Rainfall monitoring with opportunistic sensors

Lotte W. de Vos

Propositions

1. Understanding uncertainties in high-quantity low-quality observations is only of interest to scientists.
(this thesis)
2. Opportunistic sensors cannot replace traditional dedicated rainfall sensing techniques.
(this thesis)
3. Faulty scientific hypotheses are not necessarily disproven by accurate observations.
4. If a scientist makes a discovery, but is not able to communicate it to other scientists, there has been no scientific advancement.
5. The idea that academic success is measured by how busy you are is a misconception, particularly destructive to young scientists.
6. The large-scale educational exchange between strangers via internet is a proof of altruism.
7. Everyone's opinion matters, but only opinions of informed individuals should matter in decision making.

Propositions belonging to the thesis, entitled

Rainfall monitoring with opportunistic sensors

Lotte de Vos

Wageningen, 6 December 2019

Rainfall monitoring with opportunistic sensors

Lotte Wilhelmina de Vos

Thesis committee

Promotor

Prof. Dr R. Uijlenhoet
Professor of Hydrology and Quantitative Water Management
Wageningen University & Research

Co-promotors

Dr H. Leijnse
Senior researcher, R&D Observations and Data Technology
Royal Netherlands Meteorological Institute, De Bilt

Dr A. Overeem
Researcher, R&D Observations and Data Technology
Royal Netherlands Meteorological Institute, De Bilt

Other members

Prof. Dr A.A.M. Holtslag, Wageningen University & Research
Dr J.A.E. ten Veldhuis, Delft University of Technology
Prof. Dr H. Kunstmann, University of Augsburg, Germany
Prof. Dr E. Morin, The Hebrew University of Jerusalem, Israel

This research was conducted under the auspices of the Graduate School for Socio-Economic and Natural Sciences of the Environment (SENSE).

Rainfall monitoring with opportunistic sensors

Lotte Wilhelmina de Vos

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Friday 6 December 2019

at 1:30 p.m. in the Aula.

Lotte W. de Vos
Rainfall monitoring with opportunistic sensors
xii + 178 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2019)
With references, with summary in English

ISBN 978-94-6395-159-3
DOI <https://doi.org/10.18174/502629>

*“I’m singin’ in the rain
Just singin’ in the rain
What a glorious feeling
I’m happy again.”*

– Gene Kelly

Summary

In this thesis the potential of using opportunistic sensing techniques to monitor rainfall at high temporal and spatial resolution is evaluated. The techniques that have been explored are rainfall measurements by personal weather stations (PWSs) that have been crowdsourced from online weather platforms, and rainfall estimates derived from power levels in commercial microwave link (CML) networks.

In **Chapter 2** time series of rain observations spanning at least 4 months in a 17-month period from 63 PWSs in the Amsterdam metropolitan area ($\sim 575 \text{ km}^2$) are crowdsourced from the Wundermap platform and validated with a gauge-adjusted radar product. The Pearson correlation coefficient of these measurements improve from 0.48 at 5 min temporal resolution to 0.60 at hourly intervals. Double mass curves show that the median of all stations resembles the reference better than the real-time (unadjusted) radar product, which heavily underestimates rainfall. The sensor performance of three Netatmo PWSs is evaluated in an experimental set-up. The raw observations correspond well with the high-end reference rain gauge (with a Pearson correlation coefficient of ~ 0.9 at 5-min intervals), but the time series obtained via the Wundermap are less accurate due to unknown processing steps on the platform. For high accuracy at sub-hourly resolution, it is therefore advisable to make use of unprocessed data.

In **Chapter 3** a two-year dataset of PWS rainfall observations is evaluated for the same Amsterdam study area as in Chapter 2, but now crowdsourced from the Netatmo Weathermap platform. Contrary to the PWS dataset in Chapter 2, these time series have not been affected by platform processes. Four types of typical errors in PWS rainfall observations are identified. A four-module quality control (QC) is proposed which can be employed in real time, without needing auxiliary data (apart for determining a one-value proxy for overall rainfall underestimation by the PWS network in an offline exercise), that can successfully identify these four types of errors in the PWS dataset. This QC was constructed using the first year of

the dataset, and subsequently applied to the second year and to a national dataset. The QC improves the overall accuracy of the second year of hourly rainfall depths in Amsterdam from a bias of -16.7% to +0.2%, a coefficient of variation (calculated as the standard deviation of the difference divided by the mean of the reference) of 3.74 to 2.86 and a Pearson correlation coefficient of 0.68 to 0.82, while maintaining 88% of the original dataset. With an average national PWS density of 1 station per $\sim 10 \text{ km}^2$ (although highly variable, related to population density), PWS data shows great promise for operational application, even at a national scale.

Chapter 4 presents a validation study on rainfall estimates from power levels in a CML network of almost 2000 unique links in the Netherlands over a period of ~ 7 months. The RAINLINK package has been made polarization-dependent and has been extended with a preprocessing module that enables processing of instantaneously sampled CML power levels. Path-averaged rainfall intensities are validated with path-averaged rainfall values calculated from a gauge-adjusted radar product. The coefficient of variation decreases from 3.70 to 2.32 and the Pearson correlation coefficient increases from 0.30 to 0.63 from instantaneous to daily estimates of rainfall accumulations. The accuracy improves during heavy rainfall events. Errors in rainfall estimates are largest for links that are shorter than 2 km and for observations that occur late at night and early in the morning as well as in winter months. This may be due to solid/melting precipitation and dew formation on the antennas. Comparisons with results from other CML datasets in the same climate indicate that min/max sampling is preferable to instantaneous sampling at 15-min temporal resolution.

Chapter 5 reports on simulations of the rainfall measurements of PWS, CML and radar based on rainfall fields described in terms of drop size distributions (DSDs) on a 100-m grid and 30-s time intervals. The resulting time series are constructed based on their respective sampling strategies and network lay-out in the $20 \text{ km} \times 20 \text{ km}$ study area in Amsterdam, and are subsequently validated with the original rainfall fields. Radar observations are mimicked from the average reflectivity (Z), from which rainfall intensity is estimated using a Z - R relationship, as well as a dual-pol strategy where R is determined from the spatially averaged specific differential phase (K_{dp}). The accuracy of radar rainfall estimates heavily relies on the appropriate choice of parameters in these relationships. Assuming perfect measurement accuracy, the PWS network is best able to capture small-scale rainfall dynamics, with a Pearson correlation coefficient above 0.86 and a relative bias below 4% for 100-m rainfall estimates at 5-min resolution. CML accuracy drastically improves at longer time steps. Both PWS and CML networks in the study area are more limited in their temporal sampling strategy than their spatial

extent, as shown by the limited reduction of accuracy for sparser networks: similar measurement accuracies are achieved with half the number of stations and links in the PWS and CML networks.

Chapter 6 showcases how various opportunistic sensors measure weather in the Amsterdam study area (same as in Chapter 2 and 3) for a 17-day period containing a passing front and a warm period during which the urban heat island effect is significant. Three opportunistic sensing networks are explored to measure six important weather variables: (1) air temperature estimates from smartphone battery temperatures and PWSs; (2) rainfall from CMLs and PWSs; (3) solar radiation from smartphones; (4) wind speed from PWSs; (5) air pressure from smartphones and PWSs; (6) air humidity from PWSs. The outcomes are compared with measurements from dedicated traditional sensors. Both the passage of the front and the urban heat island effect can be measured with opportunistic sensors, although with large uncertainties. Raw PWS measurements and CML rainfall estimates contain many errors, but are able to describe weather phenomena with spatial variability, while smart phone-derived observations are limited to city-averaged estimations. This study shows that there is potential for hydrometeorological monitoring in cities using opportunistic sensing.

The main conclusion from this thesis is that both PWS and CML can be used to monitor rainfall at high resolutions, with large accuracy improvements by employing a QC methodology for PWS rainfall observations and carefully considering the data sampling strategy in CML. As such, they can provide useful real-time available complementary information to traditional rainfall sensors like ground-based radar and operational rain gauge networks. Opportunistic sensors should not be seen as a replacement of these traditional techniques, but rather an additional source which can help to generate high-accuracy, high-resolution (urban) rainfall fields.

Contents

	Page
Summary	vii
Chapter 1 Introduction	1
Chapter 2 PWS rainfall validation in Amsterdam	15
Chapter 3 PWS rainfall quality control	39
Chapter 4 CML rainfall validation in the Netherlands	59
Chapter 5 Simulated PWS, CML and radar rainfall	87
Chapter 6 Opportunistic sensing case studies	117
Chapter 7 Synthesis	141
References	149
Statement of author contribution	165
Acknowledgements	167
Curriculum Vitae	169
List of publications & presentations	171
Graduate school certificate	175

Chapter 1

Introduction

1.1 Measuring rainfall

The hydrological cycle encompasses the fluxes of water on, above, and below the surface of the earth, of which rainfall is arguably the most noticeable aspect. Rain is a fresh water source that brings prosperity when available in the right amounts, but destruction both in its absence (droughts) and its abundance (floods). Its historical importance is reflected among others in the deities thought to be in charge of rain, such as Freyr (Norse mythology), Tlaloc (Aztec religion), Chaahk (Mayan religion) and Zeus (Greek mythology), who were also connected with lightning and the flourishing or destruction of crops. Because rain affects day-to-day life, it is not surprising that many have taken an interest in this weather phenomenon.

The first known mention of rain measurements dates back to 4 centuries BC, and for a long period of time consisted of evaluations of the rainwater amounts collected in buckets over time (see Strangeways, 2010). The past 250 years saw an increased interest in rain monitoring, with more inventive rain sensors, careful considerations concerning instruments and placing, and the first long records of rainfall observations (Strangeways, 2010). Rain is typically measured in the following ways.

1.1.1 Rain buckets

The most straightforward way to measure rainfall is by collecting rainwater in a bucket in order to measure the quantity thereof afterwards. This is a low-tech approach, feasible with cheap devices. Emptying the gauge and logging the amount of rainwater requires manual actions, which makes this technique labor-intensive, vulnerable to data gaps and human errors (Daly et al., 2007) and limits the temporal resolution of the measurements to the frequency of emptying the gauge. Some of these non-recording devices are equipped with a (subterranean) reservoir (Strangeways, 2010), enabling the storage of cumulative rainfall over long periods of time, e.g. yearly.

Recording rain gauges do not require systematic manual interference (apart from regular inspection for maintenance) as they measure rainfall continuously. As with manual gauges, rain is collected in a collection funnel. Recording of rainfall is done in various ways (Nystuen et al., 1996; Lanza & Vuerich, 2009), namely by (1) mechanical structures that produce graphs on paper connected to the amount of collected rain water over a period of time (pluviometer, e.g. Fabian, 1966), (2) electronic signals related to either the elevation of a floater in the collector unit (e.g. Fig. 2 in Wauben, 2004) or the weight of rainwater in the collector unit (e.g.

Plank et al., 1984), or (3) a tipping bucket mechanism that relays a signal once the amount of rainfall corresponding with the tipping bucket volume is reached (e.g. Baer, 1976). When devices are equipped with heating elements, solid precipitation can be measured as well (although with possible underestimation of precipitation due to increased evaporation).

The quality of rain gauge measurements is affected by sampling (i.e. tipping bucket volume, weighing resolution, etc.) and location considerations. Errors can result from wind effects around the gauge such that especially small droplets do not fall into the funnel, in-splashing from the surrounding area and evaporation of rain before it is measured. Siting recommendations by the World Meteorological Organization (WMO) for minimizing these errors are to place the rain gauge in a level location (i.e. avoid roofs and sloped hills), where nearby obstacles are not higher (above the gauge orifice) than twice the distance to the gauge, with a set-up that allows for horizontal wind-flow above the orifice by (a) having homogeneous dense vegetation at the same elevation as the gauge orifice, (b) mimicking ‘(a)’ with fence structures, or (c) using windshields (Jarraud, 2008).

1.1.2 Disdrometers

In addition to the rain gauges that base their measurements on collected water amounts, there are devices that measure rainfall based on signals from rain drops. Disdrometers (= **d**istribution of **d**rops **m**eter) measure rain based on the sound of impact by rain drops on the device (acoustic, e.g. Wolf et al., 2017), the voltage induced in a sensing coil upon impact (impact, e.g. Joss & Waldvogel, 1967), 2D video recordings of falling droplets (light-sheet, e.g. Kruger & Krajewski, 2002) or the length and size of the interference in an laser beam created by falling hydrometeors (optical, e.g. Illingworth & Stevens, 1987). These measurements capture the terminal velocity of droplets and describe rainfall as drop size distribution (DSD), from which rainfall amounts can be derived. This additional layer of detail can be useful when evaluating physical aspects of rainfall. Also worth mentioning are two historical experiments measuring DSDs through the counting and sizing of rain drop impressions on flour (Laws & Parsons, 1943) and paper (Marshall & Palmer, 1948), although for obvious reasons these techniques are not suitable for operational implementation.

1.1.3 Ground-based weather radar

Radar (= **radio detection and ranging**) is a technique to detect airborne objects using microwave radiation signals. These signals are emitted from a radar antenna, typically installed in a tower, through scans at increasing elevations. When the signal, often transmitted in the form of pulses, interacts with obstacles in the atmosphere, part of the signal returns and is received by the radar. The delay and change of the returning signal indicate the location, velocity and size of the encountered objects in the measurement volume. The radar beam widens and reaches higher elevations with distance from the tower. The combined radar scans can describe the atmosphere around the radar, in the form of radar reflectivity factors, which is a manifestation of DSD. The measured reflectivity can be translated to rainfall amounts. One commonly used relation between reflectivity (Z in $\text{mm}^6 \text{m}^{-3}$) and rainfall intensity (R in mm h^{-1}) is the Marshall-Palmer relation $Z = 200R^{1.6}$ (Marshall et al., 1955), but as the actual Z – R -relation is dependent on rainfall type and local climate many other relations have been proposed (e.g. Uijlenhoet, 2001; Raghavan, 2013).

The attainable radar range is related to the wavelength of the transmitted signal. Signals at short wavelengths attenuate quicker and are therefore typically used to scan only nearby, where radars can achieve higher measurement resolution. For weather observation, the most common wavelengths are ~ 10 cm (S-band), ~ 5 cm (C-band) and ~ 3 cm (X-band). The X-band radar is typically used for short-range purposes, e.g. for urban rainfall monitoring. C-band radar is common for operational weather monitoring in temperate climates, and yields rainfall observations at typically 1 km^2 and 5 min resolution. S-band is suitable for events with high rainfall intensities.

Dual-pol radar indicates that two types of signals are transmitted simultaneously: horizontally and vertically polarized. The combined received signals allow to derive more information about the precipitation. For instance, the ratio of the reflectivities of the horizontally and vertically polarized signals gives information on the average shape of the hydrometeors (which in case of rain is flattened while it is mostly spherical for signatures from tumbling hail). Also, the difference in phase shift between horizontal and vertical polarization provides information on droplet shapes and is a valuable tool to identify clutter (Seliga & Bringi, 1976). Dual polarization measurements allow for more sophisticated rainfall retrieval algorithms than merely assuming a Z – R -relation (Figueras i Ventura et al., 2012).

1.1.4 Satellite

Earth-observing satellites can be equipped with sensors to detect precipitation in various ways (Kidd & Huffman, 2011; Skofronick-Jackson et al., 2018; Sun et al., 2018). Cloud top altitude is a proxy for precipitation intensity and can be measured from infrared signals (e.g. Roebeling & Holleman, 2009). Liquid hydrometeors within a cloud emit microwave signals that can be detected with a radiometer (e.g. Spencer et al., 1989). In addition to these passive sensors, satellite-borne radars can be used to measure precipitation in the same way as the ground-based weather radar (e.g. Haynes et al., 2009).

Satellites enable rainfall monitoring at remote locations with limited to no ground-based observations (e.g. oceans), but the resolution of the rainfall observation is low compared to the other techniques. Moreover, the observations are constrained by the viewpoint from orbit of the satellite. While geostationary satellites measure continuously at the same location, low earth orbit satellites pass over periodically depending on their orbit, resulting in large temporal gaps in observations. On the other hand, the far larger elevation of geostationary satellites ($\sim 20,000$ km) typically yields a lower spatial measurement resolution than the low earth orbit satellites that measure from several hundred km elevation.

1.2 Sampling and combination techniques

Rainfall fields are highly variable in space and in time, and cannot be estimated as instantaneous point values: all rainfall observations span a certain amount of time (e.g. minute, hour, daily) and area (e.g. rain gauge collection funnel, atmospheric volume). The techniques with the highest coverage, radar and satellite-borne instruments, yield estimations of rainfall averaged over significant areas. Similarly, rain gauges and disdrometers provide accurate approximations of rainfall (averaged over a given time interval) at point-locations, i.e. they exhibit limited spatial representativeness. This difference in sampling between radar and rain gauges results in unavoidable differences in measurements even when both would be measuring with perfect accuracy (Einfalt et al., 2004; Villarini et al., 2008; Peleg et al., 2013). Also, while rainfall is measured directly by gauges, it is estimated indirectly by radar and satellite, with the introduction of conversion-related errors.

Rainfall observations from different point-scale sensors can be combined, whereby the (temporal and spatial) distance between observations is filled using interpolation methods, e.g. by inverse distance weighting or geostatistical models (Diggle &

Ribeiro, 2007). More observations tend to result in higher accuracy of the inferred rainfall fields. Combining rainfall observations with different types of sampling requires some more sophisticated interpolation techniques. For instance, Schuurmans et al. (2007) and Goudenhoofd & Delobbe (2009) evaluate the accuracy of rainfall fields constructed with radar and gauges with various merging techniques.

The Royal Netherlands Meteorological Institute (KNMI) measures rainfall in the Netherlands with two C-band Doppler dual-pol weather radars, a rain gauge network of 322 manual gauges (~ 1 station per 100 km^2) and a rain gauge network of 32 automatic rain gauges (~ 1 station per 1000 km^2), see Fig. 1.1.

Figure 1.1: Operational rainfall measurement techniques in the Netherlands employed by KNMI: automatic gauge network (left), manual gauge network (center) and C-band weather radar product (right), with respective measurement locations and time intervals. (After Overeem et al., 2016b)

By combining radar and gauge measurements, one can benefit both from the large coverage of the first, and the high accuracy at the ground-based point location from the second. Such products are constructed by KNMI. The reflectivity measured

by the two radars is combined using a weighing factor as a function of distance to the radars. This yields 5-min reflectivity observations covering the Netherlands, that are converted to rainfall intensities using the Marshall-Palmer Z - R -relation (Marshall et al., 1955) and accumulated to e.g. 5-min rainfall depths, see also Beekhuis & Mathijssen (2018). A climatological radar product is produced by adjusting the radar rainfall product with a mean field bias correction calculated with the (corrected) hourly rainfall depths from the automatic gauge network, with a spatial adjustment based on daily observations from the manual gauge network (both applied on sub-daily scale) (Overeem et al., 2009a,b). This gauge-adjusted rainfall product at 5-min intervals and spatial resolution of $\sim 1 \text{ km}^2$ contains information from all dedicated sensing techniques, and is therefore assumed to be the best available approximation of the true rainfall fields at ground level. The downside of this product is that it can only be constructed with a temporal delay in the order of a month due to the validation carried out on the manual rain gauge data.

1.3 The importance of resolution

In the previous section it is explained that a rain sensor provides information on rain at a certain location, while a network of sensors provides information on rainfall fields. With a large number of sensors spread out over an area, rainfall fields can be resolved at higher resolution.

Urban areas are characterized by heterogeneous land cover with a large fraction of impervious surface, preventing the infiltration of rain water, which leads to increased runoff. In many cities, the drainage system (e.g. storm sewers) is part of the urban design. In case of heavy rainfall events where not all water can be transported through these drainage systems fast enough, flooding may occur (Smith et al., 2002, 2013). This typically affects a large number of people due to the high population density in cities, and can result in significant financial damage (Ten Veldhuis & Clemens, 2010).

Urban drainage modeling can be employed to evaluate hydrodynamics in cities. Especially for small urban catchments, valuable model output requires an optimal combination of sufficiently high temporal and high spatial resolution rainfall input (Berne et al., 2004; Gires et al., 2012; Ochoa-Rodriguez et al., 2015; Thorndahl et al., 2017; Cristiano et al., 2017). The required resolution of rainfall observations for urban applications is determined in the range of km and minutes (Schilling, 1991; Einfalt et al., 2004; Emmanuel et al., 2012; Bruni et al., 2015), also related

to the rainfall type (convective vs stratiform), and the size of the catchment. All these authors point out that high resolution rainfall products are vital to be able to accurately reproduce the hydrological responses in cities.

1.4 Opportunistic sensing

Given the notion in the previous section, it can be concluded that more rainfall measurements are desirable, especially in-situ ground observations in cities. As opposed to expanding the traditional measurement techniques that have been described so far, which would require significant financial and labour investments, these rainfall observations can also be achieved by exploiting existing infrastructures and devices. Obtaining information from devices that have not been acquired or designed to measure large scale rainfall operationally, but can be used as such, is what we call opportunistic sensing (Uijlenhoet et al., 2018).

Weather observations can be human-based, like citizen weather observations provided via smartphone apps (Elmore et al., 2014; Guo et al., 2019), twitter (De Vasconcelos et al., 2016) and volunteers actively reporting daily rainfall at their location (Cifelli et al., 2005; Illingworth et al., 2014; Reges et al., 2016). Recent years have seen various types of rainfall information retrieved from technology as well, like readings from windshield wipers and optical sensors in cars (Rabiei et al., 2013) and rainfall intensity from camera images (Allamano et al., 2015; Jiang et al., 2019). Reviews on crowdsourcing atmospheric data in geophysics are provided by Muller et al. (2015) and Zheng et al. (2018). Although promising, these techniques are usually not yet at the stage of operational implementation due to non-guaranteed/delayed data availability, unknown data quality and/or the privacy-sensitivity of data from geo-located devices (cars, camera images) that may inadvertently convey personal information. However, there are two opportunistic sensing techniques that are not privacy-sensitive for citizens and show potential for large-scale, real-time implementation.

1.4.1 Personal weather stations

Personal weather stations (PWS) are often viewed as gadgets, bought by people interested in weather in their immediate surroundings. PWSs are opportunistic not in the sense that they were not intended to measure rainfall (they are), but in the sense that the willingness of weather station owners to link their PWS to online platforms enables crowdsourcing for large-scale rainfall monitoring. There

are several weather station platforms to which personal weather station owners can link their device. Fig. 1.2 shows the websites <https://wow.knmi.nl/>, <https://www.wunderground.com/wundermap> and <https://weathermap.netatmo.com/>, on which real-time observations are shared and visualized as colored symbols on a map. The platforms differ in the time intervals of measurements, station density, accessibility to the data and the amount of meta data available, but are otherwise quite similar.

Figure 1.2: Print-screens of online weather station platforms (from left to right): WOW-NL, Wundermap and Netatmo Weathermap, where each symbol represents a personal weather station and the amount of rainfall that was measured. Note that more station locations are revealed after zooming in on the platform web-page.

PWSs are typically low-cost compared to the gauges in networks operated by meteorological services, and the measurements are more likely to be prone to errors. This is also due to the fact that weather station owners often lack expert knowledge and/or sites for optimal placement of the sensors without interference with the surroundings. Besides experimental studies exploring PWS accuracy (Jenkins, 2014; Bell et al., 2015; Meier et al., 2017), the data has been used to measure rain (Golroudbary et al., 2018; Chen et al., 2018), wind (Droste et al., 2018), temperature (Meier et al., 2017; Chapman et al., 2017; Fenner et al., 2017; Golroudbary et al., 2018; Napoly et al., 2018; Mandement & Caumont, 2019), and pressure and humidity (Mandement & Caumont, 2019). PWS temperature measurements are very useful to quantify the effect that urban areas experience more heating than rural areas, especially at night (the so-called urban heat island effect), while the interpretation of PWS rain and wind measurements has proven more challenging.

Rainfall time series, obtained via online platforms, typically have infrequent time intervals, can contain large data gaps, and are prone to errors due to station set-up that can change instantaneously or gradually over the life time of the gauge. In order to make use of this dataset, a customized and more rigorous quality control is required than for other rain gauge networks. Where the number of stations is high, typically in densely populated areas (i.e. cities) in developed countries, this data source is still substantial after eliminating a considerable fraction of the data in a quality control procedure.

1.4.2 Commercial microwave links

The other promising opportunistic sensing technique for rainfall monitoring is from power levels in commercial microwave links (CML). Networks of CMLs are installed and maintained for the purpose of telecommunication. Telecommunication is considered essential enough that telecom operators install and maintain networks in most populated areas of the world, also those areas where financial means for rainfall sensors are limited.

Rainfall can be estimated based on the loss of microwave signal over a CML-path between a transmitting and receiving antenna (see Fig. 1.3 for an example). The strength of the signal at the receiving and transmitting end of the link path are usually monitored by telecom operators for quality control purposes. Upton et al. (2005) introduced the idea to use these CML observations to determine path-averaged rainfall, which was later proven to be successful with actual CML data (Messer et al., 2006; Leijnse et al., 2007b).

The relation between additional attenuation due to rainfall divided by link path length (specific attenuation k in dB km^{-1}) and link path-averaged rainfall intensity (R in mm h^{-1}) can be described by a power law: $R = ak^b$ (Atlas & Ulbrich, 1977). The coefficients a and b depend mainly on signal frequency and polarization (Olsen et al., 1978; Jameson, 1991). Errors in rainfall retrieval in case of variable rainfall over the link path are small for frequencies where b approaches 1 (Leijnse et al., 2010a; Overeem et al., 2011). It can be challenging to successfully identify the fraction of attenuation due to rainfall, because this is in addition to signal fluctuations during dry weather which can be caused by variations in atmospheric conditions. Additionally, wet films on the antennas result in so-called wet antenna attenuation, which is non-negligible, and increases with rainfall intensity (Leijnse et al., 2008).

Theoretically, there are no restrictions to the temporal resolutions and sampling

Figure 1.3: Example of a microwave link antenna tower, used in telecommunication networks (source: Pixabay).

strategy at which the power levels are logged and with which precision, however these factors are often decided by telecom operators. These sampling decisions are usually done without prioritizing rainfall estimation. Furthermore, it can be challenging to convince said telecom operators to share this data, as it is outside of their business scope. Fortunately, many telecom operators have successfully been convinced as shown by the large number of studies on CML datasets from many international locations, e.g. in Italy (Alberoni et al., 2018), Brazil (Rios Gaona et al., 2018), Pakistan (Sohail Afzal et al., 2018), Burkina Faso (Doumounia et al., 2014), Germany (Chwala et al., 2018), Israel (Lieberman et al., 2014), and many others.

This technique is especially promising in urban areas where the link density is higher than in rural areas, and in areas of the world where means for dedicated rain sensors are limited. Although the technology is fast-changing (e.g. changes in signal frequencies in future networks and replacements with fiber-optics in some areas), it is expected that CML networks operating at hydrometeor-sensitive frequencies will continue to cover large areas of the world in the coming decades (Ericsson, 2018).

To promote knowledge-sharing and operational use, several tools to estimate rain-

fall from CML data have been developed and made freely available: ‘Rcmlrain’ (<https://github.com/fenc1mar/Rcmlrain>), ‘RAINLINK’ (<https://github.com/overeem11/RAINLINK>) and ‘pycomlink’ (<https://github.com/pycomlink/pycomlink>), as well as ‘pySNMPdaq’ (<https://github.com/cchwala/pySNMPdaq>) to obtain CML data. The accuracy of CML rainfall observations has been explored by Leijnse et al. (2008); Zinevich et al. (2010); Overeem et al. (2011); Chwala et al. (2012) (among others), which is helpful in establishing the potential for operational use of this technique. Because it is possible to log power levels in real time, this technique allows for real-time rainfall monitoring (Chwala et al., 2016; Andersson et al., 2017; Chwala et al., 2018). Messer & Sendik (2015), Uijlenhoet et al. (2018), and Chwala & Kunstmann (2019) provide reviews on the details and background of this technique.

1.5 Aim and scope of this thesis

A few things become evident from the previous sections, namely that (1) a large number of additional rainfall observations are desirable, and (2) that they can be obtained with opportunistic sensors, (3) of which PWS and CML show promise for operational implementation. The aim of this thesis is twofold. Firstly, to explore the relative accuracy of these two opportunistic sensing techniques as compared to traditional methods. Secondly, to use the obtained insights for recommendations on using these techniques operationally. This thesis consists of this introduction (Chapter 1) and the following chapters.

Online weather platforms have steadily become more popular and user-friendly, resulting in a platform network that has become substantial only in recent years. The surge in Internet of Things (IoT) has facilitated this as well. The first study on rainfall data from a large number of crowdsourced urban weather stations is presented in Chapter 2. The potential of PWS data is explored by comparing rainfall observations with a gauge-adjusted radar product, and PWS accuracy is further explored in an experimental set-up.

Chapter 3 focuses on a different PWS dataset obtained in the same study area as in Chapter 2. The insights on error types that is obtained in the study described in Chapter 2 form the basis for a quality control methodology. The proposed algorithm is applicable in real-time, uses no auxiliary or metadata and successfully identifies the observations within an urban and national dataset that contain the errors that are typical for this data source.

The potential of CML is evaluated in a validation study on rainfall estimates from CML data from a national T-Mobile CML network in the Netherlands with a gauge-adjusted radar reference. In this study presented in Chapter 4, special attention is paid to the error dependency on known conditions like signal frequency, path length, time of observation, etc., in order to gain understanding of rainfall estimation accuracy under those conditions.

The relative accuracy of PWS, CML and the traditional technique, ground-based radar, is related to their respective sampling strategies. Using a rainfall simulation in the form of high-resolution DSD fields, the relative accuracy limits of each technique is determined in Chapter 5. As opposed to the previous chapters, using a simulation study allows for a known true rainfall field for validation instead of another device observation (i.e. affected by measurement errors) to compare with. Furthermore, the accuracy of the PWS and CML networks, whose lay-outs are mimicked from existing networks, is evaluated for reduced network densities.

Two case studies are presented in Chapter 6, where PWS and CML observations are used to describe two weather events, in addition to other (opportunistic) sensors. The presentation of raw findings of several types of opportunistic sensing techniques illustrates their respective value to measure weather events in an urban study area.

In the synthesis (Chapter 7), the main conclusions and highlights derived from this investigation are presented. These provide an overview of the potential of opportunistic sensing for high-resolution operational rainfall monitoring, with directions for future research.

Chapter 2

PWS rainfall validation in Amsterdam

The high density of built-up areas and resulting imperviousness of the land surface makes urban areas vulnerable to extreme rainfall, which can lead to considerable damage. In order to design and manage cities to be able to deal with the growing number of extreme rainfall events, rainfall data are required at higher temporal and spatial resolutions than those needed for rural catchments. However, the density of operational rainfall monitoring networks managed by local or national authorities is typically low in urban areas. A growing number of automatic personal weather stations (PWSs) link rainfall measurements to online platforms. Here, we examine the potential of such crowdsourced datasets for obtaining the desired resolution and quality of rainfall measurements for the capital of the Netherlands. Data from 63 stations in Amsterdam ($\sim 575 \text{ km}^2$) that measure rainfall over at least 4 months in a 17-month period are evaluated. In addition, a detailed assessment is made of three Netatmo stations, the largest contributor to this dataset, in an experimental setup. The sensor performance in the experimental setup and the density of the PWS network are promising. However, features in the online platforms, like rounding and thresholds, cause changes from the original time series, resulting in considerable errors in the datasets obtained. These errors are especially large during low-intensity rainfall, although they can be reduced by accumulating rainfall over longer intervals. Accumulation improves the correlation coefficient

with gauge-adjusted radar data from 0.48 at 5 min intervals to 0.60 at hourly intervals. Spatial rainfall correlation functions derived from PWS data show much more small-scale variability than those based on gauge-adjusted radar data and those found in similar research using dedicated rain gauge networks. This can largely be attributed to the noise in the PWS data resulting from both the measurement setup and the processes occurring in the data transfer to the online PWS platform. A double mass comparison with gauge-adjusted radar data shows that the median of the stations resembles the rainfall reference better than the real-time (unadjusted) radar product. Averaging nearby raw PWS measurements further improves the match with gauge-adjusted radar data in that area. These results confirm that the growing number of internet-connected PWSs could successfully be used for urban rainfall monitoring.

This chapter was originally published as:

De Vos, L.W., Leijnse, H., Overeem, A., & Uijlenhoet, R. (2017). The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. *Hydrology and Earth System Sciences*, 21, 765–777. doi:10.5194/hess-21-765-2017

2.1 Introduction

Urban catchments are characterized by a high proportion of impervious surfaces, leading to a large fraction of rainfall producing direct runoff and a fast hydrological response. This makes cities especially vulnerable to flooding. The temporal and spatial resolutions of rainfall data required for urban applications exceed those needed for rural catchments (Schilling, 1991). The rainfall information at spatial and temporal resolutions of typically 1 km by 1 km and 5 min generated by most operational weather radars is considered valuable for urban hydrological analysis and forecasting (Liguori et al., 2012). However, radar has significant limitations; rainfall is determined indirectly, over an atmospheric volume with a size depending on the distance from the radar station, which may not be representative for rainfall at ground level (Einfalt et al., 2004; Peleg et al., 2013). Errors in rainfall estimates from radar due to sampling uncertainties can be significant. In addition, there is an optimum spatial resolution corresponding to a given temporal resolution (Fabry et al., 1994; Bell & Moore, 2000). Rain gauges, if well maintained, provide accurate ground-based measurements, although they are limited in their spatial representation; Villarini et al. (2008) showed that approximations of true spatial rainfall fields with rain gauges requires a dense network and/or large temporal measurement intervals.

Hydrological models, designed to deal with high-resolution input, provide the best simulation results not just when the temporal resolution or the spatial resolution is high, but particularly when the combination thereof is optimal. The required spatiotemporal resolutions for urban applications have been studied extensively. Berne et al. (2004) determined a relation between the space–time resolution required for hydrological applications as a function of the catchment size for Mediterranean conditions. It was found that for urban catchments in the order of 10 km², rainfall data are needed at a temporal resolution of 5 min and a spatial resolution of 3 km. For urban catchments of 1 km², these resolutions were 3 min and 2 km, respectively. The space–time scales of four types of rainfall are evaluated by Emmanuel et al. (2012). With the use of variograms of 24 storm events, the spatial resolutions required to capture these types of rainfall at urban scales range from 0.8 to 3 km for instantaneous monitoring and from 2.5 to 8 km for 30 min intervals.

Gires et al. (2012) found an outflow uncertainty of up to 20% in an urban catchment of 9 km² due to rainfall variability at scales smaller than the typical C-band radar resolution of 1 km by 1 km and 5 min. Bruni et al. (2015) addressed the loss in urban hydrodynamic model accuracy due to smoothing and smearing. Radar data of four storm events in 1 min temporal resolution were aggregated to various spatial and temporal resolutions (highest range resolution of 30 m) and used as precipitation input in a 3.4 km² Dutch urban catchment. Smoothing occurs when the ratio of radar resolution over catchment size becomes larger than 0.2 and storms that move near the catchment boundary are averaged partly out of the catchment. Smearing becomes significant when the ratio of the spatial

resolution of radar measurements over the rainfall correlation length exceeds 0.9, leading to averaging of rainfall over the coarse spatial grid and resulting in underestimation of rainfall rates in areas within the storm cells and overestimation in the surrounding areas. Also, a runoff peak time shift of up to 6 min was found due to temporal aggregation (from 1 min to 5 and 10 min) of rainfall input.

Ochoa-Rodriguez et al. (2015) evaluated the required spatial and temporal resolutions of rainfall in a simple spatiotemporal scaling framework. A spatial resolution of 1 km, typically found in radar, was found to give good hydrodynamic model results, although some extremes were missed. Temporal resolutions should ideally be below the 5 min intervals currently available in most operational weather radar products. Nevertheless, the accuracy of 5 min radar data can be improved with the use of an accumulation procedure that assumes constant velocity of the rainfall field and rainfall intensity to vary linearly in time (Fabry et al., 1994). Coarsening temporal resolution has more impact on the accuracy than coarsening spatial resolution. Initial results from an ongoing study by the authors indicate that this impact is reduced when temporal resolutions are coarsened through aggregation (i.e., similar to rain gauges) instead of sampling. Lobligeois et al. (2014) evaluated the circumstances where hydrological model performance is enhanced by higher spatial resolution of rainfall. They did so by comparing lumped and semi-distributed models with subcatchment sizes of 64, 16 and 4 km². From comparisons between the various model outputs and observations in 181 catchments in France, it was found that model accuracy improvement depends on scale, catchment and event characteristics, and that the spatial representation of rainfall can be a highly important factor in the model performance.

From these works it becomes evident that an increase of the number of measurements would yield a higher accuracy of rainfall fields and would improve hydrological applications. Adding sensors (rain gauges or others) to a network is costly, although there are alternatives. For instance, rain maps can be produced from received signal strength in cellular communication networks, as the microwave signals propagating over the link paths are attenuated by rainfall (Overeem et al., 2016b). Weather data can also be provided directly by crowdsourcing measurements from amateurs in various ways (Muller et al., 2015). A growing number of weather enthusiasts measure their local weather with automatic personal weather stations (PWSs). PWS accuracy on measuring temperature, relative humidity, radiation, pressure, rainfall, wind speed and direction has been evaluated for popular high-end expensive weather stations (Jenkins, 2014; Bell et al., 2015), as well as for the cheaper, user-friendly Netatmo type (temperature only) (Meier et al., 2015), which have grown rapidly in number over the past years. So far, weather stations have been used to obtain air temperature data to examine the urban heat island effect (Steenneveld et al., 2011; Wolters & Brandsma, 2012), although other meteorological variables, such as rainfall, are measured by some of these stations as well.

A large number of PWSs share data on online platforms, both on the owner's own initiative (Gharesifard & Wehn, 2016) or automatically as an intrinsic software feature of the

product (i.e., for Netatmo). Netatmo has its own online platform collecting and visualizing data from all operational Netatmo stations. The WunderMap of company Weather Underground is a similar online platform. Data from Netatmo stations are automatically linked to the WunderMap. Owners of other PWS types can actively transmit their measurements to this platform as well. A growing number of automatic weather stations are linked to these platforms; in May 2016 there were 258 personal weather stations linked to WunderMap in the Amsterdam metropolitan area ($\sim 575 \text{ km}^2$) alone (239 of type Netatmo), of which 83 stations measured rainfall (64 of type Netatmo). By contrast, the official national automatic weather station network in the Netherlands ($\sim 35,000 \text{ km}^2$) consists of 31 stations, and these are, as a rule, always located outside urban areas. Figure 2.1 shows the relative resolutions in the Netherlands of networks discussed in this paper. At many locations, the density of PWS stations collecting rainfall data far exceeds that of any realistic operational network implemented by national weather services or local authorities beyond experimental campaigns. As the online platforms collecting and sharing PWS weather data are not nation-bound, global rainfall measurements have become easily available, with especially high densities in western Europe, USA and Japan.

Although rainfall data availability with PWS networks is cause for optimism for urban hydrological applications, errors are expected to be larger than those in traditional measurements. PWSs come in many types, a large fraction of which are low cost with expected low sensor quality. In most cases, there is no information available on the PWS type, the installation setup, maintenance of the sensor or data postprocessing while transferring measurements to the online platform. Bell et al. (2013) examine the potential improvement on the UK's observational network with the real-time and local weather measurements of air temperature, relative humidity and pressure collected from WunderMap. The most critical issue was found to be the estimation of data quality. Validation procedures like range tests (i.e., a check whether the measurement is within predefined extremes limits) and internal consistency tests should be applied to precipitation data from automatic weather stations (Estévez et al., 2011). The integration of crowdsourced data with variable temporal resolutions in hydrological discharge modeling by accounting for different uncertainties for data of various sources has been addressed in recent research (Mazzoleni et al., 2017).

It becomes clear that urban applications would benefit from high-resolution rainfall measurements. The potential of crowdsourced PWS rainfall data for this purpose has not previously been explored. Using the existing PWS network requires minimal financial investment, and would therefore be an economically reasonable alternative to conventional techniques to increase measurement resolutions. This study aims to determine the added value of crowdsourcing automatic weather stations for urban rainfall monitoring. For this purpose, the most common PWS is tested in an experimental setup with a high-quality rain gauge reference. Additionally, a dataset of 63 crowdsourced PWS stations in Amsterdam is validated with a gridded dataset based on radar data, a manual network and a WMO-certified automatic rain gauge network. These combined results provide insight on the rainfall measurement accuracy of the most commonly used PWS, as well as any

Figure 2.1: Temporal and spatial resolution of unfiltered rainfall measurements in the Netherlands with PWS network obtained via Netatmo API, WunderMap API and the potential availability of Netatmo measurements, as well as the resolution of KNMI’s automatic and manual rainfall measurement network and radar product. The curve represents a relation between the temporal and the spatial resolution of rainfall measurement required for urban hydrology as determined by Berne et al. (2004) for Mediterranean climate, where the square represents the value for an urban catchment with surface area of 0.1 km^2 .

issues that occur in operational crowdsourcing of PWS rain measurements. Following this introduction is the Methods section, where Sect. 2.2.1 describes the data and Sect. 2.2.2 gives an outline to determine the achieved measurement scales and quality of PWS, respectively. The results of an experimental PWS setup, a comparison of a larger dataset in Amsterdam with gauge-adjusted radar data and an analysis on inter-gauge spatial correlation of this dataset are given in Sect. 2.3. Finally, a discussion on the state and future role of PWS networks in (urban) hydrological applications and conclusions are given in Sects. 2.4 and 2.5, respectively.

2.2 Methods

2.2.1 Data collection

Personal weather stations

From the WunderMap website, a dataset of 63 automatic weather stations located in the Amsterdam area ($\sim 575 \text{ km}^2$) has been retrieved. Stations were selected based on the availability of rainfall measurements, which should cover at least 4 months between December 2014 and April 2016. Of these stations, 49 are of brand Netatmo, 7 are of brand Davis and 7 are of other unspecified brands. No details on the devices are given. According to the product specifications provided by the manufacturer, the Netatmo rain gauges have a measurement range of $0.2\text{--}150 \text{ mm h}^{-1}$ with an accuracy of 1 mm h^{-1} . The plastic tipping buckets have a volume of 0.101 mm and a collecting funnel with a diameter of 13 cm . The rain gauge module communicates in a wireless manner to the Netatmo indoor module over distances up to 100 m . The number of tips in the previous interval is communicated every $\sim 5 \text{ min}$ from the indoor module to the online dashboard via a WiFi connection, where it can be monitored by the weather station owner. Simultaneously, the measurement is linked to the Netatmo weather map from which it is sent every $\sim 10 \text{ min}$ to the WunderMap. The WunderMap stations that contribute to the dataset are visualized in Fig. 2.2. The WunderMap platform collects the rainfall measurements and rewrites them into rainfall over the past hour and cumulative rainfall for that day. Daily rainfall is reported in increments of 0.1 inch ($= 0.254 \text{ mm}$), but provided in the mm amounts rounded to 1 decimal. Therefore, daily rainfall only becomes non-zero (i.e. 0.3 mm) once the 0.254 mm threshold is reached and subsequent rainfall is only reported as an additional 0.2 or 0.3 mm depending on the decimal rounding of the amount.

While Netatmo hardware can store measurements for a period of time in the event of bad connectivity with the server, only real-time data are automatically transferred to the WunderMap. This causes gaps in the WunderMap datasets where there may be none in the original Netatmo data, which are only accessible to the weather station owner. WunderMap time series are characterized by (large) gaps in the dataset and irregular measurement frequencies, though often 5 , 10 or 15 min . Also, the locations of Netatmo weather stations on the WunderMap are obtained from the settings at the Netatmo platform without notice to or confirmation from the PWS owner. Relocations of the station that are communicated to the Netatmo platform are not simultaneously adjusted on the WunderMap, potentially leading to large errors in sensor location.

We process the data obtained via WunderMap by calculating the difference in cumulative daily rainfall compared with the previous time step. Since these time steps are not fixed, this results in rainfall accumulations over time intervals of varying lengths. In order to obtain compatible time series, the rainfall is interpolated on a fixed timeline with constant steps, where constant rainfall within the original intervals is assumed. Original intervals

Figure 2.2: Locations and operational days (i.e., days with measurements) of Netatmo (squares) and other types (triangles) of PWSs, with the radar pixel grid in the Amsterdam metropolitan area. Inter-station distances are represented in the histogram, colored green for nearest neighbor distances. The background map is taken from OpenStreetMap (www.openstreetmap.org).

longer than 20 min are discarded. Faulty values in precipitation data from automatic weather stations can be identified with range tests and internal consistency tests (Estévez et al., 2011). As a first quality check, values of the interpolated time series are compared with the median rainfall of all stations for each time interval. Values exceeding this median by more than 50mm h^{-1} are excluded. Dry periods in the dataset are identified as periods of at least 24 h where the median of all PWS measurements indicate zero rainfall. If a PWS reports continuous zero rainfall for at least 12 h outside of this dry reference, the measurements in this dry period are considered as faulty zero rainfall measurements and are discarded. Finally, inter-gauge correlations are determined. If a low correlation (i.e., average and median < 0.21) is found between a station and all other stations, the entire time series for that station is excluded. Visual comparison with corresponding radar rainfall time series showed that a filter based on these criteria was suitable in excluding obviously incorrect data from the datasets. This filter could be applied in real time, although for operational uses beyond this dataset, adjustments will be required.

Radar

As rainfall reference, we use gauge-adjusted radar data from a climatological rainfall dataset by the Royal Netherlands Meteorological Institute (KNMI) (Overeem et al., 2009a,b, 2011), freely available as "Radar precipitation climatology" via <http://climate4impact.eu>. This dataset is based on data from two C-band Doppler weather radars in De Bilt and Den Helder and has a temporal resolution of 5 min and a spatial resolution of 0.92 km^2 , covering the entire land surface of the Netherlands. This radar makes volumetric scans in all directions, measuring instantaneous rainfall at a location every 5 min. In this product, radar composite images have been adjusted with rainfall measurements from the KNMI rain gauge networks (31 automatic and 325 manual gauges). For details on the method of adjusting, we refer to Overeem et al. (2009a,b, 2011). It should be noted that, due to their different representativeness, there can be significant differences between radar pixel areal rainfall and point rainfall (Schilling, 1991; Einfalt et al., 2004; Villarini et al., 2008; Peleg et al., 2013). Using a radar product that is adjusted with ground measurements will likely reduce this difference.

Netatmo experimental set-up

As the majority of the weather stations linked to the WunderMap is of type Netatmo, we examine the quality of Netatmo rain gauges in a dedicated experimental setup; see Fig. 2.3, photo inset. As reference, we use a high-quality KNMI pit gauge at the Cabauw Experimental Site for Atmospheric Research (CESAR) (Leijnse et al., 2010a), that measures cumulative rainfall in intervals of 12 s. This electronic rain gauge is placed in a so-called pit gauge configuration: a small hill of diameter 6.2 m with a circular pit with diameter 3m and a depth of 40 cm in the middle. Precipitation is collected in the instrument (collecting funnel with a diameter of 16 cm, i.e., 200 cm^2) and in the event of solid precipitation melted by a heating element in the funnel. The amount of liquid water is measured by the position of a floating unit connected to a potentiometer. Rainfall is measured every 12 s within the range of 0–0.7 mm with a resolution of 0.1 mm and an accuracy of 0.2 mm. The Netatmo sensors are placed at $\sim 40 \text{ cm}$ around the electronic sensor in the center of the pit in such a way that the top of each sensor is level with the rim of the pit. The period considered is from 12 February to 25 May 2016. The datasets, as collected directly from the Netatmo personal account in millimeters of rainfall per interval of typically 5 min, as well as via the WunderMap platform, are compared to the pit gauge reference. One of the stations was offline between 20 April and 1 May, and one station could not be accessed via Weather Underground.

Figure 2.3: (a) Cumulative rainfall according to reference pit gauge, gauge-adjusted radar, Netatmo stations (N1, N2 and N3) and Netatmo stations obtained via WunderMap (W2 and W3). N2 (and, as a consequence, W2) was offline between 20 April and 1 May. The photo shows the experimental setup of the rain gauges in the pit gauge configuration. (b) Scatter plots of 10 min rainfall and linear fits of rainfall according to N1, N2 and N3 as compared to reference pit gauge (orange) and gauge-adjusted radar (blue).

2.2.2 Analysis

Station measurement density

As mentioned previously, the original PWS data temporal resolution from WunderMap is quite irregular. The number of stations containing rainfall measurements for time series per 5 and 10 min shows that the data availability is quite variable; see Fig. 2.4. Moreover, the fraction of the measurements over the period that is filtered out does not seem to vary significantly in time. It is not straightforward how to attribute a certain measurement resolution to a network that has highly irregular measurement frequencies and station locations at irregular distances from one another. When the Amsterdam area is divided into grid cells, or pixels, of a certain size, the number of pixels that contain at least one measurement is an indication of the network resolution. The fraction of total pixels that contain at least one measurement has been calculated for all time steps over the entire period, for various combinations of pixel sizes and time step lengths in the scale range relevant for urban applications. It is found that for the Amsterdam dataset (before filter has been applied), the fraction of pixels containing at least one measurement is more limited by the number of stations than the measurement frequency; see Fig. 2.5. Only

when dividing the period in time steps shorter than 10 min, an increase of measurement frequency will result in a higher fraction. This is unsurprising as most stations in the dataset link their measurements to the Weather Underground platform approximately every 10 min. Adding stations will result in an increase in fraction at all time step sizes in this range. The PWS network consists of more stations than the number examined in this dataset and continues to grow, which will have a positive effect on the PWS network measurement resolution.

Figure 2.4: Number of stations with rainfall data from the PWS dataset, before and after applying filter, for every 5 and 10 min interval over the entire period, smoothed per day. The two indicated dips correspond to complete outage of stations, the third with a longer period of fewer measurements in all stations.

Station measurement quality

With the Netatmo experimental setup, the performance of this type of PWS and the consequences of transferring its data to the online platform are examined. The measurements are compared to the high-resolution pit gauge as well as to the radar rainfall at the corresponding pixel. These two comparisons should give an indication of the differences due to sensor performance and those due to differences in representativeness of radar and rain gauges.

Rainfall measurements of the PWS dataset in Amsterdam are compared with the radar rainfall measurement at their corresponding radar pixels. When comparing station data

Figure 2.5: Indicated with curves as well as colors are the fractions of pixels containing at least one measurement of the unfiltered dataset for combinations of time step length and pixel grid size over the Amsterdam area between December 2014 and April 2016.

with gauge-adjusted radar data, the coefficient of variation of the residuals (CV) is calculated. The standard deviation of the differences between the datasets is divided by the mean of the gauge-adjusted radar data. A low value of CV indicates a good match between the datasets. Additionally, spatial correlations between stations are estimated with the use of Pearson correlation coefficient (r):

$$r = \frac{E[XY] - E[X]E[Y]}{\sqrt{(E[X^2] - E[X]^2) \cdot (E[Y^2] - E[Y]^2)}}, \quad (2.1)$$

where $E[\cdot]$ is the expectation (estimated as the arithmetic mean) and (X, Y) are corresponding time series of rainfall measurements. Because of the spatial and temporal variability of rainfall, the correlation of two point locations decreases with distance between these points. A three-parameter exponential function is suggested by Habib et al. (2001a) to describe this spatial dependency relation between inter-station correlation (r) and distance (d):

$$r = r_0 \exp \left[- \left(\frac{d}{X_0} \right)^{S_0} \right], \quad (2.2)$$

where r_0 is the nugget parameter, X_0 is the correlation distance and S_0 is the shape factor.

The nugget parameter r_0 is a measure of small-scale variability and/or measurement error and is equal to 1 for perfect zero-distance correlation. Correlation distance X_0 indicates the distance at which the rainfall decorrelates (i.e., the distance beyond which the correlation drops below e^{-1}), which should be interpreted with caution when it exceeds the investigated spatial extent.

The relationship in Eq. 2.2 is sensitive to rainfall extremes (Habib et al., 2001a), climatic regimes (Krajewski et al., 2003) and seasonality (Van de Beek et al., 2011; Tokay & Öztürk, 2012) as well as strongly dependent on time interval (Krajewski et al., 2003; Ciach & Krajewski, 2006; Van de Beek et al., 2011; Tokay & Öztürk, 2012; Van de Beek et al., 2012; Peleg et al., 2013). For the PWS dataset in Amsterdam, correlograms are constructed and compared with spatial dependencies found in literature. Special consideration is given to the correlations between Netatmo stations as compared to the other types of rain gauges.

2.3 Results

2.3.1 Netatmo comparison with pit gauge

The original data of three Netatmo stations (measurement frequency of ~ 5 min) are compared with pit gauge data (measurement frequency of 12 s) and gauge-adjusted radar data (measurement frequency of 5 min), over the period February–May 2016. Over this period, the cumulative rainfall of station 2 was lower than that of the others; see Fig. 2.3a. This was the result of station outage. In general, the Netatmo stations measure less rainfall than the pit gauge and radar reference over this period. The scatter plots in Fig. 2.3b do not include the intervals where one or both of the time series contain no measurements (in the event of station outage), and show a good r^2 of 0.94 between Netatmo measurements and the pit gauge reference. Even though this r^2 suggests a small measurement error in Netatmo, the comparison with radar shows significant scatter away from the perfect fit. This is inherent to comparisons between point locations and pixel averages, and the scatter plot resembles those reported in Peleg et al. (2013), though the radar value used there was an average value of 12 pixels instead of 1.

The correlation between Netatmo and the pit gauge is calculated for a multitude of accumulation intervals; see Fig. 2.6. This correlation reflects small-scale rainfall variability and thus is closely related to the nugget parameter in Eq. 2.2. As expected, an increase of correlation is found for larger accumulation intervals. However, the correlations of data from the same devices obtained via WunderMap with the same pit gauge reference show far lower values; see Fig. 2.6. The original Netatmo data have typical time steps of 5 min against 10 min for the WunderMap data. If this was the only difference between the time series, the correlation graphs should overlap for accumulation intervals above 10 min. As

they only approach one another for hourly accumulations, it can be concluded that besides this effect, additional information is lost in the transfer of data between platforms.

Figure 2.6: Correlation between rainfall measurements by Netatmo stations as obtained via personal dashboard (N1, N2 and N3), as well as those obtained via WunderMap (W2 and W3), and the pit gauge reference for various accumulation time steps.

Besides the Netatmo dashboard (available to the station owner) and WunderMap, Netatmo data are also accessible from the Netatmo weather map platform. In this research, real-time measurements from the three stations in the experimental setup were obtained with from this platform with an application programming interface (API). It was found that rainfall measurements from this dataset were attributed with a time stamp of the moment the data were collected, instead of the time stamp of the measurement itself. In the event of sensor outage, the last available measurement was collected repeatedly. These artifacts resulted in faulty interval attribution of rainfall and negatively affected the correlations with the original dataset as well as with gauge-adjusted radar data. An API containing such processing errors will result in datasets that contain considerable errors, though these errors are easily overlooked without the original data. Fortunately, the original data can also be obtained from the Netatmo platform. These time series are identical to the data from the Netatmo dashboard (N1, N2 and N3 from the experimental setup) and can be obtained in real time.

2.3.2 Amsterdam weather station comparison with radar

Figure 2.7 shows the double mass plot of the filtered PWS dataset in Amsterdam, as well as the unadjusted (real-time) radar with the gauge-adjusted radar reference at the same locations. The only intervals considered are those where both time series contain measurements. Even though individual stations often do not follow the diagonal line representing a perfect match, the median of all available stations only shows a slight underestimation as compared to the gauge-adjusted radar rainfall data. This underestimation is far greater in the real-time radar product. Though large deviations occur, the median of the stations resembles the reference quite well.

Figure 2.7: Double mass plots of station filtered rainfall measurements and real-time radar data with gauge-adjusted radar rainfall at the corresponding location in the period between December 2014 and March 2016. Only intervals where both radar and station contain measurements are taken into account. Colored regions indicate the range between double mass plot of stations with minimum and maximum steepness and dashed lines represent the median of the combined datasets..

When comparing station rainfall against corresponding gauge-adjusted radar rainfall data over the entire period with the condition that the radar measures non-zero rainfall, a better correspondence is found for longer time steps; see Fig. 2.8. A similar scatter as in Fig. 2.8 is found by Peleg et al. (2013). At longer accumulation intervals, the averages resemble each other more, the CV decreases and the r increases, indicating a better resemblance

between gauge-adjusted radar and station datasets.

Figure 2.8: Scatter density plots of all station rainfall measurements against the gauge-adjusted radar rainfall data in the corresponding radar pixel when radar reported non-zero rainfall (> 0.1 mm). The $\bar{R}_{radars}$, $\bar{R}_{stations}$, CV , r and n values in the panels represent the average rainfall according to the gauge-adjusted radar data, the average rainfall according to the stations, the coefficient of variation of the residuals, the correlation and the number of intervals, respectively. Graphs are made for 5 min, 30 min and hourly accumulation intervals.

2.3.3 Amsterdam center average comparison

In order to investigate whether the generally poor quality of individual PWS measurements can (partly) be compensated by the generally high quantity of measurements, averages of unfiltered PWS measurements are compared with radar pixel averages over a small area in Amsterdam. The selected area is the region with highest parking rates; the densely populated and touristic area of the city center and Museum square, as floods in this area will heavily impact residents, businesses and tourism alike. This region of ~ 20 km² is shown in Fig. 2.9, where the cumulative rainfall of each station relative to the mean of the 12 stations is shown. From Fig. 2.9 the variation between station measurements becomes evident. Some stations measure highly unlikely values considering the measurements of their nearby stations, such as station 3, 9 and 12.

The means of all possible subsets of the 12 PWSs are compared with the average of the 20 radar pixels over the selected Amsterdam center region. For each subset, the correlation, standard deviation and coefficient of variation of the residuals (CV) of rainfall intensity is calculated over all intervals where each station contains measurements. The resulting outcomes of each subset are represented with boxplots in Fig. 2.10 per number of stations contributing to the PWS-mean. The correlation increases and the standard deviation and CV decrease when averaging multiple stations, even when some of the station time series consist of obviously faulty measurements, see Fig. 2.10. By averaging the unfiltered measurements of a dozen stations, crowdsourced measurements seem to be

Figure 2.9: (a) Locations of 12 stations and 20 radar pixels in the city center of Amsterdam, where symbol size represents the number of unique days with measurements by the station (range of 371–514 days). The background map is taken from OpenStreetMap (www.openstreetmap.org). (b) Double mass plots of the station measurements as compared to the mean of all 12 stations over the intervals where all stations contain measurements.

able to describe rainfall in the city center. As expected, the values based on 60 minute rainfall intensities show a better correspondence with gauge-adjusted radar data than 5 minute rainfall intensities.

2.3.4 Amsterdam weather station spatial correlations

Rainfall variability is often described with correlograms; see Sect. 2.2.2, describing Pearson correlation coefficient between station pairs as a function of distance. Correlograms of PWS data at longer accumulation intervals show higher inter-station correlations and the decrease with distance is not as steep; see Fig. 2.11. This is similar to the results reported by Villarini et al. (2008), Peleg et al. (2013) and Tokay & Öztürk (2012). Especially in winter (see upper panels of Fig. 2.11) and for short accumulation intervals, the non-Netatmo pairs show higher correlation with one another. However, the goodness of fit of the correlograms differs significantly from those found by Villarini et al. (2008), Peleg et al. (2013) and Tokay & Öztürk (2012).

The correlations of all station pairs in the dataset are fitted with the relation in Eq. 2.2. Fitting was done by determining the nonlinear (weighted) least-squares estimates of the parameters with the Gauss–Newton algorithm. The resulting parameters for the total

Figure 2.10: Box plots of correlation, standard deviation and coefficient of variation of residuals (CV) of averaged rainfall intensity time series. The box plots contain the outcomes for all possible subsets within the 12 stations in the Amsterdam city center, as compared to the gauge-adjusted radar rainfall intensity 20-pixel mean for interpolated 5 min and hourly time series.

dataset, as well as winter and summer individually, are given in Fig. 2.12. The graphs for winter show the most deviating response, suggesting irregularities in this subset in particular. The nugget parameter r_0 of the total dataset varies between 0.50 and 0.67 for this accumulation interval range. Villarini et al. (2008) found a similar nugget parameter of 0.51 for 1 min accumulations, though with far larger values at higher accumulation intervals. The nuggets found by Krajewski et al. (2003) (0.95–0.97 for 15 min and longer), Ciach & Krajewski (2006) (0.995 and higher for 1 min and longer), Tokay & Öztürk (2012) (0.97 and higher for 5 min and longer) and Peleg et al. (2013) (0.92 and higher for 1 min and longer), are all considerably higher than the nugget parameters found here. This is not surprising as the gauges in the networks evaluated in those papers are carefully controlled and of higher sensor quality than typical PWSs.

The correlation distance of the total PWS dataset increases with interval size in a similar manner as in previous research; see Fig. 2.12. The erratic response of the winter graphs suggests a poor fit resulting from other factors than rainfall variability. Likely the correlation distance of stratiform winter rainfall is larger than the spatial scale examined here. The shape parameters do not seem to follow an obvious dependence, similar to Peleg et al. (2013), though other research found this parameter to increase with interval size (Krajewski et al., 2003; Ciach & Krajewski, 2006; Villarini et al., 2008; Tokay & Öztürk, 2012).

Figure 2.11: Correlograms of all stations after filtering at various accumulation intervals for winter (top panels) and summer (bottom panels). The red and blue areas represent the interquartile range of the Netatmo stations and non-Netatmo stations, respectively. The areas are constructed with a moving window of width 5 km. The scatter plots are fitted with the exponential relation of Eq. 2.2, the parameters of which are given in the panels.

Figure 2.12: Timescale dependency of nugget (a), correlation distance (b) and shape factor (c) parameters from fit described in Eq. 2.2, for the total PWS dataset, as well as winter and summer only. Dotted lines represent values found in previous research by Peleg et al. (2013) (violet), Villarini et al. (2008) (orange), Tokay & Öztürk (2012) (brown) and Ciach & Krajewski (2006) (purple), where the dashed line in the first panel shows the timescale dependency of the Netatmo station nugget found in the experimental setup as previously shown in Fig. 2.6.

2.4 Discussion

In the experimental setup in Cabauw, the immediate overlying radar pixel that was first considered as reference turned out to show a significant bias as compared to gauge-adjusted radar rainfall data in all neighboring pixels. The next nearest pixel to the setup was then used as reference instead. The distance between radar pixel center and experimental setup thereby increased slightly from 428.9 to 473.5 m. Faulty measurements can occur in the gauge-adjusted radar dataset, which should be kept in mind when it is used as a reference. When comparing the Amsterdam area radar pixels used in this research to their combined mean value over the 17-month period, individual time series showed up to 10% consistent higher or lower values. Biases in gauge-adjusted radar could result in a larger spread in Fig. 2.8, although they have a far smaller influence on the results found in Fig. 2.10 as, in that case, the values are averaged.

Each aspect of this research, i.e., the Netatmo experimental setup, the analysis of the station data obtained with the Netatmo API and the Amsterdam PWS dataset from WunderMap, concerned time series over a different, though partly overlapping, time period. As the shorter time series were examined with the purpose of identifying artifacts in the data, those conclusions can be carried over to the longer, more robust analyses. The results on PWS data availability (see Figs. 2.1 and 2.5) do not take measurement quality into account. Because of the faulty attribution of rainfall to measurement intervals due to rounding in the data transfer, the measurements in the current form should be accumulated to larger intervals to reduce errors, although this reduces the temporal resolution appreciably. It would be more desirable to address the collection method of the PWS data in the platforms in order to maintain the quality of the original PWS rainfall measurements before data transfer.

The filter applied on the PWS dataset in this paper was based on all stations in the dataset. For operational purposes, the median value that is used as a selection criterion should be based on nearby stations only. Large rainfall values were excluded based on a limit on maximum rainfall of 50 mm h^{-1} above the median rainfall at all PWSs at that interval. This potentially excludes rainfall with plausible return times: we take the example of a 10 min interval during which the median rain intensity of the stations is 4 mm h^{-1} . A measurement of 54 mm h^{-1} and higher would then be excluded, though this corresponds to an event that would occur statistically every 1.5 years (Buishand & Wijngaard, 2007). Because of the small spatial scales and the lack of extremely heavy precipitation in this dataset, the current filter was applicable, as confirmed by visual comparison with gauge-adjusted radar data.

Although a large fraction of the PWS networks consists of Netatmo stations, this does not imply similar performance of these datasets. Factors like placement and maintenance are unknown and not necessarily equally interfering with the measurements. Even less metadata is available on the other PWS types in the dataset, since information on data

transfer and the sensors used is not provided for those PWSs. It is expected that there is a positive correlation between the purchase costs of the PWS and the importance of maintenance and high-quality measurements to its owner, although this assumption could not be examined with our dataset. Furthermore, the location of the station is based on the setting provided by the PWS owner, although these may be faulty due to inaccurate localization, rounding of the longitude and latitude or relocation of the station at a later time. Even when relocations of PWSs are accurately provided to the Netatmo platform, this is not automatically communicated to WunderMap, resulting in inaccurate time series for that location. This issue is found to arise in the PWS dataset, though the filter criterion regarding minimum correlation with the other stations excludes time series of those stations entirely.

Different spatial correlation parameters between studies are to be expected due to different climates, rainfall types, gauge network density and quality. However, the nugget-parameter r_0 (1 for perfect correlation between time series) found here is significantly lower than in other studies. Additionally, the nugget values of the Amsterdam dataset are significantly lower than the correlation found between the Netatmo datasets with the electronic rain gauge reference in the experimental setup when the data were obtained via the WunderMap platform; see also Fig. 2.6. This suggests the interference of additional factors besides sensor measurement errors and data transfer rounding when rainfall measurements are gathered in a less controlled manner. Such factors could be measurement errors due to station placement and poor maintenance.

It is important to note that, even though gauge-adjusted radar rainfall is used as a rainfall reference, differences with point measurements are to be expected because of representativeness errors. Ideally, a high-density gauge network could be used to improve this rainfall product in the future. A non-identical match should therefore not directly be interpreted as negative. However, as the nugget parameter from the station analyses was considerably lower than could be explained by rainfall variability alone, differences with gauge-adjusted radar data here are likely mainly caused by errors in the PWS dataset. Besides data transfer errors that heavily influence the nugget parameter, the installation errors (e.g., due to shielding), that are minimized in the experimental setup, further decrease the nugget in the Amsterdam dataset. When comparing nuggets from the experimental setup and the Amsterdam dataset in the left panel in Fig. 2.12, the correlations found in the former do indeed reach higher values than those influenced by installation errors in the latter.

2.5 Conclusions

The resolution and quality of crowdsourced PWS rainfall measurements from the platform with the most dense PWS network were analyzed to establish whether this data source allows urban hydrological applications. Although the required resolutions (as described by

Schilling (1991), Berne et al. (2004), Emmanuel et al. (2012) and Ochoa-Rodriguez et al. (2015)) are not yet achieved by the current PWS networks, the density of these networks is expected to increase. As the resolution of the current network in Amsterdam is more limited in the spatial resolution than the temporal resolution, the expected continued growth of PWSs that share rainfall measurements via online platforms will yield a network approaching the desired resolutions. This offers a vast contrast compared to KNMI's automatic rain gauge network which, in the Amsterdam metropolitan area, only measures rainfall at one location outside of the city (at Schiphol airport).

From comparisons between Netatmo rainfall time series in an experimental setup that reduces the errors due to faulty installation to a minimum, the measurements closely resemble those from the high-resolution electronic rain gauge. Larger differences are found with radar rainfall, likely due to differences in representativeness between pixels and point measurements. Although the sensor performance of this largest contributor of data in the PWS network considered in this research looks promising, there is a significant loss in accuracy due to transfer of data to the online platform WunderMap. In this study, the daily cumulative rainfall values as obtained from WunderMap are rewritten as the difference in rainfall as compared to the previous time step. WunderMap cumulative daily rainfall can only become non-zero when at least 0.254 mm (reported as 0.3 mm) rainfall has been collected. Especially in the event of light rain, rainfall could occur for a longer period than the interval length in which the daily cumulative rainfall increases. The rainfall is then attributed to a single interval instead of all previous intervals in which it may have been raining as well. This causes significant errors at small timescales. These errors result in inter-station correlations that were considerably poorer than those found in literature, especially in winter and at short accumulation intervals.

The median rainfall of the Amsterdam PWS dataset shows less systematic bias than the real-time available radar product. Averaging PWS time series further improves correlation, standard deviation and coefficient of variation with the averaged gauge-adjusted radar rainfall in a certain region ($\sim 20 \text{ km}^2$). Provided that the degree and likelihood of overestimation of rainfall by PWSs is similar to the degree and likelihood of rainfall underestimation, as was the case in our Amsterdam city center dataset, a dense subset of PWSs can provide good rainfall estimation over a small area, even for intervals of 5 min and without applying a quality filter.

The largest obstacles for the use of crowdsourced PWS datasets are the errors resulting from data transfer, errors due to poor maintenance and faulty installations (i.e., at shielded locations). The rounding of cumulative daily rainfall measurements occurring in the WunderMap platform and the time stamp uncertainty of measurements obtained from platforms with faulty APIs can lead to considerable errors in the time series, which are only reduced at large accumulation intervals. For the purpose of a high-quality rainfall measurement network with PWS data, these issues need to be addressed first. Processing errors can be avoided by obtaining raw data from the Netatmo weather map platform, though the station density is slightly lower than that of the network linked to the Wun-

derMap. When the processing of data is no longer interfering with the quality of the datasets, the potential of PWS platforms becomes significant. It provides rainfall measurements from all over the world that are easy to collect, located in rural areas as well as in cities, with station densities and coverage exceeding those from national weather services, and growing towards a level matching the reported resolutions that are required for urban hydrological applications.

Chapter 3

PWS rainfall quality control

Automatic personal weather stations owned and maintained by weather enthusiasts provide spatially dense in-situ measurements that are often collected and visualized in real-time on online weather platforms. While the spatial and temporal resolution of this data source is high, its rainfall observations are prone to typical errors, currently preventing its large-scale, real-time application. This study proposes a quality control methodology consisting of four modules targeting these errors, applicable in real-time without requiring auxiliary measurements. The quality control improves the overall accuracy of a year of hourly rainfall depths in Amsterdam to a bias of -11.3% (0.2% when a proxy for overall rainfall underestimation by PWSs is used), a Pearson correlation coefficient of 0.82 and a coefficient of variation of 2.70, while maintaining 88% of the original dataset. Application on a national scale (average 1 station per $\sim 10 \text{ km}^2$) yields high-resolution nationwide rainfall maps, hence showing the great potential of personal weather stations for complementing existing often sparse traditional rain gauge networks.

This chapter was originally published as:

De Vos, L.W., Leijnse, H., Overeem, & Uijlenhoet, R., (2019). Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring. *Geophysical Research Letters*, 46, doi:10.1029/2019GL083731

The quality control R-code has been made freely available as ‘PWSQC’ on <https://github.com/LottededeVos/PWSQC> under cc-by-sa 4.0 license.

3.1 Introduction

Accurate rainfall monitoring is vital in understanding hydrological and meteorological processes. Rain gauge networks provide direct point-scale measurements. However, the combined orifices of gauges routinely used to produce global precipitation products span an area smaller than a soccer field, and most of these gauge measurements consist of daily observations (Kidd et al., 2017). This is problematic as high-resolution (in space and time) flood forecasting requires high-resolution precipitation input in order to produce meaningful results, especially in urban areas (Berne et al., 2004; Emmanuel et al., 2012; Ochoa-Rodriguez et al., 2015; Cristiano et al., 2017). Crowdsourcing has been investigated as a strategy to obtain more rainfall observations, ranging from studies exploring citizen observations collected via smartphone apps (Elmore et al., 2014; Guo et al., 2019), active daily rainfall amounts reported by volunteers (Cifelli et al., 2005; Illingworth et al., 2014; Reges et al., 2016), rainfall intensities from camera images (Allamano et al., 2015; Jiang et al., 2019), rain intensity and occurrence from car sensors (Rabiei et al., 2013), derived weather information from twitter messages (De Vasconcelos et al., 2016) to simulation studies incorporating those techniques (e.g. Mazzoleni et al., 2017; Yang & Ng, 2017). Muller et al. (2015) describe some of these and other crowdsourcing strategies to gain atmospheric data. Zheng et al. (2018) present a recent overview and state of the art of crowdsourcing data collection methods in geophysics.

Recent developments enable owners of automatic weather stations to easily monitor their environment and share weather observations in real time on online platforms. Popular online platforms such as Netatmo and Weather Underground collect and visualize measurements from personal weather stations (PWSs) every ~ 5 to 10 minutes. The average density of Netatmo PWSs measuring rainfall in the Netherlands is 1 per $\sim 10 \text{ km}^2$, while the national networks employed by the Royal Netherlands Meteorological Institute (KNMI) consist of a manual gauge every $\sim 100 \text{ km}^2$ and an automatic gauge every $\sim 1000 \text{ km}^2$. As PWS density is correlated with population density, this provides weather observations at high temporal and spatial resolution in urban areas particularly.

Previous studies investigated the accuracy of common PWS devices (Jenkins, 2014; Bell et al., 2015; Meier et al., 2017, Chapter 2) and made use of this data source to quantify the urban heat island effect (Meier et al., 2017; Chapman et al., 2017; Fenner et al., 2017; Golroudbary et al., 2018; Napoly et al., 2018). Rain measurement from Weather Underground PWSs are explored in Chapter 2 (63 PWSs in Amsterdam), Golroudbary et al. (2018) (11 PWSs in the Netherlands) and Chen et al. (2018) (11 PWSs in Norfolk, Virginia).

Inaccurate rain observations can be due to (1) instrumental errors, (2) compromised set-up, and (3) data processing issues. Quality control (QC) methods are designed to exclude inaccurate measurements. QC can consist of comparisons with auxiliary data (e.g. Qi et al., 2016), pre-set (dynamic) thresholds to exclude unlikely values (e.g. Estévez et al.,

2011), or internal consistency between stations and/or in time (e.g. Zahumenskỳ, 2004; Chen et al., 2018). PWS rainfall data is arguably highly prone to errors as the typically low-cost devices are often installed without knowledge of or access to optimal set-up locations, and are not regularly maintained. Measurement accuracy can change suddenly, e.g. due to hindrance of tipping bucket mechanisms by clogging or tilted set-ups after windy weather, which can be resolved just as suddenly. Notwithstanding their enormous potential for operational rainfall monitoring, these sources of error currently prevent the large-scale, real-time application of PWSs in meteorology and hydrology.

This paper, for the first time, explores an unprecedented large dataset in terms of length (2 years), covered area and density, and shows that accurate, nationwide rainfall maps can be constructed from crowdsourced PWS rainfall measurements. For this purpose we propose a real-time applicable QC method, consisting of a set of quality filters that excludes inaccurate observations, requiring no auxiliary data source or metadata (besides station location). We show the ability of this filter to correctly flag measurement intervals with typical errors for this data source, and unflag once the PWS produces reliable values again. Nationwide rainfall maps constructed from the filtered dataset show remarkable similarities with a reference radar/rain gauge dataset.

3.2 Method

3.2.1 PWS dataset

Two extensive datasets of PWS rainfall observations were obtained from the Netatmo Weathermap (<https://weathermap.netatmo.com/>). Netatmo PWSs consist of an indoor and an outdoor module (~ 160 euro) measuring temperature, relative humidity and (indoor) sound, barometric pressure and CO_2 levels, with optional additional modules for wind and rain (~ 65 euro each). The rain module is a plastic tipping bucket with a collection funnel, 13 cm in diameter, that reports the number of tips via a wireless connection of up to 100 m to the indoor module. This indoor module broadcasts all observations to the platform every ~ 5 min from the moment the station becomes operational resulting in ~ 5 min time series. From here measurements can be accessed via smartphone or tablet and visualized on the online Weathermap. Netatmo gives rain observations as multiples of 0.101 mm, or, in multiples of the tipping bucket volume that is determined by the weather station owner using the calibration feature of the device. Approximately 13.5% of Dutch rain gauges are manually calibrated, as shown by deviating tipping bucket volumes. Netatmo rain gauges have a measurement range of $0.2\text{--}150\text{ mm h}^{-1}$ and an accuracy of 1 mm h^{-1} according to the manufacturer specifications. It was shown in Chapter 2 that Netatmo devices in an experimental set-up with a collocated well-calibrated operational reference rain gauge measure rainfall with high accuracy when installed properly and using unrounded measurements, as in this dataset.

Two PWS datasets are analyzed in this work:

- Urban dataset: All PWSs with a rain module in the Amsterdam metropolitan area, defined as the area between $4.67^\circ - 5.05^\circ$ longitude and $52.24^\circ - 52.44^\circ$ latitude ($\sim 575 \text{ km}^2$) between 1 May 2016 – 1 June 2018.
- National dataset: All PWSs with a rain module within the Netherlands for the month May 2018.

From PWS observations we construct time series at 5-min and at hourly intervals. The first year of the urban dataset is used as calibration dataset (CAL) to design the QC algorithm. The QC was subsequently applied on the second year (VAL) and on the national dataset (NL) to illustrate the ability of the filters to independently identify inaccurate observations. For both CAL and VAL the QC starts one month before the study period of one year to allow for the warm-up period in the filters.

3.2.2 Reference dataset

The reference is a climatological dataset that covers the entire land surface of the Netherlands in pixels of $\sim 1 \text{ km}^2$ at 5 min temporal resolution, freely accessible on https://data.knmi.nl/datasets/rad_nl25_rac_mfbs_em_5min/2.0. This rainfall product is based on two C-band radars, adjusted with two rain gauge networks (31 automatic and 325 manual gauges). Detailed information on radars and processing are provided by Beekhuis & Mathijssen (2018), and on the methodology by Overeem et al. (2009a,b, 2011). This radar product, adjusted with quality controlled high-end daily and sub-daily ground observations, is considered the most accurate rainfall reference but only becomes available with a delay of 1–2 months.

3.2.3 Validation

In order to validate observations, the Pearson correlation coefficient (r), the relative bias (bias from now on) and the coefficient of variation of the errors (CV) are calculated using the following equations:

$$r = \frac{\text{cov}(R_{\text{PWS}}, R_{\text{ref}})}{\text{sd}(R_{\text{PWS}}) \text{sd}(R_{\text{ref}})} \quad (3.1)$$

$$\text{bias} = \frac{\overline{\Delta R}}{\overline{R}_{\text{ref}}} \quad (3.2)$$

with:

$$\Delta R = R_{\text{PWS}} - R_{\text{ref}} \quad (3.3)$$

$$CV = \frac{sd(\Delta R)}{\bar{R}_{ref}} \quad (3.4)$$

where R_{PWS} are rainfall time series aggregated at 5-min or hourly temporal resolutions from PWSs (mm) and R_{ref} are the corresponding reference time series in the overlying radar-pixel (mm) where the bar indicates the average of observations of all stations at all time intervals.

3.3 Quality control

3.3.1 Types of errors

Sampling and representativeness error

The crowdsourced rainfall time series have variable time intervals in which the number of tipping bucket tips since the last timestamp are reported. In addition to the intrinsic tipping bucket error where rain can be attributed to a later timestamp (Habib et al., 2001b), additional errors result from gaps in the time series during connectivity problems. The sampling error in rain gauges with tipping bucket volumes of 0.101 mm, measuring in ~ 5 min intervals, has been determined in a simulation exercise with 12-s resolution electronic rain gauge observations as basis and ground-truth. This yields an r of ~ 0.96 and a CV of ~ 2.29 (see Sect. 3.5, Fig. 3.5). Using the radar product for validation introduces an additional error due to gauge-pixel discrepancy (and sources of error in radar rainfall retrieval), which reduces the similarity to an r of ~ 0.75 and a CV of ~ 5.0 .

Bias

Unbiased PWS measurements rely on an unshielded set-up where all raindrops reach the collection funnel, and a level gauge so that the tipping bucket mechanism is not hindered. However, completely exposed rain gauges are known to suffer from wind-induced underestimation (Pollock et al., 2018). Bias can also be due to the actual tipping bucket volume of the gauge not corresponding with the reported value due to manufacturing variability or faulty calibration. Netatmo PWS owners can calibrate their gauge by pouring a known amount of water through and calculating the tipping volume from the number of tips. If water is poured too quickly, some water bypasses the tipping mechanism during each tip, resulting in overestimation of the tipping volume. The majority of calibrated Netatmo PWSs in the nationwide dataset has an estimated tipping bucket volume larger than the default ($11.5\% > 0.101$ mm and $2.0\% < 0.101$ mm). While bias differs greatly between stations, the PWS-network has an overall tendency to underestimate rainfall (bias of -11.1% in VAL, see Table 3.2).

Faulty zeroes (FZ)

When the tipping bucket mechanism is obstructed completely, due to tilted rain gauge or physical obstructions (i.e. leaves, insects, solid precipitation, etc.), no tip will occur. Thus, only zero amounts are communicated to the platform, even during rain events.

High influx (HI)

A PWS can also report large amounts of rainfall unrelated to weather, e.g. by people pouring liquids through the rain gauge for cleaning, handling of the device with tilting movements, or sprinklers in the vicinity.

Station outlier (SO)

Sometimes PWS measurements do not correspond with local rainfall dynamics. This is true when the reported station location is incorrect. Also, some rare occasions have been observed where, for a period of time, rainfall is recorded in repeated daily cumulative amounts, thus resulting in far too high values.

3.3.2 Filter design

The CAL dataset was used to design the FZ-, HI-, SO-filter and bias correction (Sect. 3.3.1). Detailed flow charts of each filter are provided in Sect. 3.5. The general concepts are explained in this section, with the names of the 11 parameters underlined. Each PWS time interval receives a flag of either 0, 1 or -1 for ‘no error’, ‘error’, or ‘not enough information available to determine error’, respectively.

FZ-filter

All stations within a range ($\underline{d}$) around a given station are selected to compute the median rainfall over the surrounding area. If fewer than $\underline{n}_{\text{stat}}$ neighboring stations with rainfall measurements are available, the median cannot be calculated and the FZ-flag is set to -1 . The FZ-flag is set to 1 if this median rainfall is larger than zero for at least $\underline{n}_{\text{int}}$ time intervals while the station itself reports zero rainfall. The FZ-flag remains 1 until the station reports nonzero rainfall.

HI-filter

Unrealistically high rainfall amounts are determined based on a comparison with the median rainfall amount from all stations within a range ($\underline{d}$) around a given station. If the median does not exceed a threshold value ($\underline{\phi}_A$), the HI-flag is set to 1 for any rainfall value from the station itself above threshold $\underline{\phi}_B$. When the surrounding stations report moderate to heavy rainfall, the threshold becomes variable: for a median of $\underline{\phi}_A$ or higher, the stations' HI-flag is set to 1 when its measurements exceed median times $\underline{\phi}_B/\underline{\phi}_A$. HI-flag is set to -1 if fewer than $\underline{n}_{\text{stat}}$ neighboring stations report observations.

Bias correction & SO-filter

First, a default bias correction factor (DBC) is determined to address the fact that the Netatmo rain gauges have a general tendency to underestimate rainfall. DBC is a single-value one-off proxy of the correction needed for the overall PWS-network bias, and can be determined a priori by comparing network measurements over a period with typical rainfall for the local climate. In this study we base it on the median of all PWS bias values of 5-min observations during the month preceding the start of the dataset, when compared with the gauge-adjusted radar values, excluding intervals where FZ or HZ were not 0. DBC becomes:

$$\underline{\text{DBC}} = \frac{1}{1 + \text{median}(\text{bias})} \quad (3.5)$$

These DBCs (Table 3.1) correspond with the negative bias that was found for Wunderground PWS rain data in the same area (Fig. 2.7). Note that although we use another dataset for this step, it occurs in an offline exercise, meaning no auxiliary data is needed for the bias correction in the operational, real-time QC. If no reference is available, DBC could be set to 1 or to a value that was found elsewhere for this particular instrument in a similar climate. Initially, the bias correction factor (BCF) for each PWS is equal to DBC. In order to test the performance without the availability of any reference data, DBC=1 is also applied.

To determine whether a station yields nonsensical measurements for that location, it is compared with time series of neighboring stations within a range ($\underline{d}$). A previous period of $\underline{m}_{\text{int}}$ intervals, or any longer interval where the station has at least $\underline{m}_{\text{rain}}$ intervals of nonzero rainfall measurements, is evaluated. There need to be at least $\underline{n}_{\text{stat}}$ stations with at least $\underline{m}_{\text{match}}$ intervals overlapping with the evaluated station to compute the SO-flag. The r (Eq. 3.1) and bias (Eq. 3.2) with all neighboring stations are calculated. If the median of the r -values falls short of threshold $\underline{\gamma}$, the SO-flag is set to 1. If this threshold is exceeded, BCF_{new} is computed from the median of the bias values with the neighboring stations. If $|\log(\text{BCF}_{\text{new}}/\text{BCF}_{\text{prev}})| > \log(1+\underline{\beta})$, this is deemed a systematic change for that station and BCF_{prev} is replaced with the new value. This is hence a way to dynamically update BCF for individual stations.

Table 3.1: Parameter settings for QC, in detail explained in Sect. 3.3.2. The independent default bias correction (DBC) values were determined from the bias values in the urban dataset in the preceding month, i.e. May 2016 for CAL (June 2016 – June 2017), May 2017 for VAL (June 2017 – June 2018) and April 2018 for the NL dataset (May 2018).

Filter parameter	Value
d (m)	10,000
n_{stat}	5
n_{int}	6
ϕ_A (mm)	0.4
ϕ_B (mm)	10
m_{int}	4,032
m_{rain}	100
m_{match}	200
γ	0.15
β	0.2
DBC [CAL]	1.24
DBC [VAL]	1.13
DBC [NL]	1.13

Parameter choices

The chosen values for the 11 parameters are given in Table 3.1. Several sets of parameters were evaluated, and the best one was chosen based on the achieved improvement (see Sect. 3.2.3) for the CAL dataset after QC, while aiming for large applicability (i.e. fraction of flags = -1 small) and without flagging abundantly (i.e. fraction of flags = 1 small).

The QC principle applies to gauge networks in general, although the parameter values should be considered carefully for each network separately. For a sparser network, a larger d parameter and/or lower n_{stat} are needed to select enough neighbor stations. The number of values used to construct the median can be limited if n_{stat} is small, possibly resulting in outlier values. Higher values for n_{int} , m_{int} , m_{rain} and m_{match} result in more robust subsets of data on which flags and bias corrections are determined, at the cost of a longer unflagged warm-up period and more cases in which flags cannot be attributed. Most rainfall observations that should be targeted by the HI-filter were found to be very high, thus small variations in ϕ_A and ϕ_B hardly affect the results. Higher γ yields more SO-flags and lower β results in more frequent bias correction factor adjustments (and possibly overfitting).

3.4 Filter performance

When cumulative rainfall over a full year as measured by the PWSs is plotted against the cumulative amount according to the reference, FZ, HI and SO errors are shown as horizontal line segments, vertical line segments and fluctuating lines deviating from the diagonal, respectively. Figure 3.1 shows that the QC attributes flags to the time intervals causing these horizontal, vertical and fluctuating line segments. Stations can be susceptible to bias, seen in Fig. 3.1 as lines with slopes differing from the gray diagonal line.

Figure 3.1: Double mass plots of PWS observations against their respective reference, for year 1 (CAL) and year 2 (VAL) of the Amsterdam dataset, for raw data, subsets where intervals with FZ, HI or SO flag = 1 are excluded (“Filtered Flex”) and subsets where intervals with FZ, HI or SO flag not equal to 0 are excluded (“Filtered Strict”), including boxplots indicating the spread of the correlations and the fractions of the measurement intervals remaining for the stations after filtering.

The dataset can be filtered in two ways: retaining only intervals where no flag is 1 (“Flex”), or, retaining only intervals where all flags are 0 (“Strict”). After QC is applied, which includes station-specific bias correction, the remaining measurements correspond far better with the reference, i.e. the lines resemble the gray diagonal, especially for the Strict version, while the Flex version still includes some inaccurate measurements. The 87.2% and 88.0% of all intervals of the first and second year of the urban dataset (CAL and VAL, respectively) without any error flag show a dramatic improvement in accuracy

regarding bias, CV and r (Table 3.2). Each filter yields accuracy improvement, with the largest effect in the HI-filter given the small number of flagged intervals (Table 3.2). The comparison of the 5-min Strict-filtered VAL dataset with the gauge-adjusted radar yields huge improvements as compared to the raw data metrics, with a bias from -0.111 to 0.023 , a CV from 53.24 to 7.19 and an r from 0.07 to 0.58 , thus more closely resembling the upper limit of accuracy of rainfall data sampled in this manner of CV of 5.0 and r of 0.75 (see Sect. 3.5). The scatter density plots of all hourly raw and filtered (Flex and Strict) rainfall observations show this improvement as well (Fig. 3.2).

Figure 3.2: Scatter density plots of hourly PWS observations against their respective reference, for year 1 (CAL) and year 2 (VAL) of the Amsterdam dataset, for raw data, subsets where intervals with FZ, HI and/or SO flag = 1 are excluded (“Filtered Flex”) and subsets where intervals where FZ, HI or SO flag were not 0 are excluded (“Filtered Strict”).

The filters can be applied on a national scale. Rainfall patterns found by the PWSs correspond well with those from gauge-adjusted radar (Fig. 3.3). As the filters rely on neighbor checks, the QC is best applicable in the urban areas in the west where the PWS-network is densest (Fig. 3.4).

Table 3.2: Validation metrics and remaining fraction of original observations of 5 min and hourly PWS time series, before (Raw) and after QC of the individual filters (in Strict manner) as well as all combined filters applied (in both Flex & Strict manner), also when considering only the subset where reference exceeds a threshold of 0.1 mm and 0.5 mm for 5-min and 1-hour values, respectively. Metrics of VAL filtered (Strict) results are included where bias correction factors are calculated with $\underline{DBC}=1$.

Time interval	Dataset	Filter type	bias	CV	r	Remaining
5 min	CAL	Raw	1.39	147.08	0.04	100 %
		FZ-filtered	1.545	153.12	0.04	95.6 %
		HI-filtered	-0.061	12.95	0.35	99.9... %
		SO-filtered	0.003	18.1	0.27	88.3 %
		bias-corrected	0.052	51.52	0.11	100 %
		All filters - Flex	0.056	8.93	0.58	89.0 %
		All filters - Strict	0.053	8.79	0.59	87.2 %
	VAL	Raw	-0.111	53.24	0.07	100 %
		FZ-filtered	-0.044	55.46	0.08	94.0 %
		HI-filtered	-0.133	7.57	0.50	99.9... %
		SO-filtered	-0.076	55.86	0.08	89.8 %
		bias-corrected	-0.030	17.81	0.24	100 %
		All filters - Flex	0.021	7.22	0.58	89.2 %
		All filters - Strict	0.023	7.19	0.58	88.0 %
		All filters - Strict; DBC = 1	-0.095	6.68	0.58	88.0 %
	VAL, Ref > 0.1mm	Raw	-0.372	1.26	0.45	100 %
		FZ-filtered	-0.324	1.26	0.46	92.9 %
		HI-filtered	-0.373	1.25	0.45	99.9... %
		SO-filtered	-0.329	1.24	0.47	90.4 %
		bias-corrected	-0.299	1.35	0.45	100 %
		All filters - Flex	-0.236	1.32	0.48	88.9 %
		All filters - Strict	-0.234	1.32	0.48	88.0 %
1 hour	CAL	Raw	1.302	144.37	0.03	100 %
		FZ-filtered	1.475	152.03	0.03	95.6 %
		HI-filtered	-0.107	9.29	0.38	99.9... %
		SO-filtered	-0.024	15.91	0.25	88.5 %
		bias-corrected	0.031	16.54	0.27	100 %
		All filters - Flex	0.042	3.74	0.81	89.1 %
		All filters - Strict	0.040	3.62	0.82	87.3 %
	VAL	Raw	-0.167	3.74	0.68	100 %
		FZ-filtered	-0.099	3.67	0.71	94.0 %
		HI-filtered	-0.168	3.69	0.69	99.9... %
		SO-filtered	-0.131	3.25	0.75	89.9 %
		bias-corrected	-0.064	3.50	0.74	100 %
		All filters - Flex	0.001	2.88	0.82	89.3 %
		All filters - Strict	0.002	2.86	0.82	88.1 %
		All filters - Strict; DBC = 1	-0.113	2.70	0.82	88.1 %
	VAL, Ref > 0.5mm	Raw	-0.291	0.79	0.67	100 %
		FZ-filtered	-0.231	0.74	0.71	92.4 %
		HI-filtered	-0.291	0.79	0.67	99.9... %
		SO-filtered	-0.246	0.73	0.71	90.2 %
		bias-corrected	-0.199	0.83	0.68	100 %
		All filters - Flex	-0.126	0.74	0.74	88.3 %
		All filters - Strict	-0.125	0.73	0.74	87.5 %

Figure 3.3: Daily rainfall accumulations between 2018-05-29 08:00 UTC and 2018-05-30 08:00 UTC according to the gauge-adjusted radar product (left) and Flex filtered PWSs (right), interpolated using Ordinary Kriging with fitted variograms and nugget set to zero. Only stations with at least 95% data availability after QC during that day are included.

Figure 3.4: Density maps of the number of 5 min station measurements within 10 km per day, calculated between 2018-05-29 08:00 UTC and 2018-05-30 08:00 UTC, for all measurements (“Raw”), excluding intervals with FZ, HI or SO flag = 1 (“Filtered Flex”) and excluding intervals where FZ, HI and SO flags are not equal to 0 (“Filtered Strict”).

3.5 Discussion & conclusions

This study proposes a real-time applicable QC algorithm that does not require auxiliary or metadata. Rather than stations, time intervals are identified and flagged for errors related to faulty zero observations, high influxes, and station outliers. Additionally, dynamic bias correction is performed. The QC was designed on a full year of PWS measurements in the Amsterdam metropolitan area, and applied on measurements in the same study area during the subsequent year, as well as on a month of all PWS measurements covering the Netherlands. Results show large improvements of filtered data over raw measurements in the calibration, and even more in the validation dataset (bias of 0.023, CV 7.19 and r of 0.58, while retaining 88% of the original dataset at 5-min resolution), likely due to the higher data availability in the second year.

The QC is successful in flagging observations that are inaccurate, although as it relies on neighbor comparison it is better applicable on the urban dataset in Amsterdam than for other areas in the Netherlands with fewer PWSs. Also, it inherently assumes that nearby PWSs measure the same rainfall dynamics, which may be true in a relatively flat country like the Netherlands, but is far less likely in areas with systematic spatial rainfall gradients, e.g. due to mountains. This could be addressed by introducing elevation dependent neighbor selection, which only selects neighbor PWSs at similar elevations or orientations with respect to mountain ranges. This would require additional metadata.

Depending on the requirements of the resulting rainfall dataset, one may choose to make the QC more selective (at the expense of observation density) by decreasing $\underline{d}$, increasing n_{stat} , n_{int} , m_{int} , m_{rain} and m_{match} , or more inclusive the other way around (at the expense of accuracy). These parameters are related to the spatial and temporal scales of rainfall events, and should therefore correspond to typical rainfall variability in the local climate.

Although the QC does not need auxiliary data operationally, the $\underline{\text{DBC}}$ parameter was determined offline, based on bias between PWSs and gauge-adjusted radar reference. Ideally, $\underline{\text{DBC}}$ should be redetermined with reliable local rainfall information $\sim$ annually to address changes in overall network bias (due to loss and additions of PWSs, and accuracy changes over lifetime). If the bias correction module is implemented without this proxy ($\underline{\text{DBC}} = 1$) the QC results in rainfall estimates that are as good as with a calibrated value of $\underline{\text{DBC}}$, except, of course, for the overall bias (see Table 3.2).

In this study the filter was applied on crowdsourced PWS rainfall observations every 5 minutes, although the QC (with adjusted settings) will also be applicable at other time scales, resulting in time series with that time interval. The QC targets the errors that are typical for crowdsourced PWS networks measuring at variable time intervals. However, it can be applied successfully on any gauge networks with active periodic measurements. Sparser networks may be processed employing a longer comparison range, thus at longer time intervals. The next phase for this work is to make raw data, collected on platforms

maintained by commercial organizations, accessible in real-time for PWS networks in order for them to become a viable data source for rainfall monitoring. If this issue is addressed, a huge number of in-situ rainfall observations available in real time (mostly in developed regions of the world) can be used for various (operational) purposes, e.g. for PWS-adjusted radar products that, contrary to our validation dataset, can be available in (near) real time.

Background: default sampling errors and QC flow charts

This section consists of the extended simulation results on the measurement error due to sampling of a tipping bucket rain gauge (Fig. 3.5), scatter density plots of the hourly raw and filtered datasets against the gauge-adjusted radar reference product (Fig. 3.2) and the flowcharts of the QC filters for errors of types faulty zero (FZ), high influx (HI), and station outlier (SO) and bias correction (Fig. 3.6, 3.7 and 3.8).

Figure 3.5 illustrates the error in various crowdsourced rainfall measurement time series due to tipping bucket rain gauge sampling alone, by simulating time series from high-resolution rainfall observations. Our starting point was 1 year of KNMI electronic rain gauge observations at the Cabauw Experimental Site for Atmospheric Research (CESAR) with a resolution of 0.006 mm/h, at 12 s time intervals (Leijnse et al., 2010a, see also photo inset in Fig. 2.3: rain gauge in the center of the pit). This dataset was used to simulate new measurement time series by rounding and/or aggregating values to longer time intervals, as well as a ground truth. We simulated measurements from crowdsourced tipping bucket observations with various tipping bucket volumes (V_{tip}) and various sampling intervals, meaning the time between transmitting the number of tips since the previous call to the platform. These sampling intervals are typically ~ 5 min for time series obtained from the Netatmo platform. Crowdsourced rainfall observations from Weather Underground often have sampling intervals from ~ 10 min and sometimes ~ 5 min (Chapter 2). Temporal sampling errors are caused by tips attributed to later time intervals due to (much larger) V_{tip} compared to resolution of ground truth rain gauge. These errors are most evident when evaluating the time series at high temporal resolution. The accuracy of all simulated time series (with varying interval lengths) is calculated when the measurements are temporally interpolated at time series with fixed time intervals for various time interval lengths.

The benefit of using simulated time series to quantify sampling errors is that there is a known ground-truth, namely the original high-resolution data from the electronic KNMI rain gauge. The rainfall according to the 5-min gauge-adjusted radar product at the radar pixel in which a gauge is located is used as reference in this paper. To illustrate the additional errors due to the intrinsic mismatch between these techniques, the same

experiment is repeated using this reference dataset as the ground truth. Figure 3.5 shows that the error due to sampling at ~ 5 min resolution with tipping bucket volumes of 0.101 mm, as is the case for this study, yields a correlation (r) of 0.96 and coefficient of variation (CV) of 2.3 when evaluating at 5 min resolution, which deteriorate to an r of 0.75 and a CV of 5.0 when using the gauge-adjusted radar product as ground truth. Therefore, these second metrics should be seen as the upper limit of accuracy that can be achieved in the Netatmo time series after QC in the paper ($r = 0.58$ and $CV = 7.19$ over the year of the validation dataset in Amsterdam; see Table 3.2).

Figures 3.6, 3.7 and 3.8 show flow-charts of the QC attributes flags for FZ, HI and SO errors, and adjusts the bias correction factor. The framework visualized here forms the basis of QC that can be applied in real-time, and attributes a flag indication of 1 (error), 0 (no error) or -1 (not enough information to determine error) for each time interval. Depending on how severe one would want to apply the QC one may choose to include intervals with flag = -1 (“Filtered Flex”), or exclude them (“Filtered Strict”). The setting input (purple circles in the flow charts) should be adjusted for the typical climate, specific network density and sampling frequency on which the filter is applied.

Figure 3.5: Sampling error dependence of correlation (r) and coefficient of variation (CV) on tipping bucket volume (V_{tip} [mm]) and call frequency to platform (sampling intervals), evaluated at multiple temporal resolutions, validated both with known ground truth (left panels) and radar reference (right panels).

FZ filter

Figure 3.6: Flow chart of faulty zero intervals (FZ) filter.

HI filter

Figure 3.7: Flow chart of high influx intervals (HI) filter.

Chapter 4

CML rainfall validation in the Netherlands

Commercial microwave links are installed and maintained for the purpose of telecommunication. Hydrometeors between transmitting and receiving antennas cause the microwave signal to be attenuated. From signal attenuation, the path-averaged rainfall intensity can be calculated. A 7-month dataset of instantaneously logged signal powers from almost 2000 unique links in the Netherlands is analyzed. Rainfall intensities are calculated with the RAINLINK-package with a novel preprocessing module, enabling the package to be applied on instantaneously logged data from now on. Rainfall intensities per link are validated with the path-averaged rainfall intensities according to a gauge-adjusted radar product. Both the overall performance and the dependence of errors on link characteristics and measurement conditions are evaluated. The coefficient of variation decreases from 3.70 to 2.32 and the correlation increases from 0.30 to 0.63 from instantaneous to daily estimates of rainfall accumulations. The coefficient of variation is also smaller during heavy rainfall. Errors are largest for path lengths shorter than 2 km, for observations during the late night and early morning, and observations during colder months (when solid or melting precipitation could occur and dew is more likely to form on the antennas). Comparison of our results with those of earlier studies show that min/max sampling (widely employed in network management systems) outperforms instantaneous sampling regarding detection of both quantity and occurrence of rain at a 15 minute sampling rate in the Dutch climate.

This chapter was originally published as:

De Vos, L.W., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2019). Rainfall estimation accuracy of a nation-wide instantaneously sampling commercial microwave link network: error-dependency on known characteristics. *Journal of Atmospheric and Oceanic Technology*, 36, 1267–1283. doi:10.1175/JTECH-D-18-0197.1

4.1 Introduction

Rain gauge networks and weather radar are well-known and prevalent rainfall monitoring technologies. However, rainfall observations are sparse or lacking in large areas of the world. Sensors are costly to install and maintain. Especially developing countries and urban areas are likely to benefit from denser rainfall measurements than currently available (Schilling, 1991; Berne et al., 2004; Kidd et al., 2017).

Microwave links consist of a transmitting and a receiving antenna at several tens of meters above the ground and typically several km apart, between which microwave signals propagate. The links, owned and maintained by commercial telecom operators for the purpose of cellular communication, are abundant in populated areas. Microwave links covered 20% of the world's land surface in 2007 (GSMA, 2012). Upton et al. (2005) first suggested to make use of signal attenuation over commercial microwave links (CMLs) to determine rainfall, notably in urban areas. It was demonstrated that this could be achieved with actual commercial microwave link data (Messer et al., 2006; Leijnse et al., 2007b). This is promising as CMLs are widespread even in sparsely gauged areas (GSMA, 2016) and a CML rainfall product is found to outperform some satellite rain products (for the Netherlands) (Rios Gaona et al., 2017).

CML data is obtained by logging the power levels of mobile operators, which is typically done by determining the minimum and maximum power levels over a time interval (min/max sampling) or by logging instantaneous power levels every time interval. There have been numerous validation studies on several types of microwave link data in various climates (e.g. Goldshtein et al., 2009; Zinevich et al., 2009; Schleiss & Berne, 2010; Chwala et al., 2012; Doumounia et al., 2014; Kim & Kwon, 2018), and the technique is maturing towards becoming a real-time rainfall monitoring tool (Chwala et al., 2016). A good understanding of the accuracy of rainfall estimations from realistic CML networks is needed for its operational use. Several studies explored rainfall observations from large CML networks in the Netherlands with min/max 15 minute sampling (Overeem et al., 2011, 2013a; Rios Gaona et al., 2015; Overeem et al., 2016a,b; Van het Schip et al., 2017; Rios Gaona et al., 2017). Instantaneous sampling of signal levels is another common strategy for telecom operators, but to our knowledge, large validation studies were only conducted with min/max sampled datasets. This study, for the first time, validates CML observations of a country-wide CML network (almost 2000 links) over a long period (~7 months), where the data is sampled instantaneously.

The general CML rainfall retrieval principle is explained in Sect. 4.2. Many factors can influence the accuracy of rainfall estimates. This study aims to determine the overall accuracy of rainfall observations from instantaneously sampled CMLs and to identify the factors that affect that accuracy. Section 4.3 provides an overview of the known CML characteristics and their expected impact on rainfall estimation accuracy based on previous studies. Contrary to experimental or simulation studies, (e.g. Berne & Uijlenhoet,

2007; Leijnse et al., 2008, 2010b; Van Leth et al., 2018) these characteristics cannot be completely isolated in this dataset, but findings are indicative for large-scale occurrence. This is followed by the Methods in Sect. 4.4 detailing specifics of the instantaneous CML dataset and how rainfall estimates are calculated from it. The results in Sect. 4.5 are related to the expected behavior (from Sect. 4.3) in the discussion in Sect. 4.6, followed by the conclusions on the value and prospects of this measurement technique for rainfall retrieval in Sect. 4.7.

4.2 General rainfall retrieval principle

As telecommunication companies strive for a fully functional network at all times, avoiding too low signal-to-noise ratios is a point of attention. The transmitted signal levels (TSL) and received signal levels (RSL) are typically stored in the mobile operators' network management systems for the purpose of network quality monitoring. CMLs usually operate at frequencies where the signal is sensitive to hydrometeors, resulting in increased attenuation (TSL – RSL) during rainfall compared to the attenuation due to free-path loss only. The more raindrops and the larger their sizes, the larger the signal attenuation. From this signal attenuation with respect to dry weather, path-averaged rainfall intensity can be obtained via this power law (Atlas & Ulbrich, 1977):

$$R = ak^b \quad (4.1)$$

between specific signal attenuation (k in dB km^{-1}) and link path-averaged rainfall intensity (R in mm h^{-1}), where coefficients a (in $\text{mm h}^{-1} \text{dB}^{-b} \text{km}^b$) and b (-) are dependent on signal frequency and polarization (Olsen et al., 1978; Jameson, 1991). Leijnse et al. (2010b) determine these a and b values experimentally for Dutch conditions for a range of frequencies based on measured raindrop size distributions. Because $b \approx 1$ for typical operational microwave frequencies, the R - k relation is considered almost linear. That means that, in principle, the path-averaged rain rate can be determined irrespective of the variability of the rain rates along the path (Overeem et al., 2011).

Rain-induced specific attenuation k can be isolated from the total signal loss over the link path by subtracting the attenuation (TSL – RSL) representative of dry conditions (free-space loss) and the attenuation due to wet antennas due to rain (A_a in dB), and dividing that attenuation by the path length (L in km):

$$A = \text{TSL} - \text{RSL}; \quad k = \frac{A_{\text{wet}} - A_{\text{dry}} - A_a}{L} \quad (4.2)$$

Note that the wet antenna attenuation (A_a) will lead to considerable rainfall overestimation if not accounted for (Leijnse et al., 2008). Identification of dry intervals to determine the baseline attenuation A_{dry} (to avoid nonzero rain estimations during dry weather), the

so-called wet-dry classification, has been constructed in various ways: based on auxiliary data from e.g. ground-based radars or satellites (Overeem et al., 2011; Van het Schip et al., 2017), time series of a single link at high sampling frequencies (Upton et al., 2005; Schleiss & Berne, 2010; Kaufmann & Rieckermann, 2011; Chwala et al., 2012; Wang et al., 2012), or on observations of multiple links (Rahimi et al., 2003; Overeem et al., 2011; Rayitsfeld et al., 2012; Overeem et al., 2013a). Alternatively, Fenicia et al. (2012) obtain the fraction of the total attenuation that is caused by rain without a wet-dry classification, but calculate a variable baseline with a 1-parameter linear low-pass filter.

We refer to Messer & Sendik (2015); Uijlenhoet et al. (2018) and Chwala & Kunstmann (2019) for a comprehensive overview of research on microwave links for rainfall monitoring. Section 4.4.2 details the method to retrieve rainfall estimates from the dataset that was chosen in this study.

4.3 Error dependency on link characteristics and observation conditions

As stated in Sect. 4.2, path-averaged rainfall intensity can be calculated after extracting k from the total signal power loss, calculated as $\text{TSL} - \text{RSL}$. The accuracy of the resulting rainfall observations is affected by several factors whose impacts depend on each other. This section highlights these dependencies as reported in previous studies and formulates the expected impact for the CML dataset evaluated in this study.

4.3.1 Sampling strategy

In min/max sampling, the minimum and the maximum RSL over the time interval are recorded, where the TSL is often constant. Instantaneous sampling yields snapshots of RSL and TSL. Telecom providers may choose to vary TSL depending on RSL to ensure signal transfer. This does not negatively affect the rainfall retrieval as long as the changing TSL are logged, in the case of instantaneous sampling at the same time as the RSL, with similar accuracy and precision as the RSL (although in practice the TSL are often logged with less accuracy and precision, hence leading to a degradation in the quality of rainfall retrievals).

Different types of CML sampling strategies result in different errors. Leijnse et al. (2008) simulate microwave link signals from radar data. Three sampling strategies have been compared: (1) continuous powers are converted to rainfall intensities and averaged over 15 minutes, (2) averaged signal powers over 15 minutes are converted to rainfall intensities (somewhat similar to min/max sampling) and (3) instantaneous powers every 15 minutes are converted to rainfall intensities (i.e. instantaneous sampling). Based on the overall

simulation experiment, Leijnse et al. (2008) show the limitations of instantaneous sampling as compared to the other two strategies.

Ostrometzky et al. (2017) show that min/max sampling in combination with rounding of logged signal levels will yield a positive bias equal to the quantization interval in TSL plus twice the quantization interval in RSL. This bias is the same for both wet and dry periods. Therefore, by correcting with a baseline based on dry weather intervals, this bias is canceled out completely. Nevertheless, rounding of power levels will lead to errors that are most severe in cases where the rain-induced attenuation over the link path is small and hence the effects of rounding are large. This is the case for short links and low frequencies, especially for low rainfall intensities (Leijnse et al., 2008).

Although instantaneous sampling may be preferable at high sampling frequencies, at 15 minute intervals it is expected to miss some of the rainfall dynamics over the interval and to be outperformed by the min/max sampling strategy that uses information over the complete interval.

An extensive number of min/max sampled CML datasets have been evaluated in the Netherlands (Overeem et al., 2011, 2013a; Rios Gaona et al., 2015; Overeem et al., 2016a,b; Van het Schip et al., 2017; Rios Gaona et al., 2017). As the current instantaneously sampling network is employed in the same climatic conditions, this allows for a fair comparison between the two sampling strategies. The discussion includes a table with an in-depth overview of the previous CML studies carried out in the Netherlands, in which the dataset of the study by Overeem et al. (2016b) has been revisited to improve comparability with the current study.

4.3.2 Link characteristics

The accuracy of link rainfall estimates is influenced by frequency and path length in three distinctive ways.

1. Frequency dependent choice of a and b (Eq. 4.1)

When estimating R from k (Eq. 4.1), one needs to obtain a and b values that describe that relation for a certain polarization and frequency in that climate. The actual relation may not be as deterministic in reality, and assuming the relation as such can yield an error in estimated R . Berne & Uijlenhoet (2007) quantify this error in a simulation experiment to be nearly negligible (Mean Relative Error < 2%) for links paths longer than 5 km. Secondly, climatological standardized parameters, e.g. from the International Telecommunication Union (2005), are often the only available information, although this can yield very different rainfall estimates than from fitted a and b values using data from the region under study. Berne & Uijlenhoet (2007) show that assuming these climatological parameters can result in a Mean Relative Error up to 20% for frequencies of 15 GHz.

The a and b parameters used in this study were determined experimentally by Leijnse et al. (2010b) for a range of frequencies for both horizontal and vertical polarization in the Dutch climate. Leijnse et al. (2008, 2010b) showed a nearly perfect fit between k and R at 30 GHz and considerable scatter at 10 GHz. This suggests an expected larger error for the links at the lower end of the frequency range in this study.

2. Path length and rainfall variability

Path-averaged rain rate estimation relies on a near-linear R - k relation, where k is constant over the link path. The errors due to variable rain rates along the path will be largest for frequency and polarization combinations where b deviates from 1 (Leijnse et al., 2010b; Overeem et al., 2011), which is at frequencies considerably higher or lower than 32 GHz in Dutch conditions (Leijnse et al., 2008). Overeem et al. (2011) estimate that the error due to rainfall variability over the link path varies between $\sim +2.2\%$ for 13 GHz and $\sim -0.4\%$ for 39 GHz.

3. Path length and relative measurement error

Rainfall can only be estimated as path-averaged rainfall intensities, meaning that the spatial scale of the observation increases with path length. Under constant rainfall intensities, absolute attenuation due to rain is larger over long link paths than over short links (given the same signal frequency). Measurement errors in total attenuation over long links are divided by longer path lengths resulting in a relatively smaller error in k . Also the contribution of wet antenna attenuation is relatively larger for shorter link paths.

These three mechanisms show the intricate way link characteristics may impact the uncertainty in a deterministic R - k relationship and thereby in CML rainfall estimations. The CML dataset evaluated in this study contains links with frequencies varying between 39.3 – 12.8 GHz and path lengths between 0.5 – 18 km, where short links typically correspond with high signal frequency values and vice versa. The majority of the links is vertically polarized (72%).

Based on the previous research described in this section, we expect best results in our dataset for links longer than 5 km transmitting signals at frequencies near 32 GHz.

4.3.3 Meteorological conditions

Heavy rainfall

Rainfall intensity is related to the uncertainty in the assumptions made in wet antenna attenuation and to the effect of power level rounding. The wet antenna attenuation results in a larger overall attenuation than merely due to k , which could lead to a considerable overestimation if not accounted for. In the simulation study by Leijnse et al. (2008) this A_a depends on rainfall intensity and to a lesser extent on frequency (over the range of the

current dataset). At 20 GHz, the A_a of a single antenna varies between ~ 1 dB at 1 mm h^{-1} and ~ 2 dB at 10 mm h^{-1} (Leijnse et al., 2008). Nevertheless, Overeem et al. (2016b) show good rainfall observations from a Dutch min/max sampling CML network where A_a has a constant value of 2.3 dB (for two antennas). This corresponds with Schleiss et al. (2013) who find the dependency of wet antenna attenuation on time after the start and end of the rainfall is larger than the dependence on rainfall intensity. Secondly, the error due to rounding of RSL and TSL becomes relatively smaller when the total attenuation increases, i.e. for larger rainfall intensities.

Time of day

Rainfall estimations are prone to errors due to dew formation on the antenna (Van Leth et al., 2018) and changes in the atmosphere that affect the refractive index of air and thereby the microwave signal propagation (i.e., super refraction and ducting). Valtr et al. (2011) report how the attenuation of a microwave signal may vary during the day due to these effects on beam propagation over the link path. Both the dew-formation and the atmospheric conditions related to beam propagation mostly occur during late night or early morning, and they therefore impact the attenuation most severely then.

Time of year

The rainfall retrieval principle described in Sect. 4.2 only applies for liquid precipitation. Solid precipitation along the link path attenuates the signal far less than rain droplets. However, melting precipitation along the link path or on the antenna heavily attenuates the signal, possibly over a long period. This may lead to considerable errors in microwave rainfall estimation during and after solid or melting precipitation events if not properly accounted for (Paulson & Al-Mreri, 2011). In the Netherlands these events are most likely to occur in winter (Dec - Feb).

4.4 Methods

4.4.1 Dataset

This study uses a CML dataset from telecom provider T-Mobile NL, consisting of instantaneous transmitted and received powers at 15 minute resolution over 1936 unique microwave link paths in the Netherlands. The signal frequency ranges between 12.8 and 39.3 GHz. Transmitted powers (dBm) are truncated to integers and typically remain constant over long periods of time. Received powers (dBm) are provided in one decimal accuracy. The period from 2016-02-18 to 2016-10-16 is evaluated, chosen because of the

high data availability during this period. Metadata of the links consist of location and height of receiving and transmitting antennas, link ID, signal polarization and microwave frequency. Figure 4.1 shows the layout and data availability of the link network.

Figure 4.1: Map of the Netherlands with data availability and network lay-out of the CML dataset.

For validation a climatological rainfall dataset was obtained from the Royal Netherlands Meteorological Institute (KNMI) based on data from two C-band radars (in Den Helder and de Bilt, and a new radar in Herwijnen during the last 4 weeks of the study period) adjusted with data from two rain gauge networks (31 automatic and 325 manual gauges) (Overeem et al., 2009a,b, 2011), freely available via https://data.knmi.nl/datasets/rad_n125_rac_mfbs_em_5min/2.0. Rainfall amounts are given over the entire land surface of the Netherlands in pixels of $\sim 1 \text{ km}^2$ with a temporal resolution of 5 minutes.

4.4.2 RAINLINK processing

The software package RAINLINK (Overeem et al., 2016a), written in scripting language R, has been designed to calculate CML rainfall estimates from minimum and maximum RSL with a constant TSL. A simple preprocessing module replaces the minimum and maximum RSL columns (in dBm) with $\text{RSL} - \text{TSL}$ (negative attenuation in dB). The unit change and the negative values have no consequences for the following calculations. This approach

enables the package in its current form to calculate rainfall from instantaneous power measurements. Also, it no longer requires constant TSL and makes use of polarization-dependent a and b parameters (Eq. 4.1). The R-package, its documentation (including documentation on the additional preprocessing module for instantaneous sampling) and a 2-day sample dataset of over 2000 CMLs are freely available on GitHub (via <https://github.com/overeem11/RAINLINK>).

The RAINLINK-package is used to convert RSL – TSL into rainfall observations following 6 steps:

1. Duplicated link IDs, IDs with varying metadata, and links with frequencies outside the range 12.5 – 40.5 GHz are excluded from the analysis. Note that all the links in the dataset operate within this frequency range.
2. The wet-dry classification relies on the spatial correlation of rainfall. When more than half of nearby links (within a radius of 15 km) experience reduced RSL – TSL levels, the interval is labeled as wet. These reductions are calculated as the difference between the RSL – TSL as compared to the maximum (i.e. least negative) RSL – TSL values of the respective links over the past 24 hours, both as difference and as difference divided by the link path length. If the median of all nearby links is more negative than predefined thresholds (Q_{mP} in dB for difference and Q_{mPL} in dB km⁻¹ for difference per km link path), the link is labeled as wet for that interval. This step is optional in the software, which when omitted results in all intervals to be labeled as dry in step 3.
3. The median RSL – TSL of all dry intervals during the previous 24 h is considered the dry weather reference signal level.
4. In the outlier filter, intervals of a link are excluded if the cumulative difference between its specific attenuation and that of the surrounding links over the previous 24 h becomes lower than the outlier filter threshold.
5. From the attenuation during wet intervals, the dry weather reference and a fixed value for wet antenna attenuation (A_a) are subtracted. This value divided by the path length yields k .
6. R is calculated from k (with Eq. 4.1), using frequency-dependent a and b values provided in the package. In case of min/max sampled data the weighted average of the R_{min} and R_{max} rainfall would be calculated with weighing factor α . However, the α parameter becomes obsolete in case of instantaneous sampling.

An extensive explanation of the rainfall retrieval algorithm is provided by Overeem et al. (2016a). Note that in the default version of the RAINLINK-package step 2 and 3 are applied on power levels instead of attenuation values.

As parameter values may differ between climates and sampling strategies, it is recommended to calibrate RAINLINK parameters on a subset of the dataset of interest (Overeem et al., 2016a). The default parameter values for Q_{mP} , Q_{mPL} in step 2 and A_a in

step 5 in RAINLINK were based on min/max sampled data (Overeem et al., 2016a). These parameters have been newly determined according to an optimization analysis:

A subset of 10 rainy days spread over the study period were used to calculate hourly rainfall estimates for all combinations of Q_{mP} between -2 and -0.2 , Q_{mPL} between -1.4 and -0.2 and A_a between 0 and 3 (steps of 0.2). Validation with gauge-adjusted radar ($R_{\text{rad},15\text{min}}$) cumulated to hourly values yields metrics for each of the 10 days on: Coefficient of Variation (CV), Pearson correlation coefficient (r), number of intervals where rainfall could be estimated (n), relative bias in the mean (bias), probability of detection (POD in %) and false alarm ratio (FAR in %). The best combination of parameters was determined as the set where the cost function was minimal:

$$\text{cost}_j = \sum_{i=1}^d (V_{\text{CV},i,j} + V_{r,i,j} + V_{n,i,j} + 2V_{\text{bias},i,j} + 2V_{\text{POD},i,j} + 2V_{\text{FAR},i,j}) \quad (4.3)$$

with:

$$V = \begin{cases} x, & \text{if } x \leq 1 \\ 3, & \text{if } x > 1 \end{cases}, \quad (4.4)$$

$$\begin{aligned} x_{\text{CV}} &= \frac{\text{CV}}{6}; & x_r &= 1 - r; & x_n &= 4 \left(1 - \frac{n}{n_{\text{max}}} \right); & x_{\text{bias}} &= \frac{|\text{bias}|}{2}; \\ x_{\text{POD}} &= \frac{100 - \text{POD}}{100}; & x_{\text{FAR}} &= \frac{\text{FAR}}{100} \end{aligned} \quad (4.5)$$

Here j indicates the combination of parameters, with i a day (d) in the subset of 10 rainy days. The outlier filter threshold (used in step 4 of the RAINLINK-package) was kept at its default value. The resulting values for Q_{mP} and Q_{mPL} (-0.6 dB and -0.4 dB km^{-1} , respectively) became less negative, and the A_a value (1.4 dB) was lower than the default. Note that this default value was determined in a previous analysis on min/max data by jointly optimizing α and A_a (Overeem et al., 2016a), which may have resulted in values of α compensating for the higher values of A_a (or vice versa).

4.4.3 Validation

In order to obtain a link-based validation dataset from radar pixel measurements, the rainfall intensities of all pixels overlying the link path are averaged, weighted by the length of the link path segment through the pixel. Two types of radar rainfall intensity references (in mm h^{-1}) are used in this research. The first one is used for the validation of instantaneous CML observations, and is constructed from radar data of a single time interval of 5 minutes closest to the instantaneous CML timestamp: $R_{\text{rad},5\text{min}}$.

Secondly, the temporal average of the three path-averaged radar rainfall intensities within the 15 minutes for which the instantaneous CML measurement is assumed to be representative: $R_{\text{rad},15\text{min}}$. The $R_{\text{rad},15\text{min}}$ includes all rainfall information between CML instantaneous observations, and is used to examine how well CML describes rainfall over longer periods when aggregating measurements over longer time intervals. Whenever $R_{\text{rad},15\text{min}}$ or rainfall intensity from link observations were aggregated over longer durations, any interval with less than 80% availability of 15 minute intervals was discarded.

Path-averaged link rainfall estimates (R_{link}) are determined over ~ 7 months (excluding the 10 day calibration subset used in the optimization analysis). In order to validate R_{link} with R_{radar} , where R_{radar} can be either $R_{\text{rad},5\text{min}}$ or (aggregated) $R_{\text{rad},15\text{min}}$, the following diagnostics were calculated. Intervals where both R_{link} and R_{radar} observe zero rainfall (i.e. dry weather) are excluded. The Pearson correlation coefficient (r) is defined as:

$$r = \frac{\text{cov}(R_{\text{link}}, R_{\text{radar}})}{\text{sd}(R_{\text{link}}) \text{sd}(R_{\text{radar}})}, \quad (4.6)$$

where $\text{cov}(x, y)$ is the covariance between x and y , and $\text{sd}(x)$ is the standard deviation of x . The relative bias (bias from now on) is defined as:

$$\text{bias} = \frac{\overline{\Delta R}}{\overline{R_{\text{radar}}}} \quad (4.7)$$

with $\overline{R}$ indicating the mean of R and:

$$\Delta R = R_{\text{link}} - R_{\text{radar}} \quad (4.8)$$

The coefficient of variation (CV) is calculated as the standard deviation of the difference divided by the mean of the reference:

$$\text{CV} = \frac{\text{sd}(\Delta R)}{\overline{R_{\text{radar}}}} \quad (4.9)$$

Two measures of how often link observations correctly estimate the occurrence of rainfall are probability of detection (POD) and false alarm ratio (FAR).

$$\text{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}} \times 100\% \quad (4.10)$$

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \times 100\% \quad (4.11)$$

Here, hits are the number of intervals R_{link} and R_{radar} both detect rainfall, misses are the number of intervals when R_{link} detects dry weather while R_{radar} is rainy, and the false alarms are the number of intervals where R_{link} is rainy when R_{radar} is dry. An interval can be labeled as ‘wet’ when the rainfall is above zero, or above a certain rainfall minimum (where ‘dry’ intervals measure zero rainfall / rainfall amounts below this minimum).

The entire dataset was subdivided based on known features in the CML dataset to determine their impact on the overall accuracy:

- Link geographical location
- Link path length
- Link signal frequency
- Link signal polarization
- Time of day
- Time of year
- Rain type, e.g. during heavy rainfall events

Intervals of heavy rainfall were determined based on $R_{\text{rad},15\text{min}}$; Intervals in the link dataset during all events where $R_{\text{rad},15\text{min}}$ reported at least an average of 6 mm h^{-1} over 2.5 hours or longer were selected. Intervals at the start and end of an event with rainfall intensities below 3 mm h^{-1} were excluded, and the complete event was discarded if it became shorter than 1 hour after excluding start and end intervals, i.e. if it contained only a few or a single high rainfall intensity value(s). R_{link} was compared with $R_{\text{rad},5\text{min}}$ for all selected intervals for all links. These criteria were found to provide a substantial subset while still containing continuous periods of heavy rainfall.

4.5 Results

4.5.1 Overall comparison

Figure 4.2 shows time series of the number of links in the dataset before and after applying the RAINLINK algorithm. The link density decreases over time. The gaps in the time series are in principle due to data storage issues in the network management system of the cellular telecommunication company rather than the CML network being out of service. After processing the data there are on average 1458 links available per 15 min interval over the total period. Data availability reduced to an average of 1451 links per 15 min interval when missing values in the radar reference were taken into account.

The performance of all links over the entire period, disregarding any characteristic of the links or time of observation, is visualized in Fig. 4.3. The links overestimate rainfall and there is a considerable spread around the diagonal (grey line) that would indicate perfect agreement between link and reference. The comparison between links and instantaneous radar ($R_{\text{rad},5\text{min}}$) during heavy rainfall intervals shows a far smaller bias and reduces the CV from 3.70 to 1.59 in the total dataset (Fig. 4.3).

Figure 4.4 shows the comparison with the $R_{\text{rad},15\text{min}}$ for 15 minutes, hourly, 3-hourly, and daily accumulations. This demonstrates the ability of the link observations, once every 15 minutes, to describe rainfall dynamics at various temporal aggregation scale. The accuracies of the links increase for longer accumulation intervals, as shown by a decrease

Figure 4.2: Availability per 15 minute interval of raw (‘original dataset’) and processed (‘processed dataset’) CML data in time.

Figure 4.3: Validation of path-averaged 15-min CML rainfall depths against gauge-adjusted radar data of (a) all instantaneous rainfall values at 15 minute temporal resolution and (b) all instantaneous intervals during heavy rainfall events.

in CV (4.15, 3.43, 3.24 and 2.32 respectively) and increase of r (0.28, 0.52, 0.57 and 0.63 respectively). Link rainfall estimates are higher than the radar reference on average for all accumulation intervals.

To explore whether over- or underestimation occurred similarly for all links, all double-mass curves (i.e, total link accumulation vs. total radar accumulation) are evaluated in a density plot (Fig. 4.5). Intervals where either the link or the reference contained missing data were excluded. This figure shows that the densest region of the double mass curves is parallel to the diagonal, indicating agreement between link and reference. However, there is overestimation by the links as shown by the large density area above the diagonal. A considerable fraction of the overestimation by CMLs occurs near the origin, corresponding in most cases with the colder earlier months of the dataset when melting precipitation may occur.

The accuracy of the links in determining rainfall occurrence, as opposed to rainfall amounts, is visualized in Fig. 4.6. The values of POD and FAR are displayed for a rain occurrence threshold of 0 mm, as well as 0.1 and 0.5 mm during the time interval. Although the POD increases and the FAR decreases for longer time intervals, the occurrence of rainfall in the reference was missed in $\sim 60\%$ of the hourly intervals by the links with a threshold of 0.1 mm when we assume $R_{\text{rad},15\text{min}}$ to be accurate.

Figures 4.5 and 4.6 indicate that the links tend to overestimate rainfall when they detect it but often miss rainfall that occurs according to the reference, respectively.

4.5.2 Link characteristics

Known link characteristics are geographic location, link path length, signal frequency and signal polarization. Geographic location was found not to be a significant factor on accuracy in this dataset, as the validation metrics of links showed no spatial dependence over the Netherlands (not shown).

Each link has its individual signal frequency. Shorter links generally operate at higher frequencies. In Fig. 4.7 the comparisons of link rainfall and radar reference per bin of frequency and path length show relative large bias, low POD and high CV for the shortest links, especially up to 2 km. The longer links show mixed results and a poor performance in terms of POD values, although this may in some cases (e.g. bin A2) be due to outliers in the relatively small number of values (Fig. 4.7). The error seems to depend more on path length than frequency and the dataset is evaluated based on path length only. Fig. 4.8 shows similar results for all links except for links shorter than 2 km and between 12 and 13 km (corresponding with bin A2 in Fig. 4.7) where particularly the bias and CV are largest. Further examination showed that the poor performance of links with path lengths of 12 – 13 km is due to 2 of the total 17 links of this length class that consistently overestimate rainfall during the complete study period.

Figure 4.4: Validation of path-averaged CML rainfall depths against gauge-adjusted radar data. Scatter density plots of links with reference (based on $R_{\text{rad},15\text{min}}$) over (a) 15 minutes, (b) hourly, (c) 3 hourly and (d) daily accumulation intervals.

Figure 4.5: Density plot of the double mass curves of all individual links with the radar reference $R_{\text{rad},15\text{min}}$.

Figure 4.6: Validation of the ability of CMLs to detect rainy periods against gauge-adjusted radar data. POD and FAR calculated for instantaneous path-averaged comparison as well as various longer accumulation intervals for three thresholds for rain occurrence.

Figure 4.7: Validation of CML path-average rainfall estimates for combinations of path length and microwave frequency against a gauge-adjusted radar dataset. Plots reflecting the (a) bias (sign indicates whether the bias is positive or negative), (b) r (c) CV, (d) POD, (e) FAR and (f) number of values per path length and frequency bin. The values are calculated based on instantaneous rainfall intensity at 15 minute temporal resolution. Brighter colors represent a better match with the reference.

Figure 4.8: Comparison between instantaneous rainfall intensities of links and radar at temporal resolution of 15 minutes for various path length bins. Axes are oriented such that low values indicate higher accuracy for all lines.

The metrics of horizontally and vertically polarized links were also compared (not shown). Measurements from the horizontally polarized links (26% of observations in dataset) have a smaller relative bias (0.25 against 0.33), similar r (0.29 against 0.30) and smaller CV (3.6 against 3.7) compared with measurements from vertically polarized links. The subsets are different in the sense that link paths of over half of horizontally polarized links are shorter than 2 km, against only one third of the vertically polarized links.

4.5.3 Time of observation

Because of the weak signal effect of snow and heavy attenuation during melting precipitation on antennas and along the link path, it is likely that the link rainfall retrieval error is larger in winter months. The CV and bias were particularly large in the first month, with also a correlation lower than in other months (Fig. 4.9). Further evaluation showed that solid precipitation did occur in the Netherlands in February and March.

Figure 4.9: Monthly comparison between instantaneous rainfall intensities of links and radar at temporal resolution of 15 minutes. Axes are oriented such that low values indicate higher accuracy for all lines.

The metrics show high positive bias, large values of CV and FAR particularly between 4 and 6 UTC. This corresponds with the end of night / early morning (local time), and may be related to dew formation and atmospheric changes affecting beam propagation (Fig. 4.10).

Figure 4.10: Comparison between instantaneous rainfall intensities of links and radar at temporal resolution of 15 minutes for each hour of the day (Dutch local time is UTC+2 for most of the considered period). Axes are oriented such that low values indicate higher accuracy for all lines.

4.6 Discussion

4.6.1 RAINLINK parameter choices

The optimization analysis of RAINLINK on instantaneous data results in less negative thresholds than the parameters which were obtained in min/max sampled data in the same climate: $Q_{\text{mP}} = -0.6$ dB instead of -1.4 dB and $Q_{\text{mPL}} = -0.4$ dB km⁻¹ instead of -0.7 dB km⁻¹. This means that an interval is more easily classified as wet with the new parameters. However, the relatively low POD values in Fig. 4.6 imply that a significant fraction of rainfall is missed, even with the new thresholds. The Dutch min/max sampled CML dataset from Overeem et al. (2016b) (evaluated in the same manner as the current dataset) indeed shows a slightly higher POD and a lower FAR than the instantaneously sampled dataset (Fig. 4.11, showing the differences in POD and FAR between this study and Overeem et al. (2016b) for two different thresholds above which the interval is considered as wet, with higher POD and FAR for higher thresholds).

This suggests that less rainfall is missed in min/max sampling than in instantaneous sampling, even with adjusted parameter settings. Min/max sampling enables the detection of rainfall at any time in the interval, where instantaneous sampling may miss it.

The optimized A_a parameter changed from 2.3 dB to 1.4 dB, which is more in the range of the values found by Overeem et al. (2011) (1.2 – 1.9 dB), although the authors find an

A_a of 2.3 dB more suitable in a later study (Overeem et al., 2013a). In both studies, the A_a value is determined together with another parameter α , the weighting factor between the calculated rainfall intensity from the minimum and maximum RSL, which is obsolete for instantaneous sampled data. These two parameters (α and A_a) are highly dependent, where higher values of α yield similar accuracy as long as A_a increases as well (Overeem et al., 2016b). This highlights the need for determining A_a in the absence of α in the processing of the instantaneous dataset, and that such a lower value would make more sense for use in the algorithm where min/max values are available.

A fixed A_a of 1.4 dB (effectively 0.7 dB per antenna if both are wet) could be an underestimation of the actual A_a for rainfall intensities above 1 mm h^{-1} (Leijnse et al., 2008). This would result in a rainfall overestimation during heavy rainfall. This overestimation was not found for the current dataset (see Fig. 4.3), which suggests that other factors overshadow this effect.

4.6.2 Accuracy dependence on link and rain characteristics

The figures in Sect. 4.5.2 and 4.5.3 should not be mistaken as a means to identify the effect of isolated characteristics, but a means to identify the most prominent influences on rainfall observations experiencing multiple error sources. It should be noted that the validation dataset used here is a radar product that measures rainfall at an elevation of typically $\sim 1.5 \text{ km}$ because of the beam sampling by radar, which leads to uncertainties regarding actual rainfall at the surface. As short links typically operate at high frequencies, causality of the relatively larger errors in this subset can be difficult to determine. Horizontally polarized links perform slightly better than vertically polarized links especially concerning bias. Horizontally polarized links are more sensitive to rainfall along the link path (Ruf et al., 1996).

The graphs in Fig. 4.8, 4.9 and 4.10 were also constructed on hourly values (not shown). Although, as expected, the values indicated better agreement with the reference (see also panel (b) vs (a) in Fig. 4.4), the dependence of the values on path length, time of day and time of year were the same as for the instantaneous comparison.

The error seems to depend more on path length than on frequency in Fig. 4.7. Not all bins of path length and frequency are equally well represented in the dataset, and the low number of values on which the calculations are based may be responsible for cases of poor accuracy. The poor accuracy of the 12 – 13 km path length links is due to two links with continuous rainfall overestimation. Excluding these two links would result in validation metrics similar to those of bin 11 – 12 and 13 – 14 km. The accuracy is lowest for links shorter than 2 km, which was not due to outliers in this subset. It could however be caused by larger relative measurement errors (see also Sect. 4.3), or the large uncertainty in the ground truth that is constructed from pixel averaged rainfall intensities.

Previous research described in Sect. 4.3.2 indicated that links longer than 5 km and

transmitting signals with frequencies near 32 GHz are expected to be most accurate. Figure 4.7 shows that the corresponding bin (E6) indeed shows relatively high r and CV, and it is the only bin where rainfall is not overestimated. However, the POD and FAR are not as good as for some other path lengths and frequencies.

The temporal variation in the error corresponds with the possible influence of atmospheric processes, dew and solid precipitation (Valtr et al., 2011; Paulson & Al-Mreri, 2011; Van Leth et al., 2018), but this cannot be verified in our study set-up. Additional measurements in an experimental set-up are needed to confirm a causal relation.

4.6.3 Performance compared to other studies

The current dataset originates from a network in the same climate (the Netherlands) as several 15 minute min/max sampled datasets explored by Overeem et al. (2011); Rios Gaona et al. (2015); Overeem et al. (2016a,b); Van het Schip et al. (2017); Rios Gaona et al. (2017). Although these studies differed slightly from the current one in terms of validation approach (e.g. the decision to exclude intervals below a certain threshold, map-based or path-based comparison), network (e.g. number of links) and accumulation interval, they enable us to draw conclusions on the relative performance of instantaneous datasets as compared to min/max sampled CML data. Table 4.1 summarizes the results found in these studies that most closely resemble the current study, and the dataset in Overeem et al. (2016b) has been re-analyzed in the same manner as the current approach for better comparison. These values show a smaller bias, lower CV and higher correlations in rainfall observations from min/max sampled data at 15 minute intervals than from instantaneously sampled observations, as was hypothesized in Sect. 4.3.1.

Figure 4.11: POD and FAR calculated on 15 minute path-averaged rainfall observations based on the min/max sampled data described in Overeem et al. (2016b) for total period and the months March – October (indicated as summer) in the 2.5-year dataset, as well as the POD and FAR values found in the current study. The values are calculated for two thresholds of nonzero rainfall and compared with gauge-adjusted radar data.

Table 4.1: CML validation studies in the Netherlands, including current results where r values are squared for better comparison.

Study	Dataset	Comparison	Threshold	Bias [%]	CV	r^2
Rios Gaona et al. (2015) [same dataset as Overeem et al. (2013a, 2016a) and Van het Schip et al. (2017)]	Average 1514 link paths over 12 rainy days in 2011 (June – September)	Map-based comparison with gauge-adjusted radar	radar > 0.1 mm	-14.3	1.216	0.366
		1 km ² , 15 min				
		81 km ² , 15 min	radar > 0.1 mm	-9.1	0.995	0.496
		1 km ² , daily	radar > 0.1 mm	+1.6	0.523	0.720
		Link-based comparison with gauge-adjusted radar, 15 min	link or radar > 0.1 mm	~ -13.3	1.437	0.437
Overeem et al. (2016b)	Average 3383 links (2044 unique link paths) over 2.5 years (14 January 2011 – 30 July 2013)	Map-based comparison with gauge-adjusted radar data	no threshold	+1.6	5.58	0.40
		74 km ² , 15 min, entire period				
		74 km ² , 15 min, winter	no threshold	+10.9	5.8	0.26
		74 km ² , 15 min, summer	no threshold	-4.1	4.81	0.53
		0.9 km ² , daily, entire period	radar > 0.1 mm	-4.8	1.05	0.36
		0.9 km ² , daily, winter	radar > 0.1 mm	+5.8	1.53	0.20
		0.9 km ² , daily, summer	radar > 0.1 mm	-7.8	0.59	0.67
Rios Gaona et al. (2017)	Average 1404 links over ~ 7 months (12 March 2014 – 1 November 2014)	Map-based comparison with gauge-adjusted radar data, with dew-filter	radar ≥ 0.1 mm (not for bias)	-8.88	1.16	0.37
		0.1° × 0.1°, 30 min				
		0.1° × 0.1°, daily	radar ≥ 1.0 mm (not for bias)	-8.88	0.58	0.66
Overeem et al. (2016b) dataset re-evaluated in current study	Average 3492 links over 2.5 years, only considering period 18 February – 16 October (613 days)	link-based comparison with path-averaged gauge-adjusted radar	link or radar > 0 mm	+5.75	2.84	0.27
Current study	Average 1451 links over ~ 7 months (18 February 2016 – 16 October 2016), instantaneously sampled	15 min				
		Link-based comparison with path-averaged gauge-adjusted radar	link or radar > 0 mm	+30%	3.70	0.09
		15 min				
		hourly	link or radar > 0 mm	+23%	3.43	0.27
		daily	link or radar > 0 mm	+25%	2.32	0.40

4.7 Conclusion

4.7.1 Validation outcome

This study explores the accuracy of rainfall observations obtained from an instantaneously sampling CML network covering the Netherlands, operating at frequencies between 12.8 and 39.3 GHz. A preprocessing module was successfully added in order to enable the open source R-package RAINLINK to process data from this sampling format. Instantaneous rainfall observations were found to have a CV of 3.70 and r of 0.30, with improved results during heavy rainfall and at longer time intervals, with a CV of 2.32 and r of 0.63 for daily observations. The links tend to miss a considerable fraction of rainfall that is captured by the ground truth (POD of $\sim 40\%$ to detect rainfall exceeding 0.1 mm), and overestimate rainfall intensities when they do detect it (overall bias of $\sim +30\%$).

There are various ways in which signal frequency, polarization and geographic location could potentially influence the accuracy of CML rainfall observations, but for this dataset the effect of path length and temporal variability was most prevalent. This was evident from the larger errors in links shorter than 2 km, in the first month of the dataset and between 4 – 6 UTC. These errors could be related with large relative measurement errors at small spatial sampling and/or uncertainty in the ground truth over short links, solid or melting precipitation in February/March and dew formation and/or beam propagation changes in the early morning.

Comparisons with the performances of CML networks as explored in previous studies in the same climate indicate that min/max sampling, which effectively includes information over the whole time interval, is preferable over instantaneous sampling at 15 minute time intervals. The instantaneous sampling approach is expected to improve at shorter time intervals as more information on rainfall dynamics can be captured that way. It should be noted that CML networks are not installed or maintained for rainfall monitoring, but can successfully be used as such.

4.7.2 Future prospects

Understanding the potential of obtaining rainfall estimates from CML datasets captured in commonly used sampling strategies is vital for future operational use. Communicating the benefits for improved rainfall observation accuracy by slight modifications in the manner that RSL levels are logged in network management systems, without deteriorating the information that telecommunication companies obtain from these observations, may be the way forward for future CML collaborations. These modifications would comprise more frequent logging of TSL and RSL, as well as making these values real time available.

In this work the gauge-adjusted radar product was considered as ground truth, even though radar estimates rainfall based on observations of an atmospheric volume aloft. The differences in sampling compared to ground-based sensors can lead to considerable differences in observations that in our current work is interpreted as errors in the CML estimate. It may therefore be interesting to validate in the future with collocated ground-based sensors i.e. by comparing CML link estimates with nearby WMO rain gauges. This can also be done to investigate in further detail the CML accuracy under various weather conditions.

Even though CML networks have not been designed for environmental sensing, they are successful in capturing rainfall signals. However, the introduction of fiber optical communication technology has led to a reduction in the number of operational commercial microwave links in the Netherlands. The T-Mobile CML network in the Netherlands is currently far sparser than represented in Fig. 4.1. The use of CML observations for rainfall monitoring has most potential for those areas in the world where fiber optical cable communication is not favorable (e.g. because of local topography, limited benefits or financial constraints) and other rainfall sensing techniques are lacking. According to Ericsson, key-player in the construction of commercial microwave links, ~40% of global radio sites will be connected by CMLs in 2023 (Ericsson, 2018).

This study shows that the RAINLINK-package is applicable to instantaneously sampled datasets. Apart from studies in Brazil (Rios Gaona et al., 2018), Italy (Alberoni et al., 2018) and Pakistan (Sohail Afzal et al., 2018) RAINLINK has mainly been applied to Dutch datasets. To advance the process toward operational application and upscaling of this technique, long time series should be analyzed for other networks and climates. It can be particularly useful to blend rainfall data from CMLs with other rainfall observations (e.g. Haese et al., 2017).

Chapter 5

Simulated PWS, CML and radar rainfall

Many applications in urban areas require high-resolution rainfall measurements. Typical operational weather radars can provide rainfall intensities at 1-km² grid cells every five minutes. Opportunistic sensing with commercial microwave links yields path-averaged rainfall intensities (typically 0.1–10 km) within urban areas. Additionally, large amounts of urban in-situ rainfall measurements from amateur weather observers are obtainable in real-time. The accuracy of these three techniques is evaluated for an urban study area of 20 km × 20 km, taking into account their respective network layouts and sampling characteristics. We use two simulated rainfall events described in terms of drop size distributions on a 100-m grid and with a temporal resolution of 30 s. Accurate radar rainfall estimation with the Z – R relationship relies heavily on an appropriate choice of parameters, and a dual-polarization (dual-pol) strategy is more suitable for higher intensities. Under ideal measurement conditions, the weather station network is the most promising, with a Pearson correlation coefficient above 0.86 and a relative bias below 4% for 100-m rainfall estimates at 5-min resolution. Microwave link rainfall observations contain the largest error, shown by a consistently larger coefficient of variation. The accuracy of all techniques improves when considering rainfall at larger scales, especially by increasing time intervals, with the strongest improvements found for microwave links for which errors are largely caused by their spatial and temporal sampling. Sparser networks are examined, showing that the decline in measurement accuracy only becomes significant when the link and station network density are reduced to less than half their levels in Amsterdam.

This chapter was originally published as:

De Vos, L.W., Raupach, T.H., Leijnse, H., Overeem, A., Berne, A., & Uijlenhoet, R. (2018). High-resolution simulation study exploring the potential of radars, crowd-sourced personal weather stations, and commercial microwave links to monitor small-scale urban rainfall. *Water Resources Research*, 54, 10293–10312. doi:10.1029/2018WR023393

5.1 Introduction

Urban catchments are typically small, with a short response time in terms of rainfall-runoff dynamics. This makes cities vulnerable to extreme rainfall events, which are expected to occur more frequently and with higher intensities in Europe due to climate change (IPCC (2012); Madsen et al. (2014); Nissen & Ulbrich (2017); other areas in the world are also affected). For hydrological modelling, especially in an urban context, rainfall information is required at a higher space-time resolution than current ground measurement networks usually provide (Schilling, 1991; Berne et al., 2004; Ochoa-Rodriguez et al., 2015; Thorndahl et al., 2017).

Radar is a valuable rainfall information source due to its extensive coverage. The resolution of operational weather radar, typically 1 km^2 every 5 min, suggests suitability for urban rainfall monitoring. As radar estimates rainfall from a measurement of scattering properties of an atmospheric volume aloft, and not at the surface, significant errors are introduced (Villarini & Krajewski, 2010). Also, even these resolutions fall short of capturing the small-scale rainfall variability required for urban rainfall-runoff models (Gires et al., 2012; Bruni et al., 2015). High-density X-band radar networks like those described by Chen & Chandrasekar (2015) can measure at even higher resolutions and closer to the ground at a shorter range, but operational weather radars are usually much more widely spaced.

Adjustment of radar observations with ground measurements is a common approach to improve the accuracy of retrieved rainfall fields (Einfalt et al., 2004; Wang et al., 2013). For this purpose a dense network of ground observations is vital, especially for geostatistical merging methods (Goudenhoofd & Delobbe, 2009). As the installation and maintenance of a measurement network involves large investments, dense in situ urban measurement networks are often lacking. This study explores the potential of opportunistic sensing to provide high-resolution ground rainfall observations. The techniques are opportunistic in the sense that the creation of a rainfall data set is a by-product of the original purpose of the device owners.

Many types of atmospheric information can be obtained via crowdsourcing (Muller et al., 2015). A growing number of people own low-cost automatic weather stations that share their measurements in real-time on online platforms. For some types of personal weather stations (PWSs), this is even a preprogrammed default feature of the product. The measurements are collected and visualized on a website, from which they can be collected. Contrary to other dense measurement networks like those described by Basara et al. (2011), the stations are acquired and maintained by individual citizens. This data source has been investigated by Bell et al. (2013) regarding pressure, relative humidity, and temperature measurements and used in multiple studies regarding the Urban Heat Island effect (Chapman et al., 2017; Meier et al., 2017). Rainfall measurements obtained from PWS platforms have been evaluated in Chapter 2, where it was shown that PWS

rainfall data may include significant errors on small time scales due to numerical rounding processes at the platform. It is therefore recommended to collect rainfall data from PWS platforms where no rounding or time delays are introduced in the raw measurements (e.g. the Netatmo Weathermap platform). PWS networks are dense in heavily populated areas like cities and in Chapter 2 good sensor quality was found for PWS rain gauges, although it was also shown that the measurements contain errors likely due to poor maintenance, shielded locations, and faulty installation.

Another method to obtain rainfall information in cities is to make use of commercial microwave links (CMLs). Telecom operators have CML networks for the purpose of data transfer. CMLs exchange microwave signals from a transmitting to a receiving antenna through the atmosphere. Signals at the frequencies typically used in CMLs are sensitive to hydrometeors. Telecom operators monitor the loss of signal strength over the link path as a quality check. Messer et al. (2006) first demonstrated that these signal attenuations in CMLs can be used for rainfall estimation. Later research has further explored the potential of CMLs for rainfall estimation (e.g. Leijnse et al., 2007a,b; Zinevich et al., 2010; Overeem et al., 2011; Chwala et al., 2012). By combined processing of the signal attenuations of a dense network of links, rainfall maps can be constructed (Zinevich et al., 2008; Overeem et al., 2013a). See Messer & Sendik (2015) and Uijlenhoet et al. (2018) for comprehensive overviews of literature regarding this technique.

Both PWS and CML data can be collected in (near) real-time via existing infrastructures, making them interesting additional data sources to traditional networks for (urban) hydrological applications. Methods to combine multiple data sources for an improved rainfall product have been introduced by, for example, Bianchi et al. (2013), Scheidegger & Rieckermann (2014); (gauges, radar and CML), Liberman et al. (2014), Trömel et al. (2014); (radar and CML), Haese et al. (2017) and Fencel et al. (2017); (gauges and CML). As CML and PWS have not been designed for operational rainfall monitoring and their rainfall estimations likely contain errors, validation is needed before they may be implemented in the future. A major limiting factor in determining the accuracy of rainfall measurements is the lack of ground truth. Direct comparison between rainfall observations using different techniques will show differences if their sampling is dissimilar, even if both techniques measure with perfect accuracy. Rain gauge measurements and radar-based rainfall estimates are intrinsically different because of the small-scale variability and intermittency of rainfall, especially for short time steps (Kitchen & Blackall, 1992).

To circumvent the problem of a lacking reference for validation, a simulated rainfall event can be used. Yang & Ng (2017) do so in order to construct combinations of synthetically generated rain gauge measurements and crowdsourced rainfall observations by citizens and to determine the resulting accuracy of modeled stormwater flows by an urban stormwater model. Different from Yang & Ng (2017), the rainfall fields used in this study consist of drop size distributions (DSDs) instead of rainfall intensities. This makes it possible to derive radar and CML-based rainfall measurements as well as rain gauge observations and subsequently enables a fair comparison of the three techniques. Thus, this work yields

vital information to establish how these (combined) remote and opportunistic sensing techniques may be used for urban rainfall monitoring.

5.2 Methods

Figure 5.1: Amsterdam metropolitan area with indicated $20 \text{ km} \times 20 \text{ km}$ domain for rainfall simulations, including the $1 \text{ km} \times 1 \text{ km}$ radar grid (green), the 82 PWS locations (purple dots) and 58 CMLs (orange lines).

Figure 5.2: Cumulative rainfall over the study area for two events. The red crosses indicate the locations of pixels with high (blue), medium (sea green) and low (green) rainfall amounts for which cumulative rainfall depths are visualized.

5.2.1 Background

Rainfall simulator

In this paper we make use of a DSD-retrieval technique that estimates DSDs based on polarimetric radar data. The fields were generated using a modified version of the DSD-

retrieval technique shown in Raupach & Berne (2017), using 5-min radar data from France, where fields at 30-s resolution were computed using advection-based temporal interpolation. The distances from the radar being greater than 30 km for some pixels, it is probable that the radar sampled solid as well as liquid precipitation in these events. The noise correction routine and DSD-retrieval technique ensure that resulting DSDs are reasonable for rain, and comparisons of retrieved DSD moment distributions with those measured on the ground in a nearby region show that the results are sufficiently representative of liquid precipitation for this application. The aim of the simulated DSDs is not to accurately describe the rainfall event on which they were based. Instead, they are designed to represent realistic rainfall dynamics for the study area in Amsterdam. The interpolation method required smoothing of the fields. More details are given in Sect. 5.6. Two rainfall events of 23.5 and 5.25 hr, respectively, are studied, with DSD estimations made on a grid of points at 100-m resolution and 30-s time intervals.

The rainfall simulation domain spans an area of 20×20 km that in our controlled experiment framework covers the Amsterdam metropolitan area as visualized in Fig. 5.1. From the simulated DSDs rainfall amounts can be calculated (Fig. 5.2). The total rainfall maxima of Events 1 and 2 are ~ 50 mm in 1 day and ~ 15 mm in 5 hr, respectively. Rainfall intensities of 50 mm/day have a return period of about 5 years in the Dutch climate, while the rainfall amounts such as in Event 2 occur on average several times per year (STOWA, 2015). Spatial variability of rainfall can be described with correlograms that illustrate the reduction of Pearson correlation coefficient between measurement locations with distance. For the two events, correlations (r) of time series of rainfall intensity, computed from the simulated DSDs, at all combinations of 500 randomly chosen measurement locations in the 20×20 km study area are plotted against their respective distance (d in km). The relation in Eq. 5.1 (see e.g. Villarini et al. (2008)):

$$r = \exp \left[- \left(\frac{d}{d_0} \right)^{s_0} \right] \quad (5.1)$$

was fitted by nonlinear least-squares regression with a Gauss-Newton algorithm. The parameter d_0 indicates the decorrelation distance in kilometers where correlation decreases to a value of e^{-1} and s_0 is the shape factor. Rainfall fields at longer accumulation intervals are more constant in space than at short intervals. The decline of correlation with distance is most steep in Event 2 (Fig. 5.3). We therefore expect that the more homogeneous rainfall in Event 1 requires fewer measurement locations to describe its spatial variability than for the more convective rainfall in Event 2.

Radar

The Royal Netherlands Meteorological Institute (KNMI) monitors Dutch rainfall with two polarimetric C-band weather radars in Herwijnen and Den Helder at distances of,

Figure 5.3: Fitted correlograms of Event 1 and 2, accumulated at various accumulation intervals, according to the relation in Eq. 5.1, with decorrelation distance d_0 indicated for all correlograms. Related fitted s_0 values for 30 s, 5 min, 15 min and 30 min are 0.73, 0.83, 0.95 and 1.02 respectively for Event 1 and 0.93, 1.33, 1.64 and 1.51 for Event 2.

respectively, 62.5 km to the south and 64.4 km to the north of the study area. The product that is operationally used for rainfall monitoring is a result of compositing 1500-m Constant-Altitude Plan Position Indicators from the two radars. The resulting radar rainfall product consists of instantaneous rainfall intensities over 1 km^2 pixels every 5 min nationwide (see Beekhuis & Mathijssen (2018) for specifics on the Dutch national radar product). Currently, the operational rainfall product of KNMI is constructed based on reflectivity only, but future additional rainfall products will be constructed from polarimetric variables. The focus of this study is the traditional radar rainfall retrieval, but an analysis with dual-polarization (dual-pol) radar rainfall retrieval is included for comparison.

From two polarimetric C-band radars a nationwide rainfall product is constructed on a grid at 1 km^2 pixels. Because of the increase in beam volume and height the sampling of radar is not constant in space. Figure 5.4 shows the effective spatial sampling of the radars in Herwijnen and Den Helder for each assumed Cartesian pixel of 1 km^2 in the study area, based on the distance between the radars and the Amsterdam metropolitan area. From this we conclude that using a 1-km cartesian grid for simulating radar data is sufficiently close to real radar sampling over the study area.

Figure 5.4: Effective spatial sampling of radars in study area Cartesian grid.

CML

As CML reference we use the layout of the link network employed by T-Mobile NL (see Chapter 4), consisting of 1387 unique operational links in May 2017 in the Netherlands. This specific network has reduced significantly in size since then, but we use its past locations and signal frequencies and polarizations to reflect realistic CML network layouts in urban areas.

Instantaneous transmitted and received signal strengths over all link paths are logged every 15 min. The links operate with both horizontal and vertical polarization and various transmit signal frequencies over path lengths of typically several kilometers (Chapter 4). From the T-Mobile NL network, 58 unique links are located within the Amsterdam study area, 7 of which cross the boundaries of the 20×20 km domain. In the derivation of data sets, only the fractions of these link paths that lie within the simulation area are considered (indicated in orange in Fig. 5.1), with the same polarizations and transmit frequencies as the original links. The links are mostly vertically polarized (39 links) with signal frequencies between 18.0 and 39.2 GHz with most links having frequencies around 32 GHz.

PWS

There is a large number of citizens who own a PWS and share its data on online platforms (Chapter 2). In Netatmo devices (<http://weathermap.netatmo.com>), this is an automatic feature of the product that requires no additional actions from the weather station owner. Although there are many PWS types operational in Amsterdam, this paper only considers the Netatmo devices. They make up a large fraction of PWS devices that can be

crowdsourced in Amsterdam (see also Chapter 3) and in many areas of the world. In contrast to other platforms, rainfall measurements from the Netatmo platform have not been processed/rounded and are updated frequently. Approximately every 5 min the rainfall amount since the last update is provided as a multiple of 0.101 mm, corresponding to one Netatmo tipping bucket tip volume. The update moments are not synchronized between stations, and the interval lengths between updates vary around 300 s, with peaks at 295 s and at 308 s. Station outages are common and result in gaps in the time series.

The number of stations is steadily growing; the selection of station locations used in this paper is based on the 82 stations linked to the Netatmo platform in the study area on 19 December 2016, whereas 93 stations were already available on 10 May 2018. This steadily growing data source is therefore increasingly relevant, although it is expected that PWS accuracy decreases over the life time of these low-cost devices in case of poor maintenance.

5.2.2 Derivation of rainfall datasets

Data sets of simulated radar, CML, and PWS rainfall intensities are constructed by calculating how they would register a rainfall event as described in the high-resolution rainfall simulation, based on their respective measurement principles, locations of the sensors and typical sampling intervals. The method is visualized in Fig. 5.5.

The ground truth rainfall intensity (R in mm h^{-1}) with a resolution of 100 m and 30 s is derived directly from the simulated DSDs according to Eq. 5.2:

$$R = 6\pi \cdot 10^{-4} \sum_{i=1}^{N_D} D_i^3 N_i v(D_i) \Delta D_i \quad (5.2)$$

N_D is the number of diameter bins, D_i is the class center in mm, N_i is the simulated concentration for bin i in per mm m^{-3} , and ΔD_i is the class width in mm. Here $v(D_i)$ (in m s^{-1}) is the functional relationship between the raindrop terminal fall speed (in still air) and the equivolume raindrop diameter belonging to a certain drop size bin, as described by Atlas & Ulbrich (1977).

Radar derivation

For the radar derivation, first the radar reflectivity (Z in $\text{mm}^6 \text{m}^{-3}$) is calculated. Given the Rayleigh scattering assumption, Z can be determined from the sixth-order moment of the DSDs (Bringi & Chandrasekar, 2001):

$$Z = \sum_{i=1}^{N_D} D_i^6 N_i \Delta D_i \quad (5.3)$$

Figure 5.5: Schematization of methods to obtain ground truth and derived rainfall fields in 100-m grids at 30-s resolution of radar, CML, and PWS. The assumptions and the equations used for each step in the derivation methods are given in Sect. 5.2.2. DSD = drop size distribution; CML = commercial microwave link; PWS = personal weather station.

For each time step, simulated Z values in the 100 pixels that correspond with a single 1 km^2 pixel are replaced by their averaged value. Subsequently the rainfall for each $100 \text{ m} \times 100 \text{ m}$ grid cell is set to the rainfall intensity computed from this averaged Z using the Z - R relationship for stratiform rainfall by Marshall et al. (1955):

$$R = \left(\frac{Z}{200} \right)^{\frac{1}{1.6}} \quad (5.4)$$

Calculating R directly from the DSD-fields as compared to via Eq. 5.3 and 5.4 will yield different outcomes. There has been extensive research in the past regarding the Z - R relationship for various rain types and climates (e.g. Stout & Mueller, 1968; Raghavan, 2013). The often-used Marshall-Palmer relationship is chosen here, because this is the relation operationally used for radar rainfall estimation at KNMI and is, on average, representative for rainfall in a temperate climate (Uijlenhoet, 2001). Figure 5.6 shows

how this relation differs from the fitted relationships between Z and R as derived from the DSDs. The Z – R relations were constructed by nonlinear least-squares fits with a Gauss-Newton algorithm, where R was the dependent and Z the independent variable.

In addition to the Marshall-Palmer Z – R relation, rainfall can be estimated by making use of polarimetric variables measured by dual-pol radars. Figueras i Ventura et al. (2012) compared the skill of several dual-pol algorithms applied to several C-band radars in France. This study uses the dual-pol radar rainfall approach that was found to perform best, namely, an algorithm that combines the Marshall-Palmer relationship with the R – K_{dp} algorithm by Beard & Chuang (1987):

$$R = c \left(\frac{K_{dp}}{f} \right)^d \quad (5.5)$$

$$K_{dp} = \frac{180 \, v \, 10^{-9}}{\pi f} \sum_{i=1}^{N_D} \text{Re} [f_{H,f,i} - f_{V,f,i}] N_i \Delta D_i \quad (5.6)$$

where R is in mm h^{-1} , K_{dp} is the specific differential phase in $^{\circ} \text{km}^{-1}$, f is the frequency of the radar (5.6 GHz), v is the speed of light in m s^{-1} , $f_{H,f,i}$ and $f_{V,f,i}$ are the forward scattering amplitude function in mm in , respectively, horizontal and vertical orientation, with the parameters $c = 129$ and $d = 0.85$ as determined by Beard & Chuang (1987). The dual-pol algorithm applies Eq. 5.4 when $K_{dp} < 0.5^{\circ} \text{km}^{-1}$ and Eq. 5.5 above this threshold.

Radar measurements are based on a single radar scan every 5 min. To mimic this temporal sampling, a single 30-s derived rainfall field is compared to multiple ground truth rainfall fields. The radar rainfall field is calculated every tenth time step. This radar rainfall field of time step t is then compared to all ground truth rainfall fields from time steps $t-5$ to $t+4$ (see also Fig. 5.7).

CML derivation

For each link the path-averaged DSD is calculated by making a weighted average of DSDs in all overlying pixels, based on the fraction of the link path running through that pixel. The specific attenuation ($k_{f,pol}$ in dB km^{-1}) value per link is determined with Eq. 5.7.

$$k_{f,pol} = \frac{10^{-2}}{\ln(10)} \sum_{i=1}^{N_D} Q_{ext,f,i,pol} N_i \Delta D_i \quad (5.7)$$

Here $Q_{ext,f,i,pol}$ is the extinction cross-section of a raindrop (in mm^2) dependent on frequency, drop size diameter, and polarization. We have assumed here that the temperature

Figure 5.6: Simulated values of R (Eq. (5.2)) and Z (Eq. (5.3)) for each simulation pixel (purple), as well as the mean values per 1 km² pixel (green) and their respective fitted relationships, together with the Marshall-Palmer Z – R relationship (Eq. (5.4)). Note that due to the size of the dataset only every 10th data point is plotted for Event 1 (left panel), which had no impact on the fitted relationship.

dependence of the extinction cross-sections is negligible. Path-integrated rainfall estimates are calculated from the specific attenuation of each link according to Eq. 5.8 (Atlas & Ulbrich, 1977).

$$R = ak_{f,pol}^b \quad (5.8)$$

where R is in mm h^{-1} and $k_{f,pol}$ is in dB km^{-1} . Coefficients a and exponents b depend largely on microwave frequency and polarization and vary therefore per link. The a and b values used here, originally determined by Leijnse et al. (2008), are also provided by the RAINLINK software package (Overeem et al., 2016a) designed to estimate R from CML data. The R values obtained from k with predetermined a and b values from the RAINLINK package are slightly higher than R values obtained with fitted a and b values on the k and R from the DSDs. This effect, leading to an overestimation of R , is small for low rainfall intensities ($< 4 \text{ mm h}^{-1}$) in all links. The path-integrated rainfall intensity estimates are spatially interpolated on the same grid as the ground truth using Ordinary Kriging with a climatological semivariogram. The rainfall intensity over a link is attributed to the center of the link path, identical to the approach of the RAINLINK package (Overeem et al., 2016a).

To mimic instantaneous measurements every 15 min, rainfall values of a single 30-s time step are repeatedly compared with all ground truth rainfall fields during the 15-min

interval. A CML rainfall field at time t is compared to the ground truth values from time step $t-15$ to $t+14$ (see also Fig. 5.7).

Figure 5.7: Method of simulating instantaneous measurements in radar and CML from 30-s simulated drop size distributions. Colors of the 30-s time steps in the CML and radar time lines represent the time steps of the drop size distributions used to derive those rainfall fields. Time steps colored in black are not used in the derivation of radar and CML. CML = commercial microwave link.

PWS derivation

The derivation of PWS time series is visualized in Fig. 5.8. Netatmo stations measure rainfall with a tipping bucket, meaning that rainfall is only registered once the amount corresponding with a tipping volume, here 0.101 mm, is reached. At each PWS location, a time series of cumulative rainfall is constructed by assuming that the simulated rainfall intensity is constant in the 30-s interval. From this time series, the moments where cumulative values of 0.101 mm are reached are calculated. For each station, a time series is constructed of randomly varying interval lengths typical for those on the Netatmo platform. Next, the number of tipping moments during an interval in the PWS time series are counted and multiplied by 0.101 mm. This method mimics the way rain gauges may capture rainfall from a previous interval in a later time step.

The derived PWS time series are integrated on a fixed time grid at 30-s resolution by assuming constant rainfall within the original intervals. For each 30-s interval the resulting rainfall intensities are then interpolated using Ordinary Kriging on the same spatial grid as the ground truth. Sampling at such small time steps with the 0.101 mm threshold will lead to large uncertainties, especially during light rain, which is explored by evaluating the performance without spatial interpolation at multiple time steps.

5.2.3 Validation

The resulting derived interpolated data sets have the same resolution as the rainfall ground truth. In order to determine the accuracy of rainfall maps obtained by the derived data sets, the relative bias in the mean (referred to throughout the paper as “bias” from now

Figure 5.8: Method of personal weather station time series derivation from (cumulative) simulated time series (t_{sim}), via time lines of tipping moments (t_{tip}), to a timeline with variable interval lengths (t_{var}) and rainfall quantities of (multiples of) 1 tipping volume (Δtip), which is then interpolated on a fixed time grid (t_{fix}): schematic example.

on), Pearson correlation coefficient (r) and the coefficient of variation (CV) are calculated. Bias is calculated as

$$\text{bias} = \frac{\hat{\mu}_{[R-R_{ref}]}}{\hat{\mu}_{[R_{ref}]}} \quad (5.9)$$

Here R is the derived dataset, R_{ref} is the ground truth, and $\hat{\mu}$ is the sample mean. The Pearson correlation coefficient is calculated as

$$r = \frac{\widehat{\text{Cov}}(R, R_{\text{ref}})}{\hat{\sigma}_{[R]}\hat{\sigma}_{[R_{\text{ref}}]}} \quad (5.10)$$

Here the covariance of R and R_{ref} is divided by the multiplied standard deviations of R and R_{ref} , represented by $\hat{\sigma}_{[R]}$ and $\hat{\sigma}_{[R_{\text{ref}}]}$ respectively. CV is defined as

$$\text{CV} = \frac{\hat{\sigma}_{[R-R_{\text{ref}}]}}{\hat{\mu}_{[R_{\text{ref}}]}} \quad (5.11)$$

Where $\hat{\sigma}_{[R-R_{\text{ref}}]}$ is the standard deviation of the difference between R and R_{ref} . The observation approach that is capable of accurate monitoring of small-scale rainfall will yield a relative bias and CV near 0, and an r close to 1.

5.2.4 Sensitivity study

Radar relationship sensitivity

For the radar-derived rainfall fields the Marshall-Palmer relation is assumed. To show the dependency of the accuracy of radar rainfall estimation on the chosen Z - R relationship, the analysis is repeated where Eq. 5.4 is replaced by a range of power law relationships with varying exponents and coefficients. For comparison the validation of dual-pol radar rainfall fields is included as well.

Network density

The analysis described so far focuses particularly on the network layouts at the Amsterdam study area. For sparser networks, the accuracy of resulting rainfall fields is expected to decrease. The validation study of CML and PWS data sets is repeated for all numbers of links/stations fewer than the existing layout. For n links/stations, 50 random subselections of n out of the total network are determined. The validation is performed over the complete period of Events 1 and 2 for each subselection, yielding 50 outcomes per network density. This thinning analysis shows the dependency of the accuracy of PWS- and CML-derived rainfall fields on their respective network densities.

5.2.5 Isolated sampling errors

The PWS-derived rainfall fields include errors due to the temporal sampling, which is largely affected by accumulation interval, as well as spatial sampling directly related to the network density. To isolate the error due to the way the tipping buckets are able

to sample rainfall, the derived time series are compared with the ground truth without spatial interpolation.

When comparing the derived data sets with the ground truth as specified in Fig. 5.5 both the error due to uncertainties in the Z - R or k - R relationship used (for radar and CML, respectively) and the error due to spatial and temporal sampling are included. In order to isolate the sampling error, the validation is repeated for derived rainfall fields of radar and CML that are obtained directly from the ground truth rainfall fields without the use of any intermediate Z - R or k - R relationship.

5.3 Results

5.3.1 Overall comparison of datasets

Comparing the derived data sets of the two events with their respective ground truth data sets (at 5-min intervals) yields the scatter density plots from Fig. 5.9. Based on these 5-min values, the radar and PWS perform better than the CML, shown by the higher value of r and lower value of CV. The CV is highest in Event 2, where the rainfall variability is higher than in Event 1 (see Fig. 5.3).

The Marshall-Palmer Z - R relationship used to derive radar rainfall fields (Eq. 5.4) can also be described as

$$Z = \alpha R^\beta \quad (5.12)$$

where $\alpha = 200$ and $\beta = 1.6$ when Z is expressed in $\text{mm}^6 \text{m}^{-3}$ and R in mm h^{-1} . Figure 5.6 already showed that the optimized α and β values for both events differ from those in the Marshall-Palmer relation. The fitted Z - R relationship in Event 1 deviates mostly in the value of α , while the fitted relationship in Event 2 differs mostly in the β value from the Marshall-Palmer relationship. Although the radar bias in Event 2 is small, this seems to be caused by overestimation of small values and underestimation at high rainfall intensities (Fig. 5.9). To explore the dependence of the validation results on the Z - R relationship, the radar derivation and validation are repeated for a range of α and β values (Fig. 5.10). The patterns in Fig. 5.10, especially for bias, differ between the more homogeneous Event 1 and the convective Event 2. The bias from assuming the Marshall-Palmer relation is high in Event 1 (0.17), although in Event 2 the results are comparable with those from the optimized values. The variability in r is not included in Fig. 5.10 as it remained almost constant over these α and β ranges for both events.

Figure 5.9: Scatter density plots of radar, CML, and PWS-derived simulated observations for the two events, where all time series were aggregated to 5-min values of averaged rainfall intensities over the 5-min intervals. CV = coefficient of variation; CML = commercial microwave link; PWS = personal weather station.

Figure 5.10: Dependency of CV and bias on radar derived rainfall fields at 30-s temporal resolution for various combinations of α and β values in Eq. 5.12, indicated with crosses in the figure. Darker colors represents higher accordance with the ground truth. Note that the color scales are different for the different panels. The gray triangles indicate the Marshall-Palmer relationship, and the gray square shows the optimized values found in Fig. 5.6. CV = coefficient of variation.

5.3.2 Scale dependency of performance

Figure 5.9 only shows comparisons of the data sets at 5-min temporal and 100-m spatial resolution. It is expected that the correspondence between measurements and ground truth improves at larger aggregation scales. To show the dependency of the accuracy of each technique on spatial and temporal aggregation scale, the comparison has been repeated for various combinations of pixel sizes and time intervals. The rainfall intensities within a larger pixel and longer time interval are averaged. Figure 5.11 shows how CV and r change for both events. The bias is independent of scale and remains the same as

the values from Fig. 5.9 for each resolution.

Figure 5.11 confirms that overall agreement with the ground truth increases at larger scales. CML has a stronger improvement by temporal aggregation than radar and PWS. As the CML data set consists of snapshots at 15-min intervals, this is not surprising. Although the CV and r of CML resemble those of radar and PWS in Event 1, especially at slightly larger scales, the performance of CML during Event 2 is the poorest, indicating that highly variable precipitation is difficult to correctly estimate from CML at high space-time resolutions.

5.3.3 Isolation of sampling errors

PWS temporal sampling

The error in the PWS rainfall maps is partly due to spatial representativeness and partly due to the sampling strategy of the weather stations (Fig. 5.8). The current study enables us to evaluate the exact error due to the sampling strategy of PWS stations. Figure 5.12 shows the CV and r of the PWS-derived time series at the 84 locations compared with their respective ground truth values. Especially on time intervals smaller than 5 min the mismatch is considerable. This is expected as the PWS time series only contain multiples of tipping bucket volumes and report their measurements in varying intervals of ~ 5 min. The sampling error is larger in Event 2 than in Event 1, likely because Event 2 is much more variable (see Fig. 5.3). The bias is almost zero in both cases.

Radar and CML temporal and spatial sampling

Table 5.1 includes the r , CV and bias values found in the overall comparison from Fig. 5.11, as well as those where radar and CML are derived directly from the DSDs. This isolates the error from the spatial and temporal sampling strategy from errors due to assumptions in the respective $Z-R$ and $k-R$ relations. For comparison, the values from dual-pol derived radar are included. All comparisons are made on time series that were aggregated to 5-min values.

Table 5.1 shows that the error in radar rainfall is largely reduced without $Z-R$ uncertainty, while CML observations without $k-R$ uncertainty do not necessarily improve. Spatial interpolation in CML values (CML-A) yields a smooth field that resembles the ground truth rainfall field even better than the comparison at the link path centers only (CML-B). This shows that the error in CML rainfall retrievals is dominated by the spatial and temporal sampling, with only small errors due to assumptions in the $k-R$ relationship. The results for Rad-dual-pol and Rad-A are very similar for Event 1, because K_{dp} rarely exceeds the threshold value of $0.5^\circ \text{ km}^{-1}$ and the Marshall-Palmer approach was used for a large fraction of the event. In Event 2, dual-pol radar rainfall retrieval outperforms the

Table 5.1: Performance of radar and CML (5 min resolution)

	Rad - A	Rad - dual-pol	Rad - B	CML - A	CML - B
bias Event 1	0.17	0.174	-0.005	0.076	0.003
bias Event 2	0.012	0.198	0.002	-0.003	-0.073
CV Event 1	0.565	0.574	0.399	0.78	0.748
CV Event 2	2.137	1.931	1.609	2.946	2.962
r Event 1	0.92	0.921	0.953	0.825	0.826
r Event 2	0.85	0.88	0.911	0.655	0.653

A = all errors are considered as in Fig. 5.11

B = only sampling errors

$Z-R$ rainfall retrieval approach. The small bias of Rad-A in Event 2 (0.012) is caused by overestimation at small rainfall intensities compensated by underestimation of large values (Fig. 5.11). The scatter density plot of radar obtained with the dual-pol approach shows no systematic underestimation at higher rainfall intensities (not shown here), resulting in a larger overall bias.

5.3.4 Impact of network density

To determine how well the PWS and CML networks can capture rainfall patterns at lower network density, the validation study is repeated for subsets of the total networks. For each number of links and stations, 50 random subsets of the total network are used in the derivation for the entire event and validated.

Figure 5.13 shows that the difference between the 5th and 95th percentiles is largest for the sparsest networks, as the location of the chosen station/link subset may lead to overrepresentation or underrepresentation of rainfall peaks. As expected, the results are poorer for sparser networks. However, the increase in accuracy when more stations or links are included becomes less dramatic with network density. According to Fig. 5.13, approximately 20 links and 40 stations or more are required to describe the spatial patterns of rainfall fields in the 20×20 km study area for both events.

Figure 5.11: Variation of CV and r of derived data sets for radar, CML, and PWS compared with the ground truth as a function of spatial and temporal scales (indicated with black crosses). Darker colors correspond to better agreement with the ground truth. CV = coefficient of variation; CML = commercial microwave link; PWS = personal weather station.

Figure 5.12: Correlation (r , red) and CV (blue) values from the comparison between the 82 PWS time series with the ground truth (i.e., without interpolation) for multiple temporal aggregation scales. CV = coefficient of variation; PWS = personal weather station.

Figure 5.13: CV, r , and bias for the CML and PWS-derived rainfall fields at 30-s resolution for all numbers of links/stations between 1 and the total data set. The colored regions indicate the area between the 5th and 95th percentile of the 50 results per network density and the median is indicated with a circle (for Event 1) and square (for Event 2). The light colored circle and square indicate the values for the complete data set. CV = coefficient of variation; CML = commercial microwave link; PWS = personal weather station.

5.4 Discussion

This study investigates the error in rainfall measurements due to sampling and network characteristics. The measurement error is disregarded entirely, which influences the outcomes considerably. The following sections list further assumptions made per measurement technique.

5.4.1 Radar assumptions

In the derivation of radar data, it is assumed that the Marshall-Palmer Z - R relation applies. Figure 5.6 shows that the optimal Z - R relations found for the two events are different. The Z - R relationships from Event 1 ($Z = 253R^{1.60}$) and from Event 2 ($Z = 226R^{1.49}$) are within the range of the Z - R relations found in literature (e.g. Stout & Mueller, 1968; Raghavan, 2013). The underlying measurements used to simulate the DSDs were obtained during the HyMeX campaign (Ducrocq et al., 2014) in the Ardèche (France), which is a midlatitude region subject to Mediterranean rainfall. Nevertheless, the Z - R relationships derived from the resulting DSD patterns are not uncommon for typical Dutch weather (see also Fig. 4.14: graph in third row, first column in Uijlenhoet (2008)). The spatial sampling (1 km averaging instead of 100 m pixels) has a large impact on the Z - R fit, especially in Event 2.

When evaluating the dependency of radar results on the choice of Z - R relationship, the Marshall-Palmer relationship yielded similar results as the optimized Z - R relation for both events in terms of CV. However the effect in bias may be large for some events. Optimized Z - R relationships may differ per event.

Overestimation of low rainfall intensities and underestimation at high rainfall intensities leads to a small bias in Z - R derived rainfall fields, especially for Event 2. The dual-pol approach outperforms the Z - R approach in Event 2 due to more accurate measurements of high rainfall values, which results in a positive bias in the dual-pol comparison. In Event 1, the dual-pol approach and the Z - R approach perform equally well. This is because $K_{dp} < 0.5^\circ/\text{km}$ for most of the event, which implies that mainly the Z - R approach is applied.

5.4.2 CML assumptions

To obtain R from k we use a and b parameters as supplied from the RAINLINK software package. This R - k relation is slightly steeper for all microwave links than when R and k are calculated directly from the DSDs, resulting in a slight overestimation of R by using the RAINLINK parameters (not shown in figures). In the CML derivation, only DSDs on link paths are considered, and the resulting R values are interpolated with Ordinary Kriging. The error resulting from that approach is likely far larger than the error resulting

from assuming slightly inaccurate a and b values. The relatively small bias and high CV in the CML validation in Fig. 5.9 are in accordance with that statement.

The approach to derive CML measurements mimics the instantaneous sampling method at 15-min intervals that was used in the T-Mobile NL CML network. Often multiple CML networks are overlapping in the same area, operated by different telecom operators. Combining those data sets would increase the total link density, and therefore the capacity to detect small-scale rainfall patterns. However, in some areas of the world CML networks are replaced with fiber optic technology. The availability of CMLs and the willingness of local telecom organizations to share the data sets should therefore be evaluated per location.

The sample strategy to log CML data may vary between networks. A simulation study performed by Leijnse et al. (2008) showed the large discrepancy between rainfall estimations based on various sampling strategies, with unfavorable results for instantaneous sampling at 15-min intervals. Therefore, the results of CML would likely improve if measurements were obtained with an altered sampling strategy, for example, min/max sampling that logs min and max signal levels within an interval, and/or at a higher logging frequency. This is also clearly shown by Fig. 5.13, where the increase of links does not improve the measurement accuracy much after a certain density, as the errors due to temporal sampling dominate.

Interpolating with Ordinary Kriging produces fields that are likely to be much smoother than real fields, especially as the rainfall intensity over the link paths are attributed to the center of the CML paths. This is a simplified approach that has been applied before (Overeem et al., 2013a; Rios Gaona et al., 2015; Overeem et al., 2016a,b). In future research, the interpolation could be done in a more sophisticated manner, so that the entire link path is considered (e.g. Zinevich et al., 2008, 2009; Roy et al., 2016).

5.4.3 PWS assumptions

In order to obtain the rainfall for the personal weather stations, the rainfall intensities from the simulated DSDs of the overlying pixel were considered. It is assumed that the rainfall was constant over the entire $100\text{ m} \times 100\text{ m}$ pixel and the total 30 s of the interval. In reality, a point measurement is likely to differ from a pixel average. This is examined by Raupach & Berne (2016) by comparing point-scale DSDs with pixel-averaged DSDs, which showed considerable differences for pixels of $2.8\text{ km} \times 2.8\text{ km}$ and larger. For simplicity, and because pixels in this study are far smaller, we did not take this into account. As the temporal and spatial scales of the simulated rainfall event are already quite small, we expect that this effect is limited.

The PWS comparison without spatial interpolation shows the error resulting from the PWS sampling strategy. At short accumulation intervals, especially those smaller than the measurement update frequency of PWS stations, this makes up a considerable fraction of

the overall error. Spatial representativeness constitutes the other part of the error found in this study. Figure 2.10 (Chapter 2) illustrates that the error in rainfall estimation decreases for higher density of the PWS network in a small region in Amsterdam. In this simulation study the sampling error also decreases when more PWSs contribute to the rainfall field estimate. However, above a certain density, the improvement is limited as the temporal sampling errors dominate there.

5.5 Conclusion

This simulation study aims to investigate the highest possible rainfall monitoring accuracy of radar, CML, and PWS networks in realistic layouts. Validation studies have been performed on simulated radar, CML, and PWS data sets in configurations with realistic counterparts in Amsterdam. Simulations were done based on DSD-fields at high spatial (100 m) and temporal (30 s) resolutions. In order to determine the errors due to sampling, perfect measurement accuracy is assumed.

The results from the comparison with the known ground truth show that the PWS network is best able to capture small-scale rainfall. The bias is close to zero for both events and the CV and r values show best agreement with the ground truth for Event 2. Radar also performs well with the highest r coefficient and lowest CV for Event 1. However, a relative bias of 17% was introduced primarily due to the fitted $Z-R$ relationship that deviates from the Marshall-Palmer relation. The optimal radar $Z-R$ relation differs per event, and the impact of a poorly chosen $Z-R$ relation is larger on the bias than on CV and r . In Event 2 more accurate results are achieved with the dual-pol radar rainfall retrieval method.

More scatter occurs in the event with more spatial variability of rainfall, resulting in larger CV values. Although the bias of CML observation is small, the CV is higher and the r is lower than the other measurement techniques for both events. The overall error in CML rainfall retrievals is mostly caused by the temporal sampling strategy and far less by uncertainties associated with the parameters of the $k-R$ relation. Accuracy increases when considering larger aggregation scales. Resolution dependency is larger in time than in space for CML. The scale-dependency of radar observations is comparable to those of PWS.

With the layouts of the CML and PWS sensors, the error due to temporal sampling dominates over the error due to sparsity of the network. The temporal sampling error is especially large at aggregation intervals below the sampling interval. The error in PWS measurements is largest at aggregation intervals smaller than 5 min, even without spatial interpolation (Fig. 5.12). The accuracy of rainfall estimates is determined for sparser networks. Similar measurement accuracy would be achieved with a CML network of ~ 20 of the 58 links. The same is true for PWSs, which yield similar results until the network

is reduced to ~ 40 of the 82 stations in our 20×20 km study area. This corresponds with the correlograms of the two events (Fig. 5.3) where decorrelation distances are far higher than the typical distance between links and stations. This means that the CML and PWS data set may yield rainfall fields of similar accuracy when a large part of the data set is not (always) available or when part of the data set is excluded, for example, as a result of quality control.

This research has shown the theoretical potential of these techniques when each is assumed to measure with perfect accuracy. In realistic conditions, the measurements will be affected by considerable measurement and retrieval errors. The next step in this research will include these errors typically found in these techniques. This will give insight in their realistic potential for urban rainfall monitoring.

5.6 Background: Drop size distribution retrieval method

Estimation of DSDs

DSDs were estimated for points at 100-m resolution from radar data. The radar and data are described in Raupach & Berne (2017) as the HyMeX dataset. In this study we assume that although the radar data used were collected in France, they can be transposed over Amsterdam as reasonable estimates of conceivable Dutch rain events.

Radar scans were available every 5 to 6 minutes. To estimate the fields between these scans, advection was estimated for fields of Z_H (normalized fields with missing values set to zero), using the optical flow technique of Kroeger et al. (2016) (see also https://github.com/tikroeger/OF_DIS). Parameters were modified to produce a smoother flow (coarsest scale 6, patch size 80, forward-back consistency enabled). Radar scans were read for point data, as in Raupach & Berne (2017), for a $40 \text{ km} \times 40 \text{ km}$ grid of points at 100 m resolution within the radar domain. Each point on the 100 m by 100 m grid was assigned radar values from horizontally overlapping radar volumes, defined using the radar's 3-dB beamwidth (multiplied by two to produce a smooth field) and 75 m range resolution. Optical flow was calculated between consecutive observations on this grid. To further smooth the flow, the average flow was found using a moving window of three observed time steps.

The center $20 \text{ km} \times 20 \text{ km}$ area of the grid was used for DSD-retrieval. For two observations at time t_1 and t_2 , intervening fields for Z_H , Z_{DR} , and K_{dp} were calculated. Observations at t_1 were advected forward proportionally in time using proportions of the optical flow vectors for that pair of observations. Observations at t_2 were advected proportionally backward, using reversed flow vectors, to simulate the same intervening point. The two fields (forward and backward) for each variable were smoothed using a 3×3 pixel

moving window mean function, then combined as a weighted average with the weight favoring the closest observation in time (if one field had a missing value the unweighted value of the other field was used). The output field was smoothed again using a 3×3 pixel moving window mean. Observed fields were smoothed twice to match the advected fields. Advection and smoothing of Z_H and Z_{DR} fields was performed with the variables transformed to linear scale.

Improvements to DSD-retrieval algorithm

The DSD-retrieval technique used to generate fields of DSDs for this paper is that of Raupach & Berne (2017), improved in the following ways:

- The method was reparameterized using updated disdrometer data (improved instrument sampling area calculation, improved handling of canting angles in raindrop scattering property calculation). The parameters of the generalized gamma model used for the normalized DSD are $\mu = 2.23$ and $c = 1.68$. Moment three of the DSD is retrieved using parameters shown in Table 5.2, and the moment six estimate $\widehat{M}_6$ ($\text{mm}^6 \text{ m}^{-3}$) is retrieved using

$$\widehat{M}_6 = \begin{cases} Z_h^{1.01} & \text{if } Z_H \leq 28 \\ 2.68 Z_h^{0.86} & \text{if } Z_H > 28, \end{cases} \quad (5.13)$$

where Z_h is horizontally polarized radar reflectivity ($\text{mm}^6 \text{ m}^{-3}$) and $Z_H = 10 \log_{10}(Z_h)$ is the radar reflectivity in dBZ.

- If predicted values of mass-weighted mean raindrop axis ratio are greater than “Max Z_{DR} ” in Table 5.2, those values are replaced by 0.8.
- The noise correction routine for differential reflectivity Z_{DR} (dB) in Raupach & Berne (2017) overestimated the predicted value $\widehat{Z}_{DR}$ for high values of Z_H . An improved estimate was found by making two fits, forcing their continuity, and fitting to median values of linear Z_{DR} (ξ_{dr} [-]) per class of Z_H (in dBZ), using orthogonal least squares in log space. The “expected” value of linear Z_{DR} is estimated as

$$\widehat{\xi}_{dr} \sim \begin{cases} 0.979 Z_h^{0.018} & \text{if } Z_H \leq 27.01 \\ 0.599 Z_h^{0.097} & \text{if } Z_H > 27.01. \end{cases} \quad (5.14)$$

- The noise correction routine for specific differential phase shift K_{dp} was replaced with a simpler model trained as for Z_{DR} . “Expected” K_{dp} ($^\circ \text{ km}^{-1}$) is

$$\widehat{K}_{dp} \sim \begin{cases} 0.00007 \times Z_h^{1.152} & \text{if } Z_H \leq 21.61 \\ 0.00066 \times Z_h^{0.701} & \text{if } Z_H > 21.61. \end{cases} \quad (5.15)$$

- Hard thresholds on when to apply noise correction produced unrealistic spatial patterns. To improve this aspect of the performance, weighted means between the measured and expected values were used when measured values could have been noisy. Weights are determined as the proportion of the distance between two thresholds.
 - For Z_{DR} , we expect values under 0.2 dB to be noisy (determined for S-band by Bringi et al., 2002). In our training data, this level corresponds to Z_H between 11.6 dBZ (1st percentile) and 29.3 dBZ (99th percentile). Z_{DR} values for records with Z_H below 11.6 dBZ are replaced with the expected value, values above 29.3 dBZ are kept as the measured value, and those with Z_H values between the thresholds are replaced with a (linear) weighted mean of measured and expected values, with the weight proportional to the distance between the two thresholds in linear scale.
 - For K_{dp} , we expect values under $0.3^\circ \text{ km}^{-1}$ to be noisy (determined for S-band by Bringi et al., 2002). K_{dp} values are replaced as for Z_{DR} , with thresholds at 33.8 dBZ and 43.6 dBZ.
 - To deal with unreasonably low values of Z_{DR} and K_{dp} , a second replacement is made; using this time as the thresholds 10% and 40% of the expected value of Z_{DR} (in dB) and K_{dp} . Z_{DR} and K_{dp} are again replaced using these thresholds.

Table 5.2: Table 3 from Raupach & Berne (2017), with updated parameters $\hat{C}$ (Equation Raupach & Berne (2017, Eq. 16)) and c_i (Raupach & Berne (2017, Eq. 14)). The maximum value of Z_{DR} [dB] each relationship can handle is shown by “Max Z_{DR} ”, and drop axis ratios are those of Thurai et al. (2007), Andsager et al. (1999), Brandes et al. (2002), and Beard & Chuang (1987) (as written in Kalogiros et al. (2013)).

Ratio.	$\hat{C}$	c_1	c_2	c_3	c_4	c_5	c_6	Max Z_{DR}
Thurai	3.472	1	-0.073273	0.041236	-0.016342	0.002186	-0.000058	4.73
Brandes	3.297	1	-0.079489	0.052234	-0.023047	0.004255	-0.000285	7.99
Andsager	3.242	1	-0.091394	0.072726	-0.035161	0.007148	-0.000534	7.05
Beard	3.240	1	-0.087146	0.053131	-0.020315	0.002910	-0.000122	5.09

Chapter 6

Opportunistic sensing case studies

The ongoing urbanisation and climate change urges further understanding and monitoring of weather in cities. Two case studies during a 17-day period over the Amsterdam metropolitan area, the Netherlands, are used to illustrate the potential and limitations of hydrometeorological monitoring using non-traditional and opportunistic sensors. We employ three types of opportunistic sensing networks to monitor six important environmental variables: (1) air temperature estimates from smartphone batteries and personal weather stations; (2) precipitation from commercial microwave links and personal weather stations; (3) solar radiation from smartphones; (4) wind speed from personal weather stations; (5) air pressure from smartphones and personal weather stations; (6) humidity from personal weather stations. These observations are compared to dedicated, traditional observations where possible, although such networks are typically sparse in urban areas. First we show that the passage of a front can be successfully monitored using data from several types of non-traditional sensors in a complementary fashion. Also we demonstrate the added value of opportunistic measurements in quantifying the Urban Heat Island (UHI) effect. The UHI can be clearly determined from personal weather stations, though UHI values tend to be high compared to records from a traditional network. Overall, this study illustrates the enormous potential for hydrometeorological monitoring in urban areas using non-traditional and opportunistic sensing networks.

This chapter was submitted after revision in May 2019 to *Bulletin of the American Meteorological Society*, as:

¹De Vos, L.W., Droste, A.M., Zander, M.J., Overeem, A., Leijnse, H., Heusinkveld, B.G., Steeneveld, G.J. & Uijlenhoet, R., “Hydrometeorological monitoring using opportunistic sensing networks in the Amsterdam metropolitan area”,
of which a revised version was accepted for publication on 25 October 2019.

¹The first two authors contributed equally to this work and share co-first authorship.

6.1 Introduction

Traditionally, hydrologists and meteorologists, scientists and practitioners alike, have relied on dedicated measurement equipment in their research and operations. Such instruments are typically owned and operated by governmental agencies. Installed and maintained according to (inter)national standards, they offer accurate and reliable information about the state of environment we study, monitor and manage. Standard instruments are often based on novel measurement techniques that originate in the research community and have been tested extensively during dedicated field campaigns.

Unfortunately, the operational measurement networks available to the hydrometeorological community today often lack the required spatial and/or temporal density for high-resolution monitoring or forecasting of rapidly responding environmental systems. Apart from the high installation and maintenance costs of such dedicated networks, it can be challenging, if not impossible, to install meteorological monitoring instruments according to the official requirements in urban areas (Oke, 2006).

Yet, sensors are omnipresent in our environment nowadays, often related to the rapid development in wireless communication networks (e.g. McCabe et al., 2017; Balsamo et al., 2018; Tauro et al., 2018; Zheng et al., 2018). To make use of such opportunistic sensors could be greatly beneficial to (meteorological) science and environmental monitoring and management operations. Opportunistic sensors are devices that were not installed with the intention to generate large-scale (meteorological) observations, but can be used as such. They may not be as accurate or reliable as the dedicated equipment we are used to, let alone meet official international standards. However, they typically come in large numbers and are often accessible online. Hence, combined with smart retrieval algorithms and statistical treatment, opportunistic sensors may provide a valuable complementary source of information regarding the state of our environment.

This article surveys recent opportunistic sensing techniques in meteorology, from (1) rainfall monitoring using commercial microwave links (CML) from cellular communication networks, via (2) crowdsourcing urban air temperature, pressure and solar radiation using smartphones to (3) high-resolution urban monitoring of air temperature, pressure, humidity, wind speed, and rainfall using personal weather stations (PWS). We present a 17-day analysis for the Amsterdam metropolitan area, the Netherlands, where these opportunistic sensors are employed in a complementary fashion, in particular to provide detailed monitoring of the passage of a front, as well as to demonstrate the potential of such sensors to quantify the Urban Heat Island (UHI) effect.

This study aims to showcase the availability of several opportunistic sensing techniques

and their ability to capture meteorological events. Rather than exploring only a single technique and/or variable, the performance of multiple techniques for a range of meteorological variables during two events is evaluated. The results should therefore be seen as indicative of what can be achieved with opportunistic sensing for hydrometeorological monitoring. Details on the study area and period are provided in Sect. 6.2. The opportunistic techniques used in this study are described in Sect. 6.3, covering their background as well as how the datasets were obtained and processed. While other studies have developed extensive quality control algorithms, this work uses limited processing to showcase the raw potential of the techniques. In Sect. 6.4 the results are presented for the two cases, followed by a discussion and conclusions in Sect. 6.5, which presents an assessment of the relative performance of the opportunistic sensing techniques for describing meteorological events.

6.2 Case selection & study area

We selected Amsterdam (capital of The Netherlands) and its surroundings and the period between 6 June 2017 00:00 UTC and 23 June 2017 00:00 UTC as case study period (local time is UTC+2 hours). This period contains both sufficient data from opportunistic sensing techniques, and interesting meteorological events to illustrate the potential of the opportunistic sensing techniques. The selected region is bound by 4.67–5.05°E & 52.24–52.44°N (26 km × 22 km). To be able to distinguish between the inner city and suburbs, the study area was divided into two parts, i.e. the urban center dataset: 4.83–4.95°E. & 52.34–52.385°N and the suburban dataset 4.67–5.05°E. & 52.24–52.44°N, excluding the urban center area (see Fig. 6.1a).

The Netherlands has a temperate maritime climate (Köppen Cfb). With a mean temperature of 18.0°C and 50.5 mm of rainfall June 2017 was about 2.5°C warmer and 10.5 mm drier than the climatological mean (based on the past 30 years of observations at station WMO 06240 Amsterdam airport, from hereon referred to as “Amsterdam airport”). The month had eight summer days and two tropical days (max. temp. above 25 & 30°C respectively).

On June 6, a small low-pressure system developed over the North Sea off the coast of the Netherlands and passed over the country, resulting in a substantial pressure drop to 992 hPa, an hourly maximum wind speed of 54 km h⁻¹ (7 Bft) and 12 mm of rainfall measured at Amsterdam airport. In the morning of June 9, an active cold front brought in relatively cold air which resulted in 27 mm of rainfall. A clear-sky episode occurred 9–11 June, while another cold front passed in the early morning of June 12 (Fig. 6.2i). In the following period, no rainfall occurred, and temperatures were mild (daily maximum temperatures below 25°C), followed by a warm episode between June 16 and June 19.

Figure 6.1: Map of Amsterdam metropolitan area and city center with locations of all sensor networks: Personal Weather Stations (PWS) and Commercial Microwave Links (CML) (a), and of smartphone battery temperature readings and Amsterdam Atmospheric Monitoring Supersite (AAMS) stations (b).

On June 19 the maximum air temperature reached 29.8°C at Amsterdam airport. This warm episode ended with the passage of a cold front and associated precipitation and thunderstorms on June 22. For the remainder of the paper we will focus on two cases, i.e. case A, describing the passing front and resulting rainfall at the start of the study period, and case B, containing the hot summer period ending with a thunderstorm, with a focus on UHI detection.

For this study, the UHI is defined as the instantaneous urban air temperature difference between the city and the countryside (Stewart, 2011). The UHI develops as a result of the relatively low albedo of cities, high heat capacity of the urban fabric, thermal radiation trapping, and low surface evapotranspiration. The UHI is favored by weather conditions with high solar insolation (low cloud cover) and low wind speeds (Oke, 1982; Theeuwes et al., 2017). Earlier crowdsourcing observations indicated that Dutch urban areas experiences a mean daily maximum UHI of 2.3°C and the 95th percentile amounts to 5.3°C (Steenefeld et al., 2011). Ronda et al. (2017) found a mean evening UHI of ~1°C, and a maximum of 4.5°C in Amsterdam for the summer of 2015 as a whole.

Figure 6.2: Time series of opportunistic measurements of weather variables. PWS observations of temperature (a), cumulative rainfall (c), wind (f), pressure (g) and dewpoint depression calculated from humidity and temperature (i), smartphone battery derived air temperature (b), light (e) and pressure (g). The colored areas indicate the interquartile range (IQR is mean 25–75 percentile) of all observations at that time; lines show the median values, except for (b) where the line shows mean temperature. Shaded areas indicate night-time. Pearson correlation (r), standard deviation of the difference (SD) and absolute bias (bias) are calculated based on hourly values compared with WMO observations at Schiphol.

6.3 Sampling techniques

6.3.1 Traditional sensing methods

Gauge-adjusted radar dataset

The Royal Netherlands Meteorological Institute (KNMI) operates two C-band Doppler weather radars. The 5-min reflectivity data from these radars are combined into one composite using a weighing factor as a function of distance from the radar. Beekhuis & Mathijssen (2018) provide detailed characteristics on the radars and the processing of their data. Reflectivity factors Z ($\text{mm}^6 \text{ m}^{-3}$) are converted to rainfall intensities R (mm h^{-1}) with a fixed Z - R relationship (Marshall et al., 1955), $Z = 200R^{1.6}$, and, subsequently, accumulated to rainfall depths for different durations. The two KNMI rain gauge networks are employed to adjust the radar-based accumulated rainfall depths: an automatic network with 1-h rainfall depths for each hour (~ 1 station per 1000 km^2) and a manual network with 24-h 08:00–08:00 UTC rainfall depths (~ 1 station per 100 km^2). A daily spatial adjustment utilizing the manual gauge data is combined with an hourly mean-field bias adjustment employing the automatic gauge data. The resulting gauge-adjusted radar rainfall dataset has a spatial resolution of 0.9 km^2 , with no missing data for the study period. Overeem et al. (2011) provide a more detailed description of this radar dataset, which largely uses the methodology developed by Overeem et al. (2009a,b). Finally, 15-min path-averaged rainfall intensities are derived from the radar pixels covering each link path of the CML dataset (described in Sect. 6.3.2). The gauge-adjusted radar rainfall dataset is used as a reference to validate rainfall estimates from CMLs and PWSs.

WMO station Amsterdam airport

The WMO station Amsterdam airport, WMO 06240 (4.78°E , 52.32°N ; Fig. 6.1a) provides hourly air temperature and cloud cover observations. This surface synoptic station is operated by KNMI, situated in a polder (4.18 m below MSL) and surrounded by meadows, arable land, and buildings as well as infrastructure from Amsterdam airport. Air temperature is observed at 1.5-m height above short mowed grass. The sensor is covered by a radiation screen and well ventilated. Cloud cover aloft is obtained from a LD40 ceilometer, which uses LIDAR to detect the height and concentration of particles, such as cloud droplets. KNMI (2000) provides more information on the temperature observation.

Amsterdam Atmospheric Monitoring Supersite

As an urban reference network we utilize the observations from the Amsterdam Atmospheric Monitoring Supersite (AAMS; Ronda et al., 2017), which consists of 30 weather stations across the city. The network consists of temperature and humidity sensors (Decagon VP-3, U.S.A.) mounted inside a 184 mm aspirated radiation shield (Davis, U.S.A.). The ventilation fan is powered by 2 small solar panels mounted on top of the shield. The fans work at global radiation levels $>100 \text{ W m}^{-2}$. The radiation screens are mounted onto lantern posts using a boom to mount the center of the radiation screen 0.46 m away from the edge of the lantern post at a height of 4.0 m above ground level. The sonic anemometer (Decagon DS-2, U.S.A.) has an accuracy of 0.30 m s^{-1} or 3% (whichever is larger). The anemometers were mounted above the radiation screens 0.50 m away from the lantern post edges and at heights of 4.30 m (from ground level to center of the anemometer).

6.3.2 Opportunistic sensing methods*Smartphone data*

Smartphones contain many sensors to support their functionality, including sensors for light levels to adjust screen brightness, pressure sensors to complement the GPS for an accurate (vertical) location estimation, and thermometers for the battery to avoid damage from overheating. Readings from such sensors can be used for opportunistic environmental sensing by collecting them through mobile applications ('apps'). Through these apps, the sensor readings are sampled with a certain frequency, along with the last stored GPS coordinates. Examples of apps that collect and store smartphone sensor readings include Pressurenet (<http://www.cumulonimbus.ca/>) (Mass & Madaus, 2014; Madaus & Mass, 2017), OpenSignal (<https://opensignal.com/>), and Atmos (Niforatos et al., 2014, 2017).

In Mass & Madaus (2014); Madaus & Mass (2017), smartphone pressure data was shown to improve representation of convective events when assimilated into NWP models. The raw smartphone pressure readings were quality controlled to only include one value per smartphone per assimilation time step, and were also corrected for the terrain elevation and checked for spatial and statistical consistency (Madaus & Mass, 2017). Niforatos et al. (2017) compared smartphone light sensor readings with manually reported classifications of weather, which showed light readings to be indicative of present weather conditions. City-wide air temperatures can be estimated from smartphone battery temperature readings, as has been shown for eight major cities (Overeem et al., 2013b), for the city of Birmingham (Muller et al., 2015) for daily temperatures, and for São Paulo for hourly and daily temperatures (Droste et al., 2017). Statistical training with independent temperature measurements was performed based on a steady-state heat transfer model: a

smartphone is typically, although not always, carried close to the user's body. The thermal energy generated by the smartphone must be balanced by heat exchange to the body and the environment. The conductive heat flow between two adjacent systems is assumed to be proportional to their temperature difference, and depends on the thermal insulation between smartphone and environment, and between smartphone and body. This principle allows us to estimate hourly-averaged air temperatures from hourly-averaged battery temperatures (Overeem et al., 2013b):

$$\bar{T}_{e,j,h}^{A,\text{hour}} = m_j^h (\bar{T}_{\text{bat},j,h}^{A,\text{hour}} - T_0) + T_0 + \epsilon_{j,h}, \quad (6.1)$$

where $\bar{T}_{e,j,h}^{A,\text{hour}}$ is the hourly mean urban air temperature, $\bar{T}_{\text{bat},j,h}^{A,\text{hour}}$ is the hourly-averaged battery temperature (both in space A and time), and T_0 a constant equilibrium temperature. m_j^h is a coefficient, $\epsilon_{j,h}$ is a random disturbance, and h denotes the hour. Recent work has shown improved results when using both in- and out-of-pocket models, but the evaluation occurred in a lab-experiment with two smartphones only (Chau, 2019). In this analysis the in-pocket model approach (Eq. 6.1) that was successfully applied on large datasets (Overeem et al., 2013b; Droste et al., 2017) is therefore used.

For this study a large dataset of observations is obtained from the Android application OpenSignal, which crowdsources data relevant to wireless connectivity along with the aforementioned sensor readings. Compared to the previously mentioned studies, readings were obtained at a far higher frequency, i.e. 15-s intervals whenever the smartphone screen is active, not requiring the app to be opened by the user. A total of 3.14 million smartphone observations are available for the entire study period for the Amsterdam metropolitan region (Fig. 6.1a). The OpenSignal dataset includes self-reported accuracy scores of the light and pressure readings, as determined by the sensor management software in the smartphones (Android, 2019). Only readings with the highest possible accuracy were included in our analysis.

A smartphone light sensor measures illumination in lux (lumen m^{-2}), i.e. irradiance weighted for the visible part of the electromagnetic spectrum, so a measure for the perceived brightness for the human eye. To estimate the equivalent solar radiation, an empirical factor of 0.0079 lux per W m^{-2} is used, based on the spectral distribution of sunlight (Chua, 2009). By applying this transformation, the readings are treated as if they were measurements of solar radiation. This is a fairly large assumption, as it is expected that most readings will not be made in a representative manner: with the smartphone perpendicular towards the sun and in direct sunlight. User behavior plays a large role (e.g. indoor versus outdoor measurements), so one may expect that most light readings will underestimate the solar radiation, resulting in a skewed distribution. The smartphone light sensor has a limited view angle ($<180^\circ$) and not a good cosine response. Additionally, the sensor can oversaturate at high light intensities. Therefore it is desirable to have many readings to increase the probability of observations taken in favorable conditions (unshaded and perpendicular to direct sunlight). Only light sensor readings

above 0 lux are taken into account, leaving 2.32 million readings in the whole study period.

All smartphone pressure sensor readings below 950 hPa are excluded, based on the lowest recorded pressure in the Netherlands, 954.2 hPa (De Haij, 2009), which results in a dataset of 2.06 million readings. We only include battery temperature readings between 10 – 47°C when the smartphone is not charging. Hourly battery temperature readings are averaged spatially over the city center domain (Fig. 6.1b), light and pressure are averaged over the entire region for each hour (Fig. 6.1a).

Ambient air temperatures are estimated from battery temperature (Eq. 6.1); the value of equilibrium temperature (T_0) as optimized by Overeem et al. (2013b), 39°C, is used. Figure 6.1b shows the positions of the underlying 0.4 million battery temperature readings. Two different datasets are derived: one without and one with optimizing the coefficients of the heat transfer model for the available dataset. The first dataset uses a fixed value of m_j^h for all hours, 2.4, as found for a summer period in London based on daily averages (Overeem et al., 2013b). These results, without further model calibration, are presented in Fig. 6.2b, which also shows the 25th and 75th percentile. For the second dataset, records from 1 June 00:00 UTC – 15 June 00:00 UTC are employed to calibrate a value of m_j^h for each clock-hour (24 in total). These optimized values, found using a least squares regression, are applied to the validation dataset from 16 June 2017 00:00 UTC – 23 June 2017 00:00 UTC.

Commercial microwave links

Cell phone communication relies on a telecommunication link network that consists of transmitting and receiving antennas, typically several km apart, between which radio signals propagate. Telecom operators commonly use signal frequencies that are sensitive to hydrometeors. This causes attenuation of the microwave link signals when liquid precipitation occurs between the antennas. Upton et al. (2005) first suggested to use signal attenuation in CML networks, which is typically monitored for quality control purposes, to determine rainfall. Soon after, this was shown to be successful with actual CML data (Messer et al., 2006; Leijnse et al., 2007b). This was promising as microwave link networks are widespread, also in areas of the world with limited to no traditional rainfall sensors.

The relation between rainfall attenuation and rainfall intensity can be described with a power law between path-averaged specific signal attenuation (k in dB km⁻¹) and link path-averaged rainfall intensity (R in mm h⁻¹) (Atlas & Ulbrich, 1977):

$$R = ak^b, \quad (6.2)$$

where

$$A = \text{TSL} - \text{RSL}; \quad k = \frac{A_{\text{wet}} - A_{\text{dry}} - A_a}{L}. \quad (6.3)$$

Coefficients a (in $\text{mm h}^{-1} \text{ dB}^{-b} \text{ km}^b$) and b (-) are dependent on signal frequency and polarization (Olsen et al., 1978; Jameson, 1991). TSL and RSL are the transmitted and received signal level (dB) respectively, A_a is the attenuation due to wet antennas (dB), A_{wet} and A_{dry} are the attenuation under wet and dry weather conditions respectively (dB) and L is the length of the link path (km). The specific attenuation due to rainfall is what remains when the attenuation due to other causes (i.e. dry weather conditions and wet antennas) are subtracted.

Subsequent research has focused on improving the techniques to obtain accurate rainfall estimates from these datasets, (e.g. Leijnse et al., 2008; Zinevich et al., 2010; Overeem et al., 2011; Chwala et al., 2012) and produce rainfall maps (Overeem et al., 2013a, 2016b) with real-time applicability (Chwala et al., 2016; Andersson et al., 2017; Chwala et al., 2018). Comprehensive overviews of literature on this technique were provided by Messer & Sendik (2015), Uijlenhoet et al. (2018), and Chwala & Kunstmann (2019). Several tools have been developed, documented and made (freely) available for users to construct rainfall observations with CML data: ‘Rcmlrain’ (<https://github.com/fencimar/Rcmlrain>), ‘RAINLINK’ (<https://github.com/overeem11/RAINLINK>), ‘pycomlink’ (<https://github.com/pycomlink/pycomlink>), and ‘pySNMPdaq’ (<https://github.com/cchwala/pySNMPdaq>).

The time series shown in Fig. 6.2d originate from the T-Mobile CML network visualized in Fig. 6.1a. Between 2017-06-06 00:00 UTC and 2017-06-10 14:00 UTC, 74 links were operational in the study area. Power levels were instantaneously sampled every 15 min. Due to data transfer issues, no power levels were available at the end of the study period. Rainfall time series for each link were constructed with the open source package RAINLINK (Overeem et al., 2016a), using the approach and optimized parameters from Chapter 4. A constant value is assumed for A_a . The wet antenna attenuation makes up a larger fraction of the total attenuation for short links, meaning that a small error in A_a will result in a relatively large error in k for short links, and the effect on the estimated value of R would subsequently be larger than for long links given the same error in A_a .

Crowdsourced personal weather stations

PWSs allow anyone to measure weather variables in their direct environment. Many automatic PWSs can upload their measurements directly to on-line platforms where they can be visualized and shared. Weather Underground

(<https://www.wunderground.com/wundermap>), WOW-NL (<https://wow.knmi.nl/>) and the Netatmo Weathermap (<https://weathermap.netatmo.com/>) are examples of platforms where weather observations are visualized in real time. Ideally, weather variables can be crowdsourced from such platforms in far higher spatial and temporal resolution than from traditional sensor networks.

The devices are often low-cost with a lower expected measurement accuracy than typical sensors from meteorological institutes. The PWSs are installed by citizens without expert knowledge on sensor placement requirements and/or lacking available measurement site without interference from surroundings. Hence we expect that many of the PWSs generate compromised measurements. For tipping bucket rain gauges, obstructions (e.g. insects, twigs) and the device not being completely level with the ground, could hinder the tipping mechanism. A shielded location will also lead to underestimation of rainfall. Overestimation of rainfall can result from PWS owners cleaning or handling the device, resulting in tipping bucket tips, creating measurements of artificial rain. PWS wind measurements are also largely affected by their position in relation to obstacles and the shielding effect of buildings. Furthermore, PWSs with a sonic anemometer are sensitive to rain blocking the path of the sound waves, so data quality might be compromised during rain events. Urban wind is highly variable in space, and is often measured as profile using e.g. LIDAR (Drew et al., 2013), so spatial averaging of PWS wind measurements is needed to obtain useful data. Temperature readings are highly affected by direct radiation: the lack of a proper radiation screen in most PWSs can result in overestimation of temperature by several degrees when positioned in direct sunlight (Bell et al., 2015; Chapman et al., 2017). Finally, the updates of measurements to the platform can be infrequent, and connectivity problems will result in large gaps in the time series.

Only a few studies compared PWSs with high-end sensors; temperature, relative humidity, radiation, pressure, rainfall, wind speed and direction: Jenkins (2014); Bell et al. (2015), temperature: Meier et al. (2015), rainfall: Chapter 2. Other studies have benefited from available PWS temperature records in cities. The UHI is then defined as the difference between PWS temperatures and a rural reference station (Meier et al., 2017; Chapman et al., 2017; Fenner et al., 2017; Golroudbary et al., 2018; Napoly et al., 2018). Cornes et al. (2019) suggest a statistical correction for short wave radiation bias in PWS temperature data. Preliminary work has been performed on crowdsourced wind (Droste et al., 2018) and rainfall measurements (Chapter 2, Golroudbary et al., 2018; Chen et al., 2018) (and explored with simulated PWS rainfall measurements in Chapter 4 as well). In other studies code has been developed and made available to apply quality control on crowdsourced PWS data (the CrowdQC R-package for PWS temperature observations <https://depositonce.tu-berlin.de/handle/11303/7520.3> and TITAN <https://github.com/metno/TITAN/>) but in this study we focus on the raw observations.

Measurements from all Netatmo weather stations in the Amsterdam study area (Fig. 6.1a) are evaluated. All devices measure temperature, pressure and humidity. Additionally, rain and/or wind are measured in case those optional modules are installed for that PWS. In order to standardize the variable time intervals, all measurements are attributed to the timestamp of the 5-min interval in which it occurred. If multiple measurements occurred within the 5 min interval they are averaged (or accumulated in case of rainfall). The PWS network is visualized in Fig. 6.1a. The raw measurements over the study period are shown in Fig. 6.2 (panels (a), (c), (f) and (h)), where panel (i) indicates the dewpoint depression (*DPD*) as calculated from the temperature and humidity measurements from the PWS. *DPD* is here preferred over dewpoint temperature itself to identify the frontal passage in case A.

6.4 Results

6.4.1 Case A: Weather front

First we focus on the passage of a cold front over the study area on June 9. At 6 UTC the operational model analysis provided by KNMI locates the frontal zone to the west of Amsterdam (not shown), and by 12 UTC the front has passed the city. Prior to the frontal passage itself, an upper air disturbance moved over Amsterdam between 3 and 4 UTC, bringing strong convection and rainfall. Such frontal zones cause distinctive behavior in various meteorological variables, which we expect to be distinguishable in the crowdsourced data (Fig. 6.2).

The passage of the front is clearly visible in the observed dewpoint depression and the wind speed (Fig. 6.3a). The dewpoint depression steadily drops during the approach and passage of the cold front, reaching a minimum of 1.4°C at 9 UTC. Between 10 and 11 UTC, when the front has passed, the dewpoint depression increases again up to 6.8°C , indicating the cold and dry air mass brought in by the cold front. Crowdsourced and reference wind speed steadily increase as the front passes (from 2 to over 4 km h^{-1}), before reaching its maximum (5 km h^{-1}) directly after the passage. The convection associated with the upper air disturbance at around 3–4 UTC generates a strong peak in the wind speed. Despite the unknown measurement setup of the PWS anemometers, the average signal of all PWSs corresponds well to that of the quality-controlled reference AAMS network (mean bias of 0.4 km h^{-1}), which shows the same behavior for the upper air disturbance and the front passing. However, the AAMS signal indicate a delayed onset of the wind speed increase (at around 9 UTC) and takes longer to reach a higher maximum wind speed.

The ambient air pressure (Fig. 6.3b), measured by PWSs and smartphones, starts

Figure 6.3: (a) Hourly average wind speed measured by PWS (solid line) and AAMS (dotted line), as well as the PWS dewpoint depression (red line, right y-axis). (b) Hourly averaged air pressure measured by PWS (solid line) and smartphone (dashed line) on June 9. The vertical lines indicate the boundaries of the frontal passage. At 6 UTC the front is located to the west of Amsterdam; at 12 UTC the front has passed over the city.

increasing at the moment the front passes (8 UTC). Typically, air pressure decreases before a cold front, then rapidly increases during the passage, and increases at a slower rate afterwards. The expected drop before the front is not very pronounced in the measurements: there is a slight decrease in pressure between 0 and 2 UTC (1.7 hPa decrease for PWS; 3.5 hPa for smartphone). The latter is more likely associated with the upper air disturbance. After the frontal passage at 8 UTC, the pressure rises, from 1006–1008 hPa (PWS–smartphone) up to a maximum of 1013–1016 hPa at midnight. The pressure tendency (change in pressure over time) remains roughly 1 hPa hour^{-1} after the front has passed.

The light intensity as measured by smartphones shows a distinct diurnal pattern in Fig. 6.2e, following the trend in solar radiation. The measured data is strongly skewed, so the median light intensity values are low (Fig. 6.2e). Figure 6.4 shows the 99th percentile of

Figure 6.4: Hourly 99th percentile of smart phone radiation (green lines) with Amsterdam airport cloud cover in oktas on June 9 (blue, circles) and June 18 (orange, diamonds). June 9 (blue solid line) and June 18 (orange solid line) are a cloudy and cloud-free day, respectively.

light readings to capture the readings made in the most favorable light conditions (see Sect. 6.3). The sky on June 9 is overcast (8 oktas) until 11 UTC, at which time the front has passed over Amsterdam and the sky clears up to scattered cloudiness (Fig. 6.4). The light intensity is also very low until 10 UTC, even though this is well within daylight hours. Compared to June 18 (a clear day) the light intensity is roughly halved, and the shape of the line is not as symmetrical (as we would expect from the diurnal cycle of global radiation). The green lines in Fig. 6.4 indicate the other days over the study period, showing the strong variability in the daily course of light intensity. The light intensity measured by smartphones not only depends on incoming radiation, but also strongly on user behavior (indoors vs outdoors, the angle of the phone) and the type of light sensor in the smartphone, which can differ between brands. The light sensor may also be oversaturated during high light intensities, resulting in flattened peak values.

The light intensity peak at 4 UTC coincides with the the upper air disturbance seen in Fig. 6.3, but is actually an artifact of the low number of observations. The number of available observations is higher during the day than during night and early morning, since it is related to user activity whether the smartphone logs an observation (as detailed in Sect. 6.3.2). At 4 UTC there are only 502 smartphone observations, compared to 10,373

at 15 UTC (17:00 local time, the typical end of the working day), so the data is more sensitive to outliers.

Figure 6.5: Double mass plots of commercial microwave link derived rainfall observations (a) and PWS rainfall observations (b) on 9 June 2017 against the reference of respectively the path-averaged and overlying pixel gauge-adjusted radar rainfall observations. The red lines in (a) indicate the shorter links (< 2 km). Information on the polarization (H or V), path length and frequency of the CML network is provided in the inset.

The upper air disturbance, and subsequent frontal passage, of June 9 results in 27 mm precipitation as measured by the gauge-adjusted radar reference. Figures 6.2c and 6.2d show that the peak of rainfall occurs after sunrise, coinciding with the timing of the frontal passage. Figures 6.5a and 6.5b depict the cumulative rain over June 9, measured by CML and PWS, against the reference. Total amounts differ between the two methods, but both show the same time response. The relatively short links (< 2 km) overestimate rainfall, with the majority reporting > 30 mm precipitation (relative bias is 87%). The longer links (≥ 2 km) also tend to overestimate, but much less extreme (relative bias is 12%). Although the expected uncertainty in rainfall estimates is higher for short links, the larger systematic bias (54% for all links) indicates that the methods to derive rainfall (RAINLINK) were not ideal for this rainfall event, especially for short links.

PWS measurements tend to underestimate the rain as measured by the reference, with some occurrences of large reported rainfall values that are not otherwise captured (Fig. 6.5b). Nevertheless, the majority of PWSs seem to agree overall with the reference

Figure 6.6: Map of 60-min rainfall depths over the Amsterdam metropolitan area based on gauge-adjusted radar data (pixels; 100% availability), CML data (paths; only CMLs with 100% availability are shown), and PWS data (circles; only PWSs with at least 83.3% availability are shown).

(Fig. 6.5b). The spatial distribution of rainfall (Fig. 6.6) measured by PWS and CML corresponds to that of the gauge-adjusted radar reference. We find that areas with high precipitation in the reference also yield high accumulations in the CML and PWS data in these areas. The overestimation by short links up to 8 mm is visible to the northwest of the band with high precipitation. The precipitation observations by PWSs correspond well to the spatial pattern of rain, although a number measure little (< 1 mm) rain during the hour represented in Fig. 6.6. These stations are mainly clustered in the city center. The large amount of obstructions inside the city center could reduce the rainfall received by the stations, which may explain part of the underestimation tendency already seen in Fig. 6.5b.

Figure 6.7: (a) Time series of temperature measurement according to the median of all (blue), center (red) and suburban (green) PWSs, as well as the median of the AAMS network (orange) and the mean smartphone derived air temperatures (purple). (b) The difference between the median PWS center and suburban temperatures (red, dashed); the AAMS and WMO (blue, dashed); and the PWS center and WMO (green, dashed). Shaded areas indicate night-time. (c) Scatter plot of hourly median PWS and smartphone temperatures against median AAMS station data, with Pearson correlation (r), standard deviation of the difference (SD) and absolute bias (bias).

6.4.2 Case B: Urban Heat Island

The last days of the study period is characterized by high temperatures and generally clear, sunny weather, leading to higher urban temperatures (PWS median air temperature up to 30°C on June 19, Fig. 6.7). Air temperature is measured by PWS, and derived through the smartphone battery temperature using the second, calibrated dataset (Sect 6.3.2). The AAMS network serves as urban reference, and the Amsterdam airport measurements are used as rural reference for the UHI (Fig. 6.7b). The smartphone-derived air temperature differs clearly from the PWS and AAMS measurements, with more erratic behavior and strong minimum values at night and early morning (as low as 7°C when the AAMS values are above 16°C). During daytime the smartphone-derived temperatures correspond better with the PWS and AAMS measurements. The diurnal cycle is clearly visible: the low values at night are most likely due to a low number of measurements available, increasing the sensitivity to outliers. Despite these occasional large deviations, the bias amounts to -0.1°C compared to AAMS, which is relatively small. A large systematic overestimation (2.0°C) is found when a fixed literature value for m_j^b is used, for the time series shown in Fig. 6.2b.

The diurnal pattern of air temperature between center and suburban PWSs is similar,

although the center stations tend to be warmer at night, and colder during the day (Fig. 6.7a). The AAMS air temperature is typically about 2 to 3°C lower during the day: this could partially be caused by the lack of shielding of the Netatmo temperature sensor, which makes it sensitive to radiation errors. In Fig. 6.7b the UHI has been plotted, including the UHI estimated by subtracting the center and suburban PWS (red dashed line). This particular PWS–UHI shows spatial variability within the PWS data, which is most pronounced during daytime, where the difference can be up to -1.5°C (i.e. the center is 1.5°C colder than the suburban area). Higher urban shading in the morning, and the faster heating rate of the relatively thin rural boundary layer compared to the deeper urban boundary layer cause this urban cool island in the morning (Theeuwes et al., 2015).

The other two UHI estimates are constructed using Amsterdam airport as rural background, showing that the city center is indeed much warmer at night than the rural surroundings. Urban cool islands typically form in the morning, persisting for several hours before the city heats up more. A remarkable 6°C UHI peak is visible on June 22, in the afternoon (13–14 UTC). This seems to be mainly caused by the Amsterdam airport temperature, since the PWS–UHI (which has no true rural reference) shows a value close to 0°C at that time. This is visible in Fig. 6.7a, where temperatures rapidly decrease in the course of a few hours on June 22 afternoon. Thunderstorms were reported on this day, and several mm of rain were measured at Amsterdam airport (according to radar) between 14 and 15 UTC. The UHI in this case is likely caused by the sudden cooling of the rural reference, rather than strong urban heating.

The spatial variability present in the PWS is captured in Fig. 6.8, which is an hourly snapshot of the UHI of PWS and AAMS stations against the Amsterdam airport reference, between 2 and 3 UTC, the hour of the day where the UHI is typically largest, on June 18. The cluster of stations in the center yields higher values than the suburban stations, although in both areas many stations deviate from this trend. The center PWS report an average UHI of 4.0°C, the AAMS UHI is 3.6°C, whereas the suburban areas have an average UHI of 2.7°C. Variability between measurement sites is high: there are stations with a temperature difference of up to 12.4°C, and even a few that have 0 or negative UHI (up to -0.6°C).

Figure 6.8: Urban Heat Island map for the Amsterdam metropolitan area (black rectangle) showing difference between hourly averaged air temperature for AAMS network (squares) and PWSs (circles) with respect to 1.5-m air temperature at WMO station Amsterdam airport observed at 3 UTC (triangle). The blue rectangle represents Amsterdam city center, the remainder of the metropolitan area is suburban. Only stations with at least 80% availability are shown. Of the 309 PWSs only 4 are colder than WMO, at most 0.6°C, and 24 are at least 6.0°C warmer than WMO, at most 12.4°C.

6.5 Discussion and conclusions

We have shown the potential and limitations of using meteorological data derived from several non-traditional sources. These data were compared to more traditional observations where possible. We explicitly consider observations that can be obtained near-directly from the opportunistic sensors, without applying many correction schemes, to illustrate their inherent potential: validation using the available quality assurance schemes was not the aim of this research. We use temperature from smartphone batteries and personal weather stations (PWS), precipitation from commercial microwave links (CML) and PWS, solar radiation from smartphones, wind speed from PWS, air pressure from smartphones and PWS, and humidity from PWS. Two case studies in a 17-day period over the city of Amsterdam, the Netherlands, are explored. In the first case study we show how the passage of a front is apparent from many of the data sources. The second case study shows that these measurements can be valuable in monitoring the Urban Heat Island (UHI) effect, especially given the fact that WMO stations in urban areas are very rare.

The passage of a front is visible in all of the opportunistic sensing data sources that we have considered. The dynamics of the temperature (especially from PWS, less so from smartphones), precipitation, solar radiation, wind speed, air pressure, and humidity all show the passage of the front. However, not every aspect of the weather events is sufficiently captured by the data: techniques using smartphone observations can only estimate a variable as a spatial average over the city and cannot be used to describe detailed spatial variability. Information from opportunistic sensing should be thoroughly quality-controlled, and when possible used as a complement to high-end traditional sensing techniques.

The PWS are suitable for monitoring the UHI. When compared to the AAMS urban reference network, these PWS show an UHI of the same order of magnitude (2–4°C), especially during the night. We observe urban cool islands during the period between sunrise and local noon. Air temperatures derived from smartphone battery temperatures exhibit much more noise than PWS temperatures, which limits their use for UHI measurement. It should be noted that the PWS thermometers are not shielded from solar radiation, whereas the AAMS are. This is clearly visible in Fig. 6.7, where the PWS temperature (and derived UHI) increases much more quickly than the AAMS temperature. This is in line with the findings of Bell et al. (2015). A rural reference station is needed to quantify the UHI. The AAMS network does not have a rural station that can be used for this, and the PWS network has most sensors where most people live: urban areas. Even the PWS locations outside the city center of Amsterdam (suburban, see Fig. 6.1 and 6.8) are mostly in built-up areas, and are hence expected to experience the UHI, although less severely. This is supported by Fig. 6.7b, where the temperature difference between

the city center and the suburban areas is much less pronounced than when the WMO station is used as rural reference. The temporal UHI pattern is visible in the PWS data, but exact quantification of the UHI requires a rural reference. In using such a rural reference, careful analyses of differences in weather conditions between the city and the rural locations should be done. This is clear from the very high apparent UHI on June 22 (see Fig. 6.7), which was caused by local cooling at the rural site due to a thunderstorm. Finally, we note that the distinction between center and suburban in this study was made rather crudely. In future UHI studies we recommend a more sophisticated partitioning of the stations into different classes (such as Local Climate Zones (Stewart & Oke, 2012)).

Average wind velocities from the PWSs are very low compared to what would be expected during the passage of a front. This may in part be due to how the PWS anemometers are installed. However, the carefully installed AAMS stations show average wind velocities of the same order of magnitude (around 5 km h^{-1}), indicating that the placement of the anemometers does not play a large role here. The reasons for these wind speeds to be lower than expected lies in the fact that the urban wind measurements are made at a lower level, and the urban fabric greatly reduces wind speeds at street level (Macdonald, 2000). This also means that wind speeds are expected to be highly variable across the city, which is clear from Fig. 6.2f. Spatial averaging over the city is therefore needed in order to see clear signals in the wind speed. Spatial averages of wind speed show the same behavior between PWS and AAMS, indicating their use to measure the urban wind as a whole. It should be noted that the wind sensor on the Netatmo PWS is a sonic anemometer, which are negatively affected by precipitation, hence wind observations during rainfall can be less reliable.

Data from PWS and CML are shown to provide useful information on both rainfall amount and space-time variation. Their ability to show detailed variations in space and time makes them useful for qualitative use in rainfall monitoring. The CML network overestimated rainfall in case A (Fig. 6.5a), although the relative bias of long links ($\geq 2 \text{ km}$) was 71% smaller than that of short links ($< 2 \text{ km}$). This is likely related to the larger error contribution wet antennas have for shorter links, and that the correction was calibrated on a different dataset, possibly with more long links (Chapter 4).

The PWS also show a good agreement with the reference, although most stations underestimate precipitation (Fig. 6.5b). This may be due to the higher wind speed above the urban fabric which could cause buildings to act as a shield for the PWS rain gauges. It is also apparent from Fig. 6.5b that some PWSs report either zero rainfall when there is clearly rain or large amounts of rainfall where there was none. Such errors could be corrected by using automated filters (Chapter 2 & 3). The difference in accumulations between the city-averaged CML and PWS rainfall data (see Fig. 6.2) is partly caused by the overestimation by CML. However, differences may also be due to the spatial variation

of rain and the respective locations of the PWS and CML. Figure 6.6 shows that for the examined hour the CML are more abundant in high-rainfall areas, whereas the PWS are more clustered in the city center, where less rainfall was observed.

The method used to derive rainfall estimates from CML data (RAINLINK) is one of many possible methods (see Sect. 6.3.2). Our dataset consists of instantaneously sampled CML data, which is more prone to errors than CML data obtained with other sampling strategies and/or more frequently than every 15 min (Chapter 4). Furthermore, CML retrieval results also depend on the spatial density of these networks, which could decline in those areas where microwave links are being replaced by fiber optic cables.

We have shown that even though each technique has considerable limitations regarding accuracy, the data from opportunistic sources can be used to monitor meteorological phenomena. The potential of these techniques lies in the high spatial density of such observations, especially in urban areas. Many PWSs are found in densely populated areas, where also many smartphones are operational. This is mainly true for urban areas in parts of the world where people have funds to invest in these devices (although smartphones are considered so important that they are essentially ubiquitous, independent of living standards). CML networks differ as well, in sampling strategy and frequency (which affects the accuracy of rainfall estimates) and in network density (depending on replacement by fiber optic technology). The availability of opportunistic sensing observations should be explored in order to judge their usefulness, especially as their accuracy heavily relies on the quantity of observations. Because traditional meteorological measurements are generally absent in urban areas, these new data provide a welcome addition. This is particularly important for monitoring the effect of cities on temperature (UHI), and wind and rainfall at street level. Careful interpretation and filtering of the data is required because of the highly heterogeneous nature of the effect of the disturbances caused by the urban fabric. But not only urban areas can benefit from these opportunistic data. There are large areas on the globe where meteorological measurements are absent. It is in these areas where these new data sources can provide much-needed information, even if there are large uncertainties involved.

In this paper we have discussed several opportunistic meteorological data sources. Beside the techniques that we have investigated, there are several other promising sources of meteorological data that could be exploited. For example, using cameras as rain gauges (Allamano et al., 2015), rainfall information from sensors in driving cars (Rabiei et al., 2013), deriving the UHI from measurements of gradients of shallow groundwater (Buik et al., 2004), for example using fiber-optic cables (Bense et al., 2016), using airplanes to measure upper-air wind and temperature (De Haan, 2011), using hot-air balloons to measure boundary-layer winds (De Bruijn et al., 2016), using networks of solar panels for radiation monitoring, and many more. We therefore urge the scientific community

to keep investigating new sources of data, and to study the uncertainties therein. In combination with reference networks of meteorological measurements or stand-alone, these new sources will provide much needed complementary information.

Chapter 7

Synthesis

7.1 Aim of the present research

This work addressed the need for dense, accurate, in-situ ground-based rainfall observations, especially in urban areas, which are usually insufficiently provided by traditional rainfall sensors. Opportunistic sensors, such as crowdsourced personal weather station (PWS) observations and rainfall estimates from commercial microwave links (CML), can theoretically bridge this gap. This thesis explores the relative measurement accuracy of these techniques, and determines their potential for operational large-scale rainfall monitoring.

Although gauges provide reliable information on the rainfall amounts at a certain location, their spatial representativeness is limited. Weather radars have a good coverage, but determine rainfall indirectly based on the reflectivity, and in case of dual-pol also the specific differential phase shift, from hydrometeors. Additionally, radar samples an atmospheric volume aloft, which may differ from the actual rainfall amounts at the ground. Rainfall fields that are constructed from information from both radar and rain gauge networks yield improved results. The Royal Netherlands Meteorological Institute (KNMI) provides a so-called climatological gauge-adjusted radar rainfall product, constructed from measurements from two (single-pol, later replaced with dual-pol) C-band doppler radars, a network of manual gauges for daily rainfall amounts (~ 1 gauge per 100 km^2) and an automatic rain gauge network measuring rainfall in 10-min intervals (~ 1 gauge per 1000 km^2). Although highly useful (and used as a ground truth throughout this thesis), this product cannot be constructed in real time, as the manual rain gauge data become available a posteriori and are manually validated before they are combined with the radar rainfall fields. Additionally, these rain gauges are typically installed in rural areas because of siting criteria to limit the influence from surroundings.

In order to determine whether opportunistic sensing can yield useful real-time high-resolution rainfall fields, the accuracy of actual observations is explored, both using measurements and with simulation studies. The following section highlights the results for the Amsterdam metropolitan area and the Netherlands, followed by a section with broader discussion on the general potential of these two techniques and recommendations for future developments.

7.2 Opportunistic sensors: relative accuracy

7.2.1 Intrinsic limitations due to sampling effects

As rainfall observations describe average rainfall over space (atmospheric volume, rain gauge collection funnel, path between receiving and transmitting antennas) and time, radar, PWS and CML will differ even when measuring with perfect accuracy. In order to

know if a technique is accurate enough, one must first determine the spatial and temporal resolution at which it should be able to resolve the rainfall fields. These two aspects are related because rainfall is correlated in space and time, with more variability for convective storms than for stratiform rain events (Fabry, 1996; Berne et al., 2004; Emmanuel et al., 2012).

Chapter 5 disregards sources of error in rainfall retrieval and determines the accuracy of rainfall observations from radar, CML and PWS based on sampling strategy and the sensor lay-outs in the Amsterdam study area. The results show that PWS best captures the small-scale variability of the rainfall. However, when rainfall dynamics need to be resolved at lower resolutions than 100 m and 30 s, the other measurement techniques quickly become more comparable in terms of sampling errors (Fig. 5.11). In the study area, CML and PWS networks are so dense that the accuracy of the resulting rainfall fields are more limited by their temporal sampling than their spatial sampling. Similar accuracies by both opportunistic sensing techniques would be achieved by half the respective networks (Fig. 5.13). Given that 100% data availability is unlikely, a positive consequence of this is that small data gaps (spread randomly in time and space) in CML and PWS networks will not have an immediate big impact in this study area, when measurement errors are not taken into account.

The level of detail of rainfall fields that can be resolved by PWS and CML networks is affected by their spatial lay-out and temporal sampling. Radar accuracy is also related to the choices made with respect to the rainfall retrieval algorithm. Chapter 5 concludes that PWS outperforms CML and radar at the 100-m spatial scale, assuming that raw PWS measurements at ~ 5 min resolution are available. This is the case when crowdsourcing PWS measurements from the Netatmo weather platform. Reporting the number of tipping bucket tips within a certain time interval (e.g. 5 min) introduces an error as some of the collected rain could have fallen outside this interval. This error logically increases with larger tipping bucket volumes and less frequent updates to the platform (see also Fig. 3.5). Another sampling error is introduced when the timestamp of the crowdsourced measurement is not the update moment of that measurement, but of a later call from an intermediate server. This additional temporal sampling effect occurred for Netatmo rain observations obtained via the Wundermap, the platform of Weather Underground (Fig. 2.6), resulting in reduced accuracy for sub-hourly resolutions. Although Netatmo PWSs are no longer automatically linked to the Wunderground platform since an update from Weather Underground in July 2018, it is important to understand the effects of re-attributing timestamps of PWS rainfall measurements on rainfall-retrieval accuracy.

7.2.2 Accuracy of PWS rainfall measurements

In order to describe rainfall fields, unprocessed PWS observations are preferable over measurements affected by processes on the platform, although this effect cancels out above

hourly time scales (Chapter 2). However, even unprocessed PWS observations are notoriously noisy. Raw PWS rainfall observations can suffer from four types of error. Faulty zeroes (FZ) and negative bias likely result from tilted PWSs or otherwise obstructed tipping bucket mechanisms. High influx (HI) errors result from pouring liquids through gauges, handling the device with tilting movements or other erroneous large values reported unrelated to precipitation. Station outlier (SO) errors contain all occurrences when rainfall dynamics reported are not consistent with the measurement location of the PWS, either because the gauge has been moved or because it is significantly affected by elements in its surroundings. Besides the previously mentioned cause for negative bias, under- or overestimation of rainfall by PWSs can be caused by shielded locations and/or inaccurate assumed values for tipping bucket volumes.

Because of the large number of PWSs in the study area of Amsterdam, these errors can be recognized with a quality control (QC) based on neighbor checks. FZ-, HI-, SO-errors and biases can be identified based on comparisons with nearby stations. A real-time QC algorithm is proposed that is successful in improving the accuracy of hourly rainfall observations from a bias of -16.7% to +0.2%, a coefficient of variation of 3.74 to 2.86 and a Pearson correlation coefficient of 0.68 to 0.82, by excluding only 12% of the raw data.

The proposed QC method is successful in the Amsterdam study area, where the density of the PWS network is high, temporal sampling is ~ 5 min, and the assumption that nearby stations should measure similar rainfall dynamics is safe. Adjustments to the QC algorithm are needed in case of sparser networks and also for mountainous areas, where the elevation of a gauge is related to the expected rainfall measured at that location. Some approaches to these issues are proposed in Chapter 3.

7.2.3 Accuracy of CML rainfall measurements

The CML-studies in this thesis have focused on rainfall observations from instantaneously sampling CMLs at 15-min temporal resolution, corresponding with the dataset from the national T-Mobile NL network. Chapter 4 details many factors that affect the accuracy of CML rainfall observations (Sect. 4.3 provides an overview from literature). In the 7-month CML dataset from the Netherlands, some dependencies of measurement errors on link characteristics and observation conditions were found, namely increased errors for links shorter than 2 km, during late night and early morning, and in (winter) months with possible solid precipitation events. However, the sampling strategy seemed to be a more dominant factor, as CML rainfall observations from other networks obtained with min/max sampling in the same climate were more accurate (Chapter 4). Large overestimates of rainfall by CMLs are also found for the event described in Chapter 6, which is especially dire for link paths shorter than 2 km (Fig. 6.5).

The simulation study in Chapter 5 also shows that instantaneous CML observations at 15-min resolution are by far more limited by their temporal sampling than by the spatial

extent of the network in the Amsterdam study area. Based on these findings, this CML network does not outperform the PWS rainfall observations. A more frequent temporal sampling of CML rainfall observations would be a good strategy to increase their accuracy and yield similar applicability of CML for operational rainfall monitoring as the other techniques. Further improvements in the calculation of rainfall from attenuation may also be made, such as a wet antenna attenuation that is related to rainfall intensity instead of assumed constant (e.g. Schleiss et al., 2013). In the current RAINLINK version the wet antenna attenuation is assumed as a constant value.

7.3 Outlook

7.3.1 Expectations for the future

KNMI has replaced their single-pol radars with dual-pol radars in 2016. Information from the combined horizontal and vertical reflectivity will be used for improved radar rainfall estimates in the operational radar quantitative precipitation estimation (QPE) product in the near future. No changes are expected for the automatic rain gauge network, but the manual rain gauge network heavily relies on the willingness and ability of numerous volunteers to keep executing these measurements. With the possibilities available to automatize such observations, KNMI may opt to update this network and make it less dependent on manual actions.

If the manual rain gauge network will be replaced by an automatic one, quality control may be applied to these measurements in an operational manner. This would allow for real-time quality controlled rainfall observations. An adjusted version of the PWS QC proposed in this thesis may be used for the KNMI automatic rain gauge network(s) for this purpose. By extension, the quality controlled rain gauge observations can then be merged with operational radar rainfall estimates, generating a (semi) real-time gauge-adjusted radar rainfall product, which is currently not yet available. Future research should be focused on optimizing quality control for these automatic rain gauge networks, and designing a real-time merging methodology to generate high-resolution rainfall fields from gauges and radar, particularly for urban applications, using existing techniques as a starting point (Ochoa-Rodriguez et al., 2019).

Automatic personal weather stations are popular and their numbers on PWS platforms have grown over the past decade to a sizable and significant network. The ease of linking a PWS to online weather platforms has facilitated the process of sharing weather observations in real time enormously. The network may grow further in time with more people acquiring stations, possibly also in areas of the world where professional networks have been very sparse (e.g. in Africa, South-America and the Middle-East). The vast number of rainfall measurements made accessible by weather platforms has not gone unnoticed by the companies behind these platforms. While PWS owners altruistically share their

data via these platforms, it is important to remember that companies behind the platforms have a commercial purpose. Once shared, the data becomes the property of an organization that is driven by financial incentives and may make drastic changes in data availability without consulting users of said data. One example hereof is the Weather Underground company, that revoked their previously free API access to PWS weather data in May 2018. Free availability of PWS measurements cannot be assumed, and becomes less likely once the value of this data becomes evident.

Commercial microwave links are installed for the purpose of telecommunication, which seems to be a priority in more areas in the world than rainfall monitoring is. Many countries lack dedicated operational rainfall sensors, but do have CML networks. As CML networks can be used for rainfall monitoring, this technique can significantly increase the extend of global rainfall monitoring. Also, because of the high density of CML networks in urban areas, CML rainfall observations can be employed for increased rainfall measurement resolution in cities. However, telecommunication demands are rapidly changing, resulting in changes in backhaul networks. In some areas, fiber optic technology will replace CMLs, and with a move from 4G to 5G, new CMLs operating at other frequency (E-band; 60–90 GHz) ranges will be introduced (Ericsson, 2018), requiring adjusted CML rainfall retrieval methods with expected changes in rainfall estimation accuracy. Prospects are that the global number of CMLs will remain large enough for rainfall monitoring, but changes will vary per location. CMLs that operate on those high frequency ranges need to be relatively short as compared to the current links to achieve stable signal transmission. Therefore, it is expected that CMLs operating at the 7–40 GHz range will continue to exist in the near future. Obtaining CML data in real time remains a large obstacle in employing this technique and requires fruitful collaborations with mobile network operators, possibly requiring financial rewards for making power levels available.

7.3.2 Recommendations

This thesis shows the potential of PWS and CML for operational rainfall monitoring. In the Netherlands, especially the PWS rainfall observations are available at high resolution (in both space and time). The PWS rainfall measurement accuracy is promising after quality control is applied. The CML network in the Netherlands has been declining in density over the past years and the 15-min instantaneous sampling strategy yields poorer rainfall estimation accuracy than other common sampling strategies. In areas where CML data is sampled in a different manner, or where dedicated rainfall sensors are sparse, the added benefit for rainfall monitoring will be larger. Also, the potential of CMLs for the Netherlands should be re-evaluated in case CML infrastructure changes drastically (e.g. in case of nation-wide implementation of 5G).

In the Netherlands, the PWS network is substantial, with an average density of ~ 1 gauge per 10 km^2 , against ~ 1 in 100 km^2 and 1000 km^2 for the operational gauge networks from KNMI. However, using PWS measurements creates a liability when PWS data is owned by

companies with commercial motives. Constructive collaborations between PWS platform companies and end-users are a possible approach. An example of such a collaboration is the Norwegian Meteorological Institute that makes use of Netatmo PWS measurements (Norwegian Meteorological Institute, 2018).

CML and PWS are valuable techniques to provide weather information in urban areas, with far better spatial resolution than other opportunistic sensing techniques can offer (Chapter 6). While opportunistic sensors can definitely provide high-resolution rainfall observations, they should not be seen as a replacement of traditional techniques. This is partly because it is unwise to rely on external organizations for something as essential for public safety and well-being as rainfall monitoring, especially organizations whose market interests differ vastly from those of meteorological institutes. Additionally, climatological studies need long time series of observations of high quality that are measured under constant conditions. PWS and CML are too recent, contain too many errors, and are too variable in time to provide that.

Traditional dedicated rainfall measurement techniques provide a comfortable starting point, which may be extended upon with opportunistic sensing. The combined high accuracy of the traditional measurements with the high-quantity-low-quality opportunistic sensor observations can yield real-time rainfall fields at higher resolution and with higher accuracy than either one of them would have achieved individually. Future work should therefore be focused on the combination of these data sources. Several studies have explored the merging of gauges, links and/or radar (e.g. Bianchi et al., 2013; Scheidegger & Rieckermann, 2014; Liberman et al., 2014; Trömel et al., 2014; Haese et al., 2017; Fencil et al., 2017; Yang & Ng, 2019), but much work is yet to be done, especially concerning merging in an operational setting. Openly sharing data and tools will facilitate this process.

References

- Alberoni, P. P., Fornasiero, A., Roversi, G., Pasetti, S., Folegani, M., & Porcù, F. (2018). Comparison between different QPE based on: Microwave links, radar adjusted and gauges. (pp. 851–860). doi:10.18174/454537 extended abstract for ERAD2018 conference.
- Allamano, P., Croci, A., & Laio, F. (2015). Toward the camera rain gauge. *Water Resources Research*, 51, 1744–1757.
- Andersson, J. C. M., Berg, P., Hansryd, J., Jacobsson, A., Olsson, J., & Wallin, J. (2017). Microwave links improve operational rainfall monitoring in Gothenburg Sweden. In *Proceedings of the 1st International Conference on Environmental Science and Technology* (pp. 1–4).
- Android (2019). Android developer guide for SensorManagement. <http://developer.android.com/reference/android/hardware/SensorManager.html>. Last accessed 2019-01-29, last updated 2019-01-23.
- Andsager, K., Beard, K. V., & Laird, N. F. (1999). Laboratory measurements of axis ratios for large raindrops. *Journal of the Atmospheric Sciences*, 56, 2673–2683.
- Atlas, D., & Ulbrich, C. W. (1977). Path-and area-integrated rainfall measurement by microwave attenuation in the 1–3 cm band. *Journal of Applied Meteorology*, 16, 1322–1331.
- Baer, J. (1976). Rain gauge. US Patent 3,943,762.
- Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Bousserez, N., Boussetta, S., Brown, A., Buizza, R. et al. (2018). Satellite and in situ observations for advancing global earth surface modelling: A review. *Remote Sensing*, 10, 1–72.
- Basara, J. B., Illston, B. G., Fiebrich, C. A., Browder, P. D., Morgan, C. R., McCombs, A., Bostic, J. P., McPherson, R. A., Schroeder, A. J., & Crawford, K. C. (2011). The Oklahoma City Micronet. *Meteorological applications*, 18, 252–261.
- Beard, K. V., & Chuang, C. (1987). A new model for the equilibrium shape of raindrops. *Journal of the Atmospheric Sciences*, 44, 1509–1524.

- Beekhuis, H., & Mathijssen, T. (2018). From pulse to product, Highlights of the upgrade project of the Dutch national weather radar network. (pp. 960–965). doi:10.18174/454537 extended abstract for ERAD2018 conference.
- Bell, S., Cornford, D., & Bastin, L. (2013). The state of automated amateur weather observations. *Weather*, 68, 36–41.
- Bell, S., Cornford, D., & Bastin, L. (2015). How good are citizen weather stations? Addressing a biased opinion. *Weather*, 70, 75.
- Bell, V. A., & Moore, R. J. (2000). The sensitivity of catchment runoff models to rainfall data at different spatial scales. *Hydrology and Earth System Sciences*, 4, 653–667.
- Bense, V. F., Read, T., Bour, O., Le Borgne, T., Coleman, T., Krause, S., Chalari, A., Mondanos, M., Ciocca, F., & Selker, J. S. (2016). Distributed temperature sensing as a downhole tool in hydrogeology. *Water Resources Research*, 52, 9259–9273. doi:10.1002/2016WR018869.
- Berne, A., Delrieu, G., Creutin, J.-D., & Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*, 299, 166–179.
- Berne, A., & Uijlenhoet, R. (2007). Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability. *Geophysical Research Letters*, 34. doi:10.1029/2007GL029409.
- Bianchi, B., Jan van Leeuwen, P., Hogan, R. J., & Berne, A. (2013). A variational approach to retrieve rain rate by combining information from rain gauges, radars, and microwave links. *Journal of Hydrometeorology*, 14, 1897–1909.
- Brandes, E. A., Zhang, G., & Vivekanandan, J. (2002). Experiments in rainfall estimation with a polarimetric radar in a subtropical environment. *Journal of Applied Meteorology*, 41, 674–685.
- Bringi, V. N., & Chandrasekar, V. (2001). *Polarimetric Doppler weather radar: principles and applications*. Cambridge university press.
- Bringi, V. N., Huang, G.-J., Chandrasekar, V., & Gorgucci, E. (2002). A methodology for estimating the parameters of a gamma raindrop size distribution model from polarimetric radar data: Application to a squall-line event from the TRMM/Brazil campaign. *Journal of Atmospheric and Oceanic Technology*, 19, 633–645.
- Bruni, G., Reinoso, R., Van de Giesen, N. C., Clemens, F. H. L. R., & Ten Veldhuis, J. A. E. (2015). On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrology and Earth System Sciences*, 19, 691–709.
- Buik, N., Stolk, P., & Willemsen, G. (2004). Analyse van temperatuurmetingen in de Nederlandse ondergrond. *Stromingen*, 10, 13–22. In Dutch.
- Buishand, T. A., & Wijngaard, J. (2007). *Statistiek van extreme neerslag voor korte neerslagduren [Statistics of extreme rainfall for short durations]*. Royal Netherlands Meteorologic Institute.

- Chapman, L., Bell, C., & Bell, S. (2017). Can the crowdsourcing data paradigm take atmospheric science to a new level? A case study of the urban heat island of London quantified using Netatmo weather stations. *International Journal of Climatology*, *37*, 3597–3605.
- Chau, N. H. (2019). Estimation of air temperature using smartphones in different contexts. *Journal of Information and Telecommunication*, (pp. 1–14).
- Chen, A. B., Behl, M., & Goodall, J. L. (2018). Trust me, my neighbors say it's raining outside: Ensuring data trustworthiness for crowdsourced weather stations. In *Proceedings of the 5th Conference on Systems for Built Environments* (pp. 25–28). ACM.
- Chen, H., & Chandrasekar, V. (2015). The quantitative precipitation estimation system for Dallas–Fort Worth (DFW) urban remote sensing network. *Journal of Hydrology*, *531*, 259–271.
- Chua, S. (2009). Light vs. distance. University Lecture, Powerpoint Slides. http://bccp.berkeley.edu/o/Academy/workshop_09/pdfs/InverseSquareLawPresentation.pdf, accessed on 31 January 2019.
- Chwala, C., Gmeiner, A., Qiu, W., Hipp, S., Nienaber, D., Siart, U., Eibert, T., Pohl, M., Seltmann, J., Fritz, J. et al. (2012). Precipitation observation using microwave backhaul links in the alpine and pre-alpine region of Southern Germany. *Hydrology and Earth System Sciences*, *16*, 2647–2661. doi:10.5194/hess-16-2647-2012.
- Chwala, C., Keis, F., & Kunstmann, H. (2016). Real-time data acquisition of commercial microwave link networks for hydrometeorological applications. *Atmospheric Measurement Techniques*, *9*, 991–999.
- Chwala, C., & Kunstmann, H. (2019). Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges. *Wiley Interdisciplinary Reviews: Water*, *6*, e1337.
- Chwala, C., Smiatek, G., & Kunstmann, H. (2018). Real-time rainfall from a country-wide network of commercial microwave links in Germany. (pp. 265–265). doi:10.18174/454537 abstract for ERAD2018 conference.
- Ciach, G. J., & Krajewski, W. F. (2006). Analysis and modeling of spatial correlation structure in small-scale rainfall in central oklahoma. *Advances in Water Resources*, *29*, 1450–1463.
- Cifelli, R., Doesken, N., Kennedy, P., Carey, L. D., Rutledge, S. A., Gimmestad, C., & Depue, T. (2005). The community collaborative rain, hail, and snow network: Informal education for scientists and citizens. *Bulletin of the American Meteorological Society*, *86*, 1069–1078.
- Cornes, R. C., Dirksen, M., & Sluiter, R. (2019). Correcting citizen-science air temperature measurements across the netherlands for short wave radiation bias. *Meteorological Applications*, (pp. 1–16). doi:0.1002/met.1814.
- Cristiano, E., ten Veldhuis, M.-C., & Van de Giesen, N. (2017). Spatial and temporal

- variability of rainfall and their effects on hydrological response in urban areas—a review. *Hydrology and Earth System Sciences*, 21, 3859–3878.
- Daly, C., Gibson, W. P., Taylor, G. H., Doggett, M. K., & Smith, J. I. (2007). Observer bias in daily precipitation measurements at united states cooperative network stations. *Bulletin of the American Meteorological Society*, 88, 899–912.
- De Bruijn, E. I. F., De Haan, S., Bosveld, F. C., Schreur, B. W., & Holtslag, A. A. M. (2016). Observing boundary-layer winds from hot-air balloon flights. *Weather and Forecasting*, 31, 1451–1463. doi:10.1175/WAF-D-16-0028.1.
- De Haan, S. (2011). High-resolution wind and temperature observations from aircraft tracked by mode-s air traffic control radar. *Journal of Geophysical Research: Atmospheres*, 116, 1–13. doi:10.1029/2010JD015264.
- De Haij, M. (2009). *Automatische validatie van druk- en windwaarnemingen op het KNMI - een verkenning*. Technical Report, . URL: <http://bibliotheek.knmi.nl/knmipubIR/IR2009-03.pdf>.
- De Vasconcelos, L. E. G., dos Santos, E. C. M., Neto, M. L. F., Ferreira, N. J., & De Vasconcelos, L. G. (2016). Using tweets for rainfall monitoring. In *Information Technology: New Generations* (pp. 1157–1167). Springer International Publishing, Cham, Switzerland volume 448. doi:10.1007/978-3-319-32467-8_100.
- Diggle, P. J., & Ribeiro, P. J. (2007). *Model-based Geostatistics*. New York, USA: Springer Science + Business Media.
- Doumounia, A., Gosset, M., Cazenave, F., Kacou, M., & Zougmore, F. (2014). Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a West African test bed. *Geophysical Research Letters*, 41, 6016–6022. doi:10.1002/2014GL060724.
- Drew, D. R., Barlow, J. F., & Lane, S. E. (2013). Observations of wind speed profiles over Greater London, UK, using a Doppler lidar. *Journal of Wind Engineering and Industrial Aerodynamics*, 121, 98–105. doi:10.1016/j.jweia.2013.07.019.
- Droste, A. M., Heusinkveld, B. G., Fenner, D., & Steeneveld, G. J. (2018). Crowdsourcing the urban wind. In *EGU General Assembly Conference Abstracts* (p. 5671). volume 21.
- Droste, A. M., Pape, J. J., Overeem, A., Leijnse, H., Steeneveld, G. J., Van Delden, A. J., & Uijlenhoet, R. (2017). Crowdsourcing urban air temperatures through smartphone battery temperatures in São Paulo, Brazil. *Journal of Atmospheric and Oceanic Technology*, 34, 1853–1866.
- Ducrocq, V., Braud, I., Davolio, S., Ferretti, R., Flamant, C., Jansa, A., Kalthoff, N., Richard, E., Taupier-Letage, I., Ayrat, P.-A. et al. (2014). HyMeX-SOP1: The field campaign dedicated to heavy precipitation and flash flooding in the northwestern Mediterranean. *Bulletin of the American Meteorological Society*, 95, 1083–1100.
- Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.-E., Quirnbach, M., Vaes, G., & Vieux, B. (2004). Towards a roadmap for use of radar rainfall data in urban drainage.

- Journal of Hydrology*, 299, 186–202.
- Elmore, K. L., Flamig, Z., Lakshmanan, V., Kaney, B., Farmer, V., Reeves, H. D., & Rothfus, L. P. (2014). MPING: Crowd-sourcing weather reports for research. *Bulletin of the American Meteorological Society*, 95, 1335–1342.
- Emmanuel, I., Andrieu, H., Leblois, E., & Flahaut, B. (2012). Temporal and spatial variability of rainfall at the urban hydrological scale. *Journal of Hydrology*, 430, 162–172.
- Ericsson (2018). Ericsson microwave outlook. <https://www.ericsson.com/assets/local/microwave-outlook/documents/ericsson-microwave-outlook-report-2018.pdf>, Accessed March 14, 2019.
- Estévez, J., Gavilán, P., & Giráldez, J. V. (2011). Guidelines on validation procedures for meteorological data from automatic weather stations. *Journal of Hydrology*, 402, 144–154.
- Fabian, N. O. (1966). Pluviometer. US Patent 3,229,519.
- Fabry, F. (1996). On the determination of scale ranges for precipitation fields. *Journal of Geophysical Research: Atmospheres*, 101, 12819–12826.
- Fabry, F., Bellon, A., Duncan, M. R., & Austin, G. L. (1994). High resolution rainfall measurements by radar for very small basins: The sampling problem reexamined. *Journal of Hydrology*, 161, 415–428.
- Fencl, M., Dohnal, M., Rieckermann, J., & Bares, V. (2017). Gauge-adjusted rainfall estimates from commercial microwave links. *Hydrology and Earth System Sciences*, 21, 617–634.
- Fenicia, F., Pfister, L., Kavetski, D., Matgen, P., Iffly, J.-F., Hoffmann, L., & Uijlenhoet, R. (2012). Microwave links for rainfall estimation in an urban environment: Insights from an experimental setup in Luxembourg-City. *Journal of Hydrology*, 464, 69–78.
- Fenner, D., Meier, F., Bechtel, B., Otto, M., & Scherer, D. (2017). Intra and inter local climate zone variability of air temperature as observed by crowdsourced citizen weather stations in Berlin, Germany. *Meteorologische Zeitschrift*, 26, 525–547.
- Figueras i Ventura, J., Boumahmoud, A.-A., Fradon, B., Dupuy, P., & Tabary, P. (2012). Long-term monitoring of French polarimetric radar data quality and evaluation of several polarimetric quantitative precipitation estimators in ideal conditions for operational implementation at C-band. *Quarterly Journal of the Royal Meteorological Society*, 138, 2212–2228.
- Ghariesifard, M., & Wehn, U. (2016). To share or not to share: Drivers and barriers for sharing data via online amateur weather networks. *Journal of Hydrology*, 535, 181–190.
- Gires, A., Onof, C., Maksimovic, C., Schertzer, D., Tchiguirinskaia, I., & Simoes, N. (2012). Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study. *Journal of Hydrology*, 442, 117–128.

- Goldshtein, O., Messer, H., & Zinevich, A. (2009). Rain rate estimation using measurements from commercial telecommunications links. *IEEE Transactions on Signal Processing*, 57, 1616–1625. doi:10.1109/TSP.2009.2012554.
- Golroudbary, V. R., Zeng, Y., Mannaerts, C. M., & Su, Z. B. (2018). Urban impacts on air temperature and precipitation over the Netherlands. *Climate Research*, 75, 95–109.
- Goudenhoofdt, E., & Delobbe, L. (2009). Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrology and Earth System Sciences*, 13, 195–203.
- GSMA (2012). How mobile can bring communications to all. <https://www.gsma.com/publicpolicy/wp-content/uploads/2012/03/universalaccessfullreport.pdf>, Accessed November 13, 2017.
- GSMA (2016). mAgri Weather forecasting and monitoring: Mobile solutions for climate resilience. https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2016/02/GSMA_mAgri_Weather_forecasting_and_monitoring_mobile_solutions_for_climate_resilience.pdf, Accessed December 19, 2017.
- Guo, H., Huang, H., Sun, Y.-E., Zhang, Y., Chen, S., & Huang, L. (2019). Chaac: Real-time and fine-grained rain detection and measurement using smartphones. *IEEE Internet of Things Journal*, 6, 997–1009.
- Habib, E., Krajewski, W. F., & Ciach, G. J. (2001a). Estimation of rainfall interstation correlation. *Journal of Hydrometeorology*, 2, 621–629.
- Habib, E., Krajewski, W. F., & Kruger, A. (2001b). Sampling errors of tipping-bucket rain gauge measurements. *Journal of Hydrologic Engineering*, 6, 159–166.
- Haese, B., Hörning, S., Chwala, C., Bárdossy, A., Schälge, B., & Kunstmann, H. (2017). Stochastic reconstruction and interpolation of precipitation fields using combined information of commercial microwave links and rain gauges. *Water Resources Research*, 53, 10740–10756. doi:10.1002/2017WR021015.
- Haynes, J. M., L’Ecuyer, T. S., Stephens, G. L., Miller, S. D., Mitrescu, C., Wood, N. B., & Tanelli, S. (2009). Rainfall retrieval over the ocean with spaceborne w-band radar. *Journal of Geophysical Research: Atmospheres*, 114.
- Illingworth, A., & Stevens, C. (1987). An optical disdrometer for the measurement of rain-drop size spectra in windy conditions. *Journal of Atmospheric and Oceanic Technology*, 4, 411–421.
- Illingworth, S. M., Muller, C. L., Graves, R., & Chapman, L. (2014). UK citizen rainfall network: a pilot study. *Weather*, 69, 203–207.
- IPCC (2012). *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and

- New York, NY, USA, 582 pp.
- Jameson, A. R. (1991). A comparison of microwave techniques for measuring rainfall. *Journal of Applied Meteorology*, 30, 32–54.
- Jarraud, M. (2008). *Guide to meteorological instruments and methods of observation (WMO-No. 8)*. World Meteorological Organization: Geneva, Switzerland.
- Jenkins, G. (2014). A comparison between two types of widely used weather stations. *Weather*, 69, 105–110.
- Jiang, S., Babovic, V., Zheng, Y., & Xiong, J. (2019). Advancing opportunistic sensing in hydrology: A novel approach to measuring rainfall with ordinary surveillance cameras. *Water Resources Research*, 55. doi:10.1029/2018WR024480.
- Joss, J., & Waldvogel, A. (1967). A raindrop spectrograph with automatic analysis. *Pure Applied Geophysics*, 68, 240–246.
- Kalogiros, J., Anagnostou, M. N., Anagnostou, E. N., Montopoli, M., Picciotti, E., & Marzano, F. S. (2013). Optimum estimation of rain microphysical parameters from X-band dual-polarization radar observables. *IEEE Transactions on Geoscience and Remote Sensing*, 51, 3063–3076.
- Kaufmann, M., & Rieckermann, J. (2011). Identification of dry and rainy periods using telecommunication microwave links. In *12nd International Conference on Urban Drainage, Porto Alegre/Brazil* (pp. 10–15). doi:10.1109/LGRS.2010.2043052.
- Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, how much of the Earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, 98, 69–78. doi:10.1175/BAMS-D-14-00283.1.
- Kidd, C., & Huffman, G. (2011). Global precipitation measurement. *Meteorological Applications*, 18, 334–353. doi:10.1002/met.284.
- Kim, M.-S., & Kwon, B. (2018). Rainfall detection and rainfall rate estimation using microwave attenuation. *Atmosphere*, 9, 287.
- Kitchen, M., & Blackall, R. M. (1992). Representativeness errors in comparisons between radar and gauge measurements of rainfall. *Journal of Hydrology*, 134, 13–33.
- KNMI (2000). *Handbook for the Meteorological Observation*. Koninklijk Nederlands Meteorologisch Instituut. Available via http://projects.knmi.nl/hawa/pdf/Handbook_H01_H06.pdf.
- Krajewski, W. F., Ciach, G. J., & Habib, E. (2003). An analysis of small-scale rainfall variability in different climatic regimes. *Hydrological Sciences Journal*, 48, 151–162.
- Kroeger, T., Timofte, R., Dai, D., & Van Gool, L. (2016). Fast optical flow using dense inverse search. In *European Conference on Computer Vision* (pp. 471–488). Springer.
- Kruger, A., & Krajewski, W. F. (2002). Two-dimensional video disdrometer: A description. *Journal of Atmospheric and Oceanic Technology*, 19, 602–617.

- Lanza, L. G., & Vuerich, E. (2009). The wmo field intercomparison of rain intensity gauges. *Atmospheric Research*, *94*, 534–543.
- Laws, J. O., & Parsons, D. A. (1943). The relation of raindrop-size to intensity. *Eos, Transactions American Geophysical Union*, *24*, 452–460. doi:10.1029/TR024i002p00452.
- Leijnse, H., Uijlenhoet, R., Van de Beek, C. Z., Overeem, A., Otto, T., Unal, C. M. H., Dufournet, Y., Russchenberg, H. W. J., Figueras i Ventura, J., Klein Baltink, H. et al. (2010a). Precipitation measurement at CESAR, the Netherlands. *Journal of Hydrometeorology*, *11*, 1322–1329.
- Leijnse, H., Uijlenhoet, R., & Berne, A. (2010b). Errors and uncertainties in microwave link rainfall estimation explored using drop size measurements and high-resolution radar data. *Journal of Hydrometeorology*, *11*, 1330–1344. doi:10.1175/2010JHM1243.1.
- Leijnse, H., Uijlenhoet, R., & Stricker, J. N. M. (2007a). Hydrometeorological application of a microwave link: 2. Precipitation. *Water Resources Research*, *43*.
- Leijnse, H., Uijlenhoet, R., & Stricker, J. N. M. (2007b). Rainfall measurement using radio links from cellular communication networks. *Water Resources Research*, *43*. doi:10.1029/2006WR005631.
- Leijnse, H., Uijlenhoet, R., & Stricker, J. N. M. (2008). Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation. *Advances in Water Resources*, *31*, 1481–1493. doi:10.1016/j.advwatres.2008.03.004.
- Liberman, Y., Samuels, R., Alpert, P., & Messer, H. (2014). New algorithm for integration between wireless microwave sensor network and radar for improved rainfall measurement and mapping. *Atmospheric Measurement Techniques*, *7*, 3549–3563.
- Liguori, S., Rico-Ramirez, M. A., Schellart, A. N. A., & Saul, A. J. (2012). Using probabilistic radar rainfall nowcasts and NWP forecasts for flow prediction in urban catchments. *Atmospheric Research*, *103*, 80–95.
- Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., & Loumagne, C. (2014). When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrology and Earth System Sciences*, *18*, 575–594.
- Macdonald, R. W. (2000). Modelling the mean velocity profile in the urban canopy layer. *Boundary-Layer Meteorology*, *97*, 25–45. doi:10.1023/A:1002785830512.
- Madaus, L. E., & Mass, C. F. (2017). Evaluating smartphone pressure observations for mesoscale analyses and forecasts. *Weather and Forecasting*, *32*, 511–531. doi:10.1175/WAF-D-16-0135.1.
- Madsen, H., Lawrence, D., Lang, M., Martinkova, M., & Kjeldsen, T. R. (2014). Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *Journal of Hydrology*, *519*, 3634–3650.
- Mandement, M., & Caumont, O. (2019). Contribution of personal weather stations to the

- observation of deep-convection features near the ground. *Natural Hazards and Earth System Sciences Discussions*, (pp. 1–35). doi:10.5194/nhess-2019-229.
- Marshall, J. S., Hitschfeld, W., & Gunn, K. L. S. (1955). Advances in radar weather. In *Advances in Geophysics* (pp. 1–56). Academic Press Inc., New York volume 2.
- Marshall, J. S., & Palmer, W. M. K. (1948). The distribution of rain-drops with size. *Journal of Meteorology*, 5, 165–166. doi:10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2.
- Mass, C. F., & Madaus, L. E. (2014). Surface pressure observations from smartphones: A potential revolution for high-resolution weather prediction? *Bulletin of the American Meteorological Society*, 95, 1343–1349. doi:10.1175/BAMS-D-13-00188.1.
- Mazzoleni, M., Verlaan, M., Alfonso, L., Monego, M., Norbiato, D., Ferri, M., & Solomatine, D. P. (2017). Can assimilation of crowdsourced data in hydrological modelling improve flood prediction? *Hydrology and Earth System Sciences*, 21, 839–861.
- McCabe, M. F., Rodell, M., Alsdorf, D. E., Gonzalez Miralles, D., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N., Franz, T. E. et al. (2017). The future of earth observation in hydrology. *Hydrology and Earth System Sciences*, 21, 3879–3914.
- Meier, F., Fenner, D., Grassmann, T., Jänicke, B., Otto, M., & Scherer, D. (2015). Challenges and benefits from crowd-sourced atmospheric data for urban climate research using Berlin, Germany, as testbed. In *ICUC9 - 9th International Conference on Urban Climate jointly with 12th Symposium on the Urban Environment*.
- Meier, F., Fenner, D., Grassmann, T., Otto, M., & Scherer, D. (2017). Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Climate*, 19, 170–191.
- Messer, H., & Sendik, O. (2015). A new approach to precipitation monitoring: A critical survey of existing technologies and challenges. *IEEE Signal Processing Magazine*, 32, 110–122. doi:10.1109/MSP.2014.2309705.
- Messer, H., Zinevich, A., & Alpert, P. (2006). Environmental monitoring by wireless communication networks. *Science*, 312, 713–713. doi:10.1126/science.1120034.
- Muller, C. L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., & Leigh, R. R. (2015). Crowdsourcing for climate and atmospheric sciences: current status and future potential. *International Journal of Climatology*, 35, 3185–3203.
- Napoly, A., Meier, F., Grassmann, T., & Fenner, D. (2018). Development and application of a statistically-based quality control for crowdsourced air temperature data. *Frontiers in Earth Science*, 6.
- Niforatos, E., Vourvopoulos, A., & Langheinrich, A. (2017). Understanding the potential of human-machine crowdsourcing for weather data. *International Journal of Human-Computer Studies*, 102, 54–68. doi:10.1016/j.ijhcs.2016.10.002. Special Issue on Mobile and Situated Crowdsourcing.
- Niforatos, E., Vourvopoulos, A., Langheinrich, M., Campos, P., & Doria, A. (2014).

- Atmos: a hybrid crowdsourcing approach to weather estimation. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (pp. 135–138). ACM. doi:10.1145/2638728.2638780.
- Nissen, K. M., & Ulbrich, U. (2017). Increasing frequencies and changing characteristics of heavy precipitation events threatening infrastructure in Europe under climate change. *Natural Hazards and Earth System Sciences*, 17, 1177–1190.
- Norwegian Meteorological Institute (2018). Private weather observations improve temperature forecasts on Yr. <https://www.met.no/en/archive/private-weather-observations-improve-temperature-forecasts-on-yr>. Last accessed 2019-05-15.
- Nystuen, J. A., Proni, J. R., Black, P. G., & Wilkerson, J. C. (1996). A comparison of automatic rain gauges. *Journal of Atmospheric and Oceanic Technology*, 13, 62–73.
- Ochoa-Rodriguez, S., Wang, L.-P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., van Assel, J. et al. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology*, 531, 389–407.
- Ochoa-Rodriguez, S., Wang, L.-P., Willems, P., & Onof, C. (2019). A review of radar-rain gauge data merging methods and their potential for urban hydrological applications. *Water Resources Research*, . doi:10.1029/2018WR023332. In press.
- Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, 108, 1–24. doi:10.1002/qj.49710845502.
- Oke, T. R. (2006). Initial guidance to obtain representative meteorological observations at urban sites. *World Meteorological Organization, Geneva, ION report 81*.
- Olsen, R. L., Rogers, D. V., & Hodge, D. (1978). The aR^b relation in the calculation of rain attenuation. *IEEE Transactions on Antennas and Propagation*, 26, 318–329. doi:10.1109/TAP.1978.1141845.
- Ostrometzky, J., Eshel, A., Alpert, P., & Messer, H. (2017). Induced bias in attenuation measurements taken from commercial microwave links. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 3744–3748). IEEE. doi:10.1109/ICASSP.2017.7952856.
- Overeem, A., Buishand, T., & Holleman, I. (2009a). Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar. *Water Resources Research*, 45. doi:10.1029/2009WR007869.
- Overeem, A., Holleman, I., & Buishand, A. (2009b). Derivation of a 10-year radar-based climatology of rainfall. *Journal of Applied Meteorology and Climatology*, 48, 1448–1463. doi:10.1175/2009JAMC1954.1.
- Overeem, A., Leijnse, H., & Uijlenhoet, R. (2011). Measuring urban rainfall using microwave links from commercial cellular communication networks. *Water Resources Research*, 47.

- Overeem, A., Leijnse, H., & Uijlenhoet, R. (2013a). Country-wide rainfall maps from cellular communication networks. *Proceedings of the National Academy of Sciences*, *110*, 2741–2745.
- Overeem, A., Leijnse, H., & Uijlenhoet, R. (2016a). Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network. *Atmospheric Measurement Techniques*, *9*, 2425–2444. doi:10.5194/amt-9-2425-2016.
- Overeem, A., Leijnse, H., & Uijlenhoet, R. (2016b). Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks. *Water Resources Research*, *52*, 8039–8065. doi:10.1002/2016WR019412.
- Overeem, A., Robinson, R. J. C., Leijnse, H., Steeneveld, G. J., Horn, P. B. K., & Uijlenhoet, R. (2013b). Crowdsourcing urban air temperatures from smartphone battery temperatures. *Geophysical Research Letters*, *40*, 4081–4085.
- Paulson, K., & Al-Mreri, A. (2011). A rain height model to predict fading due to wet snow on terrestrial links. *Radio Science*, *46*. doi:10.1029/2010RS004555.
- Peleg, N., Ben-Asher, M., & Morin, E. (2013). Radar subpixel-scale rainfall variability and uncertainty: lessons learned from observations of a dense rain-gauge network. *Hydrology and Earth System Sciences*, *17*, 2195–2208.
- Plank, V. G., Crist, S. D., & LaGross, D. L. (1984). Rain rate meter. US Patent 4,476,718.
- Pollock, M. D., O'Donnell, G., Quinn, P., Dutton, M., Black, A., Wilkinson, M. E., Colli, M., Stagnaro, M., Lanza, L., Lewis, E. et al. (2018). Quantifying and mitigating wind-induced undercatch in rainfall measurements. *Water Resources Research*, *54*, 3863–3875.
- Qi, Y., Martinaitis, S., Zhang, J., & Cocks, S. (2016). A real-time automated quality control of hourly rain gauge data based on multiple sensors in MRMS system. *Journal of Hydrometeorology*, *17*, 1675–1691.
- Rabiei, E., Haberlandt, U., Sester, M., & Fitzner, D. (2013). Rainfall estimation using moving cars as rain gauges-laboratory experiments. *Hydrology and Earth System Sciences*, *17*, 4701–4712.
- Raghavan, S. (2013). *Radar Meteorology* volume 27. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Rahimi, A., Holt, A., Upton, G., & Cummings, R. (2003). Use of dual-frequency microwave links for measuring path-averaged rainfall. *Journal of Geophysical Research: Atmospheres*, *108*. doi:10.1029/2002JD003202.
- Raupach, T. H., & Berne, A. (2016). Small-scale variability of the raindrop size distribution and its effect on areal rainfall retrieval. *Journal of Hydrometeorology*, *17*, 2077–2104.
- Raupach, T. H., & Berne, A. (2017). Retrieval of the raindrop size distribution from polarimetric radar data using double-moment normalisation. *Atmospheric Measurement Techniques*, *10*, 2573–2594.

- Rayitsfeld, A., Samuels, R., Zinevich, A., Hadar, U., & Alpert, P. (2012). Comparison of two methodologies for long term rainfall monitoring using a commercial microwave communication system. *Atmospheric Research*, 104, 119–127. doi:10.1016/j.atmosres.2011.08.011.
- Reges, H. W., Doesken, N., Turner, J., Newman, N., Bergantino, A., & Schwalbe, Z. (2016). COCORAHs: The evolution and accomplishments of a volunteer rain gauge network. *Bulletin of the American Meteorological Society*, 97, 1831–1846.
- Rios Gaona, M. F., Overeem, A., Brasjen, A. M., Meirink, J. F., Leijnse, H., & Uijlenhoet, R. (2017). Evaluation of rainfall products derived from satellites and microwave links for the Netherlands. *IEEE Transactions on Geoscience and Remote Sensing*, 55, 6849–6859. doi:10.1109/TGRS.2017.2735439.
- Rios Gaona, M. F., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2015). Measurement and interpolation uncertainties in rainfall maps from cellular communication networks. *Hydrology and Earth System Sciences*, 19, 3571–3584. doi:10.5194/hess-19-3571-2015.
- Rios Gaona, M. F., Overeem, A., Raupach, T. H., Leijnse, H., & Uijlenhoet, R. (2018). Rainfall retrieval with commercial microwave links in São Paulo, Brazil. *Atmospheric Measurement Techniques*, 11, 4465–4476.
- Roebeling, R. A., & Holleman, I. (2009). Sevir rainfall retrieval and validation using weather radar observations. *Journal of Geophysical Research: Atmospheres*, 114. doi:10.1029/2009JD012102.
- Ronda, R. J., Steeneveld, G. J., Heusinkveld, B. G., Attema, J., & Holtslag, A. A. M. (2017). Urban finescale forecasting reveals weather conditions with unprecedented detail. *Bulletin of the American Meteorological Society*, 98, 2675–2688. doi:10.1175/BAMS-D-16-0297.1.
- Roy, V., Gishkori, S., & Leus, G. (2016). Dynamic rainfall monitoring using microwave links. *EURASIP Journal on Advances in Signal Processing*, 2016, 77.
- Ruf, C. S., Aydin, K., Mathur, S., & Bobak, J. P. (1996). 35-ghz dual-polarization propagation link for rain-rate estimation. *Journal of Atmospheric and Oceanic Technology*, 13, 419–425.
- Scheidegger, A., & Rieckermann, J. (2014). Bayesian assimilation of rainfall sensors with fundamentally different integration characteristics. *Paper presented at International Symposium Weather Radar and Hydrology (WRaH) Washington, DC*, (pp. 1–10).
- Schilling, W. (1991). Rainfall data for urban hydrology: what do we need? *Atmospheric Research*, 27, 5–21.
- Van het Schip, T. I., Overeem, A., Leijnse, H., Uijlenhoet, R., Meirink, J. F., & van Delden, A. J. (2017). Rainfall measurement using cell phone links: classification of wet and dry periods using geostationary satellites. *Hydrological Sciences Journal*, 62, 1343–1353. doi:10.1080/02626667.2017.1329588.
- Schleiss, M., & Berne, A. (2010). Identification of dry and rainy periods using telecommu-

- nication microwave links. *IEEE Geoscience and Remote Sensing Letters*, 7, 611–615. doi:10.1109/LGRS.2010.2043052.
- Schleiss, M., Rieckermann, J., & Berne, A. (2013). Quantification and modeling of wet-antenna attenuation for commercial microwave links. *IEEE Geoscience and Remote Sensing Letters*, 10, 1195–1199.
- Schuurmans, J., Bierkens, M., Pebesma, E., & Uijlenhoet, R. (2007). Automatic prediction of high-resolution daily rainfall fields for multiple extents: the potential of operational radar. *Journal of Hydrometeorology*, 8, 1204–1224.
- Seliga, T. A., & Bringi, V. N. (1976). Potential use of radar differential reflectivity measurements at orthogonal polarizations for measuring precipitation. *Journal of Applied Meteorology*, 15, 69–76.
- Skofronick-Jackson, G., Kirschbaum, D., Petersen, W., Huffman, G., Kidd, C., Stocker, E., & Kakar, R. (2018). The global precipitation measurement (gpm) mission's scientific achievements and societal contributions: reviewing four years of advanced rain and snow observations. *Quarterly Journal of the Royal Meteorological Society*, 144, 27–48.
- Smith, B., Smith, J. A., Baeck, M., Villarini, G., & Wright, D. (2013). Spectrum of storm event hydrologic response in urban watersheds. *Water Resources Research*, 49, 2649–2663.
- Smith, J. A., Baeck, M. L., Morrison, J. E., Sturdevant-Rees, P., Turner-Gillespie, D. F., & Bates, P. D. (2002). The regional hydrology of extreme floods in an urbanizing drainage basin. *Journal of Hydrometeorology*, 3, 267–282.
- Sohail Afzal, M., Shah, S. H. H., Cheema, M. J. M., & Ahmad, R. (2018). Real time rainfall estimation using microwave signals of cellular communication networks: a case study of Faisalabad, Pakistan. *Hydrology and Earth System Sciences Discussions*, (pp. 1–20). doi:10.5194/hess-2017-740. In review.
- Spencer, R. W., Goodman, H. M., & Hood, R. E. (1989). Precipitation retrieval over land and ocean with the SSM/I: Identification and characteristics of the scattering signal. *Journal of Atmospheric and Oceanic Technology*, 6, 254–273.
- Steenveld, G. J., Koopmans, S., Heusinkveld, B. G., Van Hove, L. W. A., & Holtslag, A. A. M. (2011). Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in the Netherlands. *Journal of Geophysical Research: Atmospheres*, 116.
- Stewart, I. D. (2011). A systematic review and scientific critique of methodology in modern urban heat island literature. *International Journal of Climatology*, 31, 200–217. doi:10.1002/joc.2141.
- Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93, 1879–1900. doi:10.1175/BAMS-D-11-00019.1.
- Stout, G. E., & Mueller, E. A. (1968). Survey of relationships between rainfall rate and

- radar reflectivity in the measurement of precipitation. *Journal of Applied Meteorology*, 7, 465–474.
- STOWA (2015). *Actualisatie meteogegevens voor waterbeheer 2015*. Stichting Toegepast Onderzoek Waterbeheer. Beersma, J., J. Bessembinder, T. Brandsma, R. Versteeg & H. Hakvoort, Rapport nr. 2015-10, STOWA, Amersfoort, Netherlands (in Dutch).
- Strangeways, I. (2010). A history of rain gauges. *Weather*, 65, 133–138.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56, 79–107.
- Tauro, F., Selker, J., Van de Giesen, N., Abrate, T., Uijlenhoet, R., Porfiri, M., Manfreda, S., Caylor, K., Moramarco, T., Benveniste, J. et al. (2018). Measurements and observations in the XXI century (MOXXI): innovation and multi-disciplinarity to sense the hydrological cycle. *Hydrological Sciences Journal*, 63, 169–196.
- Ten Veldhuis, J. A. E., & Clemens, F. H. L. R. (2010). Flood risk modelling based on tangible and intangible urban flood damage quantification. *Water Science and Technology*, 62, 189–195.
- Theeuwes, N. E., Steeneveld, G. J., Ronda, R. J., & Holtslag, A. A. M. (2017). A diagnostic equation for the daily maximum urban heat island effect for cities in Northwestern Europe. *International Journal of Climatology*, 37, 443–454. doi:10.1002/joc.4717.
- Theeuwes, N. E., Steeneveld, G. J., Ronda, R. J., Rotach, M. W., & Holtslag, A. A. M. (2015). Cool city mornings by urban heat. *Environmental Research Letters*, 10, 114022. doi:10.1088/1748-9326/10/11/114022.
- Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., ten Veldhuis, J. A. E., Arnbjerg-Nielsen, K., Rasmussen, M. R., & Molnar, P. (2017). Weather radar rainfall data in urban hydrology. *Hydrology and Earth System Sciences*, 21, 1359–1380.
- Thurai, M., Huang, G. J., Bringi, V. N., Randeu, W. L., & Schönhuber, M. (2007). Drop shapes, model comparisons, and calculations of polarimetric radar parameters in rain. *Journal of Atmospheric and Oceanic Technology*, 24, 1019–1032.
- Tokay, A., & Öztürk, K. (2012). An experimental study of the small-scale variability of rainfall. *Journal of Hydrometeorology*, 13, 351–365.
- Trömel, S., Ziegert, M., Ryzhkov, A. V., Chwala, C., & Simmer, C. (2014). Using microwave backhaul links to optimize the performance of algorithms for rainfall estimation and attenuation correction. *Journal of Atmospheric and Oceanic Technology*, 31, 1748–1760.
- Uijlenhoet, R. (2001). Raindrop size distributions and radar reflectivity–rain rate relationships for radar hydrology. *Hydrology and Earth System Sciences*, 5, 615–628.
- Uijlenhoet, R. (2008). Precipitation physics and rainfall observation. Chapter 4. In *Climate and the Hydrological Cycle* (Marc Bierkens, Peter Troch, and Han Dolman, editors) IAHS Special Publications 8 (pp. 59–97). IAHS Press, Wallingford, UK.

- Uijlenhoet, R., Overeem, A., & Leijnse, H. (2018). Opportunistic remote sensing of rainfall using microwave links from cellular communication networks. *WIREs Water*, 5, 1–15. doi:10.1002/wat2.1289.
- Upton, G. J. G., Holt, A. R., Cummings, R. J., Rahimi, A. R., & Goddard, J. W. F. (2005). Microwave links: The future for urban rainfall measurement? *Atmospheric Research*, 77, 300–312. doi:10.1016/j.atmosres.2004.10.009.
- Valtr, P., Pechac, P., Kvicera, V., & Grabner, M. (2011). Estimation of the refractivity structure of the lower troposphere from measurements on a terrestrial multiple-receiver radio link. *IEEE Transactions on Antennas and Propagation*, 59, 1707–1715.
- Van de Beek, C. Z., Leijnse, H., Torfs, P. J. J. F., & Uijlenhoet, R. (2011). Climatology of daily rainfall semi-variance in the Netherlands. *Hydrology and Earth System Sciences*, 15, 171–183.
- Van de Beek, C. Z., Leijnse, H., Torfs, P. J. J. F., & Uijlenhoet, R. (2012). Seasonal semi-variance of dutch rainfall at hourly to daily scales. *Advances in Water Resources*, 45, 76–85.
- Van Leth, T. C., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2018). A measurement campaign to assess sources of error in microwave link rainfall estimation. *Atmospheric Measurement Techniques*, 11. doi:10.5194/amt-11-4645-2018.
- Villarini, G., & Krajewski, W. F. (2010). Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surveys in Geophysics*, 31, 107–129.
- Villarini, G., Mandapaka, P. V., Krajewski, W. F., & Moore, R. J. (2008). Rainfall and sampling uncertainties: A rain gauge perspective. *Journal of Geophysical Research: Atmospheres*, 113.
- Wang, L.-P., Ochoa-Rodríguez, S., Simões, N. E., Onof, C., & Maksimović, Č. (2013). Radar–raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology. *Water Science and Technology*, 68, 737–747.
- Wang, Z., Schleiss, M., Jaffrain, J., Berne, A., & Rieckermann, J. (2012). Using Markov switching models to infer dry and rainy periods from telecommunication microwave link signals. *Atmospheric Measurement Techniques*, 5, 1847–1859. doi:10.5194/amt-5-1847-2012.
- Wauben, W. M. F. (2004). *Precipitation amount and intensity measurements with the Ott Pluvio*. Koninklijk Nederlands Meteorologisch Instituut.
- Wolf, L. A., Siegfried, B. J., & Smith, A. L. (2017). Disdrometer having acoustic transducer and methods thereof. US Patent 9,841,533.
- Wolters, D., & Brandsma, T. (2012). Estimating the urban heat island in residential areas in the Netherlands using observations by weather amateurs. *Journal of Applied Meteorology and Climatology*, 51, 711–721.
- Yang, P., & Ng, T. L. (2017). Gauging through the crowd: A crowd-sourcing approach to

- urban rainfall measurement and storm water modeling implications. *Water Resources Research*, 53, 9462–9478.
- Yang, P., & Ng, T. L. (2019). Fast Bayesian regression kriging method for real-time merging of radar, rain gauge, and crowdsourced rainfall data. *Water Resources Research*, 55, 3194–3214. doi:10.1029/2018WR023857.
- Zahumenskÿ, I. (2004). Guidelines on quality control procedures for data from automatic weather stations. *World Meteorological Organization, Switzerland, WMO-No 955*.
- Zheng, F., Tao, R., Maier, H. R., See, L., Savic, D., Zhang, T., Chen, Q., Assumpção, T. H., Yang, P., Heidari, B. et al. (2018). Crowdsourcing methods for data collection in geophysics: State of the art, issues, and future directions. *Reviews of Geophysics*, 56, 698–740. doi:10.1029/2018RG000616.
- Zinevich, A., Alpert, P., & Messer, H. (2008). Estimation of rainfall fields using commercial microwave communication networks of variable density. *Advances in Water Resources*, 31, 1470–1480. doi:10.1016/j.advwatres.2008.03.003.
- Zinevich, A., Messer, H., & Alpert, P. (2009). Frontal rainfall observation by a commercial microwave communication network. *Journal of Applied Meteorology and Climatology*, 48, 1317–1334. doi:10.1175/2008JAMC2014.1.
- Zinevich, A., Messer, H., & Alpert, P. (2010). Prediction of rainfall intensity measurement errors using commercial microwave communication links. *Atmospheric Measurement Techniques*, 3, 1385–1402.

Statement of author contribution

AB = Alexis Berne (EPFL)	HL = Hidde Leijnse (KNMI)
AD = Arjan Droste (WUR)	LV = Lotte de Vos (KNMI/WUR)
AO = Aart Overeem (KNMI/WUR)	MZ = Marjanne Zander (WUR)
BH = Bert Heusinkveld (WUR)	RU = Remko Uijlenhoet (WUR)
GS = Gert-jan Steeneveld (WUR)	TR = Timothy Raupach (UniB)

The introduction (Chapter 1) and the synthesis (Chapter 7) have completely been written by LV only. The other chapters were written with contributions from co-authors.

The data of the work in Chapter 2 was collected and analyzed by LV. HL, AO and RU assisted LV with the work design and interpretation of the results. The article was drafted by LV, revised by HL, AO and RU, and approved for publication by all.

The work in Chapter 3 was designed by LV, who also collected the data. LV performed the data analysis and interpretation in consultation with HL, AO and RU. AO has generated all map-figures. LV drafted the article that was revised by HL, AO and RU, and approved for publication by all.

The design of the work in Chapter 4 was done by LV, HL, AO and RU. The data was collected by AO, analyzed by LV, in consultation with AO, HL and RU. LV drafted the article that was revised by HL, AO and RU, and approved for publication by all.

The work in Chapter 5 was designed by LV, assisted by HL and RU. TR constructed the dataset, assisted by AB. The background section (Sect. 5.6, originally an appendix to the paper) detailing the construction of this dataset was written by TR. Data analysis was performed by LV, in consultation with HL, RU and TR. LV drafted the article that was revised by TR, HL, RU, AO and AB, and approved for publication by all.

The idea for the work in Chapter 6 was proposed by RU, in consultation with LV, AD, AO and GS. The dataset collection was performed by LV (PWS data for rain, temperature, relative humidity and pressure), AO (CML data, smartphone data and radar reference), BH (AAMS data) and AD (PWS data for wind and AWS reference). LV and AO were responsible for the rain results of the study. AD produced the UHI results assisted by AO, as well as the wind and smartphone pressure results. MZ was in charge of the smartphone results on light. GS, RU and HL assisted in the interpretation of the results. LV and AD were responsible for the final manuscript in which RU wrote the abstract and introduction, GS provided text on case selection and study area and information on AAMS, AO wrote sections on the traditional measurement techniques and smartphone data assisted by MZ, LV provided text on CML and PWS, AD took the lead in writing the results sections assisted by LV and MZ, and HL wrote the conclusion. Critical revision was performed by LV, AD, AO, MZ, GS, RU and HL, and all (including BH) approved the article for publication.

Dankwoord

Stephen Fry, who I greatly admire for his enormous knowledge of both useful and useless facts, stated that having a great intellect is no path to being happy. During my PhD track I have worked towards increasing my intellect, with the help of the often high IQ's of the people around me. For the sake of my happiness however, I also benefited from their various other traits, like their kindness, wit, creativity and empathy. Therefore, in addition to 188 pages of intellectual material, allow me to gush these 2 pages in a most empirical and biased manner, listing the people who have assisted me in making this a happy path as well.

Een PhD traject is een lange, roerige tijd met ups en down. Hier wil ik graag stilstaan bij alle mensen die de ‘ups’ mogelijk maakten en de ‘downs’ minder diep.

Allereerst wil ik mijn begeleiders bedanken. Hoewel ik tijdens mijn master stage op het KNMI met klem betoogde dat een PhD niets voor mij zou zijn wist Hidde Leijnse me te overtuigen om dit toch te overwegen. Ik ben enorm blij dat hij mij dit academische pad opgestuurd heeft, en me in het hele traject als co-promotor heeft begeleid. Hidde, ik ben dankbaar voor de vele inspirerende discussies, goede adviezen, en vooral je onafneembare optimisme en enthousiasme. Samen met mijn andere co-promotor, Aart Overeem, werd ons KNMI kantoor een prettige werkplek waar evenveel ruimte was voor inhoudelijke discussies als voor slechte woordgrappen. Aart, in de samenwerking met jou waardeer ik vooral je oog voor detail, je scherpe blik en nauwkeurige uitleg. Desalniettemin zal ik voornamelijk je humoristische anglicismen missen. Daarnaast heb ik het enorm getroffen met Remko Uilenhoet als mijn promotor. Remko, meer nog dan jouw ervaring en wetenschappelijke kennis, bewonder ik jouw betrokkenheid en vrolijkheid in je werk. Je enthousiasme werkt aanstekelijk, en een overleg met jou leidde altijd tot meer goede ideeën en hernieuwde energie om aan de slag te gaan. Bedankt voor jullie onvermoeibare bereidheid mee te denken, jullie expertise en humor, en voor alle kansen die ik kreeg om naar mooie plekken te gaan en nieuwe dingen te leren.

Ik heb het voorrecht gehad om op twee locaties te werken, en had daarmee fijne collega's van zowel het KNMI en van Wageningen Universiteit. Wat voor mij karakteriserend werd

voor het KNMI is de vriendelijkheid en wetenschappelijke nieuwsgierigheid die ik zag bij mijn collega's, al sinds ik daar begon aan mijn master stage. Bedankt voor alle leuke gesprekken bij de afdelingskoffie, sinterklaascommissie, sportdagcommissie, dragon's dens, borrels, en spontane gesprekjes bij de koffieautomaat. Ook met mijn Wageningse HWM-collega's heb ik veel mooie momenten gedeeld, op excursies in de Ardennen en door IJsland, tijdens de EGU in Wenen, op borrels, en wederom de spontane gesprekjes bij de koffieautomaat. Met name dank aan (ex-)kamerogenoten Sony, Karl en Claudia, die laatste ook voor haar waardevolle hulp bij onderwijstaken. Special mention to all those with whom I have worked together on projects or publications: Tim Raupach, Alexis Berne, Arjan Droste, Marjanne Zander, Gert-Jan Steeneveld, Bert Heusinkveld, Marcel Brinkenbergh, Marie-claire ten Veldhuis, Marc Schleiss, Elena Cristiano, Pierre Kirstetter, Rinske Hutten, Hedy Wessels, Bingyu Zhang and Tamara Schalkx, thank you all for the pleasant collaborations. Also a shout-out to my lovely paranympths Julia and Antonello who are great to have by my side.

Buiten het werk kreeg ik ook veel steun, in de eerste plaats van mijn ouders. Lieve Arie en Els, ongeacht de routes die ik in mijn leven kies staan jullie onvoorwaardelijk in mijn team. Bedankt voor jullie meelevens, lieve woorden en trots. Ook dank aan Pim, Jettie, Juriën, Joris en Ilse, die laten zien op hoeveel manieren je succesvol door het leven kunt gaan. Ik weet dat ik altijd op jullie humor en liefde kan rekenen.

Volgens een mooie quote website zei Charles Darwin dat iemands waarde afgemeten kan worden aan diens vriendschappen. In dat geval had ik geen beter gezelschap kunnen kiezen.

Veel dank aan de WZ-chicks, Berber, Amarens, Liesbeth, Martina, Leonie en Mandy, wat zijn we een eind gekomen sinds we brugwuppen waren.

Mijn Bèta-Gamma studie leverde naast een diploma voornamelijk een grote groep fantastische vrienden op, waaronder Nes, Nus, Remy en Duifje, en natuurlijk de B&D-vriendinnen. Tessa, Agnes, Julia, Jesminne, Emma, Kim, Shayla, Anna, Anneke, Bonnie, Camilla, Danne en Hannah, wat heerlijk dat ik altijd op jullie kan rekenen om een fles wijn open te trekken bij goed of bij slecht nieuws. Naast goede vriendinnen zijn jullie ook nog eens een indrukwekkende groep badasses die me stuk voor stuk inspireren met hoe jullie in het leven staan.

Mijn MPOC-buddies Wytze, Fiona, Sjoerd, Eddy, Sharon, Timo, Steffie en Leon mag ik ook zeker niet vergeten voor hun steun tijdens de soms lastige overgang van scheikunde bachelor naar natuurkunde master, en ook voor hun hulp om te integreren in Utrecht. Met name de donderdag-eetgroep ben ik hier veel dank voor verschuldigd. Daarbij wil ik Leon, en later Sanne, bedanken voor alle huiselijke gezelligheid en oneindige kopjes koffie waarmee Overvecht echt een fijne thuishaven werd.

Aan al deze mensen wil ik zeggen dat ik me gelukkig prijs met jullie aandeel in mijn leven, en ik hoop dat dit ook ver na mijn promotie zo zal blijven. Hora est!

Curriculum Vitae

Lotte Wilhelmina de Vos
26 May, 1989, Gouda, the Netherlands

Education

2001–2007	Pre-university education (VWO), Willem de Zwijger, Schoonhoven
profile <i>Nature & Health</i>
2008	DELFI B2, Institut International d'Études Françaises, Strasbourg
language <i>French for non-native speakers</i>
2008–2011	BSc Béta-gamma, Universiteit van Amsterdam
major <i>Chemistry</i>
2012	High school teaching qualification, Universiteit van Amsterdam
subject <i>Chemistry teacher, 2nd degree</i>
2014	Tesla minor, Universiteit van Amsterdam
track <i>Consultancy at interface of science, business and society</i>
2012–2015	MSc Meteorology, Physical Oceanography & Climate, Universiteit Utrecht
thesis project at <i>Royal Netherlands Meteorological Institute</i>

List of publications & presentations

Peer-reviewed journal publications

- De Vos, L.W., Leijnse, H., Overeem, A., & Uijlenhoet, R. (2017). The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. *Hydrology and Earth System Sciences*, 21, 765–777. doi:10.5194/hess-21-765-2017 (Chapter 2)
- De Vos, L.W., Raupach, T.H., Leijnse, H., Overeem, A., Berne, A., & Uijlenhoet, R. (2018). High-resolution simulation study exploring the potential of radars, crowdsourced personal weather stations, and commercial microwave links to monitor small-scale urban rainfall. *Water Resources Research*, 54, 10293–10312. doi:10.1029/2018WR023393 (Chapter 5)
- De Vos, L.W., Overeem, A., Leijnse, H., & Uijlenhoet, R. (2019). Rainfall estimation accuracy of a nation-wide instantaneously sampling commercial microwave link network: error-dependency on known characteristics. *Journal of Atmospheric and Oceanic Technology*, 36, 1267–1283. doi:10.1175/JTECH-D-18-0197.1 (Chapter 4)
- De Vos, L.W., Leijnse, H., Overeem, A. & Uijlenhoet, R., (2019). Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring. *Geophysical Research Letters*, 46, doi:10.1029/2019GL083731 (Chapter 3)
- De Vos, L.W., Droste, A.M., Zander, M.J., Overeem, A., Leijnse, H., Heusinkveld, B.G., Steeneveld, G.J. & Uijlenhoet, R., (in press). “Hydrometeorological monitoring using opportunistic sensing networks in the Amsterdam metropolitan area”. *Bulletin of the American Meteorological Society*. (Chapter 6)

Other scientific publications

- De Vos, L.W., (2016). Van de regen in de stad. *Nederlands Tijdschrift voor Natuurkunde*, 1. 4–6.
- De Vos, L.W. (2017). Veelbelovend: dichters meetnet in stad dankzij amateur weerstations, Retrieved April 17, 2017, from <https://www.riool.net/veelbelovend-dichter-meetnet-in-stad-dankzij-amateur-weerstations>
- De Vos, L.W., (2017). Na regen komt data-analyse. *Nederlands Tijdschrift voor Natuurkunde*, 9.
- De Vos, L.W., “MOOC II: Co-creating Sustainable Cities” [week 7: Urban Water: Dry Feet for Citizens – 7.3.1 Rainfall Measurements in Cities & 7.3.2 Crowdsensing Rainfall Data]. MOOC offered by the Amsterdam Institute for Advanced Metropolitan Solutions. Retrieved on Nov 22, 2017 from <https://www.edx.org/course/co-creating-sustainable-cities>
- Book of Abstracts: 10th European Conference on Radar in Meteorology and Hydrology (ERAD 2018), 1–6 July 2018, Ede-Wageningen, The Netherlands, Editors: De Vos, L.W., Leijnse, H., & Uijlenhoet, R., Lay-out: Zhang, B., (pp. 983). doi:10.18174/454537

Presentations at conferences & symposia

- GeoBuzz symposium, *Verbetering van stedelijke klimaatmonitoring*, 's-Hertogenbosch, the Netherlands, November 2015. [oral presentation]
- 98th Dies Natalis symposium, *Metropolitan solutions*, Ede-Wageningen, the Netherlands, March 2016. [pitch presentation]
- Vereniging Weerkunde & Klimatologie ALV, *Crowdsourced regenmetingen: Een validatie studie in Amsterdam*, de Bilt, the Netherlands, April 2016. [invited talk]
- EGU2016 conference, *Validation of crowdsourced automatic rain gauge measurements in Amsterdam*, Vienna, Austria, April 2016. [poster presentation]
- ERAD2016 conference, *The potential of crowdsourced automatic weather stations for improved rainfall measurement resolution in Amsterdam*, Antalya, Turkey, October 2016. [poster presentation]
- EGU2017 conference, *Rainfall monitoring with microwave link networks – state of the art*, Vienna, Austria, April 2017. [poster presentation]
- EGU2017 conference, *Urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam*, Vienna, Austria, April 2017. [oral presentation]
- ICUD2017 conference, *High-resolution urban rainfall monitoring by opportunistic sensing*, Prague, Czech Republic, September 2017. [oral presentation]

-
- Boussinesq symposium, *High-resolution urban rainfall monitoring by opportunistic sensing*, Delft, the Netherlands, October 2017. [oral presentation]
 - EGU2018 conference, *Determining the ability of radar, commercial microwave links and crowdsourced personal weather stations to capture small-scale urban rainfall*, Vienna, Austria, April 2018. [poster presentation]
 - EGU2018 conference, *Validation study of nationwide rainfall observations from commercial microwave links in the Netherlands*, Vienna, Austria, April 2018. [oral presentation]
 - ERAD2018 conference, *Commercial microwave links for rainfall monitoring: A nation-wide validation study in the Netherlands*, Ede-Wageningen, the Netherlands, June 2018. [oral presentation]
 - Themadag Weer & Waterbeheer, *Internationaal Radar Composiet (IRC)*, Amersfoort, the Netherlands, November 2018. [invited talk]
 - Nederlandse Hydrologische Vereniging symposium / Nederlandse Vereniging ter Bevordering van de Meteorology najaarssymposium, *Het weer in je achtertuin: crowdsourcen in de stad*, Wageningen, the Netherlands, November 2018. [invited talk]
 - UrbanRain18 workshop, *Opportunistic sensing for urban rainfall monitoring: what spatial and temporal resolutions can be resolved by personal weather stations and commercial microwave links?*, Pontresina, Switzerland, December 2018. [oral presentation]
 - 101st Dies Natalis symposium, *Urban rainfall monitoring: the potential of crowd-sourced weather measurements*, Wageningen, the Netherlands, March 2019. [oral presentation]
 - EGU2019 conference, *Toward operational rainfall monitoring with crowdsourced personal weather stations: suggested real-time quality control method*, Vienna, Austria, April 2019. [poster presentation]
 - IMAU colloquium, *Monitoring rainfall with opportunistic sensing*, Utrecht, the Netherlands, May 2019. [invited talk]
 - Nederlandse Hydrologische Vereniging symposium, *Opportunistic sensoren voor regenmonitoring*, Amersfoort, the Netherlands, June 2019. [invited talk]
 - IPC12 conference, *Dense crowdsourced rainfall observations from personal weather stations: proposed real-time quality control methodology*, Irvine, USA, June 2019. [poster presentation]

Graduate school certificate

*Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment*

D I P L O M A

For specialised PhD training

The Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

Lotte Wilhelmina de Vos

born on 26 May 1989 in Gouda, The Netherlands

has successfully fulfilled all requirements of the
Educational Programme of SENSE.

Wageningen, 6 December 2019

The Chairman of the SENSE board

Prof. dr. Martin Wassen

the SENSE Director of Education

Dr. Ad van Dommelen

The SENSE Research School has been accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW)

K O N I N K L I J K E N E D E R L A N D S E
A K A D E M I E V A N W E T E N S C H A P P E N

The SENSE Research School declares that **Lotte Wilhelmina de Vos** has successfully fulfilled all requirements of the Educational PhD Programme of SENSE with a work load of 49.0 EC, including the following activities:

SENSE PhD Courses

- o Environmental research in context (2015)
- o Research in context activity: Contributing to Massive Open Online Course (MOOC) 'Co-creating sustainable cities', module 7 - Urban Water: Dry Feet for Citizens (2017)
- o Geostatistics (2015)

Selection of other PhD and Advanced MSc Courses

- o Radar QPE, ERAD2016 short course (2016)
- o Modelling of integrated urban drainage-wastewater systems, PhD course organized by DTU and Université Laval at Czech Technical University (2017)
- o Masterclass Git, GitHub and Markdown in a R-environment, KNMI (2018)

Selection of Management and Didactic Skills Training

- o Supervising three BSc students and one MSc student with thesis (2016-2018)
- o Invited guest lecturer at TU Delft MSc course 'Urban Climate and Hydrology' (2017 and 2018)
- o Assisting in the BSc/MSc courses 'Water Quantity and Quality', 'Catchment Hydrology' and 'Urban Hydrometeorology' (2016-2019)
- o Member of 10th European conference on radar in meteorology and hydrology (ERAD) conference organizing committee (2018)
- o Member of coordinating team of the Iceland field trip of the MSc course Catchment Hydrology (2016 and 2017)
- o Participation in selection committee of several open job positions (2017)
- o Three publications in popular media in Dutch (2016-2017)

Selection of Oral Presentations

- o Invited talk: *Crowdsourced regenmetingen: Een validatie studie in Amsterdam*, Vereniging Weerkunde & Klimatologie ALV, April 2016, de Bilt, the Netherlands.
- o *High resolution urban rainfall monitoring by opportunistic sensing*, ICUD2017 conference, September 2017, Prague, Czech Republic.
- o *Validation study of nationwide rainfall observations from commercial microwave links in the Netherlands*, EGU2018 conference, April 2018, Vienna, Austria.
- o *Opportunistic sensing for urban rainfall monitoring: what spatial and temporal resolutions can be resolved by personal weather stations and commercial microwave links?*, UrbanRain18 workshop, December 2018, Pontresina, Switzerland.

SENSE Coordinator PhD Education

Dr. ir. Peter Vermeulen

Special mentions

The research described in this thesis was performed at Wageningen University & Research and the Royal Netherlands Meteorological Institute (KNMI).

KNMI provided funding in the form of MSO project 2013.09. Additional financial support was received from the Amsterdam Institute for Advanced Metropolitan Solutions (RainSense project & project VIR17006, as well as support for AAMS in Chapter 6), Het Waterschapshuis (project Onderzoek Neerslagmetingen) and the Netherlands Technology Foundation (SMART city project 13760). The Netherlands Organisation for Scientific Research provided financial support for Gert-Jan Steeneveld and Arjan Droste (project 864.14.007), who co-authored Chapter 6.

Radar data that was used in the analysis described in Chapter 5 were obtained from the HyMeX program (<https://doi.org/10.14768/mistrals-hymex.721>), sponsored by grants MISTRALS/HyMeX and ANR-2011-BS56-027 FLOODSCALE project. The gauge-adjusted radar rainfall dataset from KNMI is freely available in netCDF4 format, ‘Radar precipitation climatology’ via <http://climate4impact.eu> or in HDF5 format at the KNMI Data Centre https://data.knmi.nl/datasets/rad_nl25_rac_mfbs_em_5min/2.0?q=radar.

The opportunistic sensing data was collected with help from Briac Blanquart and colleagues (Netatmo, Chapter 3 & 6), Ronald Kloeg and Ralph Koppelaar (T-Mobile NL, Chapter 4 & 6) and James Robinson (OpenSignal, Chapter 6). Marcel Brinkenberg of KNMI provided assistance with the experimental set-up of the weather stations (Chapter 2) at the Cabauw Experimental Site for Atmospheric Research (CESAR).

Financial support from Wageningen University and KNMI for printing this thesis is gratefully acknowledged.

Cover design by Lotte W. de Vos

The squares, dots and lines represent the rainfall measurement sampling in the form of pixel-averages, point values and path-averages by weather radar, rain gauges and microwave links, respectively. The skyline refers to the Amsterdam metropolitan area that was used as our study area to evaluate these measurement techniques.