.9: TU Delft data vs radar, only pixels containing both types of citizen stations

The TU Delft stations have a much higher correlation with the radar than the Netatmo stations. This validates the expectation that the TU Delft station are of a high quality than the Netatmo stations and therefore fitting the model by imposing the sill (fig. 3.7) cannot be used.

Interestingly, when comparing the KNMI to the radar, for the pixel containing the KNMI station, it gives an unexpected results (fig. 3.10). The radar-product is bias-corrected by ground stations, like the KNMI weather station which is used in this research. If the rainfall data from the KNMI station is compared to the radar, a high correlation is expected. However, it gives a lower correlation than the comparison between the TU Delft stations and radar (fig. 3.9). This clearly shows that there are large discrepancies between radar and ground stations because this doesn't mean that the KNMI station is of a lower quality than the TU Delft stations. It shows that there is no linear relation between radar and ground stations. Therefore, validating the quality of the ground stations based on this method might not be meaningful.

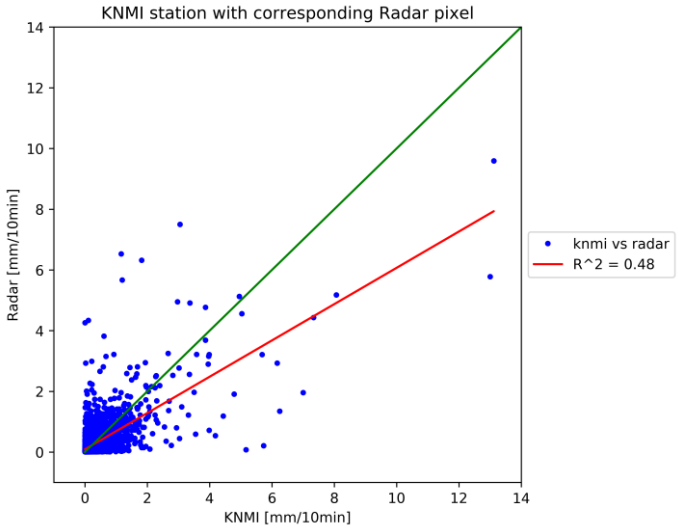


Figure 3.10: KNMI data vs radar, pixel that contains the KNMI station

But only Netatmo and TU Delft stations located in the same radar pixels are compared to each other. Therefore, it will still be assumed that the TU Delft stations have a higher quality than the Netatmo stations. Unfortunately, there are no Netatmo or TU Delft stations in the same radar pixel as the KNMI station. This could have given a better indication of the relation between radar and ground stations.

3.3.4 TU Delft:KNMI cross-variogram, imposing boundaries by using the estimations from the Netatmo:KNMI cross-variogram

Knowing that the TU Delft stations are of a higher quality can be used to estimate the nugget to fit the variogram. The TU Delft:KNMI cross-variogram should have a nugget which is lower than the nugget estimated in the Netatmo:KNMI cross-variogram. In order to avoid an arbitrary estimation of the nugget, this study assumes that both type of stations have the same nugget.

The estimation of the nugget (eq. 3, with an imposed nugget) resulted in a sill of 0.12 mm^2 and a range of 6.8 km. The range corresponds with the expectation that all the points on the variogram are already within the levelled-out section of the variogram. Noticeably, the range of the TU Delft fit is less than halve of the Netatmo fit. Just like the issue with the nugget in section 3.2.3, the range indicates that the TU Delft stations are worse than the Netatmo stations. Ideally, the range from the TU Delft fit should at least be similar to the Netatmo fit.

However, if the range of the TU Delft fit is increased to match the Netatmo, it is expected that the sill of the TU Delft fit will also increase. This is shown in table 5, where both the nugget and sill are estimated to be similar to the Netatmo fit (eq.3 with an imposed nugget and range). Additionally, table 5 shows the parameters estimated by the fit without boundaries and the parameters estimated by the fit where the nugget of the TU Delft:KNMI cross-variogram is imposed similar to the Netatmo:KNMI cross-variogram.

		Nugget [mm^2]	Sill [mm^2]	Range [m]
No boundaries	TU Delft	0.09	128	151487805
	Netatmo	0.04	644	242713908
Imposed nugget	TU Delft	0.02	0.12	6834
	Netatmo	0.02	0.12	15890
Imposed nugget & range	TU Delft	0.02	0.15	15890
	Netatmo	0.02	0.12	15890

Table 5: Variogram variables for the two station-types when the fit has no imposed boundaries, after imposing the nugget of the TU Delft:KNMI cross-variogram to be similar to the Netatmo:KNMI cross-variogram and after imposing the nugget and range o the TU Delft:KNMI cross-variogram to be similar to the Netatmo:KNMI cross-variogram.

From table 5, it becomes clear that the lack of TU Delft stations poses a problem in finding a proper fit, meaning a lower nugget, a lower sill and a higher range than the Netatmo fit. Choosing either of the options, will mean that we underestimate the quality of the TU Delft stations. Based on the method with which the merged rainfall maps are calculated (eq. 2), the impact of the range is considered to be higher and therefore the variables from the fit where both the nugget and rang are imposed will be used to calculate the merged rainfall maps. Figure 3.11 show this fit.

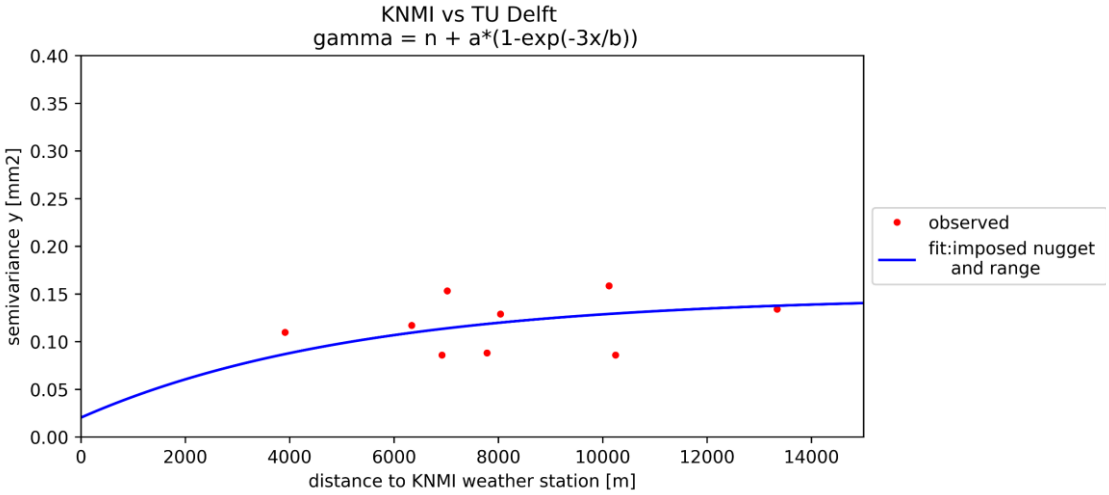


Figure 3.11: Exponential fit to the TU Delft:KNMI cross-variogram by imposing the nugget and range to be similar as the Netatmo:KNMI cross-variogram

3.4 Creating merged rainfall maps

Based on eq. 2 and the three variogram variables found from the fit of the cross-variograms (table 5), a single rainfall map can be calculated based on the two station-types. Since the range is the distance at which the stations no longer have a spatial correlation, if a station is further away from a pixel then the range, it will not be considered in the calculation of the rainfall.

Because the rainfall data from the KNMI weather station has a 10-minute temporal resolution, the rainfall maps based on the citizen stations will also have a 10-minute resolution, as well as the radar.

Based on the TU Delft and Netatmo stations, three different rainfall maps are created. The first map is based on just the Netatmo stations, the second map based on just the TU Delft stations and the third map based on a combination of the two. By doing so, it becomes clear how each station type performs and if a combination of multiple station types can produce a better map.

Figure 3.12 shows a 10-minute timestep of the rainfall radar, with the corresponding rainfall maps based on the citizen stations for the event of June 2016. What needs to be noticed is the range of the legends. The rainfall map based on the Netatmo stations has a much lower rainfall depth compared to the radar. The map based on the TU Delft stations also has a lower rainfall depth, albeit less than the Netatmo. There are a few reasons that can explain this: 1) the inherent difference between radar and ground stations and 2) the method of interpolation to calculate the rainfall maps.

The difference between radar and ground stations is discussed in section 3.3.3 and shows that the ground stations are measuring a lower rainfall intensity than the radar, which explains why the interpolated maps have a lower rainfall depth.

The interpolation process considers the decorrelation-range from the variogram in which the nugget and range are imposed (table. 5). However, the range from the Netatmo and TU Delft stations is still quite large, which means that many stations will be integrated in the

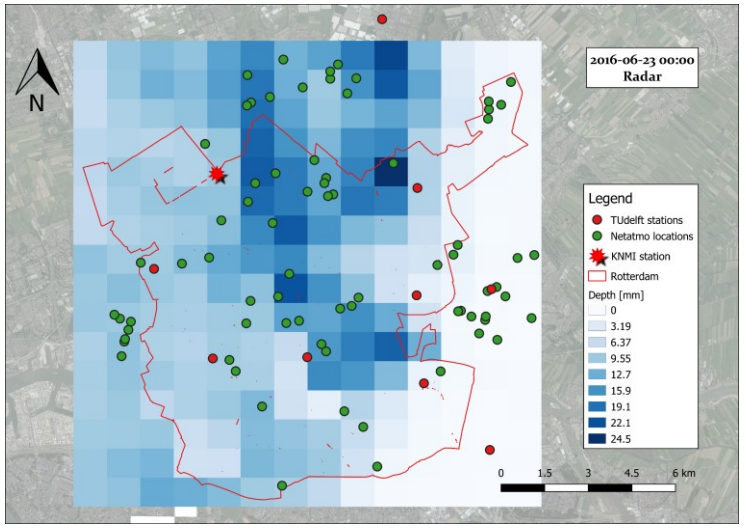


Figure 3.12a: Rainfall radar, 2016-06-23 00:00

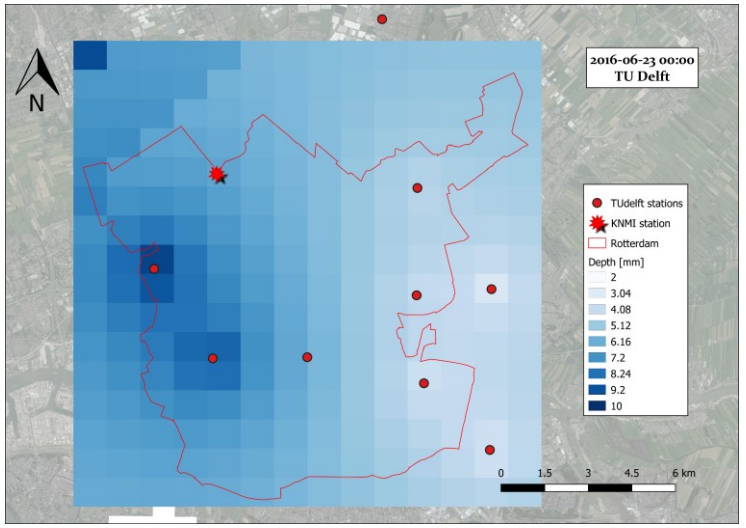


Figure 3.12b: Rainfall map based on TU Delft stations, 2016-06-23 00:00

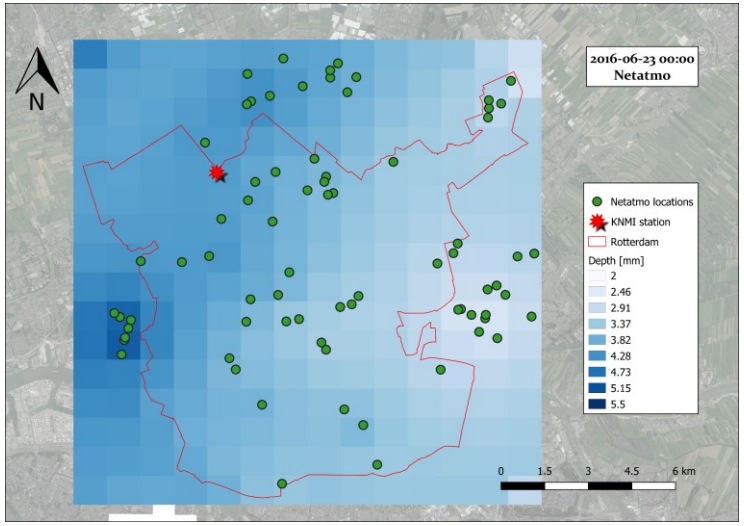


Figure 3.12c: Rainfall map based on Netatmo stations, 2016-06-23 00:00

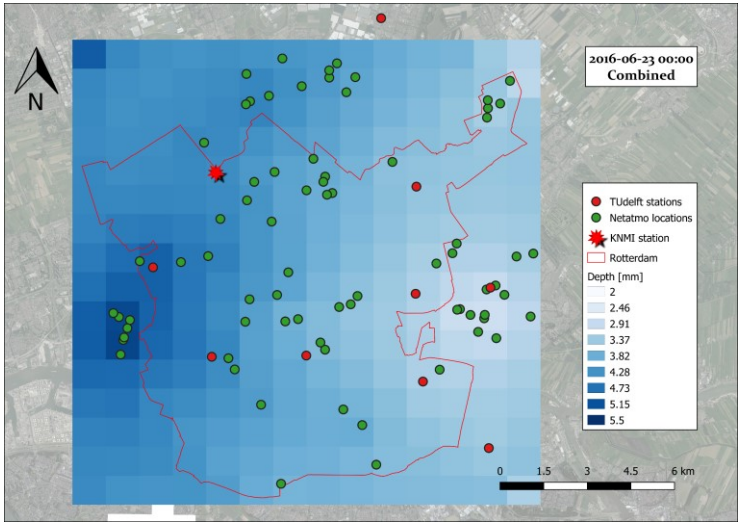


Figure 3.12d: Rainfall map based on both station-types, 2016-06-23 00:00

interpolation process and stations that are not within the boundaries of the rainfall event also influence the calculation. For this reason, the interpolation process will give, on average, lower intensities than the radar within the boundaries of the rainfall event and higher intensities outside the boundaries of the rainfall event.

Another thing that can be noticed is the peak rainfall pixel from the radar and the corresponding pixel in the Netatmo map. Figure 3.12a shows that there is a Netatmo station on top of the peak rainfall pixel from the radar. Therefore, it is expected that the corresponding pixel in the Netatmo map will give a high value as well, as the station is expected to capture this rainfall. Unfortunately, the Netatmo station is measuring a depth of 0 mm. With the large weighing factor based on the distance to the centre of the pixel, this measurement greatly reduces the calculated rainfall depth for that pixel.

Because of the boundaries of the manual quality control, it is unknown whether this measurement is a faulty zero. However, the zeros are not taken out in calculating the rainfall maps, since the data is used as is, in calculating the rainfall maps. Removing the zeros in this step, will also remove correct zero-values which results in maps that will give an overestimation of the rainfall. Keeping the potential faulty zeros will cause an underestimation of the calculated rainfall maps, but this in turn shows whether the data can be used as is.

In order to see the impact of possible faulty zeros, appendix C will include the filter created by Hutten et al. (2018) for the Netatmo stations and compare results. However, the filter did not recognize the zero measurement from the Netatmo station in the peak rainfall pixel in figure 3.12a as a faulty zero. Furthermore, the introduction of the filter did not introduce new faulty Netatmo stations that needed to be removed. Therefore, if there are only a few stations that show faulty zeros out of the seventy stations used for interpolation, the impact will not be that significant. The impact of the filter becomes even less during assessment on the event scale (10-minute timesteps across the entire event), since a larger dataset is used. The impact might have been higher if the filter introduced new faulty stations that needed to be removed.

3.5 Assessment of rainfall maps

In order to assess the rainfall maps, they will be compared to the single professional station baseline, assuming the KNMI rainfall is uniform over the entire study area. Table 6 gives descriptive statistics for maps of each rainfall event based on the different type of stations.

From table 6, it is clear that none of the rainfall maps capture the peak intensity measured by the radar. But looking at the median, it seems that the rainfall maps come close to the radar on the event-scale. Furthermore, if the citizen stations are compared separately to each other, it appears that the TU Delft stations measure values closer to the radar than the Netatmo stations, even though there are much fewer stations. This shows that the quality of the TU Delft stations is better, as was also concluded from figure 3.8 and figure 3.9. Moreover, the Netatmo variogram (fig. 3.6) shows that this many stations of citizen-level quality introduces a lot of noise. This greatly influences the fitted model, especially the range and sill. On the other hand, having fewer higher quality citizen stations (fig. 3.7) introduces a lack of spatial information. The TU Delft variogram could not have been modelled without the Netatmo variogram.

	Radar			KNMI			Netatmo			TU Delft			Combined		
Event (yyyy/mm/dd)	min	max	med	min	max	med	min*	max	med	min*	max	med	min*	max	med
Unit	mm/10 min			mm/10 min			mm/10 min			mm/10 min			mm/10 min		
2015-11-29	0.0	8.81	0.05	0.0	5.68	0.02	0.0	2.84	0.07	0.0	3.27	0.06	0.0	2.78	0.07
2016-06-23	0.0	24.5	0.23	0.0	13.12	0.02	0.0	6.45	0.71	0.0	10.09	0.70	0.0	6.44	0.67
2016-11-17	0.0	5.1	0.15	0.0	0.82	0.15	0.0	1.04	0.11	0.0	1.24	0.13	0.0	1.03	0.12
2017-07-25	0.0	6.56	0.01	0.0	3.5	0.01	0.0	1.22	0.07	0.0	1.92	0.04	0.0	1.13	0.06

Table 6: Minimum, maximum and median for the rainfall maps for each rainfall event

**The minimum is close to zero. Because of the spatial distribution of the stations and the estimated decorrelation-range, all pixels are within the range of Netatmo and TU Delft stations.*

3.5.1 Comparison with the baseline, 10-minute rainfall maps

In this section, some 10-minute rainfall maps will be compared to the KNMI station. This will be done be comparing the KNMI and citizen rainfall maps to the radar images. This will result in a difference map. By comparing these to each other, it can be checked whether the citizen stations improve information on the spatial variability of rainfall compared to the KNMI station.

Figure 3.13 shows four different 10-minute timesteps, one from each event. These timesteps were chosen based on the radar images, showing different spatial rainfall structures. The timesteps show the rainfall in a straight horizontal line over the KNMI weather station (fig. 3.13a), spread over the entire area (fig. 3.13b), covering the northern part of the area (fig. 3.13c) and concentrated in the middle of the area (fig. 3.13d).

No events were chosen where there was no rain on top of the KNMI station.

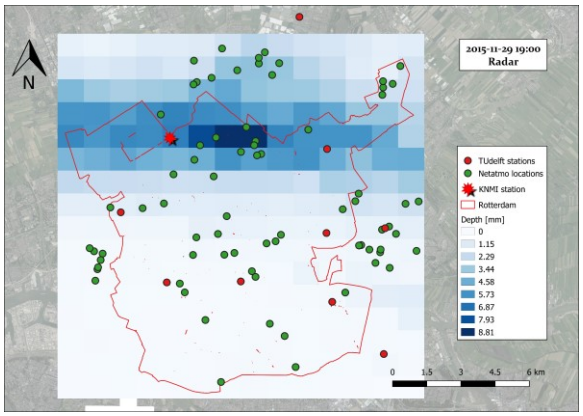


Figure 3.13a: Rainfall radar, 2015-11-29 19:00

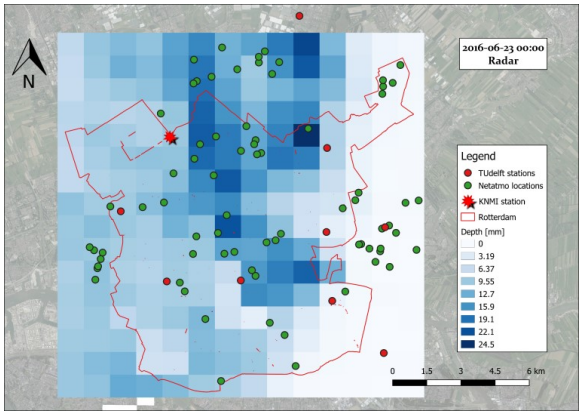


Figure 3.13b: Rainfall radar, 2016-23-06 00:00

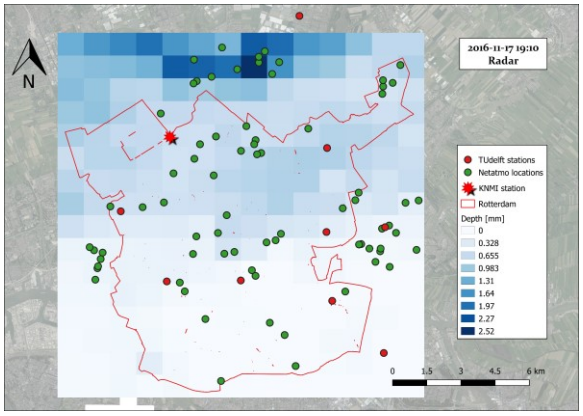


Figure 3.13c: Rainfall radar, 2016-11-17 19:10

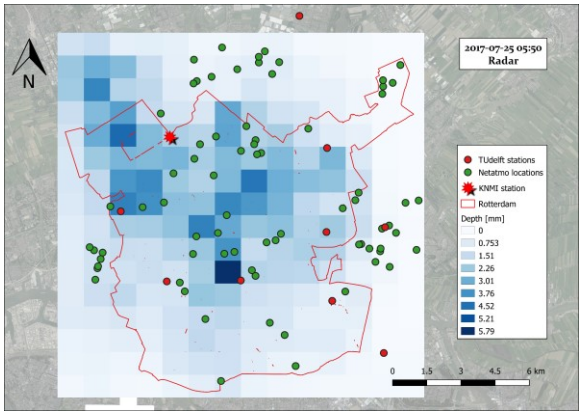


Figure 3.13d: Rainfall radar, 2017-07-25 05:50

The difference maps are given in figure 3.14a/d through 3.17a/d. They show the absolute value of the difference.

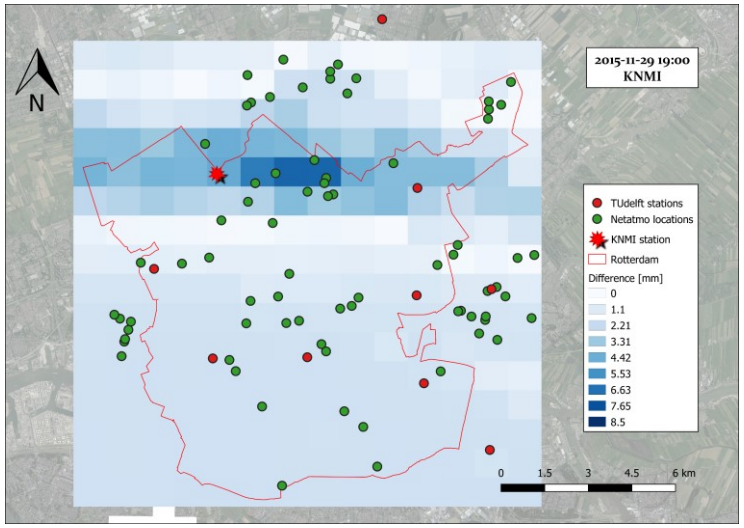


Figure 3.14a: Difference between radar and KNMI, 2015-11-29-19:00

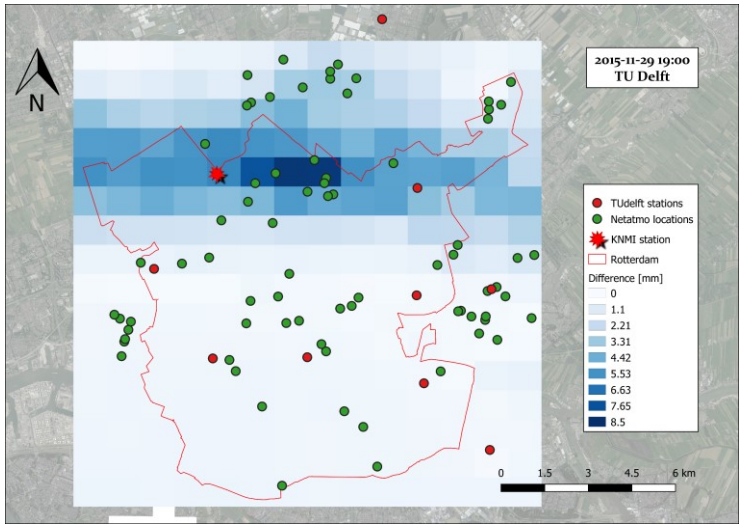


Figure 3.14b: Difference between radar and TU Delft, 2015-11-29-19:00

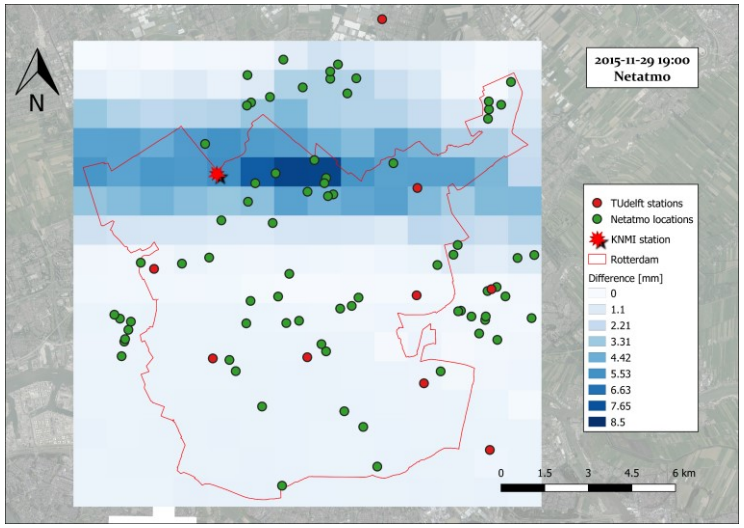


Figure 3.14c: Difference between radar and Netatmo, 2015-11-29-19:00

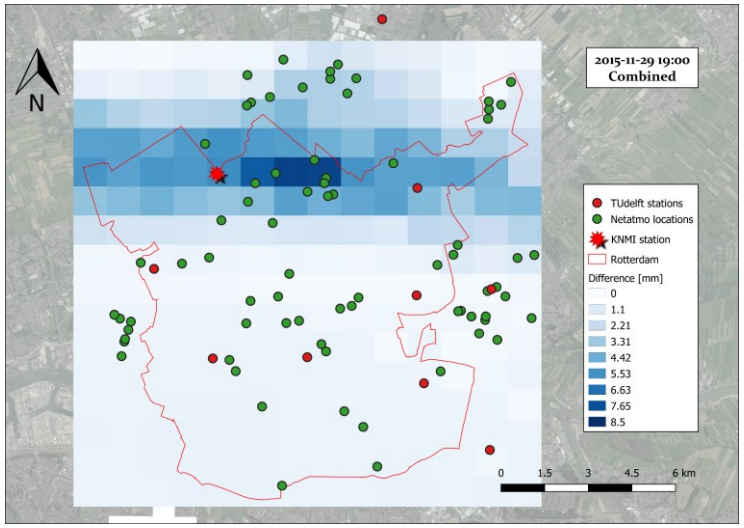


Figure 3.14d: Difference between radar and both station-types, 2015-11-29-19:00

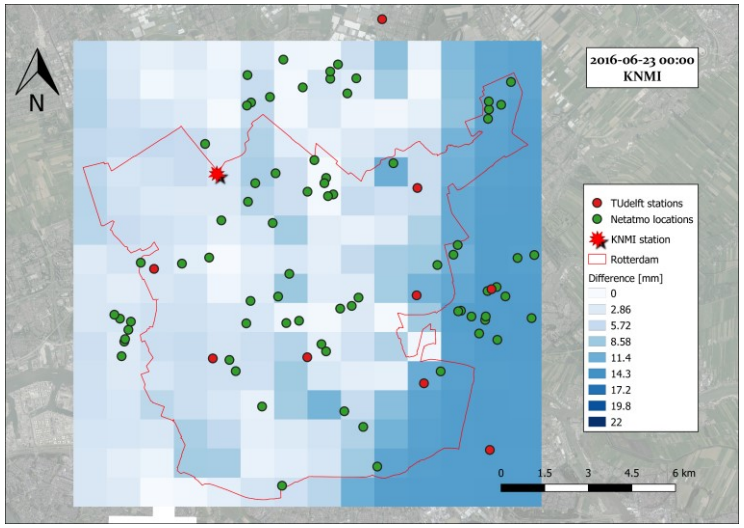


Figure 3.15a: Difference between radar and KNMI, 2016-06-23 00:00

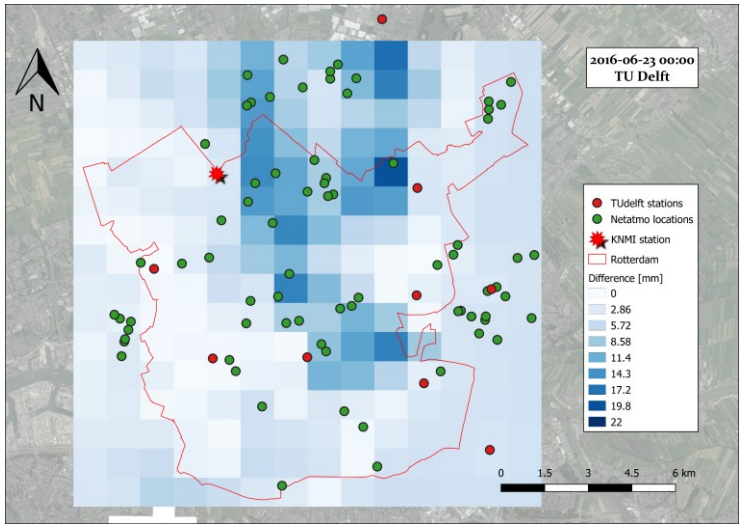


Figure 3.15b: Difference between radar and TU Delft, 2016-06-23 00:00

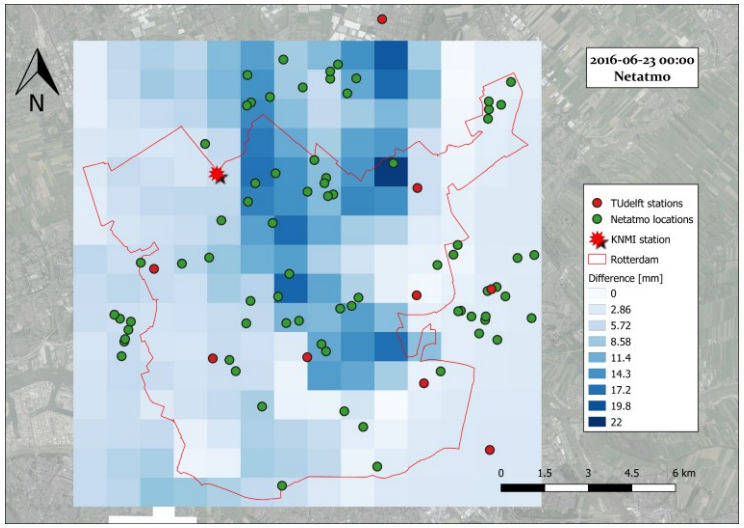


Figure 3.15c: Difference between radar and Netatmo, 2016-06-23 00:00

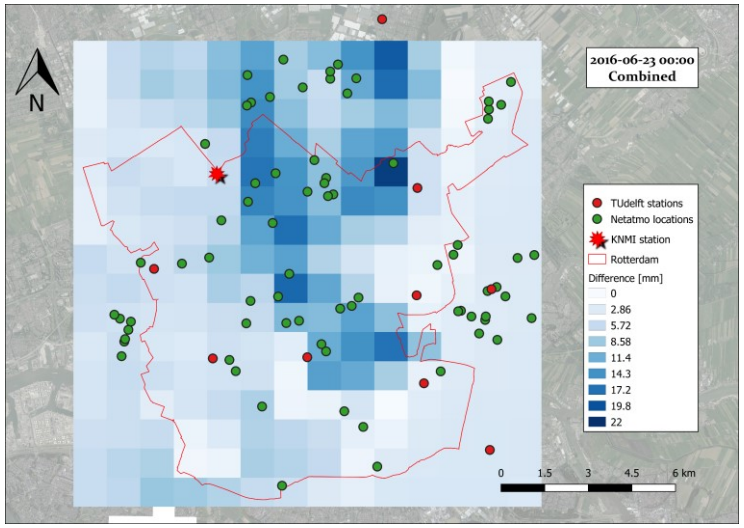


Figure 3.15d: Difference between radar and both station-types, 2016-06-23 00:00

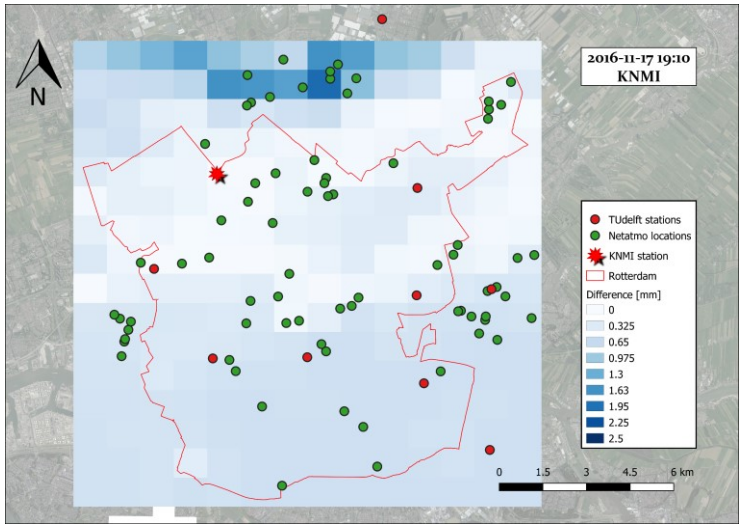


Figure 3.16a: Difference between radar and KNMI, 2016-11-17 19:10

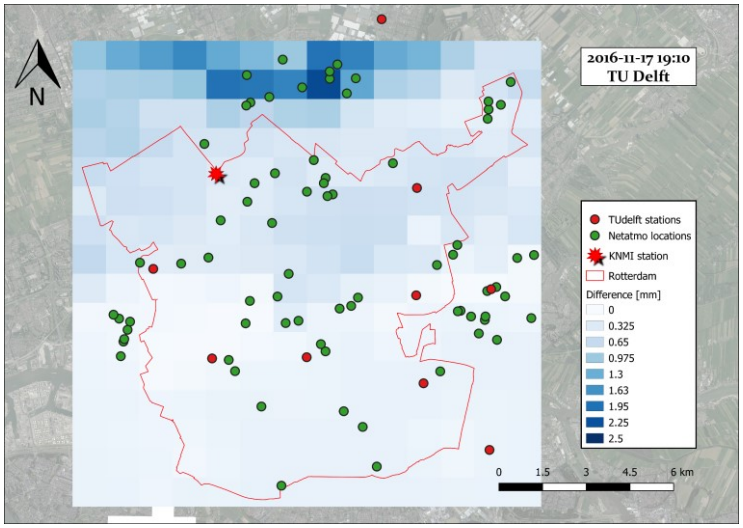


Figure 3.16b: Difference between radar and TU Delft, 2016-11-17 19:10

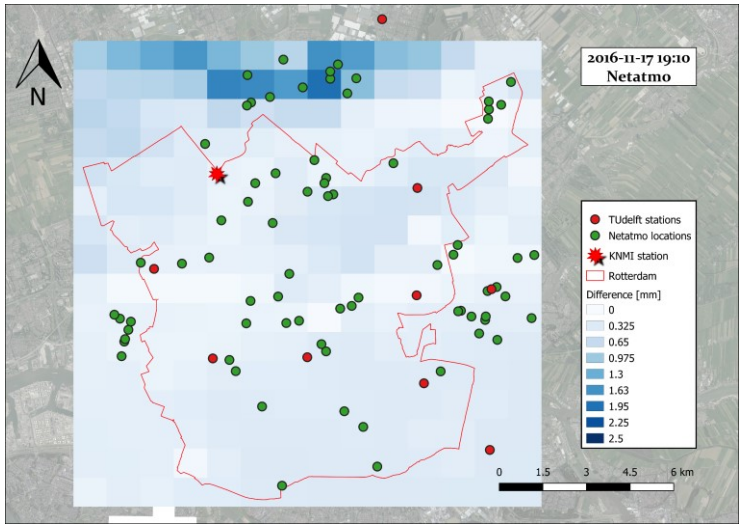


Figure 3.16c: Difference between radar and Netatmo, 2016-11-17 19:10

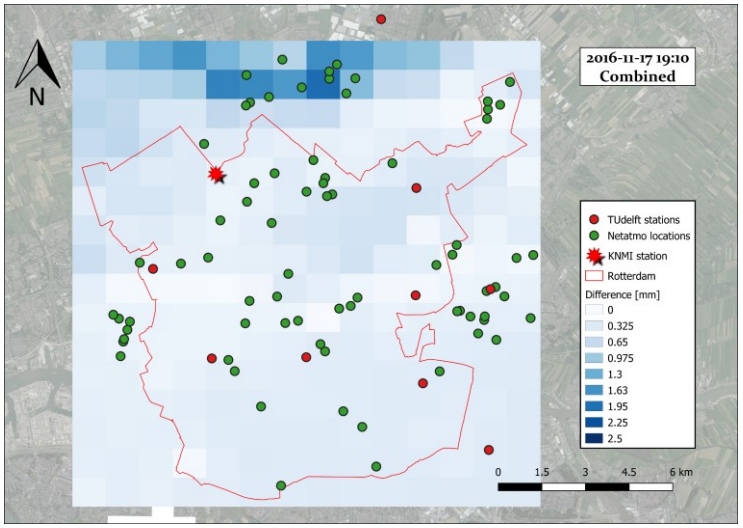


Figure 3.16d: Difference between radar and both station-types, 2016-11-17 19:10

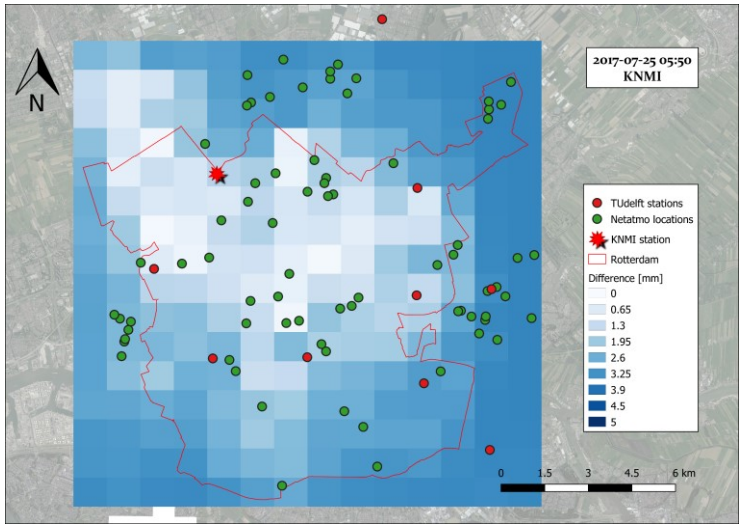


Figure 3.17a: Difference between radar and KNMI, 2017-07-25 05:50

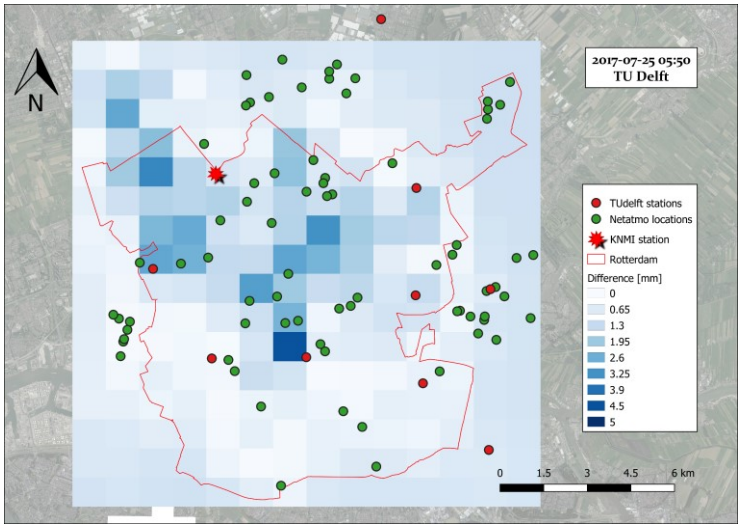


Figure 3.17b: Difference between radar and TU Delft, 2017-07-25 05:50

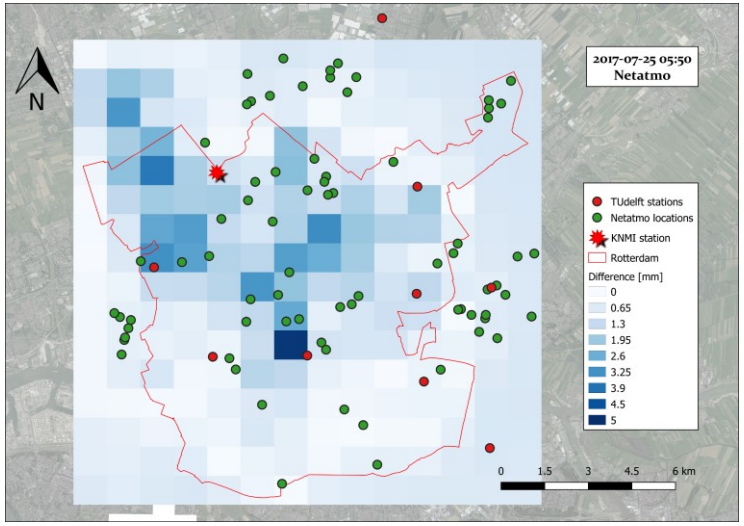


Figure 3.17c: Difference between radar and Netatmo, 2017-07-25 05:50

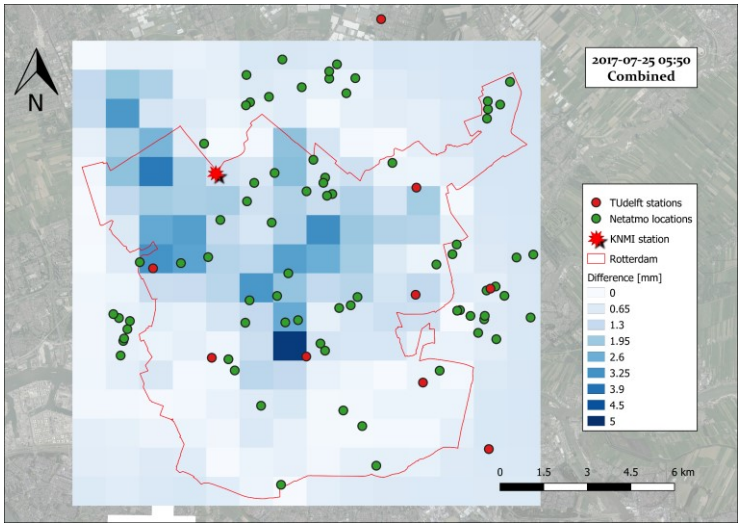


Figure 3.17d: Difference between radar and both station-types, 2017-07-25 05:50

By calculating the RMSD (eq.4) of the difference map and the Pearson correlation (eq.5) between the radar maps and interpolated maps, we can assess which map gives an overall better picture of the spatial variability of rainfall. This can be seen in table 7.

	RMSD [mm]				Pearson [-]			
Event	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
2015-11-29 19:00	0.013	0.012	0.011	0.011	0.00	0.64	0.70	0.73
2016-06-23 00:00	0.315	0.224	0.265	0.263	0.00	0.38	0.47	0.45
2016-11-17 19:10	0.001	0.001	0.001	0.001	0.00	0.65	0.77	0.79
2017-07-25 05:50	0.006	0.004	0.005	0.005	Nan	0.52	0.46	0.50

Table 7: RMSD and Pearson-correlation for the four 10-minute events

Table 7 shows that there is no correlation between the radar and the KNMI for the 10-minute timestep. This makes sense as the point-measurement from the KNMI weather station is assumed uniform over the entire area. More interesting is the difference is correlation between the TU Delft and Netatmo. Even though the TU Delft ground stations have a higher correlation with the radar than the Netatmo ground stations (fig. 3.8 and 3.9), the correlation of the interpolated maps based on TU Delft have a lower correlation than the interpolated maps based on Netatmo. This is influenced by two things: 1) the underestimation of the quality of the TU Delft stations in the fit of the variogram and 2) the structure of the network of Netatmo stations compared to the TU Delft stations.

The underestimation of the TU Delft stations is explained in section 3.2.4, in which the number of TU Delft stations was too little to create a proper fit. This resulted in a value for the sill which was too high, meaning the quality is estimated worse than it actually is. The structure is about the spread of the stations over the area in addition to the number of stations. Since there are many more Netatmo stations with a better spatial distribution, they are better capable at capturing the rainfall over the area. Does this mean that having many more lower quality stations is better? This can be answered by looking at the RMSD. The RMSD for the TU Delft map is very similar to that of the Netatmo map. This means that even with a less representation of the rainfall structure, the quality of the stations is able to balance out the lack of spatial representation of the rainfall structure.

If both maps are compared to the KNMI, it seems that the interpolated maps outperform the KNMI weather station. However just by looking at the difference maps, it seems that the pixels in which the radar detects rainfall, have a lower difference when compared to the KNMI than to the citizen stations. This would mean that even though the citizen stations better capture the structure of the rainfall event, the KNMI station still gives a more accurate value for the rainfall when assumed uniform over the area. In order to verify this, the absolute difference is split into two categories: 1) the absolute difference between the radar and ground station for radar pixels that measure rainfall and 2) the absolute difference for radar pixels that don't measure rainfall. The boundary for measuring rainfall is set to 0.2 mm, which is the capacity of the TU Delft tipping buckets.

Table 8 gives an overview of the median for these two categories.

	Median [mm] with Radar >= 0.2 mm				Median [mm] with Radar < 0.2 mm			
Event	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
2015-11-29 19:00	1.33	1.45	1.2	1.24	1.75	0.33	0.53	0.5
2016-06-23 00:00	4.43	3.69	5.37	5	13.06	4.21	3.08	3.24
2016-11-17 19:10	0.18	0.43	0.29	0.3	0.5	0.14	0.28	0.26
2017-07-25 05:50	2.34	0.61	0.54	0.54	3.46	0.93	0.76	0.77

Table 8: Median of the absolute difference when radar pixels measure rainfall or doesn't measure rainfall

From table 8, it is clear that the citizen stations are better when it comes to detecting the boundaries of the rainfall field. For the radar pixels that measure rainfall, no clear conclusion can be drawn as the median doesn't conclusively show which map is better. This means that the KNMI station doesn't necessarily give a more accurate value for the rainfall when assumed uniform over the area. However, we expect the absolute difference between the radar and the KNMI station to become lower when the distance between the pixels and the KNMI station becomes less. This should not be the case for the absolute difference between the radar and the citizen stations, as these stations are spread over the entire area. In order to see this behaviour, figure 3.18 a/d shows a boxplot, with a bin size of 2 km, on how the absolute difference changes in relation to the distance to the KNMI station for the first category (radar pixel measures rainfall) for each type of difference map (KNMI, TU Delft, Netatmo and Combined).

A difference can be detected because of the shape of the rainfall event. When the event is covering the entire area (fig. 3.18b) or centralized in the area and dissipating outwards (fig. 3.18d), the absolute difference between KNMI and radar starts to increase further away from the KNMI station. However, in the area wide event (fig. 3.18b), the spread of the difference is becoming larger when moving away from the KNMI station, while the spread of the centralized event (fig. 3.18d) is becoming smaller when moving away from the KNMI station. This is because for the centralized event, the rainfall is decreasing evenly when moving away from the KNMI station. For the area wide event, the rainfall is measurements are varying a lot, thus causing a larger spread in the difference. When the event is covering only a part of the area (fig. 3.18a and 3.18c), the absolute difference is unexpectedly decreasing when moving away from the KNMI station. This can be explained by looking at the radar images for those events (event 2015-11-29 19:00, fig. 3.13a and event 2016-11-17 19:10, fig. 3.13c). For event 2016-11-17 19:10, the rain event covers the northern halve of the area. Furthermore, outside of the peak rainfall rate, the intensities are quite similar for the entire event, causing pixels farther away to have the same absolute difference as pixels close by. For event 2015-11-29 19:00, the event is in a narrow horizontal band over the KNMI station. This means that there are not that many radar pixels containing rainfall. In addition, the radar pixel covering the KNMI station has very similar rainfall values as the pixels further away while the peak intensity pixels are close by. This causes the absolute difference to decrease when going further away from the KNMI station.

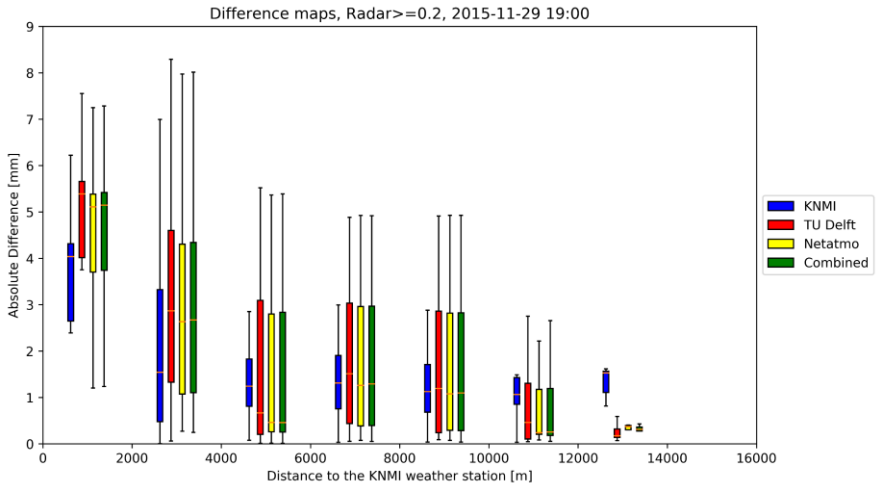


Figure 3.18a: Difference, radar vs interpolated maps when the radar is measuring rain, 2015-11-29 19:00

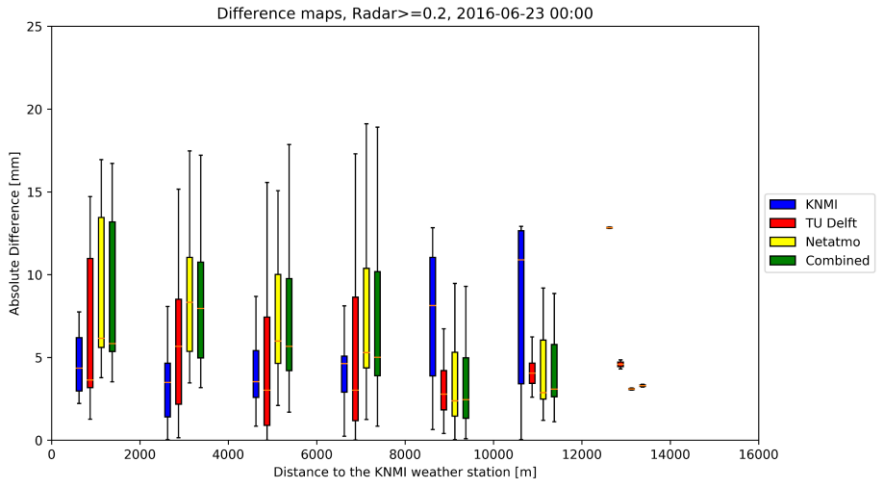


Figure 3.18b: Difference, radar vs interpolated maps when the radar is measuring rain, 2016-06-23 00:00

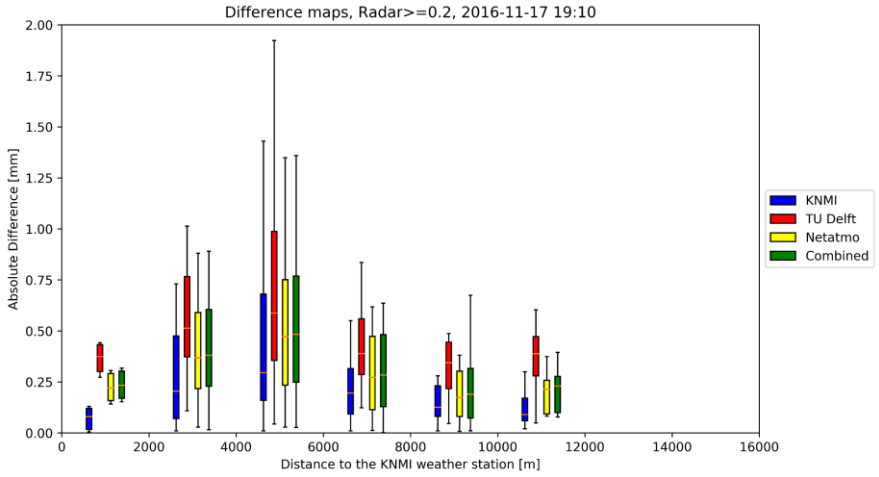


Figure 3.18c: Difference, radar vs interpolated maps when the radar is measuring rain, 2016-11-17 19:10

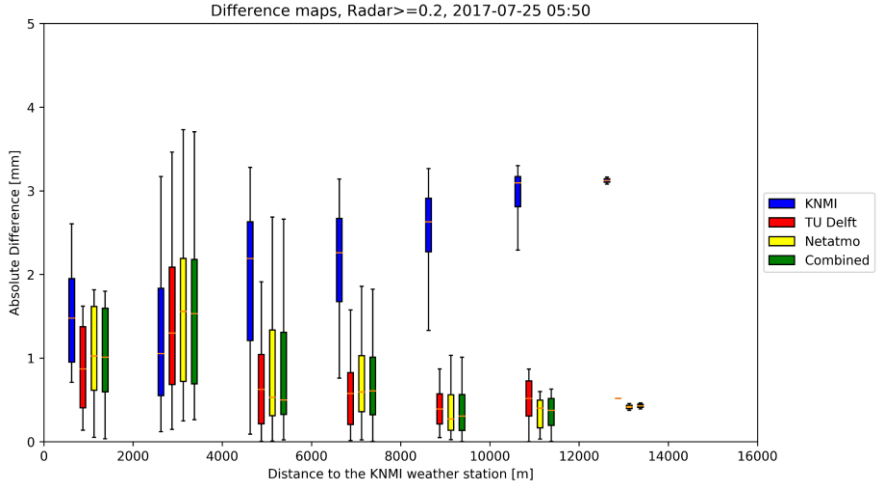


Figure 3.18d: Difference, radar vs interpolated maps when the radar is measuring rain, 2017-07-25 05:50

The next step is to compare the absolute difference of the KNMI station to the citizen stations. In three of the events (fig. 3.18a,b,c), when the pixels are close to the KNMI station, the KNMI has a lower absolute difference than the citizen stations, which was expected. Going further away from the KNMI station, the citizen stations start the become able to capture the rainfall better or equal to the KNMI station. The more interesting case is event 2017-07-25 05:50 (fig. 3.18d), where the rainfall is centralized in the study area. This event shows that the absolute difference from the citizen stations is already less or equal to that of the KNMI station. This would mean that for this particular event, the citizen stations are better than the KNMI, not just at capturing the structure of the event over the entire study area, which was already concluded from table 7, but also at capturing the rainfall values.

Considering the entire radar image, not just the pixels containing rainfall, a clear turning point can be seen after which the citizen stations begin to capture the rainfall better than the KNMI (fig. 3.19a/d). After 4 to 6 km, the citizen stations are starting to become better while after 8 km, they are conclusively better than the KNMI station.

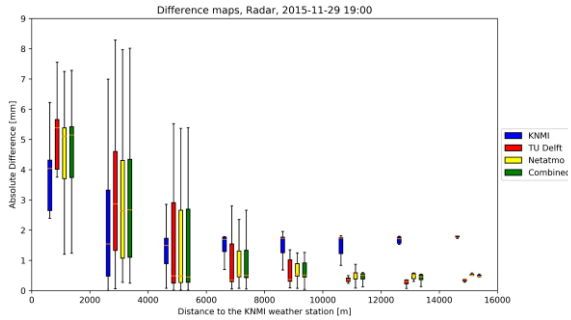


Figure 3.19a: Difference, radar vs interpolated maps, 2015-11-29 19:00

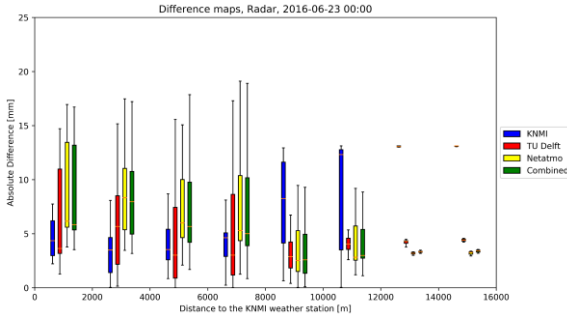


Figure 3.19b: Difference, radar vs interpolated maps, 2016-06-23 00:00

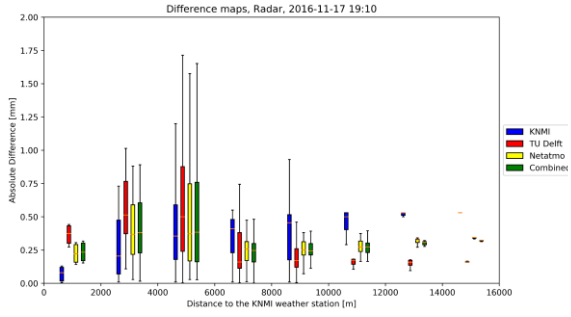


Figure 3.19c: Difference, radar vs interpolated maps, 2016-11-17 19:10

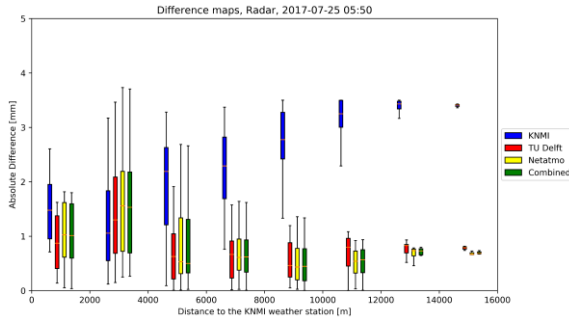


Figure 3.19d: Difference, radar vs interpolated maps, 2017-07-25 05:50

3.5.2 Comparison with the baseline, entire rainfall event

The previous section showcased the performance of the citizen stations to the KNMI on the smallest scale, four 10-minute timesteps. However, the four events (table 2), have a total of one-hundred 10-minute timesteps. So, it is also important on the event-scale, to compare performances.

This is done by taking the RMSD (eq. 4) and the correlation (eq. 5) per pixel over the entire rainfall event. Table 9 contains the median of the RMSD and correlation.

	Median of the RMSD [mm/10 min]				Median of the Pearson correlation [-]			
Event nr.	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
1 2015-11-19	1.32	0.66	0.66	0.65	0.53	0.74	0.74	0.75
2 2016-06-23	3.35	2.55	2.9	2.84	0.66	0.71	0.59	0.61
3 2016-11-17	0.28	0.21	0.22	0.22	0.58	0.79	0.74	0.75
4 2017-07-25	0.84	0.48	0.47	0.47	0.44	0.59	0.66	0.65

Table 9: Median of the RMSD and Pearson correlation for the four events

On the event scale, the 10-minute values across the entire event, it is clear that the citizen stations capture the structure of the rainfall event better than the KNMI station, as the RMSD is lower and the correlations are higher for the citizen stations. This means that citizen stations better represent the spatial variability of rainfall than the KNMI station.

In comparing the 10-minute timesteps, it became clear that the citizen stations gave a better representation of the rainfall than the KNMI station, when the pixels moved further from the KNMI station. Figure 3.20 a/d shows a boxplot, with a bin size of 2 km, on how the RMSD changes in relation to the distance to the KNMI station to verify if a similar trend can be seen on the event-scale. Figure 3.21 a/d show a similar boxplot for the correlation.

The RMSD of the KNMI behaves according to expectation, which is an increasing RMSD the further the pixels are from the KNMI station. This shows the importance of also taking the event-scale into account instead of only the 10-minute timesteps (fig 3.18) as they gave some counterintuitive results. Comparing the KNMI to the citizen stations, once again a trend can be seen in the distance at which the citizen stations start to give a better representation of the rainfall than the KNMI. As expected, close to the KNMI station, the KNMI gives a better representation of the rainfall than the citizen stations. However, after 4 km, this starts to shift and after 8 km, the citizen stations are conclusively better. This is a similar result as the 10-minute timesteps.

Looking at the correlation (fig. 3.21), the distance at which the citizen stations have a better correlation with the radar than the KNMI station is in the 6 to 10 km range, which is a little higher than the RMSD.

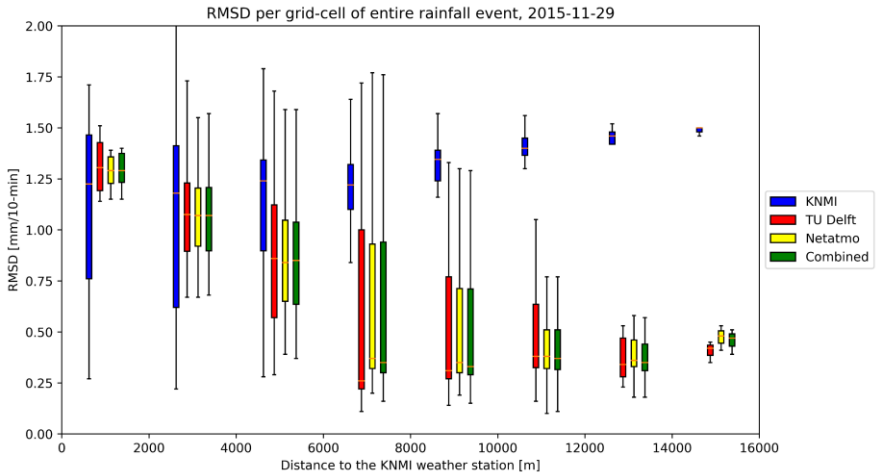


Figure 3.20a: RMSD per pixel in relation to the distance to the KNMI station, event 1: 2015-11-29

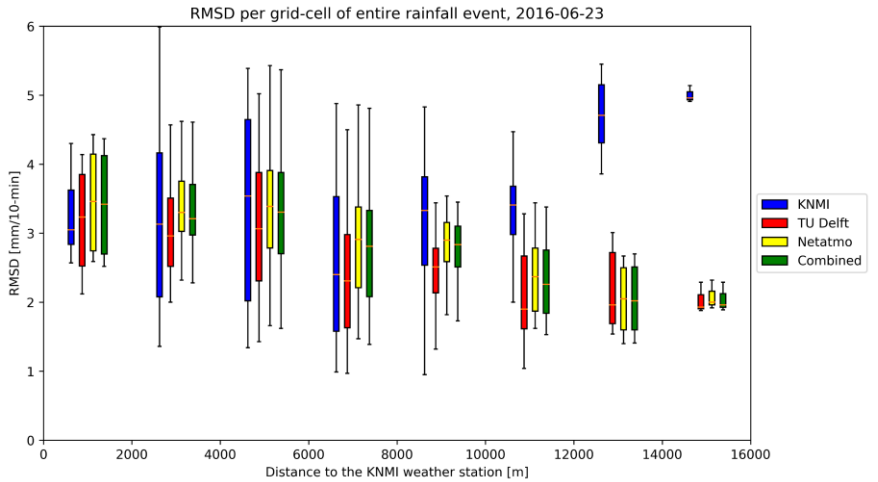


Figure 3.20b: RMSD per pixel in relation to the distance to the KNMI station, event 2: 2016-06-23

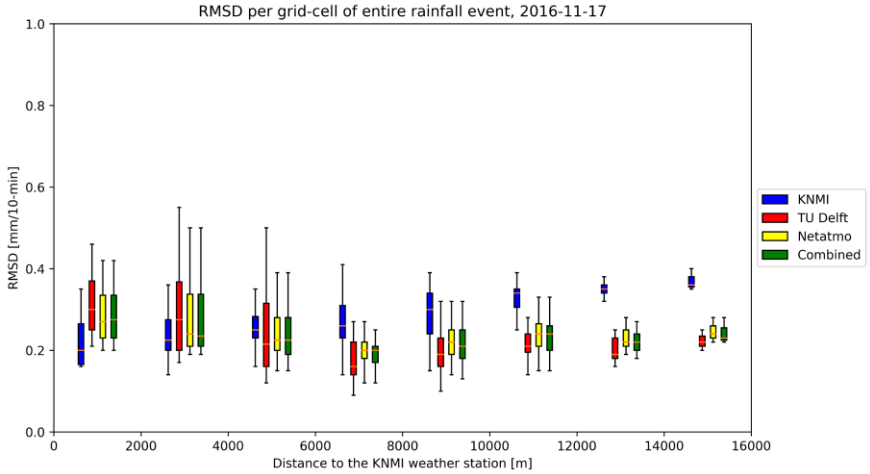


Figure 3.20c: RMSD per pixel in relation to the distance to the KNMI station, event 3: 2016-11-17

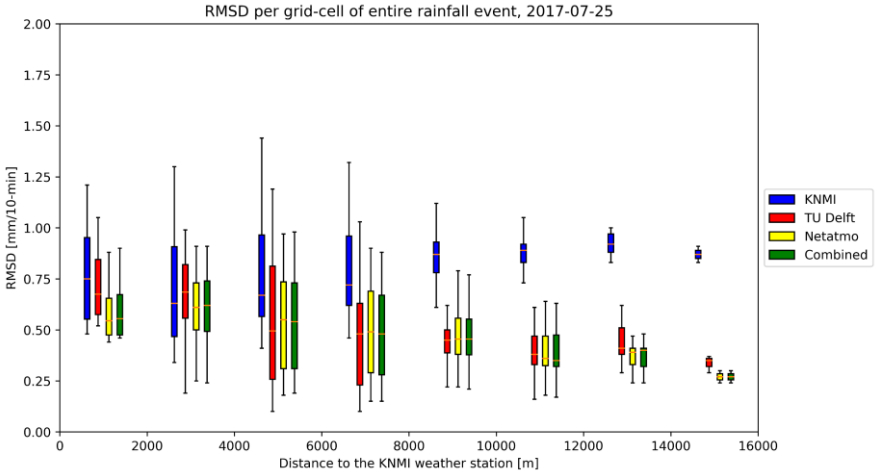


Figure 3.20d: RMSD per pixel in relation to the distance to the KNMI station, event 4: 2017-07-25

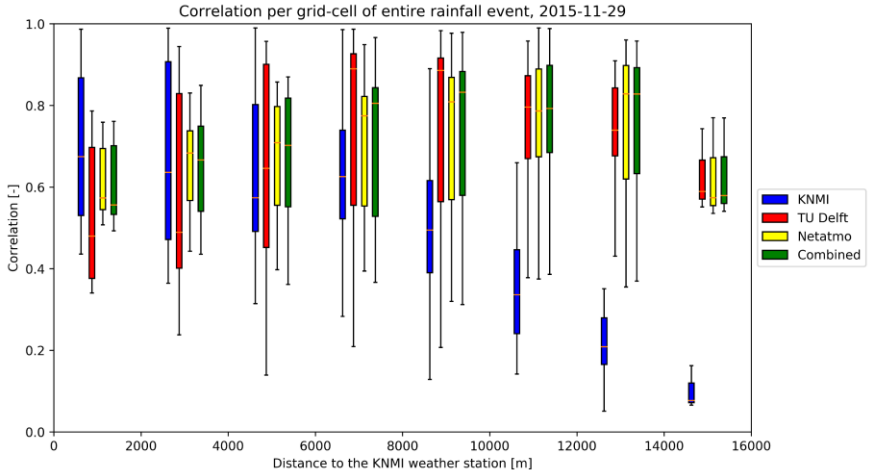


Figure 3.21a: Correlation per pixel in relation to the distance to the KNMI station, event 1: 2015-11-29

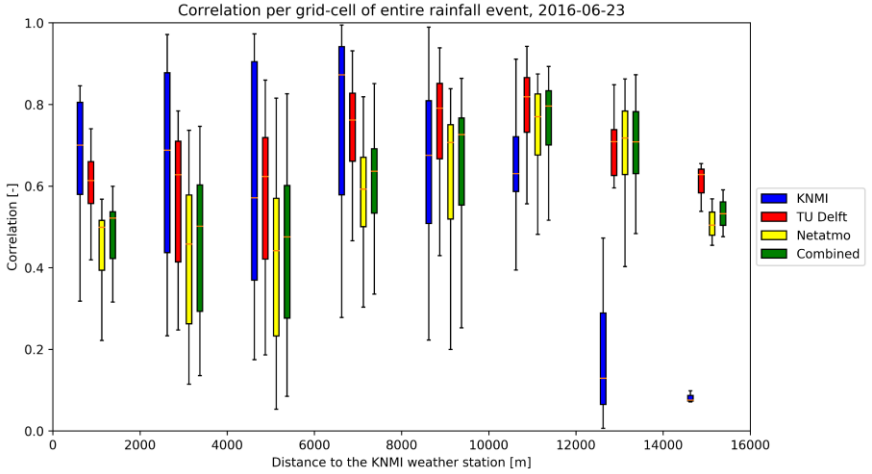


Figure 3.21b: Correlation per pixel in relation to the distance to the KNMI station, event 2: 2016-06-23

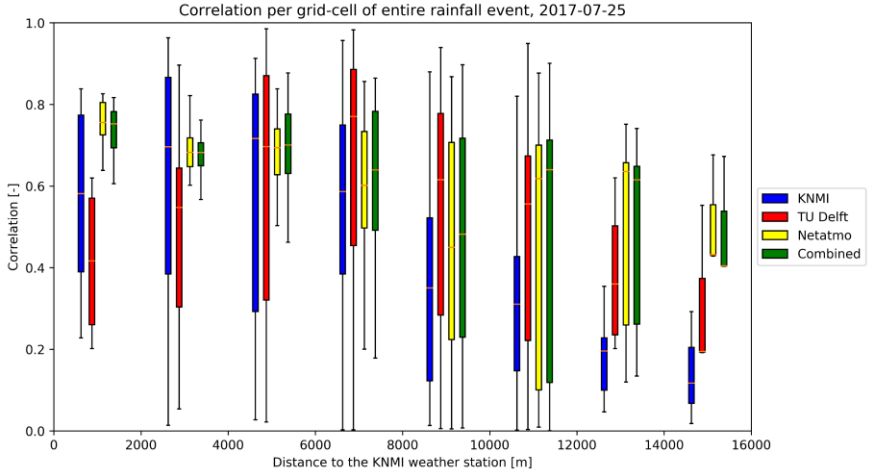


Figure 3.21c: Correlation per pixel in relation to the distance to the KNMI station, event 3: 2016-11-17

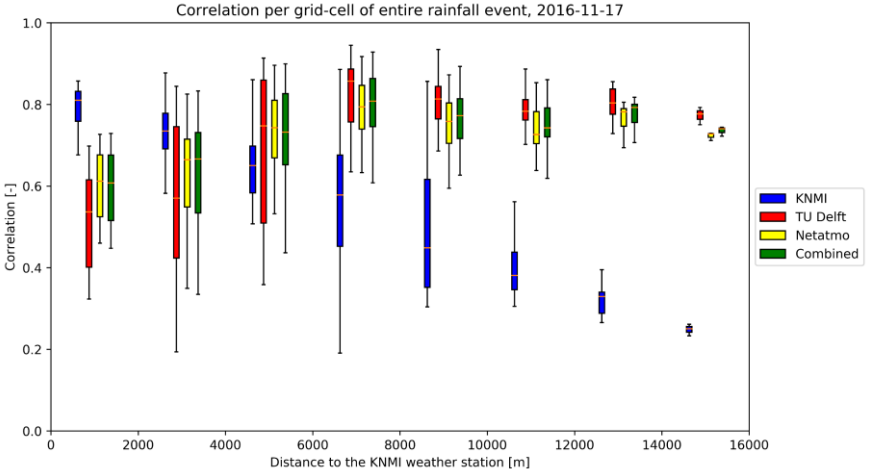


Figure 3.21d: Correlation per pixel in relation to the distance to the KNMI station, event 4: 2017-07-25

4. Discussion

In this section, the limitations of the study are discussed as well as possible improvements.

1. The Netatmo and KNMI data series used in this study ranged from the period of October 2015 to October 2017. However, the TU Delft stations had a longer available timeseries (in most cases), namely April 2013 to April 2018. In calculating the cross-variogram, the timeseries of the KNMI station limited the use of the full data from the TU Delft stations. Having a longer KNMI timeseries could have given a better semivariance but would not have fixed the issues that existed with the TU Delft:KNMI cross-variogram, a lack of stations and distribution in distance to the KNMI station.

2. In calculating the cross-variogram, only non-zero values were considered (from both the citizen stations as well as the KNMI station) to remove the possibility of faulty zero measurements in addition to giving a better picture of the spatial correlation when rainfall is measured. However, when creating the interpolated rainfall maps for the four chosen events, the zeros are considered. Had the zeros been removed, none of the pixels could potentially become zero. When compared to the radar, which has many zero pixels, this could negatively influence the assessment of these pixels. In order to verify the influence of faulty zeros during this study, appendix C shows the implementation of the filter by Hutten et al. (2018) in removing these faulty zeros and compares results. However, this did not result in significant changes of the results.

3. The fit of the Netatmo:KNMI cross-variogram resulted in a high decorrelation range of 16 km. This means that 16 km away from a Netatmo station, there still is spatial correlation in rainfall. This range was found by imposing the sill to be the median of the levelled-out part of the variogram. However, it was difficult to establish when the variogram started to level-out because of the noise. In fitting the TU Delft:KNMI cross-variogram and using the nugget from the Netatmo:KNMI cross-variogram, the decorrelation range was found to be 7 km, which seemed more realistic. But since the quality of the TU Delft stations is higher than the Netatmo stations, the range should not be lower than found in the Netatmo fit. Ideally, the range from the Netatmo fit would be changed to 7 km, as it seemed more realistic. However, this would mean that the Netatmo fit would be adjusted according to the TU Delft fit, while the TU Delft fit was estimated using the Netatmo fit. This seemed rather arbitrary, and therefore the range of the TU Delft fit was changed to match the range of the Netatmo fit as well as the nugget.
   However, this meant that two out of the three variogram variables for the TU Delft:KNMI cross-variogram fit were imposed as estimated in the Netatmo:KNMI cross-variogram fit. Since the weights are determined based on the fitted model, it begs the question of the significant difference between the two fitted models. Will there be a significant difference if the TU Delft stations get weighing factors according to the Netatmo:KNMI cross-variogram fit? This should be tested by having a larger network of semi-professional stations, in order to find a proper fit.

5. Conclusion

This research studied the potential of citizen observatories on improving the spatial measurement of rainfall in urban areas. This was done by using a network of semi-professional weather stations and a network of citizen weather stations to create interpolated rainfall maps and comparing these to a homogeneous field based on a single professional station (KNMI weather station) radar images. By doing so, it can be analysed whether a single professional station is better than multiple citizen stations spread across the area.

Looking at the small-scale (10-minute timesteps), there is no significant correlation between assuming the rainfall measured by the KNMI station uniformly over the area and the radar. The interpolated maps based on the citizen stations have a medium to large correlation with the radar, as the correlation is between 0.3 and 1.0. Assessment based on the RMSD show that the citizen stations give a better picture of the spatial measurements. However, based on the difference maps, it seemed that the KNMI gave a better representation of the actual rainfall values in the radar pixels that measured rain and the citizen stations are better in the case when radar pixels don't measure rain. This would mean that even though the citizen stations are better at capturing the spatial structure of the rainfall event over the area, the actual measurements for the rainfall are worse. By calculating the median of the absolute difference in the radar pixels that contain rainfall, it showed that this was not the case and that the rainfall measurements calculated in the interpolated maps are not necessarily worse than that of the KNMI station.

Furthermore, the performance of the rainfall maps was assessed by looking at the absolute difference per pixel in relation to the distance at the 10-minute scale. It became clear that the radar pixels within 4 km radius from the KNMI station are better represented by the KNMI station but that after a distance of 8 km, the citizen stations are conclusively better.

Looking at the event-scale, once again it became clear that the interpolated maps based on the citizen stations provide a better picture of the spatial measurements of rainfall in both the RMSD and correlation. Looking at the RMSD per pixel and correlation in regard to the distance to the KNMI station, for the RMSD showed that once again within a 4 km range from the KNMI station, the KNMI gave a better representation and after a distance of 8 km, the citizen stations are better. For the correlation the shift from KNMI to citizen station happened after 6 km with the citizen stations becoming conclusively better after 10 km.

Overall, the interpolated rainfall maps based on the citizen stations captured the structure of the rainfall over the area better than the professional station and giving a better representation of the actual rainfall values when the pixels are 8 to 10 km away from the KNMI station.

References

(UN), United Nation. 2018. *World Urbanization Prospects, The 2018 Revision, key facts.* United Nations.

AghaKouchak, A., Habib, E., Bárdossy, A. 2010. "Modeling radar rainfal estimation uncertainties: Random Error Model." *Journal of Hydrologic Engineering.*

Barnston, A.G. 1992. "Correspondence among the Correlation, RMSE, and Heidke Forecast verification measures; Refinement of the Heidke Score." *Weather and Forecasting.*

Biniyam, Y., Kemal, A. 2017. "The impacts of climate change on rainfall and flood frequency: the case of Hare watershed, southern rift valley of Ethiopia." *Earth Science & Climate Change.*

Bouwens, C.J.L, ten Veldhuis, J.A.E., Schleiss, M.A., Tian, X., Schepers, J. 2018. "Towards identification of critial rainfall thresholds for urban pluvial flooding prediction based on citizen flood observations."

Brouwer, T., Eilander, D., Loenen an, A., Booij, M.J., Wijnberg, K.M., Verkade, J.S., Wagemaker, J. 2017. "Probabilistic flood extent estimates from social media flood observations." *NHESS EGU.*

Buytaert, W., Zulkafli Z., Grainger, S., Acosta, L., Alemie, T.C., Bastiaensen, J., Bièvre de, B., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D.M., Hergarten, C., Isaeva, A., Karpouzoglou, T., Pandeya, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Hecken van, G., Zhumanova, M. 2014. "Citizen science in hydrology and water resouces: opportunities for knowledge generation, ecosystem service management, and sustainable development." *frontiers in earth science.*

Eijgenraam, Rolf. 2018. *Algemeen Dagblad.* 31 May. Accessed September 2018. https://www.ad.nl/rotterdam/massale-vissterfte-in-vijvers-en-singels~a683a426/.

Eilander, D., Trambauer, P., Wagemaker, J., Loenen van, A. 2016. "Harvesting social media for generation of near real-time flood maps." *Elsevier.*

Espinosa, B., Hromadka, T.V., Perez, R. 2015. "Comparison of radar data versus rainfall data." *Elsevier.*

Fung, A. 2006. "Varieties of Participation in Complex Governance." *Public Administration Review.*

Goodchild, M.F. 2007. "Citizens as sensors: the world of volunteerd geography." *GeoJournal.*

Houston, D., Werritty, A., Bassett, D., Geddes, A., Hoolachan., A., McMillan, M. 2011. *Pluvial (rain-related) flooding in urban areas: the invisible hazard.* Joseph Rowntree Foundation.

Howe, J. 2006. *Crowdsourcing.* 2 June. Accessed May 25, 2018. http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html.

Hutten, R.K., ten Veldhuis, J.A.E., Schleiss, M.A., de Vos, I.W., Leijnse, H., van de Giesen, N.C. 2018. "Automated quality-control filters for undetected rainfall in citizen rain gauge data."

IPCC, Field, C.B., Barros, V., Stocker, T.F., Dahe, Q., Dokken, D.J., Ebi, K.L., Mastrandrea, M.D., Mach, K.J., Plattner, G-K., Allen, S.K., Tignor, M., Midgley, P.M.,. 2012. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation.* A Special Report of the Intergovernmental Panel on Climate Change., Cambridge University Press.

Kidd, C., Becker, A., Huffman, G.J., Muller, C.L., Joe, P., Skofronick-Jackson, G., Kirschbaum, D. 2017. "So, how much of the earth's surface is covered by rain gauges?" *American Meteorological Society.*

Lopez, V., Napolitano, F., Russo, F. 2005. "Calibration of a rainfall-runoff model using radar and raingauge." *Advances in Geosciences.*

Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., Leigh, R.R. 2015. "Crowdsourcing for climate and atmospheric science: current stats and future potential." *International journal of climatology.*

Overeem, A., Leijnse, H., Uijlenhoet, R. 2016. "Two and a half years of country-wide rainfall maps using radio links from commercial cullular telecommunication networks." *AGU publications.*

Paul, J.D., Buytaert, W., Allen, S., Ballesteros-Cánovas, J.A., Bhusal, J., Cieslik, K., Clark, J., Dugar, S., Hannah, D.M., Stoffel, M., Dewulf, A., Dhital, M.R., Liu, W., Nayaval, J.L., Neupane, B., Schiller, A., Smith, P.J., Supper, R. 2018. "Citizen science for hydrological risk reduction and resilience building." *WIREs Water.*

Piccolo, F., Chirico, G. 2005. "Sampling errors in rainfall measurements by weather radar." *Advances in Geosciences.*

RCI. 2008. *Gemeente Rotterdam.* Accessed September 2018. https://www.010duurzamestad.nl/wat-wij-doen/lopende-projecten/rotterdamse-adaptatiestra/.

Restrepo-Estrada, C., Camargo de Andrade, S., Abe, N., Fava, M.C., Mendiondo, E.M., Porto de Albuquerque, J. 2017. "Geo-social media as a proxy for hydrometeorological data for streamflow estimation and to improve flood monitoring." *Elsevier.*

Starkey E., Parkin G., Birkinshaw S., Large A., Quin P., Bibson C.,. 2017. "Demonstrating the value of community-based ('citizen science') observations for catchments midelling and characterisation." *Elsevier.*

ten Veldhuis, J.A.E., Schleiss, M.A. 2017. "Statistical analysis of hydrological response in urbanising catchments based on adaptive sampling using inter-amount times." *Hydrology and Earth System Sciences.*

van de Giesen, N.C., Hut, R., ten Veldhuis, J.A.E. 2017. "Raindrop intervalometer." *Geophysical Research Abstracts.*

Wehn U., Rusca M., Evers J., Lanfranchi V. 2015. "Participation in flood risk management and the potential of citizen observatories: A governance analysis." *Elservier.*

Zimmerman, B., Zehe., E., Hartmann, N.K., Elsenbeer, H. 2008. "Analyzing spatial data: An assessment of assumptions new methods, and uncertainty using soil hydraulic data." *Water Resources Research.*

Appendix A

This section gives an overview the citizen stations and the arguments for adjusting the data during the manual quality control.

Table A1 shows the citizen stations that were adjusted. Table A2 and A3 show information on the TU Delft and Netatmo stations that were used during this study.

Results from manual quality control	
Station ID	Reason for adjustment
Ridderkerk	This station had a few days in which the sum of the daily rainfall depth exceeded 100 mm. These days were removed from the dataset
Oost	This station had a few days in which the sum of the daily rainfall depth exceeded 100 mm. These days were removed from the dataset
id354	This station did not measure rainfall for 8 months straight. This hints to a clogged funnel or tipping bucket. The station has been removed entirely.
id385	This station has a lot of gaps in the data-series. The station has been removed entirely.
id485	This station has a lot of gaps in the data-series. The station has been removed entirely.

Table A 1: Overview of adjusted stations.

ID	Start Date [yyyy-mm-dd hh-mm]	End Date [yyyy-mm-dd hh-mm]	N-Datapoints	% zero
Bolnes	2015-10-01 00:00	2017-10-01 00:00	86197	94.86
Capelle	2015-10-01 00:00	2017-10-01 00:00	105250	94.89
Delfshaven	2015-10-01 00:00	2017-10-01 00:00	105261	94.42
Lansingerland	2015-10-01 00:00	2017-10-01 00:00	105259	94.88
Ommoord	2015-10-01 00:00	2017-10-01 00:00	105258	94.82
Oost	2015-10-01 00:00	2017-10-01 00:00	105253	94.77
Ridderkerk	2015-10-01 00:00	2017-10-01 00:00	105251	94.91
Rijnhaven	2015-10-01 00:00	2017-10-01 00:00	105261	96.57
Spaanse Polder	2015-10-01 00:00	2017-10-01 00:00	105244	94.77

Table A 2: Overview of all TU Delft stations

ID	Start Date [yyyy-mm-dd hh-mm]	End Date [yyyy-mm-dd hh-mm]	N-Datapoints	% zero	ID	Start Date [yyyy-mm-dd hh-mm]	End Date [yyyy-mm-dd hh-mm]	N-Datapoints	% zero
154	2015-10-01 00:20	2017-10-01 00:00	105256	92.92	369	2015-12-21 14:20	2017-10-01 00:00	91420	93.84
155	2016-10-30 16:00	2017-09-30 07:00	46552	93.33	370	2016-07-20 14:30	2017-10-01 00:00	60677	94.00
156	2016-12-14 15:00	2017-10-01 00:00	41794	93.10	372	2016-08-31 18:40	2017-10-01 00:00	56873	94.13
159	2016-12-02 18:00	2017-10-01 00:00	43187	92.85	373	2015-10-01 04:20	2017-10-01 00:00	100511	93.51
161	2015-10-01 00:20	2017-10-01 00:00	105192	93.13	375	2015-10-28 16:20	2017-10-01 00:00	100669	92.83
162	2015-10-01 00:20	2017-10-01 00:00	104238	92.46	377	2015-12-08 16:30	2017-10-01 00:00	95038	94.00
165	2015-10-01 00:20	2017-10-01 00:00	103813	92.53	380	2015-10-28 14:40	2017-10-01 00:00	99449	92.57
167	2015-10-01 04:20	2017-10-01 00:00	103612	92.71	381	2017-08-09 21:20	2017-09-30 17:50	5161	90.87
168	2015-10-01 00:20	2017-10-01 00:00	103444	93.25	382	2015-10-01 00:20	2017-10-01 00:00	97176	94.08
169	2015-11-05 18:10	2017-10-01 00:00	95529	91.88	384	2017-01-14 11:30	2017-10-01 00:00	37361	93.43
187	2015-10-01 00:20	2017-10-01 00:00	100812	92.56	387	2016-06-10 13:30	2017-07-19 14:00	58068	94.33
340	2015-10-01 00:20	2017-03-08 11:00	73873	91.30	388	2015-10-01 00:20	2017-10-01 00:00	104007	92.78
341	2015-10-01 06:50	2017-10-01 00:00	103551	93.68	389	2015-10-01 00:20	2017-10-01 00:00	102759	93.24
342	2015-10-01 00:20	2017-10-01 00:00	104513	91.63	477	2016-09-17 09:50	2017-10-01 00:00	54041	93.47
343	2016-02-09 16:10	2017-10-01 00:00	74951	95.03	480	2016-03-25 15:00	2017-10-01 00:00	79380	93.47
344	2015-10-01 00:20	2017-10-01 00:00	98987	92.12	481	2015-10-01 00:20	2017-10-01 00:00	97183	92.01
345	2017-03-18 11:40	2017-10-01 00:00	26933	97.00	482	2017-05-30 10:30	2017-10-01 00:00	17227	86.72
346	2015-10-01 00:20	2017-10-01 00:00	101700	94.58	483	2015-10-01 00:20	2017-10-01 00:00	105098	92.67
347	2015-10-01 00:20	2017-04-28 18:20	82111	92.60	484	2017-08-21 10:40	2017-10-01 00:00	5787	90.18
348	2015-10-01 00:20	2017-10-01 00:00	105188	92.60	487	2015-12-25 12:00	2017-10-01 00:00	92277	92.79
349	2016-07-13 18:10	2017-10-01 00:00	60317	94.68	489	2017-08-18 20:00	2017-10-01 00:00	6213	90.63
350	2015-10-01 00:20	2017-10-01 00:00	103606	91.51	491	2015-12-03 16:50	2017-10-01 00:00	95893	93.03
352	2017-07-03 19:50	2017-10-01 00:00	12842	88.79	492	2015-10-01 00:20	2017-07-21 08:10	94935	92.40
353	2015-10-01 00:20	2017-10-01 00:00	101529	93.84	493	2015-10-01 00:20	2017-10-01 00:00	105006	93.39
355	2017-01-02 12:30	2017-10-01 00:00	39083	93.19	494	2015-10-01 00:20	2017-10-01 00:00	100432	92.39
356	2016-03-12 10:10	2017-10-01 00:00	75095	93.87	496	2017-07-26 13:50	2017-10-01 00:00	9505	91.15
358	2015-10-01 00:20	2017-10-01 00:00	96064	92.88	498	2016-06-11 08:30	2017-10-01 00:00	68559	94.36
359	2015-10-01 00:20	2017-04-03 06:20	68072	92.59	499	2015-10-01 00:20	2017-10-01 00:00	103134	92.72
360	2015-10-01 00:20	2017-10-01 00:00	100471	94.42	500	2015-10-01 00:20	2017-10-01 00:00	105263	92.27
361	2015-10-01 00:20	2017-10-01 00:00	104943	91.59	506	2016-07-05 16:50	2017-10-01 00:00	64971	93.66
363	2015-12-05 14:40	2017-10-01 00:00	94073	92.93	507	2015-10-01 00:20	2017-10-01 00:00	104516	94.97
365	2017-06-20 16:50	2017-10-01 00:00	14725	92.88	509	2017-02-18 10:30	2017-10-01 00:00	27249	94.37
366	2015-10-15 17:50	2016-01-23 09:10	13656	89.86	510	2015-12-05 19:40	2017-10-01 00:00	84189	92.94
367	2015-10-01 00:20	2017-10-01 00:00	81638	93.90	511	2015-10-01 00:20	2017-04-20 17:20	57816	95.03
368	2015-10-01 00:20	2017-09-25 04:50	95106	94.17	514	2016-05-01 11:20	2017-10-01 00:00	65974	93.38

Table A 3: Overview of all Netatmo stations.

Appendix B

Visualization of the four events chosen to compare the radar to the rainfall maps based on the citizen stations. These events were chosen visually by looking at intervals of the raw radar data.

Event 2015-11-29

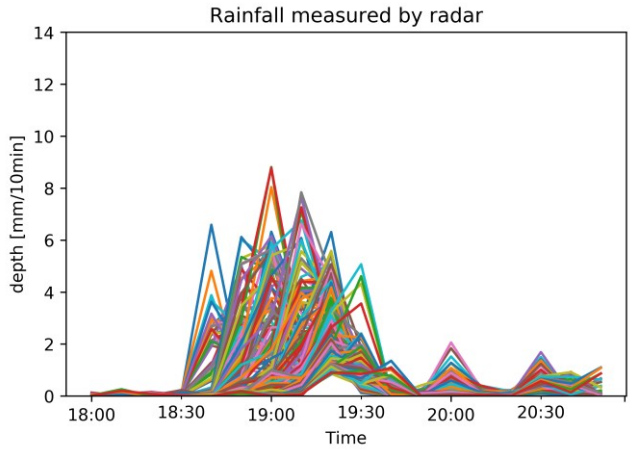


Figure B1a: Rainfall measured by the radar, event 1

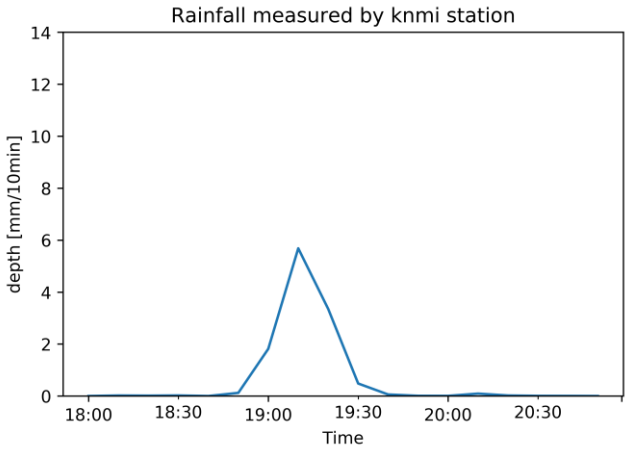


Figure B1b: Rainfall measured by the KNMI station, event 1

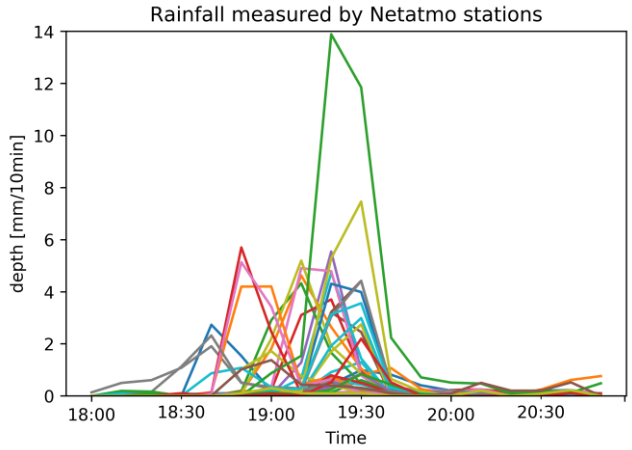


Figure B1c: Rainfall measured by the Netatmo stations, event 1

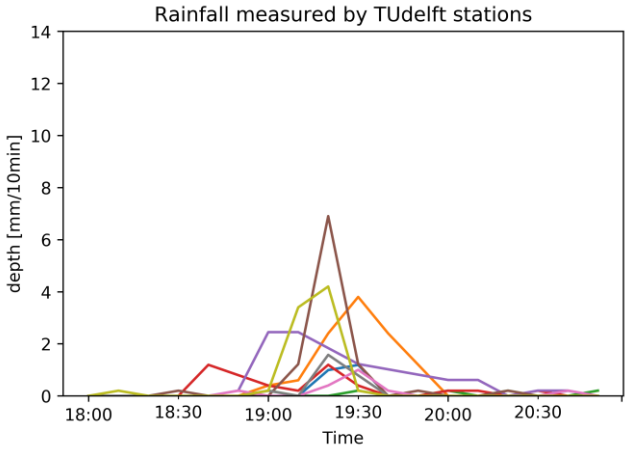


Figure B1d: Rainfall measured by the TU Delft stations, event 1

Event 2016-06-23

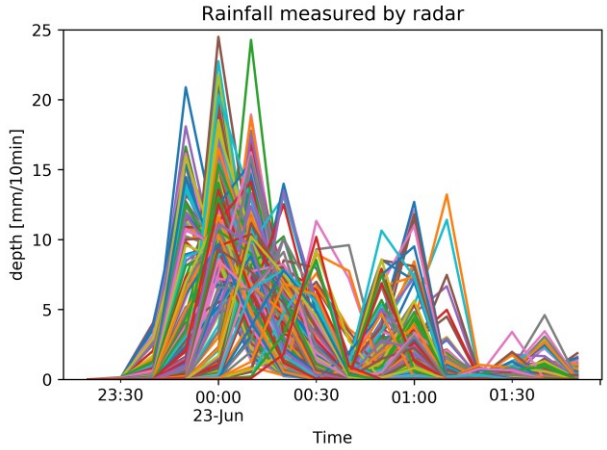


Figure B2a: Rainfall measured by the radar, event 2

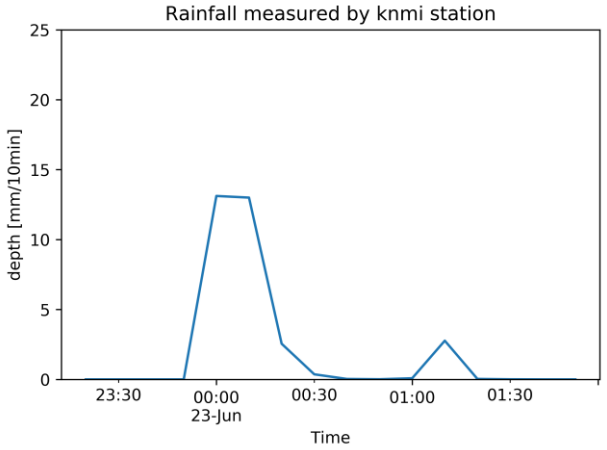


Figure B2b: Rainfall measured by the KNMI station, event 2

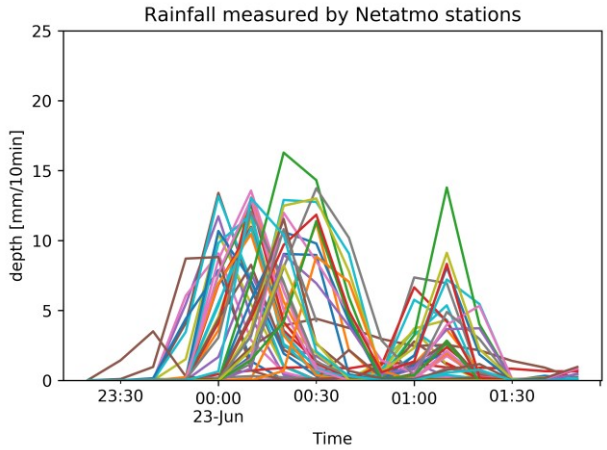


Figure B2c: Rainfall measured by the Netatmo stations, event 2

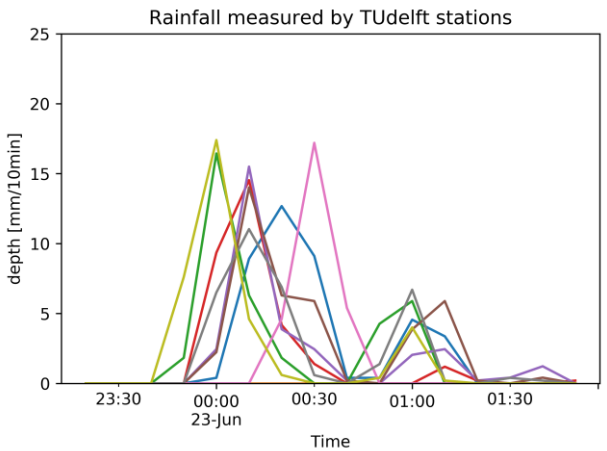


Figure B2d: Rainfall measured by the TU Delft stations, event 2

Event 2016-11-17

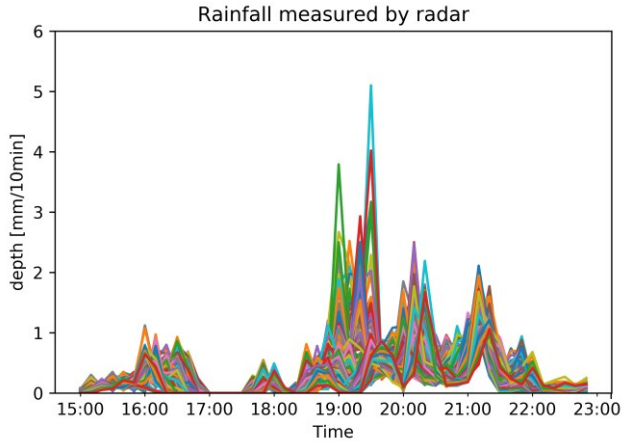


Figure B3a: Rainfall measured by the radar, event 3

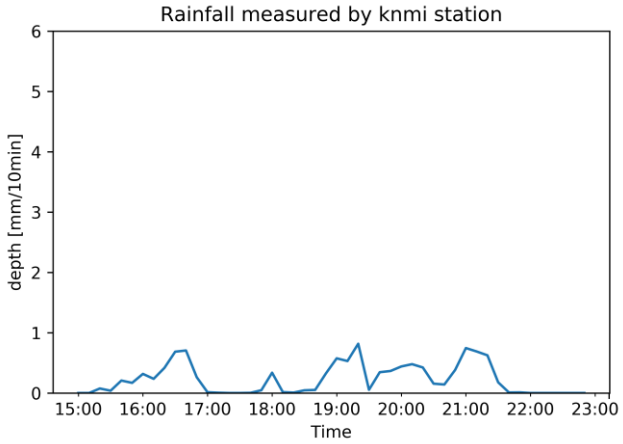


Figure B3b: Rainfall measured by the KNMI station, event 3

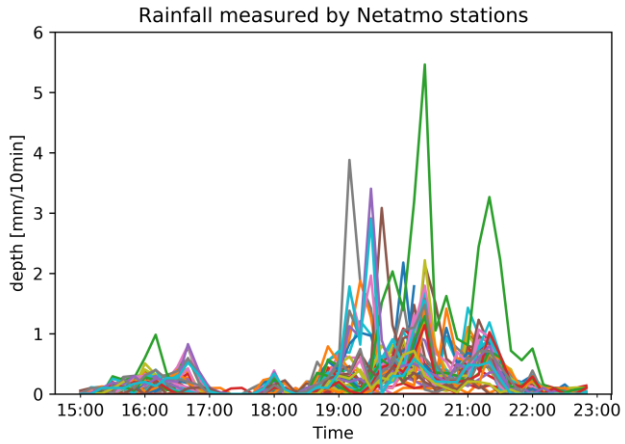


Figure B3c: Rainfall measured by the Netatmo stations, event 3

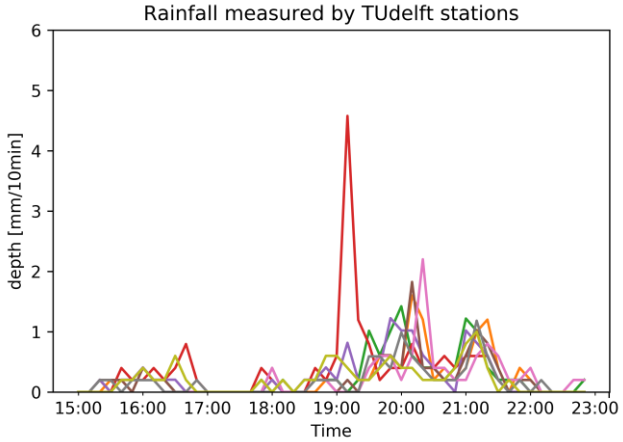


Figure B3d: Rainfall measured by the TU Delft stations, event 3

Event 2017-07-25

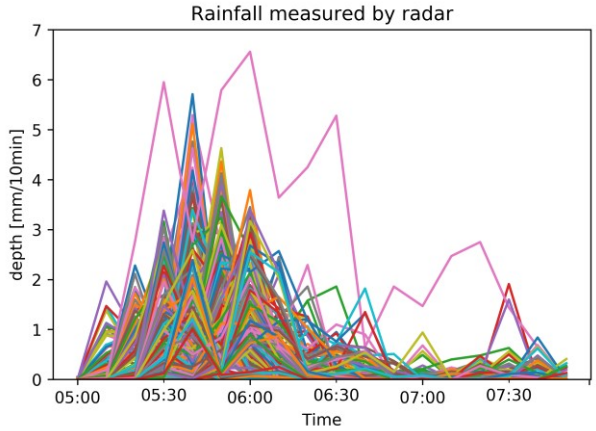


Figure B4a: Rainfall measured by the radar, event 4

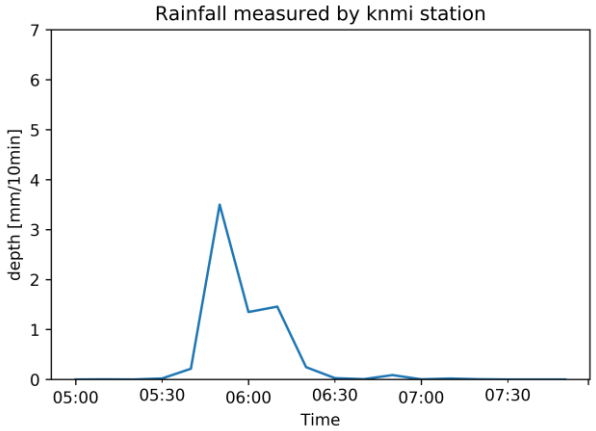


Figure B4b: Rainfall measured by the KNMI station, event 4

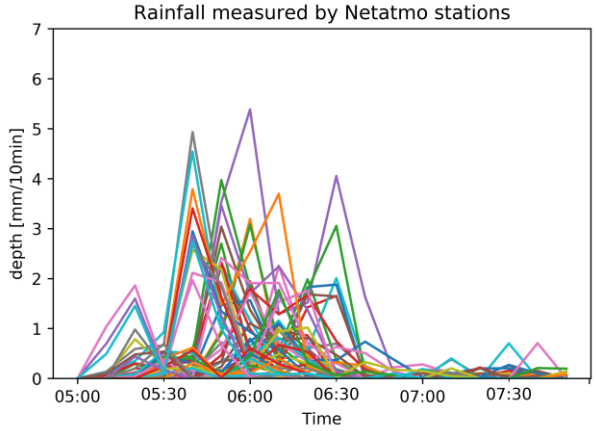


Figure B4c: Rainfall measured by the Netatmo stations, event 4

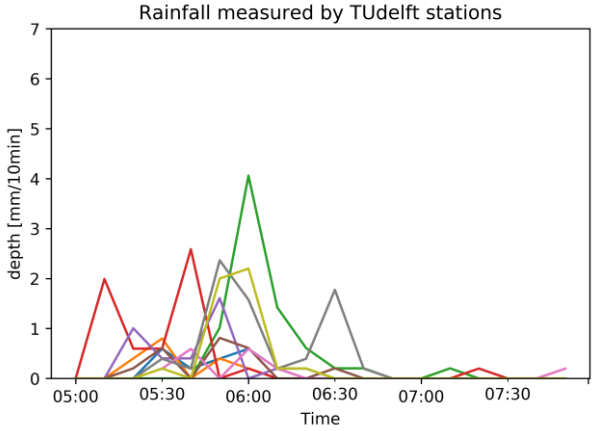


Figure B4d: Rainfall measured by the TU Delft stations, event 4

Appendix C

In this section, the filter that was created in the study by Hutten et al. (2018) is applied on the data to filter put faulty zero values from the Netatmo stations. Then the manual quality control is applied once more to check if more stations turn out to be faulty. The results are then compared to the results from section 3.4 to see whether this additional form of quality control contributed to an improvement of the rainfall maps.

Section C1 will discuss the differences in the step for the quality control and the impact on the fits and interpolated rainfall maps. Section C2 will compare the basic statistics, section C3 the 10-minute rainfall maps with corresponding RMSD and correlation and section C4 the RMSD and correlation of the entire events.

C1. Quality control and variogram fit

Before the manual quality control was applied on the Netatmo stations, the data went through the faulty-zero filter. However, in removing stations and data that are obviously wrong during the manual quality control, the filter did not reveal new stations. This means that with or without the filter, the manual quality control removed the same stations from the dataset.

This meant that the fit for the Netatmo:KNMI cross-variogram did not change, as the cross-variogram was already calculated by taking out all the zeros. Consequently, the fit for the TU Delft:KNMI cross-variogram remained the same as well. Therefore, the same weighing factors could be applied in creating the interpolated rainfall maps.

The filter will have an impact on the values of the interpolated rainfall maps, as some zero-values are now removed from the interpolation, which should result in overall higher rainfall values.

C2. Descriptive statistics

Table C1 shows the original statistics while table C2 show the statistics after applying the filter on the Netatmo stations. As expected, the interpolated Netatmo and Combined maps have higher maximum and median values when the filter is applied. But the difference is very small. Consequently, the TU Delft stations still give a better representation of the peaks and performs better on the event-scale.

	Radar			KNMI			Netatmo			TU Delft			Combined		
Event (yyyy/mm/dd)	min	max	med	min	max	med	min*	max	med	min*	max	med	min*	max	med
Unit	mm/10 min			mm/10 min			mm/10 min			mm/10 min			mm/10 min		
2015-11-29	0.0	8.81	0.05	0.0	5.68	0.02	0.0	2.84	0.07	0.0	3.27	0.06	0.0	2.78	0.07
2016-06-23	0.0	24.5	0.23	0.0	13.12	0.02	0.0	6.45	0.71	0.0	10.09	0.70	0.0	6.44	0.67
2016-11-17	0.0	5.1	0.15	0.0	0.82	0.15	0.0	1.04	0.11	0.0	1.24	0.13	0.0	1.03	0.12
2017-07-25	0.0	6.56	0.01	0.0	3.5	0.01	0.0	1.22	0.07	0.0	1.92	0.04	0.0	1.13	0.06

Table C1: Minimum, maximum and median for the rainfall maps for each rainfall event using only the manual quality control

	Radar			KNMI			Netatmo			TU Delft			Combined		
Event (yyyy/mm/dd)	min	max	med	min	max	med	min*	max	med	min*	max	med	min*	max	med
Unit	mm/10 min			mm/10 min			mm/10 min			mm/10 min			mm/10 min		
2015-11-29	0.0	8.81	0.05	0.0	5.68	0.02	0.0	3.10	0.08	0.0	3.27	0.06	0.0	2.99	0.08
2016-06-23	0.0	24.5	0.23	0.0	13.12	0.02	0.0	7.02	0.75	0.0	10.09	0.70	0.0	7.12	0.73
2016-11-17	0.0	5.1	0.15	0.0	0.82	0.15	0.0	1.17	0.15	0.0	1.24	0.13	0.0	1.07	0.15
2017-07-25	0.0	6.56	0.01	0.0	3.5	0.01	0.0	1.45	0.07	0.0	1.92	0.04	0.0	1.41	0.07

Table C2: Minimum, maximum and median for the rainfall maps for each rainfall event using the faulty zero filter by Hutten et al. (2018) and the manual quality control

C3. Comparison with the baseline, 10-minute timesteps

The filter has some effect, but not a lot. Because the interpolation process has 70 available stations, if only a few of these stations have faulty zeros, the impact will not be that significant.

	RMSD [mm]				Pearson [-]			
Event	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
2015-11-29 19:00	0.013	0.012	0.011	0.011	0.00	0.64	0.70	0.73
2016-06-23 00:00	0.315	0.224	0.265	0.263	0.00	0.38	0.47	0.45
2016-11-17 19:10	0.001	0.001	0.001	0.001	0.00	0.65	0.77	0.79
2017-07-25 05:50	0.006	0.004	0.005	0.005	Nan	0.52	0.46	0.50

Table C3: RMSD and Pearson correlation for the four 10-minue events using only the manual quality control

	RMSD [mm]				Pearson [-]			
Event	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
2015-11-29 19:00	0.013	0.012	0.010	0.011	0.00	0.64	0.71	0.73
2016-06-23 00:00	0.315	0.224	0.259	0.257	0.00	0.38	0.48	0.45
2016-11-17 19:10	0.001	0.001	0.001	0.001	0.00	0.65	0.82	0.83
2017-07-25 05:50	0.006	0.004	0.005	0.004	Nan	0.52	0.42	0.47

Table C4: RMD and Pearson correlation for the four 10-minute events using the faulty zero filter by Hutten et al (2018) and the manual quality control

	Median [mm] with Radar >= 0.2 mm				Median [mm] with Radar < 0.2 mm			
Event	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
2015-11-29 19:00	1.33	1.45	1.2	1.24	1.75	0.33	0.53	0.5
2016-06-23 00:00	4.43	3.69	5.37	5	13.06	4.21	3.08	3.24
2016-11-17 19:10	0.18	0.43	0.29	0.3	0.5	0.14	0.28	0.26
2017-07-25 05:50	2.34	0.61	0.54	0.54	3.46	0.93	0.76	0.77

Table C5: Median of the absolute difference when the radar pixels measures rainfall or doesn't measure rainfall, using only the manual quality control

	Median [mm] with Radar >= 0.2 mm				Median [mm] with Radar < 0.2 mm			
Event	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
2015-11-29 19:00	1.33	1.45	1.14	1.19	1.75	0.33	0.58	0.54
2016-06-23 00:00	4.43	3.69	4.74	4.42	13.06	4.21	3.43	3.56
2016-11-17 19:10	0.18	0.43	0.2	0.22	0.5	0.14	0.35	0.32
2017-07-25 05:50	2.34	0.61	0.58	0.95	3.46	0.93	0.88	0.88

Table C6: Median of the absolute difference when the radar pixels measures rainfall or doesn't measure rainfall, using the faulty zero filter by Hutten et al (2018) and the manual quality control

C4. Comparison with the baseline, entire rainfall event

On the event-scale the effect is even lower. However, this is too be expected since a larger dataset reduces the impact of outliers (which would be the faulty zeros). The filter could have been very impactful had it revealed an increased number of faulty Netatmo stations, therefore causing the interpolation to use fewer stations. But since this was not the case, the overall impact of the filter during this research was not significant.

	Median of the RMSD [mm/10 min]				Median of the Pearson correlation [-]			
Event nr.	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
1 2015-11-19	1.32	0.66	0.66	0.65	0.53	0.74	0.74	0.75
2 2016-06-23	3.35	2.55	2.9	2.84	0.66	0.71	0.59	0.61
3 2016-11-17	0.28	0.21	0.22	0.22	0.58	0.79	0.74	0.75
4 2017-07-25	0.84	0.48	0.47	0.47	0.44	0.59	0.66	0.65

Table C7: Median of the RMSD and the Pearson correlation of the four events using only the manual quality control

	Median of the RMSD [mm/10 min]				Median of the Pearson correlation [-]			
Event nr.	KNMI	TU Delft	Netatmo	Combined	KNMI	TU Delft	Netatmo	Combined
1 2015-11-19	1.32	0.66	0.64	0.64	0.53	0.74	0.75	0.75
2 2016-06-23	3.35	2.55	2.88	2.8	0.66	0.71	0.58	0.6
3 2016-11-17	0.28	0.21	0.22	0.22	0.58	0.79	0.74	0.75
4 2017-07-25	0.84	0.48	0.47	0.46	0.44	0.59	0.66	0.66

Table C8: Median of the RMSD and the Pearson correlation of the four events using the faulty zero filter by Hutten et al (2018) and the manual quality control